BEAMFORMING ON ACOUSTIC TRANSIENTS IN THE DISCRETE ORTHONORMAL STOCKWELL BASIS

William Garth Frazier
frazier@olemiss.edu
National Center for Physical Acoustics
University of Mississippi

Abstract

This paper reports the development of a method for beamforming on transient acoustic signals using a set of basis vectors known as the Discrete Orthonormal Stockwell Basis (DOSB). The DOSB is a set of orthonormal basis vectors in C^N that is derived from a specific sampling of the continuous S-transform kernel. The DOSB, and the S-transform as well, is like wavelet bases in that it inherently provides a means to represent a length-N time series at a range of scales and frequencies. However, it provides the additional benefit of retaining an absolute time reference with respect to the beginning of the data record which is like the discrete-time Fourier basis. This feature, along with the sparse nature of the DOSB's frequency-domain representation and the ability to apply the transform in $O(N \log_2 N)$ operations, enables beamforming in much the same way that is performed in the Fourier domain. The difference being that the beamforming occurs on each Stockwell basis function which has a specific time-frequency localization. Thus, the ability to beamform on multiple short transients or a transient I strong clutter within a single data record is possible, which makes it particularly useful for processing of battlefield events such as gunfire and explosions. Moreover, it is essentially an automatic method for estimating the direction of arrival of transients without the explicit need for a separate transient wave detector. Theory and an example applied to synthetic data are provided.

Keywords: Stockwell Transform, Array Signal Processing, Beamforming

The Stockwell Transform

The Stockwell Transform (ST), or S-Transform [Stockwell, (1996)], as it is sometimes referred to, is a time-frequency transform much like the Gabor transform in that a Gaussian window is applied to the temporal waveform prior to transformation to the frequency domain. However, an additional parameter is included to control the size of the window at each location of temporal focus. Specifically, the ST of a function $x(t)$ is

$$S_x(t,\, f) = \int_{-\infty}^{\infty} x(\tau)|f|e^{-\pi(t-\tau)^2 f^2} e^{-i2\pi f\tau}\, d\tau, \tag{1}$$

where the shape of the window focused on time t is controlled by the parameter f—frequency. Of course, for periodically sampled, finite-duration time-series a related transformation is desirable. Such a transformation was presented in Stockwell (2007) and is referred to as the Discrete Orthonormal Stockwell Transform (DOST). Details of its formulation are provided in the reference, and the definition is given by

$$S_x(j,n) = \sum_{m=0}^{N-1} H_x(m+n)e^{-2\pi^2 m^2/n^2} e^{i2\pi mj/N}, \tag{2}$$

where x is an N-point sequence, H_x is the DFT of x. (Note: $n = 0$ is a special case not stated here.) Clearly, this is just the inverse-DFT of a weighted DFT. The indices j and n play the roles of time and frequency, respectively. Like the DFT, the DOST can be calculated in $O(N \log_2 N)$ operations. The basis vectors of the DOST can be obtained by applying the transform to discrete-time Dirac deltas. A few of the Discrete Orthonormal Stockwell Basis (DOSB) vectors for a 4096-point sequence are provided in Figure 1. It is immediately apparent that the temporal localization is limited, unlike most discrete wavelets. But there is a redeeming feature that makes the DOSB desirable from a beamforming perspective. The DFT of each DOSB vector only has spectral content in a subset of the total set of bins and at each scale these bins do not overlap with the bins at other scales. For example, referring to scale 0 as DC, scale 1 has one frequency only (the lowest), scale 2 has two frequencies, scale 3 has 4, etc.—a power of two sequence. Thus, at the highest numbered scale, the non-zero spectral components are all the upper half of the total. This equivalent to the band from half-Nyquist to Nyquist.

Beamforming with the Stockwell Basis

There are many approaches to beamforming to find the directions-of-arrival (DOA) of multiple plane waves impinging on an acoustic array, and the word sometimes means different things to different people. In this manuscript the meaning is any algorithm used to search for plane wave signals in simultaneously, periodically measured finite-duration multiple time series. The type of algorithm usually depends on the types of signals that are anticipated, such as deterministic sinusoids (e.g., rotating machines), broadband random processes (e.g., jet engines) and impulsive transients (e.g., explosions). One common way for handling transients when there is persistent clutter is to use a detector to frame the transient in time, and then apply a frequency-domain beamforming such as maximum-likelihood (non-linear least-squares data fitting over slowness vector space) on the DFT'ed data from that frame to identify the directions-of-arrival of the

multiple plane waves and separate the waveforms. This can be done one frequency at a time or over a specified band.

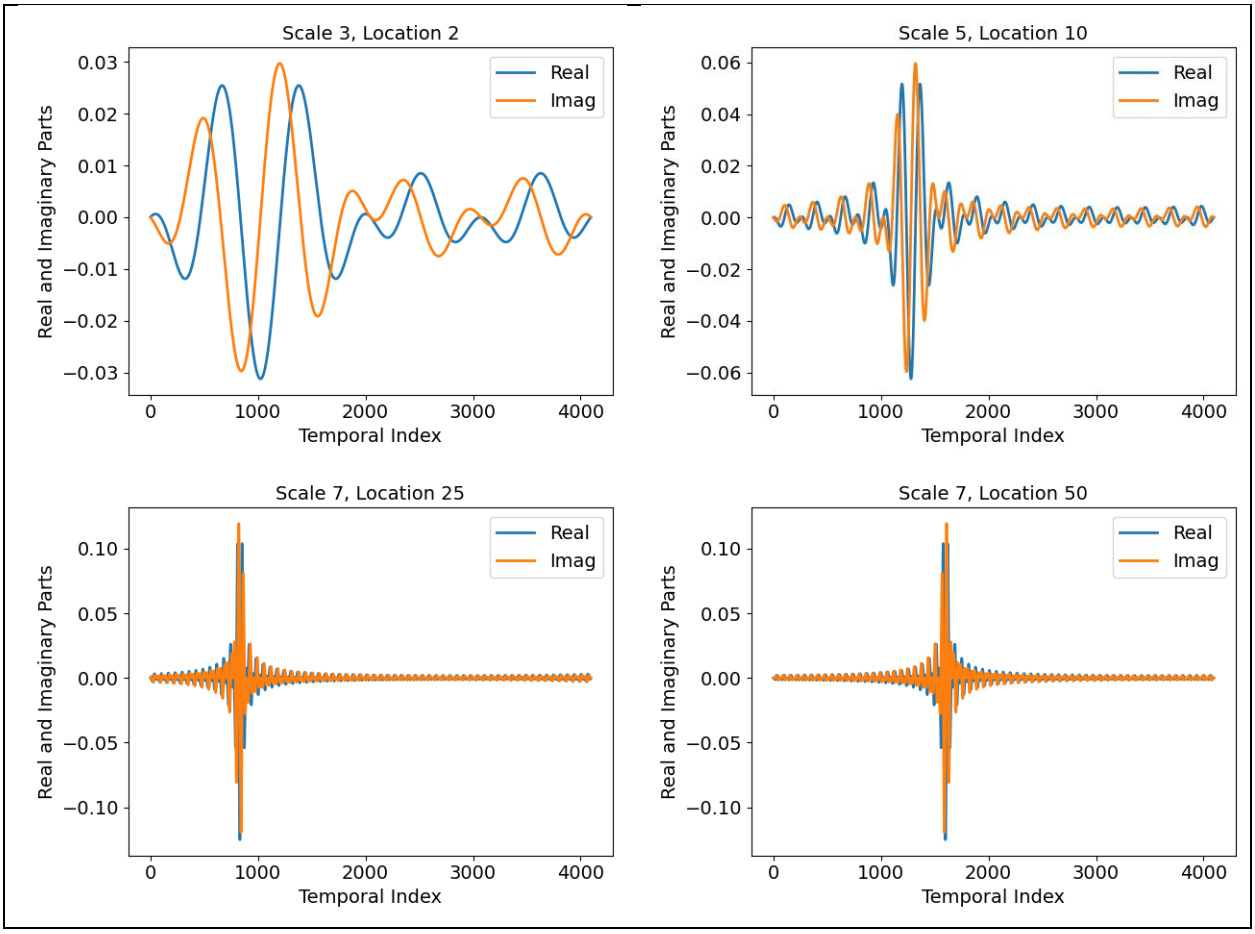


Figure 1. Examples of Stockwell Basis vectors. Note that they are not strictly localized in time as are typical discrete wavelets. However, this enables strictly localized, non-overlapping frequency bands.

Based on the properties of the DOSB, it is reasonable to consider applying it to a window of data one basis vector at a time. To formulate the approach, consider the usual model that the measurements are the result of a few signals plus noise. The pressure fluctuations at sensor m located at location r_m are

$$y_m(t) = \sum_{k=1}^{N_k} x_k(t - r_m \cdot w_k) + n_m(t), \tag{3}$$

where w_k is the slowness vector for signal k. Rewrite this in terms of a measurement vector (finite time series). This is what is available for analysis:

$$y_m = \sum_{k=1}^{N_k} d_m(w_k)[x_k] + n_m, \tag{4}$$

where $d_m(w_k)[x_k]$ denotes a time-shift operator being applied to the signal vector x_k. Now suppose it is desired to perform direction-of-arrival estimation on the data at a set of temporal locations and scales specified by the symbol $\mathcal{K}$. This will correspond to a particular set of columns of the DOSB *matrix S*. Denote the matrix formed from these columns as $S_{\mathcal{K}}$. The most obvious case is simply one vector that will correspond to just one time and scale. This is much like selecting one frequency bin for analysis when using standard frequency-domain beamforming. Now, project the measurement vectors onto this subspace

$$(S_{\mathcal{K}} S_{\mathcal{K}}^H) y_m = \sum_{k=1}^{N_k} (S_{\mathcal{K}} S_{\mathcal{K}}^H) d_m(w_k)[x_k] + (S_{\mathcal{K}} S_{\mathcal{K}}^H) n_m, \tag{5}$$

which just linear filtering. Notice that the orthonormal property of S has been exploited. This is still in the time-domain, so apply the DFT to this, which is equivalent to matrix multiplication on the left:

$$F^H (S_{\mathcal{K}} S_{\mathcal{K}}^H) y_m = \sum_{k=1}^{N_k} F^H (S_{\mathcal{K}} S_{\mathcal{K}}^H) d_m(w_k)[x_k] + F^H (S_{\mathcal{K}} S_{\mathcal{K}}^H) n_m, \tag{6}$$

where F a matrix that corresponds to the complete DFT basis, and $[\cdot]^H$ means complex-conjugate transpose. Now let us cheat a little bit by assuming that F^H and $(S_{\mathcal{K}} S_{\mathcal{K}}^H)$ commute (we do this when using standard DFT techniques applied to ordinary (e.g., Hann) windowed data. This approximation is very good for large N—the length of the sequences. This leads to

$$F^H (S_{\mathcal{K}} S_{\mathcal{K}}^H) y_m = \sum_{k=1}^{N_k} D_m(w_k) F^H (S_{\mathcal{K}} S_{\mathcal{K}}^H) x_k + F^H (S_{\mathcal{K}} S_{\mathcal{K}}^H) n_m, \tag{7}$$

where $D_m(w_k)$ is just a diagonal matrix of complex phase shifts. Write this in terms of Stockwell coefficients and (discrete) Fourier transformed quantities

$$\tilde{S}_{\mathcal{K}} y_m^s(\mathcal{K}) = \sum_{k=1}^{N_k} D_m(w_k) \tilde{S}_{\mathcal{K}} x_k(\mathcal{K}) + \tilde{S}_{\mathcal{K}} n_m^s(\mathcal{K}), \tag{8}$$

where $\widetilde{[\cdot]}$, $[\cdot]^s$, and $[\cdot](\mathcal{K})$, indicate DFT transformation, DOST transformation, and restriction to the set $\mathcal{K}$. While this equation is a vector with dimension equal to the length of the time-series being analyzed, *at least half* of the entries in the left side vector $\tilde{S}_{\mathcal{K}} y_m^s(\mathcal{K})$ are *identically equal to zero*. This dramatically reduces computations when using a non-linear least-squares method to estimate the slowness vectors of the sources. To simply matters further let $\tilde{x}_k(\mathcal{K}) = \tilde{S}_{\mathcal{K}} x_k(\mathcal{K})$ and rewrite the equation in terms of a data fitting residual assuming only non-zero terms are retained.

$$\tilde{r}_m(\mathcal{K}) = \tilde{S}_{\mathcal{K}} y_m^s(\mathcal{K}) - \sum_{k=1}^{N_k} D_m(w_k) \tilde{x}_k(\mathcal{K}). \tag{9}$$

At this point the data fitting problem *is the same* as used when applying DFT-based approaches. The only difference in the formulation is that special temporal windows (filters) have been applied

to the data prior to processing, and the spectral properties of those windows have been exploited to achieve a measure of time-frequency localization and computational advantages.

Application to Synthetic Data

Consider a 5-element square array having a center element and an aperture of 0.3 meters. Suppose the sampling frequency is 4096 Hz, and one second of data is to be analyzed. There is a bandpass (100-400 Hz) random process plane wave signal arriving from 90 degrees and a transient pulse (center frequency of 200 Hz) arriving from 180 degrees. Plots of the measurement data, the random process and transient are provided in Figure 2 through Figure 4.

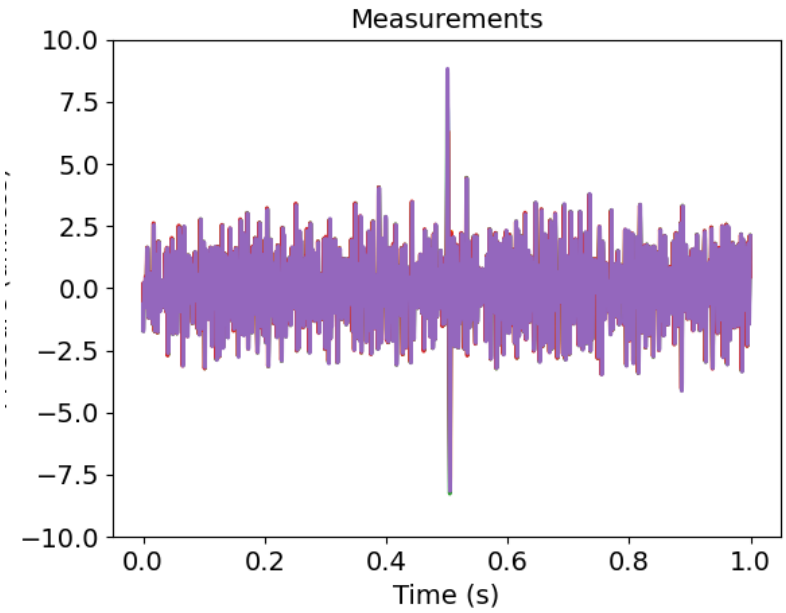


Figure 2. Measurement data from 5-sensor array. The sharp transient at 0.5 s is clear.

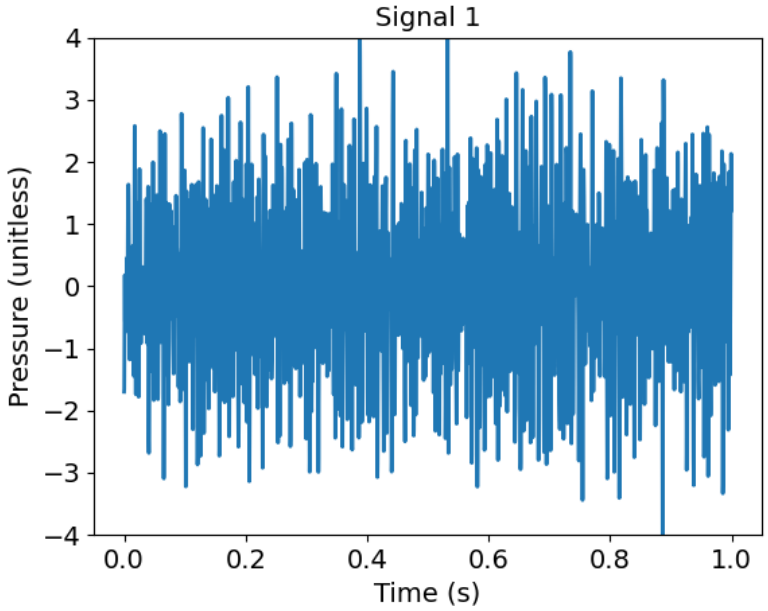


Figure 3. The bandpass random process

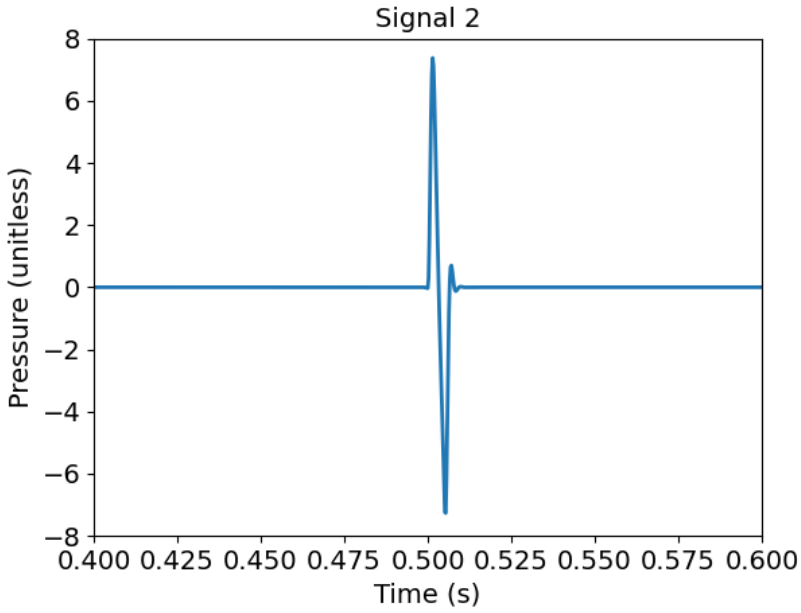


Figure 4. The transient. Temporal scale has been zoomed.

Prior to performing direction-of-arrival estimation, the DOST was applied to the data from center element of the array. DOST coefficients were calculated, and a few of these are provided in Figure 5. They appear to indicate the presence of a transient in the time-series. To limit the total computational effort, only these scales were used to look for sources. Two analyses were performed. One under the assumption that only one signal was present; the other under the

assumption two signals were present. No attempt was made to apply a signal count metric such as BIC, but this should be implemented in practice.

The results are presented in Figure 6 through Figure 9 for scales 6 through 9. At scale 6 (includes only frequency bins between 32 and 64 Hz) the results are inconclusive other than to say that the scatter is mostly limited to 100 and 150 degrees. However, at scale 7 (frequency bins between 64 and 128 Hz) a source near 90 degrees is apparent throughout the period except for anomalies near the center. The two-signal assumption appears to separate the signals somewhat during the anomaly. Increasing the resolution to scale 8 reveals that the transient is narrower in duration than revealed at scale 7, and the two-signal assumption is better. Finally, at scale 9, the narrowness of the transient is fully revealed, the two-signal assumption accurately separates the direction-of-arrival.

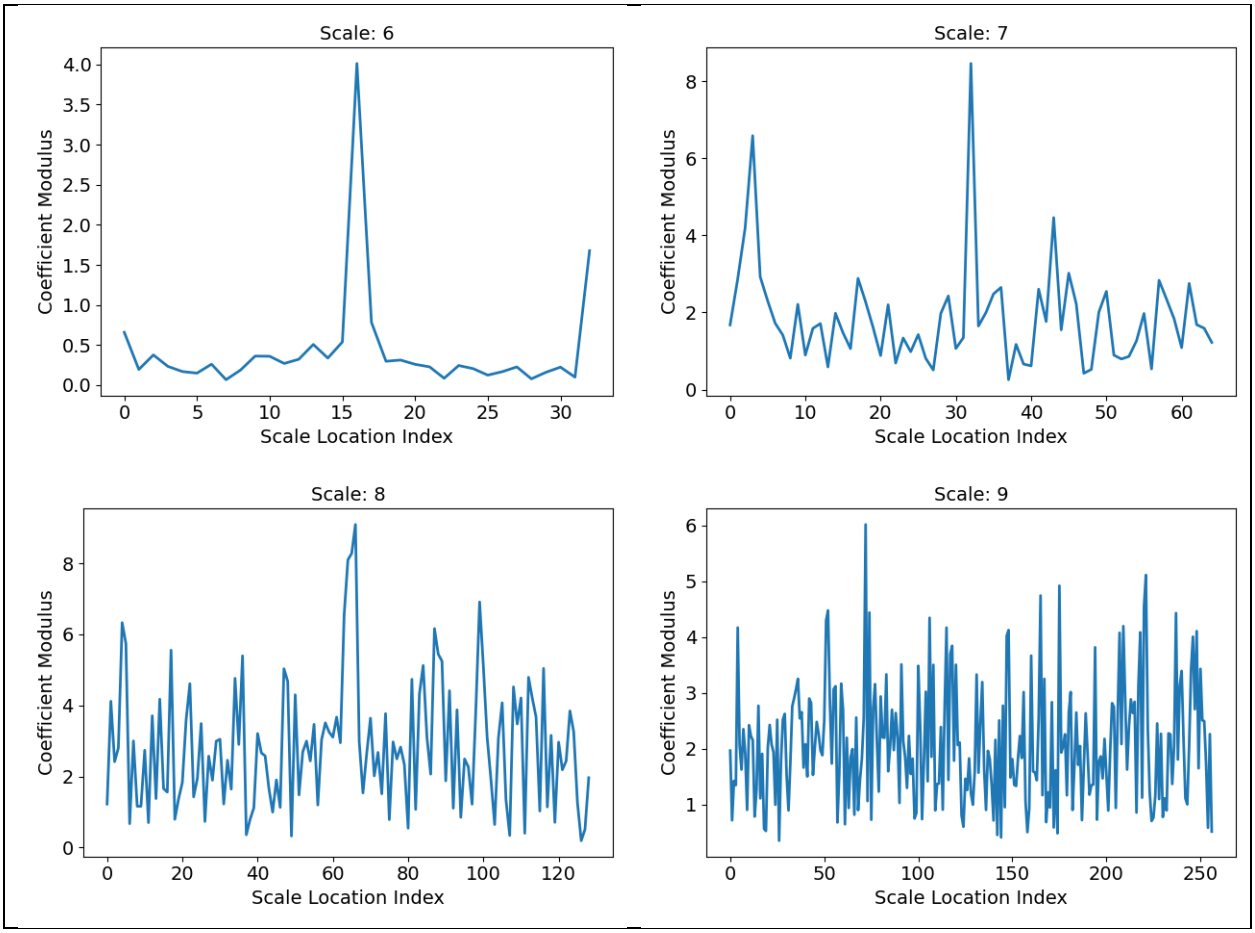
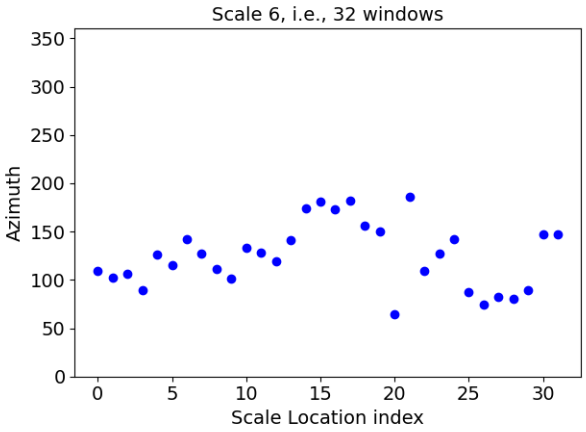


Figure 5. DOST coefficients for several scales. Scales 6-8 appear to indicate a transient near the center (temporal) at each scale. Notice that as the scale index increases, temporal resolution increases (more coefficients).

One Signal Assumption



Two Signals Assumption

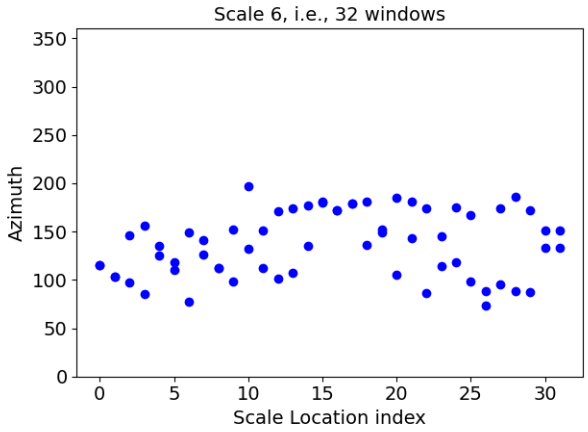
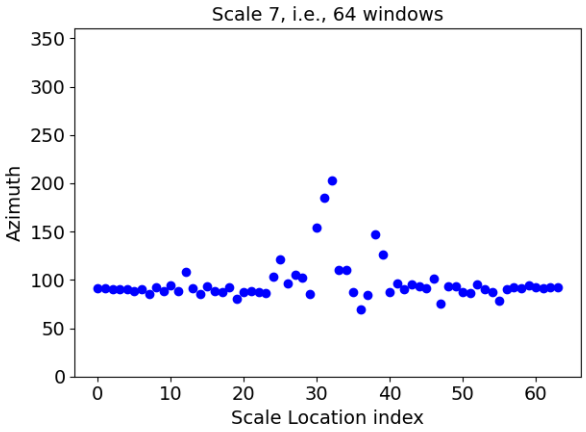


Figure 6. Source azimuth estimates at scale 6. There appears to be something between 100 and 150 degrees, but it is not distinct.

One Signal Assumption



Two Signals Assumption

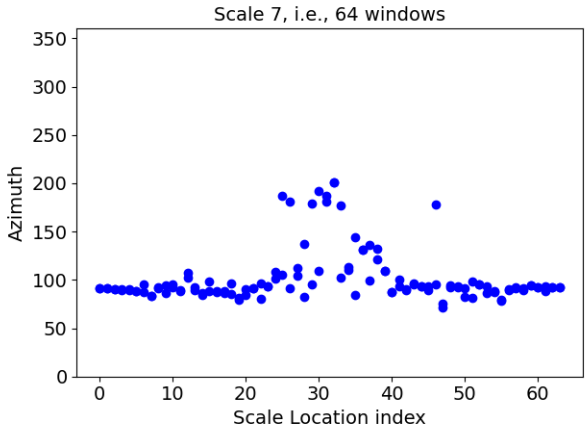


Figure 7. Source azimuth estimates at scale 7. Now there appears to be a signal at 90 degrees throughout the analysis period with an anomaly near the center.

One Signal Assumption

Two Signals Assumption

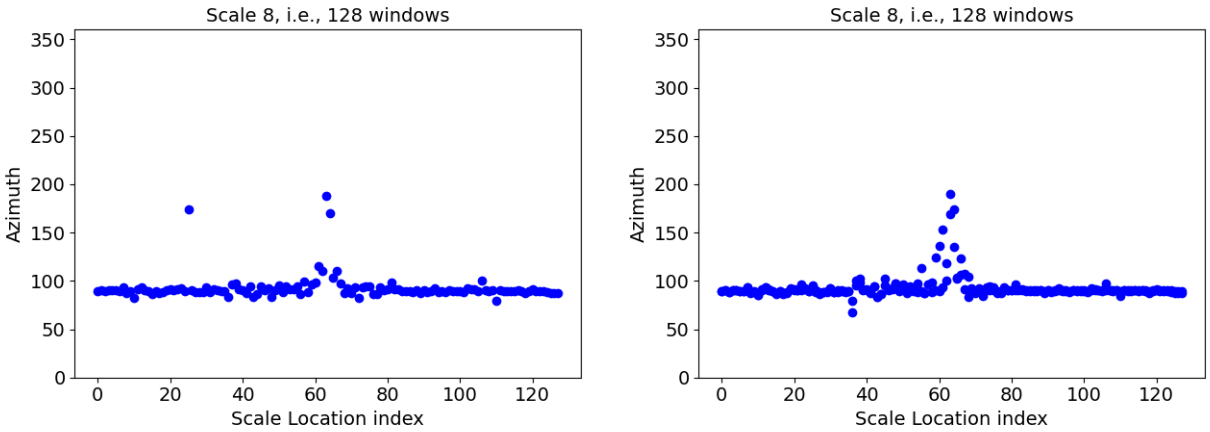


Figure 8. Source azimuth estimates at scale 8. The signal at 90 degrees is still there, but the anomaly near the center appears to be narrower than at scale 7.

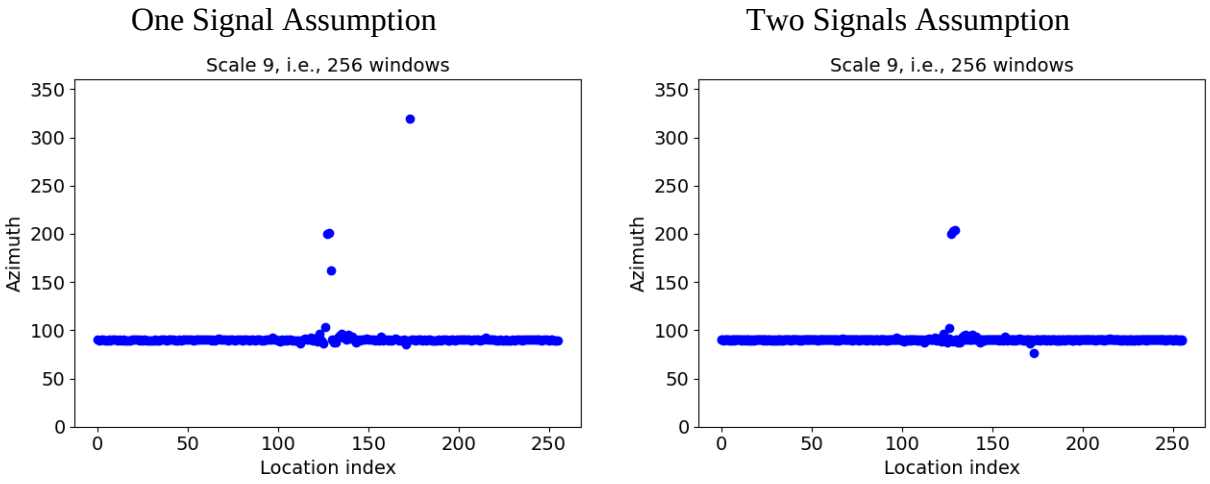


Figure 9. Source azimuth estimates at scale 9. The signal at 90 degrees is still there, and the anomaly appears to be only a few temporal frames wide. Also, under the two-signal assumption, the signals are separated better.

Conclusions

It appears that using the DOST and the DOSB vectors has merit for detecting transients in strong clutter (not surprising given its wavelet nature), but it is also effective for performing direction of

arrival estimation under the same circumstances. Based on the development presented in this manuscript, it should be clear that the primary advantage of using a DOST-based approach over simply applying standard windows to the data at different scales and time is the natural way that it breaks up the frequency spectrum into a very sparse structure. The price paid is that temporal resolution is not as sharp as can be achieved with most discrete wavelets. To produce a general software tool for analysis, automatic signal count estimation will need to be developed and inversion of DOSB coefficients to produce time-domain estimates of the multiple signals estimated will need to be implemented. Although the results presented here corresponding to analysis at each Stockwell Basis vector individually, there is no reason (other than computational burden) not to utilize more than one basis vector at a time depending on the prior analyses—adaptive processing. Another application that is of interest is when a source changes its location rapidly during a single analysis window. In principle, the proposed approach should be able follow it by estimating the slowness vector at the appropriate scale over time.

Acknowledgements

This material is based upon work partially supported by U.S. Army Corps of Engineers Engineer, Research, and Development Center under prime contract W912-HZ21-C0003 through subcontract from Hyperion Technology Group, Inc. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Army Corps of Engineers. The author would like to thank both organizations for supporting this work.

References

Stockwell, R. G. (1996); Mansinha, L.; Lowe, R. P. "Localization of the complex spectrum: the S transform". IEEE Transactions on Signal Processing. **44** (4): 998–1001.

Stockwell R. G. (2007), "A basis for efficient representation of the S-transform", *Digital Signal Processing*, vol. 17, no. 1, pp. 371-393

Wang, Y. (2009) and Orchard, J., “Fast Discrete Orthonormal Stockwell Transform,” SIAM Journal of Scientific Computing, vol. 31, no. 5, p.4000-4012