A SIMULTANEOUS SEGMENTATION AND RECONSTRUCTION
MODEL FOR ECT IMAGES

Valen E. Johnson
ISDS, Duke University

James E. Bowsher
Department of Radiology, Duke University
Medical Center

DP# 91-A02
A SIMULTANEOUS SEGMENTATION AND RECONSTRUCTION MODEL FOR ECT IMAGES

Valen E. Johnson† and James E. Bowsher‡

ABSTRACT. We propose a Bayesian model for the reconstruction of ECT images. The prior density effectively reduces the dimension of the estimation problem by merging pixels in the image with similar intensities, thus making EM type algorithms feasible. Two specific models are proposed; one imposes no prior constraint on the shapes of regions except that regions must be at least four contiguous pixels in size, and the other imposes minor shape constraints and requires that regions be at least five pixels in size. The model provides an automatic segmentation of the image which may be useful in quantization and visualization applications. Additionally, mean intensities and region sizes are immediately available as output from the algorithm. Computationally, the models are easy to implement and require approximately a 5% increase in computation over standard EM/MLE algorithms.

Key Words. Bayesian inference, Gibbs distributions, positron emission tomography, emission computed tomography, estimation-maximization algorithm, Bayesian data augmentation.

† Valen E. Johnson is an Assistant Professor, Institute of Statistics and Decision Sciences, Duke University. ‡ James E. Bowsher is a Research Associate, Department of Radiology, Duke University Medical Center, Durham, NC 27706.
1. INTRODUCTION

Reconstruction of emission computed tomography images (ECT) using Bayesian methods in conjunction with estimation-maximization (EM) type algorithms (e.g. [10],[11]) has been the topic of extensive discussion in recent years. The reason for the interest in Bayesian methods derives primarily from the the failure of the maximum likelihood estimate (MLE) to provide adequate representations of the image scene. As numerous authors (e.g. [2]) have pointed out, the MLE fails because of the high dimension of the estimation problem; each pixel intensity represents a parameter, and the number of pixels in reconstructed images is typically on the order of \(10^4\). The topic of this paper is the presentation of a Gibbs prior density that effectively reduces the dimension of the estimation problem by adaptively merging pixels with similar values. An additional benefit of the model is that the image is automatically segmented into regions of interest, and the size and mean intensity of each region of interest are available for display and quantization.

We restrict attention to Gibbs priors because they have an important advantage not shared by other prior models for image scenes; they are specified in terms of local properties of images. Thus, prior information concerning global characteristics regarding size and shape of image regions are not required.

Previous models utilizing Gibbs priors have fallen into two basic categories. The first class of Gibbs priors may be characterized by the restriction of potential functions to potentials dependent on only first order neighborhood (or nearest neighbor) systems. To avoid smoothing across sharp edges in the image, these models typically employ bounded potential functions, although some loss of contrast is incurred with such limited spatial information. Examples of such models are [4],[5], and [8].

The second category of models is typified by the use of extended neighborhood systems and the inclusion of line sites in the field associated with the model. The expanded neighborhood system permits more accurate identification of homogeneous regions, and the line sites permit correlations between low contrast regions to be severed. Unfortunately, models with line sites are typically not convex, and so point estimation is troublesome. Also, care must be taken to prevent the introduction of artifacts associated with the configuration of line sites (e.g. boundaries tend to be horizontal and vertical if the line sites appear in a rectangular grid). Examples of these models may be found in [3],[8],[9].

The motivation for the models considered here derives from properties exhibited by ML reconstructions. By following the progress of the EM algorithm through successive iterations, it becomes clear that a small number of pixels are assigned an increasingly disproportionate share of the bin counts as the algorithm iterates, and that at convergence reconstructions consist of a small number of "hot" spots. This phenomenon can be avoided by reducing the scale of the image, for example from 128 × 128 to 32 × 32 arrays. However,
the penalty for this reduction is that it effectively defines image regions as being large rectangular pixels. The model described here accomplishes a similar reduction, but instead of arbitrarily collapsing pixels into square regions, our model has the effect of collapsing pixels according to current estimates of their intensities.

2. MODEL DESCRIPTION

To specify our Gibbs prior, let \( \lambda = \{\lambda_{ij}\} \) denote the array of image intensities defined on a graph \( \Xi = \{\xi_{ij}\} \), and let \( n = \{n_{ij}\} \) denote the array of unobserved pixel counts obtained from the EM algorithm (see []). Noting that \( n_{ij} \) has a Poisson distribution given \( \lambda_{ij} \), and that \( n_{ij} \) is independent of \( n_{kl} \) given \( \lambda_{ij} \) and \( \lambda_{kl} \), Johnson et al [9] noted that a square root transformation of the unobserved counts yielded a likelihood function that was approximately Gaussian with constant variance 1/4. For this reason, we specify our model on this transformed scale, and so \( \lambda \) and \( n \) actually represent the square roots of the image intensities and unobserved counts.

In order to segment the image into homogeneous regions, we augment the graph associated with the image to include an array of region identifiers \( r = \{r_{ij}\} \), where \( r_{ij} \) are integer labels denoting the region identification of pixel \( \xi_{ij} \). Also, let \( \mu = \{\mu_K\} \) denote a random vector containing the mean intensity of each region, where \( K \) is also a random variable representing the number of regions contained in the scene. In other words, \( \mu_{r_{ij}} \) is the mean intensity of the pixels in region \( r_{ij} \). Of course, \( r, \mu, \) and \( K \) are estimated simultaneously with \( \lambda \) and \( n \).

The Gibbs prior density is then determined by defining a set of cliques and potential functions on the expanded graph system. The first such clique is defined for each pixel \( \xi_{ij} \) in the graph and is a function of \( \lambda_{ij}, r_{ij}, \) and \( \mu \). The potential function for this clique, \( C_1 \), is defined to be

\[
V_{C_1}(\lambda_{ij}, r_{ij}, \mu) = \frac{1}{2\sigma^2} (\lambda_{ij} - \mu_{r_{ij}})^2.
\]  

(1)

Note that this implies that \( \mu \) is contained in the neighborhood of each pixel in the image. \( \sigma^2 \) is an arbitrary smoothing parameter, which when assigned a small value forces all intensities \( \lambda_{ij} \) within a given region to be arbitrarily close together.

The second clique in the system contains only region identifiers and is pictured in Figure 1. The purpose of this clique is to restrict the minimum size of a pixel region. We consider two such restrictions. In the first, we set the potential of this clique so that regions that contain less than four contiguous pixels are prohibited. Aside from this constraint, the first clique places no restrictions on shape. In the second, we prohibit regions of size less than five, and also impose certain shape constraints.
In Figure 1, suppose that pixel $\xi_{ij}$ has been estimated as being in region one, or that $r_{ij} = 1$. Then the first potential function for $C_2$ that we consider is assigned the value zero unless either (a) none of the region identifiers superscripted with (*) have value one, or (b) none of the region identifiers labeled (a) and less than three identifiers labeled (a) are in region one. Then if either condition (a) or (b) is met, the potential assigned to $C_2$ is some suitably large constant, say $\phi$. For practical purposes, $\phi$ may be set to infinity, provided that the Gibbs sampler (or other deterministic procedures used to draw values from the posterior) are modified so that regions containing four pixels or more are permitted to form simultaneously. As stated, this potential function, which we label $V_{C_2}^1$, has the effect of preventing the formation of regions of less than four contiguous pixels.

The second potential function assigned to $C_2$ is similar to $V_{C_2}^1$. It is defined to be zero whenever there are at least four other identifiers, including at least two (*) sites, that have the same value of $r_{ij}$. It too is assigned an arbitrarily large value if this condition is not satisfied. The effect of this potential function is to prohibit the formation of regions containing less than five pixels, as well as certain other configurations of five pixels or more. For example, identifiers \{$(r_{i-1,j-1}, r_{i-1,j+1}, r_{ij}, r_{i+1,j-1}, r_{i+1,j+1})$\} cannot form their own region since the potentials would be infinite for the cliques centered at all of the sites except $r_{ij}$. Denote this potential function $V_{C_2}^2$.

![Figure 1. Region identity clique $C_2$.](image)

The final clique in the graph is simply the set of all region identifiers. The potential of this clique is defined as $\alpha K$, where $K$ is the number of distinct region identifiers. The effect of this clique is to impose a penalty of $exp(\alpha)$ on the formation of new regions.

Combining the prior density defined by these potentials with the likelihood function on the transformed scale leads to particularly simple forms for the conditional posterior densities of the random variables $\lambda_{ij}$, $r_{ij}$, and $\mu$. Given the transformed count $n_{ij}$ and conditionally on the current values of all other variables in the graph, the posterior density of $\lambda_{ij}$ has a Gaussian distribution with mean and variance given by

$$E(\lambda_{ij}) = \frac{4n_{ij} + \mu r_{ij}/\sigma^2}{4 + 1/\sigma^2}, \quad \text{Var}(\lambda_{ij}) = \frac{1}{4 + 1/\sigma^2}. \quad (2)$$

Likewise, the conditional posterior distribution of $\mu r_{ij}$ given all region identifiers and all intensities $\lambda_{ij}$ is
Gaussian with mean and variance given by

\[
E(\mu_{r_{ij}}) = \sum_{r_{kl} = r_{ij}} \lambda_{kl}/m, \quad \text{Var}(\mu_{ij}) = \sigma^2/m. \tag{3}
\]

Here \(m\) is the number of pixels in region \(r_{ij}\).

The conditional posterior density of a region identifier \(r_{ij}\) given the values of all neighboring variables is similarly straightforward. For each integer \(h \leq K\) define

\[
p_k = \begin{cases} 
\exp \left[ -\frac{(\lambda_{ij} - \mu_k)^2}{2\sigma^2} \right], & \text{if } V_{C_0}^1 \text{ or } V_{C_0}^2 = 0; \\
0, & \text{otherwise.} 
\end{cases} \tag{4}
\]

Then the conditional posterior probability that \(r_{ij}\) takes the value \(k\) is given by \(p_k / \sum_j p_j\) (under the implicit assumption that \(p_{K+1}\) is negligible).

If procedures similar to Gibbs sampling are used to update the region identifiers, a problem arises in the formation of new regions and in the elimination and combination of existing regions. Formally, by keeping \(\phi\) finite, there is a small probability that any given set of pixels will eventually form their own region. However it is not practical to wait for this to happen, so in our updating procedure we systematically identify groups of pixels whose intensities have large deviations from their region means. If the reduction in the likelihood function and potentials \(V_C\), obtainable by forming a new region offset the penalty \(\alpha\), we allow a new region to form. Likewise, if two regions border one another and have region means that are sufficiently close, we compute whether the net energy of the system can be decreased by combining the regions.

With these comments in mind, the estimation algorithm proceeds as described in [8],[9],[11]. Initially, the image is estimated to be one contiguous region with a common mean. One iteration of the EM algorithm is used to update the unobserved counts \(n\). These counts are then used to update \(\lambda\), which in turn are used to update \(\mu\). Finally, the region identifiers \(r\) are updated and a check is made for the elimination and formation of new regions. The updated \(n\) is then returned to the EM step, and the process repeated until the reconstructions converge.

In updating each of these random variables, we impute the modes of the conditional distributions, making the algorithm equivalent to Besag's ICM method [1]. It should be noted that conditionally on \(r\), the conditional posterior distributions of all other variables are concave. Thus, the ICM method converges to a global maximum if the optimal region identifiers are known, but in general converges only to a local maximum.
3. EXAMPLES

We tested our model using simulated PET data from the Hoffmann phantom, using both a low count observation ($\approx 350,000$ counts) and a high count observation ($\approx 3,500,000$ counts). The mean bin intensities were generated by incrementing the bin in each projection plane closest to a given pixel by the intensity of that pixel divided by the number of projection planes. This procedure is quite similar to incrementing each mean bin intensity by a factor proportional to the length of intersection of the bin centerline and the pixel, but avoids costly multiplications (see [7]). To generate the data, a Poisson observation was then simulated for each bin with the given mean intensity. This process does not account for scatter, attenuation, or the imprecise spatial resolution of sensors, and so is not a realistic method for simulating patient data. However, the manner in which the data was simulated makes it ideal for FBP reconstruction, although in more realistic simulations the factors previously mentioned are not easily handled using FBP algorithms. However, in theory account can be taken of such factors in both the EM and Bayesian reconstruction algorithms. The number of bins and projection planes were 108 and 72, respectively.

Figure 2 depicts reconstructions of the low count observation using several different algorithms with varying parameters. Figure 2a illustrates the phantom, Figures 2b and 2c reconstructions obtained using unsmoothed and smoothed FBP, Figure 2d the EM reconstruction after 30 iterations, Figures 2e and 2f reconstructions using our model with potential $V_{\alpha}$, and Figures 2g and 2h reconstructions using potential $V_{\alpha}^2$. Similar illustrations are provided in Figure 3 for the high count data. For the reconstructions obtained using our model, we have provided several images corresponding to different values of $\alpha$, the penalty term for the formation of new regions.

As the pictures indicate, both the EM and the unsmoothed FBP reconstructions appear to preserve indications of subtle, low contrast regions, particularly in the lower portion of each image. However, it is not clear if these regions would be identifiable without having the phantom available or if they would be confused with noise. In comparison to the Bayesian reconstructions, the high intensity regions in both the EM and unsmoothed FBP reconstructions appear less prominently, and the size of these regions are more accurately identified in the Bayesian reconstructions. In our opinion, the Bayesian reconstructions more accurately identify regions of interest, although in high count situations the FBP seems to provide adequate reconstructions. The EM algorithm and the Bayesian reconstructions require essentially the same amount of time, and in our view the Bayesian method provides superior representations of the true scene. It also provides region information that may prove valuable in computer visualization and in quantification.
4. DISCUSSION

A number of aspects concerning the implementation of our model remain unresolved. First, we have assumed that $\sigma^2$ is known a priori. However, this parameter itself may be of interest in some applications, as it represents the within region variation of intensities. It also plays a prominent role in estimating the posterior variance of the region intensities, $\mu$. Currently, we are examining techniques for estimating $\sigma^2$ from projection data and are also considering an extension of the model to include a distinct value of $\sigma^2$ for each region.

Another unexplored feature of the model is its potential for estimating uncertainty in image qualities. Using the Gibbs sampler or other analytic approximations, we hope to obtain indications of the precision to which region sizes and mean intensities are determined. Such results have obvious implications in applications requiring absolute quantification of image attributes.

Finally, we are currently experimenting with our model on actual SPECT data generated from both physical phantoms and patient studies.

ACKNOWLEDGEMENTS

The work of Dr. Bowsher was supported by NIH grant CA46856.

REFERENCES


Figure Captions

Figure 2. Results from estimation of Hoffmann phantom from 350,000 count observation.

a) True scene. b) Reconstruction using FBP with “ramp” filter. c) Reconstruction using FBP with ramp filtering in conjunction with a Hanning window (cutoff frequency equal to 0.7 Nyquist frequency).

d) EM reconstruction after 30 iterations. e) Bayesian reconstruction with $\sigma^2 = 0.18$ and $\alpha = 1.2$ using the potential $V_C^1$. f) Bayesian reconstruction with $\sigma^2 = 0.18$ and $\alpha = 0.6$ using the potential $V_C^1$. g) Bayesian reconstruction with $\sigma^2 = 0.18$ and $\alpha = 1.2$ using the potential $V_C^2$. h) Bayesian reconstruction with $\sigma^2 = 0.18$ and $\alpha = 0.6$ using the potential $V_C^2$.

Figure 3. Results from estimation of Hoffmann phantom from 3,500,000 count observation.

a) True scene. b) Reconstruction using FBP with “ramp” filter. c) Reconstruction using FBP with ramp filtering in conjunction with a Hanning window (cutoff frequency equal to 0.97 Nyquist frequency).

d) EM reconstruction after 30 iterations. e) Bayesian reconstruction with $\sigma^2 = 0.25$ and $\alpha = 1.2$ using the potential $V_C^1$. f) Bayesian reconstruction with $\sigma^2 = 0.25$ and $\alpha = 0.6$ using the potential $V_C^1$. 