AFRL-AFOSR-VA-TR-2015-0390

Algorithms on Flag Manifolds for Knowledge Discovery in N-way Arrays

Michael Kirby
COLORADO STATE UNIVERSITY

11/20/2015
Final Report

DISTRIBUTION A: Distribution approved for public release.

Air Force Research Laboratory
AF Office Of Scientific Research (AFOSR)/ RTB1
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
20-11-2015	Final Performance	01-08-2012 to 31-07-2015

4. TITLE AND SUBTITLE
Algorithms on Flag Manifolds for Knowledge Discovery in N-way Arrays

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-12-1-0408

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)
Michael Kirby

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
COLORADO STATE UNIVERSITY
601 S HOWES ST
FORT COLINS, CO 805212807 US

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AF Office of Scientific Research
875 N. Randolph St. Room 3112
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/AFOSR RTB1

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

12. DISTRIBUTION/AVAILABILITY STATEMENT
A DISTRIBUTION UNLIMITED: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
We proposed an approach for hyperspectral imagery classification that exploits the geometric framework of the Grassmann manifold i.e.,a parameterization of k dimensional subspaces of n-dimnsional space. The algorithm is particularly well suited to applications where sets of pixels are to be classified. Multiple pixels from a data class characterize the variability of the class information using a subspace representation. We use two metrics defined on the Grassmannian, chordal and geodesic, and one pseudometric, to compute pairwise distances between the points--subspaces. Once a distance matrix is generated, we use the classical multidimensional scaling to find a configuration of points with preserved or approximated original distances, thus realizing an embedding of the Grassmannian into Euclidean space. A sparse support vector machine (SSVM) trained in the embedding space simultaneously classifies embedded subspaces and selects a subset of optimal dimensions of the embedding for subsequent model reduction and data visualization.

15. SUBJECT TERMS
flag manifolds

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Michael Kirby
Unclassified	Unclassified	Unclassified	UU		**19b. TELEPHONE NUMBER** *(Include area code)* 970-491-6850

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

DISTRIBUTION A: Distribution approved for public release

Final Report

Algorithms on Flag Manifolds for Knowledge Discovery in N-way data arrays.

Award Number: FA9550-12-1-0408 P00001

Program Manager Years I-II: Dr. Tristan Nguyen
Program Manager Year III: Dr. Arje Nachman

Principal Investigator: Dr. Michael Kirby, Department of Mathematics, Colorado State University[1]

CO-Principal Investigator: Dr. Chris Peterson, Department of Mathematics, Colorado State University

[1]joint appointment Department of Computer Science, Colorado State University

1 Objectives

This project concerns the development of new algorithms for the analysis of N-way data sets that arise in diverse applications such as video surveillance and hyperspectral imaging of chemical and biological agents. We address two distinct pattern classification problems. First, given an N-way data set or a set of N-way data sets, identify subsets that are *similar* to a given set. Secondly, given an N-way data set or a set of N-way data sets, identify subsets that are *dissimilar* to a given set. The first problem concerns pattern matching while the second deals with anomaly detection, the goal of which is to determine novel instances of patterns.

The basic theme of this proposal is the design and implementation of mathematical algorithms that exploit the geometric structure of matrix manifolds, in particular, the flag manifolds. This is the natural extension of prior work on Grassmann and Stiefel manifolds which has proven to be very effective for knowledge discovery in 3-way data sets including face recognition at ultra low-resolution with variations in illumination. The long-term objectives of this research include data fitting using Schubert varieties, the contruction of mappings (and their inverses) from flag manifolds to their tangent spaces and the application of these ideas to computing statistics on flag manifolds. Further, we propose to use this geometric setting to compute curvature on N-way data sets and to develop a geometric approach for classifying points on flag manifolds.

2 Accomplishments

2.1 Year I

One of the challenges of analyzing N-way array data is the determination of boundary points. We consider a specific application to 3-way array data and the identification of vertices of a convex hull, as well as the distance of points to a boundary. The convex hull of a set of points, C can be used to expose extremal properties of C and to help identify elements of C of high interest. For many problems, particularly in the presence of noise, the true vertex set (and facets) may be difficult to determine and one should expand the list of high interest candidates to points lying near the boundary of the convex hull. We propose a quadratic program for the purpose of stratifying points in a data cloud based on proximity to the boundary of the convex hull. A quadratic program is solved for each data point to determine an associated weight vector. We show that the weight vector encodes geometric information concerning the point's relationship to the boundary of the convex hull. The computation of the weight vectors can be carried out in parallel, and for a fixed number of points and fixed neighborhood size, the overall computational complexity of the algorithm grows linearly with dimension. As a consequence, meaningful computations can be completed on reasonably large, high dimensional data sets [5,26].

Sparse representations that arise from the solution of an optimization with an ℓ_1-norm penalty have proven to be very powerful for characterizing data in 3-way arrays. For example, the Local Linear Embedding Algorithm has proven useful for determining structure preserving dimension reducing mappings of data on manifolds. We developed a modification to the Linear Embedding Algorithm optimization problem that serves to minimize the number of neighbors required for the representation of each data point using a sparsity inducing penalty term. The algorithm is shown to be robust over wide ranges of the sparsity parameter producing an average number of nearest neighbors that is consistent with the best performing parameter selection for LLE. Given the number of non-zero weights may be substantially reduced in comparison to LLE, sparse LLE can be applied to substantially larger data sets. We illustrated the approach using three numerical examples including the swiss roll and a gene expression data set to illustrate the behavior of the method in comparison to LLE [6,26].

Many computer vision tasks such as action classification and object recognition employ subspace models to represent data. These tasks often benefit from the ability to create an average or a prototype for a set of subspace data points. The most widely used method for averaging subspaces is the Karcher mean, also known as the Riemannian center of mass. However, this approach can be very slow and has substantial storage requirements. To overcome the shortcomings of subspace means found in the literature, we have developed several algorithms for averaging point clouds of subspaces on Grassmann and Stiefel manifolds as described in [7,28].

In [7,27] we explored the Split Bregman algorithm for solving ℓ_1-norm optimization problems. In particular, we examined several multivariate analytic techniques including Sparse Principal Components Analysis, Bisparse Singular Value Decomposition and Bisparse Singular Value Decomposition. For each of these problems we construct and solve a new optimization problem using these Bregman iterative techniques. Each of the proposed optimization problems contain one or more regu- larization terms to enforce sparsity in the solutions. We applied the Bisparse Singular Value Decomposition to the Hyperspectral Image denoising problem.

Another source of applications of 4-way array data is the numerical simulation of atmospheric dynamics. Each snapshot is a 3D profile of temperatures and velocities and these evolve in time. Algorithms for extracting information from these large data sets allow the discovery of weather processes. For example, the structure of a tropical cyclone eye and eyewall plays an important role in intensification. While the eyewall is usually defined in terms of instantaneous velocity and derived quantities such as vorticity, or thermodynamic variables such as equivalent potential temperature, or pressure, a Lagrangian eyewall definition is based on the transport of particles. In this study [4], we analyse a Lagrangian eyeeyewall interface (LEEI), which is defined as a surface that acts as barrier to particle motion. The surface is then analysed over varying initial time, and structural differences in time and height show that differences in Lagrangian structure and the degree of axisymmetry correspond to changes in intensity.

In [9,19] we propose a flag manifold representation as a framework for exposing geometric structure in a large data set. We illustrate the approach by building pose flags for pose identification in digital images of faces and action flags for action recognition in video sequences. These examples illustrate that the flag manifold has the potential to identify common features in noisy and complex datasets.

MOSSE lters provide a method for creating a model of a desired, or target, object from real data [30]. Given a target feature in a set of images (2-way arrays) or target actions in video sequences (3-way arrays), we can generate a MOSSE lter that can be used to detect these features or actions in new data sets. The lter strongly weights the frequencies local to the identied point while weakly averaging the rest of the signal. We found detection in 2-way arrays to be substantially better than 3-way arrays but this may be due to the inaccuracy of labeling the 3D locations in the training phase.

We address the problem of subclassification of rare circulating cells using data driven feature selection in 3-way arrays [13,31]. The data set consists of images of circulating tumor cells and three marginal cell populations from patients with diagnosed breast, prostate, and lung cancers. We determine a set of low level features which can structurally differentiate between different cell types of interest to contribute to the treatment and monitoring of cancer patients. We have implemented an image representation based on the characterization of a cell in terms of its concentric Fourier rings. The Fourier Ring Descriptors (FRDs) exploit the size variations and morphological differences between rare cell events while being rotationally invariant. Additionally, FRDs are invertible and allow us to visualize specific structural information pertinent to a given classification task. Using the low level descriptors, FRDs, as a representation with a linear support vector machine decision tree classifier we have been able to obtain good quantifiable accuracy on our data set. We discuss the applications of the results to clinical use in context of metastatic cancer patients.

2.2 Year II

Given a finite set of subspaces of $\mathbb{R}^n$, perhaps of differing dimensions, we describe a flag of vector spaces (i.e. a nested sequence of vector spaces) that best represents the collection based on a natural optimization criterion and we present an algorithm for its computation. The utility of this flag representation lies in its ability to represent a collection of subspaces of differing dimensions. When the set of subspaces all have the same dimension d, the flag mean is related to several commonly used subspace representations. For instance, the d-dimensional subspace in the flag corresponds to the extrinsic manifold mean. When the set of subspaces is both well clustered and equidimensional of dimension d, then the d-dimensional component of the flag provides an approximation to the Karcher mean. An intermediate matrix used to construct the flag can also be used to recover the canonical components at the heart of Multiset Canonical Correlation Analysis. Two examples utilizing the Carnegie Mellon University Pose, Illumination, and Expression Database (CMU-PIE) serve as visual illustrations of the algorithm, see [9] for details.

Let $\mathcal{C} = \{V_1, \dots, V_k\}$ be a collection of subspaces of a finite-dimensional real vector space V. Let L denote a one-dimensional subspace of V and let $\theta(L, V_i)$ denote the principal angle between L and V_i. Motivated by a problem in data analysis, we seek an L that maximizes the function $F(L) = \sum_i \cos\theta(L, V_i)$. Conceptually, this is the line through the origin that best represents $\mathcal{C}$ with respect to the criterion $F(L)$. A reformulation shows that L is spanned by a vector $v = \sum_i v_i$ which maximizes the function $G(v_1, \dots, v_k) = ||\sum_i v_i||^2$ subject to the constraints $v_i \in V_i$ and $||v_i|| = 1$. In this setting, v is seen to be the longest vector that can be decomposed into unit vectors lying on prescribed hyperspheres. A closely related problem corresponds to finding the longest vector that can be decomposed into vectors lying on prescribed hyperellipsoids. Using Lagrange multipliers, the critical points of either problem can be cast as solutions of a multivariate eigenvalue problem. We employ homotopy continuation and numerical algebraic geometry to solve the problem and obtain the extremal decompositions. See [10] for details.

Determining solutions to polynomial systems has been shown to be related to determining optimal representations on Grassmann manifolds. Further, we can use the solutions of such systems to construct a flag of best fit (see references [2, 3, 7, 10, 28] for further details and examples). Given a polynomial system $f : \mathbb{C}^N \to \mathbb{C}^n$, the methods of numerical algebraic geometry produce numerical approximations of the isolated solutions of $f(z) = 0$, as well as points on any positive-dimensional components of the solution set, $\mathbf{V}(f)$. Some of these methods are guaranteed to finnd all isolated solutions (nonsingular and singular alike), while others may miss singular solutions. One of the most recent advances in this field is regeneration, an equation-by-equation solver that is often more efficient than other methods. We consider the use of perturbed homotopies for solving polynomial systems. In particular, we propose solving a perturbed version of the polynomial system, followed by a parameter homotopy to remove the perturbation. Such perturbed homotopies are sometimes more efficient than regular homotopies. Second, a useful consequence is that the application of this perturbation to regeneration will yield all isolated solutions, including all singular isolated solutions. This version of regeneration (perturbed regeneration) can decrease the efficiency of regeneration but increases its applicability. See [3] for further details.

In [13], we explore the development of a low-level rotationally invariant feature selection method that addresses the problem of subclassification of rare circulating cells using data driven feature selection. The data set consists of images of circulating tumor cells and three marginal cell populations from patients with diagnosed breast, prostate, and lung cancers. We determine a set of low level features which can structurally differentiate between different cell types of interest to contribute to the treatment and monitoring of cancer patients. We have implemented an image representation based on the characterization of a cell in terms of its concentric Fourier rings. The Fourier Ring Descriptors (FRDs) exploit the size variations and morphological differences between rare cell events while being rotationally invariant. Additionally, FRDs are invertible and allow us to visualize specific structural information pertinent to a given classification task. Using the low level descriptors, FRDs, as a representation with a linear support vector machine decision tree classifier we have been able to obtain good quantifiable accuracy on our data set. We discuss the applications of the results to clinical use in context of metastatic cancer patients.

In [21] we propose an ℓ_1-norm penalized sparse support vector machine (SSVM) as an embedded approach to the hyperspectral imagery band selection problem. SSVMs exhibit a model structure that includes a clearly identifiable gap between zero and non-zero weights that permits important bands to be definitively selected in conjunction with the classification problem. The SSVM Algorithm is trained using bootstrap aggregating to obtain a sample of SSVM models to reduce variability in the band selection process. This preliminary sample approach for band selection is followed by a secondary band selection which involves retraining the SSVM to further reduce the set of bands retained. We propose and compare three adaptations of the SSVM band selection algorithm for the multiclass problem. Two extensions of the SSVM Algorithm are based on pairwise band selection between classes. Their performance is validated by using one-against-one (OAO) SSVMs. The third proposed method is a combination of the filter band selection method WaLuMI in sequence with the (OAO) SSVM embedded band selection algorithm. We illustrate the perfomance of these methods on the AVIRIS Indian Pines data set and compare the results to other techniques in the literature. Additionally we illustrate the SSVM Algorithm on the Long-Wavelength Infrared (LWIR) data set consisting of hyperspectral videos of chemical plumes.

Many computer vision algorithms employ subspace models to represent data. Many of these approaches benet from the ability to create an average or prototype for a set of subspaces. The most popular method in these situations is the Karcher mean, also known as the Riemannian center of mass. The prevalence of the Karcher mean may lead some to assume that it provides the best average in all scenarios. However, other subspace averages that appear less frequently in the literature may be more appropriate for certain tasks. The extrinsic manifold mean, the L 2-median, and the ag mean are alternative averages that can be substituted directly for the Karcher mean in many applications. This paper evaluates the characteristics and performance of these four averages on synthetic and real-world data. While the Karcher mean generalizes the Euclidean mean to the Grassman manifold, we show that the extrinsic manifold mean, the L 2-median, and the ag mean behave more like medians and are therefore more robust to the presence of outliers among the subspaces being averaged. We also show that while the Karcher mean and L 2-median are computed using iterative algorithms, the extrinsic manifold mean and ag mean can be found analytically and are thus orders of magnitude faster in practice. Finally, we show that the ag mean is a generalization of the extrinsic manifold mean that permits subspaces with different numbers of dimensions to be averaged. The result is a cookbook that maps algorithm constraints and data properties to the most appropriate subspace mean for a given application. See [22] for details.

We propose an approach for capturing the signal variability in hyperspectral imagery using the framework of the Grassmann manifold. Labeled points from each class are sampled and used to form abstract points on the Grassmannian. The resulting points on the Grassmannian have representations as orthonormal matrices and as such do not reside in Euclidean space in the usual sense. There are a variety of metrics which allow us to determine a distance matrices that can be used to realize the Grassmannian as an embedding in Euclidean space. We illustrate that we can achieve an approximately isometric embedding of the Grassmann manifold using the chordal metric while this is not the case with geodesic distances. However, non-isometric embeddings generated by using a pseudometric on the Grassmannian lead to the best classification results. We observe that as the dimension of the Grassmannian grows, the accuracy of the classification grows to 100% on two illustrative examples. We also observe a decrease in classification rates if the dimension of the points on the Grassmannian is too large for the dimension of the Euclidean space. We use sparse support vector machines to perform additional model reduction. The resulting classifier selects a subset of dimensions of the embedding without loss in classification performance. See [23] for details.

2.3 Year III Accomplishments

In this application [24], we return to 4-way arrays, i.e., hyperspectral movies. Specifically, we present an application of persistent homology to the detection of chemical plumes in hyperspectral movies. The pixels of the raw hyperspectral data cubes are mapped to the geometric framework of the real Grassmann manifold $G(k, n)$ (whose points parameterize the k- dimensional subspaces of $\mathbb{R}^n$) where they are analyzed, contrasting our approach with the more standard framework in Euclidean space. An advantage of this approach is that it allows the time slice in a hyperspectral movie to be collapsed to a sequence of points in such a way that some of the key structure within and between the slices is encoded by the points on the Grassmann manifold. This motivates the search for topological structure, associated with the evolution of the frames of a hyperspectral movie, within the corresponding points on the Grassmann manifold. The proposed framework affords the processing of large data sets, such as the hyperspectral movies explored in this investigation, while retaining valuable discriminative information.

We developed an algorithm for detecting anomalies in video sequences, i.e., a 3-way array [25]. One of the nice features of this algorithm was the fact we were able to integrate our flag of best fit work which provided a much faster option to the Karcher mean for computing means of subspaces. Given the goal of anomaly detection, we used video data of *nominal* activity for constructing a representation of the data. The resulting model produces alarm notifications when anomalous activity is observed. The approach involves characterizing segments of video as subspaces and invoking the geometric framework of Grassmann manifolds, i.e., the space of k-dimensional subspaces of n-dimensional space, $Gr(k, n)$. With subspaces treated as abstract points together with a suitably chosen metric on the Grassmann, i.e., the manifold of such points, one can exploit novel aspects of the geometry of the data for the purpose of anomaly detection. This mathematical framework is used to extend the Multivariate State Estimation Technique to the context of Grassmann manifolds. We present an application to the ETHZ Living Room Data Set for detecting anomalous activities.

Extedning initial work in [23] in [17] an approach for hyperspectral imagery classification that further exploits the geometric framework the Grassmann manifold (or the Grassmannian), i.e., a parameterization of k-dimensional subspaces of $\mathbb{R}^n$. The algorithm is particularly well suited to applications where sets of pixels are to be classified. Multiple pixels from a data class characterize the variability of the class information using a subspace representation. We use two metrics defined on the Grassmannian, chordal and geodesic, and one pseudometric, to compute pairwise distances between the points, i.e., subspaces. Once a distance matrix is generated, we use the classical multidimensional scaling to find a configuration of points with preserved or approximated original distances, thus realizing an embedding of the Grassmannian into Euclidean space. A sparse support vector machine (SSVM) trained in the embedding space simultaneously classifies embedded subspaces and selects a subset of optimal dimensions of the embedding for subsequent model reduction and data visualization. The pseudometric framework allows for as low as one-dimension SSVM-based selection. We analyze frameworks and compare binary classification results for the three distances. Lastly, we provide multiclass results, realizing a higher-dimensional embedding of the encoded points from multiple data classes.

In [18] we present a data array analysis of the human immune response to respiratory viruses including influenza, respiratory syncytia virus, and human rhinovirus, and compare this with the response to Lipopolysaccharides (LPS). Using an anomaly detection framework we identified 16 pathways that achieve a minimum cutoff accuracy for predicting outcomes across the four different respiratory viruses H1N1, H3N2, RSV and HRV. A subset of 8 of these pathways were identified as early warning pathways including inflammatory bowel disease, toll-like receptor signaling, Influenza A, lysosome, intestinal immune network for IgA production, HIVNEF, and NF-kappa B signaling. These early warning pathways correctly predict for H1N1 and H3N2 that almost half of the subjects will become symptomatic in less than forty hours of monitoring and that three of 18 subjects will become symptomatic after only 8 hours. Host pathway analysis of a human endotoxin gene expression data set revealed a 14 pathway signature that identified symptomatic subjects within 2-3 hours post exposure. Comparative analysis between the prognostic bacterial and viral pathway signatures showed a single pathway, IL-22BP, that overlapped between the signatures. These results suggest that there are strong pathway signatures that characterize the immune system's response to infection at its earliest stages. The identification of prognostic respiratory virus biomarkers has the potential to provide an early warning system that is capable of predicting that subjects will become symptomatic at the earliest stages of infection expanding medical diagnostic capabilities and treatment options. The immune system's response to disease may be viewed as a deterministic, carefully orchestrated signaling network responsible for maintaining the health of the host organism. The initial response of the immune system may be viewed as a "canary in a coal mine" as the host deviates from the healthy state. We are motivated to identify pathway signatures that reflect the very earliest perturbations in the host response to acute infection. We contend that these pathways, once identified, can be used to monitor the health state of the host by using the anomaly detection workflow to quantify and predict health outcomes to pathogens.

Data in N-way arrays where one of the dimensions is time may be viewed as a trajectory in a very high dimensional space. The shape of this trajectory may be characterized by generalized curvatures. Let γ be a sufficiently smooth non-degenerate curve in $\mathbb{R}^n$. The Frenet-Serret apparatus of γ consists of a frame and generalized curvature values $\kappa_1, \ldots, \kappa_{n-1}$ at each point of γ. The local singular value apparatus of γ consists of an ordered sequence of n mutually orthogonal lines and local singular values $\sigma_1, \ldots, \sigma_n$ at each point of γ. In [16] we define the local singular value apparatus, show how it can be computed, and describe how it relates to the Frenet-Serret apparatus.

3 Personnel Supported

Year I

Michael Kirby, PI. Chris Peterson, CO-PI.

Year II

Michael Kirby, PI. Chris Peterson, CO-PI. Sofya Chepustanova, graduate research assistant.

Year III

Michael Kirby, PI. Chris Peterson, CO-PI. Sofya Chepustanova, graduate research assistant. Kun Wang, postdoctoral research assistany.

4 Technical Publications

4.1 Journal Publications

4.1.1 Year I

1) J. Chang and C. Peterson and M. Kirby, Feature Patch Illumination Spaces and Karcher Compression for Face Recognition via Grassmannians, Advances in Pure Mathematics, Vol. 2, No. 4, 226-242, 2012.

2) D. Eklund, C. Jost, C. Peterson, A method to compute Segre classes of subschemes of projective space, Journal of Algebra and its Applications Vol. 12, no. 2, (2013).

3) D. Bates, J. Hauenstein, T. McCoy, C. Peterson, A. Sommese, Recovering exact results from inexact numerical data in algebraic geometry, Experimental Mathematics, Vol. 22, Issue 1, pg 38-50 (2013).

4) B. Rutherford, G. Dangelmayr, and M. Kirby, *A time-dependent Lagrangian eyewall*, Quarterly Journal of the Royal Meteorological Society, Vol. 138, No. 669, John Wiley & Sons, Ltd., pp 2009–2018, 2012.

5) Lori Ziegelmeier, Michael Kirby, Chris Peterson, *A quadratic program to stratify high dimensional data based on proximity to the boundary of the convex hull*, (under review).

6) Lori Ziegelmeier, Michael Kirby and Chris Peterson, *Sparse Locally Linear Embedding*, (under revision).

7) J. Marks, M. Kirby, and C. Peterson, *A Normal/Tangent Bundle Algorithm for Averaging Point Clouds on Grassmann and Stiefel Manifolds*, (under revision.)

8) N. Rohrbacher and M. Kirby, *Sparse Principal Component Analysis via Bregman Iterations with Applications to Face Recognition*, (under revision).

4.1.2 Year II

9) Draper, B., Kirby, M., Marks, J., Marrinan, T., and Peterson, C. (2014). *A flag representation for finite collections of subspaces of mixed dimensions.* Linear Algebra and its Applications, 451, 15-32.

10) Daniel Bates and Davis, Brent and Kirby, Michael and Marks, Justin and Peterson, Chris (2015) *The max-length-vector line of best fit to a set of vector subspaces and an optimization problem over a set of hyperellipsoids*, Numerical Linear Algebra with Applications, Vol. 22 pp 453–464.

11) Stephen O'Hara, Kun Wang, Richard A Slayden, Alan R Schenkel, Greg Huber, Corey S O'Hern, Mark D Shattuck and Michael Kirby *Iterative Feature Removal Yields Highly Discriminative Pathways*, BMC Genomics 14:832, 2013.

12) Kun Wang, Vineet Bhandari, Sofya Chepustanova, Greg Huber, Stephen OHara, Corey S. OHern, Mark D. Shattuck, Michael Kirby *Which Biomarkers Reveal Neonatal Sepsis?* PLoS ONE 8(12): e82700. doi:10.1371/journal.pone.0082700, Dec 18, 2013.

13) Emerson, Tegan and Kirby, Michael and Bethel, Kelly and Kolatkar, Anand and Luttgen, Madelyn and OHara, Stephen and Newton, Paul and Kuhn, Peter, (2015), *Fourier-Ring Descriptor to Characterize Rare Circulating Cells from Images Generated Using Immunofluorescence Microscopy*, Computerized Medical Imaging and Graphics, Vol. 40, pp 70–87.

4.1.3 Year III

14) Mai M, Wang K, Huber G, Kirby M, Shattuck MD, OHern CS (2015), *Outcome Prediction in Mathematical Models of Immune Response to Infection.* PLoS ONE 10(8): e0135861.

15) Arta Jamshidi and Kirby, Michael, (2015) *A Radial Basis Function Algorithm with Automatic Model Order Determination,* SIAM Journal of Scientific Computation (SISC), Vol. 37., No.3, ppA1319-A1341.

16) R. Arn, B. Draper, M. KIrby and C. Peterson, *The Frenet-Serret Apparatus and Local Singular Value Decomposition of curves in* $\mathbb{R}^n$, (submitted).

17) Sofya Chepushtanova and Michael Kirby, Sparse Grassmannian Embeddings for Hyperspectral Image Classification.

18) K. Wang, S. Langevin, J. Morrison, S. Ogle, C. O'Hern, M. Shattuck, R. Slayden, M. Katze, M. Kirby, *Anomaly Detection in Host Signaling Pathways for the Early Prognosis of Acute Infection* (submitted)

4.2 Reviewed Conference Proceedings

4.2.1 Year I

19) T. Marrinan, R. Beveridge, B. Draper, M. Kirby and C. Peterson, (2015) *Flag Manifolds for the Characterization of Geometric Structure in Large Data Sets*, A. Abdulle et al. (eds.), Numerical Mathematics and Advanced Applications ENUMATH 2013, Lecture Notes in Computational Science and Engineering 103, Springer International Publishing. 457-464. European Numerical Mathematics and Advanced Applications, Lausanne, Switzerland 8/2013

20) E. Hanson, F. Motta, C. Peterson, L. Ziegelmeier, On the Strengthening of Topological Signals in Persistent Homology through Vector Bundle Based Maps, Proceedings of the Canadian Conference on Computational Geometry (2012), pg 303-308.

4.2.2 Year II

21) Sofya Chepushtanova, Christopher Gittins, and Michael Kirby, *Band Selection in Hyperspectral Imagery Using Sparse Support Vector Machines*, submitted SPIE conference on Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XX, 5/2014.

22) B. Draper and M. Kirby and J. Marks and T. Marrinan and C. Peterson, *Finding the Subspace Mean or Median to Fit Your Needs*, to appear Computer Vision and Pattern Recognition (CVPR) June, 2014.

23) Sofya Chepushtanova and Michael Kirby, *Classification of Hyperspectral Imagery on Embedded Grassmannians*, to appear 6th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, June 2014, Lausanne, Switzerland WHISPERS 2014.

4.2.3 Year III

24) Sofya Chepushtanova, Michael Kirby, Chris Peterson and Lori Ziegelmeier, *An Application of Persistent Homology on Grassmann Manifolds for the Detection of Signals in Hyperspectral Imagery*, In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 2015

25) K. Wang, J. Thompson, C. Peterson, and M. Kirby, (2015) *Identity maps and their extensions on parameter spaces: Applications to anomaly detection in video*, Proceedings Science and Information Conference, pp. 345-351, London, July 28-30, 2015.

4.3 Ph.D. Dissertations, M.S. Theses, Undergrad Honors Theses

4.3.1 Year I

26) Lori Ziegelmeier, *Exploiting Geometry, Topology and Optimization for Knowledge Discovery in Big Data*, Ph.D. 5/2013

27) Nicholas Rohrbacker, *Sparse multivariate analyses via ℓ_1-regularized optimization problems solved with Bregman iterative techniques*, Ph.D. 10/2012

28) Justin Marks, *Mean Variants on Matrix Manifolds*, Ph.D. 8/2012

29) Tim Marrinan, *The Flag of Best Fit as a Representative for a Collection of Subspaces*, M.S. 7/2013

30) Robert Arn, *Object and Action Detection Methods Using MOSSE Filters*, M.S. 11/2012

31) Tegan Emerson, *Automated Detection of Circulating Cells Using Low Level Features*, M.S. 6/2013

32) Matt Heine, Undergraduate Honors Thesis.

4.3.2 Year II

33) Silvia Osnaga, *Low Rank Representations for Matrices and Tensors*, Ph.D. defended 7/8/2014

34) Justin Hughes, *Group Action on Neighborhood Complexes of Cayley Graphs*, Ph.D. defended 5/2014

35) Cory Previte, *The $\mathcal{D}$-Neighborhood complex of graphs*, Ph.D. defended 5/2014

36) Kelly Shick, Undergraduate Honors Thesis, 5/14.

4.3.3 Year III

37) Sofya Chepushtanova, Ph.D., *Algorithms for feature selection and pattern recognition on Grassmann manifolds*, 6/2015.

38) Drew Schwickerath Ph.D., *Linear Models, Signal Detection, and the Grassmann Manifold*, 12/2014

5 Interactions/Transitions

5.1 Presentations

Below is a list of presentations by the PI, CO-PI and their students during the year of the annual report.

5.1.1 Year I

- Invited Lecture, Infectious Disease Workshop, Michael Kirby, Yale University, 7/2013
- 12th International Conference for Complex Acute Illness, Michael Kirby, Budapest, Hungary, 8/2013
- Regina, Canada, CMS special session, Chris Peterson speaker
- San Jose, Costa Rica, Presented minicourse on Persistent Homology, Chris Peterson
- Minneapolis, Minnesota, SIAM special session, Chris Peterson - speaker
- Lincoln, Nebraska, University of Nebraska - Immerse Summer Program, Chris Peterson - speaker - 2 talks
- Catania, Italy - Fifteen Years of Pragmatic Conference - Chris Peterson, speaker
- San Diego, California, AMS National Meeting special session, Chris Peterson - speaker
- Boulder, Colorado, AMS Sectional Meeting special session, Chris Peterson,- speaker
- Two talks at the Accademia Peloritana dei Pericolanti, Chris Peterson, Messina, Italy

- Seminar talk at Villa Pace, Chris Peterson, Università di Messina, Messina, Italy
- Fort Collins, Colorado, SIAM Conference on Applied Algebraic Geometry, Chris Peterson, speaker, organizer of three special sessions
- Manifold Analysis for Hyperspectral Imagery: A Collaboration with MIT Lincoln Lab (oral presentation) Justin Marks, 2012 DTRA/NSF Algorithms Workshop, San Diego, CA, 11/2012 -speaker
- Classication of Data on Embedded Grassmannians, Sofya Chepustova, 2012 DTRA/NSF Algorithms Workshop, San Diego, CA, 11/2012, -poster presentation
- Solution to Sparse Locally Linear Embedding using Split Bregman Lori Ziegelmeier, 2012 DTRA/NSF Algorithms, San Diego, CA, 11/2012 - speaker
- Tools and Techniques in Geometric and Topological Data Analysis, Lori Ziegelmeier, Colorado State May 15, 2013 - Ph.D. defense
- Sparse Nearest Neighbor Selection for the Locally Linear Embedding Algorithm, Lori Ziegelmeier, AMS Spring Western Sectional Meeting, Boulder, CO April 13, 2013 - speaker
- Robust Geometric Structure from High Dimensional Data using Sparse LLE, Lori Ziegelmeier, Joint Mathematics Meetings, San Diego, CA January 12, 2013 - speaker
- On the Strengthening of Topological Signals in Persistent Homology through Vector Bundle Based Maps, Lori Ziegelmeier, September 6, 2012 Greenslopes Seminar, Colorado State - speaker
- Comprehensive Analysis of Hyperspectral Data using Band Selection based on Sparse Support Vector Machines, Sofya Chepustanova, March 2013 Front Range Applied Mathematics (FRAM) Student Conference, Denver, CO - speaker
- Hyperspectral Band Selection Using Sparse Support Vector Machines, Sofya Chepustanova, Joint Mathematics Meetings, San Diego, CA, January 2013 - speaker

5.1.2 Year II

- Fort Collins, Colorado, SIAM Conference on Applied Algebraic Geometry, Chris Peterson, speaker 2 sessions, organizer one special session August 2013:
  - Special session on *Numerical Perspectives on Classical Themes in Algebraic Geometry* - speaker
  - Special session on *Algebraic Geometry of Tensor Decompositions* - speaker
  - Special session on *Algebro-geometric Approaches to Tensor Spaces, Tensor Decomposition, and Identifiability"* - co-organizer with Hirotachi Abo, Giorgio Ottaviani, Luke Oeding

- Louisville, Kentucky, AMS Sectional Meeting special session - Peterson speaker (October 6, 2013)
- Colloquium talk in The Department of Applied & Computational Mathematics & Statistics, Notre Dame University, Notre Dame, Indiana, Peterson - speaker (Apri, 2014)
- Bilbao, Spain, First Joint International Meeting RSME-SCM-SEMA-SIMAI-UMI - special session *Applications of Algebraic Geometry" - Peterson speaker (June/July 2014)*
- SPIE DSS 2014, Baltimore, MD Poster Presentation Band Selection in Hyperspectral Imagery Using Sparse Support Vector Machines, Sofya Chepustanova, speaker.
- March 2014 Algorithms for Threat Detection Program Review, Boulder, CO Oral Presentation Exploring Uses of Persistent Homology for Hyperspectral Remote Sensing, Sofya Chepustanova, speaker.
- March 2014 Conference on Data Analysis (CoDA) 2014, Santa Fe, NM Poster An Application of Persistent Homology on Grassmann Manifolds to the Detection of Signals in Hyperspectral Imagery, Sofya Chepustanova, speaker.
- February 2014 Argonne National Laboratory Oral Presentation Data Analysis Methods and Applications: Hyperspectral Band Selection and Data Classi- cation on Embedded Grassmannians, Sofya Chepustanova, speaker.
- February 2014 Topological Data Analysis Workshop, SAMSI, NC Poster Set-to-Set Pattern Recognition on Grassmann Manifolds, Sofya Chepustanova, speaker.
- January 2014 2014 Joint Mathematics Meetings, Baltimore, MD Oral Presentation Pattern Classication by Ellipsoidal Machines Using Semidenite Programming
- September 2013 IMA Hot Topics Workshop on Imaging in Geospatial Applications, Minneapolis, MN Poster Sparse SVMs for Hyperspectral Band Selection, Sofya Chepustanova, speaker.
- Chemical Signature Detection Using Flag Representations in Hyperspectral Images, DTRA/NSF Algorithms Workshop, Boulder, CO, 3/2014 - Timothy Paul Marrinan, speaker
- Linear Models, the Grassmann Manifold, and Signal Detection, DTRA/NSF Algorithms Workshop, Boulder, CO, 3/2014, Anthony Schwickerath, speaker
- Schubert Varieties and their relation to Linear Models, the Grassmann Manifold, and Signal Detection, 5/2014, Pattern Analysis Laboratory Seminar, Colorado State University, Anthony Schwickerath, speaker
- *Flag Manifolds for Characterization of Information in Video Sequences*, European Numerical Mathematics and Advanced Applications, Lausanne, Switzerland 8/2013, M. Kirby speaker.

- Classification of Hyperspectral Imagery on Embedded Grassmannians, 6th Workshop on Hyperspectral Image and Signal Processing Evolution in Remote Sensing, Lausanne, Swizterland, 6/2014, M. Kirby speaker.

5.1.3 Year III

- *Pathway Monitoring as a Methodology for the Early Diagnosis of Infection*, 2015 Chemical and Biological Defense Science and Technology Conference, St. Louis, May 2015 (Kirby poster presentation).

- *Detecting Threats in Data: from Euclidean Space to Grassmannians*, NSF/DTRA ATD Workshop, Washington, D.C., July 13, 2015. (Kirby oral presentation)

- *Anomaly Detection in Host Signaling Pathways for the Early Prognosis of Acute Infection*, Lipari School on BioInformatics and Computational Biology, July 2015 (Kirby poster presentation).

- *An Application of Persistent Homology on Grassmann Manifolds for the Detection of Signals inHyperspectral Imagery*, International Geoscience and Remote Sensing Symposium 2015 (IGARSS 2015), July 26, Milan (Kirby poster prentation).

- *Identity maps and their extensions on parameter spaces: Applications to anomaly detection in video*, Science and Information Conference, July 30, 2015. (Kirby oral presentation)

- Tegan Emerson, The Split Bregman Algorithm with Application to Aerosol Unmixing." Algorithms for Threat Detection Workshop, National Science Foundation. July 14, 2015

- Tegan Emerson, Statistical Signal Processing in Hyperspectral Images: a Framework for Dimensionality Reduction in Detection." Mini-workshop, Statistical and Computational Interface of Big Data Conference, Hong Kong, January 12, 2015

- Sofya Chepustanova, "Sparse Grassmannian embeddings for hyperspectral image classification", January 2015 Joint Mathematics Meetings, San Antonio, TX.

- Tim Marinnan, (Ph.D. student), "Detecting Weak Signals in Linear Subspace Data", 2nd Annual Signature Discovery Workshop, University of Washington. November 2014

- Tim Marinnan, Detecting weak signals in hyperspectral images and videos by spanning variation, DTRA/NSF Workshop on Algorithms for Threat Detection, Arlington, VA, July 2015

- Chris Peterson, special session "Computational Algebraic Geometry", Montevideo, Uruguay - FOCM -(December, 2014)

- Chris Peterson, Department of Mathematics Seminar, Oklahoma State University (February, 2015)

- Chris Peterson, Department of Mathematics Seminar, University of Chicago (May, 2015)

- Tegan Emerson, Topics in Geometric and Topological Data Analysis." Heidelberg Laureate Forum, Heidelberg University, Germany September 25, 2014 (poster presentation)

- Sofya Chepustanova, "Geometric data analysis: Grassmannian framework for set-to-set pattern recognition", Amazon Graduate Research Symposium, Seattle, WA.November 2014

- Sofya Chepustanova, "Persistent Homology for Hyperspectral Data Analysis under the Grassmannian Framework", DTRA/NSF Workshop on Algorithms for Threat Detection, Arlington, VA, July 2015 (poster presentation)

5.2 Transitions

5.2.1 Year I

None.

5.2.2 Year II

In [4] we introduce Iterative Feature Removal (IFR) as an unbiased approach for selecting features with diagnostic capacity from large data sets, i.e., sets of matrices or 2-way arrays. The algorithm is based on our recently developed tools (see [7]) that are driven by sparse feature selection goals. When applied to genomic data, our method is designed to identify genes that can provide deeper insight into complex interactions while remaining directly connected to diagnostic utility. We contrast this approach with the search for a minimal best set of discriminative genes, which can provide only an incomplete picture of the biological complexity. Our results challenge the paradigm of using feature selection techniques to design parsimonious classifiers from microarray and similar high-dimensional, small-sample-size data sets.

An additional transition involves looking at sepsis biomarker data in the framework of Grassmannians emerging from [1,8,9] and sparse support vector machines [7]. We address the identification of optimal biomarkers for the rapid diagnosis of neonatal sepsis. We employ both distances on Grassmannians and sparse support vector machine (SSVM) classifiers to select the best subset of biomarkers from a large hematological data set collected from infants with suspected sepsis from Yale-New Haven Hospital's Neonatal Intensive Care Unit (NICU). Grassmann manifold distances are shown to be related to canonical correlation analysis (CCA) and are used to select sets of biomarkers of increasing size that are most highly correlated with sepsis infection. The effectiveness of these biomarkers is then validated by constructing a sparse support vector machine diagnostic classifier. We find that the following set of five biomarkers capture the essential diagnostic information (in order of importance): Bands, Platelets, neutrophil CD64, White Blood Cells, and Segs. Further, the diagnostic performance of the optimal set of biomarkers is significantly higher than that of isolated individual biomarkers. These results suggest an enhanced sepsis scoring system for neonatal sepsis that includes these five biomarkers.

Anthony Schwickerath passed his Ph.D. preliminary examination entitled *Linear Models, the Grassmann Manifold, and Signal Detection.*

During Year 2 of this award we applied for and received other grants related to the analysis and processing of N-way arrays including:

- ATD: Detection and Classification of Threats Using Subspace Manifold Geometry (NSF)
- Compressed Sensing for Wide Area Chemical and Biological Early Warning (DOD)
- LWIR Compressive Sensing Hyperspectral Imager (DOE)

5.3 Transitions

5.3.1 Year III

The application of techniques for analyzing N-way arrays is finding fruitful applications in the analysis of biological data, in particular to gene expression data sets. One of the key questions is how to integrate data from many experiments, animal types, diseases, and data types. Every variable adds a dimension to the array. We have proposed to use data bundles (prior AFOSR research) and our algorithms for N-way arrays developed here to explore this problem. We have recently been informed that the proposal below will be awarded by DARPA.

THUNDER: TOLERANT HOSTS USING NOVEL DRUG-ENHANCED RESILIENCE. The goal of this study is to compare the effects of two substance abuse interventions on health outcomes in an urban population of older opiate addicts.

Tolerance to infection is the ability of a host to remain healthy during infection with a pathogen. Many examples of this intriguing phenomenon can be found in nature, but the biological mechanisms underlying tolerance during infection are understudied. The goal of project THUNDER is to discover mechanisms of tolerance and to identify and validate interventions to induce tolerance to infection. We propose to establish that the mechanisms of tolerance can be discovered by data-driven approaches for quantifying the shape and structure of biological signatures characteristic of tolerance. We will extend and apply recently developed mathematical tools to identify common traits of tolerance by analyzing the output of the data collection. We will seek to identify additional, and potentially unique, mechanisms for tolerance using the integrated data set combining pre-existing data sets with project THoR data. Our analysis will focus on characterizing the temporal evolution of the tolerant response as high-dimensional feature trajectories across multimodal biomarkers. We will use geometric, topological, and dynamical systems tools to inform the bioinformatic analysis with the goal of establishing multi-signature mechanisms of tolerance. Role: PI (Colorado State University subcontract).

Finally, in Year III, two Ph.D. students completed the requirements for their degrees. Sofya Chepushtanova, Ph.D., *Algorithms for feature selection and pattern recognition on Grassmann manifolds*, 6/2015 and Drew Schwickerath Ph.D., *Linear Models, Signal Detection, and the Grassmann Manifold*, 12/2014

1.

1. Report Type

Final Report

Primary Contact E-mail
Contact email if there is a problem with the report.

kirby@math.colostate.edu

Primary Contact Phone Number
Contact phone number if there is a problem with the report

970-481-1416

Organization / Institution name

Colorado State University

Grant/Contract Title
The full title of the funded effort.

Algorithms on Flag Manifolds for Knowledge Discovery in N-way data arrays.

Grant/Contract Number
AFOSR assigned control number. It must begin with "FA9550" or "F49620" or "FA2386".

FA9550-12-1-0408

Principal Investigator Name
The full name of the principal investigator on the grant or contract.

Michael Kirby

Program Manager
The AFOSR Program Manager currently assigned to the award

Dr. Arje Nachman

Reporting Period Start Date

08/01/2012

Reporting Period End Date

07/30/2015

Abstract

We proposed an approach for hyperspectral imagery classification that exploits the geometric framework the Grassmann manifold (or the Grassmannian), i.e., a parameterization of k-dimensional subspaces of n-dimnsional space. The algorithm is particularly well suited to applications where sets of pixels are to be classified. Multiple pixels from a data class characterize the variability of the class information using a subspace representation. We use two metrics defined on the Grassmannian, chordal and geodesic, and one pseudometric, to compute pairwise distances between the points, i.e., subspaces. Once a distance matrix is generated, we use the classical multidimensional scaling to find a configuration of points with preserved or approximated original distances, thus realizing an embedding of the Grassmannian into Euclidean space. A sparse support vector machine (SSVM) trained in the embedding space simultaneously classifies embedded subspaces and selects a subset of optimal dimensions of the embedding for subsequent model reduction and data visualization. The pseudometric framework allows for as low as one-dimension SSVM-based selection. We analyze frameworks and compare binary classification results for the three distances. Lastly, we provide multiclass results, realizing a higher-dimensional embedding of the encoded points from multiple data classes.

We demonstrated an application of persistent homology to 4-way arrays, i.e., the detection of chemical

plumes hyperspectral movies. The pixels of the raw hyperspectral data cubes are mapped to the geometric framework of the
real Grassmann manifold G(k,n) (whose points parameterize the k-dimensional subspaces of n-dimensions) where they are analyzed, contrasting our approach with the more standard framework in Euclidean space. An advantage of this approach is that it allows the time slice in a hyperspectral movie to be collapsed to a sequence of points in such a way that some of the key structure within and between the slices is
encoded by the points on the Grassmann manifold. This motivates the search for topological structure, associated with the evolution of the frames of a hyperspectral movie, within the corresponding points on the Grassmann manifold. The proposed framework affords the processing of large data sets, such as the hyperspectral movies explored in this investigation, while retaining valuable discriminative information.

We developed an algorithm for detecting anomalies in video
sequences, i.e., a 3-way array. One of the nice features of this algorithm was the fact we were able to integrate our flag of best fit work which provided a much faster option to the Karcher mean for computing means of subspaces. Given the goal of anomaly detection, we used video data of {\\it nominal} activity for constructing a representation of the data. The resulting model produces alarm notifications when anomalous activity is observed.
The approach involves characterizing segments of video as subspaces and invoking the geometric framework of Grassmann manifolds, i.e., the space of k-dimensional
subspaces of n-dimensional space, Gr(k,n). With subspaces treated as abstract points together with a suitably chosen metric on the Grassmann, i.e., the manifold of
such points, one can exploit novel aspects of the geometry of the data for the purpose of anomaly detection. This mathematical framework is used to extend the Multivariate State Estimation Technique to the context of Grassmann manifolds. We present an application to the ETHZ Living Room Data Set for detecting anomalous activities.

We present a data array analysis of the human immune response to respiratory viruses including influenza, respiratory syncytia virus, and human rhinovirus,
and compare this with the response to Lipopolysaccharides (LPS). Using an anomaly detection framework we identified 16 pathways that achieve a minimum cutoff accuracy for predicting outcomes across the four different respiratory viruses H1N1, H3N2, RSV and HRV. A subset
of 8 of these pathways were identified as early warning pathways including inflammatory bowel disease, toll-like receptor signaling, Influenza A, lysosome, intestinal immune network for IgA production, HIVNEF, and NF-kappa B signaling. These early warning pathways correctly predict for H1N1 and H3N2 that almost half of the subjects will become symptomatic in less than forty hours of monitoring and that three of 18 subjects will become
symptomatic after only 8 hours. Host pathway analysis of a human endotoxin gene expression data set revealed a 14 pathway signature that identified symptomatic subjects within 2-3 hours post exposure. Comparative analysis between the prognostic bacterial and viral pathway signatures showed a single pathway, IL-22BP, that overlapped between the signatures. These results suggest that there are strong pathway signatures that characterize
the immune system\'s response to infection at its earliest stages. The identification of prognostic respiratory virus biomarkers has the potential to provide an early warning system that is capable of predicting that subjects will become symptomatic at the earliest stages of infection expanding medical diagnostic capabilities and treatment options. The immune system\'s response
to disease may be viewed as a deterministic, carefully orchestrated signaling network responsible for maintaining the health of the host organism.

Data in N-way arrays where one of the dimensions is time may be viewed as a trajectory in a very high dimensional space. The shape of this trajectory may be characterized
by generalized curvatures. Let g be a sufficiently smooth non-degenerate curve in n-dimensions. The

Frenet-Serret apparatus of g consists of a frame and generalized curvature values k1, k2, kn-1 at each point of g. The local singular value apparatus of g consists of an ordered sequence of n mutually orthogonal lines and local singular values at each point of g. We define the local singular value apparatus, show how it can be computed, and describe how it relates to the Frenet-Serret apparatus.

Distribution Statement

This is block 12 on the SF298 form.

Distribution A - Approved for Public Release

Explanation for Distribution Statement

If this is not approved for public release, please provide a short explanation. E.g., contains proprietary information.

SF298 Form

Please attach your SF298 form. A blank SF298 can be found here. Please do not password protect or secure the PDF The maximum file size for an SF298 is 50MB.

AFD-070820-035KIRBY.pdf

Upload the Report Document. File must be a PDF. Please do not password protect or secure the PDF . The maximum file size for the Report Document is 50MB.

FinalReportKirbyAFOSR.pdf

Upload a Report Document, if any. The maximum file size for the Report Document is 50MB.

Archival Publications (published) during reporting period:

1) J. Chang and C. Peterson and M. Kirby, Feature Patch Illumination Spaces and Karcher Compression for Face Recognition via Grassmannians, Advances in Pure Mathematics, Vol. 2, No. 4, 226-242, 2012.

2) D. Eklund, C. Jost, C. Peterson, A method to compute Segre classes of subschemes of projective space, Journal of Algebra and its Applications Vol. 12, no. 2, (2013).

3) D. Bates, J. Hauenstein, T. McCoy, C. Peterson, A. Sommese, Recovering exact results from inexact numerical data in algebraic geometry, Experimental Mathematics, Vol. 22, Issue 1, pg 38-50 (2013).

4) Rutherford, Blake, Dangelmayr, Gerhard and Kirby, Michael, A time-dependent Lagrangian eyewall, Quarterly Journal of the Royal Meteorological Society, Vol. 138, No. 669, John Wiley & Sons, Ltd., pp 2009--2018, 2012.

5) Draper, B., Kirby, M., Marks, J., Marrinan, T., and Peterson, C. (2014). A flag representation for finite collections of subspaces of mixed dimensions. Linear Algebra and its Applications, 451, 15-32.

6) Daniel Bates and Davis, Brent and Kirby, Michael and Marks, Justin and Peterson, Chris (2015), The max-length-vector line of best fit to a set of vector subspaces and an optimization problem over a set of hyperellipsoids, Numerical Linear Algebra with Applications, Vol. 22 pp 453--464.

7) Stephen O\'Hara, Kun Wang, Richard A Slayden, Alan R Schenkel, Greg Huber, Corey S O\'Hern, Mark D Shattuck and Michael Kirby, Iterative Feature Removal Yields Highly Discriminative Pathways, BMC Genomics 14:832, 2013.

8) Kun Wang, Vineet Bhandari, Sofya Chepustanova, Greg Huber, Stephen O´Hara, Corey S. O´Hern, Mark D. Shattuck, Michael Kirby, Which Biomarkers Reveal Neonatal Sepsis?} PLoS ONE 8(12): e82700. doi:10.1371/journal.pone.0082700, Dec 18, 2013.

9) Emerson, Tegan and Kirby, Michael and Bethel, Kelly and Kolatkar, Anand and Luttgen, Madelyn and O'Hara, Stephen and Newton, Paul and Kuhn, Peter, (2015), Fourier-Ring Descriptor to Characterize Rare Circulating Cells from Images Generated Using Immunofluorescence Microscopy, Computerized Medical Imaging and Graphics, Vol. 40, pp 70--87.

10) Mai M, Wang K, Huber G, Kirby M, Shattuck MD, O'Hern CS (2015), Outcome Prediction in Mathematical Models of Immune Response to Infection. PLoS ONE 10(8): e0135861.

11) Arta Jamshidi and Kirby, Michael, (2015), A Radial Basis Function Algorithm with Automatic Model Order
Determination,} SIAM Journal of Scientific Computation (SISC), Vol. 37., No.3, ppA1319-A1341.

12) T. Marrinan, R. Beveridge, B. Draper, M. Kirby and C. Peterson, (2015), Flag Manifolds for the Characterization of Geometric Structure in Large Data Sets, A. Abdulle et al. (eds.), Numerical Mathematics and Advanced Applications
ENUMATH 2013, Lecture Notes in Computational Science and Engineering 103, Springer International Publishing. 457-464. European Numerical Mathematics and Advanced Applications, Lausanne, Switzerland 8/2013

13) E. Hanson, F. Motta, C. Peterson, L. Ziegelmeier, On the Strengthening of Topological Signals in Persistent Homology through Vector Bundle Based Maps, Proceedings of the Canadian Conference on Computational Geometry (2012), pg 303-308.

14) Sofya Chepushtanova, Christopher Gittins, and Michael Kirby, Band Selection in Hyperspectral Imagery Using
Sparse Support Vector Machines, submitted SPIE conference on Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XX, 5/2014.

15) B. Draper and M. Kirby and J. Marks and T. Marrinan and C. Peterson, Finding the Subspace Mean or Median to Fit Your Needs, Computer Vision and Pattern Recognition (CVPR) June, 2014.

16) Sofya Chepushtanova and Michael Kirby, Classification of Hyperspectral Imagery on Embedded Grassmannians, 6th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, June 2014, Lausanne, Switzerland WHISPERS 2014.

17) Sofya Chepushtanova, Michael Kirby, Chris Peterson and Lori Ziegelmeier, An Application of Persistent Homology on Grassmann Manifolds for the Detection of Signals in Hyperspectral Imagery, In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 2015

18) K. Wang, J. Thompson, C. Peterson, and M. Kirby, (2015)
Identity maps and their extensions on parameter spaces: Applications to anomaly detection in video},
Proceedings Science and Information Conference, pp. 345-351, London, July 28-30, 2015.

Changes in research objectives (if any):

Change in AFOSR Program Manager, if any:

Program Manager Years I-II: Dr. Tristan Nguyen
Program Manager Year III: Dr. Arje Nachman

Extensions granted or milestones slipped, if any:

AFOSR LRIR Number

LRIR Title

Reporting Period

Laboratory Task Manager

Program Officer

Research Objectives

Technical Summary

Funding Summary by Cost Category (by FY, $K)

	Starting FY	FY+1	FY+2
Salary			
Equipment/Facilities			
Supplies			
Total			

Report Document

Report Document - Text Analysis

Report Document - Text Analysis

Appendix Documents

2. Thank You

E-mail user

Nov 09, 2015 14:03:57 Success: Email Sent to: kirby@math.colostate.edu