PAPER F

TOMOGRAPHIC TRAVELTIME INVERSION USING NATURAL PIXELS

Reinaldo J. Michelena and Jerry M. Harris
Seismic Tomography Project

ABSTRACT

Traditionally in the problem of tomographic traveltime inversion, the model is divided into a number of rectangular cells of constant slowness. Inversion consists of finding these constant values using the measured traveltimes. The inversion process can demand a large computational effort if a high resolution result is desired.

We show in this paper how to use a different kind of parametrization of the model based on beam propagation paths. This parametrization is obtained within the framework of reconstruction in Hilbert spaces by minimizing the error between the true model and the estimated model. The traveltimes are interpreted as the projections of the slowness along the beam paths. Although the actual beam paths are described by complicated spatial functions, we simplify the computations by approximating these functions with functions of constant width and height, i.e., "fat" rays, which collectively form a basis set of natural pixels.

With a simple numerical example we demonstrate that the main advantage of this parametrization, compared with the traditional decomposition of the model in rectangular pixels, is that 2D reconstructed images of similar quality can be obtained with considerably less computational effort. This result suggests that the natural pixels can provide considerable computational advantage for 3D problems.

INTRODUCTION

The process of reconstructing an image using line integrals through it is called tomography. In traveltime tomography the image to be reconstructed is the slowness model $S(\mathbf{r})$. The reconstructed model $\tilde{S}(\mathbf{r})$ is usually represented as a linear combination of functions $\beta_n(\mathbf{r})$ in the form

$$\tilde{S}(\mathbf{r}) = \sum_{n=1}^{M} a_n \beta_n(\mathbf{r}). \tag{1}$$

The problem consists of determining the unknown coefficients a_n from the measured traveltimes. Once these coefficients have been calculated, the computation of the sum (1) is straightforward.

The representation (1) has two important degrees of freedom that influence decisively the kind of results obtained. These are the number M and kind of functions

$\beta_n(\mathbf{r})$ to be used. The most common choice for the functions $\beta_n(\mathbf{r})$ is orthogonal cells (square or cubic pixels) and in that case the coefficients a_n represent the slowness within each cell (McMechan, 1983; Ivansson, 1985). Although this is the most popular basis function used for estimating the slowness model, others have been suggested recently. Harlan (1989) defines the velocity function as a sum of smooth basis functions (Gaussians), and Van Trier (1988) defines the functions $\beta_n(\mathbf{r})$ as cubic B-splines multiplied by functions that reproduce the expected structure of the model. The number of functions M is also arbitrary but is usually wanted to be "small" to avoid having to solve a huge system of equations.

The kind and number of functions used for expanding the slowness model determine many of the general features of the final image. With the same data set it is possible to obtain different results just because different parametrizations have been used. However, the goal is to obtain a reconstructed model free from these artifacts derived from the parametrization. This means that the selection of the basis function is a critical step in the inversion process and then should be considered more carefully, as described below.

There is no general criteria for deciding which representation is the best, although some may have clear advantages for solving specific problems. Our selection of the basis function will be based on the minimization of the expression that estimates the norm of the null space of the problem

$$\|S(\mathbf{r}) - \sum_{n=1}^{M} a_n \beta_n(\mathbf{r})\| \tag{2}$$

where $S(\mathbf{r})$ is the true slowness model. Due to the nature of the measurements in traveltime tomography (integral along beam paths) we show in this paper that the minimum of (2) can be reached when the functions $\beta_n(\mathbf{r})$ describe the beam paths and when M equals the number of measurements available (because there is only one measurement per beam path). In the first part of the paper, this fact is demonstrated within the framework of reconstruction in Hilbert spaces. The remainder of the paper presents a comparison of the inversion of synthetic data using the traditional representation of the model in square pixels and the proposed representation in constant regions along the beam paths called natural pixels.

RECONSTRUCTION IN HILBERT SPACES

A Hilbert space is a linear space on which an inner product is defined. For example, the inner product for the Hilbert space L^2 of the Lebesque square-integral functions of support Ω is

$$< f(x), \beta(x) > = \int_{\Omega} f(x)\beta(x)dx. \tag{3}$$

We can assume that the particular function $f(x)$ that we want to estimate belongs to a Hilbert space H. Let's assume also that the information we have about $f(x)$, i.e., data, is a set of inner products of the function $f(x)$ with a finite set of known

functions $\beta_m(x) \epsilon H$

$$d_m = < f(x), \beta_m(x) > \qquad m = 1, ..., N. \tag{4}$$

In this context, the data can be interpreted as the projections of the unknown function $f(x)$ onto the "sampling" functions $\beta_m(x)$.

If F_1 is a closed linear subspace of the Hilbert space H, then $H = F_1 \oplus F_1^{\perp}$ (Berberian, 1976), where $F_1^{\perp}$ is called the orthogonal complement of F_1. From the projection theorem (Stakgold, 1979), we can always decompose $f(x)$ into $f_1(x)+f_2(x)$ (Fig. 1) where $f_1(x) \epsilon F_1$ and $f_2(x) \epsilon F_1^{\perp}$. $f_1(x)$ is called the orthogonal projection of $f(x)$ in F_1.

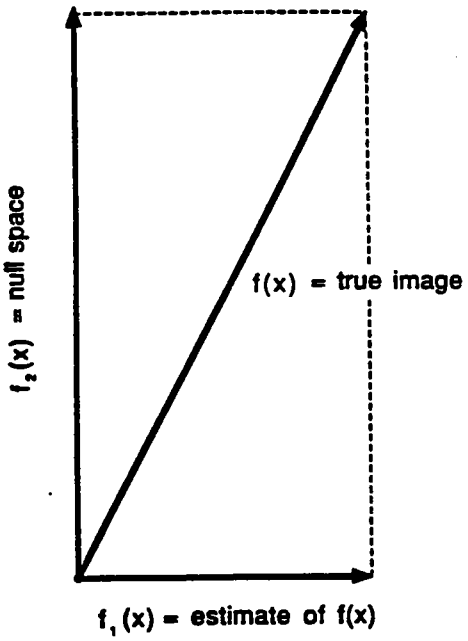

Figure 1: Orthogonal projections of the true image.

If we assume that the functions $\beta_n(x)$ form a basis of the space F_1, we can write

$$f(x) = \sum_{n=1}^{N} a_n \beta_n(x) + f_2(x). \tag{5}$$

We can understand the meaning of the function $f_2(x)$ by multiplying both sides of (5) by $\beta_m(x)$ and integrating in Ω

$$d_m = \sum_{n=1}^{N} a_n < \beta_n(x), \beta_m(x) > + < f_2(x), \beta_m(x) > . \tag{6}$$

Since $< f_2(x), \beta_m(x) >= 0$, we can say that $f_2(x)$ contains the information about $f(x)$ that does not affect the measurements made by the sampling functions $\beta_m(x)$.

Finally, the estimate $\tilde{f}(x)$ of $f(x)$ can be written as

$$\tilde{f}(x) = f_1(x) = \sum_{n=1}^{N} a_n \beta_n(x), \tag{7}$$

where the coefficients a_n are calculated from the forward equation for the data d_m

$$d_m = \sum_{n=1}^{N} a_n < \beta_n(x), \beta_m(x) > \qquad m = 1, ..., N. \tag{8}$$

Note that the system of equations (8) is square. If all the sampling functions are independent, the problem is well conditioned and therefore, the system (8) has a unique solution. A unique solution is expected because the orthogonal projection of $f(x)$ in the space F_1 expanded by the functions $\beta_m(x)$ is unique (Fig. 1). If any of the sampling functions can be expressed approximately as a linear combination of the others, the problem becomes ill-conditioned since the rows (or columns) in the matrix of Eqn. (8) are not independent any more.

Minimum Norm Solution

This same result for $\tilde{f}(x)$ can be obtained through minimization of the norm $\|f_2(x)\|$ with respect to the unknown coefficients a_n (Darling et al., 1983),

$$min \ \|f_2(x)\|^2 = min \ \|f(x) - \sum_{n=1}^{N} a_n \beta_n(x)\|^2. \tag{9}$$

For this reason, the estimate $\tilde{f}(x)$ is called the minimum norm estimate of the unknown function $f(x)$. This estimate $\tilde{f}(x)$ is unique and consistent with the data (Eqn. (8)). It is also strongly related to the way the modeled data are generated, because the unknown function $f(x)$ is expressed as a linear combination of the sampling functions $\beta_m(x)$ used to compute the forward modeling (Eqn. 4). This means that each experiment will suggest "naturally" the reconstruction procedure which produces the minimum norm solution.

Examples of different sampling functions in different problems are the beam paths in the problem of traveltime tomography, complex exponentials when the measurements are the frequency components of $f(x)$, etc.. When the measurements are the frequency components of $f(x)$ the minimum norm estimate is simply the Fourier series expansion of $f(x)$ (Stakgold, 1979). In fact, Eqn. (7) can be interpreted as a generalized Fourier series for basis sets not necessarily orthogonal.

If we expand $\tilde{f}(x)$ in any other set of basis functions $\{\alpha_n(x, y), \ n = 1, ..., M\}$, the norm of $f_2(x)$ is

$$\|f_2(x)\| = \|f(x) - \sum_{n=1}^{M} c_n \alpha_n(x)\|. \tag{10}$$

Minimizing this expression with respect to the unknown coefficients c_n, we obtain

$$< f(x), \alpha_m(x) > = \sum_{n=1}^{M} c_n < \alpha_n(x), \alpha_m(x) > \qquad m = 1, ..., M. \tag{11}$$

Note that the independent term on the left-hand side in the system of equations (11) is formed by the inner products of the selected basis function with the unknown function $f(x)$. Note also that the matrix elements are the inner products among the different elements of the selected basis. The independent term is equal to the measurements only if the basis set used for expanding $\tilde{f}(x)$ ($\{\alpha_n(x,y),\ n = 1, ..., M\}$) is the same one used for generating the data, i.e., $\alpha_n = \beta_n$. If any other basis set is used, $\alpha_n \neq \beta_n$, then the independent term must be computed from the measurements. Therefore, the choice $\alpha_n = \beta_n$ is a convenient one among many others basis sets because the solution obtained is still minimum norm and the independent term in Eqn. (11) represents directly the measurements. We will see later that in the problem of traveltime inversion the functions $\beta_n(x,y)$ are the beam paths and then the minimum norm solution can be obtained easily from Eqn. (7).

Let's illustrate these ideas with a simple example. Although this example is not a geophysical one, it can help to understand why square orthogonal pixels are a convenient basis set when the measurements represent the average of the unknown over the same square regions. When the measurements have a different meaning, it might be computationally easier to get the minimum norm solution by using a different basis set. Consider a two dimensional object $O(x,y)$ (a photograph, for example). The process of spatial discretization of the object can be interpreted as the convolution of $O(x,y)$ and the sampling function $R(x,y)$ which describes the shape of the pointer of the digitizer (for simplicity the function $R(x,y)$ is assumed to have unit area)

$$d(x,y) = O(x,y) * R(x,y), \tag{12}$$

where $d(x,y)$ is the digitized image. This is equivalent to superimposing a square grid, for example, over the object and calculating the function $d(x,y)$ from the volume of the object in the support of each pixel surrounding the grid point. This procedure can be expressed as

$$d_i = \int_{\Omega_i} O(x,y) R_i(x,y) dx dy \tag{13}$$

where

$$R_i(x,y) = \begin{cases} 1 & \text{if } (x,y) \text{ is in the pixel } i \qquad (i = 1, ..., N) \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

$d_i,\ i = 1, ..., N$ are the data points and Ω_i is the support of the i-pixel.

Given the inner product (13), the minimum norm estimate of $O(x,y)$ according to (7), is

$$\tilde{O}(x,y) = \sum_{n=1}^{N} a_n R_n(x,y). \tag{15}$$

The coefficients a_n can be found from the equation

$$d_m = \sum_{n=1}^{N} a_n < R_n(x,y), R_m(x,y) > \qquad m = 1, ..., N. \tag{16}$$

According to the definition of R_i

$$< R_n(x,y), R_m(x,y) > = \delta_{nm}. \tag{17}$$

Then, $a_n = d_n$ and the estimate becomes

$$\tilde{O}(x,y) = \sum_{n=1}^{N} d_n R_n(x,y). \tag{18}$$

As expected, the reconstructed object is formed with a superposition of N cells, each with constant height d_n and located where the measurements were taken. If we use any other basis set instead of $R_i(x,y)$, it is possible to get an estimate that does not reproduce the data, or requires of more model parameters to get a better representation.

As we said before, Fourier analysis is another example where the sampling functions $\exp(ik_n x)$ are the same ones used to expand the estimate of the unknown. The result is also a minimum norm estimate.

The theory of reconstruction in Hilbert spaces generates consistent estimates of the unknown in the sense that the same basis set used to sample the function is used to expand it. The two previous examples (orthogonal sampling and Fourier reconstruction) confirm that in some situations this might be a convenient choice among many other possibilities. In the following sections we will exploit this idea of consistency in the problem of traveltime tomography where the data are generated in a very specific way (integrals along beam paths).

TOMOGRAPHIC TRAVELTIME INVERSION

The traveltime along a ray l_m in a medium where the slowness is $S(x,y)$, is traditionally given as

$$t_m = \int_{l_m} S(x,y) dl_m \qquad m = 1, ..., N, \tag{19}$$

where dl_m is the incremental distance along the ray path l_m. In general, the ray path depends on the slowness distribution. For sake of simplicity, let's assume that the variations in the slowness are just a few precent. Then we can safely consider that the ray paths are straight lines and independent of the slowness. The general case will be discussed later.

Although the expression (19) simplifies the mathematics considerably, it fails to convey the fact that the traveltimes between two points are affected by velocities in the region called the Fresnel zone, which is infinitely narrow only when the wavelength

λ is infinitely small, $\lambda \rightarrow 0$ (Nolet, 1987). To account for the finiteness of this effect, we can say that the traveltime between two points can be better described by the equation

$$t_m = \int_\Omega S(x,y)\phi_m(x,y)dxdy, \tag{20}$$

where $\phi_m(x,y)$ is a two dimensional function or "beam" of finite support centered along the ray path and Ω is the support of $S(x,y)$. The functions $\phi_m(x,y)$ can be interpreted as the wavepaths introduced by Woodward (1989).

With the forward modeling equation written in this way, the estimation of the slowness from the traveltimes can be seen as a reconstruction problem in a Hilbert space where the inner product is defined by (20). According to (7), the minimum norm estimate of the slowness $S(x,y)$ is

$$\tilde{S}(x,y) = \sum_{n=1}^{N} a_n \phi_n(x,y), \tag{21}$$

where N is the number of traveltimes.

We can transform (19) into a two dimensional integral of the form of (20), if we describe the ray path with a two dimensional delta function $\delta_m(x,y)$. However, the problem of reconstruction of the slowness from such an expression cannot be seen as a reconstruction problem in a Hilbert space because the inner product $< \delta_i(x,y), \delta_i(x,y) >$ is not defined.

From Eqn. (8), we find that the coefficients a_n can be calculated through the system of equations

$$t_m = \sum_{n=1}^{N} a_n < \phi_n(x,y), \phi_m(x,y) > \qquad m = 1, ..., N \tag{22}$$

where

$$< \phi_n(x,y), \phi_m(x,y) > = \int_\Omega \phi_n(x,y)\phi_m(x,y)\, dxdy. \tag{23}$$

In contrast with the traditional reconstruction using square pixels as basis function (Eqn. (18)), the reconstruction described above is based on a discretization of the model along the beam paths. In the example discussed in the previous section we showed that the square orthogonal pixels are a convenient discretization that lead to minimum norm estimators when the data are "point" orthogonal samples of the two dimensional function that we want to reconstruct. The discretization along the beam paths comes from the fact that they are the regions sampled with each measurement in traveltime tomography.

When the beam paths are used, the discretization of the model will depend in general on the particular data set that is going to be inverted, because it will reflect the propagation of the energy in the medium. It does not have the advantage of other parametrizations that can reflect some prior knowledge we may have about the model. In that sense, some flexibility is lost. The discretization along the beam paths is similar to the discretization of the model in square pixels in the way the prior

information is handled, since both assume no prior information about the model. The difference is that when the problem is nonlinear and it is solved as a sequence of linearized steps, the discretization along the beam paths adapts progressively to the real model.

It is important to note that although the solution of the system of equations (22) is unique, the null space of the problem has not been suppressed. It has been just separated at the beginning of the formulation in the form of an space orthogonal to the beam paths (function $f_2(x)$, Eqn. (5)). The same situation occurs in discrete Fourier reconstruction, where for a given data set the estimate of the unknown is unique even though the null space (Fourier components above the Nyquist frequency) is not zero.

Natural Pixels

As a first approximation, we can describe the basis function $\phi_i(x,y)$ as functions of width λ' and heigth $1/\lambda'$

$$\phi_i(x,y) = \begin{cases} 1/\lambda' & \text{if } (x,y) \text{ is in the region of width} \\ & \lambda' \text{ centered along the ray path } i \\ 0 & \text{otherwise.} \end{cases} \tag{24}$$

Therefore, the matrix coefficients $< \phi_n(x,y), \phi_m(x,y) >$ are

$$< \phi_n(x,y), \phi_m(x,y) > \lambda'^2 = \begin{cases} \textit{area of the beam path} & \text{if } m = n \\ \textit{area of the intersection} & \text{if } m \neq n. \end{cases} \tag{25}$$

A natural pixel for a single ray is shown in Fig. 2. Even when the rays curve or when reflections are included, the natural pixels are "tubes" centered on the ray path. The shape of the tubes may vary. For example, they can have variable cross section that gives more weight to the center than to the sides.

Buonocore, et al., (1981) and Buonocore (1981), without working within the framework of reconstruction in Hilbert spaces, define an estimator identical to (21) and call it "natural pixel" decomposition of the two dimensional image, where the natural pixels are the functions $\phi_m(x,y)$. They study extensively the properties of such a reconstruction and the theoretical advantages of it compared with the traditional reconstruction using square pixels. They show that the matrix of coefficients $< \phi_i(x,y), \phi_j(x,y) >$ represent the measurement covariance matrix if there is no measurement noise. If the measurement noise is nonzero but uncorrelated, only the diagonal elements of the matrix are different from those of the measurement covariance matrix. According to Buonocore et al., (1981), square pixels errors are caused by the inaccurate estimation of the measurement covariance. These errors can be eliminated only if the size of the pixels could be made infinitely small.

An example of a set of natural pixels is shown in Fig. 3, for the case of a cross borehole geometry in a medium of constant slowness.

The number of cells in the square-pixel-based inversion is commonly determined by a trade off between the required resolution and cost of the inversion. With square

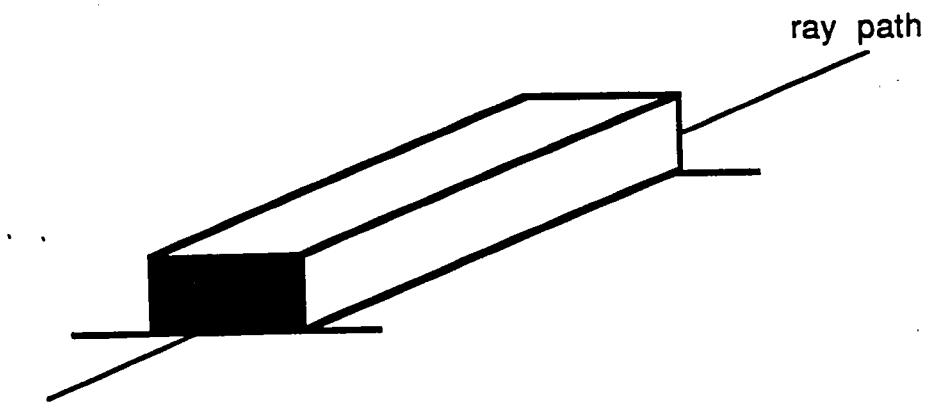

Figure 2: Natural pixel for a single ray.

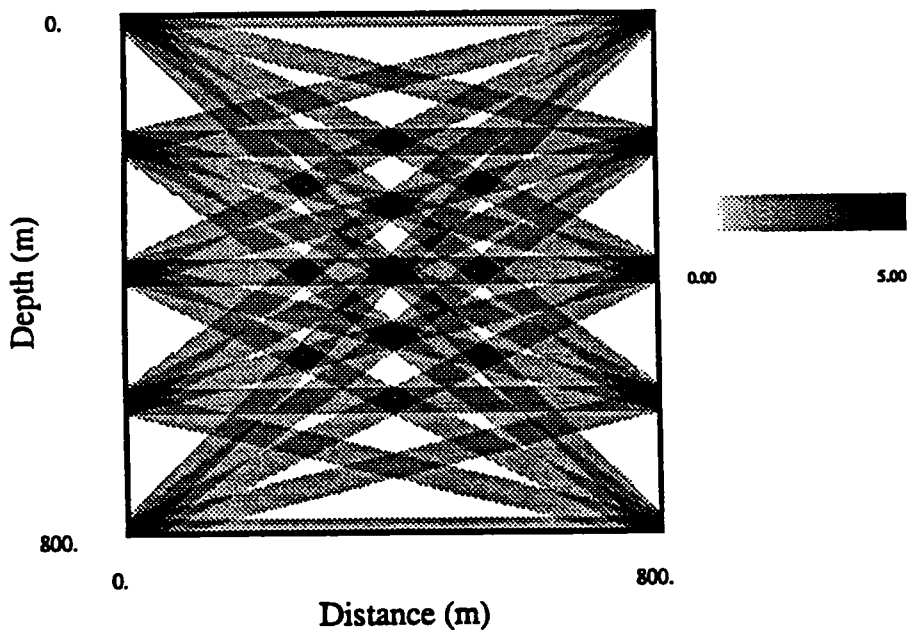

Figure 3: Natural pixels in a constant slowness medium for a cross-well configuration of five sources and five receivers.

pixels we are forced to establish that compromise because they do not differentiate between model parameters and display parameters, although both have opposite purposes: we want many display parameters for an accurate representation but, at the same time, we want few model parameters for an inexpensive inversion. This trade off is particular of the square pixels. An example of this will be explained below. If we decide to use any other basis function instead of square pixels, we always have to discretize fine enough to ensure the *numerical* accuracy in the summation (21), but this fact does not change the number of model parameters.

The coefficients a_n in Eqn. (21) represent all the information gained from the measurements, which can be seen easily when only one measurement is available. In this case, the coefficient a_1 is proportional to the average slowness in the region of the beam path. If we want to invert that measurement using square pixels, the problem in general will be underdetermined or in the best case (using only one pixel in the inversion) we can obtain information not present in the data, e.g., information outside the spatial support of the observation.

Each traveltime measurement *does not* contain information about variations in the slowness along the beam path. The variations are averaged into a single number. It is the combination of them (if they overlap) that gives the information about these variations along the path. This basic fact is contradicted when the model is discretized into square pixels that potentially introduce variations along each beam path and demand from the data more information than it contains. Introducing in the problem more degrees of freedom might be convenient if there is enough information to resolve all of them. The extra information needed to solve the problem is introduced in the form of constraints, some of them result of independent data and others are simply "reasonable" ones. Smoothness is an example of a reasonable constraint that can help to solve the problem of the extra information needed. However, it is not clear how the solution may depend on the various ways of introducing the smoothness (Claerbout, 1976) or other reasonable constraints that don't come from independent observations.

To this point, the inversion is strictly linear, this is, the sampling functions do no depend on the slowness. This is analogous to Fourier reconstruction where the sampling functions (complex exponentials) do not depend on the properties of the unknown. No iterations are needed after the estimate is found. This is not the situation in traveltime tomography where the sampling functions may strongly depend on the unknown slowness. The next section addresses this topic.

Iterative inversion

As we said before, the traveltime along a ray in a medium of slowness $S(x,y)$ is

$$t = \int_l S(x,y)dl, \tag{26}$$

where the traveltimes as well as the ray paths depend on the slowness. When the ray paths are straight lines, like in X-ray tomography or when the variations in slowness are small, we can derive $S(x,y)$ from the traveltimes using this expression.

In geophysical applications, however, straight rays are rarely found and as a result, the problem (26) becomes highly nonlinear since the unknown $S(x,y)$ is also implicitly present in the ray path (Nolet, 1987).

If the medium is perturbed to $S'(x,y) = S(x,y) + \Delta S(x,y)$, the new traveltime calculated along the new ray path l' is

$$t' = \int_{l'} S'(x,y)dl', \tag{27}$$

Using the Fermat principle, it can be shown (Aki and Richards, 1980) that the difference in traveltime between the two media is

$$\Delta t = \int_{l} \Delta S(x,y)dl \tag{28}$$

where $\Delta t = t' - t$.

The nonlinear problem is then solved as a sequence of linearized steps that seeks to minimize the difference between real and calculated traveltimes.

If the perturbations in traveltimes are calculated as integrals along the beam paths, (28) becomes

$$\Delta t_m = \int_{\Omega} \Delta S(x,y)\phi_m(x,y)dxdy \tag{29}$$

The beam paths are centered in the rays traced in the unperturbed model. Substituting $\Delta S(x,y)$ by $\tilde{S}(x,y)$ in Eqn. (21) and Δt_m by t_m in Eqn. (22), we can get the estimate of the slowness perturbation $\Delta S(x,y)$ after solving (22). This estimate reproduces the perturbations in traveltimes and because of this, when the problem is linear, it converges in one iteration. This is the situation that is going to be studied in the examples that follow.

NUMERICAL EXAMPLES

We will now show synthetic inversion examples comparing natural pixels and square pixels as basis functions. Our aim is to compare the results of the inversion when both are used with the same data set. This goal can be achieved with synthetic data for a cross borehole geometry generated from the model shown in Fig. 4. The example is simplified considerably by assuming that the contrast between the circular disc ($S = 2.02$) and the background ($S = 2.00$) is 1%. Therefore, straight rays adequately describe the propagation of the energy in the medium.

The data are generated from strip integrals across the model of Fig. 4. The integrals are calculated from the analytical expressions of the intersection of the strips with the circle. In this way the numerical errors in the forward modeled data have been minimized. In Appendix A it is explained how to calculate numerically strips integrals across any two dimensional slowness model. For these examples 289 traveltimes were computed, which corresponds to the 17 sources and 17 receivers used. Another simplification is made assuming that the width of the strips $\lambda' = 40\ m$ is the same during both the forward modeling and the inversion.

When the model is discretized into square pixels the estimate of $\Delta S(x,y)$ (Eqn. 29) is obtained after solving a system of linear equations where the matrix coefficients represent the area of the intersection of the strip with each pixel. We are going to solve this system and the one obtained with the natural pixels (Eqn. 22) using the LSQR variant of the conjugate gradient method (Nolet, 1987) that has been proved to be faster than SIRT methods (Nolet, 1985; Van der Sluis and Van der Vorst, 1987).

Fig. 5 and 6 show the results of the inversion when the model is discretized with two different pixel sizes. The starting model has a constant slowness $S_0(x,y) = 2$. The inversion produces directly the slowness value in each pixel, and therefore, reducing the size of the pixels (for better resolution) increases the number of model parameters and consequently the size of the system of equations to solve. In the examples shown, the size of the system of equations solved is 289 X 1681 (grid size = 41 X 41; Fig. 5) and 289 X 25921 (grid size = 161 X 161; Fig. 6) respectively. Evidently, the quality of the reconstruction and the amount of information about the model contained in the image increases with the number of model parameters. Note that no interpolating or smoothing process has been applied to the images. The coarse discretization and the limited view of the data are the causes of the artifacts in Fig. 5.

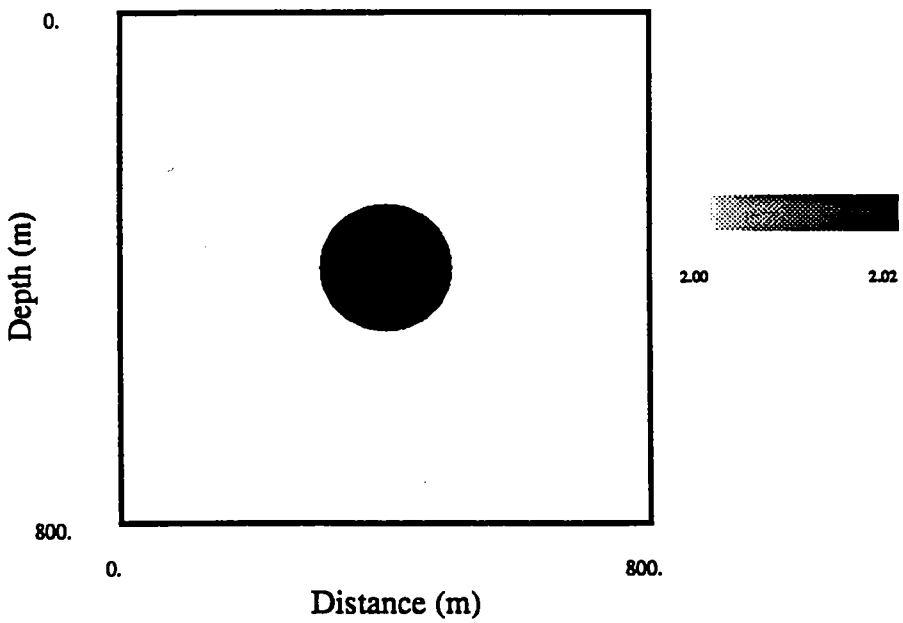

Figure 4: Slowness perturbation. 17 sources are located on the right hand side of the model and 17 receivers are located on the opposite side. The radius of the disc is $r = 100\ m$. The width of the natural pixels is $\lambda' = 40\ m$. The vertical separation between adjacent sources and/or receivers is 50 m.

The result of the inversion using natural pixels is shown in Fig. 7. This image is represented with a grid identical to the one used in Fig. 6 (161 X 161) and then, both results can be compared directly. The system of equations solved with the natural

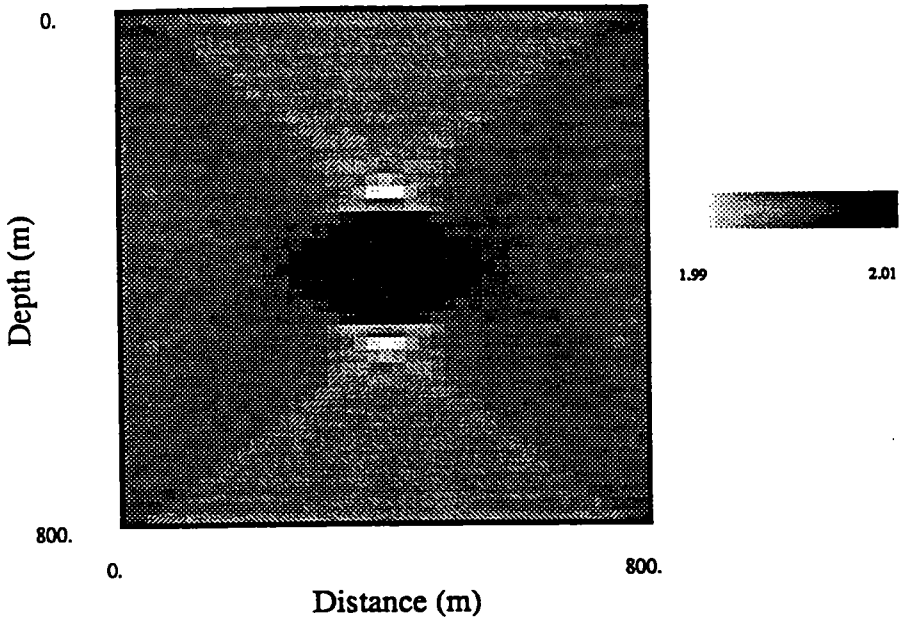

Figure 5: Inversion when a grid of 41 X 41 square pixels is used.

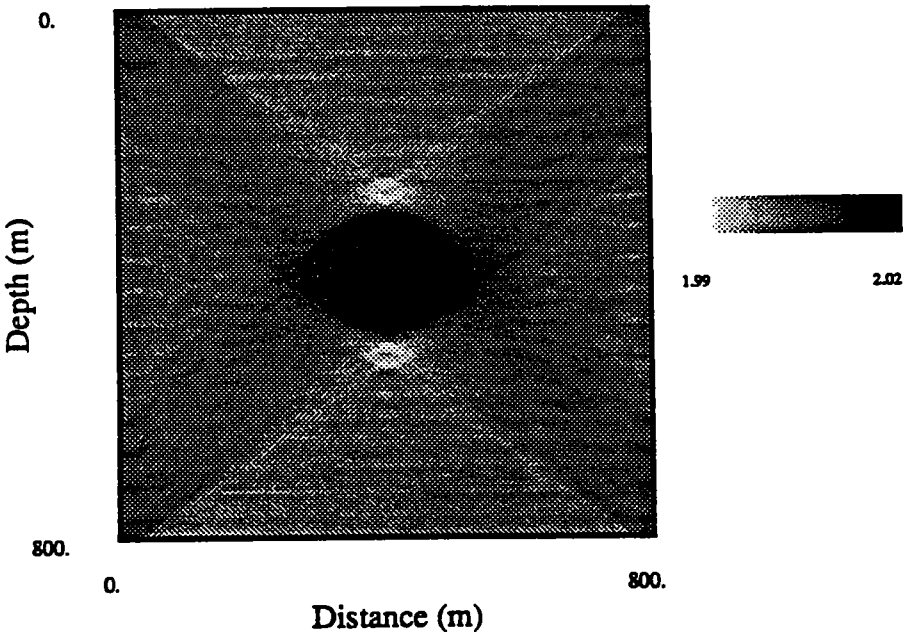

Figure 6: Inversion when a grid of 161 X 161 square pixels is used.

pixels is 289 X 289 and these dimensions are independent of the level of resolution of the image.

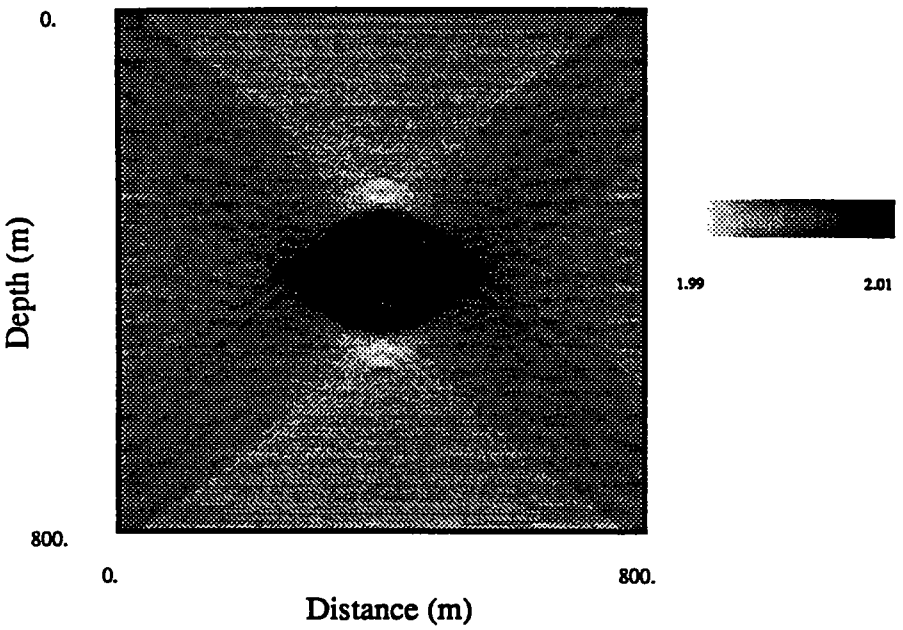

Figure 7: Inversion when the model is discretized in natural pixels. The image is displayed in a grid of 161 X 161 cells.

Both images look almost identical in terms of resolution. The artifacts produced by a coarse sampling of the model (Fig. 5) has been reduced. We observe that the discretization along the natural pixels does not contribute to eliminate limited view problems in the inversion, since we can see them in both cases. The main difference between the two solutions is related with the smoothness of the image. The reconstruction with the square pixels produces an slightly smoother image than the reconstruction with the natural pixels.

For comparing the results of the inversion with square and natural pixels, we plotted the absolute value of the difference between the original image (Fig. 4) and the inverted ones (Fig. 6 and 7). The results are shown in Fig. 8 and 9. The maximum error obtained with natural pixels is the same as obtained with square pixels. We expect both images to be roughly the same since both discretizations, minimize the expression (2). This expression represents the norm of the null space of the problem $f_2(x, y)$. If we discretize the original image very densely (grid size = 889 X 889), we can calculate this norm in both reconstructions. For the square pixels the result is $\|f_2\| = 2.383$ and for the natural pixels $\|f_2\| = 2.397$. The norm of the null space for the inversion with less square pixels (Fig. 5) is $\|f_2\| = 2.471$. This means that although the dimensions of the null space increases by sampling the image more densely in the square pixels based inversion, its norm decreses.

The noisy appearance of the inversion with natural pixel reconstruction can be

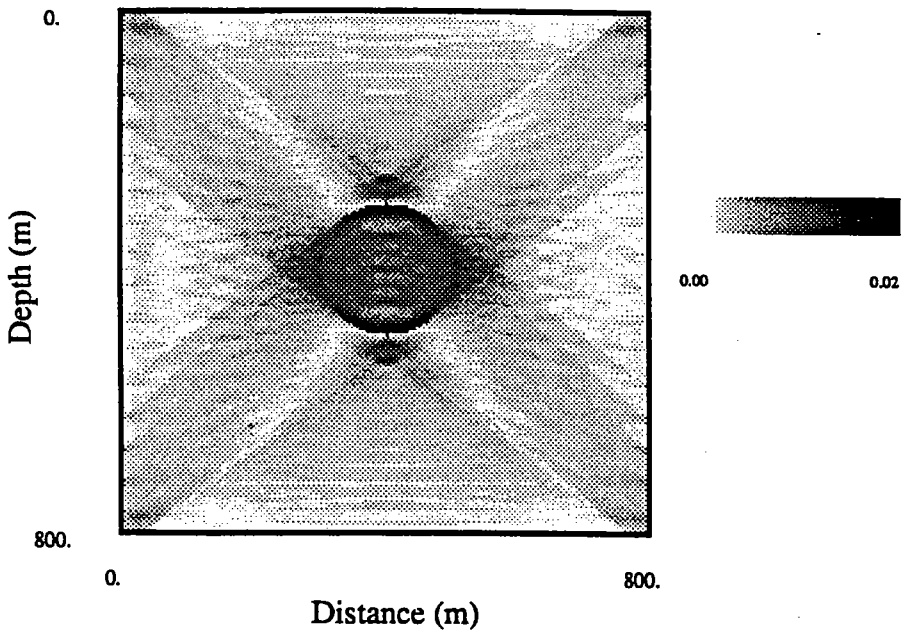

Figure 8: Error in the inversion with square pixels. This image represents the absolute value of the difference between the original model (Fig. 4) and the inverted one (Fig. 6).

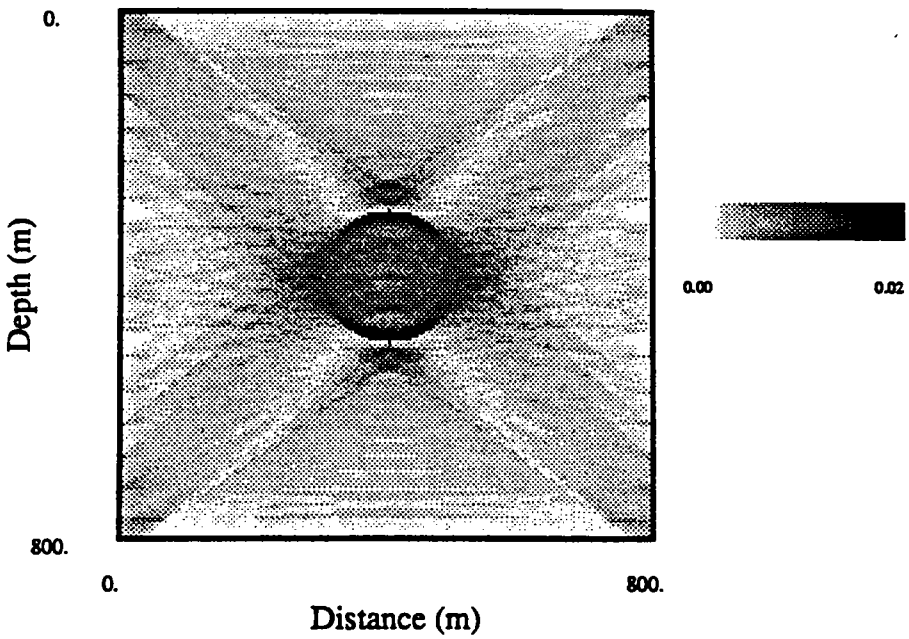

Figure 9: Error in the inversion using natural pixels. This image represents the absolute value of the difference between the original model (Fig. 4) and the inverted one (Fig. 7).

reduced by sampling the beam paths more densely. The result is shown in Fig. 10, where a grid of 401 X 401 points has been used. The norm of the null space of this image is $\|f_2\| = 2.394$. We have to say here that the extra computations necessary for producing the image from the digitized beam paths (Eqn. (21)), is negligible compared with the computation of the matrix elements and the solution of the system of equations for the coefficients of the square pixels. Once the coefficients a_n for the natural pixels are obtained, the image can be displayed using different grid sizes without having to build a new matrix and solve the system of equations again.

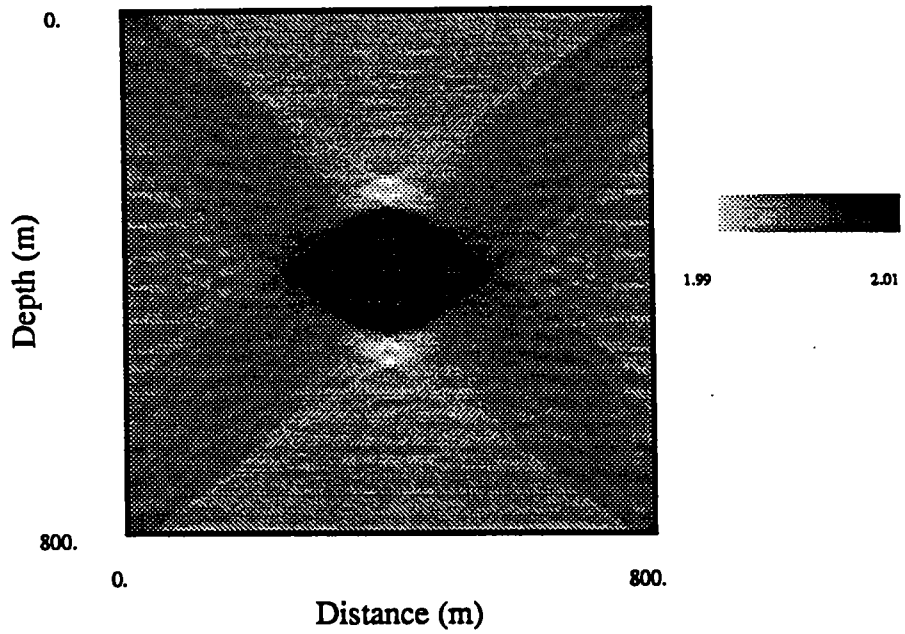

Figure 10: Inversion when the model is discretized in natural pixels. The image is displayed in a grid of 401 X 401 cells.

We can compare also the mean absolute error in both images from the following expression

$$error = \frac{1}{N}\sum_{i=1}^{N} | \, (original\ pixel)_i - (reconstructed\ pixel)_i \, |, \qquad (30)$$

where N in this case represents the total number of cells.

For the natural pixels as well as for the square pixels, the mean error is $2.0\ 10^{-3}$. However, remember that although the quality of the inversion is basically the same for both basis functions, the computational effort necessary in the whole process is roughly two orders of magnitude smaller when natural pixels are used and both images are densely sampled with the same number of points.

As we said before, the sizes of the matrices involved in the previous inversions are 289 X 1681 and 289 X 25921 for the square pixels and 289 X 289 for the natural

pixels. The first and the last matrices allow the computation of the singular values in a reasonable time. The results are shown in Fig. 11, were the curve labeled with "3" (upper curve) refers to the natural pixels discretization and the curve "1" to the 289 X 1681 matrix obtained when square pixels are used. Curve "2" represents the singular values of the matrix obtained when the model is discretized in a grid of 17 X 17 square pixels (size of the matrix 289 X 289, the same as the matrix in curve "3"). We notice that the matrix computed as intersections of natural pixels (curve "3") is significantly better conditioned that the other two matrices. Sampling the image more densely also makes the square pixels matrices (curves "1" and "2") better conditioned.

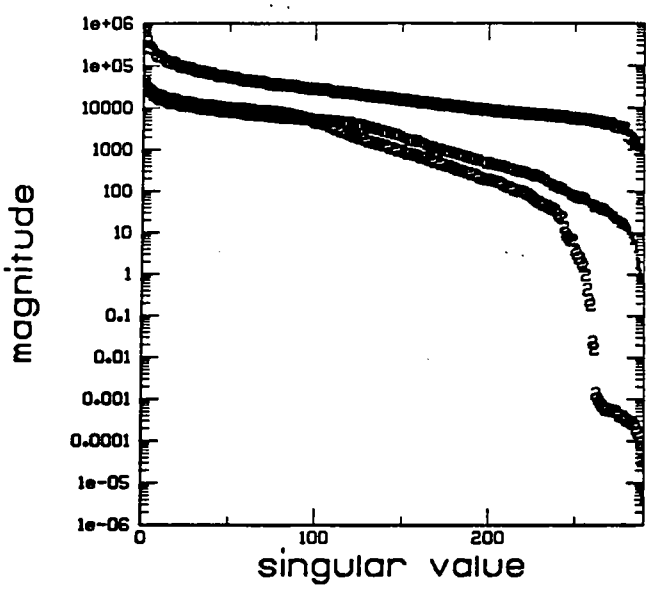

Figure 11: Singular value decomposition for the matrices obtained with different the discretizations: 1) 41 X 41 square pixels, 2) 17 X 17 square pixels and 3) natural pixels.

CONCLUSIONS

We have shown that the natural pixels provide an efficient way of discretizing the slowness model in the problem of traveltime tomographic inversion. In the examples studied, images of similar quality were obtained using natural pixels compared with the traditional reconstruction of square pixels. The main advantage of the natural pixels is that the number of model parameters needed is two orders of magnitude smaller, which means a proportional reduction on the computational effort. Besides that, the inversion with the natural pixels is better conditioned than the inversion with square pixels when comparable amount of parameters are used.

For obtaining the estimate $\tilde{S}(x, y)$ with natural pixels, two different minimizations problems are involved. The first one is implied by Eqn. (9), where the function to be minimized is the error between the true *model* and the estimated one. The second minimization problem is related to the solution of the system of equations (22), where the function to be minimized is the error between observed and calculated *data*. We have shown in this paper the importance of the first minimization in terms of the computational effort required to obtain the the final estimate.

The number of natural pixels equals the number of data points. It means that the number of model parameters in the inversion remains *constant* for a fixed amount of data, regardless the spatial dimensions of the problem or the resolution of the display. Consequently, the natural pixels provide a direct procedure for inversion in three dimensions, problems that can be computationally impossible to attack if the model is described with orthogonal three dimensional pixels (boxes).

The discretization of the model along the beam paths will change in general from one experiment to another and from one iteration to another one within the same inversion, since the beam paths may change in each of these situations. In this sense, we can say that the discretization is flexible because it depends on the given data. However, some flexibility is lost if we want to introduce in the model information not described by beam paths such as known boundaries or slowness in some areas.

More research has to be done for determining the most appropriate function that approximates the beam paths (instead of the natural pixels described herein), depending on the characteristics of the data set. Woodward (1989) gives important indications about this problem describing the beam paths (wavepaths) as elliptical, multiple-Fresnel-zone patterns, analogous to the migration ellipses. The width of her wavepaths is inversely proportional to the bandwidth and it is independent on the central frequency. Woodward's wavepaths are calculated using finite differences which might not be convenient when large data sets are inverted with our method of discretization.

ACKNOWLEDGMENTS

We would like to thanks Spyros Lazaratos for many interesting discussions and important suggestions. The first author thanks INTEVEP, S.A. for the financial support.

REFERENCES

Aki, K., and Richards, P. G., 1980, Quantitative seismology: W. H. Freeman and Co.

Berberian, S. K., 1976, Introduction to Hilbert space: Chelsea Publ. Co.

Buonocore, M. H., 1981, Fast minimum variance estimator for limited angle computed tomography image reconstruction: Ph.D thesis, Stanford University.

Buonocore, M. H., Brody, W. R., and Macosvski, A., 1981, A natural pixel decomposition for two dimensional image reconstruction: IEEE Trans. Biomedical Engineering, **BME-28**, 69-78.

Claerbout, J. F., 1976, Fundamentals of geophysical data processing: McGraw-Hill.

Darling, A. M., Hall, T. J., and Fiddy, M. A., 1983, Stable noniterative object reconstruction from incomplete data using a priori knowledge: J. Opt. Soc. Am., **73**, 1466-1469.

Harlan, W. S., 1989, Tomographic estimation of seismic velocities from reflected raypaths: presented at the 59th Ann. Internat. Mtg., Soc. Expl. Geophys.

Ivansson, S., 1985, A study of methods for tomographic velocity estimation in the presence of low velocity zones: Geophysics, **50**, 969-988.

McMechan, G. A., 1983, Seismic tomography in boreholes: Geophys. J. Roy. Astr. Soc., **74**, 601-612.

Nolet, G., 1985, Solving or resolving inadequate and noisy tomographic systems: J. Comp. Phys., **61**, 463-482.

Nolet, G., 1987, Seismic wave propagation and seismic tomography, *in* Nolet, G., Ed., Seismic Tomography: D. Reidel Publ. Co., 1-23.

Stakgold, I., 1979, Green's functions and boundary value problems: John Wiley & Sons, Inc.

Van der Sluis, A., and Van der Vorst, H. A., 1987, Numerical solution of large, sparse linear algebraic systems arising from tomography problems, *in* Nolet, G., Ed., Seismic Tomography: D. Reidel Publ. Co., 49-84.

Van Trier, J., 1988, Migration velocity analysis using geological constraints: presented at the 58th Ann. Internat. Mtg., Soc. Expl. Geophys.

Woodward, M. J., 1989, Wave equation tomography: Ph.D thesis, Stanford University.

APPENDIX A: TRAVELTIME CALCULATION ALONG NATURAL PIXELS

We explain in this appendix how to compute accurately the integrals along the

natural pixels in any slowness model after the rays has been traced.

Equation (20) is a surface integral (in 2D). Surface integrals are in general difficult to evaluate, specially if there is no analytical expression for the integrand. A conceptually simple way of evaluating this integral is based on a fine discretization of the model, evaluating and summing the contribution of points that belong to the support of each particular natural pixel. The final sum is divided by the width of the natural pixel. The main disadvantage of this method is that a very fine discretization of the model is needed to ensure accuracy in the traveltimes, even when the medium is homogeneous. Fig. A1 (thick curve) shows an example of the traveltimes calculated with this method in an homogeneous medium.

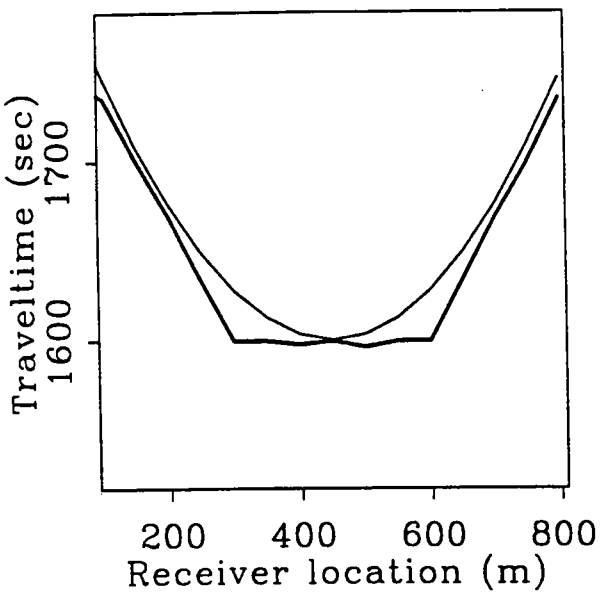

Figure A1: Traveltimes in a homogeneous medium calculated using natural pixels. The thin curve is calculated multiplying the step of the ray by the average velocity in a perpendicular segment. The thick curve is calculated by discretizing and summing for all the points that belong to the support of each natural pixel.

A more precise way of evaluating numerically Eqn. (20) is simply by summing the products of the length of the segment between two consecutive points of the ray by the average slowness in a line perpendicular to that segment. The line has *length = width of the tube.* This is simply a generalization of the traveltimes calculated from expression (19) as the sum of products *slowness x length.* An example of this technique is shown in Fig. A1. Note that the shape of the curve is the expected hyperbola (thin curve), which is not the case (thick curve) in the previously explained method, where a very fine discretization is needed for the hyperbolic shape to be reproduced. In the example of Fig. A1 the discretization used for calculating the thick curve is approximately 60 times denser than the one used for the other curve.

Fig. A2 compares the traveltimes anomalies (i.e. *travetime in heterogeneous medium - traveltime in homogeneous medium*) calculated using Eqn. (19) and Eqn. (20). The thick curve represents anomalies calculated as integrals along lines (Eqn. (19)) and the thin curve represents anomalies calculated as integrals along natural pixels. The source is located at 400 m in the model of Fig. 4. Note that the anomaly calculated with the natural pixels is smoother than the one calculated integrating along the ray path. Usually, the model is smoothed before tracing the rays, and this difference might be eliminated. Another difference between the two cases (which cannot be solved by smoothing the model) is that the traveltimes calculated with natural pixels affect a larger range of receivers than the ones calculated with rays. This behavior better resembles the real measurements. The thin curve is not completely symmetric because inaccuracies in the interpolation for calculating the borders of the natural pixels and also due to roundoff errors.

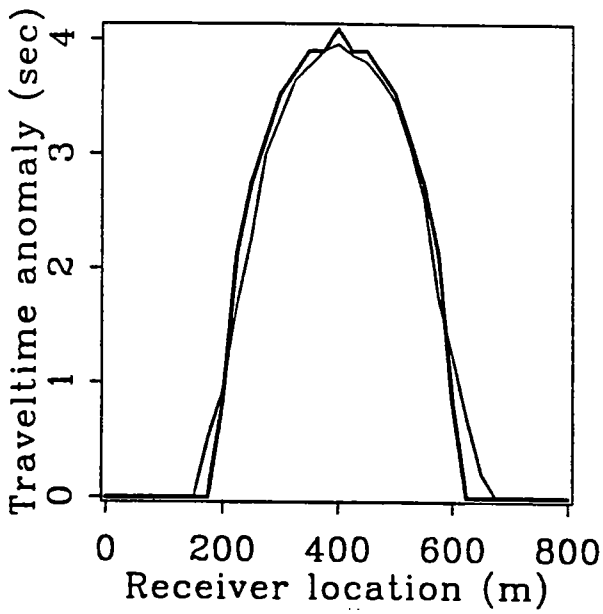

Figure A2: Traveltime anomalies calculated for the model of Fig. 4. The source is in the left of the model at 400 m depth. Receivers are in the opposite well. The thick curve is calculated with conventional ray tracing. The thin curve is calculated multiplying the step of the ray by the average velocity in a perpendicular segment.