Deblending of simultaneous-source data using a 
structure-oriented space-varying median filter\textsuperscript{a}

\footnotesize{\textsuperscript{a}Published in Geophysical Journal International, 216, 1214-1232, (2019)}

Yangkang Chen\textsuperscript{*}, Shaohuan Zu\textsuperscript{†‡}, Yufeng Wang\textsuperscript{‡}, and Xiaohong Chen\textsuperscript{‡}

\textbf{ABSTRACT}

In seismic data processing, the median filter is usually applied along the structural direction of seismic data in order to attenuate erratic or spike-like noise. The performance of a structure-oriented median filter highly depends on the accuracy of the estimated local slope from the noisy data. When local slope contains significant error, which is usually the case for noisy data, the structure-oriented median filter will still cause severe damages to useful energy. We propose a type of structure-oriented median filter that can effectively attenuate spike-like noise even when the local slope is not accurately estimated, which we call structure-oriented space-varying median filter. A structure-oriented space-varying median filter can adaptively squeeze and stretch the window length of the median filter when applied in the locally flattened dimension of an input seismic data in order to deal with the dipping events caused by inaccurate slope estimation. We show the key difference among different types of median filters in detail and demonstrate the principle of the structure-oriented space-varying median filter method. We apply the structure-oriented space-varying median filter method to remove the spike-like blending noise arising from the simultaneous source acquisition. Synthetic and real data examples show that structure-oriented space-varying median filter can significantly improve the signal preserving performance for curving events in the seismic data. The structure-oriented space-varying median filter can also be easily embedded into an iterative deblending procedure based on the shaping regularization framework and can help obtain much improved deblending performance.

\textbf{INTRODUCTION}

Simultaneous shooting has been a long-standing research topic in the industry, aiming at boosting the acquisition efficiency by significantly shortening the acquisition time period as well as improving subsurface sampling by increased shot density (Beasley, 2008; Wapenaar et al., 2012; Abma, 2014; Amundsen et al., 2018; Zu et al., 2017a; Asgedom et al., 2017; Zhang et al., 2017; Kim et al., 2017; Wang et al., 2017; Wu and Bai, 2018b). Because of the reduced acquisition cost, it supports more frequent data
recording in time-lapse seismic monitoring. Due to the improved spatial shot sampling, it makes the high-resolution subsurface imaging and inversion more achievable. All the aforementioned benefits, however, are compromised by the intense interference noise introduced during simultaneous shooting (Doulgeris et al., 2012; Mahdad, 2012; Qu et al., 2014; Wu, 2014; Wu et al., 2015; Chen et al., 2016; Zhou et al., 2018; Bai and Wu, 2018; Zhou and Li, 2018; Zhou and Han, 2018). To deal with the strong blending noise in simultaneous source data, one can either develop new algorithms to separate the blended records so that the subsequent processing and workflows do not need to be changed, which is called deblending, or to design new workflows for migrating the blended data directly regardless of the intense interference, which is called direct imaging (Verschuur and Berkhout, 2011; Gan et al., 2016b; Zu et al., 2017b; Bai and Wu, 2017; Wu and Bai, 2018a; Bai et al., 2018b,a; Zhang et al., 2018b,a).

There are two main types of deblending approaches. The first is based on filtering, which simply treats the blending interference as noise and uses denoising algorithms to remove it (Mousavi et al., 2016; Mousavi and Langston, 2016b,a, 2017). These methods include the median filtering based approaches (Huo et al., 2012; Chen, 2015; Gan et al., 2016a; Huang et al., 2018), adaptive subtraction methods (Spitz et al., 2008; Kim et al., 2009), and dip filtering based methods (Hampson et al., 2008). The second is based on inversion. The inversion based algorithms take advantage of iterative solvers to solve the ill-posed inverse problem in deblending. Regularization strategies should be used to constrain the model to be the desired solution. A number of methods based on different types of constraints have been developed in the literature, e.g., the seislet constraint (Chen et al., 2014), the curvelet constraint (Kumar et al., 2015; Qu et al., 2016; Zu et al., 2016), the Fourier constraint (Abma et al., 2015), the rank constraint (Cheng and Sacchi, 2014; Zhou et al., 2017), the hybrid rank-sparsity constraint (Zu et al., 2017c), and the hybrid mask-sparsity constraint (Zhou, 2017). Here “seislet constrint” denote “seislet-domain sparsity constraint”. While the filtering based methods are more computational efficient and easier to implement, the inversion-based approaches usually lead to better deblending performance (van Borselen et al., 2012; Zu et al., 2016).

To date, there are also many reported trials for directly migrating the simultaneous source data. Because of the strong interference noise, the migrated images might contain distinct artifacts if no extra implementation step is introduced into the migration process. Berkhout et al. (2012) discussed the illumination aspect of direct imaging of simultaneous source data. Xue et al. (2016) proposed a new direct imaging method that is based on least-squares reverse time migration and structural enhancing operator (Liu et al., 2010). Different from the preconditioning constraint that was used in Dai and Schuster (2011), Xue et al. (2016) developed a new iterative solver to solve the least-squares reverse time migration related inverse problem, which is based on the shaping regularization framework. To tackle the smearing problem of structural smoothing operator in the least-squares reverse time migration method proposed by Xue et al. (2016), Chen et al. (2017) developed a local singular spectrum analysis (SSA) constraint to preserve the edges and discontinuities on the image. While those imaging methods are promising to make the future imaging framework
more flexible, most reported successes are stuck on the synthetic test. A lot of researchers are still working hard to push the direct imaging technique from simulation to practical applications (Dutta, 2017; Li et al., 2018).

In this paper, we are aiming at introducing a novel median filtering method for either separating the simultaneous sources in a fast one-step way, or iteratively improving the traditional shaping regularized deblending approaches by constructing a more powerful shaping operator. Since the new method is by nature a type of median filter, it is worth mentioning some existing methods that use median filtering for deblending. Huo et al. (2012) proposed a multi-dimensional vector median filter for separating simultaneous-source data. A dip scanning step is first performed so that a vector median filter can be applied along the structural direction. Chen (2015) developed a space-varying median filter to deal with curved events in seismic data. The filter length varies spatially so that noise is removed using a long filter length and the signal is preserved using a relatively smaller filter length. Gan et al. (2016a) proposed a structure-oriented median filter to maximize the effectiveness of a median filter when processing a complicated dataset. Both methods in either Huo et al. (2012) or Gan et al. (2016a) can be considered as structure-oriented filtering, which requires an sufficiently-accurate slope estimation. However, when the slope is not properly estimated, e.g., in the case of strong ambient noise and blending noise, the structural filtering strategy would fail because those filters will tend to damage signal energy.

We investigate the reason why the structural filtering fails when the slope is not accurate enough. The principle of the structural filtering is to create a local window along the structural direction, or in other words, to create a locally flattened gather to apply the 1D median or mean filter. When the slope is accurate, e.g., using the plane-wave destruction algorithm (Fomel, 2002), the local gather is exactly flattened. However, when the slope is not precise enough, the local gather still contains curved events. In this case, we take advantage of the spatially varying median filter to deal with the slightly curved events in the flattened dimension. The proposed method is thus capable of preserving more useful energy than the traditional structural filtering method.

The paper is organized as follows, we first introduce the basics of median filtering and structure-oriented filtering. Then, we focus on introducing the proposed structure-oriented space-varying median filtering method. We also introduce the way we imbed the proposed filtering method in the shaping regularized iterative deblending framework. Next, we use several synthetic and real data examples to show the potential of the proposed method, either in obtaining a fast simultaneous source separation, or in improving the deblending performance in an iterative fashion. Finally, we draw some key conclusions at the end of the paper.
STRUCTURAL-ORIENTED SPACE-VARYING MEDIAN FILTER (SOSVMF)

Median filter

A median filter can be used to obtain successful performance in rejecting impulsive (spiky) noise, which has anomalous value compared with neighboring values. Compared with the commonly known smoothing filter, the median filter enjoys several advantages such as preserving edges, being less sensitive to anomalous values and being able to be applied repeatedly. Given a 2D signal, e.g., the waveforms recorded by geophone arrays, a 1D median filter is applied along the spatial direction and each row of the waveform matrix is filtered one by one. Mathematically, a median filter can be expressed using the following formula (Chen, 2015) when applied to a 2D signal:

$$\min_{u_m \in U_{i,j}} \sum_{l=1}^{L} \|u_m - u_l\|_1,$$

The solution of minimization problem 1 is the output value for location $x_{i,j}$. $U_{i,j} = \{u_1, u_2, \ldots, u_L\}$ denotes the 1D filtering window. $i, j$ are the position indices in a 2D profile. $l$ and $m$ are both indices in the filtering window. Both $u_l$ and $u_m$ are temporary variables. $L$ is the length of the filtering window. Simply speaking, for each point in the 2D data, the median filter substitutes it with the median of neighbor points included by the window with size $L$. The process of $\min_{u_m \sum_{l=1}^{L} \|u_m - u_l\|_1$ corresponds to finding the median in the window constructed by $\{u_1, u_2, u_3, \ldots, u_L\}$.

Median filter along the structural direction

Median filtering has a remarkable capability for attenuating spike-like noise. However, the median filter might still cause some damage to useful components when the data is highly non-stationary in the space direction. Considering that the recorded waveforms may have structural patterns, if the median filter is applied along the structural direction, the damage can be minimized. We utilize a way that is used in reflection seismic data processing to flatten the recorded waveform data in a local manner. The flattening process is based on the predictability between neighbor traces. In the following context, we will introduce in details the flattening process. Then, the median filter can be applied in the flattened dimension to maximize its effectiveness.

Flattening the waveforms is basically a data mapping process. We perform the data mapping by a recursive prediction strategy. We recall that the target of the data mapping is to create a flattened gather in each local window centered at each trace. Let the width of the window to be $2N + 1$, then the central trace $d_i$ can be predicted from its $j$th ($j \leq N$) right trace $d_{i+j}$ in such a recursive manner,

$$d_i = P_{i+1,i} \cdots P_{i+j-1,i+j-2} P_{i+j,i+j-1} d_{i+j},$$
where $P_{i,j}$ denotes a prediction operator to predict trace $d_i$ from trace $d_j$. For predicting trace $d_i$ from its left trace, we use a similar recursive formula,

$$d_i = P_{i-1,i} \cdot \cdot \cdot P_{i-j+1,i-j+2} P_{i-j,i-j+1} d_{i-j}.$$  \hfill (3)

Figure 1 shows a sketch of the trace prediction process for $N = 1$. Each trace is associated with a local flattened window. The prediction process transforms the curving event in the local window to a flattened event. More details regarding the trace prediction and the flattening operator can be found in Liu et al. (2010). Figure 2 shows a demonstration of the structure-oriented processing. Figure 2a is the data with clear structural patterns. Figure 2b shows the locally flattened data in the third dimension by predicting each trace from its left and right neighboring traces. The median filter is then applied along the third dimension (prediction axis) to remove erratic or spike-like noise. The 2D section defined by the axes “Time (ms)” and “Position (m)” indicates the original profile so it contains curving events. It is the same as the profile shown in Figure 2a. The 2D section defined by the axes “Time (ms)” and “Prediction (trace)” indicates the flattened data. The 2D section defined by the axes “Prediction (trace)” and “Position (m)” indicates the constant Time slice. The data shown in Figure 2 can be considered as a common offset data or a migrated image. The curving events denote the curving subsurface reflectors. It is worth noting that the flattening strategy also works for crossing events, but the median filters may damage the useful energy due to the anomalous amplitude when crossing events exist.

### Median filter along the structural direction with a variable window length

When the data is not flattened well, e.g., when the local slope of the waveform profile is not estimated accurately due to the strong spike-like noise (e.g., using the plane-wave destruction algorithm (Fomel, 2002)), the traditional median filter with a constant filter length will still cause some damages. Figures 3b and 3e show two estimated slope sections from the clean data and noisy data respectively. It is obvious that the two local slope maps have distinct differences. The color indicates the slope value. Warm (red) color indicates high positive slope and cool (blue) color indicates high negative slope. Positive slope and negative slope mean opposite dipping directions. Figures 3c and 3f show the flattened data using true and estimated local slopes. For a better comparison, we select three flattened gathers for different space locations from Figures 3c and 3f, respectively, and show them in Figure 4. When using the true local slope, we can clearly observe that the data are exactly flattened and events in the prediction dimension are exactly horizontal. However, when we use the inaccurate local slope that is estimated from the observed noisy data, the prediction based flattening operator cannot obtain a perfect performance. The events in the prediction dimension are not exactly horizontal, but are more or less curved. Note that in Figures 4c and 4f, the first blank trace is caused by the boundary trace prediction.
For example, if a seismic section contains 98 traces, then the predicted 90th trace from its 9th right neighbor trace is a zero trace.

The space-varying median filter can be used in such case where the events are not perfectly flattened. For space-varying median filter, $L$ becomes $L_{i,j}$, varying with respect to location $x_{i,j}$. The new filtering expression is,

$$
\min_{v_m \in U_{i,j}} \sum_{l=1}^{L_{i,j}} \|v_m - v_l\|_1.
$$

The solution $\hat{v}_{i,j}$ is the output value for location $x_{i,j}$ after applying a space-varying median filter, $U_{i,j} = \{v_1, v_2, \ldots, v_{L_{i,j}}\}$ denotes the local filtering window in the case of space-varying median filter. The filter length $L_{i,j}$ can be adaptively chosen according to the signal reliability $s_{i,j}^L$ (Chen, 2015). Signal reliability is a local attribute to indicate how reliable a data point is considered as a signal point. The signal reliability can be defined as the local similarity (Fomel, 2007b) between the initially filtered data $u_{i,j}^L$ with a window length $L$ and the original data $u_{i,j}$,

$$
s_{i,j}^L = S(u_{i,j}^L, u_{i,j}).
$$

Here, $S(x,y)$ denotes the local similarity between $x$ and $y$. Here, the operator $S$ denotes the process of calculating the local similarity between two data sets: $u_{i,j}^L$.
Figure 2: A simple demonstration of structure-oriented processing. (a) Input data. (b) Locally flattened domain. The 2D section defined by the axes “Time (ms)” and “Position (m)” indicates the original profile so it contains curving events. The 2D section defined by the axes “Time (ms)” and “Prediction (trace)” indicates the flattened data. The 2D section defined by the axes “Prediction (trace)” and “Position (m)” indicates the constant Time slice.
and $u_{i,j}$. The selection of the initial filter length requires some human experience and trials. The initially filtered data means the filtered data using a conventional median filtering method. Later, we will illustrate how the signal reliability/local similarity looks like. The criterion for defining the filter length is expressed as follows:

$$L_{i,j} = \begin{cases} L + l_1, & 0 \leq |s_{i,j}^L| \leq s_1 \\ L + l_2, & s_1 < |s_{i,j}^L| < s_2 \\ L, & s_2 \leq |s_{i,j}^L| \leq s_3 \\ L - l_3, & s_3 < |s_{i,j}^L| < s_4 \\ L - l_4, & s_4 \leq |s_{i,j}^L| \leq s_{max} \end{cases},$$  

(6)

where $l_1$, $l_2$, $l_3$, and $l_4$ are predefined parameters corresponding to the increments or decrements of the length of filter window and are generally chosen as 4, 2, 2, and 4 in default, respectively. $s_1$, $s_2$, $s_3$, and $s_4$ are four thresholds, and are empirically chosen as $s_1 = 0.15s_{max}$, $s_2 = 0.25s_{max}$, $s_3 = 0.75s_{max}$, and $s_4 = 0.85s_{max}$. $s_{max}$ denotes the maximum value of the similarity map. The appendix provides a brief review of the calculation of local similarity.

To facilitate an easier understanding of the space-varying median filtering, more specifically the relation among initially filtered data, local similarity, variable filter length, we draw an illustrative picture shown in Figure 5. Figure 5(a) denotes the noisy data. Figure 5(b) shows the map of constant filter length. Figures 5(a) and 5(b) are used to generate the initially filtered data shown in 5(c). Using equation 5, from the original data and the initially filtered data, we can calculate the local similarity map. Utilizing equation 6, we can calculated the map of variable filter length. With the map of variable filter length, we can obtain the finally filtered data shown in Figure 5(f). From this illustration, it is clear that the local similarity clearly detect the main distribution of the useful reflection energy. Those areas with better signal preservation after the initial median filtering are revealed by higher local similarity. The map of variable filter length clearly shows that those noise areas are dealt with very longer filter length while those signal areas correspond to relatively shorter (safer) filter length.

Applying the space-varying median filter in the imperfectly flattened gathers, we can preserve much useful energy that remains curved. Figure 6 shows a comparison between median filter and space-varying median filter in attenuating the spike-like noise when applied to the flattened gathers as shown in Figure 4 from (d) to (f). It is obvious that space-varying median filter helps preserve the energy much better than the median filter. To further confirm this conclusion, we also plot the removed noise sections in Figure 7. It is more salient that the median filter damages more useful energy than the space-varying median filter.
Figure 3: (a) Clean data. (b) Slope estimated from (a). (c) Flattened domain of (a). (d) Noisy data. (e) Slope estimated from (d). (f) Flattened domain of (d).
Figure 4: A comparison of “flattened” local windows (predicted dimension) for different space locations. (a)-(c) Flattened gathers for traces 9, 50, 90, using the accurate slope. (d)-(f) “Flattened” gathers for traces 9, 50, 90, using the estimated slope.
Figure 5: An illustration of the space-varying median filter. (a) Raw noisy data. (b) A map of constant filter length. (c) Initially filtered data using the constant filter length shown in (b). (d) Calculated local similarity map from (a) and (c) using equation 5. (e) Calculated map of variable filter length using equation 6. (f) Filtered data using the variable filter length shown in (e).
Figure 6: A denoising comparison of “flattened” local windows (predicted dimension) using estimated slope for different space locations. (a)-(c) Denoised gathers for traces 9, 50, 90, using the median filter. (d)-(f) Denoised gathers for traces 9, 50, 90, using the space-varying median filter. Note that although (c) is cleaner, the curving signals are serious damaged.
Figure 7: A comparison of removed noise. (a)-(c) Removed noise for traces 9, 50, 90, using the median filter. (d)-(f) Removed noise for traces 9, 50, 90, using the space-varying median filter. Note that the median filter removes more coherent energy.
Including the structure-oriented space-varying median filter into an inversion framework

The blending process can be summarized as the following equation:

\[ d = \Gamma m, \]  

(7)

where \( d \) is the blended data, \( \Gamma \) is the blending operator, and \( m \) is the unblended data. The formulation of \( \Gamma \) has been introduced in Mahdad (2012) in detail. When considered in time domain, \( \Gamma \) corresponds to blending different receiver signals into one channel according to the shot schedules of different shots. Deblending amounts to inverting equation 7 and recovering \( m \) from \( d \).

Because of the ill-posed property of this problem, all inversion methods require some constraints. Chen et al. (2014) proposed a general iterative deblending framework via shaping regularization Fomel (2007a). The iterative deblending is expressed as:

\[ m_{n+1} = C[m_n + B[d - \Gamma m_n]], \]  

(8)

where \( C \) is the shaping operator, which provides some constraints on the model, and \( B \) is the backward operator, which approximates the inverse of \( \Gamma \). The shaping regularization framework offers us much freedom in constraining an under-determined problem by allowing different types of constraints. In this paper, the backward operator is simply chosen as \( \lambda \Gamma^* \), where \( \lambda \) is a scale coefficient closely related with the blending fold, and \( \Gamma^* \) stands for the adjoint operator of \( \Gamma \) (or the pseudo-deblending operator). For example, \( \lambda \) can be optimally chosen as 1/2 in a two-source dithering configuration (Chen et al., 2014; Mahdad, 2012).

Given a sparsity or rank constraint based shaping operator, we are proposing the hybrid shaping operator, which is constructed by cascading the structure-oriented space-varying median filter and the rank constraint based shaping operators. During the iterations, we need to vary the filter length \( L \) with each iteration, e.g., starting from a relatively large \( L \) and decreasing it to 1. It is worth noting that the median filter based methods can only be effective in filtering but not effective in inversion if not combined with other constraints. It is because the median filter is very powerful in rejecting impulsive noise but is not convenient to control for signal preservation.

**EXAMPLES**

We first test the structure-oriented space-varying median filter method on a simple synthetic example. Figure 8a shows a simple hyperbolic synthetic example. We create the data using the Kirchhoff 2-D/2.5-D modeling method with analytical Green's functions, introduced in Haddon and Buchen (1981). Figure 8b shows the noisy data with some blending noise. Figure 8c shows the slope estimated from the clean data and Figure 8d shows the slope estimated from the noisy data. The strong noise makes the slope estimation inaccurate. Figures 9a and 9b show a comparison
between the flattened unblended data using accurate slope and flattened blended data using inaccurate local slope, respectively. From Figure 9a, we can observe that the data is flattened well locally using the accurate local estimation (from the unblended data). From Figure 9b, it is obvious that the events in the flattened dimension are not horizontal, which indicates an imperfect local flattening. Figures 9c and 9d show a comparison between the filtered data using the median filter and space-varying median filter, respectively. Both median filter and space-varying median filter are applied along the flattened dimension (along the “prediction” axis). It is clear that some energy is damaged using the median filter and the signals are preserved well using the space-varying median filter method, as indicated by the white arrow in Figure 9. For a clearer comparison, we show the zoomed data of Figure 9 in Figure 10. Figure 11 shows a deblending comparison between the median filter, space-varying median filter, structure-oriented median filter, and structure-oriented space-varying median filter methods, respectively. From Figure 11, we can see that the traditional median filter results in the largest damages, which is followed by space-varying median filter. Structure-oriented median filter preserves most signals but still causes some damages, as indicated by the white arrow. Among the four results, the structure-oriented space-varying median filter result is the best in preserving dipping energy and rejecting spike-like blending noise. A zoomed comparison of Figure 11 is shown in Figure 11. Figure 13 shows the removed blending noise corresponding to the four aforementioned methods, which confirms the observation from Figure 11 that median filter, space-varying median filter, and structure-oriented median filter all result in more or less signal leakage (Chen and Fomel, 2015) to the noise sections. In this example, we use the signal-to-noise ratio (SNR) to quantitatively compare the deblending performance using different methods:

$$SNR = 10 \log_{10} \frac{\|m_{\text{clean}}\|^2}{\|m_{\text{clean}} - m\|^2},$$

where $m_{\text{clean}}$ denotes the exact solution, or the unblended data. $m$ denotes the deblended data. Using this metric, we calculate the SNRs of the blended data, deblended data using median filter, space-varying median filter, structure-oriented median filter, and structure-oriented space-varying median filter methods as 1.62 dB, 5.46 dB, 8.63 dB, 12.46 dB, and 16.55 dB, respectively. The structure-oriented space-varying median filter method obtains the highest SNR.

The next example is a field data test. Figure 14a shows the a common receiver gather that contains blending noise. The data is a benchmark testing data shared by BGP, China. Figure 14b shows the local slope map calculated from the raw data using the plane-wave destruction algorithm. The deblended results using four methods are shown in Figure 15. The removed blending noise corresponding to the four methods are shown in Figure 16. Combining Figures 15 and 16, it is easy to conclude that both median filter and space-varying median filter methods cause a large amount of damaged energy, the structure-oriented median filter method causes relatively less damaged energy around the large-offset areas, and the structure-oriented space-varying median filter method causes negligible damaged energy. Note that the
Figure 8: (a) Clean data. (b) Noisy data with spike-like blending noise. (c) Estimated slope from the clean data. (d) Estimated slope from the noisy data.
Figure 9: (a) Flattened dimension of the clean data using the accurate slope. (b) Flattened dimension of the noisy data using the inaccurate slope. (c) Filtered data using median filter in the flattened domain. (d) Filtered data using space-varying median filter in the flattened domain.
Figure 10: Zoomed comparison of Figure 9. (a) Flattened dimension of the clean data using the accurate slope. (b) Flattened dimension of the noisy data using the inaccurate slope. (c) Filtered data using median filter in the flattened domain. (b) Filtered data using space-varying median filter in the flattened domain.
Figure 11: (a) Deblended data using median filter. (b) Deblended data using space-varying median filter. (c) Deblended data using structure-oriented median filter. (d) Deblended data using structure-oriented space-varying median filter. Note that since all median filters are only applied once, so there are several spikes left. The one-step processing is efficient but can be improved through multiple iterations (e.g., in iterative deblending).
Figure 12: Zoomed comparison of Figure 11. (a) Deblended data using median filter. (b) Deblended data using space-varying median filter. (c) Deblended data using structure-oriented median filter. (d) Deblended data using structure-oriented space-varying median filter.
Figure 13: (a)-(d) Removed blending noise sections corresponding to Figures 11 (a)-(c), respectively.
Figure 14: (a) Simultaneous source data. (b) Slope estimated from the raw data.
Figure 15: (a) Deblended data using median filter. (b) Deblended data using space-varying median filter. (c) Deblended data using structure-oriented median filter. (d) Deblended data using structure-oriented space-varying median filter.
Figure 16: (a) Removed blending noise using median filter. (b) Removed blending noise using space-varying median filter. (c) Removed blending noise using structure-oriented median filter. (d) Removed blending noise using structure-oriented space-varying median filter.
Figure 17: A comparison of local similarity between deblended data and removed blending noise. (a) Local similarity using median filter. (b) Local similarity using space-varying median filter. (c) Local similarity using structure-oriented median filter. (d) Local similarity using structure-oriented space-varying median filter.
signal damages are highlighted by the white arrows in Figure 16. In this example, the exact solution is unknown, thus we cannot use the quantitative metric that is used previously to compare the performance quantitatively. But we use the local similarity to measure the signal damage. The local similarity between the denoised (deblended) and removed noise sections are calculated to detect the signal damage for a denoising (deblending) algorithm, as proposed in Chen and Fomel (2015). High anomalies in the local similarity map indicates a more severe signal damage. The most distinct advantage of the local similarity measure is that it can measure the signal damage in a local and vivid way. Figure 17 show the four local similarity maps for the four methods. The comparison of local similarity shows vividly that all median filter, space-varying median filter, and structure-oriented median filter methods cause a large amount of damaged signals in the noise and the signal damage using the structure-oriented space-varying median filter method is the least. It is worth mentioning that some shallow reflections are not well predicted. This is likely because of the curvature of the events and the linearity of the process.

Next, we simulate a multi-receivers 2D simultaneous source separation test. In this test, we want to test how the structure-oriented space-varying median filter method performs when embedded into an iterative deblending framework. Figure 18a shows the acquisition geometry of the 2D test. The red dots denote 120 evenly spaced shot locations on the surface. The blue inverted triangles denote 120 evenly spaced receivers. Note that in this acquisition geometry, the simultaneous shooting and traditional shooting only differ in the shot schedule (Abma, 2014). Figure 18b shows the shot schedules for the traditional and the faster simultaneous source shooting. It is worth mentioning here that although there is only one source in this test, the faster shooting scheduling results in an overlapping subsurface reflection responses in time, which is still considered as simultaneous shooting (Abma et al., 2015). From the shot schedule, it is obvious that the shooting time for the simultaneous source acquisition is almost half of the traditional acquisition and thus saves about half of the traditional acquisition cost. In this example, we use a relatively realistic field dataset to mimic the unblended data (m in equation 7), as shown in Figure 19a. The data is a marine data set provided by BGP China. The size of the data is $1500 \times 120 \times 120$ (in time-shot-receiver domain). Figure 19b shows the blended data (d in equation 7). It is clear that the blending noise appears to be spatially incoherent in the common receiver domain (along the shot axis) and spatially coherent in the common shot domain (along the receiver axis). Figure 20a shows iteratively deblended data using the rank-constrained shaping regularization framework (Chen et al., 2014; Zu et al., 2017c). The iterative deblending is carried out in each common receiver gather based on the coherency difference between useful signals and interfering noise. For the rank-constrained inversion, we use local windows with the size of $200 \times 20$ (time $\times$ space). A constant rank strategy is used for each window with rank $= 2$. We obtain the result shown in Figure 20a with 40 iterations. Comparing the rank-constrained deblended data with the blended data, it is clear that most blending noise is removed. However, there is still much residual noise. By cascading the structure-oriented space-varying median filter and the rank filter based on the iterative shaping
regularization framework, as expressed by equations 8, we obtain a much improved deblended result, as shown in Figure 20b. During the inversion, we linearly decrease the filter length from 11 points to 0. A comparison between the two deblended data in Figure 20 confirms the effectiveness of the structure-oriented space-varying median filter method in improving iterative deblending. Figure 21 shows the deblending error for the traditional method and the proposed method. The deblending error means the difference between the unblended data (exact solution) and the deblended data. From Figure 21, it is obvious that the deblending error for the proposed method is negligible while the error for the rank-reduction method is very large. Figure 22 shows the removed noise corresponding to the two methods, note that the interfering shots are coherent in the second spatial dimension. For evaluating the signal damages of this example, we calculate the local similarity between the deblended data and the noise and show them in Figure 23, where much larger local similarity anomalies are shown in the similarity map corresponding the traditional method. To compare the amplitude preservation in detail, we plot a trace-by-trace comparison in Figure 24a. The black line is from the unblended data. The red line is from the blended data. The magenta line corresponds to the rank-constrained method. The green line corresponds to the proposed method. It is salient that the black and green lines are very close to each other, thus the deblending error using the proposed approach is much less than the traditional method for most parts. In order to compare the hybrid method with the structure-oriented space-varying median filter and the hybrid with the structure-oriented median filter, we also have an inversion test using the structure-oriented median filter. The deblended data is shown in Figure 25. It is obvious that the hybrid method with the structure-oriented median filter causes much higher deblending error.

**DISCUSSION**

Two main research tracks of deblending are the filtering method, which is fast and easy to control, and the inversion method, which is more computationally demanding but can obtain better result. Two main existing challenges for deblending are 1) damaging signals for the filtering method and 2) lacking extra constraints during inversion in order to deal with high blending fold (Mahdad, 2012). The paper mainly discusses the filtering based deblending methods, so let us first focus on the drawbacks of the filtering method. To date, the filtering based deblending methods are mainly based on some types of median filtering, e.g., Huo et al. (2012), Chen (2015), Gan et al. (2016a), and etc. Median filtering is a classic method. However, when we dig into the median filtering method further and further, we find that the implementation of a traditional median filter affects the result significantly. For example, Huo et al. (2012) propose to apply the median filter to the dip direction of the seismic events, which can preserve the useful energy well. Chen (2015) proposes to apply the median filter with different filter lengths, which can also preserve the useful energy well. Gan et al. (2016a) propose a similar method as Huo et al. (2012) but with a different structural filtering strategy and a different slope/dip estimation strategy. However, both Huo
Figure 18: (a) Acquisition geometry. The red dots denote shot locations and blue inverted triangles denote receivers. The simultaneous shooting and the traditional shooting only differ in the shot schedule. (b) Shot schedule. Note that the shot schedule is just for testing purpose and does not represent a real field deployment. In the test, the speed of a shooting vessel is assumed to be sufficiently large. In practice, the shot sampling would be much higher than the testing case, which is 0.05 km.
Figure 19: Shot-receiver data example. (a) Unblended data. (b) Blended data.
Figure 20: A comparison between iterative deblending performance. (a) Iteratively deblended data using the rank-constrained method. (b) Deblended data using the proposed method.
Figure 21: A comparison between deblending error. (a) Error corresponding to the traditional rank-constrained method. (b) Error corresponding to the proposed method.
Figure 22: A comparison between removed blending noise. (a) Removed blending noise corresponding to the traditional rank-constrained method. (b) Removed blending noise corresponding to the proposed method.
Figure 23: A comparison between local similarity comparison. (a) Local similarity corresponding to the traditional rank-constrained method. (b) Local similarity corresponding to the proposed method. Note the much larger similarity anomaly shown in (a).
Figure 24: Single trace comparison. The black line is from the unblended data. The red line is from the blended data. The magenta corresponds to the rank-constrained method. The green line corresponds to the proposed method. Note that the black and green lines are very close to each other, thus the deblending error using the proposed approach is much less than the traditional method for most parts. (a) Comparison in the original size. (b) Comparison in a zoomed view.
Figure 25: Deblending performance using the hybrid method with structure-oriented median filter. (a) Deblended data. (b) Deblending error. Note that the deblending error in this case is much higher than the hybrid method with structure-oriented space-varying median filter (Figure 21b).
et al. (2012) and Gan et al. (2016a) depend on accurate slope/dip estimation. What if slope/dip is not estimated accurately, e.g., in extremely noisy blended data or not in CMP gathers (as required by Gan et al. (2016a))? Then, both methods fail. In addition, for Chen (2015)'s method, what if there exists steeply dipping events? No matter what thresholds of the method we use, it fails. The structure-oriented space-varying median filter is not a seemingly simple concatenation. It relieves the influence of slope/dip estimation on the following median filtering (traditional). So the dependency of an accurate slope/dip estimation no longer exists. Besides, because of the flattening (or approximate flattening when slope/dip is not accurate), the events are guaranteed to be not steeply dipping. The combination of the two methods is elegant, which solves two long-existing and fundamental problems for the “median filtering” based deblending methods. The straightforward benefit is that we can deal with the massive blended data in a very fast and amplitude-preserving way.

The structure-oriented space-varying median filter is not computationally expensive. To understand the computational cost of the structure-oriented space-varying median filter better, we can first split structure-oriented space-varying median filter into two components: space-varying median filter and structure-oriented median filter. Compared with the traditional median filter, the space-varying median filter uses a median filter twice, each with different filter length criterion, thus the computational cost of space-varying median filter is about the twice of the cost of the median filter plus the calculation of the local similarity between the initially filtered data and a removed noise section. Calculation of the local similarity is also an efficient algorithm thanks to the fast implementation of the shaping regularization framework (Fomel, 2007b). The structure-oriented median filter adds an extra flattening (or trace prediction) step into the traditional median filter and the extra computational cost is composed of local slope calculation and spatial neighbor trace prediction. According to our experience, the 2D calculation is still acceptable. More details of slope calculation are referred to Fomel (2002) and more explanations regarding the trace prediction can be found in Liu et al. (2010). The main computational cost of structure-oriented space-varying median filter comes from calculation of the local slope, trace prediction, and local similarity calculation required by space-varying median filter in each flattened local gather. All these extra computational steps can be implemented via efficient algorithms based on shaping regularization (Fomel, 2007a), thus the application of the structure-oriented space-varying median filter is not expensive. Here, we compare the computational cost in time for four different filtering methods in Table 1. The computation is done on a PC station equipped with an Intel Core i7 CPU clocked at 3.1 GHz and 16 GB of RAM. With a small data size, the computational cost for four methods are comparable. For a large data size (e.g., 800 × 400), the computational cost for the structure-oriented space-varying median filter is roughly 3 times that of the structure-oriented median filter. It is also worth noting that the actual calculation time is usually data dependent. The exact calculation time for the local slope estimation and similarity computation may vary according to the input data but are not expensive.

The structure-oriented space-varying median filter is close to an adaptive method.
Table 1: Comparison of computational time. MF denotes median filter. SVMF denotes space-varying median filter. SOMF denotes structure-oriented median filter. SOSVMF denotes structure-oriented space-varying median filter.

<table>
<thead>
<tr>
<th>Model size</th>
<th>50×50</th>
<th>100×100</th>
<th>200×200</th>
<th>800×400</th>
</tr>
</thead>
<tbody>
<tr>
<td>MF (s)</td>
<td>0.442</td>
<td>0.460</td>
<td>0.471</td>
<td>0.591</td>
</tr>
<tr>
<td>SVMF (s)</td>
<td>0.473</td>
<td>0.491</td>
<td>0.527</td>
<td>0.962</td>
</tr>
<tr>
<td>SOMF (s)</td>
<td>0.491</td>
<td>0.508</td>
<td>0.600</td>
<td>1.525</td>
</tr>
<tr>
<td>SOSVMF (s)</td>
<td>0.660</td>
<td>0.848</td>
<td>1.769</td>
<td>4.237</td>
</tr>
</tbody>
</table>

The only parameter that needs to be chosen is the filter length. Once the filter length \( L \) is fixed, the filter can adaptively adjust according to the complicated seismic structure via the two-step cascaded structure-oriented filtering and variable window strategy. Traditionally, the performance of the structure-oriented filtering highly depends on the local slope calculated from the noisy data. While adjusting the the parameters for the plane-wave destruction algorithm, e.g., the smoothing radius, regularization parameters, and the number of iterations, can somewhat improve the performance of slope estimation, it is still almost impossible to output an optimal slope map. As shown in the beginning part of the paper, a roughly calculated local slope map will result in unflattened local gathers and are not suitable for a simple median filtering. The structure-oriented space-varying median filter eases the dependency of the performance on the input slope field in that the following space-varying median filter can deal with the unflattened energy via the variable window strategy. From a different aspect, the traditional space-varying median filter does not account for the structure patterned existing in seismic data and is only effective to events with smaller dip angle. The local flattening process in structural filtering helps the space-varying median filter prepare events with small dip angle. Combining the strategies of structural filtering and variable window length, the structure-oriented space-varying median filter is adaptive to almost any input data with an arbitrary level of structural complexity. For example, in the supplementary data, we show the results of a crossing-event dataset, where we can find that the proposed method also adapts to dataset containing crossing events.

**CONCLUSION**

The signal preserving ability of a median filter has been improved from the classic implementation to a space-varying median filter, and further to a structure-oriented median filter. However, the performance of structure-oriented median filter highly depends on the accuracy of estimated local slopes of the input seismic data. We have proposed a variant of MF, named structure-oriented space-varying median filter, that preserves signal energy even in cases with complicated structure, and can adapt to the inaccurate slope estimation in a structure-oriented filtering framework by applying a
space-varying median filter along the imperfectly flattened dimension. The proposed method can be applied to remove spike-like noise in datasets that contain structural patterns. Structure-oriented space-varying median filter can also be conveniently used as an extra constraint during iterative deblending to improve the deblending performance based on the shaping regularization framework. Both synthetic and field data examples demonstrate the great advantages and potential of the proposed method.

ACKNOWLEDGEMENTS

The authors would like to thank Sergey Fomel, Shan Qu, Yaru Xue, Wei Chen, Min Bai, Dong Zhang, Jianyong Xie, Josef Paffenholtz, Araz Mahdad, Ray Abma, Min Zhou, Qie Zhang, and Zhiyong Jiang for inspiring discussions on the simultaneous source separation problem. Yangkang Chen appreciates FairfieldNodal and BP America for offering summer internships during his PhD study at UT Austin. The authors also thank Herve Chauris, S. Mostafa Mousavi, and an anonymous reviewers for constructive comments. The research is partially supported by the “Thousand Youth Talents Plan”, the starting fund from Zhejiang University, National Natural Science Foundation of China (Grant No. U1262207), National Engineering Laboratory of Offshore Oil Exploration, 973 Program of China (Grant No. 2013CB228603), and the National Science and Technology Program (Grant No. 2016ZX05010002-002).

APPENDIX: LOCAL SIMILARITY

Local similarity between vectors $a$ and $b$ is defined as:

$$c = \sqrt{c_1 \circ c_2},$$

where $\circ$ denotes element-wise product. $c_1$ and $c_2$ come from two least-squares minimization problems:

\begin{align}
\min_{c_1} \|a - Bc_1\|^2_2, \\
\min_{c_2} \|b - Ac_2\|^2_2.
\end{align}

Here, $A$ is a diagonal operator composed of the elements of $a$, $B$ is a diagonal operator composed of the elements of $b$. Note that in equations 10-12, $a$, $b$, and $c$ denote vectorized 2D matrices. Equations 11 and 12 can be solved using shaping regularization with a local-smoothness constraint:

\begin{align}
&c_1 = [\lambda_1^2 I + T(B^T B - \lambda_1^2 I)]^{-1} T B^T b, \\
&c_2 = [\lambda_2^2 I + T(A^T A - \lambda_2^2 I)]^{-1} T A^T a,
\end{align}

where $T$ is a smoothing operator, $\lambda_1$ and $\lambda_2$ are two parameters controlling the physical dimensionality and enabling fast convergence when inversions for $c_1$ and $c_2$ expressed in equations 13 and 14 are implemented iteratively. $\lambda_1$ and $\lambda_2$ can be chosen as $\lambda_1 = \|B^T B\|_2$ and $\lambda_2 = \|A^T A\|_2$. 

REFERENCES


a deblurring filter: Geophysics, 76, no. 5, R135–R146.
Kumar, R., H. Wason, and F. J. Herrmann, 2015, Source separation for simultaneous towed-streamer marine acquisition—a compressed sensing approach: Geophysics, 80, no. 6, WD73–WD88.


——–, 2018b, Stacking seismic data based on principal component analysis: Journal of Seismic Exploration, 27, 331–348.


Zhang, Q., W. Mao, and Y. Chen, 2018a, Attenuating crosstalk noise of simultaneous-source least-squares reverse time migration with gpu-based excitation-amplitude


