LS/MCB100 | Computational Ethology

WEEK 03 | INTRODUCTION TO COMPUTATIONAL ETHOLOGY

Souvik Mandal, PhD
Post-doctoral fellow
Center for Brain Science / Molecular and Cellular Biology
Harvard University

Introduction to

COMPUTATIONAL ETHOLOGY

- **Like most of the other modern fields of research, ethology is also taking advantage of the rapid progress occurring in the field of computation (most prominently, in machine vision, artificial intelligence, and data science).**
- **Computational ethology is an interdisciplinary field in which practitioners conduct automated measurement of animal behavior using advanced electronic and computational tools to acquire and process vast amounts of digital data, and make sense of it much quicker than executing these tasks manually.**

Usage of modern computation in ethological studies generally include

- **Data collection**
  - high-throughput - can be done reliably for long duration (without worrying about fatigue and associated errors)
  - precise across time
  - without bias and consistent across animals and human observers.
- **Data arrangement/processing and analysis**
  - faster and accurate - saves researchers from the boredom of doing repetitive tasks (and allows them more time for creative tasks),
  - real-time - therefore, creating a feedback loop of “behavior shown” and “next experimental manipulations” is possible. In other words, one can design an experiment in which dynamic experimental manipulations (e.g. giving different stimuli) can be implemented immediately to animals depending on their current behavior.
- **New discoveries**
  - new behaviors or behavioral elements (especially those that are too fast for the human visual system, and so not recorded before) can be detected.
  - Micro-structure of behavior - a single behavior can be deconstructed into a spatio-temporal assemblage of finer components of body postures called **behavioral syllables,** which in turn constitute higher-order **action patterns**. These **behavioral elements** are often repeated during a single bout of a behavior (so, stereotypical), and similar across individuals from a species. So, a single behavior can be defined high-dimensionally. When synchronized with neural recordings, knowledge of behavioral elements is extremely useful to understand the role of a certain part of the neural system generating certain body movements as well as a behavior. This talk by Benjamin De Bivort from Center for Brain Science at Harvard will give you an idea about this.

Ethologists use a vast range of electronic devices for collecting data on behavior - starting from different types of sensors like thermometers, RFID chips, IR beam-breaks, to recording devices like automatic audio recorder, video cameras etc. For estimating the body poses of an animal/subject, electronic probes/markers were used on the body of the subject even in the 2010s. However, tools for extracting quantitative behavioral data from videos using "marker-less pose estimation" emerged in the last five years - and this is what our course is about. Let's have an elementary understanding of digital videos.

BASICS OF DIGITAL VIDEOS

- Videos are **temporal sequence of images/image frames**
- **Digital images**
  - a matrix of cells - each cell corresponds to a so-called "pixel" on the sensor in the camera, the physical light-sensitive units on the sensor.
    - so, from the viewpoint of data, an image clicked with a camera with a 1-megapixel sensor has one million cells.
    - each image frame of a 1080p HD video is saved as a matrix of 1,920 columns (or pixels arranged horizontally) and 1,080 rows (or pixels arranged vertically), therefore with a total of 2073600 or about 2 million cells.
  - When light falls on the sensor/ pixels
    - an image is created by populating each cell with a certain value.
    - The value each pixel gets is determined by the light intensity it receives and stores it in the matrix.
  - for an 8-bit per pixel (interchangeably called bpp, or just n-bit) monochrome or black & white/single-channel image, each cell in the matrix has a value ranging from 0 to 255 for (a gradient of greyscale, 0 being absolute black and 255 is absolute white).

- Similarly, a 1-bit image has two values (0 for black and 1 for white), a 2-bit image has four values (00, 01, 10, and 11), and an *n*-bit image will have 2^n values.
- For RGB color images, each cell has three values (one value for each color channel - Red, Green, Blue). So, a 24-bit color image, often called a true-color image, has 8 bit assigned for each color and potentially can have 2^{24} = 16777216, or about 16.7 million colors. Click here if you want to know more about this concept.

Challenge: Can you guess the size (in Megabyte) of an image clicked with a 8-bit 1 megapixel camera? (8 bit = 1 byte)

- **Important features of a video file**
  - Spatial resolution of the video frames (in pixels)
    - More the resolution, or bigger the frames - you have more pixels dedicated to your subject, which typically increases the quality of the data for behavioral analysis.
    - However, bigger frame means larger file size and that incurs slower computation.
    - We often either "downsample" or crop the video to exclude information that are unnecessary for our experiment.
  - Temporal resolution
    - shutter speed - the time for which light is allowed to fall on the sensor to create an image. Typically expressed as a fraction of a second (looks like ¼, 1/125) for artistic photography and videography, and in milliseconds for scientific purposes. Longer the duration is (or slower the shutter speed is), the sensor receives more light and thus, produces a brighter image. But it creates motion blurriness if the subjects or the camera move. If you are shooting a video for measuring behavior, you want to get razor sharp images of your focal animal in every frame. In other words, you want to avoid motion blurriness and therefore, need to have fast shutter speed. Most of the advanced video recorders or digital still-image cameras allow you to control the shutter speed manually. If you are shooting with a device with automatic shutter speed controls (like most of the smartphones/ action cameras), shoot at a well-lit place and in action/sports mode to avoid motion blurriness.

Challenge: Can you identify the cells in an image matrix with a certain (or a range of) pixel values?

VIDEO DATA RELEVANT TO MEASURING BEHAVIOR

- The amount/number/movement of any visual stimuli can be “data”.
- Spatial position of an animal/subject/body part in a frame.
- One coordinate value for the whole animal (one x, y value in a 2D space and, one x, y, z value in a 3D space, generally corresponding to the center of mass of the animal’s image) from each frame.
- Similarly, spatial position of different body parts of an animal - therefore, one x, y or x, y, z value for each body-part and several such coordinate values for one animal from each frame. This is what we are going to do for this course.

Challenge: Can you get the coordinates of the light blob in this video?

How to define a behavior?

- Definition should be descriptive, objective, and as complete as possible. Describe the behavior empirically - what are the body parts involved in generating the movements? How many elements are there in that behavior? Describe the form, velocity, duration, amplitude, orientation of constituting elements/body parts.

  E.g., behavior 1: when the stimuli were presented, the animal was not moving for *s* seconds, face was at least *n* degrees away vertically/horizontally from the stimuli, no tail wagging etc. Avoid functional description (e.g., the animal didn’t pay attention to the stimuli) or attempting to explain proximate or ultimate causes of the behavior, as such a description can bias the conclusion of the experimenter. In fact, ethologists collect data to infer the function or causation of a behavior through hypothesis testing.

- Do not worry if you are not sure about the exact elements you have in a behavior you are interested in. We can “explore” behaviors and its elements using computation.
- It is always preferable to have a description of the surrounding parameters (time of the day, weather, location, presence of other animals/objects etc), and associated context (e.g., age, sex of the animal) or stimuli (food, smell etc). Try to make sure that if you give the description to someone else, s/he will denote the same behavior. For computational ethology, you need to describe the behavior in a way that a computer can “read” it.

FROM VIDEO TO BEHAVIOR: OVERVIEW OF THE WORKFLOW

- **Pinpoint your question** - what do you want to know?
- **Finalize your experiment design** - What data do you need to answer the question? **Define** the behavior you need to record (see next page).
- **Capture the videos**. In an ideal case, you will have
  - many videos - more (number and duration) is better,
  - the focal animal(s)/subject in sharp focus,
  - No motion blur - can be achieved by recording the video in high shutter speed/in bright light/ sports mode,
  - good contrast between the subject vs background, and among different body parts of the animal,
  - more or less uniform lighting conditions across frames and videos
- **Train a "neural network"** to extract the coordinates/positions of body-parts in each frame
  - Extract some sample image frames from the videos
  - Manually label the body parts of your interest in the image frames extracted from the videos - can be done on your personal computer
  - Let the neural network "learn" the pattern - thus, get a trained model. Computationally heavy, therefore a computer with a GPU or to access the GPU on a cloud-computing facility is preferred. For this course, we will use Google Colab.
  - The first trained model is often needed to be refined or "retrained" for a better prediction power on novel frames.
  - Repeat the retraining until you get a satisfactory model.
  - There are several software that can be used for this step, but we will use a home-grown and world's leading tool "DeepLabCut", or in short, DLC.

- **Generalizing the model**
  - Extract positions of body-parts from all the frames from all the videos using the final model. We will use DLC for this step as well.
- **Analyzing movement from the body point data**
  - Now you can estimate/analyze movements, especially if you are not interested in extracting behavior.
- **Behavioral profiling: Supervised models**
  - If you "really know" the movement structure or the elements of the behavior you want to measure, define that behavior as a temporal sequence of body poses/movements with mathematical parameters (e.g., as changes of angles/distances among different body parts during the movement).
  - Next, you can use a supervised method to find the sequences of body poses throughout all videos from a data set.
  - There are a few recent software you can use for this task.
- **Behavioral profiling: Unsupervised models**

  A bit lengthy process but worth it when you want to "explore unknown behaviors".
  - The amplitude of movement for different body parts may be different to produce a behavior. To take into account the differences, using some kind of wavelet transformation.
  - Dimension reduction of the extracted coordinates of different body-parts - convert all the body parts in a frame to just one point. Methods like PCA, t-SNE, UMAP are most popular for this step.
  - Clustering the reduced dimension points. Here is an introductory article that gives a good idea about clustering.
  - Ideally, each cluster should represent one movement pattern or syllable. However, in real situations you (can) get more than one cluster for a single type of movement, or different syllables in one cluster. In such situations, calibrate the clustering parameters until you get satisfactory results.
  - Examine the behavior corresponding to each cluster - define them.

- Keep in mind that **each behavior** typically **consists of several syllables.**
- **Organize the ethogram** - the library or catalog of the behaviors (and behavioral elements) shown by an animal/ species.
- Use the behavioral data for your next level of analysis/work - try to correlate the behavior with different stimuli given.

Task: Polish your experimental design considering

How will you acquire the videos? Give as much detail as possible while reporting about the devices you intend to record.

Computation (lab part)	Click here to go to the course computation page and go through the Week 1 guide. Fill up your lab notebook as well.
Before the next meeting	Define your question, the behaviors you need to record to answer the question, prepare a written proposal (maximum 500 words) with a literature survey about the behavior.