Diffusion semigroups: A diffusion-map approach to nonlinear decomposition of seismic data without predetermined basis

August Lau*, Chuan Yin, Apache Corporation; Ronald R. Coifman, Yale University; Anthony Vassiliou, GeoEnergy, Inc

Summary

We present a nonlinear processing methodology designed to separate out various structures in the seismic data. This methodology builds on the strategy of decomposing seismic data into simple and complex parts (see Lau et al 2008) in which a variational minimization approach extracts the simple structure and complex structure. This paper presents an alternative construction using diffusion semigroups. It is a nonlinear method like the variational method. But it does have the advantage of controlling the components in the nonlinear decomposition . We show that by using various diffusion geometric tools we can build nonlinear filters which enable a decomposition of the data into various intrinsic substructures designed to facilitate interpretation tasks.

Introduction

It is very useful to decompose seismic data (prestack or poststack) into different components. The applications of such decompositions can be found in signal processing, data compression, coherent energy useful for structural interpretation, complex stratigraphy, etc. Seismic decompositions can roughly be divided into two categories of algorithms.

The first category is using a predetermined basis. Examples of such method are Fourier decomposition and wavelet transform. The basic building blocks have definite geometric shapes. It yields fast decomposition. But in general, these are linear decompositions where the superposition of the components yields the original input data.

The second category is projection methods like SVD (singular value decomposition) which defines the projections (basis elements) based on the input data. The geometric shape of each projection is not predetermined. The approach here is that complex objects cannot be fitted into a predetermined basis. The connectivity and conductivity of objects like network or images determine the projections (basis elements) to be used. The geometry of each projection is data dependent.

Diffusion geometry using the Markov semigroup approach belongs to the second category. It employs the concept of affinity in the seismic data to determine the local geometry. Then the local geometry determines the diffusion semigroup which in turn determines the global geometry of the data.

Methodology

To motivate diffusion geometry and semigroup, we will start with a 2-D image first. The left side of figure 1 shows a crossplot which looks like an inner dense circle and outer ring of dense points. The input crossplot is nonlinear in the sense that no line could separate the inner and outer circle. The first thing to do with a diffusion geometry is to define affinity. Affinity is analogous to a generalized distance function. Once that is done, we will use a transformation of the affinity to a Euclidean geometry (see right side of figure 1). The mathematics can be found in Coifman et al (Proceedings of National Academy of Science 2005). Notice that in the transformed space on the right, the inner and outer circles can be separated roughly into linear components (left and right sides of the transformed plot of the first 3 eigenfunctions).

We want to show another crossplot example in figure 2 to further illustrate the change in the geometry of the original input data versus the transformed diffusion coordinates. The long term diffusion of heterogeneous material is remapped. The left circular region in the input data has a higher proportion of heat conducting material, thereby reducing the diffusion distance among points (top to bottom). In contrast, the bottle neck in the middle increases that distance between the two lobes, and shows that the right lobe is more uniformly distributed. The idea here is that Euclidean distance in the original crossplot does not tell us how close things are in a diffusion geometry, while the right image provides an intrinsic assessment. To achieve this we need to define an appropriate affinity (or distance) and transform it to a domain where Euclidean distance is more informative.

In particular we show that by considering for a given seismic volume the collection of all 3-D data blocks of size 9x9x9, as a data base, to be organized through diffusion geometry we can extract and separate reflections and various other structures. A simple Markov matrix that defines the diffusion on the collection of blocks can be defined as follows:

Diffusion semigroup

$$A_{p,q} = \frac{\exp(-\left\| v(p) - v(q) \right\|^2 / \varepsilon)}{\sum_q \exp(-\left\| v(p) - v(q) \right\|^2 / \varepsilon)} \quad \text{where } v(p)$$

is the 9x9x9 volume around location p.

The eigenvectors of A enable the construction of nonlinear filters on the seismic volumes.

Processing steps

The dataset is a marine data set after PSTM (prestack time migration) and the output gathers are offset gathers. A 4-dimensional volume is input to the diffusion semigroup program. The 4 dimensions are inline, crossline, offset and time. We will show what the diffusion geometric eigenfunctions look like in the stack domain. The diffusion is done on all 4 dimensions in the prestack domain but it is easier to show the results in the stacks.

.

Step 1. Check conventional decomposition like Fourier domain or wavelet domain. See figure 3 for different bandwidths of the input stack. This gives us some idea of how to drive the higher eigenfunctions. It is an important QC plot to show when high frequencies will break up in terms of interpretation.

Step 2. Determine the distance definition (affinity) for the seismic data. This method is not restricted by dimensionality. It can be used for gathers to find subtractions or similarities in a prestack domain. But we will illustrate the eigenfunctions in the stack domain after diffusion semigroup decomposition as in figure 4. Only 3 eigenfunctions are displayed.

Step 3. Calculate the cumulative eigenfunctions. This is to organize the various eigenfunctions (projections) so that progressive cumulations of the eigenfunctions give progressively more details. Figure 5 shows only three cumulations. So cum-2 would mean the cumulation of the first two eigenfunctions, cum-5 the cumulation of the first 5 eigenfunctions, etc.

Step 4. To calculate the complex part of the cumulations, test various differences between the cumulations. The complex part is the difference cube which allows detailed stratigraphic interpretation or detail fault patterns. See figure 6 right most panel.

Discussion

We can compare the diffusion semigroup decomposition with previous nonlinear decomposition method. One such decomposition is the variational decomposition mentioned in the beginning. This is the complex decomposition of seismic data into simple and complex part where the simple part is the part of the data which obeys the VN (variational norm which is a mixed L1 and L2 norm). The result from VN is shown in figure 8. The complex part (rightmost panel) of figure 8 should be compared with the "complex part" of the diffusion semigroup which is the rightmost part of figure 6.

Another point of discussion is that the cum-2 (first 2 eigenfunctions) in figure 5 carries almost all the salient features of the seismic image. It is extremely powerful to use only 2 eigenvalues to characterize the full picture. Other projection methods like SVD cannot comprehend the geometric nature of data. Fourier decomposition (low pass in figure 3) cannot capture the salient features.

For completeness , we show Figures 8-10 which are the time slices for different cubes with different groupings of the eigenfunctions. They are analogous to time slices of low, mid and high "bandwidths" which comprise of the low, mid and high grouping of eigenfunctions.

Application to processing parameters

Instead of 1000 points in a crossplot as in figure 1 and figure 2, we could generate 1000 cubes very easily in seismic processing. We can use 3 parameters like velocity, mute and filtering. We can parametrize velocity (10 different velocity functions) and mute (10 different inside/outside mutes) and filter (10 different bandwidths). Then we have 1000 stack cubes to view since we have combinations of 10x10x10. Diffusion semigroup will organize these 1000 cubes in the transform space so that one could view just 10 to 20 cubes rather than 1000 cubes.

Our traditional decision process is to fix one parameter at a time. We can start with velocity. Make our pick. Then go to the next parameter like mute. Make our pick. Then finally bandwidth. Make our pick. We cannot see the various combinations of velocities, mutes and bandwidths simultaneously if we use the traditional sequential decision process.

Conclusion

Diffusion semigroup is a nonlinear decomposition method which is extremely fast and is amenable to handling high dimensional seismic data. Let us take the example of time lapse data. If we have surveys which are parametrized by inline, crossline, offset, azimuth, depth and time lapse, that would be a 6-D problem. Diffusion semigroup will project the 6-D volume into different eigenfunctions which can be used to do subtractions or similarities. If other attributes are involved outside the normal seismic reflectivities, there will not be any re-coding of the diffusion semigroup. These other attribute cubes could be added with the same algorithm.

Acknowlegments

We want to thank Apache Corporation for continued support internally as well as support for Yale research. We want to thank PGS for permission to show the seismic data. Last but not least, we want to thank Gregg Barker of Apache for his insights in seismic interpretation.

References

Chiu, S. K. and J. E. Howell, 2008, Attenuation of coherent noise using localized-adaptive eigenimage filter, SEG Annual meeting expanded abstracts, 2541-2545.

Coifman, R.R., S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, S. W. Zucker, 2005, Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps, Proceedings of the National Academy of Sciences of the USA, 7426-7431.

Kambites, M. and B. Steinberg, 2005, Wreath product decomposition for triangular matrix semigroups, arXiv:math.RA/0512024 v1. 1 Dec 2005

Lau, A., 1979, Images of compact 0-dimensional semigroups, Colloquium Mathematicum, No. 2, Vol XL, P.199-222.

Lau, A., M. Roque-Sol, C. Lapilli, J. Perdomo, C. Shih, A. Gonzalez, 2008, Imaging with complex decomposition: Numerical applications to seismic processing in difficult areas, SEG Annual meeting expanded abstracts, 1986-1990.

Li, C-F and C. Liner, 2008, Wavelet-based detection of singularities in acoustic impedances from surface seismic reflection data, Geophysics, vol. 73, No. 1 (Jan-Feb 2008) p. V1-V9.

Diffusion semigroup

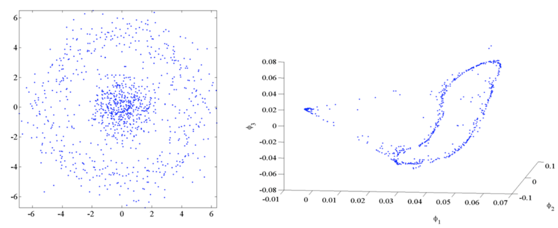

Figure 1
Left = input crossplot with inner circle and outer circle
Right = transforming the diffusion geometry into Euclidean space

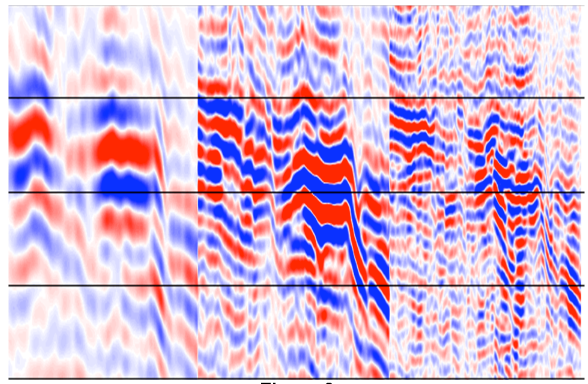

Figure 3.

low pass mid high

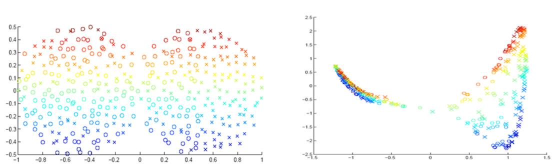

Figure 2
Left = input crossplot (two circles and a bottleneck in the middle)
Right = transforming the diffusion geometry into Euclidean space

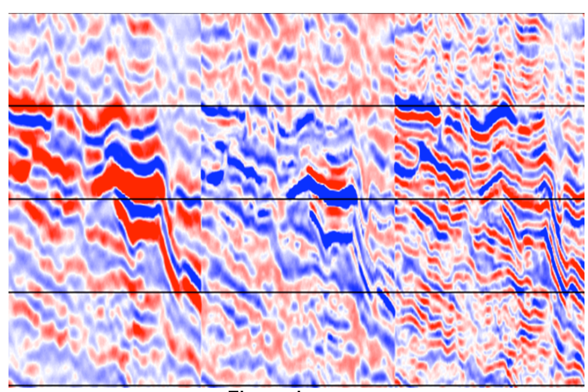

Figure 4.

Eigenfunction #2 eigenfunction #5 eigenfunction #8

Diffusion semigroup

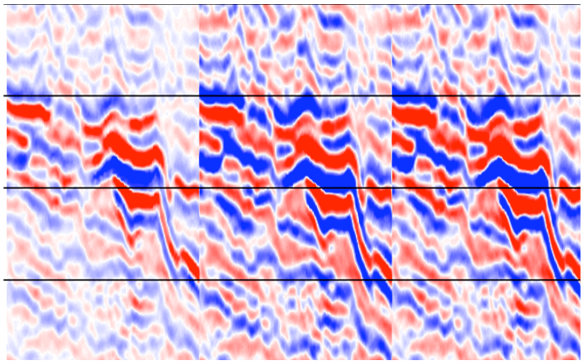

Figure 5.

Cum-2 Cum-5 Cum-8

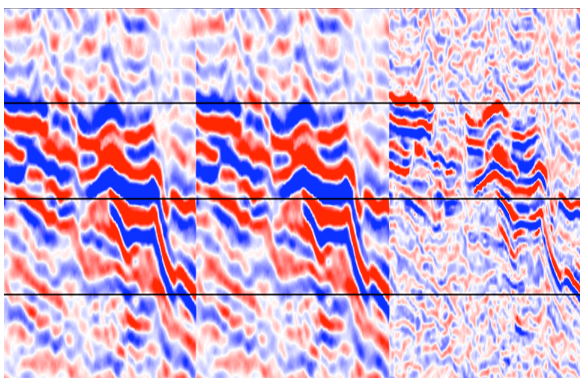

Figure 6. DIFFUSION METHOD

Cum-11 Cum-14 Complex part (14-11)

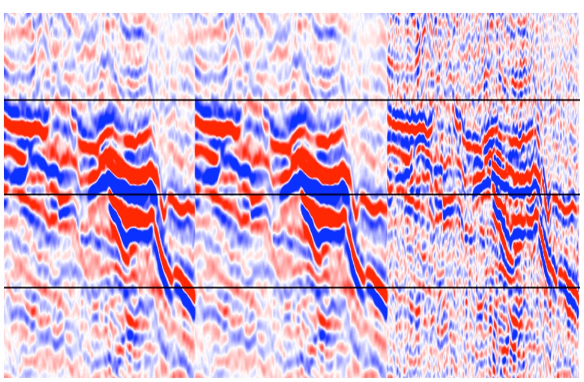

Figure 7. VARIATIONAL METHOD

Input stack Simple part Complex part

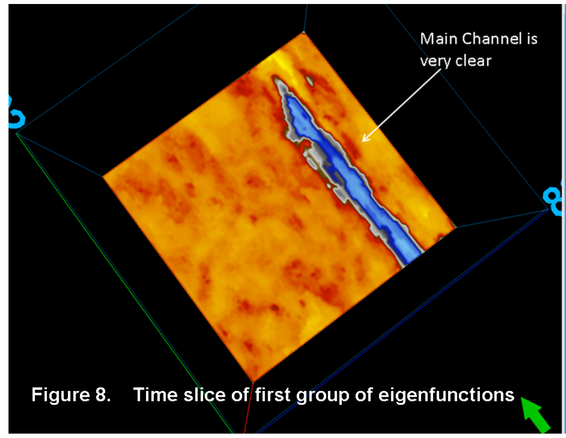

Figure 8. Time slice of first group of eigenfunctions

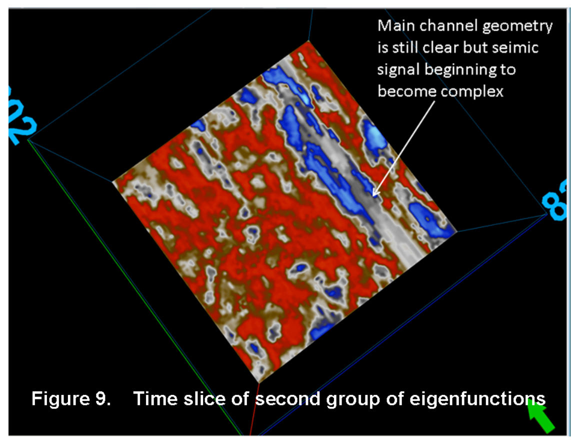

Figure 9. Time slice of second group of eigenfunctions

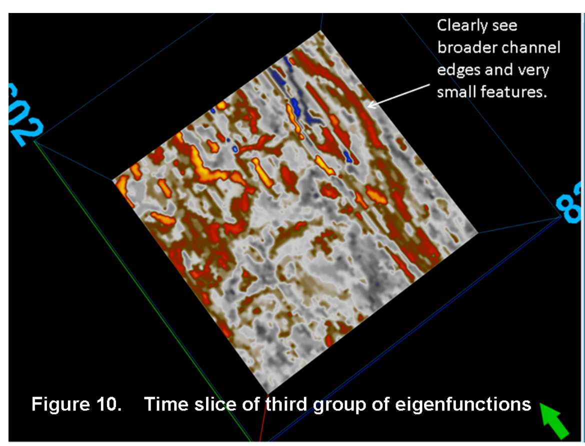

Figure 10. Time slice of third group of eigenfunctions