Time windowing passive seismic data in the frequency domain

Brad Artman, Stanford University

SUMMARY

Collecting incredible amounts of passive seismic data in a very short time is both the strenght and weakness of the technique. Processing the data to extract the subsurface structure can be a very costly endeavor. Whether cross-correlating traces to produce synthesized shot-gathers or migrating the data directly, subsampling the data in the Fourier-domain provides orders of magnitude time and computational savings without any loss of information. In fact, the process effectively prunes spurious correlations between different subsurface sources that we do not understand and are incapable of correctly processing.

INTRODUCTION

One of the principle obstacles to the utility of passive seismic data is its bulk. With several hundred to thousands of geophones, we are able to generate mountains of data in a very short time. The simplest method of trimming down this volume is to keep only the recorded wavefield around times when usable source energy is known to be present. In the case of teleseismic imaging, or when utilizing unconventional, but known sources, this is easily done. However, if one hopes to image with the truly ambient noise field, time windowing amounts to removing needed signal.

After cross-correlation of passive traces, one can produce shot-gathers equivalent to an active source experiment (Wapenaar et al., 2004). Having done so, we realize that most of the recording time axis can be discarded and maintain only enough time/lag samples to keep the reflection from the deepest subsurface structure of interest. As most of the processing of seismic data is accomplished in the Fourier-domain, I will consider the procedure and ramifications of time windowing after data have been transformed to frequency.

1D EXAMPLE

Windowing in the time domain is equivalent to subsampling in the dual domain. The first trace in Fig. 1 is a time signal that might correspond to several identical subsurface sources below a single layer in the earth. It is many thousands of samples long. The second trace is its autocorrelation calculated using all data in the frequency domain multiplication. We recognize the auto-correlation to be symmetric, and could have truncated the result after half of the samples. Also, because each event correlates with all the others, many correlation peaks appear at late time. These late events contain information about similarity of the earth's impulse response recovered by each independent source function. Mathematically, this is a chain of many convolutions of complicated functions and very difficult to understand or utilize. They do not fit into the framework of passive seismic imaging by representing the kinematics of a reflection gather. So I discard them. In this case, the "deepest subsurface reflector of interest" is the second peak of the output correlation. The correlations after this could be what Schuster et al. (2004) call *other terms*.

Instead of truncating the auto-correlation in the time-domain, the last trace was calculated by decimating the Fourier transform of the signal before multiplication. The trace was padded with zeros to facilitate plotting with the previous traces. The number of frequencies used to produce the second autocorrelation was 8 times fewer than the Fourier representation of the input. As long as the level of decimation maintains support for the time window desired for the final result, we can subsample the input after transforming to the frequency domain and before subsequent processing.

SYNTHETIC SEISMIC EXAMPLE

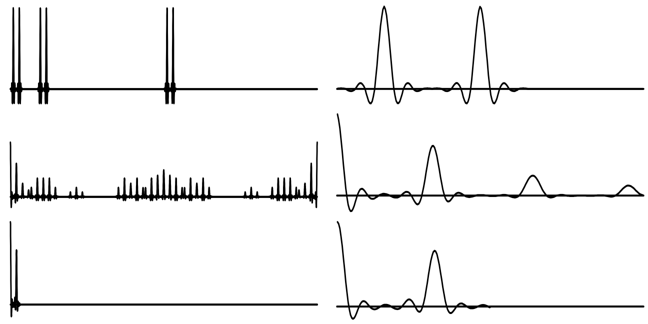

Figure 1: Idealized signal of three identical subsurface sources, its autocorrelation, and autocorrelation after subsampling the frequency axis by a factor of 4. Right panel is zoomed in view.

The seismic example was designed to mimic the passive seismic experiment and reflects the injection of a few simple but very important assumptions about the nature of the subsurface noise-field. These assumptions immediately move us away from the truly random ambient noise experiment. Strictly, I am assuming that the length of any particular subsurface source is fairly short in duration *i.e.* less than a few seconds) and randomly distributed throughout the recording time of the experiment. Cross-talk between sources and their reflections about the subsurface will be introduced under these assumptions if the subsurface sources are not completely uncorrelated. Some degree of correlation will arise if they are not separated in time by at least the two-way traveltime to the deepest reflector and if the sources have correlable waveforms.

Previously, I have manufactured synthetic passive data by convolving transmission wavefields from individual sources with different, long, random traces. The signal to noise ratio of the result can in this case be shown to improve with the square root of the number of time samples in the random source function and the square root of the number of transmission wavefields, that is number of sources, used. There are however important physical implications to modeling in this manner. Specifically, it implies many sources exploding simultaneously with infinite duration and perfect distribution of energy during this entire period. A more reasonable experiment instead is one where the sources have durations of less than a few seconds, and are randomly distributed during the course of the recording interval.

The synthetic examples presented here are generated with random source time functions no longer than 3 seconds and a random bulk time shift. This facilitates the exploitation of this subsampling strategy to its limit. Under these assumptions, increasing the signal-to-noise ratio of the image will require recording more subsurface sources or allowing the sources to ring for longer than 3 seconds. In either case, the signal to noise ratio will increase with the square root of the quantity considered.

Figure 2 explores the use of Fourier subsampling on the migration results of the synthetic passive data. Three data volumes were created and imaged without first cross-correlating the traces (Armtan et al., 2004). All three panels used 280 subsurface energy sources. The first panel was modeled with a total experimental duration of 8 seconds. The second and third panels assumed the recording time was 260 seconds. The first and second panel directly migrated all

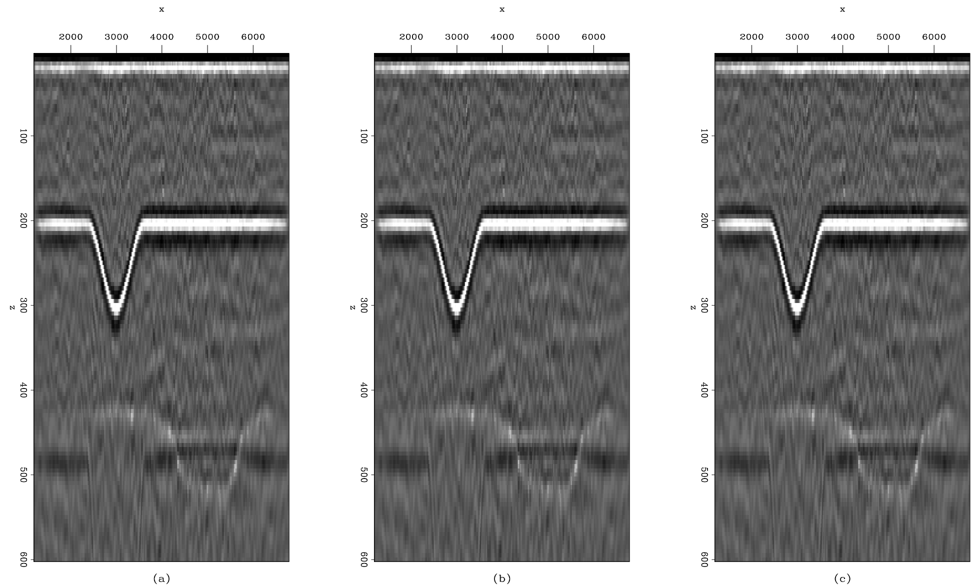

Figure 2: Identical imaging results from direct migration of modeled passive data from 280 subsurface sources (of random duration between 50 and 3000 ms) collected over 8 (panel (a)) and 260 (panels (b)) seconds and a version of the long data subsampled by a factor of 32 in the frequency domain (panel (c)).

the data from the entire recording duration. They are identical to within machine precision. The third panel subsampled the 260 second data set in the Fourier domain to a level that will only support 4 seconds of an inverse Fourier transform. This image is identical to the others.

COST IMPLICATIONS

Considering a data set with n_x receivers with n_t time samples, Table lists the number of operations required to construct shot-gathers by cross-correlating the traces of the transmission wavefield and migration of either the raw volume or the data volume produced by correlation. Under the conditions presented above, the n_{ts} time samples needed for the shot-gathers is the same as the number of frequencies required for direct migration. This will be orders of magnitude less than the n_t time samples collected during a passive seismic experiment in most cases. Finally, it is possible that not all receiver stations need be correlated when processing a large, high frequency experiment. ν will be a subsampling factor, less than 1, that controls the number of traces in, or aperture of, the correlated shot-gathers.

Migration costs scale according to the size of the input data set and the size of the image domain through which the data is extrapolated. X will represent a scalar multiplier due to the computer overhead costs of the migration strategy used. This will vary from a factor of 5 to n_x depending on the algorithm and accuracy required, but will be common to either direct migration or the migration of the correlated shot-gathers.

The size of the image space is assumed to be n_x samples areally by n_z samples in depth. No inverse Fourier transforms are required to prepare for migration, as the shot-gathers are needed to be functions of frequency for many migration algorithms. Also, I assume that the field passive seismic data fulfills the model of short source functions unevenly dispersed along the time axis of the duration of data collection.

With these costs in mind, the ratio of the sum of the first two rows to the last must be balanced to decide which choice requires the least amount of computation operations. When $\nu = 1/n_x$, the costs

	operations
Multiply	$\nu n_x^2 n_{ts}$
Gather Mig.	$(X n_z n_x) n_{ts} \nu n_x^2 \log n_x$
Raw Mig.	$(X n_z n_x) n_{ts} n_x \log n_x$

Table 1: Operation counts for migrating passive seismic data. X represents the a scalar multiplier of migration overhead. n_x is the number of receivers. n_t is the length of the time axis of the passive experiment. n_{ts} is the length of the time axis associated with the two-way traveltime to the deepest reflector of interest.

of producing an image by either method is the same. The meaning of this situation is shot-gathers of one trace, *i.e.* a constant-offset (post-stack sized) migration. Thus, the passive seismic experiment acts similarly to a natural phase-encoding of active seismic shot-gathers that (Romero et al., 2000) explains as a method to reduce the cost of shot-profile migration schemes. This can be thought of as performing a full prestack migration for the cost of a zero-offset migration.

Therefore, if it is appropriate to assume that potential subsurface seismic sources are reasonably short in duration, and that the length of the passive experiment is dictated by the requirement to collect a sufficient number of them to illuminate our image space, we can save substantial computation cost, orders of magnitude, by migrating the raw data directly. Even if one is concerned that such a severe decimation of the frequency axis might be detrimental, many safety multiples can be carried without affecting the speed-up of direct migration of the data.

CONCLUSION

I have used Fourier sampling theory to the apply a time window to correlated passive seismic data. Under a reasonable (I believe) set of assumptions, the initial Fourier transform of the raw data acts to redistribute the energy of late time arrivals toward the origin of the frequency axis, which effectively handles the zero-time problem of passive imaging. The multiplication (by conjugate traces) handles the waveform comparison problem.

In practice, we can avoid many calculations associated with the long time axis of transmission wavefields by exploiting the fact that the fine increments of the Fourier domain contain the information in the late times of the input data. Acknowledging that we are incapable of using the late time correlations within the framework of conventional seismic processing algorithms, this information can (and probably should) be removed after correlation. Conveniently, since the correlation is performed in the frequency domain with point-to-point multiplications, this extra information can be removed immediately upon its initial transform to the Fourier domain, and before any further processing is performed. This holds true for direct migration or manufacture of modeled shot-gathers from a transmission wavefield.

REFERENCES

Armtan, B., Draganov, D., Wapenaar, C., and Biondi, B., 2004, Direct migration of passive seismic data: 66th Conferance and Exhibition, EAGE, Extended abstracts, P075.

Romero, L. A., Ghiglia, D. C., Ober, C. C., and Morton, S. A., 2000, Phase encoding of shot records in prestack migration: Geophysics, **65**, no. 02, 426–436.

Schuster, G., Yu, J., Sheng, J., and Rickett, J., 2004, Interferometric/daylight seismic imaging: Geophysics Journal International, **157**, 838–852.

Wapenaar, K., Thorbecke, J., and Draganov, D., 2004, Relations between reflection and transmission responses of three-dimensional inhomogeneous media: Geophysical Journal International, **156**, 179–194.