Adaptive subtraction using 3D curvelets: A linear optimisation framework

Amarjeet Kumar, Rajiv Kumar, Gary Hampson, Mike Hartley and Troy Thompson*
DownUnder GeoSolutions, Perth, Australia

SUMMARY

A very common geophysical task is the optimum subtraction of a noise model (e.g., multiples) from input data that is assumed to be composed of both signal and noise. If the noise model is accurate it may be removed by direct subtraction. However, noise models are typically not perfect and so it is common to adaptively subtract noise in a carefully constrained manner. The parameterisation of the adaption strongly influences the quality of outcome. Unconstrained adaption usually damages the signal. However, with judicious constraints, noise subtraction can be successful without signal damage. In this paper, we propose a linear optimisation framework for complex curvelet domain adaptive subtraction that compensates for errors in the noise models in an optimum way.

INTRODUCTION

The attenuation of noise often involves two stages; estimation of the noise model and subsequent subtraction from the input data. Imperfections in the noise estimate can be compensated for during the subtraction stage by using an adaptive approach to subtraction. For example, predicted multiple models often include wavelets and directivity patterns that differ from those present in the input data (Ikelle et al., 1997). In the case of 2D-SRME, errors in the predicted multiples can result from the 3D complexity of the earth. In 3D-SRME, imperfections are commonly due to incomplete acquisition sampling (Dragoset and Jericevic, 1998). To compensate for these kinds of imperfections in the noise models, the subtraction step must be performed adaptively such that the adapted models collectively better resemble the noise in the input dataset. Adaptive subtraction is most often performed by finding a short linear filter in the space-time domain that modifies the noise model so that the energy in the subtracted result is minimised. If more than one noise model has been estimated, (e.g., a Radon multiple model and a surface related multiple model) they can be adaptively subtracted simultaneously so that the components of each model are selected according to their ability to minimise the energy in the resulting subtraction.

There is no reason to restrict the domain in which adaptive subtraction takes place. Indeed, there is a range of possibilities, each with associated advantages and disadvantages. One domain which is particularly attractive is the complex curvelet domain (Herrmann et al., 2008) in which the seismic data is decomposed into short wavefront-like components. These components have dip, scale, amplitude and phase. A number of authors have considered curvelet domain subtraction methods in recent years (Saab et al., 2007; Herrmann, 2008; Yu and Yan, 2011; Wu and Hung, 2015; Nguyen and Dyer, 2016; Zhai et al., 2017). Many adaptive methods are very powerful because they have the ability to remove *all* of the energy from the input data if they are used without constraint. The curvelet domain method is no exception in this regard, therefore it is essential to moderate its behaviour appropriately.

Neelamani et al. (2010) proposed an adaptive framework in the complex curvelet domain to subtract estimated multiples from seismic data. In their approach, each complex curvelet coefficient of the noise model is rotated and scaled to match the corresponding coefficient in the input data. The subtracted results are then inverse transformed into the space-time domain. Their method can also be used to subtract a number of noise models simultaneously. Although this is an elegant method, it has two disadvantages. First, it acts on each coefficient in the curvelet domain independently, i.e., its action is not coupled to any other local coefficient. This results in rapid variations of the filter coefficients which can over-fit the noise to the data. Therefore, the user needs to specify a constraint on amplitude and phase variations for every single curvelet coefficient in order to avoid perfect fitting of the noise to the data. Additional control of the rapid change of filter coefficients can be achieved by coupling the solution of neighbouring curvelet coefficients (Herrmann, 2008). Second, since the unknowns (amplitude and phase of the filter) are in polar coordinates, this method requires the solution of a non-linear optimisation problem.

To overcome these disadvantages, we reformulate this problem as a linear constrained optimisation problem that exploits neighbouring curvelet coefficient information in the 3D curvelet domain. We pose the problem in a standard least squares manner and demonstrate how the use of polar co-ordinates turns a linear problem into a more difficult non-linear problem. We find that this linear approach is not only very efficient but by suitable regularisation and windowing, it is also easily constrained. We demonstrate our approach on both synthetic and real data before drawing our conclusions.

THEORY

Let us consider the problem of subtracting N different modeled estimates of noise in an optimum manner from an input dataset in the complex curvelet domain. In what follows, the number of curvelets, L, may be all, or some subset, of the curvelet domain. Let the data be $\mathbf{d} \in \mathbb{C}^{L\times 1}$, the k^{th} of N noise models be, $\mathbf{m}_k \in \mathbb{C}^{L\times 1}$ and the adaptive filter coefficients be $\mathbf{f} \in \mathbb{C}^{N\times 1}$. The collection of all N noise models is the matrix $\mathbf{M} \in \mathbb{C}^{L\times N}$. The least squares adaptive subtraction objective function is given by,

$$J = \frac{1}{2}\left\|\mathbf{d} - \mathbf{M}\mathbf{f}\right\|_2^2, \qquad (1)$$

in which, for notational simplicity, we have omitted regularisation terms. The elements of the complex filter, $\mathbf{f}$, dictate the linear combination of columns of $\mathbf{M}$ to subtract from the data, $\mathbf{d}$. Since $\mathbf{f}$ is complex, objective function (1) does not satisfy

10.1190/segam2019-3216021.1

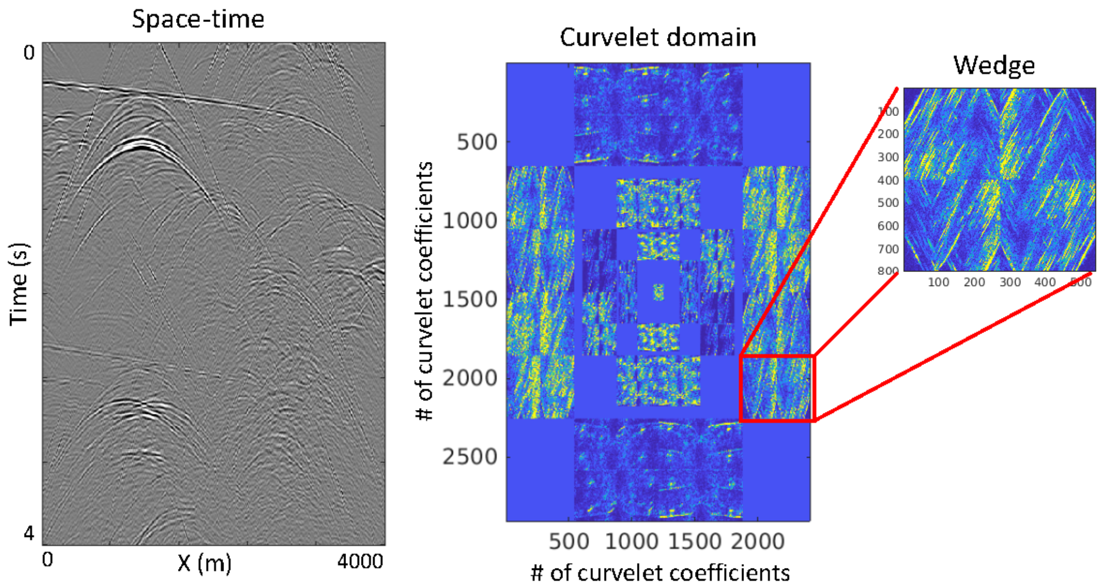

Figure 1: An example of data in the TX domain and the corresponding curvelet domain. The zoom section shows a wedge for the last scale and the last angle.

the Cauchy-Riemann equations, therefore, it is a not an analytic function and $\partial J/\partial \mathbf{f}$ is not defined. As a result, we write the gradient with respect to $\mathbf{f}$ using the generalized complex derivative (Brandwood, 1983) :

$$\frac{\partial J}{\partial \mathbf{f}} = \frac{\partial J}{\partial \mathrm{Re}(\mathbf{f})} - i\frac{\partial J}{\partial \mathrm{Im}(\mathbf{f})} = (\mathbf{Mf}-\mathbf{d})^H\mathbf{M}, \tag{2}$$

which is linear in the real and imaginary parts of $\mathbf{f}$. If instead, we had considered parameterising the filter in polar form, with $f_k = a_k e^{i\theta_k}$, which we can write as either, $\mathrm{diag}(a_k)e^{i\theta_k}$ or $\mathrm{diag}(e^{i\theta_k})a_k$, we would find that,

$$\frac{\partial J}{\partial \mathbf{a}} = \mathrm{Re}\left[\frac{\partial J}{\partial \mathbf{f}}\frac{\partial \mathbf{f}}{\partial \mathbf{a}}\right] = \mathrm{Re}\left[(\mathbf{Mf}-\mathbf{d})^H\mathbf{M}\mathrm{diag}(e^{i\theta_k})\right]. \tag{3}$$

$$\frac{\partial J}{\partial \theta} = \mathrm{Re}\left[\frac{\partial J}{\partial \mathbf{f}}\frac{\partial \mathbf{f}}{\partial \theta}\right] = \mathrm{Re}\left[i(\mathbf{Mf}-\mathbf{d})^H\mathbf{M}\mathrm{diag}(a_k e^{i\theta_k})\right]. \tag{4}$$

These equations ((3), (4)) are non-linear in the unknowns. Each equation contains both sets of unknowns. Therefore, solving for polar co-ordinates turns out to be a non-linear optimisation problem, which is typically more difficult and time consuming to solve than the linear problem. As a result, we choose to solve the linear problem, (2), in rectangular co-ordinates. This is easily solved with a choice of efficient methods. Note that, Neelamani et al. (2010) solved equation (3) and (4) for $L = 1$ with both amplitude and phase constraints.

Constraints

We have described how, in the curvelet domain, we can derive N complex filter coefficients, $\mathbf{f}$, to fit N L-length noise estimates, $\mathbf{m}_k$ to the L-length data vector, $\mathbf{d}$. We use two ways to constrain the adaption filter; a damping term, such as Tikhinov regularisation, added to (1) and the degree to which this system is over-determined by adjusting L. If $N \geq L$, perfect fitting can be achieved, whereas, if $N < L$, the equations would only be approximately satisfied. Therefore, L can control the degree of fitting that takes place. The freedom to vary L allows considerable flexibility. For example, L could be chosen to be the whole of the curvelet domain, individual wedges or partitions of wedges. It even means that the L coefficients do not need to be from a contiguous region of the transform if that curious choice demonstrated some advantage.

The particular transform that we use is the 3D curvelet transform in which the curvelet coefficients are divided into several *scales*. The coefficients within each scale are further partitioned into several directions or dips. Each scale/direction pair is known as a *wedge*. This arrangement is shown in Figure 1.

Of the many choices to windowing the curvelet domain, we have found that partitioning each wedge into windows that are proportional to the wedge size works well. That is, we use the same number of windows in every wedge. This means that L is larger at higher scales and smaller at lower scales. However, we can conceive of many other arrangements that could have advantages and disadvantages. This algorithm can be succinctly written in pseudo-code,

```
curvelet transform d, m_1, ..., m_N
for each wedge
      for each window, j
            solve M_j^H M_j f_j = M_j^H d_j for f_j
            apply b_j = d_j - M_j f_j
            sum b_j into output wedge
      end window
end wedge
inverse curvelet transform
```


10.1190/segam2019-3216021.1

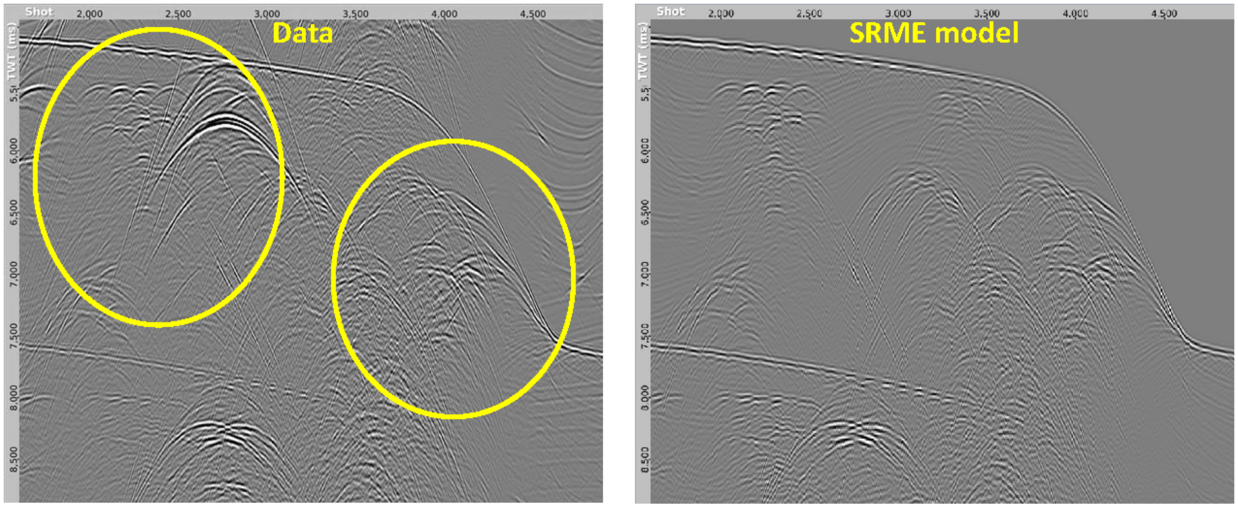

Figure 2: A common offset synthetic section (left) over the Sigsbee model and its corresponding SRME multiple model (right). The yellow circles indicate areas of interest shown zoomed in Figure 3.

DATA EXAMPLES

We demonstrate the application of this adaptive subtraction technique on a synthetic data example and a real data example. We generated synthetic data using the Sigsbee model (SMAART JV, 2001). Figure 2 shows a common offset section and the corresponding SRME multiple model. After a bulk time shift and scalar we applied the curvelet domain method with appropriate constraints to minimise the residual energy in the output. The two areas circled are shown in more detail after adaptive subtraction in Figure 3. In Figure 3 we show, from left to right, the results of subtraction using a single global time domain filter, adaptive subtraction using 2D curvelets and addaptive subtraction using 3D curvelets. The upper row and lower rows are from the left and right circled areas, respectively, in Figure 2. The single time shift plus scaling is sub-optimum, with complex parts of the multiples still present (as indicated by the red arrows). The 2D adaptive curvelet subtraction is a substantial improvement, although there is still residual multiple energy in some areas (yellow arrows). The 3D adaptive curvelet subtraction is significantly better than the 2D version, with residual multiples being very hard to detect.

Our real data example is shown in Figure 4. The left hand section is the CMP stack of the input data with complex multiple energy indicated by the circle. The middle section is the subtraction of the SRME model, after conventional windowed space-time domain 1D adaptive filtering. It clearly has residual complex multiple energy remaining (in the circled area) as well as a residual low frequency imprint of the multiple (shown by yellow arrows). The right hand section is the result of using the curvelet domain adaptive subtraction described in this paper. The complex multiple energy and the low frequency imprint has been very effectively attenuated, helping to reveal the underlying primary structure.

CONCLUSIONS

There has been a rapid adoption of curvelet domain adaptive subtraction techniques over the last few years. We have shown that the non-linear optimisation method of Neelamani et al. (2010) has some drawbacks that can be avoided by using a linear approach where the complex filter coefficients are in rectangular form (real and imaginary parts), rather than in polar co-ordinates (amplitude and phase). We constrain our optimisation with a damping term and by controlling the degree to which systems of equations are over-determined. Although we have considerable flexibility in how we partition the curvelet domain, we find that windowing inside each wedge with windows proportional to wedge size is a very effective approach. As a result we reduce unhelpful fluctuations in the curvelet coefficients and avoid over-fitting. We have shown, using the Sigsbee synthetic data and a real dataset, that this curvelet domain adaptive subtraction is very effective, out-performing conventional time-space domain approaches. This is particularly marked for the 3D curvelet approach.

ACKNOWLEDGMENTS

We would like to express our appreciation to Spectrum Geo Ltd for permission to use the field data example and to DownUnder GeoSolutions for allowing us to present this work.

10.1190/segam2019-3216021.1

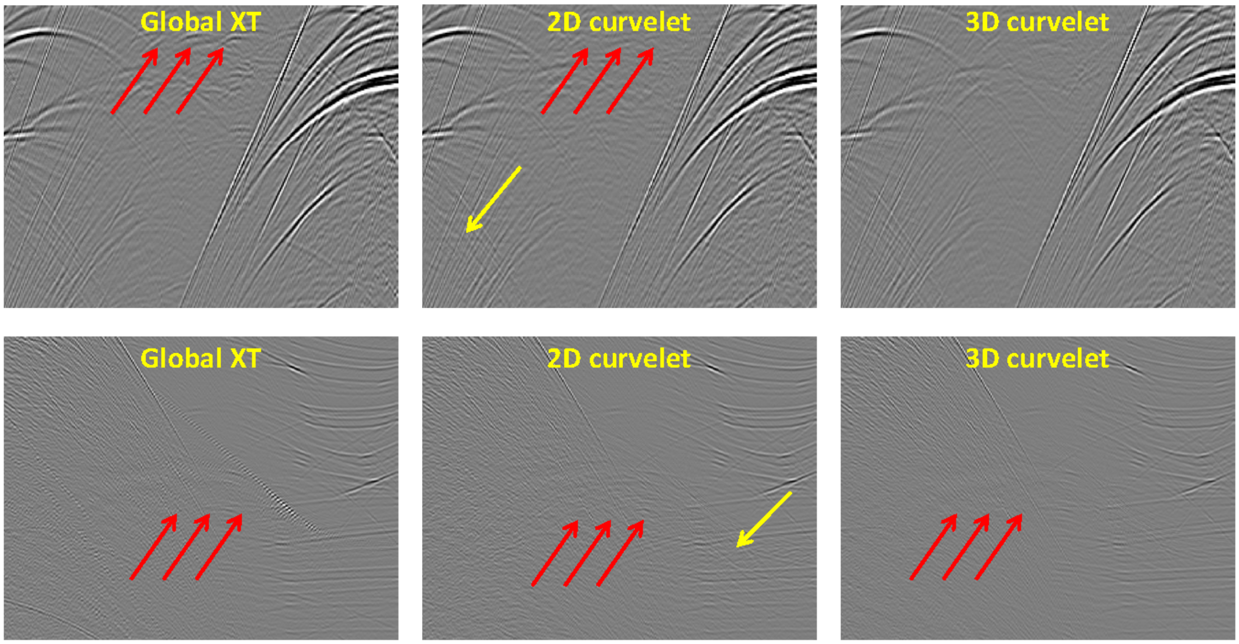

Figure 3: The results of adaptive subtraction. The columns show, from left to right, adaptive subtraction using a single global time domain filter, adaptive subtraction using 2D curvelets and adaptive subtraction using 3D curvelets. The upper and lower rows are from the left and right circled areas, respectively, in Figure 2.

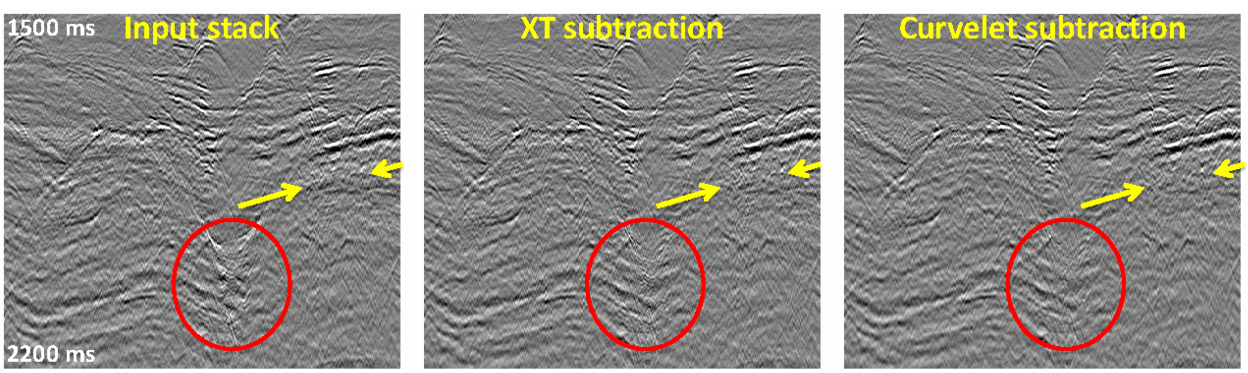

Figure 4: A real data example of an input stack (left), subtraction using conventional space-time domain adaptive subtraction (middle) and curvelet domain adaptive subtraction (right). (Data courtesy of Spectrum Geo Ltd.)

REFERENCES

Brandwood, D., 1983, A complex gradient operator and its application in adaptive array theory: IEE Proceedings H-Microwaves: Optics and Antennas, IET, 11–16.

Dragoset, W., and Z., Jericevic, 1998, Some remarks on multiple attenuation: Geophysics, **63**, 772–789, doi: https://doi.org/10.1190/1.1444377.

Herrmann, F. J., 2008, Curvelet-domain matched filtering: 78th Annual International Meeting, SEG, Expanded Abstracts, 3643–3649, doi: https://doi.org/10.1190/1.3255624.

Herrmann, F. J., D., Wang, and D. J., Verschuur, 2008, Adaptive curvelet-domain primary-multiple separation: Geophysics, **73**, no. 3, A17–A21, doi: https://doi.org/10.1190/1.2904986.

Ikelle, L., G., Roberts, and A., Weglein, 1997, Source signature estimation based on the removal of first-order multiples: Geophysics, **62**, 1904–1920, doi: https://doi.org/10.1190/1.1444291.

Neelamani, R., A., Baumstein, and W. S., Ross, 2010, Adaptive subtraction using complex-valued curvelet transforms: Geophysics, **75**, no. 4, V51–V60, doi: https://doi.org/10.1190/1.3453425.

Nguyen, T., and R., Dyer, 2016, Adaptive multiple subtraction by statistical curvelet matching: 86th Annual International Meeting, SEG, Expanded Abstracts, 4566–4571, doi: https://doi.org/10.1190/segam2016-13710124.1.

Saab, R., D., Wang, O., Yilmaz, and F. J., Herrmann, 2007, Curvelet-based primary-multiple separation from a Bayesian perspective: 77th Annual International Meeting, SEG, Expanded Abstracts, 2510–2514, doi: https://doi.org/10.1190/1.2792988.

SMAART JV, 2001, Sigsbee 2a acoustic velocity model: SEG Research Workshop on Velocity Model Independent Imaging for Complex Media.

Wu, X., and B., Hung, 2015, High-fidelity adaptive curvelet domain primary-multiple separation: First Break, **33**, 53–59.

Yu, M., and Z., Yan, 2011, Flexible surface multiple attenuation using the curvelet transform: 81st Annual International Meeting, SEG, Expanded Abstracts, 3485–3489, doi: https://doi.org/10.1190/1.3627922.

Zhai, Y., Z., Liu, J., Sheng, B., Wang, and J., Heim, 2017, A hybrid crossgather curvelet domain multiple elimination method and its application: 87th Annual International Meeting, SEG, Expanded Abstracts, 4752–4757, doi: https://doi.org/10.1190/segam2017-17741149.1.

10.1190/segam2019-3216021.1