Practical Aspects of Non Local Means Filtering of Seismic Data

Carlo De Gaetani, Samuel Winters, Janet Barnes, Sergio Grion, Dolphin Geophysical Ltd.

SUMMARY

Non Local Means (NLM) filtering is a well-known image-processing algorithm for random noise attenuation. It is based on the assumption that coherent and non-coherent features can be identified and separated using a measure of similarity between adjacent samples. In this paper we review and extend previous work on the use of NLM in seismic data processing. Our objective is to improve the computational efficiency of the method and investigate practical aspects of its implementation and application. Synthetic and real data examples demonstrate the achieved improvements.

Introduction

The reduction of non-coherent noise is a key aspect of seismic data processing. While *f-x* and *f-x-y* deconvolution are extensively used for this purpose and known to be robust and effective, the increased bandwidth of broadband data prompts for more sophisticated algorithms, in particular with edge-preservation capabilities. The Non Local Means (NLM) algorithm is, in this respect, a promising option. Originally developed by Buades et al. (2005) for digital image processing applications, it was later applied by Bonar and Sacchi (2012) to seismic data processing. In their paper, Bonar and Sacchi prove the reliability of the algorithm on time slices of a spectrally-decomposed volume and on post-stack cross-sections. However, while the quality of the obtained results is indisputable, the high computational cost of the algorithm makes its use problematic, in particular for 3D applications. Maraschini and Turton (2013) suggest a windowed application to partially address this issue. We propose a revised NLM algorithm aimed at improving computational efficiency while maintaining the noise reduction capability.

Non Local Means filtering

We begin with a concise description of the NLM algorithm. If $v(i)$ is the amplitude of sample i and $v(j)$ the amplitude of one of m samples j not necessarily in the vicinity of i, a filtered value $\hat{v}(i)$ is obtained as:

$$w(i,j) = exp\left(\frac{-D^2(i,j)}{h^2}\right) \quad \text{with weights:} \quad \hat{v}(i) = \frac{\sum_j w\,(i,j)v(j)}{\sum_j w\,(i,j)}$$

In other words, $\hat{v}(i)$ is a weighted mean of sample values in a given window. The weights $w(i,j)$ depend on a measure of relative distance (or “similarity”) $D^2(i,j)$ between the reference sample i and the m samples j. This leads to assigning a higher weight to samples that have consistent amplitudes. As consequence, structures and coherent features are preserved while random noise is attenuated. The filtering harshness is controlled by the parameter h. The similarity between patches is computed as:

$$D^2(i,j) = \sum_l \left[K(l)\left(v\big(N_i(l)\big) - v\left(N_j(l)\right)\right)\right]^2$$

The term $K(l)$ is the $l-th$ element of a normalized convolutional kernel K that balances the overall similarity between the reference neighbourhood patch N_i around sample i and the corresponding comparison patch N_j around sample j. Standard NLM filtering is applied to 2D images using square neighbourhoods and square windows. In this case, the complexity of this algorithm is $n \times m^2 \times k^2$ where n is the number of samples of the image to be filtered, m is the size of the window and k is the size of the patches. NLM is well suited to be parallelized but further adjustments to the algorithm are required in order to make its 3D application to 3D data feasible in practice. We therefore discuss the following three key issues:

- use of specific convolutional kernels;
- anisotropy of the kernel and use of non-square comparison patches and windows;
- reference patch weight.

The convolving kernel smooths the similarity measure between patches and indirectly determines their size, and therefore the computational cost of the algorithm. Previous authors base their work on a Gaussian kernel. Such a kernel balances the similarity measure in favour of the elements closer to the central part of the patch. In practice however, a similar effect can be obtained using smaller patch sizes in combination with triangular or uniform kernels. We illustrate this in Figure 1 where Gaussian, triangular and uniform kernels are superimposed. Additionally, for an n-D extension of the algorithm, kernel anisotropy can also be introduced. A 1D kernel can be viewed as an anisotropic kernel in which smoothing occurs in just one out of many dimensions available. On cross-line or in-line gathers we exploit this behavior to better address signal recognition. In seismic data, the signal to be preserved is essentially a repetition of an elementary 1D wavelet along 2D or 3D structures that are

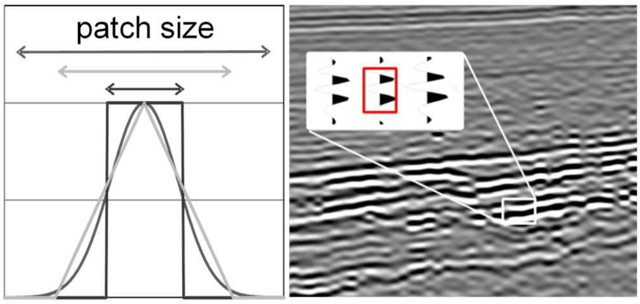

Figure 1 Various kernels and associated patch sizes (left); reference patch in red and search window in white (right).

sometimes horizontal or close to horizontal. We therefore choose 1D patches fully encompassing a reference wavelet or wavelet train (Figure 1, red box) in combination with rectangular search windows extended in the 2D or 3D dipping structure direction (Figure 1, white box). The size and orientation of the search window depends on geological dip as well as on faults and discontinuities that must be preserved. Finally, we investigate the weight to be assigned to the reference patch. During the comparison process within the search window, the similarity is a maximum when the reference patch slides exactly over itself. In this case, the assigned weights are artificially over-scaled with respect to all the others and can potentially compromise the ability of the NLM algorithm to separate coherent and non-coherent features. To avoid this problem, we exclude from the comparison process the reference patch and we assign to it a default value $w_{ii} = max(w_{ij})$. With this approach the reference patch is used in the weighted averaging but assumes a non-dominant weight.

The three modifications just described reduce the complexity of the algorithm from $n \times m^2 \times k^2$ to $n \times m_1 \times m_2 \times k$, where the rectangular window is of size $m_1 \times m_2$ with $m_1 \ll m_2 = m$, and improve the S/N enhancement thanks to a now optimized exploitation of the anisotropic convolution kernel, a properly sized search windows and a proper re-balancing of the weight assigned to the reference patch. Examples of the obtained advancements are described in the following section.

Examples

First, we show a synthetic example. This test is inspired by the example shown by Bonar and Sacchi (2012). A synthetic gather comprising crossing, curved, faulted and linear events (see Figure 2) is corrupted with Gaussian white noise, with a resulting SNR of about -5.5dB. *f-x* deconvolution and NLM are then compared in terms of noise removal and signal preservation. The developed modifications to the NLM algorithm are individually evaluated. With NLM$_0$ we refer to the standard form of the algorithm, i.e. with square patches and search windows (19 traces x 19 samples wide) combined with a Gaussian kernel and $w_{ii} = 1$. With NLM$_1$ we compare the effect of considering $w_{ii} = max(w_{ij})$ and with NLM$_2$ we additionally evaluate the use of a uniform kernel with rectangular search windows 31 traces x 11 samples wide and a 1D reference patch 13-samples long. The NLM smoothing parameter is h=0.2 for all cases.

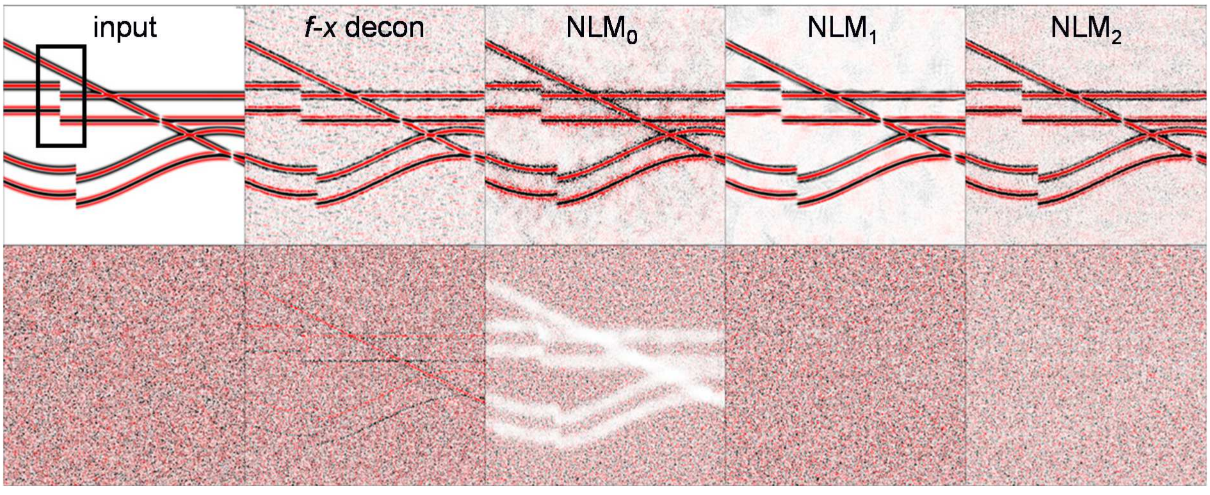

***Figure 2** Synthetic data test. Signal (top) and noise (bottom). White corresponds to zero amplitude.*

In Figure 2, *f-x* deconvolution filters most of the noise but signal is also partially removed. NLM_0 obtains a better noise removal but noise overlapping coherent events is excluded from the filtering process. Signal and noise separation is highly improved with NLM_1. A similar result but with a 50% computational cost reduction can be obtained using a uniform kernel and a smaller (13 x13) reference patch (not shown in Figure 2). The computational cost is further reduced with the NLM_2 test, about 29 times less than the previous NLMs. In this case the noise removal capability is similar to *f-x* deconvolution but signal preservation is much better.

Furthermore, NLM preserves edges and discontinuities. In Figure 3 we show a close-up of one of the discontinuities (black box in Figure 2). *f-x* deconvolution partially removes coherent energy and introduces smearing effects. NLM_0 and NLM_1 give improved results and NLM_2 is smearing-free due to the use of a 1D reference patch.

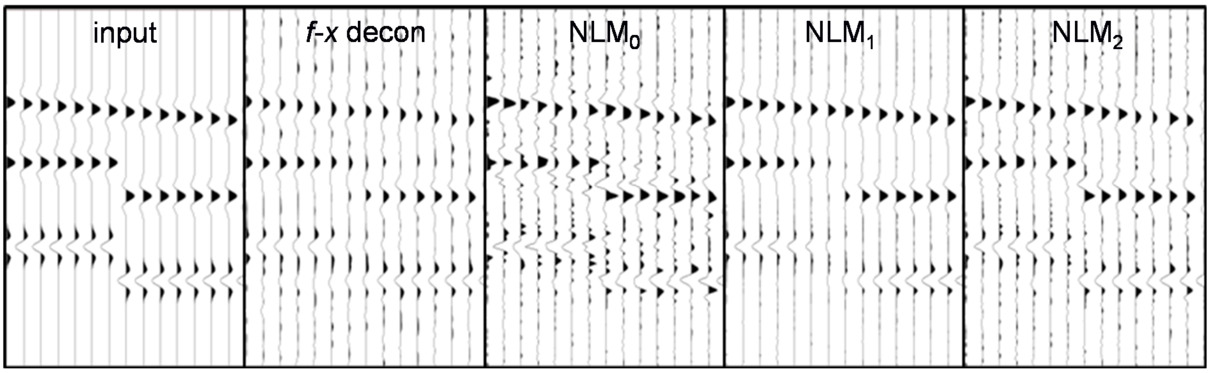

Figure 3 *Synthetic data test. Close-up from the top row of Figure 2.*

We provide now a second test case on real data. We apply our revised NLM on a 3D pre-migration common-offset (135m) section and we evaluate the effects of extending the NLM algorithm to three dimensions. Benchmark of these tests is the result obtained on an inline section using a 1D reference patch of 15 time samples with a uniform kernel, 2D search windows 31x1x7 wide (cross-line x in-line x time samples) and a smoothing parameter h=0.5. We call it 1D2D test and we show a comparison between a detail of the input and the corresponding NLM result in Figure 4. Incoherent energy is well removed and the continuity of the events appears properly maintained. Reduction of noise content is quite homogeneous across the entire section (not shown for brevity).

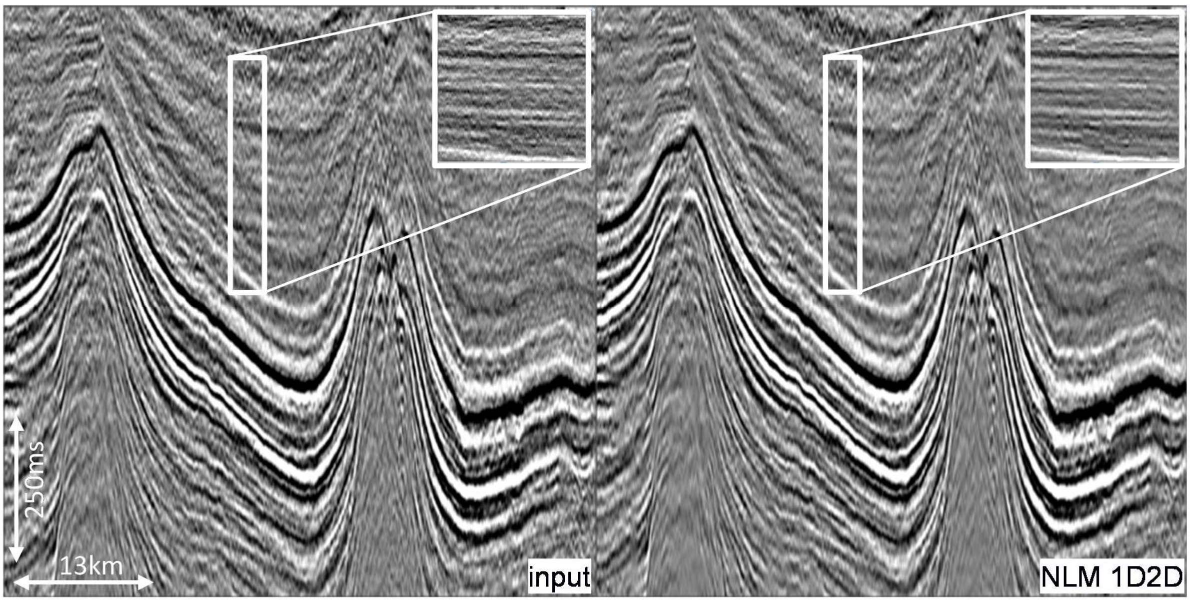

Figure 4 *Real data common-offset test. Detail of input and NLM output. Close-up to better visualize the noise reduction (white box).*

We compare this result with a search window extension to 31x3x7 including 3 in-lines (1D3D test) and further extension of the reference patch to size 1x3x15 (2D3D test). We evaluate the results by comparing the noise detected by the different configurations as the difference between the input and the output, scaled by a factor of 10 for better visualization. Results are shown in Figure 5.

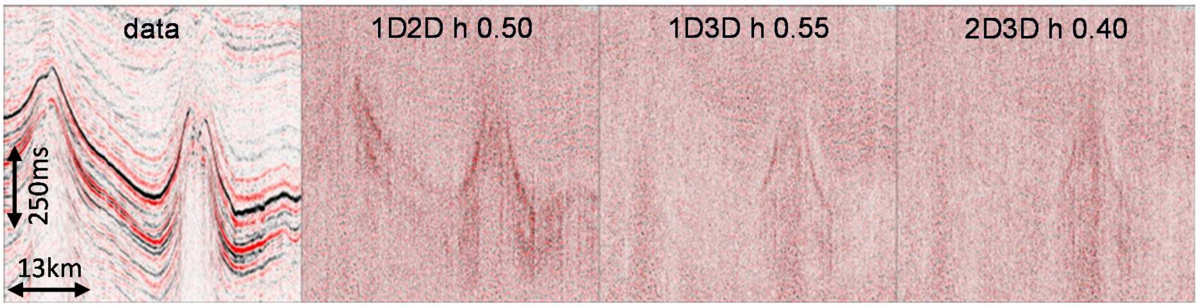

Figure 5 *Common-offset test. Data and detected noise (×10). White corresponds to zero amplitude.*

In 1D2D we detect noise homogeneously but the main structure is clearly visible in the difference section. In 1D3D, despite a 10% increment of the smoothing parameter, the signal is better preserved, but the computational speed is three times slower. In 2D3D the signal-noise separation is only slightly improved, but at great computation cost increase. The increase in filtering harshness requires a milder smoothing parameter (h=0.40) and the computational cost of this test is nine times bigger than the benchmark.

Conclusions

Our revised NLM algorithm gives improved results in seismic data processing applications. We focus on the improvements obtainable on sections rather than on time slices. We achieve better noise reduction and signal preservation, both with a reduced computational cost. The key elements of the proposed changes are a proper re-balancing of the weight assigned to the reference patches, an optimized combination of reference patch and search window sizes and the use of anisotropic uniform kernels. The extension of the algorithm to more dimension improves further the SNR enhancement capability but with a consequent reduction of the computational speed. Our tests show that the extension of the search windows to 3D while maintaining the reference patch 1D gives satisfactory results. The signal and its discontinuities are well preserved and incoherent noise is attenuated.

Acknowledgements

We thank Dolphin Geophysical Multi Client department for the permission to publish our tests carried out on real data.

References

Bonar, D. and Sacchi, M. [2012] Denoising seismic data using the Nonlocal means. Geophysics, vol. 77, no. 1, 2012, pp A5-A8, doi: 10.1190/GEO2011-0235.1.

Buades, A., Coll, B., Morel, J.M. [2005] A review of image denoising algorithms, with a new one. SIAM Journal on Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal, vol. 4, no. 2, 2005, pp.490-530, doi: 10.1137/040616024.

Maraschini, M. and Turton, N. [2013] Random Noise Attenuation Preserving Geological Detail – A Fast and Effective Non-local-means Filter. 75th Conference & Exhibition, Extended Abstracts, doi: 10.3997/2214-4609.20130881.