AFRL-AFOSR-JP-TR-2018-0049

Discovering patterns in human-robot interaction: New tools for complex adaptive social systems - UQ part

Janet Wiles
THE UNIVERSITY OF QUEENSLAND

05/28/2018
Final Report

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

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
31-05-2018

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*
24 Jan 2017 to 23 Jan 2018

4. TITLE AND SUBTITLE
Discovering patterns in human-robot interaction: New tools for complex adaptive social systems - UQ part

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA2386-17-1-4007

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)
Janet Wiles, Andrea Chiba

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
THE UNIVERSITY OF QUEENSLAND
UNIVERSITY OF QUEENSLAND
BRISBANE, 4072 AU

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AOARD
UNIT 45002
APO AP 96338-5002

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

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-AFOSR-JP-TR-2018-0049

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

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The purpose of this project is to develop a set of complex systems analysis tools for analyzing human-robot interaction and social neuroscience data. Topological data analysis (TDA) is an exciting new set of techniques which allow for the quantification of topological features of data that is not easily captured by traditional geometric analyses. The AFOSR grant has ended, but the collaborative project with UCSD and AFRL/RH is ongoing, with further funding to this team provided by the Australian government. In the first year, the research team developed a novel TDA pipeline for analyzing and visualizing the simplicial structure of delay embedded time series data. This TDA pipeline was used to analyze both a social neuroscience and human-robot interaction dataset. Recurrence quantification analysis (RQA) is a time series analysis technique which allows for the visualization and quantification of the temporal patterning of when a dynamical system re-enters a similar part of phase space. The team generated RQA vectors for neural data and showed improved classification for some classes when compared with the classification of power spectral features. The Chiba lab (UCSD) and the Wiles lab (UQ) collaborated to develop a TDA and RQA pipeline and apply these methods to neuroscience, biomechanics, and humanrobot interaction datasets.

15. SUBJECT TERMS
human-robot interaction, adaptive social system, interaction dynamics, robotics, social neuroscience, cooperative teaming

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			ROBERTSON, SCOTT
Unclassified	Unclassified	Unclassified	SAR	10	19b. TELEPHONE NUMBER *(Include area code)* +81-042-511-7008

Final Report for AOARD Grant FA2386-17-1-4007

Discovering patterns in human-robot interaction:
New tools for complex adaptive social systems

Report for 23 January 2018

Name of Principal Investigators (PI and Co-PIs)
Professor Janet Wiles

- e-mail address : j.wiles@uq.edu.au
- Institution : The University of Queensland
- Mailing Address : School of Information Technology and Electrical Engineering
- Phone : +61 7 3365 2902
- Fax : +61 7 3365 4999

Professor Andrea Chiba

- e-mail address : achiba@ucsd.edu
- Institution : University of California, San Deigo
- Mailing Address : 9500 Gilman Dr., La Jolla, CA 92093
- Phone : 1 (858) 246-2508
- Fax : 1 (858) 534-1128

Collaborator: Dr Greg Funke,
Wright-Patterson Air Force Base

Period of Performance: 01/24/2017 – 01/23/2018

Abstract

The purpose of this grant is to develop a set of complex systems analysis tools for analysing human-robot interaction and social neuroscience data. Topological data analysis (TDA) is an exciting new set of techniques which allow for the quantification of topological features of data that is not easily captured by traditional geometric analyses. We developed a novel TDA pipeline for analysing and visualizing the simplicial structure of delay embedded time series data. This TDA pipeline was used to analyze both a social neuroscience and human-robot interaction dataset. Recurrence quantification analysis (RQA) is a time series analysis technique which allows for the visualization and quantification of the temporal patterning of when a dynamical system re-enters a similar part of phase space. We generated RQA vectors for neural data and showed improved classification for some classes when compared with the classification of power spectral features. In the first year of the grant, Chiba lab (UCSD) and Wiles lab (UQ) collaborated to develop a TDA and RQA pipeline and apply these methods to neuroscience, biomechanics, and human-robot interaction datasets.

Introduction

Complex adaptive systems, such as human-robot teams or interacting brain regions, are often analyzed with methods that ignore variation and the high-dimensional temporal properties that characterize the dynamics of the system of interest. An aim of this grant is to develop an Interactive Systems Toolkit in Matlab incorporating novel qualitative and quantitative analysis techniques that emphasize variation and high-dimensional temporal structure to analyse multiple data sets from the fields of human-robot interaction and social neuroscience. Many traditional geometric analyses of time series data, such as Fourier analysis, allow for the quantification of geometric features of time series into frequency, amplitude, and phase features. This year, our team has explored the application of topological data analysis (persistent homology) and recurrence quantification analysis for the dynamical analysis of time series data of human-robot interaction and social neuroscience data.

The key foundation of nonlinear dynamical systems analysis is phase space reconstruction by time delay embedding. Takens' embedding theorem is a method for reconstructing a shadow manifold in phase space from a single time series observation because of the nonlinear deterministic relationship between the variables in a dynamical system (Takens, 1981). These shadow manifolds have been shown to maintain topological similarity to the dynamical system's actual generating dynamics. When applying Takens' theorem to actual time series data, it is assumed that the observed signals are generated by attractors within the phase space of a dynamical system and encode information about other variables to which they are non-linearly related.

Recurrence quantification analysis (RQA) is a method for quantifying when a dynamical system has entered a similar part of phase space (Riley & Van Orden, 2005). A recurrence plot is a method for visualizing the recurrent patterns in phase space generated from a time delay embedding. Recurrence plots are created by putting a binary threshold on a distance matrix generated by subtracting each timestep in the time series with every other timestep. RQA allows for the quantification of relevant geometric and statistical features of recurrence plots, such as: determinism, entropy, trend, laminarity, longest recurrence diagonal, and trapping time. These RQA statistics might prove to be valuable for classifying high dimensional temporal properties of data that are otherwise not captured by traditional geometric or statistical methods.

Topological data analysis encompasses a broad set of techniques for quantifying the topological structure of data (Carlsson, 2009). Topology is concerned with connectivity of a space which, for a discrete set, forms a simplicial complex. The base unit of a simplicial complex is the simplex, a generalization of a triangle which consists of a set of points fully connected by edges. These individual simplicies are joined at their faces to form a full complex.

Point clouds, like the data used in this report, present a challenge because their connectivity is not predefined. Fortunately, it has been shown that the Čech Complex, which is formed by connecting points whose open balls intersect, is homomorphic to the manifold on which the data points lie. This means that it preserves the topology of the original manifold. Unfortunately, this complex is extremely difficult to compute and requires a very dense point cloud. To overcome these challenges, we used a technique called Vietoris–Rips (Rips) filtration which finds persistent homologies in the data (Ghrist, 2008). In a Rips filtration, points are connected if they are within a certain distance, epsilon, of one another. Then the n-dimensional holes in that complex are identified. Finally, it sweeps through values of epsilon to see at what distance these holes are created and destroyed, with the assumption that holes that last the longest are most representative of real features in the data. Holes are a useful measure of the topology of the data because they are invariant to continuous transformations. The output of this algorithm is a barcode plot or persistence diagram, which allows the analyses and visualization of when the holes come in and out of existence as the filtration radius increases (Ghrist, 2008).

Between the Wiles lab (Ferris, Back, and Wiles), the Chiba lab (Leonardis, Breston, and Chiba) and Funke lab (Tolston and Funke) our team has collaboratively analyzed data from running biomechanics, the systems neuroscience of social interaction and spatial navigation, and the intention-to-interact between a human and a humanoid robot. For description and summary of the data sets tested by the Chiba lab, see the Annual Report for AOARD Grant 17IOA006. These include social neuroscience data sets consisting of consisting of time series of simultaneously recorded local field potentials (LFP) from the olfactory bulb, amygdala, and hippocampus; RQA analysis of both actual data and a Wilson-Cowan model of comparable olfactory bulb dynamics; and an open data set on human-robot interaction (Kato, Kanda, & Ishiguro, 2015).

The remainder of this report will detail the methodological contribution of the Wiles lab to the systems algorithmic development.

Methodology

Two primary branches of topological data analysis were explored and a range of existing packages tested, as well as new algorithms developed and coded in Matlab and Python:

1. Persistent Homology: Given a point cloud, a distance function, and a maximum dimensionality as its input, it calculates the Betti numbers of the point cloud's topological structure as it becomes locally connected on different scales. By constructing a filtration (a strictly ordered sequence of simplicial complexes), topological features can be observed to appear and disappear at particular filtration radii. These patterns can be visualised using Barcode Plots or Persistence Diagrams, or fed into machine learning algorithms using Persistence Landscapes. Computing Persistent Homology is computationally expensive, especially in analysis of dense clouds which can produce many high-dimensional simplexes.

2. The Mapper algorithm (https://research.math.osu.edu/tgda/mapperPBG.pdf) is a more computationally tractable summary of topological structure, which takes as input a point cloud, one or more level set functions and ranges over those points, and a clustering algorithm, and returns a reduced topological summary of the point cloud. Mapper sections the point cloud into overlapping regions determined by the level set functions, within which it then performs clustering. Where Persistent Homology quickly becomes computationally intractable as the maximum simplex dimensionality and number of points increase, Mapper can operate on larger datasets and higher dimensions with a more reasonable computational overhead. A major advantage of Mapper over Persistent Homology's Betti number summaries is that we can visualise the structures involved and identify "flares" in the structure. In time series data, these flares may correspond to different kinds of actions or states which are non-cyclical and which algebraic topology is uninterested in but geometric topology is.

JavaPlex: The JavaPlex MATLAB package (https://appliedtopology.github.io/javaplex/) was used for computation of Persistent Homology and Persistent Cohomology. The code is written in Java and is accessed by MATLAB through external connections. While useful for smaller analysis tasks or initial investigation of topological properties, JavaPlex is inherently limited in that it is implemented in Java, a slow language which runs on a virtual machine. See Figure 1.

Initial investigations involved analysing the topological structure of high-dimensional time-delay embeddings of low-dimensional neural data (from Chiba lab). This revealed structure in the Wilson-Cowan model data. However, it was ineffective for the data from neural recordings, which has considerable real world noise, and did not initially reveal any regularity, due to the significant variation both in frequency and amplitude. This essentially "filled in" the space which would have been identified as a topological hole (or temporal loop), making it impossible to identify. Further investigation indicated that while JavaPlex is useful for certain tasks, it is not the most useful tool in TDA for analysing the kinds of time series datasets we are working with. See Figure 2.

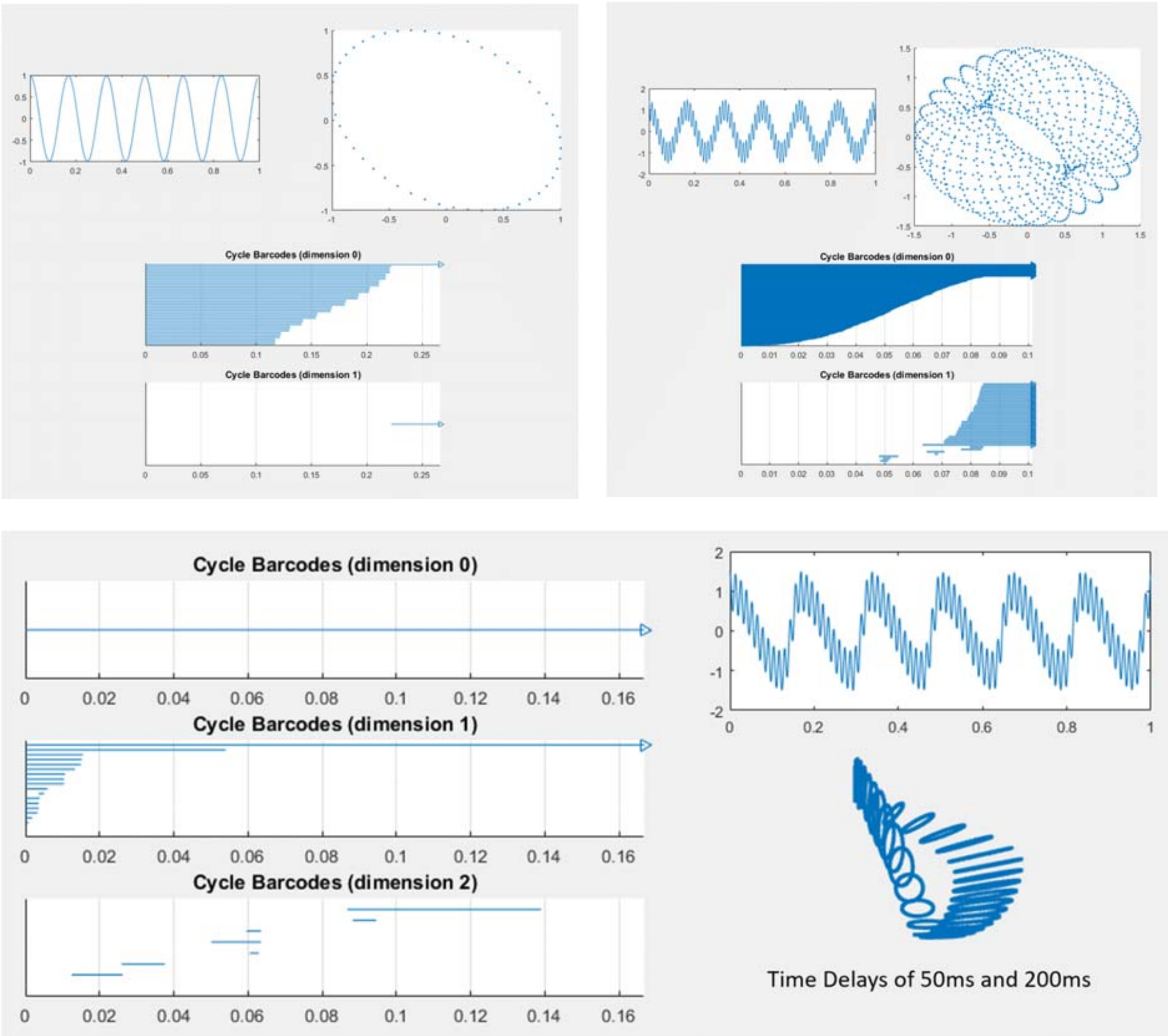

Fig. 1. Example signals with cyclic structure, and barcodes. (top, left) an individual cycle is recovered and has a single persistent barcode of dimension 1; (top, right) nested cycles are not recoverable from 2D time delay embeddings; (bottom) embedding in 3D with two time delays recovers the nested cycle with a barcode of dimension 2 (a witness filtration was used for computation efficiency).

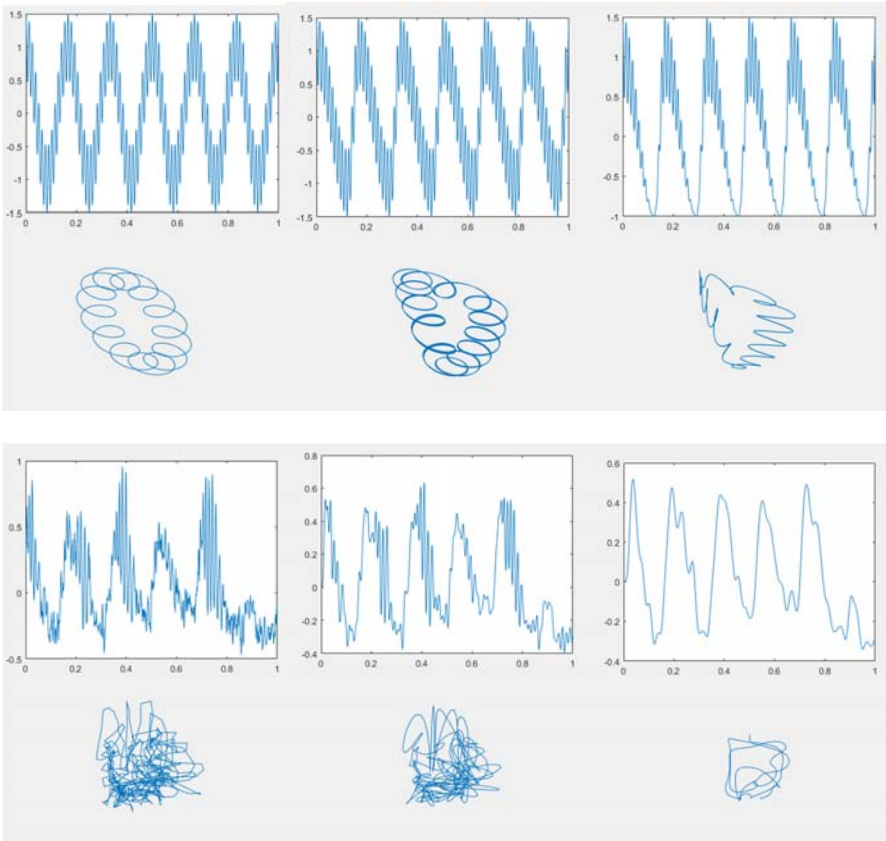

Figure 2. Nested cycles and toroids; (top) Different forms of idealised neural cycles, with additive and multiplicative nested structures, and the recovered toroidal forms; (bottom) Neural recordings with low pass filtering.

A novel finding (TDE-PCA): Despite the noise in the neural recordings, temporal structure was clearly revealed using recurrence plots with time delay embeddings. Taken's theorem indicates that high dimensional structure can be recovered from a minimal number of embedding dimensions. Due to the structure of the data, we experimented with oversampling the embedding dimensions. These extended forms of embedding revealed an interesting finding, that pairing particular PCA axes of a sliding window time delay embedding of a signal resulted in extremely structured circular patterns. We called this technique *time delay embedding principal components analysis* (TDE-PCA). From projections onto these principal components, we could calculate an amplitude and a frequency for the pattern being revealed by that "kernel" as it moved over a sliding window. This technique was further investigated and found to be highly effective at separating amplitude-modulated waves, and also able to track frequency modulation within a reasonable tolerance. This is being written up as an exploratory paper.

Dionysus 2: The Dionysus 2 Python package (http://www.mrzv.org/software/dionysus2/) was used for computation of Persistent Cohomology. Dionysus 2 is highly optimised and written in C++ (a faster, lower-level language than Java), with Python handles into the C++ code. Dionysus 2 is also able to calculate circular coordinates from particular cocycles found by persistent homological analysis (using the technique from De Silva, V., & Vejdemo-Johansson, 2009 in https://arxiv.org/pdf/0905.4887.pdf). For a point cloud which is generated from time series data, this allows us to track the behaviour of a coordinate around a particular 1-cycle over time, and use the acquired topological insight to examine time series behaviour. See Figure 3.

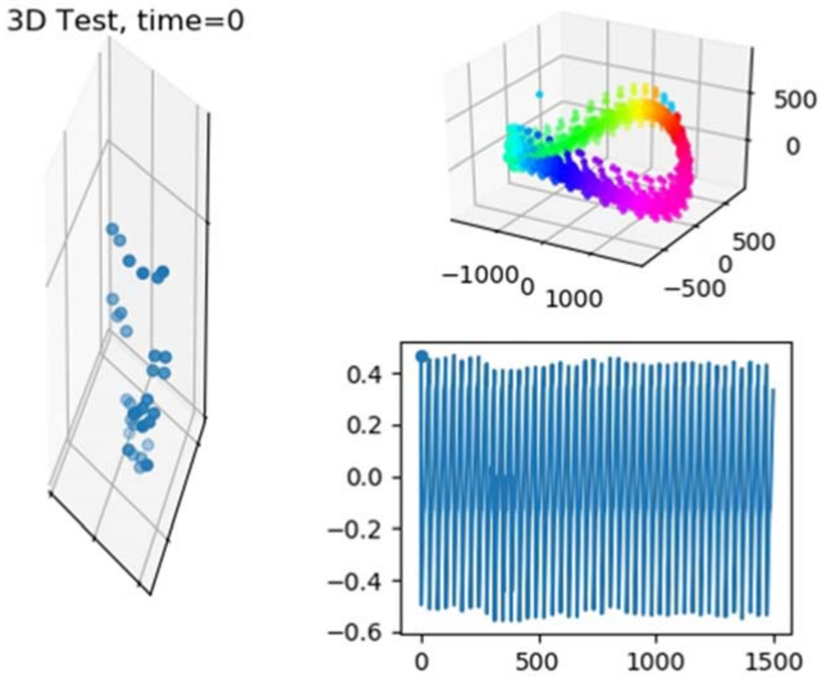

Figure 3. Coordinatisation of cyclic movement. (left) One frame from a time series of a motion captured runner showing the sampled data; (right top) Once a cycle has been identified, we can parameterise a cyclic coordinate around it; (right bottom) Moving back into the time domain shows how this coordinate varies over time.

Cocycle parameterisation was successfully applied to human walking gait data. By finding the most persistent cocycle, parameterising around it, and then looking at how that parameter changes over time, we can automatically determine statistics such as the relative speed of each step, or by unwrapping the radial coordinate in time, track the number of steps that have been taken.

Cocycle parameterisation was also used to identify and parameterise around recovered cocycles in a 351-dimensional neural data set of a rat running a spiralling maze. This recovered 185 unique circular coordinates. Using a subset of four coordinates, the shape of the maze was recovered.

This approach of cocycle smoothing and parameterisation seems especially applicable to identifying and monitoring cyclical temporal behaviours, but can also be used for localisation of high-dimensional information as well. This change to cyclical coordinates (which may be distributed among many traditional spatial dimensions) has the potential for many difficult results in standard analysis techniques to fall out trivially from the construction methodology.

Mapper in Matlab: The Mapper algorithm was implemented in MATLAB, allowing for interactive examination of datasets and the structures output from the algorithm. This code is being further developed as an open source Mapper Toolbox for MATLAB, allowing for analysis of arbitrary datasets in an interactive and intuitive way. This will serve a similar purpose to an existing package called Python Mapper, but will be implemented in MATLAB rather than Python. Python Mapper's GUI is primarily focused on one-dimensional topological summaries of data rather than multidimensional ones, and not designed for analysis of time series datasets. In the future we will continue to develop this software with the intention of using it to analyse our own data and open-sourcing it for others to use as well.

Finding cycles in non-uniformly distributed data: Real world data has statistical properties and is often clustered. Complementing the "pure" TDA approaches mentioned above, we have also been exploring probabilistic persistent homology. This approach results in a significant reduction in computational complexity, deriving the Vietoris-Rips Complex directly from the probability density function (pdf), and easily integrates with traditional approaches using Gaussian mixture models and expectation maximization (EM). This work is currently being written up for publication.

Discussion

Working with our international collaborators, the year-long program has resulted in the development of RQA, TDA and visualization pipelines, and our explorations have shown the use cases of each of these forms of analysis, and developed novel algorithms. While TDA has been shown to be useful for highly specific applications to point clouds with a low number of data points, its usefulness for time series requires further analysis, particularly of appropriate filtering. RQA is a promising method for generating previously unaccounted for features of the data that have relevant information for classifying behavior.

In this one-year project (the first year of a three year collaboration), strong links were established between two teams of researchers in the USA and the team in Australia, who took complementary approaches to the preparation of data sets, development and evaluation of algorithms based on recurrence analysis and topological data analysis. Reciprocal visits and regular online research presentations were held throughout the year. The Australian team will continue to be funded for another two years under DST funding.

List of Publications and Significant Collaborations

Papers are in preparation describing the algorithms and analyses of a range of data sets.

Significant Collaborations

For the month of November 2017, cognitive science and neuroscience Ph.D. students from the Chiba Lab worked at the University of Queensland (Brisbane, AU) as visiting research scholars in Wiles' Complex and Intelligent Systems (CIS) lab under the AFOSR International Supplemental Student Exchange Program. During this training period, the Chiba lab and the Wiles lab were able to examine the topological features of well-understood datasets, develop approaches and plan future analyses. Breston and Leonardis worked together with Ferris to analyze Nitz's neural data of single unit-recordings from parietal cortex, hippocampus, and premotor cortex during the navigation of a spiral maze. Breston and Leonardis also worked with Ferris to analyze a dataset in human biomechanics and to study the topology of human running gait at different speeds.

Throughout the year, the Chiba lab (Chiba, Leonardis, and Breston) have had bi-weekly Skype meetings with Gregory Funke and Michael Tolston from the Air Force Research Laboratory (AFRL) as well as Ferris and Wiles from the Wiles Lab. In these meetings, participants from the Chiba, Wiles and AFRL have been discussing the nuances of their applications of both TDA and RQA, while keeping note of the idiosyncrasies that come along with applying these exploratory methods to specific datasets. They have been invaluable for their guidance in properly conducting the recurrence quantification analysis and exploring topological data analysis. Having their input and expertise has been highly beneficial to the graduate students funded on the training portion of the grant, and our team looks forward to future collaborations.

The Australian team also met with members of Robert Bolia's team from DST group, and will continue the project under DST funding for another two years.

References

Bauer, U. (2015). c Ripser github. com. *Ripser/ripser*.

Carlsson, G. (2009). Topology and data. *Bulletin of the American Mathematical Society*, *46*(2), 255-308.

De Silva, V., & Vejdemo-Johansson, M. (2009). Persistent cohomology and circular coordinates. SCG’09, June 8–10, 2009, Aarhus, Denmark. Retrieved from https://arxiv.org/pdf/0905.4887.pdf

Ghrist, R. (2008). Barcodes: the persistent topology of data. *Bulletin of the American Mathematical Society*, *45*(1), 61-75.

Kato, Y., Kanda, T., & Ishiguro, H. (2015, March). May I help you?: Design of human-like polite approaching behavior. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction* (pp. 35-42). ACM.

Riley, M. A., & Van Orden, G. C. (Ed.). (2005). *Tutorials in contemporary nonlinear methods for the behavioral sciences*. National Science Foundation.

Wilson, H. R., & Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical journal*, *12*(1), 1-24.