Noise attenuation by 4D greedy Radon transform

Juefu Wang
Primary GeoSerivices Ltd.

Summary

I present a noise attenuation method using a frequency domain 4D greedy Radon transform that simulates local seismic data with limited number of dips in three directions. The three directions correspond to CDP-x, CDP-y and offset for 3D data. The optimal solution is computed by a greedy approach. Particularly events with larger amplitude are solved earlier than those with smaller amplitude. A real data example shows that the algorithm can effectively suppress random noise and preserve signal. It can also reduce high-dip linear noise because only a limited number of moveout parameters are used in the inversion. Due to the flexibility of Radon transform in handling the survey geometry, the algorithm can process data with irregular spatial sampling, and it does not lose seismic traces due to binning as required by many conventional methods.

Theory

Seismic data contains both random noise and coherent noise. Usually they are treated in different ways in seismic data processing. In this paper I mainly focus on removing random noise, although the proposed method can also reduce aliased linear noise as I show later in this paper. Because random noise is not predictable, it cannot be modelled by linear systems. We can reduce random noise by simulating seismic data with linear equations. Naturally the residual is the random noise. We can have different random noise attenuation methods based on different linear systems. For example, the classic prediction filtering method (Canales, 1984) assumes that a seismic trace can be predicted by summing nearby traces with proper weights in the FX (frequency and space) domain. The task is to optimize the weights with redundant linear equations, then predict clean data with the weighted sum of nearby traces using the optimized weights. Another example is the polynomial approach (Wang, 2004). The method tries to fit data with polynomials based on observed data and its spatial coordinates by solving for coefficients of polynomials. The clean data is then modelled with the optimal coefficients. A recent popular method called Cadzow filtering (Trickett, 2002; Trickett, 2015) assumes that coherent signal is contained in the first few strong eigenvectors of a Henkel matrix. Note that both the prediction filtering method and Cadzow filtering method need to have data on a regular grid. In reality, this requirement is compromised due to limited financial budget and complex surface conditions. A practical solution is to bin the data on a predefined grid. The binning process introduces errors in the data prediction, and if there are multiple traces located in a bin, only one trace will be processed. The polynomial method can handle irregular geometry, but it struggles to handle structured data.

To overcome the shortcomings of conventional methods, I use a 4D Radon transform to setup a linear system to predict seismic data. Radon operators can honor true locations of seismic traces, and they can handle spatially aliased data with sparseness constraints (Thorson and Claerbout, 1985; Sacchi and Ulrych, 1995; Herrmann et al., 2000; Trad et al, 2003; Hugonnet and Boelle, 2007). In the past, Radon transform has mainly been used as a tool to remove coherent noise (e.g., linear noise and multiples). The idea is to focus seismic events in the Radon domain, and then the model of noise is extracted and transferred back to the space and time domain. Finally the predicted coherent noise is subtracted from the input. Here for random noise attenuation, I simply subtract the data misfit of inversion from the input.

The 4D Radon transform is much more expensive than the conventional 2D Radon transform because the model space grows by two dimensions. To make the algorithm practical, I limit the model space by limiting the number of dips of seismic events in each local spatial and temporal window.

In the space and temporal frequency (FX) domain, noise-free seismic data can be modelled using the following equation:

$$d(x, y, h, \omega) = \sum_{p_x, p_y} m(p_x, p_y, p_h, \omega) e^{i\omega(p_x x + p_y y + p_h h^2)}, \quad (1)$$

where *x* and *y* are two spatial coordinates, h is the offset, p_x and p_y are slopes (normalized moveout parameters) in crossline and inline directions, p_h is the slope in the offset direction, $d(x, y, h, \omega)$ is seismic data in the FX domain, $m(p_x, p_y, p_h, \omega)$ is the Radon model in the p and temporal frequency domain, The adjoint of Equation (1) can be expressed as

$$m(p_x, p_y, p_h, \omega) = \sum_{x,y} d(x, y, h, \omega)\, e^{-i\omega(p_x x + p_y y + p_h h^2)}. \quad (2)$$

Equation (2) describes a slant-stacking process in the FX domain to transform seismic data to a Radon model.

Considering noise I modify Equation (1) in matrix form as below

$$d = Lm + n, \quad (3)$$

where d is the seismic data as a vector, L is the forward operator as a matrix, m is the Radon model as a vector, and n is the random noise as a vector. To attenuate the noise is equivalent to minimize the following cost function:

$$J(m) = \|d - Lm\|^2. \quad (4)$$

To avoid overfitting the noise, I regularize the model using the following cost function instead

$$J(m) = \|d - Lm\|^2 + \mu R(m), \quad (5)$$

where R is a regularizing function based on *a priori* information of the model, and μ is a trade-off parameter to determine the amount of regularization. A lot of work has been devoted to imposing sparseness to the Radon model. For example, L_1 norm or Cauchy norm can be used to enhance model resolution (Sacchi & Ulrych, 1995). Alternatively, a greedy approach (Ng and Perz, 2004) can be adopted to obtain a model with even higher resolution. There is no explicit regularizing function for the greedy method. The sparse feature is achieved by a prioritized inversion based on the amplitude of the Radon model. Below is a greedy implementation for the frequency domain 4D Radon transform:

1. Convert data from $d(x, y, h, t)$) to $d(x, y, h, \omega)$ (i.e. from TXY to FXY domain) by 1D Fourier transform along the time direction.
2. Loop over frequencies to find an optimal Radon model for each frequency using the following algorithm in pseudo code.

$d_{resi} = d, \quad m = 0;$
$for\ iter = 1, nitr$
  $m_e = L' d_{resi};$
  $sort\ \ m_e\ \ in\ amplitude\ descending\ order\ and\ keep\ the\ sorting\ indices\ of\ first\ N$
  $samples$;
  for i=1,N
    optimize the weights of the i-th model element using the steepest descent method;
    update the data residual d_{resi}*by sutracting the predicted data using the optimal model;*
  end
end
$d_{clean} = d - d_{resi};$

Note that L' is the adjoint operator described by Equation (2).

3. Convert clean data from $d_{clean}(x, y, h, \omega)$ to $d_{clean}(x, y, h, t)$ by 1D inverse Fourier transform.

The algorithm is similar to the anti-leakage Fourier transform (Xu et al, 2005) except that it uses a different operator and solver. First of all, Radon transform instead of Fourier transform is used so that it is more straightforward to control the dip range of seismic data. Second of all, in this algorithm there is no need to recalculate the full adjoint model after optimizing a single set of moveout parameter (p_x,p_y, p_h), A series of similar moveout parameters are solved after each amplitude prioritization, which can improve convergence and significantly reduce the computational cost without scarifying the resolution. Typically 6-8 iterations of the outer loops are enough to obtain a satisfactory result. The number of dips (*N* as described in the pseudo code) to solve in the inner loop is usually set to 20-40, a small fraction of the total number of the Radon model elements. Therefore the main cost of the algorithm is the step of computing the full adjoint model in the outer loop, which is used to find dips with maximum amplitude.

In practice, we often limit the ranges of moveout parameters to be big enough to model the signal but not linear noise with large dips. Consequently the linear noise is not modelled, and it remains in the residual. Therefore the algorithm can naturally remove linear noise with large dips (often aliased).

Field data example

I tested the algorithm with a real 3D land dataset. Figure 1 compares a CDP gather before and after Radon denoise. It can be seen that there is a lot of noise including random noise and erratic noise present in the input gather due to using a vibrator as the source. To make the inversion stable, I precondition the input with Automatic Gaining Control (AGC). After noise is suppressed, the AGC scalar is removed for each time sample. The algorithm significantly reduced the random and erratic noise from the input (see Figure 1b). In addition, a lot of linear noise is also removed. The difference plot indicates that there is no obvious signal leakage in the process. The result is preserving the amplitude, which is helpful for AVO compliant processing.

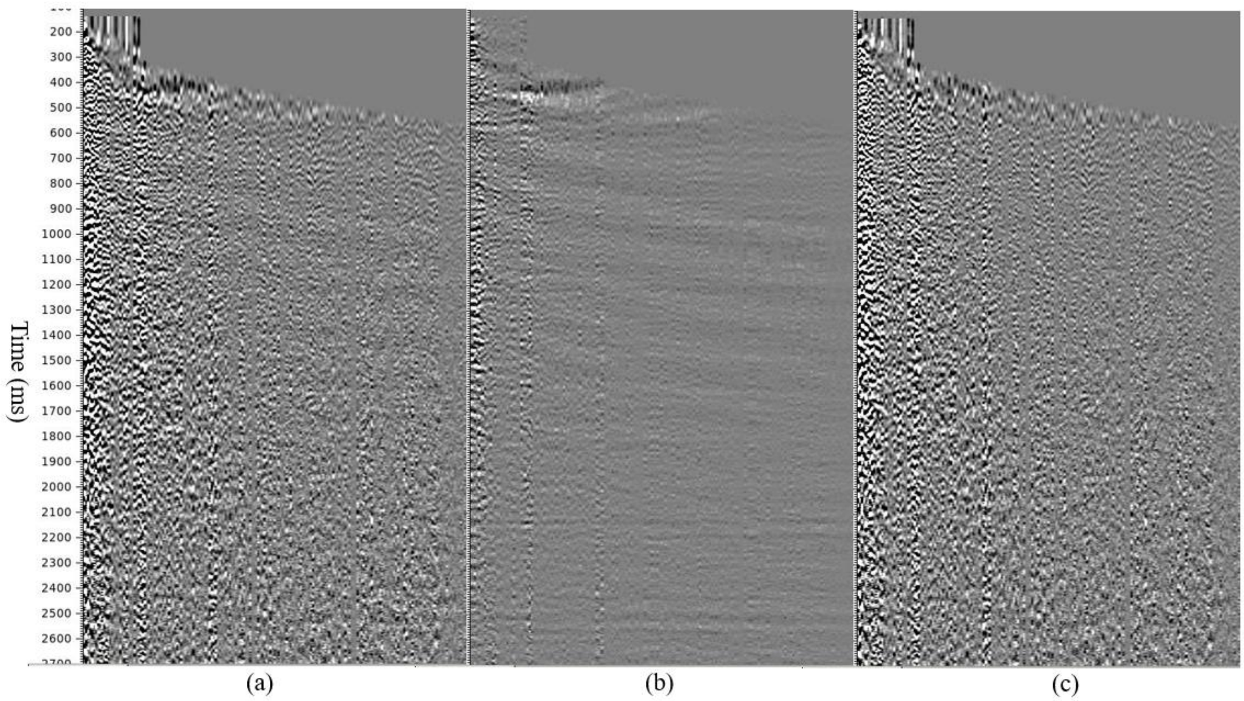
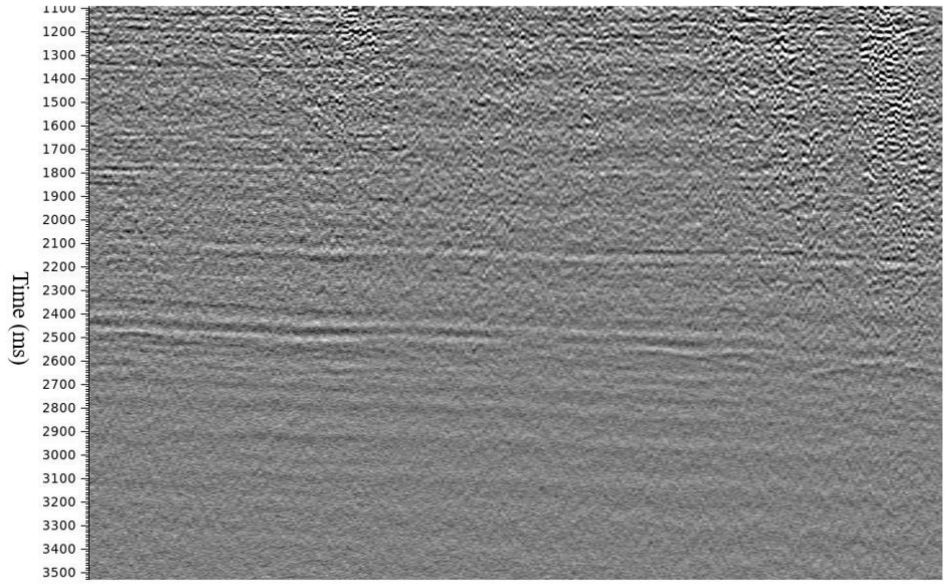
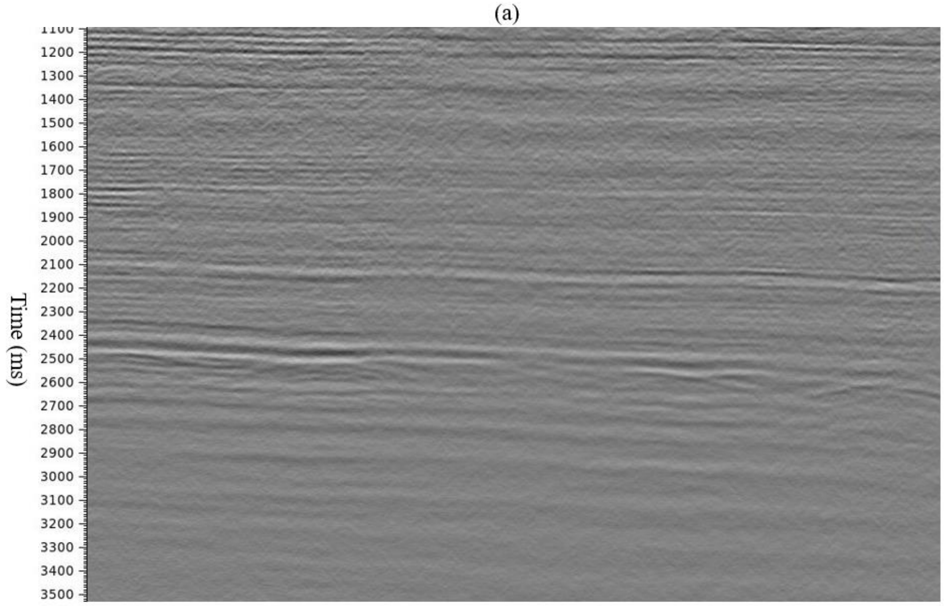


Figure 1: Comparison of CDP gathers before and after greedy Radon denoise. (a) Input gather. (b) Output gather. (c) Difference.

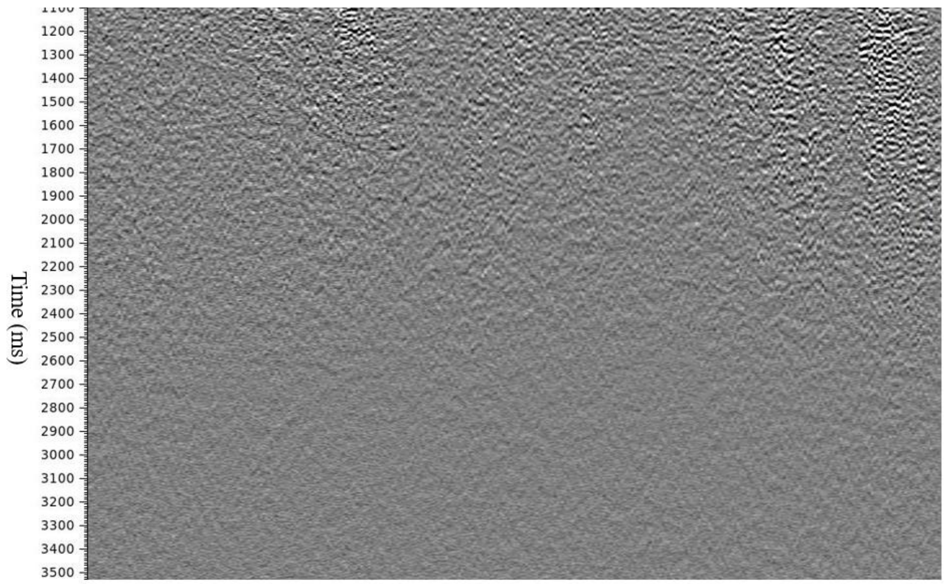
To further check the effectiveness of the method, I compare the stacks before and after greedy Radon transform in Figure 2. It can be seen that the algorithm significantly improves the quality of the structural stack without obvious signal leakage. Both random noise and linear noise is suppressed. It is interesting that the all diffractions are well preserved in the result. Note that curved diffractions often pose a challenge to methods that assume local linear events.


Time (ms)
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000
2100
2200
2300
2400
2500
2600
2700
2800
2900
3000
3100
3200
3300
3400
3500

(a)


Time (ms)
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000
2100
2200
2300
2400
2500
2600
2700
2800
2900
3000
3100
3200
3300
3400
3500

(b)



(c)

Figure 2: Comparison of stacks before and after greedy Radon denoise. (a) Input stack. (b) Output stack. (c) Difference.

Conclusions and discussion

I have implemented a versatile noise attenuation algorithm using a frequency domain 4D greedy Radon transform. The method can effectively reduce random noise and residual linear noise and preserve signal even when data is highly structured. The algorithm can be applied to 3D data in the CDP domain, which allows for fast testing on target lines without sorting data to different domains. This is an advantage that helps to reduce processing time. Since true geometry of the seismic survey is honored, the method can preserve detailed structure and accurate amplitude better than many conventional methods that require binning of seismic data.

The idea of optimizing multiple dips in the inner loop after sorting all events based on amplitude can reduce the computational cost of greedy algorithms in higher dimensions. A straightforward application is to develop 4D/5D interpolation algorithms to improve spatial sampling of seismic data.

Acknowledgements

I would like to thank Primary GeoServicies and Absolute Imaging for allowing me to publish this paper. I would also like to acknowledge the processing team at Absolute Imaging for their effort in preparing and reviewing the real data example. Special thanks to Professor Mauricio Sacchi for the discussion about noise attenuation.

References

Canales, L. L., 1984, Random noise reduction: 54th Annual International Meeting, SEG, Expanded Abstracts, 525-527.

Herrmann, P., T. Mojeskey, M. Magesan and P. Hugonnet, 2000, De-aliased, high-resolution Radon transforms: 70th Annual International Meeting, SEG, Expanded Abstracts, 1953-1956.

Hugonnet, P. and J. L. Boelle, 2007, Beyond aliasing regularization by plane wave extraction: 69th Conference & Exhibition, EAGE, Extended Abstracts, 114-117.

Ng, M. and M. Perz, 2004, High resolution Radon transform in the t-x domain using "intelligent" prioritization of the Gauss-Seidel estimation sequence: 84th Annual International Meeting, SEG, Expanded Abstracts, 2160-2163.

Sacchi, M. D. and T. J. Ulrych, 1995, High-resolution velocity gathers and offset space reconstruction: Geophysics, **60**, 1169-1177.

Thorson, J. R. and J. F. Claerbout, 1985, Velocity stack and slant stack stochastic inversion: Geophysics, **50**, 2727-2741.

Trad, D., T. Ulrych and M. D. Sacchi, 2003, Latest views of the sparse Radon transform: Geophysics, **68**, 386-399.

Trickett, S., 2002, F-xy eigenimage noise suppression: 72nd Annual International Meeting, SEG, Expanded Abstracts, 2166-2169.

Trickett, S. 2015, Preserving signal: Automatic rank determination for noise suppression: 85th Annual International Meeting, SEG, Expanded Abstracts, 4703-4707.

Wang, X., 2004, 3D prestack seismic trace interpolation in time domain with input optimization: CSEG National Convention.

Xu, S., Y. Zhang, D. Pham and G. Lambare, 2005, Antileakage Fourier transform for seismic data regularization: Geophysics, **70**, v87-v95.