Migration using short-time Fourier transforms

Biondo Biondi

INTRODUCTION

Fourier transforms are widely used in geophysics, and are basic to most of signal-processing algorithms. They are used in linear filtering and more generally in the application of convolutional operators since convolution becomes simple multiplication in the Fourier domain. The most significant disadvantage of Fourier methods is that their application is limited to time-invariant or space-invariant operators.

In this paper I present a generalization of Fourier methods that permits their application to time-variant or space-variant operators. This generalization is known in signal-processing literature as short-time Fourier transform theory (Portnoff, 1980; Crochiere and Rabiner, 1983). The basic idea on which short-time Fourier transform methods are based is to represent a one-dimensional function, for example one defined in the time domain, by a two-dimensional function defined in both the time and frequency domain.

In geophysics Fourier transforms are often used for solving the wave equation in migration or modeling of seismic data. In the solution of partial-differential equations such as the wave equation, methods based on Fourier transforms have a lot of potential advantages over methods based on finite difference. The derivatives of a function are correctly computed in the Fourier domain, but finite-differences approximations of derivatives lead to numerical dispersion. In reflection seismology it is often convenient to solve equations containing the square root of differential operators; the classic example is the one-way wave equation used in migration. In these cases a finite-difference solution requires one to approximate the square root by some rational expression whereas expressing the derivatives in the Fourier domain enables an exact computation of the square root. Solving partial-differential equations in the Fourier domain also allows one to avoid the stability problems inherent to finite-difference schemes. Moreover, the Fourier transform decomposes data into frequency components that can be processed separately; this is very convenient for parallel processing and for large datasets as 3-D data.

The major drawbacks of solving the wave equation in the Fourier domain are the computa-

tional cost and the assumption of a constant wave-propagation velocity along the spatial axes that are transformed into the wavenumber domain. The computational cost is becoming a less important issue with the widespread use of highly vectorized computers and hardware optimized for computing Fast Fourier Transforms (FFTs). The constant velocity assumption is a restricting one, and I would like to lift it by introducing the short-time transforms to generalize the Fourier methods to the cases in which velocity is varying. Short-time Fourier transforms are more expensive to apply than classical Fourier methods; in general the additional cost for properly handling velocity variation depends on the rate of variation of the velocity model and on the grade of accuracy required in the final result.

The short-time Fourier analysis can have others useful applications in geophysics, such as spectrum analysis or deconvolution of nonstationary signals (Rabiner and Allen, 1980), but in this paper I will focus on the application of short-time Fourier transforms to the wave-equation operators.

SHORT-TIME FOURIER TRANSFORMS

The goal of short-time Fourier theory is to extend the use of the Fourier transforms to the application of variant linear operators. In this section I describe the fundamental concepts of the short-time Fourier transforms theory; the reader interested in more particulars can find them in Portnoff (1980) and Crochiere and Rabiner (1982).

First I present the analysis and the synthesis procedures: the analysis gives us a representation of the data in both the time and frequency domains; the synthesis is the transpose of the analysis and transforms the data back to the original time domain. Following this I consider the application of variant linear operators in the mixed time-frequency domain, and I describe a fast algorithm for this application.

Analysis and synthesis of a signal using short-time Fourier transforms

The Fourier transform gives a representation of a discrete sequence $x(t)$ defined in the time domain, in terms of its frequency components $X(\omega)$. The short-time Fourier transform represents $x(t)$ as a two-dimensional function $X(t,\omega)$ defined in both time and frequency. This is achieved by windowing $x(t)$ using a sliding *analysis window* $h(t)$, and then Fourier transforming the windowed data. The windowed data $x_2(t_w, m)$ is a two-dimensional function of the time t_w of the window, and of the position m inside the window:

$$x_2(t_w, m) = h(t_w - m)x(m). \tag{1}$$

The short-time Fourier transform is given by the equation

$$X_2(t_w,\omega) = \sum_{m=-\infty}^{\infty} x_2(t_w,m)\exp(-i\omega m) = \sum_{m=-\infty}^{\infty} h(t_w-m)x(m)\exp(-i\omega m). \tag{2}$$

The recovering of the original function from its short-time transform can be performed in two steps: the first is to transform $X_2(t_w,\omega)$ back to $x_2(t_w,m)$, and the second is to weight $x_2(t_w,m)$ with a *synthesis window* $f(t)$ and sum all the windows. The synthesized signal $x_s(t)$ is given by

$$x_s(t) = \frac{1}{2\pi}\sum_{t_w=-\infty}^{\infty}\int f(t-t_w)X_2(t_w,\omega)\exp(i\omega t)d\omega = \sum_{t_w=-\infty}^{\infty} f(t-t_w)x_2(t_w,t). \tag{3}$$

From equation (3) we can compute the recovered signal $x_s(t)$ as a function of the analysis and synthesis windows, and of the original signal $x(t)$:

$$x_s(t) = \sum_{t_w=-\infty}^{\infty} f(t-t_w)h(t_w-t)x(t) = \left[\sum_{m=-\infty}^{\infty} f(-m)h(m)\right]x(t). \tag{4}$$

The synthesized function $x_s(t)$ is equal to the original $x(t)$ if and only if

$$\sum_{m=-\infty}^{\infty} f(-m)h(m) = 1. \tag{5}$$

The short-time transform depends on the local characteristics of the function $x(t)$, and the reconstructed signal $x_s(t)$ depends locally on the short-time transform $X(t_w,\omega)$. Because the short-time Fourier transform is not a global transform like the conventional Fourier transform but a local one, it can be used with variant operators.

Variant linear operators

The behavior of a linear operator is described by its Green function, or equivalently by its impulse response.

If the operator is invariant the impulse response $r(m)$ depends only on the distance m from the impulse; the result $y(t)$ of the application of the operator to the function $x(t)$ is given by the convolution

$$y(t) = \sum_{m=-\infty}^{\infty} r(m)x(t-m), \tag{6}$$

or in the frequency domain by the multiplication

$$Y(\omega) = R(\omega)X(\omega). \tag{7}$$

If the operator is variant the impulse response $r_2(t,m)$ also depends on the impulse time t. The convolution of equation (6) becomes

$$y(t) = \sum_{m=-\infty}^{\infty} r_2(t,m)x(t-m). \tag{8}$$

I want to find an equivalent expression for this convolution in the in the mixed time-frequency domain defined by the short-time Fourier transforms, just as for invariant operators equation (7) is the frequency domain equivalent of equation (6).

Transforming the impulse response $r_2(t,m)$ over m we find the time and frequency representation of the impulse response:

$$R_2(t,\omega) = \sum_{m=-\infty}^{\infty} r_2(t,m)\exp(-i\omega m). \tag{9}$$

Multiplying the short-time transform of a signal $x(t)$ by $R_2(t,\omega)$ leads to a modified short-time transform

$$Y_2(t,\omega) = R_2(t,\omega)X_2(t,\omega). \tag{10}$$

This expression is a multiplication in the mixed time-frequency domain; however to be equivalent to equation (8) the synthesized function $y_s(t)$ from $Y_2(t,\omega)$ must be equal to the result $y(t)$ of the convolution of $x(t)$ with $r_2(t,m)$. The synthesized function $y_s(t)$ from $Y_2(t,\omega)$ is

$$y_s(t) = \frac{1}{2\pi}\sum_{t_w=-\infty}^{\infty}\int f(t-t_w)R_2(t_w,\omega)X_2(t_w,\omega)\exp(i\omega t)d\omega. \tag{11}$$

It can be shown (Portnoff, 1980) that $y_s(t)$ is only approximately equal to $y(t)$, and that it is equal to the result of the convolution of $x(t)$ with a modified impulse response $\bar{r}_2(t,m)$, that is

$$y_s(t) = \sum_{m=-\infty}^{\infty} \bar{r}_2(t,m)x(t-m), \tag{12}$$

where the modified impulse response $\bar{r}_2(t,m)$ is equal to

$$\bar{r}_2(t,m) = \sum_{t_w=-\infty}^{\infty} f(t_w)h(m-t_w)r(t-t_w,m) = \sum_{t_w=-\infty}^{\infty} w(t_w,m)r(t-t_w,m). \tag{13}$$

The two-dimensional function $w(t_w,m)$ is the *effective window* and is given by the product of the synthesis window with a delayed version of the analysis window:

$$w(t_w,m) = f(t_w)h(m-t_w). \tag{14}$$

Equation (13) tell us which accuracy we can achieve when performing a variant convolution using the short-time transforms; the accuracy of the result depends on the characteristics of the windows $h(t)$ and $f(t)$, and of the impulse response $r_2(t,m)$. The original impulse response $r_2(t,m)$ is smoothed along t, and is windowed along m by the effective window $w(t_w,m)$. This effect can also be seen in the frequency domain where the equivalent of equation (13) is

$$\bar{R}_2(\psi,m) = W(\psi,m)R_2(\psi,m), \tag{15}$$

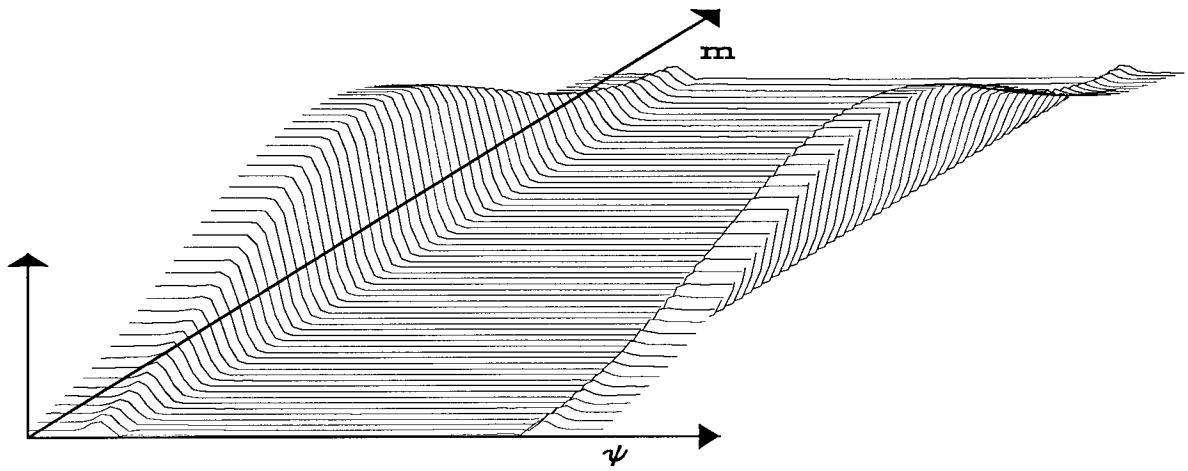

FIG. 1. Absolute value of the effective window $W(\psi, m)$. The cutoff frequency in ψ is at 1/4 of the Nyquist. The window is nicely sharp in ψ but is tapered down in m.

where ψ is the frequency dual of t, that is the spectrum in ψ depends on the variation of the impulse response with time. In equation (15) $R_2(\psi, m)$ is windowed in m and filtered in ψ by $W(\psi, m)$. The effective window $w(t_w, m)$ is a two-dimensional function given by the product of two one-dimensional functions; we can freely design the two windows $h(t)$ and $f(t)$ but $w(t_w, m)$ is the result of the product in equation (14) and its choice is constrained. Owing to this constraint on the choice of the effective window the modified impulse response $\bar{r}(t, m)$ can be long in time or rapidly varying, but cannot be at the same time long and rapidly varying. In general, to get an effective window that allows for significant time variation the analysis window must be longer than the synthesis window.

Figures 1 and 2 show the absolute value of $W(\psi, m)$ for two different choices of windows. In both cases the analysis window is longer than the synthesis window and their length is chosen for positioning the cutoff frequency in the ψ domain at one fourth of the Nyquist frequency. The differences between the two examples is in the way the windows are tapered; for the first figure I used a Hamming window and for Figure 2 a boxcar window. In the first case $W(\psi, m)$ is nicely sharp at the cutoff frequency on the ψ axis but is tapered in the m direction; this effect decreases the amplitude of the larger lags of the impulse response. In the second case $W(\psi, m)$ is, on the contrary, almost constant for all the lags of the impulse response but shows a lot of ringing along the ψ direction. These examples show some of the problems that can be encountered designing $h(t)$ and $f(t)$. The best solution should be found using some procedure to optimize the result as

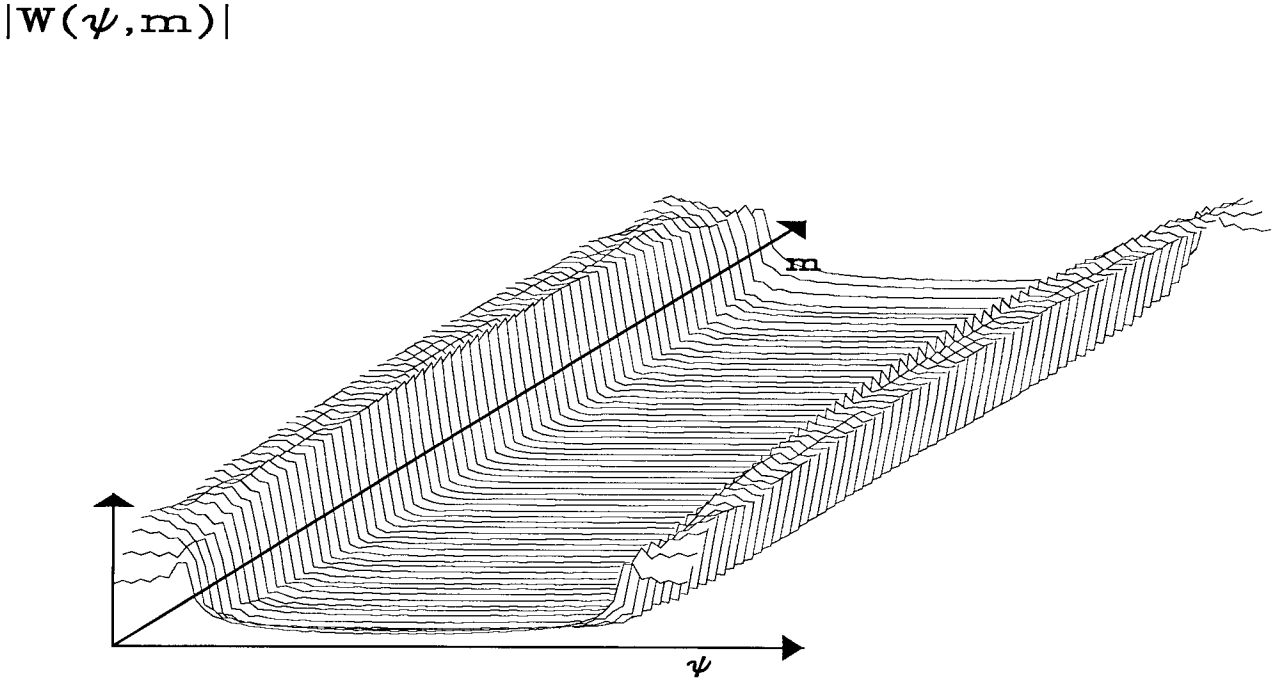

FIG. 2. Absolute value of the effective window $W(\psi, m)$. The cutoff frequency ψ is at 1/4 of the Nyquist. The window is almost constant along m but shows a lot of ringing in ψ.

a function of the constraints given by the particular application.

Fast algorithms for applying short-time Fourier transforms

Applying variant convolutions as described in the previous section leads to expensive algorithms because the short-time transform is a function of the continuous variable ω and of each value of the time index t. However the short-time transform contains redundant information about the signal; that is, the values at different times and at different frequencies are not independent. Therefore, a short-time representation of a signal can be undersampled.

There are two possible approaches to computing efficiently the undersampled short-time transform of a signal. They are equivalent in computational requirements and in the accuracy of their results. The first is using polyphase structures (Crochiere and Rabiner, 1983; Portnoff, 1980), and the second is using an overlap-add structure (Crochiere and Rabiner, 1983). In this paper I describe and use in my examples an overlap-add structure, because it is simpler.

The overlap-add approach works as follows: to undersample in time we use only some of the windows of the data; to undersample in frequency we use finite-length windows and a Discrete Fourier Transform (DFT) to transform the windowed data in the frequency domain. Then, if T is the undersampling period, that is, the delay between two consecutive windows, and $\hat{t}$ is the undersampled time index, such that $t = \hat{t}\,T$, the undersampled windows $\hat{x}_2(\hat{t}_w, m)$ of the data are

equal to

$$\hat{x}_2(\hat{t}_w, m) = x_2(\hat{t}_w T, m) = h(\hat{t}_w T - m)x(m). \tag{16}$$

If the analysis window is M samples long, the undersampled angular frequency $\hat{\omega}$ is sampled every $\Omega = 2\pi/M$ radians and the undersampled short-time transform $\hat{X}_2(\hat{t}_w, \hat{\omega})$ is equal to

$$\hat{X}_2(\hat{t}_w, \hat{\omega}) = X_2(\hat{t}_w T, \hat{\omega}\Omega) = \sum_{m=0}^{M-1} \hat{x}_2(\hat{t}_w, m) \exp(-i\hat{\omega}\Omega m). \tag{17}$$

The undersampled transform of the impulse response of the operator with which we want to convolve the sequence $x(t)$ is

$$\hat{R}_2(\hat{t}, \hat{\omega}) = R_2(\hat{t}T, \hat{\omega}\Omega) = \sum_{m=0}^{M-1} r(\hat{t}T, m) \exp(-i\hat{\omega}\Omega m). \tag{18}$$

Convolution is performed in the short-time transform domain by the multiplication $\hat{Y}_2(\hat{t}_w, \hat{\omega}) = \hat{R}_2(\hat{t}_w, \hat{\omega})\hat{X}_2(\hat{t}_w, \hat{\omega})$ (equation (10)), and the synthesis formula of an undersampled short-time transform $\hat{Y}_2(\hat{t}_w, \hat{\omega})$ using a synthesis window M samples long is

$$y_s(t) = \frac{1}{M} \sum_{\hat{t}_w = \frac{M}{2}-1}^{\frac{M}{2}} \sum_{\hat{\omega}=0}^{M-1} f(t - \hat{t}_w T)\hat{Y}_2(\hat{t}_w, \hat{\omega}) \exp(i\hat{\omega}\Omega t). \tag{19}$$

The undersampling rate allowed by the overlap-add structure is limited by the characteristics of the effective window $w(t_w, m)$ and of the impulse response $r(t, m)$. It can be shown that equation (15) can be generalized to the undersampled case, becoming the equation

$$\bar{R}_2(\psi, m) = W(\psi, m)\frac{1}{T} \sum_{q=0}^{T-1} \sum_{p=-\infty}^{\infty} R_2(\psi - 2\pi q/T, m - pM). \tag{20}$$

This equation is fundamental for applying correctly the short-time transform method to variant convolution. It shows that, beside the windowing of the impulse response $R(\psi, m)$ as in the continuous case (equation (15)), in the undersampled case the impulse response is distorted by its replications on the time axis m and on the frequency axis ψ. First we can check that without undersampling (that is, with $T = 1$ and $M \Rightarrow \infty$), equation (20) becomes equation (15). Second, we find the constraints on T and M, according to the characteristics of $R_2(\psi, m)$, to prevent the overlapping of the replications of $R_2(\psi, m)$. To avoid time aliasing inside the windows, the impulse response of the operator must be shorter than M samples, for every ψ. To prevent frequency aliasing of the impulse response, the sampling frequency $2\pi/T$ must be larger than the bandwidth of $R_2(\psi, m)$, for every m (Figure 3). Equation (20) also gives us the criteria for designing the windows: to eliminate images of $R_2(\psi, m)$ in the time domain, M must be larger than the time duration of $W(\psi, m)$, and to eliminates images in the frequency domain, $2\pi/T$ must be larger

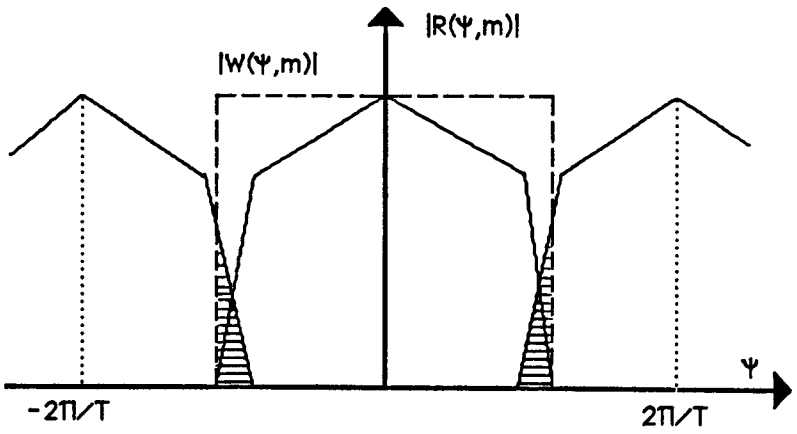

FIG. 3. Example of aliasing of the impulse response $R(\psi, m)$ in the frequency domain ψ. The undersampling period T is too large for the bandwidth of the impulse response.

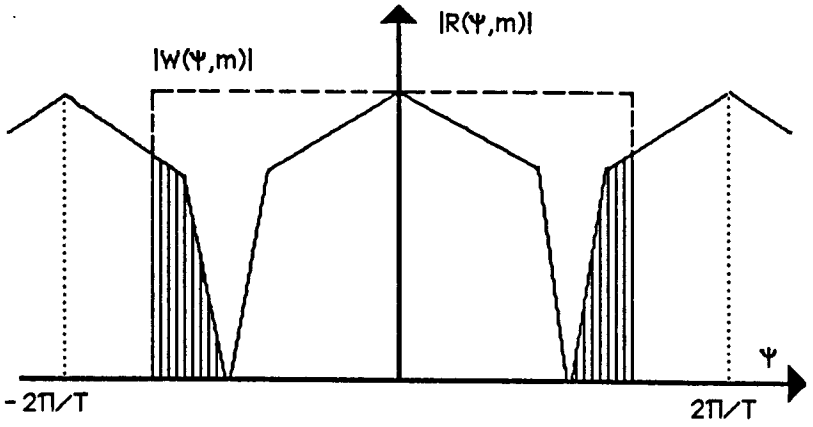

FIG. 4. Example of imaging of the impulse response $R(\psi, m)$ in the frequency domain ψ. The window $W(\psi, m)$ is too large for the undersampling period T and does not window out the replications of $R(\psi, m)$.

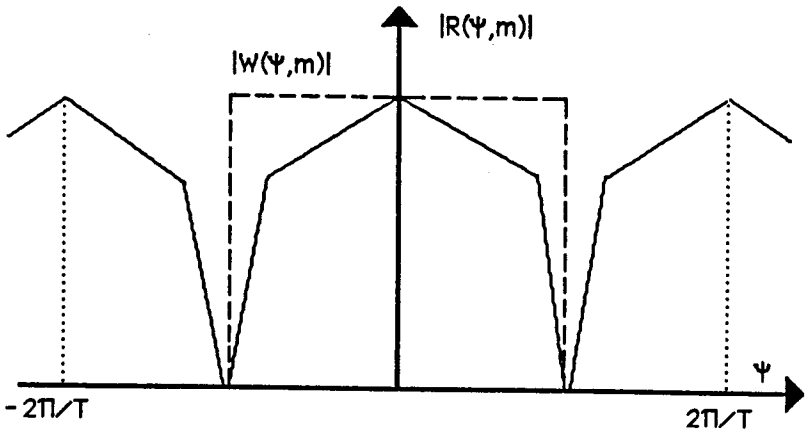

FIG. 5. Example of a correct choice of the undersampling period T and of the window $W(\psi, m)$. Neither aliasing nor imaging occurred .

than the bandwidth of $W(\psi, m)$ (Figure 4). The ideal situation, in which neither aliasing nor imaging occurs, is shown in Figure 5.

The frequency-aliasing problem of the impulse response can be also studied as a problem of sampling the two-dimensional function $r_2(t, m)$ along the t axis; the sampling period T depends on the frequency content of $r_2(t, m)$ as a function of t. In the particular case in which the variant operator is a pure time-shift operator, as for example the Normal Moveout operator (NMO), the variations of the time shift cause a dipping *event* in the function $r_2(t, m)$; the goal is then not to alias these dips. The dip α of these events is equal to the rate of variation with t of the time shift Δt, that is,

$$\alpha = \frac{d\Delta t}{dt}. \tag{21}$$

The result from two-dimensional sampling theory is that, if the maximum frequency along m is $\omega_{max} = 1/\tau$, the constraint on the undersampling period T, as a function of α and τ, is

$$T\,|\alpha| \leq \frac{\tau}{2}. \tag{22}$$

This is a useful constraint on the undersampling period T and for many geophysical applications is the relation to use to determine the correct undersampling period T.

If α is changing with time, it is possible to change T as a function of the local rate of variation of the impulse response to save computational effort. Moreover, when the impulse response is convolved with the data, ω_{max} is the maximum frequency of the input data $x(t)$. This last considerations allows us to relax the constraint in (22) and to use a larger undersampling period T.

1-D EXAMPLE: NMO

The Normal Moveout operator is an example of a one-dimensional operator that is linear and time variant. In this paper I use NMO as a simple and useful example to show some of the properties of the short-time Fourier method. I am not suggesting that the short-time Fourier method be used to apply the NMO transformation; I think NMO is more efficiently implemented by the usual interpolation methods.

NMO is a transformation of variables that transforms a seismic trace in *zero-offset* data. The transformation is dependent on the velocity of the medium V and on the offset of the trace x and is given by the equation

$$t^2 = t_0^2 + \frac{x^2}{V^2}\,, \tag{23}$$

where t is the time of the original trace, and t_0 is the *zero-offset* time. The NMO operator is thus a time-shift operator, and the time shift Δt expressed as a function of the zero-offset time

t_0 is

$$\Delta t(t_0) = t - t_0 = \sqrt{t_0^2 + \frac{x^2}{V^2}} - t_0 \ . \tag{24}$$

The time and frequency representation of the NMO operator is

$$NMO(t_0, \omega) = \exp(i\omega \Delta t(t_0)) = \exp\left[i\omega\left(\sqrt{t_0^2 + \frac{x^2}{v^2}} - t_0\right)\right], \tag{25}$$

and the NMO transformation of a trace $tr(t)$ in terms of its short-time Fourier representation $tr(t, \omega)$ is given by the multiplication in the time-frequency domain

$$tr_{NMO}(t_0, \omega) = NMO(t_0, \omega) tr(t_0, \omega) = \exp\left[i\omega\left(\sqrt{t_0^2 + \frac{x^2}{v^2}} - t_0\right)\right] tr(t_0, \omega). \tag{26}$$

The duration of the impulse response of the NMO operator is $\Delta t(t_0)$, and to avoid aliasing in the time domain the windows must have a duration of at least $\Delta t(t_0)$.

Avoiding the frequency aliasing of the impulse response is a more complicated problem. The undersampling period T depends on the rate of change of $\Delta t(t_0)$ as a function of t_0 and on the dominant period τ of the input data. From the consideration of this problem in the previous section (equation (22)) we can conclude that, to sample adequately the impulse response, T must satisfy the constraint

$$T\left|\frac{d\Delta t(t)}{dt}\Big|_{t=t_0}\right| = T\left(1 - \frac{t_0}{\sqrt{t_0^2 + \frac{x^2}{v^2}}}\right) \leq \frac{\tau}{2} \ . \tag{27}$$

I have tested an NMO algorithm using short-time Fourier transforms on a synthetic trace with different undersampling periods and different windows parameters to verify the conclusions of the theory in a simple situation. Figure 6 shows the synthetic trace and the result of NMO computed by a standard algorithm using a sinc interpolation. It also shows the amplitude spectrum of the two traces: in the trace after NMO the higher frequencies were attenuated by the NMO stretch. The input trace is a zero-phase bandpass wavelet with the cutoff frequency at 40 Hz. The NMO velocity is 3,000 m/sec, and the offset is 1,000 m.

Figure 7 shows NMO computed on the same input trace and with the same velocity and offset as before, but using a short-time transform algorithm. The undersampling period T is .032 seconds, corresponding to a distance between adjacent windows of eight samples and derived by the application of equation (27) for t_0 approximately equal to the wavelet time. The windows $f(t)$ and $h(t)$ were designed to position the cutoff frequency of $W(\psi, m)$ at 15.65 Hz, that is, 1/8 of the Nyquist frequency, to avoid imaging. The windows are 128 samples, long plus others 128 samples of zero padding, to prevent wraparound.

To produce the trace presented in Figure 8 I reduced the cutoff frequency of $W(\psi, m)$ to 1/16 the Nyquist frequency. This make the window too narrow in ψ and it smoothed too much

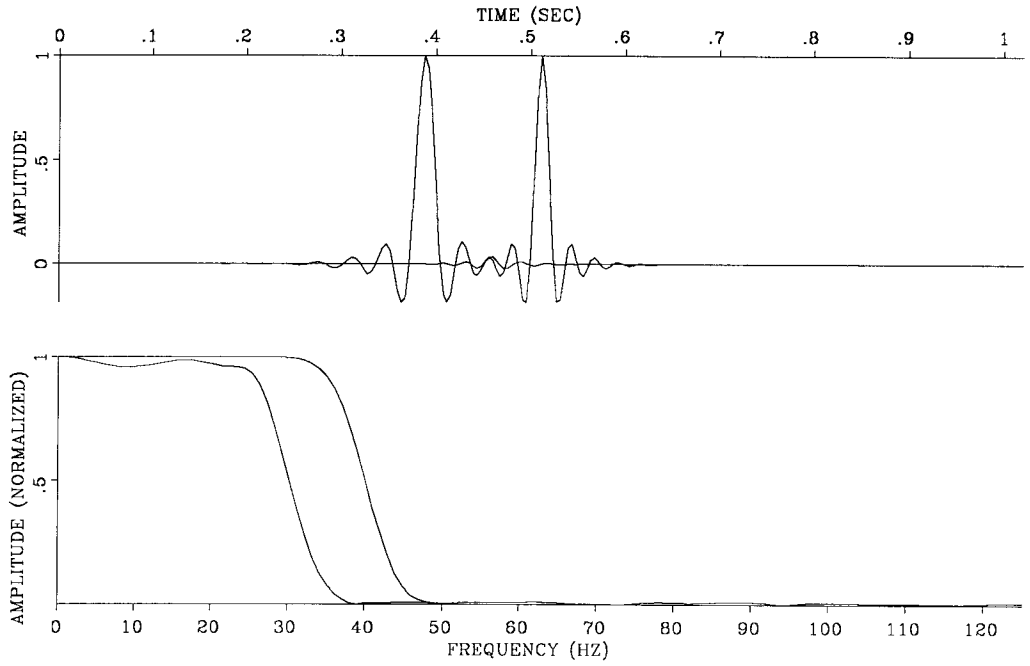

FIG. 6. A synthetic trace and its amplitude spectrum, before and after NMO using a standard interpolation algorithm. The NMO velocity is 3,000 m/sec, and the offset is 1,000 m.

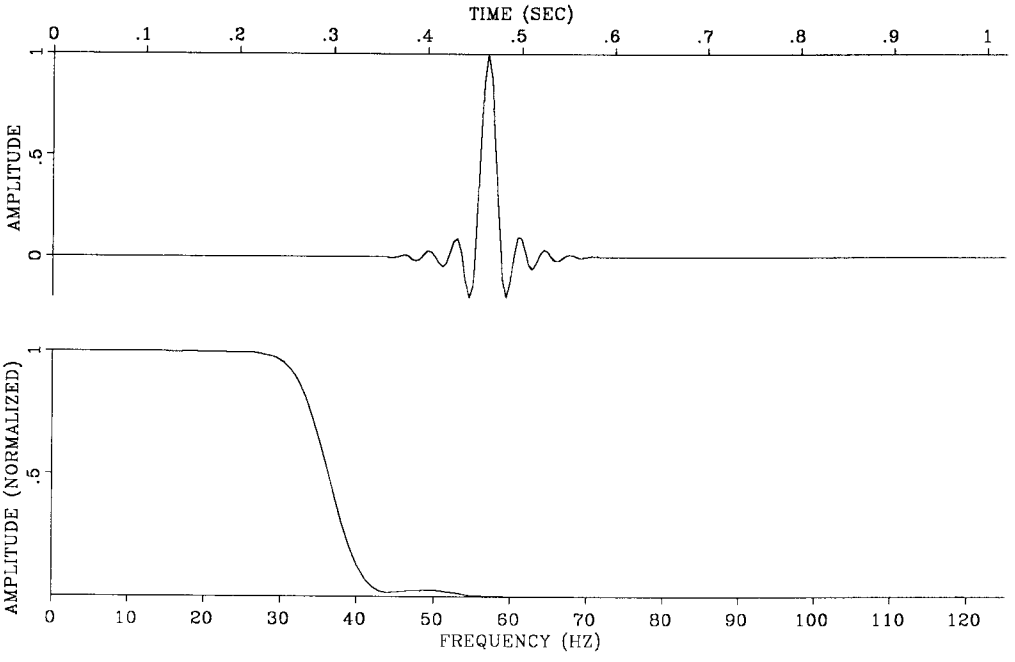

FIG. 7. The result of NMO, using a short-time Fourier transforms algorithm, on the same synthetic trace as in Figure 6. The undersampling period T was derived by equation (27).

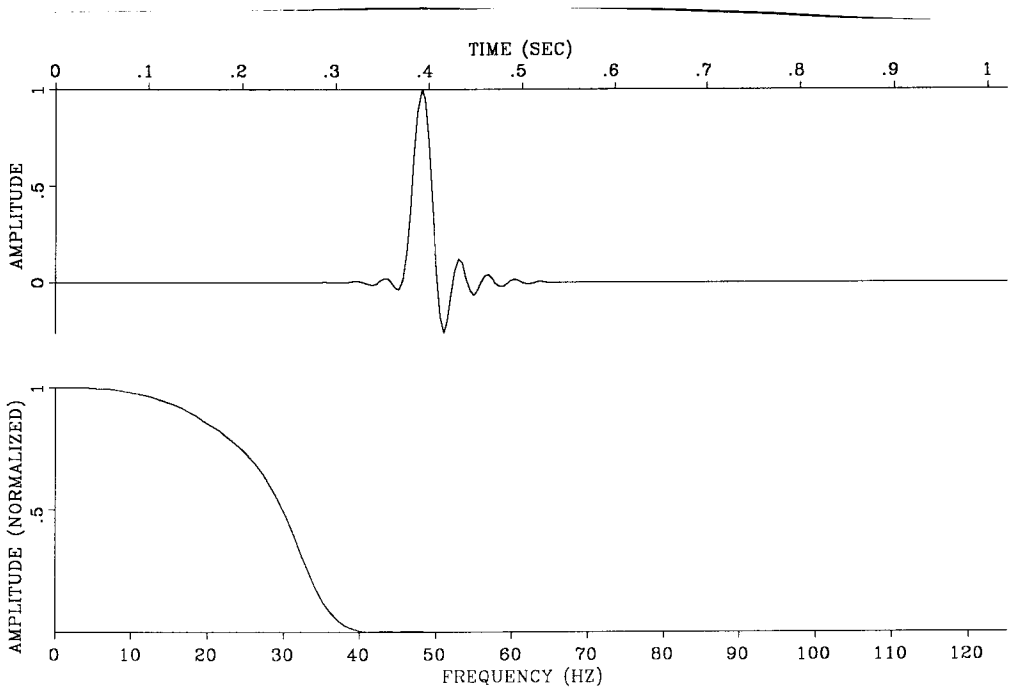

FIG. 8. The result of NMO, using a short-time Fourier transforms algorithm, when the window $W(\psi, m)$ is too narrow in the frequency ψ. The impulse response variations are smoothed, particularly at earlier times.

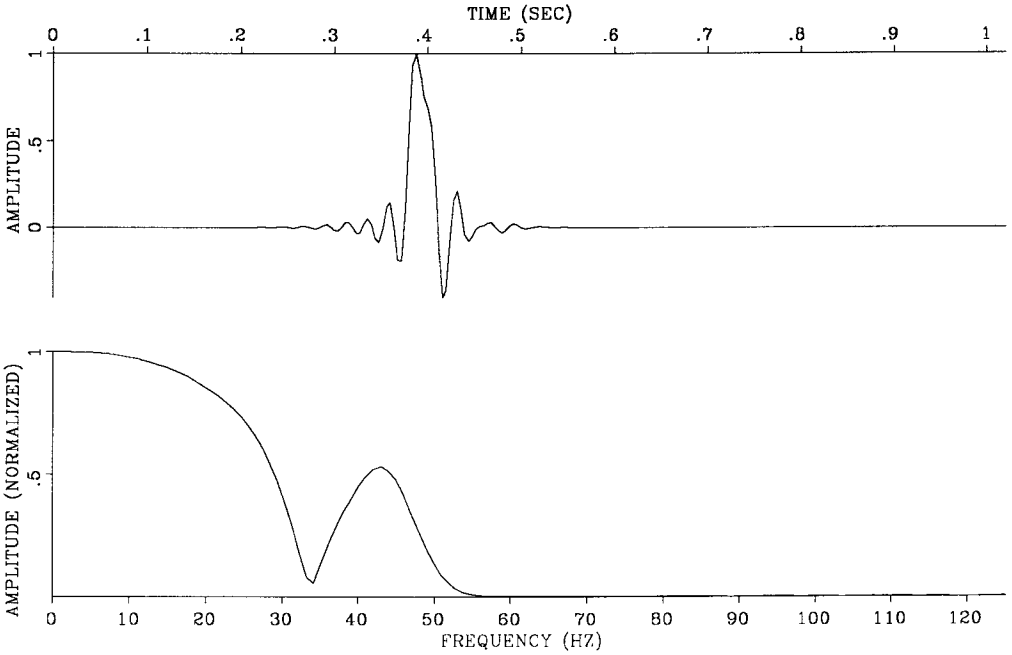

FIG. 9. The result of NMO, using a short-time Fourier transforms algorithm, when the undersampling period T is too long and aliasing of the impulse response $R(\psi, m)$ occurs.

the variation of the impulse response, particularly at earlier times when the time shift Δt was decreasing faster. When I also doubled T to 0.064 seconds, aliasing occurred on the ψ axis. The result is displayed in Figure 9. However, since aliasing depends on the rate of variation of the impulse response, and decreasing the offset make the operator less variant, changing the offset to 650 m was enough to eliminate it, as it shown in Figure 10. Aliasing also depends on the maximum frequency of the input signal, and thus decreasing the bandwidth of the input wavelet to 10 Hz removed the problem. The result was the correctly moved out trace displayed in Figure 11.

Finally I show the case of an imaging problem in Figure 12. All the processing parameters here were the same as in Figure 7 except for the cutoff frequency of $W(\psi, m)$, which was 1/4 of the Nyquist. Now the $W(\psi, m)$ is too large in ψ and does not window out the replication of the impulse response along the ψ axis. To eliminate this problem I halved the undersampling period T and produced the correct result shown in Figure 13. This result is comparable to the one shown in Figure 7 but is twice as expensive to compute because the undersampling period T was reduced.

Concluding my discussion of NMO examples, I would like to summarize a procedure for setting the parameters of a short-time Fourier transform algorithm. First an expression for the impulse response as a function of the time is needed. From this expression determine the length of the windows to avoid time aliasing and derive the rate α of variation of the impulse response. From α determine the maximum undersampling period T that does not produce frequency aliasing. The last step is to design the windows correctly to avoid both smoothing the impulse response (Figure 8) and imaging (Figure 12).

WAVE-EQUATION OPERATORS

Many of wave-equation algorithms using Fourier transforms require the assumption of constant velocity along the spatial axes that are transformed into the wavenumber domain. In principle short-time Fourier transform theory can be applied to any of these algorithms to lift this restrictive assumption.

In this section I describe a phase-shift migration algorithm generalized to handle lateral velocity variations by means of short-time transforms. I also implemented a Stolt migration algorithm for layered media and a modeling program that uses the one-way wave equation to evolve the wave field in time. I did not present the last two algorithms in this paper because the basic ideas are similar to the ideas used in the phase-shift case. The modeling algorithm can also be used for reverse-time migration, where it has the advantage, respect to finite-difference implementations of the full wave equation, that it does not produce internal reflections at velocity discontinuity since it evolves the wave field in time using a one-way wave equation. I used this modeling program to

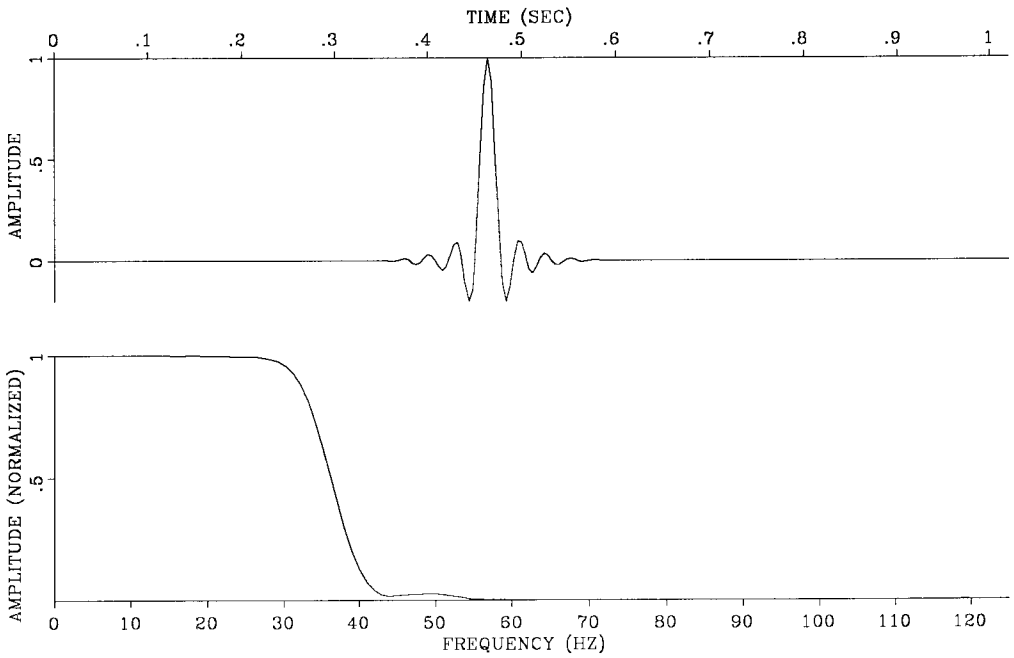

FIG. 10. The result of NMO using the same undersampling period as in Figure 9 but the offset is now 650 m. The operator was varying slower and thus the impulse response was not aliased.

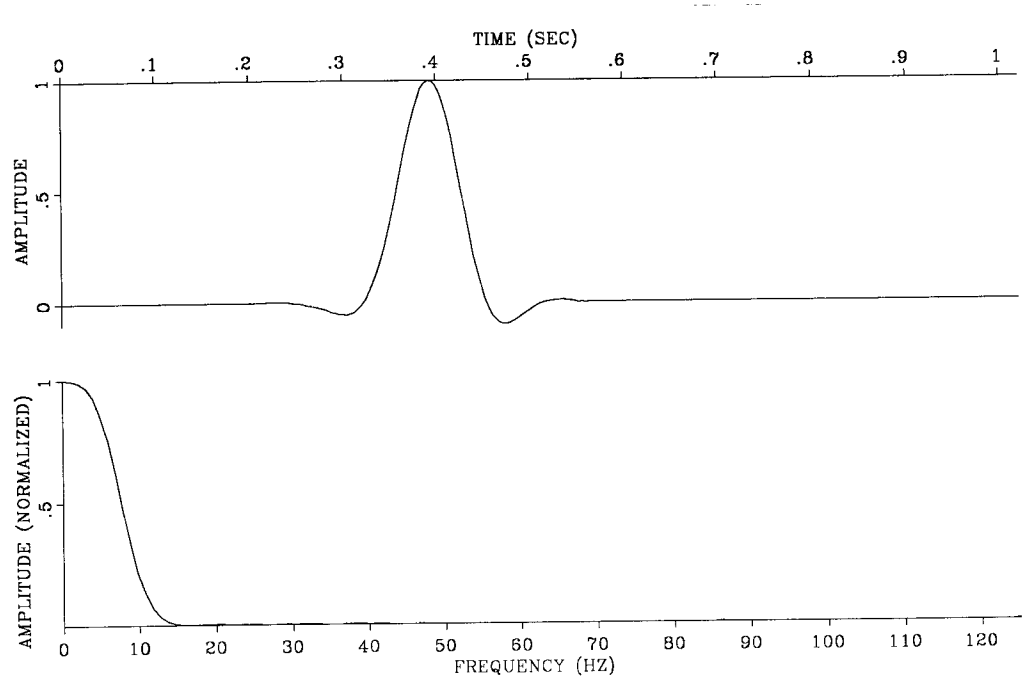

FIG. 11. The result of NMO using the same undersampling period as in Figure 9 and Figure 10. Here the result is not aliased because the input trace had narrower frequency bandwidth.

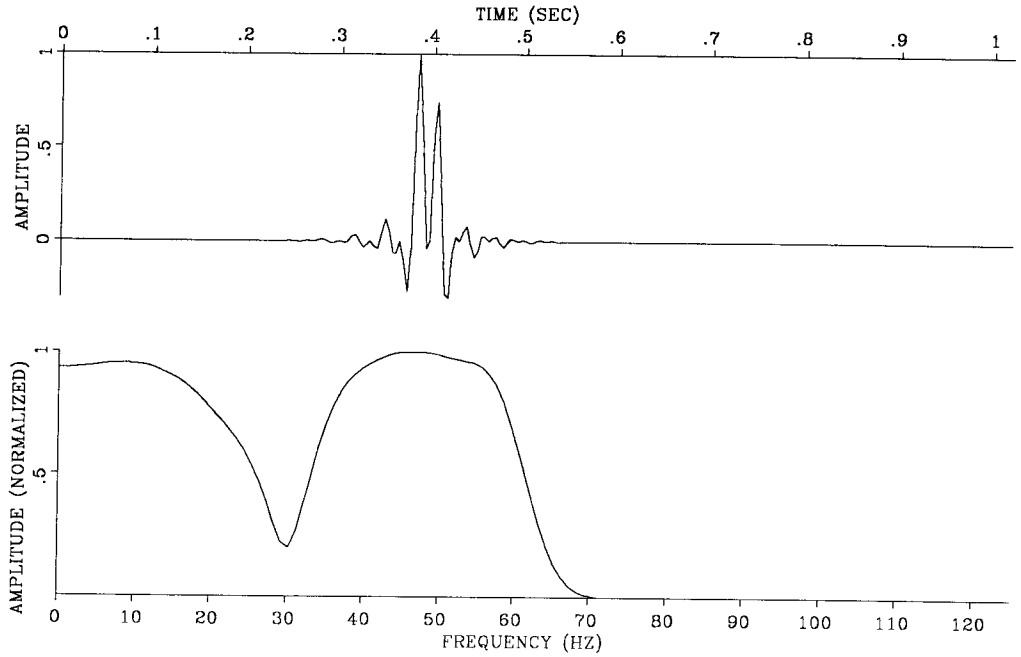

FIG. 12. The result of NMO using the same parameters as in Figure 7 except for the window $W(\psi, m)$ that was too large in the frequency ψ to window out the replications of the impulse response $R(\psi, m)$. In this case imaging of the impulse response occurred.

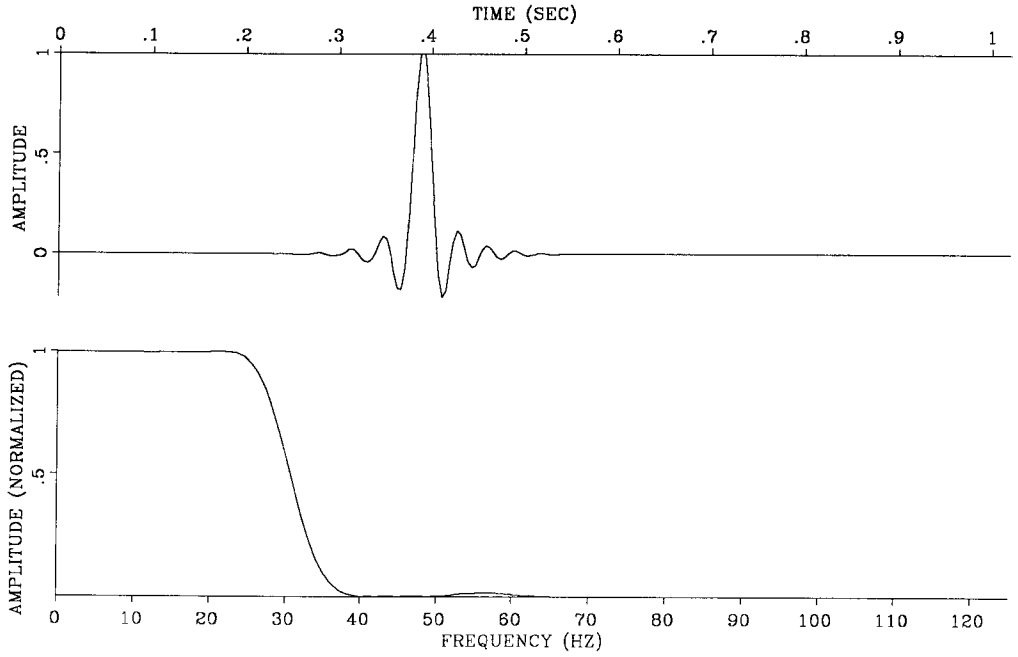

FIG. 13. The result of NMO using the same parameters as in Figure 12 except for the undersampling period T that was reduced to avoid imaging. The result is equivalent to that shown in Figure 6 but with twice the computational cost.

generate the synthetic zero-offset sections to test the phase-shift migration method.

Short-time Fourier transforms could also be used to implement Dip-Moveout (DMO) algorithms in the frequency-wavenumber domain. In the DMO case the capability of handling time variations could be exploited to operate efficiently the time varying convolution defined by the DMO operator.

Wave equation operators are two-dimensional operators but their application using short-time Fourier transforms follows the same principles of one-dimensional applications, such as NMO. The major difference is that the impulse responses are two-dimensional functions and we must avoid aliasing of the impulse response variations in both the two-dimensions, even if we used the short-time Fourier transforms only along one axis.

Phase-shift migration

Phase-shift migration is an accurate migration method but it is limited by the requirement of no lateral variation of the medium velocity. The algorithm is based on the downward continuation of the wave field by a phase-shift operator in the frequency-wavenumber domain. The operator is velocity dependent and thus when velocity is horizontally varying we need to operate short-time Fourier transforms along the horizontal axis x.

The short-time transformed wave field at the surface is

$$P(\omega, k_x, x_w, z=0) = \sum_{m=-\infty}^{\infty} h(x_w - m) P(\omega, m, z=0) \exp(-ik_x m), \tag{28}$$

where $P(\omega, m, z=0)$ is the recorded wave field transformed along the time axis. The operator used to downward continue the wave field in the frequency-wavenumber domain is the pure phase-shift filter

$$PHASE(\omega, k_x, x) = \exp\left(i\Delta z \sqrt{\frac{\omega^2}{v(x,z)^2} - k_x^2}\right). \tag{29}$$

The transformed wave field downward continued at depth Δz is

$$P(\omega, k_x, x_w, z=\Delta z) = P(\omega, k_x, x_w, z=0) \exp\left(i\Delta z \sqrt{\frac{\omega^2}{v(x_w,z)^2} - k_x^2}\right). \tag{30}$$

The same wave field defined in x and ω is

$$P(\omega, x, z=\Delta z) = \frac{1}{2\pi} \sum_{x_w=-\infty}^{\infty} \int f(x - x_w) P(\omega, k_x, x_w, z=\Delta z) \exp(ik_x x) dk_x. \tag{31}$$

Repeating the downward continuation and imaging the wave field at time equal zero, coherently with the exploding-reflector hypothesis, produces the migrated section

$$Mig(x,z) = P(t=0, x, z) = \frac{1}{2\pi} \int P(\omega, x, z) d\omega\ . \tag{32}$$

The impulse response of the phase-shift operator, used to downward continue the wave field from depth z_0 to depth $z_0 + \Delta z$, is the portion of hyperbola

$$v(x_0, z_0)^2 (t - t_0)^2 - (x - x_0)^2 = \Delta z^2, \quad 0 \leq t \leq t_0 \,, \tag{33}$$

where t_0 is the impulse time and x_0 is the impulse horizontal location.

For implementing an algorithm with the undersampled versions of the analysis equation (28) and of the synthesis equation (31) we must determine the undersampling distance X and the windows length M that do not cause aliasing.

The phase-shift operator moves the wave energy in both time and space. To determine the correct undersampling distance X between the windows we must derive an expression for the time shift Δt and for the lateral shift Δx as functions of velocity. The phase-shift operator moves the wave energy depending on its apparent dip $\beta = dt/dx$, therefore it is more convenient to express these shifts as functions of β instead of the temporal and spatial coordinates. In the appendix I derive the time shift Δt:

$$\Delta t = \frac{\Delta z}{v(x_0, z_0) \left[1 - \beta^2 v(x_0, z_0)^2\right]^{\frac{1}{2}}}, \tag{34}$$

and the lateral shift Δx:

$$\Delta x = \pm \frac{\Delta z \beta v(x_0, z_0)}{\left[1 - \beta^2 v(x_0, z_0)^2\right]^{\frac{1}{2}}}. \tag{35}$$

Differentiating equation (34) with respect to the velocity we find

$$\frac{\partial \Delta t}{\partial v} = \frac{\Delta z \left[2\beta^2 v(x_0, z_0)^2 - 1\right]}{v(x_0, z_0)^2 \left[1 - \beta^2 v(x_0, z_0)^2\right]^{\frac{3}{2}}}, \tag{36}$$

and differentiating equation (35) we find

$$\frac{\partial \Delta x}{\partial v} = \pm \frac{\Delta z \beta}{\left[1 - \beta^2 v(x_0, z_0)^2\right]^{\frac{3}{2}}}. \tag{37}$$

The constraints that X must satisfies to avoid aliasing of the impulse response, following the general rule of equation (22), are

$$X \left| \frac{\partial \Delta t}{\partial v} \frac{\partial v(x,z)}{\partial x} \right| \leq \frac{\tau}{2}, \tag{38}$$

and

$$X \left| \frac{\partial \Delta x}{\partial v} \frac{\partial v(x,z)}{\partial x} \right| \leq \frac{\chi}{2}. \tag{39}$$

The undersampling distance X depends on the lateral gradient of the velocity model $\partial v(x,z)/\partial x$, besides depending on the dominant period τ and on the dominant wavelength χ.

The fundamental requirement of a migration algorithm is that it migrates properly the low dips of the data; the part of the impulse response correspondent to the low dips should be never

aliased. However from equation (36) we see that for zero dip, that is $\beta = 0$, the derivative $\partial \Delta t / \partial v$ is different from zero, therefore a strong gradient of the velocity model would cause the aliasing of the zero dip component of the impulse response. The low dips are downward continued by the part of the impulse response that is close to the apex of the hyperbola; they can be aliased because the time t_a of the apex of the hyperbolas $t_a = t_0 - \Delta z / v(x,z)$ depends on the migration velocity. A change of time variable to the retarded time coordinates $t^* = t - \Delta z / v(x,z)$ (Claerbout, 1985) solves the problem since in retarded time the apex of the hyperbola is always at the impulse time t_0 with any migration velocity $v(x_0, z_0)$, and thus the apex cannot be aliased by velocity variations.

The aliasing problem cannot be avoided for the high dips components of the data. In both equation (36) and (37) the denominator becomes zero for $\beta = 1/v(x_0, z_0)$; this value is the dip limit for the evanescent energy region. The rate of variations of the time shift and of the lateral shift tends to infinite as the dips tends to be evanescent, therefore the high dips component of the data are likely to be aliased when the velocity gradient is significant.

The problem is more serious in space than in time. The dip expressed in frequency domain is $\beta = k_x / \omega$; the high dip components of the data have a low angular frequency ω and a longer dominant period τ, therefore the constraint in equation (38) becomes less tight because τ is longer. On the contrary χ is shorter for the high dip components.

The portion of hyperbola in equation (33) is constrained to be positive in time, thus not all the dips are present in the impulse response. The range of dips in the impulse response depends also on the impulse time t_0. The later is the impulse the wider is the range of dips, and the worse is the aliasing problem.

Another parameter that depends on the impulse time t_0 and on the velocity $v(x_0, z_0)$ is the maximum lateral shift Δx_{max}; Δx_{max} is useful to determine the length M of the windows and can be derived by equation (33) with $t = 0$,

$$\Delta x_{max} = \sqrt{v(x,z)^2 t_0^2 - \Delta z^2}. \tag{40}$$

At later times Δx_{max} may be large and may require long windows to prevent the spatial aliasing of the impulse response. The higher dips are subject to the larger lateral shift and thus also the required length of the windows depends on the dip range of the data.

One way to reduce the windows length is to use first a constant-velocity migration and then a residual migration with varying velocity; the goal of the first migration is to decrease the migration velocity of the residual migration. Unfortunately with strong velocity variations the residual migration theory breaks down and cannot be used any more.

Gazdag and Sguazzero (1984) have proposed a phase-shift method, called Phase Shift Plus Interpolation (PSPI), for handling lateral velocity variation using interpolation in the frequency

domain between sections migrated with different velocity. Using short-time Fourier transform for phase-shift migration is a similar approach to PSPI since the synthesis procedure expressed in equation (31) is an interpolation in the frequency domain between the windows. I think that the short-time transforms method is a more general approach of the problem and allows a more precise choice of the sampling parameters. Moreover the use of windows shorter than the whole section allows saving in computational effort.

With the proposed algorithm the migration cost depends on the choice of the undersampling distance X and on the choice of the windows length M. A correct choice of these parameters depends on the velocity model, on the frequency bandwidth and on the dip range of the zero offset section we want to migrate. The choice of the parameters effects the quality of the result and one advantage of migration using short-time Fourier methods is that the cost can be chosen by the user in function of the quality of the result.

Another advantage of the proposed migration method is the use of the one-way wave equation to downward continue the data. The one-way wave equation has the advantage that does not propagate the internal reflections caused by velocity discontinuity; these reflections do not fit the primary only model of a stacked section and thus are annoying. The finite-difference schemes capable of migrating steep dip reflections are based on two-way wave equation and generate this unwanted reflections.

Synthetic results

I tested the phase-shift migration algorithm with some synthetic examples. I generated a zero-offset section from the depth model shown in Figure 14 assuming a laterally varying velocity model, linearly increasing from the left to the right. At the left end of the section the velocity was 960 m/sec and on the right end the velocity was 4800 m/sec; the velocity gradient is 2.5 sec^{-1}. The dipping reflector on the left is dipping 60 degrees and the one on the right 30 degrees.

Figure 15 shows the synthetic zero-offset section; the reflection from the flat reflector is dipping because of the velocity gradient. Figure 16 shows the result of migration with an undersampling distance X of 96 meters, equivalent to a distance between adjacent windows of eight samples. The windows were 128 samples long. The migration algorithm imaged almost perfectly the reflectors, except for a decrease in amplitude of the 60 degrees dipping reflector. Notice that the non uniformity of the amplitudes in the flat bed is caused by an artifact, a reflection from the left side boundary.

The migrated section in Figure 17 was obtained using the same processing parameters as in Figure 16 except for the undersampling distance that was 240 meters. In comparison with the previous figure we can notice that the dipping beds have a lower amplitude and are not completely focused.

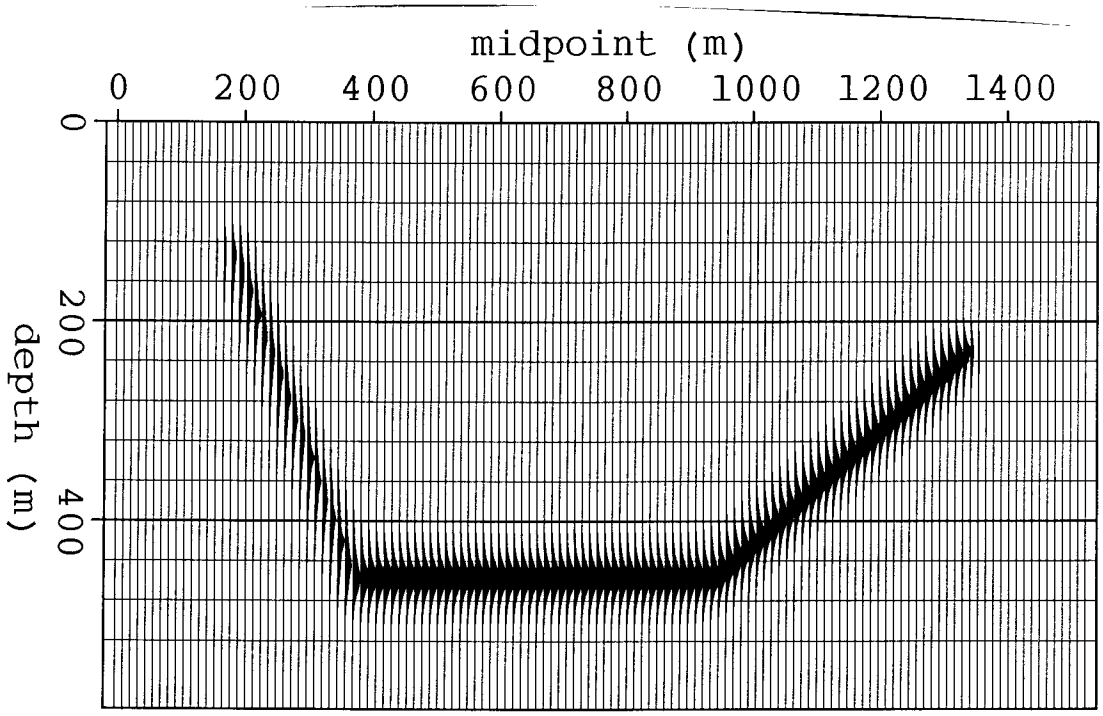

FIG. 14. Depth model for the synthetic test. The reflector on the left is dipping 60 degrees and the reflector on the right is dipping 30 degrees. The velocity was assumed to linearly increase with gradient 2.5 sec^{-1} from the left to the right. At the left end the velocity was equal to 960 m/sec and at the right end was equal to 4800 m/sec.

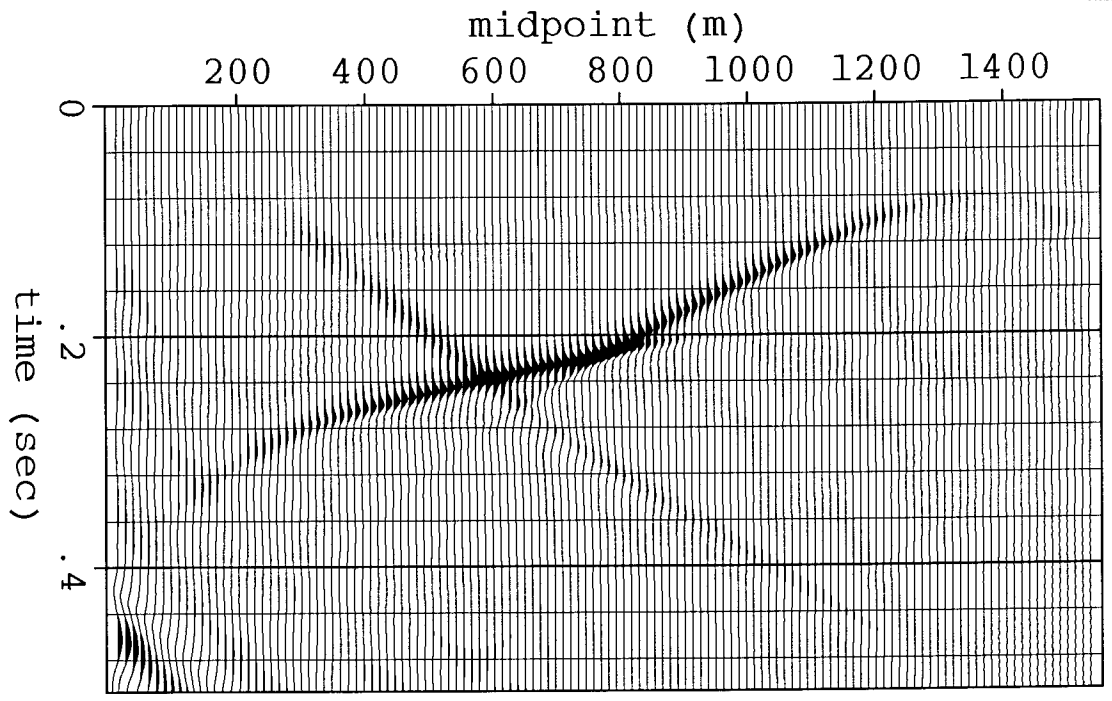

FIG. 15. Synthetic zero-offset sections modeled from the depth model in Figure 14

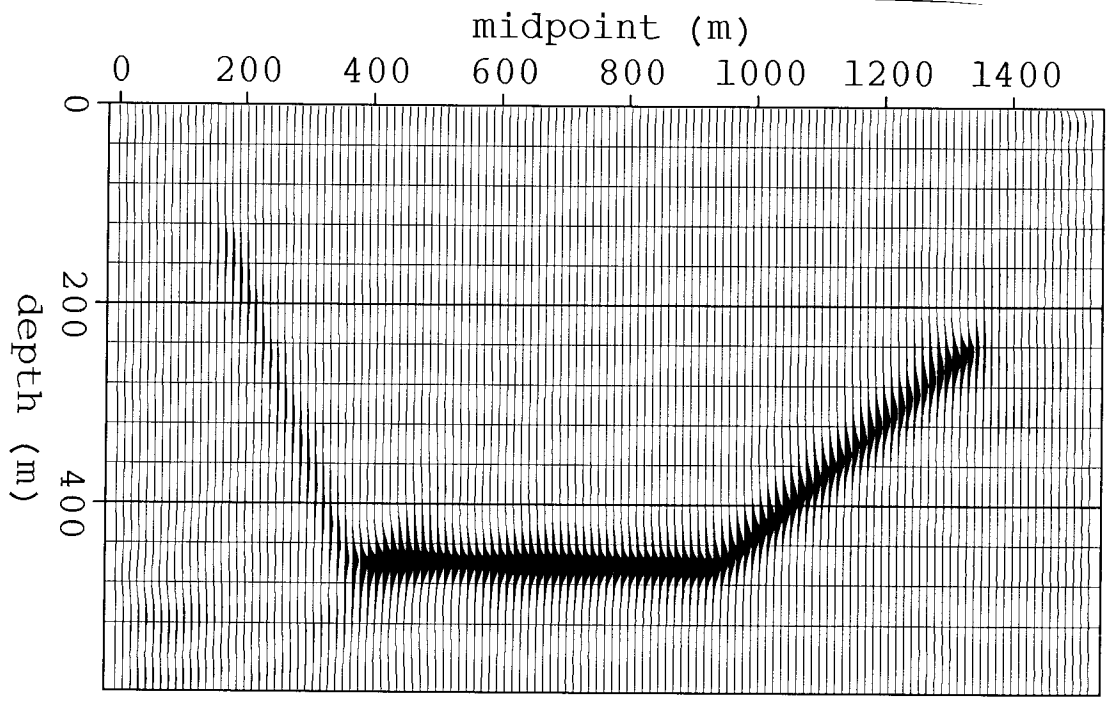

FIG. 16. Migration result using an undersampling distance of 96 meters. The reflector are well imaged except for an artifact superimposed to the flat reflector.

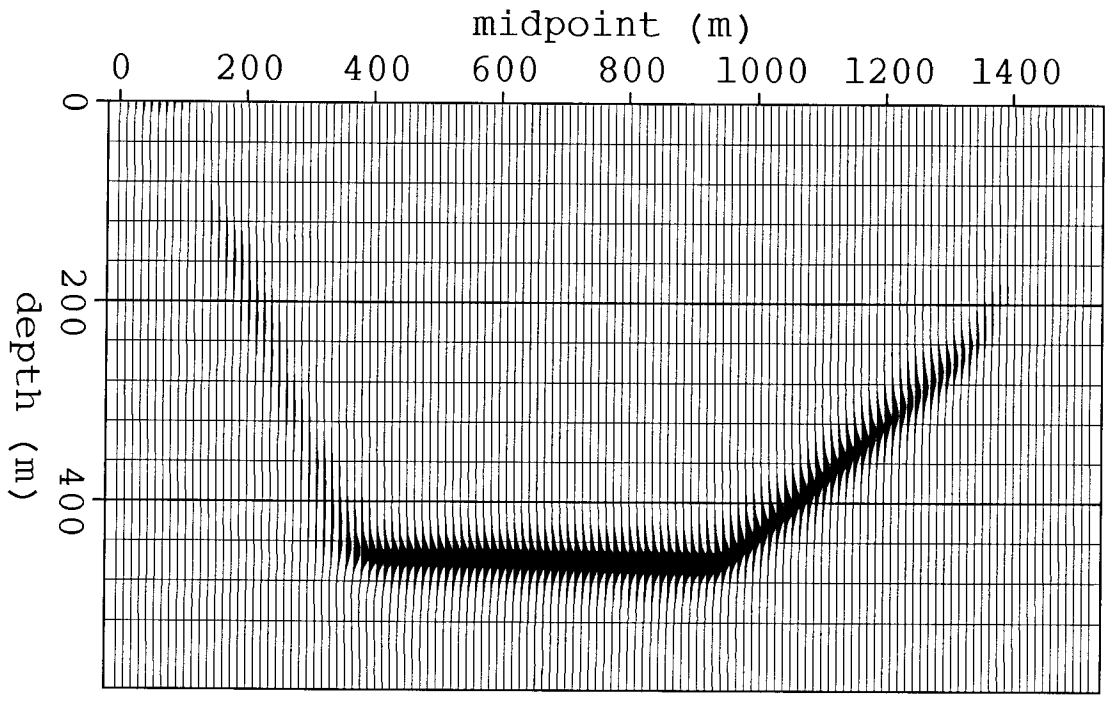

FIG. 17. Migration result using an undersampling distance of 240 meters. The result is worse compared to the more expensive processing but it is satisfactory.

Although the high velocity gradient the migration algorithm succeeded to image well the depth model in both cases. The quality of the result was superior using the more expensive algorithm, but also with a larger undersampling distance the final result is satisfactory.

CONCLUSIONS

The short-time Fourier transform is a versatile method to apply variant linear operators. In geophysics this technique can be advantageously used to generalize wave-equation operators when the velocity of wave propagations is space variant.

I analyzed in particular a phase-shift migration algorithm that can handle lateral velocity variation and shown, with some examples, the behavior of the algorithm in presence of a strong velocity gradient. The cost of the migration algorithms based on the short-time Fourier method is higher than the cost of conventional algorithms, but because the computational effort depends on the accuracy desired in the final result it is possible to vary the cost of processing according to the particular needs of the dataset.

ACKNOWLEDGMENTS

I would like to thank Kamal Al-Yahya and John Etgen for useful comments and discussions.

REFERENCES

Claerbout, J.F., 1985, Imaging the Earth Interior: Blackwell Scientific Publication, p. 85-87.

Crochiere, R., and Rabiner, L., 1983, Multirate digital signal processing : Prentice Hall.

Gazdag, J., and Sguazzero, P., Migration of seismic data by phase shift plus interpolation: Geophysics, **49** 124-131.

Portnoff, M. R., 1980, Time-frequency representation of digital signal and systems based on short-time Fourier analysis: IEEE Trans. Acoust., Speech, and Signal Processing, **28**, 55-69.

Rabiner, L. R., and Allen J. B., 1980, On the implementation of a short-time spectral analysis method for system identification: IEEE Trans. Acoust., Speech, and Signal Processing, **28**, 69-80.

APPENDIX

The goal of this appendix is to derive the time shift Δt and the lateral shift Δx caused by the phase-shift operator, as functions of the dip β and of the velocity v. The impulse response of the phase shift operator is given by the hyperbola expressed in equation (33). For the purpose of simplifying the notation I will consider the impulse time t_0 and the impulse location x_0 in the origin. I also change $v(x_0, z_0)$ to v. The hyperbola is now

$$v^2 t^2 - x^2 = \Delta z \quad -t_0 \leq t \leq 0. \tag{A1}$$

The time shift Δt and the lateral shift Δx are the amount of delay and of lateral displacement that the operator apply to the wave field. They are equal to the delay and to the distance of each of the points of the impulse response from the impulse; in this case $\Delta t = t$ and $\Delta x = x$.

Solving the equation (A1) for t we find

$$t = \frac{\sqrt{x^2 + \Delta z^2}}{v^2}, \tag{A2}$$

and solving it for x

$$x = \sqrt{v^2 t^2 - \Delta z^2}. \tag{A3}$$

The dip β is equal to the derivative dt/dx and it is found differentiating (A2) and back substituting t, that is

$$\beta = \frac{dt}{dx} = \frac{x}{v^2 t}. \tag{A4}$$

Substituting equation (A2) in equation (A4) and resolving for $\Delta t = t$ we find equation (34):

$$\Delta t = \frac{\Delta z}{v(1 - \beta^2 v^2)^{\frac{1}{2}}}, \tag{A5}$$

and substituting equation (A3) in equation (A4) and resolving for $\Delta x = x$ we find equation (35):

$$\Delta x = \frac{\Delta z \beta v}{(1 - \beta^2 v^2)^{\frac{1}{2}}}. \tag{A6}$$

波 決 汽 沢 水 気 氏 毛 比 毎 母 死 列 整 雌 此
止 欠 機 橋 横 標 構 楽 楕 極 植 森 械 根 桜 核
相 柱 柳 松 枝 林 析 木 月 題 曇 暗 最 晴 量 暑
晶 時 春 星 昨 明 昔 易 早 日 放 於 方 新 斜 対
文 数 散 教 故 支 摘 接 振 挺 指 持 押 折 手
扇 戸 或 成 戈 慣 態 感 想 惑 恒 性 心 衡 衝 微
御 術 徒 従 後 彼 径 役 形 帰 弾 強 弦 弧 引 弓
式 座 庭 度 底 応 幾 千 幅 帆 布 巻 項 左 工 川
嵐 炭 山 層 居 屈 尺 尾 尺 常 当 光 小 導 寒 密 家
室 実 定 宙 字 字 学 存 子 始 如 女 太 大 名 多
外 夕 各 処 冬 士 増 境 塩 場 塔 基 型 均 地 在
去 土 圏 国 固 図 回 因 四 器 嗅 哲 品 味 吸 号
右 口 双 又 原 圧 反 点 上 準 幹 真 南 直 古 十
北 動 加 力 割 切 刀 冷 次 周 同 円 前 其 並 典
共 公 分 八 入 先 億 像 傾 偏 側 値 係 供 例 価
何 余 作 低 体 伸 位 全 合 代 仕 他 今 化 人 商
率 変 夜 京 交 主 市 六 元 二 事 七 乙 島 重 乗
東 垂 来 卵 束 系 年 未 弗 少 午 千 及 九 単 為
半 氷 永 表 果 曲 向 出 本 甲 内 中 悪 夏 歪 昼
画 函 亜 再 両 百 正 平 可 不 天 五 下 三 万 一