

Diffuse Tomography

Alan Johnson *

July 9, 2004

Abstract

Diffuse tomography is similar to classic tomography, except that the wave being sent through the body can be diffracted and thus does not travel in a straight line. The data available to the inverse tomographer is a matrix of the probabilities of a photon entering a given border site, and exiting a given border site. The process is modelled using a rectangular grid, where the photon is able to move in straight lines between the nodes. The probability that the photon moves in a given direction is dependent on its previous position, and since there are four possible directions to travel due to the rectangular model, there are 16 parameters per pixel. Simple cases are examined. The forward problem of taking in the parameters at each pixel and producing the data is solved, and is used in an algorithm to numerically approximate solutions to two examples.

1 Introduction

Classical tomography reconstructs information about the inside of an object by sending high energy waves through the object and measuring the difference in intensities. Because of their high energy, these waves, x rays are a common example, travel in straight lines through the object, and mathematical methods exist for imaging the inside of the body using this information. However, it is desirable to use waves with less energy to image the inside of the object, because of the harm that high energy waves have on the human body. The problem that results from the use of lower energy waves, such as visible light, is that they do not travel in a straight line, which renders the existing mathematical models for imaging by x rays unusable. This problem is called diffuse tomography; attempting to image the inside of a body using waves that can diffract upon entering the body.

In understanding the inverse problem, it is necessary to understand the model used to describe the process. Consider an $M \times N$ grid of pixels, as shown in Figure 1. A device exists to shoot a photon into the system from each of the outside edges, called ports of which there are $2(M + N)$, and another device exists at each port which detects the arrival a photon at that port. As a photon enters a pixel, there is a probability that the photon turns by increments of 90 degrees, dependent upon which direction it enters the pixel. Photons do not interact inside the body. Thus the future state of the photon is dependent on both the current pixel of the photon, and the previous pixel the photon was at. The process thus described is a Markov process. There are three different types of Markov states, incoming, outgoing, and hidden. Incoming states are those paths that enter a pixel from the outside of the grid. Outgoing states are the paths which travel backward relative

*University of Utah Summer REU 2004

to the incoming states. Any other state is called a hidden state. Each type of state contains many states, each of which is given a number according to Figure 1. Incoming and outgoing states that are identical except for opposite directions are given the same number. The data from the model is the data matrix Q , a $2(M+N) \times 2(M+N)$ matrix, where the entry in cell i, j is the probability that a photon that enters the system at port i will leave the system at port j . This data matrix is the data that is available to the inverse tomographer. The probability of a photon traveling a given direction coming from a given direction at each pixel are the parameters which are being recovered.

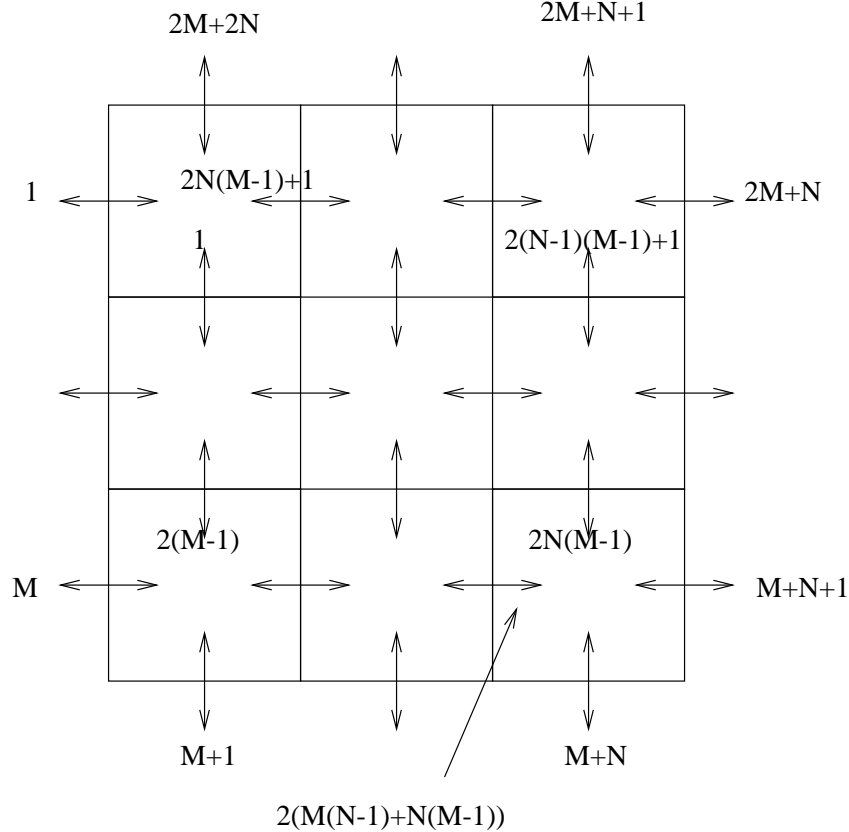


Figure 1: General Setup

2 Simple Cases

The easiest possible case is a 1×1 grid, a single pixel. In this case, the data matrix Q contains the exact probabilities for the pixel, and no further work is required. However, the 1×1 case is not very enlightening for further examination of the problem. Figure 2 shows the next simplest case, two pixels next to each other. Consider the case in which the photon's future direction is independent of its previous pixel. In this case, simple formulas exist for each entry in the data matrix Q . Name the probability of starting from pixel i, j and going down, d_{ij} , going left, l_{ij} , going right, r_{ij} , and going up, u_{ij} . Thus the probability of entering port 1 and leaving port 1 is

$$Q(1, 1) = l_{11} \sum_{i=0}^{\infty} (d_{11} u_{21})^i = \frac{l_{11}}{1 - d_{11} u_{21}} \quad (1)$$

because the sum is a geometric series. Similarly one can write an expression for the probability of entering port 2 and leaving port 1

$$Q(2, 1) = u_{21} l_{11} \sum_{i=0}^{\infty} (d_{11} u_{21})^i = \frac{l_{11} u_{21}}{1 - d_{11} u_{21}} = u_{21} Q(1, 1) \quad (2)$$

again, because the sum is a geometric series. Because all entries of Q are known, we can solve for u_{21} by using (1) and (2). Using similar methods with $Q(2,2)$ and $Q(1,2)$ will yield d_{11} , and once these are known the geometric ratio is known and all other probabilities at each cell can be computed. While the 2×1 example is more enlightening than the 1×1 case, it provides no tools for dealing with larger grids in which the expression for $Q(i, j)$ is not a simple geometric series.

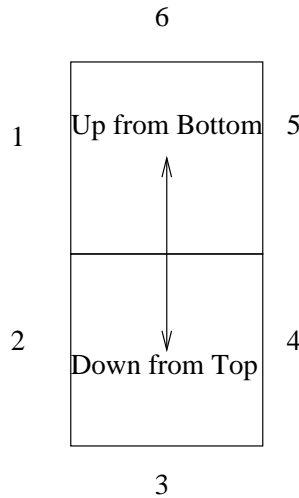


Figure 2: 2×1 case, border numbers indicate port numbers

3 The Forward Problem

The general forward problem has an explicit solution as given by Patch [1]

$$Q = P_{io} + P_{ih}(I - P_{hh})^{-1}P_{ho}, \quad (3)$$

where P_{io} is a probability matrix with $P_{io}(i, j)$ being the probability of entering an incoming state i and immediately exiting through an outgoing state j . P_{ih} is a probability matrix with $P_{ih}(i, j)$ the probability of going from incoming state i to hidden state j . P_{hh} is a probability matrix where

$P_{hh}(i, j)$ is the probability of going from hidden state i to hidden state j . P_{ho} is a probability matrix where $P_{ho}(i, j)$ is the probability of going from hidden state i to outgoing state j .

Thus, given the probabilities at all the pixels, the matrices P_{io} , P_{ih} , P_{hh} , P_{ho} can all be derived. The use of a computer makes it possible to easily calculate an arbitrary $M \times N$ case quickly, using the state numbering scheme given for Figure 1. Figure 3 is the 15×15 case, where the photon's future state is independent of its previous state, the probability of going up or left at every cell is .45, and the probability of going right or down at every cell is .15.

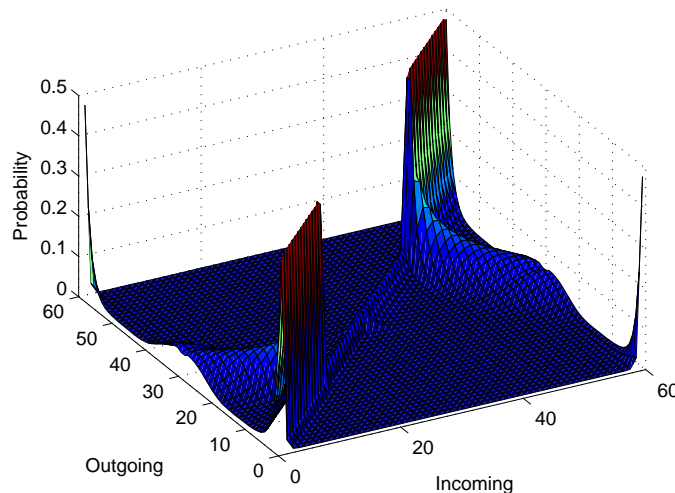


Figure 3: Data from skewed 15×15 square

4 Approximating a Solution

Having an implementation of the forward problem for any given set of probabilities, it becomes possible to recover approximations to the actual set of probabilities given a set of data Q using the gradient descent method. In approximating a solution, the assumption is made that all information about P_{io} , P_{ih} , and P_{ho} are known. This makes physical sense, as the exterior of the object in question is available to examine without the need of tomography. It also simplifies the process without losing the most interesting part of diffuse tomography, the recovery of the inner, hidden probabilities. Thus the only probabilities that are allowed to vary are those probabilities in the matrix P_{hh} .

4.1 Calculating a Gradient

In order to calculate a gradient, we must first define a function of which we can take the gradient. Let

$$J(P) = \sum_{i=1}^k \|(F(P) - Q)e_i\|_2^2,$$

where $P = P_{hh}$, and $F(P)$ is the $2(M + N) \times 2(M + N)$ matrix that represents the data calculated with an arbitrary matrix P , Q is the actual data matrix, e_i is a $2(M + N)$ vector of all zeros except a 1 in the i^{th} spot, k is the number of columns of Q , and the norm used is the l^2 norm. Minimizing J will be the approach in approximating a solution to a given data matrix. The Gâteaux derivative of J at P in the direction δP is

$$DJ(P)(\delta P) = \sum_{i=1}^k \langle DF(P)(\delta P)e_i, (F(P) - Q)e_i \rangle, \quad (4)$$

where $DF(P)(\delta P)$ is the Gâteaux derivative of F at P in the direction δP , and the inner product is the standard dot product on \mathbb{R}^n , $n \in \mathbb{Z}$, $n > 0$. Using the fact that

$$DF(P)(\delta P)e_i = P_{ih}(I - P)^{-1}(\delta P)(I - P)^{-1}P_{ho}e_i \quad (5)$$

a formula for the gradient of J is found explicitly. Combining (4) and (5) and correctly manipulating the expression one gets

$$DJ(P)(\delta P) = \langle \delta P, [(I - P)^{-1}]^T P_{ih}^T (F(P) - Q) P_{ho}^T [(I - P)^{-1}]^T \rangle \quad (6)$$

Thus the gradient of J at P is the second term of the inner product, keeping in mind that P_{ih} and P_{ho} are known by assumption.

4.2 Gradient Descent Method

Having calculated a gradient, it is now possible to approximate a solution to a given data matrix. The algorithm to be used is the gradient descent method. First a guess is made at the answer, and the gradient in (6) is calculated at that point. The guess is then modified by subtracting h , a small increment with $h > 0$, times the gradient calculated. The probabilities being recovered must be by their very nature must be in the interval $[0, 1]$. If the gradient calculated results in a probability falling outside the interval $[0, 1]$, that probability is instead left unchanged. If J calculated at the modified guess is greater than a set tolerance, the modified guess is taken as the guess, and the gradient is recalculated and a new guess generated. At this point it is important that the forward problem is already solved, because without the forward problem, J could not be calculated. This process is continued until J is less than the desired tolerance. This method is assured to result in an area surrounding a local minimum, so that with a sufficiently small tolerance (and enough patience), the probabilities to a solution can be recovered.

5 Recovering Probabilities

Now that a method exists for approximating the solution given a data matrix, two different cases are investigated. One is a situation that enjoys much symmetry and uniformity. The other is a “tumor,” or anomaly that one might be looking for inside a body.

5.1 A Symmetric Probability Distribution

Consider a 5×5 square in which the future location of a photon is independent of its previous location. The probability of going up and left is .45 and the probability of going down or right is .05. The data matrix is first calculated so that the gradient descent method can be used. As a guess, the photon is considered to travel independent of its previous position, and the probability of going up, down, left, and right is .25. Using a step size of .01, and a tolerance of .05, the results are shown in Figure 4. The graphs show the matrix P , the probability of going from a hidden state to a hidden state. The general features of the actual probabilities are recovered, but the specifics are quite poor. On average every entry is about 17 percent off of the actual value. This rather large discrepancy is not reflected in the output of the data matrix, however. The sum of the squares of the differences between the entries of the actual data matrix, Q , and the data matrix calculated from the recovered probabilities is about .05. The problem is that the farther inside the grid a pixel is, the less likely a photon is to ever reach that pixel. Thus the effect that the hidden states have on the output data matrix, Q , is rather small. In order to achieve higher accuracy, the step size and the tolerance must both be lowered, which results in greatly increased computing time. It took 736 iterations and about 10 seconds to produce the recovered results.

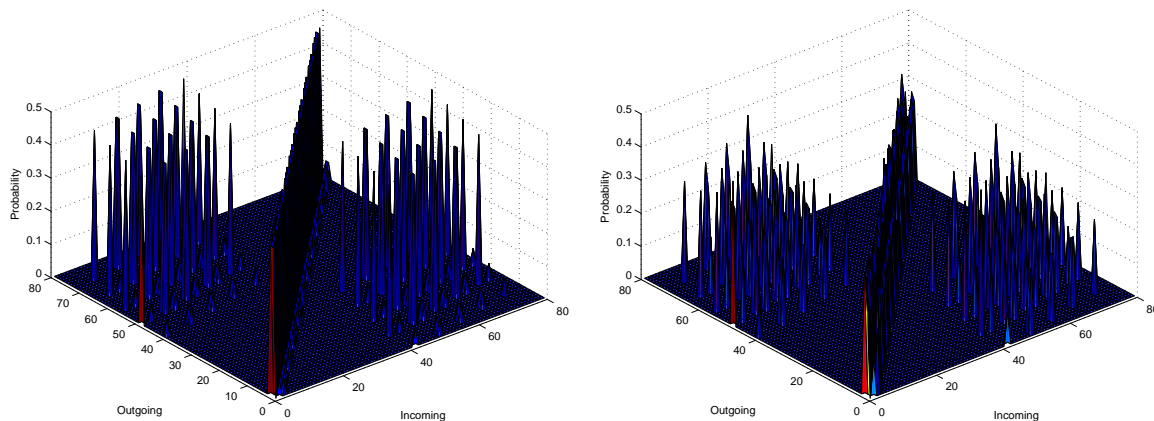


Figure 4: Left Actual, Right Recovered

5.2 Recovering an Anomaly

Again consider a 5×5 grid, but this time the photon’s future position is dependent on its previous position. Specifically, coming from any direction and at all pixels the probability that the photon goes straight is .97, and going left, right, or back all have probability of .01. In this set up, the

diffraction of the photons is small, and thus the problem approximates the classic tomography problem. The center cell is different than the rest of the cells, the future state of the photon upon entering the center cell is independent of its previous position with probability of turning .25 in all directions. Thus the center cell acts as a “tumor”, or anomaly, that one would want to be able to recover using diffuse tomography.

Using this setup, the forward problem is solved and the data matrix Q is produced. Similar to the last section, the gradient descent method is used with the same initial guess, step size, and tolerance. From the data recovered, 16 images of the grid can be formed, one image for each possible combination of previous and future state. These 16 possible images fall into three groups, where the groups share similar features. Figure 5 shows the three different groups. The groups are the probability of going into a pixel and continuing in the same direction, the probability of going into a pixel and turning left or right, and the probability of going into a pixel and going back in the opposite direction. Once again the recovered probabilities get a solution that shares large scale features with the actual solution. The left/right recovered probabilities show a spike in the middle corresponding to the anomaly. However, the spike is an about 8 percent over the actual value, and there are fluctuations around the spike that are not present in the actual values. The straight recovered probabilities show an inverse spike at the middle pixel, which corresponds to the anomaly, and again is about 8 percent over the true value, with minor fluctuations around the anomaly. The back recovered probabilities show the most error. A ridge appears where there should be a single spike. The ridge is inaccurate in the sense that the highest probability falls away from the pixel with the anomaly.

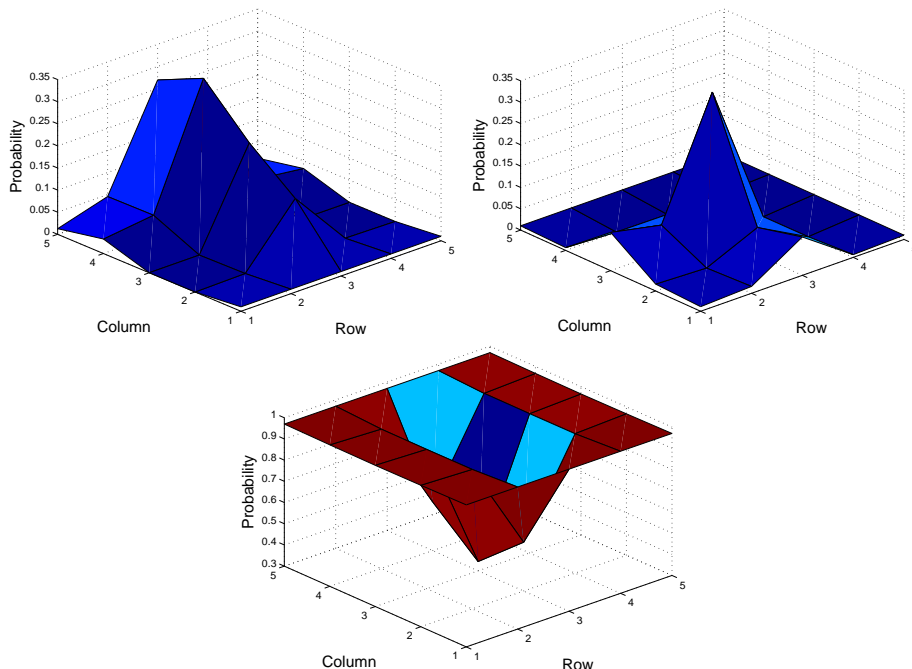


Figure 5: Left Recovered Back, Bottom Recovered Straight, Right Recovered Left/Right

6 Conclusion

Modeling the forward problem and using the gradient descent method provide a means by which the probabilities of the grid can be recovered. However, the degree to which the recovered probabilities resemble the actual probabilities is dependent on the tolerance which is used in the gradient descent algorithm. However, the matrix P_{hh} involves many paths which are highly unlikely to occur, and thus they do not affect the data in an appreciable manner. Thus, there is a high percentage of error, as seen in the two examples shown. As the system gets larger, the innermost of the hidden to hidden states become less and less likely to occur, and thus affect the data to an even smaller amount. These probabilities are going to be the most difficult to recover with a high degree of certainty.

7 Further Work

There are many areas of this problem which still can be explored. Instead of declaring the matrices P_{io} , P_{ih} , P_{ho} to be constant, they can be assumed to change, and a new expression for the gradient can be used. The Gâteaux derivative of F can be written as a matrix, and the singular value decomposition can be viewed to find the size and shape of the nullspace. This will tell what types of changes will not affect the data, and give an idea of what types of configurations are recoverable. Further computer experiments can be run to see the behavior of the recovery for larger systems, and different types of probabilities to be recovered, such as multiple anomalies. Pertinent to medical imaging is the three dimensional case, which does not vary in any fundamental sense from the two dimensional case studied here.

References

- [1] S. K. Patch. Consistency conditions in diffuse tomography. *Inverse Problems*, 10:199–212, 1994.