REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
31-03-2004	Progress Report / Final Technical Report	28-Jan-2002 - 31-Mar-2004

4. TITLE AND SUBTITLE
Integrated Sensing and Processing in Missile Systems

5a. CONTRACT NUMBER
F49620-02-C-0019

5b. GRANT NUMBER
n/a

5c. PROGRAM ELEMENT NUMBER
n/a

6. AUTHOR(S)
Dr. Harry A. Schmitt

5d. PROJECT NUMBER
n/a

5e. TASK NUMBER
CLIN 0003

5f. WORK UNIT NUMBER
n/a

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Raytheon Systems Company
P.O. Box 11337
Tucson, AZ 85734

8. PERFORMING ORGANIZATION REPORT NUMBER
IS_P.A002-001

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Defense Advanced Projects Agency / DSO
Dr. Carey Schwartz / DARPA DSO
Professor Douglas Cochran / DARPA DSO
Program Manager: Dr. Jon Sjogren / AFOSR

10. SPONSOR/MONITOR'S ACRONYM(S)
See #9

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
n/a

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
None

14. ABSTRACT

Advances in sensor technologies, computation devices, and algorithms have created enormous opportunities for significant performance improvements on the modern battlefield. Unfortunately, as information requirements grow, conventional network processing techniques require ever-increasing bandwidth between sensors and processors, as well as potentially exponentially complex methods for extracting information from the data. To raise the quality of data and classification results, minimize computation, power consumption, and cost, future systems will require that the sensing and computation be jointly engineered. ISP is a philosophy/methodology that eliminates the traditional separation between physical and algorithmic design. By leveraging our experience with numerous sensing modalities, processing techniques, and data reduction networks, we will develop ISP into an extensible and widely applicable paradigm. The improvements we intend to demonstrate here are applicable in a general sense; however, this program focused on distributed sensor networks and missile seeker systems.

15. SUBJECT TERMS
Adaptive sensing modalities, data analysis and processing, data exploitation, and complex system optimization.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. Harry A. Schmitt
U	U	U	SAR	37	19b. TELEPHONE NUMBER *(Include area code)* 520-545-9578

Raytheon

Raytheon Systems Company

31 March 2004

Progress Report

CLIN No. 0003 – Final Technical Report

Final Progress Report for Period of Performance

28 January 2002 – 31 March 2004

Integrated Sensing and Processing in Missile Systems

Program Manager: Dr. Harry A. Schmitt

Principal Investigator: Dr. Harry A. Schmitt

Sponsored By:

Defense Advanced Projects Agency/DSO

Dr. Carey Schwartz/DARPA DSO

Professor Douglas Cochran/DARPA DSO

Program Manager: Dr. Jon Sjogren/AFOSR

Issued by AFOSR under Contract # F49620-02-C-0019

Prepared By:

Raytheon Systems Company

P.O. Box 11337

Tucson, AZ 85734

Integrated Sensing and Processing in Missile Systems
Contract F49620-02-C-0019
CLIN No. 0003
Final Technical Report
31 March 2004

Table of Contents

List of Figures

List of Tables

1.0. Management Summary

This report summarizes technical and programmatic accomplishments that have occurred during the contract period of performance 16 August 2002 through 31 March 2004. This is the final submittal for the referenced contract; there have been two prior interim progress reports submitted. The program has status has remained largely "on track". Raytheon often encounters significant difficulties finding mutually acceptable *Terms and Conditions* when subcontracting with universities and the negotiations with Rice University were unusually time-consuming. Raytheon has used the experienced gained from this rather frustrating experience to significantly improve the subcontracting process with universities. Unfortunately, the delays experienced in placing the contracts with our two subcontractors resulted in Raytheon having to request a no-cost extension to the contract. Rice University (Rice) and Fast Mathematical Algorithms and Hardware (FMAH) are such major components of the contract, that staffing at Raytheon was kept at a reduced level while negotiations were completed. These problems have been resolved and Raytheon now expects to complete the contract on time and on budget. FMAH has completed their subcontract on time and on budget. The Rice contract, which is Fixed Price, currently has some unexpected funds; they will finish on budget.

2.0. Personnel Associated/Supported:

2.1 Raytheon Missile Systems

Raytheon personnel that received significant funding support under the *Integrated Sensing and Processing for Missiles* program included:

Dr. Harry A. Schmitt (PI)
Mr. Donald Waagen (Co-PI)
Dr. Nitesh Shah
Mr. David Zaugg
Mr. Wesley Dwelly
Mr. Craig Savage

2.2 Fast Mathematical Algorithms and Hardware

FMAH personnel that received significant funding support under the *Integrated Sensing and Processing for Missiles* program included:

Professor Raphy Coifman
Dr. Paolo Barbano

2.3 Rice University

Rice personnel that received significant funding support under the *Integrated Sensing and Processing for Missiles* program included:

Professor Rich Baraniuk
Professor Rob Nowak

2.4 Significant Personnel Actions

There were no significant personnel actions or changes at Raytheon or FMAH during the current period of performance. Professor Rob Nowak has left Rice University for the University of Wisconsin-Madison; however, he remains active in the program.

3.0. Program Technical Summary

Advances in sensor technologies, computation devices, and algorithms have created enormous opportunities for significant performance improvements on the modern battlefield. Unfortunately, as information requirements grow, conventional network processing techniques require ever-increasing bandwidth between sensors and processors, as well as potentially exponentially complex methods for extracting information from the data. To raise the quality of data and classification results, minimize computation, power consumption, and cost, future systems will require that the sensing and computation be jointly engineered. ISP is a philosophy/methodology that eliminates the traditional separation between physical and algorithmic design. By leveraging our experience with numerous sensing modalities, processing techniques, and data reduction networks, we will develop ISP into an extensible and widely applicable paradigm. The improvements we intend to demonstrate here are applicable in a general sense; however, this program focused on distributed sensor networks and missile seeker systems.

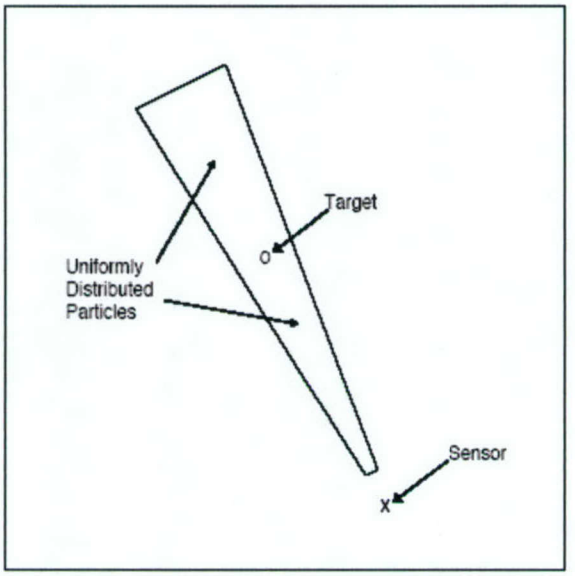
3.1. Missile Applications of Embedded Monte-Carlo Algorithms

Sequential Monte Carlo methods, or particle filters, have been investigated for the tracking of beam aspect targets, the tracking of targets obscured by altitude return, and the tracking of targets using a passive sensor. Particle filters are Bayesian tracking filters that are not constrained to the assumptions of Gaussian statistics and linearity. The strengths of particle filters were exploited to improve upon conventional tracking methods. These strengths can be exploited in all three of the previously mentioned applications, yielding performance improvements.

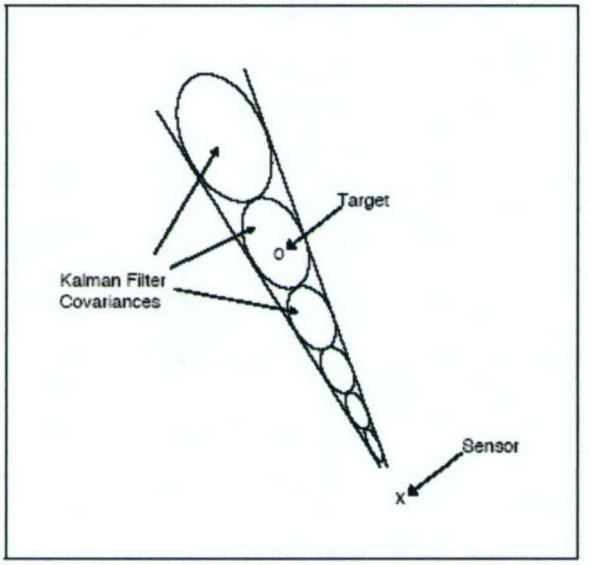
Bearings-only tracking is widely used in the defense arena. Its value can be exploited in systems using optical sensors and sonar, among others. Even though the limited information available to a passive sensor complicates the tracking problem, the advantages can be invaluable. Non-linearity and non-Gaussian prior statistics are among the complications of bearings-only tracking. Several filters have been used to overcome these obstacles, including multi-hypothesis extended Kalman filters (MHEKF), particle filters, and extended Kalman filters (EKF). A MHEKF can only approximate the prior distribution of a bearings-only tracking scenario and needs to be linearized. However, the likelihood distribution maintained for each MHEKF hypothesis demonstrates significant track memory and lends stability to the algorithm, potentially enhancing tracking convergence. Also, the MHEKF is insensitive to outliers. These characteristics may yield a smaller mean-squared error.

The initialization of a passive ranging tracking filter is critical. Due to the inherent non-linearity of the problem and the non-Gaussian prior distribution, a greater extent of the capabilities of particle filters can be exploited. Figure 1 illustrates the initialization

support associated with particle and MHEKF approaches.



(a)



(b)

Figure 1: Particle (a) and MHEKF (b) initialization

The EKF, while similar to the MHEKF, is more limited because of its necessary Gaussian approximation of the prior distribution. Because of its simplicity, it may not be as stable, but this simplicity may be a strength in terms of convergence speed. Indeed, each of these the filters have a set of advantages and disadvantages. We compared these approaches in different tracking scenarios to determine how their characteristics affect their tracking performance in a diversity of situations. The tracking scenarios included: tracking a stationary target, tracking a closing target, and tracking a crossing target. In the first two cases, the sensor's flight path is predetermined, but in the third, the sensor is

allowed to maneuver in an attempt to maximize tracking performance. For these scenarios, we compare and contrast the acquisition time and mean-squared tracking error performance characteristics of these three types of filters by means of Monte Carlo simulation. These scenarios are illustrated in Figure 2.

Each scenario includes a single target and a single tracker with an angle sensor. A Monte Carlo simulation is necessary because of the non-deterministic aspects of tracking, including process noise, measurement noise, and the random nature of the particle filter resampling algorithm.

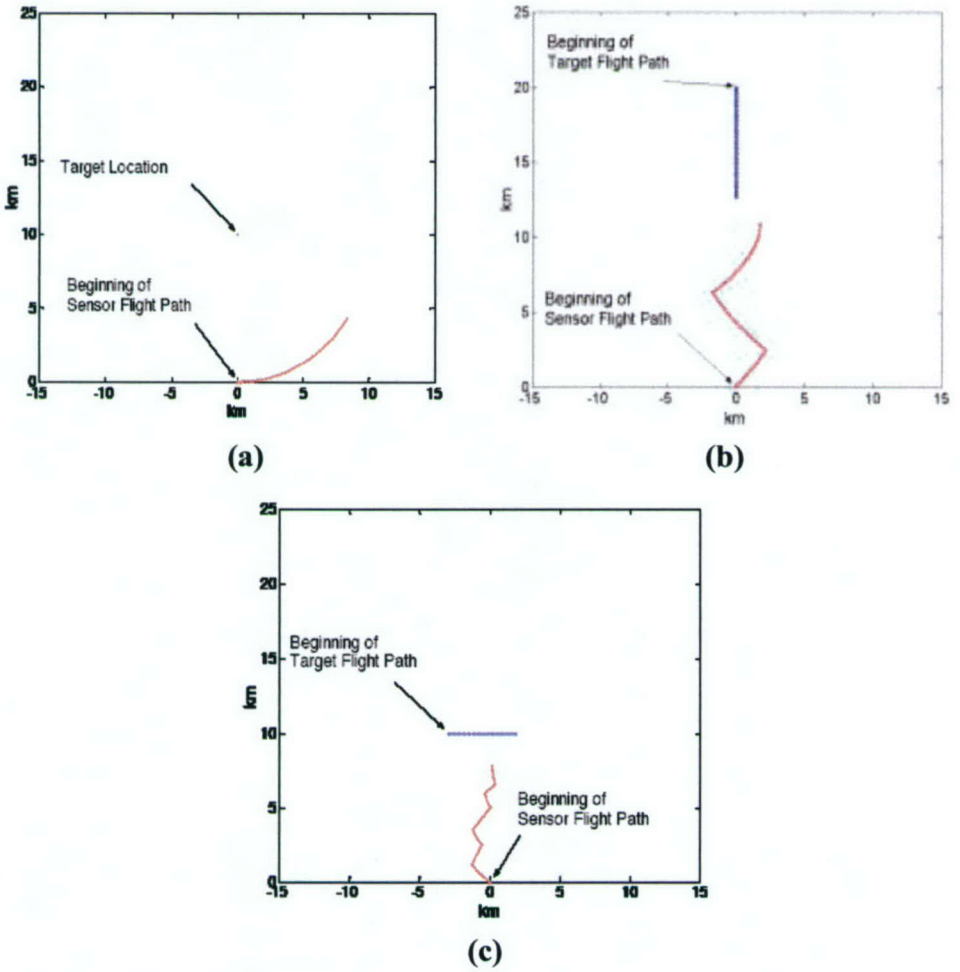
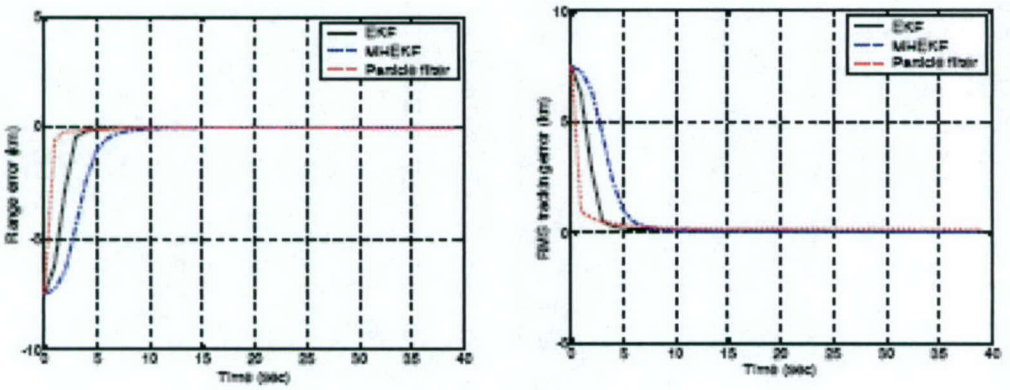


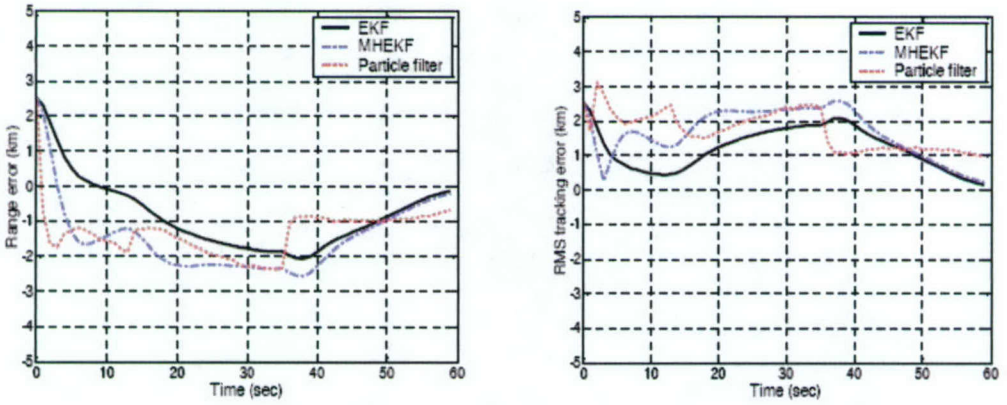
Figure 2: Stationary (a), Closing (b), and Crossing (c) target scenarios

We quantified the tracking ability of the three approaches for the scenarios described. The metrics for comparison are the range error versus time and the root mean squared (RMS) tracking error versus time. These results are the average of 100 Monte Carlo runs. The stationary, closing, and crossing target tracking results are respectively shown in Figures 3, 4, and 5.



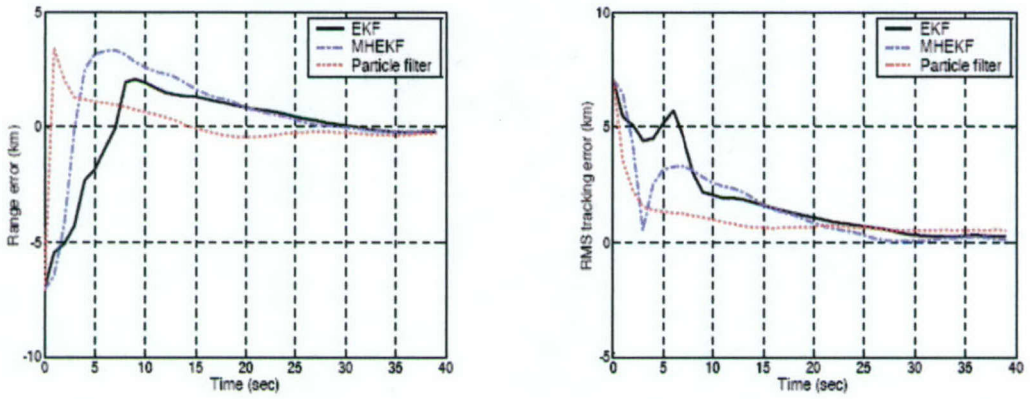
(a) **(b)**

Figure 3: Average range (a) and RMS (b) error for closing target scenario



(a) **(b)**

Figure 4: Average range (a) and RMS (b) error for closing target scenario



(a) **(b)**

Figure 5: Average range (a) and RMS (b) error for crossing target scenario

The stationary scenario is the easiest for all three filters, since the target is not moving. The filters can use less process noise, so the estimate converges tightly on the

target. The particle filter converges fastest in both range and RMS error, followed by the EKF, and then the MHEKF. However, the MHEKF has the smallest steady state error, followed by the EKF, and then the particle filter. The particle filter can be expected to be faster than the EKF and MHEKF because the particle filter is not linearized, and it does not have as much memory as the MHEKF. Because the particle filter is not linearized, it does not introduce linearization errors as it iterates. The likelihood distribution maintained for each MHEKF hypothesis demonstrates significant memory, but this penalizes the filter when it comes to convergence speed.

The closing target scenario is much more challenging than the stationary target case because the initial range is almost doubled, the target is moving, and the sensor is closing on the target. Increasing the range and closing on the target reduces the angular velocity of the sensor with respect to the target, making it less observable. Because the target is moving, the tracker must use more process noise. However, as the sensor gets closer to the target, the track starts to converge. Since this scenario requires the use of more process noise in the filter, the track cannot converge as tightly.

The results are quite different for this scenario. Considering range error, the particle filter converges the quickest at first, followed by the MHEKF, and then the EKF. They all seem to overshoot significantly, and finally converge at the end. As they converge at the end, the EKF is fastest, followed by the MHEKF, and finally the particle filter. The RMS error plot shows them converging towards the end. Again, the EKF is first, followed by the MHEKF, and finally the particle filter.

The crossing target scenario is also challenging, by considering the orientation of the uncertainty volume with respect to the target velocity. Since the sensor measures angle and has no a priori knowledge of target range, the uncertainty volume is long down range and narrow in the cross range direction. Therefore a crossing target could quickly leave the uncertainty volume causing a loss of track. For this reason, it is necessary to use significant process noise. Again in this scenario, different filters perform better in different time intervals. The range error and the RMS error show that the particle filter converges more quickly at first, but is not able to converge as tightly as the EKF or MHEKF. The ability of the sensor to adaptively maneuver improves tracking performance because the sensor is measuring with maximum ARI.

The scenarios tested the filters in a different way. The particle filter initially converges the fastest, but is then surpassed by the EKF and MHEKF in long term tracking error. Of the EKF and MHEKF, the MHEKF converges more quickly in the more difficult tracking scenarios, and maintains less steady-state error. These results indicate that the particle filter would be advantageous for track initialization, but that the EKF or MHEKF could be better for long-term tracking.

We are continuing investigating the extension of this technique to three other compelling applications: distributed sensor network, radar tracking in a range denied environment (jamming) and passive ranging for Ballistic Missile Defense.

3.2. Entropic Processing

Our goal is to develop techniques for characterizing (organizing, sorting, indexing, querying, *etc.*) the information content of data residing in high-dimensional spaces. In particular, we seek to enhance the process for *jointly* selecting features that improve class separability, rather than relying on classical margin-distribution-based feature analysis.

For characterizing high-dimensional joint data distributions, we are investigating a graph-theoretic method for estimating divergence between two sets of features. This method is based on recent work by Professor Alfred Hero *et al.*, wherein it is shown that a statistic determined from the length L of the minimal spanning tree (MST) of a graph formed from n d-dimensional feature vectors asymptotically converges to the α-Rényi Entropy, $H_\alpha(Z)$, of the feature set Z:

$$H_\alpha(Z) = \lim_{n\to\infty} \frac{1}{1-\alpha}\left[\ln\left(\frac{L(Z)|_{\{a_i,b_i\}}}{n^\alpha}\right)\right] - \left\{\lim_{n\to\infty} \frac{1}{1-\alpha}\left[\ln\left(\frac{L(Z_U)|_{\{a_i,b_i\}}}{n^\alpha}\right)\right] - \sum_{j=1}^{d} \ln w_j\right\}, \quad \alpha \in (0,1) \qquad (1)$$

Here, the data support is $\{a_i,b_i\}$ with widths $w_j=b_j-a_j$, and the second term on the RHS, known as the β parameter, contains an evaluation of an MST on data sampled from a Uniform Distribution. This direct method for estimating entropy can be applied to high-dimensional data, where classical methods typically fail. Hero *et al.* define the α-Jensen Entropy Difference, $H_\alpha(A,B) = H_\alpha(AUB) - 0.5(H_\alpha(A) + H_\alpha(B))$, as a statistic to evaluate the divergence between feature sets A and B. The individual α-Rényi Entropy terms are estimated using (1).

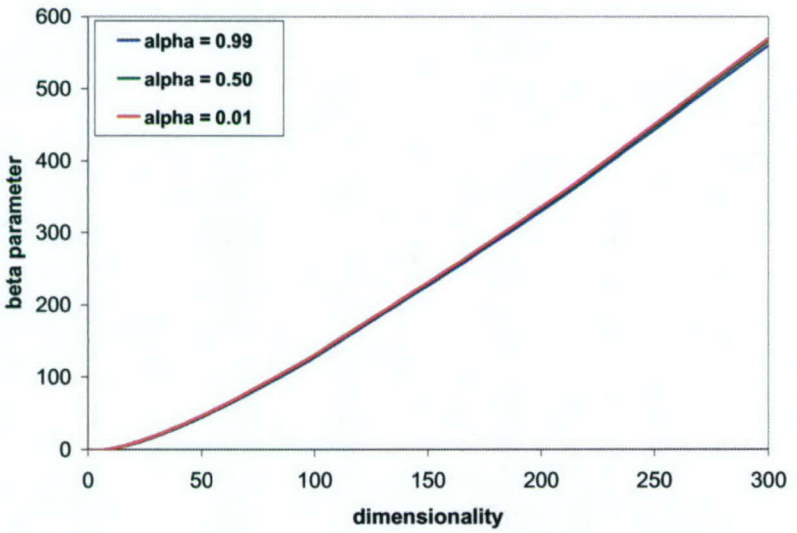


Figure 6: Dependence of β on d and α, for data support $[0,1]^d$.

Hero *et al.*'s prescription for determining α-Rényi Entropy contains a parameter, β, that depends on the data support; $0 < \alpha < 1$; and the dimensionality, d. Hero *et al.* do

not evaluate this parameter, choosing instead to calculate relative entropies among data sets sharing common values for the data support, α and d. We have extended this result by calculating values of the parameter β for $0 < \alpha < 1$ and $d < 300$, with fixed data support w_j=1, and we have shown that β is insensitive to the choice of α, and that β varies smoothly with d (Figure 6). We note that the parameter β is independent of the scale length.

For fixed data support, we are using our implementation to study joint feature distributions in a data set related to missile defense. In this data set, there are four classes: Class 1 is the target, and Classes 2 through 4 are different types of clutter. A total of 256 features are generated via wavelet-packet technique using the Kolmogorov-Smirnov test statistic for feature selection and feature ranking. Features selected by this method are highly correlated within class. We are investigating the use of ΔH_α as a technique to rank features via a *joint* density rather than the current marginal densities. In Figure 7, we show a 2-feature example. Feature 1 is taken as the first feature of the pair, and the second feature is varied over features 2 through 256. The features have already been individually ranked in terms of their class separation efficacy, *i.e.*, Feature 1 is the single best feature for separating classes and Feature 256 is the single worst feature for separating classes. We use n = 100 samples for each evaluation, and estimate ΔH_α pairwise over the classes.

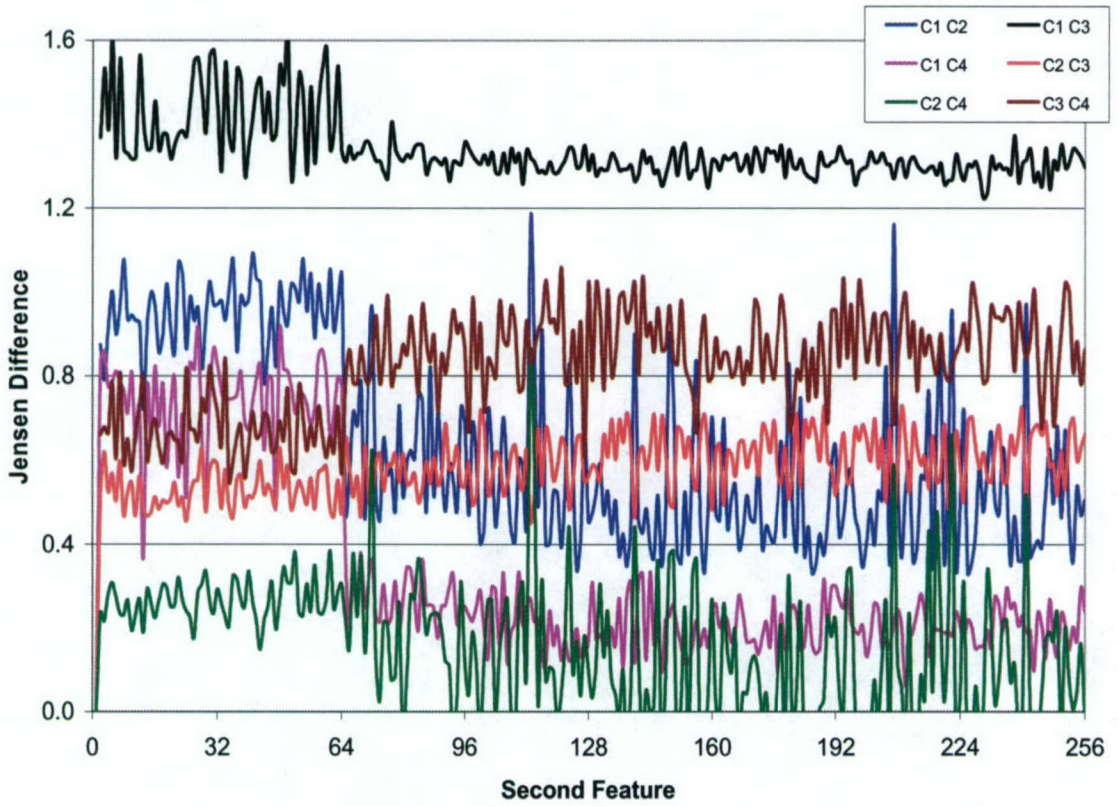


Figure 7: ΔH_α for 4-Class Data

The method of Hero *et al.* converges asymptotically. In practice, $n > 500$ samples are required to closely approach the asymptotic value. We have demonstrated that given only $n \sim 500$ samples, back-evolution by sub-sampling is a robust method for estimating the asymptotic behavior of the entropy estimate. However, in many applications, the number of available samples may be as low as $n \sim 100$. We are investigating an approach to improve entropy estimates in this sample-starved regime. In this approach, we estimate $\Delta H_\alpha(A,A)$ and $\Delta H_\alpha(B,B)$. In the asymptotic case, both of these quantities should converge to zero. In sample-starved situations, their deviation from zero should provide some information for better estimating $\Delta H_\alpha(A,B)$. Even with this possible improvement, poor asymptotic convergence remains a problem. Another problem is estimating the true data support $\{a_i,b_i\}$ given only a small data sample.

With these issues in mind, we have identified other techniques for working with high-dimensional data. These techniques include

- Friedman-Rafsky & extensions (multivariate two-sample test)
- Johnson-Lindenstrauss & extensions (low-dimensional subspace projection)
- ISOMAP (nonlinear dimensionality reduction)
- Locally Linear Embedding (nonlinear dimensionality reduction)
- Kernel-PCA (nonlinear dimensionality reduction)

Hero *et al.* have developed an approach that combines their MST-based work with ISOMAP. Several of these approaches were investigated and our results will be discussed.

We have also identified an MST-based approach for addressing the k-MST problem (determining the shortest path connecting any k nodes in a graph). In one variant, the MST is determined, and all edges not on the MST are removed from consideration. Then, using each node in turn as the starting point, use a greedy algorithm to add the next (k-1) closest nodes, and measure the k-length of the resulting edges. After all nodes have been used as a starting node, select the minimum value of the found k-lengths. In a second variant, using each node in turn as a starting point, use a greedy algorithm to develop the MST. For each nodes MST growth, keep track of the intermediate k-lengths as the next-closest nodes are added one by one. Finally, after all nodes have been used as a starting node, select the smallest k-length found for each value of k, producing the k-MST for $k = 2 \ldots n$. We have started discussion with Professor Hero on the usefulness of these two approaches to the k-MST.

3.3. Exploitation of Alternative Nonlinear Spaces:

The entropic approach of Hero *et. al.* provides a nonparametric approach for estimation of joint feature utility for classification problems. Traditional approaches for dimensionality reduction (*e.g.* Karhunen-Loeve, Principal Components, Independent Component Analysis, …) are linear in nature. Unfortunately, these latter transformations are suboptimal when the data resides in a nonlinear manifold of the original high dimensional space. Approaches, like ISOMAP or Kernel-PCA, attempt to estimate and extract the underlying nonlinear structure of the data.

We investigated exploiting nonlinear, high dimensional functional mappings of the feature/data for classification problems of interest. A tenet of kernel-based approaches to classification, including support vector machines (SVMs), is that data that is not linearly seperable in the original low-dimension feature space can often be linearly separated in a high dimensional space, if a mapping is defined by an appropriate nonlinear function. We investigated SVM's as an approach for separating low-dimensional non-separable data sets to a high (possibly infinite) dimensional alternative space. Formally, given a training set $S=((\mathbf{x}_1,y_1),\ldots,(\mathbf{x}_n,y_n))$, composed of n d-dimensional patterns $\mathbf{x}_i \in X$ and associated class labels $y_i \in \{-1,1\}$, a Support Vector Machine is a linear function of the form

$$f(\mathbf{x})=\sum_{i=1}^{m}\alpha_i y_i K(\mathbf{x},\mathbf{x}_i)+b. \tag{2}$$

The variables α_i are Lagrange multipliers, whose values are derived *via* maximizing

$$L(\alpha)=\sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j K(\mathbf{x}_i,\mathbf{x}_j) \tag{3}$$

subject to the constraints $\sum_{i=1}^{n}\alpha_i y_i=0,\ \alpha_i \geq 0 \quad \forall i=1,\ldots,n.$

A major hinderance to using SVMs is the need to determine the appropriate values for the kernel hyperparameters. The kernel parameter is frequently selected on an ad-hoc or experimental basis, in which an SVM is trained on various values of the parameter until "good enough" results are obtained. Indeed, these parameters (*e.g.* σ for the Gaussian kernel) directly effect the concept of distance in the alternative space, and have a critical performance impact. The appropriate selection of the kernel hyperparameters directly impacts the generalization and classification efficacy of the SVM. Figure 8 demonstrates the decision boundaries generated when the parameter value is too small (8a) and too large (8b).

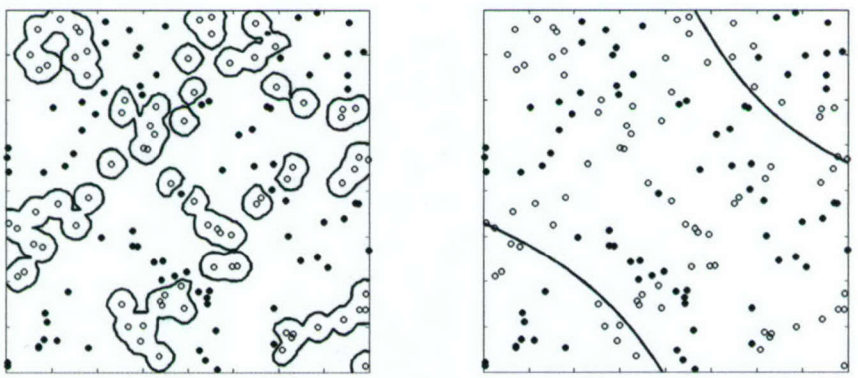


Figure 8: Decision boundaries produced via poor selection of kernel parameter (over-fit left; under-fit right)

Initial research developed an approach which can differentiate the conditions of over-fitting and under-fitting of SVM training for Gaussian kernels (Figure 3) thereby leading to a bounded range to search for an appropriate kernel parameter. A simple yet effective approach for identification of over and under-fitting training conditions was developed. This approach involved visualization of the distribution of margins values γ_k, defined by $\gamma_k = f(\mathbf{x}_k)$, which is literally the projection of the training data onto the hyperplane defined by the SVM in the alternative feature space. The probability density distribution of the margins can be estimated and visualized by simple statistical modeling techniques. We chose to use a Parzen kernel function as our density estimator. The class-conditional margin distributions and associated SVM decision boundaries for a simple checkerboard problem are shown in Figure 4. Note that the class-conditional densities when the SVM Gaussian σ value is too small are two delta functions (centered at ± 1), while the distributions overlap significantly when the σ value is too large.

By examining the class-conditional margin distributions associated with the training set mapped onto the vector defined by the SVM, an over-fit or under-fit condition is readily declared and a range for the kernel width parameter, σ, can be identified. Although the class data must be trained in this range to experimentally determine the desired value for σ, the initial search range can be significantly limited thereby decreasing the number of iterations of SVM training required. Moreover, this method provides insight into the separability of classes with the SVM.

Once a range for the search is established, we iterate the training in a fashion to minimize the number of support vectors. In practice, we set our iterations to some maximum level in order to limit the computational burden. Unfortunately, the SVM iterative training required by this approach is computationally expensive, and a more efficient automated approach for parameter selection was truly desired.

An alternative approach, developed by Cristianini et. al., defines the concept of kernel *alignment*, which effectively is a measure of the correlation of class labels and the Gram similarity matrix, and is formally defined as

$$A = \frac{\langle K, yy' \rangle_F}{\sqrt{\langle K, K \rangle_F \langle yy', yy' \rangle_F}} \quad (4)$$

In (4), K is the Gram or similarity matrix, y is the vector of class labels and F denotes the Frobenius inner product. This statistic was used by Christianini to estimate the utility of particular kernels (and their parameters) and thereby drive kernel adaptation. This simple, yet effective, statistic provides a measure for maximizing the within-class similarity (clustering) induced via the kernel parameters, while penalizing between-class similarity induced by the same kernel parameters.

An example of a Gram matrix computed for two-class checkerboard problem is given in Figure 10. The quadrants on the diagonal represent within class similarities while the anti-diagonal quadrants represent between-class similarities. Figure 11 displays

Gram (similarity) matrices for a two-class checkerboard problem using a Gaussian kernel function and four values of σ. For the first plot, the value for σ (0.1) is too small for this data set resulting in comparable within and between class similarity. For the second case ($\sigma = 0.4$), the plot shows high within class similarity while the between class similarity is much lower. For the last two plots of Figure 11, the value of σ is too large and is beginning to form large enough clusters that the all classes look "alike", resulting in similar within-class and between-class values. A key concept with this approach is that all class separability information is contained entirely in these similarity matrices rendering iterative training of the SVM unnecessary.

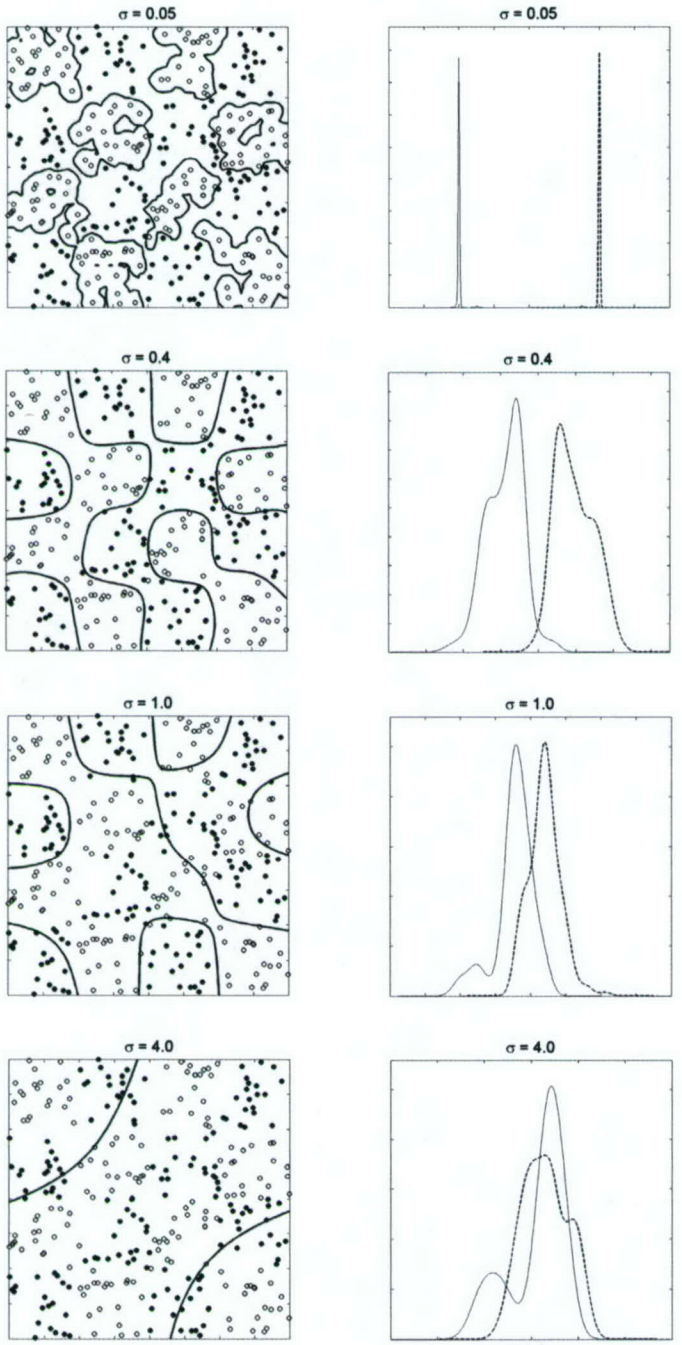


Figure 9: SVM decision boundaries and associated class-conditional margin distributions

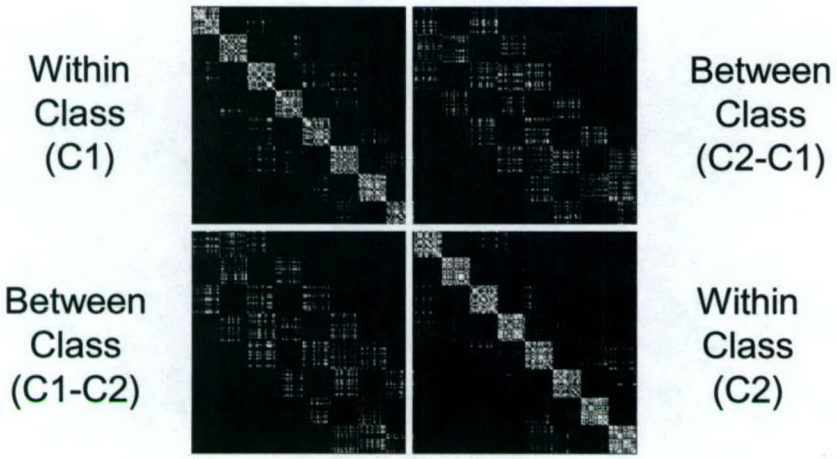


Figure 10: Sample similarity matrix for the two-class checkerboard problem. White indicates high similarity (~1) while black symbolizes low similarity (~0).

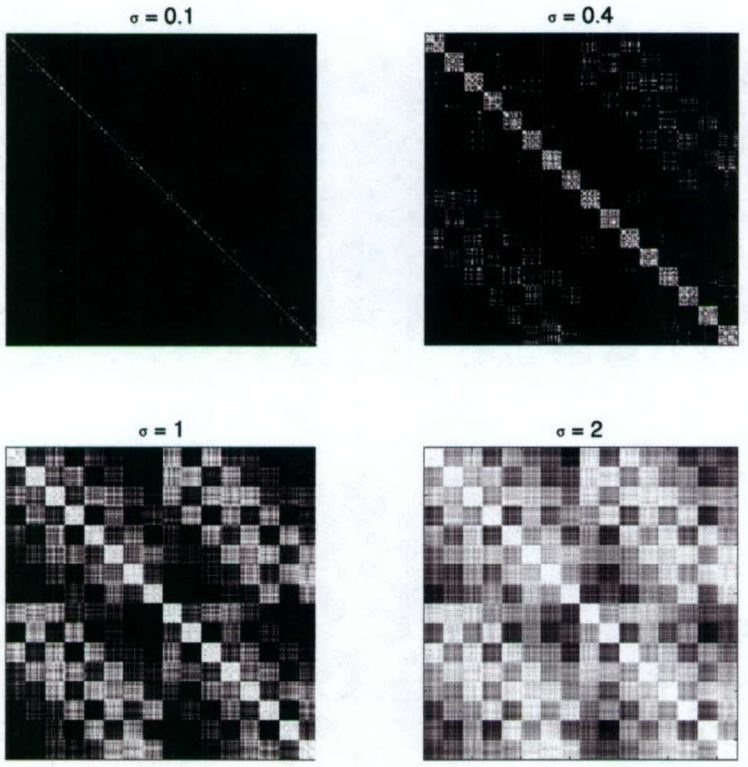


Figure 11: Similarity matrices for various kernel width selections

We noted that while this statistic is appropriate for true two class problems, in a multi-class (1 class vs. *m* classes) training environment the *alignment* statistic as defined does not differentiate between the desired within-class clustering of the class of interest

and the within-class clustering of the m alternative classes (the world). Therefore, in the multi-class case, the statistic can be biased when attempting to maximize the similarity of the world data vectors.

Our research amended the *alignment* approach, which we called *semi-alignment,* in a straightforward manner by applying a Frobenius inner product on a subset of the similarity matrix rather than on the entire matrix. By using a subset of the matrix, we remove the within class similarity of the world class from consideration. For multi-class cases (greater that two classes), the statistic will no longer encourage the collection of 'other' classes to look "alike". Although this may decrease the sample support for a true two-class case, it removes the induction of a false bias caused by the treatment of disparate classes of the world as a single class. For a Gaussian kernel with the sigma parameter, semi-alignment is defined as

$$f(\sigma) = -\langle K_\sigma, yy' \rangle_{F-lite} = -\sum_{y_i=1=y_j} K_\sigma(x_i, x_j) + \sum_{y_i=1\neq y_j} K_\sigma(x_i, x_j). \qquad (5)$$

In (5), we arbitrarily using the negative, and utilize a gradient descent approach to expedite the search for the minimum function value.

Figure 12 illustrates the subsets of the matrix used to calculate the *semi-alignment* test statistic for both two and four class cases. Class C1 is shown as the class of interest for both scenarios. In the two class case, C2 is the world while, in the four class case, classes C2-C4 are grouped together as the world. Any of the four classes could have been designated as the class of interest. As can be seen from the plots, the *semi-alignment* method uses only the class of interest and the between class data from the matrix.

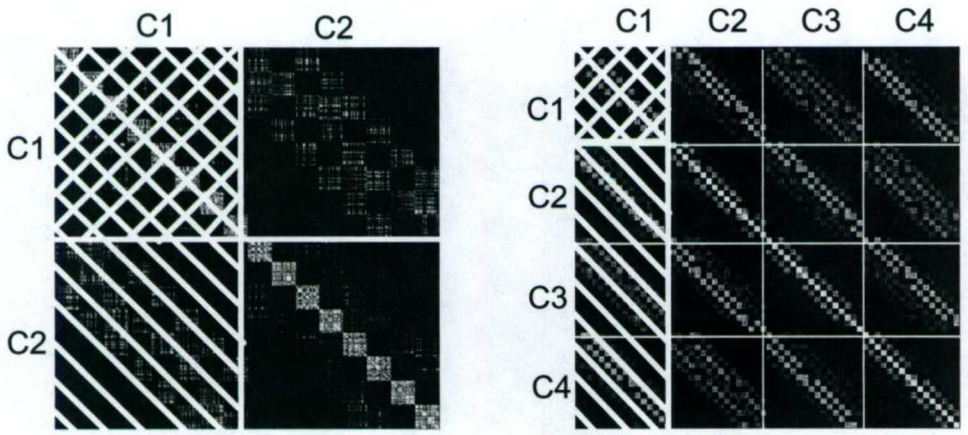


Figure 12: Subset of similarity matrix (two-class on left, four-class on right) used for semi-alignment Cross-hatched is the within-class, single hatched is the between-class

Our semi-alignment kernel parameter optimization approach is summarized Figure 13.

Step 0: Initialize σ. Starting with a "small" value of σ, use a bracketing technique to find appropriate minimum and maximum values for σ search.

Step 1: Calculate $f(\sigma)$ and $\frac{df}{d\sigma}$. Search over bracketed values found in step 0 for a minimum function value using a gradient descent approach.

Step 2: Iterate steps 0 – 1 for an estimate of the "optimal" value of σ for each class.

Figure 13: Kernel semi-alignment algorithm

3.3.1 *Kernel Parameter Optimization on Simulated Data*

To characterize the previously described approaches for kernel parameter optimization, we used linearly nonseparable two-class and four-class classification scenarios. Our first data set is a two-dimensional pattern space consisting of two classes distributed in a 4x4 cell checkerboard pattern (the checkerboard problem). The second simulated data set (the quadboard pattern), is a four class data set in an 8x8 cell pattern. Sample data sets for each of these are shown in Figure 14.

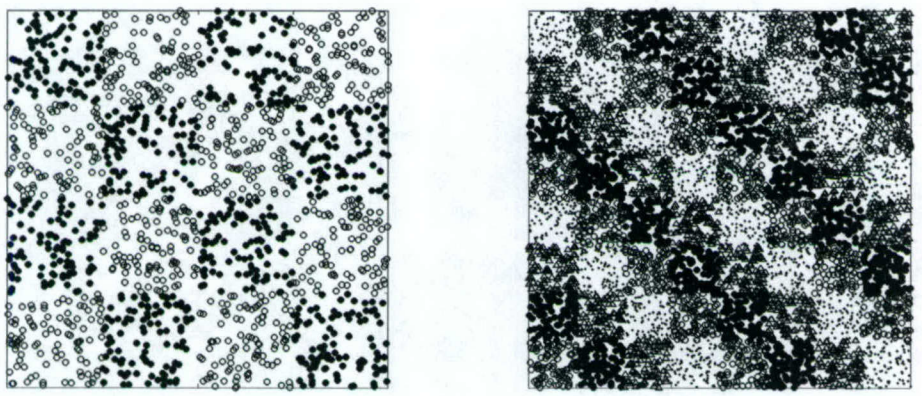


Figure 14: Sample checkerboard data set (left) and quadboard data set (right). 100 Samples are shown per cell.

Table 1 tabulates the association between the kernel parameter (σ), the mean and standard deviation of the number of support vectors in the associated SVM, and the corresponding mean and standard deviation for the classification efficacy for the two-class checkerboard problem. The iterative SVM training approach resulted in selecting σ

= 0.4, where the minimum number of support vectors was obtained for the algorithm. The classification efficacy was also near the maximum at this value. The last two rows of the table show the results for $\hat{\sigma}_{opt}$ with the *semi-alignment* and *alignment* techniques. The *semi-alignment* value of $\hat{\sigma}_{opt}$ as determined by our algorithm is in the range of $0.33 \leq \hat{\sigma}_{opt} \leq 0.39$ with a mean of 0.35 and standard deviation of 0.0099 for our random data set trials for the two class case as shown. The *alignment* algorithm resulted in essentially the same results. Note that the error rate for $\hat{\sigma}_{opt}$ for both alignment approaches is in the neighborhood of the optimal value obtained by our iterative SVM training. Remember, iterative SVM training was computationally several orders of magnitude more expensive than alignment.

Table 1: Relationship between kernel parameter σ, the SVM support vector (s.v.) statistics, and PCC statistics for two class checkerboard problem.

Kernel σ Value	# s.v. mean	# s.v. stdev	PCC mean	PCC stdev
0.05	316	2.1	81.6	0.03
0.10	279	5.1	90.4	0.87
0.20	147	5.2	91.3	0.83
0.30	101	5.1	91.6	0.89
0.40	95	5.7	92.3	0.86
0.50	99	6.1	92.4	0.89
0.60	113	6.4	91.9	0.95
0.70	135	6.7	90.8	1.10
1.0	228	6.6	79.4	1.58
2.0	298	3.4	56.0	2.14
4.0	302	3.0	51.2	0.03
semi-*alignment* $0.33 \leq \hat{\sigma}_{opt} \leq 0.39$	97	5.4	92.0	0.87
alignment $0.34 \leq \hat{\sigma}_{opt} \leq 0.39$	97	5.4	91.7	0.86

For the quadboard scenario, a boxplot of the results for the four-class quadboard problem are shown in Figure 15 and the details of the results shown in Table 2. The iterative training results in selection of 0.5 as the value for the kernel width, with an average 305 support vectors and the highest classification efficacy. *Semi-alignment* results for $\hat{\sigma}_{opt}$ obtained for our four class case were in the range $0.58 \leq \hat{\sigma}_{opt} \leq 0.67$ with a mean of 0.61, and a standard deviation of 0.015. For the *alignment* approach, the range for the optimal value of σ found over the trials varies more widely ($0.75 \leq \hat{\sigma}_{opt} \leq 1.37$).

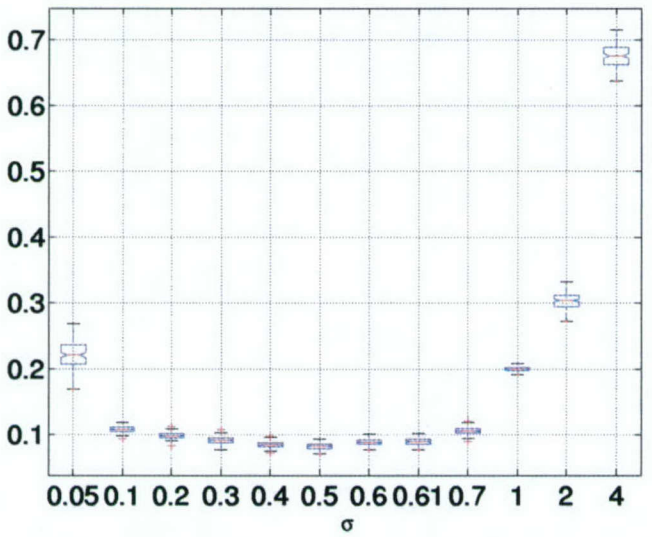


Figure 15: Classification error with semi-alignment $\sigma_{opt} \approx 0.61$ ($0.58 \leq \sigma_{opt} \leq 0.67$) and pre-specified values for σ for the four class quadboard data.

Additionally, alignment suffers a significant degradation in classification efficacy. Although the two class case results are identical for *alignment* and semi-*alignment*, we see the benefit for using the *semi-alignment* approach in a multi-class setting.

We next investigate the effect of sample support on the two techniques, *alignment* and semi-*alignment*, by considering results with reduced numbers of training samples. Table 3 shows performance results with 32, 80, 160 and 320 training samples for the checkerboard problem. The performance degrades as the sample support decreases, but the results with the two techniques are essentially identical.

The results of the quadboard (four class case) are shown in Table 4. Note that our *semi-alignment* approach consistently outperforms the *alignment* technique with this multiple class case until the sample support has decreased to 1 sample per cell. At this sample level, the *semi-alignment* approach results are similar to guessing (PCC = 26.5%), while the *alignment* approach performs slightly better (PCC = 36.3%).

Table 2: Relationship between kernel parameter σ, the SVM support vector (s.v.) statistics and PCC statistics for four class quadboard problem.

Kernel σ Value	# s.v. mean	# s.v. stdev	PCC mean	PCC stdev
0.05	1096	52.1	76.4	4.96
0.10	1134	9.1	88.8	0.47
0.20	655	7.7	89.8	0.46
0.30	418	6.0	90.5	0.50
0.40	324	6.1	91.1	0.48
0.50	305	6.2	91.3	0.49
0.60	313	6.4	90.8	0.51
0.70	350	7.7	89.2	0.56
1.0	438	9.8	79.8	0.44
2.0	644	2.5	69.4	1.28
4.0	960	142	50.1	9.85
semi-alignment $0.58 \leq \hat{\sigma}_{opt} \leq 0.67$	317	7.1	90.6	0.51
Alignment $0.75 \leq \hat{\sigma}_{opt} \leq 1.37$	511	24.4	83.5	1.54

Table 3: Results for variations in the training sample support for the two class case.

Number of training samples	Technique	# s.v. mean	# s.v. stdev	PCC mean	PCC stdev
16 (1/cell)	*semi-alignment*	16	0	51.0	0.04
16 (1/cell)	*alignment*	16	0	50.7	0.03
32 (2/cell)	*semi-alignment*	31.7	0.50	72.3	0.08
32 (2/cell)	*alignment*	31.7	0.55	71.2	0.08
80 (5/cell)	*semi-alignment*	54.0	3.6	83.2	1.8
80 (5/cell)	*alignment*	53.9	3.6	83.1	1.8
160 (10/cell)	*semi-alignment*	69.1	3.9	87.9	1.3
160 (10/cell)	*alignment*	69.0	3.9	87.9	1.3
320 (20/cell)	*semi-alignment*	97	5.4	92.0	0.87
320 (20/cell)	*alignment*	97	5.4	91.7	0.86

Table 4: Results for variations in the training sample support for the four class case

Number of training samples	Technique	# s.v. mean	# s.v. stdev	PCC mean	PCC stdev
64 (1/cell)	*semi-alignment*	64	0	26.5	0.04
64 (1/cell)	*alignment*	32	0.2	36.3	0.05
128 (2/cell)	*semi-alignment*	102	5.6	73.8	0.02
128 (2/cell)	*alignment*	43	5.3	69.5	0.05
320 (5/cell)	*semi-alignment*	130	3.6	81.9	0.01
320 (5/cell)	*alignment*	67	4.7	76.7	0.03
640 (10/cell)	*semi-alignment*	197	5.2	86.8	0.01
640 (10/cell)	*alignment*	116	7.3	80.6	0.01
1280 (20/cell)	*semi-alignment*	317	7.1	90.6	0.51
1280 (20/cell)	*alignment*	511	24.4	83.5	1.54

3.3.2 Optimization of SVM kernel parameters on measured HRR data

So what is the benefit of these techniques with a measured data set? To answer that question, we applied these techniques to a three-class (see Figure 16) measured High Resolution Radar (HRR) data set and investigate selection of the Gaussian kernel width parameter with *semi-alignment* and *alignment*. The data set used for testing the algorithms consisted of 1417 inverse synthetic aperture (ISAR) images, of which 360 samples, 120 per class, were selected for training leaving for 1057 for testing. The original ISAR images were converted to real-beam range profiles by means of frequency domain processing for algorithm performance testing. The complex target signatures were converted to real-valued magnitude profiles. The breakdown by class for the test set is shown in Table 5. Detection of the target was pre-supposed, as this study was geared towards evaluation of algorithmic performance.

Figure 16: Three class measured data set targets

Table 5: Number of test samples per class for measured data set

Class	Number of samples
BTR	352
M2	349
ZIL	356

Target pose information is not considered either in the training or testing phases, resulting in classification representative of class differences across vehicle angular aspects. Aspect angle variations present an important challenge in classification with radar signal signatures since they exhibit a high degree of aspect angle dependence. Desired classification schemes include those that exhibit little dependence on aspect angle with respect to the separation of classes. With no knowledge of aspect angle, we exploit signal characteristics that are common at all aspects thus forcing pose independence. Examples of training and test range profiles for a target at the same pose are shown in Figure 17.

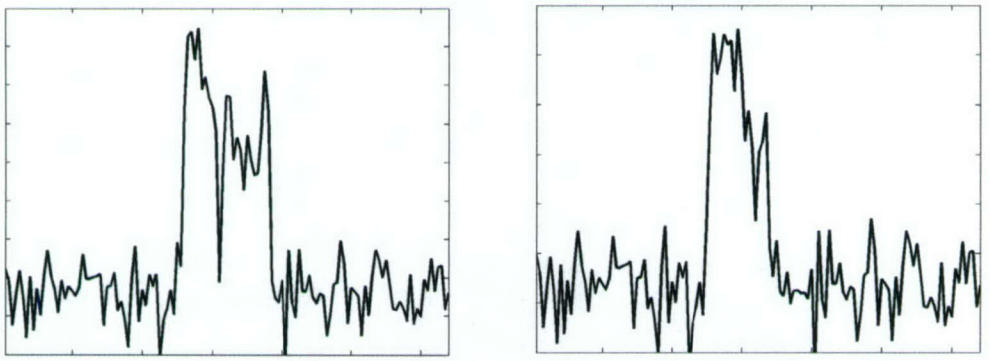


Figure 17: Training (left) and test (right) signatures for a target

We utilized wavelet based features selected by forming empirical distribution functions (EDFs) and implementing wavelet base selection via the Kolmogorov-Smirnov (KS) test statistic. In an alternative approach to wavelet base selection, Saito *et.al.* use an ASH estimate of the class probability density and implement the base selection via the Kullback-Leibler (KL) test statistic. We modified Saito *et. al.*'s KL approach to allow more flexibility in the score normalization process in a multiclass setting. To do this we form a score matrix of class pair-wise scores and select an overall node score based on a selection of a norm technique. We use the minimax and sup norms for this data set.

During previous empirical SVM training experiments with this data, we had selected an 'optimal' value (determined by empirically trying a range of values) for the SVM kernel parameter $\sigma = 0.5$. We compared our previous classification results using value with both classification results derived via *alignment* and *semi-alignment* selection processes.

The results of this comparison are shown in Figures 18–21. These figures show a comparison of the classification efficacy of the data set using the KS or KL wavelet feature selection techniques and *alignment* and *semi-alignment* optimized values of for the SVM kernel parameter (σ), as well as our previously best baseline value of $\sigma = 0.5$.

Figures 18 and 19 show the results for the KS minimax and sup scoring approaches. We see significant improvement in the performance by optimizing σ with both *semi-alignment* and *alignment* for both scoring methods. The results for the two Gram-matrix optimization techniques diverge slightly with the *semi-alignment* exhibiting superior performance at low dimensionality and *alignment* at higher dimensionality. With the optimized SVM parameters, we continue to see the inherent data dependence that must be considered when selecting the scoring approach. The performance of this data set with the KL technique results in several interesting conclusions. Figures 20 and 21 show these results. Here we see the same trends that were found with the KS approach. KL minimax and KL sup result in marked improvement over the baseline value of σ with the KL techniques performing better than KS at higher dimensions and KS performing better at lower dimensions.

Both the *semi-alignment* and *alignment* approaches provide better estimates for the value of σ as compared with training over a pre-specified range of values approach. Recall that with both of these techniques, an optimal σ value is determined for each class while the baseline approach selects an overall value for σ (all classes are restricted to a single common value). The *semi-alignment* and *alignment* approaches performed similarly with *semi-alignment* classification efficacy higher at lower dimensionality and *alignment* better at higher dimensionality. We note that this is a three class case; a trial with more classes most likely would begin to demonstrate differences in *semi-alignment* and *alignment* due to the inherent grouping of dissimilar classes into a world class by *alignment*. Indeed, we saw a false induction of similarity for the *alignment* approach with the quadboard data case.

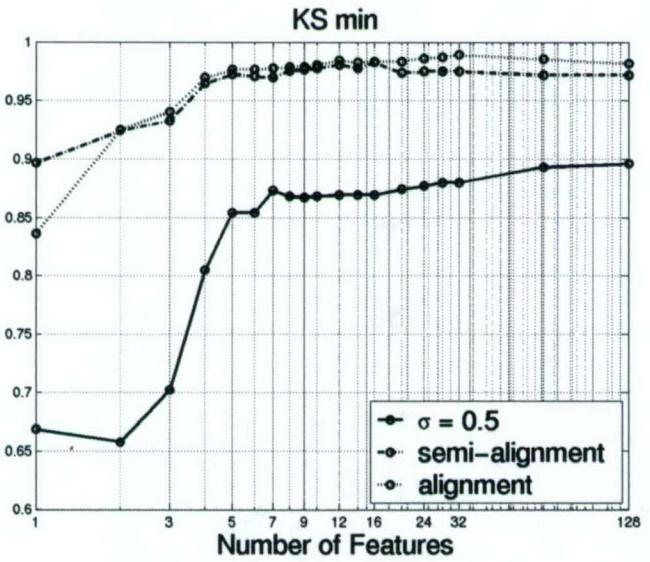


Figure 18: Comparison of classification results of σ selection techniques of three class measured data set with KS minimax scoring technique.

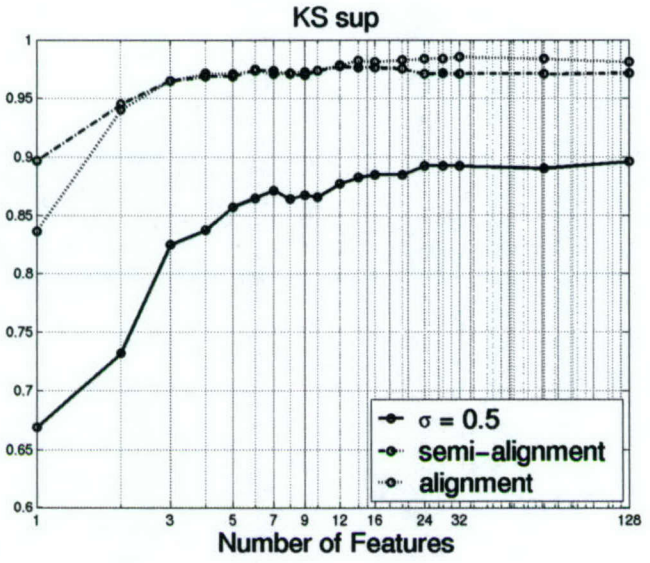


Figure 19: Comparison of classification results of σ selection techniques for three class measured data set with KS sup scoring technique.

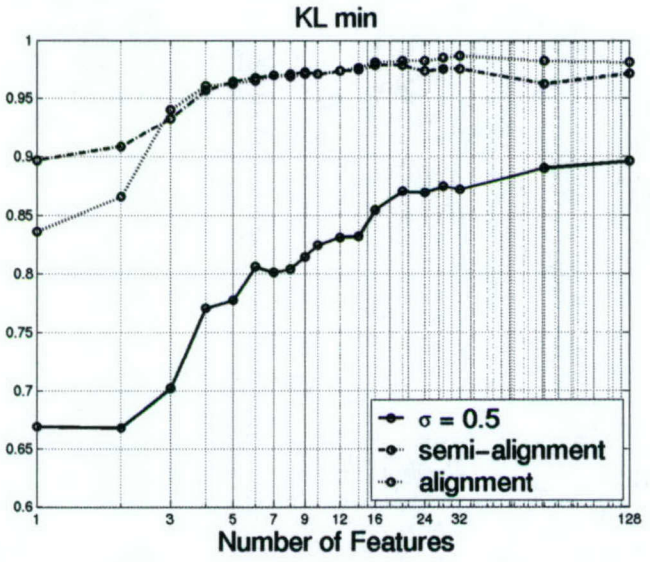


Figure 20: Comparison of classification results of σ selection techniques for three class measured data set with KL minimax scoring technique.

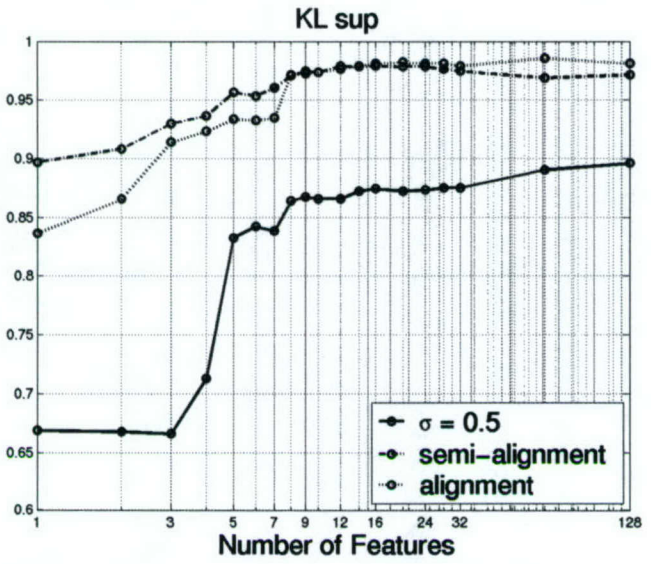


Figure 21: Comparison of classification results of σ selection techniques for three class measured data set with KL sup scoring technique.

These results indicate that *alignment* or *semi-alignment* techniques provide efficacious and efficient to estimate kernel parameters on simulated and real data sets. Our *semi-alignment* approach expected to be generally preferable to *alignment* for multi-class data, as simulations demonstrated better classification efficacy.

3.4 Sensor Scheduling Against Swarms/TBM

As a preliminary study in sensor scheduling, we examined scheduling algorithms against a number of targets converging on a central sensor. Such a scenario may represent a number of engagements, including a ballistic missile attack or a number of small, explosive-laden boats converging on an aircraft carrier.

3.4.1 Model Assumptions

The swarm scenario is modeled under the following assumptions.

- The system has a priori knowledge as to the number of objects, and their approximate position and velocity vectors.
- All objects are moving in a straight line towards the sensor.
- The objects do not accelerate.
- Each object must be tracked before it is engaged by a weapon.
- Objects are "friendly" with a certain probability. Note that an object need not be identified before weapon deployment; however, if a "friend" is engaged, a penalty is paid.
- Sensor usage is divided into a series of "dwells" during which it may attempt to sense N objects.
- The scenario is completely observed. That is, while it is not guaranteed that a sensed object will be identified, its state of being identified or not is known. That is, there are no false alarms. Objects can have the following states:

1. X: The null state. This represents objects that have not been examined by the sensor.
2. D: Detected. For objects that have been detected, but are not under track or identified.
3. T & I: Tracked and identified. If not hostile, this is the terminal case.
4. T & ~I: Tracked, not identified.
5. ~T & I: Not tracked, but identified.
6. K: Killed. For objects after successful weapon deployment.

- State transitions are handled via probability estimates. Each object, not the sensor, has a set of probabilities centered about some nominal value. In this manner, there is inhomogeneity between objects. If an object is not viewed by a sensor, its state is left unchanged. Otherwise, state transitions are modeled by the following mode-transition matrix.

Table 6: Mode-transition matrix for Sensor Scheduling against SWARM/TBM

Old \ New	X	D	T&~I	~T&I	T&I	K
X	1-Pd	Pd(1-Qi)	0	PdQi	0	0
D	0	(1-Pd)+Pd(1-Pt)(1-Qi)	PdPt(1-Qi)	PdQi(1-Pt)	PdPtQi	0
T&~I	0	0	1-PdPi (-Pk)	0	PdPi	(Pk)
~T&I	0	0	0	1-PdPt	PdPt	0
T&I	0	0	0	0	1-Pk	Pk
K	0	0	0	0	0	1

Note that for the transition T&~I-> K, in most cases, no weapon is deployed; hence, the mode transition is conditioned upon weapon deployment. Weapon deployment criteria are covered in more detail later. As for notation, it is assumed that Q<= P, so that identification is more likely if an object is already under track.

3.4.2 Mathematical Formulation

Under the definitions and assumptions in section 3.4.1, we consider mathematical approaches to scheduling solutions. Multi-Armed Bandits (MAB) are of particular interest, due to the congruence of the assumptions made in our model and the assumptions in the hypothesis of the MAB. Specifically,

1. Each bandit is governed by its own unknown state transition matrix.
2. The reward of examining an object is tied only to its current state.
3. States do not change when not examined.

A MAB is defined by trying to maximize discounted rewards over some time horizon. For a policy, u(t), denoting which object(s) are examined at time t, the goal is:

$$\arg\max_{u} \sum_{t=0}^{T} \gamma^{t} \sum_{i} c_i u_i(t) \sum_{x,y \in X} p(x \to y) R(x \to y)$$

where rewards are calculated according to the probability of state transition from a state x to state y. The first summation occurs over a certain time horizon, T, according to a discount factor, (gamma). The second considers which object(s) to consider, while the final is the expected reward for transitions from state x to state y. The constants, ci, are inversely proportional to the time it would take an object to reach the central sensor. In our tests, ci = v/d, which, scaled to a dwell time, is 1/ngo, or the number of expected dwells to impact.

With those similarities in mind, there are also notable differences between a "classical" MAB and the current problem. These include:

1. *Exploitation premium*: In the MAB case, the goal is to find a winning "bandit" and play it as often as possible. Conversely, in our case, we want to move objects to a terminal state and then never revisit them.
2. *Probability estimation*: Many MAB solution algorithms attempt to estimate the state transition probabilities for estimating future rewards. Conversely, we do not care about characterizing the state transition, so long as we can migrate the state into a terminal state.

Fortunately, however, these two points may be overcome by constructing a suitable reward function. Corresponding to the state transition matrix from above, we also implement a reward for moving to each state.

Table 7: Reward matrix corresponding to state transition matrix

State	Reward
X	0
D	Rd
T&~I	Rt
~T&I	Ri
T&I	Rc
K	RhPh – Cf(1-Ph)

Where Rd < (Rt, Ri) < Rc. In the kill column, Rh is the reward for killing a hostile target, Ph is the probability that the object is hostile (which, if the object is identified, is either 0 or 1; otherwise, it is an a priori estimate), and Cf and Pf are the corresponding cost and probability for friendly targets. Finally, if an object approaches the central sensor to some lethal distance, a huge cost Cb >> {Cf, R.} is incurred. Note that the rewards for moving to either T&~I or ~T&I are not strictly comparable in the model. This is somewhat offset by the lower probability of moving from D->(~T&I), so that equal rewards for the two will naturally give preference to attempting to track before identifying an object.

While the sensor progresses an object from state-to-state, weapon deployment is independent of sensor function. For the simulation, one weapon may be deployed against

any object, or no object. Because of the completely observed nature of the simulation, a weapon will be deployed against an object when it has been tracked and identified, or if it is within a dangerous range of the base. The completely observed nature is reflected in the binary nature of Ph and Pf.

3.4.3 Simulations

Armed with selected values of all P, Q, R, and C values, as well as selected discount rate, time horizon, and state information, scenes are randomly generated. Each scene is a random selection of objects according to some number of objects, range and velocity profiles, and all relevant P and Q values on each object, according to some distribution. The primary purpose of the simulation is to evaluate multiple scheduling algorithms, not evaluate system performance. Through basic simulation, the parameters of the scene (e.g. range and velocity profile) are larger drivers of base survivability than which algorithm is used. Analysis from the simulation is limited to lessons learned.

Time Horizon

The first realization was specific cases under which having a multi-epoch cost function was more advantageous than a "greedy" algorithm. Consider a case where two objects in the null state (state X), are 5 dwells from the base, and the sensor can only look at one object at a time. Even under a benign case of all probabilities equal to one, one can reach the following quandary:

Object\Dwell to go	5	4	3	2	1
1	X	D	T	?	
2	X	X	X	?	

With two dwells to go, for some set of rewards, a myopic cost function may choose to examine target 1, looking to gain the reward for identification. However, doing so neglects object 2, bringing it to only state D on the last dwell, resulting in the destruction of the base.

Looking multiple steps ahead is not strictly necessary. One could tweak the reward values, or change the constants to be inversely proportional to (ngo – 2), rather than ngo. However, tweaking the rewards will be an ongoing problem, whereas using (ngo-2) works for this case, but begins to fail for, say, three objects with ngo = 7. However, looking ahead for three dwells, the algorithm sees a greatly increasing cost associated with looking at object 2 from the base destruction. Hence, at this stage, a non-greedy algorithm deploys a weapon at the first target, and tracks the second target before lethality. That said, it is generally better for this situation to never arise in the first place, where all objects have been killed before they get this close to the base.

Multiple Examinations Within a Dwell

The policy, u, need not represent a single target of interest. Instead, the policy may be a set of objects to be examined within a dwell. While allowing more objects within a dwell to be examined increases the computational complexity, one can sort values according to their ngo values, and only consider targets which have smaller ngo values in a certain state. That is, if there are 20 objects in state X, because the probability and rewards for state transition are all identical, their respective costs may be sorted by their ngo values. A similar preordering may be done across all states, and can be further ordered according to the state into which they would transition.

Multiple Sensor Types

In addition to incorporating multiple targets within a policy, if multiple sensor types are present, the sensor utilized during a dwell can also be incorporated into the policy. Following the work of Krishnamurthy, we consider two sensors, one which has good tracking performance, while the other has good identification performance. Under such a structure, under nominal conditions, the tracking sensor tracks N objects, at which time the identification sensor tries to identify them all. Such scenarios vary under some objects getting "close" to the base, but, otherwise, the progression is fairly predictable.

3.5. Waveform Design and Scheduling

3.5.1 FMAH Spectral Analysis Codes

We propose to investigate advanced waveform coding to suppress clutter specifically for those situations where standard statistical techniques become unstable. This section discusses a novel class of multiscale waveforms that possess a number of properties that are applicable to the ISP problem. By using a completely new approach to the classical theory of Walsh functions, we have developed a series of mathematical algorithms for the design of coding sequences – Spectral Analysis Codes (SAC) – that can be utilized specifically to detect and resolve spectral characteristics of target returns buried in clutter and noise. SAC design techniques can be employed both as (1) signal processing tools at the receiver as well as (2) in the generation of modulating sequences for pulse-coded waveforms.

In order to separate the target from clutter return we are capable of producing a family of SAC codes with spectral characteristics which can be customized to respond "flatly" to the kind of clutter return determined by the application. Once the SAC family is determined it can be used as a frequency analysis filter: we identify the target by tracing any fluctuations from the statistically expected value in the Power Spectral Density picture drawn by using our SAC family as a basis for the frequency transform. It is important to note that this "clutter-customized" power-spectrum estimation can be carried at several time-scales simultaneously. Furthermore, the approach suggested above is based on algorithms whose complexity does not exceed the one of classical Fourier-based peak-position estimation methods and has the additional advantage of being a more flexible scheme to adapt to different clutter/target characteristics.

SAC families of coding sequences can be modeled to suitably comply with a variety of time-frequency analysis requirements. Both their Frequency response and Power Spectral Density can be designed rather easily to be close to AWGN or highly coherent, depending on the requirements imposed by the application. In addition, the theoretical approach developed allows for the design of SAC code families that exhibit a prescribed auto-and cross- correlation pattern. This is a valuable characteristic, enabling the customization of coded waveforms to take advantage of the specific performance of the transmitter/receiver. The characteristics of the auto-correlation path of the pulse-compressed signals are adjustable, *e.g.,* to the specific constraints dictated by the antenna pattern under consideration. Furthermore, our techniques can be implemented in a scenario where our target is illuminated by two or more radar signals in order to optimize the cross-correlation performance. Another remarkable property of the new coded waveforms is their potential to be operated at different scales whenever the need arises to provide multiple resolution modes, *e.g.,* in a ranging application. The availability of these coded waveforms affords the possibility of improved clutter suppression.

We first consider the recursive formula defining Walsh functions:

$$\begin{aligned} W_0(x) &= 1 \\ W_{2n}(x) &= W_n(2x) + W_n(2x-1) \\ W_{2n+1}(x) &= W_n(2x) - W_n(2x-1) \end{aligned}$$

And observe that the rule allowing movement from one scale to the next is in fact just one out of the many possible unitary transformations that can be used to produce a family of orthogonal functions with the same time-frequency characteristics at each scale. In a more general approach we investigate a series of multi-scale transformations giving rise – by means of the very same iterative Walsh scheme – to a whole class of new codes that exhibit the same auto- and cross-correlation characteristics at each scale.

The modified scheme can be described as:

$$\begin{aligned} C_0(x) &= \underline{v} \\ C_{2n}(x) &= S_1(C_n(2x)) + T_1(C_n(2x-1)) \\ C_{2n+1}(x) &= S_2(C_n(2x)) - T_2(C_n(2x-1)) \end{aligned}$$

Where S's and T's are suitably "well-behaved" transformations and v is the initial vector possessing the desired characteristics. In this context we chose S and T among those transformations which will preserve the auto-correlation pattern. An example of this procedure is given by the so-called Rudin-Shapiro sequence:

$$
\begin{aligned}
C_0(x) &= 1 \\
C_{2n}(x) &= (1-i)C_n(2x) + (1+i)C_n(2x-1) \\
C_{2n+1}(x) &= (1+i)C_n(2x) - (1-i)C_n(2x-1)
\end{aligned}
$$

It should be noted here that in the case of Walsh functions and Rudin-Shapiro sequences the transformations S and T are multiplications by a (real or complex) number of modulus one. This is not at all the only possible choice. The "good" choices for S and T can be efficiently described by making use of tools arising from Harmonic Analysis, so that the emphasis can be set on the space characteristics (auto- and cross-correlation, number of phases, *etc.*) or the frequency content of the resulting coded signals. The complexity of these algorithms is directly proportional to N Log(N) times the complexity of the transformations S and T. The described procedure is illustrated in Figure 10: the transformations S and T are "correlation-preserving" mappings, while the initial auto-correlation pattern is designed by computer.

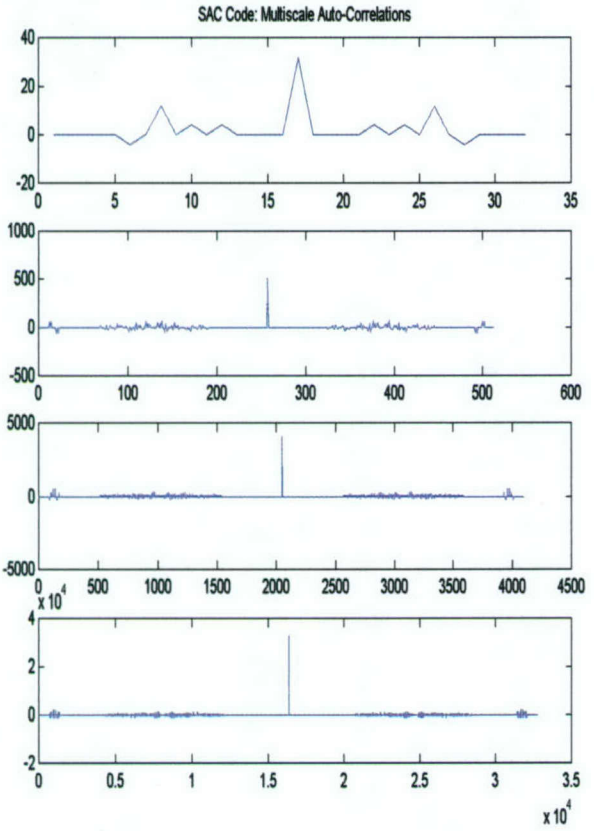


Figure 22: Illustration of SAC Correlation Properties

The case where we want to model our multi-scale SAC codes to have a pre-assigned frequency content is entirely similar. Our SAC family may, for example, be a DFT-like set with a fixed number of phases. By convolving the sampled return of a Radar receiver with an ad-hoc SAC sequence we can spot fluctuations in the Power Spectral Density of the signal, possibly due to the presence of a target.

3.5.2 Waveform Testing

Over the course of the contract, Raytheon has built up an Ka-Band radar test bed, which can support multiple advanced proof of concept (POC) engineering tests. The hardware for this set-up was procured using funds from the IR&D committed to this contract. We have worked supported two ISP Phase I subcontractors FMAH and the University of Melbourne (UniMelb). UniMelb has provided on set of binary waveforms, the so-called Prometheus Orthonormal Set or PONS, while FMAH provided a family of multiscale waveforms know as Spectral Analysis Codes or SAC.

3.5.2.1 Ka Band Radar: Facility and Test Equipment

The data sets were collected from one of the radar test towers at the Raytheon airport facilities. An Agilent E8267C Vector Signal Generator was used to replace the Direct Digital Synthesizer and up-convertor, which greatly simplifies waveform generation. Waveform I/Q data can be created from MATLAB and downloaded for transmission. Any type of waveform can be generated within an 80 MHz Bandwidth. The E8267C can be incorporated into a closed loop system as part of an integrated signal processing demonstration to evaluate waveforms, processing and waveform selection. The Agilent E8267C is limited to a 20GHz Frequency. Figure 23 shows the radar tower test set-up, while Figure 24 shows examples of a few transmit waveforms.

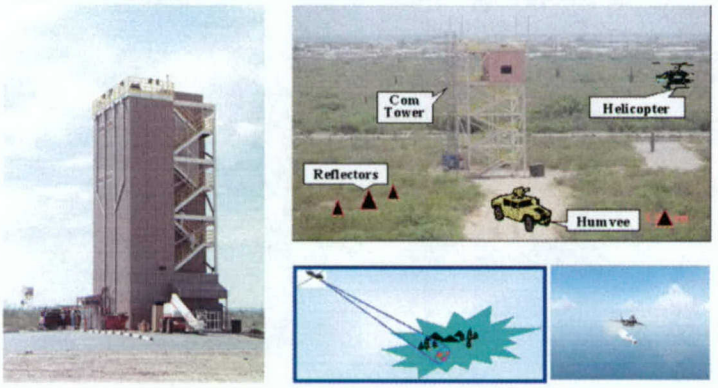


Figure 23: RF Tower and Simulated Targets

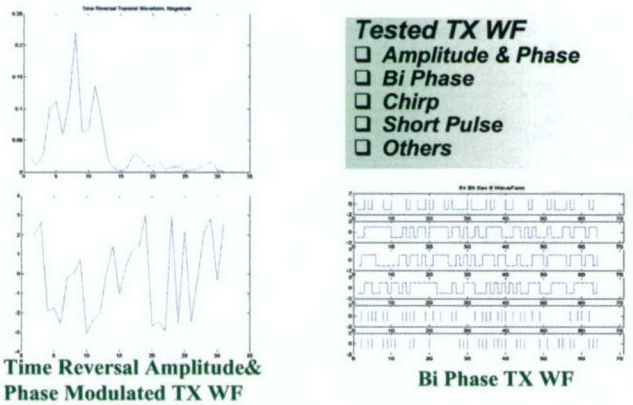


Figure 24: Sample Transmit Waveforms

3.5.2.2 Bi Phase Waveform Test

Radar targets are typically smeared both in range and Doppler space. The amount of smearing and its general shape depends critically on the waveform used as well as the subsequent processing of the return. This effect is particularly important in high clutter environments, where the clutter is smeared into regions of the range-Doppler space of targets of significance. As a result detection and tracking of such targets can be severely compromised. We investigated several new bi-phase waveforms that provide a greater degree of control over ambiguities. The three bi-phase waveforms were: PONS, Walsh and SAC. We generated ambiguity diagram for each tested waveform set and compared their range side lobes and Doppler tolerance regions. Figure 25 shows the ambiguity diagram for the three tested waveforms. Figure 26 show example test data plots of the PONS and Walsh waveforms. Figure 27 shows example of SAC waveforms test data with the target selection and null features.

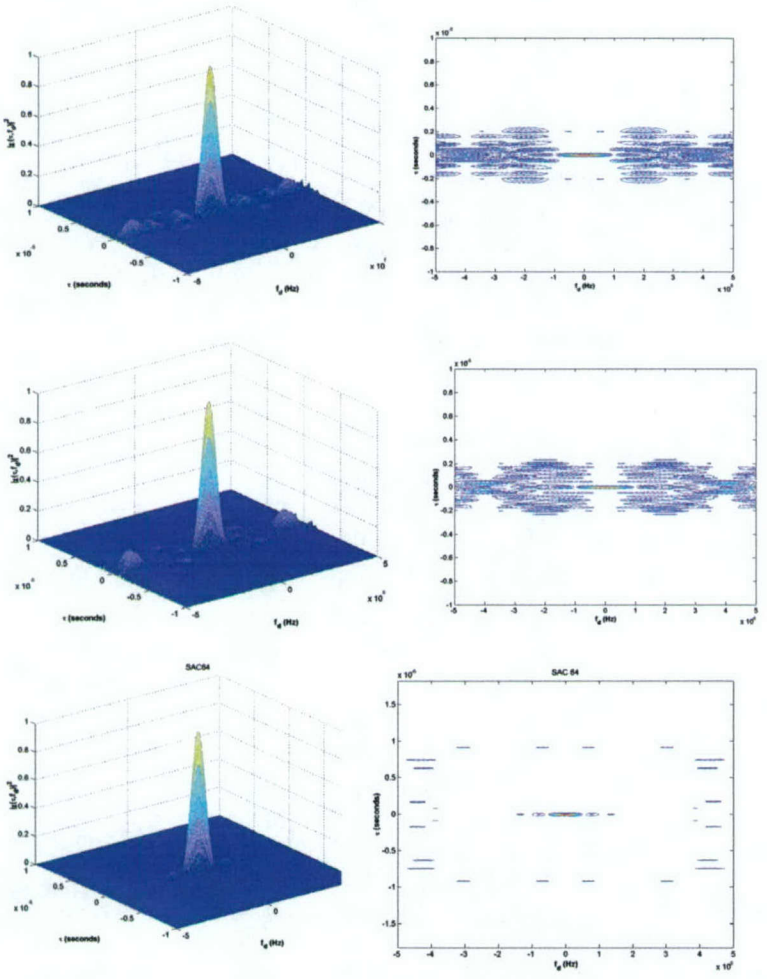


Figure 25: Waveforms Ambiguity Functions (PONS/Walsh/SAC)

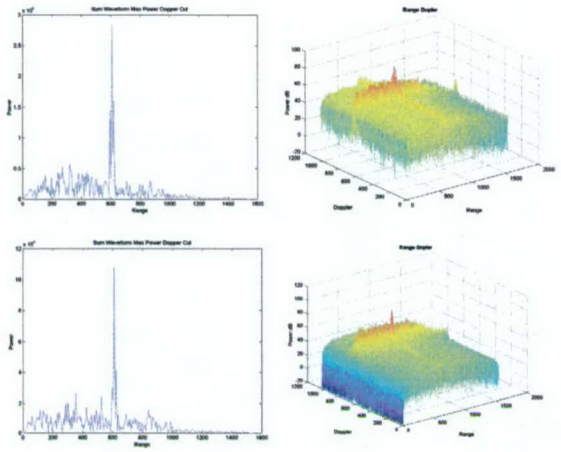


Figure 26: Test plots of PONS and Walsh waveforms

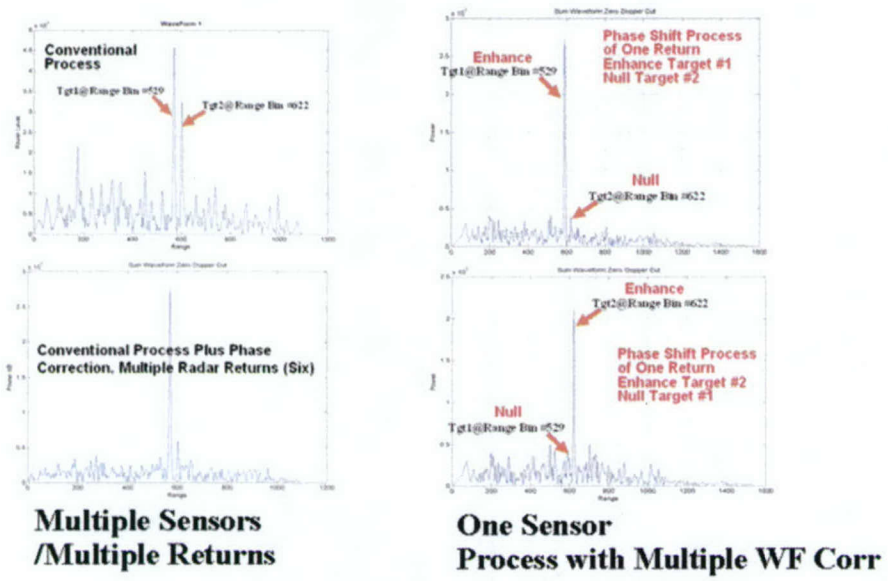


Figure 27: SAC Waveform Set, Test Data

4.0. Related Conference and Journal Articles

Over the duration of the contract, Raytheon has committed a significant amount of IR&D funds related primarily to structured materials and advanced signal processing. This work has resulted in a number of publications, which we also include here for completeness.

R (Refereed Journal Articles)
CR (Conference Proceedings, Refereed)
C (Conference Proceedings)
CI (Conference Proceedings, Invited)

1. [CR/CI] "Implementation of Distributed Networks of μUAVs with Low Power Low Bandwidth Sensing Modalities: Some Selected Challenge Problems," H. A. Schmitt and J. G. Riddle in the Proceedings of DASP 2001/02, July 2002.
2. [R] "An Object Detection Strategy for Uncooled Infrared Imagery," by H. A. Schmitt, J. G. Riddle, T. M. Brucks, R. R. Coifman and I. Cohen, J. Modern Optics, **50**, no. 9, 2003.
3. [C] "Advances in ATR Technology for Millimeter Wave Real Beam Target Identification," D. E. Waagen, M. L. Cassabaum, H. A. Schmitt and J. G. Riddle, ATR Science, Technology and Transition Symposium on *Tomorrow's Technology for Homeland Defense: Using ATR to Identify, Dismantle, Disrupt and Punish Terrorists Before They Strike,* October 2002.
4. [C] "Quantum Image Processing," R. D. Rosenwald, D. Meyer and H. A. Schmitt, 2003 Meeting of the MSS Specialty Group on Passive Sensors, 24-28 February 2003.
5. [C] "Adaptive FPA Using Photonic Band Gap Materials," D. J. Garrood, N. Shah and H. A. Schmitt, 2003 Meeting of the MSS Specialty Group on Passive Sensors, 24-28 February 2003.
6. [C] "Unsupervised Support Vector Machine Optimization via Margin Distribution Analysis," D. E. Waagen, H. A. Schmitt, M. L. Cassabaum and B. Pollock, Aerosense SPIE, Orlando, FL, 21-25 April 2003.
7. [C] "Asymptotic Performance of ATR in Infrared Images," C. Ceritoglu, D. Bitouk, M. I. Miller, H. A. Schmitt, Aerosense SPIE, Orlando, FL, 21-25 April 2003.
8. [C] "Adaptive Focal Plane Array," D.G. Garrood, N. N. Shah and H. A. Schmitt, RMS EOSTN Conference, 20-22 May 2003, Dallas, TX (Best paper award in the New and Innovative Technology Category).
9. [CR/CI] D. E. Waagen, M. L. Cassabaum, C. Scott and H. A. Schmitt, "Wavelet Basis Selection: Statistically Diverse Wavelet Bases for Multi-Class Discrimination," FUSION 2003 the 6th International Conference of Information Fusion, 8-11 July 2003.
10. [CI] "A Combined Particle/Kalman Filter for Improved Tracking of Beam Aspect Targets", D. A. Zaugg, D. E. Waagen and H. A. Schmitt, Special Session on Applications of Particle Filters in Signal Processing, 2003 IEEE Statistical Signal Processing Workshop, 28 September-1 October 2003.
11. [CR] "Simulated Bearings-Only EKF, Multi-Hypothesis EKF, and Particle Filter Performance with a Comparison to AT3 and HARM Data," D. A. Zaugg, A. A.

Samuel, D. E. Waagen, and H. A. Schmitt, The 12th Annual Workshop on Adaptive Sensor Array Processing, MIT Lincoln Laboratory, 16 - 18 March, 2004.

12. [C] "A Bearings-only Tracking Performance Comparison Using Simulated Particle and Multi-hypothesis Kalman Filters, and AT3 and HARM Data," D. A. Zaugg, A. A. Samuel, D. E. Waagen, and H. A. Schmitt, MSS Passive Sensors, Tucson, 22-26 March 2004.
13. [C] "Unsupervised Optimization of Support Vector Machine Parameters," M. Cassabaum, D. Waagen, J. Rodríguez and H. A. Schmitt, Defense and Security Symposium, Orlando, 2004.
14. [C] "A Comparison of Particle Filters and Multiple Hypothesis Extended Kalman Filters for Bearings-Only Tracking of Maneuvering Targets," D. Zaugg, D. Waagen, A. Samuel and H. A. Schmitt, Defense and Security Symposium, Orlando, 2004.
15. [C] "Incremental-adaptive support vector machine learning," D. Waagen, H. A. Schmitt and M. Palaniswam, Defense and Security Symposium, Orlando, 2004.
16. [CR] "Cognitive Nanoprobes: The Geometry of Processing and Sensing," H. A. Schmitt, *et al.*, 5th Asian Control Conference, Melbourne, Australia, 2004, accepted.
17. [CR] "Applications of Quantum Algorithms to Partially Observable Markov Decision Processes," R. D. Rosenwald, D. Meyer and H. A. Schmitt, 5th Asian Control Conference, Melbourne, Australia, 2004, accepted.
18. [CR/CI] "Unsupervised Optimization of Support Vector Machine Parameters," M. Cassabaum, D. Waagen, H. A. Schmitt, Defense Applications of Signal Processing, 1-5 November 2004, accepted.
19. [CI] "Sensor Scheduling Approaches for SWARMS and Ballistic Missile Defense," C. O. Savage, W. Moran, D. E. Waagen and H. A. Schmitt, Thirty-Eighth Annual Asilomar Conference on Signals, Systems, and Computers, Special session on "Signal Processing for Agile Sensors, Pacific Grove, CA, 7-10 November 2004, accepted.
20. [CI] "Computational Origami for Sensor Configuration and Control," H. A. Schmitt, D. E. Waagen, I. Streinu and G. Barbastathis, Thirty-Eighth Annual Asilomar Conference on Signals, Systems, and Computers, Special session on "Signal Processing for Agile Sensors, Pacific Grove, CA, 7-10 November 2004, accepted.
21. [C] "Novel Bi-Phase Waveform for Next Generation Radar", V. Adams and W. Dwelly, Raytheon 2003 RF Symposium, FL, May 2003.
22. [C] "Transmit Waveforms as part of the Integrated Signal Processing", V. Adams and W. Dwelly, Raytheon 2003 Processing Technology Symposium, CA, Sept 2003
23. [C] "Three Novel Sensing Transmit Waveforms and Cognitive Processing Ka Band Radar", V. Adams and W. Dwelly, Raytheon RF Symposium, MA, May 2004.
24. [C] "Simple & Low Cost Complex Waveforms Generation and Targets Simulation for Ka-Band Radar Tests", V. Adams and W. Dwelly, Raytheon RF Symposium, MA, May 2004.
25. [C] "New Radar Adaptive Transmit Waveform and Cyclic Processing", V. Adams and W. Dwelly, Raytheon RF Symposium, MA, May 2004.

5. New Discoveries, Inventions or Patent Disclosures:

System and Method for Tracking Beam-Aspect Targets with Combined Kalman and Particle Filters, D. A. Zaugg, A. A. Samuel, D. E. Waagen and H. A. Schmitt

The following patent application was presented to the Raytheon Patent Committee. The Committee has elected to defer processing of the patent application and has requested further technical and programmatic information.

Adaptive Waveform and Cyclic or Permuted Processing, P. Barbano, D. Healy, V. Adams and W. Dwelly

6. Interactions/Transitions:

6.1. Meetings

1. Unimodular Sequences Workshop, University of Maryland, June 2003.
2. Raytheon personnel have given ISP overview briefings to Customers on over fifty occasions. Audiences have included military and civilian personnel from AFRL/Rome, AFRL/Eglin, AFRL/Wright-Patterson, NSWC/China Lake, US Army Fort Huachuca, Special Operations Forces and the Border Patrol, as well as Customers for a number of proprietary programs.

6.2. Consultative and Advisory Functions

No consultative or advisory services were provided during this period of performance.

6.3 Honors/Awards:

No honors or awards were received during this period of performance.