Random noise reduction using Bayesian inversion

This article has been downloaded from IOPscience. Please scroll down to see the full text article.
2012 J. Geophys. Eng. 9 60
(http://iopscience.iop.org/1742-2140/9/1/007)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 137.82.25.41
The article was downloaded on 13/12/2011 at 02:55

Please note that terms and conditions apply.
Random noise reduction using Bayesian inversion

Sanyi Yuan, Shangxu Wang and Guofa Li

State Key Laboratory of Petroleum Resource and Prospecting, China University of Petroleum, Beijing 102 249, People’s Republic of China
and
CNPC Key Laboratory of Geophysical Exploration, China University of Petroleum, Beijing 102 249, People’s Republic of China
E-mail: wangsx@cup.edu.cn

Received 9 January 2011
Accepted for publication 3 November 2011
Published 12 December 2011
Online at stacks.iop.org/JGE/9/60

Abstract
Enlightened by the classical total variation (TV) model, we present a novel random noise reduction method for seismic data based on Bayesian inversion, called Bayesian inversion filtering. The method regards 2D ‘clean’ seismic data as model parameters, and the noise reduction is equivalent to inverting these parameters from the observed data. The inversion is implemented by maximizing a posterior distribution, which is replaced by the product of a priori distribution and a likelihood function. The performance of this method mainly depends on the choice of a priori information. Based on statistical knowledge or assumption that coherent events oscillate slightly and random noise strongly, TV is used as the prior information to control the noise reduction. A variety of synthetic and real data examples show that the method can clarify amplitude, phase and frequency variations in the spatial direction. Moreover, the geometrical properties of the events, such as the curvature and slope, can be retained. In particular, this method preserves the edges of discontinuous events, which usually correspond to important geologic features.

Keywords: Bayesian inversion, noise reduction, prior information, total variation, edge-preserving

1. Introduction

The basic principle of seismic exploration is to reveal the structural features and physical properties of subsurface reflectors by using seismic records. However, random disturbances always smear the response from reflectors. Therefore, reducing random noise with high signal fidelity is critical to the following seismic processing, interpretation and inversion.

So far, a number of random noise reduction methods have been proposed, such as $f-x/f-xy$ deconvolution (Canales 1984, Chase 1992), $f-x/f-xy$ projection filtering (Soubaras 1994, Ozdemir et al 1999), prediction filtering in $t-x$ domain (Abma and Claerbout 1995), (local) singular value decomposition (Ulrych et al 1988, Bekara and van der Baan 2007), (partial) Karhunen–Loeve transform (Jones and Levy 1987, Al-Yahya 1991), $f-x/f-xy$ Cadzow (Trickett 2008, Yuan and Wang 2011), wavelet transform (Ioup and Ioup 1998), curvelet transform (Wang et al 2010), empirical mode decomposition (Bekara and van der Baan 2009), (time-varying) median filtering (Bednar 1983, Liu et al 2009) and edge-preserving smoothing (AlBinHassan et al 2006, Liu et al 2010). Every method has its own advantages and limitations. For example, $f-x$ deconvolution can deal with linear events directly and has a good noise reduction effect. But for nonlinear events, it needs to subdivide the data into smaller panels where the events are linear. Thus $f-x$ deconvolution has to take the windows into account.

Bayesian inversion (Tarantola 2005) has been widely applied to retrieve the structural features and physical properties of subsurface reflectors. The basic idea is to invert model parameters by maximizing a posterior probability, which is replaced by the product of a likelihood function and a priori distribution. Generally, the likelihood function is readily determined. Therefore, the prior distribution, which usually plays an important part in reducing ambiguity of
According to Bayes’ rule (Bayes 1763), the conditional posterior distribution, can be written as

\[ p(S|D) = \frac{p(D|S)p(S)}{p(D)} \propto p(D|S)p(S) \]  

Assuming that the observed seismic data \( D \) are the simple superposition of the signal \( S \) with random noise \( N \):

\[ D = S + N. \]  

According to Bayes’ rule (Bayes 1763), the conditional probability of the signal given the observed data, also called posterior distribution, can be written as

\[ p(S|D) = \frac{p(D|S)p(S)}{p(D)} \propto p(D|S)p(S), \]  

where \( p(D|S) \) is the conditional probability of the observed data given the signal, also called likelihood function; \( p(S) \) is the prior probability of the signal, and \( p(D) \) is the prior probability of the observed data. The symbol \( \propto \) denotes that \( p(S|D) \) is proportional to \( p(D|S)p(S) \) since \( p(D) = \int p(D|S)p(S)dS \) is a constant (Ulrych et al. 2001).

Assuming that \( N \) can be approximated as uncorrelated zero-mean Gaussian noise with unknown variance \( \sigma^2 \), the likelihood function in equation (2) can be expressed as (Theune et al. 2010)

\[ p(D|S) = C_1 \exp \left[ -\frac{|D - S|^2}{2\sigma^2_n} \right]. \]

where \( C_1 \) is a constant; \( |D - S|^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} (d_{ij} - s_{ij})^2 \), \(|\bullet||\bullet|\) denotes the Frobenius norm, \( d_{ij} \) and \( s_{ij} \) are the elements of matrices \( D \) and \( S \), respectively; \( m \) and \( n \) are time sample number and trace number, respectively. The magnitude of \( p(D|S) \) is used to measure the misfit between the observed data and the noise reduction data.

In a word, the idea of Bayesian inversion filtering is to find an estimate \( \hat{S} \) of the signal \( S \) by maximizing the posterior distribution \( p(S|D) \). The process of inversion is the process of noise reduction. However, before the \( \hat{S} \) is estimated, it is crucial to obtain a priori information about \( S \).

In this paper, the prior information is assumed as a function \( \varphi(S) \); thus, the prior distribution of \( S \) can be given by (Sacchi and Ulrych 1995)

\[ p(S) = C_2 \exp \{-\varphi(S)\}, \]

where \( C_2 \) is a constant. Substituting equations (3) and (4) into equation (2) yields

\[ p(S|D) \propto p(D|S)p(S) = C_1C_2 \exp \left[ -\left( \frac{|D - S|^2}{2\sigma^2_n} + \varphi(S) \right) \right]. \]

After taking the negative logarithm from both sides of equation (5), we have

\[ -\log p(S|D) \propto -\log (C_1C_2) + \left( \frac{|D - S|^2}{2\sigma^2_n} + \varphi(S) \right). \]

Since the logarithm function is monotonic, the maximization of \( p(S|D) \) is identical to the minimization of \( -\log p(S|D) \). Thus, we wish to minimize the objective function that establishes a tradeoff between data misfit and random noise reduction

\[ O(S) = \frac{1}{2} ||D - S||^2_F + \varphi(S), \]

where \( \lambda = \frac{1}{2\sigma^2_n} \) is the trade-off parameter. If \( \varphi(S) = 0 \), \( \hat{S} = S \). There is no noise reduction. Thus, in order to attenuate random noise, \( \varphi(S) \) must be a reasonable function.

2. Theory

2.1. Noise reduction framework based on Bayesian inversion

Assuming that the observed seismic data \( D \) are the simple superposition of the signal \( S \) with random noise \( N \):

\[ D = S + N. \]  

According to Bayes’ rule (Bayes 1763), the conditional probability of the signal given the observed data, also called posterior distribution, can be written as

\[ p(S|D) = \frac{p(D|S)p(S)}{p(D)} \propto p(D|S)p(S), \]  

where \( p(D|S) \) is the conditional probability of the observed data given the signal, also called likelihood function; \( p(S) \) is the prior probability of the signal, and \( p(D) \) is the prior probability of the observed data. The symbol \( \propto \) denotes that \( p(S|D) \) is proportional to \( p(D|S)p(S) \) since \( p(D) = \int p(D|S)p(S)dS \) is a constant (Ulrych et al. 2001).

Assuming that \( N \) can be approximated as uncorrelated zero-mean Gaussian noise with unknown variance \( \sigma^2 \), the likelihood function in equation (2) can be expressed as (Theune et al. 2010)

\[ p(D|S) = C_1 \exp \left[ -\frac{|D - S|^2}{2\sigma^2_n} \right], \]  

where \( C_1 \) is a constant; \( |D - S|^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} (d_{ij} - s_{ij})^2 \), \(|\bullet||\bullet|\) denotes the Frobenius norm, \( d_{ij} \) and \( s_{ij} \) are the elements of matrices \( D \) and \( S \), respectively; \( m \) and \( n \) are time sample number and trace number, respectively. The magnitude of \( p(D|S) \) is used to measure the misfit between the observed data and the noise reduction data.

In a word, the idea of Bayesian inversion filtering is to find an estimate \( \hat{S} \) of the signal \( S \) by maximizing the posterior distribution \( p(S|D) \). The process of inversion is the process of noise reduction. However, before the \( \hat{S} \) is estimated, it is crucial to obtain a priori information about \( S \).

In this paper, the prior information is assumed as a function \( \varphi(S) \); thus, the prior distribution of \( S \) can be given by (Sacchi and Ulrych 1995)

\[ p(S) = C_2 \exp \{-\varphi(S)\}, \]  

where \( C_2 \) is a constant. Substituting equations (3) and (4) into equation (2) yields

\[ p(S|D) \propto p(D|S)p(S) = C_1C_2 \exp \left[ -\left( \frac{|D - S|^2}{2\sigma^2_n} + \varphi(S) \right) \right]. \]  

After taking the negative logarithm from both sides of equation (5), we have

\[ -\log p(S|D) \propto -\log (C_1C_2) + \left( \frac{|D - S|^2}{2\sigma^2_n} + \varphi(S) \right). \]  

Since the logarithm function is monotonic, the maximization of \( p(S|D) \) is identical to the minimization of \( -\log p(S|D) \). Thus, we wish to minimize the objective function that establishes a tradeoff between data misfit and random noise reduction

\[ O(S) = \frac{1}{2} ||D - S||^2_F + \varphi(S), \]  

where \( \lambda = \frac{1}{2\sigma^2_n} \) is the trade-off parameter. If \( \varphi(S) = 0 \), \( \hat{S} = S \). There is no noise reduction. Thus, in order to attenuate random noise, \( \varphi(S) \) must be a reasonable function.

2.2. Prior information

The prior information about the seismic signal \( S \) should be independent of the observed data \( D \). It can be derived from geological knowledge, statistical knowledge or assumptions.

In general, seismic signal \( S \) is smooth because the high frequency components are absorbed by the layers. That is to say, \( S \) always displays less oscillation. In contrast, random noise oscillates strongly. From the viewpoint of probability, the larger the noise energy in \( D \), the stronger \( D \) oscillates. Consequently, the \( L_1 \) norm of the first derivative of the data, namely TV, can be used to measure the strength of oscillation, and could be considered as a priori information.
Here, two kinds of TV are introduced. One is in the time direction named TTV, defined as the sum of absolute value of every element in time difference matrix \((\nabla S)_t\), and the other is in the spatial direction named STV, defined as the sum of absolute value of every element in space difference matrix \((\nabla S)_x\). They are expressed as

\[
TTV = \sum_{i=1}^{m} \sum_{j=1}^{n} |((\nabla S)_t)_{ij}|, \tag{8}
\]

\[
STV = \sum_{i=1}^{m} \sum_{j=1}^{n} |((\nabla S)_x)_{ij}|, \tag{9}
\]

where

\[
(\nabla S)_t = \begin{bmatrix}
-1 & 1 & 0 & \cdots & 0 \\
0 & -1 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & -1 & 1 \\
0 & \cdots & 0 & \cdots & 0
\end{bmatrix}_{m \times m}
\]

and

\[
(\nabla S)_x = \begin{bmatrix}
-1 & 1 & 0 & \cdots & 0 \\
0 & -1 & 1 & \cdots & 0 \\
0 & \cdots & 0 & -1 & 1 \\
0 & \cdots & 0 & \cdots & 0
\end{bmatrix}_{n \times n}
\]

Like the idea behind TV denoising in image processing (Rudin et al. 1992), \(\varphi(S)\) is formulated as

\[
\varphi(S) = \sum_{i=1}^{m} \sum_{j=1}^{n} |V_{ij}|, \tag{10}
\]

where \(V = \alpha (\nabla S)_t + i\beta (\nabla S)_x\) is a complex matrix; \(\alpha \geq 0\) and \(\beta \geq 0\) are two weights; \(i = \sqrt{-1}\) is the imaginary unit. If \(\alpha = 0\), \(\varphi(S)\) is STV. If \(\beta = 0\), \(\varphi(S)\) is TTV. If \(\alpha = \beta = 1\), \(\varphi(S)\) is the conventional TV. If \(\alpha > \beta\), TTV is predominant; otherwise, STV is predominant.

2.3. Algorithm

By substituting equation (10) into equation (7), we obtain the objective function

\[
\varphi(S) = \frac{\lambda}{2} ||S - D||_2^2 + \sum_{i=1}^{m} \sum_{j=1}^{n} |V_{ij}|, \tag{11}
\]

where the first term is a fidelity term used to preserve the structure of seismic data and lower distortion. The second term is the statistical prior information used for noise reduction. Equation (11) is a nonlinear coupled inversion problem including the \(L_1\) and \(L_2\) norms. For this problem some fast methods, such as the nonlinear primal-dual method (Chan et al. 1999), splitting-and-penalty method (Wang et al. 2007) and split Bregman method (Goldstein and Osher 2009), were proposed to solve it. In this paper, the split Bregman method is adopted, whose main idea is to introduce a new variable to split the \(L_1\) and \(L_2\) norms.

For simplicity, we use \(||V||_1\) as the representation of \(\sum_{i=1}^{m} \sum_{j=1}^{n} |V_{ij}|\). Then the objective function is

\[
O(S) = \frac{\lambda}{2} ||S - D||_2^2 + ||V||_1 \\
\text{subject to } V = \alpha (\nabla S)_t + i\beta (\nabla S)_x. \tag{12}
\]

In fact, the Bregman iteration method solves equation (12) by splitting it into two sub-problems. One is to minimize the fidelity term \(\frac{\lambda}{2} ||S - D||_2^2\) with respect to \(S\), and the other is to minimize the TV term \(||V||_1\) with respect to \(V\). The
minimization of equation (12) is equivalent to the minimization of the following 'split Bregman' formulation:

\[
O(S, V, B) = \frac{\lambda}{2}||S - D||^2_F + \frac{\eta}{2}||V||_1
+ \frac{\eta}{2}||V - (\alpha (VS) + i\beta (VS)) - B||^2_F,
\]

where \( B = B_1 + iB_2 \) is a complex matrix variable; \( B_1 \) and \( B_2 \) denote the real part and imaginary part matrices of \( B \), respectively; \( \eta \) is a positive constant and the third term \( \frac{\eta}{2}||V - (\alpha (VS) + i\beta (VS)) - B||^2_F \) stands for a quadratic penalty term.

Equation (13) can be solved by the following three steps in an iterative manner:

\[
\begin{align*}
S^{k+1} &= \arg \min_S \frac{\lambda}{2}||S - D||^2_F + \frac{\eta}{2}||V^k||_1 \\
&\quad - (\alpha (VS)_1 + i\beta (VS)_1)_1 - B^k||^2_F \quad (a) \\
V^{k+1} &= \arg \min_V ||V^k||_1 + \frac{\eta}{2}||V^k||_2 \\
&\quad - (\alpha (VS^{k+1}) + i\beta (VS^{k+1}))_1 - B^{k+1}||^2_F \quad (b) \\
B^{k+1} &= B^k + (\alpha (VS^{k+1}) + i\beta (VS^{k+1}))_1 - V^{k+1} \quad (c)
\end{align*}
\]

where equations \((14(a)–(c))\) are three independent problems. Equations \((14(a)\) and \((b)\)) are solved by the Gauss–Seidel method (Tavakoli and Davami 2007) and the shrinkage method (Goldstein and Osher 2009), respectively. \( B \) in equation \((14(c))\) is updated by substituting the solutions of equations \((14(a)\) and \((b)\)) into it. Iterate the three equations in sequence until \( ||S^{k+1} - S^k||^2_F < \varepsilon \), where \( \varepsilon \) is a small positive constant. The final \( S \) is just the noise reduction result \( \hat{S} \).

The whole algorithm is summarized as follows:

1. Set parameters \( \lambda, \eta, \alpha, \beta \) and \( \varepsilon \);
2. Initialize \( S^0, V^0 \) and \( B^0 \);
3. Update \( S \) using equation \((14(a))\);
4. Update \( V \) using equation \((14(b))\);
5. Update \( B \) using equation \((14(c))\);
6. Calculate \( ||S^{k+1} - S^k||^2_F \). If \( ||S^{k+1} - S^k||^2_F < \varepsilon \), go to step (3); otherwise stop and output \( \hat{S} = S^{k+1} \).

In all the following examples, we set \( S^0 = D, V^0 = 0 \), \( B^0 = 0 \), and \( \eta = 8 \) by experience.

3. Examples

3.1. Numerical simulation

In order to test the performance of Bayesian inversion filtering, we make noise-free synthetic seismic data (figure 1(a)), consisting of 50 traces with a sample interval of 1 ms. This profile contains some interesting features: two curved events with different curvatures (A), an isolated event (B), a broken event (C), a lateral frequency varying event (F), a lateral amplitude varying event (G), and a lateral phase varying event (H). By adding 50% (NSR = 0.5, defined as the ratio of noise energy to signal energy) Gaussian noise to the noise-free profile, we generate a noisy profile (figure 1(b)).

We tried to use the prior information with different weights to control the noise reduction, and define the relative mean square error (RMSE) as \( ||\hat{S} - S||^2_F/||S||^2_F \). The three curves in figure 2 are RMSE-\( \lambda \) curves after filtering by using TTV norm with \( \beta = 0, \alpha = 1 \), STV norm with \( \alpha = 0, \beta = 1 \)

![Figure 2. RMSE variations versus \( \lambda \) by using TTV, STV and TV as the prior information. Bayesian inversion filtering using TV norm as the prior information has the best noise reduction effect.](image)

![Figure 3. Noise reduction result for the profile in figure 1(b). (a) Noise reduction result, and (b) difference profile between figures 3(a) and 1(b). Bayesian inversion filtering can preserve the edges (B, C, D and F), clarify amplitude (G), phase (H) and frequency (F) variations in the spatial direction, and keep the curvatures of the curved events (A).](image)
Figure 4. Another synthetic profile used to make comparisons between Bayesian inversion filtering and the other five noise-reduction methods. (a) Noise-free synthetic data and (b) noisy data with 30% Gaussian noise. The profile contains a discontinuous event (A), a quasi-linear event (B), two conflicting events (C), and a weak-stationary event with small amplitude and phase variation in the spatial direction (D).

Figure 5. Results of different methods: (a) $f-x$ deconvolution, (b) local SVD, (c) $f-x$ Cadzow, (d) Bayesian inversion filtering, (e) median filtering, and (f) edge-preserving smoothing filtering.
and TV norm with $\alpha = \beta = 1$, respectively. As the figure shows, the parameter $\lambda$ plays a key role in Bayesian inversion filtering. When $\lambda$ is very small, such as 1 or 2, the variation norm is dominant; thus, more noise is removed. However, RMSE is larger because of the signal distortion. In addition, there are two main risks caused by small $\lambda$. One is that the peaks and troughs of the events become blocky or flat after filtering, and the other is that some non-horizontal features are lost. When $\lambda$ is quite large, such as 29 or 30, the fidelity term is dominant. Large $\lambda$ ensures higher fidelity for seismic profiles, but only removes a little noise. In general, $\lambda$ corresponding to the minimum RMSE should be considered as a reference. According to the trend of the three curves, Bayesian inversion filtering using TV norm as a priori information has the best noise reduction effect.

Figure 3(a) is the noise reduction result with $\lambda = 8$, $\alpha = \beta = 1$, and figure 3(b) is the difference profile between figures 3(a) and 1(b). As figures 3 and 1 show, Bayesian inversion filtering using TV norm as the prior information is good at suppressing noise while preserving signal. The curvature and discontinuity of the seismic events are retained with considerable fidelity.

Similarly, TV norm with $\alpha = \beta = 1$ is utilized as the prior information to control noise attenuation for the following examples.

3.2. Comparisons of noise reduction methods

To compare Bayesian inversion filtering with other noise reduction methods, such as $f_x$ deconvolution, local SVD filtering, $f_x$ Cadzow filtering, median filtering and edge-preserving smoothing filtering, we make another synthetic seismic profile (figure 4(a)), consisting of 21 traces with a sample interval of 1 ms. This profile contains a discontinuous
Figure 7. Real seismic data used to test the application potential of Bayesian inversion filtering. (a) Original seismic data, (b) noise reduction result and (c) difference profile between (b) and (a).

Figure 4(b) is the noisy profile with 30% (NSR = 0.3) Gaussian noise. Figures 5(a)–(f) are the results of $f-x$ deconvolution, local SVD filtering, $f-x$ Cadzow filtering, Bayesian inversion filtering, median filtering and edge-preserving smoothing.
Random noise reduction using Bayesian inversion

Figure 8. $f$–$k$ spectra for the three data sets in figure 7: (a) original seismic data, (b) noise reduction result and (c) difference profile.

filtering, respectively. For $f$–$x$ deconvolution, the final noise reduction result is the average of the forward and backward prediction filtering with the filter operator length 4. For the local SVD method, the two horizontal events are reconstructed by the first weighted eigenimage of the original noisy profile. Then they are subtracted from the original data, and a difference profile containing the dipping event is obtained. After that, the dipping event is flattened by shifting the difference profile, reconstructed by the first weighted eigenimage of the difference profile, and rotated to its original position. The sum of the two reconstructed horizontal events and the dipping event is the final noise reduction profile. For $f$–$x$ Cadzow, the first three weighted eigenimages are used to recover ‘clean’ seismic data. For Bayesian inversion filtering, the parameters are set as $\lambda = 25$ and $\alpha = \beta = 1$. For median filtering, the filter length is 3, but it is 1 on the boundaries. For edge-preserving smoothing, a $5 \times 5$ window is utilized. These parameters of the six methods are determined by trial and error. Figures 6(a)–(f) are the corresponding difference profiles. As figures 5 and 6 show, although $f$–$x$ deconvolution and SVD filtering are classic noise reduction methods, they do not have the desired effect for this example. However, Bayesian inversion filtering can preserve edges of discontinuous events and clarify amplitude and phase variations of weak-stationary events. Even if Bayesian inversion filtering, edge-preserving smoothing filtering and median filtering all use statistical features of seismic data to remove noise, Bayesian inversion filtering better preserves the signal. No doubt many other factors affect the results of these six methods, but we have reasons to believe that Bayesian inversion filtering has the potential to preserve the edges of discontinuous events and clarify amplitude and phase variations of non-stationary events.

3.3. Real data

To further test the performance of Bayesian inversion filtering, the method is applied to a real post-stack seismic data set (figure 7(a)), consisting of 200 traces with a sample interval of 2 ms. The data set mainly includes discontinuous events, non-stationary events, linear events, quasi-linear events, and background noise. From the figure, it is obvious that the signal is blurred by noise. In particular, the discontinuity and non-stationarity of the events are hardly discovered.

We set the parameters as $\lambda = 3$ and $\alpha = \beta = 1$. After 32 iterations, the noise reduction result (figure 7(b)) is obtained. After filtering the signal, even discontinuous and non-stationary events can be identified clearly. In addition, the edges of most events are clarified. In the difference profile (figure 7(c)), there are no obvious indications of coherent events. We also display $f$–$k$ spectra (figures 8(a)–(c)) of the three data sets in figures 7(a)–(c), respectively. As figures 8(a)–(c) show, Bayesian inversion filtering not only reduces noise, but also preserves signal with high fidelity.

4. Conclusions

As a novel random noise reduction method, Bayesian inversion filtering utilizes the minimization of TV norm as a priori information to control seismic noise reduction. Because this method does not require continuity, stationarity and linearity assumptions about seismic events, it is applicable to most seismic data sets. The method not only reduces random noise, but also clarifies the structures and physical properties of coherent events. In particular, it has the ability to preserve the edges of discontinuous events without quantitative awareness of edges’ location.

For this method, there is an important parameter $\lambda$ to balance the noise reduction and signal fidelity. The smaller the parameter $\lambda$, the better the noise attenuation, but the worse
the signal fidelity. Moreover, small $\lambda$ probably leads to two main risks after filtering. One is that the peaks and troughs of the events become blocky or flat, and the other is that some non-horizontal features are lost. However, the larger the $\lambda$, the better the signal fidelity, but the worse the noise attenuation. Specially, when $\lambda = \infty$, no noise is reduced. In fact, this parameter can be adjusted by trial and error.

For the prior information TV, different weights have different influences on filtering results. For examples in this paper, TV norm with equivalent weight is superior to TTV or STV. Detailed study about the weights is our next research topic.

The running speed of this method is fast, since only linearization technique is adopted. For our experiments, the optimal seismic noise reduction results can be achieved after about 20 or 30 iterations.

Acknowledgments

We are grateful to the anonymous reviewers for their constructive comments on this paper. This research was supported in part by National Key Basic Research Development Program (grant no 2007CB209600).

References

Abma R and Claerbout J 1995 Lateral prediction for noise attenuation by $t$--$x$ and $f$--$x$ techniques

Geophysics 60 1887–96

AlBinHassan N M, Luo Y and Al-Faraj M N 2006 3D edge-preserving smoothing and applications

Geophysics 71 P5–P11

Al-Yahya K M 1991 Application of the partial Karhunen–Loeve transform to suppress random noise in seismic sections

Geophys. Prospect. 39 77–93

Bayes T 1763 Essay towards solving a problem in the doctrine of chances (reprinted 1958 in Biometrika 45 298–315)

Bednar J B 1983 Applications of median filtering to deconvolution, pulse estimation, and statistical editing of seismic data

Geophysics 48 1599–610

Bekara M and van der Baan M 2007 Local singular value decomposition for signal enhancement of seismic data

Geophysics 72 V59–V65

Bekara M and van der Baan M 2009 Random and coherent noise attenuation by empirical mode decomposition

Geophysics 74 V89–V98


Chase M K 1992 Random noise reduction by FXY prediction filtering

Explor. Geophys. 23 51–6

Goldstein T and Osher S 2009 The split Bregman method for L1-regularized problems SIAM J Imaging Sci. 2 323–43


Jones I F and Levy S 1987 Signal-to-noise ratio enhancement in multichannel seismic data via the Karhunen–Loeve transform

Geophys. Prospect. 35 12–32


Geophysics 74 V17–V24

Oldenburg D W, Scheuer T and Levy S 1983 Recovery of the acoustic impedance from reflection seismograms

Geophysics 48 1318–37


Portniaguine O and Zhulanov M S 1999 Focusing geophysical inversion images Geophysics 64 874–87

Rudin L I, Osher S and Fatemi E 1992 Nonlinear total variation based noise removal algorithms Physica D 60 259–68

Sacchi M D 1997 Reweighting strategies in seismic deconvolution

Geophys. J. Int. 129 651–6

Sacchi M D and Urych T J 1995 High-resolution velocity gathers and offset space reconstruction Geophysics 60 1169–77

Soubaras R 1994 Signal-preserving random noise attenuation by the $f$--$x$ projection 64th Annu. Int. Meeting, SEG, Expanded Abstracts pp 1576–9


Theune U, Jensas I and Eidsvik J 2010 Analysis of prior models for a blocky inversion of seismic AVO data Geophysics 75 C25–C35

Tikhonov A N and Arsenin V Y 1977 Solution of Ill-posed Problems (Washington, DC: Winston and Sons)


Yuan S Y and Wang S X 2011 A local f--x Cadzow method for noise reduction of seismic data obtained in complex formations Petrol. Sci. 8 269–77