

AN INTRODUCTION TO THE BIDA LAB

by

Caroline Pappas

A Senior Honors Project Presented to the

Honors College

East Carolina University

In Partial Fulfillment of the

Requirements for

Graduation with Honors

by

Caroline Pappas

Greenville, NC

December, 2019

Approved by:

(Dr. Sunghan Kim)

(College of Engineering)

An Introduction to the BIDA Lab

Resources Compiled by Caroline Pappas and Lana Wang

Updated: Nov. 24, 2019

Table of Contents

1. The BIDA Lab Overview
 - 1.1. Dr. Kim
 - 1.2. Your Role in the Lab
2. Research Focus
 - 2.1. Event Related Potentials
3. Measuring Methods
 - 3.1. BCI2000
 - 3.2. g.SAHARA
 - 3.3. USBamp
4. Software
 - 4.1. MATLAB
 - 4.1.1. EEGLab Toolbox
 - 4.1.2. FieldTrip Toolbox
 - 4.2. Virtual Private Network and Virtual Machine
 - 4.2.1. VMWare – Student
 - 4.2.2. Cisco AnyConnect
 - 4.3. The Donders Institute
5. The Donders Institute
6. Basic Vocabulary and Concepts
 - 6.1. Event Related Potential
 - 6.2. Step-Wise Linear Discriminant Analysis
 - 6.3. Receiver Operating Characteristic Curve
 - 6.4. Area Under Curve
 - 6.5. Ensemble Sparse Classifier
 - 6.6. Independent Component Analysis
 - 6.7. Principal Component Analysis
 - 6.8. Scalp Current Density
 - 6.9. Time-frequency Analysis
7. Getting Started

Thank you for your interest in joining Dr. Kim's Lab! We all appreciate your commitment and hope you can begin your own study very soon.

This manual is a compilation of resource materials and instructions via text, images, websites, and videos you can utilize to become familiar with the softwares you will be using. This is only an introduction and will not cover 100% of the knowledge needed for further experiments, but it will give you a strong start. If you have any questions as you go through this material, please feel open to ask the current lab members as well as Dr. Kim. If you cannot access any link in this manual, please contact your lab partners for assistance.

1. The BIDA Lab Overview

1.1. Your Principal Investigator

Dr. Sunghan Kim is a Biomedical Engineering researcher who specializes in the interest of Event Related Potentials. His lab has access to patients from the Brody School of Medicine and Vidant Health in order to study different cognitive functioning deficits within the nervous system. Cognitive functions in this case are considered events such as object or facial recognition. The main lab is currently located in The Science and Technology Building in Room 134.

1.2. Your Role in the Lab

Your role in the lab is to help collect, analyze, interpret, and present electroencephalogram (EEG) data through different software methods. New tasks will be assigned to you each week and it is critical to have a grasp on time management. Weekly action items cannot be done in one night; they should be performed in short time increments, i.e. a couple hours every day as needed.

2. Research Focus

2.1. Before starting, please read through the PowerPoint in the same folder as this manual for information on bioelectric signaling and EEGs.

2.2. Event Related Potentials

Event Related Potentials (ERPs) are bursts of activity in the brain by the neurons as a result of a stimulus. They are often indicative of object recognition, counting tasks, spelling trials, and many other methods. There are multiple ERP components that occur at their own respective time marks or latency. Each component is thought to correlate to a difference function. Common ERP components are the N170, P300, and the P600. Components are labeled based on where they occur and the type of deflection.

For example, the N170 occurs as a negative deflection around the 170ms time tick while the P600 occurs as a positive deflection around the 600ms time tick. The P300 is a common ERP component that is analyzed typically appearing around the 300ms mark of an EEG recording and measures the subject's reaction to a stimulus. Figure 1 shows a

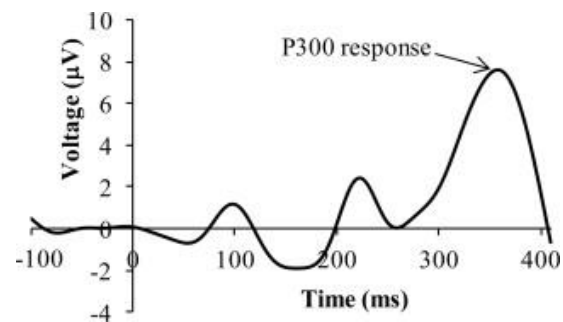


Figure 1: P300 Response behavior as shown by voltage over time.

4. Software and Toolboxes

4.1. MATLAB

MATLAB is a statistical software that you will be using for a majority of data analysis . [Here](#) is a major video tutorial with timestamps in the description about MATLAB and its basic functions.

4.1.1. EEGLab

The EEGLab toolbox is used for pre-processing and Independent Component Analysis (ICA) of EEG data (See Independent Component Analysis). Pre-processing involves the utilization of filters to remove signal noise as seen in (Figure 4). Common filters that are used are the high pass, low pass and common average reference (CAR) filter. After filtering, the data is epoched. Specific time windows in the raw EEG signal are extracted and are averaged per stimulus type. The time windows of interest are the segments following a presentation of stimulus. An example of epoching is shown in (Figure 5). In this example, X and O mark the post-stimulus frames for a specific stimulus. X can represent the familiar face stimulus while O can represent the unfamiliar face stimulus. All the X epochs are extracted and averaged to produce an averaged ERP for the familiar face stimulus. All the O epochs are extracted and averaged to produce an averaged ERP for the unfamiliar face stimulus. Epoching is performed for raw EEG signal collected for each electrode site. After pre-processing, an averaged ERP is produced for each stimulus type per channel.

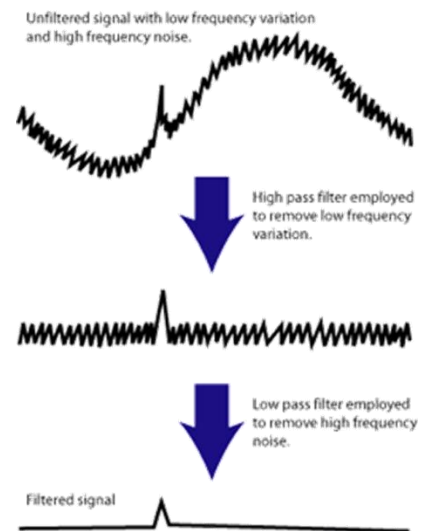


Figure 4: High-pass and Low-pass filter application.

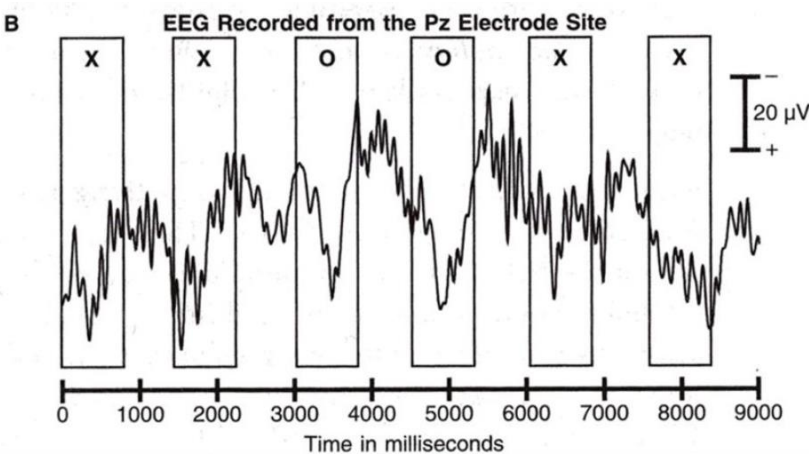


Figure 5: Data epoching where X represents a familiar face and O represents an unfamiliar face.

4.1.2. FieldTrip

The FieldTrip toolbox is used for post-processing of EEG data. [Here](#) is a video explaining the functions of FieldTrip and EEGLab combined. Post-processing involves many methods that can be used to analyze the pre-processed data. Data analysis can be performed in the time domain or in the frequency domain.

4.2. Virtual Private Network and Virtual Machine

A Virtual Private Network (VPN) is a type of software that allows sensitive information to be encrypted and sent over networks that may not be the safest. A Virtual Machine on the other hand, is a secondary desktop set up on your personal computer that can access files through a VPN. Virtual Machines are essentially the computers used in the lab that can be remotely accessed through your laptop at home. Medical professionals often use virtual machines to do patient charts at home securely. All data stored on a virtual machine can be retrieved from another device later. The VPN software should already be downloaded onto your laptop if you have taken ENGR 1012.

4.2.1. VMWare – Student

This is the Virtual Machine that will allow you to have access to shared lab files from your personal computer. In the BIDA Lab, you will be working with sensitive health information that must be kept private from those outside the lab. First, you will need access to a virtual private network (VPN) before accessing the Virtual Machine itself.

4.2.2. Cisco AnyConnect

This is the VPN you will use to run all MATLAB code and save files to. A virtual machine is like having a second computer space on your laptop. It allows you to run programs safely and securely on a cloud server away from the hardware of your real PC. Before you can run the code for yourself, Dr. Kim will have ECU's ITCS grant access to his lab's database to you. The process can take up to a week. Once you are granted access, you will get an email from Dr. Kim with instructions on how to access the Virtual Machine. The Virtual Machine will not run unless you have your VPN installed and activated first.

5. The Donders Institute

The Donders Institute is a research team that provides the methods we focus on using in the lab. Their codes go through the preprocessing aspects of Time-Domain preprocessing, Frequency-Domain preprocessing, Time-Frequency Domain processing, Linear Constrained Minimum Variance Beamforming (LCMV), and finally Dynamic Imaging of Coherent Source filtering (DICS). The main goal of you as a new lab recruit is to successfully run the Donders Institute codes and understand the basics of code functions. You will acquire the knowledge to edit and run this code to your own data. These codes explain different methods that are used for pre-processing and post-processing. Here is the [ultimate reference](#) for the type of work you will be doing in the lab.

6. Basic Vocabulary and Concepts

- 6.1. **Event-Related Potential (ERP)**: To elaborate on ERPs, the more intense the peak, the more recognition is said to have occurred.
- 6.2. **Stepwise Linear Discriminant Analysis (SWLDA)**: Achieving an overall linear model to represent all stimuli through linear regression and pick out significant variables.
- 6.3. **Receiver Operating Characteristic Curve (ROC)**: How good a model or method is at distinguishing between two different classes. A curve of probability.
- 6.4. **Area Under Curve (AUC)**: Used to discuss the effectiveness of a method when examining its ROC. An AUC of 1 is the best outcome; the method being used for distinguishing between one group of subjects and another is near perfect. An AUC of 0.5 indicates that the measure being used has no statistical significance. An AUC of 0 indicates that the measure being used results in the reciprocal of its prediction.
- 6.5. **Ensemble Sparse Classifier (ESC)**: Data comes with a high number of features, most of which, may not be significant. ESC can be used to pick out the most important features for their significance. Like trying to spot the difference between two datasets.
- 6.6. **Independent Component Analysis (ICA)**: ICA is used to separate sources from EEG waves that can become mixed together. The Gaussian trend, or showing the function of a bell curve, is another important factor. For example, random plotted points can be put on a graph to show signals from two different EEG channels being bell curves. ICA rotates this graph to give all areas on the graph statistical significance. This method is used over others, such as Principal Component Analysis because the order of inputs makes no difference on the outcome, and they don't have to be based on right angles (orthogonal). Each vector is an independent component of your data.
- 6.7. **Principal Component Analysis (PCA)**: PCA is used for smaller sets of data to determine if a correlation exists between two variables. This process depends on the input order of data and is based off of orthogonal trends. We do not use this method in the lab for data analysis.
- 6.8. **Scalp Current Density (SCD)**: A topographical map of the head that shows heightened activity according to condition of the subject or time recorded. For example, this paper on the advancement of neuroimaging shows that alcoholics have less ERP activity in the brain when compared to controls. The redder an area is, in this case, the more activity is being recorded by electrodes.
- 6.9. **Time-Frequency Analysis**: This tool looks at the time and frequencies of EEG waves. It identifies a time where various frequencies are present and helps you identify times where activity is occurring.

7. Getting Started

MATLAB: <https://itcs.ecu.edu/matlab-statistical-software/>

EEGLab toolbox: <https://sccn.ucsd.edu/eeglab/downloadtoolbox.php>

FieldTrip toolbox: <http://www.fieldtriptoolbox.org/download.php>

NOTE: This toolbox link will take you to a database of all updated packs of FieldTrip. You will need to download the most current dated link.

VPN: <https://itcs.ecu.edu/services/network-wireless/virtual-private-network/>

Virtual Machine: https://my.vmware.com/web/vmware/info/slug/desktop_end_user_computing/vmware_horizon_clients/3_0#win

NOTE: I have compiled videos for steps 1-6 below in [this playlist](#).

1. Download all links above in order. (NOTE: You will only need to download MATLAB on your personal laptop, it comes pre-installed on your virtual machine later).
2. Enable Multi-Factor Authentication on your ECU Outlook account.
3. Wait to be granted access to the BIDA lab data via VMWare (this can take a few days, you will be notified through an email.)
4. Open your VPN and enter both your credentials and an MFA code.
5. Open VMWare and sign into the Student Dedicated server.
6. Map the BIDA lab network drive: `\\cet-mako\kims`
7. Go to the Lab Tutorial folder on the shared network.
8. Open the Codes and Data folder and read through the Annotated Code to learn the basics of coding language.
 - a. If you get an error message (See to the right) when trying to open a code file, follow [these](#) instructions.
9. After reading through parts A-C, open the Practice code and begin to work with the one data file given to you in the same folder.
 - a. You may have to [add FieldTrip or EEGLab to your path at the beginning](#).
 - b. If you get an error message about a certain function not being found, go to the Auxiliary Files folder and add whichever function was not previously available. Talk to Dr. Kim about missing these functions later.
 - c. Some newer versions of FieldTrip or EEGLab may have the same function but with a changed name.
 - d. Refer to the Annotated Code side-by-side if you have further issues.
10. Save images of generated images from the Practice Code and show Dr. Kim.
 - a. After running Code A, your Workspace should look like **(Figure 7)**.
 - b. After running Code B, you should get an image like **(Figure 8)**.
 - c. After running Code C, with the Event set to (1), you should get an image like **(Figure 9)**.

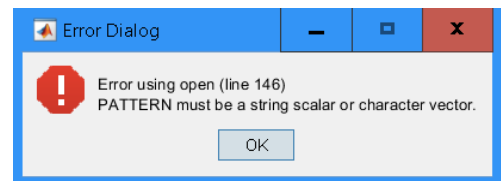


Figure 6: An error message you may see.

NOTE: The explanation and purpose of these codes is within the code itself.

