

# Resumen metodológico

Rodrigo Pérez-Ordoyo Bellido

01 de Marzo 2020

En resumen, tras la recogida de datos y el reordenarlo para poder manejarlo (al final son N sujetos x aproximadamente 200 ensayos cada x (60 canales + 1 audio) ), podemos filtrarlo por bandas, calcular su señal analítica y posteriormente discretizarlo para calcular la información mutua entre señales.

## 1 Recogida de datos e importación

Los datos recogidos son de actividad electroencefálica muestreada a 250 Hz. La actividad se recoge con, en este experimento, 60 electrodos, teniendo 4 más para los movimientos oculares. Estos últimos 4 no son necesarios para nuestro análisis.

Los datos que David exporta desde BrainVision son datos ya con correcciones de baseline, revisados para la mayoría de artefactos, con un archivo que contiene la segmentación y filtrados entre 0 y 100 Hz.

Cada sujeto que acude al experimento realiza una serie de ensayos. Cada segmento de datos electroencefalográficos contendrá la información recolectada de un ensayo.

## 2 Pre-procesado

Para preprocesar los datos nos encargamos de cargar los ficheros y segmentarlos de acuerdo a los marcadores puestos en el experimento. En este experimento, se reproduce un sonido (unas frases), con lo cual contamos con unos marcadores extra que nos indican cuándo se ha empezado a reproducir el sonido y cuándo ha terminado de reproducirse. Teniendo en cuenta esto, se podrá recortar la parte del ensayo de interés para el análisis (actualmente la parte en la que se reproduce la pista de audio).

La manera en la que segmento yo los datos es mediante el uso de una primera línea conteniendo toda la información posible del ensayo (punto en el que empieza el audio, en el que acaba, fichero de audio, imagen mostrada durante el ensayo...) y posteriormente se encuentran en líneas sucesivas los 64 canales. No se tiran los canales oculares todavía por si fuera necesario manejarlos.

Estos segmentos se guardan en archivos .dat cuyo nombre de archivo hace referencia al número de sujeto y al número de segmento en orden cronológico. Si en cualquier momento debido a la consistencia interna de los ficheros de datos algún ensayo tuviera información incongruente entre dos ficheros, ese ensayo se deshecha.

Los canales oculares se deshechan antes de pasar al siguiente paso.

## 3 Procesado de la señal

El procesado de la señal comienza filtrando por bandas la señal. Utilizamos un filtro de Butterworth the tercer orden (ya que el de cuarto orden para frecuencias bajas se comporta de manera extraña), filtrando la señal *forward and reverse*. Tras esotransformada de Hilbert obteniendo los parámetros de las señales analíticas y podemos hacer un estudio pormenorizado de cada parte de la señal.

En concreto por ahora estamos realizando el estudio de la fase (el ángulo de la señal analítica), pero posteriormente podríamos estudiar la envolvente (el valor absoluto), o la frecuencia inmediata (derivada respecto al tiempo de la fase).

## 4 Cálculo de la información mutua

Dado que la teoría de la información estudia sistemas discretos, lo primero que vamos a tener que hacer será discretizar más la señal para que tenga sentido el uso de esta perspectiva. Si la discretización es muy fina, no podremos saber cómo de «ordenada» está nuestra señal o cómo de coherente es. Si es muy gruesa, todo será igual y tampoco podremos ver nada. La discretización que usamos es la que han realizado otros grupos en el pasado, como punto de referencia: dividimos la fase en los cuatro cuadrantes:  $[-\pi, -\frac{\pi}{2})$ ,  $[-\frac{\pi}{2}, 0)$ ,  $[0, \frac{\pi}{2})$ ,  $[\frac{\pi}{2}, \pi)$ .

Gracias a esta discretización, podemos calcular para cada señal las probabilidades de cada fase de la manera expuesta en la ecuación 1, siendo  $p_i$  la probabilidad de que la fase esté en el espacio  $i$ ,  $n_i$  el número de veces que la fase se encuentra en ese espacio y  $N$  el número total de datos.

$$p_i = \frac{n_i}{N} \quad (1)$$

A partir de esto, podemos calcular la entropía de todas las señales que tenemos. Esta entropía nos será útil para calcular la información mutua. La entropía se define como descrito en la ecuación 2 (en bits). La información mutua se calcula según la ecuación 3, siendo  $H_{i|j}$  la entropía condicional de dos señales (en nuestro caso un canal de EEG y la señal de audio),  $p_{ij}$  la probabilidad conjunta de las señales.

$$H = - \sum_i p_i \log_2 p_i \quad (2)$$

$$I_{i;j} = H_i - H_{i|j} = H_i + \sum_i \sum_j p_{ij} \log_2 \frac{p_{ij}}{p_i p_j} \quad (3)$$

Realizamos una comparación relativa de la información mutua en los ensayos en los que se muestra cara y en los que no, como se muestra en la ecuación 4.

$$\eta = \frac{I_{cara} - I_{scrambled}}{I_{cara}} \quad (4)$$

Con esto, y dado que es una comparación valor a valor, podemos realizar una  $t$  de student para obtener la significancia.

## 5 Desarrollo futuro

- Realizar bandas de frecuencias combinadas
- Realizar el análisis de límites que se realiza en Cogan y Poeppel 2011
- Realizar análisis con ensayos incorrectos (vs correctos)
- Análisis de distintas partes de la señal y combinaciones
- Variar la discretización de las señales