Non-Local Means Denoising of Seismic Data

David Bonar, Department of Physics, University of Alberta
bonar@ualberta.ca
and
Mauricio Sacchi, Department of Physics, University of Alberta

GeoConvention 2012:Vision

Summary

The non-local means algorithm was originally developed for the random noise attenuation of images and has recently been applied in other fields such as medical imaging. To denoise each pixel or point of the image, the non-local means algorithm utilizes other similar pixels within the image regardless of their spatial proximity, making the process non-local. Through assuming that the redundancy of the structures within any image can be applied for denoising, we propose to adopt the non-local means algorithm to attenuate random noise in seismic data.

Introduction

Random noise attenuation within seismic data has generally been performed using several methods based upon different assumptions about the data. For example, band-pass, $f-k$, and k_x-k_y filtering (Yilmaz, 2001) all transform seismic data into the Fourier domain to mute undesirable portions of the signal, i.e., noise, based on the assumption that the signal and noise are separated in this new domain. Other methods such as $f-x$ deconvolution (Canales, 1984), $t-x$ prediction filtering (Abma and Claerbout, 1995), and Cadzow filtering (Trickett, 2008) or Singular Spectrum Analysis (Oropeza and Sacchi, 2011) attempt to remove random noise based on assumptions such as the linearity of seismic events. Originally developed for image processing (Buades et al., 2005), the Non-Local Means (NLM) algorithm is a random noise attenuation filter that assumes the degree of redundancy present within an image can be utilized to reinforce the structures within any small window, or neighborhood, within the image from the many similar windows that are also within that image. Thus, the data itself is employed for denoising. However, since the NLM algorithm uses the entire image or data set to denoise a single pixel or location, it can become computationally demanding (Buades et al., 2010). Therefore, several variations of the algorithm have arisen to decrease the computational time such as utilizing its highly parallelizable nature and attempting to decrease the computational time for a single pixel or location (Coupé et al., 2008; Mahmoudi and Sapiro, 2005; Brox et al., 2008). The NLM method has been shown to successfully denoise medical data such as MRI, radar data, speech and audio data, and microscopy images. We propose to adopt the NLM algorithm for attenuating the random noise within seismic data. In particular, we use the traditional NLM algorithm originally developed by Buades et al. (2005) for random noise attenuation on two synthetic data sets, one containing curved events and sharp discontinuities and the other containing an AVO response.

Theory

We describe the NLM denoising algorithm based on the description provided by Buades et al. (2010). Let a discrete noise contaminated image v be defined by

$$v = u + n, \tag{1}$$

or simply the summation of the original noise free image u with random noise n. At the pixel i, the non local means denoised pixel, $\hat{v}(i)$, is simply the weighted average of all of the pixels within the noisy image,

$$\hat{v}(i) = \sum_j w(i,j)v(j), \tag{2}$$

where the weights $w(i,j)$ depend upon the similarity between the pixels i and j and must satisfy the conditions $0 \le w(i,j) \le 1$ and $\sum_j w(i,j) = 1$. Note that each pixel i of the image has its own independent weights of the other j pixels within the image. To quantify the similarity between the pixels i and j, a neighborhood or window, $\mathcal{N}_i$, around the pixel of interest is defined to allow for information about local structures and textures to be incorporated. The neighborhood of a pixel is generally chosen to be a square or cube, depending upon the dimensionality of the image, with a dimension size of 3 to 9 (Awate and Whitaker, 2006; Coupé et al., 2008; Dowson and Salvado, 2011) centered upon the pixel of interest, however, the size and shape of the neighborhood can vary. The similarity between the pixels i and j is then computed using a Gaussian weighted Euclidean distance, $D(i,j)$, between the neighborhood around the pixel i, $v(\mathcal{N}_i)$, and the neighborhood around the pixel j, $v(\mathcal{N}_j)$,

$$D^2(i,j) = ||v(\mathcal{N}_i) - v(\mathcal{N}_j)||^2_{2,a} = \sum_l^{nl} [G_a(l)(v(\mathcal{N}_i(l)) - v(\mathcal{N}_j(l)))]^2 \tag{3}$$

Here the operator $||.||^2_{2,a}$ denotes the squared Gaussian weighted Euclidean distance $D^2(i,j)$, G_a represents the Gaussian kernel with standard deviation a, and l represents one of the total nl elements within a neighborhood. For a two dimensional image, the Gaussian kernel, G_a, can be defined by,

$$G_a(x,y) = \exp\left(-\frac{(x-x_o)^2 + (y-y_o)^2}{2a}\right), \tag{4}$$

where x_o and y_o denote the center of the Gaussian kernel with x and y corresponding to the coordinates of the element l in equation 3. By weighting the Euclidean distance between the neighborhoods of pixels i and j with a Gaussian kernel, smaller weights are assigned to distant pixels within a neighborhood allowing structures closest to the pixel of interest to be more likely preserved. Given the Gaussian weighted Euclidean distance, $D(i,j)$, between the pixels i and j, the weights $w(i,j)$ are computed according to

$$w(i,j) = \frac{1}{Z(i)} \exp\left(\frac{-D^2(i,j)}{h^2}\right), \tag{5}$$

where $Z(i)$ is the normalizing factor defined by

$$Z(i) = \sum_j \exp\left(\frac{-D^2(i,j)}{h^2}\right) \tag{6}$$

to ensure $\sum_j w(i,j) = 1$. The parameter h is a constant which controls the decay of the exponential function as a function of the Euclidean distance. For example, a large value for h will provide similar weight for all j pixels in the image while a small value for h will provide a significant weight for only a few of the j pixels in the image. Proper estimation of the parameter h can be carried out using methods such as the χ^2 criterion for the goodness of fit if the standard deviation of the noise is known or heuristically, like the estimation of the prediction filter length in $f-x$ deconvolution, if little is known about the noise. For consistency, the parameter h was chosen heuristically throughout this article. In other words, we visually inspected the denoised image for both noise attenuation and structure preservation to determine the optimal value for the parameter h. Typically this value was found to be around one order of magnitude less than the largest amplitude present within the data.

Examples

One of the specific advantages that the NLM algorithm can have for attenuating random noise in seismic data is its ability to handle sharp discontinuities, such as faults, and events with curvature. To illustrate these capabilities, a synthetic data set containing linear, curved, and crossing events with

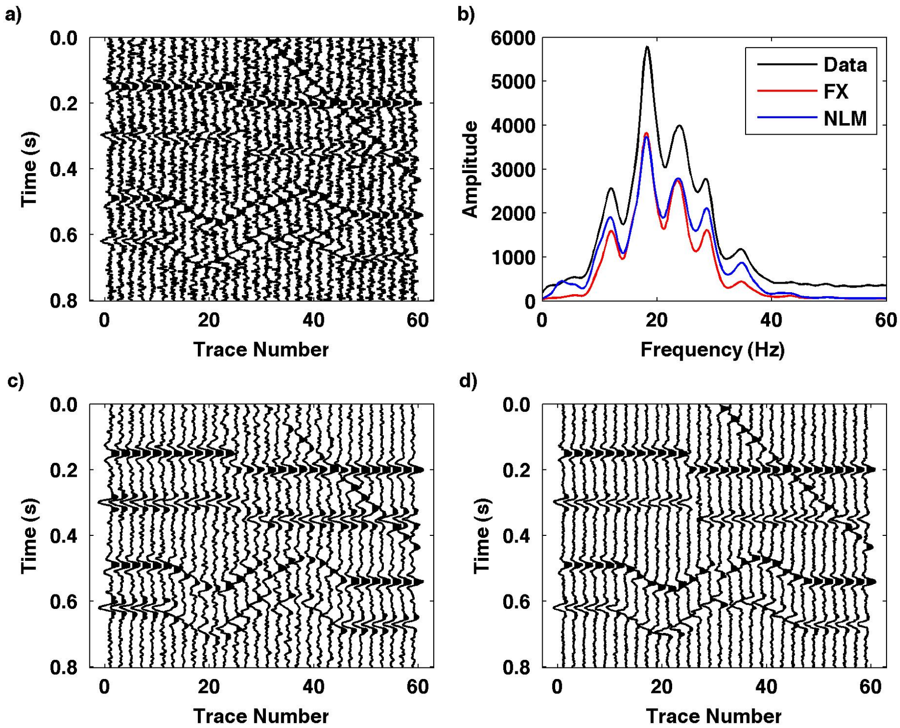

Figure 1 *a) Noisy synthetic data with linear, curved, and crossing events with sharp discontinuities. b) Amplitude spectra for a), c) and d). c) Denoising with $f-x$ deconvolution. d) Denoising with NLM.*

sharp discontinuities was created and contaminated with random noise such that the signal to noise ratio (SNR) was 1.2. Here, the SNR is defined as the variance of the data divided by the variance of the noise. The denoising of this synthetic data set, as shown in Figure 1, was performed using NLM with $h = 0.15$, $a = 0.25$, and an $11\text{x}11$ square window and $f-x$ deconvolution that was applied in small spatial windows, 11 traces wide, to help validify its plane wave assumption. Unlike $f-x$ deconvolution, the NLM algorithm preserves the amplitude of the curved events and does not smear seismic energy across the sharp discontinuities. Figure 1b depicts the Fourier amplitude spectrum of the synthetic data and its denoised versions and highlights the ability of NLM to adequately maintain high frequencies as compared to $f-x$ deconvolution, even though it is an averaging process. However, it should be noted that to achieve these results, the simple implementation of the NLM algorithm had a computational time that was greater than two orders of magnitude larger than that of $f-x$ deconvolution. To become more feasible for application with seismic data, further research into decreasing the computational time of NLM needs to be conducted. As a second example, the NLM algorithm was applied to a synthetic data set of SNR = 2 that contains an AVO signature, a constant gradient between an amplitude of 1 to -1 across the section, as seen in Figure 2. For optimal results, h was chosen to be 0.15 with a window size of $11\text{x}11$ and a was chosen to be 0.25. While significantly decreasing the noise level of this synthetic data set, NLM was also able to preserve and denoise the AVO response of the horizon depicted with the red line.

Conclusions

The NLM algorithm is a random noise attenuation filter that utilizes the redundancy of structures within a data set to denoise each location within it. Unlike other common denoising methods for seismic data, the concepts governing NLM enables it to handle AVO responses, curvature and faults without losing the resolution of these features. In its basic implementation there are three parameters to be controlled: the window or neighborhood size, the standard of deviation a for the Gaussian kernel, and the parameter h. While the algorithm is relatively insensitive to the parameter a since it only controls

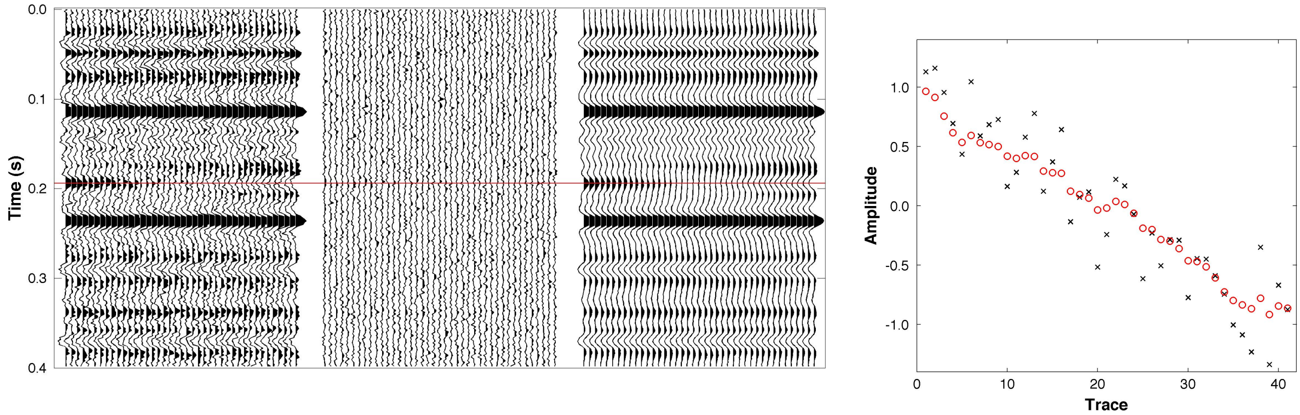

Figure 2 *NLM denoising of a noisy synthetic data set containing an AVO response. From left to right: noisy data, estimated noise, NLM denoised data, amplitude of the horizon of the red line (x - noisy data, o - NLM denoised data).*

the weights at the edges of a neighborhood, the neighborhood size must be chosen such that it is large enough to encapsulate the structures of interest within the data. For reflection seismic data, this means that the neighborhood size, in the time dimension, should larger than the wavelet length. The parameter h controls the level of denoising the NLM algorithm performs and for the examples provided was found to be optimally set to be about one order of magnitude less than the maximum amplitude of the data. The NLM algorithm can be easily expanded to incorporate multiple dimensions and has the potential to become a common denoising method in either prestack or poststack seismic data because of its abilities to not smear seismic energy at termination points or sharp discontinuities (i.e., faults).

Acknowledgements

We would like to thank the sponsors of the Signal Analysis and Imaging Group (SAIG) for sponsoring this research.

References

Abma, R. and J. Claerbout, 1995, Lateral prediction for noise attenuation by $t-x$ and $f-x$ techniques: Geophysics, **60**, 1887–1896.

Awate, S. and R. Whitaker, 2006, Unsupervised, information-theoretic, adaptive image filtering for image restoration: IEEE Transactions on Pattern Analysis and Machine Intelligence, **28**, 364 –376.

Brox, T., O. Kleinschmidt, and D. Cremers, 2008, Efficient nonlocal means for denoising of textural patterns: IEEE Transactions on Image Processing, **17**, 1083 –1092.

Buades, A., B. Coll, and J. M. Morel, 2005, A review of image denoising algorithms, with a new one: Multiscale Modeling and Simulation, **4**, 490–530.

——— 2010, Image denoising methods. A new nonlocal principle: SIAM Review, **52**, 113–147.

Canales, L. L., 1984, Random noise reduction: SEG Technical Program Expanded Abstracts, **3**, 525–527.

Coupé, P., P. Yger, S. Prima, P. Hellier, C. Kervrann, and C. Barillot, 2008, An optimized blockwise nonlocal means denoising filter for 3-d magnetic resonance images: IEEE Transactions on Medical Imaging, **27**, 425–441.

Dowson, N. and O. Salvado, 2011, Hashed nonlocal means for rapid image filtering: IEEE Transactions on Pattern Analysis and Machine Intelligence, **33**, 485–499.

Mahmoudi, M. and G. Sapiro, 2005, Fast image and video denoising via nonlocal means of similar neighborhoods: IEEE Signal Processing Letters, **12**, 839–842.

Oropeza, V. and M. Sacchi, 2011, Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis: Geophysics, **76**, V25–V32.

Trickett, S., 2008, F-xy cadzow noise suppression: SEG Technical Program Expanded Abstracts, **27**, 2586–2590.

Yilmaz, Ö., 2001, Seismic data analysis: processing, inversion, and interpretation of seismic data: Society of Exploration Geophysicists, 2 edition.