A Cookbook for RAVE Users

Zhengjia Wang

2017-11-07

## Introduction

This is a guide book for RAVE users. This "cookbook" will cover the following topics

* A brief pipeline
* Global options
* Data hierarchy
* Modules hierarchy
* Pre-process

The first part Global options will launch a toy application.

Data/Modules hierarchy will define format for data and modules.

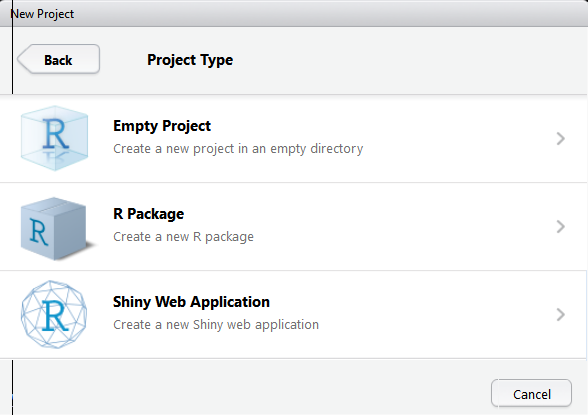
## A brief pipeline

Let's get RAVE to work first.

#### Step 1: Create an R project environment

Download [R](https://www.r-project.org/), and [RStudio](https://www.rstudio.com/products/rstudio/download/) and install them first.

After installation, open RStudio > File > New Project. Then create an empty project, save it in your local machine. In this guide book, we name it ecogr. (Everytime double-click ecogr.Rproj to open this R project)



Create an empty project - ecogr

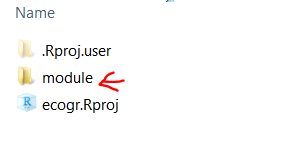
#### Step2: install RAVE

In your RStudio, TODO(Download from github)

Now, load RAVE in your RStudio:

library(rave)  
system.file('example/', package = 'rave')

Run the code and copy the result above and enter the path in your finder/browser, then copy module, data folder to ecogr. (Theoratically speaking, your data folder can be anywhere on the machine, but module folder *MUST* stays at your project folder).



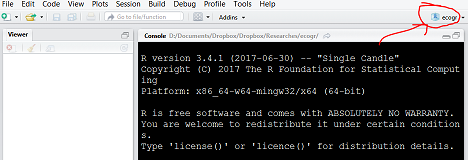
Project folder - ecogr

#### Step3: Modify modules.csv

Go to module folder, open modules.csv, modefy column source\_path to relative path (relative to project root directory). For example, your ecogr.Rproj folder path is [PROJECT\_ROOT\_DIR], and example\_module.R (inside of modules.csv) has relative path "./module/example\_module.R", change source\_path for this module to ./module/example\_module.R.

To test if this module can be found, run in RStudio (make sure that you are at ecogr project):

modules <- read.csv('./module/modules.csv', stringsAsFactors = F)  
if(!file.exists(modules$source\_path[1])){  
 message('Cannot find module.')  
 message('Please make sure that you can find ')  
 file.path(getwd(), modules$source\_path[1])  
}else{  
 message("Congratz! I found the module")  
}



Make sure your see "ecogr" on the top-right of RStudio

#### Step4: Change options

In order to let RAVE know where it can find data directory as well as modules, we need to set options

rave\_opts$set\_options(  
 data\_dir = './data', # Path to data directory  
 module\_lookup\_file = './module/modules.csv' # Path to modules.csv  
)  
rave\_opts$save\_settings() # Save settings if you are   
 # using MacOS/Linux  
 # Your settings will be loaded  
 # automatically next time

#### Step5: Setup your SUMA, Matlab, Unbuffer path

(Optional) If you want to enable SUMA, Matlab interface, you need to change some options.

#### Run app

init\_app()

## Global options

Similar to system variable, or AFNI global environment, RAVE also has an option rave\_opts that controls the global variables such as data directory SUMA location, and output path etc. It's easy to change and save.

Here are the options (*all values should be characters*). You can get their values by rave\_opts$get\_options("[KEY]") and set values by rave\_opts$set\_options([KEY]="[VALUE]")

debug: Reserved for development. (Default 'FALSE')

data\_dir: Path to data directory containing subjects' ECoG data

big\_object\_size: [*Deprecated*] Parallel computing support if folk-clusters are disabled. (Default '500000')

server\_time\_zone: Time zone. (Default 'America/Chicago')

module\_export: Path where the modules will be exported to. (Default './export')

content\_regex: Regular expression of name for electrodes. Default value is suggested. (Default 'e([0-9]+)[^0-9]\*')

content\_format: Postfix for electrodes, by default is *mat*. Other formats are under development. (Default 'mat')

module\_lookup\_file: Index file path for modules. This file is used to manage modules. (Default './modules.csv')

delay\_input: Measured in milliseconds. How long does RAVE wait to update outputs after inputs are changed. This option prevents RAVE from firing updates too frequently. Suggested value is from '20' to '200'. (Default '200')

max\_worker: How many active background jobs are allowed at the same time. (Default '20')

export\_path: Path to export modules/results. (Default './export')

temp\_dir: Path to temporary directory. (Default 'temp')

suma\_monitor\_dir: Directory to monitor inputs from SUMA. (Default 'temp/monitor')

suma\_to\_niml: Command to generate NIML data. (Default '~/abin/ConvertDset -o\_niml -input %s -i\_1D -node\_index\_1D %s -prefix %s')

suma\_send\_niml: [*Deprecated*] Command to send NIML generated to SUMA (*WARNING: this command is currently diabled due to some bugs in X11. Those bugs might crush SUMA*). (Default '~/abin/DriveSuma -com surf\_cont -load\_dset %s')

suma\_nodes\_per\_electrodes: How many nodes in electrodes\_\*.gii stands for one electrode. (Default '42')

suma\_parallel\_cores: For parallel computing, how many CPU cores will be used to calculate SUMA values. (Default '2')

suma\_gifti\_name\_regex: Regular expression to find GIFTY files that store electrodes. (Default 'electrode\_[a-zA-Z0-9]+.gii')

dyld\_library\_path: Google *"AFNI dyld\_library\_path"*". (Default '/opt/X11/lib/flat\_namespace')

suma\_path: Where to find suma command. (Default '~/abin')

suma\_spec\_file: Name for spec file in subjects' folder. (Default 'test.spec')

unbuffer\_path: Where to find unbuffer command (On MacOS, please install expect package. Default '/usr/local/bin')

matlab\_path: Where to find Matlab. (Default '/Applications/MATLAB\_R2016b.app/bin')

use\_rhdf5: This feature is still under development. (Default 'FALSE')

batch\_bytes: This number controls memory usage, only valid if use\_rhdf5=FALSE. Set this option as large as possible. However, it should not exceed half of the RAM. For example, you are running RAVE on a 64GB-RAM machine, this value is suggested to be 32000000000. (*32000000000 Bytes = 32GB*). (Default '5000000000')

## Data Hierarchy

Let's check the toy example at data folder

[ DATA\_DIRECTORY ] `./data`  
+-- SUBJECT\_ID\_01 (dipterix)  
| +-- ecog  
| | +-- cached (Optional, automatically generated)  
| | +-- meta  
| | | +-- electrodes.csv  
| | | +-- frequencies.csv  
| | | +-- time\_points.csv  
| | | +-- trials.csv  
| | +-- e1.mat  
| | +-- e2.mat  
| | +-- e3.mat  
| | +-- ...  
| +-- suma  
| | +-- rave (Optional, automatically generated)  
| | | +-- \*.csv  
| | | +-- \*.niml.dset  
| | +-- electrodes.gii  
| | +-- lh.pial.gii  
| | +-- rh.pial.gii  
| | +-- test.spec  
+-- SUBJECT\_ID\_02  
...

In our example, there is only one subject with ID dipterix. In RAVE, this folder name will be the IDs for subjects. Within each subject folder, there are two sub-folders ecog and suma.

ecog contains ECoG signals, with ei.mat for each electrode i. Within each .mat file, data should be a three-mode tensor (trials x frequencies x time). meta folder contains experiment meta information, whose format will be defined in part *Pre-process*. Notice that you don't have to create cached folder because once data is prepared, this folder will be generated by function cache\_subject

suma folder contains three GIFTY files and one test.spec. RAVE will launch SUMA via command suma -spec test.spec. In the setting-up phase, there's no need to create rave folder. It will be automatically created in the future.

## Modules hierarchy

Let's check the toy example at module folder. You can see that this folder is quite simple compared to data directory. There is one modules.csv file, as well as several R scripts. Each R scripts is a module. modules.csv is an index file that manages modules.

Modules.csv

Modules.csv

It is highly recommended that you **copy module directory in your project folder, and use relative path (relative to project folder) inside of modules.csv.**

category: Reserved for future use

module\_id: unique ID for each modules

label: Module's name

source\_path: Relative path to ecogr project, where to find module script.

active: Default TRUE. FALSE if the module is de-activated

packages: Additional packages needed to run the module. Use "," to seperate packages. For example stringr,MASS. Those packages will be downloaded automatically if not installed. However, it's highly suggested that you install them manually since this feature is still under test.

author: Who wrote the module

version: Reserved

is\_univariate: Reserved

suma\_enabled: Reserved

order: Order of modules

## Pre-process

Let's feed the real ECoG data to RAVE

TODO (Need patrick's doc)