# INTERIM PROGRESS REPORT

Period covered by report: 1 August 2012 to 31 July 2013

Proposal Title: (MURI) Value-centered Information Theory for Adaptive Learning, Inference, Tracking, and Exploitation

Contract/Grant #: W911NF-11-1-0391

Authors of report: Alfred O. Hero (PI), U of Michigan
Douglas Cochran, Arizona State U
Emre Ertin, Ohio State U
John W. Fisher III, MIT
Jonathon How, MIT
Michael Jordan, UC Berkeley
Randolph L. Moses, Ohio State U
Raj Rao Nadakuditi, U of Michigan
Stefano Soatto, UCLA
Alan Willsky, MIT
Angela Yu, UCSD

ARO proposal number: W911NF-11-1-0391

Principal Investigator:
Alfred O. Hero, Professor
Department of Electrical Engineering and Computer Science
The University of Michigan
1301 Beal Ave
Ann Arbor, MI 48109-2122
Tel: 734 763 0564
Email: hero@umich.edu

Submitted to: Dr. Liyi Dai
Program Manager
Army Research Office
Raleigh, NC

**Abstract**

This interdisciplinary project is developing a comprehensive set of principles for task-specific information extraction and information exploitation that can be used to design the next generation of autonomous and adaptive sensing systems. The significance of this research is that it addresses the widespread and longstanding problem of defining, assessing, and exploiting the value of information in active sensing systems. This year we report progress in twenty areas organized around three main thrusts: (1) learning and representation of high dimensional data, (2) distributed information fusion, and (3) active information exploitation. In the learning and representation thrust, progress ranges from assessing value of Kronecker representations of high-dimensional covariance matrices to learning to rank user preference data, an important task for human-in-the-loop decision systems. In the distributed information fusion thrust, progress is reported in assessing value of information in distributed information gathering and dimensionality reduction systems with application to sensor networks. In the active information exploitation thrust, progress is reported in information geometric trajectory planning, adversarial information collection, active learning in Bayes nets, and multistage adaptive estimation of sparse signals. Our future plans are to continue to develop linkages between these thrust areas, to further our development of fundamental theory for designing and evaluating distributed active information collection systems, and to account for human interactions in the sensing and processing loop.

# 1 Overall objective of project

Sensing and actuation systems are inundated with diverse and high volumes of data. Much of this data is uninformative and irrelevant to the end tasks of the system, which can evolve over the mission. The problem of extracting and exploiting the relevant and informative portion of sensor data has been an active area of research for several decades. Despite some progress, notably in information-driven tracking and data fusion, a general solution framework remains elusive, especially for autonomous and distributed sensing systems. The aim of this MURI is to develop a comprehensive set of principles for task-specific information extraction and information exploitation that can be used to design the next generation of autonomous and adaptive sensing systems. These principles will go beyond the standard information theoretic approaches that fail to account for non-classical information structures due to factors such as small sample size, poorly-specified target and clutter models, feedback control actions, hostile or adversarial environments, computation/ communication constraints, distributed sensing resources, and time-critical decision making.

# 2 Approach

Our research program aims to lay the foundations for a new systems theory that applies to general controlled information gathering and inference systems with mission planning. The research approach comprises three inter-related research themes that collectively address the most critical research challenges. These thrusts are: (1) information-driven structure learning and representation; (2) distributed information fusion for fast-paced uncertain environments; and (3) active information exploitation for resource management. We aim to develop an end-to-end framework that will result in better raw sensor data acquisition and processing, improved fusion of multiple sources and modalities, and more effective sensor management and control that accounts for human intervention.

# 3 Scientific barriers

This research addresses several challenges:

1. Reliable value-of-information (VoI) measures for active multi-modal sensing systems are not available. Existing approaches to learning and representation of information do not account for the sequential nature of data collection. This arises in active sensing systems such as autonomous maneuvering robots with vision/IR/LIDAR capabilities. Quantifying the value of information collected from active sensing systems is essential but there exists no suitable theory to do so. Classical Shannon information theory is inadequate as it was not designed for learning in active sensing systems; rather it was designed for data transmission in communications systems. A new theory for learning the value of information is needed that accounts for real-time feedback and control of the sensor, applies to signals that are non-linearly embedded in high-dimensional spaces, accounts for models with complex structural components (e.g.,

hierarchical graphical models of interactions in the scene), has scalable computation even in large distributed sensor systems, and accounts for the economic or human cost of acquiring data or fielding a new sensor.

2. There is no broadly applicable theory of information fusion for fast-paced uncertain environments. The design and operation of sensing systems must accommodate collection and delivery of a wide range of data at different times, spatial locations, and often with severe bandwidth and delay constraints. These systems must not have too many user-defined tuning parameters that could overwhelm the human operator. There is no generally applicable theory of multi-modal information fusion that accounts for all of these factors. Existing information theoretic measures and associated surrogates are often only weakly predictive of information fusion performance, and they usually require careful tuning when used as objective functions to drive the fusion algorithm. Reliable measures are needed for fusion in compromised environments having high background/clutter variability and spotty situational awareness coverage.

3. Most information exploitation algorithms do not accurately predict the ultimate value of a current sensing or navigation action in the presence of uncertain hostile environments. The sensor manager plans ahead and controls the degrees-of-freedom (actions) of the sensor and platform in order to achieve system objectives. These degrees of freedom include: region of focus of attention, choice of modality and mode (e.g., EO vs LIDAR), transmit waveform selection, and path planning actions (platform maneuvering). The manager must predict the value of information resulting from each of the candidate sensing actions. This prediction must account for the uncertainty of the environment, time-varying visibility constraints (e.g., target obscuration), erratic or adversarial target behavior, and sensor resource constraints. To date, most plan-ahead sensing and navigation approaches have been based on heuristics, like maximizing Shannon information-gain, and do not account for the value of information measure as a function of the end task or the uncertainty in the environment.

4. Information collection systems very often involve human intervention at some point in the collection process. Examples are annotation through Mechanical Turk, validation of contextual data, or curation of relations that have been imputed by machine into database. A basic challenge is how to mathematically model human-machine interaction in such as way as to be predictive of the value the intervention. Mathematical modeling is challenging since it must account for fatigue, latency, and biases that a human may unwittingly contribute to the corpus. There has been very little theory developed for human-in-the-loop processing for adaptive sensing that accounts for these factors and uses human cognition models from experimental psychology. This past year we have pursued several research directions in this area, described below. Furthermore, in recognition of the importance of the human-in-the-loop problem, we have added a co-PI to our team: Angela Yu, a faculty member in Cognitive Science at the University of California, San Diego, is an expert in human cognition, cooperative human behaviors models, and mathematical optimization paradigms (Markov decision processes) for human-human and human-machine interaction.

# 4 Significance

The significance of this research is that it addresses the longstanding problem of defining, assessing, and exploiting the value of information in active sensing systems. By defining new information measures that account for the future value of data collection, we can design better sensing, fusion, and planning algorithms that come with performance guarantees; e.g., tight value-specific bounds and performance approximations. By developing scalable and accurate methods to assess the value of information from empirical data, we can better design active sensor fusion and sensor planning to exploit the information collected thus far. The impact of the research is summarized by the following five points:

1. The research will result in more accurate prediction of performance using a new class of information measures that account for both quality and value of information.
2. The research will provide a foundational "systems theory" for active information gathering systems that use these new measures.
3. The research will use this foundational theory to develop highly adaptive and learning-based sensing strategies with significantly enhanced performance having reduced user tuning requirements.
4. The research will apply these sensing strategies to improve sensor signal processing, information fusion, and sensor platform navigation and control,
5. The research will uncover new strategies for involving a human-in-the-loop and assessing the intrinsic value of such involvement for different sensing and situational awareness tasks.

# 5 Specific accomplishments over the period 8/1/12 — 7/31/13

Our efforts remain organized around the three research thrusts defined in our proposal: (1) information-driven structure learning and representation, (2) distributed information fusion, and (3) active information exploitation for resource management. These thrusts are interdependent and most of our efforts fall across the boundaries between them. However, for clarity of presentation, in what follows we associate each reported progress and accomplishment with one of these thrusts.

## 5.1 Information-driven structure learning and representation

Because learning and feature representation are basic building blocks of fusion and resource planning, this thrust is crucial for maximizing the value of information collected by a sensing, processing, and decision-making system. Our effort in information-driven learning and representation encompasses three areas: (i) learning and representation of high-dimensional data, (ii) VoI performance quantification and tradeoffs, and (iii) human-in-the-loop processing.

### 5.1.1 Learning and representation of high-dimensional data

Contributors: Hero UM, Jordan UCB, Nadakuditi UM
Publications: [40], [41], [25]

Feature representation and learning for high-dimensional data is the starting point for studying value of information. Poorly chosen features will deprive the system of the ability to extract useful information for fusion, inference, or resource planning. Poor learning rates will make the system unable to adapt to fast-changing scenarios. We have made significant progress in both of these domains, as summarized below.

**Progress 1: Kronecker sum models for representing covariance matrices (Hero UM)** Information in spatial-temporal processes is often buried in their covariance functions. For example, in a network of wind velocity sensors, the speed and direction of atmospheric disturbance is encoded in the spatial dependencies between sensors, while the nature of the disturbance is encoded in the temporal dependencies within sensors. A first-order model for such spatio-temporal dependency is a Kronecker-product covariance model, which decouples the spatial and temporal spectra into product form. Such a first-order model benefits from having few parameters. For example, for 100 sensors collecting blocks of 100 time samples the spatio-temporal covariance matrix is of dimension $10,000 \times 10,000$ and has 100 million unknown entries (parameters). Under the Kronecker product model the covariance is completely specified by two smaller $100 \times 100$ covariance matrices corresponding to only 20,000 unknown parameters, providing three orders of magnitude reduction in the number of model parameters to be estimated. Last year we studied the problem of covariance and inverse covariance estimation in Kronecker product models when there is an additional requirement that the model be sparse; i.e. one or both factors are sparse matrices. We introduced an Kronecker covariance estimation algorithm, called the Kronecker Glasso (KGlasso) [40], and showed that, in terms of estimation root mean-squared error, use of this algorithm results in several orders of magnitude increase in VoI-per-sample [35], [38], [41].

This year, we extended our original first-order Kronecker model to arbitrary orders, an approach we call the "Kronecker sum covariance matrix decomposition" in analogy to PCA. The Kronecker sum covariance matrix decomposition finds the Frobenius norm best-fit matrix sum $\hat{\mathbf{\Sigma}}_r$ to the $p \times p$ sample covariance matrix $\Sigma$, where $\hat{\mathbf{\Sigma}}_r = \sum_{i=1}^r \alpha_i \mathbf{A}_i \bigotimes \mathbf{B}_i$ and $\mathbf{A}_i$ and $\mathbf{B}_i$ are non-negative definite matrices of row dimensions $m$ and $n$, respectively, with $mn = p$. The number of terms $r$ in the decomposition is called the separation rank. The decomposition has the following remarkable properties: (a) it is a complete representation; i.e., any covariance matrix can be represented as a sum of Kronecker products for some separation rank; (b) the decomposition is guaranteed to be positive definite; (c) the Kronecker spectrum, defined as the sequence $\{\alpha_i\}_{i=1}^r$, can be defined similarly to the quadratic spectrum of standard PCA; and (d) using the Kronecker sum decomposition results in VoI-per-sample that is several orders of magnitude greater than using the standard covariance matrix, where the VoI gain decreases proportionally to $r$ [37]. We have shown [41] that when applied to weather wind speed data the Kronecker sum decomposition gives a spectrum that is much more concentrated than the PCA spectrum; i.e. many fewer components are necessary to represent the covariance. The method is currently being applied to prediction and anomaly detection in video streaming and human activity modeling. These activities will be reported next year.

**Progress 2: Spectral measures and subspace detection from random matrices (Nadakuditi UM)**
In the area of non-commutative information theory, we are seeking to establish fundamental limits on the information that can be extracted from non-commutative observations, such as random matrices and tensors. For symmetric matrices these limits are governed by the asymptotic behavior of eigenvalues and eigenvectors of the matrix, and they specify phase transition thresholds of SNR and matrix dimension for which these eigen-quantities cannot be reliably estimated empirically. Such phase transition thresholds are key for developing the non-commutative information theory of dimensionality reduction, which is relevant, for example, to variable selection in sensor fusion.

This year we have made substantial progress in developing data-driven algorithms for low-rank signal matrix denoising using non-commutative (or free) probability theory. Specifically, we have developed an algorithm [25] for denoising a low-rank signal matrix buried in noise by optimal singular value shrinkage. The algorithm[1] explicitly utilizes information in the "noise portion" of the singular value spectrum to compute these shrinkage coefficients and returns an estimate of the approximation MSE that is provably consistent and that can serve as a new VoI metric.

Specifically, the new VoI metric more naturally encodes the inherent "noisiness" of the matrix-valued dataset than commonly used heuristic metrics, such as the singular value gap statistic. We have utilized this VoI metric to develop a scheme for improving estimation of the low-rank signal matrix in a multi-view, multi-modality setting where the modalities have different, time-varying intrinsic SNRs. Our new algorithm uses this VoI metric to optimally weight the individual signal estimates before fusion and thereby realizes significant performance gains to other averaging schemes based on *ad hoc* proxies for the intrinsic risk. In related work, we have described settings where middle components might be more informative than principal components and can hence be exploited to extract additional processing gain [26].

We have extended our characterization [27] of the universality (or genericity) of the square-root decay at the edge for free convolutions of compactly supported measures and shown that signal-plus-noise matrices even when compressed retain the square-root decay at the edge. This will facilitate the characterization of phase transition thresholds in the detectability of low-rank signals buried in noise and the determination of data-driven noncommutative VoI metrics analysis that can be subsequently exploited for inference and classification.

Additionally, we have analyzed the characterization of the local spectrum of truncations of the Kronecker products of random Haar distributed unitary matrices [10]. This work was inspired by the work by co-PI Hero, discussed under Progress 1 above, and his students on Kronecker structured graphical model estimation. Our work provides a first step in the direction of Kronecker structured compressed sensing.

We have initiated and have made progress in collaborations with co-PIs Cochran and Hero on extensions of this work to the sensing and detection of correlated signals in multi-modal signal processing problems. We also plan to develop a new eigen-VoI based rank detection algorithm that can be used to detect the number of signal and the intrinsic dimensionality of signals buried in noise. We are actively collaborating with MURI co-PIs to investigate extensions. In the upcoming

[1] Software available at `www.eecs.umich.edu/~rajnrao/optshrink`

year, we plan to validate our algorithms on real-world datasets.

#### 5.1.2 VoI performance bounds and tradeoffs

Contributors: Ertin OSU, Jordan UC Berkeley
Publications: [9], [7], [8]

Value of information is closely connected to performance quantification. Performance quantification is divided into (a) offline performance benchmarking and analysis and (b) online empirical performance prediction from measurements. We have made significant advances in both of these areas. A new area for this MURI this year is the evaluation of the effects of privacy constraints on performance. Privacy and security are closely related and have become increasingly relevant design parameters in sensing and learning systems in an era where much of the information infrastructure of modern defense systems serves dual-use civilian and military purposes.

**Progess 3: Performance bounds for inference using high-dimensional data (Ertin OSU)** Many sensor systems, such as camera networks or EO/RF sensors mounted on airborne platforms, are able to interrogate a scene persistently over a large range of aspect angles. Learning and exploiting the additional information provided by wide-aspect target signatures is key to developing successful automatic target recognition algorithms and characterizing their performance. The sensor data naturally resides in a high-dimensional space due to the number of sensors and their wide-aspect interrogation of the scene. Making probabilistic inferences for the class and pose of targets in clutter is a particularly challenging problem. This is due to the unknown noise statistics in the high-dimensional sensor data space. While performing feature extraction provides certain noise immunity, clutter can lead to false detections of local features and also to occlusion of some target features. The resulting perturbations in the high-dimensional data are extremely hard to parameterize and learn since they are akin to the "shot noise" commonly observed in optical systems, but with non-uniform spatial correlation unlike the optical system. As an example, characterization of the Hausdorff distance between two sets in high dimensions in the presence of spatial point process noise is an open problem. Without a statistical model for clutter perturbations, accurate performance bounds for pose and class estimation cannot be computed.

In our recent work [9], we employ manifold learning based dimensionality reduction for modeling and learning the clutter perturbations in a low-dimensional embedding space. In the embedding space, correlated perturbations at various dimensions are accumulated to small number of variables having multivariate normal distributions, which are easily parameterized and learned by the covariance matrix. For the signal model, the geometry of the low-dimensional embedding is learned to compute gradients and differential area elements. To validate this promising direction, we performed simulation experiments with data from wide-angle SAR sensors. To learn the local geometry in the embedding space from unlabeled training data, we employed a coordinate-free graph Laplacian operator. Next, Cramér-Rao Bound (CRB) analysis was used to quantify information provided by the wide-aspect target signatures. Our performance bounds learned from the data are in close agreement with state-of-the art ATR algorithms developed in independent work.

**Progress 4: Computation/statistics tradeoffs (Jordan UCB)**
The increasingly large scale of modern datasets creates a fundamental problem at the intersection of the computational and statistical sciences: how to provide guarantees on the quality of statistical inference given bounds on computational resources such as time or space. In recent work [7], we have developed a new approach to this problem based on the notion of "algorithmic weakening," in which a hierarchy of algorithms is ordered by both computational efficiency and statistical efficiency, allowing the growing strength of the data at scale to be traded off against the need for sophisticated processing. We have illustrated this approach in the setting of denoising problems, using convex relaxation as the core inferential tool. We provide a precise characterization of both the estimation performance and the computational complexity of employing a particular relaxation of a convex set by appealing to convex geometry and to results on the complexity of solving convex programs. In this manner, convex relaxations provide a principled mechanism to weaken inference algorithms in order to reduce the runtime in processing larger datasets. In ongoing work we are investigating extensions to model selection problems, where data are allocated non-uniformly to models as a function of their risk/computation tradeoff profiles.

**Progress 5: Local privacy and statistical minimax rates (Jordan UCB)**
The counterpart of the desire to exploit information is the need to keep some aspects of an information stream private, so that the information is exploited only in part. That is, the overall goal is often that of minimizing a statistical risk subject to a constraint on privacy. Working under a model of privacy in which data remains private even from the centralized statistician, we study the tradeoff between privacy guarantees and the utility of the resulting statistical estimators [8]. We prove bounds on information-theoretic quantities, including mutual information and Kullback-Leibler divergence, that influence estimation rates as a function of the amount of privacy preserved. When combined with standard minimax techniques, such as Le Cam's and Fano's methods, these inequalities allow for a precise characterization of statistical rates under local privacy constraints. We provide a complete treatment of three canonical problem families: mean estimation in location family models, parameter estimation in fixed-design regression, and convex risk minimization. For all of these families, we provide lower and upper bounds that match up to constant factors, giving privacy-preserving mechanisms and computationally efficient estimators that achieve the bounds. Our next step in this direction will involve the study of the effect of imposing privacy constraints within "neighborhoods of trust."

### 5.1.3 Human-in-the-loop processing systems

Contributors: Hero UM, Jordan UC Berkeley, Yu UCSD
Publications: [39], [36], [43]

It is self-evident that a human can provide essential contextual information to an automated sensing algorithms and aid to estimation, tracking, classification and situational awareneness, among other tasks. The challenge that we are addressing is how to mathematically model human machine interaction in such a way that the human-in-the-loop can be incorporated into our value-of-information framework for adaptive learning. We are making progress on human-in-the-loop processing in several different directions: a "twenty questions" game with a controller that asks questions of human and a machine who cooperate to localize a target in a noisy image; theory on the feasibility of reconstructing the partial ordering (directed acyclic graph) of human preference patterns from partial and incomplete observations; and modeling multi-agent (human) systems for foraging (wide area cooperative search) using Markov decision process models.

**Progress 6: Cooperative human-machine processing (Hero UM)**
This year, we started a new effort that explores the potential improvement in VoI when a human is included in the processing loop. In this work, we adopted a 20-questions framework. Consider a human and a computer cooperating to locate the position of a target based on sensing inputs; e.g. noisy imagery or ranging data. A controller asks the computer and the human a sequence of binary questions about the location of the target. In [39] we formulated this problem as a Markov decision process (MDP) and, under the assumption that the human and computer answer the questions correctly according to their own binary symmetric communications channels, we established properties of the optimal controller strategies; i.e., what questions to ask and in what order. We showed single-stage optimality of a sequential bisection policy that formulates the next question based on the median of the posterior distribution of target location at the current time. Under a simple acuity-limited human spatial resolution model we establish that the human adds more value in the beginning than at the end of the 20 questions game and that the VoI curve has a closed form Pareto shape. In [36], this framework is extended to the case of unknown response error (BSC crossover) probabilities. This work is in collaboration with Brian Sadler at ARL, and experimental validation of the model through an experiment at ARL is being considered for next year.

**Progress 7: Efficient ranking from pairwise comparisons (Jordan UCB)**
Existing approaches to the problem of learning to rank are often unrealistic, requiring all $n(n-1)/2$ comparisons to be made by a user or requiring repeated or active sampling. In many human-in-the-loop applications, it is instead possible to only obtain a subset of all $n(n-1)/2$ comparisons that is passively and noisily observed. While centralized optimization algorithms can be shown to learn rankings with a complexity that matches a lower bound of $\Omega(n)$ comparisons, it has been an open question as to whether such a bound could be matched in the distributed setting. In recent work [43], we have developed two new distributed ranking algorithms that match this lower bound in expectation. Furthermore, if an average of $O(nlog(n))$ binary comparisons are measured, then one of the two algorithms recovers the true ranking in a uniform sense, while the other predicts the ranking more accurately near the top than the bottom.

**Progress 8: Computational models of human cognition in cooperative search problems (Yu UCSD)**
This new new activity under the MURI started in summer 2013 when Angela Yu joined the project team. Her research group focuses on understanding the computational processes underlying human cognition, in particular how the brain represents and seeks out information from the environment as it tries to achieve behavioral goals in varying contexts. This work is important both for advancing human cognitive neuroscience and for improving human-machine interactions in hybrid or multi-agent settings. Currently, Yu's group is using a combination of theoretical and experimental tools to understand selective attention and active learning in human agents. They have found that humans adopt near-optimal but computationally efficient algorithms to actively acquire information from the environment, both in the context of active sensing and multi-arm bandit problems. In the next

year, they plan to build on this research to examine human behavior, and the interaction between humans and machines, in jointly exploring and exploiting novel or changing environments. On the theoretical side, Yu's group plans to expand their current Markov Decision Processes models for learning and control to multi-agent settings. Specifically, they will build on state-of-the-art models in active learning and multi-arm bandit problems, as well as incorporating elements from game theory, to study these multi-agent problems. On the experimental side, they will conduct human behavioral experiments both in more traditional experimental psychology paradigms and also develop online crowd-sourcing paradigms based on Amazon's Mechanical Turk infrastructure.

## 5.2 Distributed information fusion

Depending on the fusion algorithm implemented, fusion of information across a network of sensors can either enhance or degrade the value of information collected at each sensor. Furthermore, centralized fusion at a fusion center is often not possible due to limited communications bandwidth and throughput of the network. Thus the study of effective distributed information fusion methods is a key part of the MURI team's activities. Our progress in distributed fusion falls in three areas: (a) local information aggregation in networks, (b) VoI for image fusion, and (c) multiple-agent systems distributed fusion and planning.

### 5.2.1 Local information aggregation in networks

Contributors: Hero UM, Moses OSU, Cochran ASU, Ertin OSU
Publications: [17], [3], [47], [34].

We are developing inference algorithms that can be implemented in a distributed manner over a network of sensors. Inference tasks of interest include parameter estimation, detection, and tracking. Our progress in these areas is reported below.

**Progress 9: Distributed learning of Gaussian graphical models via marginal likelihoods (Hero UM)**

We have discovered a new fast and accurate algorithm for accelerating distributed inference over networks, significantly extending and improving on our progress reported last year [16]. Adopting the Gaussian graphical model (GMM) setting, we have shown that a modified ensemble of local maximum-likelihood (ML) estimators can come very close to achieving the same performance as the global ML estimator *without performing message passing.* The key is to expand the definition of local neighborhood beyond the standard first order (one hop) neighborhood of each node, including in addition the second order neighbors (two hop) of each node. This simple expansion of the neighborhood gives a statistically consistent and accurate ML estimator without any iterative message passing. Since message passing incurs considerable slowdown and delay, the practical implication is that our new algorithm can be more rapidly performed in a networked setting without sacrifice in estimation accuracy. This advantage is inherited by other algorithms that act on the output of our ML estimator; e.g., distributed PCA, prediction, or anomaly detection in the sensor network. This was published as a paper in a premier machine learning conference [17], where it won a Notable Paper Award.

**Progress 10: Aggregating local information under communication constraints for decision-level fusion (Moses OSU)**
In the area of distributed inference in sensor networks, we are analyzing the interplay between local decision, global inference, performance, and communication. In our earlier work, we analyzed how the performance of a large-scale sensor network scales with the density of sensor nodes in a random sensor network. We analyzed the detection performance as a function of sensor density subject to a constant network bandwidth constraint, where we assumed an idealized communication network where messages are transmitted reliably if the total rate does not exceed the available bandwidth. Under this model, we showed that the policy of desensitizing sensors (i.e., increasing their thresholds) leads to better detection performance with increasing sensor density while satisfying constant network bandwidth and global false alarm probability constraints.

In work undertaken this year [47], we extended our results to the case of realistic random access channel models with collisions resulting in packet losses. Specifically we consider a randomly deployed large-scale sensor network whose nodes follow a random sleep/wake schedule. Each node awakens with some probability at each time interval, and if active, senses its environment, computes a local decision, and communicates any detections to the fusion node using a Slotted ALOHA protocol. When multiple nodes attempt to communicate over a shared wireless communications medium, transmitted messages are subject to errors due to collisions with other ongoing message transmissions. Therefore, the fusion node must take number of detected collisions into account when forming a decision rule. The fusion center can detect both successful communications and communication collisions in the channel. We show that the optimum fusion rule is a weighted sum of the number of detections received and the number of collisions detected by the fusion node. We derive analytical expressions that characterize the performance of the system. Our results suggest that the optimal fusion rule that utilizes information in detected collusions have comparable performance to the case of "ideal" communications with no collisions. Additionally, the gap in performance can be reduced by increasing the number of slots, but at the expense of added delay.

**Progress 11: Value of information sharing in networks (Cochran ASU)**
Another vein of research in the past year has considered the value of information sharing in a networked system. In the context of multiple-channel detection of an unknown signal source using networked receivers, we have been able to quantify losses in operational performance metrics (e.g., probability of correct detection for a fixed probability of false alarm) incurred when actual measured data on a particular link in the network is replaced by a minimally informative (maximum entropy) surrogate measurement. This loss characterizes the value of the particular link between nodes in the network, hence information sharing, in terms of the information collaboration objective [3].

A recently initiated collaboration between two MURI team members (Nadakuditi and Cochran) is investigating the relative efficacy of multiple-channel detection with a network of sensors in a standard way ("fuse then analyze") versus a scheme that eliminates uninformative single-channel measurements before engaging a multi-channel detector ("analyze then fuse").

**Progress 12: Quantizing information for sequential decisions (Ertin OSU)**
Spatially distributed collections of micro-sensor nodes can provide surveillance and monitoring of large-scale structures. To detect and localize low signal-to-noise ratio events, sensors have to integrate their information across time. Successful adoption of large-scale distributed sensor systems is only possible if such networks can provide long lifetime. As a result, temporal fusion of the sensor data has to be performed using low-complexity/low-power algorithms. Sequential decision procedures rely on computing, aggregating, and communicating likelihood information at high precision. But this may be unsuitable for low-power sensor nodes with limited computation and communication capabilities. Our work this year [34] has considered design of quantized likelihood algorithms in the form of finite-state machines that are suitable for implementation in low-complexity devices. We have derived necessary and sufficient conditions for optimal likelihood quantization for sequential testing. Based on these results, we introduced an iterative algorithm, based on policy iteration, that converges to optimal quantization rules. We show that a simple finite-state machine decision rule with a small number of states can closely approximate optimal sequential test performance. We also studied the performance of quantized likelihood decision rules as the number of memory states and sampling cost is varied. For a small number of quantization levels, the value-aware quantization that is optimized for sequential detections was shown to outperform uniform quantizers of likelihood information.

### 5.2.2 VoI for image fusion

Contributors: Fisher MIT, Soatto UCLA
Publications: [6], [48]

Exploitation of image and video data is one of the focus applications guiding the basic research in this MURI. Active vision (Soatto) and spatio-temporal video analysis (Hero) are image applications described elsewhere in this report. The specific problem of image and video fusion have been addressed in terms of scene reconstruction and fusion of higher-order dependencies in video. Progress in these two areas is described below.

**Progress 13: Value of information for LIDAR versus EO in scene reconstruction (Fisher MIT)**
We have considered the problem of 3D scene analysis inferred from multi-modal data sources. Wide-area motion imagery (WAMI) and LIDAR provide rich information about large geographic areas at relatively low cost. The goal of this line of research is to investigate VoI analysis on real-world problems in order to quantify the relative values of different sensing modalities in the context of latent variable models. One challenge to principled multi-modal approaches is that one must consider the joint statistical properties across modalities in order to combine within a consistent mathematical framework. We investigated a generative Bayesian formulation in which joint properties are encoded implicitly via a latent representation of the scene and which allows for VoI computations. In doing so, we are able to demonstrate both mathematically and empirically the trade-off between image-only approaches versus those which exploit both imagery and LIDAR measurements. The results demonstrate comparable (and at times superior) reconstructions with fewer measurements and at lower computational cost by exploiting the expected information of each incremental measurement.

This work is the subject of a completed Master's thesis[6] and will be submitted to an upcoming technical conference.

**Progress 14: Inference in Markov random fields (Soatto UCLA)**
In [48], we have developed analytical and computational tools to exploit high-order relationships between nodes in a graph for the purpose of inference in a Markov-Random Field framework. The relevance of this work in the context of the MURI arises from the fact that remote sensing data (e.g., imaging) contain high-order dependencies ("context") that must be taken into account in order to make the inference unambiguous. One of the limitations of low-order dependencies (e.g., co-occurrence) is that no geometric context is enforced; with higher-order dependencies, these can be captured, modeled, and learned from data. Inference in high-order MRFs is notoriously difficult, but in [48] this is simplified by considering nonlinear relations that simplify the derivation of the posterior.

Each of the nodes in the MRF typically represents local properties of the sensor field domain, such as color, gradient orientation distribution, or the response to a bank of filters. The temporal dimension plays an important role, and should not be treated just like one more spatial dimension; in [31] we have developed low-level bottom-up spatio-temporal partitions of the images for the purpose of video analysis.

### 5.2.3 Multiple agent systems distributed fusion and planning

Contributors: How MIT
Publications: [48], [21], [23]

**Progress 15: VoI-based distributed inference and planning (How MIT)**
In networked multiple agent systems, agents typically have limited on-board resources, such as computation ability, communication bandwidth, and fuel. Hence, efficient distributed algorithms are needed to ensure that agents can optimally utilize these limited resources to execute the missions, especially in dynamic, uncertain environments. Our recent work has focused on VoI-based sensing and planning algorithms for multi-agent distributed systems to address these issues.

This year, we extended our previous work on VoI-based distributed estimation (see [20]) by developing theoretical bounds and testing it on temperature data with more depth. With VoI-based Distributed Sensing (VoIDS) and Adaptive VoIDS (A-VoIDS) algorithms, agents communicate with neighbors only when the local information exceeds a value threshold measured by appropriate information metrics. These algorithms are designed to overcome known shortcomings (e.g., excessive communication cost and slow convergence rate) of traditional consensus-based algorithms. Furthermore, VoIDS and A-VoIDS do not require the knowledge of network topology, which also alleviates known limitations of inference algorithms based on graphical models. We proved that, within exponential family distributions, the communication cost of VoIDS asymptotically converges to zero and that the estimation error is bounded. We also showed that the adaptive variant of the approach (A-VoIDS) is guaranteed to almost surely converge to zero estimation error while attempting to exploit all of the available communication bandwidth.

We further evaluated the approach using a real dataset that has been used previously to analyze other distributed sensing algorithms [5]. In this dataset, there are 54 sensors around the Intel Berkeley Research lab that collect timestamped information such as humidity, temperature, light, and voltage values every half minute. This dataset reflects real effects such as sensor noise, sensor bias, and time varying (slowly) drifting parameters. The performance is compared with consensus based algorithms HPC [11] and Random Broadcast [32] in Figure 1. VoIDS significantly reduces the communication cost compared to HPC, but also has a larger error. The Random Broadcast reduces cost by randomly censoring agents, but its performance has high variance and is no better than VoIDS on average. A-VoIDS with more communication bandwidth (reflected by $c$) is situated closest to the bottom-left corner of the error-cost figure, indicating that this adaptive approach can give an estimate while using significantly less communication cost than the other algorithms that have similar error. These results are published in [21, 23].

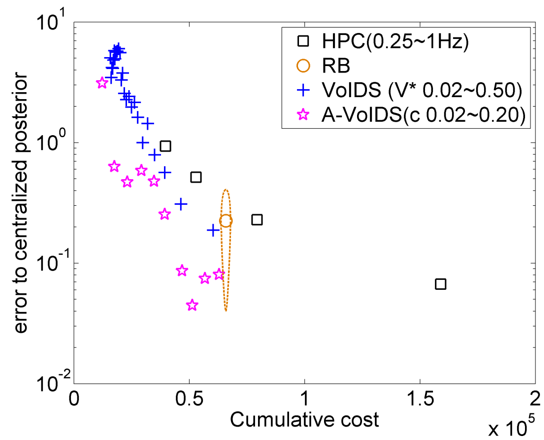


Figure 1: Performance Comparison of different algorithms for estimating average room temperature at a given time of the day from the Intel dataset. The total communication cost incurred is on the X-axis and KL-divergence to centralized (ideal) posterior is plotted on the Y axis. Adaptive distributed fusion algorithm which adaptively adjusts the VOI threshold $V^*$ based on available comm. bandwidth $c$ outperforms HPC both on accounts of accuracy and cost.

Multi-agent missions also require high-level autonomous planning algorithms that coordinate heterogeneous agents with different abilities. We have seen robust optimization techniques being applied to agent planning problems in uncertain environments [1, 15, 4]. When agents can actively take observations, it is well-studied in sensor management literature how to allocate agents to optimally reduce uncertainties in targets/environment [2, 42, 49, 13, 14]. However, in heterogeneous planning problems, it is often the case that the main goal is to perform specific missions instead of purely reducing uncertainty. Therefore assignments only based on uncertainty reduction can potentially lead to an inefficient use of resources (i.e., reduce high uncertain low reward tasks). Previous work at MIT by Bertuccelli and How proposed a planning algorithm based on integer programming [4] that establishes an explicit coupling between the exploration and exploitation of missions, but that work is limited to uncertainties that have a Gaussian distribution of the potential mission reward. While Gaussian distribution is a good model of uncertainty for measurement noise of continuous states such as positions, orientations and velocities, it does not provide a good representation of discrete uncertainties, such as the uncertainty in the category of the mission type.

Additional work has developed a *Value of Information based Distributed Planning framework* that couples the value of exploration activities into the mission return of tasking activities in a distributed system with heterogeneous agents. More specifically, two algorithms are compared: the decoupled approach, which independently allocates exploration agents and tasking agents based on the uncertainty reduction and mission return, respectively; and the coupled approach, which


Figure 2: Hardware testbed in Aerospace Controls Lab., MIT. Targets are patched pictures glued on ground, with dominant color representing its classification. Exploration is done by security cameras hanging on the wall. Tasking is done by robots visiting targets.

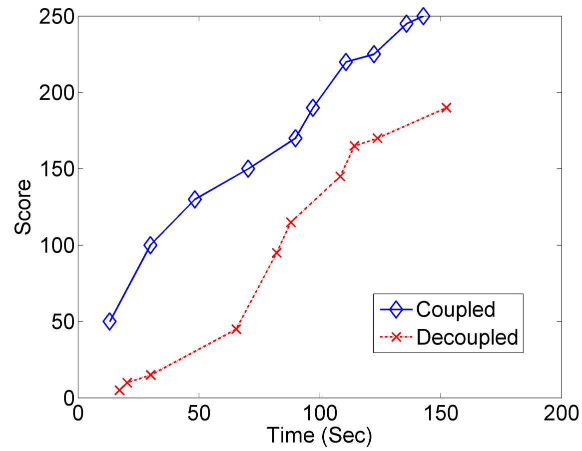


Figure 3: Cumulative scores during the mission showing the improved performance of the coupled approach.

allocates exploration agents based on their value in helping improve the tasking agents' mission return. Detailed numerical comparisons of these two approaches have been done with a categorical uncertainty in tasks. A hardware testbed was also built in Aerospace Control Laboratory at MIT to evaluate these planning algorithms. An overview of the lab is shown in Figure 2. Targets are patched pictures glued on ground, with a dominant color representing its classification (but other colors representing the uncertainty). Exploration is done by PTZ cameras mounted on the wall. Cameras can tune their pan/tilt angles and zoom levels to take pictures of targets. Since the picture is not pure in color, the measurement of classification is noisy. When a robot visits a target, it earns a score corresponding to the target class, and thus the uncertainty in the target class couples in tightly with the planning problem. In this experiment, there are 10 targets, 2 exploration agents (cameras) and 2 tasking agents (robots). Figure 3 shows the cumulative scores obtained by the two approaches over time, and the results show that the decoupled approach starts the mission with tasks that have higher uncertainty, which are not necessarily the most rewarding ones, while the coupled approach takes observations of potentially more rewarding targets first and thus goes to most rewarding targets from the beginning. This explains why the coupled approach has a higher score in the beginning at Figure 3. After 100 seconds, the gap between decoupled and coupled approach narrows because the decoupled approach discovers more rewarding targets during this time. Overall the score of coupled approach is always higher than the decoupled approach because coupled approach encourages cooperation between agents. Exploration agents will take observations of a target before tasking agents arrive there, thus the uncertainty (measured by entropy) is lower and score is higher when tasks are being done in the coupled framework. More specifically, the mean entropy of the targets of the coupled approach is 0.0695, which is significantly less than that of decoupled approach (0.2947). This work is published in [22, 24].

This work contributes significantly to the goal of this MURI because it uses VoI concepts to provide a more efficient framework for performing distributed parameter estimation and task planning. Furthermore, it is significant to the distributed system literature because it introduces the notion of *value of information* instead of only qualify of information. The algorithms discussed here, and their possible variants, would translate to significant resource savings in real-world distributed applications by preventing gathering and processing irrelevant and marginally useful information. This work therefore makes progress towards the objective of developing the next generation efficient and accurate intelligent systems. Both works are presented in more detail in graduate student Beipeng Mu's masters thesis [19].

## 5.3 Active information exploitation for resource management

The flow of information through a controllable sensor system affects decisions on control actions that can enhance target detection or tracking performance. The active information exploitation thrust lies at the interface of estimation and control and we have made progress in two subareas: information-driven sensor planning and Robust adaptive planning and policy approximation.

### 5.3.1 Information-driven sensor planning

Contributors: Fisher MIT, Ertin OSU, Cochran ASU
Publications: [28], [29], [33], [18]

Information can be used as a metric for sensing in situations where there are non-classical information flows. Non-classical information flow occurs when there is an adversary who tries to hinder the system's attempts to maximize the flow of information from the target to the decision maker. Information flow is also non-classical when control actions are introduced into the process of sensing, processing and decision making. In the latter case, we think that differential geometry of the statistical manifold can capture the effect of these control actions on the value of information. We report several areas of progress: MI-based planning in Markov chains, adversarial information structures, and information-geometric planning.

**Progress 16: Efficient computation of information rewards in Markov chains (Fisher MIT)**
We have studied the problem of information planning for active inference in graphical models. A wide variety of inference problems are commonly formulated using such models. Sensing in distributed sensor networks is a relevant example. An important aspect of such problems is that the measurement process is often subject to selection constraints as a means of conserving resources. The goal then is to maximize information gain subject to the given selection constraints. Mutual information (MI) is a commonly used objective criterion as it quantifies uncertainty reduction. The problem of choosing an optimal subset of sensors/measurements is often cast as a combinatorial optimization problem which becomes intractable as the number of measurements grows. Prior work exploiting the submodular property of conditional mutual information resulted in theoretical performance guarantees and a variety of both off-line and on-line bounds when comparing tractable

greedy measurement selection to combinatorial (i.e., intractable) optimal measurement selection in the context of inference in graphical models. These bounds were invariant to the order in which measurements sets are visited (as described by a "walk"). As the reward of each individual walk will, in general, be different, analysis of the computational complexity of evaluating the information rewards of different walks subject to the same selection constraints may yield higher information rewards. In the context of Gaussian models, we show that evaluation complexity is dependent on a number of critical factors including the number of observations sets, the sparsity of the measurement model and the structure of the walk itself. The analysis is specialized to latent variable models where the latent structure is a Markov chain, tree, or poly-tree and measurements are independent conditioned on the latent structure. We demonstrate empirically that both the information reward and evaluation complexity are largely decoupled and as such, exploration of low complexity walks yields higher information rewards tighter upper bounds. Additionally, this analysis has led to an efficient implementation of incremental inference in Markov Chains which exploits the same sparsity properties leading significant reduction in computational complexity. This work is currently under review for publication [28, 29].

**Progress 17: Value of information on statistical manifolds (Cochran ASU)**
During this year, we have continued our investigations into how classical information measures, the Fisher information in particular, can be exploited in sensor management. Specifically, we consider situations in which the controllable degrees of freedom in a configurable sensor suite are constrained to vary smoothly in time. This assumption is realistic for a broad class of controllable sensor parameters, such as the position of a vehicular platform which cannot instantaneously jump from one position to another. But some situations allow virtually instantaneous change between very distinct sensor configurations (e.g., the beam pattern of an electronically steered antenna array). We consider the exploitation of such a sensor suite in a fairly general class of estimation problems, which incorporates many detect/classify/track scenarios, in which a parameter $\theta$ on a smooth parameter manifold $\Theta$ is to be estimated from data collected by the sensors.

We have developed a formalism that identifies each sensor configuration with the Riemannian manifold defined by regarding the Fisher information for estimation of $\theta$ from sensor data collected with the particular sensor configuration as a Riemannian metric on $\Theta$ [18]. In this way, we identify each sensor configuration with a Riemannian metric on $\Theta$. The collection of all Riemannian metrics on a smooth manifold is known to constitute an (infinite-dimensional) Riemannian manifold $M$, and we have shown that the family of metrics corresponding to sensor configurations is a finite-dimensional manifold $S \subset M$ that inherits a Riemannian structure from $M$. We have described an algorithm for changing sensor configuration with time by following geodesic paths in $S$, and we have shown that this strategy accumulates information, in a sense that has both Shannon (symmetrized mutual information) and Fisher interpretations, optimally quickly.

The computational complexity of this information-geometric sensor management scheme has proven formidable for even small problems. Consequently, one aspect of our current work is examining more efficient ways to realize conceptually similar approaches; e.g., following approximate geodesic flow fields in the sensor configuration manifold $S$.

An additional information-geometric scheme for navigation directly on a parameter manifold has emerged this year from a collaboration between two MURI investigators (Hero and Cochran) in consultation with and ARL researcher (Brian Sadler). This idea seeks to capitalize on relationships between global measures of distance (e.g., Hellinger distance) between probability distributions in a parametric class and distance measures (e.g., information distance) that support local strategies for following optimal (geodesic) trajectories. This work is still at a formative stage, with results restricted to specific families of distributions (multinomial). Additional collaborative investigation of this idea is planned for the coming year.

**Progress 18: Adversarial information structures (Ertin OSU)**
We are continuing to investigate the degree to which an intelligent target can reduce information collection efficiency of a sensor system. In particular, we are developing information-theoretic uncertainty measures to provide a foundation for quantifying value of information in adversarial situations. In turn, these measures are used to devise information-driven sensor management/placement algorithms applicable to adversarial scenarios.

In our early work, we considered the detection problem when the conditional density of the sensor readings can be affected by the target's adversarial actions. A typical application of the proposed setup is surveillance with spatially distributed sensors, where the adversary is changing locations to evade detection. We considered the setting where the fusion center typically can access only a subset of sensors at any time, to satisfy bandwidth or energy constraints, and seeks to maximize a performance metric for an inference task, such as tracking or detection. Using asymptotic error exponents as the payoff function, we derived optimal strategies for the target and the observer, found the Nash equilibrium of the corresponding surveillance game, and characterized the value function as a VoI measure applicable to adversarial settings. Subsequently, we applied this VoI measure in deriving optimal sensor placement algorithms [33]. We formulated this sensor placement problem as a two-stage game between the observer and the target: in the first stage sensor locations are chosen by the observer, in the second stage the observer and the target choose open-loop control strategies to optimize probability of detection with opposing interests. The saddle point of the second-stage game is parameterized as a function of the sensor configuration. With this set-up, the problem of finding an optimal configuration for the sensors that maximizes the value of the saddle point problem is proposed and addressed. We provide a numerical solution to the saddle point problem along with a computationally efficient gradient based method.

#### 5.3.2 Robust adaptive planning and policy approximation

Contributors: Hero UM, Soatto UCLA
Publications: [46], [44], [12], [30]

By closing the loop between sensing, processing, and decisions one can exploit measurements and models to significantly enhance performance under resource constraints. Planning ahead using predictive models results in sensor actions that are informed by previous measurements and which use all available resources most efficiently. We have made progress in two areas of robust adaptive sensor planning, as described below.

**Progress 19: Multi-stage adaptive estimation of sparse signals (Hero UM)**
We have made significant progress on multi-stage scheduling of agile sensors for estimation of a few targets over large search space, the sparse signal scenario. This problem arises in wide-area

search and tracking, sensor selection, waveform selection, and other relevant scheduling problems. Building on our previous convex optimization approach [45] to multi-stage planning under resource constraints, we have been able to provide theoretical guarantees on the task-related VoI of our policies. By restricting the policies to the class of open-loop feedback controls (OFLC), we show that the multi-stage optimal policy can be found by dynamic programming. These policies improve monotonically with the number of stages, and the improvement can be quantified by an information divergence measure. This allows the VoI of multistage planning to be evaluated quantitatively for different scenarios and observation models. This work will soon appear in a journal [46] and it has been applied to spectrum sensing with application to cognitive radio [44]. We are actively collaborating with co-PI Jon How on incorporating this convex-optimization framework into wide search applications where target classification is the primary goal and correct identification of different targets may have different value payoffs relative to the mission.

**Progress 20: Control authority and performance in active vision (Soatto UCLA)**
In [12], we have continued our attempts, described in the original proposal, to quantify the tradeoff between control authority and performance in a decision and control task. Control authority refers to the policy for sensing and navigation actions over the measurement campaign. This tradeoff is central to active perception and active learning, and to the development of representations that are invariant to nuisance factors that cannot be eliminated in pre-processing. In [12] We have performed some analysis and an empirical study of the performance in a simple binary classification task (detection of an object previously seen) as a function of the control authority of the sensor, that involves both energy (mobility, sampling frequency) as well as the geometry of the reachable space (planar versus spatial robots). There does not appear to be any minimal sufficient statistic for this problem but an maximal invariant sufficient statistic can be derived that quotients out nuisance parameters that characterize appearance and rotations. This maximal invariant statistic gives a mathematical representation of the objects in the scene whose form depends on the desired task: e.g., object recognition (detection/classification) versus object reconstruction. The dichotomous nature of the representation for different, but complementary, creates challenges for active vision that will be the focus of our future efforts.

Previously reported work on the development of an active vision system for fire detection [30] is now on-line and working during the fire season.

# 6 Future research plans and anticipated scientific accomplishments

In the past two years we have made significant progress in developing fundamental theory to assess and optimize value-of-information for adaptive sensing, distributed fusion, and information exploitation problems. Aspects that distinguish our research from other VoI approaches include: development of theory that provides performance guarantees; use of the theory to predict of system requirements, e.g., bandwidths or sample sizes; and development of new algorithms that provably and practically outperform the state-of-the-art. For the coming year we will continue to develop fundamental theory in current areas of focus in addition to developing theory in some new areas of focus.

The current areas of focus that we will continue to address include the following. 1) Theory of Kronecker product representations of spatio-temporal process covariance and associated information gain achievable with such representations in video and other imaging applications. 2) Refinement of the theory of cooperative human-machine processing that includes decentralized teams of human-machine decisionmakers and its application to target search and classification with imaging sensors. 3) Theory of human-in-the-loop processing that accounts for privacy and performs predictive inference of human behavior from partially observed data, in particular preference ranking data and web search data. 4) Random matrix and non-commutative information theory approaches to signal subspace detection and information fusion that can extend the signal detection threshold by several dBs; 5) Exploitation of information sharing by second and higher-order neighborhoods for decentralized state estimation in sensor networks. 6) Value-of-information driven fusion and sensor planning that accounts for a target-dependent reward in addition to the standard reward for correct target detection or classification. 7) Value-of-information theory that captures tradeoffs between algorithmic complexity and algorithm performance. 8) The continued use of submodularity and adaptive submodularity for bounding the loss due to use of greedy sensor fusion and planning strategies. 9) Enhancing the value-of-information by adaptive sensing strategies including navigation, waveform selection, and beam-scheduling for radar, sonar and active vision modalities.

In addition to the research plans described above in the context of individual co-PI research projects, there are several new research directions that we may pursue as collaborative projects among the MURI co-PIs. These research directions have come out of extensive discussions, meetings, and visits between co-PI's and their students over the past year. Projects that co-PI's have discussed for next year include the following.

1. **Validation of our performance predictions and fusion algorithms using a software defined radar testbed.** In 2012, the team applied for a DURIP grant to build a software defined radar testbed for use by the MURI co-PIs. This DURIP grant was awarded in mid-2013. OSU is lead on the grant and plans to build several X-band radars with 1 square meter spatial resolution and range of approximately 100 meters. These will be portable with small form factor (breadbox size), and they will be software configurable. Several members of the team have plans to use these radars. At MIT, John Fisher and Jon How plan to mount them on robots and semi-autonomous vehicles to supplement imaging modalities that they have already instrumented. At Michigan and at Arizona State, Hero and Cochran plan to use them to validate algorithms for agile-waveform sequencing (Cochran), beamforming (Hero), and programmable PRI (Hero). Cochran and Hero have previously derived such algorithms under this MURI but, while they have been validated mathematically and by simulation, they have not yet been validated on actual hardware for real radar environments. This will likely lead to more realistic statistical models, and quite possibly inspire the creation of new algorithms to meet the constraints of the hardware. OSU plans to assemble sufficient numbers of these portable software radars to be distributed to any MURI co-PIs who wish to use them.

2. **Convex value-of-information criteria for mission planning**. The UM (Hero) and MIT (How) groups have been collaborating on adapting UM's convex optimization approach to mission-aware planning problems being pursued in How's laboratory. The advantage of UMs convex optimization approaches is that they are easy to analyze, come with performance guarantees and yield tractable optimal planning strategies. To date, the UM convex opti-

mization approach has been limited to target localization and tracking tasks. Together with MIT graduate student Beipeng Mu, Hero and his students (Newstadt) have extended the convex approach to the mission-aware classification problem where identification of different targets have different associated value. Over the next year we will be performing simulations and doing a multi-agent experiment to demonstrate this approach and its ability to optimize value of information.

3. **Information geometric value of information for sensor planning**. UM (Hero) and ASU (Cochran) have been independently looking at formulation of value-of-information for sensor planning using information geometry. The advantage of information geometric theory of VoI is that it may provide the same simple design rules that come with other geometric theory, such as projection theorem of associated with optimal Wiener filtering, Kalman estimation, and matched filtering for detection. Two different but complementary theories of information geometry for sensor planning have emerged – one a dynamic theory based on adaptive filtrations and the other a static theory based on the geometry of sufficient statistics (probability distributions and likelihood ratios). Brian Sadler at ARL has also been involved in these discussions. We anticipate making progress on bridging these two theories over the next year.

4. **Improving information collection via human-in-the-loop processing**. Several on the team have been studying aspects of human-in-the-loop processing in the context of adaptive sensing (Hero), human-machine interfaces (Hero, Jordan), and interacting multi-agent learning systems (Yu). The field of human-factors in information theory and adaptive learning is early in its gestation and most of our activities have been on the theoretical side to date (see Progress descriptions above). Three collaborative efforts have been discussed and may be pursued in the coming year: inference of preference graphs from incomplete human preference data (Jordan and Hero) and competitive foraging using POMDP and approximate POMDP strategies (Yu and Hero). In addition, Hero and Sadler (ARL) have been discussing doing a human subjects experiment to test a UM Pareto image database indexing and retrieval method. While this Pareto method was developed by Hero and his students under a different ARO grant on database indexing and retrieval (W911NF-09-1-0210) this human experiment will generate data and test human-machine interaction models relevant to our human-in-the-loop processing efforts on this MURI grant.

5. **Empirical non-parametric estimation of Fisher information**. In a collaboration between Hero and Ertin, we have formulated a method for estimation of Fisher information by estimation of locally perturbed f-divergences. The Fisher information is a key quantity that captures the value of information for estimation of parameters in a likelihood model; e.g., parameters determining target state dynamics. The D-optimal experimental design approach uses Fisher information to optimize sensor actions for maximizing parameter estimation accuracy. By empirically estimating the Fisher information one can reduce sensitivity to model mismatch and model drift. Over the next year we will investigate f-divergence estimation strategies in the context of Fisher information.

# 7 Publications

The following is a list of papers, theses, and other publications of research that have been cited in the above narrative.

[1] M. Alighanbari and J. P. How, "A robust approach to the UAV task assignment problem," *International Journal of Robust and Nonlinear Control*, vol. 18, no. 2, pp. 118–134, January 2008.

[2] L. Arienzo, "An information-theoretic approach for energy-efficient collaborative tracking in wireless sensor networks," *EURASIP J. Wirel. Commun. Netw.*, vol. 2010, pp. 13:1–13:14, January 2010.

[3] K. Beaudet, L. Crider, and D. Cochran, "Detection in networked radar," in *Proceedings of the SPIE Defense, Security, and Sensing Conference*, May 2013.

[4] L. Bertuccelli and J. How, "Active exploration in robust unmanned vehicle task assignment," *AIAA Journal of Aerospace Computing, Information, and Communication*, vol. 8, pp. 250–268, 2011.

[5] P. Bodik, W. Hong, C. Guestrin, S. Madden, M. Paskin, and R. Thibaux, "Intel lab data," Technical report, February 2004.

[6] R. Cabezas, *Aerial Reconstructions via Probabilistic Data Fusion*, Master's thesis, Massachusetts Institute of Technology, 2013.

[7] V. Chandrasekaran and M. I. Jordan, "Computational and statistical tradeoffs via convex relaxation," *Proceedings of the National Academy of Sciences*, vol. 110, pp. E1181–E1190, 2013.

[8] J. Duchi, M. I. Jordan, and M. Wainwright, "Local privacy and statistical minimax rates," in *54th Annual Symposium on Foundations of Computer Science (FOCS)*, Berkeley, CA, 2013.

[9] E. Ertin, "Manifold learning methods for wide-angle SAR ATR," in *Proceedings of International Conference on Radar*, Adelaide, September 2013.

[10] B. Farrell and R. Nadakuditi, "Local spectrum of truncations of Kronecker products of haar distributed unitary matrices,", 2013. Under review at the Electronic Communications in Probability.

[11] C. S. Fraser, L. F. Bertuccelli, H.-L. Choi, and J. P. How, "A hyperparameter consensus method for agreement under uncertainty," *Automatica*, vol. 48, no. 2, pp. 374–380, February 2012.

[12] V. Karasev, A. Chiuso, and S. Soatto, "Controlled recognition bounds for visual learning and exploration," in *Proc. of NIPS*, Dec. 2012.

[13] C. Kreucher, M. Morelande, K. Kastella, and A. Hero, "Particle filtering for multitarget detection and tracking," in *IEEE Aerospace Conference*, pp. 2101 –2116, march 2005.

[14] A. Logothetis, A. Isaksson, and R. Evans, "An information theoretic approach to observer path design for bearings-only tracking," in *Decision and Control, 1997., Proceedings of the 36th IEEE Conference on*, volume 4, pp. 3132 –3137 vol.4, dec 1997.

[15] W. M. McEneaney and B. Fitzpatrick, "Control for UAV operations under imperfect information," *Proc. 1st AIAA UAV Symposium*, 2002.

[16] Z. Meng, A. Wiesel, and A. Hero, "Distributed principal component analysis on networks via directed graphical models," in *IEEE Intl Conf. on Speech, Acoustics and Signal Processing (ICASSP)*, Prague, April 2012.

[17] Z. Meng, D. Wei, A. Wiesel, and A. Hero III, "Distributed learning of Gaussian graphical models via marginal likelihoods," in *AISTATS*, 2013. arXiv preprint arXiv:1303.4756.

[18] W. Moran, S. D. Howard, D. Cochran, and S. Suvorova, "Sensor management via Riemannian geometry," in *Proceedings of the Allerton Conference on Communications, Control, and Computing*, pp. 358–362, October 2012.

[19] B. Mu, *Value of Information Based Distributed Inference and Planning*, Master's thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge MA, June 2013.

[20] B. Mu, G. Chowdhary, and J. P. How, "Efficient distributed information fusion using value of information based censoring," Aerospace control laboratory technical report, Massachusetts Institute of Technology, http://hdl.handle.net/1721.1/71875, July 2012.

[21] B. Mu, G. Chowdhary, and J. P. How, "Efficient distributed information fusion using adaptive decentralized censoring algorithms," in *American Control Conference (ACC)*, Washington DC, July 2013, IEEE.

[22] B. Mu, G. Chowdhary, and J. P. How, "Value-of-information aware active task assignment," in *SPIE Defense Security and Sensing*, Baltimore, May 2013, SPIE.

[23] B. Mu, G. Chowdhary, and J. P. How, "Efficient distributed inference using adaptive censoring algorithms," *Automatica*, 2013 (provisionally accepted).

[24] B. Mu, G. Chowdhary, and J. P. How, "Value-of-information aware active task assignment," in *AIAA Guidance Navigation and Control*, Boston, August 2013 (to appear).

[25] R. R. Nadakuditi, "Optshrink: An algorithm for improved low-rank signal matrix denoising by optimal, data-driven singular value shrinkage," *arXiv:1306.6042*, 2013.

[26] R. Nadakuditi, "When are the most informative components for inference also the principal components?," *Arxiv preprint arxiv.org:1302.1232.*

[27] S. Olver and R. Nadakuditi, "Numerical computation of convolutions in free probability theory," *arXiv:1203.1958*, 2012.

[28] G. Papachristoudis and J. W. Fisher III, "Efficient information planning in Markov chains," Submitted for publication.

[29] G. Papachristoudis and J. W. Fisher III, "Incremental belief propagation," Submitted for publication.

[30] A. Ravichandran and S. Soatto, "Modeling stationarity at multiple time-scales for applications in fire detection and traffic monitoring," *Proc. of the Eur. Conf. on Comp. Vision*, 2012.

[31] A. Ravichandran, C. Wang, and S. Soatto, "Superfloxels: A mid-level representation for video sequences," in *Proc. of ARTEMIS*, October 2012.

[32] V. Saligrama and M. Alanyali, "A token-based approach for distributed computation in sensor networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 4, pp. 817 –832, August 2011.

[33] N. Sugavanam and E. Ertin, "Sensor selection and placement in adversarial environment," in *IEEE Global Conference on Signal and Information Processing (GLOBALSIP)*, Austin, TX, Dec 2013. (Submitted).

[34] D. Teng and E. Ertin, "Optimal quantization of likelihood for low complexity sequential testingsensor selection and placement in adversarial environment," in *IEEE Global Conference on Signal and Information Processing (GLOBALSIP)*, Austin, TX, Dec 2013. (Submitted).

[35] T. Tsiligkaridis and A. Hero, "Sparse covariance estimation under sparse Kronecker product structure," in *IEEE Intl Conf. on Speech, Acoustics and Signal Processing (ICASSP)*, Prague, April 2012.

[36] T. Tsiligkaridis and A. Hero, "A collaborative 20 questions model for target search with human-machine interaction," *arXiv:1302.5828*, June 2013.

[37] T. Tsiligkaridis and A. Hero, "Low separation rank covariance estimation using Kronecker product expansions," in *Proc. of IEEE ISIT*, Istanbul, 2013, IEEE.

[38] T. Tsiligkaridis, A. Hero, and S. Zhou, "Convergence properties of Kronecker graphical lasso algorithms," in *Proceedings of the IEEE Statistical Signal Processing Workshop*, Ann Arbor MI, August 2012.

[39] T. Tsiligkaridis, B. Sadler, and A. Hero, "A collaborative 20 questions model for target search with human-machine interaction," in *Proc. of IEEE ICASSP*, Vancouver, 2013, IEEE.

[40] T. Tsiligkaridis, A. Hero, and S. Zhou, "On convergence of Kronecker graphical lasso algorithms," *IEEE Trans. on Signal Processing*, vol. 61, no. 9, pp. 1743 –1755, 2013.

[41] T. Tsiligkaridis and A. O. Hero III, "Covariance estimation in high dimensions via Kronecker product expansions," *IEEE Trans. on Signal Processing*, p. To appear, 2013. arXiv preprint arXiv:1302.2686.

[42] H. Wang, K. Yao, and D. Estrin, "Information-theoretic approaches for sensor selection and placement in sensor networks for target localization and tracking," *Communications and Networks, Journal of*, vol. 7, no. 4, pp. 438 –449, dec. 2005.

[43] F. Wauthier, M. I. Jordan, and N. Jojic, "Efficient ranking from pairwise comparisons," in *International Conference on Machine Learning (ICML)*, New York, 2013, ACM Press.

[44] D. Wei and A. O. Hero, "Adaptive spectrum sensing and estimation," in *IEEE ICASSP*, 2012.

[45] D. Wei and A. O. Hero, "Multistage adaptive estimation of sparse signals," in *Statistical Signal Processing Workshop (SSP), 2012 IEEE*, pp. 153–156. IEEE, 2012.

[46] D. Wei and A. O. Hero, "Multistage adaptive estimation of sparse signals," *IEEE Trans on Signal Processing*, To appear 2013. Available as arXiv 1210.1473.

[47] G. T. Whipps, E. Ertin, and R. L. Moses, "Distributed detection with collisions in a random, single-hop wireless sensor network," in *Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vancouver, May 2013.

[48] Y. Zeng, C. W. adn S. Soatto, and S.-T. Yau, "Nonlinearly constrained MRFs: Exploring the intrinsic dimensions of higher-order cliques," in *IEEE Intl. Conf. on Comp. Vis and Patt. Recog.*, June 2013.

[49] F. Zhao, J. Shin, and J. Reich, "Information-driven dynamic sensor collaboration," *Signal Processing Magazine, IEEE*, vol. 19, no. 2, pp. 61 –72, mar 2002.

# 8 Statistics

1. Submissions or publications under ARO sponsorship during this reporting period. List the title of each and give the total number for each of the following categories:

   a. Papers published in peer-reviewed journals

      1. V. Chandrasekaran and M. I. Jordan, "Computational and statistical tradeoffs via convex relaxation," *Proceedings of the National Academy of Sciences*, 110, E1181–E1190, 2013.
      2. J. Duchi, L. Mackey, and M. I. Jordan, "The asymptotics of ranking algorithms," *Annals of Statistics* (in press).
      3. M. I. Jordan, "On statistics, computation and scalability," *Bernoulli* (in press).
      4. A. Kleiner, A. Talwalkar, P. Sarkar, and M. I. Jordan, "A scalable bootstrap for massive data," *Journal of the Royal Statistical Society, Series B* (in press).
      5. B. Mu, G. Chowdhary, and J. P. How, "Efficient distributed inference using adaptive censoring algorithms," *Automatica*, 2013 (provisionally accepted).
      6. T. Tsiligkaridis and A. O. Hero III, "Covariance estimation in high dimensions via Kronecker product expansions," *IEEE Trans. on Signal Processing*, 2013 (in press). arXiv preprint arXiv:1302.2686.
      7. T. Tsiligkaridis, A. Hero, and S. Zhou, "On convergence of Kronecker graphical lasso algorithms," *IEEE Trans. on Signal Processing*, vol. 61, no. 9, pp. 1743–1755, 2013.
      8. D. Wei and A. O. Hero, "Multistage adaptive estimation of sparse signals," to appear in *IEEE Journal of Selected Topics in Signal Processing*, 2013.

   b. Papers published in non-peer-reviewed journals

      1. R. R. Nadakuditi, "When are the most informative components for inference also the principal components?," arXiv:1302.1232 2013.

2. B. Farrell and R. R. Nadakuditi, "Local spectrum of truncations of Kronecker products of Haar distributed unitary matrices," UM Technical Report June 2013. Submitted to *Electronic Communications in Probability.*
3. T. Tsiligkaridis and A. Hero, "A collaborative 20 questions model for target search with human-machine interaction," (submitted for publication at IEEE Trans. on Information Theory). Available as *arXiv:1302.5828*, June 2013.

c. Presentations

i. Presentations at meetings, but not published in Conference Proceedings

1. S. Ahmad and A. J. Yu, "Active sensing as Bayes-optimal sequential decision-making," *Twenty-ninth Annual Conference on Uncertainty in Artificial Intelligence*, 2013.
2. D. Cochran, "An operator-theoretic view of radar," *SIAM Annual Meeting*, San Diego, CA, July 2013
3. E. Ertin, "Geometry of SAR Imagery," *SIAM Annual Meeting*, San Diego, CA, July 2013
4. A. Hero, "High dimensional Kronecker product correlation models," seminar presented at ARL during Feb 2013 visit.
5. A. Hero, "Value-centered information theory for adaptive learning, inference, tracking, and exploitation," MURI overview presented at ARL during Feb 2013 visit.
6. A. Hero, "Winnowing signals from massive data: SP for Big Data and its Relation to Systems Engineering," invited presentation at the NSF Big Data Symposium April 2013, Washington DC.
7. A. Hero, "Cooperative man-machine system modeling", invited presentation at the SILO meeting in June 2013, Univ of Wisconsin.
8. A. Hero, "Kronecker covariance decompositions for high dimensional data," invited presentation at the CIMI Colloquium, July 2013, Univ. of Toulouse, France.
9. A. Hero, "Value-centered information theory for adaptive learning, inference, tracking, and exploitation," presented at the OSD MURI Program Review, July 2013, Washington DC.
10. J. How, "Distributed Information Fusion Under Uncertainty and Communication Constraints," WISeNet Seminar at Duke University, April 2013.
11. M. Jordan, Invited Speaker, SAMSI Workshop on Massive Data Analysis, 9/9/12.
12. M. Jordan, Invited Speaker, Méthodes Bayésiennes non Paramétriques pour le Traitement du Signal et des Images, Telecom ParisTech, Paris, France, 9/8/12.
13. M. Jordan, Invited Speaker, Seminaire Parisiens de Statistique, Paris, France, 9/17/12.
14. M. Jordan, Invited Speaker, Workshop on Random Matrices and their Applications, Paris, France, 10/9/12.
15. M. Jordan, Invited Speaker Colloquium, Department of Informatique, Ecole Normale Superieure, 10/2/12.

16. M. Jordan, Invited Speaker, Workshop on Optimization and Statistical Learning, Les Houches, France, 1/8/13.
17. M. Jordan, Invited Speaker, Simons Workshop on Big Data, New York, 1/24/13.
18. M. Jordan, Invited Speaker, Lecture Series, Ecole Nationale de la Statistique et de l'Administration, Paris, 5/13.
19. M. Jordan, Distinguished Lecture, Department of Statistics, University of Oxford, 5/7/13.
20. M. Jordan, Invited Speaker, Colloquium, Department of Statistics, University of Cambridge, 5/10/13.
21. M. Jordan, Invited Speaker, GdR ISIS Conference, Telecom ParisTech, Paris 5/16/13.
22. M. Jordan, Invited Speaker, Workshop on High-Dimensional Statistics, Moscow, 6/26/13.
23. P. Shenoy and A. J. Yu, "A rational account of contextual effects in preference choice: What makes for a bargain?," *Cognitive Science Society Conference*, 2013.
24. S. Zhang and A. J. Yu, "Cheap but clever: Human active learning in a bandit setting," *Cognitive Science Society Conference*, 2013. R. Moses, Plenary speaker at IEEE Workshop on Statistical Signal Processing, 8/12.

ii. Non-Peer-Reviewed Conference Proceeding publications (other than abstracts)

iii. Peer-Reviewed Conference Proceeding publications (other than abstracts)

1. W. Moran, S. D. Howard, D. Cochran, and S. Suvorova, "Sensor management via Riemannian geometry," in *Proceedings of the Allerton Conference on Communications, Control, and Computing*, pp. 358–362, October 2012.
2. K. Beaudet and D. Cochran, Multiple-channel detection in active sensing, *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, April 2013.
3. K. Beaudet, L. Crider, and D. Cochran, Detection in networked radar, *Proceedings of the SPIE Defense, Security, and Sensing Conference*, May 2013.
4. J. Duchi, M. I. Jordan, and M. Wainwright, "Local privacy and statistical minimax rates," *54th Annual Symposium on Foundations of Computer Science (FOCS)*, Berkeley CA, 2013.
5. J. Duchi, M. I. Jordan, and M. Wainwright, "Privacy aware learning," in P. Bartlett, F. Pereira, L. Bottou and C. Burges (Eds.), *Advances in Neural Information Processing (NIPS)*, 25, Red Hook, NY: Curran Associates.
6. E. Ertin, "Manifold learning methods for wide-angle SAR ATR," *Proceedings of the International Conference on Radar*, September 2013 (to appear).
7. J. Gao and E. Ertin, "Contactless sensing of physiological signals using wideband RF probes," *Proceedings of the Forty Sixth Asilomar Conference on Signals, Systems and Computers*, November 2013 (to appear).
8. V. Karasev, A. Chiuso, and S. Soatto, "Controlled recognition bounds for visual learning and exploration," *Proceedings of NIPS*, December 2012.
9. A. Kleiner, A. Talwalkar, S. Agarwal, M. I. Jordan, and I. Stoica, "A general bootstrap performance diagnostic," *Proc. 19th ACM Conference on Knowledge Discovery and Data Mining (SIGKDD)*, Chicago, 2013.

10. Z. Meng, D. Wei, A. Wiesel, and A. Hero III, "Distributed learning of Gaussian graphical models via marginal likelihoods," *AISTATS*, 2013. arXiv preprint arXiv:1303.4756.(Notable Paper Awardee).
11. B. Mu, G. Chowdhary, and J. P. How, "Efficient distributed information fusion using adaptive decentralized censoring algorithms," *Proceedings of the American Control Conference*, Washington DC, July 2013.
12. B. Mu, G. Chowdhary, and J. P. How, "Value-of-information aware active task assignment," *SPIE Defense, Security and Sensing Symposium*, Baltimore, May 2013.
13. B. Mu, G. Chowdhary, and J. P. How, "Value-of-information aware active task assignment," *Proc. AIAA Guidance Navigation and Control Conference*, Boston, August 2013.
14. G. Papachristoudis and J. W. Fisher III, "Efficient information planning in Markov chains," (submitted for publication).
15. G. Papachristoudis and J. W. Fisher III, "Incremental belief propagation," (submitted for publication).
16. A. Ravichandran and S. Soatto, "Modeling stationarity at multiple time-scales for applications in fire detection and traffic monitoring," *Proc. of the Eur. Conf. on Comp. Vision*, 2012.
17. A. Ravichandran, C. Wang, and S. Soatto, "Superfloxels: A mid-level representation for video sequences," *Proc. of ARTEMIS*, October 2012.
18. M. R. Riedl, L. C. Potter, and E. Ertin, "Augmenting synthetic aperture radar with space time adaptive processing," *Proceedings of SPIE Conference on Algorithms for SAR*, vol. 8746 (Baltimore, MD), May 2013.
19. C. Rossler, E. Ertin and R. L. Moses, "Waveform diversity and optimal change detection," *Proceedings of the Forty Sixth Asilomar Conference on Signals, Systems and Computers*, November 2012.
20. N. Sugavanam and E. Ertin, "Sensor selection and placement in adversarial environment," *IEEE Global Conference on Signal and Information Processing (GLOBALSIP)*, December 2013 (to appear).
21. D. Teng and E. Ertin, "Optimal quantization of likelihood for low complexity sequential testing," *IEEE Global Conference on Signal and Information Processing (GLOBALSIP)*, December 2013 (to appear).
22. T. Tsiligkaridis, A. Hero, and S. Zhou, "Convergence properties of Kronecker graphical lasso algorithms," in *Proceedings of the IEEE Statistical Signal Processing Workshop*, Ann Arbor MI, August 2012.
23. T. Tsiligkaridis and A. Hero, "Sparse covariance estimation under sparse Kronecker product structure," in *IEEE Intl Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Prague, April 2012.
24. T. Tsiligkaridis, B. Sadler, and A. Hero, "A collaborative 20 questions model for target search with human-machine interaction," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2013.

25. T. Tsiligkaridis and A. Hero, "Low separation rank covariance estimation using Kronecker product expansions," *Proc. IEEE Intl Symposium on Information Theory (ISIT)*, Istanbul, 2013.
26. F. Wauthier, M. I. Jordan, and N. Jojic, "Efficient ranking from pairwise comparisons," in S. Dasgupta and D. McAllester (Eds.), *International Conference on Machine Learning (ICML)*, New York: ACM Press.
27. D. Wei and A. O. Hero, "Adaptive spectrum sensing and estimation," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2013.
28. G. Whipps, E. Ertin, and R. L. Moses, "Distributed detection with collisions in a random, single-hop wireless sensor network," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2013.
29. Y. Zeng, C. Wang, S. Soatto, and S.-T. Yau, "Nonlinearly constrained MRFs: Exploring the intrinsic dimensions of higher-order cliques," in *IEEE Intl. Conf. on Comp. Vis and Patt. Recog.*, June 2013.

d. Manuscripts

1. R. Cabezas, *Aerial Reconstructions via Probabilistic Data Fusion*, Master's thesis, Massachusetts Institute of Technology, 2013.
2. B. Mu, *Value of information based distributed inference and planning*, Masters thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge MA, June 2013

e. Books

1. E. Ertin, "Three Dimensional Imaging of Vehicles from Sparse Apertures in Urban Environment," (In review) Book Chapter in *Compressive Sensing for Urban Radars*, edited by Moeness Amin, CRC Press, 2013.

f. Honors and Awards

1. Co-PI Hero and his co-authors received a Notable Paper Award from the AISTATS conference for the paper: Z. Meng, D. Wei, A. Wiesel, and A. Hero III, "Distributed learning of Gaussian graphical models via marginal likelihoods," in *AISTATS*, 2013. arXiv preprint arXiv:1303.4756 (work funded by this MURI).
2. Co-PI Hero was awarded a Distinguished Lectureship at Wayne State University in February 2013.
3. Co-PI Hero was invited to give keynote "Small sample community detection in massive data sets," IEEE CAMSAP Workshop, Dec 2013.
4. Co-PI Hero was invited to give keynote "Resource constrained adaptive sensing," New Sensing and Statistical Inference Methods Symposium, IEEE GlobalSIP Conference, Dec 2013.
5. Co-PI Hero was invited to give keynote "Spatio-temporal graphical models for high dimensional network data," Network Theory Symposium, IEEE GlobalSIP Conference, Dec 2013.
6. Co-PI Hero was invited to give keynote "Information-driven multimodality fusion," AFRL/UES Workshop on Data Fusion for the Detection of Rare and Anomalous Events, Dec 2012.

7. Co-PI Hero was invited to give keynote "Graphical modeling for high dimensional data analysis," (Keynote 1) Solar Information Processing Workshop (SIP) VI, MSU Bozeman, Aug. 14 2012.
8. Co-PI Hero was invited to give keynote "Learning with entropic graphs," (Keynote 2) Solar Information Processing Workshop (SIP) VI, MSU Bozeman, Aug. 15 2012.
9. Co-PI Jordan was awarded a Chaire d'Excellence, Fondation Sciences Mathématiques de Paris
10. Co-PI Jordan was made Fellow, Society for Industrial and Applied Mathematics (SIAM)
11. Co-PI Jordan was Elected Member, International Statistical Institute (ISI)
12. Co-PI Jordan gave the Harry Nyquist Distinguished Lecture, Yale University
13. Co-PI Jordan gave the Vincent Meyer Colloquium, Israel Institute of Technology
14. Co-PI Jordan was Keynote Speaker, Bayesian Nonparametrics Workshop, Amsterdam, 6/10/13
15. Co-PI Jordan was Keynote Speaker, Workshop on Nonsmooth Optimization in Machine Learning, Liege, Belgium, 3/4/13
16. Co-PI Jordan was Keynote Speaker, StatLearn Workshop, Bordeaux, France, 4/8/13
17. Co-PI Jordan was Keynote Speaker, ACM Conference on Knowledge Discovery and Data Mining (SIGKDD), Beijing, China, 8/15/12
18. Co-PI Jordan was Keynote Speaker, 21st Century Computing Conference, Tianjin, China, 10/25/12
19. Co-PI Jordan was Keynote Speaker, ICONIP, Doha, Qatar, 11/12/12
20. Co-PI Jordan was Keynote Speaker, Amazon Machine Learning Conference, Seattle, 4/28/13
21. Co-PI Moses was Plenary Speaker, IEEE Statistical Signal Processing Workshop, Ann Arbor, 8/12

g. Title of Patents Disclosed during the reporting period

h. Patents Awarded during the reporting period

2. Student/Supported Personnel Metrics for this Reporting Period

   a. Graduate Students

      1. Doctoral students
         (a) OSU student Nithin Sugavanatham supported at 50% annualized FTE
         (b) OSU student Siddarth Baskar supported at 50% annualized FTE
         (c) OSU student Gene Whipps supported at 0% annualized FTE (Whipps is a researcher at ARL on temporary assignment to OSU to complete his PhD relevant to this MURI)
         (d) MIT student Georgios Papachristoudis supported at 50% FTE
         (e) UM student Pin-Yu Chen supported at 33% FTE
         (f) UM student Hamed Firouzi supported at 8.33% FTE
         (g) UM student Theodoros Tsiligkaridis supported at 37.5% FTE

(h) UM student Tianpei Xie supported at 50% FTE
(i) UM student Nick Azendorf supported at 16.5% FTE
(j) UCLA student Tai-Hee Lee supported at 15% FTE.
(k) UCLA student Georgios Georgiadis supported at 15% FTE.
(l) UCLA student Nikos Karianakis supported at 7.5% FTE.
(m) UCLA student Jingming Dong supported at 7.5% FTE.
(n) UCLA student Jonathon Shih supported at 7.5% FTE.
(o) UC Berkeley student Nicholas Boyd supported May at 29.13% and June at 67%
(p) UC Berkeley student Fabian Wauthier supported May at 29.13% and June at 67%

2. Masters students
   (a) ASU student Kaitlyn Beaudet, 25% FTE
   (b) MIT student Beipeng Mu supported at 50% FTE
   (c) MIT student Randi Cabezas 37% FTE
   (d) MIT student Sue Zheng supported at 12.5% FTE

b. Post Doctorates
   1. UM postdoc Dennis Wei supported at 50% annualized FTE
   2. MIT postdoc Girish Chowdhary supported at 15% annualized FTE
   3. UCLA postdoc Wang Chao-Hui supported at 7.5% annualized FTE
   4. UCSD postdoc Shunan Zhang supported at 10% annualized FTE

c. Faculty
   1. ASU faculty and co-PI Cochran supported at 12% annualized FTE
   2. OSU faculty and co-PI Ertin supported at 20% annualized FTE
   3. OSU faculty and co-PI Moses supported at 0% annualized FTE
   4. MIT faculty and co-PI Fisher supported at 10% annualized FTE
   5. MIT faculty and co-PI How supported at 10% annualized FTE
   6. UM faculty and co-PI Hero supported at 10% annualized FTE
   7. UCBerkeley faculty and co-PI Jordan supported at 22% annualized FTE
   8. UM faculty and co-PI Nadakuditi supported at 10% annualized FTE
   9. UCLA faculty and co-PI Soatto supported at 10% annualized FTE
   10. UCSD faculty and co-PI Yu supported at 0% annualized FTE

d. Undergraduate Students
   1. OSU Student Christopher Dean 10% FTE
   2. ASU student Lauren Crider, 15% FTE

e. Graduating Undergraduate Metrics (funded by this agreement and graduating during this reporting period):
   i. Number who graduated during this period: 2
   ii. Number who graduated during this period with a degree in science, mathematics, engineering, or technology fields: 2

iii. Number who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields: 2
iv. Number who achieved a 3.5 GPA to 4.0 (4.0 max scale): 2
v. Number funded by a DoD funded Center of Excellence grant for Education, Research and Engineering
vi. Number who intend to work for the Department of Defense
vii. Number who will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields: 2

f. Masters Degrees Awarded (Name of each, Total #)

* MIT student R. Cabezas, *Aerial Reconstructions via Probabilistic Data Fusion*, Master's thesis, Massachusetts Institute of Technology, 2013.
* MIT student B. Mu, *Value of information based distributed inference and planning*, Masters thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge MA, June 2013

g. Ph.D.s Awarded (Name of each, Total #)

1. UC Berkeley student Fabian Wauthier

h. Other Research staff (Name of each, FTE)

3. Technology transfer (any specific interactions or developments which would constitute technology transfer of the research results). Examples include patents, initiation of a start-up company based on research results, interactions with industry/Army R&D Laboratories or transfer of information which might impact the development of products.

1. Technology transitions
   (a) Co-PI Hero and student Mark Hsiao have transitioned Pareto web image search engine to ARL for evaluation and adaptation into an in-house interface (ARL POC: Brian Sadler).
   (b) Co-PI Fisher has transferred implementations of information planning algorithms to DARPA sub-contractor (Systems Technology Research) under the All Source Positioning and Navigation (ASPN) program. The method is being adapted to sensor planning and anomaly detection. STR POC: Joel Douglas.
   (c) Co-PI Hero and student Kristjan Greenwald worked on transitioning Kronecker sum decompositions to video sources at AFRL over the summer 2013.
2. Student interns at Service Labs
   (a) UM doctoral student Ted Tsiligkardis spent six weeks at ARL during Summer 2013, supported by this award.
   (b) UM doctoral student Brandon Oselio spent 3 months at ARL during Summer 2013, supported by this award.
   (c) UM masters student Kristjan Greenwald spend two and a half months at AFRL during Summer 2013.
   (d) ASU students Lauren Crider and Shih-Ling Phuong each spent 10 weeks at AFRL during Summer 2013.

   (e) ASU student Kaitlyn Beaudet spent two months at the Australian DSTO during Summer 2013.
   (f) OSU student Gene Whipps spent summer 2013 at ARL.

3. Co-PI interactions with ARL and other Federal research labs
   (a) Co-PIs Cochran, Ertin and Hero visited ARL in February 2013 and met with ARL researchers N. Nasrabadi, B. Sadler, L. Kaplan, T. Pham, and others.
   (b) Co-PI Moses served on the Technical Advisory Board for the Computational and Information Sciences Directorate of the U.S. Army Research Laboratory.
   (c) Co-PI Fisher visited ARL in July 2012 to discuss further collaborations with ARL researchers B. Sadler & L. Kaplan.
   (d) Co-PI Fisher visited MassPort (managers of Logan Airport and other transportation infrastructure in the Boston Area) to discuss collaboration for information management in their security systems in August 2012
4. Other relevant co-PI activities on national committees
   (a) Co-PI Hero served on the National Academy of Sciences Committee on Applied and Theoretical Statistics (CATS), 2012-.
   (b) Co-PI Jordan chaired the National Academy of Sciences Committee on Frontiers in Massive Data Analysis, 2012-2013.
   (c) Co-PI Moses served on the National Academy of Science Committee on Science and Technology for Defense Warning, 2012-
5. Internal MURI visits
   (a) Co-PI Fisher, Doctoral students Papachristoudis, Zheng visited co-PI Hero's group in June 2012 to discuss project collaboration and share tenchnical results.
   (b) Co-PIs Cochran (ASU) and Ertin (OSU) visited co-PIs Hero and Nadakuditi (UM) in July.
   (c) Co-PI Jordan's (UCB) student John Duchi visited UM and co-PI Hero's (UM) student Jeff Calder visited UCB in April and May.
   (d) Co-PI How's (MIT) student Beipeng Mu visited UM and co-PI Hero's (UM) student Greg Newstadt visited MIT in June and July.