LBL-25775
Preprint

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

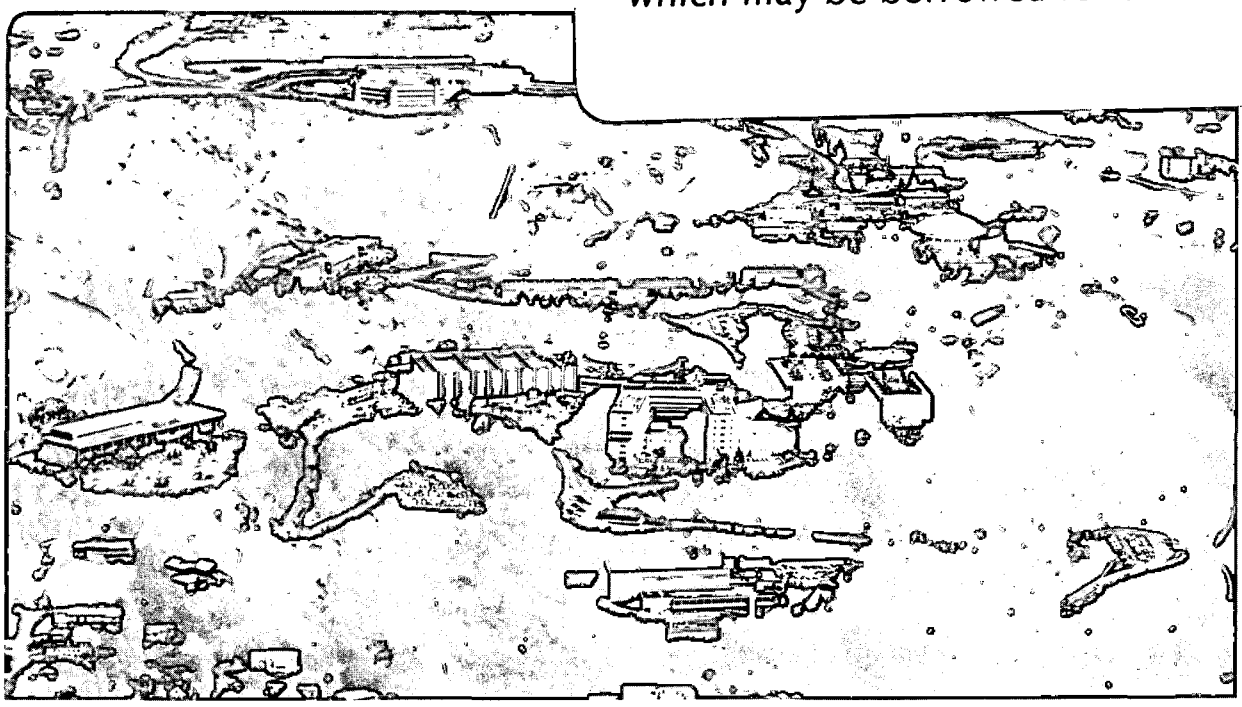
Engineering Division

To be published in IEEE Transactions on Medical Imaging,
June 1989

Feasible Images and Practical Stopping Rules for Iterative Algorithms in Emission Tomography

J. Llacer and E. Veklerov

August 1988



Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098.

DISCLAIMER

LBL-25775

Feasible Images and Practical Stopping Rules for Iterative Algorithms in Emission Tomography

Jorge Llacer and Eugene Veklerov
Engineering Division
Lawrence Berkeley Laboratory
1 Cyclotron Road
Berkeley, CA 94720

Abstract — **In this paper we continue the discussion of the causes for image deterioration in the Maximum Likelihood Estimator (MLE) method of tomographic image reconstruction that we initiated with the publication of a stopping rule for that iterative process. We introduce the concept of a feasible image, which is a result of a reconstruction that, if it were a radiation field, could have generated the initial projection data by the Poisson process that governs radioactive decay. From the premise that the result of a reconstruction should be feasible, we examine the shape and characteristics of the region of feasibility in projection space. Although MLE reconstructions from computer simulated data pass through a feasibility region when started from a uniform intensity image field, as determined by our previously published stopping rule, attempts at using that rule to detect feasibility in reconstructions with real PET data failed consistently. We examine the reasons for that failure and design a more relaxed stopping rule that takes into account the fact that the prob-**

ability matrix defining a true tomographic instrument can only be known within some error margin. With the new rule, reconstructions from real data can be tested for feasibility. Results of the test and reconstructed images for the Hoffman brain phantom are shown. We conclude with a comparative examination of the current methods of dealing with MLE image deterioration and we endeavor to put the minds of current workers in the field at ease about having to stop MLE iterations when the images look acceptable.

1 Introduction

Iterative methods, such as the Maximum Likelihood Estimator (MLE) method for image reconstruction in Emission Tomography, are an attractive alternative to more conventional filtered back-projection methods. However, several aspects of these methods are not fully understood and are actively discussed in the literature. In particular, several researchers have noticed that, as the MLE method passes a certain point, the reconstructions it yields begin to deteriorate. To counter this unfavorable feature, several remedies have been proposed. These remedies fall into three categories:

- those which modify the original criterion (likelihood);

- those which search for solutions among a set of 'smooth' images; and
- those which halt the iterative process before the deterioration occurs.

The last approach has been proposed by Veklerov and Llacer [1], Hebert, Leahy and Singh [2] and Defrise [3], who devised certain 'stopping rules'. They are derived from some fundamental principles which any reasonable reconstruction should satisfy, such as 'the image should not fit the data with an accuracy greater than the accuracy of the data' or 'the image should be capable of generating the data' in a statistical sense.

In this paper we define two versions of the concept of *feasibility* which we call weak and strong. This concept is closely related to the stopping rule described in Ref. 1. We continue by showing that successive images obtained by iterating with the MLE algorithm of Shepp and Vardi [4] pass through a region of feasibility, although the stopping rule developed in Ref. 1 may not be a good indicator of that condition when the data are from real PET scans.

Although we feel that the most important benefit of developing the concepts of feasible images and stopping rules may be in allowing an understanding of some essential characteristics of iterative image reconstruction, we feel it is useful to conclude this paper by

describing a relaxation of the strict demands of the stopping rule of Ref. 1 so that it can be applicable to the MLE or other iterative reconstruction methods with real emission tomography data.

On a historical note, similar stopping critera have been known in the case of additive noise. Trussell [5] proposed a stopping criterion based on the variance of the residual and, therefore, analogous to our weak feasibility for the Poisson case, as will be shown below. Trussell and Civanlar [6] discussed possible ways of strengthening the previous criterion by adding other constraints. Each constraint defines a set in a finite-dimensional space and the intersection of all sets corresponds to the feasible solutions. Numerical algorithms for finding such feasible solutions are based on the work by Sezan and Stark [7] which is further discussed by Trussell, Orun-Ozturk and Civanlar [8]. These works treat the subject from the viewpoint of a more general concept, namely projections onto convex sets. They show that sequential projections onto member sets converge to their intersection. Medoff, Brody, Nassi and Macovski [9] proposed an operator framework for image reconstruction algorithms using limited and missing data.

Situations in which the system's response function is not known exactly have also been extensively studied by researchers in the case of additive noise. Thus, Ward and Saleh [10] considered the

situiation where the response function, which correspond to our transition matrix, is itself random, in which case they proposed a deblurring technique.

The idea of feasibility was used in radio astronomy by Skilling and Bryan [11] in 1984. Several variations on the same theme in case of Gaussian data have been reported by other authors working in radio astronomy, see Ables [12], Gull and Daniels [13], Narayan and Nityananda [14]. The fact that not any weakly feasible image statistically fits the data was noted by Reiter and Pfleiderer [15].

2 The Concept of a Feasible Image

A feasible image is one that could have caused or produced the observed data by the known statistical process that governs the measurement. Let us use the standard notation introduced by Shepp and Vardi [4]:

$n^*(d), (d = 1, \ldots, D)$ – the projection data or the number of coincidences detected in tube d;

$\lambda(b), (b = 1, \ldots, B)$ – the emission density;

$p(b, d)$ – the transition matrix;

$\lambda^*(d) = \sum_{b=1}^{B} \lambda(b)p(b, d)$ – the means or forward projections;

where B and D are the number of pixels and the number of projections, respectively. We will consider specifically the case of emis-

sion tomography (PET or SPECT) in which disintegration data follow Poisson statistics.

Definition 1: The image $\lambda(1), \lambda(2), \ldots, \lambda(B)$ is said to be a feasible image with respect to data, $n^*(1), n^*(2), \ldots, n^*(D)$, if and only if the statistical hypothesis that $n^*(1), n^*(2), \ldots, n^*(D)$ are a Poisson sample with the means $\lambda^*(1), \lambda^*(2), \ldots, \lambda^*(D)$, respectively, can be accepted (not rejected) at a given significance level.

In order to test this hypothesis, we proposed a procedure in [1] which reduces the original problem to testing the hypothesis that a certain random variable is uniformly distributed between 0 and 1. The latter hypothesis can be tested following Pearson's procedure by computing the standard χ^2 statistic which we called the H parameter. The hypothesis is accepted when the H parameter dips below a certain value which depends on the significance level. The detailed procedure for obtaining the parameter H has been given in [1]. The main ideas are summarized here:

The first step of the test consists of scaling the differences $\lambda^*(d) - n^*(d)$ for each data pair to a new variable x which is uniformly distributed between 0 and 1 if $n^*(d)$ is a Poisson realization with mean $\lambda^*(d)$. Next, a histogram of x with N bins is generated. We conclude by testing the hypothesis that x is uniformly distributed between 0 and 1 by Pearson's procedure with $N - 1$

degrees of freedom. The histogram testing function H is defined as:

$$H = \sum_{i=1}^{N} \frac{(h_i - D/N)^2}{D/N}$$

where h_i is the observed frequency of bin i and D/N is the expected frequency if x is uniformly distributed.

Objections can be raised to using Pearson's test of goodness of fit as indicated above for testing the hypothesis of Definition 1 when the means $\lambda^*(d)$ are obtained from the data $n^*(d)$ by a reconstruction process. The objections have now been examined and our preliminary findings indicate that the contribution of the above is negligible in practice, although a theoretical justification still remains an open problem. The question will be treated in a forthcoming separate paper.

To visualize the concept of feasibility, let us imagine that we were able to recover an image for which all forward projections are exactly equal to the corresponding data. This 'ideal' image would yield *the* maximum likelihood. However, it would not be feasible, because, taken as a source, it would not be capable of generating the data from which it was reconstructed in the first place at a reasonable level of confidence. This follows from the fact that the Poisson assumption imposes a certain deviation between the data and the means.

Feasibility is a general concept in image reconstruction not limited to any specific algorithm. It allows us to formulate the following stopping rule for any iterative algorithm: stop the algorithm as soon as a feasible image has been recovered. It was shown in Ref. 1 that feasible images obtained with the MLE algorithm are visually clean, with a good compromise between sharpness and noise in uniform high intensity areas.

Since feasibility is one, though not necessarily the only, goal of image reconstruction, it is natural to ask these two questions: how large is the set of feasible images and where are they with respect to the maximum likelihood image? To do that, let us introduce another definition of feasibility by relaxing Definition 1. Since the new definition of feasibility requires less than Definition 1 does, it is appropriate to call this new form of feasibility 'weak feasibility' to distinguish it from 'strong feasibility' defined above.

Definition 2: The image $\lambda(1), \lambda(2), \ldots, \lambda(B)$ is said to be a weakly feasible image with respect to data, $n^*(1), n^*(2), \ldots, n^*(D)$, if and only if the second moments of $n^*(1), n^*(2), \ldots, n^*(D)$ are consistent with the Poisson hypothesis, namely:

$$\sum_{d=1}^{D} \frac{[n^*(d) - \lambda^*(d)]^2}{\lambda^*(d)} \approx D \tag{1}$$

Indeed, the expected value of the numerator of each term in Eq. (1) is the variance, while that of the denominator is the mean.

Therefore, Eq. (1) must be satisfied if the Poisson hypothesis holds. However, the reverse statement is not true, because the equality of the variance and the mean is necessary but not sufficient for a distribution to be Poisson or, put another way, Definition 2 is less demanding than Definition 1.

It is interesting to note that, since (1) has the form of a χ^2 function, it is tempting to state that we are carrying out a χ^2 test and that $(D - B - 1)$ should appear in the right hand side of (1). That is the number of degrees of freedom when the parameters $\lambda^*(d)$, projections of $\lambda(b)$, are estimated from $n^*(d)$ by allowing the MLE procedure to converge. Definition 2 does not, however, call for a χ^2 test. We stop the MLE procedure when (1) is satisfied for the stated reason that (1) must hold if the Poisson hypothesis is true.

Note that whereas Definition 1 allows for a certain margin of error for any given significance level, Definition 2 does not give a clue as to how much the left-hand side of (1) may differ from D in order for the image to remain feasible. We leave this statistical problem open at this time although we recognize that this is an important issue.

Let us consider the projection space and imagine that any set of projections $\lambda^*(d), d = 1, \ldots, D$ is "achievable"; that is, it can

be obtained from a distribution of intensities in the pixel space. If any point is achievable, then the maximum likelihood solution is such that $\lambda^*(d) = n^*(d)$ for all $d = 1, \ldots, D$.

We assert that all weakly (and, therefore, strongly) feasible images belong to an area resembling an ellipsoid and surrounding the maximum likelihood point. Indeed, if the letf-hand side of Eq. (1) must equal the right-hand side to the accuracy of ϵ, all points satisfying (1) form the space between two egg-shaped surfaces. They would be exact ellipsoids if the denominators were constants. They are not constants in the D-dimensional projection space of $\lambda^*(d), d = 1, \ldots, D$ and the set of points satisfying (1) will be referred to as a shell. The parameter ϵ reflects the expectation that the region of feasibility has a certain width. Figure 1, which is just a schematic 2-D illustration, shows a shell as well as lines along which likelihood is constant, which are regular ellipsoids for large values of data.

The shell defined by Eq. (1) contains all weakly feasible images. The strongly feasible images form a subset of the shell with gaps between them. The gaps are large, as can be deduced from the following reasoning: A point may satisfy (1) even if $n^*(d) > \lambda^*(d)$ for all $d = 1, \ldots, D$. But $n^*(d)$ are measurements and $\lambda^*(d)$ are their means; therefore some $n^*(d)$ should be greater than $\lambda^*(d)$,

while others should be less. The simple sign test, see e.g. Freund [16], shows that the hypothesis that $\lambda^*(d)$ are the means of $n^*(d)$ should be rejected if for $D = 1000$, for example, there are less than 450 measurements below their averages or if there are less than 450 measurements above their averages. Needless to say that not all points passing this test will be truly feasible but this test eliminates most of the points of the shell (1).

Let us now return to the assumption that all points in the projection space are achievable. If we consider that the matrix $p(b,d)$ maps a B-dimensional space of pixel intensities into a D-dimensional space of measurements, D being typically greater than B, it is clear that not all sets of projections $\lambda^*(d)$ are achievable from a distribution of intensities in the pixels. This fact further limits the region of feasibility (1) by placing holes into the remaining parts of the region. It should be pointed out, however, that the trajectory of likelihood in projection space obtained during the iterative reconstruction process, if it crosses the feasibility region, will escape those holes, since the result of the iterations is an image whose projections have to be in the solid part of the feasibility region.

Figure 1 shows hypothetical achievable regions and a trajectory of likelihood points for a reconstruction starting from the arbitrary

vector $\lambda^*(1) = \lambda^*(2) = 13$. The trajectory passes through the feasibility region and continues toward the maximum likelihood.

It is not *a priori* evident that the MLE solution invariably lies inside the feasibility shell. It was reported in Ref. 1 that the MLE solution did indeed lie inside the shell and this was confirmed in subsequent experiments, see e.g. Llacer and Veklerov [17], so long as we used computer simulated data starting from a uniform image. However, all attempts to obtain the same result while using data generated by a true PET instrument (ECAT-III of UCLA) failed to provide any reconstructed image that could be classified as feasible, which is another way of saying that the MLE solution appears to lie outside the feasibility shell. This observation was reported by us [18] and it led to further experiments. The results of these experiments helped us clarify the notion of feasibility and come up with a more practical stopping rule.

3 Analysis of the Problem

In simulation experiments, the same transition matrix $p(b, d)$ is used both in the generation and in reconstruction phases. Hence, the only error handled by the reconstruction algorithm is the one caused by statistical fluctuations of the emission and detection processes. This is not the case in real tomography, where the

two transition matrices can be expected to be different. The data generation matrix, reflecting the true characteristics of the tomography instrument is not known precisely. In addition, it is usually necessary to correct the emission data with normalization factors to compensate for variations in detector gains, which are inherently unstable, and for absorption in the patient.

Applying the corrections either to the transition matrix or to the data leads to the same results in linear reconstruction schemes. In a non-linear scheme, as in the MLE method, one can expect different images for the two correction methods, as we indeed find, except for noiseless data. The correct approach, which we have used in all the reconstruction experiments reported in this paper, is to apply the corrections to the matrix terms since this is consistent with the definition of the matrix elements $p(b, d)$. Since the correction coefficients can be any given numbers, the symmetry of the transition matrix [19, 20] which allows reduced storage space is lost and larger storage space and increased computation time have become necessary for the computations reported.

Two series of experiments are described in this section. These experiments are based on computer simulated source images which allowed us a careful measurement of the effect of some parameters on the final results, as described below. Reconstructions with real

data are presented in a later section. The first series demonstrates that, if the transition matrix used in the reconstruction process is not accurate enough, the process appears to bypass the feasibility region. The second series allows us to gain an insight into this situation and suggests a solution. The source image that was used in both series was a simulated brain phantom described in Ref. 11, with 1 million counts total, shown in Fig. 2.

1st Series of Experiments. This series consisted of a generation of projection data and subsequent reconstruction using two slightly different matrices, $p_1(b,d)$ and $p_2(b,d)$, derived as follows. First, a transition matrix for the 512 crystal geometry of the ECAT-III tomograph was obtained by the Shepp-Vardi prescription [4]. Then we assumed that the individual detectors had gains uniformly distributed in the range (0.5, 2.0). The tube gains were obtained as products of the corresponding detector gains and incorporated into the transition matrix. The result was matrix $p_1(b,d)$ which was used to generate all the source projection data and to reconstruct the 'perfect' case (case A). Matrix $p_2(b,d)$ was obtained by simulating small drifts in the individual detector gains with respect to the initial gains. Three ranges of random multiplicative drifts were studied: (0.95, 1.05) (case B), (0.93, 1.07) (case C) and (0.90, 1.10) (case D), corresponding to rms drift val-

ues of 2.88, 4.04 and 5.77 %, respectively.

The parameter H quantifying the degree of feasibility in these three cases and in the 'perfect' case is shown in Fig. 3 as a function of the iteration number. The expected pattern whereby H reaches the feasibility region and then leaves it is exhibited only in case A. The two horizontal lines represent the 90 and 99% confidence levels of an image being feasible, according to our test. In case B, the parameter H reaches the feasibility region, although somewhat later than in case A, and then remains there. In cases C and D, H remains at higher levels, which is the pattern observed in experiments with real PET data in Ref. 12.

2nd Series of Experiments. The purpose of this series was to find out whether the images generated in cases C and D are indeed not feasible or some of them are feasible but this fact is obscured by the matrix $p_2(b,d)$ providing erroneous forward projections which, in turn, determine the value of H. To this end, the reconstruction algorithm using the matrix $p_2(b,d)$ was stopped at iteration 25, at which point the current images were projected with the correct matrix $p_1(b,d)$ and the parameters H were then immediately calculated. The low values of H obtained demonstrated that the images were actually feasible, but that this fact was obscured by having used the distorted matrix $p_2(b,d)$ for the

projection needed for calculating H. The fact that the whole reconstruction was being carried out with the distorted matrix did not make the image unfeasible. Alternatively, images classified by matrix $p_2(b,d)$ as feasible in case B as late as iteration 85, were reclassified by matrix $p_1(b,d)$ as unfeasible.

The second part of this series was the reverse of the previous. Data generated by matrix p_1 were initially reconstructed by p_1 and, after iteration 25, matrix p_2 was switched in. The value of H immediately jumped above the feasible region.

All these results may be summarized as follows: If the transition matrix contains relatively small errors, the MLE algorithm will still produce feasible images, although this fact is not detected by the stopping rule formulated in Ref. 1, unless the forward projection is carried out with the exact matrix which is not available in practice. The images at iteration 25 showed slight differences among themselves, depending on the matrix used for the reconstruction. These differences were, however, very slight and no systematic effects could be discerned.

4 A robust stopping rule

The previous experiments demonstrate that the stopping rule is too sensitive to small inaccuracies in the transition matrix and suggest

a need for a more robust rule. The stopping rule described below is a natural generalization of the one in Ref. 1. It is motivated by the fact that, in practice, the transition matrix can only be known approximately. In effect, the new robust stopping rule is equivalent to increasing the feasibility region.

Let us associate an interval:

$$[\lambda^*(d)(1-\epsilon(d)), \lambda^*(d)(1+\epsilon(d))]$$

with each $\lambda^*(d)$, and suppose that the true, but unknown, value of the forward projection, say $\mu^*(d)$, lies somewhere within this interval. The following statistical hypothesis will be tested: there exists a set $\mu^*(d), (d = 1, \ldots, D)$, such that $n^*(d), (d = 1, \ldots, D)$ could be a Poisson sample with the means $\mu^*(d)$. In other words, we will require that the data could have been generated not by the available forward projections as means of Poisson distributions but by some other projections lying in their vicinity and the test will prove or disprove their existence. The vicinity is defined by the value of $\epsilon(d)$ which quantifies the accuracy of the transition matrix. The question of selection of $\epsilon(d)$ will be discussed in the next section.

The implementation of the test of this statistical hypothesis is straightforward and, for brevity, we will assume that the reader is familiar with the implementation of the test in Ref. 1. The

procedure described in Ref. 1 generates a histogram of $n^*(d)$ appropriately scaled by $\lambda^*(d)$. The procedure was such that if the distribution of $n^*(d)$ is indeed Poisson with the means $\lambda^*(d)$, the histogram will contain almost the same number of tubes d in each class and the exact meaning of the phrase 'almost the same number' was defined in Ref. 1 in terms of the χ^2 distribution.

In our case, the exact values of the means are unknown. It is known that if the hypothesis is true, they lie somewhere between $\lambda^*(d)(1-\epsilon(d))$ and $\lambda^*(d)(1+\epsilon(d))$. Therefore, we are unable to assign each tube d to a specific class of the histogram. Rather, we can only determine the lowest and the highest classes of the histogram, $h_1(d)$ and $h_2(d)$, respectively, corresponding to $\lambda^*(d)(1+\epsilon(d))$ and $\lambda^*(d)(1-\epsilon(d))$, respectively, while the true, but unknown, class lies somewhere between $h_1(d)$ and $h_2(d)$.

Thus, we associate a pair of numbers, $h_1(d)$ and $h_2(d)$, with each tube d. Then, instead of generating the histogram used in Ref. 1, we have to generate a two dimensional histogram, $m(i,j)$, as follows: if the lowest and the highest classes associated with a tube are h_1 and h_2, respectively, then we will increment the element $m(h_1,h_2)$ by 1. It is easy to see that all elements of the resulting histogram $m(i,j)$ below its diagonal will be zero.

The two dimensional histogram $m(i,j)$ contains more uncer-

tainty than the one dimensional histogram used in Ref. 1, because any tube assigned, for example, to the two dimensional class (2, 5) can actually belong to any one dimensional class from 2 through 5. Our goal is to find out whether among all possible ways of resolving this uncertainty there exists one leading to a one dimensional histogram consistent with the Poisson hypothesis. In other words, if each tube can be assigned to a specific class between its minimum and maximum classes in such a way that each class gets almost the same number of tubes, then the hypothesis should be accepted.

Hence, the problem at hand reduces to the following problem. A two dimensional triangular matrix N by N is given and the value of any of its non-diagonal elements $m(i,j)(i<j)$ can be distributed arbitrarily among the classes of a one dimensional histogram from i through j by splitting into $x_i \geq 0$, $x_{i+1} \geq 0, \ldots, x_j \geq 0$, such that $\sum_{k=i}^{j} x_k = m(i,j)$. We try to distribute them in such a way that the total number received by each i is as close to

$$a = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} m(i,j)$$

as possible and then find out whether the resulting one dimensional distribution can pass the test described in Ref. 1.

This problem has an efficient solution. Consider the following algorithm.

1. begin with row i of matrix $m(i,j)$ equal to 1

   (a) begin with column j of matrix $m(i,j)$ equal to i

   (b) if the element $m(i,j)$ is greater than a, proceed to Step 2; otherwise proceed to Step (c)

   (c) increment j by 1

   (d) check whether transferring the entire amount $m(i,j)$ to $m(i,i)$ will increase $m(i,i)$ to or above a:

   if not, transfer the entire amount, and proceed back to Step (c), unless j already equals N in which case go to Step 2;

   if yes, transfer only as much as necessary to bring $m(i,i)$ up to a

2. row i has been resolved; if $i = N$, stop;

   if not, transfer each remaining nonzero nondiagonal elements of row i to row $i+1$ within the same column

3. increment i by 1 and proceed back to Step 1.

The algorithm resolves the uncertainty of the two dimensional histogram by distributing its elements as equitably as possible. Then we can use the same test as the one described in Ref. 1 to see if the resulting one dimensional histogram is flat enough, in

which case the hypothesis must be accepted; otherwise it is rejected. As we pointed out in Section 2, this is a limited test based on the simplifying assumption that the image and the data are statistically independent. The algorithm does not add a measurable amount of time to the reconstruction algorithm.

5 Experiments with PET data

Data from the UCLA ECAT-III tomograph for the Hoffman brain phantom described in Ref. 12 have been used to test the proposed 'practical' stopping rule. Data sets totaling 1 million (1M), 3 million (3M), 12 million (12M) and 55 million (55M) counts per image were obtained. Data corresponding to real coincidence counts plus prompt random coincidences were obtained in files separate from the delayed random coincidences data [21]. The delayed coincidences were not subtracted dynamically from the data stream, in order to avoid using a non-Poisson data set. A normalization file of detector–pair gains as well as transmission data files were also available. Figure 4 shows the brain phantom, in which the ratio of activity between the black vs. white areas is approximately 4:1.

The normalization and transmission corrections were incorporated into the transition matrix. All reconstructions were carried out with the Shepp-Vardi prescription for the transition matrix [4],

with an image plane of 128 x 128 pixels of 2.03 mm sides.

An important question was the selection of the values for $\epsilon(d)$ which quantify the inaccuracy of the transition matrix. The experiments showed reasonable results for 1M count sets if $\epsilon(d)$ was around 0.065.

However, the data sets with more counts required a slightly larger value for ϵ. This fact follows from a peculiarity of the Poisson distribution whereby the larger its mean, the smaller its standard deviation in relative terms. Therefore, if ϵ is 10% of the mean and the mean is 100, for example, then the relaxed interval for the mean is (90, 110). However, if the real value of n^* should be 115, it will still be accepted, because it is consistent with the mean of 110 and a standard deviation of 10.49. If the mean is 10000, the relaxed interval for the mean is (9000, 11000), but the effective interval or the interval of acceptable n^* will be narrower in relative terms. If the real value of n^* should be 11500, it would not be likely to be consistent with a mean of 11000 and a standard deviation of 104.88. In other words, in order to achieve the same effective interval, we have to relax the means more when they are high.

The contribution of this effect is secondary compared to the main effect caused by the error in the transition matrix, but it should be taken into account when the total number of counts in

data sets varies within a wide range. We found that the following empirical values for ϵ result in reasonable stopping rule: 0.108 for 3M counts, 0.171 for 12M counts, 0.234 for 55M counts. This is admittedly an ad hoc procedure and we hope other researchers can find a more elegant solution.

Figure 5 depicts the behavior of the H function as a function of the iteration number. As could be expected, it takes more iterations to trigger the stopping rule for data sets with more counts: approximately 29 iterations for a 1M set, 39 iterations for a 3M set, 55 for a 12M set and 73 for a 55M set.

In Fig. 6 we show images resulting from the reconstructions with the Hoffman brain phantom. The four rows correspond to the 1M, 3M, 12M and 55M count images, respectively, while the first column shows an image at an iteration number substantially before the feasibility region, the second column shows the images at the approximate entrance to the region and in the third column we show iteration 100 which is past the entrance to that region in all cases. The images in the center column have good sharpness and the image deterioration process has not yet set in.

6 Conclusions

At the beginning of the paper we have indicated that there are three different approaches to remedy the observed deterioration of MLE images at high number of iterations:

- The likelihood criterion can be modified. This should be a very promising avenue that is not yet very well explored. It is principally based on Bayesian reconstruction techniques in which the function to be maximized is the product of the likelihood with an a priori distribution. Geman and McClure [22] have been working on this idea with interesting results using a Gibbs energy prior distribution. Skilling and Gull [23] are using an entropy prior distribution with apparent success in a variety of image restoration and reconstruction modalities.
- Searching for solutions among a set of 'smooth' images. The principal exponent of this approach is the method of sieves of Snyder and Miller [24].
- Halting the iterative process before the deterioration occurs.

With our initial work on the stopping rule for MLE method [1] and the description of the feasibility region in the present paper we have shown unequivocally that the MLE method of reconstruction, left alone, will lead to images that are not consistent with

the physical process that generated the data. The reason for this effect is most easily grasped if one realizes that maximizing the likelihood of the data given an image (the reconstructed one) is only one part of the Bayesian reconstruction criterion that requires the maximization of the probability of an image given the data. This idea has been the driving force behind the work of the groups working on Bayesian reconstruction methods. The search for solutions among a set of smooth images by the method of sieves can be successful in practice and is known to converge to a unique image. One objection can be raised: there is a parameter that determines the degree of smoothness accepted and the wrong choice of that parameter can lead to images that are inconsistent with the data, i.e., that are unfeasible.

What our work proposes is a method to obtain useful images from the MLE alone by the use of the prior knowledge that the image at some point in the iteration procedure should be feasible, i.e., it is physically consistent with the initial data. This prior knowledge cannot readily be cast in the form of a prior distribution for a Bayesian reconstruction. For that reason we test for feasibility after each iteration of the unmodified MLE.

The prospect of stopping a maximization problem with a unique maximum before that maximum has been reached can certainly

raise some objections. The most important argument against such procedure is that the resulting image could then depend on the initial value of the image field. We do understand that objection and we claim, at this time, only that starting from a uniform image field, the MLE passes through a region of feasibility with useful images and that the passage through the region of feasibility can be tested.

We feel that as a result of our work, medical research groups that are using the MLE method for image reconstruction and are uneasy about having to stop the process after a number of iterations can feel at ease, whether they implement the stopping rule or not. They may already be choosing to stop at a point where the images are feasible, i.e., consistent with the data. Whether these images are the 'best' that can be obtained from the available data remains to be established and we are continuing work on the subject.

Acknowledgments

The authors are grateful to an anonymous referee for his or her valuable comments.

This work has been supported by a grant from the National Cancer Institute (CA-39501) and the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

References

1. E. Veklerov, J. Llacer, "Stopping rule for the MLE algorithm based on statistical hypothesis testing," *IEEE Trans. on Med. Imaging*, vol. MI-6, pp. 313 – 319, 1987.

2. T. Hebert, R. Leahy, M. Singh, "Fast MLE for SPECT using an intermediate polar representation and a stopping criterion," *IEEE Trans. Nucl. Sci.*, Feb., 1988.

3. M. Defrise, "Possible criteria for choosing the number of iterations in some iterative reconstruction methods," Proc. NATO Advanced Summer Inst. on Mathematics and Computer Science in Med. Imag., El Ciocco, Italy, Springer-Verlag, 1986.

4. L. A. Shepp, Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imaging*, vol. MI-1, pp. 113 – 122, 1982.

5. H. J. Trussell, "Convergence criteria for iterative resoration methods," *IEEE Trans. ASSP*, Feb., 1983.

6. H. J. Trussell, M. R. Civanlar, "The feasible solution in signal restoration," *IEEE Trans. ASSP*, Apr., 1984.

7. M. I. Sezan, H. Stark, "Image restoration by the method of convex projections: part 2 — Applications and numerical results," *IEEE Trans. Med. Imaging*, vol. MI-1, Oct., 1982.

8. H. J. Trussell, H. Orun-Ozturk, M. R. Civanlar, "Errors in back-projection methods in computerized tomography," *IEEE Trans. Med. Imaging*, vol. MI-6, Sep., 1987.

9. B. P. Medoff, W. R. Brody, M. Nassi, A. Macovski, "Itera-

tive convolution backprojection algorithms for image reconstruction from limited data," *J. Opt. Soc. Am*, Nov., 1983.

10. R. K. Ward, B. E. A. Saleh, "Deblurring random blur," *IEEE Trans. ASSP*, Oct., 1987.

11. J. Skilling, R. K. Bryan, "Maximum entropy image reconstruction: general algorithm," *Mon. Not. R. Astr. Soc*, 211, pp. 111 – 124, 1984.

12. J. G. Ables, "Maximum entropy spectral analysis," *Astron. Astrophys. Suppl.*, 15, pp. 383 –393, 1974.

13. S. F. Gull, G. J. Daniell, "Image reconstruction from incomplete and noisy data," *Nature*, 272, pp. 686 – 690, 1978.

14. R. Narayan, R. Nityananda, "Maximum entropy image restoration in astronomy," *Ann. Rev. Astron. Astrophys.*, 24, pp. 127 – 170, 1986.

15. J. Reiter, J. Pfleiderer, "Improvement of MEM-deconvolution by an additional constraint," *Astron. Astrophys*, 166, pp. 381 – 392, 1986.

16. J. Freund, Mathematical Statistics, Prentice-Hall, 1971

17. J. Llacer, E. Veklerov, The maximum likelihood estimator method of image reconstruction: its fundamental characteristics and their origin. Proc. of the Xth Information Processing in Medical Imaging (IMPI) International Conference, Utrecht, the Netherlands, 1987.

18. E. Veklerov, J. Llacer, E. Hoffman, "MLE reconstruction of a brain phantom using a Monte Carlo transition matrix and a statistical stopping rule," *IEEE Trans. Nucl. Sci.*, Feb., 1988.

19. J. Llacer, S. Andreae, E. Veklerov, E. Hoffman, "Towards a practical implementation of the MLE algorithm for positron emission tomography," *IEEE Trans. Nucl. Sci.*, Feb., 1986.

20. L. Kaufman, "Implementing and accelerating the EM algorithm for positron emission tomography," *IEEE Trans. Med. Imaging*, vol. MI-6, Mar., 1987.

21. C. W. Williams, M. C. Crabtree and S. G. Burgiss, "Design and performance characteristics of a position emission computed axial tomograph – ECAT-II," *IEEE Trans. Nucl. Sci.*, NS–26,No. 1, 619–627 (1979)

22. S. Geman and D.E. McClure, "Statistical methods for tomographic image reconstruction", Proceedings of the 46th Session of the ISI, Bulletin of the ISI, vol 52, 1987.

23. J. Skilling and S.F. Gull, to be published in Proceedings of the AMS-IMS-SIAM Joint Summer Conference of Spatial Statistics and Imaging, June 1988.

24. D.L. Snyder and M.I. Miller, "The use of sieves to stabilize images produced with the EM algorithm for emission tomography", *IEEE Trans. Nucl. Sci.*, NS-32, No. 5, 1985.

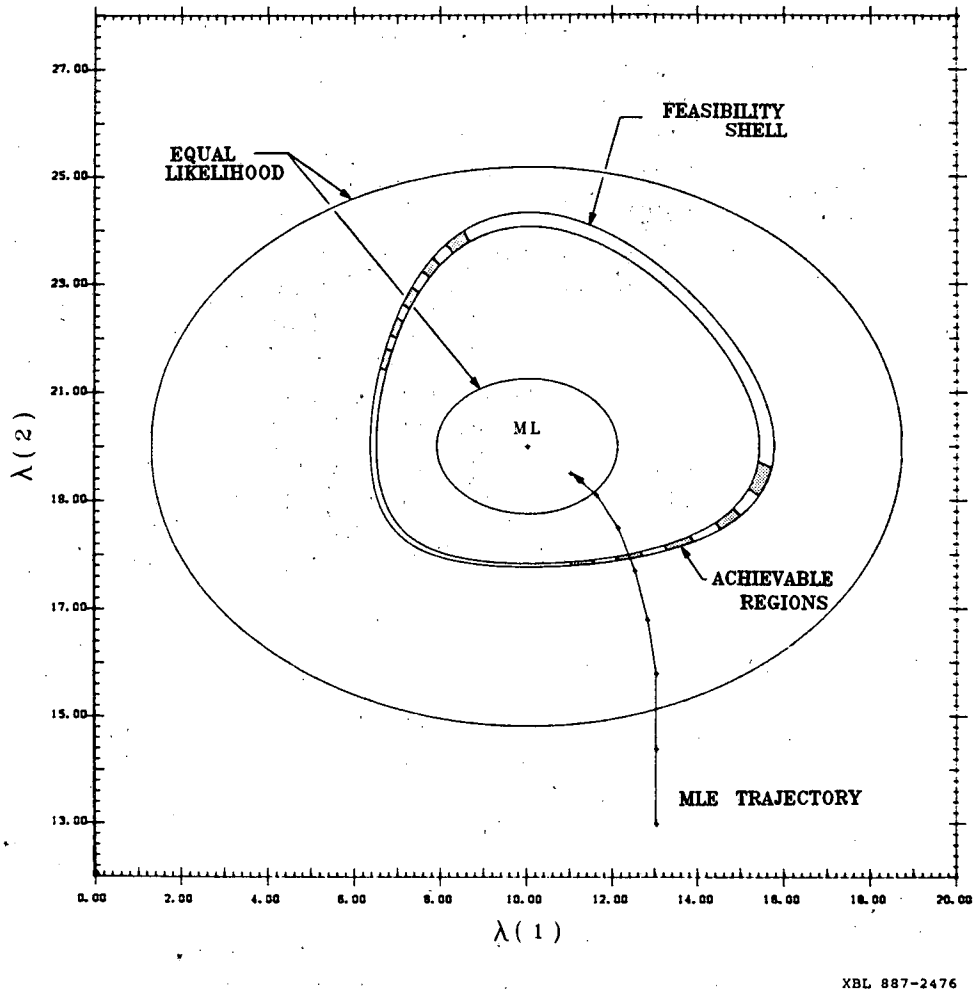
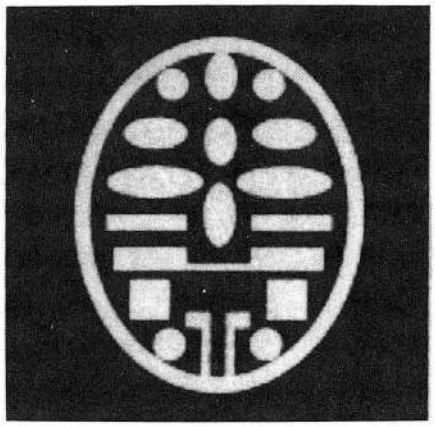

FEASIBILITY SHELL
EQUAL LIKELIHOOD
ML
ACHIEVABLE REGIONS
MLE TRAJECTORY
λ(2)
λ(1)
27.00
25.00
23.00
21.00
18.00
17.00
15.00
13.00
0.00
2.00
4.00
6.00
8.00
10.00
12.00
14.00
16.00
18.00
20.00
XBL 887-2476

Figure 1



XBB 888-7769

Figure 2

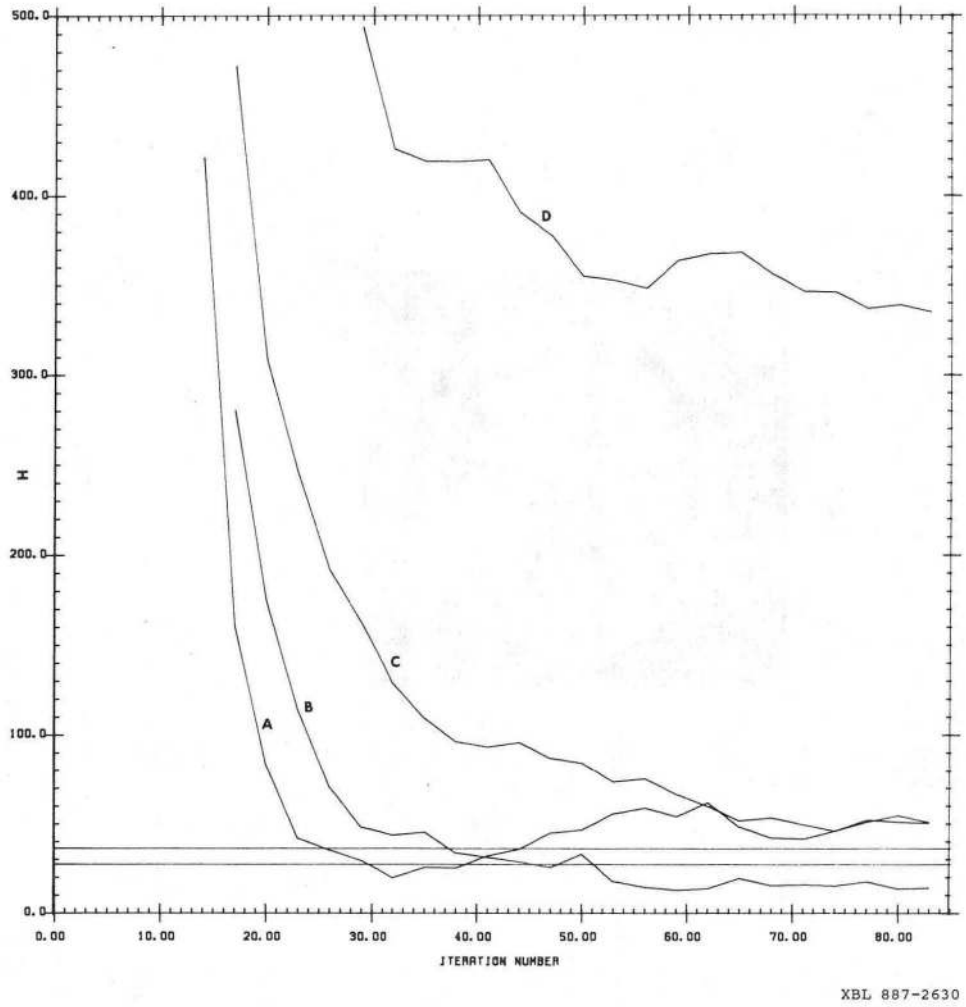

500.0
400.0
300.0
200.0
100.0
0.0
H
A
B
C
D
0.00
10.00
20.00
30.00
40.00
50.00
60.00
70.00
80.00
ITERATION NUMBER
XBL 887-2630

Figure 3

XBL 888-2850

Figure 4

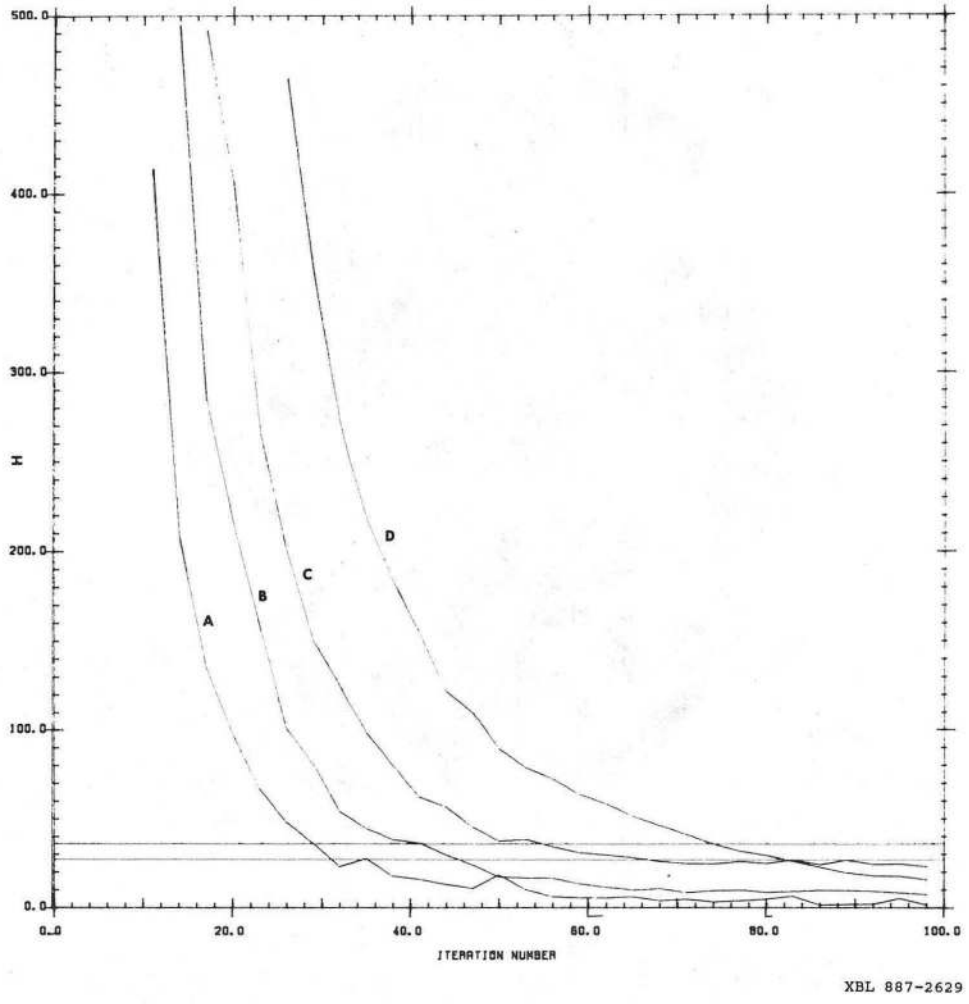

500.0
400.0
300.0
200.0
100.0
0.0
H
A
B
C
D
0.0
20.0
40.0
60.0
80.0
100.0
ITERATION NUMBER
XBL 887-2629

Figure 5

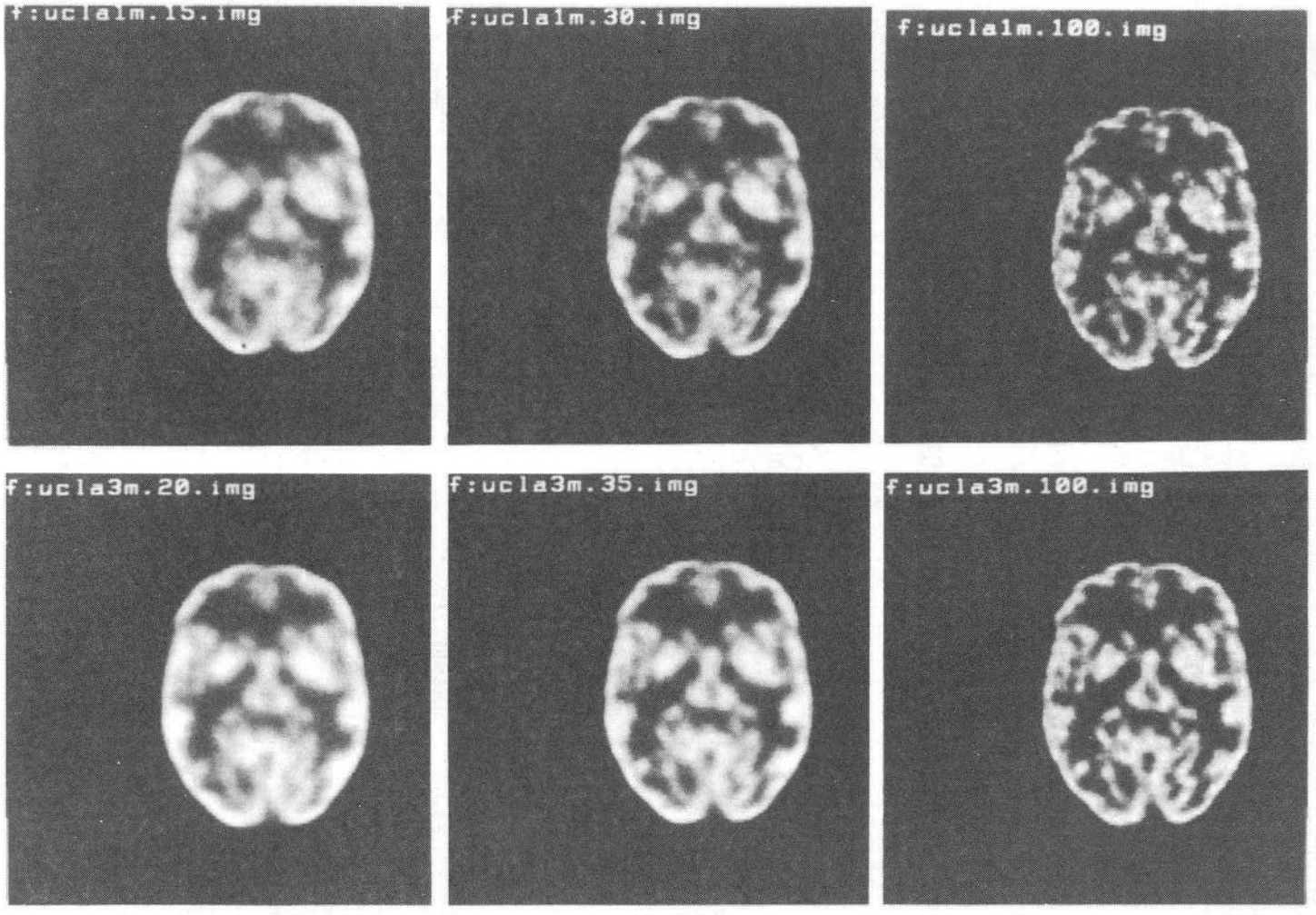

f:ucla1m.15.img
f:ucla1m.30.img
f:ucla1m.100.img
f:ucla3m.20.img
f:ucla3m.35.img
f:ucla3m.100.img

Figure 6

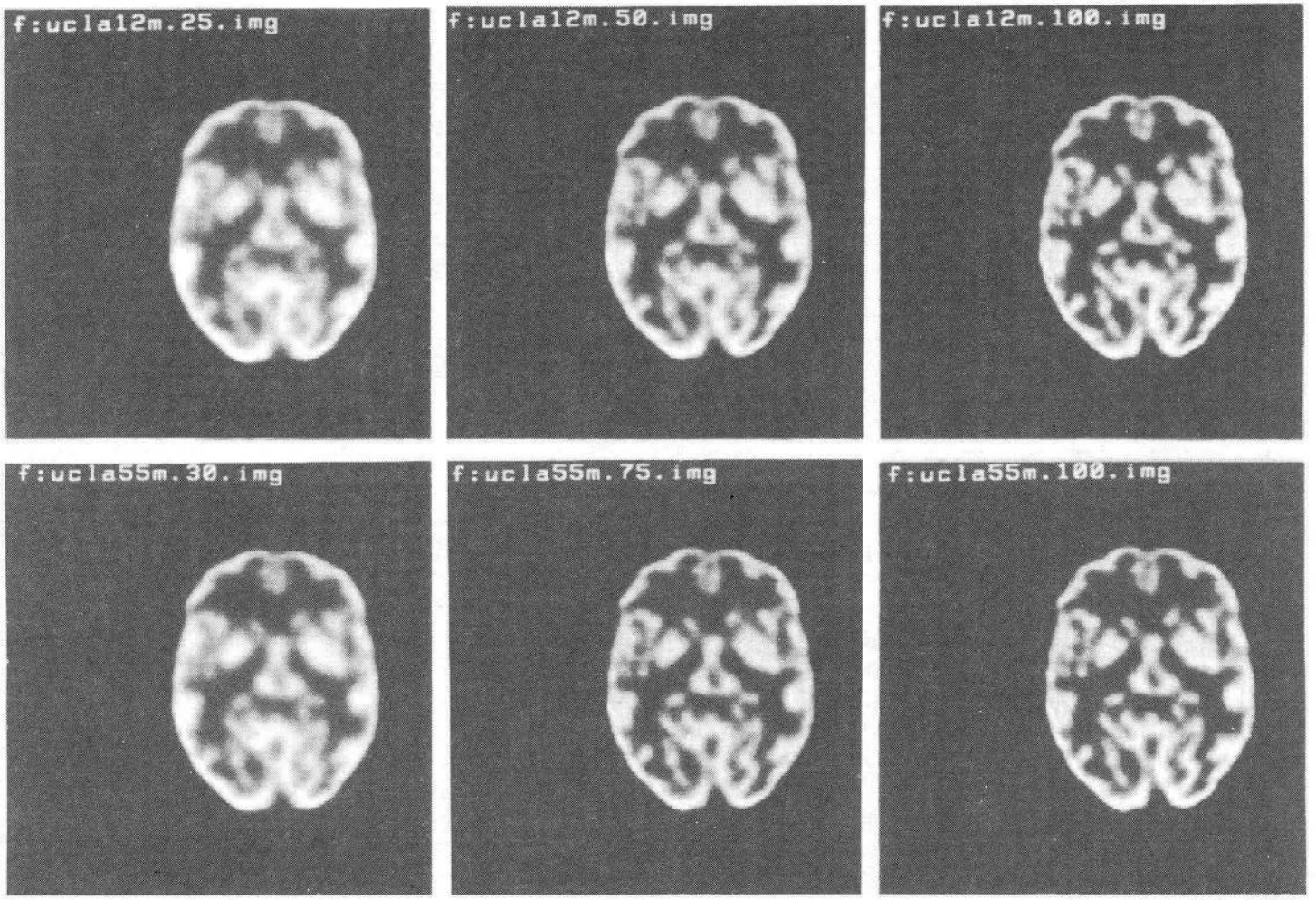


Figure 6 cont.

LAWRENCE BERKELEY LABORATORY
TECHNICAL INFORMATION DEPARTMENT
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720