Robust Tau-P denoise

Zhigang Zhang, TGS*

Summary

This abstract describes a denoise algorithm using the linear Radon transform. Instead of separating the signal and noise after the transformation, the proposed method simultaneously builds a noise and a signal model with the transform. Explicitly including the noise model in the computation reduces artifacts, and the algorithm is more robust to erratic noise than the transform-separation method. Working in the intercept time slowness (Tau-P) domain improves the sparsity of the solution and reduces the impact of aliasing for under-sampled data. The extra cost of calculating the noise model is negligible because only copy and simple multiplication-addition are involved in each iteration. Both synthetic and field data examples prove the effectiveness of the filter.

Introduction

The importance of denoise cannot be overstated for seismic imaging. In practice, noise is classified into coherent and random, depending on continuity. The transform-separation method, the rank-reduction method, and the prediction error filter are three standard denoising tools.

The transform-separation method explicitly decomposes the data in an appropriate domain, selectively picks amplitudes, and reassembles the signal after muting and tapering. Depending on the nature of the noise and signal, various domains could be used, and the transform could be a Fourier transform (Stewart and Schieck, 1989), wavelet transform (Yu et al., 2022), Tau-P transform, or others. The success of this type of method depends on the distribution of the signal and noise in the transform domain and the sharpness of the filter applied. The separation cannot be clear when the noise and signal overlap in the transform domain. This is exacerbated by the fact that seismic data is mostly spatially under-sampled. Aliasing presents another serious challenge for the transformation. An aggressive denoise risks hurting signals, while a less aggressive denoise leaves more artifacts and noise.

When noise is random and follows the Gaussian distribution, the prediction error filters, such as the *f*-x filter (Canales, 1984), remove the noise effectively. However, the prediction filters' success depends on the noise's deviation from the Gaussian distribution or the orthogonality between the noise vector and the signal vector. Erratic noise does not have a Gaussian distribution and can cause strong artifacts in the results. Rank-reduction methods, such as the Cadzow filter (Stewart, 2008), experience the same difficulties in the presence of erratic noise.

Researchers have put great effort into making the filter robust to erratic noise during the past decade. Sternfels et al. (2015) propose explicitly building a noise model while searching a low-rank Hankel/Toeplitz matrix for each temporal frequency-space (*f*-x) slice. Including the L1 norm of the erratic noise model in the object function mitigates the effect the erratic noise has on the signal. Chen and Sacchi (2015) suggest the rank-reduction filter can also become robust to erratic noise if a bisquare function replaces the quadratic error criterion function. They use the iteratively reweighted least-squares method to solve the optimal robust factorization.

Both strategies mentioned above work in the *f*-x domain and assume the erratic noise is present only in a small subset of the traces. However, in some cases, intense erratic noise can contaminate most of the data and spread over most traces, which is especially common for blended data. The *f*-x domain does not support the sparsity of the erratic noise.

In this abstract, I propose to denoise in the Tau-P domain using the linear Radon transform. I compute the noise model and the Tau-P transform of the signal simultaneously by solving a convex optimization problem. My work is inspired by the joint low-rank and sparse inversion method proposed by Sternfels et al. (2015) but differs in several ways. Firstly, we set up our work in the Tau-P domain instead of the *f*-x domain. When erratic noise exists in multiple traces, the solution in the Tau-P domain has better sparsity, and the computation is more stable. Secondly, the optimization problem for the joint low-rank and sparse inversion method involves three norms: the nuclear norm of the Hankel matrix, the L1-norm of the erratic noise, and the L2-norm of the Gaussian noise. The multiple norm types complicate the problem, and the authors use the alternating direction method of multiplier (ADMM), which is expensive. Our objective function only has one L1 norm term, which is much simpler. Following the introduction of the method, we present both synthetic and field data tests.

Method

We assume the observed data is the sum of a signal part and a noise part, and the signal has a sparse representation on a complete basis or a subset of the basis. Let d be the observed data, we can write it in the matrix form as

$$d = Ax + n, \quad (1)$$

where x is the coefficient vector of the signal representation. The matrix A consists of the necessary basis vectors to represent the signal, which may be a subset of a complete basis. The noise n is assumed to be sparse, but tests demonstrate satisfying results in non-sparse scenarios.

Different from the transform-separation method, I solve both x and n simultaneously. Firstly, we rewrite equation (1) as

$$d = [A \quad I]\begin{bmatrix} x \\ n \end{bmatrix} = \hat{A}\hat{x}, \tag{2}$$

where I is the identity matrix, $\hat{A} = [A \quad I]$, and $\hat{x} = \begin{bmatrix} x \\ n \end{bmatrix}$. The problem is highly under-determined.

To find the solution to equation (2), I seek the combination of x and n, which solves the optimization problem given by

$$\min_{x,n} |x|_1 + \alpha |n|_1 \quad \text{subject to } d = \hat{A}\hat{x}, \tag{3}$$

where α balances the signal and noise.

The optimization problem minimizes the cost of representing the observed data using a frame that combines two bases. One basis consists of all the columns of A, and the other is the standard basis. A is carefully chosen so that the cost to represent the desired signal in terms of the columns of A is much less than that using the standard basis. A satisfying basis mostly leads to a sparse solution of x. In the extreme case, only one of the elements of x is non-zero, and the desired signal matches one column of A. On the other side, the representation of the noise part, including random noise and undesired coherence events, in terms of the columns of the matrix A should be more expensive when compared to the standard basis, which is their L1 norm.

The optimization problem is a standard base pursuit problem (Berg and Friedlander, 2008), and I solve it using the spectral projected gradient method for the L1 norm (SPGL1). SPGL1 (Birgin, Martínez, and Raydan, 2000) is a fully developed solver widely used by the imaging community, with a proven record of being fast and stable. I do not separate erratic noise from Gaussian noise either. It is not practical for a processor to specify the noise level of the Gaussian noise.

I chose Tau-P domain in the computation because the transform is easy to implement. The algorithm itself is not limited to this domain. As discussed previously, a working domain providing better sparsity is always preferred. When the noise has some pattern and can be represented sparsely in a particular domain, we should use this prior by replacing the identity matrix with a basis from the specific domain.

If a source signature is available, we can expand equation (1) as

$$d = ASx + n, \tag{4}$$

where S is the source signature. A source signature improves the computation as the solution x becomes sparser.

Implementation of the operator $\hat{A}$ and its adjoint operator is the most time-consuming step of the algorithm. It includes two parts. One is the regular forward or backward transform between the Tau-P domain and the t-x domain, and the other involves the noise term n. The latter part only needs a copy operation or a multiplication-addition operation. Both operations take much less time than the linear Radon

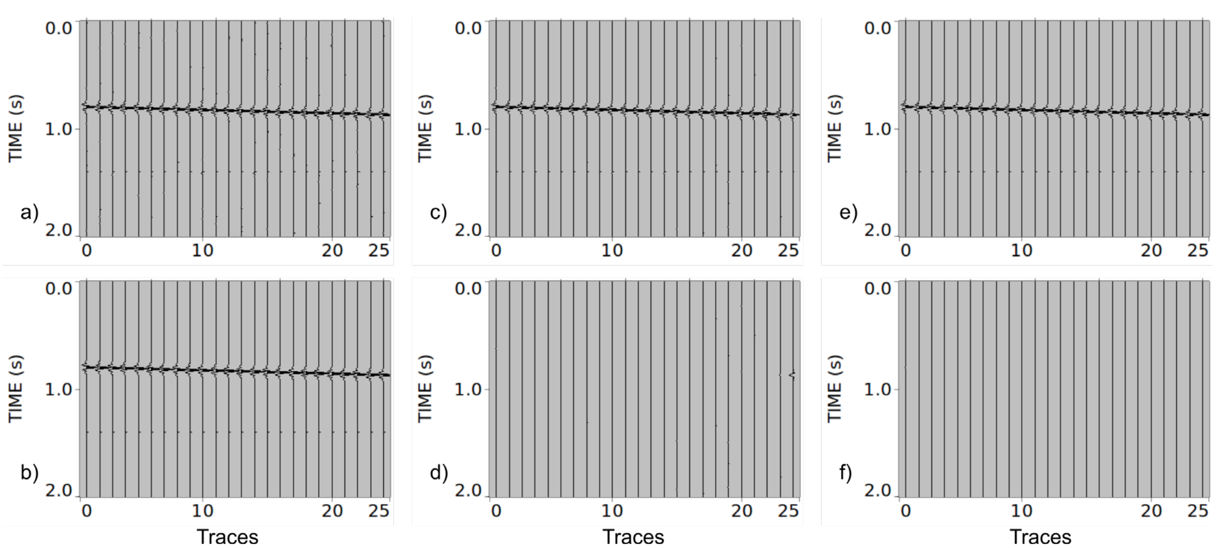

Figure 1: Comparison between the proposed method and the Cadzow filter with weak noise. a) and b): input data with noise and the noise-free data; c) and d): filtered result using the Cadzow filter and the difference between the filtered result and the true solution; e) and f): filtered result using the proposed method and the difference between the filtered result and the true solution.

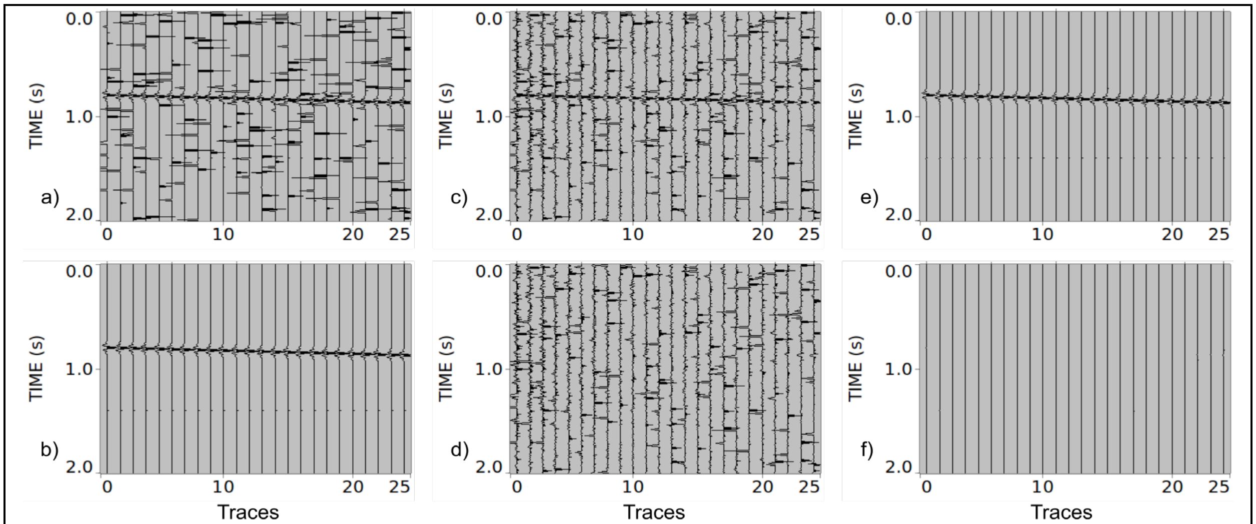

Figure 2: Comparison between the proposed method and the Cadzow filter with strong noise. a) and b): input data with noise and the noise-free data; c) and d): filtered result using the Cadzow filter and the difference between the filtered result and the true solution; e) and f): filtered result using the proposed method and the difference between the filtered result and the true solution.

transform. Fast Fourier transforms can be used to speed up the computation further.

Synthetic examples

I tested the algorithm using synthetic data and compared the proposed algorithm with the Cadzow filter. My synthetic data has 24 traces, and the recording length is 2 seconds. The distance between neighboring traces is 12.5 meters, and the offset ranges from 0 to 287.5 meters. The noise-free data consists of two hyperbolic events. One starts from 800 ms at zero offset, and the other starts from 1400 ms. Both use a normal move-out velocity of 4000 m/s. The first event uses an Ormsby wavelet of amplitude 1. The Ormsby wavelet used by the second event has an amplitude of 0.2 and a different bandwidth.

Noise spikes are randomly added to the data at 200 points. The amplitude of the spikes is also random, with a uniform distribution and zero expected value. I generated three data sets using different noise strengths. In the weak noise case, the value of the noise ranges from -0.5 to 0.5. In the strong noise case, the value of the noise ranges from -10 to 10. In the median noise case, the range is between -2 to 2.

To compare the filtered result, I define the signal/noise ratio as follows,

$$SNR = 10\log\frac{Power\ of\ the\ noise\ free\ data}{Power\ of\ the\ error}.$$

The error is defined as the difference between the filtered result and the true signal. Table 1 shows the comparison between the proposed method and a Cadzow filter. The Cadzow filter is applied on time windows of 1 s length with 500 ms overlap, and we choose the truncation rank as 2. The proposed method is applied to the whole data set without dividing the data into windows, and the noise weight factor α is set to 1. The maximum number of iterations for SPGL1 is 250.

Table 1: Signal/noise ratio comparison

Input data (dB)	After denoising (dB)	
	Proposed method	Cadzow filter
10.7 (weak noise)	26.3	13.7
-1.61 (median noise)	20.4	1.20
-16.0 (strong noise)	9.71	-12.7

The proposed method is superior to the Cadzow filter regarding the signal/noise ratio in all three cases. Although the signal/noise ratio of the input data decreases from 10.7 dB to -16.0, it only changes from 26.3 dB to 9.71 dB after our proposed method. The proposed method raises the S/N ratio about 25 dB in the strong noise case, while the Cadzow filter only improves about 3 dB. This proves the strategy's robustness to erratic noise.

Figures 1 and 2 show the seismic data before and after filtering for the weak and strong noise cases. When the noise is weak, both methods achieve satisfying results. The weak event is mostly successfully recovered except for the edge effect caused by the Cadzow filter. However, when the noise becomes very strong, the Cadzow filter deteriorates quickly. Much noise remains after denoising, and the weak event disappears. The proposed method still recovers the weak event, although slight noise residue becomes obvious.

Field example

Encouraged by the success of the synthetic data, the method was tested to remove blending noise on simultaneously shot data. Our field data example is part of a shot line from a common receiver gather of an ocean bottom node (OBN) acquisition. There are 615 traces in the data set, and the distance between neighboring traces is around 50 m. The record length is 8 s, with a sample rate of 12 ms.

The proposed robust Tau-P filter is applied on small overlapping windows. Each widow has 24 sequential traces and 100 samples in the time direction with some padding. The next window moves 12 traces in the space direction or 50 samples in the time direction.

Figure 3 presents the raw input data, the filtered result after applying the robust Tau-P filter, and the difference between the previous two. Most of the noise is removed, with signal leakage being minimal. The result could be further improved by applying local linear move out to increase sparsity.

When the noise is dense, the chance for the algorithm to create false events increases. Depending on the acquisition design of simultaneous shooting, it could become a serious problem. Inversion-based deblending, which utilizes knowledge of blending time, is expected to perform better for deblending purposes and is commonly used in the industry.

Conclusion

Explicitly modeling the noise better describes the noise term, helps reduce artifacts, and makes the filter more robust to erratic noise. By solving the basis pursuit problem, we can effectively remove unwanted noise with minimal signal leakage. The solution in the Tau-P domain has stronger sparsity than the *f*-x domain and fits better the under-determined problem. Both synthetic and field data prove the proposed Tau-P filter is effective and efficient. The cost of the extra computation of the noise term is negligible compared to the Tau-P transform, and the cost could be further reduced in practice.

Acknowledgment

The author wants to thank TGS for their permission to publish this paper and Mr. Zhaojun Liu and Dr. Carlos Calderon for their support in this work. I also want to thank my colleagues for their inspiration and the reviewers for their effort and suggestions to make this paper better.

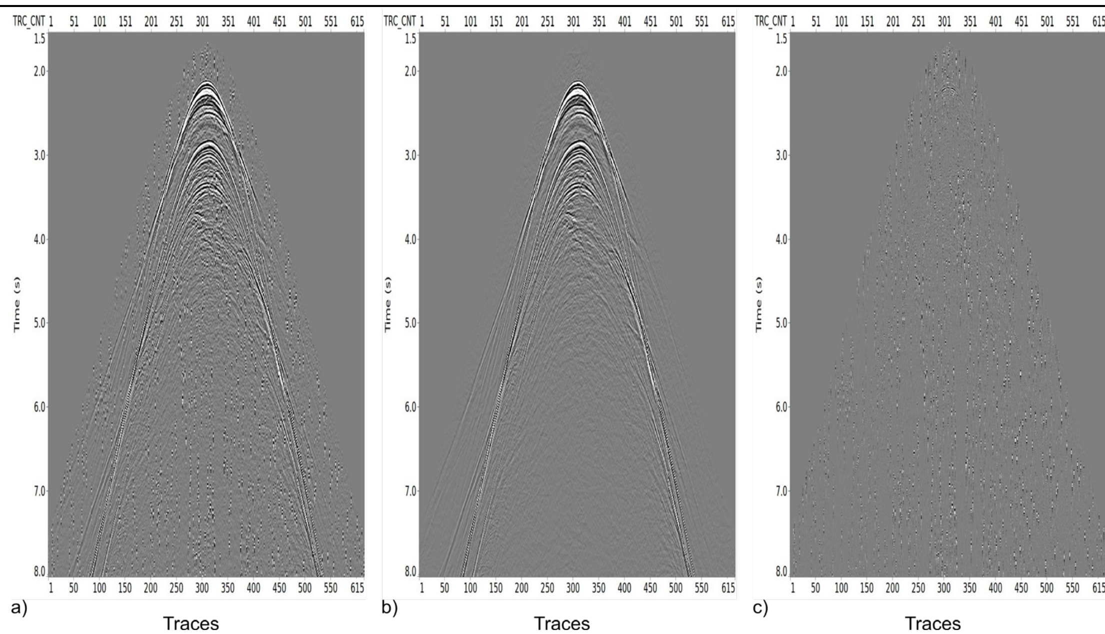

Figure 3: Field data test. The horizontal axis shows the index of traces. a) input data; b) after denoise using the proposed method; c) the difference between the input and the filtered result.