Conference Record of the 1992 IEEE Nuclear Science Symposium and Medical Imaging Conference

Segmented Attenuation Correction for PET

Jeffrey A. Fessler*
Division of Nuclear Medicine, University of Michigan

Abstract

We describe a hybrid measured/calculated method for attenuation correction in positron emission tomography (PET). This unified reconstruction/segmentation method is based on a penalized weighted least-squares (PWLS) objective function that is minimized using iterative coordinate-descent (ICD). Two penalty functions are compared; one for a discrete object parameterization, the other for a continuous parameterization. Simulations demonstrate that the methods can reduce the additional emission image variance typically introduced by noisy attenuation correction factors (ACFs).

I. Introduction

Correction for attenuation in PET is essential for both quantitative and visual tasks. In principle, ACFs computed from the ratio of a blank scan to a patient transmission scan should accurately compensate for attenuation. In practice, measurement noise in both scans produce noisy ACFs, which propagates additional variance into the reconstructed emission images.

The conventional solution to this noise problem is to acquire disproportionately lengthy transmission scans and to smooth the blank and transmission scans. Indiscriminate smoothing can produce systematic errors in the emission image. Nevertheless, some sort of smoothing is both necessary and desirable. To illustrate, consider a high-resolution system such as a CTI 931 with its 50,000 lines of coincidence. The transmission coincidence events are divided among this large number of rays, so the relative accuracy of each measurement is low. Since it is unlikely that the attenuation map requires 50,000 degrees of freedom to describe adequately, there must be redundancy in the measurements. Linear smoothing is one simple but suboptimal attempt to exploit this redundancy. The reconstruction-reprojection method is a somewhat more sophisticated approach that also reduces the degrees of freedom. (If the attenuation map is reconstructed within a circle in a 128^2 image, then there can be at most about 13,000 degrees of freedom in the reprojected ACFs.) One hopes to reduce the noise by eliminating degrees of freedom without inducing the systematic biases inherent to linear smoothing.

The method proposed in this paper reduces the degrees of freedom by exploiting two properties of attenuation maps: (i) they are composed of a relatively small number of fairly homogeneous tissue classes, such as air, lung, soft tissue, and bone, and (ii) except at boundaries between tissues, neighboring voxels tend to be composed of the same tissue class.

These properties have been used implicitly and explicitly by several investigators, almost exclusively through a *sequential* three-step process: (1) reconstruct using FBP an attenuation map from the logarithm of the ACFs, (2) process that attenuation map, and (3) reproject the attenuation map to form new ACFs. Huang *et al.* laid the groundwork for these methods by demonstrating the efficacy of segmenting an attenuation map into discrete classes. However, the manual adjustments of that segmentation method are impractical for routine use.

Recent attempts to automate the segmentation still suffer from the fundamental limitation of such sequential approaches: FBP produces streak artifacts when applied to low-count transmission data. Therefore, we propose a *unified* reconstruction/segmentation method that iteratively estimates a segmented attenuation map directly from the transmission data. This method makes better use of the statistical information in the transmission measurements.

In order to explore the bias-variance tradeoffs of these methods, we have performed a preliminary investigation using two extremes for the attenuation parameterization. One method is based on a discrete parameterization in which each voxel takes one of K tissue types, i.e., a literal interpretation of property (i) above. The other method allows each voxel to vary over the continuous set of nonnegative attenuation coefficients, thereby accommodating partial volume effects, but not making use of property (i). Understanding the limitations of these two extremes should facilitate designing intermediate methods that exploit property (i) while allowing for partial volume effects.

II. Theory

A statistical approach to image reconstruction requires five components: (i) a finite parameterization of the object (attenuation map), (ii) a system model that relates the attenuation map to ideal measurement values, (iii) a statistical model that describes how the actual measurements vary about their ideal values, (iv) an objective function that is to be maximized to estimate the attenuation map, and (v) an algorithm, typically iterative, for maximizing the objective function, including an initial estimate and stopping criterion. This section describes the reconstruction/segmentation methods in terms of these components.

*This work was supported in part by a DOE Alexander Hollaender Postdoctoral Fellowship.

A. Object Parameterization

Let $\mu(\underline{x})$ denote the spatial distribution of attenuation coefficients within the patient. We assume this distribution can be approximately decomposed into rectangular voxels:

$$\mu(\underline{x}) \approx \sum_j \mu_j I_j(\underline{x}), \tag{1}$$

where μ_j denotes the mean attenuation coefficient in the jth voxel, and $I_j(\underline{x})$ is the indicator function with the jth voxel as its support. For this approximation to be valid, the voxels must be sufficiently small. An natural generalization of this parameterization to account for inhomogeneity within voxels would be to adopt a "mixel" model.

B. System Model

For a transmission scan, the patient is surrounded by a ring containing a positron-emitting radioisotope. If the ring emits a pair of annihilation photons along a chord that intersects the ith detector pair, then a coincidence event occurs provided that both photons escape unabsorbed. If the detectors are small relative to attenuation map inhomogeneities, then this survival probability is approximately

$$\alpha_i \approx \exp(-l_i), \tag{2}$$

where l_i is the effective attenuation path length along the chord between the ith pair of detectors. We assume this length can be expressed:

$$l_i = \int_{S_i} \mu(\underline{x})\ d\underline{x} \tag{3}$$

where S_i denotes the strip integral over the ith detector chord. Other authors have used line integrals rather than strip integrals, but the finite width of PET detectors makes strip integrals more plausible. An even more accurate system model would account for noncollinearity of the annihilation photon pairs and for inter-crystal mispositioning errors in block detectors.

Under the discretization (1), we can rewrite (3) as

$$l_i = \sum_j a_{ij}\mu_j, \tag{4}$$

where

$$a_{ij} = \int_{S_i} I_j(\underline{x})\ d\underline{x}. \tag{5}$$

We precompute and store the nonzero elements of $\mathbf{A} = \{a_{ij}\}$ using the ASPIRE software library, which significantly reduces the reconstruction time.

C. Statistical Model

Ideally, the statistical model describes the distribution of each measurement about its mean, and consequently determines a measure of similarity between the actual measurements and the calculated projections of image estimates from (4). The CTI 931 at our institution uses real-time subtraction of delayed-window coincidences to correct for accidental coincidence (AC) events in its routine operation. This subtraction process produces negative sinogram bins that clearly invalidate the usual Poisson model. We present a more realistic Poisson-difference model below. Because the likelihood for this Poisson-difference model is intractable, we propose a simple weighted, transformed, least-squares similarity measure.

Let B_i denote the ith detector measurement for the blank scan. For a system employing real-time delayed-window AC event correction, the following statistical model is reasonable:

$$B_i \sim \text{Poisson}\{\tau^b(b_i + r_i^b)\} - \text{Poisson}\{\tau^b r_i^b\},$$

where τ^b denotes the time duration of the blank scan, b_i denotes the product of detector efficiency and photon flux on the ith detector pair, and r_i^b denotes the AC event rate during the blank scan (typically very small). Similarly, after the patient is put into the PET scanner, a reasonable model for the transmission scan measurement is:

$$T_i \sim \text{Poisson}\{\tau^t(b_i\alpha_i(1+s_i) + r_i^t)\} - \text{Poisson}\{\tau^t r_i^t\}$$

where τ^t denotes the time duration of the transmission scan, α_i was defined by (2), and s_i denotes the fractional increase in direct coincidence events due to scatter within the patient.

Unfortunately the individual AC event rates r_i^b and r_i^t are not available, which appears to preclude using an exact likelihood as a similarity measure. For this reason, and for computational simplicity, we propose to forgo a likelihood approach in favor of a quadratic approach based on the first two moments. The result is an approximate similarity measure that nevertheless performs remarkably well.

The first step is to apply a logarithmic transformation. Let $\{b_i\}$ and $\{t_i\}$ denote the measured realizations of the random variables $\{B_i\}$ and $\{T_i\}$. Then a (noisy) estimate of l_i, the strip integral of attenuation between the ith detector pair, is given by

$$y_i = \begin{cases} \log\left(\frac{b_i}{\tau^b}\right) - \log\left(\frac{t_i}{\tau^t}\right) + \log(1+\hat{s}_i), & t_i > 0,\ b_i > 0 \\ 0, & \text{otherwise} \end{cases}$$

where $\hat{s}_i$ is an estimate of the scatter fraction. For simplicity, in the remainder of this paper we assume $s_i = 0$. Using a Taylor's expansion for the logarithmic transformation of a Poisson variate, one can show that

$$\sigma_i^{-2} = \begin{cases} (b_i t_i)/(b_i + t_i), & t_i > 0,\ b_i > 0 \\ 0, & \text{otherwise} \end{cases}$$

is an estimate of the Fisher information (1/variance) of y_i. Let $\boldsymbol{\mu} = [\mu_1, \mu_2, \ldots]'$ denote the vector of unknown attenuation coefficients. We propose to use the following weighted least-squares (WLS) similarity measure:

$$D(\boldsymbol{\mu}) = (\mathbf{y} - \mathbf{A}\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{A}\boldsymbol{\mu}) \tag{6}$$

where $\mathbf{y} = [y_1, y_2, \ldots]'$, and $\boldsymbol{\Sigma}^{-1}$ is a diagonal matrix with elements σ^{-2}. The similarity measure proposed by Sauer is similar, except that we have included the blank scan variance b_i. That variance can be non-negligible for poor efficiency detectors.

D. Objective Function

Although one could minimize (6) to estimate $\boldsymbol{\mu}$ from the transformed data $\mathbf{y}$, it is well known that objective functions based solely on similarity measures such as (6) perform poorly due to the ill-conditioned nature of tomographic reconstruction. To remedy this problem, we have investigated two regularization methods. Both methods exploit the spatial correlation of attenuation maps through smoothness penalties, but they are based on different parameterizations.

The first method is based on a discrete parameterization of the attenuation map, i.e., we reduce the domain of $\boldsymbol{\mu}$. Assume that there are K classes of attenuation coefficients with nominal values $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_K]'$. We assume that μ_j, the attenuation coefficient in the jth voxel, is one of the K values $\{\theta_1, \ldots, \theta_K\}$. Let x_j indicate the class of the jth voxel, i.e., $\mu_j = \theta_{x_j}$, where x_j takes values $1, 2, \ldots, K$. An explicit notation for this parameterization is $\boldsymbol{\mu}(\mathbf{x}, \boldsymbol{\theta})$ where

$$\mu_j(\mathbf{x}, \boldsymbol{\theta}) = \theta_{x_j}. \tag{7}$$

We assume K is known, although information theoretic approaches can in principle be used to determine K.

To exploit the second property, we use a penalty function that encourages neighboring pixels to be of the same class. Specifically,

$$R_1(\mathbf{x}) = \frac{1}{2} \sum_j \sum_{j' \in N_j} w_{jj'} 1_{\{x_j \neq x_{j'}\}}, \tag{8}$$

where N_j is the set of eight neighbors of the j pixel. The weights w_{jk} equal 1 for horizontal and vertical neighbors, and $1/\sqrt{2}$ for diagonal neighbors. This type of penalty function is often presented as a Bayesian "prior" for the ensemble of voxel classes. We do not adopt that philosophy here since we have no evidence that the Gibb's distribution corresponding to (8) resembles the ensemble statistics of attenuation maps.

In the presence of object-dependent scatter, etc., one may wish to simultaneously estimate the attenuation coefficients $\boldsymbol{\theta}$ of the K classes from $\mathbf{y}$. A penalty function for this parameter estimation is also useful:

$$R_2(\boldsymbol{\theta}) = \sum_{k=1}^{K} w_k (\theta_k - \bar{\theta}_k)^2,$$

where $\bar{\theta}_k$ denotes the ideal attenuation coefficients and w_k denotes weights that reflect their uncertainties.

Combining the WLS similarity measure with the above penalty functions yields the following objective function:

$$\Phi_u(\mathbf{x}, \boldsymbol{\theta}) = D(\boldsymbol{\mu}(\mathbf{x}; \boldsymbol{\theta})) + \beta R_1(\mathbf{x}) + R_2(\boldsymbol{\theta}), \tag{9}$$

where β controls the influence of the smoothness penalty, in analogy with the filter window that must be chosen for FBP reconstruction.

Having defined this objective, our goal is to estimate $\mathbf{x}$ and $\boldsymbol{\theta}$ from $\mathbf{y}$:

$$(\hat{\mathbf{x}}, \hat{\boldsymbol{\theta}}) = \arg\min_{\mathbf{x}, \boldsymbol{\theta}} \Phi_u(\mathbf{x}, \boldsymbol{\theta}).$$

The second method is based on a continuous parameterization for $\boldsymbol{\mu}$, and is described by the following objective function:

$$\Phi_c(\mathbf{x}, \boldsymbol{\theta}) = D(\boldsymbol{\mu}) + \beta \frac{1}{2} \sum_j \sum_{j' \in N_j} w_{jj'} (\mu_j - \mu_{j'})^2, \tag{10}$$

where the weights $w_{jj'}$ are defined as in (8).

E. Iterative Algorithm

The objective functions Φ_u and Φ_c are minimized by iterative coordinate-descent (ICD). The ICD algorithm updates each image parameter individually by minimizing the objective function over that parameter while holding the other parameters fixed. Nonnegativity is easily enforced in this paradigm. The algorithm is globally convergent for the continuous parameterization, but only locally convergent for the discrete parameterization. We have had best results when initializing the discrete algorithm using the results from the following image-domain segmentation method.

III. Sequential method

If the transmission counts are high enough, a sequential reconstruction/segmentation method may be adequate. One approach to this is the following. First perform a FBP reconstruction of the attenuation image from $\mathbf{y}$. Let $\mathbf{f}$ denote this (noisy) image. The attenuation image $\mathbf{f}$ is then segmented to form an attenuation map $\boldsymbol{\mu}$ in image space using the following objective function:

$$\Phi_s(\mathbf{x}, \boldsymbol{\theta}) = \frac{1}{2} \|\mathbf{f} - \boldsymbol{\mu}(\mathbf{x}; \boldsymbol{\theta})\|^2 + \beta R_1(\mathbf{x}) + R_2(\boldsymbol{\theta}), \tag{11}$$

where R_1 and R_2 were defined above. This objective function is also minimized using ICD. Since the system matrix $\mathbf{A}$ is not used, the sequential method requires less computation.

IV. Epilogue

The above sections outline the methods presented at the conference. Due to the limited space, full details, including results and bibliography, will be presented in a separate manuscript. The author gratefully acknowledges helpful discussions with Neal Clinthorne, Ken Koral, and Les Rogers.