Chapter 4

Interval-velocity estimation for 2-D prestack depth migration

4.1 Velocity analysis: an optimization problem

I pose interval-velocity estimation as an optimization problem and try to find the interval-velocity model that gives the best prestack-depth-migrated image of a data set. As discussed in Chapter 1, there are many possible definitions of "best image." I define the best image to be the one whose events stack most coherently after prestack depth migration.

The relation between an interval-slowness model and the images of events after prestack depth migration with that slowness model is complicated and nonlinear. Rather than optimizing the interval slowness directly by varying images with prestack depth migration, I migrate constant-offset sections with an initial interval-slowness model w_o. I then measure residual moveout using residual NMO+DMO described in Chapter 2. Governed by the theory of Chapter 3, I solve an optimization problem to find a change to the interval-slowness model Δw that explains the measured residual slownesses.

This update to the interval-slowness model can be combined with the initial slowness model and used to depth-migrate the data again. Since residual prestack time migration and the relations between interval slowness and residual slowness I use are approximations to residual prestack depth migration, the new interval-slowness model might not be optimal when used to migrate the data. As long as the new model is closer to the correct model, the process can be iterated because the various approximations become more accurate.

4.1.1 Objective function

Data component

Residual-NMO+DMO correction after prestack depth migration builds a velocity-analysis cube consisting of semblance versus residual slowness for all events. The value of residual slowness that has the maximum semblance for a given event describes the kinematics of a residual-migration operator that best fits the event's residual moveout. One possible approach to formulating an optimization problem is to pick the peaks of the semblance for each event. The optimal change to the interval slowness is the one that best predicts the picked residual slownesses using the operator **G** from Chapter 3. The goodness or badness of a change to the interval-slowness model is evaluated by the distance between the peaks it predicts and the picked peaks; this gives a least-squares objective function.

As discussed in Chapter 1, I prefer not to pick, either moveout or peaks of semblance. Toldi (1985) showed that the semblance function itself is a more robust objective function than measuring distance from picks for two reasons. First, the pick itself is uncertain and requires a judgment call. Second, once far enough away from a peak, the semblance values approach zero and do not "feel" the peak. The quadratic function fit through the peak implied by a least-squares objective function feels the peak infinitely far away. The effect of a bad or inconsistent pick can never be ignored in a least-squares algorithm, but it is easily ignored using semblance itself as an objective function.

Build the data component of the objective function by finding the residual slowness for each event given a change in interval slowness (using **G**) and summing the semblance values at the appropriate residual slowness for all events.

$$Q_d(w) = \sum_{\xi,\eta} S(\gamma(\xi,\eta;w),\xi,\eta) \quad . \tag{4.1}$$

S is the semblance at a given location (ξ,η) in the data space for a given γ, and γ at a particular event (ξ,η) is a function of the interval-slowness model w.

Rather than using all possible locations in the migrated image, it is possible to select the most prominent events and only sum the semblance for those events. Using only prominent events saves time and storage when computing the operators of Chapter 3. Furthermore, by avoiding multiples or other undesired events, the velocity information of the objective function is more consistent with the primary-reflection model used by prestack

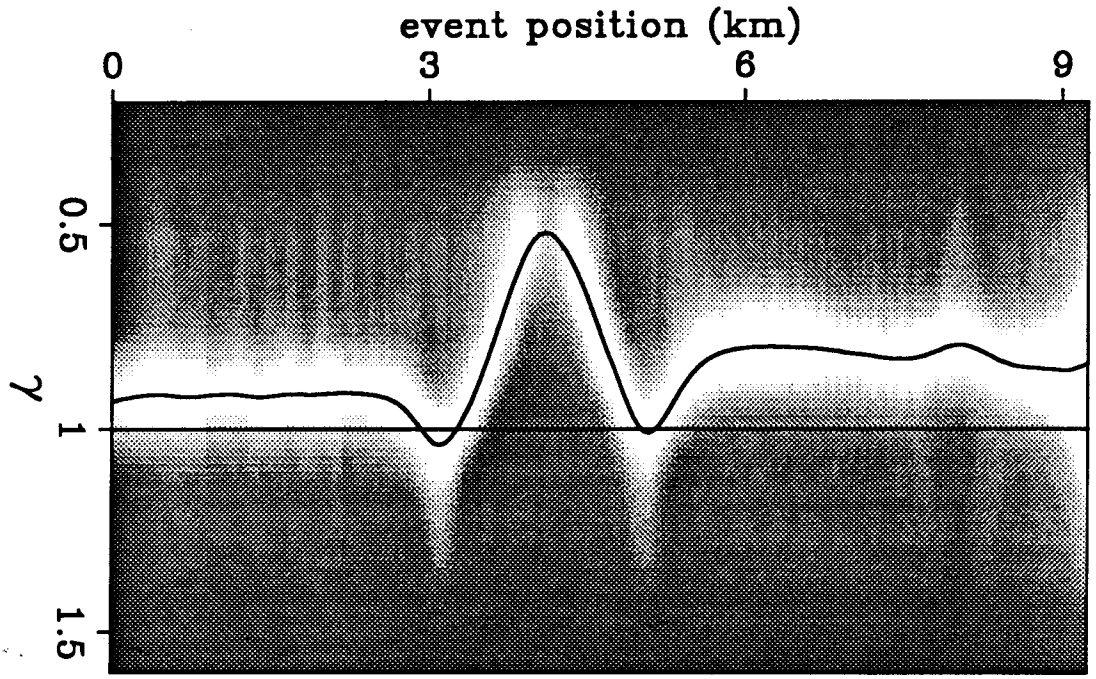

FIG. 4.1. The objective function is evaluated by summing semblance along a trajectory of γ vs. reflector. Summing along the curved line will give a higher value than summing along the straight line. The maximum possible value of the objective function is obtained when γ vs. reflector lies on the peaks of the semblance panel.

depth migration. Parameterize ξ and η as functions of picked reflector r, $\xi(r), \eta(r)$.

$$Q_d(w) = \sum_{r=\text{reflectors}} S(\gamma(\xi(r), \eta(r); w), \xi(r), \eta(r)) \ . \tag{4.2}$$

Figure 4.1 shows semblance versus residual slowness extracted along a single reflector. The dark curves are the γ's predicted by **G** for two different changes to the interval-slowness model. Changes to the interval-slowness model that predict residual slownesses near the peaks of the semblance (curved line) will give a large value for the sum in equation (4.1). Changes that predict residual slownesses away from the peaks (straight line) give smaller values for Q_d. The maximum value of Q_d occurs when the curve of predicted γ's lies exactly along the peaks of the semblance.

Semblance versus residual slowness can be sharply peaked when velocity resolution in the migrated data is good; other more sophisticated velocity measures (Key and Smithson,

1990) can have even higher resolution than semblance. If the initial velocity model does not put the peaks of semblance (or other coherency measure) near $\gamma = 1$, it may be difficult for an optimization method to find the peaks. This problem is easily solved by smoothing the semblance over γ until the optimization method finds the peaks. After the peaks are found, the smoothing can be removed in pursuit of higher resolution.

Model component

If reflectors are sparse, it is possible for several different interval-slowness models to predict the same residual slownesses (caused by the null space of $\mathbf{G}_\gamma$; see Figure 3.12). The objective function given by equation (4.2) has no way to discriminate between two different changes to the interval-slowness model if they produce the same residual slowness. We can add terms to the objective function that prefer some types of interval-slowness models over others. To favor smooth interval-slowness models add terms to penalize to the norm of the x and z derivatives of the slowness model.

$$Q_m(w) = \epsilon_x \left\| \frac{\partial w}{\partial x} \right\| + \epsilon_z \left\| \frac{\partial w}{\partial z} \right\| \quad . \tag{4.3}$$

$\|\cdot\|$ implies integration of the norm over the whole model. Take ϵ_x and ϵ_z to be negative to decrease the objective function when the norms are non-zero. It is important to penalize the entire slowness model $w = w_o + \Delta w$ and not just the change to the slowness model, otherwise the total slowness model may not be smooth if the penalty only applies to the change to the interval slowness. To keep the interval-slowness model near an a priori model, add a term that penalizes the distance of the computed model from the a priori model.

$$Q_m(w) = Q_m + \epsilon_a \|w - w_a\| \quad . \tag{4.4}$$

As above, ϵ_a is negative to decrease the objective function as the computed model deviates from the a priori model.

Different constraints are appropriate for different a priori information about the velocity model for a given data set. Equation (4.3) would be more appropriate when little is known about the velocity or the velocity information in the data is incomplete or self-conflicting. Equation (4.4) is appropriate for including information from check shots or well logs.

The total objective function to be maximized is obtained by adding the data component and model component.

$$Q(w) = Q_d(w) + Q_m(w) \quad . \tag{4.5}$$

The strengths of the constraints relative to each other and to the data component of the objective function are modified by varying the sizes of ϵ_x and ϵ_z in equation (4.3) and ϵ_a in equation (4.4). The values of the various ϵ's can be made a function of position if desired.

4.1.2 Model parameterization

Computations can be performed in terms of slowness or velocity. Although I often refer to the computations or to the entire thesis as "velocity analysis," I use slowness in computations since the traveltime-tomography calculations at the heart of the method are more nearly linear when slowness is used to describe the medium (see equation (3.2)).

Velocity or slowness models are often represented on a fine grid of points with the same sampling interval as the output sampling intervals of migration programs or other processing programs. Although this parameterization allows almost any model to be described, it is difficult to use for interval-velocity estimation. The migration algorithms I use are based on high-frequency assumptions, so the effects on wave transmission of high wavenumbers in the velocity model are neglected. Furthermore, since I assume the movement of reflectors caused by changes in the interval-velocity model can be approximated by a residual time migration, high wavenumbers in the update to the interval-slowness model will not be accounted for properly. Finally, a smooth parameterization of the interval-slowness model results in fewer unknowns to estimate. This makes solving the optimization problem less expensive and better conditioned than using a fine-grid representation of the model.

Cubic splines are a convenient parameterization for the model. If a point (x, z) in the model falls between nodes i and $i+1$ in the x direction and between j and $j+1$ in the z direction, the slowness at the point can be evaluated as a linear combination of the 16 surrounding nodes.

$$w(x,z) = \sum_{l=-1}^{2} \sum_{k=-1}^{2} C_l(x - x_i) C_k(z - z_j) \mathbf{w}_{i+l,j+k} \quad . \tag{4.6}$$

The slowness model is described by a vector of node values $\mathbf{w}$, where a single node is

denoted $\mathbf{w}_{i,j}$. C_l and C_k are the spline coefficients and are a function of distance between nodes in x and z.

Smooth parameterizations imply that "blocky" velocity models, the kind often used as interpreted models, can't be described; because they explicitly couple both high and low-wavenumber information about the velocity model. Although smooth velocity models may not look like geologic interpretations, they can contain most of the traveltime information that can be resolved from seismic reflection data. High-wavenumber information such as locations of formation boundaries or sharp changes in interval velocity do not affect the traveltime of waves transmitted through those features.

4.2 Updating the interval-slowness model

Nonlinear optimization problems like the one posed in the previous section can be attacked with a variety of methods such as gradient-ascent methods or simulated-annealing methods (Rothman, 1985). I chose to use a gradient-ascent method rather than simulated annealing because nonlinear forward modeling using the operator of Chapter 3 is expensive and gradient-ascent methods usually require fewer function evaluations and often give worthwhile partial results.

4.2.1 Gradient of the objective function

A key ingredient of all gradient-ascent methods is the gradient of the objective function with respect to the model. In classical tomography problems this is called the back-projection operator. For the optimization problem of the previous section, the gradient has two components: the data component which comes from the semblance cube and the model or penalty component.

$$\nabla_{\mathbf{w}} Q \;=\; \nabla_{\mathbf{w}} Q_d \;+\; \nabla_{\mathbf{w}} Q_m \;. \tag{4.7}$$

The gradient of the model component of the objective function can be calculated analytically using the cubic-splines formula or by simple differencing formulas applied to the node values. For example, the term involving penalties on the x derivative can be

approximated with

$$\nabla_{i,j} Q_m = -2\epsilon_x D_x^+ \mathbf{w}_{i,j} - 2\epsilon_x D_x^- \mathbf{w}_{i,j} \quad ; \tag{4.8}$$

where D_x^+ and D_x^- are forward and backward differences in x.

Applying the chain rule, write the i,j component of the gradient of the data part of the objective function with respect to the interval-slowness model.

$$\nabla_{i,j} Q_d = \sum_{\text{reflectors}} \frac{\partial S(\gamma,\xi,\eta)}{\partial \mathbf{w}_{i,j}} = \sum_{\text{reflectors}} \frac{\partial S(\gamma,\xi,\eta)}{\partial \gamma} \frac{\partial \gamma(\xi,\eta)}{\partial \mathbf{w}_{i,j}} \quad . \tag{4.9}$$

ξ, η are functions of reflector position as in equation (4.2). Thus, there are two steps involved in computing the gradient of the data component of the objective function. The first deals only with the semblance data space $S(\gamma,\xi,\eta)$. Since we have precomputed values of S for each ξ, η, and γ, the derivative of the objective function with respect to γ can be calculated numerically, for example with a finite-difference formula.

$$\frac{\partial S(\gamma,\xi,\eta)}{\partial \gamma} = \frac{1}{\Delta\gamma}\Big[S(\gamma+\Delta\gamma,\xi,\eta) - S(\gamma,\xi,\eta)\Big] \quad . \tag{4.10}$$

Remember that γ is a function of reflector position. The second step deals only with the relation between interval slownesses and residual slownesses. Multiplying $\delta S/\delta\gamma$ by $\partial\gamma/\partial \mathbf{w}_{i,j}$ for all i,j is equivalent to multiplying by the adjoint or transpose of the linearized operator of Chapter 3.

Write the total gradient of the objective function with respect to the interval-slowness model as

$$\nabla_{\mathbf{w}} Q = \left[\mathbf{G}_\gamma^T - \frac{\delta\gamma}{\delta\tau}\mathbf{G}_\tau^T\right]\nabla_\gamma Q_d + \nabla_{\mathbf{w}} Q_m \quad . \tag{4.11}$$

Since $\mathbf{G}_\gamma$ and $\mathbf{G}_\tau$ contain a tomography operator, the gradient of Q as written in equation (4.11) is a filtered tomographic back-projection (Fowler, 1988). The filter converts the changes in stack semblance that improve the objective function into changes in traveltimes and finally into an update for the interval-slowness model.

4.2.2 Iterative velocity-analysis algorithm

The optimization problem posed in the previous section is nonlinear in several ways. First, semblance versus residual slowness for each event and thus for the image as a whole is non-quadratic. As discussed in the previous section, using semblance itself can be more robust than trying to least-squares fit picked values of residual slowness; but the non-quadratic nature of the objective function can lead to slow convergence of the solution to the optimization problem if methods for strictly quadratic functions are applied. Second, the relation between changes in interval slowness and residual slowness is nonlinear due to the movement of reflectors and the change in ray paths when interval-slowness perturbations accumulate. Finally, approximating residual depth migration with residual time migration is only valid for a certain range of perturbation to the interval-slowness model. When the interval-slowness model changes significantly it may be necessary to remigrate the data.

Taking into account the hierarchy of nonlinear effects, the following algorithm finds the interval-slowness model that minimizes residual moveout of prestack-depth-migrated data.

```
Obtain initial slowness model w_0, set i = 0
Outer loop over interval-slowness models w_i {
    Migrate the data with w_i
    Check for convergence:  no remaining residual moveout → done.
    Apply residual NMO+DMO: compute S(γ,τ)
    Pick important reflectors
    Compute the operator G
    Solve for Δw_i using conjugate-gradient method
    w_{i+1} = w_i + Δw_i
    i = i + 1
}
```

To solve for $\Delta\mathbf{w}_i$, the update to the interval-slowness model, I use the PARTAN variant of the conjugate-gradient method discussed by Luenberger (1984).

```
Find gradient:  ∇_w Q = [G_γ − (δγ/δτ) G_τ] ∇_γ Q_d + ∇_w Q_m
Line search for α that maximizes Q(α ∇_w Q)
Δw_0 = α ∇_w Q
Loop over conjugate-gradient iterations:  j = 1 {
```

Find gradient $\nabla_{\mathbf{w}} Q = [\mathbf{G}_\gamma - \frac{\delta\gamma}{\delta\tau}\mathbf{G}_\tau]\nabla_\gamma Q_d + \nabla_{\mathbf{w}} Q_m$
line search for α that maximizes $Q(\Delta\mathbf{w}_{j-1} + \alpha\nabla_{\mathbf{w}} Q)$
$\Delta\mathbf{w}' = \Delta\mathbf{w}_{j-1} + \alpha\nabla_{\mathbf{w}} Q$
line search for β that maximizes $Q(\Delta\mathbf{w}_{j-1} + \beta(\Delta\mathbf{w}' - \Delta\mathbf{w}_{j-1}))$
Update $\Delta\mathbf{w}_j = \Delta\mathbf{w}_{j-1} + \beta(\Delta\mathbf{w}' - \Delta\mathbf{w}_{j-1})$
Update reflector position map $x(\tau)$, $z(\tau)$, $\delta\gamma/\delta\tau$
Check for convergence $\nabla_{\mathbf{w}} Q \approx 0$: $\Delta\mathbf{w}_i = \Delta\mathbf{w}_j$; Break
Else $j = j + 1$
}

4.3 Synthetic example

The synthetic data of Chapter 3, section 4 were migrated with a constant interval-slowness model. We can use the algorithm presented above to solve for the interval-slowness model that best migrates the data. Figure 4.2 is the same as Figure 3.19 and shows the initial migrated and stacked image. The lower reflector has pull-up caused by ignoring the high velocity beneath the first reflector and a push-down caused by the local anomaly above the first reflector. The intensity plot shown in Figure 4.3 is the same as the intensity plot of Figure 3.20 and shows horizon residual-slowness analyses for the two reflectors in the initial image. These semblance panels serve as the data part of the objective function. I ran the inner iterations of the velocity-analysis algorithm to find a change to the interval-slowness model that predicts the γ values of the peaks of semblance. The conjugate-gradient method was run until the γ's predicted by the change to the interval-velocity model were on top of the peaks as shown by the dark line in Figure 4.3. Figure 4.4 shows the new model obtained by adding the estimated Δw to the previous constant-slowness model. This represents the result of one outer iteration of the velocity-analysis algorithm.

To verify the correctness of the new model, migrate the data with it. If the semblance peaks do not line up at $\gamma = 1$, further outer iterations are needed. Figure 4.5 shows the migrated and stacked image obtained with the updated interval-slowness model after the first iteration of velocity analysis. The new migrated reflector positions are closer to their true positions than before but still not correct. There are places on both reflectors that are still affected by the shallow anomaly, and there is long-wavelength structure on the

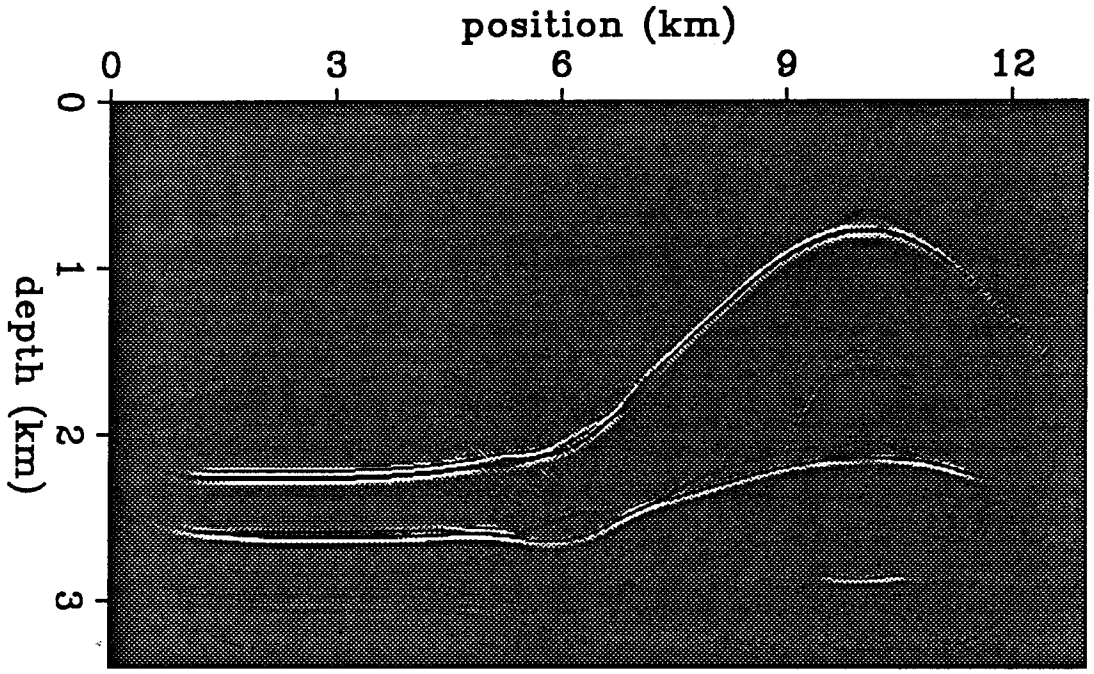

FIG. 4.2. Migrated and stacked section obtained using a constant-velocity model. The goal of velocity analysis is to estimate an interval-slowness model that migrates the data correctly.

deeper reflector. To see if there is any more velocity information in the data, follow the algorithm and build the residual-slowness analyses again. The intensity plot of Figure 4.6 shows horizon residual-slowness analyses for the two reflectors after migration with the new slowness model. Much of the long-wavelength trend in γ on the lower reflector has been removed but both reflectors still have short-wavelength variations in γ and in migrated position.

Now that the reflectors are closer to their true positions we have a better chance for estimating the short-wavelength anomaly. Since the updated model already predicts the long-wavelength features of the interval-slowness model, there is less hope that further iterations can move the lower reflector much closer to its true depth. The long-wavelength (about 10 km) error most likely corresponds to the null space of $\mathbf{G}_\gamma$ since its wavelength is about the same as the zero of the spectrum $\mathbf{G}_\gamma$ shown in Figure 3.12. If there were more reflectors, the effect of the null space of $\mathbf{G}_\gamma$ would be less dramatic since the 10 km wavelength would fall in different regions of the spectra of the operators for other reflectors.

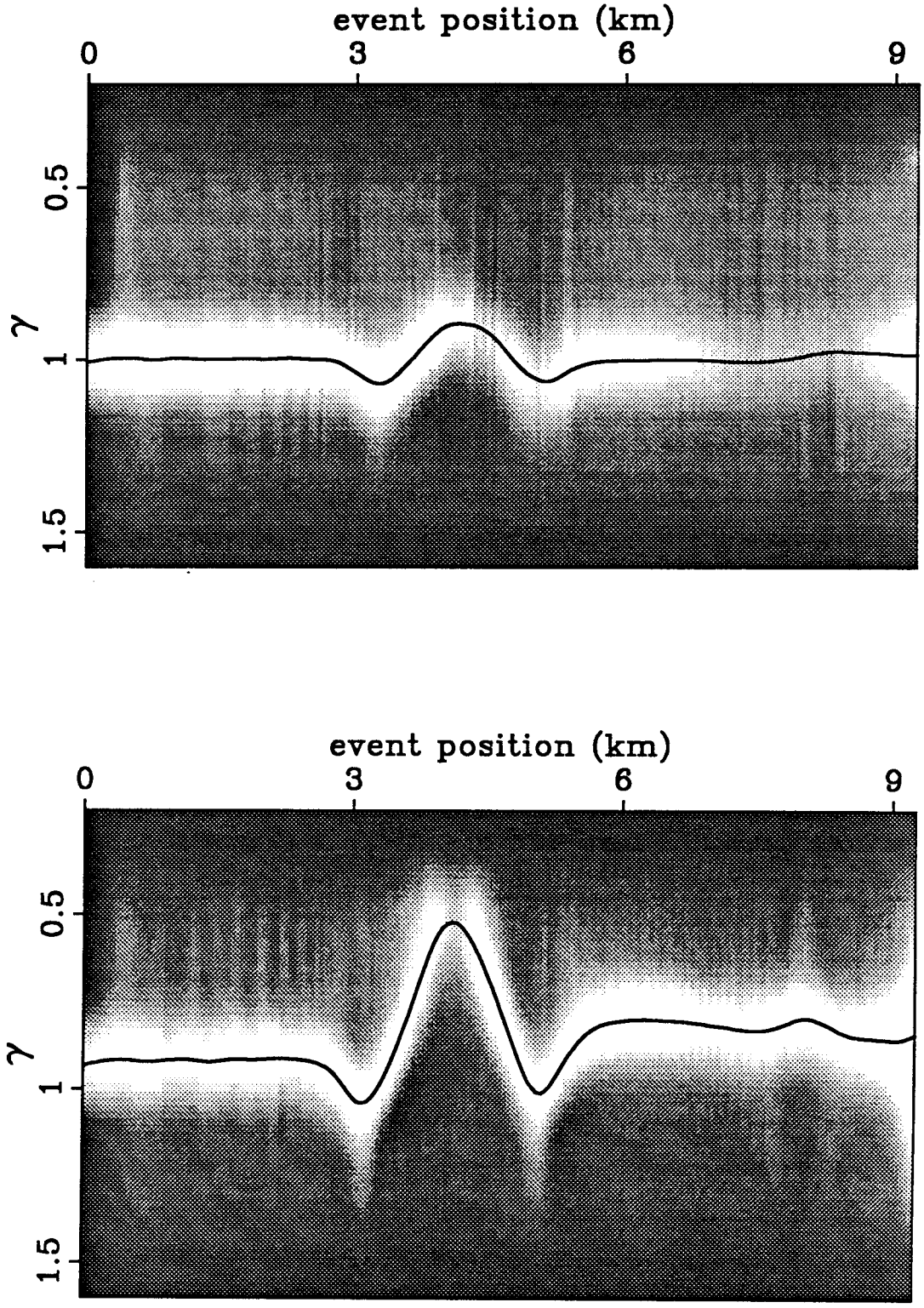

FIG. 4.3. Horizon residual-slowness analyses for the first outer iteration. The line shows the γ vs. reflector predicted by the change to the interval-slowness model.

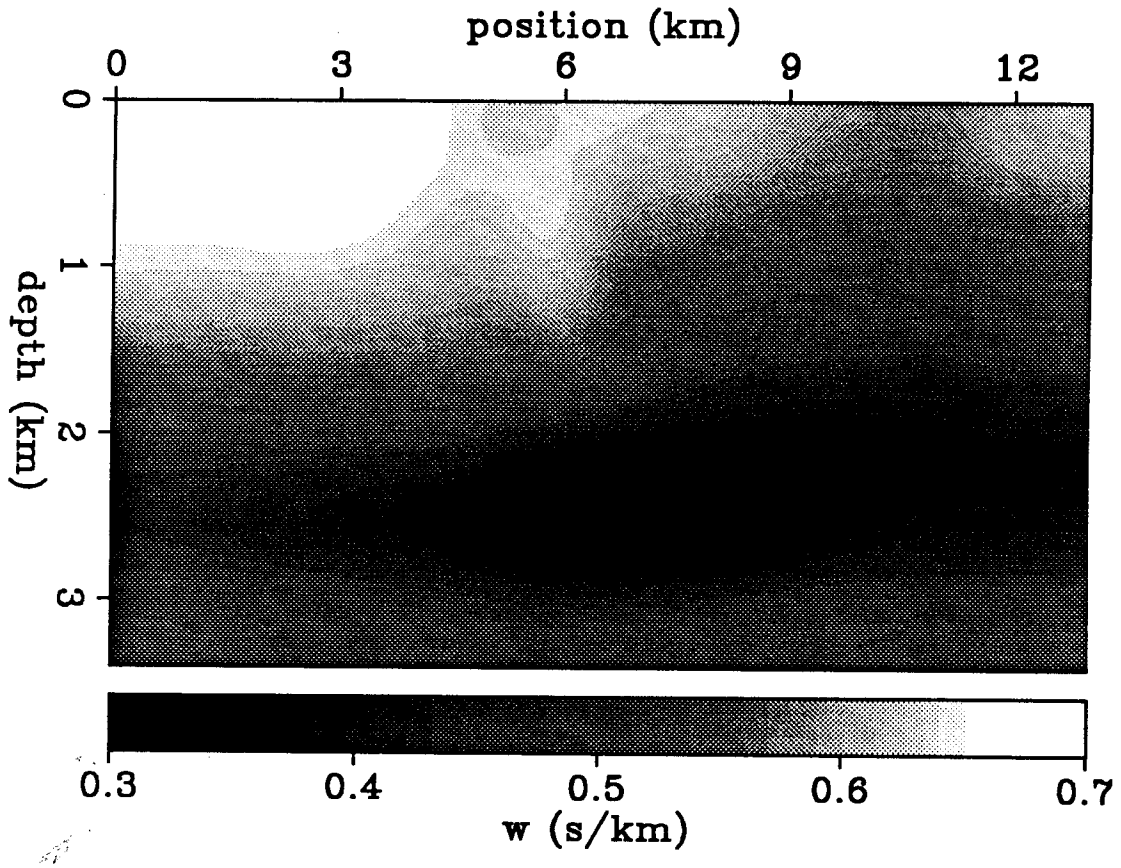

FIG. 4.4. Change to the interval-slowness model Δw estimated from residual-slowness analysis applied the reflectors of Figure 4.2

To remove the effect of the short-wavelength variations in γ and to try to make the bottom reflector flatter, I ran another outer iteration. Figure 4.7 shows the final interval-slowness model obtained after two iterations of the velocity-analysis algorithm. The curves overlaid on the semblances of Figure 4.6 show the γ's predicted by the second outer iteration of the velocity-analysis algorithm. Figure 4.8 shows the stacked section after prestack depth migration using the new (final) interval-slowness model. Some of the long-wavelength structure on the lower reflector is removed. The short-wavelength anomaly has been estimated and its effects on the two reflectors removed. Figure 4.9 shows horizon residual-slowness analyses after the final migration; the peaks of the semblance are almost exactly at $\gamma = 1$ indicating that the final model accounts for the velocity information that can be resolved using moveout over offset.

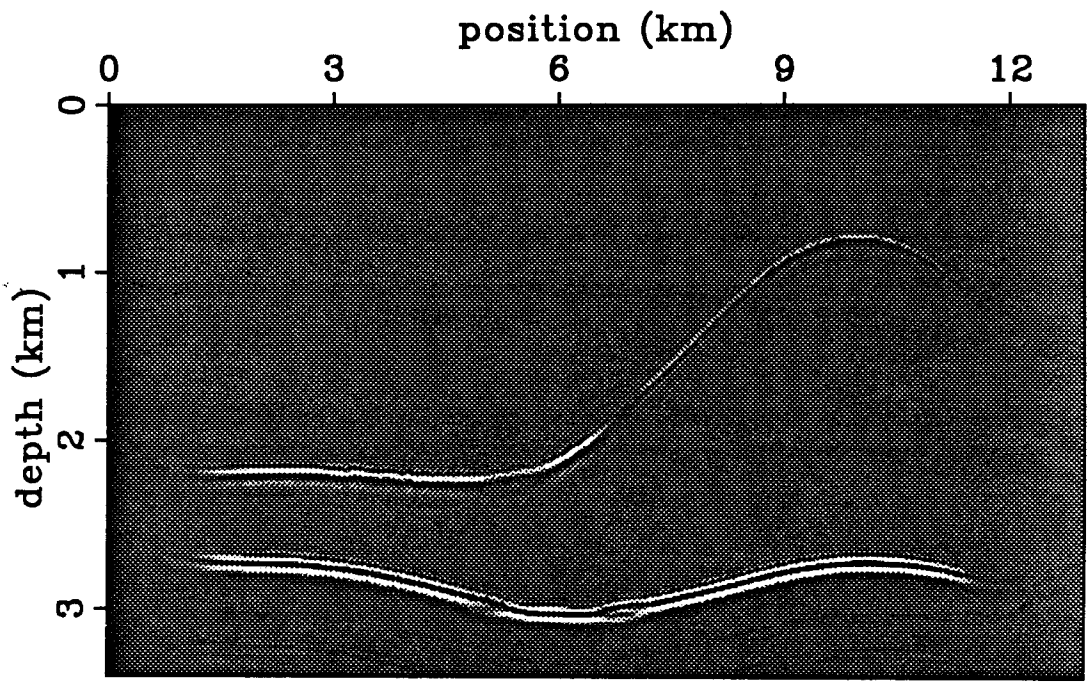

FIG. 4.5. Stacked image after prestack depth migration using the interval-slowness model obtained by one outer iteration of velocity analysis.

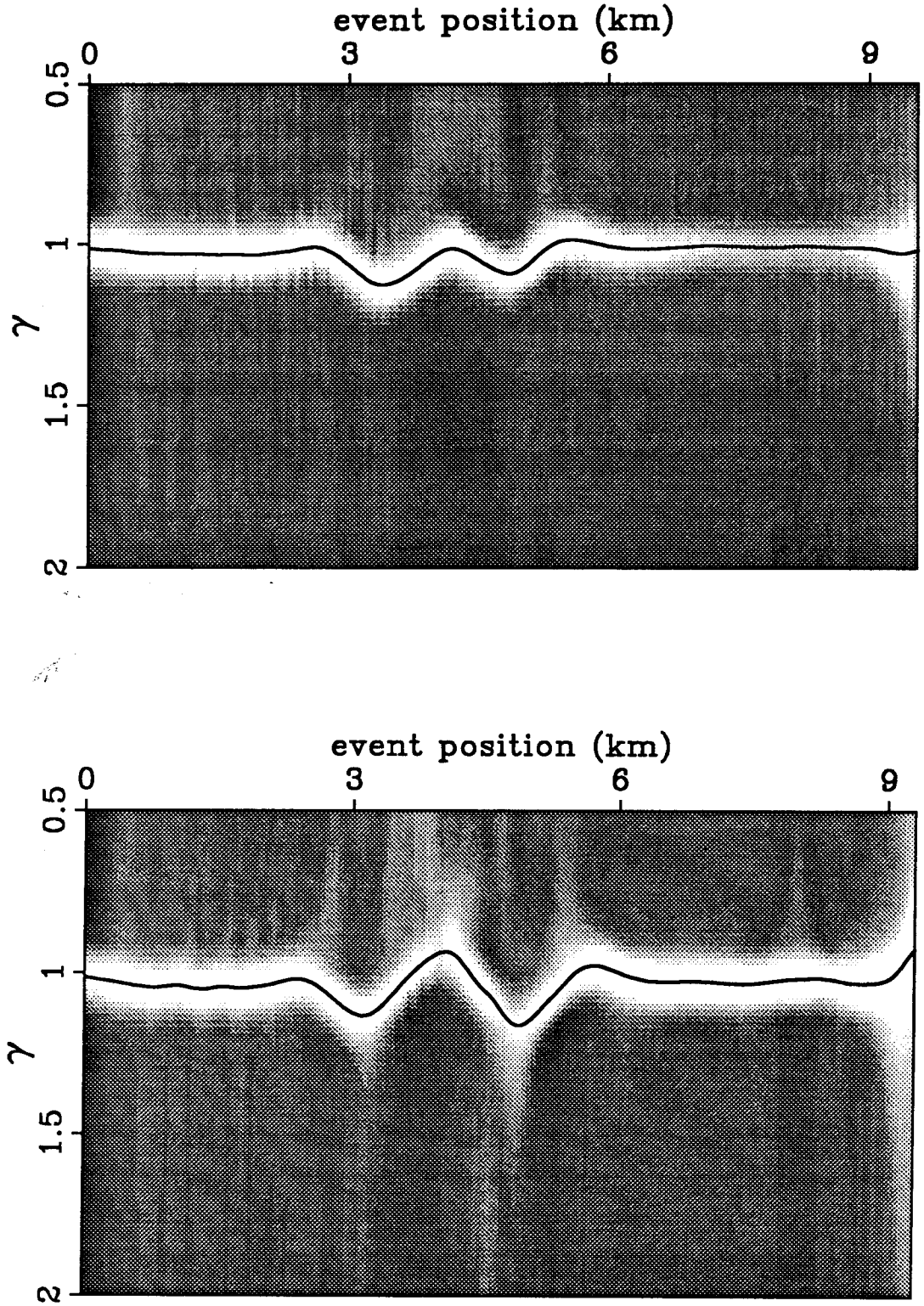

FIG. 4.6. Horizon residual-slowness analysis after migration with the first updated model. Both reflectors only show short-wavelength departures from $\gamma = 1$ indicating the long-wavelength features of the slowness model are accounted for. The curve shows the γ's predicted by the second outer iteration of velocity analysis.

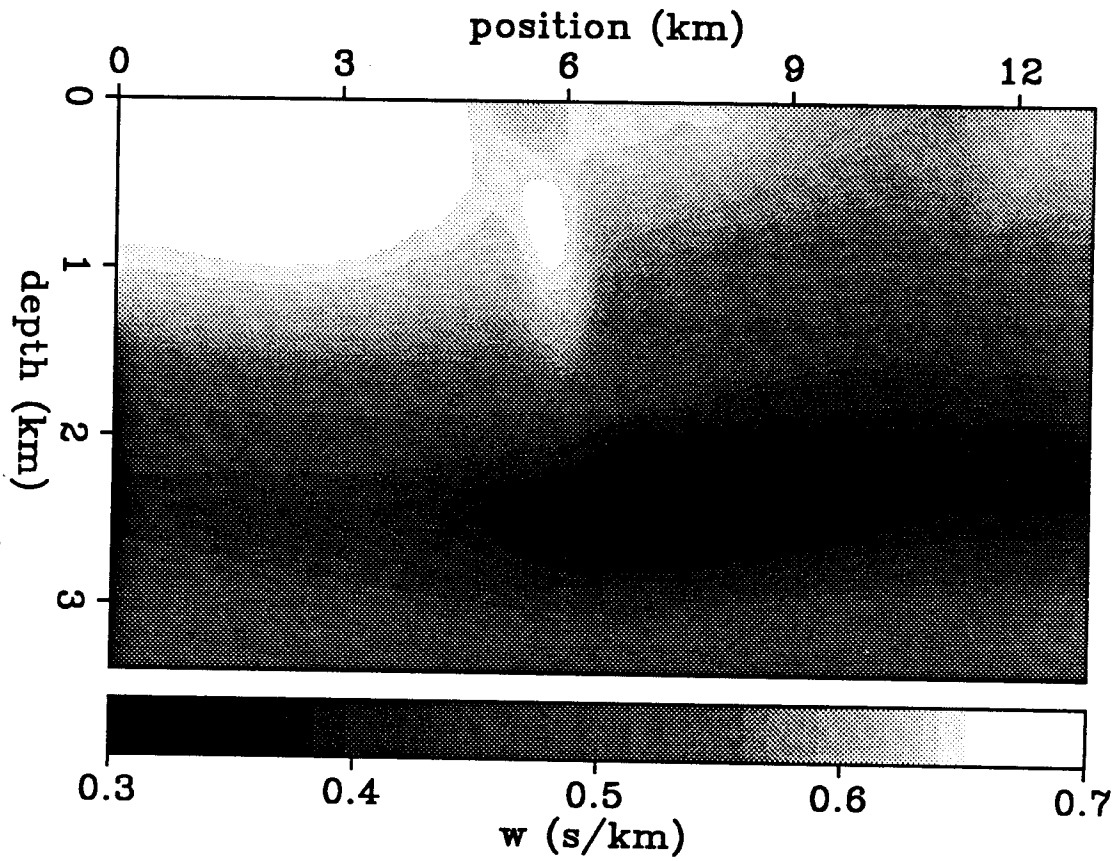

FIG. 4.7. Interval-slowness model obtained after two iterations of the velocity-analysis algorithm. The dark values are low slowness or high velocity; the light values are high slowness or low velocity.

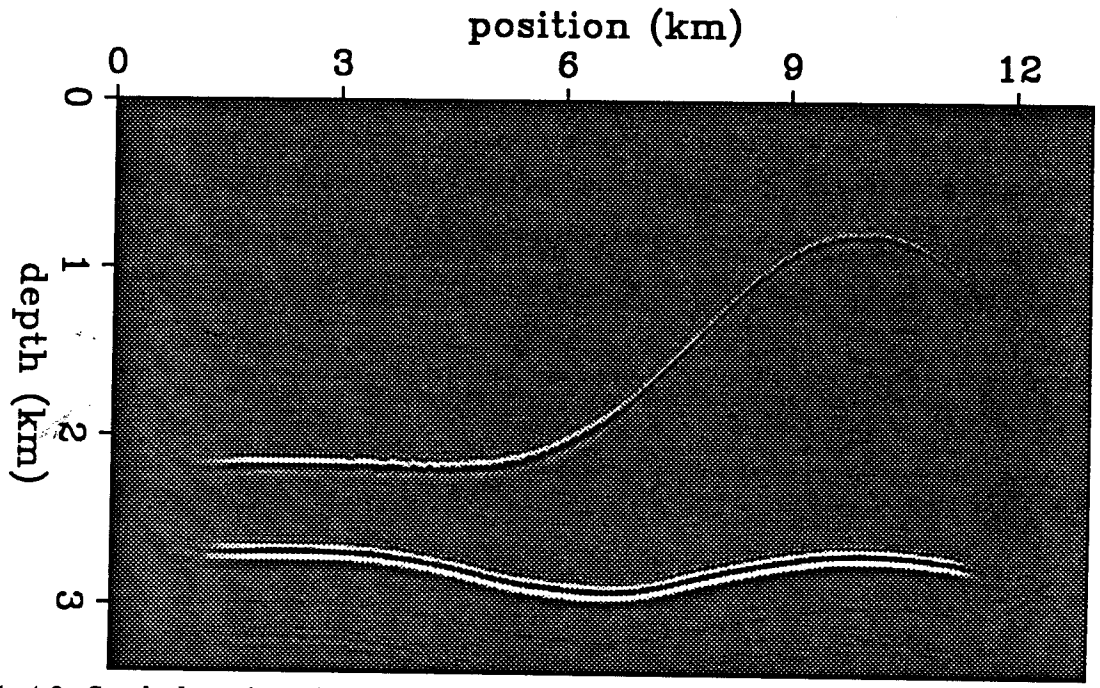

FIG. 4.8. Stacked section after prestack depth migration using the model obtained after two iterations of the velocity-estimation algorithm. Some incorrect long-wavelength structure remains on the lower reflector; most of the short-wavelength effects on both reflectors have been removed.

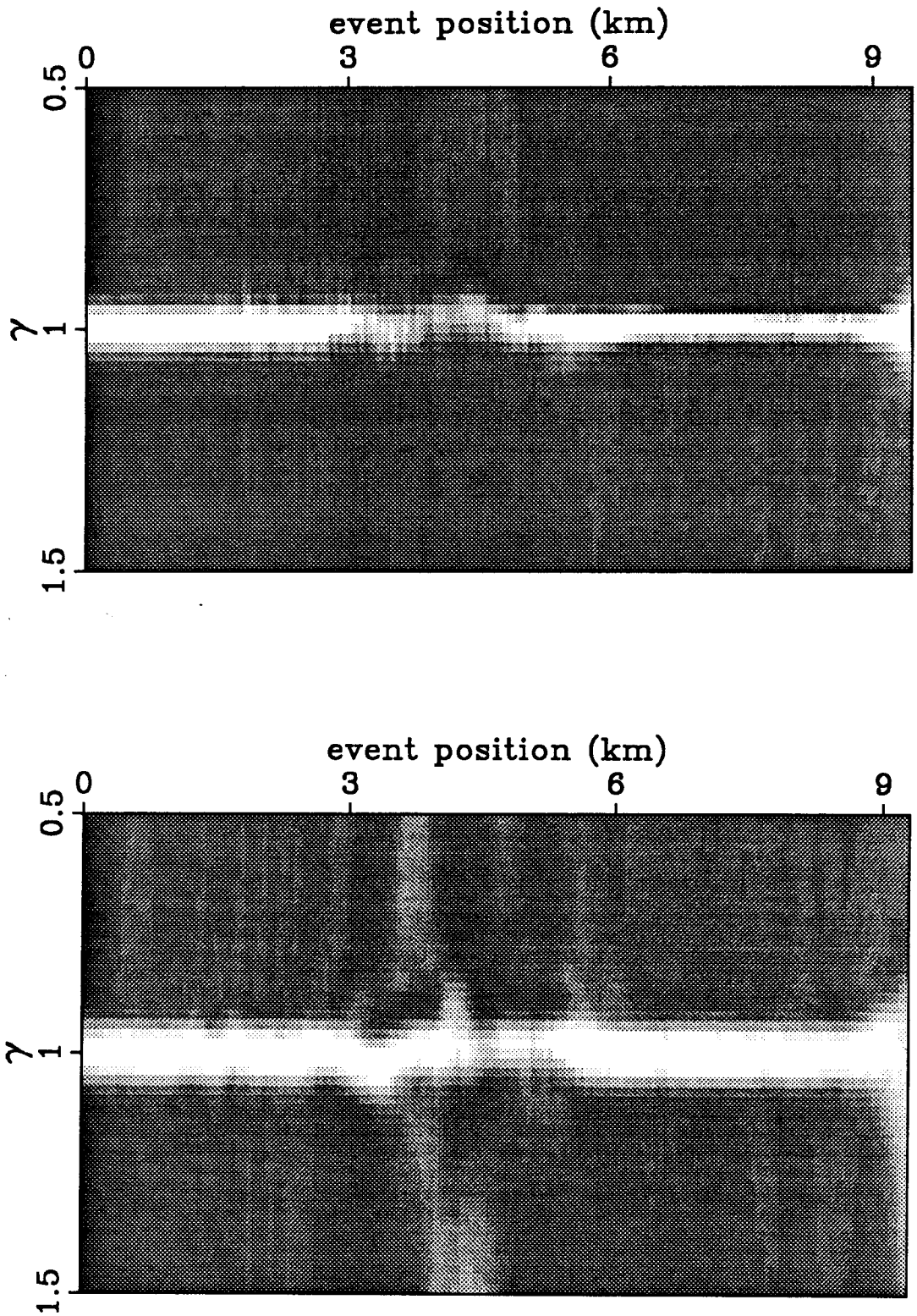

FIG. 4.9. Horizon residual-slowness analyses of the two reflectors after the final iteration of the velocity-analysis algorithm. The semblance peaks occur at $\gamma = 1$ indicating that all moveout information in the data is predicted by the final interval-slowness model.