Seismic trace interpolation beyond spatial aliasing

Lin Zhang, Jon Claerbout and Jun Ji [1]

ABSTRACT

Trace interpolation is often needed for processing spatially-aliased seismic data. One existing algorithm for trace interpolation is the linear two-step method. The method first finds prediction filters from known traces, and then estimates the missing traces by minimizing the output energy of the prediction filters. This paper presents a new method for finding prediction filters from known, spatially-aliased data. This method has three major advantages over previous methods. First, it can correctly interpolate data whose dip structure varies with respect to frequencies. Second, it does not require the amplitudes of the wavelets along an event to be constant. And third, it correctly determines prediction filters even when the data are completely aliased. Examples with synthetic and field data confirm that this method is superior to the prior algorithms in handling seriously aliased data.

INTRODUCTION

Steep events on seismic sections are often aliased because of the insufficient spatial sampling of wave fields by receivers. Such aliased events may create many kinds of coherent noises in the output of seismic event separation and seismic imaging processes. The solution to this problem is to interpolate seismic traces before these processes are applied.

A linear method of seismic trace interpolation usually consists of two steps, finding prediction filters from the known traces and then using the prediction filters to generate the missing traces. Two such methods have been developed. Claerbout (1991a) formulated the problem in the t-x domain. With his method, two-dimensional prediction filters are computed from known traces that are subsampled in time. Squeezing these filters both in time and in space gives the filters for generating the missing traces.

[1]**email:** lin@sep.stanford.edu, jon@sep.stanford.edu and jun@sep.stanford.edu

Spitz (1991) noticed that seismic data are sufficiently sampled in time, and simplified the 2-D interpolation problem into a 1-D interpolation in the ω-x domain. For each frequency, the filter that predicts data along the x-axis is computed. Stretching the coefficients of the prediction filters along the frequency axis then gives the filters that are used to generate the missing traces in the frequency domain. Spitz applied his algorithm to field data. His results showed that the linear two-step approach works well and that the trace interpolation improves the resolution of migrated images and reduces coherent noises.

However, these two methods have a common weakness. Both methods use the low-frequency components of data to estimate the prediction filters for the high-frequency components of data, which assumes that the dip structure of the seismic events are the same for all frequency components of data. This approach may fail to find the correct filters when data are completely aliased or when the dip structures of the high-frequency components of the seismic events are different from those of the low-frequency components.

Claerbout (1991b) observed that if a seismic section only contains linear events, the aliased components are distinguishable from the unaliased components in the ω-k_x domain. Using this observation, he proposed a nonlinear optimization scheme for interpolating missing data. We have applied Claerbout's principle to the process of finding the correct prediction filter and developed a new algorithm for interpolating seismic traces. With our method, the prediction filter at one frequency is obtained from the component of data at that frequency; no squeezing or stretching of the filter is required. The effects of aliased data on the prediction filters are removed by selecting the zeros of the prediction filters with a neural net. Consequently, this method does not assume a correlation among the dip structures of data within the different frequency bands. It can interpolate the missing traces even when all the data contained in the known traces are completely aliased.

This paper is organized as follows: In the first section, we define linear events and summarize the linear two-step algorithm for seismic trace interpolation. This section also introduces the notation that is required in the subsequent sections. The second section describes our method of finding the prediction filters from the aliased data and clarifies the assumptions that the method involves. The third section first compares the results of Spitz's algorithm and ours with synthetic data. It then shows how the interpolation process improves the results of removing tube waves in cross-well data.

THE LINEAR TWO-STEP METHOD

This section defines linear events and summarizes the linear two-step method in the context of Spitz's algorithm of seismic trace interpolation in the ω-x domain.

Linear events

A seismic event is called a linear event if it follows a linear trajectory. Usually such an event is mathematically described as:

$$w(t - px),$$

where w is the wavelet of the event, and p is the dip of the event. This description implies that the amplitudes of the wavelets are constant along the event. However field data show that wavelets along seismic events have slowly varying amplitude. Therefore, without loss of generality, we can describe a linear event by allowing its amplitude to vary exponentially as follows:

$$e^{\sigma x} w(t - px),$$

where σ is limited to being a real number that determines the rate of amplitude variation. An event with a complicated amplitude function is represented as a superposition of multiple linear events with the same dip. Since these amplitude variations are usually smooth, we expect that the magnitude of σ is small, and that the amplitude functions can be represented by a few exponential functions.

The predictability of linear events

The events on a real seismic section are generally not linear events, but their trajectories can be approximated by piecewise lines. Therefore, if we divide the section into many small subsections, we expect to see only linear events in each of these subsections. Let $\{g_n(t); \quad n = 1, 2, \ldots, N\}$ denote N traces that form a subsection, and suppose that this subsection contains L linear events. This subsection can be modeled in the time domain as

$$g_n(t) = \sum_{l=1}^{L} e^{(n-1)\sigma_l \Delta x} w(t - (n-1)p_l \Delta x), \tag{1}$$

and in the frequency domain as

$$G_n(\omega) = \sum_{l=1}^{L} W_l(\omega) e^{(n-1)(\sigma_l + ip_l\omega)\Delta x}, \tag{2}$$

where $W_l(\omega)$ is the Fourier transform of the wavelet of the lth event and Δx is the spatial sampling interval. With a proof similar to Canales' (Canales, 1984), we can show that for each frequency ω, $G_n(\omega)$ can be predicted by a one-step linear prediction filter, as follows:

$$P(\omega, s) = 1 - \sum_{l=1}^{L} P_l(\omega) e^{-sl\Delta x}, \tag{3}$$

where $s = \sigma_x + ik_x$ is the generalized complex wave-number. The coefficients of the filter can be computed from the following linear system:

$$
\begin{aligned}
G_n(\omega) &= \sum_{l=1}^{L} P_l(\omega) G_{n-l}(\omega) \qquad && n = L+1, \ldots, N \\
G_n(\omega) &= \sum_{l=1}^{L} P_l^*(\omega) G_{n+l}(\omega) \qquad && n = 1, \ldots, N-L,
\end{aligned}
\tag{4}
$$

where $*$ stands for the complex conjugate. If we interpolate $(M-1)$ traces between each pair of known traces, the output section should be

$$
\hat{G}_k(\omega) = \sum_{l=1}^{L} W_l(\omega) e^{(n-1)(\sigma_l + i p_l \omega)\Delta\hat{x}}, \tag{5}
$$

where $k = 1, 2, \ldots, MN$, and $\Delta\hat{x} = \Delta x / M$ is the new spatial sampling interval after trace interpolation.

Like $G_n(\omega)$, $\hat{G}_k(\omega)$ can be predicted by a prediction filter, as follows:

$$
\hat{P}(\omega, s) = 1 - \sum_{l=1}^{L} \hat{P}_l(\omega) e^{-sl\Delta\hat{x}}, \tag{6}
$$

whose coefficients similarly satisfy the linear system as follows:

$$
\begin{aligned}
\hat{G}_k(\omega) &= \sum_{l=1}^{L} \hat{P}_l(\omega) \hat{G}_{k-l}(\omega) \qquad && k = L+1, \ldots, MN \\
\hat{G}_k(\omega) &= \sum_{l=1}^{L} \hat{P}_l^*(\omega) \hat{G}_{k+l}(\omega) \qquad && k = 1, \ldots, MN-L.
\end{aligned}
\tag{7}
$$

Since the interpolation process should not change the known traces, we have

$$
\hat{G}_{(n-1)M+1}(\omega) = G_n(\omega) \qquad n = 1, \ldots, N. \tag{8}
$$

For each ω, equations (7) and (8) define a constrained linear system whose solution determines the missing traces.

Spitz's ω stretching method

To find the missing traces, we need to invert the constrained linear system defined in equations (7) and (8), which requires that we have the prediction filter $\hat{P}(\omega, \hat{k}_x)$. Spitz showed that the coefficients of the prediction filter $\hat{P}(\omega, \hat{k}_x)$ are related to the coefficients of prediction filter $P(\omega, k_x)$ as follows:

$$
\hat{P}_l(\omega) = P_l(\frac{\omega}{M}), \tag{9}
$$

whereas $P_l(\omega)$ can be computed from the known traces with equation (4). Spitz also proved that the coefficients of the prediction filters are predictable along the frequency axis. Thus, if the known traces have a limited frequency bandwidth, the prediction filters at frequencies outside the signal band can be computed from the prediction filters at frequencies inside the signal band.

Spitz's derivations of equation (9) and his proof of the predictability of the prediction filters assume that the amplitude functions of events are constant functions and that the dip structure of data is invariant with respect to frequencies. The first assumption is approximately valid if the window sizes of subsections are small enough. The second one is probably valid if all the events on the subsection have the same frequency content. However, it may be invalid if the seismic waves these events represent are initiated from independent sources or have multi-propagation modes. In these cases, the dip structure within one frequency band may be different from that within another frequency band. Hence, Spitz's ω stretching method does not give correct prediction filters in these cases. Furthermore, Spitz's algorithm does not work when events are completely aliased.

DEALIASING THE PREDICTION FILTERS

Our analysis of the problems with Spitz's method suggests that the prediction filter at a given frequency should be computed from the data component of that frequency. However, prediction filters computed from high-frequency components of data may suffer from contamination by spatially aliased energy. This section first shows that spatially aliased energy creates ambiguities regarding the zeros of the z-transform of the prediction filters and describes how to use a neural net to remove these ambiguities.

The zeros of prediction filters

Because prediction filters are finite-impulse-response filters, they can be characterized by the zeros of their z-transform. From equation (2), we know that the zeros of the prediction filter $\hat{P}(\omega, s)$ are $\{e^{(\sigma_l + ip_l\omega)\Delta x}; \quad l = 1, \ldots, L\}$. Therefore, we can express this filter as follows:

$$\hat{P}(\omega, s) = \prod_{l=1}^{L} (1 - e^{(\sigma_l + ip_l\omega - s)\Delta \hat{x}}). \tag{10}$$

If we scan the amplitude spectrum of this filter over the s plane, we can find L notches at

$$s = \sigma_l + ip_l\omega, \qquad\qquad l = 1, \ldots, L \tag{11}$$

that locate all the zeros of the z-transform of this filter. Similarly, we can express the prediction filter $P(\omega, s)$ as follows:

$$
\begin{aligned}
P(\omega, s) &= \prod_{l=1}^{L}(1 - e^{(\sigma_l + ip_l\omega - s)\Delta x}) \\
&= \prod_{l=1}^{L}(1 - e^{(\sigma_l + ip_l\omega - s)M\Delta \hat{x}}) \\
&= \prod_{l=1}^{L}\prod_{m=1}^{M}(1 - e^{(\sigma_l + ip_l\omega - s)\Delta \hat{x} - i\phi_m}),
\end{aligned}
\tag{12}
$$

where ϕ_m denotes the phases of the Mth order complex roots of the unity. Now, if we scan the amplitude spectrum of $P(\omega, s)$ over the s plane, we can find $M \times L$ notches at

$$
s = \sigma_l + ip_l\omega + i\frac{\phi_m}{\Delta \hat{x}}, \qquad l = 1, \ldots, L \quad \text{and} \quad m = 1, \ldots, M, \tag{13}
$$

M times as many notches as that of $\hat{P}(\omega, s)$. Comparing equation (13) with equation (11), it is apparent that these two equations become identical when ϕ_m is equal to zero. Thus, L out of $M \times L$ zeros of $P(\omega, s)$ are the zeros of $\hat{P}(\omega, s)$. Our goal is to identify these L zeros when $P(\omega, s)$ is known. If the component of data at frequency ω is not spatially aliased, then $P(\omega, s)$ has L zeros between two vertical lines $s = -i\pi/\Delta x$ and $s = +\pi/\Delta x$, which are L zeros of $\hat{P}(\omega, s)$. However, if the component of data at frequency ω is spatially aliased, the task of identifying the zeros becomes complicated and requires sophisticated algorithms.

Dealiasing prediction filters with a neural net

Dealiasing the prediction filter $P(\omega, s)$ is a process of identifying, for each frequency, L of its zeros that are also zeros of the prediction filter $\hat{P}(\omega, s)$. Let us call these L zeros of $P(\omega, s)$ "genuine" zeros and the rest $(M-1)L$ zeros "fake" zeros. Comparing equations (13) with (11) shows that, for genuine zeros, ϕ_m in equation (13) vanishes; while for fake zeros, ϕ_m does not vanish. Hence, if we project all zeros of $P(\omega, s)$ onto the ω-k_x plane, as shown in Figure 1, the location of each genuine zero as the function of ω follows a linear trajectory which, or the extension of which, passes through the origin. The location of each fake zero also follows a linear trajectory, but neither this trajectory nor its extension passes the origin. Equation (13) also shows we know that each genuine zero has $(M-1)$ fake zeros associated with it. We can predict where the associated fake zeros are if we know the position of the genuine zero. From above observations, it is apparent that dealiasing the prediction filter is equivalent to a pattern recognition process that retains the genuine zeros and rejects fake zeros. Our method uses a neural net to accomplish this goal.

To design a neural net for the problem, we must first find all zeros of the prediction filter $P(\omega, s)$ by scanning the spectrum of the filter over the s plane. Then we project these zeros onto the discretized ω-k_x plane. We define the input zero-indicator array as follows:

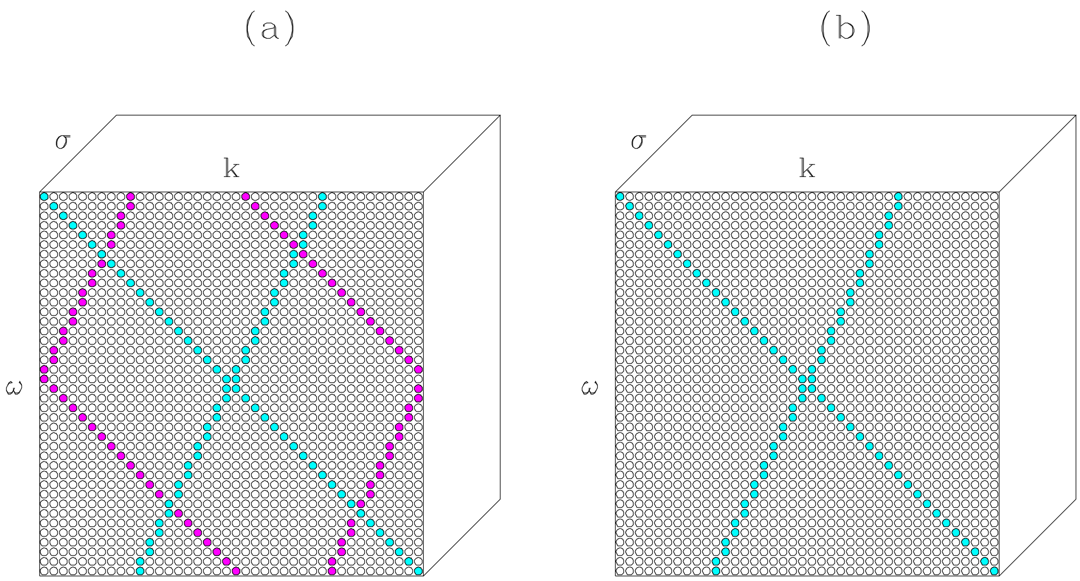

Figure 1: The principle of dealiasing the prediction filter. Each black dot indicates the projection of a zero of the prediction filter onto the ω-k plane: (a) before dealiasing; (b) after dealiasing. zind [R]

$$v_{ij} = \begin{cases} 1 & \text{if a zero at } (\omega_i, k_j) \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

and, similarly, the output zero-indicator array is defined as follows:

$$u_{ij} = \begin{cases} 1 & \text{if a “genuine” zero at } (\omega_i, k_j) \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

Now we can retain a genuine zero by setting its indicator to 1 and reject a fake zero by setting its indicator to 0. For each $v_{ij} = 1$, we define two regions, the exhibition region C_{ij} and the inhibition region D_{ij}. The exhibition region contains the indicators that are on the linear trajectory passing through both position (i, j) and the origin, and that are in the neighborhood of position (i, j). The inhibition region contains the indicators that indicate the associated fake zeros. Next, we compute

$$U_{ij} = \frac{1}{|C_{ij}|} \sum_{kl \in C_{ij}} v_{kl}, \tag{16}$$

where $|C_{ij}|$ is the number of indicators in C_{ij}. Then we use the winner-take-all (WTA) algorithm to determine the value of u_{ij}, as follows:

$$u_{ij} = \begin{cases} 1 & \text{if } U_{ij} > U_{kl} \text{ for } (k, l) \in D_{ij} \text{ and } v_{ij} = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

This algorithm is applied iteratively, as shown in Figure 2. Figure 3 displays an example of dealiasing the zeros of the prediction filters computed from field data. In spite of a few mistakes that can be identified by the human eye, the result is satisfactory. Most of the genuine zeros are retained and the fake zeros removed. Noted that the convergence of the iterative WTA algorithm is not guaranteed. However, the algorithm will never diverge because of the discrete and finite nature of the problem. Generally speaking, after several iterations, the output of the algorithm will oscillate among a few solutions that are all close to the globally optimal solution.

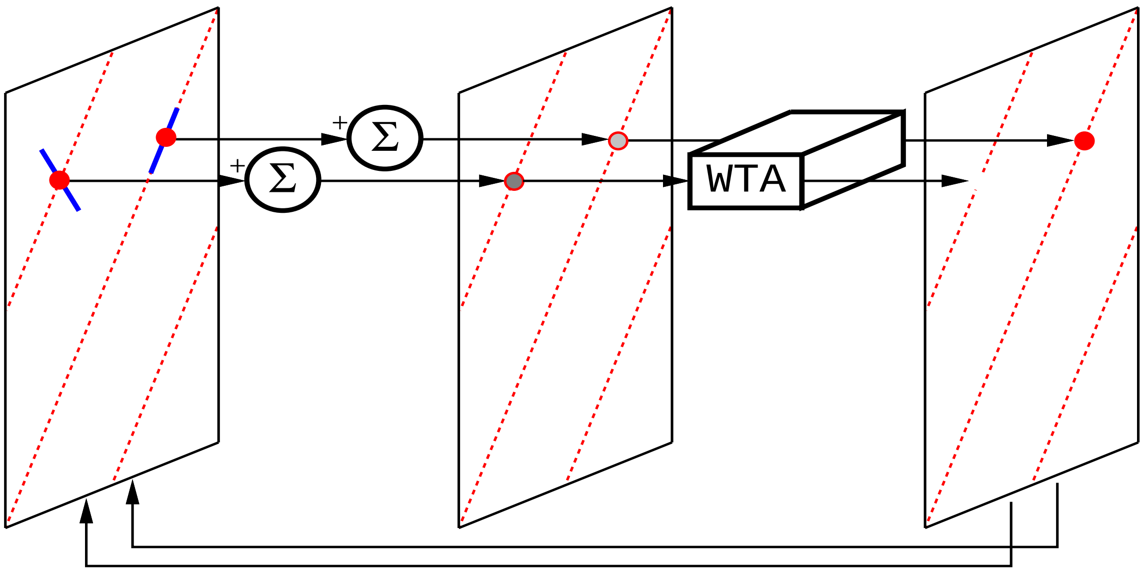

Figure 2: An illustration of the iterative winner-take-all (WTA) algorithm. neunet2 [NR]

Summary of the algorithm

Our algorithm for trace interpolation can be summarized as follows:

for each subwindow
- $G_n(\omega) = \mathrm{FFT}[g_n(t)]$
- for each frequency
  - compute prediction filter $P_l(\omega)$
  - search zeros s_l of $P_l(\omega)$
- select $\hat{s}_l$ with a neural net
- for each frequency
  - construct $\hat{P}_l(\omega)$ from $\hat{s}_l$
  - compute $\hat{G}_n(\omega)$ with least squares
- $\hat{g}_n(t) = \mathrm{FFT}^{-1}[\hat{G}_n(\omega)]$
- merge subwindows

Because of the additional computation involved, this algorithm is about twice as slow as Spitz's algorithm.

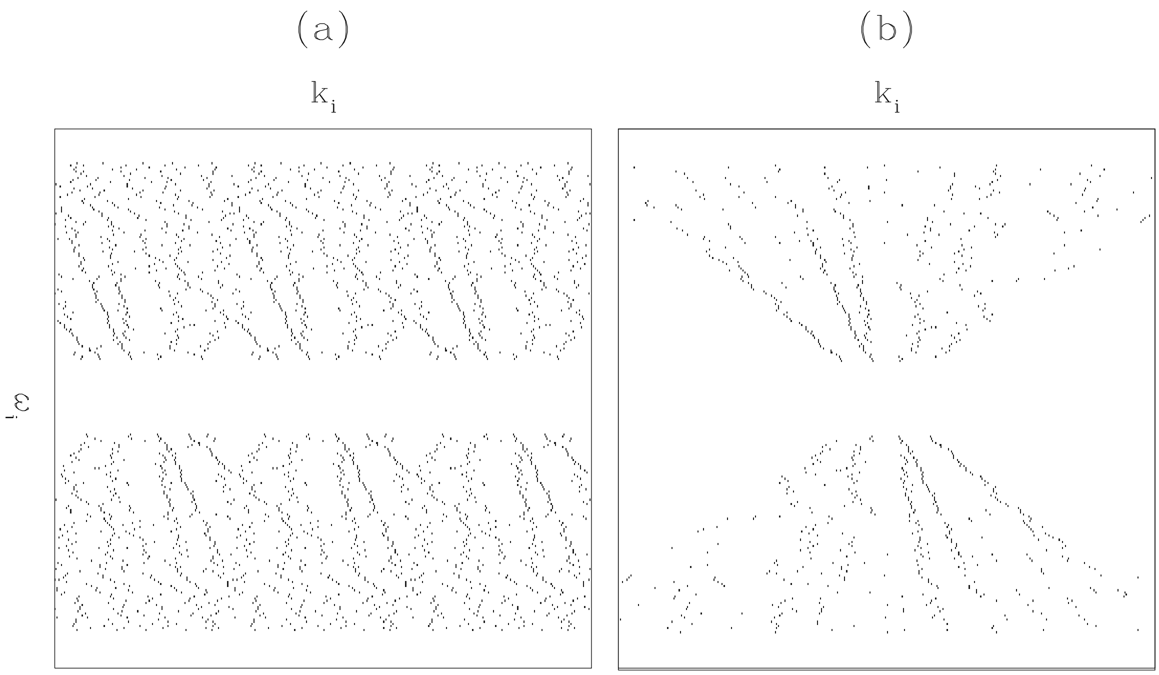

Figure 3: An example of dealiasing the prediction filter with a neural net: (a) input zero indicators, (b) output zero indicators. cwlneu [R]

NUMERICAL EXAMPLES

In this section, we first give an example with synthetic data to illustrate a potential problem with Sptiz's algorithm and demonstrate the capability of our algorithm. Then we show the results of interpolating a field record and the effects of interpolation on removing tube waves in this field record.

Synthetic data

Figure 4a shows a synthetic section containing three linear events, each of which has a different wavelet. The steepest event has a high-pass wavelet and is almost completely aliased in space. The event with its amplitude decaying along its trajectory has an all-pass wavelet. Finally, the flattest event has a low-pass wavelet. Figure 5a shows the 2-D spectrum of this section. Clearly, the dip structure in this example varies with respect to frequencies. Figures 4b and 4c show the first-order interpolation with Spitz's algorithm and our algorithm, respectively. We can see that both algorithms correctly interpolate two gentle-dip events, even though one of the events has amplitude variation. Spitz's formulation does not explicitly consider the amplitude variations of events, but it is valid when the amplitude variations do not depend on frequency; frequency-dependent amplitude variations may cause some minor problems. Spitz's algorithm fails, however, to interpolate in the correct dip direction of the steepest event that is completely aliased in space because of the frequency-dependent dip structure of data. In contrast, our algorithm interpolates the event correctly. Figures 5b and 5c show the 2-D spectra of the interpolated sections, which confirm our observations in the time-space domain. Figure 5c also displays, at frequency 40 Hz and wavenumber 8 (1/km), a gap in one event, which is a noticeable pitfall of our algorithm. This occurs because our zero-searching routine assumes a single-zero presence at each location. Hence, the neural net may reject a genuine zero if it coincides with a fake zero. Because this problem occurs at only a few frequencies, we do not expect it to have strong impact on the later processing of the interpolated data.

This synthetic example represents an extreme case that may or may not exist in practice. However, it clearly reveals the important assumptions that Spitz's algorithm entails.

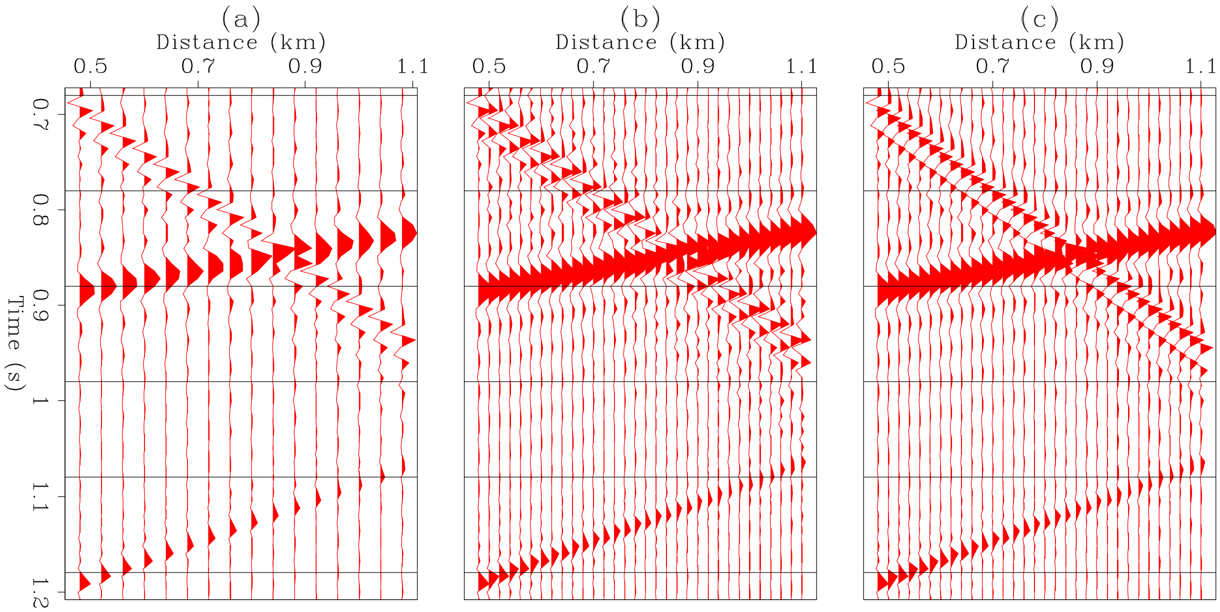

Figure 4: A synthetic example of trace interpolation: (a) before interpolation, (b) after interpolation with Spitz's algorithm, (c) after interpolation with our algorithm. syncom [R]

Field data—removing tube waves

Our field data examples are created from a shot gather recorded in a cross-well experiment. The original data have sampling intervals of 0.1 ms in time and 1 foot in depth. The source is located at a depth of 540 feet. The frequency band is from 360 Hz to 2160 Hz and the maximum dip is 0.3 ms/feet; hence neither time nor spatial aliasing exists. In order to reduce the cost of data collection, increasing the sampling interval in depth is recommended. To simulate data with a 2-foot sampling interval in depth, we decimated the original gather in space. The result is displayed in Figure 6. The steepest linear events representing tube waves that travel up and down in the receiver well become aliased in space. Because the tube waves interfere with useful events they should be removed by dip filtering during the preprocessing. However, existing dip-filtering techniques do not work for such spatially aliased data. Therefore, we need to do interpolation before dip filtering. Figure 7 displays the result of the

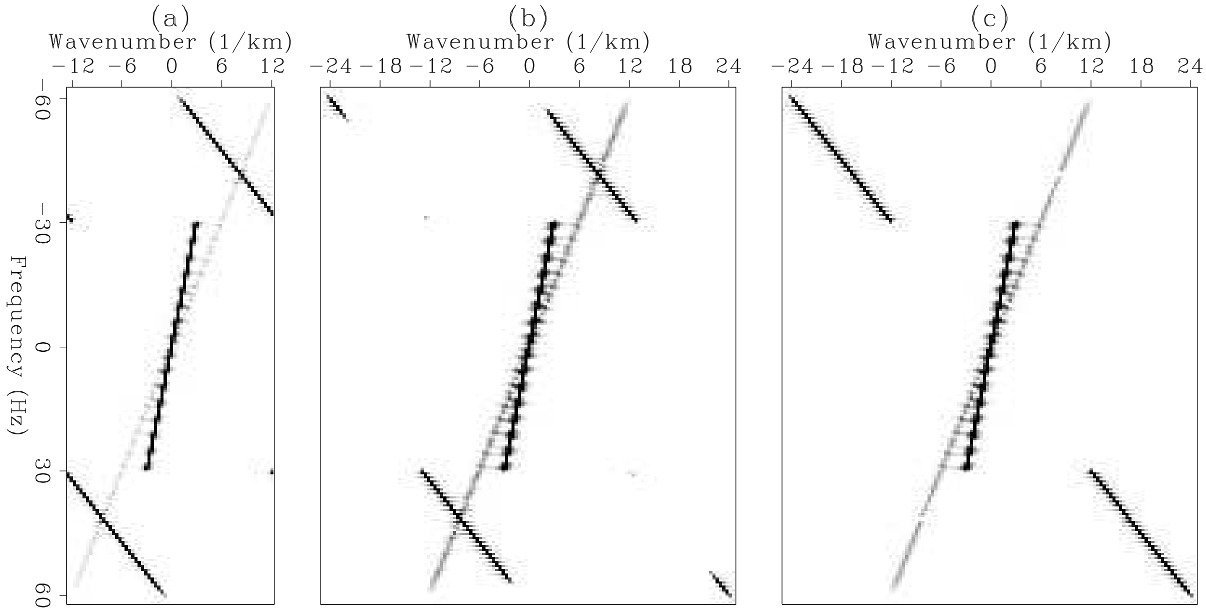

Figure 5: Comparisons of 2-D spectra: (a) before interpolation, (b) after interpolation with Spitz's algorithm, (c) after interpolation with our algorithm. synspccom [R]

first-order trace interpolation. We used overlapping subwindows of 128 samples in time and 32 samples in space. The prediction filter length is 6. A close look at two locations where strong tube-waves present—at arrival time 0.07 second and depth 970 feet, and at arrival time 1.1 second and depth 740 feet—shows that seriously aliased tube waves are correctly interpolated. To check the overall quality of interpolation, we subtract the interpolated gather from the original gather. Figure 8 shows that the interpolation errors are generally small and, more importantly, that the distribution of errors is quite random, that is, the spatial coherence of the interpolation errors is weak. Large errors appear only along the boundaries of the gather, with the exception of the trace recorded at the depth of 960 feet. After checking the original data, we found that this trace was misshifted in the original gather. Thus we consider the result satisfactory. Figure 9 displays the 2-D spectra of the gather before and after interpolation. Figure 10 compares the results of removing tube waves before and after interpolation. The depth sampling intervals are 3 feet before interpolation and 1 foot after interpolation. Dip filtering is actually performed in the slant-stack domain. Before interpolation, the steep events cannot be filtered out because the dips of the aliased event are ambiguous. After interpolation, the dips of all events are well defined. Consequently, tube waves are effectively removed. Many events that are originally blocked by tube waves are now visible.

CONCLUSIONS

We have described an algorithm for interpolating spatially-aliased seismic traces. This method can interpolate data even if the dip structure of data varies with respect to frequencies. Moreover, it does not require the amplitudes of the wavelets along

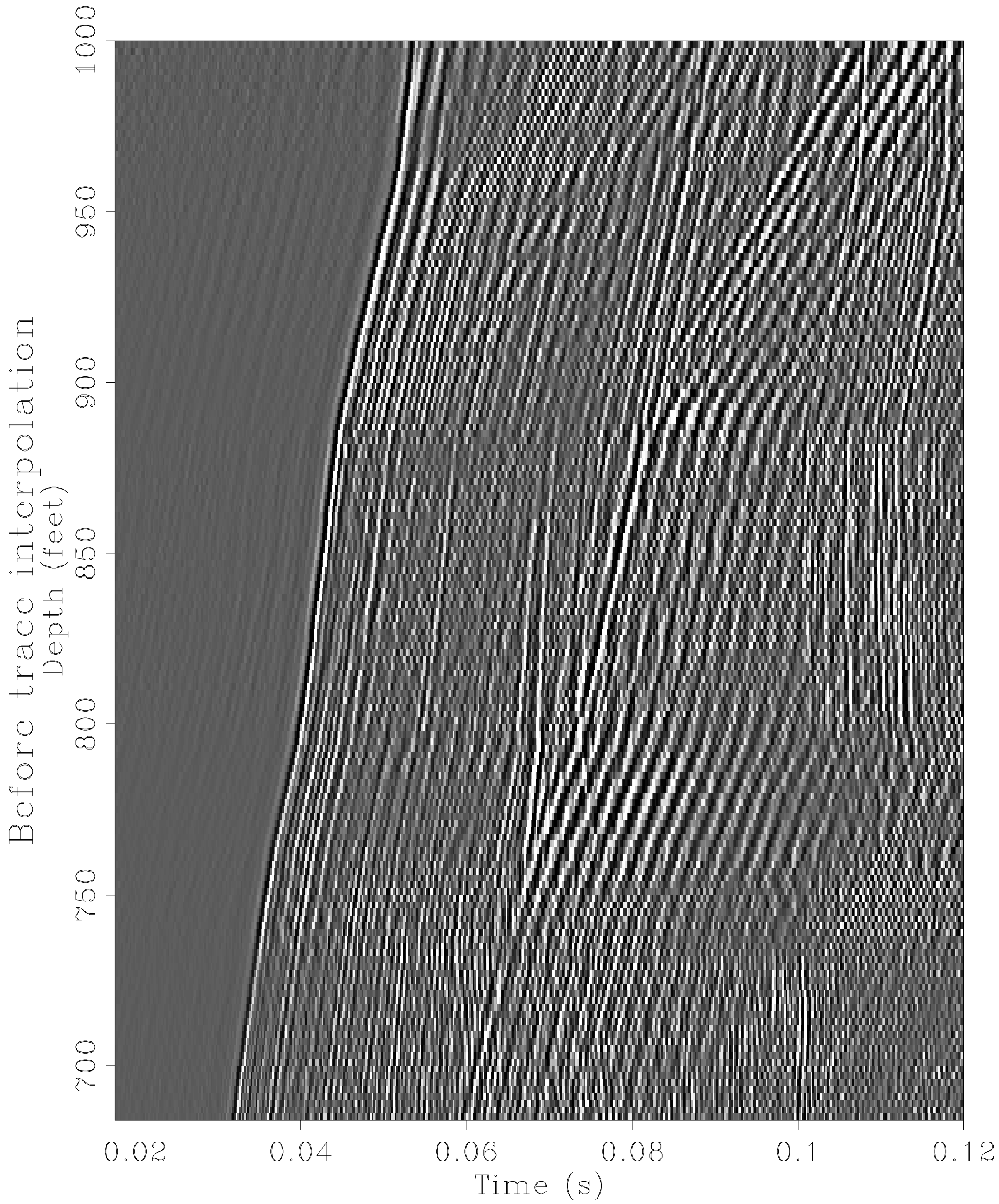

Figure 6: A decimated shot gather that was recorded in a cross-well experiment. The spatial sampling interval is 2 feet. The shot depth is 540 feet. cwlsub1 [R]

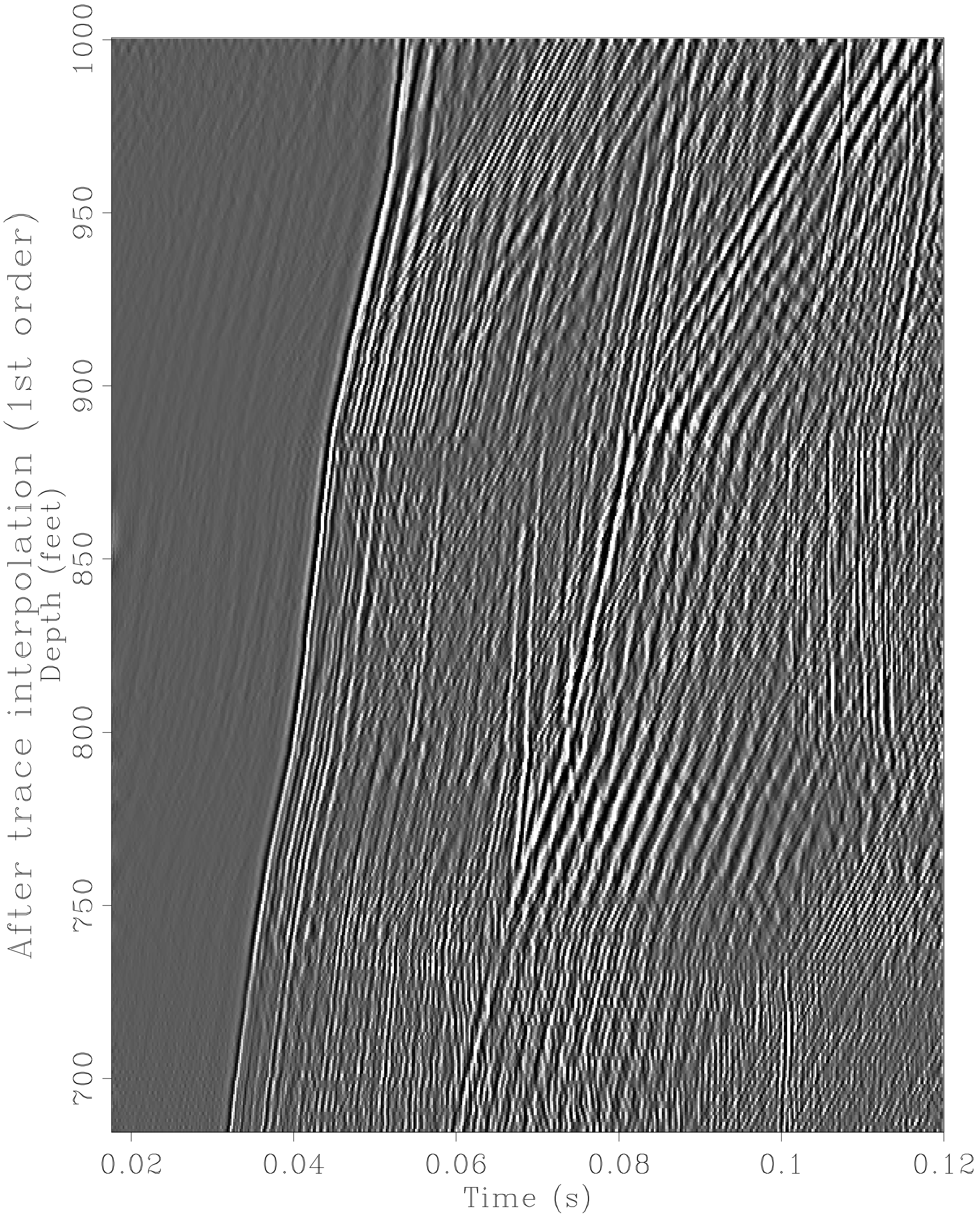

Figure 7: The shot gather after the first-order interpolation. The spatial sampling interval becomes 1 foot. cwlint1 [R]

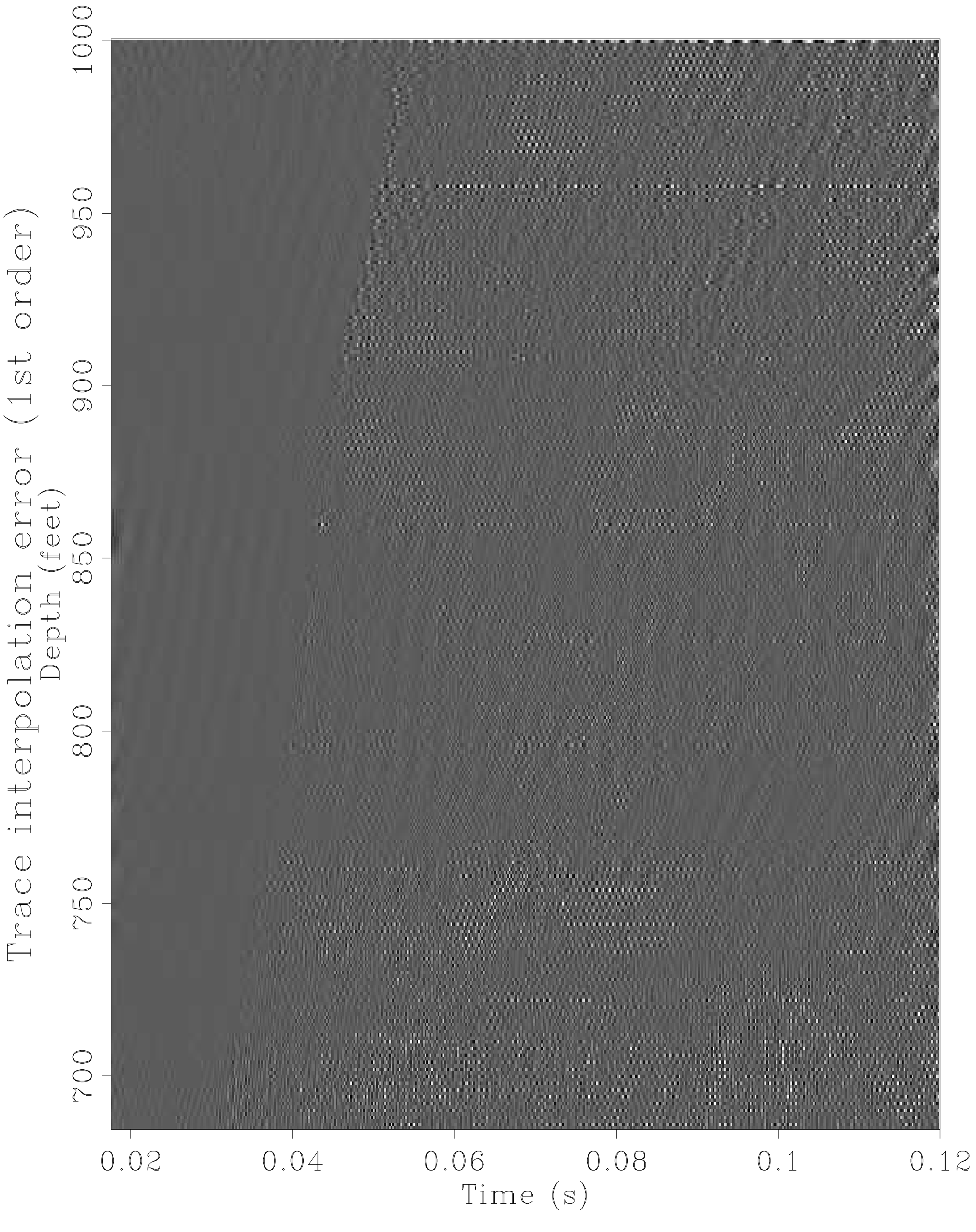

Figure 8: Interpolation errors. cwlerr1 [R]

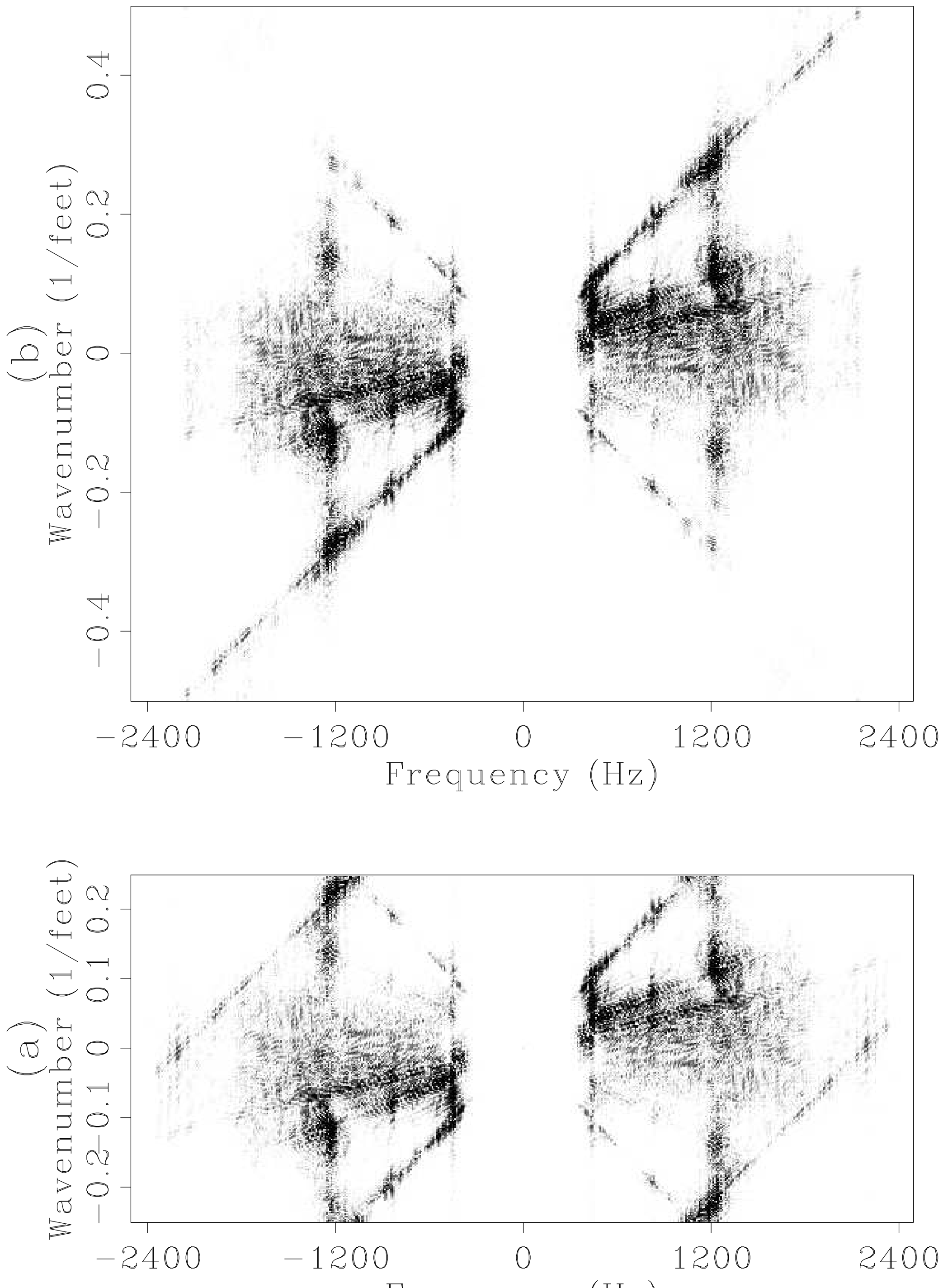

Figure 9: Comparison of 2-D spectra: (a) before interpolation, (b) after interpolation with Spitz's algorithm. cwlintspc1 [R]

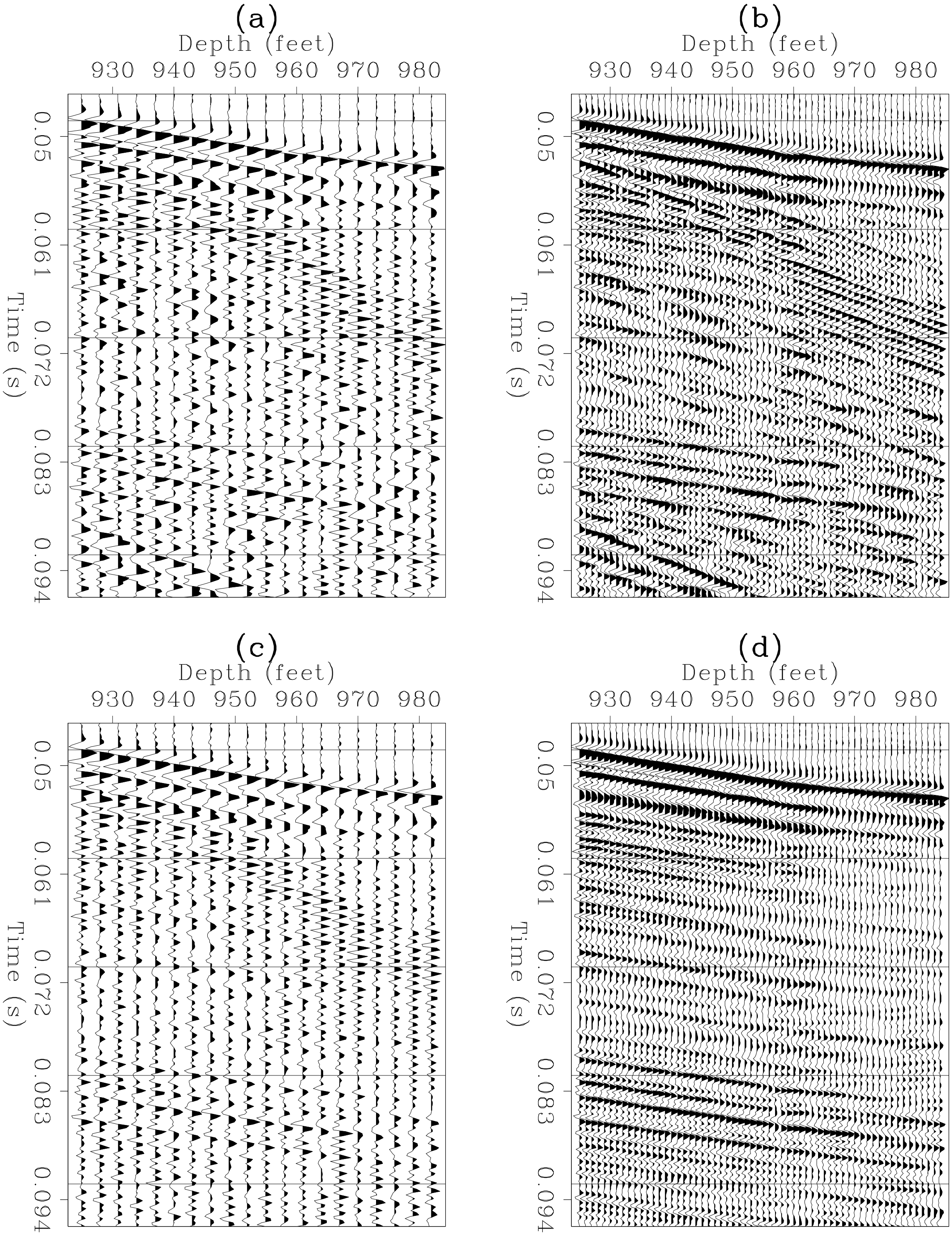

Figure 10: Results of removing tube waves: (a) data before interpolation, (b) data after interpolation, (c) output of dip filtering before interpolation, (d) output of dip filtering after interpolation. $\boxed{\text{cwlcom}}$ [NR]

an event to be constant. And finally, it finds correct prediction filters even when the data are completely aliased. Examples with synthetic and field data confirm that this method is superior to the existing algorithms in handling seriously aliased data.

Regarding its limitations, our algorithm does not, of course, work when data are random or when all events in the data have monochromatic wavelets. However, no algorithm works in these two situations. Therefore, it is to be expected that the results of our interpolation algorithm decrease in accuracy as the input data approach either of these conditions.

A less acceptable problem with our algorithm is its limited ability to handle a large number of dips. In our examples, we used prediction filters of less than 10 coefficients. Extraordinarily long filters may cause failures of the zero-searching routine and the neural net. These two parts of the algorithm need to be improved in the future.

ACKNOWLEDGEMENTS

We thank Stanford Tomography Project, especially Professor J.M., Harris for providing the cross-well data for our study.

REFERENCES

Canales, L.L., 1984, Random noise reduction: Presented at the 54th Ann. Internat. Mtg., Soc. Expl. Geophys.

Claerbout, J.F., 1991a, Earth Soundings Analysis: Processing versus Inversion: Blackwell Scientific Publications.

Claerbout, J.F., 1991b, Missing data restoration: SEP–**70**, 305–335.

Spitz, S., 1991, Seismic trace interpolation in the F-X domain: Geophysics, **56**, 785–794.