Maximal Entropy for Reconstruction of Back Projection Images

Tryphon Georgiou
Department of Electrical and Computer Engineering
University of Minnesota
Minneapolis, MN 5545

Peter J. Olver
Department of Mathematics
University of Minnesota
Minneapolis, MN 55455

Allen Tannenbaum
Depts. of Electrical & Computer and Biomedical Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250

February 24, 2003

This work was supported in part by grants from the National Science Foundation, Air Force Office of Scientific Research, and the Army Research Office.

Abstract

Maximum entropy methods have proven to be a powerful tool for reconstructing data from incomplete measurements or in the presence of noise. In this note, we apply the method to the reconstruction computed tomography data derived from backprojection over a finite set of angles. In this case, one derives quite simple formulae which may be easily implemented on computer.

1 Introduction

Computed tomography (CT) is the reconstruction of an image (2D or 3D) from its line of plane integrals. It has been an essential method in diagnostic radiology, and, with the advent of faster scanners of higher resolution, is becoming very important in image guided surgery and therapy as well.

One of the key examples of this technique is when a given cross-section of the body is scanned by an X-ray beam. The intensity loss (which is tissue dependent) is then recorded by a detector, and then computer processed to produce a two dimensional image. There are various possible geometries for the scanners [3], which is important in CT imaging but which we will ignore in this note. The point is that under generic conditions, one can exactly reconstruct a 2D image from its 1D line integrals. The problem is of course that in practice one does not have infinite number of 1D projections, but only a finite number in any given scan. Hence the problem becomes how to find the "best" reconstruction in some suitable sense of the image from such a finite set.

In this note, we propose the use of maximum entropy. Maximum entropy methods have proven to be very important for the reconstruction of data from incomplete measurements or in the presence of noise. For a very nice survey of such results see the paper [1]. Here we apply maximum entropy methods to the problem of reconstruction of images in computed tomography from a finite set of angles. As we will show, in a number of key cases, one can derive some exact formulas for the maximal entropy solution in this framework.

We now summarize the contents of this paper. In Section 2, we outline the theory of the Radon transform, and show how it may be used for image reconstruction. In Section 3, we discuss the methodology of maximal entropy with constraints. Then in Section 4, we give our formulae for optimal reconstruction using a finite set of angles in the maximal entropy sense. Here consider the cases of a continous density over the image domain, continuous densities for each pixel, and then the discrete case (sampled image and quantized density function). Finally, in Section 5, we sketch some directions for future research on image reconstruction for computerized tomography.

2 The Radon Transform and Image Reconstruction

We briefly discuss the Radon transform and its relationship to image reconstruction. The basic problem in CT (computed tomography) is how to reconstruct a 2D image from a set of 1D projections taken along various lines through the image.

More precisely, for $\rho(x, y)$ the intensity map of 2D grey-level image, we consider the integral along a line ℓ_θ which is distance s from the origin and makes an angle θ with the x-axis:

$$\begin{aligned} g(\theta, s) &= \int_{\ell_\theta} \rho(x, y) d\ell_\theta \qquad (1) \\ &= \int \rho(x, y)\delta(x \sin\theta - y\cos\theta - s)\, dx\, dy. \end{aligned}$$

This is precisely the *Radon transform.*

This leads to a simple algorithm for image reconstruction via the so-called *Fourier Slice Theorem.* Namely, it is easy to show that the 1D Fourier transorm of function $g(\theta, s)$ is the 2D Fourier transform of the intensity function $\rho(x, y)$. Thus using the the inverse Fourier transform we can reconstruct the image. This leads to a backprojection filter

$$Q(\theta, \omega) = \det J(\omega)\ G(\theta, \omega)$$

where $J(\omega)$ is the Jacobian of the change of coordinates from rectangular to polar, and $G(\theta, \omega)$ is the 2D Fourier transform of $\rho(x, y)$ evaluated at $(\omega \sin\theta, -\omega\cos\theta)$. We then have the filtered backprojection formula:

$$\rho(x, y) = \frac{1}{4\pi^2} \int Q(\theta, \omega) \exp(i\omega(x \sin\theta - y\cos\theta))\, d\omega\, d\theta.$$

Clearly, if one can compute the Radon transform over all angles θ one can reconstruct the image. In practice of course one can only making the computation only a finite sample of angles. The question we now address is what is the “best” backprojection reconstruction over such a finite sample? In the next sections, we give a notion of best reconstruction in information theoretic terms using the notion of maximal entropy.

3 Maximal Entropy with Constraints

We minimize the functional

$$\int \rho(x,y) \log \rho(x,y)\, dx\, dy, \quad \int \rho(x,y)\, dx\, dy = 1, \quad \rho(x,y) \geq 0, \tag{2}$$

subject to

$$g(\theta, s) = \int_{\ell_\theta} \rho(x,y) d\ell_\theta = \int \rho(x,y)\ \delta(x\cos\theta - y\sin\theta - s)\, dx\, dy.$$

This is equivalent to maximizing the *entropy functional*

$$-\int \rho(x,y)\ \log \rho(x,y)\, dx\, dy.$$

Accordingly we use the method of Lagrange multipliers. Here $\rho(x,y)$ is defined on some subdomain $\Omega \subset \mathbf{R}^2$ (the image domain), which we may take without loss of generality to be $\mathbf{R}^2$.

We now define the integral operator from L^2 to L^2 by

$$A[\rho](\theta, s) := \int \rho(x,y)\ \delta(x\cos\theta - y\sin\theta - s)\, dx\, dy.$$

Notice the with respect to the standard innner product on L^2 we can compute the adjoint operator as follows: For $\lambda(\theta, s) \in L^2$

$$\begin{aligned}\langle A[\rho], \lambda\rangle &= \langle \rho, A^*[\lambda]\rangle \\ &= \int \lambda \Big(\theta, s)(\int \lambda(\theta,s)\ \rho(x,y)\ \delta(x\cos\theta - y\sin\theta - s)\, dx\, dy\Big)\, d\theta\, ds \\ &= \int \rho(x,y)\ \Big(\int \lambda(\theta,s)\ \delta(x\cos\theta - y\sin\theta - s)\, d\theta\, ds\Big)\, dx\, dy\end{aligned}$$

from which we see that

$$A^*[\lambda](x,y) \ := \ \int \lambda(\theta,s)\ \delta(x\cos\theta - y\sin\theta - s)\, d\theta\, ds.$$

So we introduce the Lagrange multiplier $\lambda(\theta, s)$ and consider the minimization of

$$\int \rho(x,y)\ \log \rho(x,y)\, dx\, dy - \langle \lambda, A[\rho] - g\rangle \tag{3}$$

$$= \int \rho(x,y)\ \log \rho(x,y)\, dx\, dy - \langle \lambda, A[\rho]\rangle - \langle \lambda, g\rangle \tag{4}$$

$$= \int \rho(x,y)\ \log \rho(x,y)\, dx\, dy - \langle A^*[\lambda], \rho\rangle - \langle \lambda, g\rangle. \tag{5}$$

Taking the first variation with respect to ρ yields at a critical point,

$$1 + \log \rho(x, y) = A^*[\lambda](x, y). \tag{6}$$

This is the key relationship for which we will give an explicit solution in the finite dimensional case. Notice that (6) implies that for each $v \in \ker A$,

$$\langle 1 + \log \rho, v \rangle = 0,$$

and so

$$\int \log \rho(x, y) v(x, y)\, dx\, dy = -\int v(x, y)\, dx\, dy.$$

4 Optimal Reconstruction Using Finite Set of Angles

We consider the problem of best reconstructing an image in the maximal entropy sense using finite set of directions. We work out explicitly the cases for a continous density over the image domain, continuous densities for each pixel, and, finally, the discrete case with sampled image and quantized density function.

4.1 Continuous Density over the Image Domain

We can use the equation (6)

$$1 + \log \rho(x, y) = A^*[\lambda](x, y)$$

to the case of continuous density and finite set of sample angles $\theta_1, \ldots, \theta_n$. In this case, it is easy to see that

$$\begin{aligned} A^*[\lambda](x, y) &= \sum_{i=1}^{n} \int \lambda(\theta_i, s) \delta(x \cos \theta_i - y \sin \theta_i - s)\ ds \\ &= \sum_{i=1}^{n} \lambda(\theta_i, x \cos \theta_i - y \sin \theta_i) = \sum_{i=1}^{n} \lambda_i, \end{aligned}$$

where λ_i is a function supported on the line

$$L_i := \{y \cos \theta_i = -x \sin \theta_i\}, \qquad i = 1, \ldots, n.$$

Thus from equation (6), we see that

$$\rho(x,y) = \prod_i^n a_i, \tag{7}$$

where each a_i is a function whose support is contained in L_i for $i = 1, \dots n$.

Let us see how this argument looks for horizontal and vertical sections through the image, i.e.,

$$\int \rho(x,y)\, dy = u(x), \qquad \int \rho(x,y)\, dx = v(y).$$

Consistency (Fubini's Theorem) implies

$$\int u(x)\, dx = \int v(y)\, dy = \int \int \rho(x,y)\, dx\, dy = s.$$

Our argument above implies that the function $\rho(x,y)$ has to be of "rank one", meaning it is separable

$$\rho(x,y) = a(x)\; b(y).$$

Plugging into the constraint equations says

$$a(x) = \frac{c\, u(x)}{s}, \qquad b(y) = \frac{v(y)}{c},$$

for some constant c, and so

$$\rho(x,y) = \frac{u(x)\; v(y)}{s}$$

is the maximal entropy solution.

4.2 Finite Number of Continuous Pixel Density Distributions

We are given $\rho_i(x,y) \geq 0$, $i = 1, \dots N$, pixel density distributions

$$\sum_{i=1}^{N} \rho_i(x,y) = 1.$$

We choose a finite sample of angles

$$\theta = \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_n \end{bmatrix}$$

for distance s to get

$$A \begin{bmatrix} \rho_1 \\ \vdots \\ \rho_N \end{bmatrix} = \begin{bmatrix} g(\theta_1, s) \\ \vdots \\ g(\theta_n, s) \end{bmatrix}. \tag{8}$$

Therefore from equation (6), we see that

$$\begin{bmatrix} 1 + \log \rho_1 \\ \vdots \\ 1 + \log \rho_N \end{bmatrix} = A^* \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}.$$

We have of course $N > n$ (the number of pixels is greater than the number of measurements).

Note that

$$\text{Image of } A^* = (\text{Kernel of } A)^{\perp}.$$

Let $v_1, \ldots, v_r$ be a basis of $\ker A$,

$$v_j = \begin{bmatrix} v_{j1} \\ \vdots \\ v_{jN} \end{bmatrix}, \qquad j = 1, \ldots, r.$$

Therefore

$$v_j \cdot \begin{bmatrix} 1 + \log \rho_1 \\ \vdots \\ 1 + \log \rho_N \end{bmatrix} = 0, \qquad j = 1, \ldots, r,$$

or equivalently

$$\sum_{k=1}^{N} v_{jk}(1 + \log \rho_k) = 0, \qquad j = 1, \ldots, r.$$

We can compute that

$$-\sum_{k=1}^{N} v_{jk} = \sum_{k=1}^{N} v_{jk} \log \rho_k = \log \prod_{k=1}^{N} \rho_k^{v_{jk}}.$$

Thus we derive the system of equations for ρ_i:

$$\prod_{k=1}^{N} \rho_k^{v_{jk}} = \exp\left(-\sum_{k=1}^{N} v_{jk}\right), \qquad j = 1, \ldots, r. \tag{9}$$

We also have the n original (dependent) constraint equations for the densities (8). Note that $\dim \mathrm{Image}(A) = m \leq n$. Then $m + r = N$. From these $n + r$ equations we get $m + r = N$ independent equations for the required densities ρ_i, $i = 1, \ldots, N$.

4.3 Discrete Density Distributions

Let

$$\rho = (\rho_{ij})$$

be an $m \times n$ matrix representing an image. We impose the constraint equations

$$\sum_{j=1}^{n} \rho_{ij} = u_i, \qquad \sum_{i=1}^{m} \rho_{ij} = v_j,$$

corresponding to the row and column sums of ρ. Consistency requires

$$s = \sum_{i=1}^{m} u_i = \sum_{j=1}^{n} v_j = \sum_{i,j=1}^{n} \rho_{ij}.$$

We regard the constraint equations as a linear system $A\rho = g$ of $m + n$ equations in mn unknowns. The elements of the kernel $\ker A$ of the coefficient matrix can be identified with $m \times n$ matrices. A basis for the kernel is given by the matrices v_{ijkl} with $1 \leq i, k \leq m$ and $1 \leq j, l \leq n$, that have two entries equal to $+1$ in positions ij and kl, two entries equal to -1 in positions il and kj, and zeros elsewhere. Therefore, according to section 4, the maximal entropy solution must satisfy the equations

$$\frac{\rho_{ij}\rho_{kl}}{\rho_{il}\rho_{kj}} = 1$$

or

$$\rho_{ij}\rho_{kl} - \rho_{il}\rho_{kj} = 0$$

for all $1 \leq i, k \leq m$ and $1 \leq j, l \leq n$. The latter system of equations says that all 2×2 minors of the matrix ρ vanish. Therefore

$$\rho = a\, b^T$$

is a matrix of rank 1, i.e.

$$\rho_{ij} = a_i b_j,$$

where $a \in \mathbf{R}^m$ and $b \in \mathbf{R}^n$ are column vectors. Substituting this formula into the constraint equations, we easily find the solution

$$a = \frac{c\,u}{s}, \qquad b = \frac{v}{c}$$

where c is an arbitrary scalar. Therefore

$$\rho_{ij} = \frac{u_i\ v_j}{s}.$$

This gives the maximal entropy solution for a general $m \times n$ matrix.

5 Conclusions and Further Research

In this note, we have begun a rigorous study of the use of a maximal entropy technique for the reconstruction of imagery in computerized tomography. Maximal entropy gives a neat, elegant mathematical solution for this problem.

There are still a number of fundamental issues that must be studied. The first is to describe the procedure for all the various key scanning geometries [2]. This is essential in developing explicit computer algorithms. The next step would be then to actually apply our method to real CT imagery. Robustness to noise artifacts will of course be a major point to be carefully investigated.

References

[1] M. Miller and D. Snyder, “The role of likelihood and entropy in incomplete-data problems: applications to estimating point-process intensities and Toeplitz constrained covariances,” *Processdings of IEE* **75** (1987), pp. 892-907.

[2] R. Brooks and G. Di Chiro, ”Principles of computer assisted tomography in radiographic and radioisotopic imaging,” *Phys. Med. Biol.* **21**, pp. 689-732.

[3] F. Natterer, *The Mathematics of Computerized Tomography*, SIAM Publications, Philadelphia, 2001.

[4] W. Press, S. Teukolsky, W. Vetterling and B. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd Edition, Cambridge University Press, Cambridge U.K., 1992.

[5] G. Strang, *Introduction to Applied Mathematics*, Wellesley-Cambridge Press, Wellesley, Mass., 1986.