Compressive Sensing for Radar STAP

Michael L. Picciolo, J. Scott Goldstein, Wilbur Myrick
Advanced Missions Solutions Group
Dynetics, Chantilly, VA
Michael.Picciolo@Dynetics.com
Scott.Goldstein@Dynetics.com
Wil.Myrick@Dynetics.com

***Abstract*—We present a technique that merges Compressive Sensing and Wiener filtering (denoted as CS-Wiener) and compare it to standard Wiener filtering (denoted as Std-Wiener) in radar detection and estimation. The method leverages an inherent similarity in the Generalized Sidelobe Canceller transformation to that of the compressive sensing sampling function, namely its broadband or 'white' characteristic. We illustrate a data sampling reduction factor of 2/3 by using a CS-Wiener algorithm as compared to Std-Wiener approaches, while achieving approximately identical performance.**

I. INTRODUCTION

Intelligence, Surveillance, and Reconnaissance (ISR) systems such as radar, sonar, electro-optical, hyperspectral, infrared, electronic warfare (EW), etc. are becoming ever more capable, producing voluminous data and information per sensor system. The resolution and sensitivity continue to increase due to constant advances of the core component technologies. In addition, these systems are becoming smaller, cheaper and are being routinely deployed on numerous autonomous systems such as unmanned aerial vehicles (UAVs). As such, limited transmission channel bandwidth precludes efficient dissemination of the resulting massive data, and overwhelmed processing systems cannot keep pace with extracting useful information from the data.

To help alleviate this problem we examine some assumptions surrounding the question of what data is required to achieve the goals of the ISR mission at hand. We leverage the new theory of compressive sensing (CS) to design a processing approach that uses less data to achieve the same or similar results as traditional Wiener filtering or processing approaches. Our example application is for Space-Time Adaptive Processing (STAP) applied to radar target detection. It will be shown that CS, used in conjunction with STAP for radar detection (i.e. CS-Wiener filtering), performs similarly and perhaps better than traditional processing approaches (i.e., Std-Wiener), *while using significantly fewer data samples to achieve mission goals*. In addition, often signal reconstruction is not of interest- rather only the information about the signal is desired, for example in many radar applications. In these cases, as in this paper, we can dispense with the need for laborious CS signal reconstruction techniques all together.

Presented at the 2013 IEEE Radar Conf., Apr. 29-May3, Ottawa, Canada.

In Section II we cover the background of the field of compressive sensing. In Section III we discuss an approach to merging CS and STAP for radar. We present in Section IV the results of simulations, and we conclude in Section IV.

II. BACKGROUND

Circa 2006, the field of Compressive Sensing (CS) (or Compressive Sampling) was established by Donoho, Candes, et. al. [1-2]. This has led to a significant research exploration of the potential applications of this seminal theory, e.g. [3-8]. We continue that line of research here.

A. Compressive Sensing Overview

CS theory essentially states that a signal can be reconstructed from its samples even if sampled at a rate *below* the Nyquist sampling rate, where the well-known Nyquist rate is equal to twice the highest frequency present in the signal.

The caveat is that the signal must have a *sparse* characteristic within the basis of its sampling function. This can be interpreted to mean the desired signal must have a narrowband characteristic within a wideband spectral region. If this condition is satisfied, CS theory indicates it is not necessary to know in advance where this signal lies within this spectral bandwidth in order to reconstruct it perfectly, *even while sampling* it far below the Nyquist rate associated with the maximum frequency of the band. The key is to choose a sampling function that is broadband, so that no matter where the signal lies, it will overlap with the chosen wideband sampling function.

B. CS Theory – Discrete Time Formulation

Assume an $N\times 1$ noiseless input signal vector $\mathbf{x}$ is "K-sparse" or "K-compressible" in an $N\times N$ basis [3]. This means $\mathbf{x}$ has a representation of only K significant values, $K<N$, when projected into the basis. CS requires that $K<<N$ to fully exploit CS advantages, which is satisfied in many applications. It is desired to subsample these projections of $\mathbf{x}$ (where the subsampling function can be simply a sparse set of these projections, often chosen at random).

To apply CS, first define a measurement matrix $\mathbf{\Phi} = [\boldsymbol{\varphi}_1 \ldots \boldsymbol{\varphi}_M]^T$ where $\boldsymbol{\varphi}_i$ is an element of a set of $N\times 1$ basis vectors forming a projection matrix, where $M<N$. (Superscript

T denotes transpose operation.) Next, form the $M\times1$ subsampled vector $\mathbf{y} = \mathbf{\Phi x}$. At this point CS has been implemented since the data is reduced from $N\times1$ to $M\times1$, where often $M<<N$.

The key step is to choose a special matrix for $\mathbf{\Phi}$ such that it is a basis in which $\mathbf{x}$ is sparse. A "random matrix" choice is generally robust and has been described as providing universally good performance.

It is possible to reconstruct $\mathbf{x}$ after measuring just $\mathbf{y}$. (Note that reconstruction may not be needed for detection & estimation applications, but may be needed for other applications, e.g. generating imagery.) To reconstruct $\mathbf{x}$ from $\mathbf{y}$, define an $N\times N$ orthonormal basis $\mathbf{\Psi} = [\psi_1 \ldots \psi_N]^T$ where ψ_i are a set of $N\times1$ vectors forming a full-rank basis. The vector $\mathbf{x}$ maybe be formed as $x = \mathbf{\Psi s}$ where $\mathbf{s} = [s_1 \ldots s_N]^T$ is the coefficient vector, and it is clear that knowing $\mathbf{s}$ equates to knowing $\mathbf{x}$. The key step then is to choose $\mathbf{\Psi}$ such that $\mathbf{x}$ is K-sparse or K-compressible in $\mathbf{\Psi}$, meaning that $\mathbf{s}$ only has K significant values, ideally $K<<N$.

In CS, the K significant values of s_i are not measured directly; rather they are encoded as $M<N$ linear projections onto the 2nd set of vectors $\mathbf{\Phi} = [\varphi_1 \ldots \varphi_M]^T$.

That is, $\mathbf{y} = \mathbf{\Phi x} = \mathbf{\Phi\Psi s} = \mathbf{\Theta s}$, where $\mathbf{\Theta} = \mathbf{\Phi\Psi}$. It has been shown [1-3] that minimizing the L_1 norm solves for $\mathbf{s}$, hence $\mathbf{x}$. That is, $\mathbf{s} = min\ \|\mathbf{s}'\|_{l1}$ such that $\mathbf{y} = \mathbf{\Theta s}'$. This can be solved using linear programming and requires order N^3 operations. It was shown in [2] that only M measurements are required for $\mathbf{y}$ in order to reconstruct $\mathbf{x}$, where

$$M \geq cK\ ln(N/K) << N. \quad (1)$$

The quantity c is a small positive constant depending on each instance and ln denotes natural logarithm.

The theory developed is for the case of a signal and no additive noise. For the practical case of additive noise the formula for M and the number of operations for reconstruction are similar, where reconstruction fidelity is said to degrade gracefully with increasing noise power.

III. Approach

For this paper we design a approach to apply CS to detection and estimation where it is not required to reconstruct the original signal, but rather extract information from it, e.g., "target absent" or "target present". This of course saves the order N^3 operations required for reconstruction. Such subsampled data may be used directly as input for sufficient statistics, adaptive processing, parameter estimation, etc.

A. *Jammer Simulation*

We present a Space-Time Adaptive Processing (STAP) radar simulation with correlated barrage jamming affecting the ability to detect a target. The horizontally-oriented linear antenna array length is modeled to have $N = 70$ antenna elements with half-wavelength spacing. Each antenna element has a separate receiver, down-converter, and A/D sampling function for baseband digital output, including standard receiver thermal noise, into the STAP processor. Four barrage jammers, each with a 25 dB jammer-to-noise ratio (JNR), are modeled to be transmitting white noise and located at unique and arbitrarily placed azimuth angles relative to the array normal direction. The array steering vector searches for a target modeled to be located +25 degrees from the array normal in azimuth angle.

Since there exists only four jammers, the simulation is nominally K-sparse with $K = 4$ and the "sparsity" is defined here to be 4/70 or about 6%, indicating it is a good candidate for a CS approach. Sparsity is also closely tied to the concept of a low rank covariance matrix of the underlying signal structure. In Fig. 1 we show the eigenspectrum of the STAP jammer covariance matrix used in the simulation. As shown the effective rank of the eigendecomposition (i.e., about 4) is small relative to full rank (i.e., 70) graphically illustrating an example of the required sparsity of the data.

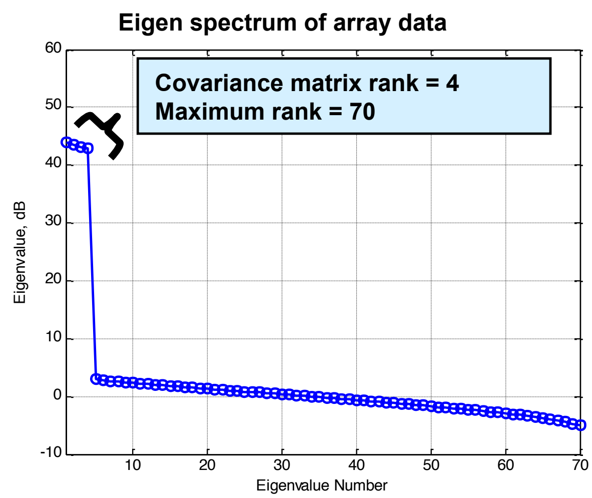

Figure 1. Eigenspectra of the jammer space-time covariance matrix used to form the data for the simulation. The small number of significant eigenvalues represents a sparse signal structure within the data, which is ideal for applying CS approaches to detection and estimation.

We note the strong jammer power must be nulled in order to detect the target. The signal to interference-plus-noise ratio (SINR) is used as the metric to determine how well the processing approach is functioning. To get a baseline performance using one of the most efficient reduced rank Wiener filters (without CS), we employ the Multistage Wiener Filter (MWF) processor [9] to efficiently learn the jammer locations from training data and to form array nulls in those locations while keeping a unity gain constraint on the target signal direction. We operate the MWF at a rank of five (5) to approximately optimize the rank choice for the most efficient use of training data. The results to be shown are for the case when the MWF weights are converged to within 3 dB of the optimum SINR, where the optimum, i.e. clairvoyant, weights are formed using the known jammer space-time covariance matrix.

To test the performance improvement of an integrated CS approach with a reduced rank Wiener filter, we substitute the Reiterative Median Cascaded Canceller with Soft Weighting (RMCC/SW) [10] as a surrogate for the reduced-rank MWF, but with an added CS process prepended, as will be described. The reason for the change of processor is due to the expanded modeling complexity required to apply the CS method to the MWF vs. the RMCC/SW. The authors believe the application of the CS method to the MWF, once accomplished, will improve the performance of the general CS-Wiener filtering approach described herein even more, as compared using the same simulation metrics shown here.

B. *A CS Method for Wiener Filtering*

Our CS method is designed to estimate the interference, not the target. To accomplish this we implement the Generalized Sidelobe Canceller (GSC) preprocessor as the first processing step. This is a benign operation because it is well known that such an invertible linear transformation of the inputs to an adaptive array suffers no performance loss to the adaptive array output in terms of SINR.

This transformation is illustrated in Fig. 2 where the $N\times1$ input vector $\mathbf{x}$ is transformed into an $N\times1$ vector $[z_0\ \mathbf{u}]^T$ by means of an $N\times1$ steering vector $\mathbf{s}$ and an $N\times(N-1)$ blocking matrix $\mathbf{B}$. The matrix $\mathbf{B}$ is formed a basis for the nullspace of $\mathbf{s}$. In the figure, $\mathbf{R}_u$ is the autocorrelation matrix of the random vector $\mathbf{u}$ and $\mathbf{r}_{uz0}$ is the cross correlation vector between the scalar z_0 and the vector $\mathbf{u}$.

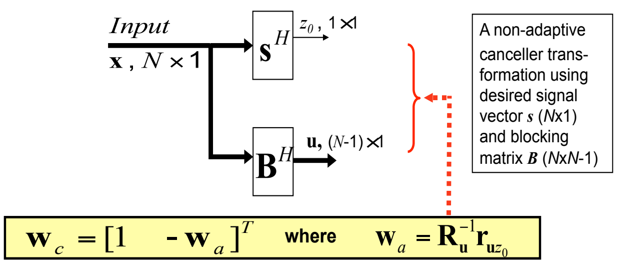

Figure 2. Generalized Sidelobe Canceller: GSC transformation and subsequent adaptive solution, $\mathbf{w}_c$. Superscript H denotes conjugate-transpose.

In Fig. 3 we show the Fourier transform of the temporal filters used for **s,** and for a random choice of two of the columns of **B**. These plots represent the filters' frequency responses. The top subfigure is the frequency response of the main channel filter **s**, and the bottom two subfigures are the frequency responses of two of the N-1 auxiliary channel filters. Though not depicted, all the GSC auxiliary channels have similar filter responses as the two shown.

Notice from Fig. 3 that the auxiliary channel filters represent the desired qualities of a CS sampling filter. They are essentially wideband, all-pass, or 'white' as is similar to a random matrix. We say 'essentially' wideband because of the narrow frequency nulls at the desired signal spectral location necessary for it to function as elements of a GSC blocking matrix. Thus, these sampling functions represent a solid approach to estimation of interference without target energy and can do so incorporating CS methodology quite naturally in this sense.

The last step needed to implement CS is to reduce the number of samples by means of decimation. Since the GSC sampling functions are wideband as required, and the jammer signal space is sparse even in the GSC transform dimension, we can simply choose a subset of the GSC channels as a decimation approach. We implement this by picking any M =

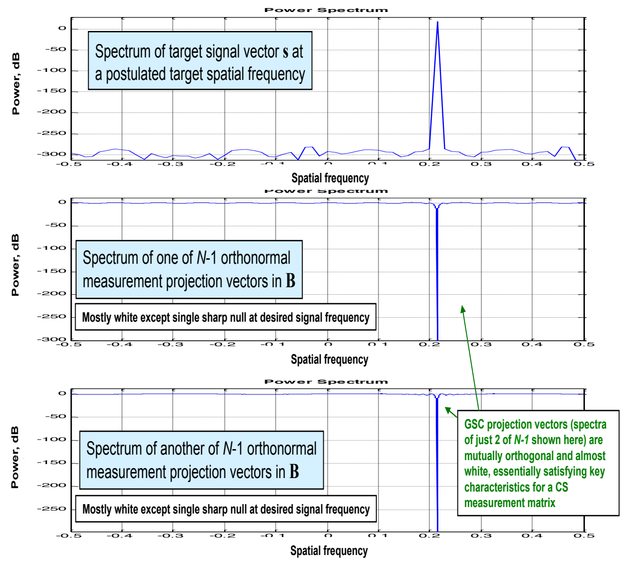

Figure 3. The top subfigure is the frequency response of the filter associated with the desired signal vector: s. The bottom two subfigures are the frequency responses associated with two of the N-1 auxiliary channel filters of the GSC. The spectra of the auxiliary filters are nearly 'white' aside from one very narrow null, representing a very good approximation to an ideal 'white' CS sampling function.

11 of the 69 auxiliary channel outputs to process by the resulting adaptive processors. The value of 11 was chosen by using the equality in (1) with c = 1, N = 69 and K = 4, matching the 6% 'sparsity' value calculated previously for our jammer STAP scenario. The chosen 11 samples happened to be the first 11 provided by the GSC calculation, but any selection should have similar performance given how robust CS is to choice of sampling functions.

IV. RESULTS

The first result has to do with a Std-Wiener approach namely a standard reduced-rank MWF STAP processor. It is operated in a nearly optimal SINR reduced rank mode with a rank of five to accommodate the four jammers, and with an extra mode to account for some uncertainty that would exist for this choice in practice. The performance is averaged over 10 Monte Carlo trials, which adequately represents the average performance for both sets of results shown.

The processor is trained using a number of snapshots, each of size $N\times1$ where N = 70, of approximately twice the interference rank. This provides an SINR that has converged to within 3dB of optimal performance. Estimating the rank to be r = 5, then $2r$ = 10 snapshots populates a STAP training matrix which totals 700 measurements of the array values in both space (70) and time (10).

The result converged adaptive array pattern, as projected back to the full dimension array size, is shown in Fig. 4 as the top curve yielding an SINR within 3dB of optimum. Underneath the adaptive pattern is the optimal antenna pattern that would produce the optimal SINR. Both patterns put nulls in the four jammer locations spread in azimuth arbitrarily. Both patterns show the target unity mainbeam constraint at +25 deg.

In Fig. 5 however, we show the adaptive and optimal STAP antenna patterns for a CS-Wiener processor. The optimal pattern is the same as in Fig. 4. However, the adaptive pattern is generated using the described CS sampling method based on the GSC transformation and integrated with an MWF surrogate (i.e. RMCC/SW). CS reduces the spatial samples from 70 to just 11, and requires only 21 temporal samples to achieve the same SINR as the MWF processor does in Fig. 4. Thus the product 11 × 21 = 231 is the total samples needed, as compared to 700 samples in Fig. 4.

The CS method thus enables a 2/3 *reduction* in the amount of space-time samples to achieve the essentially the same SINR performance. And because detection performance maps uniquely with SINR for typical Gaussian statistics, any subsequent target detection performance is also likely to be similar using far fewer data samples.

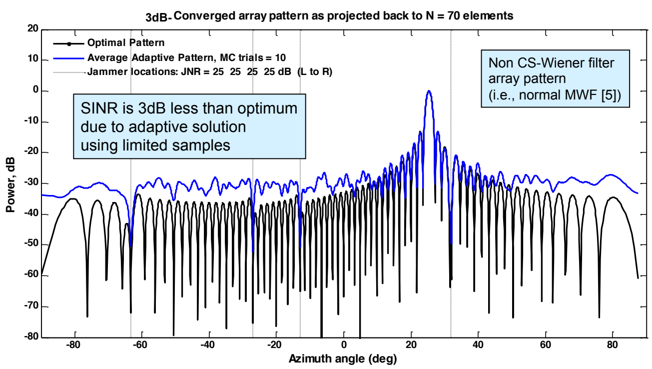

Figure 4. Adaptive (top) and optimal (lower) STAP antenna patterns. Adaptive pattern generated using MWF operated at a rank of five, and using 2r = 10 temporal samples. There are 70 spatial and 10 temporal samples, or 700 total samples.

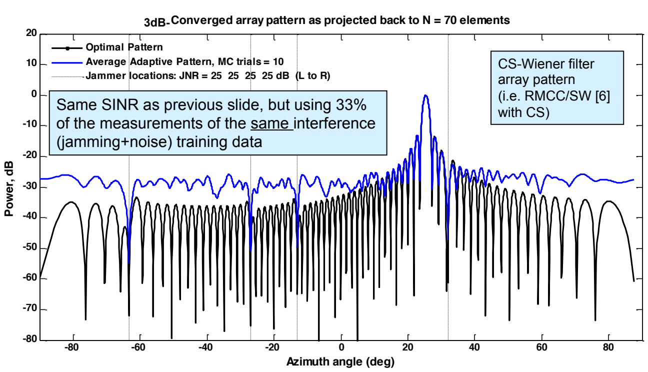

Figure 5. Adaptive (top) and optimal (lower) STAP antenna patterns. Optimal pattern is the same as in Fig. 4. Adaptive pattern is generated using CS sampling integrated with an MWF surrogate (i.e. RMCC/SW). CS reduces the spatial samples from 70 to just 11, and requires only 21 temporal samples to achieve the same SINR as in Fig. 4. Thus 11 × 21 = 231 total samples, compared to 700 in Fig. 4. The CS method thus enables a 2/3 reduction in the amount of total space-time samples to achieve the essentially the same SINR and subsequent target detection performance.

V. Conclusion

We present a method that naturally combines a compressive sensing methodology with standard Wiener filtering methods with application to detection and estimation in radar. The method leverages the GSC preprocessor form of Wiener filter methods as a means to form a good approximation to the necessary 'white' or broadband sampling function for CS. We take advantage of the fact that detection and estimation do not typically strive to reconstruct the desired signal, rather glean information about it, thus saving the computational cost, and estimation errors in additive noise, of CS signal reconstruction. A STAP radar jammer simulation is used to illustrate a 2/3 reduction in the number of total space-time samples needed to achieve very similar SINR and subsequent target detection performance. The nominal requirement for CS-Wiener filtering to perform better than Std-Wiener filtering is a relative sparseness to the underlying interference being estimated. In many applications this assumption is met, such as for the scenario described herein.

Acknowledgment

The authors thank Emmanuel Candès for past correspondence leading to insights into the practical use of compressive sensing.

References

[1] Emmanuel J. Candès, Michael B. Wakin, "An Introduction To Compressive Sampling", IEEE Signal Processing Magazine, March 2008, pp. 21-30.

[2] D. Donoho, "Compressed Sensing", IEEE Trans. Information Theory, vol. 52, no. 4, April 2006, pp. 1289-1306.

[3] Richard G. Baraniuk, "Compressive Sensing", IEEE Signal Processing Magazine, July 2007, pp. 118-124.

[4] M.A. Davenport, M.F. Duarte, D. Takhar, J.N. Laska, K.K. Kelly, and R.G. Baraniuk, "The smashed filter for compressive classification and target recognition," in Computat. Imag. V, San Jose, CA, vol. 6498, pp. 142-153, Jan. 2007.

[5] Marco F. Duarte, et. al, "Single-Pixel Imaging via Compressive Sampling", IEEE Signal Processing Magazine, March 2008, pp. 83-91.

[6] Haupt, et. al, "Compressed Sensing for Networked Data", IEEE Signal Processing Magazine, March 2008, pp. 92-101.

[7] Michael Lustig, et. al, "Compressed Sensing MRI",IEEE Signal Processing Magazine, March 2008, pp. 72-82.

[8] See http://www.dsp.ece.rice.edu/cs for an excellent and comprehensive set of compressive sensing references, seminal theory papers and links.

[9] Goldstein, J. S., Reed, I. S., Scharf, L. L., "A Multistage Representation of the Wiener Filter Based on Orthogonal Projections", IEEE Trans. on Information Theory, Vol. 44, No. 7, Nov. 1998, pp. 2943-2959.

[10] Picciolo, M.L., Gerlach, K., "Fast-Converging Adaptive Cascaded Cancellers Using A Novel Soft-Weighting and Reiteration (SWR) Technique", Proceedings of IEEE Radar Conference, Boston, MA, April 17-20, 2007, pp. 756-761.