Decentralized Learning of a Data Manifold Over a Sensor Network

**Objectives**
- Learn object structure from multiple, distributed partial views for detection and classification applications
- Estimate an underlying low-dimensional data manifold of high-dimensional observations from a spatially-distributed network of sensors; disseminate learned structure to all nodes
- Develop decentralized algorithms which are scalable and robust to sensor node or communication link failure

**Challenges**
- Current, centralized learning approaches have poor scalability properties; not robust to node or link outages
- Disseminating manifold representation out to nodes requires additional communications

**Approach**
- Develop robust algorithm that learns a low-dimensional manifold across a network and simultaneously disseminates this representation to all nodes
- Centralized models: LLE [2], Isomap [3], MFA [4]
- Model the nonlinear data manifold with a mixture of factor analyzers (MFA)
- Derive a consensus-based Expectation-Maximization (EM) algorithm from distributed measurements
- Unlike [2,3,4], our model is a soft assignment of linear mappings to each sensor node. Our algorithm is similar to [5,6]; in contrast, our work incorporates a low-dimensional structure - key to accurately modeling high-dimensional data with an intrinsic manifold structure

**The MFA Model**

\[
x_{m,i} \sim \sum_{j=1}^{J} \alpha_{m,j} N \left( \mathbf{x} \mid \mathbf{y}_j + \Lambda_j \mathbf{y}_j \right), \quad x_j = \Lambda_j y_j + \mu_j + \epsilon_j
\]

**ARL Contribution**
- Internal research with support from doctoral advisors
- Synergistic with SEDD research in automated target detection and classification from dimension-reduced features from ground-based imaging systems
- This work is in partial fulfillment of requirements toward a doctorate degree for Mr. Whipps

**Impact**
- Algorithms that are scalable and robust to node and link failures are critical features for Army ISR missions
- Fundamental understanding of data modeling across a network and properties of decentralized algorithms
- Contributes to research in dimensionality reduction, manifold learning, and reconstruction with emphasis on distributed learning systems

**Results**
- Developed a decentralized algorithm for learning and simultaneously disseminating MFA model parameters
- Communications scale linearly in the intrinsic dimensionality of the data
- Local computations scale cubically in low-dimension instead of cubically in high-dimension
- Latest results submitted to 2014 IEEE Int. Workshop on Machine Learning for Signal Processing

**Path Forward**
- Analyze convergence properties of overall algorithm and investigate speedup strategies
- Automate learning of model order and dimensionality

---

**References**
Decentralized Learning of a Data Manifold Over a Sensor Network

We consider the problem of decentralized learning of a target appearance manifold using a network of sensors. Sensor nodes observe an object from different aspects and then, in an unsupervised and distributed manner, learn a joint statistical model for the data manifold. We employ a mixture of factor analyzers (MFA) model, approximating a potentially nonlinear manifold. We derive a consensus-based decentralized expectation maximization (EM) algorithm for learning the parameters of the mixture densities and mixing probabilities.

A spatially distributed sensor network can be used to construct a rich appearance model for targets in their common field-of-view. These models can then be used to identify previously seen objects if they reappear in the network at a later time. As an example, consider a network of cameras capturing images of an object from different but possibly overlapping aspects as the object traverses through the network’s field of view. The ensemble of images captured by the network forms a low-dimensional nonlinear manifold in the high-dimensional ambient space of images. One approach to appearance modeling would be to construct independent models of a local data manifold at each sensor and share it across the network. However, such an ensemble of models suffers from discretization of the aspect space and poor parameter estimates as the number of unknown parameters necessarily scale linearly with the number of sensor nodes. Alternatively, the sensor nodes can collaborate to construct a joint model for the image ensemble. The parameter estimates of the joint model will improve with the number of sensor nodes, since the number of unknown parameters in the model is intrinsic to the object and fixed, whereas the measurements scale linearly with the number of sensor nodes. The straightforward method of pooling images to a central location for joint model construction will require large and likely impractical network bandwidth. In this paper, we develop a decentralized learning method with greatly reduced data bandwidth need and results in a global appearance manifold model shared by all sensor nodes.

Objectives
- Learn object structure from multiple, distributed partial views for detection and classification applications
- Estimate an underlying low-dimensional data manifold of high-dimensional observations from a spatially-distributed network of sensors; disseminate learned structure to all nodes
- Develop decentralized algorithms which are scalable and robust to sensor node or communication link failure

Challenges
- Current, centralized learning approaches have poor scalability properties; not robust to node or link outages
- Disseminating manifold representation out to nodes requires additional communications

Approach
- Develop robust algorithm that learns a low-dimensional manifold across a network and simultaneously disseminates this representation to all nodes
- Model the nonlinear data manifold with a mixture of factor analyzers (MFA)
  - Parametric statistical model of mixture densities
  - Each node observes the model with different mixture probabilities
- Derive a consensus-based Expectation-Maximization (EM) algorithm from distributed measurements
  - Each node calculates local statistics from its own observations
  - Locally estimate global parameters from network averaged statistics

ARL Contribution
- Internal research with support from doctoral advisors
- This work is in partial fulfillment of requirements toward a doctorate degree for Mr. Whipps
- Synergistic with SEDD research in automated target detection and classification from dimension-reduced features from ground-based imaging systems

Collaborations (doctoral advisors)
- Prof. Randolph Moses, Dept. of ECE, The Ohio State University, moses.2@osu.edu
- Prof. Emre Ertin, Dept. of ECE, The Ohio State University, ertin.1@osu.edu

Impact
- Algorithms that are scalable and robust to node and link failures are critical features for Army ISR missions
- Fundamental understanding of data modeling across a network and properties of decentralized algorithms
- Contributes to research in dimensionality reduction, manifold learning, and reconstruction with emphasis on distributed learning systems
- Contributes to Distributed Information Fusion thrust within ARO MURI on Value-centered Information Theory

Results
- Developed a decentralized algorithm for learning and disseminating a mixture of factor analyzers
- Communications scale linearly in the intrinsic dimensionality of the data
- Local computations scale cubically in low-dimension instead of cubic in high-dimension
- Presented preliminary results at 2nd-year review of ARO MURI on Value-centered Information Theory
- Latest results submitted to 2014 IEEE Int. Workshop on Machine Learning for Signal Processing

Path Forward
- Analyze convergence properties of overall algorithm and investigate speedup strategies
- Automate learning of model order and dimensionality

Synopsis

We consider the problem of decentralized learning of a target appearance manifold using a network of sensors. Sensor nodes observe an object from different aspects and then, in an unsupervised and distributed manner, learn a joint statistical model for the data manifold. We employ a mixture of factor analyzers (MFA) model, approximating a potentially nonlinear manifold. We derive a consensus-based decentralized expectation maximization (EM) algorithm for learning the parameters of the mixture densities and mixing probabilities.

A spatially distributed sensor network can be used to construct a rich appearance model for targets in their common field-of-view. These models can then be used to identify previously seen objects if they reappear in the network at a later time. As an example, consider a network of cameras capturing images of an object from different but possibly overlapping aspects as the object traverses through the network’s field of view. The ensemble of images captured by the network forms a low-dimensional nonlinear manifold in the high-dimensional ambient space of images. One approach to appearance modeling would be to construct independent models of a local data manifold at each sensor and share it across the network. However, such an ensemble of models suffers from discretization of the aspect space and poor parameter estimates as the number of unknown parameters necessarily scale linearly with the number of sensor nodes. Alternatively, the sensor nodes can collaborate to construct a joint model for the image ensemble. The parameter estimates of the joint model will improve with the number of sensor nodes, since the number of unknown parameters in the model is intrinsic to the object and fixed, whereas the measurements scale linearly with the number of sensor nodes. The straightforward method of pooling images to a central location for joint model construction will require large and likely impractical network bandwidth. In this paper, we develop a decentralized learning method with greatly reduced data bandwidth need and results in a global appearance manifold model shared by all sensor nodes.

Objectives
- Learn object structure from multiple, distributed partial views for detection and classification applications
- Estimate an underlying low-dimensional data manifold of high-dimensional observations from a spatially-distributed network of sensors; disseminate learned structure to all nodes
- Develop decentralized algorithms which are scalable and robust to sensor node or communication link failure

Challenges
- Current, centralized learning approaches have poor scalability properties; not robust to node or link outages
- Disseminating manifold representation out to nodes requires additional communications

Approach
- Develop robust algorithm that learns a low-dimensional manifold across a network and simultaneously disseminates this representation to all nodes
- Model the nonlinear data manifold with a mixture of factor analyzers (MFA)
  - Parametric statistical model of mixture densities
  - Each node observes the model with different mixture probabilities
- Derive a consensus-based Expectation-Maximization (EM) algorithm from distributed measurements
  - Each node calculates local statistics from its own observations
  - Locally estimate global parameters from network averaged statistics

ARL Contribution
- Internal research with support from doctoral advisors
- This work is in partial fulfillment of requirements toward a doctorate degree for Mr. Whipps
- Synergistic with SEDD research in automated target detection and classification from dimension-reduced features from ground-based imaging systems

Collaborations (doctoral advisors)
- Prof. Randolph Moses, Dept. of ECE, The Ohio State University, moses.2@osu.edu
- Prof. Emre Ertin, Dept. of ECE, The Ohio State University, ertin.1@osu.edu

Impact
- Algorithms that are scalable and robust to node and link failures are critical features for Army ISR missions
- Fundamental understanding of data modeling across a network and properties of decentralized algorithms
- Contributes to research in dimensionality reduction, manifold learning, and reconstruction with emphasis on distributed learning systems
- Contributes to Distributed Information Fusion thrust within ARO MURI on Value-centered Information Theory

Results
- Developed a decentralized algorithm for learning and disseminating a mixture of factor analyzers
- Communications scale linearly in the intrinsic dimensionality of the data
- Local computations scale cubically in low-dimension instead of cubic in high-dimension
- Presented preliminary results at 2nd-year review of ARO MURI on Value-centered Information Theory
- Latest results submitted to 2014 IEEE Int. Workshop on Machine Learning for Signal Processing

Path Forward
- Analyze convergence properties of overall algorithm and investigate speedup strategies
- Automate learning of model order and dimensionality