AFRL-SN-RS-TR-2004-124
Final Technical Report
May 2004

SPACE-TIME ADAPTIVE PROCESSING (STAP) FOR LOW SAMPLE SUPPORT APPLICATIONS

Syracuse Research Corporation

AIR FORCE RESEARCH LABORATORY
SENSORS DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

STINFO FINAL REPORT

This report has been reviewed by the Air Force Research Laboratory, Information Directorate, Public Affairs Office (IFOIPA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

AFRL-SN-RS-TR-2004-124 has been reviewed and is approved for publication

APPROVED: /s/

BRAHAM HIMED
Project Engineer

FOR THE DIRECTOR: /s/

RICHARD G. SHAUGHNESSY, Lt. Col., USAF
Chief, Rome Operations Office
Sensors Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	MAY 2004	Final Sep 02 – Sep 03

4. TITLE AND SUBTITLE
SPACE-TIME ADAPTIVE PROCESSING (STAP) FOR LOW SAMPLE SUPPORT APPLICATIONS

5. FUNDING NUMBERS
C - F30602-02-C-0218
PE - 62204F
PR - 762R
TA - SN
WU - 15

6. AUTHOR(S)
Harvey K. Schuman and Ping Li

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Syracuse Research Corporation
6225 Running Ridge Road
North Syracuse New York 13212

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory/SNRT
26 Electronics Parkway
Rome New York 13441-4514

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
AFRL-SN-RS-TR-2004-124

11. SUPPLEMENTARY NOTES

AFRL Project Engineer: Braham Himed/SNRT/(315) 330-2551/ Braham.Himed@rl.af.mil

12a. DISTRIBUTION / AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***
Airborne radar Space-Time Adaptive Processing (STAP) in a heterogeneous, target-rich environment is addressed. An efficient Kalman Filter implementation of the normalized form of the Parametric Adaptive Matched Filter (NPAMF) is introduced and shown to perform well against a detailed simulation of a site-specific, dense-target environment, Ground Moving Target Indication (GMTI) scenario. The number of secondary data range cells in a Coherent Processing Interval (CPI) required by NPAMF is much smaller than the product of spatial channels and pulses and, thus, NPAMF is attractive for low sample support applications. Other promising methods for low sample support applications are introduced and studied, as well. These include a Generalized Likelihood Ratio Test (GLRT)-based PAMF (ParaGLRT) [shown to perform about as well as the matched filter when used in combination with Multiple Pass Processing (MPP)], Sub-CPI Smoothing and a GLRT variant called Severely Non homogeneous Interference Processing (SNIP). These methods are shown also to perform much better than conventional STAP methods such as Joint Domain Localized (JDL). An optimized variant of MPP called T-SNIP (the "T" is for "target-rich environment") is introduced, as well. Beam space variants of the above methods also are evaluated and found to require less processing time than element space counterparts while performing at least as well.

14. SUBJECT TERMS
Space-Time Adaptive Processing, STAP, Airborne Radar, Adaptive Processing, Ground Moving Target Indication, GMTI, Parametric Adaptive Matched Filter, PAMF, Generalized Likelihood Ratio Test, GLRT

15. NUMBER OF PAGES
62

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

ABSTRACT

Airborne radar Space-Time Adaptive Processing (STAP) in a heterogeneous, target-rich environment is addressed. An efficient Kalman Filter implementation of the normalized form of the Parametric Adaptive Matched Filter (NPAMF) is introduced and shown to perform well against a detailed simulation of a site-specific, dense-target environment, Ground Moving Target Indication (GMTI) scenario. The number of secondary data range cells in a Coherent Processing Interval (CPI) required by NPAMF is much smaller than the product of spatial channels and pulses and, thus, NPAMF is attractive for low sample support applications. Other promising methods for low sample support applications are introduced and studied, as well. These include a Generalized Likelihood Ratio Test (GLRT)-based PAMF (ParaGLRT) [shown to perform about as well as the matched filter when used in combination with Multiple Pass Processing (MPP)], Sub-CPI Smoothing and a GLRT variant called Severely Non-homogeneous Interference Processing (SNIP). These methods are shown also to perform much better than conventional STAP methods such as Joint Domain Localized (JDL). An optimized variant of MPP called T-SNIP (the "T" is for "target-rich environment") is introduced, as well. Beam space variants of the above methods also are evaluated and found to require less processing time than element space counterparts while performing at least as well.

TABLE OF CONTENTS

SECTION **PAGE**

LIST OF FIGURES

FIGURE **PAGE**

LIST OF TABLES

TABLE **PAGE**

1.0 INTRODUCTION

It is well known that nonstationary clutter exacerbates the Space-Time Adaptive Processing (STAP) secondary data selection problem. This is because it limits the number of range cells suited for approximating the space-time (angle-Doppler) matched filter that corresponds to the range cell under test. Typical contributors to nonstationary clutter are severely varying terrain, high platform altitude, bistatic geometry, and conformal (non-planar) array antenna. The target-rich environment is a condition that further stresses secondary data selection. Targets contained within secondary data range cells cause degradation in the adaptive estimation procedure. Highway traffic is the typical contributor to a target-rich environment.

STAP methods that are compatible with low sample support work well in nonstationary clutter. This is because they can be applied in a sliding window architecture whereby the relatively few required secondary data cells are always in the neighborhood of the range cell under test. Variants of the Parametric Adaptive Matched Filter (PAMF) [1]-[3] and of Sub-CPI Smoothing are particularly promising such methods. This report describes a comprehensive effort with the goal of developing and evaluating such methods, as well as introducing new methods.

The methods were evaluated by application to the Knowledge Aided Sensor Signal Processing Expert Reasoning (KASSPER) Challenge data cube. This data cube was simulated by ISL, Inc. and made available at the KASSPER workshop held on 3 April 2002 in Washington, DC [4]. The data cube results from detailed modeling of site-specific clutter and ground vehicles as viewed by an airborne multichannel monostatic Ground Moving Target Indicator (GMTI) radar. The data cube spans 1000 range cells and 32 Doppler resolution cells, and contains a large number of ground targets within the transmit pattern mainlobe: 226 “cluster” (as in convoys) and 52 background (highway vehicle). Because of the large number of resolution cells (32,000) and of targets (278), the measure of performance chosen was detection count versus false alarm count. Another feature of this data set was the availability of the true interference covariance matrix for every range cell so that the matched filter detector was available for comparison.

The detectors and methods explored here assume a sliding window architecture. With this architecture, range cells under test (“test cells”) are processed consecutively each with secondary data limited to that from a “window” of range cells centered at the test cell; thus the term “sliding window.” Usually, several guard cells immediately neighboring the test cell separate the test cell from secondary data cells. The window extent often is fixed for all test cells processed. For test cells near the limits of the data cube range swath, the window may become skewed with respect to the test cell in order to maintain a fixed number of secondary data range cells with which to process the test cell.

The PAMF developed here included a Kalman Filter solution for the requisite prediction error filter (PEF). The normalized form of the PAMF (NPAMF) [13], based on the scale invariant Normalized Matched Filter (NMF), also was further developed. The NMF is referred to as the Constant False Alarm Rate (CFAR) Matched Subspace Detector [5]. In addition, a Generalized Likelihood Ratio Test (GLRT) version of the PAMF, called ParaGLRT was developed and found to be a particularly powerful detector.

The effectiveness of Sub-CPI Smoothing derives from a characteristic that is similarly found in the PAMF: temporal smoothing. With either method, one range cell provides the equivalent of more than one data vector for detector parameter estimation. Another detector that employs the Sub-CPI Smoothing architecture is Severely Nonstationary Interference Processing

(SNIP). SNIP is also based on an extended form of Kelly's GLRT detector [6]. The extended form is Kelly's GLRT modified to allow for target presence in multiple data vectors; this modification is essential because the SNIP test range cell provides several data vectors, each containing target. This modification should not be confused with the "range spread-target GLRT" formulations of Conte, et al. [7] and Gerlach, et al [8]. The extended form of Kelly's GLRT applied in SNIP retains the premise of a lone point target. The target, however, appears in multiple data vectors associated not only with the test range cell but also with neighboring range cells that otherwise would be treated as guard cells and omitted from secondary data. SNIP avoids the need for guard cells; all range cells within a small window about the test range cell are valid sources of data vectors.

Several Multiple Pass Processing (MPP) methods also were developed for dealing with target rich environments. With MPP, the secondary data selection associated with the sliding window is modified by exchanging secondary data range cells within the window that had been tagged in preceding passes as contributing to statistical outliers, either because they contain target signal, or otherwise, with untagged range cells outside of, but closest to, the window. Thus, the MPP method tries to maintain a constant number of secondary data range cells by including range cells from outside the sliding window in order to make up for those cells within the window tagged as containing outliers. One problem with MPP is that the more distant the secondary data range cells are from the test cell, the greater is the potential for degraded performance in severely nonstationary scenarios, as with bistatic radar or severely varying terrain. An alternative, GLRT-based, method to MPP that also keeps the number of secondary ranges fixed, but does so without requiring that secondary range cells be drawn from outside the sliding window, called T-SNIP, was introduced to avoid this problem.

Beam space channels as well as element space channels were applied in the evaluations. Sidelobe target rejection methods were also explored.

A maximum likelihood angle estimation method was developed that is compatible with the STAP methods presented here, and was applied to the KASSPER Challenge data, as well.

Several other potential STAP enhancements resulted from the studies conducted here. They could not be explored in great depth in this effort, but are considered worth pursuing in future efforts. One, T-SNIP, mentioned above, promises to be particularly effective in severely non-homogeneous and target-rich scenarios. Another, High Resolution STAP (HRSTAP), promises to optimize sidelobe target suppression with minimal sacrifice in Doppler resolution. Still another, Discrete Scatterer Signal Injection (DSSI), offers an optimal way of introducing known interference discretes into the STAP processing (with potential application to knowledge assisted STAP).

2.0 LOW SAMPLE SUPPORT DETECTORS

Before introducing several adaptive detectors suited for non-homogeneous scenarios, two important nonadaptive detectors are reviewed. Let the test cell data vector, denoted by $\boldsymbol{x}$, be a column matrix of NM_p elements where N is the number of spatial channels and M_p is the number of pulses in the coherent processing interval (CPI). The matched filter (MF) [9]-[11] detector for a signal with known steering vector and unknown amplitude in additive Gaussian interference is GLRT and the detection statistic is given by

$$\Lambda_{MF} = \frac{\left|\boldsymbol{s}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{x}\right|^2}{\left(\boldsymbol{s}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{s}\right)} \tag{2-1}$$

where $\boldsymbol{s}$ is the steering vector and $\boldsymbol{R}$ is the known interference plus noise covariance matrix. The elements of the steering vector usually are chosen to "match" the signal responses for a target assumed to be in the look direction and exhibiting Doppler velocity v. Usually Λ_{MF} is recomputed for a set of Doppler velocities within a Doppler ambiguity cell and the process is repeated for each test range cell. The resulting values of Λ_{MF} are compared to a threshold to determine target presence. The elements of the vectors are ordered by channel and then by pulse.

If only the structure of the interference covariance matrix is known, that is, the matrix is known within a simple scale factor, the GLRT detection statistic, for additive Gaussian interference, becomes [12]

$$\Lambda_{NMF} = \frac{\left|\boldsymbol{s}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{x}\right|^2}{\left(\boldsymbol{s}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{s}\right)\left(\boldsymbol{x}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{x}\right)} \tag{2-2}$$

where $\boldsymbol{R}$ now refers to the interference covariance matrix structure. This statistic is referred to as the CFAR Matched Subspace Detector [5] and here as the Normalized Matched Filter (NMF). One advantage, among others, of the NMF over the MF, is that, even when the exact covariance matrix is known, the NMF detector is effective at excluding high Doppler sidelobes from detection. This can be explained by the factor in the denominator of (2-2) that is absent in (2-1). If the test cell contains a strong target, this factor increases roughly in proportion to target signal power; the statistic thus decreases in the same proportion, and the chance of detecting target sidelobes is reduced. This point is further discussed in Section 3.2 and demonstrated in Section 7.2.

2.1 PAMF

The PAMF is based on a fixed-model-order representation of the upper triangular matrix form of the interference covariance matrix inverse [1]-[3]. The filter is determined by solving for the coefficient matrices of an appropriate fixed order prediction error filter (PEF) and the corresponding error covariance matrix. Excellent performance had been demonstrated with the number of secondary data range cells much less than the number of available Degrees of Freedom (DOF) [3]. Various implementations of the PAMF (and likewise for the NPAMF) are obtained using specified coefficient estimators. The version of the PAMF developed for the current effort is based on a Kalman Filter model as described below following a summary of the underlying principles of the PAMF.

The input data process, $\boldsymbol{x}(m)$, for the PAMF is a Nx1 matrix of received pulses of interference (including noise), where m denotes the pulse number. The N elements of $\boldsymbol{x}(m)$ correspond to the N spatial channels. Let $\boldsymbol{R}$ denote the NM_p x NM_p process interference covariance matrix. $\boldsymbol{R}^{-1/2}$ is a whitening operation. Let $\boldsymbol{A}^{-1}$ denote the matrix of coefficient matrices of the finite order multichannel PEF of $\boldsymbol{x}(m)$ for the test range cell, and $\boldsymbol{D}$ denote the corresponding supermatrix of error covariance matrices. Then, because it can be shown that

$$\boldsymbol{R}^{-1} = (\boldsymbol{A}^{-1})^H \boldsymbol{D}^{-1} \boldsymbol{A}^{-1} = (\boldsymbol{D}^{-1/2}\boldsymbol{A}^{-1})^H(\boldsymbol{D}^{-1/2}\boldsymbol{A}^{-1})$$

it follows by substitution into the Matched Filter (MF) detector (2-1) that $\boldsymbol{D}^{-1/2}\boldsymbol{A}^{-1}$ also is a whitening operation. The development of the PAMF recognizes this whitening function and applies two approximations to determining the PEF coefficient matrices and error covariance matrices. First, the M_p PEFs, of orders ranging from 0 through M_p-1, are replaced with one PEF of order P. Second, the coefficient matrices and error covariance matrices are estimated from secondary range cell data. The first approximation (fixed-order filter approximation) results in a slow-time pulse sequence temporal averaging of the filter parameters. This averaging augments the fast-time averaging associated with the multiple range cell processing in the second approximation. The result is that fewer secondary data range cells are required for the PAMF to achieve the same detection performance required by conventional STAP methods.

Now, $\boldsymbol{A}^{-1}$ is lower-block triangular and can be expressed as [3]

$$\boldsymbol{A}^{-1} = \begin{bmatrix} \boldsymbol{I} & [0] & [0] & \cdots & [0] & [0] \\ \boldsymbol{A}_1^H(1) & \boldsymbol{I} & [0] & \cdots & [0] & [0] \\ \boldsymbol{A}_2^H(2) & \boldsymbol{A}_2^H(1) & \boldsymbol{I} & \cdots & [0] & [0] \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{A}_{M_p-2}^H(M_p-2) & \boldsymbol{A}_{M_p-2}^H(M_p-3) & \boldsymbol{A}_{M_p-2}^H(M_p-4) & \cdots & \boldsymbol{I} & [0] \\ \boldsymbol{A}_{M_p-1}^H(M_p-1) & \boldsymbol{A}_{M_p-1}^H(M_p-2) & \boldsymbol{A}_{M_p-1}^H(M_p-3) & \cdots & \boldsymbol{A}_{M_p-1}^H(1) & \boldsymbol{I} \end{bmatrix}$$

where $\boldsymbol{I}$ is the N x N identity matrix and the block element matrices $\boldsymbol{A}_m^H(m)$, $\boldsymbol{A}_m^H(m\text{-}1)$, ..., $\boldsymbol{A}_m^H(1)$ are the matrix coefficients of the m^{th} order multichannel PEF of $\boldsymbol{x}(m)$. For future reference, the m^{th} matrix $\boldsymbol{D}_m$ of the main block diagonal of $\boldsymbol{D}$ is the error covariance matrix of the m^{th} order PEF. Following [3], the $\boldsymbol{A}_m^H(m)$ are approximated by retaining only the vector sequence for the PEF of order P, for $1 \le P \le M_p - 1$, and the moving average filtering step is replaced with a moving window. Within this approximation, the output of the PEF is given by

$$\boldsymbol{x}'(m) = \boldsymbol{D}_P^{-1/2} \sum_{k=0}^{P} \boldsymbol{A}_P^H(k)\boldsymbol{x}(m-k+P)$$

$$\boldsymbol{s}'(m) = \boldsymbol{D}_P^{-1/2} \sum_{k=0}^{P} \boldsymbol{A}_P^H(k)\boldsymbol{s}(m-k+P)$$

where $\boldsymbol{s}(m)$ is the m^{th} N x 1 spatial subvector of the NM_p x 1 steering vector $\boldsymbol{s}$, and $\boldsymbol{s}'(m)$ and $\boldsymbol{x}'(m)$ symbolize m^{th} N x 1 block vectors, and $\boldsymbol{A}_p(0)$ is given by $\boldsymbol{I}$. Application of the

corresponding approximation for $\boldsymbol{R}^{-1}$ to the MF detector (2-1) results in the PAMF detector [1]-[3]

$$\Lambda_{PAMF} = \frac{\left| \sum_{m=0}^{M-P-1} \boldsymbol{s}'^{H}(m)\, \boldsymbol{x}'(m) \right|^2}{\left[\sum_{m=0}^{M-P-1} \boldsymbol{s}'^{H}(m)\, \boldsymbol{s}'(m) \right]}$$

Several methods had been investigated for estimating the $\boldsymbol{A}_P^H$ and, of these, the Least Mean Square (LMS) algorithm had been found to be particularly effective [3]. A Kalman Filter implementation is considered that appears to offer the fast run times associated with Kalman Filter recursive algorithms. The relevant equations are given below. The Kalman Filter implementation is compared with the LMS algorithm in Section 7.3.3.

The Kalman Filter method of estimating the PAMF coefficients, for model order *P* and number of secondary range cells *K*, is given here. Define the Kalman Filter state "vector" (which is actually a matrix in this case) as

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{A}_{\mathrm{P}}(\mathrm{P}) \\ \boldsymbol{A}_{\mathrm{P}}(\mathrm{P}-1) \\ \vdots \\ \boldsymbol{A}_{\mathrm{P}}(1) \end{bmatrix}$$

where each $\boldsymbol{A}_P$ is *N* x *N*, where *N* is the number of spatial channels, and the derivative of $\boldsymbol{F}$ is the null matrix; i.e., $\dot{\boldsymbol{F}} = 0$.

The m^{th} predicted measurement is defined as

$$\boldsymbol{h}^H(m) = \left(\boldsymbol{F}^H \boldsymbol{H}^H(m)\right)^H = \boldsymbol{H}(m)\,\boldsymbol{F}$$

where, for this special case,

$$\boldsymbol{H}(m) = \begin{bmatrix} \boldsymbol{x}_1(m) & \boldsymbol{x}_2(m) & & \boldsymbol{x}_K(m) \\ \boldsymbol{x}_1(m+1) & \boldsymbol{x}_2(m+1) & & \boldsymbol{x}_K(m+1) \\ \vdots & \vdots & \cdots & \vdots \\ \boldsymbol{x}_1(m+P-1) & \boldsymbol{x}_2(m+P-1) & & \boldsymbol{x}_K(m+P-1) \end{bmatrix}^H$$

where $x_k(i)$ is the N x 1 k^{th} secondary range cell data vector for the i^{th} pulse and $k = 1, 2, \ldots K$.

The innovation becomes

$$\boldsymbol{e}^H(m) = \left[\boldsymbol{x}_1(m+p)\ \boldsymbol{x}_2(m+p) \cdots \boldsymbol{x}_K(m+p)\right]^H - \boldsymbol{h}^H(m)$$

$$\boldsymbol{R}_e(m) = \boldsymbol{H}(m)\hat{\boldsymbol{P}}(m-1)\,\boldsymbol{H}^H(m) + \boldsymbol{R}_v$$

where $\boldsymbol{R}_v$ is the measurement error covariance matrix and $\hat{\boldsymbol{P}}$ is the estimation error covariance matrix. The Kalman gain becomes

$$\boldsymbol{G}(m) = \hat{\boldsymbol{P}}(m-1)\,\boldsymbol{H}^{H}(m)\,\boldsymbol{R}_{e}^{-1}(m)$$

and the update becomes

$$\boldsymbol{F}(m) = \boldsymbol{F}(m-1) + \boldsymbol{G}(m)\,\boldsymbol{e}^{H}(m)$$

$$\boldsymbol{P}(m) = \hat{\boldsymbol{P}}(m-1) - \boldsymbol{G}(m)\,\boldsymbol{H}(m)\hat{\boldsymbol{P}}(m-1)$$

Initial values of $\hat{\boldsymbol{P}}$ and $\boldsymbol{F}$ are required for this approach. Performance was found to be insensitive to initial values for the simulations exercised to date.

2.2 NPAMF

Application of the PAMF approximation for $\boldsymbol{R}^{-1}$ to the NMF detector statistic (2-2) results in the NPAMF detector statistic [13]

$$\Lambda_{NPAMF} = \frac{\left|\sum_{m=0}^{M-P-1} \boldsymbol{s}'^{H}(m)\,\boldsymbol{x}'(m)\right|^{2}}{\left[\sum_{m=0}^{M-P-1} \boldsymbol{s}'^{H}(m)\,\boldsymbol{s}'(m)\right]\left[\sum_{m=0}^{M-P-1} \boldsymbol{x}'^{H}(m)\,\boldsymbol{x}'(m)\right]}$$

This detector exhibits the sidelobe target rejection benefits discussed in Section 2.0, among other benefits associated with the NMF.

2.3 Sub-CPI Smoothing

The basic architecture of Sub-CPI Smoothing (**Figure 2.3-1**) is that of Pre-Doppler STAP. The major difference between Pre-Doppler STAP and Sub-CPI Smoothing is that the Sub-CPI Smoothing steering vector is Doppler dependent and thus Sub-CPI Smoothing is neither a pre-Doppler method nor a post-Doppler method, but a hybrid.

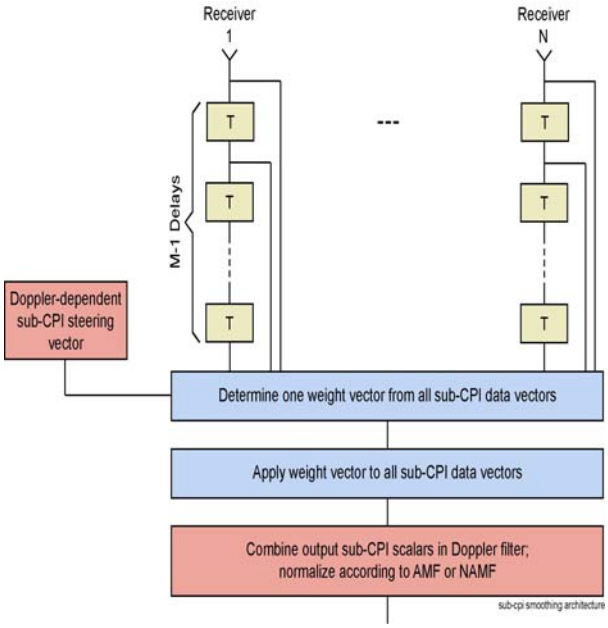

Figure 2.3-1. Sub-CPI Smoothing Architecture

In Pre-Doppler STAP, the first M pulses within the M_p-pulse coherent processing interval (CPI) are grouped into a subset of pulses called a sub-CPI. Pulses 2 through $M+1$ form a second sub-CPI that partially overlaps the first, and so on, resulting in M_p-$M+1$ partially overlapping sub-CPIs. Usually, a separate STAP procedure is conducted for each sub-CPI, or the procedure is applied to only one sub-CPI and the resulting STAP weights applied to all other sub-CPIs. Alternatively, all sub-CPIs are used as data vectors in obtaining one set of weights, which is applied, in turn, to all sub-CPIs. In either case, with Pre-Doppler STAP, a binomial weighted steering vector is employed that instills a null at the Doppler introduced by the ground in the peak of the spatial mainlobe. In a variant of Pre-Doppler STAP, called Optimized Pre-Doppler STAP (OPDS) [14], Baranoski introduces a Doppler independent steering vector that results from minimizing the clutter throughout the entire radiation sphere. In contrast with Pre-Doppler STAP, Sub-CPI Smoothing "points" the steering vector toward the Doppler of interest.

Pre-Doppler STAP usually performs well only for a small number (2 or 3) of temporal DOFs (equal to M). Otherwise, the Doppler null created by the binomial weighted steering vector is overextended and detection of subclutter targets is denied. ODPS is thought to suffer from the same condition. Baranoski found ODPS superiority to Pre-Doppler STAP limited to conditions of high backlobe clutter and antenna crab [14].) With Sub-CPI Smoothing, on the other hand, STAP performance can be improved with increased M because the steering vector is Doppler dependent and the STAP solution is recomputed for each Doppler of interest. Furthermore, because all M_p-$M+1$ partially overlapping sub-CPIs are employed as data vectors in a STAP solution in arriving at one set of weights, Sub-CPI Smoothing trades DOFs for covariance matrix temporal smoothing (slow-time averaging), which supplements the fast-time averaging afforded by averaging over range cells. Thus, Sub-CPI Smoothing requires fewer secondary data range cells than do standard STAP methods with the same number of degrees of freedom. Treating each overlapping sub-CPI as a STAP data vector effectively increases the

number of uncorrelated data vectors with identical interference spectra, by at most M_p-M+1, for each secondary data range cell. Because a data vector is composed of the complex amplitudes of the pulses in all spatial channels corresponding to one sub-CPI and one range cell, the number of secondary data vectors available for STAP parameter estimation is $L*(M_p-M+1)$ where L is the number of secondary data range cells.

Overlapping subarrays can, in principle, be formed in the spatial domain as well [15]. The subarrays must be identical, however, and this is difficult to realize because antenna array elements usually are not precisely matched due to mutual coupling and near field platform scattering. On the other hand, identical subarrays are easily formed in the temporal domain because of the excellent pulse stability achievable with modern radars.

Consider, as an example, the KASSPER Challenge data set. Here, the CPI is $M_p = 32$ pulses and the number of channels is $N = 11$. The sub-CPI size of $M = 6$ works well (Section 7); thus, M_p-M+1 = 27. The number of DOFs is $MN = 66$ and the number of secondary data range cells required to meet a "three times DOFs" criteria is $3*66/27 \approx 8$. For "six times DOFs," the number is approximately 16. These numbers of secondary data range cells for Sub-CPI Smoothing are more than twelve times smaller than the 100 to 200 required by element-space post-Doppler STAP methods such as Three-Pulse PRI-Staggered, even though the latter employs fewer degrees of freedom.

The effective number of uncorrelated data vectors per range cell, however, may be less than M_p-M+1 if the PRI is too small with respect to the Doppler spread. The degree of decorrelation is a function of several parameters, the most important being Doppler spread. It is shown in Section 2.3.1 that adjacent sub-CPIs (separated by one PRI) are nearly decorrelated if the PRI exceeds the wavelength divided by four times the platform velocity. For example, for a 1240 MHz radar aboard an aircraft cruising at 100 m/s (KASSPER example) the PRF (1/PRI) would be restricted to less than 1,650 Hz or so.

Another criterion for effective Sub-CPI Smoothing is that the interference processes in the sub-CPIs have identical spectra or at least highly correlated spectra. A necessary condition for this requirement is the excellent pulse stability achievable with modern radars. It is important to note, however, that the correlation between sub-CPI interference spectra deteriorates with severe high frequency platform vibration combined with long CPIs.

Sub-CPI Smoothing is more computationally demanding than pre-Doppler STAP methods because Sub-CPI Smoothing recomputes the STAP solution at each Doppler frequency.

Details underlying Sub-CPI Smoothing follow. Assume an array of N channels (elements or beams). As defined above, M_p is the number of pulses within the CPI, M is the number of digital time delay taps, and L is the number of range cells contributing secondary data vectors. The number of degrees of freedom is $N_{DOF} = MN$. Data vectors are formed by collecting partially overlapping subsets ("sub-CPIs") of M contiguous pulses for each channel and ordering first by channel. The indices of the pulses of one sub-CPI are offset by one from those corresponding to a neighboring sub-CPI. The number of data vectors per range cell is given by $K = M_p - M + 1$. The total number of data vectors is the product LK.

Let $\boldsymbol{x}_{lk}$ denote the k^{th} sub-CPI data vector of the l^{th} secondary data range cell. Although the data vectors are dependent, they are reasonably decorrelated under certain conditions.

Section 2.3.1 addresses this issue. The associated sample interference covariance matrix and steering vector are given by

$$\hat{\boldsymbol{R}} = \frac{1}{\boldsymbol{LK}} \sum_{\boldsymbol{l}=0}^{\boldsymbol{L}-1} \sum_{\boldsymbol{k}=0}^{\boldsymbol{K}-1} \boldsymbol{x}_{\boldsymbol{lk}} \boldsymbol{x}_{\boldsymbol{lk}}^{\boldsymbol{H}}$$

$$\boldsymbol{s} = \left[\beta_0 \boldsymbol{\gamma}^{\boldsymbol{tr}} \;\; \beta_1 \boldsymbol{\gamma}^{\boldsymbol{tr}} \cdots \beta_{\boldsymbol{M}-1} \boldsymbol{\gamma}^{\boldsymbol{tr}} \right]^{\boldsymbol{tr}}$$

where

$$\boldsymbol{\gamma} = \left[\gamma_0 \;\; \gamma_1 \cdots \gamma_{N-1} \right]^{tr}$$

$$\boldsymbol{\gamma}_{\mathrm{n}} = \exp\left(j \frac{2\pi}{\lambda} d_n \sin\theta \right) \qquad 0 \le m \le N-1$$

$$\beta_m = \exp\left(j \frac{4\pi}{\lambda} mTv \right) \qquad 0 \le m \le M-1$$

v is the Doppler velocity (negative of range rate)

T is the PRI

tr denotes the "transpose" operation

λ is the wavelength

θ is the complement of the cone angle

d_n is the location of n^{th} element along array axis $(n = 0, 1, \cdots, N\text{-}1)$

For linear uniformly spaced arrays of spacing d, d_n is given by $d_n = nd$.

The adaptive weight vector is given by

$$\boldsymbol{w} = c\hat{\boldsymbol{R}}^{-1}\boldsymbol{s}$$

where c is a constant. Let $\boldsymbol{z}_k$ denote the k^{th} sub-CPI data vector of the test cell (an NMx1 matrix). The ouput signal is given by

$$y = \sum_{k=0}^{K-1} \beta_k^* \boldsymbol{w}^H \boldsymbol{z}_k$$

or

$$y = \boldsymbol{w}^H \boldsymbol{g}$$

where * denotes the complex conjugate and

$$\boldsymbol{g} = \sum_{k=0}^{K-1} \beta_k^* \boldsymbol{z}_k$$

The normalization constant c is determined to satisfy a CFAR condition. Under the interference-only hypothesis, the test cell vectors $\boldsymbol{z}_k$ contain interference only and c is chosen to satisfy

$$E\left\{|y|^2\right\} = c^2\, \boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{B}\, \hat{\boldsymbol{R}}^{-1} \boldsymbol{s} = 1$$

where

$$\boldsymbol{B} = E\left\{\boldsymbol{g}\, \boldsymbol{g}^H\right\} = \sum_{k=0}^{K-1} \sum_{k'=0}^{K-1} \beta_k^* \beta_{k'} E\left\{\boldsymbol{z}_k\, \boldsymbol{z}_{k'}^H\right\} \tag{2-3}$$

Assuming temporal wide-sense stationarity over range cells, the double summation in (2-3) contains the covariance matrices

$$\boldsymbol{R}_{k'-k} = E\left\{\boldsymbol{z}_k \boldsymbol{z}_{k'}^H\right\} \tag{2-4}$$

Consider the approximation

$$\boldsymbol{R}_{k'-k} = \begin{cases} \hat{\boldsymbol{R}} & k = k' \\ \boldsymbol{0} & k \neq k' \end{cases} \tag{2-5}$$

It follows that $\boldsymbol{B} = K\hat{\boldsymbol{R}}$,

$$c = \frac{1}{\sqrt{K \boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{s}}}$$

and the Sub-CPI Smoothing detector is approximated by

$$|y|^2 = \frac{1}{K \boldsymbol{s}^H \hat{R}^{-1} \boldsymbol{s}} \left| \boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \sum_{k=0}^{K-1} \beta_k^* \boldsymbol{z}_k \right|^2 \tag{2-6}$$

Following (2-2), the "Normalized" Sub-CPI Smoothing detector is approximated by

$$|y|^2 = \frac{1}{K \boldsymbol{s}^H \hat{R}^{-1} \boldsymbol{s} \left(\sum_{k=0}^{K-1} \beta_k^* \boldsymbol{z}_k^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{z}_k \right)} \left| \boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1} \sum_{k=0}^{K-1} \beta_k^* \boldsymbol{z}_k \right|^2 \tag{2-7}$$

Equations (2-6) and (2-7) were applied in the simulations of Section 7.

The following derivation of $\boldsymbol{B}$ is not restricted by (2-5) but results in a more computationally demanding expression.

The covariance matrices given by (2-4) can be estimated in terms of the sub-matrices of $\hat{\boldsymbol{R}}$. First note that for $k' = k$

$$\boldsymbol{R}_0 = \boldsymbol{R} = E\left\{\boldsymbol{z}_k\, \boldsymbol{z}_k^{\boldsymbol{H}}\right\}$$

and that

$$\boldsymbol{R_0} = \begin{bmatrix} \boldsymbol{Q}_0 & \boldsymbol{Q}_1 & \cdots & \boldsymbol{Q}_{M-2} & \boldsymbol{Q}_{M-1} \\ \boldsymbol{Q}_1^H & \boldsymbol{Q}_0 & \cdots & \boldsymbol{Q}_{M-3} & \boldsymbol{Q}_{M-2} \\ \boldsymbol{Q}_2^H & \boldsymbol{Q}_1^H & \cdots & \boldsymbol{Q}_{M-4} & \boldsymbol{Q}_{M-3} \\ \vdots & \vdots & & \vdots & \vdots \\ \boldsymbol{Q}_{M-1}^H & \boldsymbol{Q}_{M-2}^H & \cdots & \boldsymbol{Q}_1^H & \boldsymbol{Q}_0 \end{bmatrix} \tag{2-8}$$

with sub-matrices given by

$$\boldsymbol{Q}_m = \boldsymbol{E}\left\{ \boldsymbol{q}_0 \, \boldsymbol{q}_m^H \right\} \qquad m = 0,1,\ldots,M-1$$

where the N element $\boldsymbol{q}_m$ vectors are sub-vectors of the general sub-CPI data vector given by

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{q}_0^{tr} \, \boldsymbol{q}_1^{tr} & \cdots & \boldsymbol{q}_{M-1}^{tr} \end{bmatrix}^{tr}$$

The covariance sub-matrices of (2-8) are estimated from

$$\boldsymbol{R}_0 = \boldsymbol{R} \approx \hat{\boldsymbol{R}}$$

By assuming that data sub-vectors separated by M or more pulses are uncorrelated, the remaining cross-correlation covariance matrices can be estimated from the sub-matrices of $\hat{\boldsymbol{R}}$. These matrices are block Toeplitz. Define $\boldsymbol{Q}_i = \boldsymbol{0}$ (null matrix) for $i > M$-1. Then, for k'-k = 1, $\boldsymbol{R}_1$ is given by

$$\boldsymbol{R}_1 = \begin{bmatrix} \boldsymbol{Q}_1 & \boldsymbol{Q}_2 & \cdots & \boldsymbol{Q}_{M-1} & \boldsymbol{0} \\ \boldsymbol{Q}_0 & \boldsymbol{Q}_1 & \cdots & \boldsymbol{Q}_{M-2} & \boldsymbol{Q}_{M-1} \\ \boldsymbol{Q}_1^H & \boldsymbol{Q}_0 & \cdots & \boldsymbol{Q}_{M-3} & \boldsymbol{Q}_{M-2} \\ \vdots & \vdots & & \vdots & \vdots \\ \boldsymbol{Q}_{M-2}^H & \boldsymbol{Q}_{M-3}^H & \cdots & \boldsymbol{Q}_0 & \boldsymbol{Q}_1 \end{bmatrix}$$

While for k'-k = 2, $\boldsymbol{R_2}$ is given by

$$\boldsymbol{R}_2 = \begin{bmatrix} \boldsymbol{Q}_2 & \boldsymbol{Q}_3 & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{Q}_1 & \boldsymbol{Q}_2 & \cdots & \boldsymbol{Q}_{M-1} & \boldsymbol{0} \\ \boldsymbol{Q}_0 & \boldsymbol{Q}_1 & \cdots & \boldsymbol{Q}_{M-2} & \boldsymbol{Q}_{M-1} \\ \vdots & \vdots & & \vdots & \vdots \\ \boldsymbol{Q}_{M-3}^H & \boldsymbol{Q}_{M-4}^H & \cdots & \boldsymbol{Q}_1 & \boldsymbol{Q}_2 \end{bmatrix}$$

In general,

$$\boldsymbol{R}_i = \begin{bmatrix} \boldsymbol{Q}_i & \boldsymbol{Q}_{i+1} & \cdots & \boldsymbol{Q}_{M-1} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{Q}_{i-1} & \boldsymbol{Q}_i & \cdots & \boldsymbol{Q}_{M-2} & \boldsymbol{Q}_{M-1} & \cdots & \boldsymbol{0} \\ \boldsymbol{Q}_{i-2} & \boldsymbol{Q}_{i-1} & \cdots & \boldsymbol{Q}_{M-3} & \boldsymbol{Q}_{M-2} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \boldsymbol{Q}_0 & \boldsymbol{Q}_1 & \cdots & & & & \\ \boldsymbol{Q}_1^H & \boldsymbol{Q}_0 & \cdots & & & & \\ \boldsymbol{Q}_2^H & \boldsymbol{Q}_1^H & \cdots & & & & \\ \vdots & \vdots & & & & & \vdots \\ \boldsymbol{Q}_{M-1-i}^H & \boldsymbol{Q}_{M-i}^H & \cdots & & & \cdots & \boldsymbol{Q}_i \end{bmatrix}; \quad i = 0,1,...,M-2$$

and

$$\boldsymbol{R}_i = \begin{bmatrix} \boldsymbol{Q}_i & \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{Q}_{i-1} & \boldsymbol{Q}_i & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{Q}_{i-2} & \boldsymbol{Q}_{i-1} & \boldsymbol{Q}_i & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \vdots & & \vdots \\ \boldsymbol{Q}_{i-M+1} & \boldsymbol{Q}_{i-M+2} & \boldsymbol{Q}_{i-M+3} & \cdots & \boldsymbol{Q}_i \end{bmatrix}; \quad i = M-1, M, M+1, ..., 2M-2$$

Note that

$$\boldsymbol{R}_{-i} = E\left\{ \boldsymbol{z}_{k+i}\ \boldsymbol{z}_k^H \right\} = \boldsymbol{R}_i^H$$

The matrix $\boldsymbol{B}$ of (2-3) is given by

$$\boldsymbol{B} = \sum_{k=0}^{K-1} \sum_{k'=0}^{K-1} \beta_k^*\ \beta_{k'}\ E\left\{ \boldsymbol{z}_k\ \boldsymbol{z}_{k'}^H \right\} = \sum_{k=0}^{K-1} \sum_{k'=0}^{K-1} \beta_{k'-k} \boldsymbol{R}_{k'-k}$$

$$\boldsymbol{B} = \underbrace{\sum_{k'=0}^{K-1} \sum_{k=0}^{k'} \beta_{k'-k}\ \boldsymbol{R}_{k'-k}}_{k \le k'} + \underbrace{\sum_{k=1}^{K-1} \sum_{k'=0}^{k-1} \beta_{k'-k} \boldsymbol{R}_{k'-k}}_{k > k'}$$

$$= \sum_{k'=0}^{K-1} \sum_{k=0}^{k'} \beta_{k'-k}\ \boldsymbol{R}_{k'-k} + \sum_{k=1}^{K-1} \sum_{k'=0}^{k-1} \beta_{k-k'}^*\ \boldsymbol{R}_{k-k'}^H$$

$$= K\boldsymbol{R}_0 + \sum_{k'=1}^{K-1} \sum_{k=0}^{k'-1} \left(\beta_{k'-k}\ \boldsymbol{R}_{k'-k} + \beta_{k'-k}^*\ \boldsymbol{R}_{k'-k}^H \right)$$

The latter equation results from the index interchange in the second double summation of the previous equation. Because $k' > k$, let $l = k' - k$, and $\boldsymbol{B}$ becomes

$$\boldsymbol{B} = K\boldsymbol{R}_0 \quad + \quad \sum_{k'=1}^{K-1} \sum_{l=1}^{k'} \left(\beta_l \ \boldsymbol{R}_l \ + \beta_l^* \ \boldsymbol{R}_l^H \right)$$

$$\boldsymbol{B} = K\boldsymbol{R}_0 \quad + \quad \sum_{l=1}^{K-1} (K-l) \left(\beta_l \ \boldsymbol{R}_l \ + \beta_l^* \ \boldsymbol{R}_l^H \right)$$

Because $\boldsymbol{R}_i = 0$ for $i > 2M\text{-}2$,

$$\boldsymbol{B} = K\boldsymbol{R}_0 \quad + \quad \sum_{l=1}^{2M-2} (K-l) \ (e^{j\frac{4\pi}{\lambda} lTv} R_l \ + e^{-j\frac{4\pi}{\lambda} lTv} R_l^H)$$

This expression for $\boldsymbol{B}$ was not exercised in this work. Instead, the simpler, less accurate expression based on (2-4) was employed and found to work reasonably well.

2.3.1 Sub-CPI Temporal Decorrelation

For Sub-CPI Smoothing, the data vectors, one for each sub-CPI, ideally should be independent samples of identically distributed interference. An incremental source of interference, at a given angle, imparts signals in two sub-CPIs that differ only by a phase shift that corresponds to the Doppler of the source and the time delay between the sub-CPIs. Thus, the distribution functions of the interference are identical between sub-CPIs, and Sub-CPI Smoothing is expected to work reasonably well if the delay between sub-CPIs is that for which the total interference (from all angles) is decorrelated.

Consider a CPI of M_p pulses transmitted with uniform PRI. Construct partially overlapping sub-CPIs of M pulses each. Consider also, element spatial channels so that the interference associated with a sub-CPI is widely distributed in angle. The problem is to determine the minimum (temporal) separation between adjacent sub-CPIs such that the interference is decorrelated.

Consider a clutter patch at angle θ from broadside (assume a linear array with axis aligned with the platform velocity vector). Consider, also, one element channel. The contribution to the m^{th} pulse is given by

$$s_m(\theta) = \begin{cases} a(\theta) \cos^{1/2} \theta \ e^{j2kmv_0 T \sin\theta} & -\pi/2 \le \theta \le \pi/2 \\ 0 & \text{otherwise} \end{cases}$$

where T = PRI, v_0 = platform velocity, $k = 2\pi/\lambda$, and $a(\theta)$ is the interference distribution. A $\cos\theta$ element gain pattern, as well as a monostatic mode of operation are assumed. No allowance is made for clutter dependence on the transmit pattern. Although the transmit pattern often has a narrow beam, its sidelobes usually are high. Furthermore, if bistatic radar was of interest, the transmit mainlobe illumination of a range cell could be especially broad.

Let $\boldsymbol{x_0}(\theta)\, d\theta$ denote the contribution at θ of clutter within an angular increment $d\theta$ to the data vector for one sub-CPI, and $\boldsymbol{x_i}(\theta)\, d\theta$ that for a sub-CPI delayed by i PRIs. Then

$$\boldsymbol{x}_i(\theta) = \exp(j2kiv_0 T \sin\theta)\ \boldsymbol{x}_0(\theta)$$

where

$$\boldsymbol{x}_0 = \begin{bmatrix} s_1 & s_2 & \cdots & s_M \end{bmatrix}^H$$

Let *R(i)* denote the correlation between sub-CPIs separated by *i* pulses. Then

$$\begin{aligned} R(i) &= E\left[\int_{-\pi}^{\pi} \boldsymbol{x}_i^H(\theta)\, d\theta \int_{-\pi}^{\pi} \boldsymbol{x}_0(\theta')\, d\theta' \right] \\ &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} E\left[\boldsymbol{x}_i^H(\theta)\, \boldsymbol{x}_0(\theta') \right] d\theta' d\theta \end{aligned}$$

Under the assumption

$$E\left\{ a(\theta)\, a(\theta') \right\} = \sigma^2 \delta(\theta - \theta')$$

where $\delta(\theta)$ is the dirac delta function,

$$\int_{-\pi}^{\pi} E\left\{ \boldsymbol{x}_i^H(\theta)\, \boldsymbol{x}_0(\theta') \right\} d\theta' = \begin{cases} C\sigma^2 e^{j2kiv_0 T \sin\theta} \cos\theta; & -\pi/2 \le \theta \le \pi/2 \\ 0; & \text{otherwise} \end{cases}$$

where *C* is a constant. Consequently,

$$R(i) = 2C\sigma^2 \frac{\sin(2kiv_0 T)}{2kiv_0 T}$$

The sub-CPI vectors are decorrelated when $R(i) = 0$ or $T = \lambda/(4v_0 i)$. Two vectors separated by one PRI are reasonably decorrelated for

$$T > \lambda/(4v_0) = 1/(2f_d)$$

where $f_d = 2v_0/\lambda$. Thus $f_d T > 1/2$. Sufficient decorrelation may require a larger separation between sub-CPI vectors if the transmit pattern, including high sidelobes, only illuminates a narrow angular extent of the range cell. Beam space channels also lead to larger separations for decorrelation, but the number of degrees of freedom is then correspondingly smaller.

2.4 SNIP

Severely Nonstationary Interference Processing (SNIP), as introduced above, is a flexible STAP detector that is an extension of Sub-CPI Smoothing. In SNIP, as in Sub-CPI Smoothing, DOFs are traded for secondary data by forming partially overlapping sub-CPIs. SNIP, however, even further reduces the size of the required secondary data range window. SNIP is based on an extended form of Kelly's Generalized Likelihood Ratio Test (GLRT) detector [6]. This extension allows for the data vectors to include target signal and may include the test range cell as well as immediate neighborhood range cells. (With most STAP methods, including Sub-CPI Smoothing, guard cells are identified within the sliding window of secondary data range cells. These cells are located in the immediate neighborhood of the test cell and are excluded from use as sources of secondary data vectors to ensure that the secondary data do not contain target signal. This is important to avoid undue cancellation of the target signal. SNIP employs a method of extracting target signal from data vectors so that guard cells are not needed and, in fact, the test cell provides secondary data vectors, as well.

Thus, SNIP further increases the effective number of STAP data vectors without increasing the secondary data range window by allowing the sub-CPIs of the test range cell and what otherwise would be guard cells to contribute interference estimation data vectors. If Sub-CPI Smoothing requires 28 secondary data range cells, and 4 (+/-2) guard cells are required to ensure that a signal from a test cell target is amply suppressed in the secondary data vectors, the secondary data range window is +/-16 cells about the test cell. For SNIP, the window reduces to +/-14 cells or, perhaps even smaller because the test cell data vectors are employed in interference estimation, as well. The penalty is that SNIP is several times more computationally intensive than Sub-CPI Smoothing. In very non-homogeneous environments, such as some urban and littoral, this may shift the problem from whether or not key targets can be detected at all to one of increased processing load.

In addition to a detection statistic, SNIP provides the maximum likelihood estimate of a signal at each range/Doppler cell within a processing window. The signal estimate is useful for target cross section measurement.

The details underlying the SNIP approach follow. As with Sub-CPI Smoothing, let N be the number of channels (elements or beams), M_p be the number of pulses within the CPI, M be the number of digital time delay taps, and L be the number of range cells contributing data vectors. However, these range cells include the test cell and its immediate neighbors. Typically, L is odd. Also, as with Sub-CPI Smoothing, the number of degrees of freedom is $N_{DOF} = MN$, and data vectors are formed by collecting partially overlapping subsets ("sub-CPIs") of M contiguous pulses for each channel. Here, however, the ordering is first by pulse and then by channel. The indices of the pulses of one sub-CPI are offset by one from those corresponding to a neighboring sub-CPI. The number of data vectors per range cell is given by

$$K = M_p - M + 1$$

The total number of data vectors is the product LK.

A target at range (r), Doppler velocity (v), azimuth and elevation direction cosines (u_a, u_e), and of complex strength (b) evokes a response in the ℓ^{th} range cell and k^{th} data vector given by $b\alpha_\ell \beta_k \boldsymbol{s}$ where

$$\boldsymbol{s} = \left[\gamma_0 \boldsymbol{\beta}^{tr}\ \gamma_1 \boldsymbol{\beta}^{tr} \cdots \gamma_{N-1} \boldsymbol{\beta}^{tr}\right]^{tr}$$

$$\boldsymbol{\beta} = \left[\beta_0\ \beta_1 \cdots \beta_{M-1}\right]^{tr}$$

$$\beta_{m'} = exp\left(j\frac{4\pi}{\lambda} m' T v\right) \qquad 0 < m' \le M_p - 1$$

$$\alpha_l = \alpha(r - \ell\rho)$$

ρ denotes the range cell width

α is the pulse compression waveform

tr denotes "transpose"

and $\gamma_n = \gamma_n (u_a, u_e)$ is either the n^{th} beam or n^{th} element (subarray) complex pattern function. These patterns are assumed to include the spatial factors associated with element (or beam) phase center offsets from a common point.

A monostatic mode of operation is assumed. If a bistatic mode is assumed, the "4" in β_m, is replaced with "2" and v is understood to be the sum of two Doppler velocities. One velocity is that of the target with respect to the transmitter platform and the other is that with respect to the receiver platform.

Under the target present hypothesis and assuming independent samples of Gaussian interference plus noise and non-fluctuating target, the likelihood function is given by

$$f_1 = f_1\left(\boldsymbol{R}, b, u_a, u_e, v, r, \{\boldsymbol{x}_{\ell k}\}\right) = \left[\frac{1}{\pi^{MN}|\boldsymbol{R}|}\exp\left(-trace\left(\boldsymbol{R}^{-1}\boldsymbol{T}_1\right)\right)\right]^{LK}$$

where $\boldsymbol{R}$= *MN x MN* interference plus noise covariance matrix; $\boldsymbol{x}_{\ell k}$ is the ℓ^{th} range cell, k^{th} sub-CPI data vector; and

$$\boldsymbol{T}_1 = \frac{1}{LK}\sum_{\ell=0}^{L-1}\sum_{k=0}^{K-1}\left(\boldsymbol{x}_{\ell k} - b\alpha_\ell\beta_k\boldsymbol{s}\right)\left(\boldsymbol{x}_{\ell k} - b\alpha_\ell\beta_k\boldsymbol{s}\right)^H$$

The likelihood function under the target absent hypothesis is given by

$$f_0 = f_0\left(\boldsymbol{R}, \{x\}\right) = \left[\frac{1}{\pi^{MN}|\boldsymbol{R}|}\exp\left(-trace\left(\boldsymbol{R}^{-1}\boldsymbol{T}_0\right)\right)\right]^{LK}$$

where

$$\boldsymbol{T}_0 = \frac{1}{LK}\sum_{\ell=0}^{L-1}\sum_{k=0}^{K-1}\boldsymbol{x}_{\ell k}\,\boldsymbol{x}_{\ell k}^H$$

The quantities $\boldsymbol{R}$, b, u_a, u_e, v, and r are unknown parameters in f_1 that ideally would be estimated by the corresponding quantities that maximize f_1, and $\boldsymbol{R}$ is the unknown parameter in f_0 that is estimated by maximizing f_0. The generalized likelihood ratio test statistic (η) is obtained as the LK^{th} root of the ratio of maximized values of f_1 and f_0. $\boldsymbol{T}_1$ is the matrix that maximizes f_1 with respect to $\boldsymbol{R}$ for fixed b, u_a, u_e, v, and r, and $\boldsymbol{T}_0$ is the matrix that maximizes f_0 with respect to $\boldsymbol{R}$. Consequently, the LK^{th} root of the likelihood ratio after maximizing f_1 and f_0 with respect to the covariance matrix becomes [6]

$$\eta_{\max_{\boldsymbol{R}}} = \frac{|\boldsymbol{T}_0|}{|\boldsymbol{T}_1|}$$

Note that this statistic is an extended form of Kelly's GLRT [6]. As indicated in Section 1, it differs from Kelly by allowing for target presence in multiple data vectors. This modification is essential because the SNIP test range cell provides several data vectors (sub-CPIs), each possibly containing target. This modification should not be confused with the "range spread-target GLRT" formulations of Conte, et al. [7] and Gerlach, et al [8]. The extended form

of Kelly's GLRT applied in SNIP retains the premise of a lone point target. The target, however, appears in multiple data vectors associated not only with the test range cell but also with neighboring range cells that otherwise would be treated as guard cells and omitted from secondary data. As indicated above, SNIP avoids the need for guard cells. All range cells within a small window about the test range cell are valid sources of data vectors.

It can be shown that the expression for b that maximizes $\eta_{\underset{\boldsymbol{R}}{max}}$ for fixed u_a, u_e, v, and r is given by [16]

$$\hat{b} = \frac{\boldsymbol{s}^H \boldsymbol{T}_0^{-1} \boldsymbol{g}}{\left(\sigma_\alpha^2 - \boldsymbol{g}^H \boldsymbol{T}_0^{-1} \boldsymbol{g}\right) \boldsymbol{s}^H \boldsymbol{T}_0^{-1} \boldsymbol{s} + \left| \boldsymbol{s}^H \boldsymbol{T}_0^{-1} \boldsymbol{g} \right|^2}$$

and that for the maximized $\eta_{\underset{\boldsymbol{R}}{max}}$ is given by

$$\eta_{\underset{\boldsymbol{R},b}{max}} = 1 + \frac{\left| \boldsymbol{s}^H \boldsymbol{T}_0^{-1} \boldsymbol{g} \right|^2}{\left(\sigma_\alpha^2 - \boldsymbol{g}^H \boldsymbol{T}_0^{-1} \boldsymbol{g}\right) \; \boldsymbol{s}^H \boldsymbol{T}_0^{-1} \boldsymbol{s}}$$

where

$$\boldsymbol{g} \;\; = \frac{1}{LK} \sum_{\ell=0}^{L-1} \alpha_\ell \sum_k \beta_k^* \boldsymbol{x}_{\ell k}$$

$$\sigma_\alpha^2 = \frac{1}{L} \sum_{\ell=0}^{L-1} \left| \alpha_\ell \right|^2$$

and * denotes "conjugate."

It is often adequate to approximate u_a, u_e, v, and r by the beam center (u_{ao}, u_{eo}), Doppler cell center (v_0), and range cell center (r_0) in determining the detection statistic. The u_a, u_e, and v that maximize f_1 can be determined after detection, thus yielding superresolution in angle ("monopulse") and in Doppler. By following this procedure, the considerable computation required in determining u_a, u_e, and v is limited only to range cells at which detections occur.

2.5 ParaGLRT

Recall that the derivation of the PAMF begins with the Matched Filter Detector (MF). From the form of the MF, $\boldsymbol{R}^{-1/2}$ is shown to be a whitening operation where $\boldsymbol{R}$ is the interference covariance matrix. Let $\boldsymbol{A}^{-1}$ denote the supermatrix of coefficient matrices of the finite order multichannel moving average (MA) linear predictor filters of the data process $\boldsymbol{x}(m)$ of the test range cell, and $\boldsymbol{D}$ denote the corresponding matrix of error covariance matrices. The data process is a Nx1 matrix of received pulses of interference (including noise) at the mth PRI, and the N elements of $\boldsymbol{x}(m)$ correspond to the N spatial channels.

Because

$$\boldsymbol{R}^{-1} = (\boldsymbol{D}^{-1/2}\boldsymbol{A}^{-1})^{H}(\boldsymbol{D}^{-1/2}\boldsymbol{A}^{-1})$$

it follows by substitution into the MF that $\boldsymbol{D}^{-1/2}\boldsymbol{A}^{-1}$ also is a whitening operation. The PAMF is based on this whitening function and the application of two approximations to determining the MA filter coefficient matrices and error covariance matrices. First, the M_p MA filters, of orders ranging from 0 through M_p-1, where M_p is the number of pulses in the coherent processing interval (CPI), are replaced with one MA filter of order P. Second, the coefficient matrices and error covariance matrices are estimated from secondary range cell data. The first approximation (fixed order filter) results in a slow-time temporal smoothing of the filter parameters. This smoothing augments the fast-time smoothing associated with the multiple range cell processing in the second approximation. The result is that the PAMF requires fewer secondary data range cells than do conventional STAP methods to achieve the same detection performance.

This whitening function underlying the PAMF and the independent-sample assumption underlying Kelly's Generalized Likelihood Ratio Test (GLRT) suggests the basis for a powerful GLRT detector, called ParaGLRT, that performs well with only a small number of secondary data range cells. First, the P^{th} order MA coefficient matrices are determined from secondary data range cells, as mentioned above, and the data vectors of the test range cell and secondary data range cells are passed through the filter yielding corresponding whitened data vectors. These whitened data vectors satisfy the independent sample assumption of Kelly's GLRT. The steering vector is passed through the filter, as well.

Next, the whitened data vectors and filtered steering vector are combined into an extended form of Kelly's GLRT. This form is identical to that developed for SNIP, as described in Section 2.4, if each sub-CPI is only one pulse long. It has been observed, however, that sometimes detection performance can be improved if more than one pulse is used to form data vectors; i.e., if overlapping sub-CPIs are formed, as in SNIP. Although overlapping sub-CPIs violates the independent sample assumption underlying the GLRT formulation applied here, it was shown in Section 2.3.1 that the sub-CPI data vectors from the same range cell of data at least contain nearly uncorrelated samples of interference. The examples of applying the ParaGLRT described later in this report employed sub-CPIs larger than one pulse.

Strictly speaking, the "whitening" resulting from the PAMF part of the ParaGLRT is not complete. It is temporal but only partially spatial. Specifically, the original data vectors and steering vector are only passed through the fixed model order filter but not premultiplied by the estimated $\boldsymbol{D}^{-1/2}$ matrix. As indicated above, the complete whitening operation is given by $\boldsymbol{D}^{-1/2}\boldsymbol{A}^{-1}$, where the $\boldsymbol{A}^{-1}$ operation reflects temporal-spatial whitening and the $\boldsymbol{D}^{-1/2}$ operation spatial whitening. The latter spatial whitening is intrinsically included in the GLRT formulation part of ParaGLRT and need not be included in the PAMF part.

Furthermore, a consequence of limiting the whitening to that of a fixed order filter, inherent in the PAMF, is that after passing M_p test cell data vectors through the P^{th} order filter, only M_p-P whitened data vectors are available on output. The CPI for the GLRT operation is effectively shortened by P pulses.

3.0 SIDELOBE TARGET REJECTION

High Doppler or angle sidelobes from strong targets can seriously degrade detection performance because of the propensity to count sidelobes as detections. Several methods for dealing with high sidelobes are briefly described below. Only Doppler sidelobes will be addressed directly, but the methods discussed apply equally well to the spatial domain.

3.1 Low Sidelobe Weighting

Low sidelobe weighting is typically applied in reduced degree of freedom (DOF) post Doppler STAP methods. One disadvantage with this method is that the 3 dB Doppler beamwidth broadens by about 50 percent resulting in decreased resolution. The signal-to-noise ratio reduces by about 1 dB, as well. Two typical such STAP methods are Joint Domain Localized (JDL) and PRI-Staggered. With JDL, Doppler bins and spatial beams are adaptively combined. Several combinations of beams and bins have been considered. A typical JDL configuration is one that employs three Doppler bins (the center bin and immediate adjacent bins) and three similarly deployed beams, for a total of nine DOFs.

The architecture of PRI-Staggered is similar to that of Sub-CPI Smoothing. The major difference is that with PRI-Staggered, the sub-CPIs are typically large and few and each sub-CPI is Doppler filtered. With both methods, the spatial channels can be beams or elements. It is interesting also to compare JDL with Beam Space PRI-Staggered. Whereas both these methods are characterized as "post Doppler beam space and small DOF," the methods are quite different. With JDL, one Doppler filter is applied to each of several adjacent beam channels and several adjacent Doppler frequency bins from each filter are combined adaptively with the corresponding bins from all the filters. With PRI-Staggered, several Doppler filters are applied to each beam channel, the filters differing by a one (or more) pulse delay applied to the pulse stream they operate on, and one corresponding Doppler bin from all the filters are adaptively combined. The point here is that JDL combines several Doppler bins (frequencies) in generating adaptive weights, whereas PRI-Staggered uses only one bin (the target bin) from Doppler filters that operate on delayed pulse streams. PRI-Staggered takes advantage of a favorable Displaced Phased Center Array (DPCA) condition, if present.

JDL was selected as the baseline method with which to compare the low sample support methods developed here. The comparisons are given in Section 7.

3.2 NPAMF and GLRT Based Detectors

The NPAMF (and the normalized form of Sub-CPI Smoothing, as well) are based upon the NMF, as given by (2-2). One advantage, among others, of the NMF over the MF, is that, even when the exact covariance matrix is known, the NMF detector is effective at excluding high Doppler sidelobes from detection. One likely explanation for this is as follows. Consider two range cells, one cell with a strong target and the other with a weak target. The NMF differs from the MF by the $\boldsymbol{x}^H\boldsymbol{R}^{-1}\boldsymbol{x}$ term in the denominator of the NMF. This term is constant for a given range cell; it simply raises or lowers the MF Doppler response curve. This term also is likely to be relatively large for a range cell with a strong target and relatively weak for one with a weak target. The consequential pronounced attenuation of the test statistic corresponding to the

strong-target range cell is the reason that a threshold can be determined that enables detection of both the strong target and the weak target without detecting the strong target sidelobes.

The GLRT, as derived by Kelly [6], also is effective at rejecting sidelobe targets [11]. This benefit is especially true when the number of secondary data range cells is not very large, as is usually the case.

3.3 High Resolution STAP (HRSTAP)

High Resolution STAP was conceived during this effort as another method of reducing the detrimental effects of sidelobes that usually are associated with uniformly weighted steering vectors. Although HRSTAP can be easily generalized to the joint angle-Doppler domain, for expediency, only the Doppler sidelobe problem is addressed here. In fact, a method similar in principle to HRSTAP is the CLEAN algorithm, and CLEAN was developed originally for addressing antenna sidelobes [17], [18]. With CLEAN, the level and Doppler (or angle) of the strongest response is assumed to be a good estimate of the strength and Doppler (or angle) of the strongest target. This information then is used to subtract the strong target from the data, unmasking the relatively weak targets. The procedure is repeated until no responses exceed threshold. The performance of CLEAN is sensitive to the accuracy of estimating the level and location of the strongest target remaining in a given step. Consider a location containing the peak of a strong target and the sidelobe of a weak target. CLEAN can be expected to perform poorly for when the weak target response is within 15 to 20 dB of the strong target response.

With HRSTAP, a STAP detection statistic is computed over a range-Doppler window of interest. A threshold is determined and applied to the resulting amplitude data. Detections are identified with the caveat of at most one detection per range cell, and the corresponding range, Doppler, angle, and RCS of the "targets" are estimated and placed in a target detection file. The detection process then is reapplied at each range cell at which detections were identified. The corresponding secondary data, however, is augmented by generating a signal corresponding to the detected target and injecting it into the secondary data. In this way, adaptivity will null the already detected target and thus allow other, perhaps smaller cross section, targets to appear. The parameters associated with these additional targets are determined and these targets are added to the target detection table. The process is repeated until no additional targets arise. The same threshold value is applied throughout the process.

It remains to be seen if the nulling that results from target injection into the secondary data will sufficiently suppress sidelobes. Otherwise, the method is optimal in that it relies on adaptivity to introduce simultaneously the optimal amount of interference suppression and injected target suppression. (This optimality is one advantage of HRSTAP over deterministic processing of CLEAN.) The potential benefits of High Resolution STAP are that it:

1. Removes false alarms due to large RCS target sidelobes,
2. Preserves maximum target resolution and gain,
3. Detects small targets that are close to large targets in Doppler (and/or angle), and
4. Is completely compatible with all STAP methods.

Implementation and validation of HRSTAP awaits a future effort.

4.0 TARGET-RICH ENVIRONMENT METHODS

The low sample support STAP methods above are particularly effective for scenarios where the selection of STAP secondary data range cells is necessarily limited to the immediate neighborhood of the test range cell (small secondary data range windows) because of non-homogeneous conditions associated with rapidly varying terrain or range dependent clutter ridge loci and/or transmit beam illumination. The target rich environment is a condition that further stresses secondary data selection. Targets are "moving discretes" that, when contained within secondary data range cells, cause errors in the noise covariance estimation. Highway vehicle traffic is a major contributor to a target rich environment.

4.1 Multiple Pass Processing (MPP)

One method developed to improve interference estimation that is seriously degraded by outliers (including target signals contained in the secondary data) is Multiple Pass Processing (MPP). It refers to a particular succession of outlier detection processing steps. The entire data set is processed in each step or "pass." Range cells at which outlier detections are found in a given pass are tagged as containing discretes (possibly targets) and are excluded from use as secondary data cells in successive passes. The range cells corresponding to strong targets, in particular, which are detected in a pass, are removed from secondary data in successive passes.

A non-homogeneity detector, such as the Generalized Inner Product (GIP) [19] or the Innovations Inner Product (IIP) [20], is applied in the first pass. A sliding window (centered about the test range cell) "detection pass" using the detection test statistic follows the first pass, whereby range cells within the secondary data sections of the sliding window that were tagged as containing an outlier in the first pass are replaced with range cells that have not been so tagged and that are nearest to, but outside of, the sliding window. The third pass is another detection pass whereby range cells within the secondary data sections of the sliding window that were tagged as containing outliers in the second pass are replaced with range cells that have not been so tagged and that are nearest to, but outside of, the sliding window. The process is repeated until no new targets are detected.

Consequently, MPP helps alleviate the target rich environment problem by simply eliminating range cells as sources of secondary data if those range cells are tagged as containing outliers in preceding passes. The number of secondary data vectors is kept constant by adding vectors, as needed, from the next available range cells that are closest to the test range cell.

If only the first two passes of MPP are applied, rather than continuing until convergence of detections is observed, the method is referred to as Two-Pass Processing (TPP). The passes are limited to outlier extraction via a nonhomogeneity detector in the first pass followed by a detection pass.

A potential problem with MPP is that the more distant the secondary data range cells are from the test cell, the greater is the potential for degraded estimation performance in severely nonstationary scenarios. An alternative, Generalized Likelihood Ratio Test-based (GLRT-based) method, to MPP that also keeps the number of secondary ranges fixed, but does so without requiring that secondary range cells be drawn from outside the sliding window, is presented in Section 4.2.

4.2 T-SNIP

A STAP method that addresses the performance deteriorating effects of a target-rich environment was developed as an extension of SNIP. As such, it is labeled "T-SNIP." To understand T-SNIP, first consider that the STAP detection statistic in SNIP is an extended form of GLRT. This statistic is most effective if the interference in the data vectors is Gaussian. Moving discretes ("movers"), which may be targets in secondary data range cells, distort the distribution. In T-SNIP the movers are appropriately treated as interference parameters in both the target present and target absent likelihood functions. In principle, the complex amplitudes, Doppler velocities, and range cells of the movers must be estimated by maximizing the likelihood functions. This is very difficult to do exactly. Alternatively, T-SNIP employs an iterative method of estimating these parameters. The method is based on MPP. In MPP, data is successively processed for targets until no additional detections arise. In each processing "pass," range cells are removed from secondary data if they contain detections identified in a previous pass. MPP, thus, essentially "avoids" the moving discretes problem by simply eliminating range cells as secondary data sources if those range cells contained detections in earlier passes. The number of secondary data vectors is kept constant by substituting vectors, as needed, from the next available range cells that are closest to the test range cell.

Although MPP, described in Section 4.1, was found to perform extremely well in a monostatic radar linear array application, as shown in the results of Section 7, MPP is not ideal for many nonlinear array, bistatic radar, or severely heterogeneous terrain scenarios because, in a target rich environment, MPP may be forced to select substitute secondary data range cells that are distant from the test cell, whereas the secondary data range cells are preferably limited to those in the near vicinity of the test cell. T-SNIP, on the other hand, is potentially more attractive here. T-SNIP, as mentioned above, is also a multiple pass process. The difference, however, is that the detections from previous passes are subtracted out of the data applied to the maximum likelihood functions of the GLRT detection statistic of the current pass. Thus, substitute secondary data range cells are not required, and the secondary data range window is not enlarged. Details follow.

Here, let $\boldsymbol{R}$ denote the true interference plus noise covariance matrix for the test range cell, N, the number of degrees of freedom, $\boldsymbol{x}(k)$, the k^{th} range cell data vector (k assumes values 0,1, …, K-1 where $k = 0$ denotes the test range cell, and K denotes the number of range cells providing data vectors for processing a given test range cell), $\boldsymbol{s}$, the steering vector, and $\boldsymbol{t}(k, p)$, the signal vector of the p^{th} mover in the k^{th} secondary data range cell where p assumes values 0,…, $P(k)$-1 and $P(k)$ denotes the number of targets within the k^{th} secondary data range cell.

The interference plus noise hypothesis likelihood function maximized with respect to $\boldsymbol{R}$ is given by

$$\max_{\boldsymbol{R}} f_0 = \left[\frac{1}{(e\pi)^N |\boldsymbol{T}_0|} \right]^K$$

where

$$\boldsymbol{T}_0 = \frac{1}{K}\left(\boldsymbol{x}(0)\boldsymbol{x}^{\boldsymbol{H}}(0) + \sum_{k=1}^{K-1}\left[\boldsymbol{x}(k) - \sum_{p=0}^{P(k)-1} \boldsymbol{t}(k,p) \right]\left[\boldsymbol{x}(k) - \sum_{p=0}^{P(k)-1} \boldsymbol{t}(k,p) \right]^H \right)$$

and e denotes the natural logarithm base. The signal plus interference plus noise hypothesis likelihood function maximized with respect to $\boldsymbol{R}$ is given by

$$\max_{\boldsymbol{R}} f_1 = \left(\frac{1}{(e\pi)^N |\boldsymbol{T}_1|} \right)^K$$

where

$$\boldsymbol{T}_1 = \frac{1}{K}\left((\boldsymbol{x}(0) - b\boldsymbol{s})(x(0) - b\boldsymbol{s})^H + \sum_{k=1}^{K-1}\left[x(k) - \sum_{p=0}^{P(k)-1} \boldsymbol{t}(k,p) \right]\left[\boldsymbol{x}(k) - \sum_{p=0}^{P(k)-1} \boldsymbol{t}(k,p) \right]^H \right)$$

where b = target complex signal voltage. $|\boldsymbol{T}_1|$ must be minimized with respect to b, and $|\boldsymbol{T}_0|$ and $|\boldsymbol{T}_1|$ must be minimized with respect to all $\boldsymbol{t}(k,P)$; $k = 1, \ldots, K\text{-}1$ and $p = 0,\ldots, P(k)\text{-}1$. Closed form solutions for the b and $\boldsymbol{t}$ are difficult to determine. An iterative solution follows.

Initially, set $\boldsymbol{t}(k,P) = 0$ for all $k = 1, \ldots, K\text{-}1$; $p = 0,\ldots, P(k)\text{-}1$. The solution for b that minimizes $|\boldsymbol{T}_1|$ is given by [6]

$$\hat{b} = \frac{\hat{\boldsymbol{s}}^H \hat{\boldsymbol{R}}^{-1} \boldsymbol{x}(0)}{\hat{\boldsymbol{s}}^H \hat{\boldsymbol{R}}^{-1} \hat{\boldsymbol{s}}}$$

where

$$\hat{\boldsymbol{R}} = \frac{1}{K-1} \sum_{k=1}^{K-1} \boldsymbol{x}(k)\, \boldsymbol{x}^H(k)$$

is the sample covariance matrix with respect to the secondary data. Now, the k^{th} root of the generalized likelihood ratio is given by

$$\zeta = \left(\frac{\max\limits_{\boldsymbol{R},b} f_1}{\max\limits_{\boldsymbol{R}} f_0} \right)^{1/K}$$

and a convenient monotonic function of ζ is given by

$$K\eta = K\left(\frac{\zeta - 1}{\zeta} \right)$$

or

$$K\eta = \frac{\left|\boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1}\boldsymbol{x}(0)\right|^2}{\left(\boldsymbol{s}^H \hat{\boldsymbol{R}}^{-1}\boldsymbol{s}\right)\left[1+\frac{1}{K}\left(\boldsymbol{x}(0)\hat{\boldsymbol{R}}^{-1}\boldsymbol{x}(0)\right)\right]}$$

The $\boldsymbol{t}$ for the second iteration is set equal to the $\hat{b}\boldsymbol{s}$ of the first iteration. The procedure continues until the n^{th} and n-1^{st} iterations result in identical values.

4.3 Closely Spaced Targets

One problem with T-SNIP is that a cluster of targets of nearly identical Doppler velocities and angles that occupy a number of adjacent, or nearly so, range cells, may avoid detection. This problem is at least partially alleviated in MPP by application of a power sorting nonhomogeneity detector, such as the GIP or the IIP. It could be alleviated in T-SNIP by applying "single-cell-processing SNIP" in the first pass. With single-cell processing, all data vectors come from the test cell. There are no secondary data range cells (or guard cells). Although single-cell processing suffers reduced nulling resolution, application of it in the first pass improves the probability of detecting cluster targets. Single-cell-processing SNIP performs similarly to direct data domain approaches [21], although it is cast in a statistical framework. MPP can be applied with a single-cell-processing first pass, and, conversely, T-SNIP can be applied with a nonhomogeneity detector first pass.

5.0 ANGLE ESTIMATION

Post adaptive sum and difference beams generally are severely distorted by the adaptive interference suppression process (i.e., STAP), rendering conventional monopulse angle estimation methods unsuitable. Two maximum likelihood azimuth angle estimators that are compatible with STAP were developed. One estimator applies to element-space STAP methods and the other to beam-space STAP methods. Sub-CPI Smoothing, SNIP, and T-SNIP can be formulated for element-space channels or for beam-space channels. For both estimators, the log likelihood function is expanded in a Taylor series and the first three terms (constant, linear and quadratic) are applied in deriving closed-form solutions for the azimuth angle estimate. These methods have yielded good results with simulated data as shown in Section 7.3.6. A detailed derivation of the beam-space angle estimator follows. The derivation for the element-space estimator is similar.

Let $\boldsymbol{z}_l'(\boldsymbol{i})$ be the i,l^{th} temporarily whitened data vector obtained in the "PAMF step" of ParaGLRT. Here, i assumes values 1, 2, …, L_P where $L_P = M_P\text{-}P$ is the number of vectors per range cell, and l assumes values 1, 2, … L where L is the number of secondary data range cells from which the vectors are obtained. Let $\boldsymbol{s}'(\boldsymbol{i})$ be the corresponding steering vectors. The number of elements in each vector equals the number of channels, N. Consider beam space STAP with $N = 4$ low-sidelobe weighted beams: sum (Σ), difference (Δ), and two offset sum beams (A_1, A_2). Define

$$\boldsymbol{s} = \begin{bmatrix} \Sigma \\ \Delta \\ A_1 \\ A_2 \end{bmatrix}$$

where

$$\Sigma = \sum_{m=1}^{M} T_m \, e^{ikX_m(u-u_0)}$$

$$D = \sum_{m=1}^{M} B_m \, e^{ikX_m(u-u_0)}$$

$$A_1 = \sum_{m=1}^{M} A_m^{(1)} \, e^{ikX_m(u-u_0)}$$

$$A_2 = \sum_{m=1}^{M} A_m^{(2)} \, e^{ikX_m(u-u_0)}$$

Here, u is the pattern direction cosine, u_0 is the boresight direction cosine, M is the number of array elements, X_m is the m^{th} element location (a linear array is assumed), and k is given by $2\pi/\lambda$. Also, the parameters T_m , B_m , $A_m^{(1)}$, and $A_m^{(2)}$ are the weights for the corresponding beams.

Let $\boldsymbol{s}' = \boldsymbol{s}'(0)$ be the result of passing $\boldsymbol{s}$ through the temporal whitening filter and $\boldsymbol{s}'(\boldsymbol{i}) = \beta_i \, \boldsymbol{s}'$ where β_i is defined in Section 2.4. The sample "noise" covariance matrix is:

$$\hat{\boldsymbol{R}} \;=\; \frac{1}{\boldsymbol{L}\boldsymbol{L}_{\boldsymbol{P}}} \sum_{\boldsymbol{l}=1}^{\boldsymbol{L}} \sum_{\boldsymbol{i}=0}^{\boldsymbol{L_P}-1} \boldsymbol{z}_{\boldsymbol{l}}'(\boldsymbol{i})\, \boldsymbol{z}_{\boldsymbol{l}}'^{\boldsymbol{H}}(\boldsymbol{i})$$

and under the target present hypothesis, the likelihood function is given by

$$\boldsymbol{f}_1(\boldsymbol{z}_0') \;=\; \left[\frac{1}{\pi^N \left|\hat{\boldsymbol{R}}\right|}\; \boldsymbol{e}^{-\frac{1}{\boldsymbol{L_P}} \sum_{i=0}^{\boldsymbol{L_P}-1} [\boldsymbol{z}_0'(\boldsymbol{i})-\boldsymbol{b}\beta_{\boldsymbol{i}}\, \boldsymbol{s}']^{\boldsymbol{H}}\; \hat{\boldsymbol{R}}^{-1} [\boldsymbol{z}_0'(\boldsymbol{i})-\boldsymbol{b}\beta_{\boldsymbol{i}} \boldsymbol{s}']} \right]^{\boldsymbol{L_P}}$$

where the $\boldsymbol{z}_{\boldsymbol{0}}'(\boldsymbol{i})$ refer to the test range cell vectors and b is the complex target signal strength. Given $\hat{\boldsymbol{R}}$, maximizing f_1 implies minimizing the function Q given by

$$\boldsymbol{Q} \;=\; \frac{1}{\boldsymbol{L_P}} \sum_{\boldsymbol{i}=0}^{\boldsymbol{L_P}-1} \left[\boldsymbol{z}_0'(\boldsymbol{i}) - \boldsymbol{b}\beta_{\boldsymbol{i}}\; \boldsymbol{s}'\right]^{\boldsymbol{H}}\; \hat{\boldsymbol{R}}^{-1}\; \left[\boldsymbol{z}_0'(\boldsymbol{i}) - \boldsymbol{b}\beta_{\boldsymbol{i}}\; \boldsymbol{s}'\right]$$

with respect to b, which results in

$$\hat{\boldsymbol{b}} \;=\; \frac{\boldsymbol{s}'^{\boldsymbol{H}}\; \hat{\boldsymbol{R}}^{-1} \boldsymbol{g}}{\boldsymbol{s}'^{\boldsymbol{H}}\, \hat{\boldsymbol{R}}^{-1} \boldsymbol{s}'}$$

where

$$\boldsymbol{g} \;=\; \frac{1}{\boldsymbol{L_P}} \sum_{\boldsymbol{i}=0}^{\boldsymbol{L_P}-1} \beta_{\boldsymbol{i}}^{-1} \boldsymbol{z}_0'(\boldsymbol{i})$$

Now, with $\hat{b}$ substituted for b in Q, minimizing Q with respect to u implies maximizing $\hat{Q}$ with respect to u, where

$$\hat{\boldsymbol{Q}} \;=\; \frac{\boldsymbol{s}'^{\boldsymbol{H}}\; \hat{\boldsymbol{R}}^{-1}\; \boldsymbol{g}\; \boldsymbol{g}^{\boldsymbol{H}}\; \hat{\boldsymbol{R}}^{-1}\; \boldsymbol{s}'}{\boldsymbol{s}'^{\boldsymbol{H}}\; \hat{\boldsymbol{R}}^{-1}\; \boldsymbol{s}'}$$

The second order Taylor series expansion of $\hat{Q}$ about u_0 is

$$\begin{aligned} \tilde{Q} \;&=\; \hat{Q}(u_0) \;+\; \left.\frac{\partial \hat{Q}}{\partial u}\right|_{u=u_0} (u-u_0) \;+\; \frac{1}{2} \left.\frac{\partial^2 \hat{Q}}{\partial u^2}\right|_{u=u_0} (u-u_0)^2 \\ &=\; \hat{Q}(u_0) \;+\; A(u_0)(u-u_0) \;+\; \frac{1}{2} B(u_0)(u-u_0)^2 \end{aligned}$$

Setting $\frac{\partial \tilde{Q}}{\partial u} = 0$ results in $u = u_0 - \frac{A(u_0)}{B(u_0)}$, where

$$\boldsymbol{A} \;=\; \left.\frac{\partial \hat{\boldsymbol{Q}}}{\partial \boldsymbol{u}}\right|_{\boldsymbol{u}=\boldsymbol{u}_0} \;=\; 2\boldsymbol{k}\left[\left|\hat{\boldsymbol{b}}_0\right|^2 \; Im\left(\boldsymbol{s}_0'^{\boldsymbol{H}} \; \hat{\boldsymbol{R}}^{-1} \; \boldsymbol{s}_{\boldsymbol{u}}'\right) - Im\left(\hat{\boldsymbol{b}}_0 \; \boldsymbol{g}^{\boldsymbol{H}} \; \hat{\boldsymbol{R}}^{-1} \; \boldsymbol{s}_{\boldsymbol{u}}'\right)\right]$$

$$\boldsymbol{B} \;=\; \left.\frac{\partial^2 \hat{\boldsymbol{Q}}}{\partial \boldsymbol{u}^2}\right|_{\boldsymbol{u}=\boldsymbol{u}_0} \;=\; 2\boldsymbol{k}^2 \left\{\frac{\left|\boldsymbol{s}_{\boldsymbol{u}}'^{\boldsymbol{H}} \; \hat{\boldsymbol{R}}^{-1} \; \boldsymbol{g}\right|^2}{\boldsymbol{s}_{\boldsymbol{o}}'^{\boldsymbol{H}} \; \hat{\boldsymbol{R}}^{-1} \; \boldsymbol{s}_0'} \;-\; Re\left(\hat{\boldsymbol{b}}_0 \; \boldsymbol{g}^{\boldsymbol{H}} \; \hat{\boldsymbol{R}}^{-1} \; \boldsymbol{s}_{\boldsymbol{uu}}'\right) \;-\; 4\,\frac{Im\left(\boldsymbol{s}_0'^{\boldsymbol{H}} \; \hat{\boldsymbol{R}}^{-1} \; \boldsymbol{s}_{\boldsymbol{u}}'\right)}{\boldsymbol{s}'^{\boldsymbol{H}} \; \hat{\boldsymbol{R}}^{-1} \; \boldsymbol{s}_0'}\right.$$

$$\left. \bullet \; Im\left(\hat{\boldsymbol{b}}_0 \; \boldsymbol{g}^{\boldsymbol{H}} \; \hat{\boldsymbol{R}}^{-1} \; \boldsymbol{s}_0'\right) + \left|\hat{\boldsymbol{b}}_0\right|^2 \left[Re\left(\boldsymbol{s}_{\boldsymbol{o}}'^{\boldsymbol{H}} \; \hat{\boldsymbol{R}}^{-1} \; \boldsymbol{s}_{\boldsymbol{uu}}'\right) - \boldsymbol{s}_{\boldsymbol{u}}'^{\boldsymbol{H}} \; \hat{\boldsymbol{R}}^{-1} \; \boldsymbol{s}_{\boldsymbol{u}}' + 4\,\frac{\left(Im\left(\boldsymbol{s}_0'^{\boldsymbol{H}} \; \hat{\boldsymbol{R}}^{-1} \; \boldsymbol{s}_{\boldsymbol{u}}'\right)\right)^2}{\boldsymbol{s}_0'^{\boldsymbol{H}} \; \hat{\boldsymbol{R}}^{-1} \; \boldsymbol{s}_0'}\right]\right\}$$

Re denotes real part, *Im* denotes imaginary part,

$$\boldsymbol{s}_0' \;=\; \boldsymbol{s}'\Big|_{\boldsymbol{u}=\boldsymbol{u}_0}$$

$$\hat{\boldsymbol{b}}_0 \;=\; \hat{\boldsymbol{b}}\Big|_{\boldsymbol{u}=\boldsymbol{u}_0} \;=\; \frac{\boldsymbol{s}_0'^{\boldsymbol{H}} \; \hat{\boldsymbol{R}}^{-1} \; \boldsymbol{g}}{\boldsymbol{s}_0'^{\boldsymbol{H}} \; \hat{\boldsymbol{R}}^{-1} \; \boldsymbol{s}_0'}$$

and $\boldsymbol{s}_{\boldsymbol{u}}'$ and $\boldsymbol{s}_{\boldsymbol{uu}}'$ are the outputs of passing $\boldsymbol{s}_{\boldsymbol{u}}$ and $\boldsymbol{s}_{\boldsymbol{uu}}$, respectively, through the temporal whitening filter where

$$\boldsymbol{s}_{\boldsymbol{u}} = \frac{1}{il}\,\frac{\partial \boldsymbol{s}}{\partial u}\Big|_{\boldsymbol{u}=\boldsymbol{u}_0} = \begin{pmatrix} \sum_{\boldsymbol{m}} \boldsymbol{T}_{\boldsymbol{m}} \; \boldsymbol{X}_{\boldsymbol{m}} \\ \sum_{\boldsymbol{m}} \boldsymbol{B}_{\boldsymbol{m}} \; \boldsymbol{X}_{\boldsymbol{m}} \\ \sum_{\boldsymbol{m}} \boldsymbol{A}_{\boldsymbol{m}}^{(1)} \; \boldsymbol{X}_{\boldsymbol{m}} \\ \sum_{\boldsymbol{m}} \boldsymbol{A}_{\boldsymbol{m}}^{(2)} \; \boldsymbol{X}_{\boldsymbol{m}} \end{pmatrix}, \qquad \boldsymbol{s}_{\boldsymbol{uu}} = \frac{1}{l^2}\,\frac{\partial^2 \boldsymbol{s}}{\partial u^2}\Big|_{\boldsymbol{u}=\boldsymbol{u}_0} = \begin{pmatrix} \sum_{\boldsymbol{m}} \boldsymbol{T}_{\boldsymbol{m}} \; \boldsymbol{X}_{\boldsymbol{m}}^2 \\ \sum_{\boldsymbol{m}} \boldsymbol{B}_{\boldsymbol{m}} \; \boldsymbol{X}_{\boldsymbol{m}}^2 \\ \sum_{\boldsymbol{m}} \boldsymbol{A}_{\boldsymbol{m}}^{(1)} \; \boldsymbol{X}_{\boldsymbol{m}}^2 \\ \sum_{\boldsymbol{m}} \boldsymbol{A}_{\boldsymbol{m}}^{(2)} \; \boldsymbol{X}_{\boldsymbol{m}}^2 \end{pmatrix}$$

6.0 KNOWLEDGE BASED STAP

Knowledge based STAP refers to the use of map data, Intelligence, assets deployment information, etc. to supplement radar data in optimizing STAP systems. Section 6.1 describes the potential of using terrain knowledge to optimize sample support size. Section 6.2 describes an optimal means of implementing knowledge of potentially strong scattering structures (that give rise to "discretes") into STAP.

6.1 Optimal Parameter Selection

A correlation between range cell sub-windows and environment type has been explored with the KASSPER Challenge data described in Section 7.1. Such a correlation would enable a knowledge processor to identify the optimal secondary data size commensurate with a given region of the environment. The example described in Section 7.5 suggests that detectability might be improved with knowledge-assisted selection of secondary data size. For the NPAMF ("One Pass"), 106 targets were detected using the optimum secondary data size for each subwindow versus 104 targets detected using K=4 sample support throughout, and for the NPAMF with Innovations Power Sorting ("Two-Pass"), 125 targets detected versus 119. Similarly, other parameters. such as PAMF model order, *P*, have considerable impact on detection and interference rejection performance and are worth investigating.

A particularly active area of research involves scenarios in which the sample support available for training the adaptive processor is limited. Monostatic radars at short range (relative to platform altitude) exhibit clutter Doppler profiles that are variable with range. Similarly, bistatic radars can have clutter Doppler profiles that may be highly variable, not only in range, but in azimuth angle as well. Thus, the support for adaptive weight training is limited to range swaths in which the clutter Doppler does not change significantly. Other related problems include scenarios in which the clutter itself is not homogeneous over significant ranges, e.g., conditions where the terrain type is highly variable, urban environments, etc. Finally, and of particular interest, is the environment in which targets are potentially so dense (relative to the sample support requirements) that they bias the weight training, causing degraded performance of the STAP processor. Such environments include those containing busy highways in a ground moving target indicator (GMTI) mode.

6.2 Discrete Scatterer Signal Injection (DSSI)

It is well known that strong localized scatterers degrade STAP performance. The associated scattered signals are often referred to as "discretes." Discretes in secondary data create additional unwanted nulls in the STAP detector. Further, discretes in the test range cells avoid being nulled by the STAP detector. Both phenomena degrade detector performance. The first phenomenon usually is dealt with by identifying range cells containing discretes and excising these from secondary data. The range cells containing discretes are identified either by data processing or from map databases. A common data processing method of identifying discretes is power sorting, whereby an inner product statistic is determined for all range cells, the statistic is amplitude sorted, a threshold is applied, and range cells corresponding to amplitudes exceeding threshold are proclaimed as unsuited for secondary data. One advantage of data processing is that statistical outliers ("clutter spikes") are excised from training data as well as

range cells containing sources of discretes. Excising statistical outliers is of benefit when the number of appropriate secondary range cells is severely limited as is usually the case with nonhomogenous terrain scenarios and where processing is limited to small "sliding" windows centered about the test cells. This procedure was discussed further in Section 4.0. The advantage of using mapping data, however, is that weaker RCS scatterer discretes are excised as well as stronger ones.

The second phenomenon is more difficult to deal with. Because a discrete in a test cell is absent from the associated secondary data cells, STAP will not suppress the corresponding interference, and the discrete, especially if strong, could mask smaller RCS targets that are not well separated in angle or Doppler from that of the discrete. The conventional approach for suppressing the discrete is to first preprocess the primary and secondary data by deterministically combining elements (or PRIs) to form beam or subarray patterns in space (or time) with nulls in the direction of the discrete (or at the Doppler of the discrete) [22]. STAP is then applied with space-time data vectors formed from the preprocessed data. One disadvantage with this approach is that it limits the applicable STAP methods to those consistent with DOFs derived from spatial (or temporal) beams or large subarrays. Another disadvantage of preprocessing with deterministic nulling is the likely excessive reduction in DOFs available for the subsequent STAP.

A new method was identified that does not suffer these drawbacks. The method is referred to as Discrete Scatterer Signal Injection (DSSI). With DSSI, first the angle, Doppler, and estimated RCS parameter values of a known potential source of discrete for a given range cell under test is determined from map databases and the scenario geometry, platform velocity, antenna beam direction, and so forth. Second, a signal is generated from the parameter values and injected into the secondary data vectors associated with the test cell. Finally, these augmented secondary data vectors are applied in the STAP processing of the test cell. Thus, a full DOF STAP method, such as PAMF, can be applied. Also, the adaptivity inherent in STAP automatically optimizes the trade between mainlobe gain loss and suppression of clutter, discrete, and other interference (jamming, etc.) within the limits of the available DOFs. DSSI was not pursued further during this effort.

7.0 PERFORMANCE EVALUATIONS

7.1 KASSPER Data Cube

The detectors developed here were evaluated by application to the KASSPER Challenge data cube. This data cube was simulated by ISL, Inc. and made available at the KASSPER workshop held on 3 April 2002 in Washington, DC [4]. The scenario included an airborne multichannel monostatic GMTI platform traveling at 100 m/s. The radar had 11 subarray channels, wherein each subarray was composed of eight vertical elements. The data cube contained extremely dense ground targets in site-specific clutter. The scenario locale was from the southwest United States. **Figure 7.1-1** displays the detailed clutter modeling and indicates the range limits of the data cube and the transmit pattern beam width. The data cube spanned 1000 range cells and 32 Doppler resolution cells. The data cube contained a large number of targets within the transmit pattern mainlobe: 226 "cluster" (as in convoys) and 52 background (highway vehicle). **Figure 7.1-2** shows the cluster and background traffic. In addition, a number of stationary discrete scatterer sources were included in the data. These pertained to buildings clustered around towns and towers located on large mountain peaks. Antenna errors were modeled, as well. Range sidelobes and adjacent range cell correlation associated with finite bandwidth waveforms were not modeled. However, guards were included in the processing reported here nevertheless. The radar parameters are listed in **Table 7.1-1**.

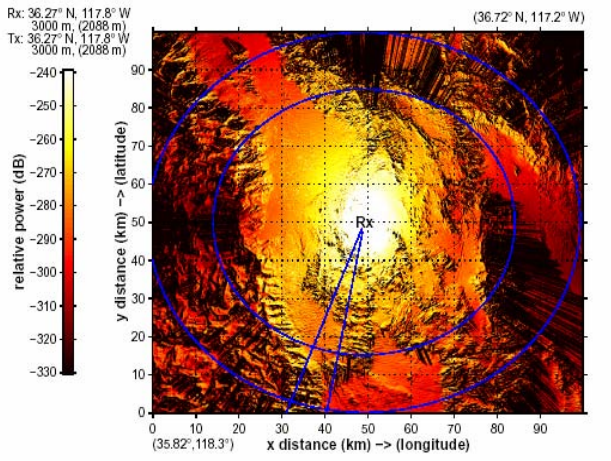

Figure 7.1-1. Clutter Map

(Courtesy ISL, Inc.)

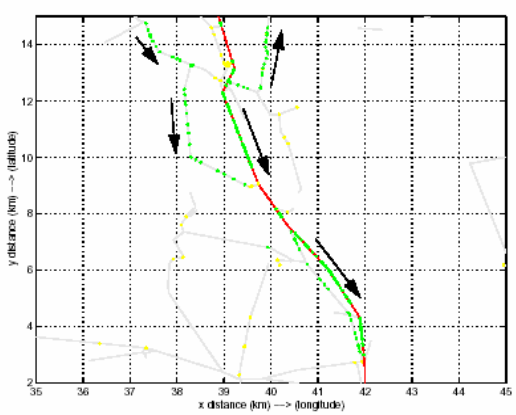

Figure 7.1-2. Close-up view of the data cube region. Yellow dots are background traffic. Green dots are clustered vehicles; arrows indicate their general direction of travel. Red line is major highway

(Courtesy ISL, Inc.)

Table 7.1-1. Radar Parameters

(Courtesy ISL, Inc.)

Parameter	Value
Carrier frequency	1,240 MHz
Bandwidth	10 MHz
Number of pulses	32
Minimum range (one way	35 km
Maximum range (one way)	50 km
Pulse repetition frequency	1,984 Hz
Peak power	15 kW
Duty factor	10%
Noise figure	5 dB
System losses	9 dB

Figures 7.1-3 and 7.1-4 show range/Doppler response diagrams resulting from processing range cells 1-250 and range cells 501-750, respectively, with a low sidelobe weighted Doppler filter. In each case, the threshold commensurate with about 55 total detections was determined and the results plotted such that black indicates exceeding threshold and yellow below threshold. Crosses indicate true cluster targets and zeros indicate true background traffic. A “t” denotes a tower and a “b” a building. Clearly, most of the targets are sub-clutter and not detectable with only deterministic filtering.

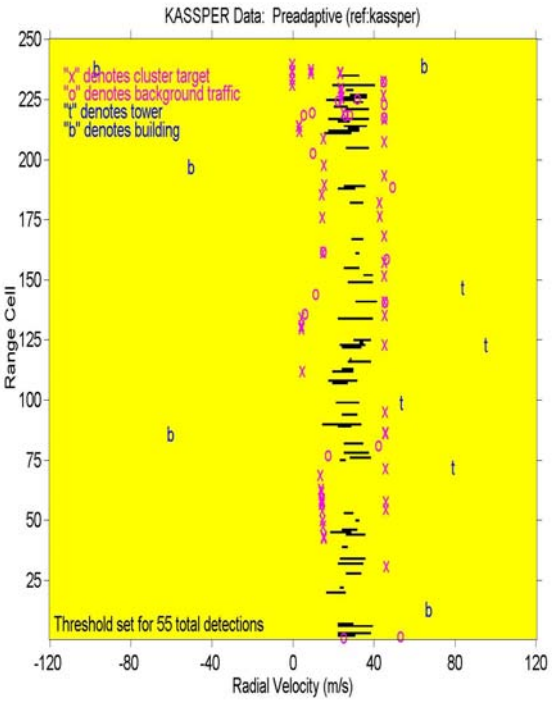

Figure 7.1-3. Nonadaptive Doppler Processed Data; Range Cells 1-250

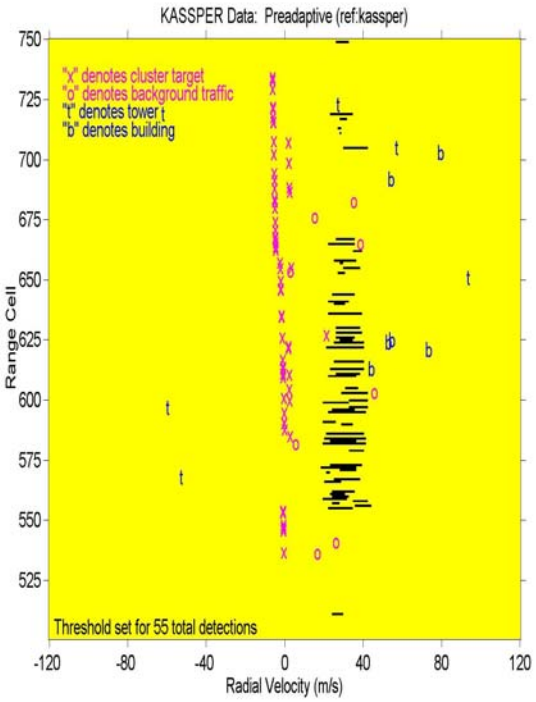

Figure 7.1-4. Nonadaptive Doppler Processed Data; Range Cells 501-750

7.2 Matched Filter Detectors

Results of applying the matched filter detector (MF) and the normalized matched filter detector (NMF) to the first 250 range cells of the KASSPER Challenge data cube are shown in **Figures 7.2-1 and 7.2-2**. These detectors are realizable because the KASSPER Challenge Data includes the true interference covariance matrix for all range cells. The threshold is set for about 55 target detections. Almost all targets are detected with both detectors. Clearly, however, the NMF does a far better job of rejecting sidelobe targets. The probable reason for this was discussed in Section 3.2.

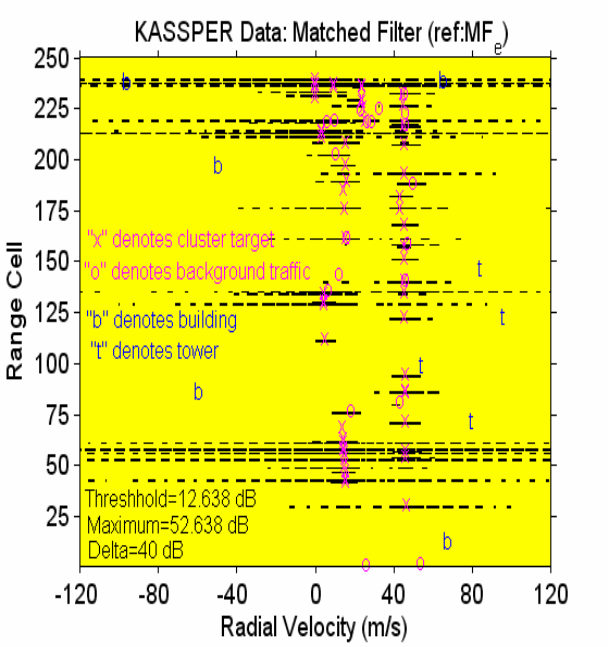

Figure 7.2-1. Matched Filter (MF) Detector Performance

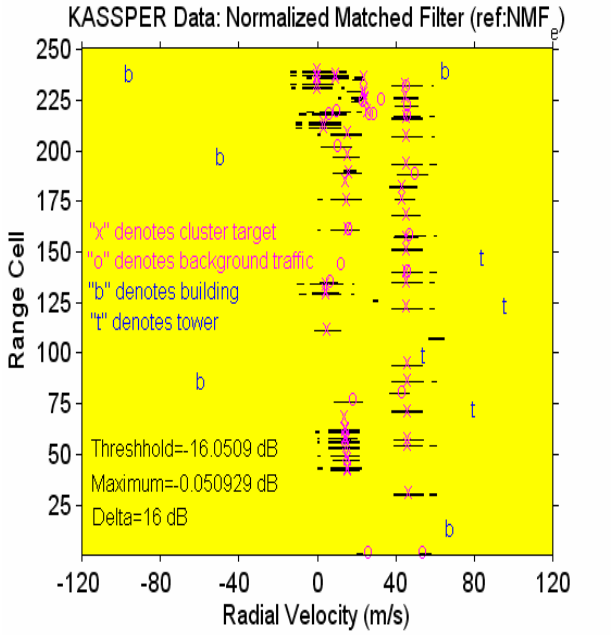

Figure 7.2-2. Normalized Matched Filter (NMF) Detector Performance

7.3 Sidelobe Target Rejection Adaptive Detectors

7.3.1 NPAMF

The Normalized Parametric Adaptive Matched Filter (NPAMF) detector was applied to the KASSPER Challenge data cube in a sliding window architecture. The Kalman filter method of computing the error predictive filter coefficients was used. The model order was 6, the number of guard cells was 4 (2 on each side of the test cell), and the number of secondary data cells was 4 (2 on each side of the test cell). For comparison, the NMF and the Normalized Joint Domain Localized (NJDL) STAP detector were also applied. NJDL was applied with three spatial beams and three Doppler beams resulting in nine degrees of freedom; spatial beams were weighted for 30 dB sidelobes and Doppler beams for 60 dB sidelobes. NJDL employed a test statistic of form similar to that of NMF. NMF was applied instead of MF because of its superior sidelobe target rejection capability. NPAMF was applied twice, once with elements as spatial

channels and again with beams. For the latter, four beams were formed: three 30 dB sidelobe Taylor weighted sum beams and one 30 dB sidelobe Bayliss weighted difference beam.

Figures 7.3.1-1, **7.3.1-2**, and **7.3.1-3** show results of applying NMF, Element Space NPAMF, and NJDL, respectively, to the first 250 range cells of the data. In each case, the threshold commensurate with 55 "total" detections (targets plus false alarms) was determined and the results plotted such that black indicates exceeding threshold and yellow below threshold. As before, crosses indicate true cluster targets and zeros indicate true background traffic. A "t" denotes a tower and a "b" a building. The NMF detector, which makes use of the exact covariance matrix, did extremely well, as expected. The NPAMF appeared to do somewhat better than NJDL in terms of reducing the number of false alarms.

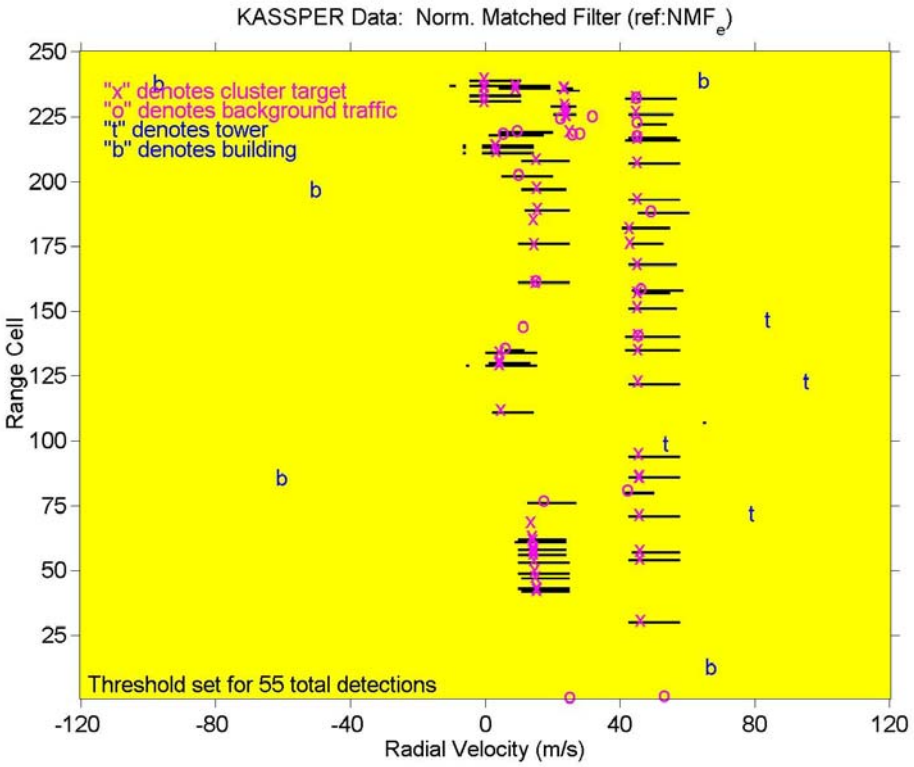

Figure 7.3.1-1. Normalized Matched Filter Processing of KASSPER Data

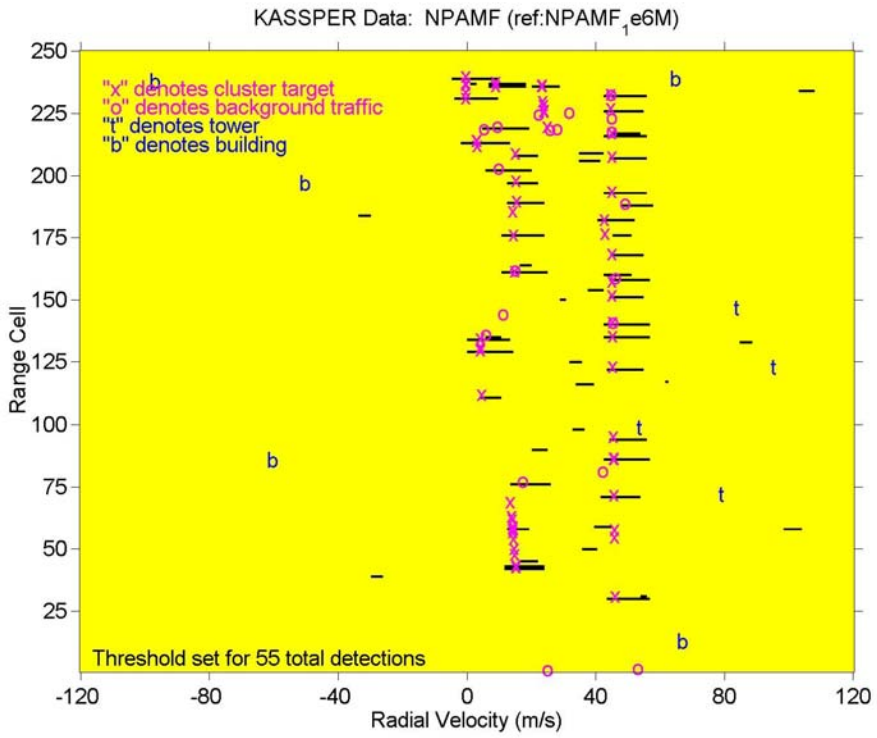

Figure 7.3.1-2. NPAMF Processing of KASSPER Data

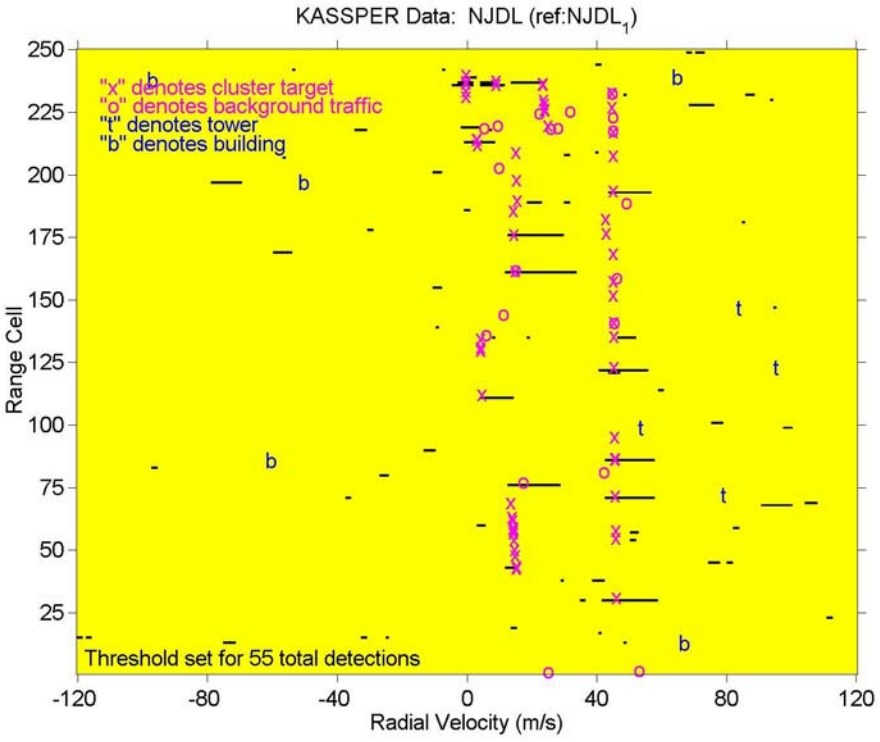

Figure 7.3.1-3. NJDL Processing of KASSPER Data

The detection ratio versus false alarm ratio for the three detectors and, also, the Beam Space NPAMF detector are shown in **Figure 7.3.1-4**. All 1000 range cells were used in generating these curves. The detection ratio is the ratio of the number of targets detected to the total number present (as provided by the KASSPER data ground truth). The false alarm ratio is the ratio of the number of false detections to the number of possible outcomes (32,000 range/Doppler resolution cells). These curves were generated to provide a quantitative measure of comparison between the various detectors. As indicated previously, this measure was feasible because of the large number of resolution cells (32,000) and targets (278). Because a target may impact more than one Doppler resolution cell, a local maxima test was applied in order to avoid over counting detections. Clearly, NPAMF performance exceeded that of NJDL, but fell substantially short of that of NMF. (The dashed line refers to the Beam Space NPAMF.)

Next, TPP (Two-Pass Processing as defined in Section 4.1) was combined with NPAMF(KF) and NJDL and applied to the data. The IIP nonhomogeneity detector was applied in the first pass for both cases. Note that a model order is associated with IIP, just as with PAMF [20]. The model order applied for the IIP pass was 2 and the number of secondary range cells for the IIP pass was 128. The results are shown in **Figure 7.3.1-5**. Little difference in performance is evident with respect to the "single-pass" results.

Next, MPP was combined with the various detectors and applied to the data. Convergence, defined as no change in detections between successive passes, was usually achieved in about four passes. The number of reference data range cells for NPAMF was increased to 20. It is important to note that without MPP, the number of secondary range cells could not be increased beyond 4 or so without observing a drop in performance and increasing the number of secondary range cells is important if the adaptive detector performance is to approach that of the matched filter detector. As is evident in **Figures 7.3.1-6**, NPAMF with MPP performs almost as well as NMF and much better than NJDL with MPP.

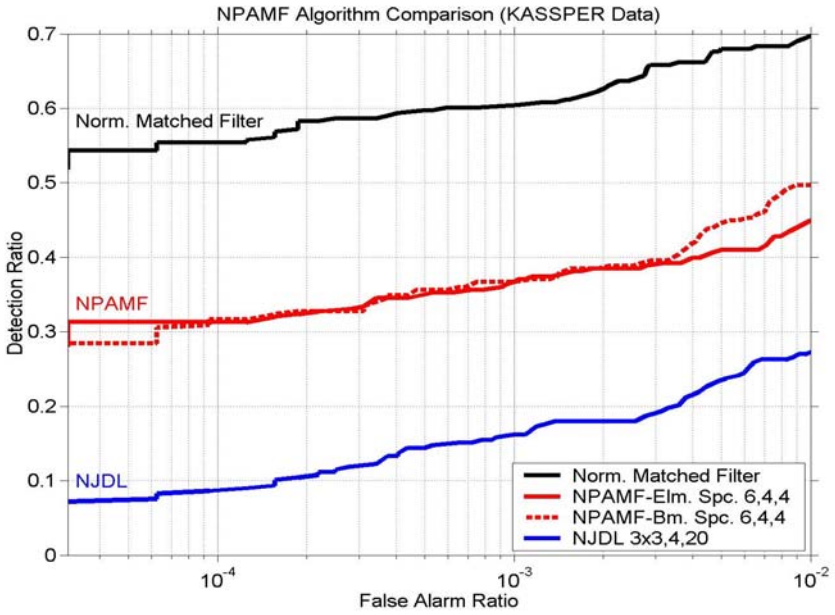

Figure 7.3.1-4. Algorithm Comparison with KASSPER Data

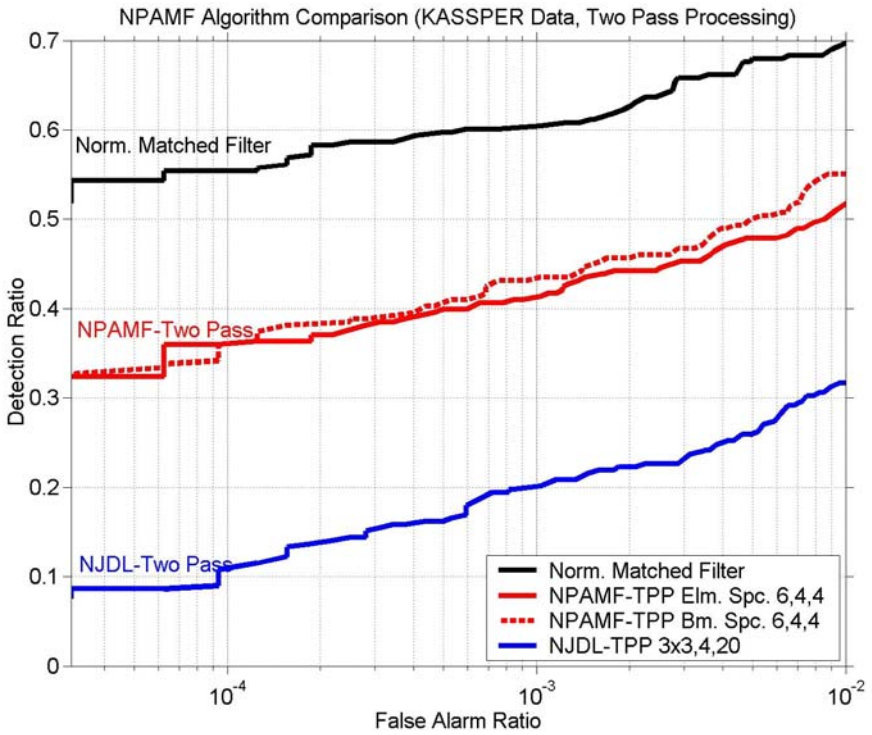

Figure 7.3.1-5. Two Pass Processing Algorithm Comparison

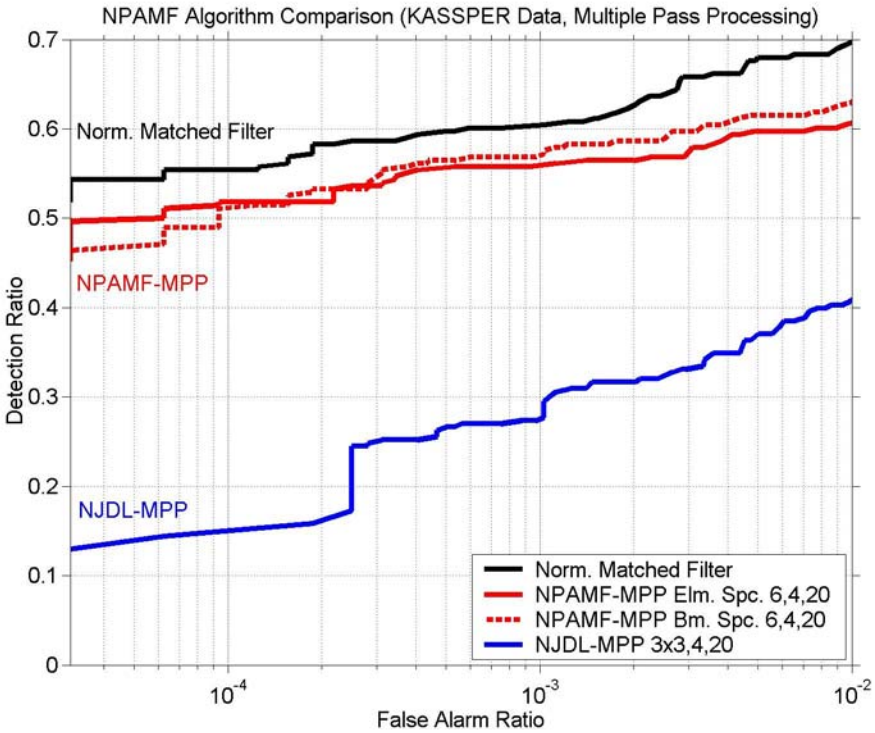

Figure 7.3.1-6. Multiple Pass Processing Algorithm Comparison

7.3.2 ParaGLRT

Figures 7.3.2-1 and 7.3.2-2 show the performance of ParaGLRT with element spatial channels and beam spatial channels, respectively. The NMF, NPAMF, and NJDL are shown for comparison. Model order 6, +/-2 guard cells, four secondary data cells, and sub-CPI size 3 are the relevant parameter values for ParaGLRT. As before, and in all results given in this report, these parameter values were determined to be nearly optimal by observing the results of varying the parameters through limited ranges. Model order 6, +/-2 guard cells, and 4 secondary data cells are the parameter values for NPAMF, and 3 spatial beams by 3 Doppler beams, +/-2 guard cells, and 20 secondary data cells are those for NJDL. As is evident from these figures, NPAMF performs a little better than ParaGLRT for element space and the performance is about the same for beam space. Both methods perform substantially better than NJDL. These comparisons may not be very meaningful because of the propensity for secondary data range cell targets and for these targets to impart signals at angles and Dopplers almost equal to those of test cell targets.

Figures 7.3.2-3 and 7.3.2-4 show the performance of these adaptive detectors in conjunction with Two-Pass Processing. As before, the Innovations Inner Product nonhomogeneity detector with model order 2 and with 128 secondary range cells was applied in the first pass for these detectors. These detectors demonstrate improved performance over the conventional "One-Pass Processing" case.

Figures 7.3.2-5 and 7.3.2-6 show the performance of these adaptive detectors in conjunction with Multiple Pass Processing. The performance improvement over conventional processing and over Two-Pass Processing is dramatic. Apparently MPP thoroughly eliminates range cells containing detections from secondary data range cell windows. The dramatic improvement is thought to be a consequence of the scenario wherein major highways are predominantly aligned parallel to radials from the radar. This results in many vehicles in different range cells imparting signals with nearly equal Dopplers and angles. Note that with MPP, Beamspace ParaGLRT is superior to the other adaptive detectors and performs about as well as the Matched Filter.

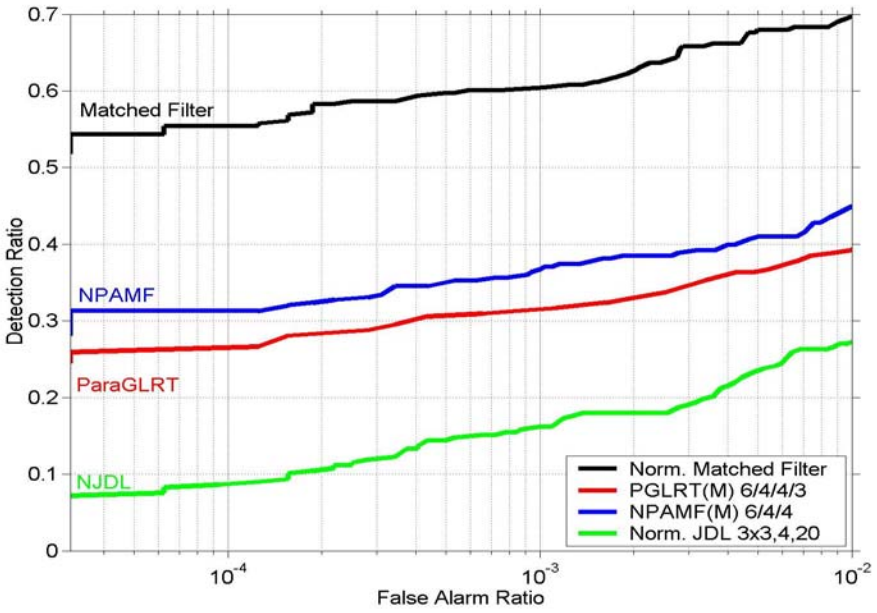

Figure 7.3.2-1. Element Space ParaGLRT

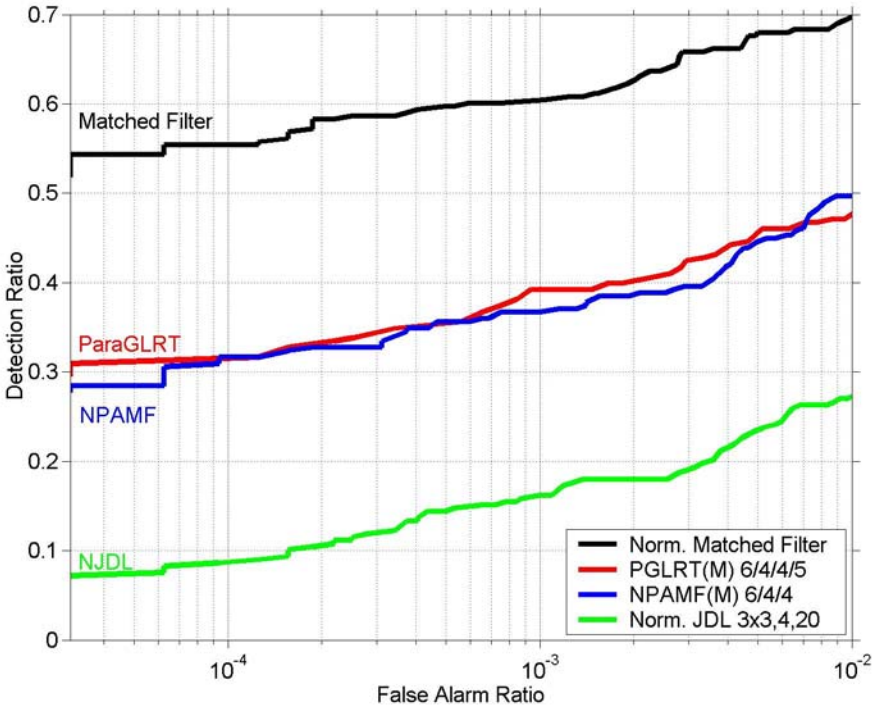

Figure 7.3.2-2. Beam Space ParaGLRT

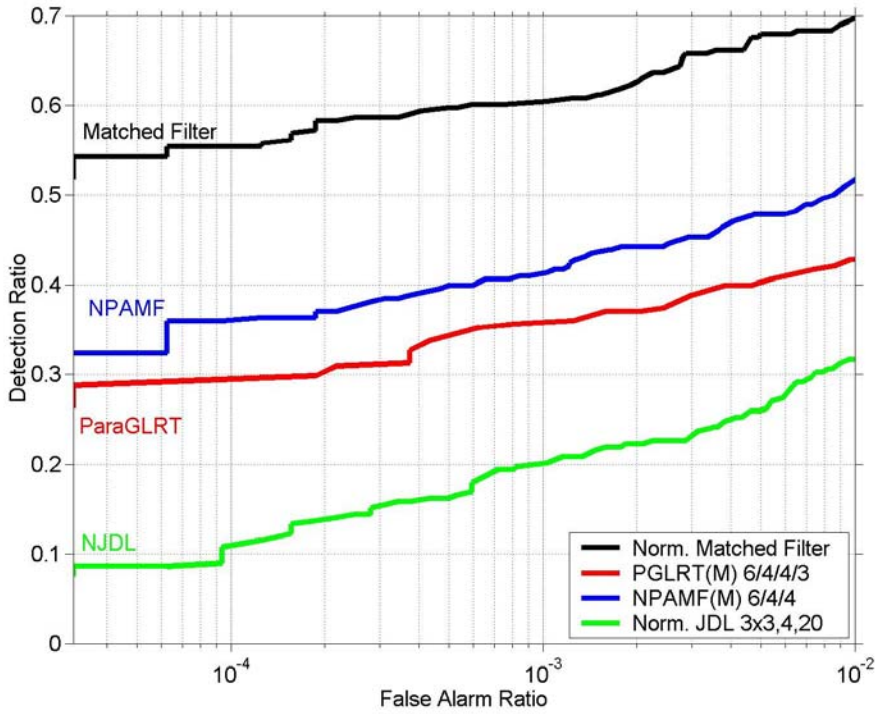

Figure 7.3.2-3. Element Space ParaGLRT with Two-Pass Processing

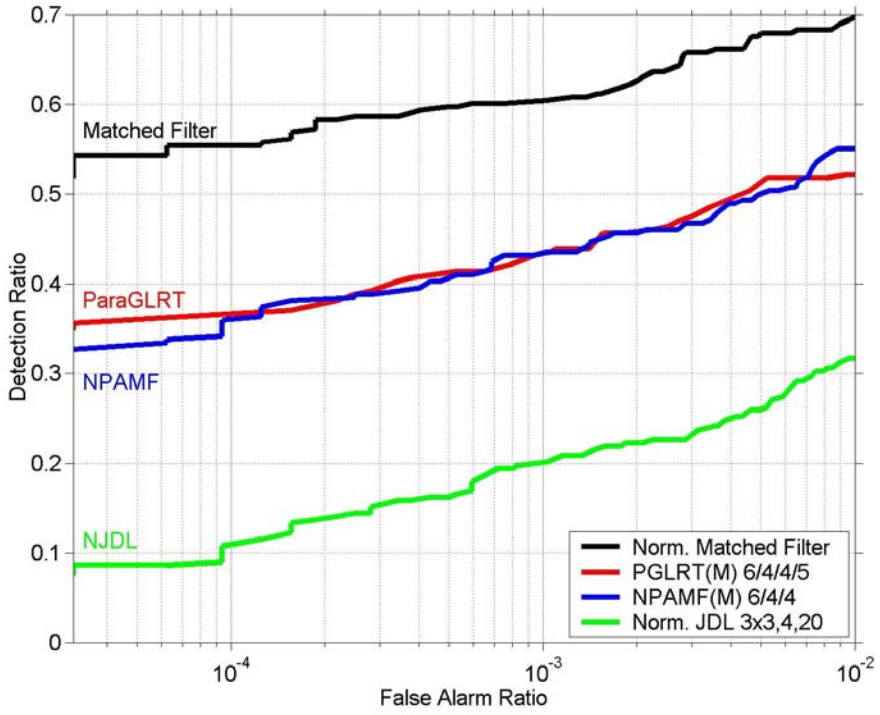

Figure 7.3.2-4. Beam Space ParaGLRT with Two-Pass Processing

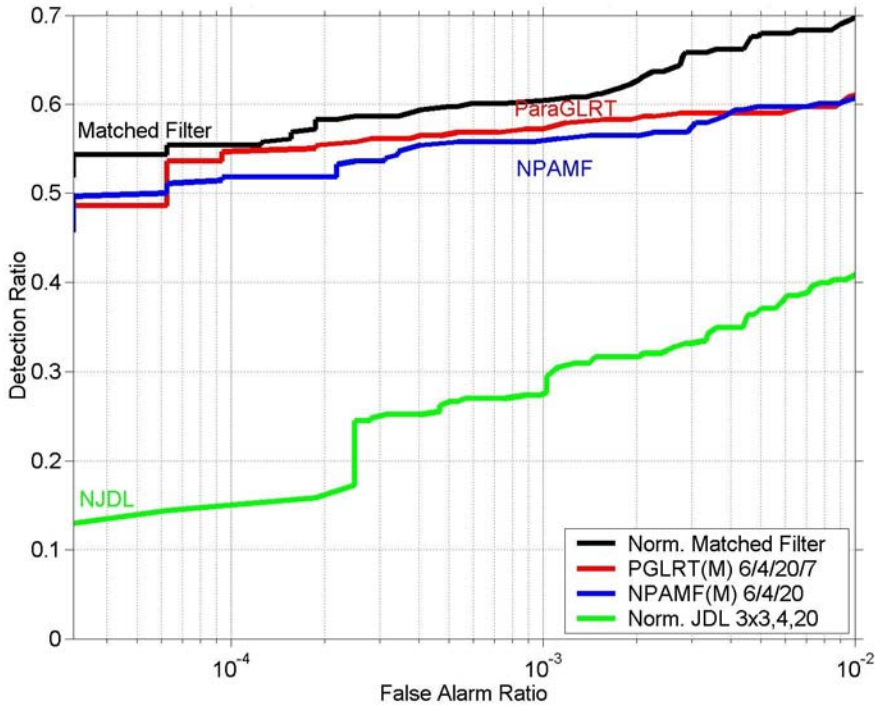

Figure 7.3.2-5. Element Space ParaGLRT with Multiple Pass Processing

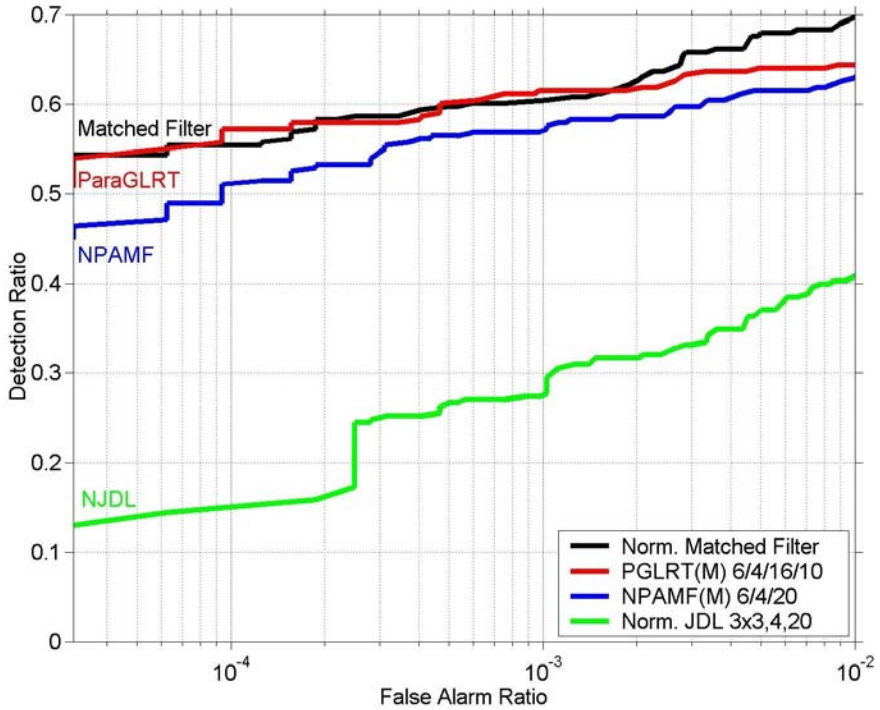

Figure 7.3.2-6. Beam Space ParaGLRT with Multiple Pass Processing

7.3.3 PAMF Model Comparison

The LMS algorithm method for computing PAMF coefficient matrices and error covariance matrix was compared with the Kalman Filter method. For LMS, correlation matrices are computed using $\boldsymbol{x}(m)\boldsymbol{x}(n)^H$, where, as before, the index refers to PRI number. These matrices are averaged over the secondary data range cells, and the resulting averaged matrices are used to estimate the coefficient and error covariance matrices. The potential advantage of the Kalman Filter method is that, because it is recursive, it can lead to a fast running code. **Figure 7.3.3-1** shows the performance of NPAMF with MPP on the KASSPER data for three methods of implementing the LMS algorithm. First, the correlation matrices are averaged over the secondary data range cells and the coefficient matrices and error covariance matrix are computed from the averaged values. Second, the coefficient matrices are averaged and the error covariance matrix is computed from the averaged coefficient matrices. Third, the coefficient matrices and error covariance matrix are averaged. Clearly, averaging the correlation matrices yields the best performance.

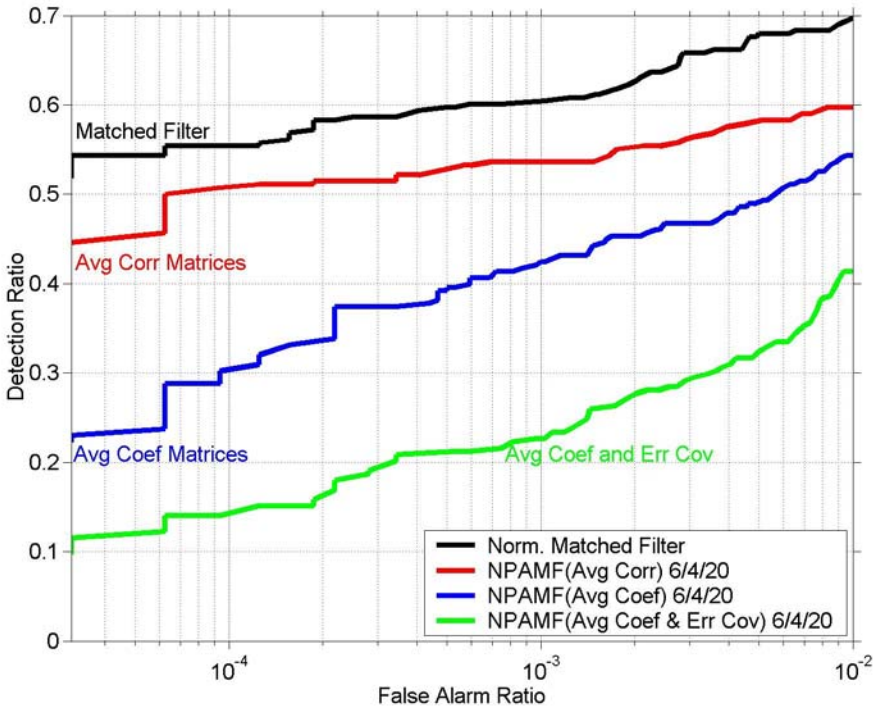

Figure 7.3.3-1. Element Space "LMS" NPAMF with MPP (KASSPER Data)

Figures 7.3.3-2, 7.3.3-3, and 7.3.3-4 compare two of the LMS variants with the Kalman Filter method for "one pass" processing, "two pass" processing, and MPP. Because of the target rich environment, the MPP result is the most meaningful. LMS (with averaging correlation matrices) and Kalman Filter perform about the same.

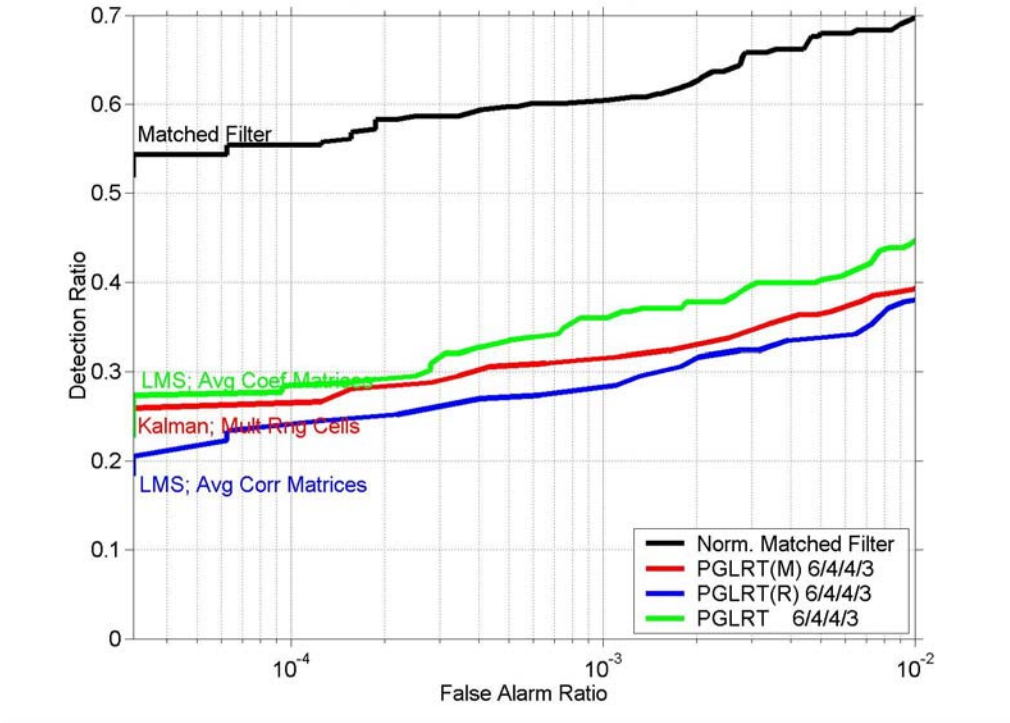

Figure 7.3.3-2. Element Space ParaGLRT "One Pass" (KASSPER Data)

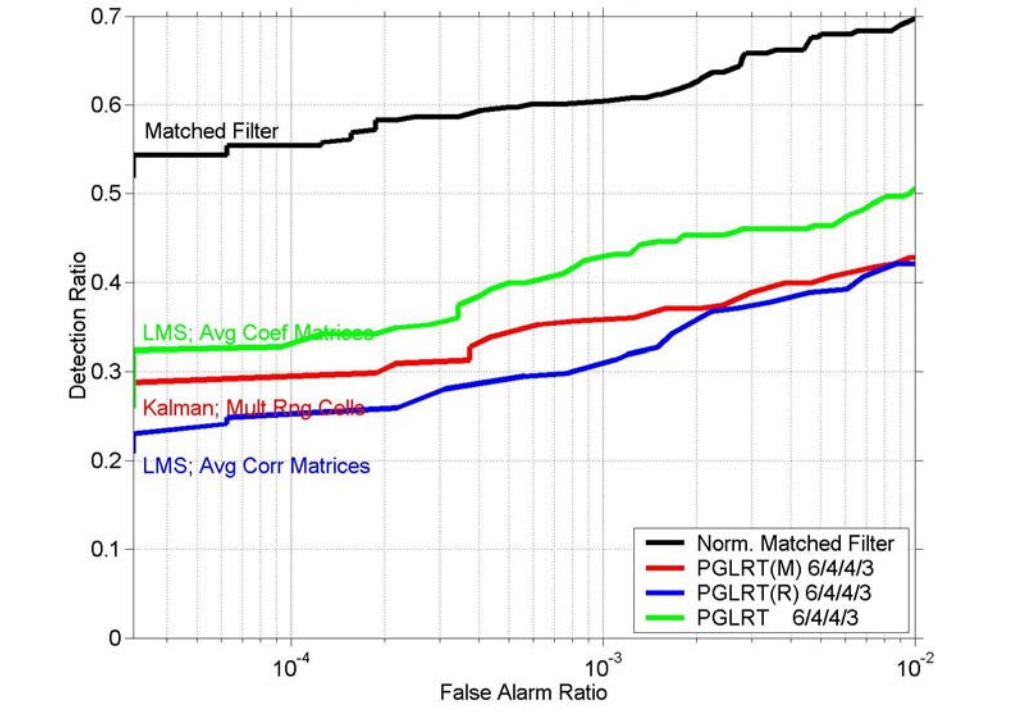

Figure 7.3.3-3. Element Space ParaGLRT "Two-Pass" (KASSPER Data)

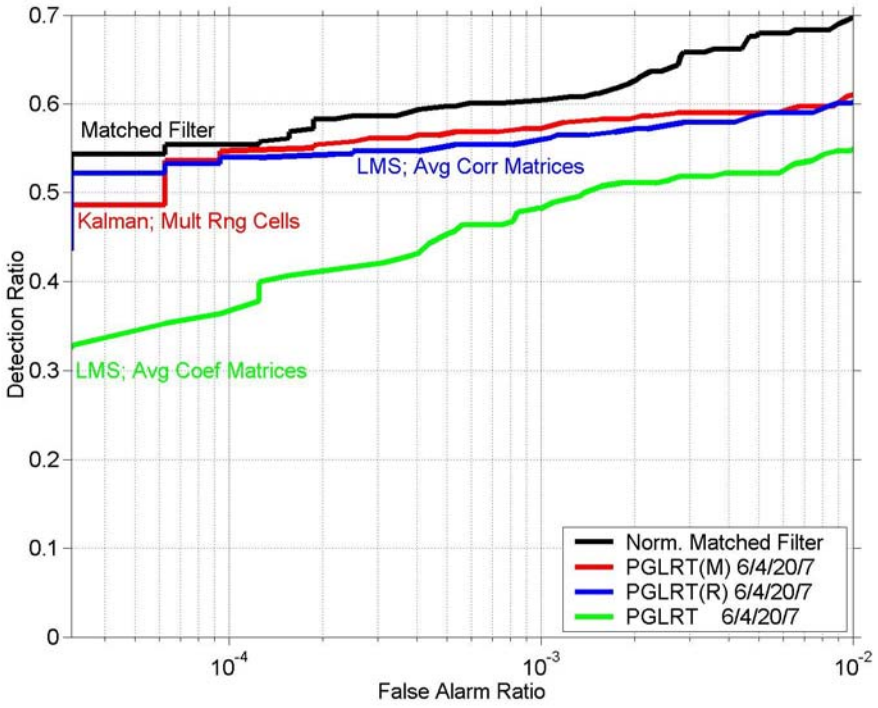

Figure 7.3.3-4. Element Space ParaGLRT with MPP (KASSPER Data)

7.3.4 Normalized Sub-CPI Smoothing

The detection ratio versus false alarm ratio for the normalized forms of the matched filter, Sub-CPI Smoothing, and the Optimal Pre-Doppler STAP (OPDS), as well as JDL, as applied to the KASSPER Challenge data are shown in **Figure 7.3.4-1**. Clearly Sub-CPI Smoothing with only four secondary data range cells is superior to OPDS with four secondary data range cells, and both are superior to JDL with 20 secondary data range cells. OPDS performance was found to deteriorate for $M>3$, when applied to the KASSPER Challenge data. Also, the number of targets detected for the same false alarm rate was almost 50 percent greater for Sub-CPI Smoothing with $M=6$ than for OPDS with $M=3$. The performance of Sub-CPI Smoothing with Multiple Pass Processing (MPP) is presented in the following subsection.

7.3.5 SNIP

Figure 7.3.5-1 shows that SNIP performs about as well as Sub-CPI Smoothing with the same number of secondary data range cells (4), but of course, the range window for SNIP is potentially smaller.

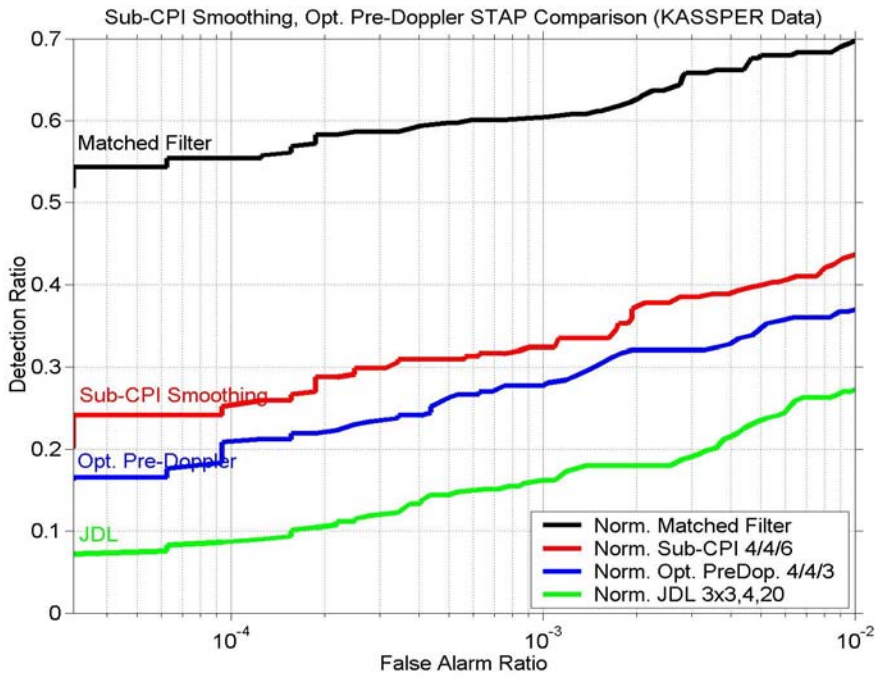

Figure 7.3.4-1. Sub-CPI Smoothing, Opt. Pre-Doppler STAP Comparison (KASSPER Data)

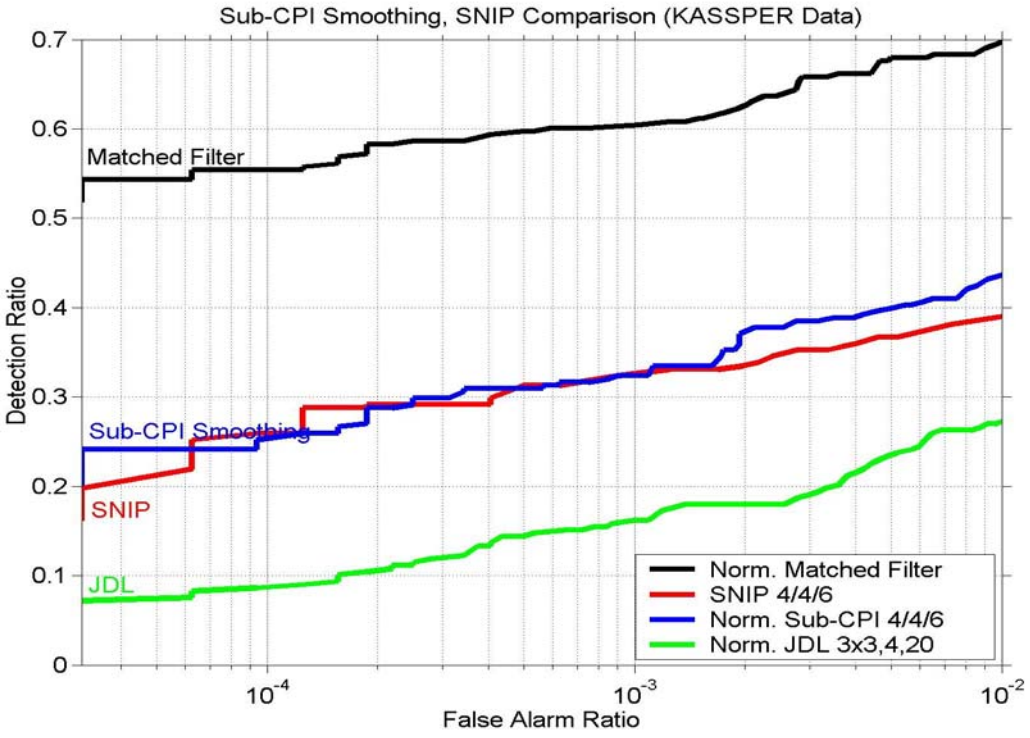

Figure 7.3.5-1. Sub-CPI Smoothing, SNIP Comparison (KASSPER Data)

The results of combining MPP with SNIP and again with JDL are shown in **Figure 7.3.5-2**. SNIP/MPP performs almost as well as the matched filter (MF), whereas JDL/MPP shows only modest improvement over JDL alone. These results are not surprising because the number of secondary data range cells for SNIP/MPP is increased to 20 (same as for JDL/MPP), which amounts to an effective number of secondary data vectors that is about eight times the number of DOFs for SNIP/MPP (as computed per the discussion in Section 2.4 for N = 11, M=6, M_p=32, and L=20) but only about twice for JDL/MPP (20 secondary data vectors divided by 9 DOFs). It is important to note that without MPP it was found that the number of secondary range cells could not be increased beyond about 4 without observing a drop in performance, and increasing the number of secondary range cells is important if the adaptive detector performance is to approach that of the matched filter detector. Convergence for MPP, defined as no change in detections between successive passes, is achieved with about 4 passes.

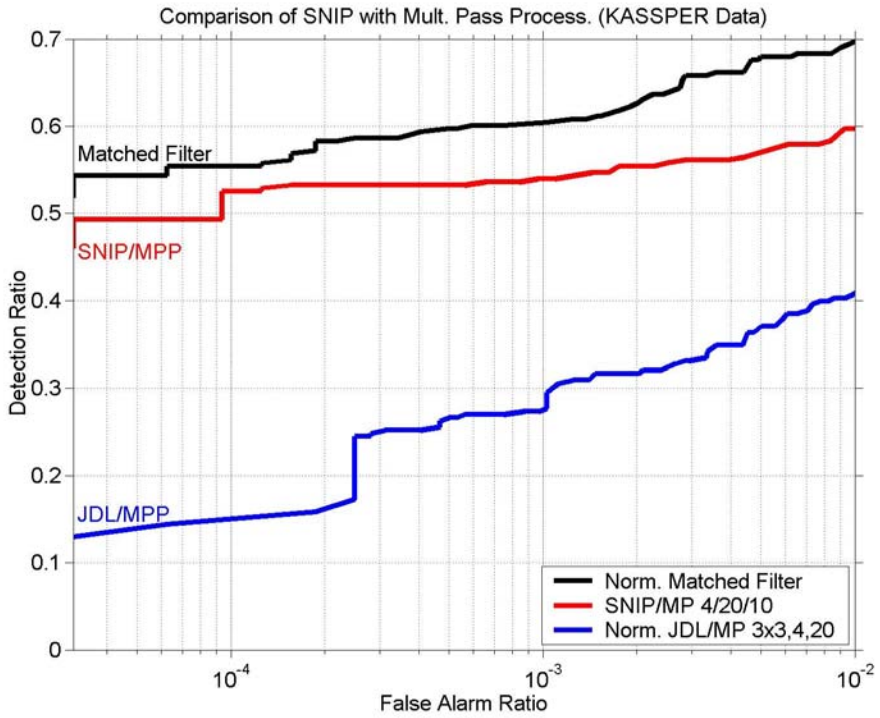

Figure 7.3.5-2. SNIP with Multiple Pass Processing (KASSPER Data)

Element space and beam space SNIP and Sub-CPI Smoothing detectors, in conjunction with MPP, are compared in **Figures 7.3.5-3 and 7.3.5-4**. SNIP and Sub-CPI Smoothing detectors perform comparably and almost as well as the matched filter. It is noted, again, however, that SNIP potentially employs a smaller secondary data range cell window for the same performance because SNIP does not depend on guard cells.

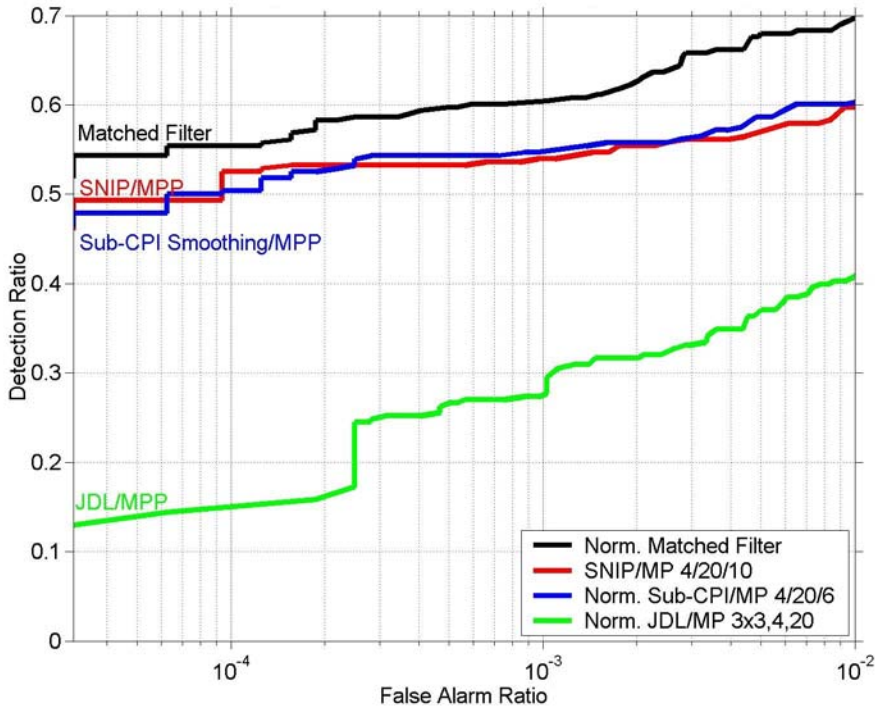

Figure 7.3.5-3. Element Space SNIP and Sub-CPI Smoothing with MPP (KASSPER Data)

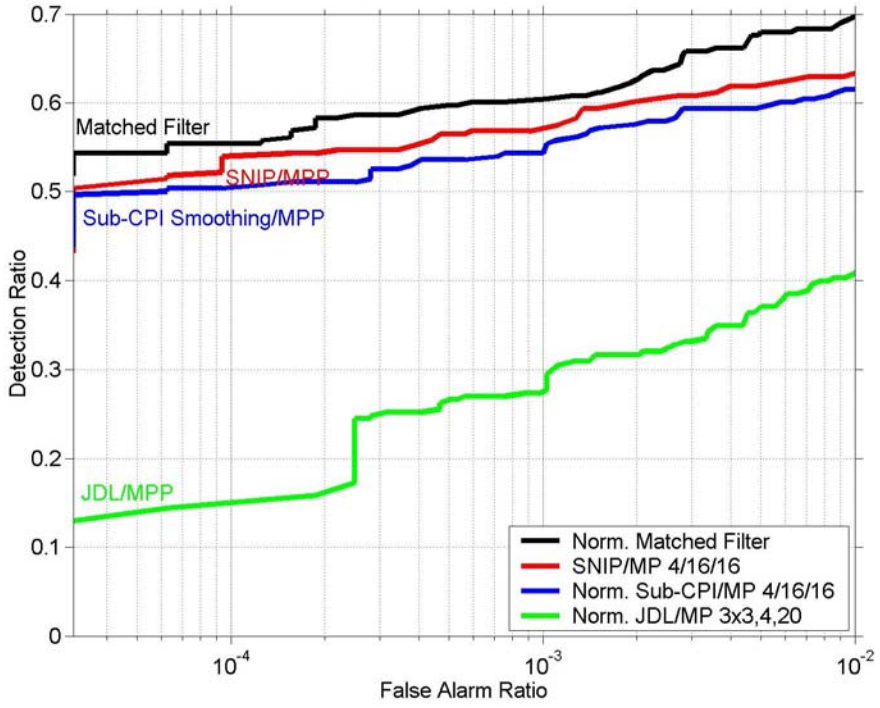

Figure 7.3.5-4. Beam Space SNIP and Sub-CPI Smoothing with MPP (KASSPER Data)

7.3.6 Angle Estimation

Application of the beam-space Maximum Likelihood Angle Estimator derived in Section 5 to the KASSPER Challenge data resulted in the detections plotted in the PPI in **Figure 7.3.6-1**, where the green dots are background traffic, the blue dots are the 15 dBsm (average) clustered targets, the cyan dots are the 5 dBsm (average) clustered targets, the black dots are the clutter discretes, and the red circles are the detections. All of the dots represent truth data. In addition, black lines indicate roadways. The detections show good agreement with the truth data. Almost all the detections can be associated with truth to within a range (15 m) and Doppler resolution cell (7.5 m/s). There does appear to be a westward bias in the location of the detections, but it is less than $1/10^{th}$ of a beamwidth.

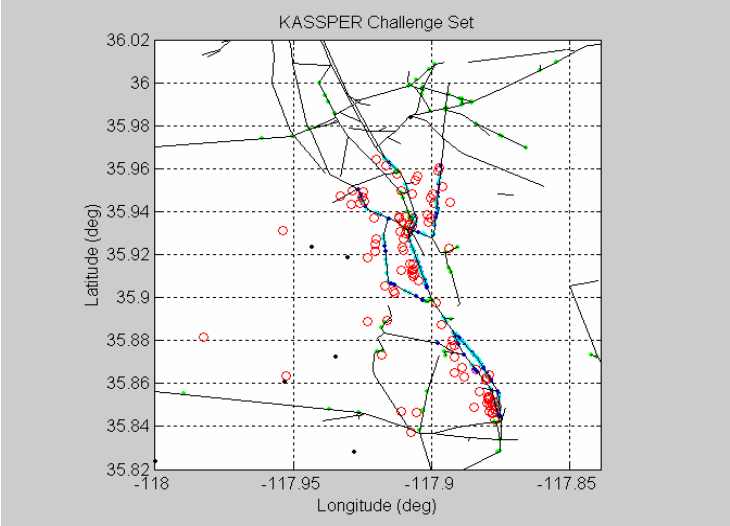

Figure 7.3.6-1. PPI of KASSPER Data Cube Detections

7.4 Sidelobe Target Sensitive Adaptive Detectors

Some detectors result in high false alarm rates because they count high sidelobes as detections. Several methods of avoiding the sidelobe problem are addressed in Section 3, the results of Section 3 pertain to sidelobe target rejection detectors. Here, results of two detectors, PAMF and Sub-CPI Smoothing, are given. Because the corresponding steering vectors are not low sidelobe weighted and the detectors are based upon the MF statistic, they are prone to high sidelobe induced false alarms. One method of evaluating the detectors, however, without having to suffer the sidelobe problem, at least in the Doppler domain is to count at most, one detection in a range cell. Although range cells with multiple targets then will be undercounted and the absolute performance of a method will be undervalued, the relative performance of methods will be accurate. Promising methods could be combined with HRSTAP (Section 3.3) as an effective means of dealing with sidelobes.

7.4.1 PAMF

Detection performance (counting at most one detection per range cell) of element space and beam space versions of PAMF are shown in **Figures 7.4.1-1** and **7.4.1-2**, respectively. The detectors are compared with the corresponding versions of MF, NPAMF, ParaGLRT, and JDL. Also, Multiple Pass Processing was employed for all detectors (except MF). Detector parameters (number of guard cells, Model Order, etc.) are similar to those of Section 7.3. Clearly, NPAMF and PAMF performed nearly identically. This result suggests that the principal advantage of introducing the normalized statistic to PAMF, thus forming NPAMF, is to alleviate the sidelobe target problem, as seen with the KASSPER Challenge data. Note, also, that ParaGLRT performed better than PAMF or NPAMF.

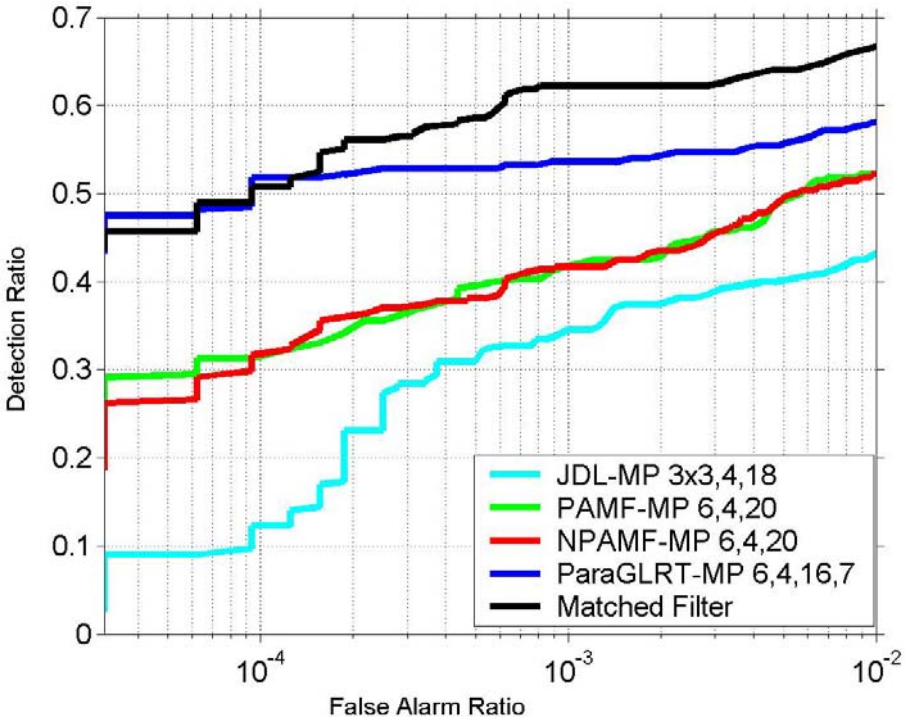

Figure 7.4.1-1 Element Space PAMF (KASSPER Data)

7.4.2 Sub-CPI Smoothing

Figures 7.4.2-1 and **7.4.2-2** show Sub-CPI Smoothing detection performance for element space and beam space implementations, respectively. As in Section 7.4.2, these results were obtained with at most one detection per range cell. Also, detector parameters (number of guard cells, sub-CPI size, etc.) were similar to those of Section 7.3. Results for both normalized and un-normalized versions of Sub-CPI Smoothing are included, as well as for SNIP, JDL, and MF. Multiple Pass Processing was employed for all detectors (except MF). Sub-CPI Smoothing and SNIP detectors performed comparably well and much better than JDL.

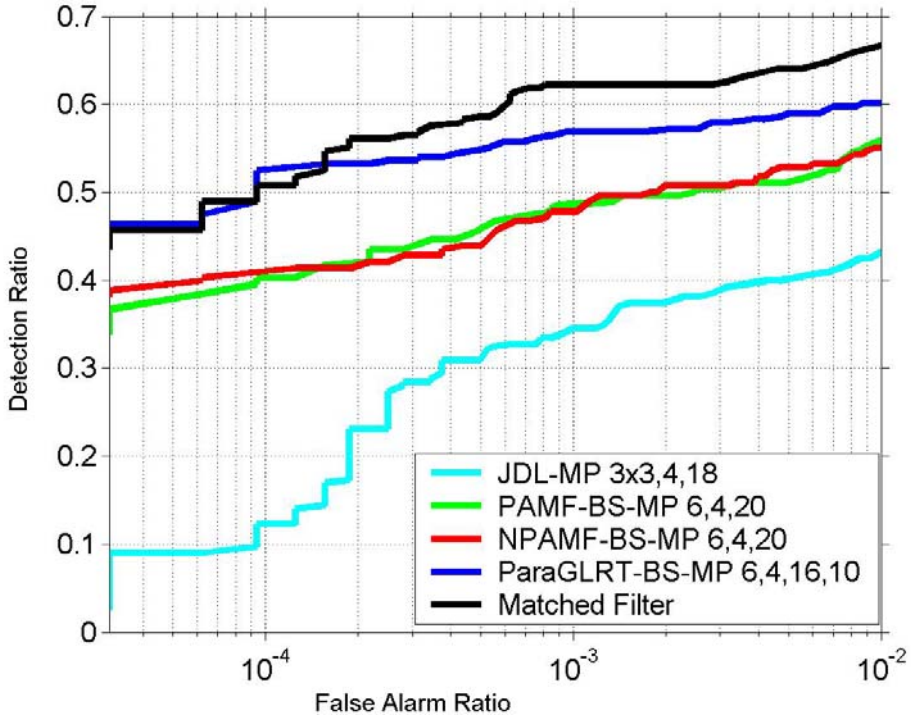

Figure 7.4.1-2 Beam Space PAMF (KASSPER Data)

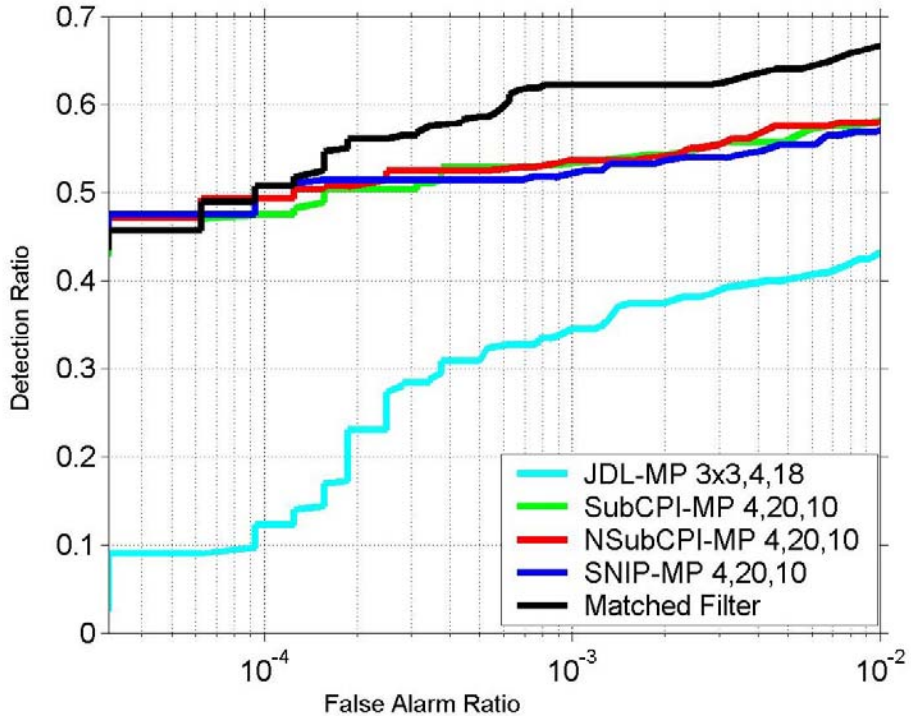

Figure 7.4.2-1 Element Space Sub-CPI Smoothing (KASSPER Data)

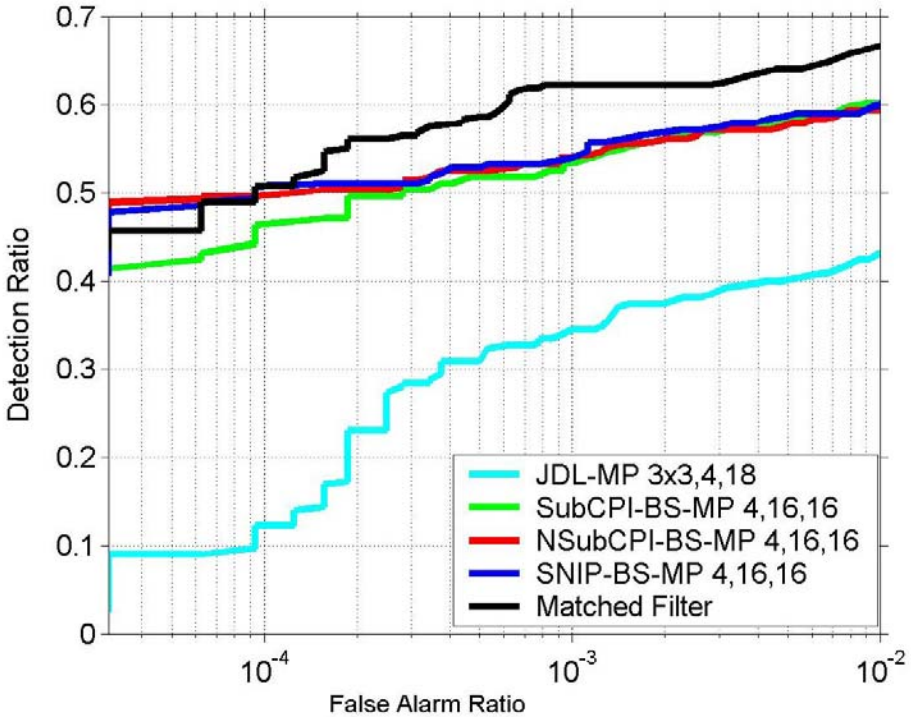

Figure 7.4.2-2 Beam Space Sub-CPI Smoothing (KASSPER Data)

7.5 Sample Support Dependence on Terrain

The NPAMF "One-Pass" and the NPAMF with Innovations Power Sorting ("Two-Pass") were applied to the KASSPER Challenge data with three sample support levels: 4, 10, and 20 secondary range cells symmetrically placed on each side of the test cell. Four guard cells were applied to separate the secondary data cells from the test cell in the sliding window processor. The 1000 range cell data window was divided into four sub-windows of 250 range cells each. For a given sub-window and a given support size, the number of targets detected commensurate with 50 total detections was determined. Multiple detections within one range cell were counted as a single detection, as discussed in Section 7.4. The results for the NPAMF (without power sorting) are shown in **Figure 7.5-1**. The smallest sample support size (K=4) yielded substantially better performance than those of the larger sizes for the first 250-cell sub-window and performance was similar for all sizes for the other sub-windows with the exception of sub-window 2 (Range Cells 251-500) where K=20 sample support performed best.

The results for the NPAMF with Innovations Power Sorting are shown in **Figure 7.5-2**. These results demonstrated a similar trend. Again, K= 4 sample support was superior for the first, third, and forth sub-windows, and K=20 was superior for the second sub-window. The impact of this observation was addressed in Section 6.1.

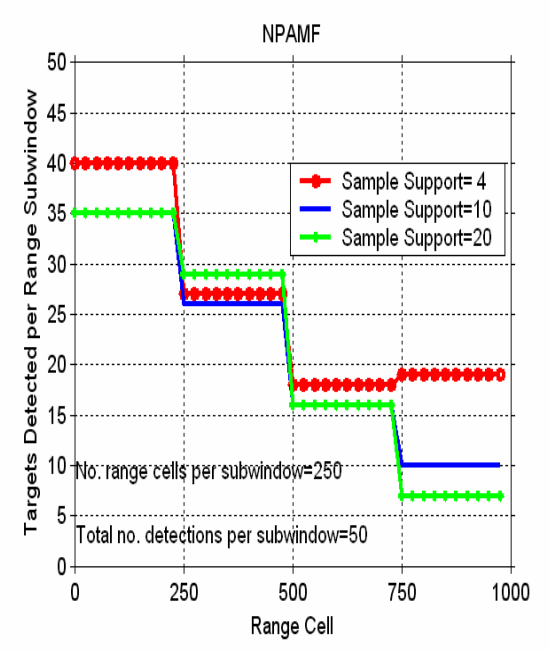

Figure 7.5-1. Sample Support Size Evaluation using NPAMF

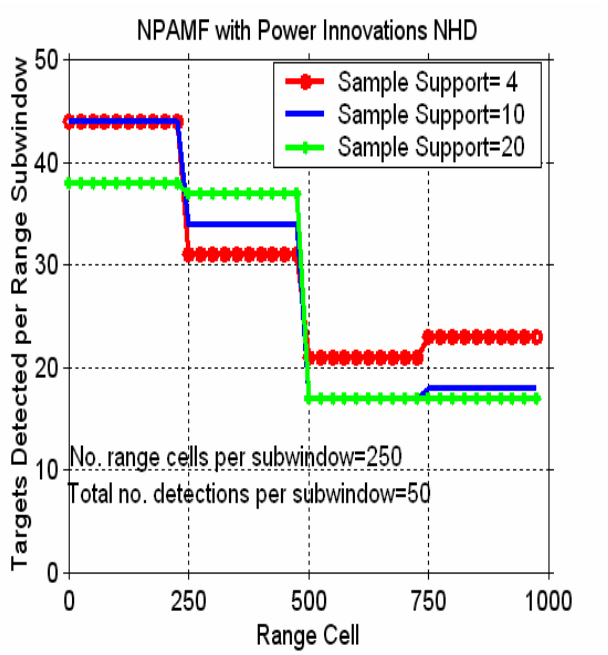

Figure 7.5-2. Sample Support Size Evaluation Using NPAMF with Innovations Power Sorting

8.0 CONCLUSIONS

The main objective of this work was to develop and evaluate Space-Time Adaptive Processing (STAP) detectors that are effective in secondary-data-limited scenarios. Such scenarios are characterized by rapidly varying terrain, localized strong scattering structures, a traffic rich environment, bistatic radar, nonlinear antenna array radar, and high PRF radar. Many new concepts were pursued. These included the Kalman Filter Parametric Adaptive Matched Filter (PAMF) (as well as the Normalized PAMF (NPAMF)), Generalized Likelihood Ratio Test PAMF (ParaGLRT), Sub-CPI Smoothing (normalized and un-normalized), Severely Nonhomogeneous Interference Processing (SNIP), Multiple Pass Processing (MPP), T-SNIP (optimizes secondary data selection of MPP), High Resolution STAP (promises to eliminate high sidelobe problems without consequence), and Discrete Scatterer Signal Injection (optimizes implementation of some knowledge based STAP approaches). Also, a STAP compatible Maximum Likelihood Angle Estimation method was introduced. Furthermore, beam space variants of the STAP detectors were explored as well as element space variants. The KASSPER Challenge simulated data cube was used exclusively for algorithm evaluation.

Tentative conclusions drawn from application to the KASSPER Challenge data cube included the following. NPAMF and the normalized form of Sub-CPI Smoothing exhibited good sidelobe target rejection. MPP with ParaGLRT performed about as well as the matched filter. MPP with Sub-CPI Smoothing and SNIP performed almost as well as the matched filter. All new methods performed substantially better than Joint Domain Localized (JDL). Beamspace variants slightly outperformed element space methods.

Much work remains. Many promising new concepts conceived as part of this effort could not be adequately explored. It is recommended that subsequent studies be initiated for this purpose. The studies should include execution against a myriad of data sets, including those involving bistatic geometries, conformal arrays, and Airborne Moving Target Indication (AMTI) scenarios, as well as, other Ground Moving Target Indication (GMTI) scenarios.

9.0 REFERENCES

1. M. Rangaswamy and J. Michels, "A Parametric Multichannel Detection Algorithm for Correlated Non-Gaussian Random Processes," *Proceeding of the 1997 IEEE National Radar Conference (NATRAD '97)*, Syracuse, NY, pp. 349-354, 13-15 May 1997.
2. J.H. Michels, et. al., *Multichannel Parametric Adaptive Matched Filter Receiver*; U.S. Patent No. 6,226,321, issued 1 May 2001.
3. J. Roman, et al., "Parametric Adaptive Matched Filter for Airborne Radar Applications," *IEEE Trans. Aerospace & Electronic. Systems*, Vol. 36, No. 2, pp. 677-691, April 2000.
4. J. S. Bergin and P. M. Techau, "High-Fidelity Site-Specific Radar Simulation: KASSPER '02 Workshop Datacube," Information Systems Laboratories, Inc. Report ISL-SCRD-TR-02-105, Vienna, VA, May 15, 2002.
5. S. Kraut and L. Scharf, "The CFAR Adaptive Subspace Detector is a Scale-Invariant GLRT," *IEEE Trans. on Sig. Proc.*, Vol. 47, No. 9, pp. 2538-2541, September 1999.
6. Kelly, E.J., "An Adaptive Detection Algorithm," *IEEE Trans. Aero. Elec. Sys.*, AES-22, pp. 115-127, March 1986.
7. E. Conte, et. al., "GLRT-Based Adaptive Detection Algorithms for Range-Spread Targets," *IEEE Trans. Sig. Proc.*, Vol. 49, No. 7, pp. 1336-1348, July 2001.
8. K. Gerlach and M. J. Steiner, "Adaptive Detection of Range Distributed Targets," *IEEE Trans. Sig. Proc.*, Vol. 47, No. 7, pp. 1844-1851, July 1999.
9. L. Cai and H. Wang, "On adaptive filtering with the CFAR feature and its performance sensitivity to non-Gaussian interference," *24th Conf. on Information Science and Systems*, Princeton University 21-23 March 1990.
10. W. Chen and I.S. Reed, "A new CFAR detection test for radar," *Digital Signal Processing*, 1, pp. 198-214, 1991.
11. F. C. Robey, D. R. Fuhrmann, E. J. Kelly, R. Nitzberg," A CFAR adaptive matched filter detector," *IEEE Trans. on Aerospace and Electronic Systems*, pp. 208-216, January 1992.
12. L.L. Scharf, *Statistical Signal Processing*, Addison-Wesley, 1991.
13. J.H. Michels, B. Himed, M. Rangaswamy, "Evaluation of the Normalized Parametric Adaptive Matched Filter STAP Test in Airborne Radar Clutter," *Proceedings of the IEEE 2000 International Radar Conference*, Arlington, VA, pp. 769-774, 7-12 May 2000.
14. E. J. Baranoski, "Constraint Optimization for Partially Adaptive Subspace STAP Algorithms," *Conference Record of the Thirty-Second Asilomar Conference on Signals, Systems and Computers*, Volume 2, pp. 1527-1531,1-4 November 1998.
15. R. L. Fante, et al., "Clutter Covariance Smoothing by Subaperture Averaging," IEEE Trans. Aero. and Elec. Sys., Vol. AES-30, No. 3, pp. 941-945, July 1994.
16. P. G. Li, "Optimal Detection-A New Adaptive Beamforming/Detection Algorithm," Syracuse Research Corporation Report TD 00-1473, 18 December 2000.
17. J. A. Hogbom, "Aperture synthesis with Non-Regular Distribution of Interferometer Baselines," *Astron. Astrophys. Suppl.*, Vol. 15, PP. 417-426, 1974.

18. J. Tsao and B. D. Steinberg, “Reduction of Sidelobe and Speckle Artifacts in Microwave Imaging: The Clean Technique,” *IEEE Trans. Antennas and Propagt.*, Vol. 36, No. 4, pp. 543-556, April 1988.

19. W. L. Melvin, “Nonhomogeneity Detection for Adaptive Signal Processing,” 1996 IASTED International Conference on Signal and Image Processing, Orlando, FL, 11-14 November 1996.

20. J. H. Michels, Discussion notes on “Innovations Inner Product,” January 2003. Also, “Data Efficient STAP Detection in a Dense Signal Airborne Radar Environment,” to be published.

21. T. K. Sarkar, et al., “A Deterministic Least Squares Approach to Adaptive Antennas,” Digital Signal Processing, Academic Press, Inc., AID DSP 0263, pp. 001-0010, 21 June 1996.

22. Salama. Y. and Senn, R., “Knowledge Base Applications To Adaptive Space-Time Processing”, Air Force Research Laboratory Final Technical Report, AFRL-SN-RS-TR-2001-146 Vol. I (of VI), July 1999.