Targeted Noise Removal by Seismic Time-Frequency Masking (STFM) and Minimum Statistics approach

Andrey Bakulin, EXPEC Advanced Research Center, Saudi Aramco; Dmitry Neklyudov, Institute of Petroleum Geology and Geophysics SB RAS, Novosibirsk, Russia; Ilya Silvestrov, EXPEC ARC, Saudi Aramco

Summary

Land seismic data are challenging for reliable seismic imaging. Even after extensive processing, desired reflections may remain weak. We present a new approach combining the seismic time-frequency masking with the minimum statistic approach to extract weak reflections from noisy data. Initial signal estimation is provided by nonlinear beamforming. We employ trace-by-trace phase and amplitude time-frequency masking to original data to refine signal estimation, correct distorted phase, and achieve noise suppression. Phase masks are responsible for the coherency of desired events. Amplitude masks aim to suppress amplitude noise. To construct amplitude masks, we adopt the Minimum Statistics approach from speech processing. New methods allow preserving the full frequency content of reflected signals and extracting reliable amplitude behavior.

Introduction

Data from modern land seismic acquisition using single sensors or small arrays requires significant prestack data enhancement. Local multidimensional stacking can effectively identify and enhance weak signals on prestack seismic data (Buzlukov and Landa, 2013; Curia et al. 2017, Bakulin et al., 2018). All these methods may be considered as versions of a time-delay-and-sum beamforming process. In this study, we use nonlinear beamforming or NLBF (Bakulin et al., 2018). NLBF was proven very effective for enhancing challenging land seismic data. While massive beamforming in seismic is efficient at finding events in the noisy original data, output (or beamformed) data suffer from undesired side effects: 1) Original amplitudes are heavily averaged; 2) higher frequencies are suppressed because of non-optimal stacking. Bakulin et al. (2020a, 2020b) demonstrated that beamformed seismic data might provide a reliable estimation of phase spectra of desired reflected signals. The phase substitution approach effectively combined time-frequency (TF) phase spectra from beamformed data with TF amplitude spectra of original data. In the presence of strong noise, such an approach fully preserves all frequencies, but amplitude noise remains uncorrected. This study addresses the amplitude noise problem. We suggest effective TF Wiener-type filters using the new version of the seismic time-frequency masking procedure. This version proposes adopting the Minimum Statistics (MS) approach well known in speech processing. The critical assumption is that beamformed data provides us an initial signal model that is further refined to get more reliable signal estimation without the undesired shortcomings mentioned above.

Seismic Time-Frequency Masking (STFM)

A wide variety of acoustic signals, such as speech or seismic data, are nonstationary. For this reason, many signal processing approaches were designed in the time-frequency domain. Transformation to the TF domain is achieved by the short-time Fourier transform (STFT). A typical assumption is that a registered single-channel time-dependent signal $x(t)$ may be represented as a superposition of the desired signal $s(t)$ and additive noise: $x(t) = s(t) + n(t)$. Processing aims to provide the best possible estimate of $s(t)$. Applying discrete STFT to $x(t)$, one obtains 2D complex-valued TF spectrum $X(k,l) = |X(k,l)|\exp[i\varphi_x(k,l)]$, where $|X(k,l)|$ is amplitude TF spectrum, $\varphi_x(k,l)$ is corresponding phase spectrum with $k,\ l$ representing the discrete frequency bin and time frame indices, respectively. Time-frequency masks (TFM) are widely used in the speech processing community for single-channel estimation of clean speech from a noisy record. TFM delivers an estimate of desired signal TF spectrum $\hat{S}(k,l)$ as a multiplication:

$$\hat{S}(k,l) = M(k,l) \cdot X(k,l)\ , \qquad (1)$$

where the filter $M(k,l)$ is usually a real-valued function, $0 \leq M(k,l) \leq 1$ (Yilmaz et al., 2004, Wang, 2008). Typically, $M(k,l)$ (time-frequency mask) is constructed based on the signal-to-noise consideration. M is close to 1 in the TF spectrum region of “signal dominance” and close to 0 in a “noise dominance” area. This TFM is called “amplitude TFM”. Note that the real-valued TFM does not modify the phase of the input signal: the output signal inherits the phase of the input signal in (1). One of the most popular “soft” amplitude TFM is the so-called Ideal Rationale Mask (IRM)

$$IRM(k,l) = \sqrt{\frac{|S_{est}(k,l)|^2}{|S_{est}(k,l)|^2 + |N_{est}(k,l)|^2}}\ , \qquad (2)$$

where $|S_{est}(k,l)|^2$ and $|N_{est}(k,l)|^2$ are power spectra estimates of desired signal and noise. One can easily observe that IRM may be rewritten in terms of the “instantaneous” signal-to-noise ratio (SNR). The IRM is the localized approximation of the Wiener filter. To construct IRM, signal and noise Power Spectrum (PS) must be estimated. In the speech processing community, many noise estimation algorithms have been established during the past decades. One of the most popular is the so-called Minimum Statistic (MS) approach proposed by Martin (1994) and modified by Doblinger (1995), Martin (2001), and many others. After estimating noise PS, the clean speech PS is approximated by spectral subtraction (Boll, 1979) as $|S_{est}(k,l)|^2 = |X(k,l)|^2 - |N_{est}(k,l)|^2$. These PS estimates allow the construction of IRM using the equation (2). This is a typical scheme used in speech signal processing. The scheme is entirely “data-driven.” The only input is contaminated speech records themselves. Note that phase spectra are usually not modified in the scheme. Seismic data have much more severe noise contamination than speech. Specifically, the original phase must be "denoised" in order to succeed. This requires the usage of reliable a priori information. Bakulin et al. (2020a, 2020b) proven that beamforming may provide a reliable estimation of phase spectra of reflections and introduced Seismic Time-Frequency Masking (STFM). Enhanced (beamformed) data is used as a "phase guide" in STFM. Two types of phase-dependent time-frequency masks were proposed: "phase substitution" mask (PSM) and "phase correction" mask (PCM):

$$PSM(k,l) = \exp[i\{\varphi_S(k,l) - \varphi_X(k,l)\}],$$
$$PCM(k,l) = sign[\cos\{\varphi_S(k,l) - \varphi_X(k,l)\}]. \qquad (3)$$

Here φ_X and φ_S are TF phase spectra of original and enhanced traces, respectively. Applying PSM or PCM in (1) provides results with considerably increased coherency of reflections. The frequency content of the recorded signals is fully preserved. The main disadvantage of phase-only masks is that the amplitude spectrum borrowed from the original trace is passed untouched to the final output. Amplitude noise propagates into the final result and remains uncorrected. Suppressing noise in the amplitude spectrum of original data is a nontrivial task because we lack knowledge about the actual noise distribution in the amplitude spectra. We propose a specific STFM version that constructs amplitude IRM using MS method exploiting an initial signal model from beamforming. Building on earlier work (Bakulin et al. 2020c), we derive all necessary implementation details and demonstrate them on a challenging field example from a desert environment.

Amplitude STFM and Minimum Statistics method

The MS method is based on two observations: 1) Clean speech (desired signal) and noise are usually statistically independent. It means that the power spectrum (PS) of the noisy signal is a superposition of PS of clean speech and noise: $|X(k,l)|^2 = |S(k,l)|^2 + |N(k,l)|^2$; 2) PS of noisy speech often becomes equal to the PS of noise. This happens during speech pauses and also between words and syllables. Hence the estimate of noise PS may be obtained by tracking the minimum of the noisy speech in each frequency bin separately. A similar analogy can be applied to a seismic trace consisting of series of target reflections separated by some quiet periods like in speech. Thus we can adopt assumption 2) used in the MS method for seismic data processing. We suggest the following workflow:

1. STFT is applied to original and enhanced traces $x(t)$, $s(t)$.
2. Residual PS is estimated by spectral subtraction $|R(k,l)|^2 = |X(k,l)|^2 - |S(k,l)|^2$. (Note that by definition, PS must be ≥ 0, so negative values that may occur in $|R|^2$ after subtraction must be removed). We assume that $|R|^2$ consists of noise and residual signal components that beamforming was unable to capture.

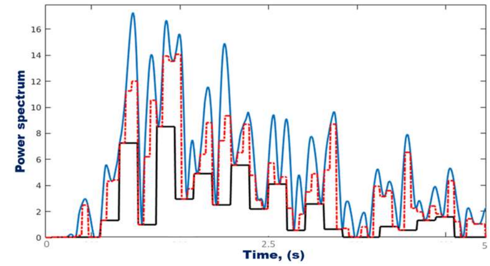

Figure 1: Residual power spectrum of a real trace at frequency 20Hz (blue) and noise PS estimation with two different time intervals used for minimum tracking: 40ms (red), 120 ms (black).

3. Noise PS is estimated using the MS approach applied to $|R(k,l)|^2$, following these simple steps. For each fixed frequency k' noise PS is estimated as $|N_{est}(k',l)|^2 = min|R(k',l \pm \Delta t)|^2$, where Δt means a half-time interval in which the search for the minimum is performed. It is an essential input parameter. As a result, noise PS is approximated as a constant value in each interval of the width $2\Delta t$. This procedure is independently performed for each frequency k.
4. Re-estimate signal PS via spectral subtraction $\left|\check{S}_{est}(k,l)\right|^2 = |X(k,l)|^2 - |N_{est}(k,l)|^2$.
5. For each frequency, , smooth $\left|\check{S}_{est}(k,l)\right|^2$ in time direction by simple rank one recursion: $|S_{est}(k,l)|^2 = \beta|S_{est}(k,l-1)|^2 + (1-\beta)\left|\check{S}_{est}(k,l)\right|^2$, with $\beta = 0.7 - 0.95$.
6. Calculate IRM (2) using estimated $|S_{est}(k,l)|^2$ and $|N_{est}(k,l)|^2$. Apply IRM separately or together with the phase masks (3) to the original trace TF spectrum using equation (1).
7. Apply inverse STFT to obtain the output filtered trace, $\hat{s}(t) = ISTFT[\hat{S}(k,l)]$.

The critical point in the workflow is that the MS method operates with the residual power spectra $|R|^2$ not the spectra of the original input data $|X(k,l)|^2$. It allows us to insert the full amount of initial signal PS (derived from NLBF) into the final signal estimation. The time interval length used for MS tracking in Item 3 is an important parameter. Figure 1 presents the residual PS of a real trace for one particular frequency, 20 Hz (shown in blue). To show the impact of time interval on IRM calculation, we plot two estimations of noise PS with different time intervals for the minimum tracking, 40 ms (red), and 120 ms (black). If a shorter time interval is used, estimated noise PS would be closer to residual $|R|^2$, implying a larger portion of it is assigned to noise. As a result, IRM with a shorter interval provides a result close to the initial guess of beamformed data. If a larger time interval is used for minimum tracking, the output result of IRM filtering would be closer to the original data. In other words, IRM will pass most of the residual PS components to the output.

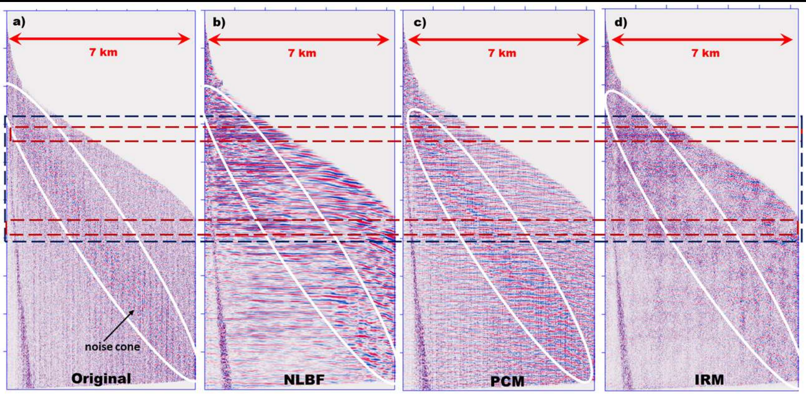

Figure 2: Real data example showing prestack NMO corrected CMP gathers obtained by different approaches: (a) Original data after conventional processing; (b) data after nonlinear beamforming; (c) data after phase correction (no amplitude filtration); and (d) data after application of amplitude mask only (no phase corrections). The white oval marks the residual noise cone zone.

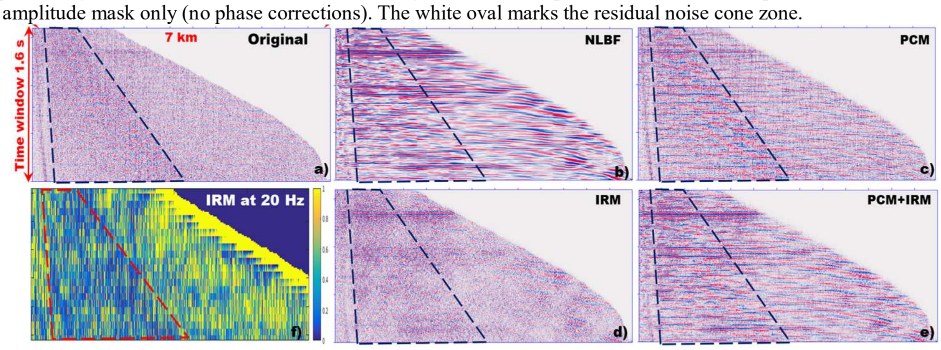

Figure 3: Zoom of the CMP gathers presented in Figure 2 (time window [1.4-3.0] s): (a) Data after conventional processing; (b) enhanced data after NLBF; (c) data after PCM (phase correction mask); (d) data after application of amplitude filtering with IRM (no phase corrections); (e) data after applying a combination of PCM+IRM; (f) IRM calculated for frequency 20 Hz. The noise cone region is marked by trapezoid.

Real data example from a desert environment

An example of common-midpoint (CMP) multi-azimuth gather after NMO correction from a challenging 3D land dataset is shown in Figure 2. Although the data have already been passed through a standard processing flow, there are almost no visible reflections in the gather. Apart from low SNR, this gather also shows an imprint of the so-called “noise cone” marked by a white oval. Figure 1b shows the same CMP gather after data enhancement by NLBF when approximately 200 neighboring traces are used to produce each output trace (apertures are 150x150 m in CMP and offset directions). Reflections are easily recognizable in the entire offset range after enhancement. The high-frequency content of the signal is suppressed due to sub-optimal stacking, and amplitudes are considerably decreased. In Figure 2d one can see what happens when amplitude only mask (IRM) is applied (the time interval for minimum search is 40 ms). Reflections start to pop up more clearly; however, wavelet and phase distortions cannot be fixed by amplitude IRM. Figure 2c shows gather after phase correction (PCM). Events become more coherent, but noise contamination in the amplitudes remains. Figure 3 zooms into details of processed CMP gathers (time window [1.4-3.0] s) to demonstrate what each of the masks can achieve. Original data after conventional processing remains distorted in both amplitude and phase (Figure 3a). One can see the noise cone's rough location with elevated amplitudes left as an imprint from prior processing marked by trapezoid. Beamforming gives an excellent estimate of the signal coming from local summation, but it is oversmoothed. The noise cone is more or less eliminated (Figure 3b). Reflection amplitudes slowly change with offsets without any noticeable jumps at the onset of the cone. Application of phase correction mask (PCM) considerably improves coherency. Amplitude imprint propagates (as it should by design) since the amplitudes remain untouched. Amplitude only mask (IRM) can suppress the noise level. Hence, signal amplitudes stand up more clearly (Figure 3d). It is unable to fix wavelet and phase variations, so reflected events continuity is not improved. The combined application of phase and amplitude masks (IRM+PCM, Figure 3e) achieves the most balanced result making reflections coherent without oversmoothing while preserving more

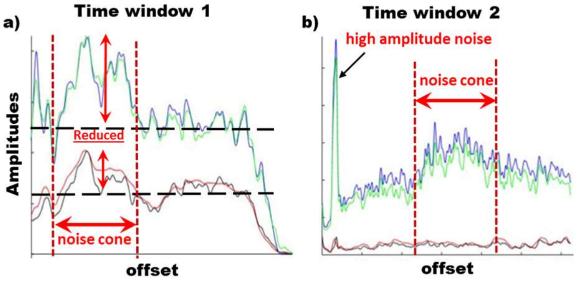

Figure 4: RMS amplitudes calculated within two time windows extracted along target reflections: (a) time window [1.63-1.7] s.; (b) time window [3.16 – 3.26] s. (blue) RMS amplitudes for original data; (green) for data after PCM; (red) for data after NLBF; (black) for data after IRM (no phase correction).

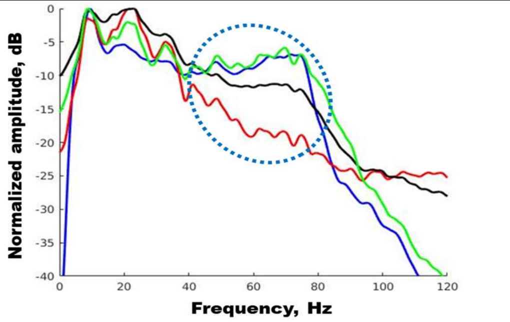

Figure 5: Normalized averaged amplitude spectra calculated within time window [1.4-3.0] of original and processed CMP gathers. (blue) original data; (red) data after NLBF; (green) phase corrected data; (black) data processed with IRM.

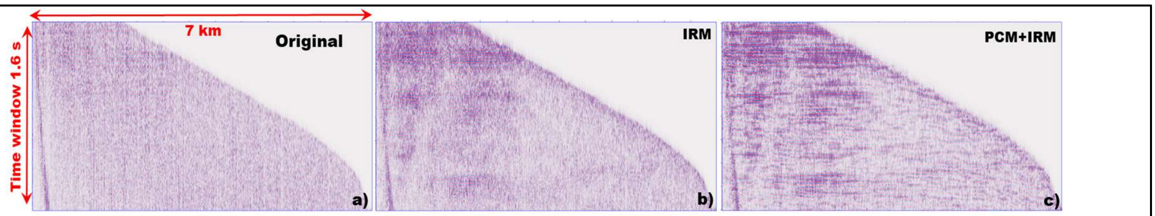

Figure 6: The same fragment of CMP gathers as in Figure 3 after band-pass filtering (40-80 Hz): (a) Original data after conventional processing; (b) data after application of amplitude filtering by IRM (no phase corrections); (c) combination of PCM+IRM. Observe uncovered coherent reflections on (b) and (c) that remain invisible in (a) covered by noise.

spatial details than beamforming itself. Figure 3f reveals IRM values for a frequency of 20Hz. By definition, IRM varies within the interval [0, 1]. Precisely as expected, we observe small values (blue color) inside the noise cone and larger values closer to 1 (yellow color) outside the cone. Further, we examine root-mean-square (RMS) amplitudes behavior within short time windows extracted along target reflections where noise cone cut them at different offset intervals (these windows are shown in Figure 2 as two red boxes). In Figure 4 one can see clear humps on original data RMS amplitudes (shown in blue), identifying the location of the noise cone at the near offset in window 1 (Figure 4a) and mid offsets in window 2 (Figure 4b). After PCM, one can see almost identical curves (shown in green) that preserve the hump. Combination PCM+IRM provides RMS behavior (shown in black), similar to the beamforming RMS amplitude itself (shown in red). Both red and black curves show either reduction (window 1) or complete elimination (window 2) of the hump. Note that amplitudes provided by two approaches (IRM and IRM+PCM) are the same. In Figure 5, we present averaged amplitude spectra of the gathers. As one can see, beamforming harshly reduces all amplitudes at frequencies above 40 Hz. In contrast, while IRM also suppresses the noise, it retains some signal amplitudes above 40 Hz. Figure 6 proves that useful reflections are recovered at higher frequencies (40-80 Hz). They are consistent with events previously recovered with NLBF (< 40 Hz, Figure 3b). As a result of IRM processing, reliable RMS amplitudes behavior with preservation of higher frequency in the band have been obtained.

Conclusions

We present a new enhancement approach of Seismic Time-Frequency Masking that is powerful enough to deal with challenging land data yet delicate enough to preserve the highest frequencies present in the data. We utilize massive beamforming with large apertures to uncover hidden reflectors. Such enhanced data serves as an approximate initial “signal model.” We exploit this model to correct original data using specially designed STFM. We demonstrate that frequency-dependent phase corrections are crucial to restoring the coherency of reflectors. Amplitude masks are designed to suppress noise. Beamformed data are used to design targeted amplitude masks suppressing noise in places of domination while not touching areas with prevalent signals. For amplitude STFM calculation, we adopt a minimum statistics approach from speech processing that allows simple estimation of the noise power spectrum. The real data example validates the effectiveness of the proposed approach. Corrected multi-channel gathers become acceptable for conventional processing by existing methods. STFM opens many new possibilities for multi-channel seismic data processing, not achievable with the current processing tools.

10.1190/segam2021-3581846.1

REFERENCES

Bakulin, A., D. Neklyudov, and I. Silvestrov, 2020a, Prestack data enhancement with phase corrections in time frequency domain guided by local multidimensional stacking: Geophysical Prospecting, **68**, 1811–1818, doi: https://doi.org/10.1111/1365-2478.12956.

Bakulin, A., D. Neklyudov, and I. Silvestrov, 2020c, Seismic data enhancement and targeted noise removal using time- frequency masking guided by beamformed data: 90th SEG Annual Meeting, Expanded Abstracts, 3134–3138, doi: https://doi.org/10.1190/segam2020-3420095.1.

Bakulin, A., I. Silvestrov, and D. Neklyudov, 2020b, Importance of phase guides from beamformed data for processing multi-channel data in highly scattering media: The Journal of the Acoustical Society of America, **147**, EL447–EL452, doi: https://doi.org/10.1121/10.0001330.

Bakulin, S. I., M. Dmitriev, D. Neklyudov, M. Protasov, and K. Gadylshin, 2018, Nonlinear beamforming for enhancing prestack seismic data with a challenging near surface or overburden: First Break, **36**, 121–126, doi: https://doi.org/10.3997/1365-2397.n0143.

Boll, S. F., 1979, Suppression of acoustic noise in speech using spectral subtraction: IEEE Transactions on Acoustics, Speech, and Signal Processing, **27**, 113–120, doi: https://doi.org/10.1109/TASSP.1979.1163209.

Buzlukov, V., and E. Landa, 2013, Imaging improvement by prestack signal enhancement: Geophysical Prospecting, **61**, 1150–1158, doi: https://doi.org/10.1111/1365-2478.12047.

Curia, D., P. Borghi, J. Noble, A. Berkovitch, D. Justo, and M. Alayón, 2017, The impact of multifocusing in the processing of land 3D seismic data in a fold and thrust belt setting: Ranquil Norte Block, Neuquén Basin, Argentina: The Leading Edge, **36**, 770–774, doi: https://doi.org/10.1190/tle36090770.1.

Doblinger, G., 1995, Computationally efficient speech enhancement by spectral minima tracking in subbands: Fourth European Conference on Speech Communication and Technology, EUROSPEECH'95, Madrid, Spain, 1513–1516.

Martin, R., 1994, Spectral subtraction based on minimum statistics: Proceedings of European Signal Processing Conference (EUSIPCO'94), Edinburgh, U.K., 1182–1185.

Martin, R., 2001, Noise power spectral density estimation based on optimal smoothing and minimum statistics: IEEE Transaction on Speech and Audio Processing, **9**, 504–512, doi: https://doi.org/10.1109/89.928915.

Wang, D., 2008, Time-frequency masking for speech separation and its potential for hearing aid design: Trends in Amplification, **12**, 332–353, doi: https://doi.org/10.1177/1084713808326455.

Yilmaz, O., and S. Rickard, 2004, Blind separation of speech mixtures via time-frequency masking: IEEE Transactions on Signal Processing, **52**, 1830–1847, doi: https://doi.org/10.1109/TSP.2004.828896.