Maximum Likelihood Methods for Reconstructing an Image In a Region-of-Interest for Transmission Tomography

Donald L. Snyder,¹ Joseph A. O’Sullivan,¹ Ryan J. Murphy,¹ David G. Politte,¹ Bruce R. Whiting,¹ and Jeffrey F. Williamson²

Abstract—Methods are described for forming a maximum-likelihood estimate of the attenuation density in a patient or object for transmission tomography in a region of interest. The methods converge monotonically to the same limit point.

I. INTRODUCTION

A desire to limit radiation exposure to only a region-of-interest motivates consideration of algorithms for reconstructing images from transmission tomographic data when the data are incomplete. A full sinogram of data includes projection lines at all orientations for every attenuating image element that affects any portion of the data. Portions of a full sinogram are missing when only a region of interest within a patient or object is scanned. For example, the missing portion can correspond to parts of a large patient that lie outside the fan-beam radiation pattern of the CT scanner, or it can correspond a high density metal object that suppresses some projection rays.

This paper is a summary of [24] in which derivations and illustrations via synthetic and real data are given for the algorithms stated herein.

Image reconstruction based on filtered backprojection is not local to a region-of-interest because the value of a reconstructed image at any site is affected not only by projection rays that pass through or near that site but also by rays that pass far away. Various methods for reconstructing an image within a region-of-interest by using only rays that are local to that region appear in the literature. Methods based on the use of wavelets for a local parametric representation of an image are developed by, among others, Berenstein and Walnut [1], A. Yagle [30], B. Sahiner and A. Yagle [20], Rashid-Farrukhi, Liu, Berenstein, and Walnut [19], and P. Das and C. Sastry [2]. Other methods described, for example, by A. Faridani, F. Keinert, F. Natterer, E. Ritman, and K. Smith [5], A. Faridani, E. Ritman, and K. Smith [6], and A. Ramm and A. Katsevich [18], reconstruct discontinuities of the image in a region-of-interest using rays that are local to that region. The need to compensate for objects with nonzero attenuation values affecting projection data but lying outside the scanner’s field of view has long been appreciated [7], [15]. B. Ohnesorge, T. Flohr, K. Schwarz, J. Heiken, and K. Bae [16] propose a “symmetric extrapolation” approach for extending projection data at scan boundaries to compensate for projections that are missing beyond the boundaries. F. Natterer [13] and F. Natterer and F. Wubbeling [14] discuss many mathematical aspects of forming reconstructions from limited data, such as uniqueness and stability of solutions. P. Seitz and P. Rüegsegger [22] also summarize known uniqueness results for reconstructions from incomplete data. Iterative reconstruction methods have also been developed. One such method is given by P. Seitz and P. Rüegsegger [21] using repeated modifications of real and synthesized projection data to account for known characteristics of metal prostheses. M. Nassi, W. Brody, B. Medoff, and A. Macovski [12] and B. Medoff [11] use a linear, least-squares formulation in formulating an iterative reconstruction method. Another iterative method is given by G. Wang, D. Snyder, and M. Vannier [25] and G. Wang, M. Vannier, and P. Cheng [26] using an information discrepancy framework.

Most methods cited for image reconstruction within a region-of-interest start with a linear model to describe projection data. One exception is the iterative method described by P. Seitz and P. Rüegsegger [21] in which beam hardening corrections are made for metal prostheses having known characteristics. A linear model is an accurate representation in some, but not all, circumstances that are encountered in practice. We will describe two methods for reconstructing images in a region-of-interest using available projection rays when the tomographic data are modeled as a Poisson process with a mean-value function that is dependent on ray integrals through Beer’s law, assuming monochromatic radiation. This is a simplified model to illustrate the ideas. It can be extended readily to a more accurate model that accounts for polychromatic source and attenuation effects, scattered photons, and energy-integrating detectors.

II. DATA MODEL

It is common to model data acquired by a transmission tomograph via a Poisson process [4], [8]. Let \(d(y_m)\) denote the data acquired at the \(m\)th source-detector position among the \(M\) such positions that are present as the scanner acquires a set of data for forming a tomographic image. The data \(\{d(y_m) : m = 1, 2, \ldots, M\}\) are mutually independent and Poisson distributed.

\[
\Pr \left[ d(y) = n \right] = \frac{q^n(y : c)}{n!} e^{-q(y : c)}, \tag{1}
\]

¹Washington University, St. Louis, MO. ²Virginia Commonwealth University, Richmond, VA
for $n = 0, 1, 2, \cdots$, and $y \in \mathcal{Y} = \{y_m : m = 1, 2, \cdots, M\}$, in which

$$q(y : c) = I_0(y) \exp \left[ - \sum_{x \in \mathcal{X}} h(y | x) c(x) \right], \tag{2}$$

where $I_0(y)$ is the source flux, $h(y | x)$ is the scanner’s point spread function, $\mathcal{Y}$ is the discretized measurement or sinogram space, and $\mathcal{X}$ is the discretized image space. The parameterization of $y \in \mathcal{Y}$ depends on the scanner’s geometry. For example, in the fan-beam geometry indicated in Fig. 1, $y = (\beta, \gamma)$, where $\beta$ is the angle of the source in its rotation about the body, and $\gamma$ is the angle of a ray within the fan, with the angular extent of the fan being $-\gamma_m \leq \gamma \leq \gamma_m$. Other parameterizations of $y$ can accommodate a three-dimensional spiral geometry and multirow detectors.

The source rotates on a circular path of radius $D$ from the system isocenter. Projection data are acquired in the sinogram space $\mathcal{Y} = \{\beta, \gamma : 0 \leq \beta < 2\pi, -\gamma_m \leq \gamma \leq \gamma_m\}$. The usual reconstruction problem is to process these data to estimate the attenuation function $c(\cdot)$ in the “circle of reconstruction” centered at the system isocenter with radius $D \sin \gamma_m$. Reconstruction algorithms that produce solutions to this problem generally perform reconstructions on a square of dimension $2D \sin \gamma_m \times 2D \sin \gamma_m$ that circumscribes the circle of reconstruction. While there are a number of such reconstruction algorithms appearing in the literature (see, for example, [4], [8], and references cited therein), the approach we will describe is most closely related to the one developed by J. O’Sullivan [17] for solving the minimization problem

$$\hat{c}(x) = \arg \min_{c \in \mathcal{C}} I(d(y), y \in \mathcal{Y} | q(y : c), y \in \mathcal{Y}), \quad x \in X, \tag{3}$$

where $\mathcal{C}$ is the set of admissible functions, $\mathcal{C} = \{c : c(x) \geq 0, x \in \mathcal{X}\}$, and

$$I(d(y), y \in \mathcal{Y} | q(y : c), y \in \mathcal{Y}) = \sum_{y \in \mathcal{Y}} d(y) \ln \left[ \frac{d(y)}{q(y : c)} \right] - d(y) + q(y : c) \tag{4}$$

is Csiszár’s I-divergence. This is equivalent to maximizing the loglikelihood functional of the Poisson distributed data $\{d(y), y \in \mathcal{Y}\}$. By expressing the I-divergence as the solution to a variational problem and using an alternating minimization method, J. O’Sullivan [17] identifies a sequence of estimates defined by

$$\hat{c}^{(k+1)}(x) = \max \left[ \hat{c}^{(k+1)}_{AM}(x), 0 \right], \tag{5}$$

where

$$\hat{c}^{(k+1)}_{AM}(x) = \hat{c}^{(k)}(x) - \frac{1}{Z(x)} \ln \left[ \sum_{y \in \mathcal{Y}} \frac{h(y | x) d(y)}{h(y | x) q(y : \hat{c}^{(k)})} \right],$$

for a simpler presentation, we have specialized the results in [17] to a monoenergetic source. The extension to include polyenergetic sources and the accompanying beam-hardening effects is a straightforward modification of the present development except for a more complicated notation.

$x \in \mathcal{X}$, that produces a convergent, nonincreasing (resp. non-decreasing) sequence of I-divergence (resp. likelihood) values when, for all $y \in \mathcal{Y}$, $Z(x)$ satisfies $\sum_{x \in \mathcal{X}} h(y | x) / Z(x) \leq 1$ but is otherwise arbitrary, with the choice affecting the rate of convergence.

Assumptions are made in the development of (5) about the attenuation function to be reconstructed and about the projection data for performing the reconstruction. These assumptions are not unique to (5) but, rather, are commonly made in the development of reconstruction approaches. They can be violated in practice, and they often are. First, it is usually assumed that image space $\mathcal{X}$ adopted for the reconstruction is large enough to support all nonzero attenuation values that affect the projection data. The patient bed and large patients, however, often lie beyond this support. Secondly, it is usually assumed that the projection data are complete in the sense that projection rays are present for all directions for all nonzero attenuation values that affect any of the rays. This assumption is often invalidated by the patient bed, large patients, high density metal objects, or limitations on view directions.

Let $\mathcal{X}$ denote an image space that is large enough to encompass all nonzero attenuation values that affect projection data, and let $\mathcal{Y}$ denote a complete sinogram space that supports projection rays that would be obtained by scanning each element of $\mathcal{X}$ from all directions. The subset of $\mathcal{Y}$ that supports the actual data, hereafter called the incomplete data, is denoted by $\mathcal{Y}_{inc}$. These incomplete data are the data acquired by the tomograph and may have missing rays. Also, let $\mathcal{Y}_{miss} = \mathcal{Y} \setminus \mathcal{Y}_{inc}$ denote the subset of $\mathcal{Y}$ that supports the missing projection rays. The image reconstruction problem is to estimate the attenuation function $\{c(x) : x \in \mathcal{X}\}$ over the entire image space $\mathcal{X}$ in terms of the incomplete projection data supported by $\mathcal{Y}_{inc}$.

The region, $\mathcal{Y}_{miss}$, that supports missing data can be identified in various ways. Here are two examples. The first is to use a mask to identify a convex region-of-interest in the image, $\mathcal{X}_{ROI} \subseteq \mathcal{X}$, as described by G. Wang, et al. [25]. Let the indicator function, $I_{ROI}(x)$, be defined by

$$I_{ROI}(x) = \begin{cases} 1, & x \in \mathcal{X}_{ROI} \\ 0, & x \in \mathcal{X} \setminus \mathcal{X}_{ROI} \end{cases}, \tag{6}$$

and define the projection, $P_{ROI}(y)$, of this function by

$$P_{ROI}(y) = \sum_{x \in \mathcal{X}} h(y | x) I_{ROI}(x), \quad y \in \mathcal{Y}. \tag{7}$$

Then, local projection rays to be acquired and used in the reconstruction are supported by the subset $\mathcal{Y}_{inc}$ of $\mathcal{Y}$ defined by

$$\mathcal{Y}_{inc} = \{y : y \in \mathcal{Y}, P_{ROI}(y) > 0\}, \tag{8}$$

and $\mathcal{Y}_{miss} = \mathcal{Y} \setminus \mathcal{Y}_{inc}$. For the second example, suppose that the angular extent of the fan in a fan-beam geometry, $-\gamma_m \leq \gamma \leq \gamma_m$, is insufficient to encompass all nonzero attenuation values for all source angles, as can occur with large patients and the patient bed. Then,

$$\mathcal{Y}_{miss} = \{\beta, \gamma : 0 \leq \beta < 2\pi, -\gamma_m \leq \gamma < \gamma_m\} \cup \{\beta, \gamma : 0 \leq \beta < 2\pi, \gamma_m < \gamma \leq \gamma_M\}, \tag{9}$$
where a fan of extent $-\gamma_M \leq \gamma \leq \gamma_M$ would be sufficient to include all relevant attenuation values. Note, also for this second example, that the image space $\mathcal{X}$ would typically, although not necessarily, be a square of dimension $2D \sin \gamma_M \times 2D \sin \gamma_M$ rather than one of dimension $2D \sin \gamma_m \times 2D \sin \gamma_m$ that is usually implemented on tomographs that perform filtered back-projection reconstructions.

III. Reconstruction When Projection Rays Are Missing

An image reconstructed from projection rays supported by $\mathcal{Y}_{\text{inc}} \subseteq \mathcal{Y}$ is defined over the complete image space $\mathcal{X}$, even though image values over only a region-of-interest subset of $\mathcal{X}$ may be of interest. We identify two methods of reconstruction depending upon how the missing projection rays are treated. In the first, rays supported by $\mathcal{Y}_{\text{miss}}$ are ignored, and in the second they are estimated.

A. Method 1: Missing Rays Ignored

The reconstruction problem is to solve the minimization

$$\hat{c}_1 (x) = \arg \min_{c \in \mathcal{C}} I [d (y), y \in \mathcal{Y}_{\text{inc}} | q (y : c), y \in \mathcal{Y}_{\text{inc}}],$$

(10)

for $x \in \mathcal{X}$, which is analogous to (3) but with the sinogram space restricted to $\mathcal{Y}_{\text{inc}}$. By following the derivation of (5) by J. O’Sullivan [17] and simply replacing $\mathcal{Y}$ with $\mathcal{Y}_{\text{inc}}$ in that derivation, it is shown in [24] that the sequence of estimates defined by

$$\hat{c}^{(k+1)}_{\text{AM1}} (x) = \max \left[ \hat{c}^{(k+1)}_{\text{AM1}} (x), 0 \right],$$

(11)

where

$$\hat{c}^{(k+1)}_{\text{AM1}} (x)$$

$$= \hat{c}^{(k)} (x) - \frac{1}{Z(x)} \ln \left[ \sum_{y \in \mathcal{Y}_{\text{inc}}} \frac{h(y|x) d(y)}{\sum_{y \in \mathcal{Y}_{\text{inc}}} h(y|x) q(y: \hat{e}^{(k)}_2)} \right],$$

$x \in \mathcal{X}$, produces a convergent, nonincreasing sequence of $I$-divergence values when $Z(x)$ is selected so that for all $y \in \mathcal{Y}_{\text{inc}}, \sum_{x \in \mathcal{X}} h(y|x) / Z(x) \leq 1$.

B. Method 2: Missing Rays Estimated

An alternative iterative reconstruction method results when the rays supported by $\mathcal{Y}_{\text{miss}}$ are not ignored but, rather, are estimated. The reconstruction problem is to solve the minimization

$$\hat{c}_2 (x) = \arg \min_{c \in \mathcal{C}} I [p (y), y \in \mathcal{Y} | q (y : c), y \in \mathcal{Y}],$$

(12)

for $x \in \mathcal{X}$, subject to the additional constraint that $p (y) = d (y)$ for $y \in \mathcal{Y}_{\text{inc}}$. For $y \in \mathcal{Y}_{\text{miss}}$, $p (y)$ must be nonnegative but is otherwise arbitrary. We prove in [24] that the sequence of estimates

$$\hat{c}^{(k+1)}_{\text{AM2}} (x) = \max \left[ \hat{c}^{(k+1)}_{\text{AM2}} (x), 0 \right],$$

(13)

where

$$\hat{c}^{(k+1)}_{\text{AM2}} (x) = \hat{c}^{(k)} (x) - \frac{1}{Z(x)} \ln \left[ \frac{\sum_{y \in \mathcal{Y}_{\text{inc}}} h(y|x) d(y) + \sum_{y \in \mathcal{Y}_{\text{miss}}} h(y|x) q(y: \hat{e}^{(k)}_2)}{\sum_{y \in \mathcal{Y}} h(y|x) q(y: \hat{e}^{(k)}_2)} \right],$$

$x \in \mathcal{X}$, produces a convergent, nonincreasing sequence of $I$-divergence values when for all $y \in \mathcal{Y}, Z(x)$ is selected so that $\sum_{x \in \mathcal{X}} h(y|x) / Z(x) \leq 1$.

IV. Conclusions

Our preliminary experiments indicate that the two maximum-likelihood methods we have described for transmission tomography do result in a reduction of artifacts caused by incomplete data. One conclusion is that it is important to perform reconstructions on an image space that is sufficiently large to support all objects having a nonzero attenuation that affect the data. This implies, for example, that performing reconstructions on an expanded image space for CT data acquired with large patients having portions of their body outside the scanner field of view does yield images with reduced artifacts.

V. Acknowledgments

This work was supported in part by the National Institutes of Health under research grant R01CA75371 from the National Cancer Institute of the National Institutes of Health, (J. F. Williamson, P. I.), the Graduate Education and Research Partnership with The Boeing Foundation, grant 34566C, and the Whitaker Foundation.

REFERENCES


