GPR data reconstruction using multichannel singular spectrum analysis

Saulo S. Martins *(University of Alberta), Mauricio Sacchi (University of Alberta), and Jandyr M. Travassos (COPPE, UFRJ)

Abstract

This works deals with problem of Ground Penetrating Radar (GPR) trace interpolation. The alias and noise is very common in seismic data and wide explored. In GPR case is not different, so why not apply seismic tools in those data? Here we apply a multichannel singular spectrum analysis (MSSA) to reconstruct the data. The idea here in the first moment is testing the methodology muting gradually a model and after a real. The strategy was muting the data randomly until 90% than try recover it back. The results shows that this method used primarily in seismic data are extremely effective in GPR data. What encourages us to apply in the next future this method in a GPR multi-offset section acquired in Santa Catarina, south of Brazil.

Introduction

Seismic trace interpolation is a fundamental step in seismic data processing. Interpolation algorithms have been widely used to correct irregularities in the spatial sampling of seismic data. Interpolation methods are utilized for data completion and for noise attenuation as well. A comprehensive discussion of single gather interpolation methodscan be found in Abma and Kabir (2005). In addition, seismic trace interpolation is often needed to improve the quality of migrated images (Spitz, 1991; Liu and Sacchi, 2004; Trad, 2009).

This work addresses interpolation of Ground Penetrating Radar (GPR) datasets. GPR surveying is often a cost effective way for environmental near surface studies and for the characterization of glacial environments (Macheret et al., 1993; Murray et al., 1996, 2000). However, it is often the case that one would like to acquire more GPR data by using coarse sampling scenarios. Undersample data can be reconstructed via mathematical algorithms. Our study aims at understanding the limits of reconstruction algorithms on GPR data and investigate potential cost effective methods to obtain large volumes of GPR in a timely fashion. Previous work in the area have studied prediction error filters to reconstruct GPR data (Martins and Travassos, 2015) following the algorithm proposed by Liu and Fomel (2011). The objective of this work is to apply the methodology called multichannel singular spectrum analysis (MSSA) introduced by Oropeza and Sacchi (2011) as a tool for denoise and reconstruct GPR data.

Methodology

Our data can be mapped to the $f-x$ domain and treated as 1D or 2D spatial series for each temporal frequency ω. Consider a 3-D data set acquired on a regular grid $S(m,n,\omega)$, $n=1,...,Ny$, $n=1,...,Nx$, where m and n are the spatial positions. For instance, n can be receiver position and m source position. If some traces are missing, a spatial slice is given by the following matrix.

$$\mathbf{S}^{obs} = \begin{pmatrix} S(1,1) & S(1,2) & 0 & S(1,4) \\ 0 & S(2,2) & S(2,3) & S(2,4) \\ S(3,1) & S(3,2) & 0 & 0 \\ S(4,1) & 0 & S(4,3) & S(4,4) \end{pmatrix} \quad (1)$$

Where unobserved samples are replaced by zero. The last equation can be represented by the element wise product of the complete data and the observed data (Oropeza and Sacchi, 2011)

$$\mathbf{S}^{obs}(m,n,\omega) = T(m,n)\mathbf{S}(m,n,\omega) \quad (2)$$

where, $T(m,n)=1$ when the data bin in the observed volume contains and observation and $T(m,n)=0$ when the bin is empty. The following numerical algorithm has been proposed to estimate the missing data (Oropeza and Sacchi, 2011).

$$\begin{aligned} \mathbf{S}^0 &= \mathbf{S}^{obs} \\ \mathbf{S}^v &= a\mathbf{S}^{obs} + (1-aP_T)\mathscr{F}\mathbf{S}^{v-1}, v=1,2,3,...,n \end{aligned} \quad (3)$$

where, $\mathscr{F} = P_A P_R P_H$ represents the MSSA filter. The operator averaging P_H indicates Hankelization, the operator P_R indicates rank reduction and the operator P_A is block diagonal averaging (to undo the Hankelization process). Hankelization entails turning the 2D spatial data into a level 2 block Hankel matrix. It is easy to show that for a supposition of dips, the the level 2 block Hankel matrix of the data is a low rank matrix with rank proportional to the number of dips. We often apply the algorithm in spatio-temporal windows to limit the number of dips in the data. The parameter a is equation 3 is used to control the reinsertion of noisy data. For noise-free data $a=1$. We adopt $a=0.4$ for real data sets. In essence, the algorithm is not only used to reconstruct data but also to enhance their SNR.

GPR Synthetic Data

To test the efficiency of the method we resort to a numerical modeling example of GPR data. We use the GprMax software (Giannopoulos, 2005) to model GPR data. We run a simple model in stratified layers following this sequence of permittivity: $\varepsilon_1 = 1$, $\varepsilon_2 = 6$, $\varepsilon_3 = 3$, $\varepsilon_4 = 20$ and $\varepsilon_5 = 12$. The data for this particular model are displayed in Figure 1. In Figure 1 (on the left side) is a schematic illustration of the model used to compete the synthetic data. Whereas on the on the right side is the result of modeling the data in terms of two way time (TWT) in nanoseconds (ns). The model simulates a Common Shot (*CS*) profile where the source (*S*) stays at $0.2m$ and the receiver (*R*) varies between $0.7m$ to $3.8m$ with a $0.005m$ step increment.

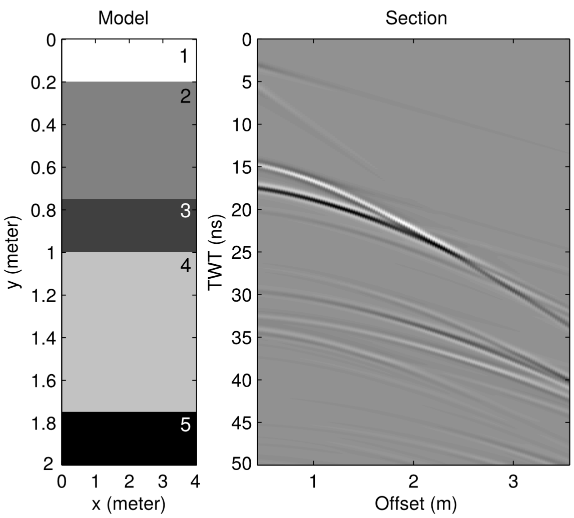

Figure 1: Illustrative model used to test our interpolation algorithm.

Model Testing

We demonstrate the effectiveness of the methodology with 2-D section. Traces where decimated via random sampling with decimation factor varying from 10% to 90%. The Figure 2A shows the data for 50% decimation and Figure 2B shows the result of the reconstruction via MSSA.

To quantify the quality of our results we define the parameter Q

$$Q = 10\log_{10}\frac{||D||_2^2}{||D - D_e||_2^2} \qquad (4)$$

Where D indicates the true data prior to decimation and D_e the estimated data. High values of Q indicate high-quality reconstructions. The Figure 2C, represents Q versus level of decimation. High-quality results are obtained for level of decimations up to 80%.

Real Case

After testing the efficiency of the MSSA method with a synthetic data example, we carried out tests with a real GPR data set. These data corresponds to fieldwork in the sedimentary formations of Santa Catarina Island, Brasil. The sediments are associated to intense sea transgression and its subsequent regression. The sediments are unconsolidated deposits of eolian, alluvial, lacustrine and marine sands with less than 5% of silt and clay.

We display results for one CMP but bear in mind that the full data was reconstructed. The data was collected with

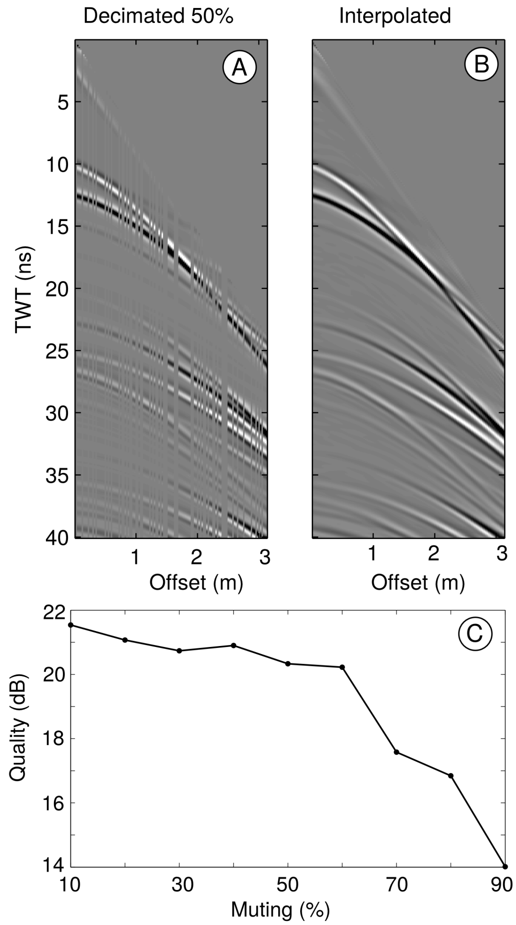

Figure 2: Model experiments. In **A** model with 50% of null traces, **B** Reconstruction and **C** a graph showing quality of reconstruction versus decimation factor.

a PULSE EKKO GPR 100 MHz unshielded antennae with trace increment of $0.2m$ and $300ns$ of time window.

The field data was decimated to test the algorithm. Again, decimation factor between 10% to 90% where utilized. Figure 3A shows the real data with 50% decimation and 3B shows recovered section via the MSSA method.

The results in Figure 3C are quite similar to the original results. Noise have attenuated because we have included noise attenuation capability in our reconstruction algorithm by adopting $a = 0.4$.

Conclusions

We discussed the application MSSA reconstruction and noise attention to georadar data. MSSA was a applied to modeled GPR data and to a field data set. We have obtained high-quality reconstruction for data that has decimated by 80%. These results smotivate us to continue to work on reconstruction methods for random acquisition of georadar data and used these type of data to test algorithms that explode randomize acquisiting to decrease acquisition cost.

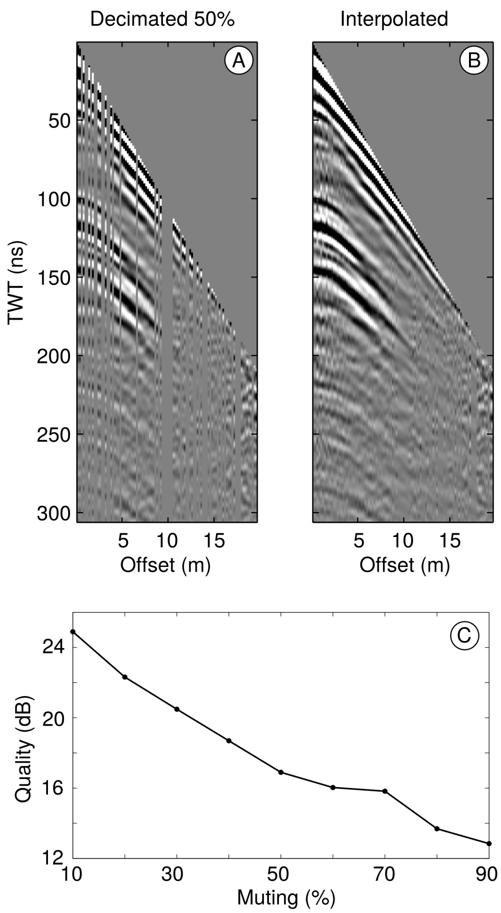

Figure 3: Real Data. In **A** model with 50% missing data, **B** reconstruction results and **C** a graph showing the quality of reconstruction versus decimation.

Acknowledgments

We thank Oropeza, V. and Naghizadeh, M. for their most appreciated help on using their MSSA methodology. S.S. Martins is sponsored by CNPq Brazil, J.M. Travassos has CNPq grants. We thank the Signal Analysis and Imaging Group at the Department of Physics of the University of Alberta for supporting this research.

References

Abma, R., and N. Kabir, 2005, Comparisons of interpolation methods: The Leading Edge, **24**, 984–989.

Giannopoulos, A., 2005, Modelling ground penetrating radar by GprMax: Construction and Building Materials, **19**, 755–762.

Liu, B., and M. D. Sacchi, 2004, Minimum weighted norm interpolation of seismic records: Geophysics, **69**, 1560–1568.

Liu, Y., and S. Fomel, 2011, Seismic data interpolation beyond aliasing using regularized nonstationary autoregression: Geophysics, **76**, V69–V77.

Macheret, Y. Y., M. Y. Moskalevsky, and E. V. Vasilenko, 1993, Velocity of radio waves in glaciers as an indicator of their hydrothermal state, structure and regime: Journal of Glaciology, **39**, 373–384.

Martins, S., and J. Travassos, 2015, Interpolating wide-aperture ground-penetrating radar beyond aliasing: Geophysics, **80**, H13–H22.

Murray, T., D. L. Gooch, and G. W. Stuart, 1996, Structures within the surge front at bakaninbreen, svalbard, using ground-penetrating radar: Annals of Glaciology, **24**, 122–129.

Murray, T., G. W. Stuart, M. Fry, N. H. Gamble, and M. D. Crabtree, 2000, Englacial water distribution in a temperate glacier from surface and borehole radar velocity analysis: Journal of Glaciology, **46**, 389–398.

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

Spitz, S., 1991, Seismic trace interpolation in the *F-X* domain: Geophysics, **56**, 785–794.

Trad, D., 2009, Five-dimensional interpolation: Recovering from acquisition constraints: Geophysics, **74**, V123–V132.