SAND98-1463J revised

Phase encoding of shot records in prestack migration *

Louis A. Romero

Applied Mathematics, Sandia National Laboratories, P.O. Box 5800, M.S. 1110, Albuquerque, NM 87185-1110 E-mail: laromer@cs.sandia.gov

Dennis C. Ghiglia

Applied Mathematics, Sandia National Laboratories, P.O. Box 5800, M.S. 1110, Albuquerque, NM 87185-1110 (Presently VEXCEL Corporation, 4909 Nautilus Ct., Boulder, CO 80301)

Curtis C. Ober

Parallel Computational Sciences, Sandia National Laboratories, P.O. Box 5800, M.S. 1111, Albuquerque, NM 87185-1111 E-mail: ccober@cs.sandia.gov

Scott A. Morton

Cray Research, A Silicon Graphics Company, (Presently Amerada Hess Corporation, One Allen Center, 500 Dallas St., Houston, TX 77002) E-mail: morton@hess.com

DISCLAIMER

(August 16, 1999)

ABSTRACT

Frequency-domain shot-record migration can produce higher quality images than Kirchhoff migration but typically at a greater cost. The computational cost of shot-record migration is the product of the number of shots in the survey and the expense of each individual migration. Many attempts to reduce this cost have focused on the speed of the individual migrations, trying to achieve a better trade-off between accuracy and speed.

Another approach is to reduce the number of migrations. We investigate the simultaneous migration of shot records using frequency-domain shot-record migration algorithms. The difficulty with this approach is the production of so-called cross terms between unrelated shot and receiver wavefields, which generate unwanted artifacts or noise in the final image. To reduce these artifacts and obtain an image comparable in quality to the single-shot-per-migration result, we have introduced a process called phase encoding which shifts or disperses these cross terms. The process of phase encoding thus allows one to trade signal-to-noise ratio for the speed of migrating the entire survey.

Several encoding functions and two application strategies have been tested. The first strategy, combining multiple shots per migration and using each shot only once,

---

*Orally presented: Ober, Oldfield, Womble, Romero and Burch, 1997, "Practical aspects of prestack depth migration with finite differences", 67^{th} Ann. Internat. Mtg of the SEG, Expanded Abstracts, pp. 1758–61. Also: Morton and Ober, 1998, "Faster shot-record depth migrations using phase encoding", 68^{th} Ann. Internat. Mtg of the SEG, Expanded Abstracts, pp. 1131-1134. Patent applied for.

provides a reduction in computation directly related to the number of shots combined. The second strategy, performing multiple migrations of all the shots in the survey, provides a means to reduce the cross-term noise through stacking the resulting images. The additional noise in both strategies may be tolerated if it is no stronger than the inherent seismic noise in the migrated image, and if the final image is achieved with less cost.

INTRODUCTION

Most companies today use Kirchhoff methods to perform their production 3-D prestack migrations, largely because they are relatively inexpensive and flexible with respect to image size and data geometry. However they are usually limited in their ability to account for multiple travel paths and finite-frequency effects, phenomena which occur in regions of complex geology.

To improve image quality and avoid these difficulties, frequency-domain migration methods have been developed and are more commonly used than in the past. However these methods can be substantially more expensive than Kirchhoff migration and the number of shot-record migrations required for prestack migration of an entire survey can make frequency-domain methods prohibitively expensive.

We could reduce the computational requirements by summing several shots together and imaging with a single migration, thus reducing the total computation by the number of shots combined. Unfortunately this results in unwanted cross terms between unrelated shot and receiver wavefields during the imaging.

If the spatial separation between the shots is large, these cross terms would be small. However, in seismic migration, one rarely migrates a single shot over the whole domain of interest, choosing instead an aperture appropriate for the problem. So adding together spatially separated shots would require enlarging the imaging domain, decreasing the number of migrations while increasing their cost, leading to

no significant change in the overall computational work.

For example, if we have two impulse responses which are spatially separated and are migrated in a constant velocity medium, an image of two hemispheres would be expected as shown in Figure 1(a). The computational expense for this run is approximately twice that required for a single shot because the computational domain has been doubled to include both shots. Thus little to no savings can be gained by using widely spaced shots. Consequently in this paper, we will consider shots which are spatially close (*i.e.*, within the migration aperture).

If two spatially close shots are combined to reduce the computational domain as shown in Figure 1(b), the resulting image has the expected response, two hemispheres, but it also has cross terms between the source wavefield of one shot and the receiver wavefield of another, and visa versa. (For this simple example, the two cross terms exactly overlay one another.) To reduce the number of migrations and yet produce a good quality image, the artifacts caused by the cross terms need to be reduced or removed.

Phase-encoding methods used in radar systems (Nathanson, 1969; Cook, 1967) can be adapted to prestack frequency-domain migration. In this paper, phase encoding methods to disperse or shift the cross-term artifacts are described and examples using synthetic datasets are shown.

PHASE ENCODING

In the standard approach to frequency-domain shot-record migration, the source wavelet and the trace data must be Fourier transformed in time to give the source and receiver wavefields at the surface, $S_n(x, y, z = 0, \omega)$ and $R_n(x, y, z = 0, \omega)$, where ω is the Fourier transform variable and the subscript n indicates the shot index. These wavefields are then propagated into the earth's interior using the one-way wave equation to determine $S_n(\underline{x}, \omega)$ and $R_n(\underline{x}, \omega)$ where $\underline{x} = (x, y, z)$ is the position vector.

To produce an image, we use a generalized form of Claerbout's (1971) correlation-imaging condition

$$I_n(\underline{x}) = \sum_{k=1}^{N_\omega} W(\omega_k) S_n^*(\underline{x}, \omega_k) R_n(\underline{x}, \omega_k), \tag{1}$$

where ω_k are the frequencies that we are processing, N_ω is the number of frequencies, W is a weighting factor used to improve the imaging condition, and S_n^* is the complex conjugate of S_n. It is important to note that the final image is the real part of $I_n(\underline{x})$. We will assume W is identically one for the remainder of paper since this does not effect the general application of phase encoding. However the images in this paper were obtained using a weighting which preserves the wavelet of the trace data (Ober et al., 1997b).

Equation (1) represents a crosscorrelation of two signals in time, similar to a matched filter pulse compression scheme used in radar systems. In this investigation we consider phase encoding only for imaging conditions of the above form. Phase encoding may be greatly complicated or not applicable for other types of imaging conditions.

The final image of the survey is then obtained by summing the images from all the shots,

$$I(\underline{x}) = \sum_n I_n(\underline{x}). \tag{2}$$

The time to image the entire survey using the standard approach is clearly proportional to the total number of shots migrated. If the number of migrations were reduced by combining shots, the overall cost to produce the final image would be decreased by nearly the number of shots per migration. A small increase of the computational domain is required to include the combined shots within the same migration therefore causing a slight increase in the cost for each migration.

To combine shots we can note that the differential equation we solve to obtain $S_n(\underline{x}, \omega)$ and $R_n(\underline{x}, \omega)$ is linear, and therefore we can compute the summed wavefields,

$$S_s(\underline{x},\omega) = \sum_n a_n(\omega) S_n(\underline{x},\omega) \tag{3}$$

and

$$R_s(\underline{x},\omega) = \sum_n a_n(\omega) R_n(\underline{x},\omega), \tag{4}$$

in nearly the same amount of time it takes to compute any of the individual terms $S_n(\underline{x},\omega)$ or $R_n(\underline{x},\omega)$. (Note: $a_n(\omega)$ is a function to be specified shortly). In order to do this we merely compute the values at the surface

$$S_s(x,y,z=0,\omega) = \sum_n a_n(\omega) S_n(x,y,z=0,\omega), \tag{5}$$

and

$$R_s(x,y,z=0,\omega) = \sum_n a_n(\omega) R_n(x,y,z=0,\omega), \tag{6}$$

then the combined wavefields at all depths, $S_s(\underline{x},\omega)$ and $R_s(\underline{x},\omega)$, are obtained by solving the one-way wave equations initialized with the summed values at the surface.

However a complication arises when $S_s(\underline{x},\omega)$ and $R_s(\underline{x},\omega)$ are processed with the imaging condition. Cross terms between unrelated shots produce artifacts in the image. For example, when two shots are summed and the imaging condition is applied, the image is given by

$$\begin{aligned} I_s(\underline{x}) = {} & |a_1|^2 I_1(\underline{x}) + |a_2|^2 I_2(\underline{x}) \\ & + \sum_k [S_1^\star(\underline{x},\omega_k) R_2(\underline{x},\omega_k) a_1^\star a_2 \\ & \quad + S_2^\star(\underline{x},\omega_k) R_1(\underline{x},\omega_k) a_2^\star a_1] . \end{aligned} \tag{7}$$

If we choose $|a_n(\omega)| = 1$, the first two terms on the right-hand side yields the stack of the individual images; however, the other terms are undesirable cross terms. By using phase encoding, we can try to eliminate or greatly reduce these unwanted cross terms. If we introduce the phase function $\gamma_n(\omega)$ and let $a_n(\omega) = e^{i\gamma_n(\omega)}$, we find

$$I_s(\underline{x}) = I_1(\underline{x}) + I_2(\underline{x}) + \sum_k \Big[S_1^*(\underline{x},\omega_k) R_2(\underline{x},\omega_k) e^{i(\gamma_2(\omega_k)-\gamma_1(\omega_k))} + S_2^*(\underline{x},\omega_k) R_1(\underline{x},\omega_k) e^{i(\gamma_1(\omega_k)-\gamma_2(\omega_k))} \Big]. \quad (8)$$

Note that the phase functions γ_n only appear in the cross terms, not in the terms producing I_1 and I_2. We now try to choose the functions $\gamma_n(\omega)$ so that when we sum over frequency, the energy from the unwanted terms will be dispersed throughout the whole image or be shifted out of the region of interest. When we want to add more than two shots together a similar procedure holds, however the cross terms become more numerous and therefore more difficult to shift or disperse.

PROPERTIES OF PHASE ENCODING

In order to understand how phase encoding can shift or disperse the cross terms, it is necessary to appreciate when an imaging condition produces a large response. It can be shown that when one does prestack migration, the image produced from a single impulsive shot can be kinematically approximated as

$$I(\underline{x}) = \int_{-\infty}^{\infty} e^{-i\omega\psi(\underline{x})} F(\omega) d\omega. \quad (9)$$

Here $F(\omega)$ is the frequency spectrum of the migrated wavefield. We are primarily interested in the fact that a large response in $I(\underline{x})$ will occur at a reflector ($\psi(\underline{x}) = 0$), and $\psi(\underline{x})$ is very nearly confined to the interval

$$-\tau_{max} < \psi(\underline{x}) < \tau_{max}, \quad (10)$$

where τ_{max} is the largest recorded travel time on the domain of interest. By this we mean the largest time to go from the shot location on the surface, to a point $\underline{x}$ in the interior, and then back to any point on the surface is τ_{max}. All of these points being in our computational domain.

In practice we replace our infinite integral by a finite and discrete sum

$$I(\underline{x}) = \sum_{k=-N_\omega}^{N_\omega} e^{-ik\Delta\omega\psi(\underline{x})} F(k\Delta\omega)\Delta\omega. \tag{11}$$

In this case we will get a large response at any point such that

$$\psi(\underline{x}) = \frac{2\pi m}{\Delta\omega}, \tag{12}$$

where m is an integer. We only want to get large responses from points with $m = 0$, the other responses lead to aliased images. Note that if

$$\frac{2\pi}{\Delta\omega} > \tau_{max}, \tag{13}$$

then we are guaranteed not to have any aliased images in our domain of interest. This follows since in that case

$$\psi(\underline{x}) - \frac{2\pi m}{\Delta\omega} > 0 \quad for \quad m < 0, \tag{14}$$

and

$$\psi(\underline{x}) - \frac{2\pi m}{\Delta\omega} < 0 \quad for \quad m > 0. \tag{15}$$

These inequalities follow from our bounds on $\psi(\underline{x})$ in the domain of interest.

When we produce an image using phase encoding, each cross term will have a form similar to the images. In particular, they can be kinematically approximated as

$$I_c(\underline{x}) = \int_{-\infty}^{\infty} e^{-i\omega\psi_c(\underline{x})} e^{i\gamma(\omega)} F(\omega) d\omega. \tag{16}$$

Here $\psi_c(\underline{x})$ is the phase produced by the cross term, and $\gamma(\omega)$ is a phase function that will hopefully shift or disperse the energy so that $I_c(\underline{x})$ is not large on the domain of interest. If it were not for the phase $\gamma(\omega)$, we would get a large response from any points where $\psi_c(\underline{x}) = 0$, and this would appear as a false reflector. When we introduce the phase function $\gamma(\omega)$, the image $I_c(\underline{x})$ does not necessarily take on its largest value at points where $\psi_c(\underline{x}) = 0$. Ideally we do not want $I_c(\underline{x})$ to take on a large value for any value of $\psi_c(\underline{x})$. This suggests that we analyze our phase function $\gamma(\omega)$ by examining the function

$$C(\alpha) = \int_{-\infty}^{\infty} e^{-i\omega\alpha} e^{i\gamma(\omega)} F(\omega) d\omega. \tag{17}$$

If the function $C(\alpha)$ takes on a large value at some point $\alpha = \alpha^*$, we could get what looks like a reflecting surface at some point $\underline{x}$ where $\psi_c(\underline{x}) = \alpha^*$. For this reason, we would like to choose a phase function $\gamma(\omega)$ that minimizes the maximum value that the function $C(\alpha)$ can assume for any value of α (*i.e.*, disperses the cross-term energy). If we actually computed our imaging condition using a continuous integral, we could choose $\gamma(\omega)$ so that the maximum value of $C(\alpha)$, and hence our cross terms, were arbitrarily small. In practice, our sampling rate imposes limits to how small we can make $C(\alpha)$. When we sample at a finite rate, we replace the integral by the discrete sum

$$C_{\Delta\omega}(\alpha) = \Delta\omega \sum_{k=-\infty}^{\infty} e^{-ik\Delta\omega\alpha} e^{i\gamma(k\Delta\omega)} F(k\Delta\omega). \tag{18}$$

To see the limits imposed by the finite sampling we use the discrete form of Parseval's equality (Churchill, 1972). This states that if we expand a periodic function $f(x)$ with period $2a$ in a Fourier series,

$$f(x) = \sum_{k=-\infty}^{\infty} a_k e^{ikx\pi/a} \tag{19}$$

then,

$$\frac{1}{2a} \int_{-a}^{a} \mid f(x) \mid^2 dx = \sum_{k=-\infty}^{\infty} \mid a_k \mid^2 . \tag{20}$$

In our particular case, the function $C_{\Delta\omega}(\alpha)$ has a period of $2\pi/\Delta\omega$, and Parseval's equality implies that

$$\frac{\Delta\omega}{2\pi} \int_{-\pi/\Delta\omega}^{\pi/\Delta\omega} \mid C_{\Delta\omega}(\alpha) \mid^2 d\alpha = (\Delta\omega)^2 \sum_{k=-\infty}^{\infty} \mid F(k\Delta\omega) \mid^2 . \tag{21}$$

This shows that it is not possible to reduce the function $\mid C_{\Delta\omega}(\alpha) \mid$ so that it is everywhere less than a constant, $H_{\Delta\omega}$, where

$$(H_{\Delta\omega})^2 = (\Delta\omega)^2 \sum_{k=-\infty}^{\infty} \mid F(k\Delta\omega) \mid^2 . \tag{22}$$

This can be approximated as

$$(H_{\Delta\omega})^2 \approx \Delta\omega \int_{-\infty}^{\infty} \mid F(\omega) \mid^2 d\omega. \tag{23}$$

No matter what phase function $\gamma(\omega)$ we choose, the function $\mid C_{\Delta\omega}(\alpha) \mid$ must be greater than or equal to $H_{\Delta\omega}$ at some point. We will now apply this formula to the case where $F(\omega) = 1$ for $\mid \omega \mid < \omega_0$, and zero elsewhere. In this case, we find

$$(H_{\Delta\omega})^2 = 2\omega_0 \Delta\omega. \tag{24}$$

If we use an optimal form for the function $\gamma(\omega)$, we can reduce the cross terms so they are all of magnitude $H_{\Delta\omega}$. We need to compare this to the results where we do not use any phase encoding then $\gamma(\omega) = 0$, and the maximum value of $C(\alpha)$ occurs when $\alpha = 0$ and is given by $H_0 = 2\omega_0$ (see Eq. (17)). It follows that the ratio of the optimally dispersed cross term, $H_{\Delta\omega}$, to the maximum value of the non-dispersed cross term, H_0, is

$$\frac{H_{\Delta\omega}}{H_0} = \sqrt{\frac{\Delta\omega}{2\omega_0}}. \tag{25}$$

This result was derived for the case where $F(\omega)$ is a square spectrum and the phase $\gamma(\omega)$ was chosen to optimally disperse the cross terms. However, the basic form of the result is quite general and crucial to understanding phase encoding. In general for a seismic source, we have $\omega_0 \delta t \approx 2\pi$, where δt is the width of the seismic pulse that is sent out. Similarly we have $\Delta\omega\tau_{max} \approx 2\pi$ where τ_{max} is the maximum recorded arrival time. Plugging into Eq. (25), it follows that if the square root of the pulse width divided by the maximum arrival time, $\sqrt{\delta t/\tau_{max}}$, is small, then we should expect to be able to do a good job of dispersing the cross terms. Assuming δt is several time samples, dt, the ratio of the dispersed cross term to the non-dispersed cross term scales as $\sqrt{\delta t/\tau_{max}} \sim 1/\sqrt{N_\omega}$. Since cross terms are comparable in magnitude to correct events, this result indicates that the effectiveness of phase encoding in dispersing the cross terms is limited by the inverse of the square root of the number of frequencies migrated.

MIGRATION METHOD

Phase encoding, while dependent on the type of imaging condition, is independent of the particular frequency-domain shot-record migration method used. However, for completeness we briefly summarize the migration method used in this study (Ober et al., 1997a). It is based on industry-standard approaches (Claerbout, 1985; Yilmaz, 1987; Li, 1991) for shot-record migration utilizing frequency-domain, implicit solutions to the paraxial wave equation. The method of fractional steps (Fletcher, 1988) is used along with Crank-Nicholson differencing and compact finite differences to produce a fourth-order finite-difference scheme. Absorbing boundary conditions similar to those of Clayton and Engquist (1980) are used. Phase-correction filters (Graves and Clayton, 1990; Li, 1991) have been implemented to correct for errors introduced by the square-root-operator and operator-splitting approximations.

PHASE-ENCODING FUNCTIONS

In this section, several phase-encoding functions are discussed and illustrated for two shots in a single migration. Their application to more than two shots per migration is illustrated in subsequent sections along with strategies for migrating entire surveys.

Linear Phase Encoding

In the following, we show that provided our sampling rate in the frequency domain is high enough (*i.e.*, $\Delta\omega$ is small enough), the unwanted cross terms can be completely eliminated by choosing the phase functions, $\gamma_i(\omega)$, to be linear. Unfortunately, the required sampling rate for two shots is twice as high as that needed to prevent aliasing, which is the one typically used in seismic imaging (for three shots the required sampling rate is three times, etc.). In other words, we usually record only

enough trace data to image the geological structures of interest. However through zero padding the traces, the cross terms can be shifted into some of the newly-created frequency domain and therefore out of our imaging domain.

Suppose we choose the phase functions $\gamma_i(\omega)$ so that

$$\gamma_2(\omega) - \gamma_1(\omega) = t_0\omega. \tag{26}$$

In this case our unwanted cross terms will be of the form

$$\int_{-\infty}^{\infty} e^{-i\omega(\psi_c(\underline{x})\pm t_0)} F(\omega) d\omega, \tag{27}$$

Assuming the function $F(\omega)$ is broadband, we will only get significant intensities from the cross terms at points where $\psi_c(\underline{x}) \pm t_0 = 0$.

We have already noted that the absolute value of the phase function $\psi(\underline{x})$ can be bounded by τ_{max}, where τ_{max} is the largest travel time in the computational domain. Assuming that our encoded shots are relatively close together, this bound also applies to the phase of the cross terms, $\psi_c(\underline{x})$. If we choose t_0 so that it is bigger than τ_{max} there will be no contribution from either of the terms $\psi_c(\underline{x}) \pm t_0$. We see that by choosing t_0 large enough we can completely eliminate one cross term, and shift the other cross term out of the region that we are imaging.

We can accomplish this only if we are sampling our data finely enough in frequency. A simple extension of our results concerning aliasing, Eq. (13), shows that in order to prevent an aliased image with negative values of m we must have

$$\frac{2\pi}{\Delta\omega} > \tau_{max} + t_0 > 2\tau_{max}. \tag{28}$$

A similar result holds for the positive values of m. We see that in order to prevent aliasing for the unwanted cross terms, we must sample at twice the rate as if we had not used a linear phase encoding. This results in twice as many frequencies used for migration and therefore no overall reduction in computational cost.

Returning to our simple example of Figure 1(b), the effect of various linear phase encodings are shown in Figure 2. The trace length is 4.096 seconds with 4 ms sampling, and the medium has a constant velocity of 3 km/s. The total travel times from source to reflectors to receiver is 1.332 seconds.

If we slightly increase t_0 to 0.064 seconds, the cross terms shift position in the image; one up and the other down as shown in Figure 2(a). Increasing t_0 to 0.512 seconds as shown in Figure 2(b), one of the cross terms is shifted completely out of the imaged domain, and the other has continued its downward shift. In Figure 2(c), the time shift, t_0, has been set to 1.792 seconds and only slight visual evidence of the cross terms can be seen. In Figure 2(d), the time shift is half the trace length, 2.048 seconds, and artifacts from the cross terms can be seen near the surface due to temporal wrap-around. If we further increase t_0, the cross terms return into the image in the reverse order in which they were shifted out (*i.e.*, Figures 2(d), 2(c), 2(b), and 2(a)). When $t_0 = 4.096$ seconds, the resulting image is identical to Figure 1(b).

To illustrate the various encoding functions on a more complex model, two shots from the Marmousi model have been selected and migrated. Figure 3 shows the standard imaging and stacking as described by Eqs. (1) and (2) (*i.e.*, a separate migration for each shot and the individual images summed). The two shots selected are located at 4250m and 4500m; the migration used 256 frequencies between 0-62.5 Hz; and the grid spacing was 12.5m by 4m.

The two shots of the Marmousi model were also combined using linear phase encoding with $t_0 = 1.792$ and the resulting image is shown in Figure 4(a) along with the difference between Figure 3 and Figure 4(a). All images in Figure 4 have the same scale and no amplitude scaling, such as automatic gain control, has been applied. The two images are very similar except near the surface where the cross-term artifacts appear as "noise" in the encoded image.

To extend linear phase encoding to more than two shots per migration, we can use $\gamma_j(\omega) = (T * j * \omega)/(K-1)$ where j is the shot number $(0 \ldots K-1)$, T is the

Fourier transform length in seconds, and K is the number of shots to be encoded.

Random Phase Encoding

While linear phase encoding moves the cross terms (potentially out of the image), we can use random phases to disperse these unwanted cross terms. We can choose the phase function, $\gamma(\omega) = \gamma_1(\omega) - \gamma_2(\omega)$, from Eq. (17) to be a sequence of random numbers between 0 and 2π. When we sum over ω to produce an image, the phases will not agree, and this should not produce a large response for any value of α.

To estimate how well the maximum cross terms can be reduced through random encoding, we again assume a flat spectrum (*i.e.,* that the function $F(\omega)$ is equal to unity for $|\omega| < \omega_0$, and zero elsewhere). Without phase encoding, we will get the largest response for the cross terms when $\alpha = 0$. In this case the maximum of a cross term $C(\alpha)$ will be equal to N_ω. However with $\gamma(\omega)$ being a uniform random variable, the theory of 2-D random walk shows us that the root-mean-square value of $C(\alpha)$ is $\sqrt{N_\omega}$. So on average, random phase encoding will reduce the magnitude of the cross terms by a factor of $1/\sqrt{N_\omega}$.

As an example of two shots per migration, the shots used in Figure 3 were randomly encoded and the resulting image is shown in Figure 4(b). Random phase encoding can be trivially extended to more than two shots per migration. Since the difference of any two uniform random variables on a periodic domain is also a uniform random variable, we can choose each $\gamma_j(\omega)$ to be a uniform random variable. This will reduce each cross term by the factor of $1/\sqrt{N_\omega}$, even though the number of cross terms increases as the square of the number of shots.

Chirp Phase Encoding

In order to disperse the cross terms, it is not necessary to choose the phases randomly. In this section we choose the function $\gamma(\omega) = \gamma_2(\omega) - \gamma_1(\omega)$ to be the

linear frequency-modulation chirp

$$\gamma(\omega) = \beta\omega^2. \tag{29}$$

In this case our unwanted cross terms have the form

$$C(\alpha) = \int_{-\infty}^{\infty} e^{-i\omega\alpha} e^{i\beta\omega^2} F(\omega) d\omega. \tag{30}$$

If β is large we can evaluate this integral using the method of stationary phase. This gives us

$$C(\alpha) \approx \sqrt{\frac{\pi}{\beta}} F(\alpha/(2\beta)) e^{-i\alpha^2/(4\beta)+i\pi/4}. \tag{31}$$

In principle, we can make the amplitude of the cross term arbitrarily small as $\beta \to \infty$. However this is only the case if we have continuous sampling in the frequency domain. If we replace the integral by a discrete sum, we see that the chirp will be aliased once $\beta > \pi/(\omega_0 \Delta\omega)$ where ω_0 is the bandwidth of our signal. This gives us an estimate of how large we can expect to choose β and still have our stationary-phase arguments hold. It follows that we can expect to reduce the signal to

$$|C(\alpha)| \approx \sqrt{\omega_0 \Delta\omega}\, |F(\alpha/(2\beta))|. \tag{32}$$

To get a rough estimate of how much we have actually reduced the cross terms, we again assume a flat spectrum for $F(\omega)$. Then without the chirp, the signal would be $2\omega_0$, and the ratio of the reduced signal to the original signal is

$$\sqrt{\frac{2\Delta\omega}{\omega_0}}. \tag{33}$$

This gives us a very similar result as the random encoding. In practice it is found that if we choose β larger than the maximum value allowed to prevent aliasing, our results do not greatly degenerate, but they do not improve.

The two shots of the Marmousi model were migrated using the chirp encoding, and the image is shown in Figure 4(c). The artifacts are spread over the entire image,

similar to the case of random encoding, however they are stronger and could easily be mistaken for a true seismic event.

In order to add more than two shots together, we can use the sequence of phases $\gamma_1(\omega) = 0$, $\gamma_2(\omega) = \beta\omega^2$, $\gamma_3(\omega) = 2\beta\omega^2$, ... $\gamma_K(\omega) = (K-1)\beta\omega^2$. The phases of the unwanted cross terms will be of the form $k\beta\omega^2$, where $|k| < K$. In order to prevent aliasing we must choose $(K-1)\beta$ small enough (*i.e.*, less than $\pi/(\omega_0\,\Delta\omega)$).

Modified-Chirp Phase Encoding

Ideally we would like to make the cross term $C(\alpha)$ in Eq. (30) equal to zero for all values of α. Parseval's equality, Eq. (21), shows us that this is not possible. However we shall show how to choose the phase function $\gamma(\omega)$ in order to make the magnitude of $C(\alpha)$ uniform over the interval $|\alpha| < \tau_{max}$. This analysis is identical to that used in Romero and Dickey (1996) to flatten out a laser beam profile.

First we assume that $\gamma(\omega) = \beta r(\omega)$ where β is a large parameter. Applying the method of stationary phase to the integral of Eq. (17), we find that

$$|C(\alpha)| \approx \sqrt{2\pi}\frac{|F(\omega(\xi))|}{\sqrt{\beta r''(\omega(\xi))}}, \tag{34}$$

where $\xi = \alpha/\beta$, and $\omega(\xi)$ is defined implicitly through the equation

$$\frac{d}{d\omega}\left(\xi\omega + r(\omega)\right) = 0. \tag{35}$$

This last condition arises as a stationarity condition for the rapidly varying phase. On any interval where we would like $|C(\alpha)|$ to be constant, Eq. (34) shows that we must have $d^2r/d\omega^2 = |F(\omega)|^2$ and hence

$$\frac{dr}{d\omega} = \int_0^\omega |F(u)|^2\,du. \tag{36}$$

Assuming that $|F(\omega)| = |F(-\omega)|$, we see that for this choice of $r(\omega)$ the phase cannot be stationary if $\alpha/\beta > \int_0^\infty |F(u)|^2\,du$, and for $|\alpha/\beta| < \int_0^\infty |F(u)|^2\,du$, $|C(\alpha)|$ will be

a constant. As with the chirp, we could get $C(\alpha)$ to be arbitrarily small by choosing very large values of β, but we are limited by the fact that we are sampling at a finite rate. In order to prevent aliasing, we must have $\beta\Delta\omega\frac{dr}{d\omega} < 2\pi$, and hence

$$\beta < \frac{2\pi}{\Delta\omega \int_0^\infty |F(u)|^2 \, du}. \tag{37}$$

Note that this requires knowing the function $F(\omega)$. However, if we only approximately know the function, we can still get an improvement over using the chirp. We see that Eq. (37) puts a limit on how big we can make β in terms of the parameters $\Delta\omega$ and the function $F(\omega)$. This expression for β is proportional to $\omega_0/\Delta\omega$.

For the special case of the first derivative of a Gaussian, $F(\omega) = \omega\lambda e^{-1/2\lambda\omega^2}$, we have $r(\omega) = q(\sqrt{\lambda}\omega)$, where

$$q(\xi) = \frac{\sqrt{\pi}}{4}\xi \, \mathrm{erf}(\xi) + \frac{e^{-\xi^2}}{2}. \tag{38}$$

In order to disperse the signal as much as possible, we must choose $\beta = 8\sqrt{\pi}/(\Delta\omega\sqrt{\lambda})$.

Using the above formulation, the two shots of the Marmousi model were migrated using the modified-chirp encoding. The resulting image is shown in Figure 4(d). Like the random and chirp encoding, the artifacts are spread across the image and have relatively large amplitudes near the surface and are small elsewhere in the image. These artifacts are periodic throughout the domain. This is due to the fact that while $|C(\alpha)|$ is constant over the domain, the image is just the real part of $C(\alpha)$.

FULL MARMOUSI RESULTS

We now apply these phase-encoding functions to all 240 shots of the Marmousi dataset using two different strategies. In the first strategy we examine two or more shots in each migration, using each shot exactly once. In the second strategy, we use all the shots in the survey in each migration, repeating this process with different phase encodings and stacking the resulting images. For comparison, the entire

Marmousi dataset was migrated using standard imaging and stacking as described by Eqs. (1) and (2), or in other words with one shot per migration. This image is shown in Figure 5.

Multiple Shots per Migration

We first compare the "correct" image in Figure 5 to phase-encoded images where two or more adjacent shots are included in each migration and each shot is used in exactly one migration. For the following phase-encoded images, the same number of frequencies was used for all phase-encoding methods, including linear phase encoding, and no amplitude scaling has been applied.

In Figure 6, the results for encoding two shots per migration are shown (*i.e.*, the 240 shots were migrated in sets of two adjacent shots for a total of 120 migrations). On the left is shown the resulting encoded images, and on the right is the difference (amplified by a factor of 10) between the encoded images and the image generated through standard imaging and stacking shown in Figure 5. Figure 6(a) shows the migrated image for linear phase encoding. The artifacts near the surface appear to be random in nature. However, the deeper artifacts appear to be coherent structures in the image, which could be mistaken for a true event.

The random-encoded image shown in Figure 6(b) is very good. The artifacts for this image are "randomly" dispersed throughout the image. This apparent noise in the image has short wavelengths near the surface and longer wavelengths at deeper locations. This is due to stretching caused by the velocity of the medium.

The chirp-encoded image is shown in Figure 6(c). The cross terms in this image have larger amplitudes than in comparison with the previous two methods shown in Figure 6. Additionally the cross terms have coherent structures which appear to be periodic. This behavior is highly undesirable.

The modified-chirp-encoded image is shown in Figure 6(d), and is very similar to

the chirp-encoded image. As with the chirp-encoded image, the artifacts are periodic and could be mistaken for true events.

Overall these encodings of two shots per migration perform very well and introduce only slight artifacts to the final image (remember the differences shown in Figure 6 have been amplified by 10). These runs took half as long as those for the standard imaging and stacking, thus substantial savings in runtime is possible.

Further savings in runtime can be obtained if more than two shots are phase encoded. To measure the amount of noise generated by these encodings, we use a relative L_2 norm of the difference between the images,

$$\frac{\left\| I_{w/\ encoding} - I_{w/o\ encoding} \right\|}{\left\| I_{w/o\ encoding} \right\|}. \tag{39}$$

Although this L_2 norm is a quantitative measure of the effectiveness of phase encoding, it can not point out qualitative features such as the packing of artifacts near the surface or the periodic nature of artifacts. In Figure 7, the relative L_2-norm differences are shown for the phase-encoding functions tested. The trivial encoding, where the combined shots are simply separated by large distances, is the best performer for less than 20 shots per migration. However this approach will not save us any computation because of the increased computational domain.

The linear phase encoding does very well for 2 and 4 shots per migration but is relatively worse at more shots per migration. This is primarily due to the trace length which limits the usefulness of linear phase encoding. The random encoding performs relatively poorly at two shots per migration, but improves relative to the other encoding methods with more shots per migration. This is encouraging because the cross-term energy for this method is randomly dispersed throughout the image and therefore will not likely introduce events which could be misinterpreted.

The chirp encoding does very poorly at two shots per migration, but performs relatively better with more shots per migration. The modified-chirp encoding does slightly better than the chirp encoding for two shots per migration, however it does

much worse with more shots per migration. After further investigation, we found that the periodic character of the chirp and modified-chirp encoding made it very difficult to interpret the final image.

The random encoding appears to be a good choice for higher number of shots per migration because of its relatively uniform distribution and lack of coherent structures, as seen in Figure 8. All images are shown at the same scale and no amplitude scaling has been applied to any of these images. As the number of shots per migration is increased, the number of migrations performed (and hence, the computational effort) decreases while the cross-term energy grows in strength. However even with all 240 shots of the Marmousi model randomly encoded in a single migration, some major features can still be seen. Yet the image in Figure 8(d) took about 1/240 of the time to generate the image in Figure 5.

Clearly we have a trade-off between adequate signal-to-noise ratio and computational speedup, and we can choose the desired trade-off by simply specifying the number of shots per migration.

Multiple Migrations of All Shots

One difficulty of the above approach is that the number of shots per migration (and hence, the total number of migrations) needs to be specified before beginning the migration. But since we do not know the relationship between the signal-to-noise ratio and the number of shots per migration, the adequacy of the final signal-to-noise ratio cannot be determined until the migrations are completed. Therefore this makes it difficult know a priori the number of shots per migration which produces an acceptable image with the fewest migrations.

Morton and Ober (1998) presented an alternative strategy for phase encoding which allows the number of migrations and the desired signal-to-noise ratio to be determined dynamically. This approach includes all shots from a survey in each mi-

gration, using different instantiations of the phase-encoding function in each migration, and stacking the resulting images. Consequently the correct image for the entire survey is present in every image, though its events are (at least) partially masked by the cross-term energy (as shown in Figure 8(d)). The addition of each subsequent migration simply improves the signal-to-noise ratio. This strategy has the advantage that the number of migrations (and hence the computational effort) is determined when the desired signal-to-noise ratio is reached.

One complication with this strategy is the data management of encoding all the shots for each migration. Though the additional computational effort of phase encoding the entire survey is small compared to each frequency-domain migration, the data must be carefully managed to ensure the additional processing time is kept small.

Though this strategy can use any phase-encoding function, the solid curve in Figure 9 demonstrates it using random phase encoding. All 240 shots were encoded, migrated and stacked, up to 240 times. The continuous reduction in the noise-to-signal ratio is apparent as more images are stacked, with the reduction following the expected inverse square-root relationship.

The squares on the dashed curve in Figure 9 show the noise-to-signal ratio for experiments using the strategy discussed in the previous section. Each square is the result of a complete migration of all 240 shots using multiple shots per migration with each shot being used only once.

To compare images for fixed computational effort, examine the encoding of 16 adjacent shots per migration (requiring 15 migrations to include the entire dataset) as shown in Figure 8(b) and 15 migrations of all 240 shots randomly encoded and then stacked as shown in Figure 10. We find that the noise in the former has an L_2 norm of 0.586 while the latter has an L_2 error norm of 0.427.

Not only does the strategy of migrating all shots in each migration have a comparable or smaller noise-to-signal ratio (for all but the trivial case of performing all 240 migrations), but it has the advantage of making frequency-domain shot-record

migration a continuous process.

SUMMARY AND CONCLUSIONS

Phase encoding is a technique applicable to any frequency-domain shot-record migration method, where a number of shot records may be migrated together, reducing the computational cost of migrating an entire seismic survey. Normally when several shot records are combined in a migration, cross terms are generated between unrelated source and receiver wavefields which cause unwanted artifacts in the image. By introducing phase functions designed for a correlation imaging condition, we can disperse (by random, chirp and modified-chirp phase encoding) or shift (by linear phase encoding) these artifacts, thereby reducing their effect on the final image.

As we increase the number of shots per migration, the computational cost decreases while the number of cross terms (and hence, the noise in the image) increases. Thus the application of phase encoding is a trade-off between computational costs and cross-term noise.

While in this paper we only considered the application of phase encoding to shot-record migration, phase encoding can also be applied to the migration of some other prestack data subsets, specifically receiver-record and plane-wave migrations. In fact, phase encoding will work with any prestack frequency-domain migration method which downward continues all of the prestack data subsets with the same operator. The current implementation of phase encoding requires a correlation imaging condition; the extension to other (*e.g.*, deconvolution) imaging conditions is not straightforward and will require further advancement. An additional limitation of phase encoding is that by combining the prestack data volumes, the redundancy typically used to perform velocity analysis is lost. We currently envision using phase encoding only to construct full-volume stacked images.

In this paper we presented two strategies for applying phase-encoding functions.

The first strategy involves migrating several shots per migration, but only using each shot once. For the Marmousi model, a synthetic dataset without seismic noise, phase encoding with two and four shots per migration produces images which have very slight artifacts, and reduces the computational costs by two and four times respectively.

In the second strategy all shots are phase encoded in each migration, producing an image containing the "correct" image partially masked by the cross terms. By repeating this process with different phase encodings and stacking the resulting images, the stack beats down the cross terms improving the signal-to-noise ratio with each additional migration. This strategy enables prestack shot-record migration of an entire survey to be a continuous process, where the number of migrations and the computational cost are determined when the desired signal-to-noise ratio is achieved.

The presented results indicate that phase encoding may enable significant reductions in the computational cost of frequency-domain shot-record migration, perhaps making it cost competitive with Kirchhoff algorithms while preserving its advantages. However further work needs to be done to make phase encoding production ready. A more optimal strategy may involve a mixture of the two strategies examined here. For real data sets, we expect (and have seen on limited runs) the inherent seismic noise to be significant, allowing more phase-encoding noise and hence a greater reduction in the computational cost. Further application to real data sets will help determine the amount of phase-encoding noise that can be tolerated.

ACKNOWLEDGMENTS

This work was supported by the United States Department of Energy under Contract DE-AC04-94AL85000, and by DOE's Office of Mathematical, Information and Computational Sciences. The authors would like to thank David Womble and Robert Gjertsen for their contributions and support to the seismic imaging efforts at Sandia

National Laboratories. Additionally we thank the partners of the ACTI-16 project, 3D Seismic Imaging of Complex Geologies, for their contributions in ideas, data sets, and discussions. Most notably we thank Charles C. Burch of Conoco Inc.

REFERENCES

Churchill, R., 1972, Operational Mathematics: McGraw Hill, New York.

Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**, 467–481.

Claerbout, J. F., 1985, Fundamentals of geophysical data processing with applications to petroleum prospecting: Blackwell Scientific Publications.

Clayton, R., and Engquist, B., May 1980, Absorbing Boundary Conditions for Wave-Equation Migration: Geophysics, **45**, no. 5, 895–904.

Cook, C., 1967, Radar signals, an introduction to theory and application: Academic Press, Orlando.

Fletcher, C., 1988, Computational Techniques for Fluid Dynamics Vol. I: Springer-Verlag, Berlin.

Graves, R., and Clayton, R., March 1990, Modeling Acoustic Waves with Paraxial Extrapolators: Geophysics, **55**, no. 3, 306–319.

Li, Z., October 1991, Compensating Finite-Difference Errors in 3-D Migration and Modeling: Geophysics, **56**, no. 10, 1650–1660.

Morton, S. A., and Ober, C. C., 1998, Faster shot-record depth migrations using phase encoding: 67th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts., 1131–1134.

Nathanson, F., 1969, Radar design principles: McGraw-Hill, New York.

Ober, C. C., Oldfield, R. A., Womble, D. E., and Mosher, C. C., 1997a, Seismic imaging in massively parallel computers: 67th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1758–61.

—— 1997b, Practical aspects of prestack depth migration with finite differences: 67th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1418–21.

Romero, L., and Dickey, F., April 1996, Lossless laser beam shaping: Optical Society of America, **13**, no. 4, 751–760.

Yilmaz, O., 1987, Seismic Data Processing, Investigations in Geophysics No. 2: Society of Exploration Geophysicists, P.O. Box 702740, Tulsa, OK 74170-2740.

FIGURES

FIG. 1. The response of two impulses which are spatially separated (a), and are spatially close (b).

FIG. 2. Linear phase encoding of image FIG. 1(b) with t_0 = (a) 0.064 (b) 0.512 (c) 1.792 and (d) 2.048 seconds.

FIG. 3. Two shots of the Marmousi Model using standard imaging and stacking (*i.e.*, one shot per migration).

FIG. 4. Two shots of the Marmousi Model using (a) linear $t_0 = 1.792$, (b) random, (c) chirp, (d) modified-chirp phase encoding. On the left is the encoded image, and on the right is the difference between the standard image in FIG. 3 and the encoded image.

FIG. 5. The Marmousi model migrated with 240 shots using standard imaging and stacking (*i.e.*, one shot per migration).

FIG. 6. The Marmousi Model combining 2 shots per migration using (a) linear, (b) random, (c) chirp, and (d) modified-chirp phase encoding. On the left is the encoded image, and on the right is the difference (amplified by a factor of 10) between the image in FIG. 5 and the encoded image.

FIG. 7. L_2-norm difference between the standard image in FIG. 5 and the various encoding functions for the Marmousi dataset.

FIG. 8. The Marmousi Model using random encoding and combining (a) 4, (b) 16, (c) 60, and (d) 240 shots per migration. On the left is the encoded image, and on

the right is the difference between the image in FIG. 5 and the encoded image.

FIG. 9. The solid curve is the L_2 norm of the difference between the standard image in FIG. 5 and the stack of multiple migrations of all randomly encoded shot records. Each data point on the dashed curve is the result of a complete migration of all 240 shots using multiple shots per migration with each shot being used only once.

FIG. 10. The Marmousi Model using random encoding 15 times on all 240 shots and stacking the resulting images. On the left is the stacked image, and on the right is the difference between the image in FIG. 5 and the stacked image. Compare this image with FIG. 8(b).

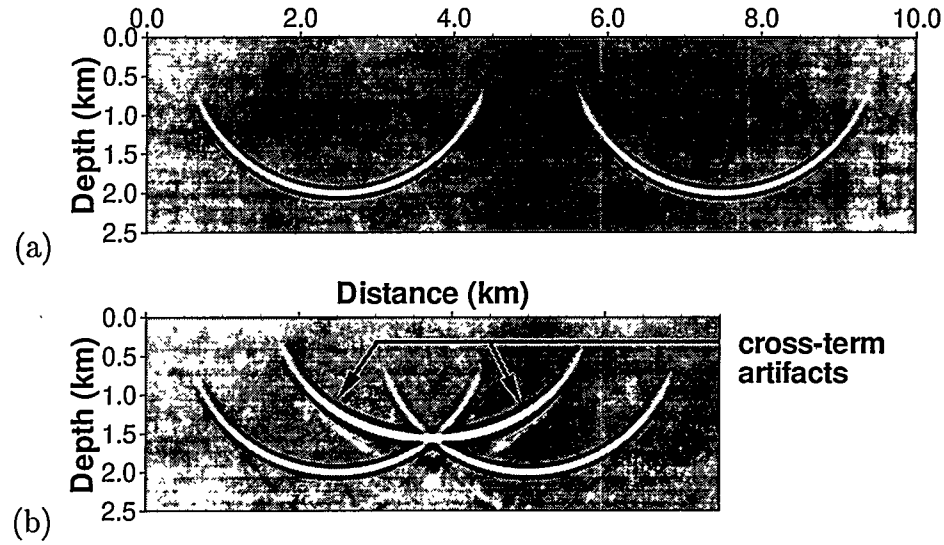

FIG. 1. The response of two impulses which are spatially separated (a), and are spatially close (b).

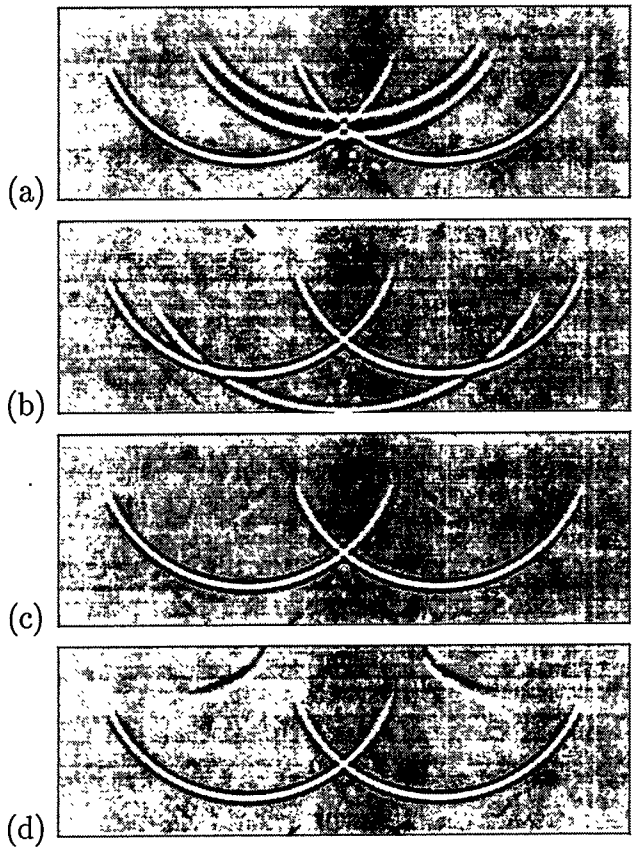

FIG. 2. Linear phase encoding of image FIG. 1(b) with t_0 = (a) 0.064 (b) 0.512 (c) 1.792 and (d) 2.048 seconds.

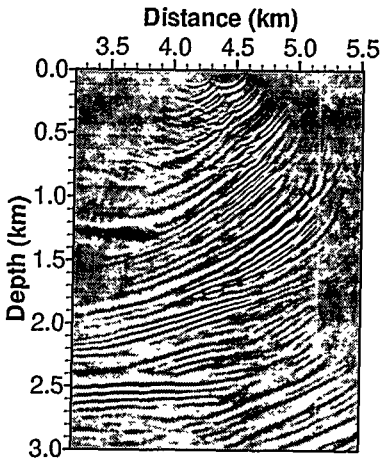

FIG. 3. Two shots of the Marmousi Model using standard imaging and stacking (*i.e.*, one shot per migration).

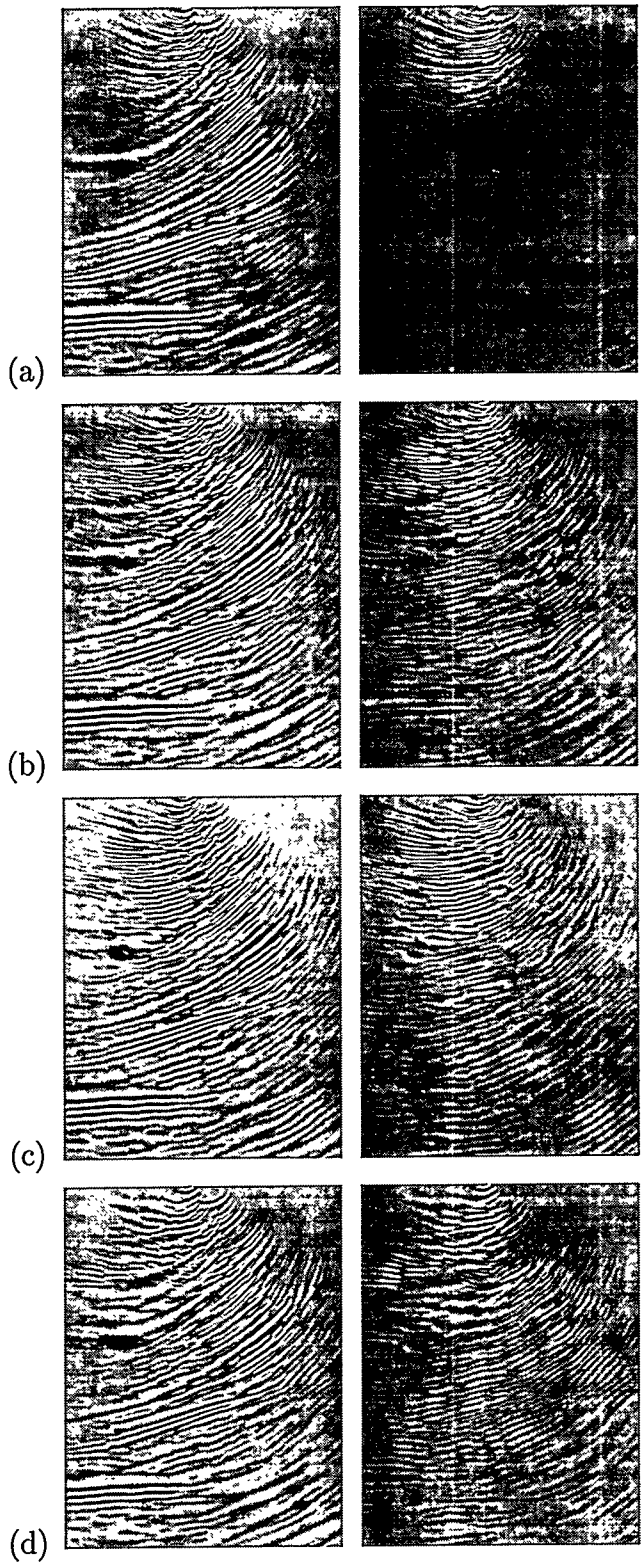

FIG. 4. Two shots of the Marmousi Model using (a) linear $t_0 = 1.792$, (b) random, (c) chirp, (d) modified-chirp phase encoding. On the left is the encoded image, and on the right is the difference between the standard image in FIG. 3 and the encoded image.

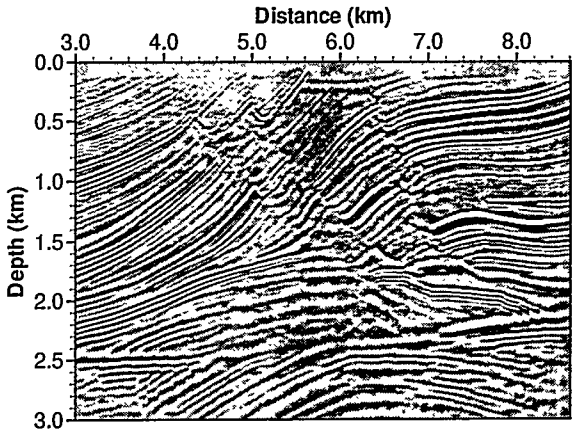

FIG. 5. The Marmousi model migrated with 240 shots using standard imaging and stacking (*i.e.*, one shot per migration).

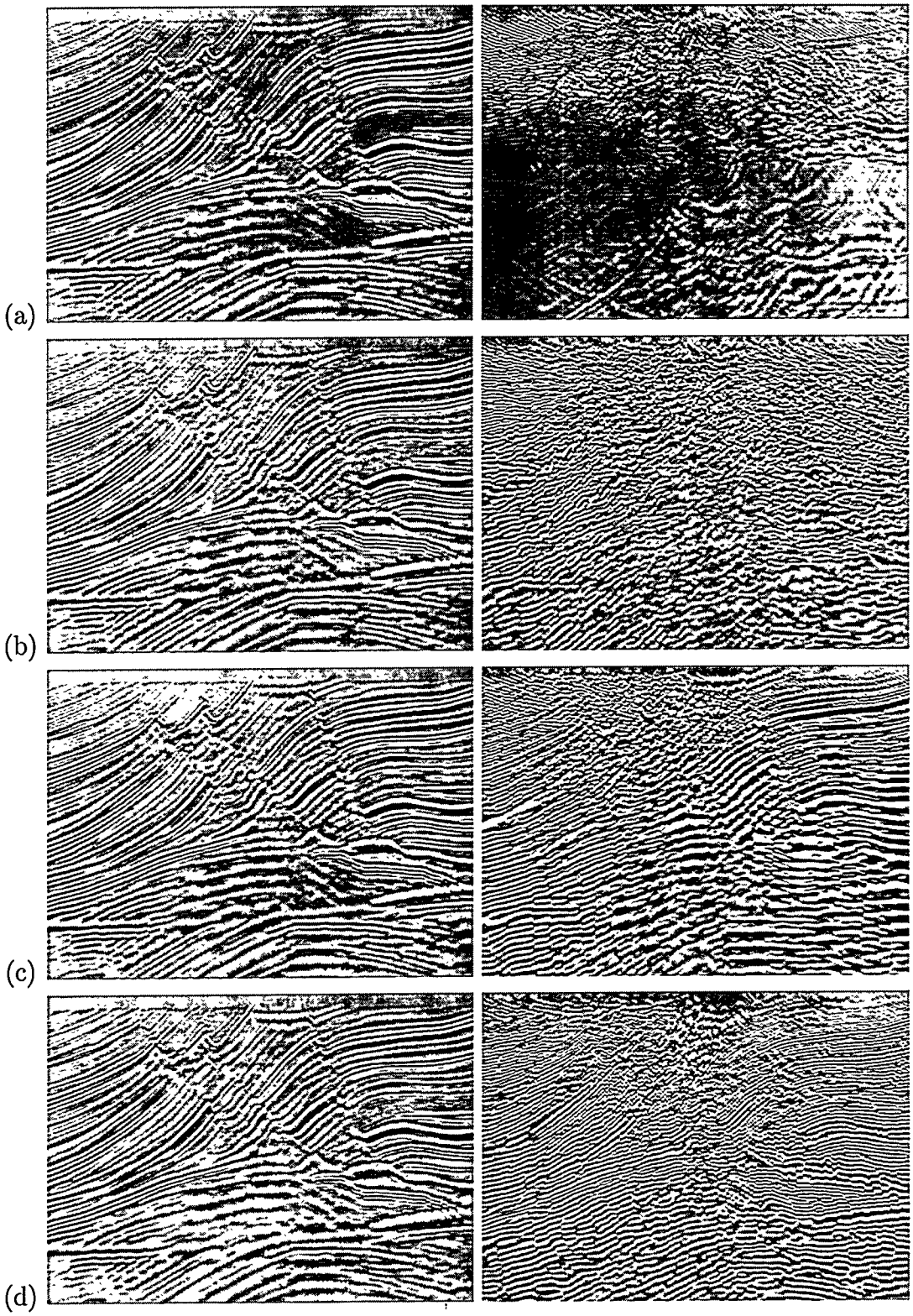
(a)
(b)
(c)
(d)

FIG. 6. The Marmousi Model combining 2 shots per migration using (a) linear, (b) random, (c) chirp, and (d) modified-chirp phase encoding. On the left is the encoded image, and on the right is the difference (amplified by a factor of 10) between the image in FIG. 5 and the encoded image.

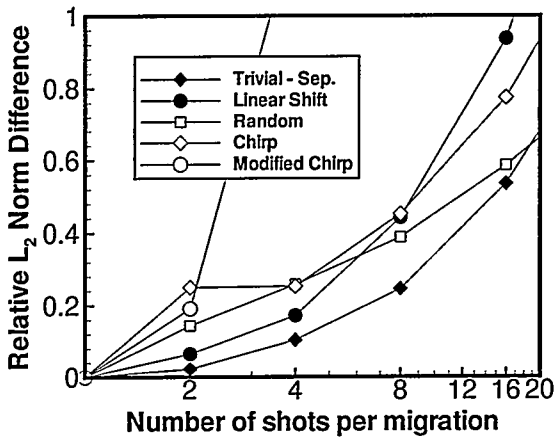

FIG. 7. L_2-norm difference between the standard image in FIG. 5 and the various encoding functions for the Marmousi dataset.

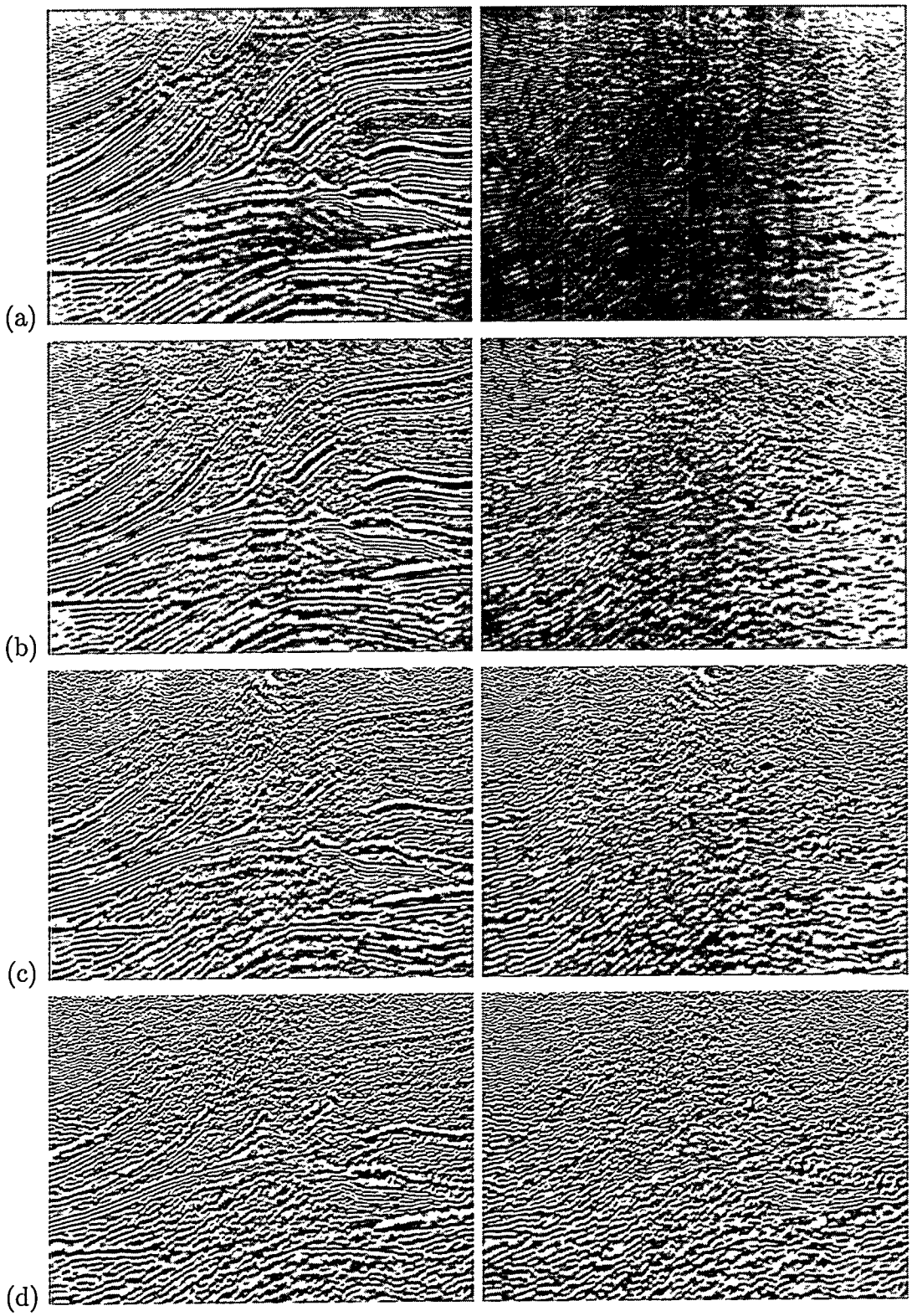

FIG. 8. The Marmousi Model using random encoding and combining (a) 4, (b) 16, (c) 60, and (d) 240 shots per migration. On the left is the encoded image, and on the right is the difference between the image in FIG. 5 and the encoded image.

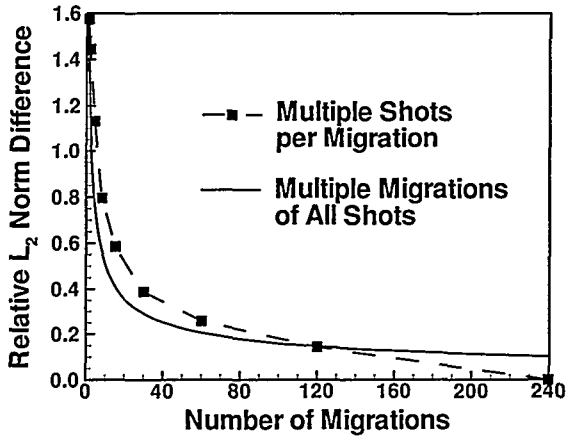

FIG. 9. The solid curve is the L_2 norm of the difference between the standard image in FIG. 5 and the stack of multiple migrations of all randomly encoded shot records. Each data point on the dashed curve is the result of a complete migration of all 240 shots using multiple shots per migration with each shot being used only once.

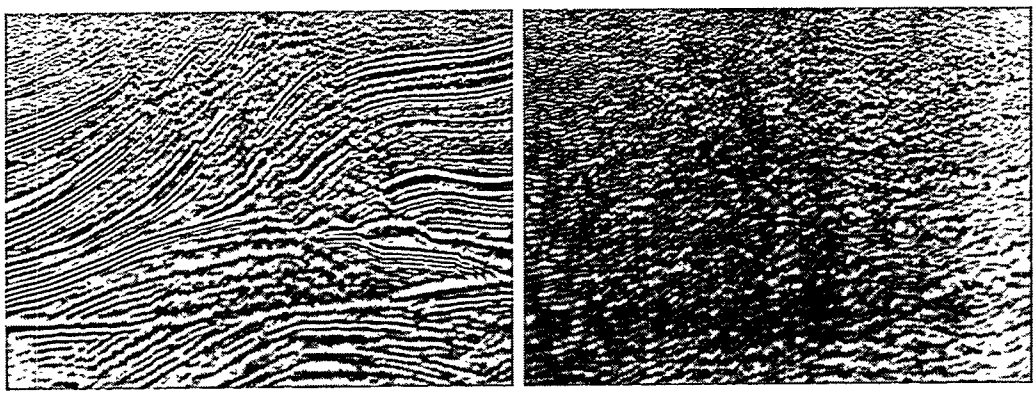

FIG. 10. The Marmousi Model using random encoding 15 times on all 240 shots and stacking the resulting images. On the left is the stacked image, and on the right is the difference between the image in FIG. 5 and the stacked image. Compare this image with FIG. 8(b).