Angle gather recovery using Iterative Soft Thresholding

Robert G. Clapp

ABSTRACT

Multi-dimensional angle gather construction for AVA and velocity analysis is a computational challenge due primarily to the accompanying increase in volume size which forces the gathers to be stored in a computationally more expensive memory level. Compressive sensing can be used to mitigate this challenge as long as the full angle gathers can be successfully recovered. Multi-dimensional wavelet transforms is a sufficiently sparse basis function to allow for 90% reduction in needed correlations. A modified version of iterative soft thresholding, which applies a different thresholding approach to different wavelet levels, proves a successful l_1 inversion scheme.

INTRODUCTION

Angle gathers are a useful tool byproduct of migration for both velocity and rock property analysis. There are several different methods for producing angle gathers (Crawley et al., 2012), but one of the most useful is based on correlating shifted version of the source and receiver fields (Sava and Fomel, 2003, 2006). The problem with this approach, particularly when creating multi-dimensional gathers, is that the volume size of the imaging domain increases by one to three orders of magnitude making the dominant cost reading/writing to distant memories (from main memory rather than a cache, across the PCI Bus, or from disk).

Donoho (2006) offers an approach termed *compressive sensing* as a potential solution to this computation and storage problem. In compressive sensing, a random subset of the desired measurements is made. An inversion problem is then set up to estimate in an ℓ_1, or preferably ℓ_0, sense, a sparse basis function that fully characterizes the desired signal. Clapp (2011) showed that sub-surface offset gathers are highly compressible using the multi-dimensional wavelet transform. Clapp (2012) showed promising recovery of angle domain gathers using a Stagewise Orthogonal Matching Pursuit algorithm (StOMP) (Donoho et al., 2006). The results were encouraging and were successful enough to be used as a tool to reduce the cost of velocity estimation(Zhang et al., 2013). Zhang et al. (2013) was only able to achieve a factor of two reduction of sub-surface offset correlations before seeing degradation in the result. Hennenfent and Hermann (2005) offers another ℓ_1 inversion approach called Iterative

Soft Thresholding. This approach characterizes the estimation problem as a non-linear problem eliminating small model components (thresholding) at each iteration. The thresholding is relaxed as a function of non-linear iteration.

In this paper I apply the IST algorithm to the angle gather estimation problem. I show that, in conjunction with a modified sampling criteria and a wavelet-level based thresholding scheme, a higher level of data reduction is possible.

COMPRESSIVE SENSING

Compressive sensing is a statistical technique whose start is usually traced back to Donoho (2006), but whose start could be placed as early as the basic pursuit work of Mallat and Zhang (1993). Compressive sensing is simply a missing data problem with a well chosen preconditioning operator (resulting in a sparse basis model/function) solved with a ℓ_0/ℓ_1 solver. Using the conventional SEP nomenclature of fitting goals it can be written as

$$\mathbf{0} \approx \mathbf{W}(\mathbf{d} - \mathbf{S}\mathbf{p}) \tag{1}$$

where $\mathbf{d}$ is the data, $\mathbf{W}$ is the masking operator (composed of 0s and 1s), $\mathbf{S}$ is the preconditioning operator, and $\mathbf{p}$ is the preconditioning variable.

Compressive sensing theory states that if you can find a transform operator which makes $\mathbf{p}$ consist of only n number of unknowns as long as you collect kn *random* data points, where k is usually in the range of 5, you can fully recover p. To reduce the amount of data points significantly $\mathbf{S^T}$ must transform into a basis function that is very sparse. One such choice, using the multi-dimensional wavelet transform (Villasenor et al., 1996), proves to produce a sparse basis function for subsurface offset gathers. Clapp (2011) showed that by throwing away up to 98% of $\mathbf{p}$, can still fully recover a sub-surface offset gather (see Figures 1 and 2). The problem that both Clapp (2011, 2012) attempted to address was what ℓ_1 approach to use to find the sparse basis function.

IST

Iterative soft thresholding is a greedy algorithm along the same lines as matching pursuit (Donoho, 2006), and to some extent the Stagewise Orthogonal Matching Pursuit (StOMP) (Donoho et al., 2006) stomp algorithm. It attempts to find the ℓ_1 solution to the problem by slowly allowing more and more non-zero elements into the basis function. There are three basic parts to the IST algorithm: find the step size, finding the gradient, and soft thresholding.

The step size calculation relies on finding the largest eigenvalue of the Hessian of the ℓ_2 problem, or $\mathbf{S^T W^T W S}$. This is done using the Rayleigh power method.

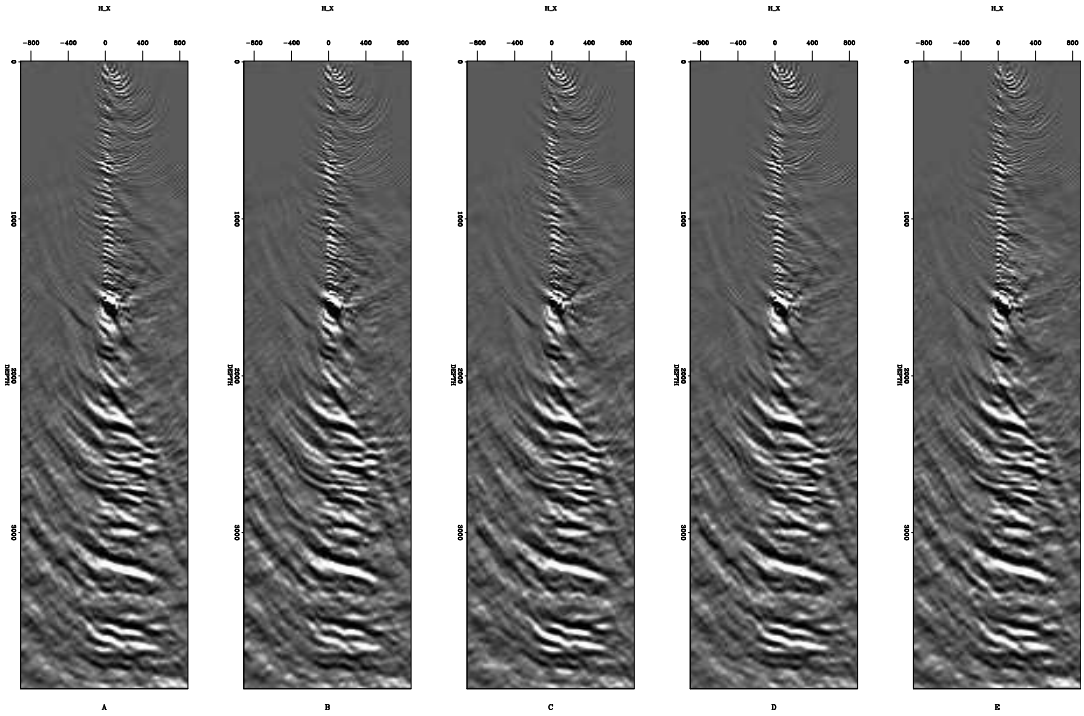

Figure 1: Five neighboring subsurface offset gathers. B and C are one midpoint in X before and after A. E is one midpoint in Y before A. Note the spatial similarity of these gathers. [**ER**]

Algorithm 1 Rayleigh power method

Fill $\mathbf{g_0}$ with random numbers
for Rayleigh iteration i **do**
  $\mathbf{g_i} = \mathbf{S^T W^T W S g_{i-0}}$
  $s = \mathbf{g^T g}$
  $\mathbf{g_i} = \frac{\mathbf{g_i}}{\mathbf{s}}$
end for

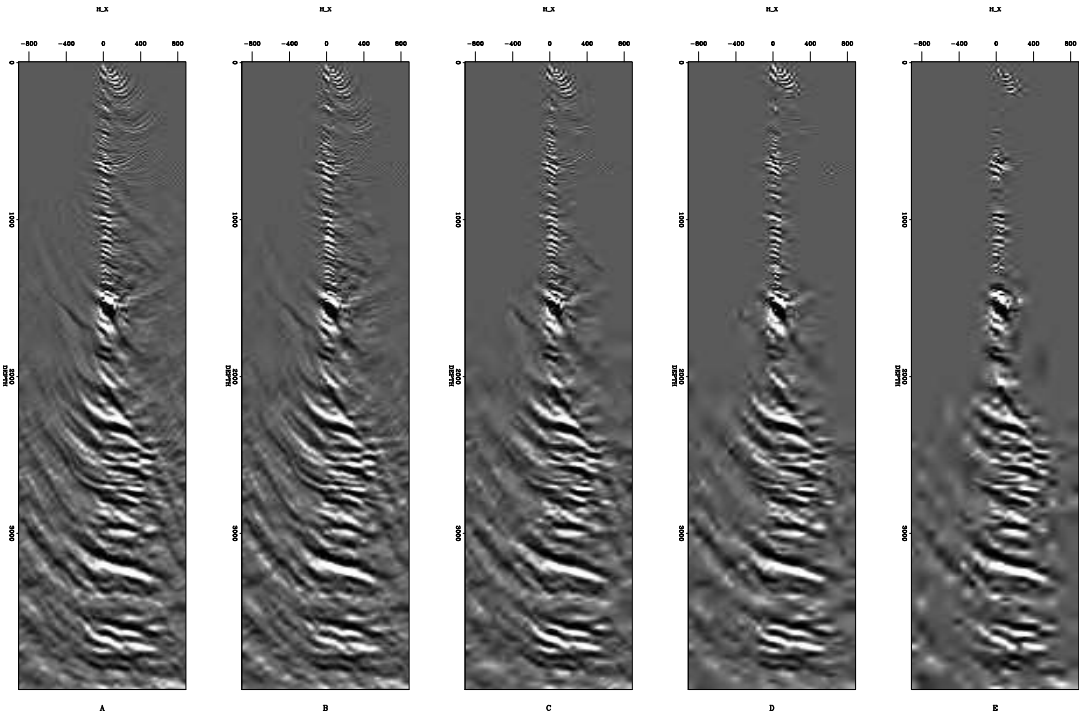

Figure 2: The result of zeroing the smallest values of the wavelet domain representation. All five panels show the same sub-surface offset gather shown in Figure 1. A shows the result of clipping 90% of the values; B, 95%; C, 98%; D,99%, and E,99.5%. Note how the reconstructed gather is nearly identical up to a 98% clip. [**ER**]

The basic algorithm can be seen in algorithm 1. After a few (20 or so) iterations s converges to largest eigenvalue and $\mathbf{g}$ to the largest eigenvector.

The quantity $\frac{1}{s}$ is used in a modified steepest descent algorithm shown in algorithm 2. The unique portion of the above algorithm is the thresholding step. The

Algorithm 2 IST algorithm

$\mathbf{x} = \mathbf{0}$
for Non-linear iteration **do**
  for Linear iteration **do**
    $\mathbf{gg} = \frac{\mathbf{WSx}}{\mathbf{s}}$
    $\mathbf{g} = \frac{\mathbf{S^T W^T gg}}{\mathbf{s}}$
    $\mathbf{x} = \mathbf{x} + \mathbf{g}$
    Threshold $\mathbf{x}$
  end for
end for

basic idea is to set $\mathbf{x}$ values below some amplitude to zero. The threshold decreases as a function of non-linear iteration *in*. Hennenfent and Hermann (2005) suggests a decreasing thresholding based on percentiles in the initial gradient. For example throwing the bottom 99%, 98%, 95%, 90%, 80%, 60%, 35%, 0% of data values at successive iterations. Figure 3 shows a portion of the multi-dimensional wavelet domain before and after thresholding at the 98% level. Note how even though nearly all coefficients are zeroed at this level the plots are very similar.

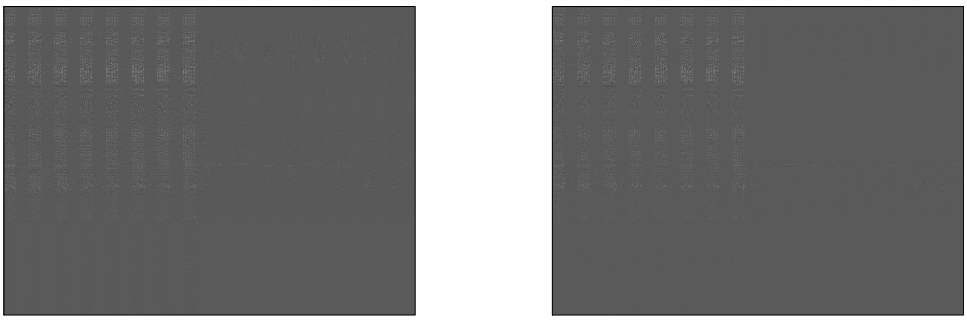

Figure 3: The left panel shows the result of performing a multi-dimensional wavelet transform on the gathers shown in Figure 1. The right panel is the result of zeroing the bottom 98% of the wavelet values. [**ER**]

Figure 4 shows the result of performing angle gather transform on the gathers seen in Figure 1. Note how most of the gathers are generally flat. Figure 5 shows an example of following the above algorithm to recover the basis function after randomly throwing away 90% of the correlations. Figures 6 and 7 show the resulting angle gathers. Note how we have nearly perfectly recovered the subsurface offset, and the corresponding angle gathers, at larger depths, but have not recovered as well shallower

depths. A closer look at the sub-surface offset gathers in Figure 1 shows the problem with a purely random sampling of the sub-surface gathers. Note how at shallower depths no coherent energy maps to large offsets and shallow depths due to an accurate velocity model and good focusing. Therefore at these depths we are not obeying the sampling criteria of compressive sensing.

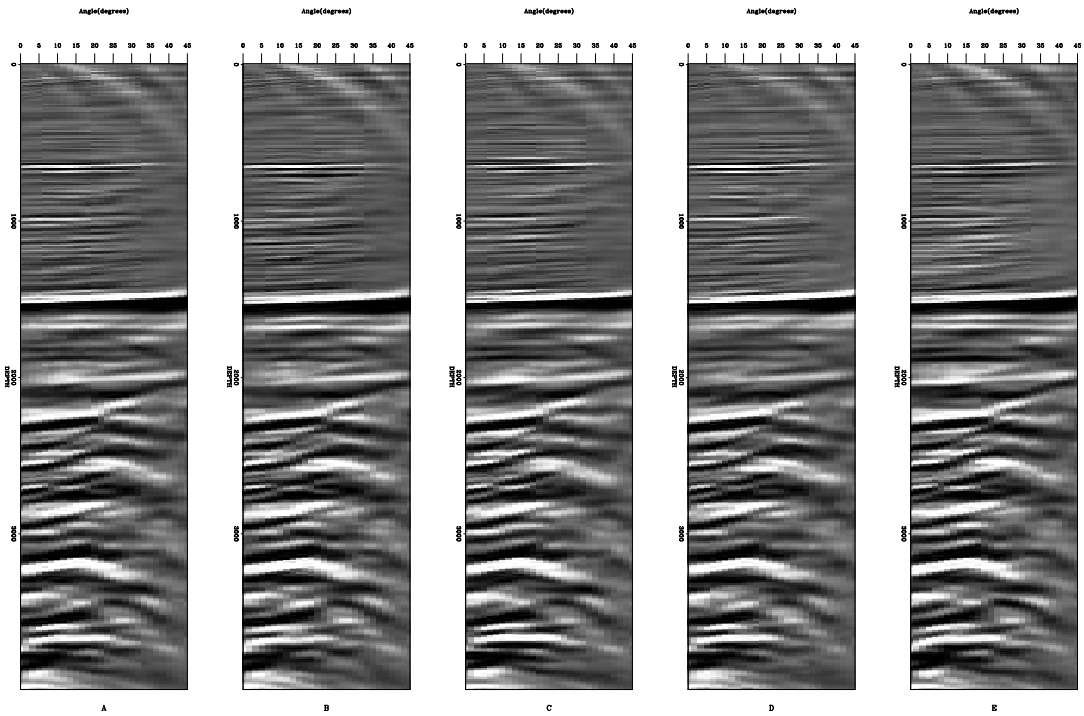

Figure 4: The result of performing an angle domain transform on the gathers shown in Figure 1.

Clapp (2012) suggested modifying the random sampling criteria so that only points within an expanding cone are selected. Figure 8 shows one such random sampling. Note how it somewhat tracks coherent energy seen in the full sub-surface offset gathers seen in Figure 1. Figure 9 shows the recovered wavelet domain. Note how more energy is less randomly focused compared to Figure 5 and closer to the wavelet basis shown in Figure 3. The resulting offset gathers, shown in Figure 10, show less random looking noise than the result shown in Figure 6. In addition the shallower depth image is better recovered than what is seen in Figure 6. The resulting angle domain gather, shown in Figure 11, shows marked improvement in the angle domain image at shallower depths. Particularly, notice the event marked 'M' in the panel which has the incorrect amplitude in Figure 6.

Figure 5: The estimated wavelet domain using the standard IST algorithm. Compared to Figure 3, note how the basic structure is recovered but with significant additional energy. [**ER**]

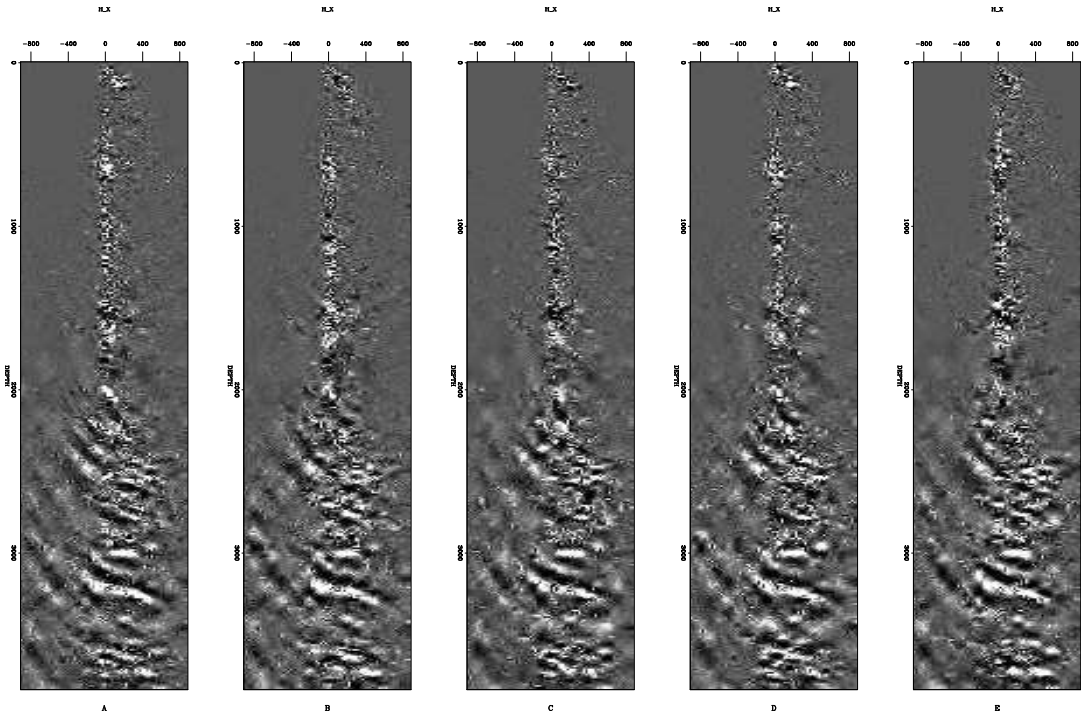

Figure 6: The result of mapping the wavelet domain representation seen in Figure 5 to sub-surface offset and showing the same gathers seen in Figure 1. Note how we have done a nearly perfect job recovering the deeper portion of the image but are not as successful in the upper portion of the image.

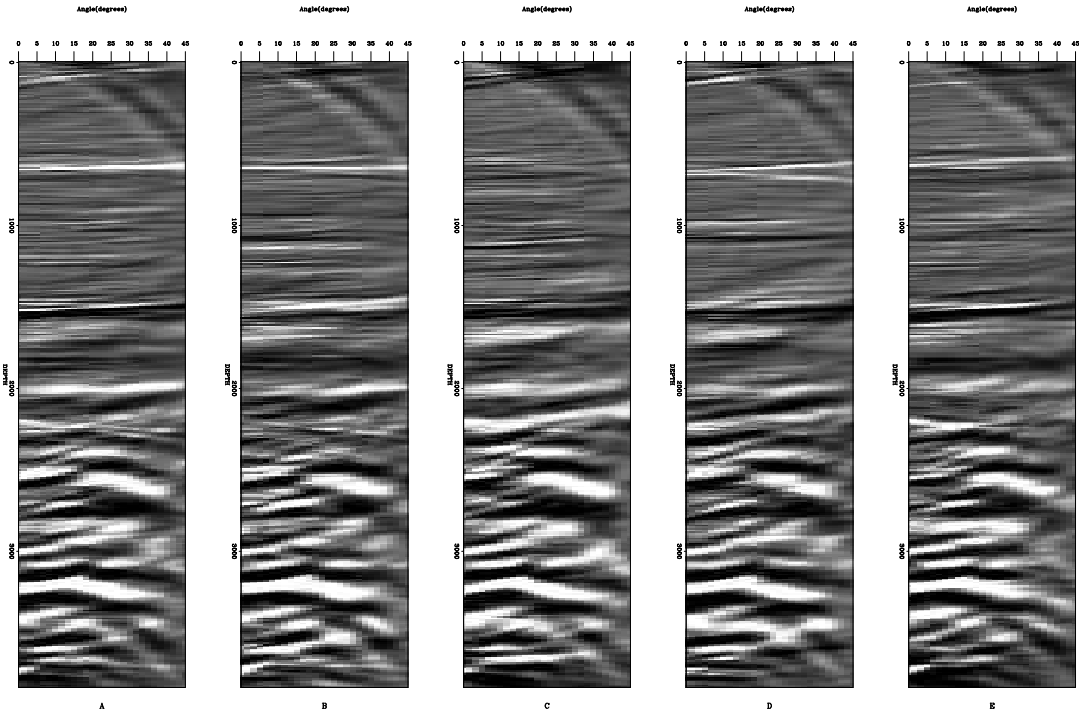

Figure 7: The angle representation of the sub-surface offset gathers shown in Figure 6. Note the errors in the upper portion of the section compared to the full sub-surface based angle gathers shown in Figure 4. [**CR**]

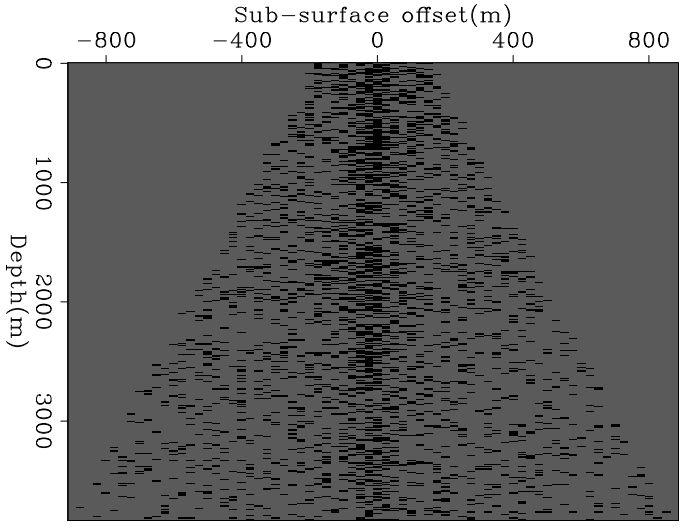

Figure 8: A cone based random sampling pattern which limits the range of possible offsets at shallow depths and tends to concentrate points at smaller offset. [**CR**]

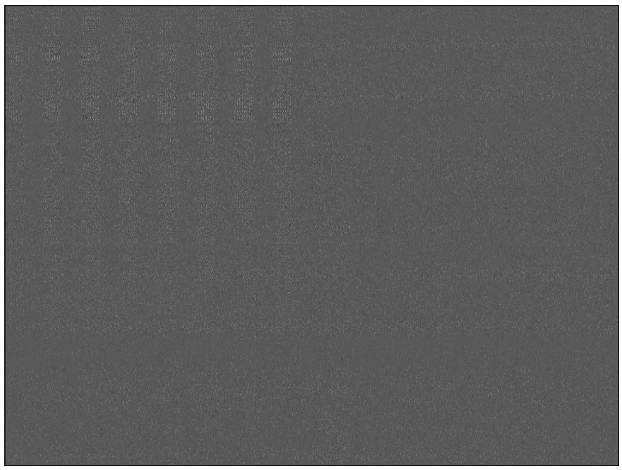

Figure 9: The wavelet basis function using the cone shaped sampling shown in Figure 8. Note how more energy is less randomly focused compared to Figure 5 and closer to the wavelet basis shown in Figure 3. [**CR**]

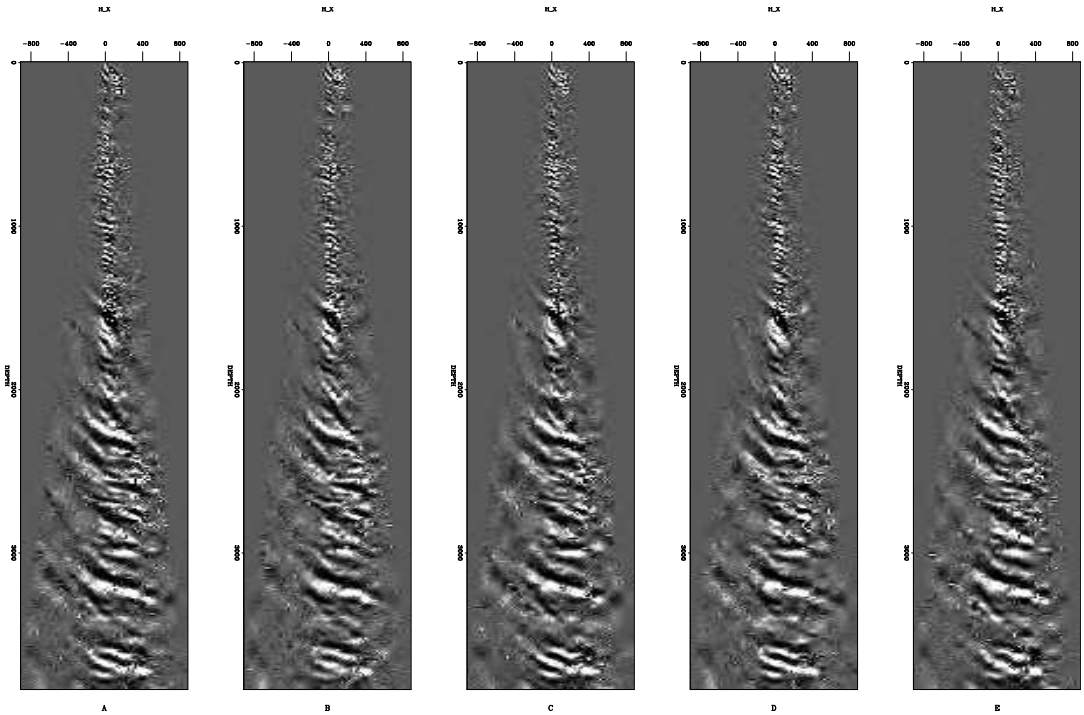

Figure 10: The result of converting to sub-surface offset the wavelet domain seen in Figure 9. Note how we see less random noise than using the standard approach seen in Figure 6. The image at small depth is also improved compared to Figure 6.

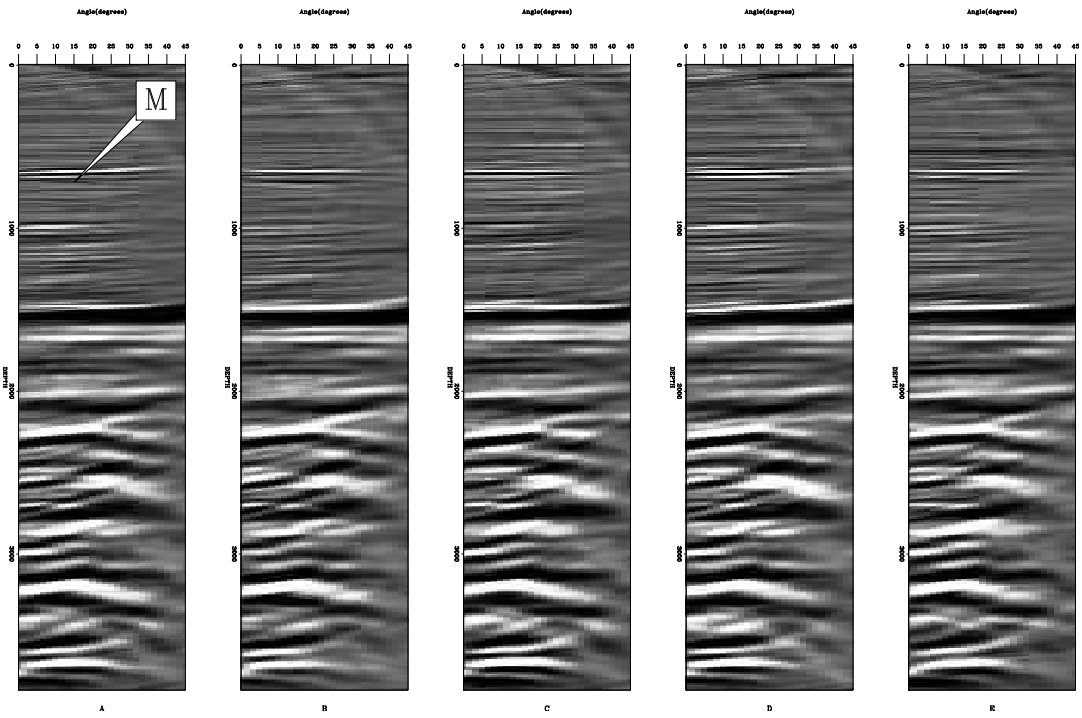

Figure 11: The result of transforming the subsurface offset gathers shown in Figure 10 to the angle domain. Note the improvement particularly at the event marked 'M'. [**CR**]

THRESHOLDING AS A FUNCTION OF WAVELET LEVEL

A wavelet transform involves applying a pair of filters, which recover the low and high frequency (or wavenumber) portion of an image in a time (or space) variant manner. The lower wavenumber portion of the transform can then be sent through the same two filters. The number of times this pair of filters is applied is referred to as a series of levels. In 1-D, each levels high-pass filter produces half as many elements as the previous level. In the multidimensional case, this size decrease pattern is modified. Figure 12[1] demonstrates this phenomenon in 2-D; right and down shows successive lower levels of the high-pass portions of the image. Note the small size of the lowest wavenumber portion of the largest level of the image compared to the first level's high pass portion.

Figure 12 and the left panel of Figure 5 demonstrate another important point of many fields, sparsity generally decreases with increasing levels. Note how the top left portion of each image are very dense with information compared to other levels. Figure 13 demonstrates this concept with a simpler model. In this case a single plane wave has been created in 4-D space panel 'A' and transformed into the

[1]http:www.wikipedia.org/wiki/Discrete_wavelet_transform

multi-dimensional wavelet domain. Panel 'B' and 'C' of Figure 13 show the lowest and highest passed portion of the wavelet domain. Panel 'D' shows a portion of the cumulative distribution function for the lowest, highest, and two intermediate wavelet levels. Note how much more energy is in the lowest levels. The difference in density level suggests that the same thresholding scheme should not be used at every wavelet level.

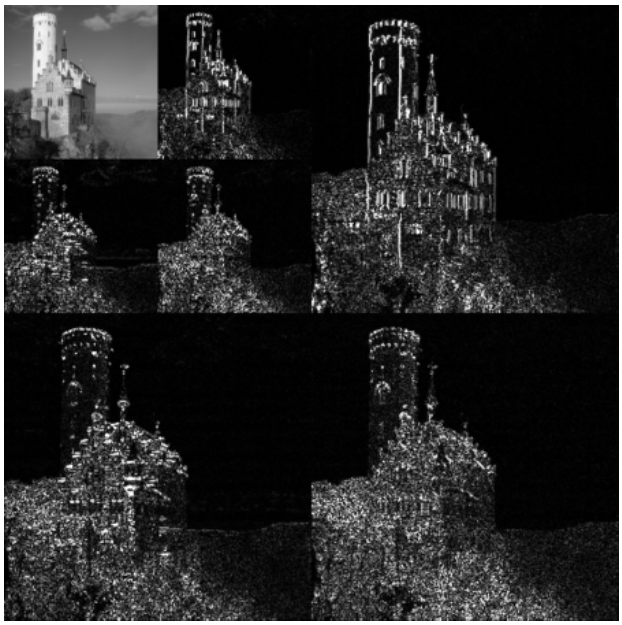

Figure 12: An example of applying a two-dimensional wavelet transform to a grayscale picture of a castle. The bottom-right portion of the image has been low-passed in both x and y; top-left, low passed in x, high passed in y; bottom-left, low passed in y, high passed in. As one moves up and to the left you see the result of higher levels of the wavelet transform. Note how the sparsity decreases as one goes up and to the left. [**NR**]

Instead of using the same thresholding scheme at all levels I modified the thresholding to be level dependent. For each different wavelet level i I then use a different thresholding scheme. For each level I choose a percentile v_i based on the number of samples in the wavelet domain n_i,

$$v_i = 100. - .6\sqrt{3}\frac{n_{\max}}{n_i} \tag{2}$$

where $n_{\max}$ is the maximum size of any level and .6 was determined through trial and error. For each non-linear iteration j, the percentile is decreased,

$$v_{i,j} = max(0., 100. - 1.6 * v_{i,j-1}), \tag{3}$$

where 1.6 was chosen imperically. These percentiles are transformed into thresholding values by taking the $v_{i,j}$ value of the ith wavelet level of $\frac{\mathbf{SWd}}{s}$. Figure 14 shows the same portion of the wavelet domain seen previously. Note how the energy is even more focused than the result seen using only cone sampling in Figure 9. Figure 15 shows the result of converting the wavelet basis function seen in Figure 14 to the subsurface offset domain. Note how we have further reduced random artifacts due to incorrect wavelet basis functions as compared to using simply the cone sampling approach seen in Figure 15. The result angle gathers also show some improvement compared to Figure 16. The amplitudes are closer to the full offset result shown in Figure 1.

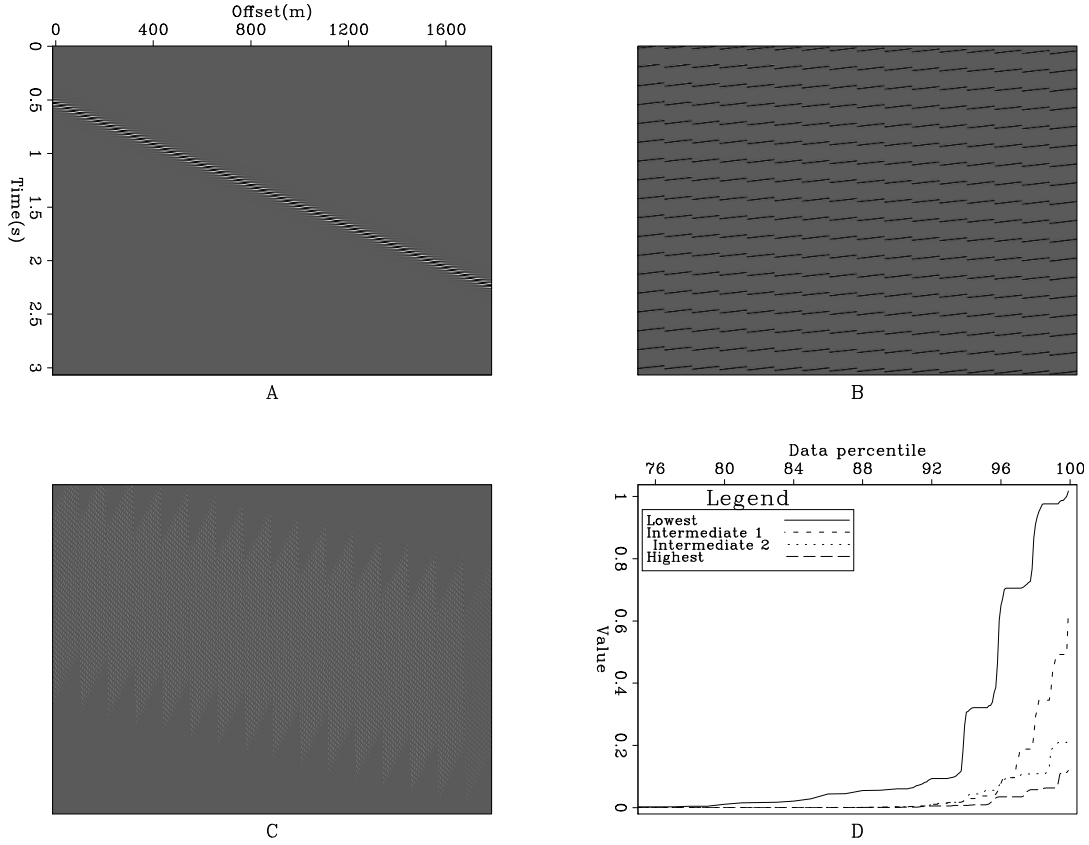

Figure 13: A slice from a plane wave created in 4-D space 'A' and transformed into the multi-dimensional wavelet domain. Panel 'B' and 'C' of Figure 13 show the lowest and highest passed portion of the wavelet domain. Panel 'D' shows a portion of the cumulative distribution function for the lowest, highest, and two intermediate wavelet levels. Note how much more energy is in the lowest levels. [**ER**]

Figure 14: The result of using different thresholding approach at different levels along with cone sampling. Compare how the non-zero coeficients of the basis function are closer to the full subsurface offset result shown in Figure 3 than the results seen in Figure 5 and 9. [**CR**]

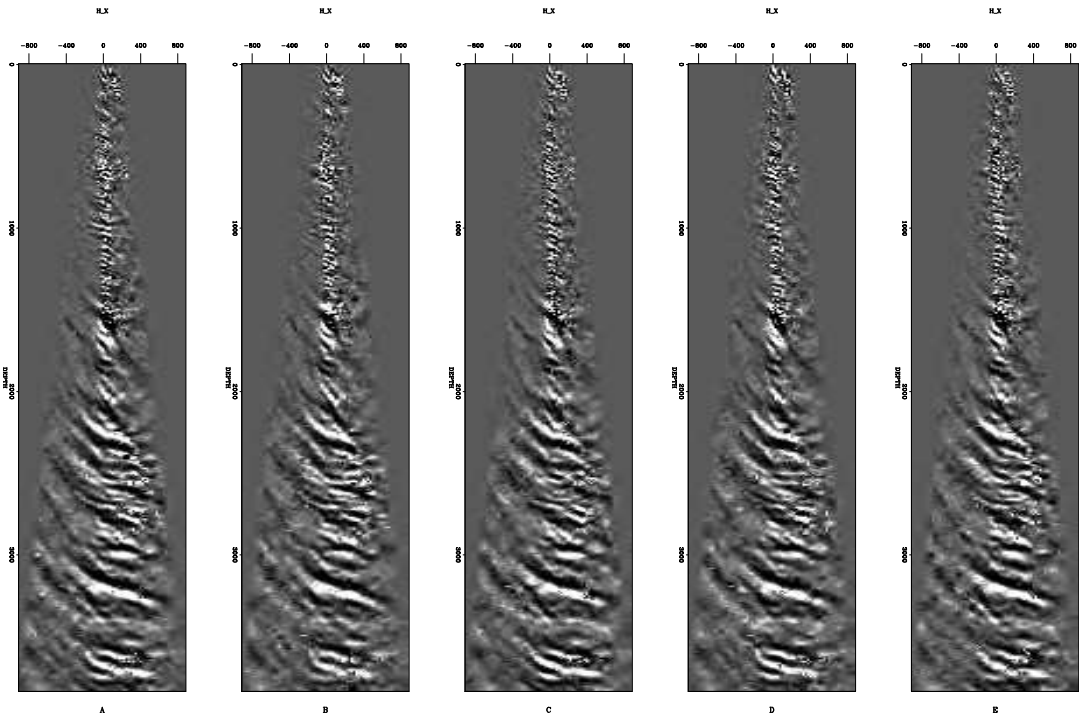

Figure 15: The result of transforming in to subsurface the wavelet basis function seen in Figure 14. Note how there is less random noise throughout the image and more continuity in shallow surface events. [**CR**]

CONCLUSIONS

I present a method for recovering sparsely sample cross correlation based angle gathers using the IST algorithm to recover a sparse multi-dimensional wavelet domain basis function. Modifying the standard random sampling criteria to limit far subsurface offset points at small depths improves model recovery. Further improvement is achieved by use of a level dependent thresholding scheme.

ACKNOWLEDGEMENTS

I would like to thank John Washbourne for his SEP sponsor talk last year, which resulted in a second attempt at this method....without a significant bug. I would like to thank Dave Nichols who suggested the plane-wave analogy for the level based thresholding approach.

REFERENCES

Clapp, R. G., 2011, Imaging using compressive sensing: SEP-Report, **143**, 149–158.
———, 2012, Image gather reconstruction using StOMP: SEP-Report, **147**, 127–138.

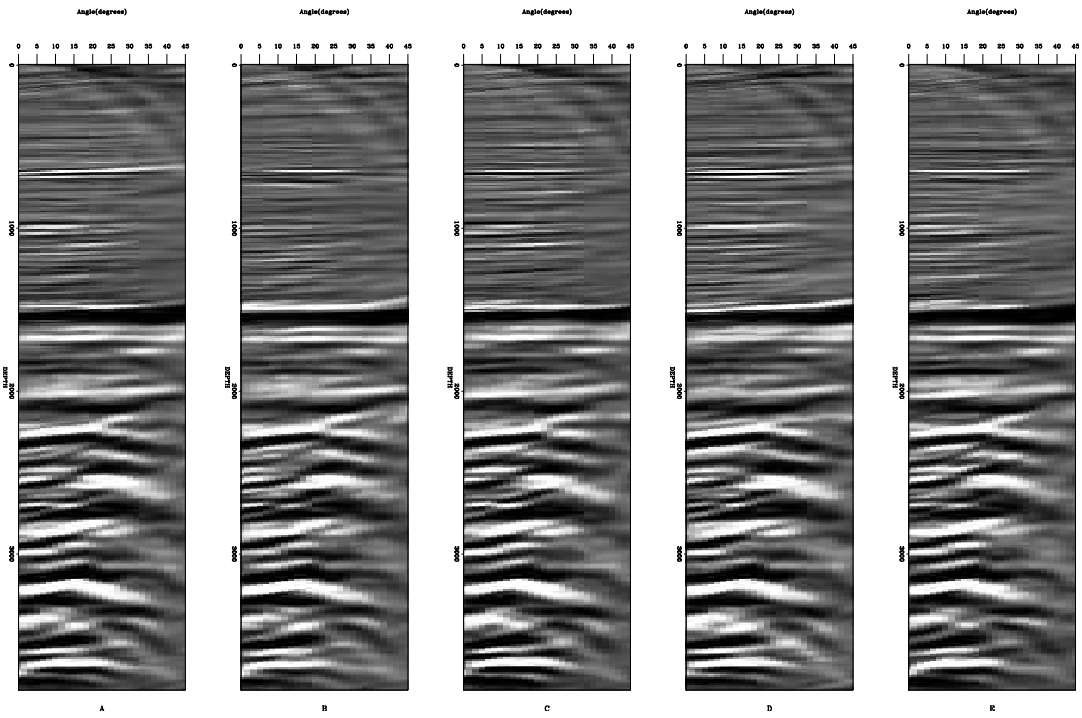

Figure 16: The result of transforming the subsurface offset gathers shown in Figure 15 to the angle domain. The amplitudes are more accurate in the shallow portion of the model than those shown in Figure 7 and 11. [**CR**]

Crawley, S., N. D. Whitmore, A. Sosa, and M. Jones, 2012, Improving rtm images with angle gathers: 82nd Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1–5, Soc. Expl. Geophys.

Donoho, D. L., 2006, Compressive sensing: Information Theory, IEEE Transactions on, **52**, 1289–1306.

Donoho, D. L., Y. Tsaig, I. Drori, and J.-L. Starck, 2006, Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit: Technical report.

Hennenfent, S. and F. Hermann, 2005, Sparsenessconstrained data continuation with frames: applications to missing traces and aliased signals in 2/3d: 75th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 2162–2165, Soc. Expl. Geophys.

Mallat, S. G. and Z. Zhang, 1993, Matching pursuits with time-frequency dictionaries: IEEE Transactions on Signal Processing, **41**, 3397–3415.

Sava, P. and S. Fomel, 2006, Time-shift imaging condition in seismic migration: Geophysics, **71**, 5209–5217.

Sava, P. C. and S. Fomel, 2003, Angle-domain common-image gathers by wavefield continuation methods: Geophysics, **68**, 1065–1074.

Villasenor, J. P., R. A. Ergas, and P. L. Donoho, 1996, Seismic data compression using high-dimensional wavelet transforms: Snowbird, UT, USA, Expanded Abstracts,

396–405, IEEE Computer Society Press.

Zhang, Y., B. Biondi, and R. Clapp, 2013, Accelerating residual-moveout-based wave-equation migration velocity analysis with compressed-sensing: SEP-Report, **149**, 95–106.