fishtoolbox

Contains a set of modules to analyze and visualize data from block1 and block2 of the experiment from 2021 September.

Dependencies

Install the following packages: - graph-tool * Installation on hpc via conda base environment `bash        conda install -c conda-forge graph-tool` or via Apptainer/Singularity - fishproviz
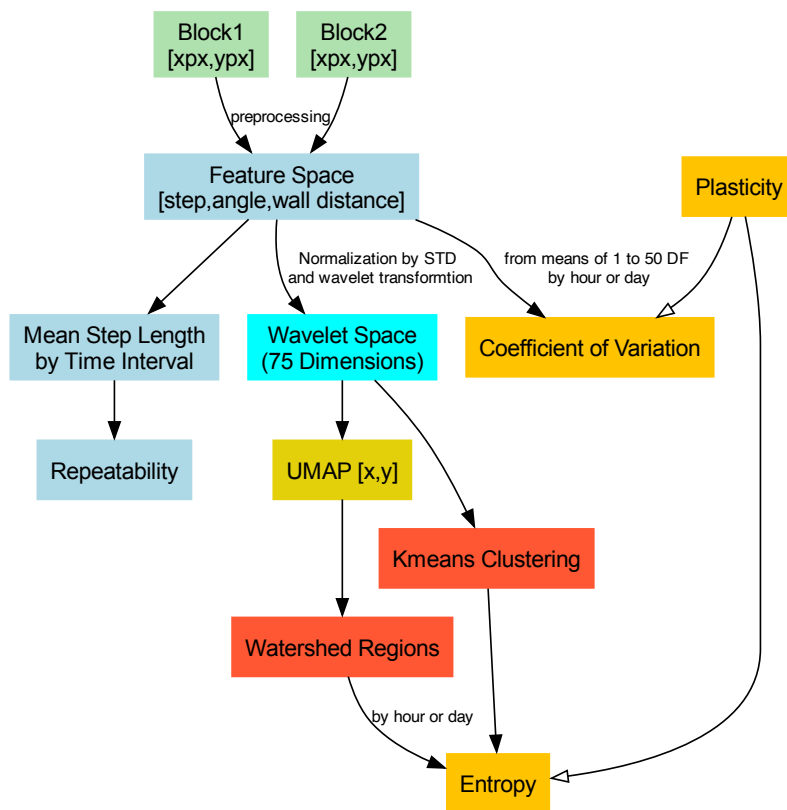
- motionmapperpy fork

Figure 1: Dataflow

Data Flow

Using the HPC cluster

1. Start on the GPU `sbatch scripts/hpc-python.sh`
2. NOTEBOOK

- `conda activate rapids-22.04`
- Type `ifconfig` and get the `inet` entry for `eth0`, i.e. the IP address of the node
- `srun --pty --partition=ex_scioi_gpu --gres=gpu:1 --time=0-02:00 bash -i` to start a new shell with a GPU
- `ssh -L localhost:5000:localhost:5000 user.name@[IP address]` on your local machine
- `jupyter-lab --no-browser --port=5000`

Start

- set the BLOCK variable to BLOCK1 or BLOCK2 in `config.py`
- set the projectPath variable to the path of a new folder in `config.py` this is where the data will be stored
- setup fishprovis with the correct paths and area configurations.
- export the preprocessed data with `python3 -m data_factory.processing`
- repeat for the other block

Parameters

- `parameters = set_parameters()` to get the parameters that are used throughout the fishtoolbox

Data Factory

Processing

- `load_trajectory_data_concat` load the x y coordinates, projections (the three features), time index, area
- `load_zVals_concat` load the umap data
- `load_clusters_concat` load cluster labels for individuals and day paramerter.kmeans = 5 to specify the clustering that you want to load.

Poltting

Caterpillar Plots

- ethnogram_of_clusters