Matrix-Based Image Reconstruction Methods for Tomography — **Source link** ↗

Jorge Llacer, John Meng

Institutions: University of California, Berkeley

Published on: 01 Feb 1985 - IEEE Transactions on Nuclear Science (IEEE)

Topics: Iterative reconstruction, Image processing, Matrix (mathematics) and Tomography

Related papers:

- Maximum Likelihood Reconstruction for Emission Tomography

- EM reconstruction algorithms for emission and transmission tomography.

- Principles of Computerized Tomographic Imaging

- Accelerated Iterative Reconstruction for Positron Emission Tomography Based on the EM Algorithm for Maximum Likelihood Estimation

- A Maximum a Posteriori Probability Expectation Maximization Algorithm for Image Reconstruction in Emission Tomography

Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

Engineering & Technical Services Division

MATRIX-BASED IMAGE RECONSTRUCTION METHODS
FOR TOMOGRAPHY

J. Llacer and J.D. Meng

October 1984

DISCLAIMER

MATRIX-BASED IMAGE RECONSTRUCTION METHODS FOR TOMOGRAPHY

Jorge Llacer and John D. Meng

Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720 U.S.A.

ABSTRACT

Matrix methods of image reconstruction have not been used, in general, because of the large size of practical matrices, ill condition upon inversion and the success of Fourier-based techniques. An exception is the work that has been done at the Lawrence Berkeley Laboratory for imaging with accelerated radioactive ions. An extension of that work into more general imaging problems shows that, with a correct formulation of the problem, positron tomography with ring geometries results in well behaved matrices which can be used for image reconstruction with no distortion of the point response in the field of view and flexibility in the design of the instrument. Maximum Likelihood Estimator methods of reconstruction, which use the system matrices tailored to specific instruments and do not need matrix inversion, are shown to result in good preliminary images. A parallel processing computer structure based on multiple inexpensive microprocessors is proposed as a system to implement the matrix-MLE methods.

INTRODUCTION

Image reconstruction methods based on matrix inversion were discussed quite early in the literature of computerized tomography[1]. Although a certain amount of work has been done in generating algorithms for the use of that basically simple method[2], the large size of the matrices resulting from practical applications, the bad behavior of the matrices under inversion and the success of Fourier based algorithms has discouraged the continuation of the initial efforts.

As part of the work carried out at the Lawrence Berkeley Laboratory in support of a heavy ion cancer therapy, it has become necessary to provide an imaging capability for annihilation radiation resulting from radioactive ions injected by the BEVATRON[3]. The circumstances are such that it is not possible to use a detector ring of conventional design. Interference with the beam delivery and with patient positioning have required the design of a detector system with an incomplete set of projections. The requirements of the project forced us to give a second look at matrix reconstruction methods with findings that lead to the design and construction of a successful imaging instrument[4,5].

Further examination of the properties of weight or response matrices resulting from accurate simulations of detectors in ring geometry indicate that there may be more merit to matrix-based reconstruction methods than was apparent initially. First, with a proper description of the positron ring geometry, it is found that the resulting matrices are well behaved (low condition numbers, CN). Second, it is not necessary to invert (or pseudo-invert) the very large matrices obtained in practice. Instead, one can use the maximum likelihood estimator (MLE) method of reconstruction described by Shepp and Vardi[6] which makes use of the weight matrices directly for

its implementation. The MLE method yields reconstructions with lower noise than filtered backprojection techniques in simulations of positron emission tomography without time-of-flight (TOF)[6] and with TOF[7]. The improved quality reconstruction with real data from a conventional BGeO ring detector has recently been confirmed by Shepp et al[8].

Substantial advantages could result from the use of matrices in reconstruction. Since the matrices can be calculated accurately by a combination of Monte Carlo and deterministic methods, or can be measured directly, the image reconstruction process is precisely tailored to each specific instrument. Resulting images can be free of spatial distortion, with the reconstructed image of a point having the same shape everywhere in the image volume. Moreover, there is added flexibility in the geometry of the instrument. A complete set of views is not a requirement and the response of the detector system to a point source (before reconstruction) does not need to be space invariant. The utilization of all the radiation detected in a multi-ring system, for example, is then possible.

This paper will present the preliminary findings on which the above assertions are based, describe computer simulations and measurements with a 96-crystal Bismuth Germanate (BGO) ring geometry, and show results of image reconstructions. It will also propose a method for implementing the MLE algorithm for practical problems with parallel computer architectures based on inexpensive microprocessors. Although the method to be discussed is not limited to positron emission tomography (PET), that specific application will be used as a framework for the ideas to be described.

THE SYSTEM MATRIX

The concept of the system, weight or response matrix of an imaging device has been published in the context of positron emission imaging by one of the authors[9,10]. The physical meaning of the "blurring" matrix obtained by multiplying the system matrix by its transpose, the meaning of the eigenvalues and eigenvectors of the blurring matrix and a description of the "pseudo" or generalized inverse for image restoration has also been discussed in detail in the publications. The basic ideas are well known in the general field of image restoration, as described, for example, by Andrews and Hunt[11]. It will suit the purpose of this section, however, to describe briefly the basic concepts indicated above.

Consider an imaging system consisting of two rows of annihilation radiation detectors that can rotate by an angle θ about the center of an $(n \times n) = N_p$ array of pixels, as shown in Fig. 1. Coincidences will only be allowed between detectors which are facing each other directly, so that there can be N_c coincidences at each position θ. This PET system is derived directly from the early X-ray CT systems, and we shall call it the X-CT system. We consider the case of uniform activity in each pixel.

1
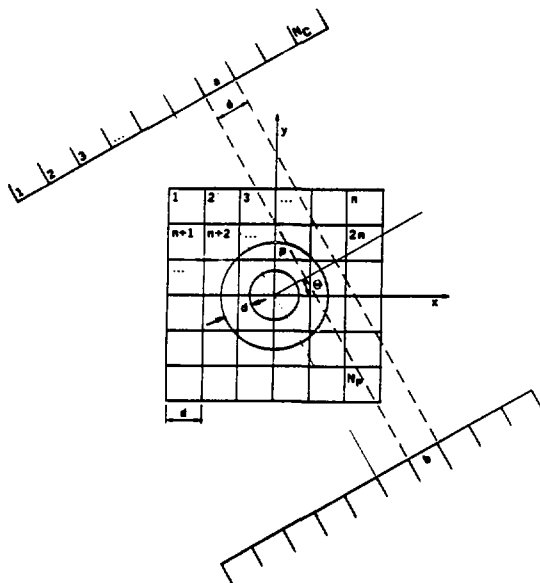
XBL 8111-12313

Fig. 1 Representation of the PET problem in terms of the traditional X-ray CT geometry. The possible coincidences in a positron ring detector system are rebinned so that they conform to one of the coincidences allowed in the X-CT problem. The casting of the problem in that fashion results in radial sampling at a distance d, equal to the detector dimensions.

The system matrix A of the X-CT system is obtained by assuming that a unit activity is placed in pixel No. 1, the detectors are made to step in N_θ angles from $0 \le \theta < 180°$ spending equal time in each angle, and recording the resulting coincidence rates as the $N_c \times N_\theta$ elements of the first column of matrix A. The unit activity is then placed in pixels Nos. 2,3, ... to N_p, each time obtaining one column of A. The resulting matrix has $N_c \times N_\theta$ rows and N_p columns.

In principle, the imaging problem consists in solving the equation

$$A x = k \qquad (1)$$

where x is an unknown vector of length N_p corresponding to some activity distribution in the pixels and k is a vector of length $N_c \times N_\theta$ which is the result of a measurement. Thus, the A matrix contains all the information about the pixel-detector system necessary to solve the linear imaging problem. The A matrix is also the set of probabilities p(b,d) that an annihilation gamma pair emitted by a pixel b will be detected in a tube d connecting a specific pair of opposing detectors. The probabilities p(b,d) provide all the information needed about a detector system for the MLE image reconstruction algorithm described by Shepp and Vardi[6].

The solution to Eq. (1) is equivalent to solving

$$A^T A x = A^T k \qquad (2)$$

where A^T is the transpose of A. We can then define $A' = A^T A$ and $k' = A^T k$ and rewrite Eq. (2) as

$$A' x = k' \qquad (3)$$

where A' is symmetric and of dimension $N_p \times N_p$ and x and k' are also of dimension N_p. Matrix A' is the blurring matrix of the system.

The solution to Eq. (3) would be straightforward if the inverse of A' could be obtained with reasonable ease and reliability for practical imaging systems. If we let B correspond to a "pseudo" inverse obtained from eigenvector decomposition of A, then the unknown activity distribution x can be found by carrying out the operation

$$x = B k' . \qquad (4)$$

the matrix A as a prescription. Multiplication by B corresponds to a filtering operation in which each particular pixel has its own filter function. Thus, image restoration using the information contained in a system matrix is equivalent to the filtering of a backprojection in which the basis functions used for the expansion of activity distributions are not the Fourier basis (as in standard filtered backprojection techniques), but are, instead, the eigenvectors of the matrix A[9,10]. The implications of the above observations are that an imaging system does not have to have a space invariant response to a point source, and does not require what is usually called a complete set of projections in order for a successful reconstruction to be feasible.

Two main difficulties appear in attempting to solve the imaging problem by the pseudo-inverse: with the exception of some particular geometries in which the determination of the point of origin of gamma pairs offers relatively little ambiguity[10], the blurring matrices of positron detector systems appear to be very badly behaved. In particular, the condition number, CN (largest divided by smallest eigenvalue) of the blurring matrix for the X-CT geometry is extremely high, implying very high sensitivity of the solution to statistical fluctuations in the experimental data. Condition numbers in the region of 250,000 have been obtained for pixel arrays as small as 8 x 8 and pseudo inverses for 13 x 13 pixel arrays have been found impossible to calculate even with a 48-bit mantissa[10].

The second difficulty is less fundamental in nature and it concerns the large size of the system and blurring matrices resulting from most practical detector systems. It would appear, however, that with the constantly improving price/performance ratios for microcomputers and memories, a solution can be found for this problem. Indeed, we want to propose one such solution in this paper.

THE POSITRON RING GEOMETRY

A close examination of the X-CT geometry of Fig. 1 has been undertaken in order to find the reason for the very high CN's encountered. Some of the findings have been surprising. For example, a 2 x 2 pixel array in which the detector lateral dimension is equal to the pixel side dimension results in a singular blurring matrix (its determinant, which can be calculated without truncation errors for such a small problem, is exactly zero). The singularity disappears when the detector side dimension is made smaller than the pixel dimension. This fact, coupled

2

to an earlier observation that the CN of a blurring matrix increases rapidly as the pixel size approaches the sampling distance[12], leads one to the preliminary conclusion that the X-CT problem is badly behaved because it is in violation of the sampling theorem. This point certainly deserves further analysis, but it suggests that a more appropriate representatation of a positron ring detector problem in terms of the actual coincidences that can occur, instead of rebinning it into an X-CT geometry, can be better behaved. In the positron ring (PR) geometry, as described in Fig. 2, the image plane is sampled at twice the radial frequency than in the X-CT geometry for detectors and pixels of the same lateral dimension.

25 x 25 pixel image array. The size of the image array was limited to 625 pixels in order to keep the lateral distance between pixel centers (1 cm) larger than the sampling distance (0.75 cm).

The calculated system matrix has been multiplied by its transpose and the blurring matrix A' has been obtained. The eigenvalues of the 625 x 625 matrix have been calculated in a VAX-11/780 and the condition number has been found to be 1,323. The PR problems is, therefore, reasonably well behaved and amenable to exact solution, without the approximations implied by the use of Fourier-based algorithms.
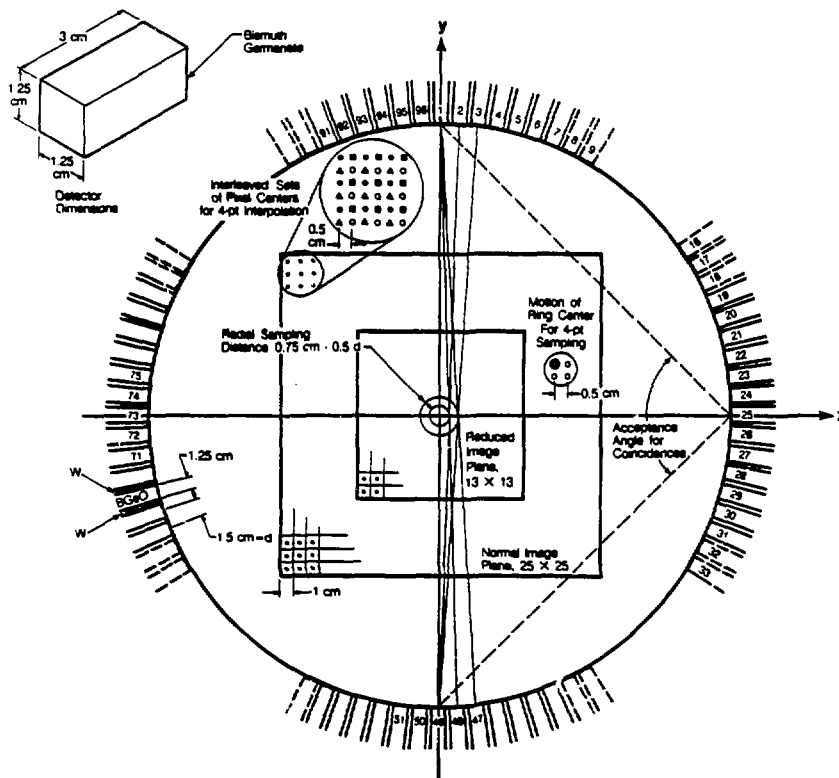

Fig. 2 Schematic drawing of a 96-crystal PET ring detector used for the computer imaging simulations and for the actual image generation experiments with moving detectors. The response matrix of the detector system was obtained for the Normal Image Plane (25 x 25 pixels) when all detectors were active. The Reduced Image Plane (13 x 13 pixels) was used when only the detectors indicated with the solid lines were active, resulting in a severely limited number of coincidences. Radial sampling distance for the system is 0.5 d. For experiments with 4-point sampling, the center of the detector ring was assumed to move in the square pattern shown. For 4-point interpolation, four interleaved sets of pixel centers were also used, as shown in the expanded section at the upper left side of the image plane.

In order to confirm the above expectation, the system matrix of a 96-crystal (BGeO) positron ring, as described in Fig. 2, has been calculated using a modified form of the computer program MATRIX, which was very successful in calculating system matrices for a two-dimensional planar positron camera (PEBA II)[4]. Figure 2 shows the dimensions of the ring and of the

Figures 3 a) and b) show the calculated blurring functions for a central and one edge pixels, respectively. The substancial differences between center and edge pixels are due to differences in solid angles, in crystal detection efficiency as a function of incidence angle and to crosstalk between neighboring crystals, effects which are included in the MATRIX

calculation. Some unsymmetries apparent in Figs. 3 are caused by small errors in the present method of computing the response matrices, which will be corrected in the future.

In order to test the correctness of the sampling concepts, we have assumed that the detector ring of Fig. 2 moves in such a way that its center comes to rest, for counting, at each of the four positions
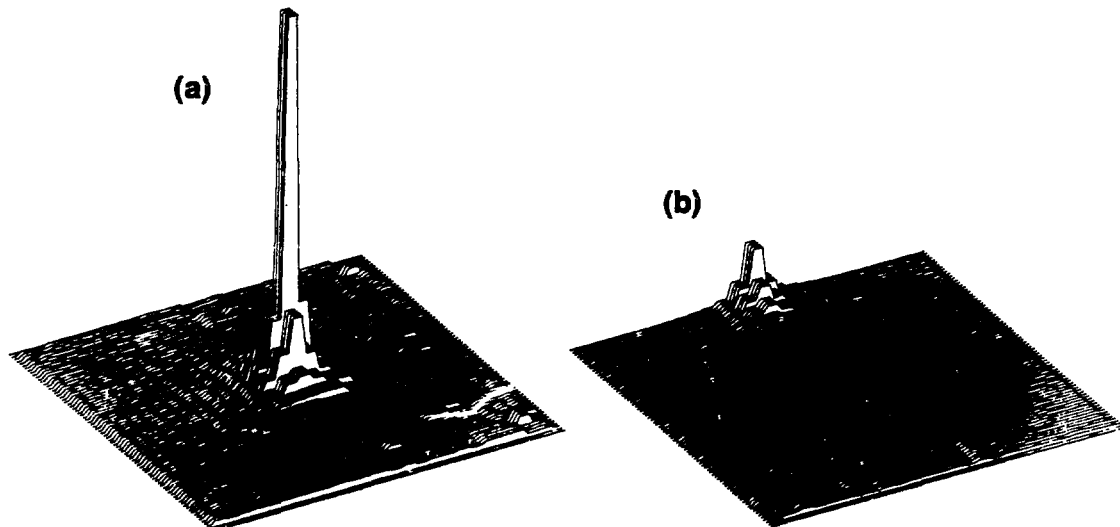
Fig. 3 a) Blurring function of the complete detector system of Fig. 2 for the center pixel of the image plane. This corresponds to the center column of the blurring matrix A', containing N_p elements, plotted on the square image plane. b) Ditto for a pixel at the edge of the image plane.

The calculation of eigenvectors and of the pseudo inverse has not been undertaken at this time. We will only demonstrate the feasibility of tomographic imaging using the calculated response matrices by means of the MLE method, which only needs the A matrix for its implementation.

SAMPLING AND INTERPOLATION

The role of sampling of the image plane in the context of matrix methods is consistent with the general understanding of the subject: Increasing the sampling frequency by using a larger number of smaller detectors and/or by incorporating motions in a detector system makes the CN of a blurring matrix smaller, for a fixed pixel dimension. In turn, this results in lower noise magnification factor. The "Signal Amplification" concept of Phelps et al[13,14], which increases the signal-to-noise ratio in PET reconstructions, is based on an equivalent argument, although cast in the more familiar spatial frequency domain.

When motions are incorporated in the design of a detector system, the columns of a system matrix become longer since they have to incorporate the detector coincidence rates for one pixel at all the detector positions considered. In the case of TOF, each element of a matrix column becomes expanded to a TOF spectrum with as many bins as TOF bins. Fortunately there is a large number of zero (or near-zero) terms in every column, so that sparse matrix techniques can be used for storage and handling of the significant terms.

forming a square with sides of 0.5 cm., as indicated in the center of the figure. This increases the sampling frequency by a factor of two in the x and y directions. Considering a column of A as a vector of a certain Euclidean norm, only elements with magnitude larger than 5% of the vector norm have been kept as significant. We find approximately 450 such elements in each column of the expanded A matrix (625 columns). The CN of the new blurring matrix is found to be 173, a substantial improvement over the value of 1323 for the initial problem. As discussed in Ref. 10, the use of TOF information also reduces greatly the CN of the tomography problem.

Independently of a sampling rate increase achieved by motion (or by skewed arrangement of detectors, as in the case of PEBA II)[4], system matrices can be generated that correspond to different interleaved sets of pixels in the imaging plane. Each one of these independent matrices can be called a submatrix. One set of measured data can then be processed with the independent submatrices providing interpolated values of activity in the image plane. We have carried out simulations and imaging experiments with four submatrices, each with 625 points, giving interpolation values every 0.5 cm in the x and y directions. An inset at the upper left hand side of the image plane in Fig. 2 shows the locations of the four sets of interleaved system points (pixel centers).

The significance of this interpolation technique is that the reconstruction of a large array of pixels can be broken down into several reconstructions of smaller interleaved arrays. The number of operations needed for a reconstruction can be expected to be proportional to the square of the number of pixels and, therefore, the smaller reconstructions can result in substantial savings in computation time.

IMAGING SIMULATIONS

For the purpose of carrying out a first test of the possibility of using system matrix information in conjunction with the MLE recontruction method, the imaging of two point sources has been simulated with the program MATRIX by calculating a set of detector responses for the positron ring of Fig. 2. Statistical fluctuations corresponding to 100 microCi in each source have been introduced into the coincidence data and probability functions obtained from system matrices and from the Shepp and Vardi (S-V) model[6] have been used with the MLE algorithm to generate images. The activity of the sources was chosen relatively high to avoid strong statistical effects in the results.

Figures 4a) and b) show the images obtained with system matrix and the Shepp and Vardi (S-V) models, respectively. The point sources were placed at (x,y) = (0,0) and (9,-9) cm in the image plane. Four-point sampling of the image plane was assumed by motion of the ring in the square pattern shown in the center of Fig. 2 and 4-point interpolation was used by reconstructing with four interleaved submatrices corresponding to the four sets of pixels also shown in the

figure (upper left image plane). Linear interpolation was finally used to generate the 128 x 128 display from the 50 x 50 image.

Measurements on the images of Figs. 4a) and b) indicate that the system matrix approach gives the correct intensity for both sources and that the shape of the corner source is almost identical to the central one. The S-V model results in a value for the corner source which is only 57 of the correct one (no detector efficiency information in the S-V model) and its shape is distorted appreciably (no detector crosstalk in the model). We can conclude from this first experience that response matrices calculated specifically for a detector system can result in more accurate imaging, particularly in the peripheral regions of the image plane.

IMAGING EXPERIMENTS

Using two BGeO detectors of dimensions shown in Fig. 2 and some simple equipment, the positron ring of Fig. 2 has been simulated by rotations in the horizontal plane. Each BGeO crystal was flanked by two identical crystals (not operating as detectors) and 0.5 mm tungsten separators were used. Energy threshold discriminators were set at approximately 125 keV matching the conditions under which the system matrices were generated. A Na-22 line source of undetermined uniformity of 14 cm length and two Na-22 point sources embedded in plastic rods 3 cm in diameter were placed in the image field of view. Their activities were approximately .078 microCi per cm, 1.1 and .185 microCi, respectively.


(a)   (b)

24 cm
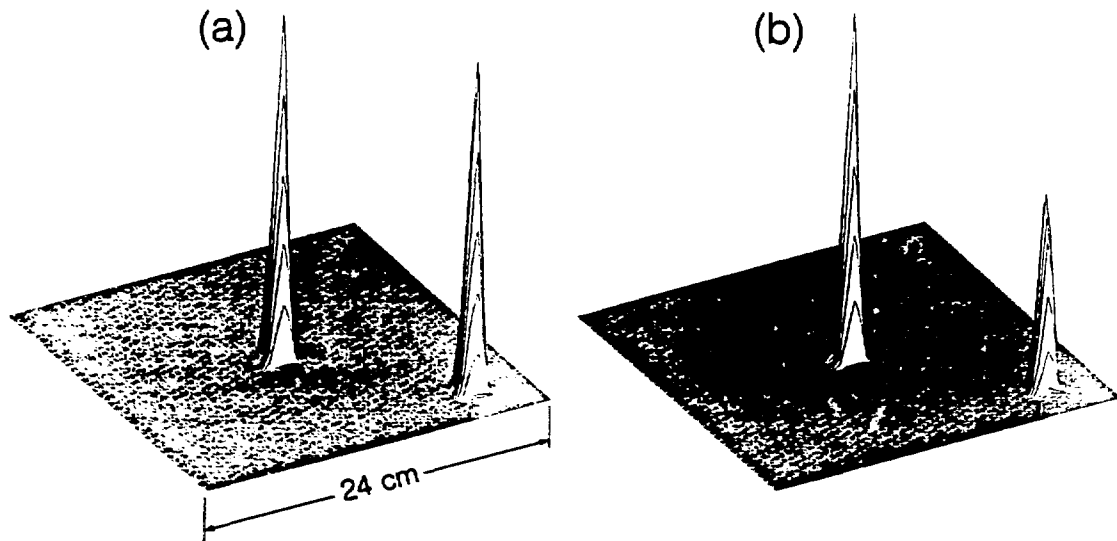
XBL 845-8432

Fig. 4 a) Results of computer simulation for the reconstruction of two point sources at the center and at (9,-9) cm in the image plane of Fig. 2. The MATRIX program was used to generate both the data and the response matrices of the instrument. Four-point sampling by simulating motion, and 4-point interpolation by assuming 4 sets of interleaved 25 x 25 pixels were used for the reconstruction by the MLE algorithm. b) Ditto using MATRIX for data generation and the Shepp-Vardi probability functions for the MLE algorithm. Since the latter does not include detector efficiency differences with detector material, incidence angle and position or detector crosstalk, the point near the corner appears to be of lower activity than the center point, and it has distortion which is somewhat observable in the figure.

5

After data acquisition by a stationary ring (no 4-point sampling), image reconstruction by the MLE algorithm using computed response matrices and also matrices obtained from the Shepp-Vardi (S-V) model[16] were carried out. Four-point interpolation was used for both the matrix and the S-V-MLE reconstructions. With the weak sources used for the experiment and a counting time of 100 sec per position, a total of 123,000 counts was gathered for the image. The existence of a 1.2 MeV gamma ray in the Na-22 source in time coincidence (but no angular relationship) with an annihilation pair results in a substantial number of background counts. The data obtained were also used for a fan-beam reconstruction with the Donner algorithm package[2].

Figures 5a), b) and c) show the images generated by the matrix-MLE, S-V-MLE and the fan-beam filtered backprojection methods, respectively. All three have been reconstructed on a 50 x 50 pixel image plane, with linear interpolation to 128 x 128. There is a large degree of similarity between the two MLE results. Expansion of the vertical scales of Figs. 5a) and b) by a factor of 4 shows no meaningful differences either. Evidently, an assessment of the advantages of using accurately calculated system matrices will require more extensive research than can be shown at this time.

(b)
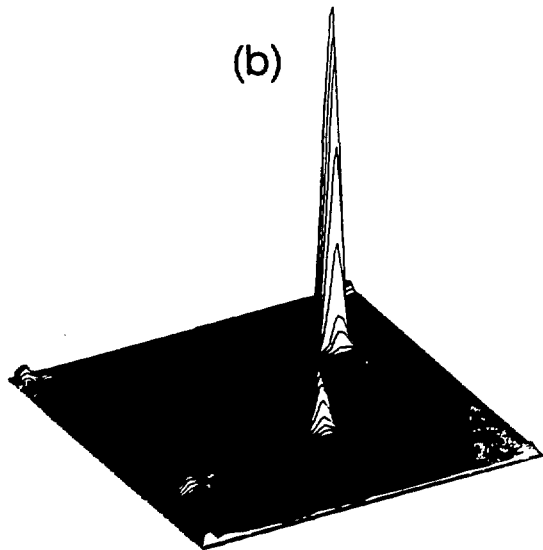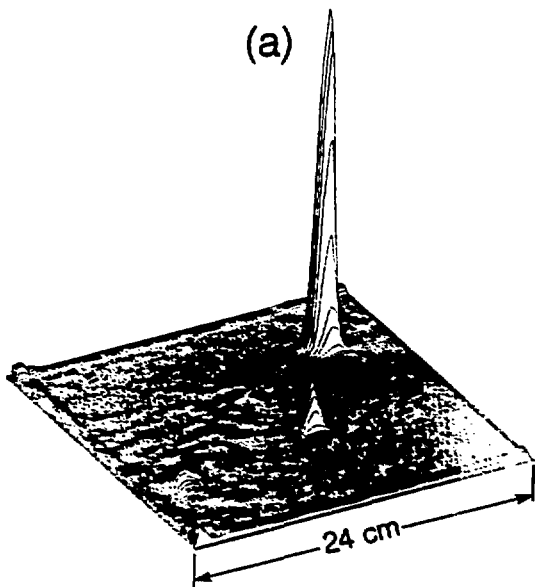
b) Same data as a), reconstructed using the Shepp Vardi calculated probability functions, also with 50 x 50 pixels by interleaving.

(a)

(c)

24 cm
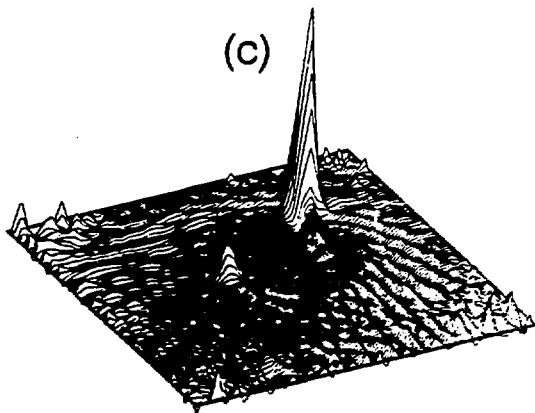
XBL 8410-4097

Fig. 5 Results of image reconstructions using rotating BGeO detectors simulating the geometry of Fig. 2 and small Na-22 sources:

a) Complete detector ring, 123,000 counts in the image. Reconstruction by using the calculated response matrix for the system and the MLE algorithm. Stationary ring, (no 4-point sampling) and 4 interleaved sets of pixel centers, for a 50 x 50 pixel reconstruction.

c) Same data as a), reconstructed using the fan beam filtered backprojection technique. The line source is almost impossible to find among the many artifacts and noise. The reconstruction was made for 50 x 50 pixels. When a 25 x 25 pixel reconstruction was carried out, noise and artifacts were of lower frequency but information in the image did not appear to be better. Notice that this image is mirrored with respect to a) and b).

There is a striking difference, however, between the MLE and the fan-beam reconstructions. Reconstruction artifacts and/or noise in the fan-beam image have obscured the presence of the line source, which is clear in the MLE reconstruction. This finding is in agreement with the expectation that the MLE algorithm would be most useful in severely count limited images, when the Poisson nature of the positron annihilation process would be most noticeable[6] and with the conclusions reached by Shepp et al by comparing images from a real PR instrument[8].

SEVERELY LIMITED NUMBER OF PROJECTIONS

For the purpose of testing the idea that the matrix-MLE method does not need a complete set of projections to give a useful reconstruction, it has been assumed that the ring of Fig. 2 has been modified by removing 44 of the 96 detectors. The remaining detectors are arranged in 4 groups of 13 detectors each (shown in solid lines in Fig. 2) and a detector of one of the groups can have coincidences only with any of the detectors of the opposing group. The number of coincidences has been reduced from 2352 to 338. A new reduced image plane of 13 x 13 cm has been defined (169 pixels) for the new instrument. The CN of the new system matrix is near 19,000, considerably worse than in the case of the complete ring. Two
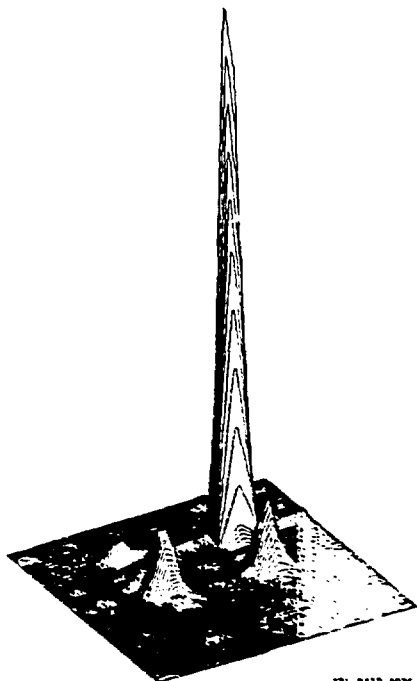
xxl 8410-4096

Fig. 6 Reconstruction of 3-point sources on the Reduced Image Plane of Fig. 2, with only the four groups of detectors indicated with solid lines in that figure. In spite of the severe cut in projections and angles, only one significant artifact appears in the upper left hand, at the conjunction of the projection of the large peak and one of the two smaller peaks. Only 20,100 counts were collected for the image.

Na-22 point sources of 1.1 and .185 microCi and one flat source of 1.6 cm diameter and activity .35 microCi embedded in plastic have been imaged. The total number of counts in the image is 20,100, and the result is shown in Fig. 6. In spite of the low number of counts and the unfavorable CN, the three point sources have been imaged reasonably well. Only one artifact of substantial magnitude has appeared, the far left peak in Fig. 6) and the related ridge connecting it to the large peak.

MULTIPROCESSING CONCEPTS

There is little doubt that the MLE method requires a very large amount of fast memory to store the probability functions for any problem of practical size. In particular, in the case of multi-ring systems with utilization of all the detected radiation, vast amounts of matrix element storage would be needed. Shepp et al[6,7] have recognized the problem and resorted to calculating the required probabilities as they were needed for the reconstructions which they have published, a method which would evidently be too slow for a practical implementation of the algorithm.

In order to find a practical and economical solution to the problem of implementing the matrix-MLE algorithm, attention has been turned towards multiprocessing in a manner that appears to be somewhat unconventional but promising: using large numbers of readily available complete microprocessor boards, and large amounts of inexpensive memory in a modular architecture such that a user can increase or decrease the number of boards depending on the size of the reconstruction problem at hand. Initial tests carried out by comparing performance of a VAX-780, a PDP-11/34 and a modest IBM-PC fitted with the 8087 floating point processor lend support to the above idea.

The tests that have been carried out fall into two categories: bench-mark speed tests and practical computing limitations tests. For bench-mark tests, we have coded a sparse vector dot product routine, which is one of the two most important loops in the MLE reconstruction and compared the performance of the three computers. The results are given in the following table and correspond to the time needed to carry out the dot product of a full 4096 element floating point vector with a 200 element vector in sparse storage form (each element contains one integer address and one floating point value):

TABLE I – SPARSE VECTOR DOT PRODUCT PERFORMANCE		
Computer	Single Precision	Double Precision
VAX-780	0.16 secs	0.25 secs
PDP-11/34	1.43 secs	1.75 secs
IBM-PC/8087	3.24 secs	3.41 secs

The VAX and PDP codes were written in FORTRAN, while the IBM-PC code was written in C-language and the resulting assembly language program modified slightly for some degree of optimization. The results indicate that in single precision, the VAX is only 20 times faster than the IBM-PC and in double precision the factor is decreased to 13.6 (notice that the 8087 processor is a double precision unit). The speed of the PDP-11/34 (with its standard floating point processor) is only a factor of two higher than the IBM-PC. This factor remained approximately constant when the tests were carried out with a Monte-Carlo Compton scattering calculation.
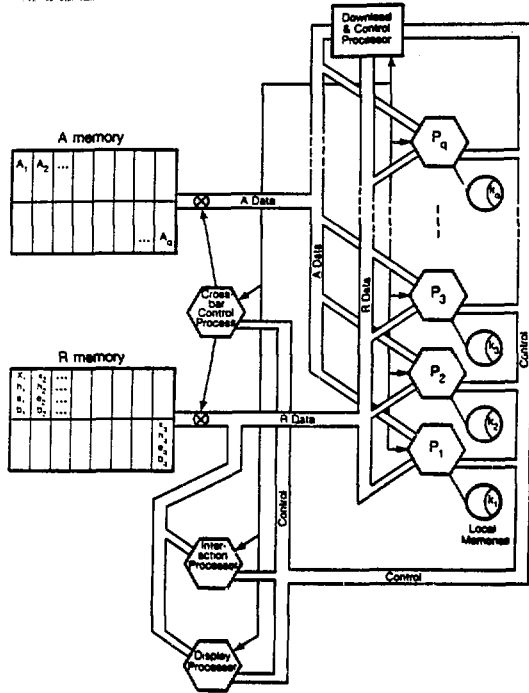
7

A memory

R memory

XBL 846-9462

Fig. 7 Preliminary form of a multiprocessor system for the efficient implementation of the MLE algorithm. Segmented memories whose contents can be connected to the different microprocessors in a few microseconds in a rotation manner allow for a linear increase in computing power with increased number of processors without memory contention or arbitration. The latter would invariably result in speed saturation effects.

The backprojection of Eq. 7 can be carried out in parallel by making each segment of vector e, $e_1 \ldots e_q$, residing in the different sections of the R-memory, simultaneously available to each processor. Each processor computes dot products of its private copy of part of vector e with its private copy of the appropriate n_{col} columns of A and the results will be partial values of the n_{col} elements of the backprojection vector b.

The segments of e are rotated upwards (P_1 connected to e_2, P_2 to e_3 ... P_q to e_1) by memory switching, new partial dot products are accumulated and the rotation continues until each processor has seen each segment of e. At that point, each processor has calculated N_b/q elements of the backprojection b, which is stored in segments of the R-memory.

The A-memory and the R-memory are formed of q independent sectors, and each sector is hardware switched to a processor, avoiding the time needed for word-by-word data flows and avoiding arbitration conflicts when all the processors are simultaneously reading their segments of the A-memory or R-memory.

With the matrix A stored by columns in sparse format, the product of Eq. 5 becomes more complicated. With individual segments of vector x available to each of the q processors, dot products of rows of A with vector x have to be computed. If each segment

of the A memory contains complete columns of matrix A, then partial dot products of x with the elements of the desired rows from one sector of the A-memory can be formed. All sectors of the A-memory are hardware switched in sequence to all the processors, until each processor has had access to all of matrix A. Simultaneously, the segments of x are rotated upwards. At the end of parallel processing with matrix and vector switching, each processor has the values of $m = N_c/q$ different elements of vector h accumulated in the R-memory.

The download, control, interaction and cross-bar processors coordinate the general input-output, carry out the necessary control and memory switching operations, monitor convergence of the iterative process and eventually activate the display processor.

The cluster of Fig. 7 can be replicated a number of times, with the control processor of each cluster responding to a central control processor which does the ultimate directing of all activities. Each cluster would process one of the interleaved sub-images into which a complete image is separated, as discussed above. The multiplication of processing capacity with increase of number of processors q or replication of clusters is assured by the lack of an arbitration need for memory read and write operations. Data needed by a processor is always available upon demand in either the A, R or local memories.

Using modern 32-bit microprocessor boards with 1.76 Megabytes per board, we estimate that a computer with 16 MLE processing boards would complete one iteration of a 32 x 32 pixel subimage from an instrument like the Columbia University PET Scanner (one ring with 63 detectors, with an effective increase in detector pairs by a factor of 36 by motion)[8] in approximately 9 seconds. Sixteen clusters would then carry out the complete 128 x 128 reconstruction (one 32 x 32 subimage per cluster) at a rate of 9 seconds per iteration. A number of factors remain to be studied with both the MLE algorithm convergence as a function of detector ring design (condition number of matrix), and with the utilization of symmetry in the response matrices. Those factors could result in substantial savings in computation times.

THE BACKPROJECTION-MLE VARIATION

There is one interesting variation on the utilization of the MLE algorithm that deserves attention. We will call this variation the backprojection-MLE method. The normal matrix-MLE solves the imaging equation $A x = k$, as described by Eq. 1. It uses the experimentally obtained vector k to obtain the image x using the system matrix values A as the probability function for the MLE process. An alternative would be to solve the problem $A' x = k'$ of Eq. 3, where $A' = A^T A$ and $k' = A^T k$. Vector k' is the backprojection of k using matrix A as a prescription and A' is the blurring matrix of the imaging system. The blurring matrix can be interpreted as a probability function $a'(i,j)$ that activity in pixel i of the true image will appear in pixel j of the backprojection. The interesting point about this development is that matrix A' is symmetric and it contains as many columns as pixels. Although it is not sparse, except to a significant extent in the case of TOF tomography, the number of matrix elements that needs to be handled by the MLE algorithm is considerably reduced. The simplification in calculation procedure for the backprojection-MLE method is significant. From the theoretical point of view, however, it is not clear at this time whether the fundamental premise of the MLE

8

Factors like memory capacity and allocation of memory and CPU time are important in determining performance of a large computer system. On the other hand, software support can be a problem with small machines designed primarily for personal use. In order to assess possible limitations in those areas, a complete matrix-MLE algorithm has been written in FORTRAN and has been run in the three computers indicated above. In each case, the system matrix for the 96 crystal ring of Fig. 2 has been used (single point sampling), and the experimental data that gave the results of Figs. 5 were loaded into the machines. The system matrix contained approximately 70,000 significant elements and the results vector allowed 2592 coincidence pairs.

For the calculations with the VAX-780 memory allocation was increased to approximately 750 kilobytes, so that no swapping between fast memory and disk would obscure the results. A 97% CPU utilization was maintained during the computations. The results converged to 0.5% accuracy in the large peak of Fig. 5a) in 16 iterations. After all the input matrices and data were loaded, each iteration took 1.75 seconds (single precision arithmetic).

The IBM-PC turned out only slightly more difficult to program and debug than the VAX, since its 16-bit main processor cannot address directly all its memory. The FORTRAN used for this particular test is quite complete, however, and provides access to all the memory through the definition of COMMON blocks. With 640 kilobytes of memory, the IBM-PC had no difficulty containing the complete system matrix and experimental data. The iteration time was 30 seconds (double precision operations from single precision data) corresponding to a factor of 17 slower than the VAX-780, a value which is consistent with the dot product benchmark tests indicated above. Results obtained with the PDP-11/34 cannot be considered to be in the same range with the above two machines. The limitation of 128 kilobytes of memory makes for considerable disk swapping and iterations take approximately three minutes, even with a fast array processor doing the calculations.

The above findings point towards a specific direction for implementation of the matrix-MLE methods at a reasonable cost: A microprocessor board like tnat of the IBM-PC with the 8088/8087 processors and 640 kilobytes of memory may cost in the order of $1500. Twenty of those boards (plus one additional control board) connected in a suitable architecture would have the computing power of a VAX-780 for that specific problem at a cost of $31,500 (excluding peripherals and development), considerably lower than a VAX-780 and 12.8 megabytes of expensive memory. Since the 8088/8087 processors are outdated, one can expect even more favorable price/performance ratios for the new 32-bit "super-micros" which are now finding application in mass produced personal and business computers.

A PRELIMINARY MULTI-PROCESSOR ARCHITECTURE FOR THE MLE ALGORITHM

Although the definition of a final, detailed architecture for the matrix-MLE multiprocessor will require substantial research, an analysis of the MLE algorithm from the point of view of multiprocessing will be given here. This will allow the definition of a preliminary processor structure which could be the starting point for more complete research.

The process of image reconstruction by the MLE algorithm[6] (matrix-MLE or S-V-MLE) can be described readily in terms of a general emission tomography problem by the use of vector algebra. We consider first the image space to be divided into N_p pixels (or voxels, for a true 3-dimensional reconstruction). Each pixel is represented by one element of a vector x whose value we want to find. Since the MLE is an iterative procedure, we let x represent an old value for the vector and x' represent a new value after one more iteration.

As a result of a measurement, an imaging instrument will yield a vector of results k. The length of this vector will be N_c. For a non-TOF positron ring system, N_c will be the total number of possible coincidences, including those between different rings in a multiring system. For TOF tomography, N_c will be the number of coincidences multiplied by the number of bins in the TOF measurement and for SPECT, it will be the number of divisions of the imaging camera times the number of positions in the camera rotation. The system matrix A of an imaging instrument will have elements $a(i,j)$ corresponding to the probabilities that a point source at the jth pixel will give a response in the ith element of the results vector k. Matrix A will have N_c rows and N_p columns and, in general, will be very sparse.

It will be convenient to define a vector h of length N_c given by

$$h = A x \qquad (5)$$

which corresponds to the results vector that the imaging instrument would yield if the activity in the imaging space were x. Further, let an error vector e with elements

$$e(i) = \begin{cases} k(i) / h(i) & \text{for } h(i) \neq 0 \\ 0 & \text{for } h(i) = 0 \end{cases} \qquad (6)$$

be defined and its backprojection

$$b = A^T e \qquad (7)$$

be obtained. Then, the MLE algorithm calculates the new values x' by

the products

$$x'(j) = x(j) \, b(j). \qquad (8)$$

In order to define a preliminary architecture for the above iterative procedure, let's consider the matrix A as being stored by columns in sparse vector format. As described in Fig. 7, assume that q processors, $P_1, P_2, \ldots P_q$, forming a cluster, are available for the calculation and that each has direct access to matrix A, n_{col} columns at a time (segments $A_1, \ldots A_q$). The two most interesting operations that will determine the success of a multiprocessor architecture are the dot products of Eqs. 5 and 7.

9

algorithm is valid for this variation. Although the numbers of coincidences counted during an imaging experiment in a particular "tube" is Poisson distributed, the linear combination of tube values contained in a backprojection element will not be a Poisson variable any longer. What effect that has on the reconstructed images will have to be investigated.

CONCLUSION

The MLE algorithm has been receiving considerable attention in the emission tomography literature during the last few years. There seems to be a general agreement about the better signal-to-noise ratios that the method will give, when compared to filtered backprojection by computer simulation techniques. The correctness of the above expectation has recently been verified by Shepp et al[8] with true data from a positron ring. The method appears to remain, however, in the "wish list" of workers in ET because it is very cumbersome to use in a practical situation.

In this paper we have confirmed the results of Shepp et al by showing the superior images created by the MLE algorithm from data with a low number of counts and we propose that using probability functions that have been calculated accurately for a specific imaging device can give a more faithful reproduction of the object imaged. The preliminary results shown here are not sufficient to determine under which conditions it is of advantage to use true response functions, and further research is needed in that direction. The expectation is that, with the advent of high resolution PET instruments, the use of true response functions will become important, since effects of gamma penetration in neighbor detectors and efficiency effects are accounted for by the response functions.

Along with the general present day realization that large ET instruments require specialized computers for image acquisition and/or reconstruction, we show that it is possible to design a multiprocessor structure based on mass produced processors that can accommodate the MLE algorithm. The fundamental characteristics of such a system are 1) the use of large amounts of inexpensive memory for storage of the probability functions and intermediate results and 2) a configuration for that memory (fast switching) that allows a rotation of portions of those probability functions and results in such a fashion that memory arbitration becomes unnecessary. The result is a structure that allows for linear increase in computation speed with increased number of processors.

Work on the two main areas of the work presented here will continue with the aim of bringing the potential of the MLE algorithm to the practical world.

REFERENCES

1. Z. H. Cho, I. S. Ahn and C. M. Tsai, "Computer Algorithms and Detector Electronics for the Transmission X-Ray Tomography", IEEE Trans. Nucl. Sci., NS-21, No. 1, pp 218-227, 1974.

2. R. H. Huesman, G. T. Gullberg, W. L. Greenberg and T. M. Budinger, "Donner Algorithms for Reconstruction Tomography", Lawrence Berkeley Laboratory, Pub. 214, 1977.

3. A. Chatterjee, E. L Alpen, C. A. Tobias, J. Llacer and J. Alonso, "High Energy Beams of Radioactive Nuclei and Their Biomedical Applications", Int. Journal Roadiation Oncology Biol. Phys., Vol 7, pp 503-507, 1981.

4. J. Llacer, A. Chatterjee, E. K. Batho and J. A. Poskanzer, "Design Analysis and Performance Evaluation of a Two-Dimensional Camera for Accelerated Positron Emitter Beam injections by computer simulation", IEEE Trans Nucl. Sci, NS-30, No. 1, pp 617-625, 1983.

5. J. Llacer, A. Chatterjee, E L. Alpen, W. Saunders, S. Andreae and H.J. Jackson, "Imaging by Injection of Accelerated Radioactive Particle Beams", IEEE Trans. Medical Imaging, MI-3, No. 3, pp 80-90, 1984.

6. L. A. Shepp an 1 Y. Vardi, "Maximum Likelihood Reconstruction for Emission Tomography", IEEE Trans, Med. Imaging, MI-1, No. 2, pp 113-121, 1982.

7. D. G. Politte and D. L. Snyder, "Results of a Comparatived Study of a Reconstruction Proceaure for Producing Improved Estimates of Radioactivity Distributions in Time-of-Flight Emission Tomography", IEEE Trans. Nucl. Sci., NS-26, No. 1, pp 596-602, 1979.

8. L. A. Shepp, Y. Vardi, J. B. Ra, S. K. Hilal and Z. H. Cho, "Maximum Likelihood Pet with Real Data", IEEE Trans. Nucl. Sci., NS-31, No. 2, pp 910-913, 1984.

9. J. Llacer, "Theory of Imaging with a Very Limited Number of Projections", IEEE Trans. Nucl. Sci., NS-26, No. 1, pp 596-602, 1979.

10. J. Llacer, "Tomographic Image Reconstruction by Eigenvector Decomposition: its Limitations and Areas of Applicability", IEEE Trans. Medical Imaging, MI-1, No. 1, pp 34-42, 1982.

11. H. C. Andrews and B. R. Hunt, "Digital Image Restoration", Prentice-Hall, 1977.

12. J. Llacer, A. Chatterjee, B. Jackson, J. Lin and V. Zunzunequi, "An imaging Instrument for Positron Emitting Heavy Ion Beam Injection". IEEE Trans. Nucl. Sci., NS-26, No. 1, 1979.

13. M. E. Phelps, S. C. Huang, E. J. Hoffman et al., "An Analysis of Signal Amplification using Small Detectors in PCT", J. Comp. Assist. Tomogr., 6, pp 551-565, 1982.

14. E. J. Hoffman, A. R. Ricci, LM. A. M. Van der Stee and M. E. Phelps, "ECAT III, Basic Design Considerations", IEEE Trans. Nucl. Sci., NS-30, No. 1, 1983 pp 729-733.