LBL-25451

JUL 2 1 1989 CONF-890115--2

Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

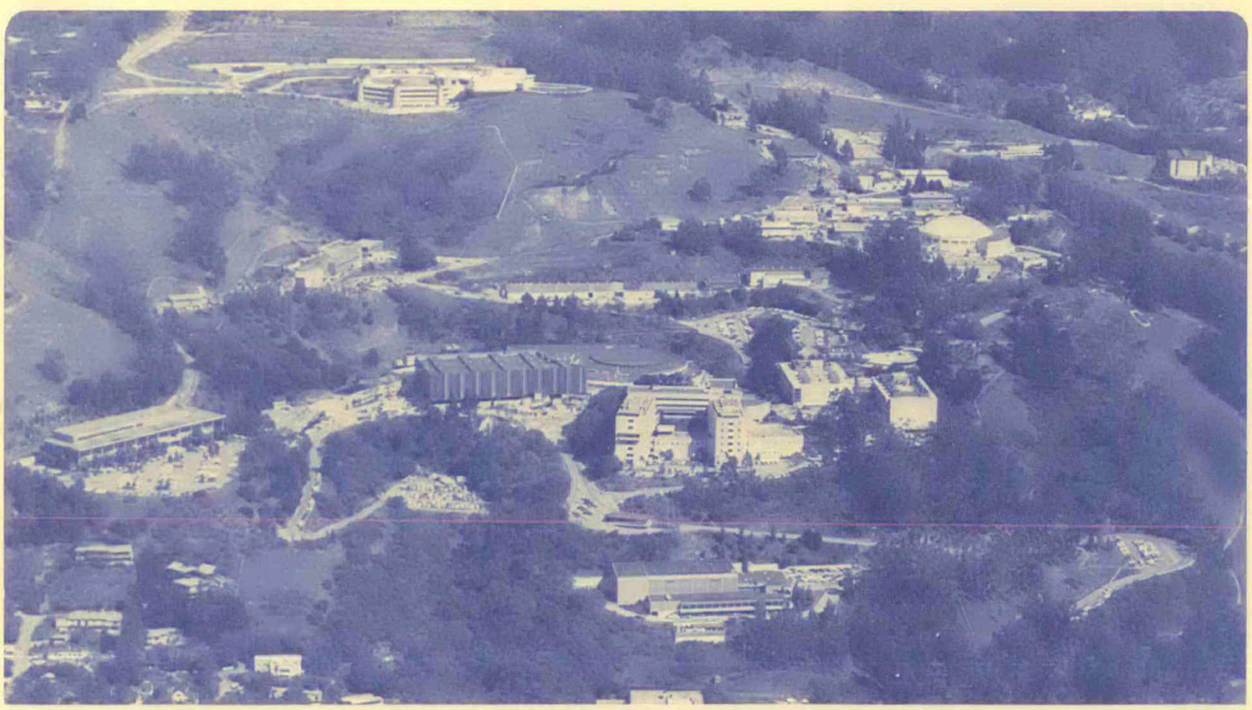
Engineering Division

Presented at SPIE Medical Imaging III Conference,
Newport Beach, CA, January 30, 1989

Maximum Entropy and the Concept of Feasility in Tomographic Image Reconstruction

J. Nunez and J. Llacer

January 1989



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

LBL-25451

Maximum entropy and the concept of feasibility in tomographic image reconstruction

Jorge Nunez* and Jorge Llacer
Engineering Division,
Lawrence Berkeley Laboratory,
1 Cyclotron Road, Berkeley, California 94720

LBL--25451

DE89 014886

ABSTRACT

Feasible images in tomographic image reconstruction are defined as those images compatible with the data by consideration of the statistical process that governs the physics of the problem. The first part of this paper reviews the concept of image feasibility, discusses its theoretical problems and practical advantages, and presents an assumption justifying the method and some preliminary results supporting it. In the second part of the paper two different algorithms for tomographic image reconstruction are developed. The first is a Maximum Entropy algorithm and the second is a full Bayesian algorithm. Both algorithms are tested for feasibility of the resulting images and we show that the Bayesian method yields feasible reconstructions in Positron Emission Tomography.

1. INTRODUCTION

Over the last two years we have been involved in a detailed study of the behavior of solutions to the image reconstruction problem in emission tomography (ET), with data generated both by computer simulation and by the ECAT-III tomograph at UCLA. In the process of studying the behavior of the Maximum Likelihood Estimator (MLE) method of image reconstruction at our Laboratory, we became aware, along with other workers, of the need to either stop the iterative procedure before the images deteriorate[1], or to choose the solution from a set of smooth images.[2]

As we progressed into the study of this phenomenon[3], we became aware of the fact that the observed deterioration is a direct consequence of the MLE criterion itself. By definition, the Maximum Likelihood method tries to find an image that would give the experimental data with the highest probability. In other words, the Maximum Likelihood criterion favors images whose forward projections are as close to the data as possible. This fact is undesirable when working with noisy data. If the reconstruction is forced to fit the data too closely, it will include features due exclusively to the noise. On the other hand, if the reconstruction fits the data too loosely, it would be a poor representation of the object.

In the case of Positron Emission Tomography (PET), it was found[4] that a true radioactive source equal to the reconstructed MLE image at convergence could not have yielded the experimental data by a Poisson disintegration

* On leave from the Department of Physics, University of Barcelona, Spain.

process. They observed, however that the images recovered after iterations within a certain range could have generated the data. That initial finding led them to formulate a stopping rule for the MLE algorithm. Upon further investigation, Llacer and Veklerov[5] realized that the stopping rule is actually based on the concept of Image Feasibility (called Causality in that paper) as discussed below.

The first part of this paper is concerned with reviewing the concept of feasibility in some detail, its application to testing the reconstructions and discussing the difficulties in carrying out a correct test for feasibility within the framework of hypothesis testing. The second and third part of the paper describe the development of two independent algorithms for image reconstruction in Emission Tomography. The first algorithm is a Maximum Entropy algorithm and the second is a full Bayesian reconstruction algorithm. Finally, the results of both algorithms are tested using the feasibility concept.

2. THE CONCEPT OF FEASIBILITY OF A RECONSTRUCTION

The concept of feasibility, as presented here, has been developed by our group for the Poisson data case as a result of the study of the image deterioration during MLE reconstruction.[4-6] Several variations of the same concept for Gaussian noise have been reported by other authors working mainly in astronomy.[7-9]

Throughout this paper we shall use the following notation:

p_j $j=1,\ldots,D$ - the projection data or the number of coincidences detected in detector pair (or tube) j;

a_i $i=1,\ldots B$ - the emission density;

f_{ji} - the point spread (transition) matrix;

$h_j = \sum_{i=1}^{B} f_{ji}\, a_i$ - the forward projection;

where B and D are the number of pixels and the number of projections, respectively. We shall consider specifically the case of emission tomography in which disintegration data follow Poisson statistics.

The basic idea of the feasibility concept is that the residuals should be distributed in accordance with the Poisson nature of the process. Two definitions have been presented:[6]

Definition 1: The reconstruction a_i $i=1,\ldots B$ is said to be strongly feasible with respect to data p_j $j=1,\ldots D$, if and only if we can accept (not reject) the statistical hypothesis that p_j $j=1,\ldots D$ is a Poisson sample with respect to the projections h_j $j=1,\ldots D$.

Definition 1 is difficult to implement as a constraint in an optimization problem. For that reason we have introduced another definition:

Definition 2: The reconstruction a_i $i=1,....B$ is said to be k-feasible with respect to data p_j $j=1,....D$, if and only if the first k moments of p_j $j=1,....D$ are consistent with the Poisson hypothesis, namely:

$$\sum_{j=1}^{D} \frac{(p_j - h_j)^n}{f_n(h_j)} \simeq D \qquad n = 1,....k \tag{1}$$

where $f_n(x)$; $n=1,...k$ is the expression for the n-th moment of the Poisson distribution with mean x.

Note that for the case of n=2 (only the second moment is considered) the left side of (1) is identical to the expression of the chi-square (χ^2) statistic and, thus, Eq. (1) is reduced to $\chi^2 = D$. This equation is precisely the constraint used by several authors in the maximum entropy method for the Gaussian noise case. The case of k=2 has been called weak feasibility.[6] Note also that if $k \to \infty$ a k-feasible reconstruction becomes strongly feasible.

3. FEASIBILITY TESTS

Several tests can be derived from the above definitions. The easiest one is to test for weak feasibility. It is important to note that this test is not a chi-square test of significance but a constraint that should be satisfied. A problem arises when we want to determine an upper and a lower limit to the constraint, as will be discussed in detail below. A test for strong feasibility has been developed[4] by means of an algorithm that, instead of computing the different moments of the distribution, analyzes the shape of the histogram of the residuals. The histogram is defined in Ref. 4 with N bins of equal probability for a Poisson process and the algorithm computes a function called H that is used to test the goodness of fit of the residuals to the Poisson hypothesis. A critical discussion of its validity is also given in Llacer, Veklerov and Nunez.[6] Let us outline the procedure for testing the feasibility.

We want to test the joint hypothesis that $p_j(1)$ is a realization drawn from a Poisson distribution with mean $h_j(1)$, $p_j(2)$ is a realization of $h_j(2)$, etc. A common method to test the hypothesis is the Pearson's goodness of fit test. The method consists of defining N mutually exclusive classes for the residuals, assigning each residual to a class and then comparing the observed number of residuals assigned to each class with the expected numbers determined by the Poisson hypothesis. The discrepancy between the two is measured by the chi-square statistic. If the chi-square statistic exceeds a certain threshold, the hypothesis is rejected. This chi-square test should not be confused with the weak feasibility test discussed above. In the weak feasibility test we are comparing the left side of Eq. (1) (for n=2) with its desired value (the number of data points) while here, we are working with the chi-square of the N classes (usually N=20) into which the residuals are divided and making a true chi-square test of significance.

A fundamental objection affecting not only the above procedure but also the computation of the upper and lower limits for any k-feasibility test stems from the fact that the image against which we are testing has been generated by the same set of data that we are testing; i.e., the data and the image are not statistically independent. On the other hand, we have observed that images generated by the MLE, when stopped according to the rule derived from the feasibility test, are visually good, with a useful compromise between edge sharpness and low noise in regions of high activity. Results similar to ours have been also obtained by Hebert et al.[10] Many other researchers (mostly in the fields of astronomy and geophysics) have used several variations of the same tests which they proposed to use as constraints for image reconstruction.[7-9,11-13] The practical usefulness of the feasibility-based tests or constraints and the theoretical difficulty stated above are two aspects that ought to be reconciled.

From statistics, we know that the hypothesis tests that can be derived from the above definitions of feasibility would have complete validity if the parameters that determine the Poisson process against which we are testing were completely and independently specified. This fact can lead us to a useful assumption that, if proven correct, may justify our feasibility-based testing:

We have an image a_i i=1,...B and its projection h_j j=1,...D. We do not know if the image was obtained during a reconstruction process or if it was generated by the proverbial team of monkeys making images by strewing white dots at random, a_i being the number that land in the i-th pixel. We have also a set of data p_j j=1,...D and we are asked to test the hypothesis that that image, as if completely specified, could have been the source that generated the data by a Poisson process. We would then proceed with testing the hypothesis by the methods developed or implied by the definitions of feasibility given above.

This assumption gives us also the way to compute the upper and lower limits for the 2-feasibility test ($\chi^2 = D$) stated above. If we are testing the image as if completely specified, the left side of Eq. (1) for n=2 possesses a χ^2 distribution with D - 1 degrees of freedom.[14] Statistical analysis indicates some upper and lower limits which χ^2 can plausibly take in a "two tail" chi-square test. The margins depend on the desired confidence levels. For χ^2 the largest acceptable value at 99% confidence is about $D \pm 3.29\ (D)^{1/2}$.[11] For D=20000 data points that means an upper limit for χ^2 of 20465. If normalized to the number of data points, the upper limit for χ^2/D is 1.023. The corresponding lower limit in a "two tail" chi-square test is about 0.977. Hebert et al.[10] have developed their stopping criterion for the MLE algorithm based also on the χ^2 test of significance. It is important to keep in mind, however, that the expression $\chi^2=D$ is a constraint and that the statistical theory has been used only to define some confidence interval around the desired value D.

The above assumption would be justified, in practice, if an analysis of the distribution of the statistic H defined by Veklerov and Llacer[4] or the chi-square function values (for Def. 2) for reconstructions of a large number of different realizations of a fixed image source distribution were found to be chi-square distributed with the expected number of degrees of freedom. We have recently initiated such a study by generating 100 different realizations of a given source and reconstructing all of them by the MLE method. Although the study is not complete, preliminary results indicate that, for the range of iteration numbers of interest, where the images appear to be good representations of the original source, the function H of Veklerov and Llacer[4], has a chi-square distribution. This finding justifies, in our opinion, the feasibility-based tests and stopping criteria developed although a theoretical justification is still missing.

The concept of feasibility has been used by our group to develop the stopping rule for the MLE algorithm both in its first version[4] and in its more robust version[15] for real data from a tomograph. In the next sections of this paper it will be used to test two different reconstruction algorithms, the first based on entropy maximization and the second on full Bayesian reconstruction.

4. MAXIMUM ENTROPY RECONSTRUCTION

Maximum entropy techniques are receiving increasing attention in the field of image reconstruction from a set of projections. The maximum entropy solution yields the image with the lowest information content compatible with the data. Thus, the criterion will give an image which avoids any bias while satisfying the constraints. The method consists of maximizing the entropy of the image and of the noise subject to constraints. We have used in this work the noise model proposed initially by Frieden[16] and the Gauss-Seidel-Newton iterative solution method described by Gullberg and Tsui.[16] The Newton-Raphson method used initially by Frieden cannot be applied to the large problems that we are dealing with. We formulate the problem as follows:

Maximize:

$$- \sum_{i=1}^{B} \frac{a_i}{N} \log \frac{a_i}{N} - \rho \sum_{j=1}^{D} \frac{n_j}{N_n} \log \frac{n_j}{N_n} \tag{2}$$

where

$$N = \sum_{i=1}^{B} a_i; \qquad N_n = \sum_{j=1}^{D} n_j; \qquad a_i \geq 0; \qquad n_j \geq 0$$

subject to:

$$\sum_{i=1}^{B} f_{ji}\, a_i + n_j - n_m = p_j \qquad\qquad j=1,\ldots.D \tag{3}$$

In the formulation (2) we use the Shannon form of entropy which was first used by Frieden[16] in image reconstruction. Parameter ρ is the weight quantifying the relative importance of the entropy of the noise vs. that of the image. The noise, computed as the difference between the data and the forward projections, can be represented as $n_j - n_m$, where n_j is a biased (always positive) noise term and n_m is a constant noise bias which insures the positivity of n_j. To be consistent with the Poisson nature of the data, we have chosen:

$$n_m = \max [\, 2 \sqrt{p_j} \,] \qquad d=1,\ldots D$$

and

$$\left[\sum_{j=1}^{D} (n_j - n_m)\right] / D = 0; \qquad N_m = \sum_{j=1}^{D} n_j = D\, n_m$$

The parameter ρ effectively controls the smoothness of the solution. The larger the constant, the less smooth an image will result for the given data. If the data contains very little noise, ρ can be made very high thus obtaining high sharpness. The bias n_m also controls the smoothness of the image. In this case the larger the bias, the smoother the image, with a resulting higher background. Note that in this model, the noise n_j is a variable in the optimization problem and the solution gives unbiased estimates of both the image and the noise. However, the model does not contain prior information about the Poisson nature of the data. It seems well suited to handle situations of little or no noise but we cannot expect it to be an accurate model for the case of Poisson noise.

In order to maximize (2) with the constraints (3), we introduce the Lagrangian function and solve the unconstrained problem of maximizing:

$$L = -\sum_{i=1}^{B} \frac{a_i}{N} \log \frac{a_i}{N} - \rho \sum_{j=1}^{D} \frac{n_j}{N_n} \log \frac{n_j}{N_n}$$

$$- \sum_{j=1}^{D} \lambda_j \left[p_j - n_j + n_m - \sum_{i=1}^{B} f_{ji}\, a_i \right] \frac{1}{N}$$

where λ_j $j=1,\ldots D$ are Lagrange multipliers. To compute the maximum of the function L we set the partial derivatives of L with respect to a_i, n_j and λ_j $i=1,\ldots B$; $j=1,\ldots D$ equal to zero, with the following solution:

$$a_i = N \exp(-1) \prod_{j=1}^{D} \exp(f_{ji}\, \lambda_j)$$

$$n_j = N_n \exp(-1) \exp(\mu_j N / \rho N_n) \qquad (4)$$

where the Lagrange multipliers are determined by the system of D nonlinear equations:

$$N \exp(-1) \sum_{i=1}^{B} f_{ji} \prod_{j=1}^{D} \exp(f_{ji} \lambda_j) + N_n \exp(-1) \exp(\mu_j N/\rho N_n) - n_m + p_j = 0 \qquad j=1,\ldots D$$

The solution of this system of D nonlinear equations for the Lagrange multipliers is obtained by using the Gauss-Seidel-Newton non-linear iterative algorithm described, for example, in Ortega and Rheinboldt.[18] Once the algorithm converges, the resulting Lagrange multipliers are subtituted in (4) and the solution is obtained. As the iterative algorithm for obtaining the Lagrange multipliers progresses, we can substitute the multipliers in (4) and monitor the formation of the image as well as its various statistics.

4.1 Results

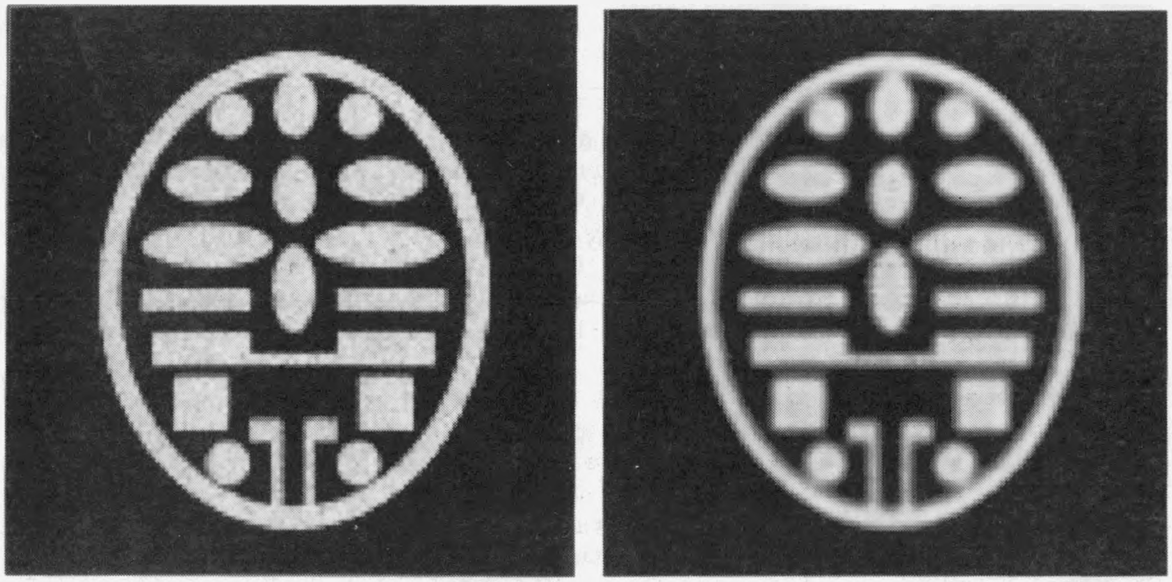
To test the algorithm we have used the computer generated brain-like phantom with one million counts shown in Fig. 1a). Figure 1b) shows the maximum entropy reconstruction of the brain phantom in the absence of noise, at convergence after 10 iterations, with $\rho = 50$. The image plane contains 128x128 pixels and the projection data are obtained by simulating the 512 detector ECAT-III of UCLA.[19] Note that the reconstruction is nearly perfect. This shows that the algorithm works very well in the case of no noise in which it is possible to use large values of ρ. In the presence of Poisson noise, the images obtained by the above algorithm are not satisfactory. Nevertheless, we have obtained reasonable results by adjusting ρ to a suitable value and smoothing the final image.

The images in Figs. 2a) and 2b) show the maximum entropy reconstructions of the brain phantom 1a) with Poisson data corresponding to 1 million counts, using $\rho = 8$. Image 2a) is after twenty iterations, at convergence. Image 2b) is obtained by filtering image 2a) by convolution with a Gaussian kernel with $\sigma = 0.75$ pixels, which improves its appearance. Note that the structures of the phantom are clearly visible in reconstructions 2a) and 2b) but the application of the feasibility concepts developed in Sects. 2 and 3 shows that none of the images pass the tests of weak feasibility nor the Veklerov and Llacer test.[4] In our opinion, the two main reasons for the failure of the algorithm to pass the feasibility tests are the lack of an adequate noise model (Poisson) in the algorithm and that the Lagrangian method requires the fitting of all the individual constraints (Eq. 3) separately, which results in a large number of Lagrange multipliers (one per tube). The solution is then very sensitive to any occasional abnormally large errors.[11]

There are several ways to reformulate the problem in an improved manner but the introduction of an appropriate Poisson data model will result in the addition of even more constraints to the model. However, the concept of maximum entropy proved useful as a prior probability function in a full Bayesian framework.

5. FULL BAYESIAN IMAGE RECONSTRUCTION

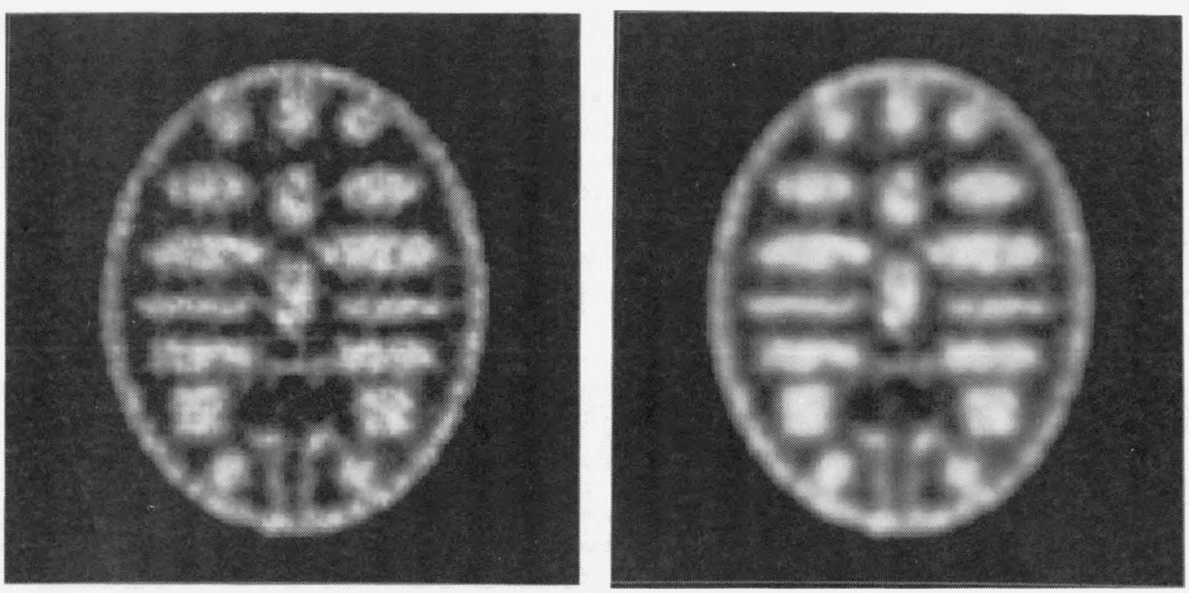
A natural mechanism for incorporating prior knowledge is through Bayesian theory. A Bayesian reconstruction seeks an image that maximizes the



a b

XBB 880-11781 XBB 880-11782

Figure 1 - a) Mathematical brain-like phantom with 1 million counts. b) Reconstruction by maximum entropy with constraints for the case with no noise in the process of assigning counts to projections.



a b

XBB 880-11783 XBB 880-11784

Figure 2 - a) ME reconstruction for the case of assignment of counts to the projection data according to a probability matrix f. b) After slight filtering. Neither of the two images are feasible.

probability of that image given the measurement data. If we consider the case of an image source **a**, an imaging instrument that transforms that source by a matrix **f** into a measurement vector **p**, then a Bayesian approach to reconstruction will seek an estimate of **a**, such that the probability $P(a|p)$ is maximized. By Bayes rule,

$$P(\mathbf{a}|\mathbf{p}) = P(\mathbf{p}|\mathbf{a})\ P(\mathbf{a})\ /\ P(\mathbf{p}) \tag{5}$$

The maximization of $P(\mathbf{a}|\mathbf{p})$ is done by maximizing the product $P(p|\mathbf{a})\ P(a)$. The first term of the product is precisely the likelihood function. Since $P(p)$ is constant, the main difference between the Bayesian approach and the MLE is the inclusion of the prior probability $P(\mathbf{a})$ function for the image. The conditional probability $P(\mathbf{p}|\mathbf{a})$ describes the projection noise and its possible object dependence. It is fully specified in the statement of the problem as the likelihood. We use the Poisson assumption and we generalize an important concept originally introduced by Frieden and Wells:[20] A Poisson image projection consists of a number of counts in each detector unit, implying the existence of a smallest intensity increment Δp_j associated with each detector to register one count. Frieden and Wells use a constant value Δp independent of the detector, but we introduce the detector dependence because, in real ET cases, the intensity increment can depend on the detector (for example, when detectors have different gains or there is attenuation in the emitting object). Under the previous assumptions, the conditional probability is:

$$P(p|\mathbf{a}) = \prod_{j=1}^{D} \exp(-h'_j)\ \frac{(h'_j)^{p_j/\Delta p_j}}{(p_j/\Delta p_j)!} \tag{6}$$

where

$$h'_j = \Big(\sum_{i=1}^{B} f_{ji}\ a_i\ \Big)/\Delta p_j \qquad\qquad j = 1, \ldots D \tag{7}$$

The prior probability $P(\mathbf{a})$ is a probability distribution function for the image we are seeking. For each pixel, $P(\mathbf{a})$ contains the probability that it takes a particular set of values. In general, and particularly in medical imaging, $P(\mathbf{a})$ is not known with any degree of accuracy. Faced with that situation, Bayesian reconstructions have aimed at using a function for $P(a)$ that corresponds to some general truth about the object being imaged. One prior knowledge function that has received substantial attention for more than 30 years is the entropy function.[16,20-23] The practical interest in using the entropy as a prior (and maximizing it in Bayes' framework) is that the resulting reconstruction has minimal configurational information, so there must be evidence in the data for any structure which is seen.[11]

Let us now describe an additional concept following Frieden.[16] Suppose that N is the total number of counts in the object and that there is an intensity increment Δa describing the finest known intensity jumps that are possible in the object. Now the values $a_i/\Delta a$ $i=1,\ldots B$ are dimensionless numbers since Δa contains the unit of radiance. It is now possible to define the prior probability of an image as proportional to the number of ways in which that image can occur.[24] The logarithm of the prior probability is then the entropy of the image:

$$\log p(a) = - \sum_{i=1}^{B} (a_i/\Delta a) \log(a_i/\Delta a) \quad + \text{const. terms.}$$

We have, then, formulated a problem with only one Lagrange multiplier (μ) that constrains the solution to a fixed number of counts N in the image. After taking the log of the product entropy and likelihood, the function to be maximized is:

$$BY = - \sum_{i=1}^{B} (a_i/\Delta a) \log(a_i/\Delta a) + \sum_{j=1}^{D} (-h'_j + (p_j/\Delta p_j) \log(h'_j)) - \mu(\sum_{i=1}^{B} a_i - N) \tag{8}$$

with h'_j as defined in (7). Parameters Δa and Δp_j $j=1,\ldots.D$, defined above, control the relative weight of the entropy vs. likelihood functions.

Parameter Δa can be either computed theoretically or adjusted. If the unknown image source were known to have 30 distinct levels of activity, for example, Δa should be set to the finest jump in the image or, approximately $\Delta a = m/30$, where m is the maximum intensity in the image. Alternatively, parameter Δa can also be used to adjust the contrast of the resulting image, and, in particular, it can be set to a value that results in convergent, feasible solutions, according to our criterion. Theoretical work is still needed to fully understand the relationship between the interpretation of the values of Δa and Δp_j proposed by Frieden and Wells[20] and our experimental findings.

The problem of maximizing Eq. (8) has been solved by the direct maximization algorithm.[8] Taking the partial derivatives of BY, Eq. (8), with respect to a_i and setting them equal to zero we obtain, for each of the pixels $i=1,2,\ldots,B$ one equation of the form

$$\frac{\partial\, BY}{\partial\, a_i} = -(1/\Delta a) \log(a_i/\Delta a) - 1/\Delta a + \sum_{j=1}^{D} [-f_{ji}/\Delta p_j + (p_j/\Delta p_j) (1/h'_j) (f_{ji}/\Delta p_j)] - \mu = 0 \tag{9}$$

Exponentiating Eq. (9) and placing a_i from the first term into the left side, we obtain an expression that can be used to derive an iterative formula.[25] Taking into account the definition of h'_j, the resulting pixel values for iteration k+1 are obtained in terms of the values of the k-th iteration and the data by

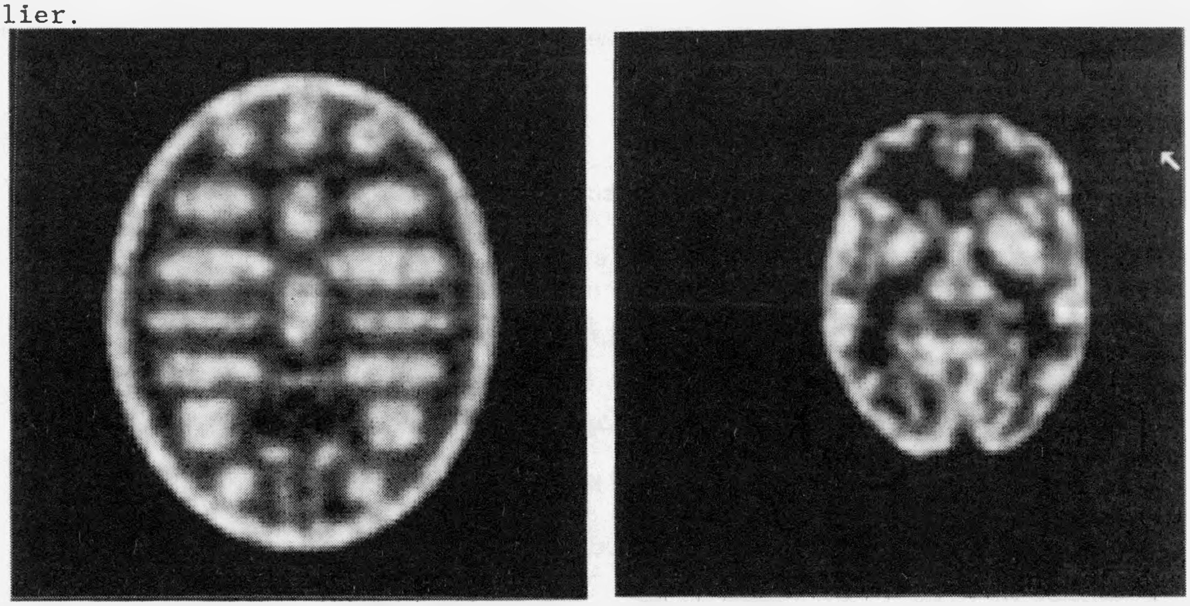
$$a_i^{k+1} = (1 - \alpha)\, a_i^k + \alpha\, K \exp\left(\Delta a \sum_{j=1}^{D} [\, f_{ji}\, (1/\Delta p_j)\, (p_j / \sum_{i=1}^{B} f_{ji}\, a_i^k - 1)\,] \right) \quad (10)$$

where K is a normalization constant that is recomputed at the end of each iteration to keep the total number of counts equal to N, and is equivalent to obtaining the Lagrange multiplier μ of Eq. (8). Note that the parameters Δp_j allow us to account for the absorption and detector gain corrections for reconstruction from real data. Eq. 10 contains a new parameter α. The correction computed by each iteration is the result of the exponentiation in that equation. These corrections correspond to a vector in a multidimensional space with the right direction but with too large a magnitude because of an instability introduced by the exponential function.[8] Only a small fraction of the correction, which is determined by the parameter α, must be used at each iteration in order to insure convergence.

5.1 Results

Simulated data: Figure 3a) shows the reconstruction of the phantom 1a) carried out with the Bayesian method at convergence with slight post-filtering with a Gaussian kernel of 0.4 pixels standard deviation. Since the data are simulated, $\Delta p_j = 1$ for all j was chosen. From previous analysis we know that the maximum of the image is approximately 200 counts in one pixel and the number of levels is 4. Thus, $\Delta a = 200/4 = 50$ was chosen. Note that the reconstruction is visually better than the one obtained with the maximum entropy method, figures 2a) and 2b). The reconstruction passes the feasibility tests described earlier.

a



b

XBB 880-11785

XBB 880-11786

Figure. 3 - a) Bayesian reconstruction with entropy prior and slight post-filtering of the simulated brain phantom. b) Bayesian reconstruction of a Hoffman brain phantom set of data with 1 million counts, with slight post-filtering.

Real data: Figure 3b) shows the results of our reconstruction with the above method for the UCLA ECAT-III Hoffman brain phantom with 1 million counts, at convergence. We know from previous computations that the maximum intensity of the image is approximately 3000 counts and that the number of activity levels is 5. Parameter Δa was, therefore, set to $\Delta a = 3000/5 = 600$. Since this is a case with real data, absorption and detector gain corrections have to be applied. The data p_j were multiplied by factors g_j. One count received by a detector pair j corresponds to an increment of g_j counts. Thus parameters Δp_j should be set to the values $\Delta p_j = g_j$. The resulting reconstruction passes the feasibility test described by Llacer and Veklerov[15] for real data. The image of Fig. 3b), has been filtered slightly with a Gaussian kernel of $\sigma = 0.6$ pixels, which improves the appearance of the image, while keeping the image feasible.

In both the reconstructions presented, parameters Δa and Δp_j have been computed theoretically and the results fit the feasibility theory. We do not have at this time, however, the theoretical knowledge to insure that this will always be the case. A practical adjustment of Δa for a given class of images could be carried out by starting with a guess for Δa (we suppose Δp_j known from the absorption and gain corrections), carrying out the reconstructions and monitoring the moments of the distribution of the residuals or the parameter H of Veklerov and Llacer[4] during the iterations. If the resulting images are not feasible, the Δa parameter should be modified.

The quality of the images of Figs. 3a) and 3b) appears quite comparable to the best of the MLE solutions, and clearly superior to the Filtered Backprojection reconstructions that we have used for comparison. The merit of our preliminary Bayesian work can be summarized, in our view, in the following two points:

1) If the choice of parameter Δa is made so that the final results are feasible, the iterative procedure converges to stable images which are visually good in a reasonable number of iterations (50 to 100).

2) The method of solution shown offers a methodology for emission tomography image reconstruction from Bayesian functions. The method is not limited to entropy prior distributions and can become the basis for successful future work.

6. CONCLUSIONS

In this paper we have reviewed the concept of image feasibility and its applicability to image reconstruction in Emission Tomography. We conclude that the feasibility based stopping criteria and residual analysis are powerful tools, but more theoretical study is needed to fully understand their basis.

We have developed two different reconstruction algorithms. The first one, based on Maximum Entropy, works well in cases of no noise but, in presence of Poisson noise, its reconstructions are inadequate due to the lack of an appropriate noise model. The second algorithm developed is a full Bayesian method in which we have incorporated a fundamental "sharpness" parameter Δa that can be computed theoretically or adjusted experimentally for feasibility, obtaining high quality reconstructions in presence of Poisson noise. Our new Bayesian reconstruction algorithm, defined by Eq. (10), has been applied to simulated and real data resulting in an iterative procedure which converges with excellent image stability. We are presently working on an improved algorithm to maximize (8) that does not require the exponentiation step or parameter α of Eq. (10).

7. ACKNOWLEDGMENTS

The authors would like to acknowledge the assistance of F. S. Goulding and E. Veklerov in reviewing the manuscript. 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. The work of J. Nunez has been supported in part by the Catalan Studies Program (U. of California and the Generalitat de Catalunya, Spain).

8. REFERENCES

1. J. Llacer and E. Veklerov, "The high sensitivity of the Maximum Likelihood Estimator method of Tomographic Image Reconstruction". SPIE Proc., Conf. on Med. Imag., vol. 767, 70-76 (1987).
2. 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.*, vol. NS-32, No. 5, 3864-3872 (1985).
3. J. Llacer and 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 (IPMI), International Conference, Utrecht, The Netherlands, 201-215, published by Plenum Pub. Co., (1987).
4. E. Veklerov and J. Llacer, "Stopping rule for the MLE algorithm based on statistical hypothesis testing", *IEEE Trans. Med. Imaging*, vol. MI-6, 313-319, (1987).
5. J. Llacer and E. Veklerov, "The use of a stopping rule in iterative image reconstruction", invited paper, June 1988 American Mathematical Society-IMS-SIAM Joint Conference on Spatial Statistics and Imaging, Brunswick, Maine, to be published in the Conference Proceedings (1988).
6. J. Llacer, E. Veklerov and J. Nunez,"The concept of causality in image reconstruction", presented at the Sept. 1988 NATO Advanced Study Institute, Povoa, Portugal, to be published in the Institute Proceedings (1988).
7. J.G. Ables, "Maximum Entropy spectral analysis", *Astron. Astrophys. Suppl. Ser.* vol. 15, 383-393 (1974).
8. S.F. Gull and G.J. Daniel, "Image reconstruction from incomplete and noisy data", *Nature*, vol. 272, 686-690 (1978).
9. R. Narayan and R. Nityananda, "Maximum entropy image restoration in astronomy", *Ann. Rev. Astron. Astrophys.* vol. 24, 127-170 (1986).
10. T. Hebert, R. Leahy and M. Singh, "Fast MLE for SPECT using an intermediate polar representation and stopping criterion", *IEEE Trans. Nucl. Sci.*, Vol. 35, No. 1, 615-619 (1988).
11. J. Skilling and R.K. Bryan, "Maximum Entropy reconstruction: general algorithm", *Monthly Notices of the Royal Astronomical Soc.*, vol. 211, 111-124 (1984).

12. J. Reiter and J. Pfleiderer, "Improvement of MEM-deconvolution by an additional constraint", *Astron. Astrophys.* vol. 166, 381-392 (1986).

13. A. Tarantola, *Inverse Problem Theory*. Elsevier (1987).

14. H. Cramer, *Mathematical Methods of Statistics*. Princeton Univ. Press. Princeton (1946).

15. J. Llacer and E. Veklerov, "Feasible images and practical stopping rules in iterative image reconstruction", accepted for publication, *IEEE Trans. Med. Imaging* (1989).

16. B.R. Frieden, "Restoring with Maximum Likelihood and Maximum Entropy", *J. Opt. Soc. Amer.*, vol. 62, 511-518 (1972).

17. G.T. Gullberg and B.M.W. Tsui, "Maximum Entropy reconstruction with constraints: iterative algorithms for solving the primal and dual programs", Proc. of the Xth Information Processing in Medical Imaging (IMPI), International Conference, Utrecht, The Netherlands, 181-199, published by Plenum Pub. Co. (1987).

18. J.M. Ortega and W.C. Rheinboldt, *Iterative solutions of nonlinear equations in several variables*. Academic Press, New York (1970).

19. E. Veklerov, J. Llacer and E.J. Hoffman, "MLE reconstruction of a brain phantom using a Monte Carlo transition matrix and a statistical stopping rule", *IEEE Trans. Nucl. Sci.*, NS-35, No. 1, 603-607 (1988).

20. B.R. Frieden and D.C. Wells, "Restoring with maximum entropy. III Poisson sources and backgrounds", *J. Opt. Soc. Am.*, Vol 68, No. 1, 93-103 (1978).

21. E.T. Jaynes, *Phys. Rev.* vol. 106, 620-630, also vol. 108, 171-190, (1957). (These papers are reprinted in *Papers on Probability, Statistics and Statistical Physics, a reprint collection*, Dordrecht, The Netherlands, D. Reidel, 1982).

22. E.T. Jaynes, "On the rationale of the Maximum-Entropy methods", *Proc. IEEE*, vol. 70, 939-952 (1982).

23. E.T. Jaynes, "Monkeys, kangaroos and N", in *Maximum Entropy and Bayesian Methods in Applied Statistics*, Proc. of the Fourth Maximum Entropy Workshop, Ed. J.H. Justice, Univ. of Calgary, 26-58 (1984).

24. E.T. Jaynes, "Prior probabilities", *IEEE Trans.* vol. SSC-4, 227-241(1968).

25. E.S. Meinel, "Origins of linear and nonlinear recursive restoration algorithms", *J. Opt. Soc. Am. A*, Vol.3, No. 6, 787-799 (1986).