Simultaneous reconstruction, segmentation, and edge enhancement of relatively piecewise continuous images with intensity-level information

Z. Liang, R. Jaszczak, and R. Coleman
Department of Radiology, Duke University Medical Center, Durham, North Carolina 27710
V. Johnson
Institute of Statistics and Decision Sciences, Duke University, Durham, North Carolina 27706

(Received 9 April 1990; accepted for publication 1 October 1990)

A multinomial image model is proposed which uses intensity-level information for reconstruction of contiguous image regions. The intensity-level information assumes that image intensities are relatively constant within contiguous regions over the image-pixel array and that intensity levels of these regions are determined either empirically or theoretically by information criteria. These conditions may be valid, for example, for cardiac blood-pool imaging, where the intensity levels (or radionuclide activities) of myocardium, blood-pool, and background regions are distinct and the activities within each region of muscle, blood, or background are relatively uniform. To test the model, a mathematical phantom over a 64×64 array was constructed. The phantom had three contiguous regions. Each region had a different intensity level. Measurements from the phantom were simulated using an emission-tomography geometry. Fifty projections were generated over 180°, with 64 equally spaced parallel rays per projection. Projection data were randomized to contain Poisson noise. Image reconstructions were performed using an iterative maximum *a posteriori* probability procedure. The contiguous regions corresponding to the three intensity levels were automatically segmented. Simultaneously, the edges of the regions were sharpened. Noise in the reconstructed images was significantly suppressed. Convergence of the iterative procedure to the phantom was observed. Compared with maximum likelihood and filtered-backprojection approaches, the results obtained using the maximum *a posteriori* probability with the intensity-level information demonstrated qualitative and quantitative improvement in localizing the regions of varying intensities.

Key words: tomography, maximum *a posteriori* probability, region of interests

I. INTRODUCTION

The importance of identifying isolated regions of interests (ROIs) (e.g., organs and/or tumors) within the body in medical image processing is well known.[1-3] For example, in cardiac blood-pool imaging, if the boundaries of the blood-pool within the ventricle can be identified at different times, then the blood volume curve within the ventricle can be determined.

Conventionally, ROIs are identified by two steps. First, an image is reconstructed from acquired data by use of a reconstruction method.[4,5] A segmentation or enhancement technique is then applied to the reconstructed image.[6,7] After the first step the reconstructed regions of different objects (or tissues) are usually interlaced and the boundaries are blurred due to low-count density of acquired data and artifacts relating to object motion, detection errors, etc. In the second step, although segmentation techniques can be used to mitigate the problem associated with the interlacing and blur, the segmented image might not be consistent with the acquired data and the artifacts in the reconstructed image remain.

In this paper, we propose a maximum *a posteriori* probability (MAP) method to simultaneously reconstruct and segment images. The method considers both the noise statistics of the acquired data and the following characteristics of images: (1) image intensity is nonnegative; (2) image-intensity distribution is conserved; (3) image intensities are relatively constant within contiguous regions over the image-pixel array; and (4) the intensity levels of these regions are determined either empirically or theoretically by selected information criteria.

The nonnegativity and conservation of image-intensity distributions are applicable in most imaging situations. The intensity-level information of piecewise continuous distributions may be valid for some imaging studies. For example, for cardiac blood-pool imaging, the intensities (or radionuclide activities) of myocardium, blood-pool, and background regions are distinct and are relatively uniform within each region. The activity levels of the muscle, blood, and background can be determined before performing image reconstruction. Empirically, the radionuclide activity of the blood can be measured by taking a blood sample from the patient. The activity of the muscle may be assumed to be negligible. The activity of surrounding tissue (or background) outside the heart may be determined from previous studies of typical blood-pool/background ratios and from the acquired data and the measured radionuclide activity of the blood. A theoretical discussion for determining the intensity levels by applying information criteria to the acquired data is included in the Appendix.

In the computer simulations presented by this paper, we assumed that the intensity levels were estimated with about 5% errors from a known phantom which consists of different ROIs with three distinct true levels. The three assumed intensity levels with 5% errors and the simulated projection

 0094-2405/91/030394-08$01.20

data from the phantom were then used to simultaneously reconstruct and segment images. The reconstructed image is expected to consist of different regions with three relatively uniform intensity levels. The three reconstructed intensities, although relatively uniform, may not necessarily be the same as the three assumed intensities, since the reconstructed image must be consistent with the projection data. The reconstructed regions should correspond to the ROIs of the phantom (i.e., the image is simultaneously reconstructed and segmented).[8] Our simulations confirmed this expectation. The reconstructed regions corresponding to the three assumed intensity levels were automatically segmented and, simultaneously, the edges of the regions were sharpened. The reconstructed levels were closer to the true levels of the phantom, as compared with the assumed intensity levels. Noise in the reconstructed images was significantly suppressed. A root-mean-square error (RMSE) criterion[4] was used to show the convergence behavior of the iterative MAP reconstruction procedure. Compared with maximum likelihood (ML) and filtered-backprojection (FBP) approaches, the simulation results using intensity-level information demonstrated qualitative and quantitative improvement in identifying ROIs. Further studies using experimental and clinical imaging data are in progress.

II. THEORY

In order to facilitate the discussion of the multinomial image model and the intensity-level information, the source region is partitioned into J small square regions (or pixels). Each pixel is assigned a value ϕ_j which represents the average of source intensity over pixel j. In emission tomography, ϕ_j reflects the average photon emission in pixel j per unit time at time = 0 (at time = 0 means that the decay of radioisotope should be considered and corrected, if necessary). Then $\Phi = \{\phi_j\}_{j=1}^{J}$ represents the average emission map over the source-pixel array at time = 0, which is the true image to be reconstructed. For our formulation, we will assume that the average photon emission ϕ_j per unit time is an integer variable in the range from zero to a finite value. The discrete value ϕ_j is proportional to the continuous radioisotope concentration in pixel j. The finite value is related to the finite total dose injected into the body.

The emission map over the source array may be viewed as a distribution of photons over the pixels. The total number of photons in the distribution is the total average emissions, $N = \Sigma_j \phi_j$. The constant N is proportional to the dose injected into the body. If the N photons can be assumed as to be distributed over the J pixels in different ways, all the distributions make up an ensemble. We attempt to find that distribution which is most probable in the ensemble and which is also consistent with the acquired data. The distribution of N photons over J pixels in the array follows the multinomial process.[9] Let p_j represent the *a priori* probability of a photon emitted from pixel j, the probability of a distribution $\{\phi_j\}$ (or an emission map) occurring over the pixel array is expressed as[9,10]

$$P(\Phi) = \frac{N!}{\Pi_j \phi_j!} \prod_{j=1}^{J} [p_j]^{\phi_j}$$

and

$$\sum_j p_j = 1, \tag{1}$$

i.e., the multinomial image model. The *a priori* probability map $\{p_j\}$ associated with the source array is developed in this paper to consider the intensity-level information.

In some imaging situations, the emission map (or source distribution) $\{\phi_j\}$ may be relatively constant within regions over the source pixel array. The emission intensities of these regions may be determined either empirically or theoretically before performing image reconstruction. Mathematically, the intensity-level information is characterized by a set of assumed intensity levels $\{\phi_s^e\}_{s=1}^{K}$, where K is the number of the levels. For simplicity, we consider only three intensity levels in the following to derive the mathematical expressions of $\{p_j\}$ and to verify these mathematical formulas by computer simulations. It is straightforward to extend the formulas to include more intensity levels. The three intensity levels are (1) ϕ_1^e for cold region, (2) ϕ_2^e for hot regions, and (3) ϕ_3^e for background. The values of ϕ_1^e, ϕ_2^e, and ϕ_3^e are assumed to have been determined empirically with about 5% errors from actual intensity levels. A theoretical determination using information criteria is given in the Appendix. If each photon has same chance of being emitted from the three intensity levels, then we have[11,12]

$$p_j = p_j(\phi_j|\phi_s^e)p_j(\phi_s^e) \approx \frac{1}{3}\sum_{s=1}^{3} \delta(\phi_j - \phi_s^e), \tag{2}$$

where $p_j(\phi_s^e) = 1$, because the levels $\{\phi_s^e\}$ and the number of levels have been determined before performing image reconstruction. Since the emission in each region is not strictly uniform, a Gaussian form, rather than delta function, may be used to reflect the relative uniform variation[11,12]

$$p_j = \frac{1}{3}\sum_s c_s \exp\left(\frac{-(\phi_j - \phi_s^e)^2}{2\sigma_s^2}\right), \tag{3}$$

where c_s is the normalization constant and σ_s is the standard deviation of ϕ_j around ϕ_s^e. In practice, since each photon would have different chances of being emitted from the three intensity levels, the more general form of p_j can be expressed as

$$p_j(\Phi) = \sum_s w_s c_s \exp\left(\frac{-(\phi_j - \phi_s^e)^2}{2\sigma_s^2}\right), \tag{4}$$

where w_s reflects the ratio of the number of pixels having the similar intensity ϕ_s^e to the total number J of pixels in the source array.[11,12] In this paper, we determine w_s using the neighbors of pixel j. For example, if only first-order neighbors of pixel j are considered, w_s is the ratio of the number of pixels (among the five pixels) having the similar intensity ϕ_s^e to five (i.e., pixel j plus the four nearest neighbors). To ensure continuity within each region of the ROIs and discontinuity at the boundaries of different intensity regions, function (4) is modified as[13]

$$p_j(\Phi) = \sum_s w_s c_s \exp\left(\frac{-(\phi_j - \phi_s^e)^2}{2\sigma_s^2} - \sum_q \frac{(\phi_{j+q} - \phi_s^e)^2}{2\sigma_s^2}\right), \tag{5}$$

where index q covers three neighbors of pixel j. It is assumed

that the three neighbors are adjacent to each other and have intensities similar to that of pixel j.[14] A hierarchical model similar to that of expressions (4) and (5) has been investigated for segmentation in Refs. 7 and 15. If the nonnegativity and conservation of image distributions are ignored, the image model expressed by (1) and (5) above is somewhat similar to the hierarchical model.[7,15]

Up to now, we have discussed the source (or image) distribution process expressed by (1) and (5). That source distribution (or configuration), which has the highest probability $P(\Phi)$ and is also consistent with the acquired data, would be the choice. This most probable distribution is usually assumed to be the representation of the actual source distribution. There may be other ways to approach the most probable distribution $\Phi^* = \{\phi_j^*\}$ consistent with the acquired data. In this paper, however, we treat $P(\Phi)$ as *a priori* probability of the source distribution and try to find that probable solution Φ^* which is consistent with the most probable likelihood of the data distribution. In emission tomography, the Poisson likelihood of a data distribution is well established as[16,17]

$$P(\mathbf{Y}|\Phi) = \prod_{i=1}^{I} \exp\Big(-\sum_j R_{ij}\phi_j\Big)\Big(\sum_j R_{ij}\phi_j\Big)^{Y_i} \Big/ Y_i!, \quad (6)$$

where R_{ij} is the probability of detecting photons from pixel j for projection ray i, Y_i is the number of photons detected for projection ray i, and $\mathbf{Y} = \{Y_i\}_{i=1}^{I}$ (I is the number of contributing rays). The assumption made in (6) is that each data element Y_i is a Poisson variable with a mean $\Sigma_j R_{ij}\phi_j$ and all elements $\{Y_i\}$ are independent. The means $\{\Sigma_j R_{ij}\phi_j\}$ of the Poisson variables $\{Y_i\}$ reflect the relation between the acquired data and the source $\{\phi_j\}$ via the probability matrix $\{R_{ij}\}$. Since $\{R_{ij}\}$ are continuously distributed over the interval $[0,1]$, the means $\{\Sigma_j R_{ij}\phi_j\}$ are continuous values over $[0,\infty)$. Since the emissions are relatively uniform over a 4π solid angle, the values $\{R_{ij}\}$ are very small for a conventional ring detector system. For parallel ray geometry, the value R_{ij} can be assumed to be proportional to the intersection length of pixel j and projection ray i.[4,5] If $\{R_{ij}\}$ are chosen as the intersection lengths of the pixels and the projection rays, the pixel values $\{\phi_j\}$ will be divided by the proportional factor, if the acquired data are used directly. Then the reconstructed image values are real over $[0,\infty)$. In practice, the proportional factor is usually neglected for image display. However, for quantitative evaluation of images, as in single photon or positron emission tomography, the proportional factor must be considered.[5]

The MAP method proposed in this paper to obtain the most probable solution Φ^* is formulated, via Bayes' law of $P(\Phi|\mathbf{Y}) = P(\mathbf{Y}|\Phi)P(\Phi)/P(\mathbf{Y})$, as

$$P(\Phi^*|\mathbf{Y}) \approx P(\mathbf{Y}|\Phi^*)P(\Phi^*) = \text{maximum}$$

or

$$\ln P(\mathbf{Y}|\Phi^*) + \ln P(\Phi^*) = \text{maximum}. \quad (7)$$

Directly solving Eq. (7) for Φ^* is very difficult due to the large size of the matrix $\{R_{ij}\}$. For example, if 50 projections each with 64 rays are acquired from an image-pixel array of 64×64, the size of the matrix is 3200×4096. An iterative approach may be therefore more practical. The expectation-maximization (EM) technique[18] is then employed to compute the solution Φ^*:[13]

$$\phi_k^{(n+1)} = \phi_k^{(n)} \frac{\Sigma_i R_{ik}(Y_i/\Sigma_j R_{ij}\phi_j^{(n)})}{\Sigma_i R_{ik} + \xi_k^{(n)} Z_k}, \quad (8)$$

where

$$Z_k(\Phi) = -\frac{\partial}{\partial\phi} \ln P(\Phi)|_{\Phi=\Phi^{(n)}+\lambda d^{(n)}}$$

and

$$\xi_k^{(n)} = [a\sqrt{n}/(b+n)] \sum_i R_{ik} > 0. \quad (9)$$

In (9), $\lambda \approx 1$ and $d_k^{(n)} = \phi_k^{(n)} - \phi_k^{(n-1)}$ can be chosen to facilitate computation; $\xi_k^{(n)}$ increases smoothly as the iterative index n increases and $\xi_k^{(n>m)}$ is set to be the maximal value $\xi_{\max}^{(m)}$ after m iterations. The values $a = 0.05$ and $b = 50$ were used in the simulation studies. A detailed discussion of the empirically determined parameter $\xi_k^{(n)}$ is available in Refs. 10 and 14. The expression of Z_k is

$$Z_k(\Phi) = \left[\left(\ln\phi_k + \frac{1}{2\phi_k}\right) - \ln\sum_s Q_s\right] + \left[\left(\phi_k/\sum_s Q_s\right)\sum_s Q_s(\phi_k - \phi_s^e)/\sigma_s^2\right], \quad (10)$$

where

$$Q_s = w_s c_s \exp\left(\frac{-(\phi_k - \phi_s^e)^2}{2\sigma_s^2} - \sum_q \frac{(\phi_{k+q} - \phi_s^e)^2}{2\sigma_s^2}\right). \quad (11)$$

This iterative Bayesian image processing (BIP) algorithm (8) has been applied to computer-simulated noise-free projections and randomized projections containing Poisson noise. The performance of algorithm (8), as compared with the iterative ML algorithm[16,17] and the FBP method,[4,5] is presented in the following sections.

III. METHODS

In this section, we describe the computer simulations for the phantom and the projections.

Figure 1 shows the mathematical phantom $\{S_j\}$ with

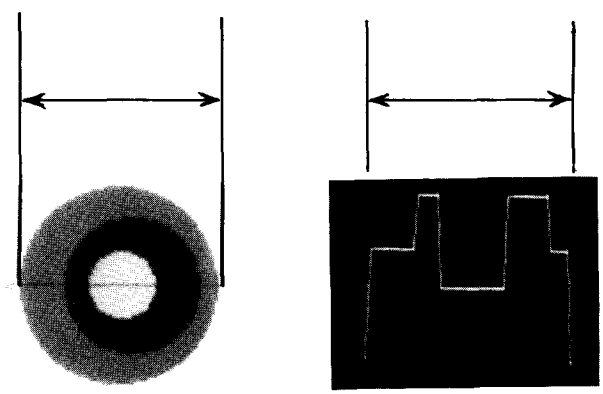

FIG. 1. The simulated phanton with three distinct intensity levels: the cold region of 2 intensity units, the hot region of 4.5 units, and the background of 3 units. The curve represents the intensity profile crossing along the center line. Note that the three regions within the phantom do not include the region exterior to the phantom, which has zero intensity.

three intensity levels, from which the values of $\{\phi_1^e, \phi_2^e$, and $\phi_3^e\}$ were assumed with 5% errors as mentioned before. The curve represents the true intensity profile through the center of the phantom. The central portion is the cold region associated with ϕ_1^e, and each pixel within this region has the true intensity of 2.0 units. The intensity units are proportional to emissions per unit time. The proportionality factor is determined by the chosen values of $\{R_{ij}\}$ as mentioned previously. The outer region is the background related to ϕ_3^e. Each pixel within this region has the true intensity of 3.0 units. The region between the cold and background regions represents the hot region (related to ϕ_2^e). This region has the true intensity of 4.5 units (4.5 was chosen to make the ratio $= 4.5/3 = 3/2$). The *a priori* assumed three intensity levels with about 5% errors from the phantom values were chosen as $\phi_1^e = 2.1$ units (for the cold region), $\phi_2^e = 3.1$ units (for the background), and $\phi_3^e = 4.4$ units (for the hot region). Fifty projections were simulated using a parallel beam geometry. Noise-free projection data were calculated by $\{\Sigma_j R_{ij} S_j\}$. Each element of the projection matrix $\{R_{ij}\}$ was determined by two-point interpolation technique.[19] By carefully normalizing the projection matrix, each element R_{ij} represents the probability of detecting photons from pixel j for projection ray i (see Sec. II). The noise-free projection data were assumed to be the means of randomized projection data, respectively. Each randomized projection datum Y_i was generated by a Poisson random number generator[20] around its mean $\Sigma_j R_{ij} S_j$. The sum of the means was 479 499.94, and the total simulated counts $\{Y_i\}$ was 479 588.

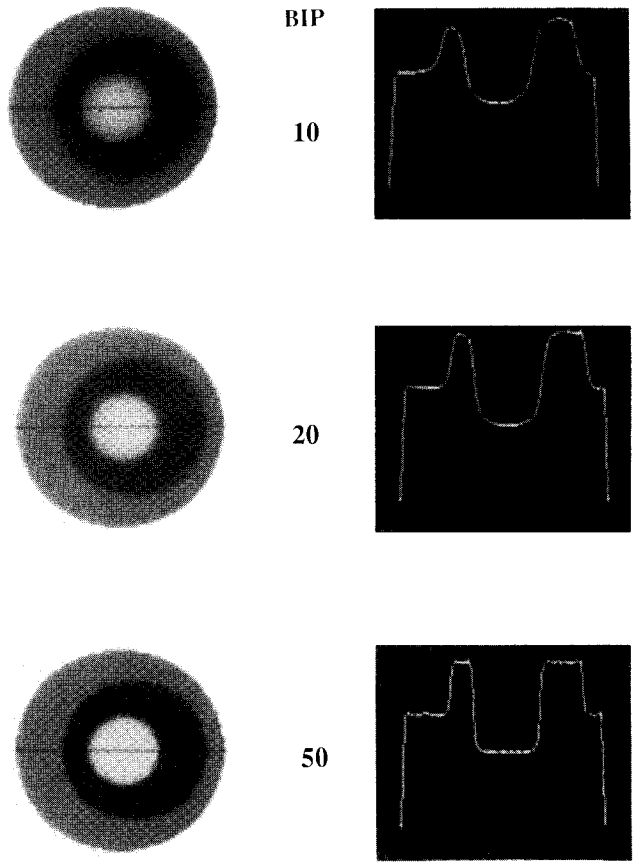

FIG. 2. The images reconstructed from the simulated noise-free projections using the BIP algorithm after 10, 20, and 50 iterations. Curves represent intensity cross profiles.

IV. RESULTS

In this section, we present the reconstructed images using the iterative BIP algorithm (8) for the simulated projection data from the mathematically specified phantom with three distinct intensity levels. The images are compared with those obtained using the iterative ML algorithm[16,17] and the FBP method.[4,5]

Figure 2 shows the images reconstructed from the simulated noise-free projection data using the iterative BIP algorithm (8) after 10, 20, and 50 iterations. The curves represent intensity profiles through the centers of the images, respectively. After 20 iterations, regions with intensities about 2, 3, and 4.5 units were obtained. Similar results were observed using the iterative ML algorithm[16,17] and FBP method[4,5] also for the noise-free projection data. The convergences of the iterative BIP and ML algorithms are compared in Fig. 3, using the RMSE criterion[4]

$$\mathrm{RMSE} = \left[\left(\sum_j (\phi_j^{(n)} - S_j)^2\right)\left(\sum_j (S_j - \bar{S})^2\right)^{-1}\right]^{1/2}, \tag{12}$$

where $\bar{S}$ is the average of $\{S_j\}$.

Figure 4 shows the reconstructed images using the BIP algorithm after 10, 20, and 50 iterations for the simulated projection data containing Poisson noise. The convergence behavior is shown in Fig. 5, compared as before with the ML algorithm. For comparison, the images reconstructed from the noisy projection data using the ML algorithm after 10, 20, and 50 iterations are shown in Fig. 6; the reconstructed images using the FBP method with a Ramp filter and with Hann filters of 1.0 and 0.5 Nyquist frequency cutoffs are shown in Fig. 7.

Quantitative comparisons in terms of mean, standard deviation, and standard deviation from the phantom are given in Table I. The calculations were carried out only using the profiles through the center of the image.

In further tests of the iterative BIP algorithm (8), we added intensity levels of $\phi_4^e = 0.0$, $\phi_5^e = 1.0$, and $\phi_6^e = 6.0$, and

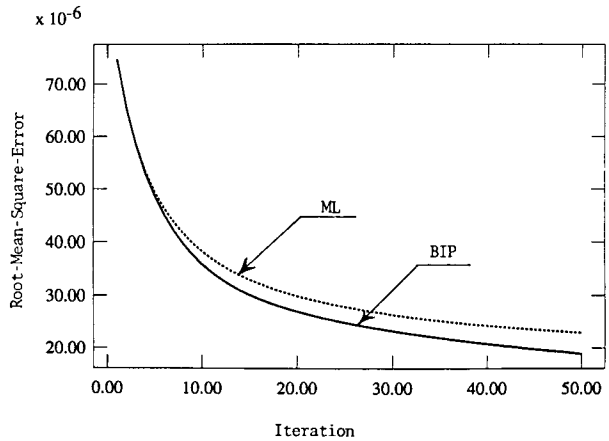

FIG. 3. The computed RMSEs as a function of iterative number for the BIP and ML algorithms in the case of noise-free projections.

APPENDIX

In this section, a theoretical discussion for determining intensity levels by applying information criteria to acquired projection data is given. Since the intensity levels are determined in image space rather than in projection space, the acquired projection data are backprojected first to the image space. This step can be accomplished with a conventional backprojection method.[4,5] Let $\mathbf{X} = \{x_r\}_{r=1}^{J_0}$ be the backprojected data vector with J_0 elements in image space. The information criteria discussed here are the Akaike information criterion (AIC)[23]

$$\mathrm{AIC}(\Theta) = -2\ln[\Psi(\mathbf{X}|\Theta)] + 2K_0 \tag{A1}$$

and the minimum description length (MDL) information criterion[24,25]

$$\mathrm{MDL}(\Theta) = -\ln[\Psi(\mathbf{X}|\Theta)] + \tfrac{1}{2}K_0 \ln J_0, \tag{A2}$$

where $\Psi(\mathbf{X}|\Theta)$ is the maximum likelihood of backprojected data vector $\mathbf{X}$ given the parameter vector Θ, and K_0 is the number of freely adjustable parameters in Θ. These criteria are usually deployed, post hoc, to select one from a number of models having different dimensions of Θ, given $\mathbf{X}$. The models are defined by the likelihood $\Psi(\mathbf{X}|\Theta)$ with different forms characterized by the parameter vector Θ, or with one and the same form but with different restrictions on the parameter vector Θ. In the former case, there are a number of models defined by the different likelihood forms. For each different form, the corresponding parameter vector Θ is determined by maximizing the corresponding likelihood. The information criterion AIC or MDL is used to select one of these models. The one selected minimizes the criterion (A1) or (A2). In the latter case, there are a number of models defined by the same likelihood form with different restrictions on the parameter vector Θ or with different parameter vectors. These parameter vectors are determined by maximizing the likelihood with the restrictions on Θ. One vector (or one model in this case) is then selected from these vectors by minimizing the criterion (A1) or (A2). The latter approach may be appropriate for selecting the intensity levels from the backprojected data. There may be a number of models of the same likelihood form with different numbers of intensity levels. The model which minimizes AIC or MDL is selected, and this model will have a specific number of intensity levels.

Using the notations of this paper, the likelihood is formulated as follows. The backprojected data vector $\mathbf{X}$ has J elements, or $J_0 = J$. Let $f(x_r|\theta_s)$ represent the Gaussian function in (4), or

$$f(x_r|\theta_s) = (2\pi\sigma_s^2)^{-1/2} \exp\left(-\frac{(x_r - \phi_s^e)^2}{2\sigma_s^2}\right), \quad r = 1,2,\ldots,J, \quad s = 1,2,\ldots,K, \tag{A3}$$

where $\Theta = \{\theta_s = (\phi_s^e, \sigma_s^2)\}_{s=1}^{K}$ is a row vector of $2K$ elements, and K is the number of intensity levels. Assume that $W = \{w_s\}_{s=1}^{K}$, $(0 < w_s < 1, \Sigma_s w_s = 1)$, are the weights of $f(x_r|\theta_s)$. The likelihood function for $\mathbf{X}$ is then expressed as[26]

$$L(\mathbf{X}|W,\Theta) = \prod_{r=1}^{J} g(x_r|W,\Theta) = \prod_{r=1}^{J}\sum_{s=1}^{K} w_s f(x_r|\theta_s). \tag{A4}$$

and

$$\Psi(x|\Theta) = \max\{L(x|W,\Theta)\}.$$

Now we have a number of models of a Gaussian likelihood form $\{\theta_s = (\phi_s^e, \sigma_s^2)\}$ with different dimensions (or levels) K. The problem of determining the intensity levels is to select the K value and therefore the values of $\{\theta_s = (\phi_s^e, \sigma_s^2)\}$. This problem is the same as that of the second case mentioned above. For different K, there are different models of a Gaussian likelihood form with different parameter vectors $\Theta(K)$. That model with K_c levels, which minimizes the information criterion AIC or MDL, will be selected. The numerical calculations for K_c can be performed as follows: (1) determine the relation between K and K_0; (2) assume a K value, which is in a prespecified range $[K_{\min}, K_{\max}]$, where $K_{\min} \leqslant K_c \leqslant K_{\max}$; (3) maximize $\Psi(\mathbf{X}|\Theta)$ with respect to Θ for the assumed K value; (4) compute the criterion value of AIC or MDL; (5) repeat steps (2)–(4) for $K \pm 1$ values; (6) if the calculated criterion value for K using AIC or MDL is less than those for $K \pm 1$, then $K_c = K$ is selected, otherwise, $K \pm 2$ are assumed and the steps above are repeated. In (A4), there are $K_0 = 3K - 1$ freely adjustable parameters: $2K$ free parameters of $\{\theta_s = (\phi_s^e, \sigma_s^2)\}$ and $K - 1$ free parameters of $\{w_s\}$; hence step (1) is accomplished. The starting value K in the step (2) is chosen arbitrarily, depending on one's prior knowledge about the backprojected data $\{x_r\}$. If the starting value K is chosen closer to K_c, then less computation time is needed. Many numerical techniques can be used to maximize the likelihood with respect to the parameters $\{\theta_s\}$.[26,27] Investigation on this topic is under progress.

Note that the information criteria provide a possible means to theoretically determine the intensity levels. Both criteria measure the Kullback–Leibler mean distance (or the directed Kullback divergence[28]) between the estimated model and the model which generates the measured data, and selected that model which minimizes the criterion (A1) or (A2).

It has been shown that when AIC is used in model selection, that model with minimum AIC gives the minimum mean-squared error of prediction.[29] If the maximum likelihood is identical for two models, AIC selects that one with less freely adjustable parameters $\{\theta_s\}$.

The difference between AIC and MDL is due to the penalty terms of $(2K_0)$ for AIC and $(K_0 \ln J_0)$ for MDL. The former assumes a uniform prior to all model candidates and proposes to select the model which yields the maximum likelihood, so the penalty term of AIC is independent of the number of measured data. The latter one, however, assigns each competing model an *a priori* probability and selects the model which yields the maximum posteriori probability. Detailed comments on the two criteria can be found in Refs. 30–32.

[1] G. Pohost, C. Higgins, J. Morganroth, J. Ritchie, and H. Schelbert, *New Concepts in Cardiac Imaging* (Year Book Medical Publishers, Chicago, 1989).

[2] A. Margulis and C. Gooding, *Diagnostic Radiology* (Lippincott, Philadelphia, 1989).
[3] M. Goris and P. Briandet, *A Clinical and Mathematical Introduction to Computer Processing of Scintigraphic Images* (Raven, New York, 1983).
[4] G. Herman, *Image Reconstruction from Projections: The fundamentals of computerized tomography* (Academic, New York, 1980).
[5] T. Budinger and G. Gullberg, "Three-dimensional reconstruction in nuclear medicine emission imaging," IEEE Trans. Nucl. Sci. **21**, 2–20 (1974).
[6] E. Hall, *Computer Image Processing and Recognition* (Academic, New York, 1979).
[7] H. Derin and H. Elliott, "Modeling and segmentation of noisy and textured images using Gibbs random fields," IEEE Trans. Pattern Anal. Mach. Intell. **9**, 39–55 (1987).
[8] Z. Liang, R. Jaszczak, and R. Coleman, "Simultaneous reconstruction, segmentation, and enhancement of simulated heart images using a multinomial model and concentration-level information," J. Nucl. Med. **31**, 858 (1990).
[9] A. Papoulis, *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, New York, 1984).
[10] Z. Liang and H. Hart, "Bayesian reconstruction in emission computerized tomography," IEEE Trans. Nucl. Sci. **35**, 877–885 (1988).
[11] H. Hart and Z. Liang, "Bayesian image processing of data from constrained source distributions, II: Valued, uncorrelated and correlated constraints," Bull. Math. Biol. **49**, 75–94 (1987).
[12] Z. Liang and H. Hart, "Bayesian image processing of data from constrained source distributions: Fuzzy pattern constraints," Phys. Med. Biol. **32**, 1481–1494, (1987).
[13] Z. Liang and H. Hart, "Source continuity and boundary discontinuity considerations in Bayesian image processing," Med. Phys. **15**, 754–756, (1988).
[14] Z. Liang, R. Jaszczak, C. Floyd, and K. Greer, "A spatial interaction model for statistical image processing," in *Information Processing in Medical Imaging* (Wiley, New York, 1989).
[15] P. Kelly, H. Derin, and K. Hartt, "Adaptive segmentation of speckled images using a hierarchical random field model," IEEE Trans. Acoust. Speech Signal Process. **36**, 1628–1641 (1988).
[16] L. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," IEEE Trans. Med. Imag. **1**, 113–122 (1982).
[17] K. Lange and R. Carson, "EM reconstruction algorithms for emission and transmission tomography," J. Comput. Assist. Tomog. **8**, 306–316 (1984).
[18] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," J. R. Stat. Soc. **39**, 1–38 (1977).
[19] T. Peters, "Algorithms for fast back- and re-projection in computed tomography," IEEE Trans. Nucl. Sci. **28**, 3641–3647 (1981).
[20] B. Carnahan, H. Luther, and J. Wilkes, *Applied Numerical Methods* (Wiley, New York, 1978).
[21] V. Johnson, "Bayesian restoration of PET images using Gibbs priors," in *Information Processing in Medical Imaging* (Wiley, New York, 1989).
[22] C. Chen, V. Johnson, X. Hu, W. Wong, and C. Metz "PET image reconstruction with the use of correlated structural images," J. Nucl. Med. **31**, 748 (1990).
[23] H. Akaike, "A new look at the statistical model identification," IEEE Trans. Auto. Control **19**, 716–723 (1974).
[24] G. Schwarz, "Estimating the dimension of a model," Ann. Stat. **6**, 461–464 (1978).
[25] J. Rissanen, "Modeling by shortest data description," Automatica **14**, 465–471 (1978).
[26] T. Lei and W. Sewchard, "A new stochastic model-based image segmentation technique for X-ray CT images," in *Visual Communication and Image Processing* (SPIE, Boston, MA, 1988).
[27] D. Luenberger, *Introduction to Linear and Non-Linear Programming* (Addison-Wesley, Reading, MA, 1973).
[28] S. Kullback, *Information Theory and Statistics* (Wiley, New York, 1959).
[29] B. Efron, "How biased is the apparent error rate of a prediction rule," J. Am. Stat. Assoc. **81**, 461–470, (1986).
[30] H. Akaike, "A Bayesian analysis of the minimum AIC procedure," Ann. Inst. Stat. Math. **30**, 9–14 (1978).
[31] M. Stone, "Comments on model selection criteria of Akaike and Schwarz," J. R. Stat. Soc. **41**, 276–278 (1979).
[32] L. Baxter, "Futures of statistics," Am. Stat. **44**, 128–129 (1990).