CS486C – Senior Capstone Design in Computer Science

Project Description

Project Title: NeuroLight: A graphical workbench for analyzing neuronal activity in brain imagery	
Sponsor Information:	**Dr. Charles Allen** Oregon Health Sciences University allenc@ohsu.edu **Dr. Eck Doerry** Collaborative Computing Lab (CCL) School of Informatics, Computing, and Cyber Systems Northern Arizona University

Project Overview:

Human society has rigid schedules for work, school, and social activities. These schedules frequently require people to work and be active when their endogenous circadian clock signals that it is time to sleep. Examples include shift workers in occupations like health care, maintenance, or air traffic control. At the same time, increasing evidence has shown that synchronized circadian rhythmicity is necessary for overall health; circadian disruption causes early death and notable changes in the brain, metabolism, and behavior. In humans, the importance of circadian rhythms is most apparent behaviorally when trying to sleep at an inappropriate circadian time and most evident clinically with the significant detrimental health effects of chronic circadian disruption. Given the importance of appropriate circadian rhythmicity to human health, one of the most critical questions in circadian biology is how rhythms are generated and synchronized in the brain. Exploring and understanding how the brain produces and adapts circadian rhythms is the primary focus of our work in the [insert lab name] focus here at OHSU.

In mammals, the neuronal timekeeping mechanism that serves as our circadian clock, called the *master circadian oscillator*, is located in a tiny part of the brain called the *hypothalamic suprachiasmatic nucleus (SCN)*, which aligns the phases of secondary circadian clocks situated throughout the body. The SCN, a small bilateral nucleus, contains neurons and astrocytes, each expressing a molecular circadian clock consisting of transcription-translation feedback loops synchronized to a common period via intercellular communication provided by the SCN neural network. Intercellular communication between individual SCN neurons, mediated by the SCN neural network activity, is critical for generating precise circadian timing signals, stabilizing the circadian clock, and determining an animal's behavioral circadian phenotype.

The long-term goal of our research program is to identify and characterize the exact neuronal signaling mechanisms that mediate the strength of coupling between individual suprachiasmatic nucleus (SCN) neurons. Understanding these mechanisms in detail would allow in-depth study of how they respond to sudden shifts of the day-night activity cycle (e.g. through shift work, or travel to distant time zones) and, ultimately, could lead to

innovative treatments (e.g. pharmaceuticals, light therapies) that allow the circadian cycle to be rapidly adjusted. Imagine taking a pill immediately after your transatlantic flight that quickly "resets" your body's circadian clock, eliminating days of jet lag!

In practical terms, the key to understanding how the SCN regulates circadian rhythms and how it responds to changes in day-night cycles is to directly examine the patterns of neuronal activity within the SCN. Working with a special genetic strains of Venus mice, and using complex and highly precise laboratory techniques, we are able to extract time-series microscope images of SCN activity, by attaching a fluorescent "reporter protein" to the genetic signature of the circadian clock gene. This effectively makes neurons expressing this gene, i.e., neurons that are currently active in regulating the circadian rhythm "light up" under black light under a microscope.

Taking snapshots through the microscope at regular (30 minute) intervals effectively creates a time-lapse movie of the SCN in action, with neurons regulating circadian rhythm slowly lighting up, peaking in a glowing constellation, and then dimming back down over the course of a seven day period.

These time-lapse movie clips are a significant achievement, a unique and high quality data resource…but they represent only the raw data starting point in our analysis. The key step in this research is to turn these movies into concrete data and this will require specialized image analysis software: the time-lapse imaging data must be processed quantitatively to unambiguously and rigorously identify rhythmic cells, investigate the circadian parameters of individual neurons, and develop analytical tools to identify networks of neurons with a high degree of synchrony. Specifically, some key features of this software tool will include:

- An analytic "workbench" to allow scientist to manage the entire analytic process, including a graphical interface (GUI) to manage the data input, choose specific types of analysis or adjust processing parameters, and quickly visualize and review resulting output.
- An analysis pipeline to determine the fluorescent intensity of individual neurons across a sequence of images taken every periodically over some time period (typically 30 minutes for seven days for this research, but the software workbench should support values/intervals).
- Output of quantitative florescence data for each neuron for further data analysis.
- A highly modular software architecture based around neuronal image processing "modules". This project will create one such module (i.e., to support our specific analysis of circadian neuronal activity), but the software workbench should allow easy addition of novel modules in the future, to be able to support any sort of time lapse image series analysis of neuronal activity useful to other researchers.
- Must include a user manual for neuroscientist end-users, as well as strong architectural design and code documentation to support easy extension by future software teams.
- Software should be made available for easy end-user download and installation via Github, PyPi or similar sites (depending on choice of implementation languages) for use by the larger circadian rhythm research community.

Fortunately, this project team will have a strong starting point for this project: various researcher-programmers over the last five years have developed some software tools to perform the required analysis. Thus, a proof-of-concept software exists. As a piece of software pieced together over time by a series of amateur scientist-programmers, however, it is inefficient, difficult to use, poorly documented, difficult to understand and maintain, and implemented as a set of command-line programs requiring extensive geek level knowledge to operate.

The primary goal of this project, therefore, is to (a) design a coherent graphical "workbench" concept easily usable by minimally trained scientist end-users; (b) analyze the existing analytic software scripts to extract the algorithmic process they implement; and (c) re-implement the software in an efficient, modular, and extensible framework embedded in the NeuroLight workbench product.

To help communicate our priorities for this project, we can roughly separate the main functional milestones of this project into three milestones:

Level 0: Minimum Viable Product: Core functional basics without which the product is useless. In this case, this would be essentially a re-write to improve the existing code, as a series of one or more command line programs.

1. Allows input of a movie clip for processing, as well as specification of key processing parameters.
2. Determines the fluorescence intensity of individual neurons during the time-lapse imaging session. This requires consistent tracking of identity and presence of each neuron across image frames in the "movie".
3. Determine the specific parameters of the circadian rhythm represented in the movie clip, including the period, the rhythmicity index, the amplitude of the rhythm, the phase, the time of each peak (acrophase) and trough of the rhythm, and the mesor. Project sponsors will provide technical support on methods to accomplish this computationally.
4. Perform statistical test(s) to determine quantitatively whether a given neuron expresses a circadian rhythm. Project sponsors will provide support for statistical algorithms to use, e.g., a Lomb-Scargle periodogram.
5. Perform a Rao and Rayleigh analysis to determine the level of synchrony between individual neurons.
6. In addition to individual cell analysis, we analyze the overall culture signal to provide an estimate of the time of day. This data represents a single signal, with the peak defined at Circadian Time 6 (CT6).

Level 1: An nicely implemented graphical workbench product: essentially this amounts to taken the key analytic processes developed for the MVP and embedding them as the analytic steps in a "toolchain" implemented in a graphical workbench concept to manage the entire workflow.

1. Graphical workspace with functions to import video clip(s) for analysis
2. Allows users to graphically construct a "workflow" of processing actions/steps, each with it's own processing parameters. Processing steps are selected from an extensible set of processing modules.
3. Allows users to set and clearly display analytic parameters, and to control and visually monitor of each processing step in a workflow.
4. Implementation of an interactive working process for users, allowing scientists to load a video clip(s), specify a (potentially multi-step) analytic workflow for processing, including configuring processing parameters, and then "step through" that workflow step-by-step, visually reviewing intermediate results, adjusting processing parameters, and re-running, until a satisfactory parameterized workflow results.
5. Allows saving of parameterized workflows; these could then be loaded later to process new movie clips in the same way. Or a scientist could share a workflow with a colleague or publish it in a publication, allowing others to process their own video clips in precisely the same way.
6. Allows exporting a "workflow results", i.e., all intermediate processed images and data readings from each step in a workflow in a coherent data package. Should allow easy further processing of data in the package using other external analysis or visualization tools.
7. Possible addition of "small world" analysis methods to identify possible intraSCN networks.

Level 2: Stretch goals. These feature goals move from practical implementation of core needs towards exploration of an innovative, next-generation vision for this scientific tool. Implementation would be impressive, but should not receive any effort until the other feature levels are completed.

1. Implementation of a "laboratory notebook" concept for scientist end-users. Allows scientists to create "Analysis Experiments", then create insert new or saved processing workflows into the experiment, then

load video-clips for analysis into the experiment. As a configured workflow is then run on a video-clip, the intermediate results/data for each step are displayed sequentially in the notebook, and the scientist is able to insert commentary or notes about each step into the sequentially evolving experiment. Inspiration for this concept can be taken from the Jupyter Notebook tool commonly used in data analysis.

2. Ability of notebooks to be securely encrypted and archived for the scientist, creating a secure but permanent record of analysis using the NeuroLight workbench.
3. Others, to be inspired by your own ideas and ours, as the project moves forward!

As a final general requirement, our aim is for this product to be accessible and usable by a maximally broad audience of scientists. Therefore, design should focus on languages, frameworks, and technologies that work seamlessly across computing platforms including Linux, Windows, and MacOS.

Knowledge, skills, and expertise required for this project:

The computer science skills required for the project will emerge over the course of early design, as the team evaluates implementation options (languages, frameworks, approaches). No specific knowledge of biology or neuroscience is needed, as guidance and education on the basics will be provided by the sponsors. However, some useful background could include:

- Knowledge of mathematics and statistics, to support the implementation of data analysis steps.
- Experience with image processing in general, including use of common image processing packages like OpenCV and similar libraries.
- Experience in design and implementation of graphical interfaces (GUIs), especially in a multi-platform context.

Equipment Requirements:

- There should be no equipment or software required other than a development platform and software/tools freely available online.

Software and other Deliverables:

- Complete software product, preferably downloadable as an easily installed software product from Github, PyPi or other appropriate online archive. Must include an automated installer or, alternatively, a clear, step-by-step readme document that ensures successful installation of the product and all required dependencies.
- A strong as-built report detailing the design and implementation of the product in a complete, clear and professional manner. This document should provide a strong basis for future development of the product.
- A nice user manual for the product, including illustrative screenshots of key dialogs, screens, or interfaces in the product, to allow scientists end-users with moderate computer use experience (but no particular technical skills) to easily learn and operate the NeuroLight workbench product.
- Complete professionally-documented codebase, delivered both as a repository in GitHub, or some other version control repository; repository ownership should be assigned to a project sponsor at completion to allow for easy further development.