University of Tartu
Faculty of Mathematics and Computer Science
Institute of Computer Science

Kristjan Korjus, Ilja Kuzovkin

Visualization of intracranial EEG data

Project in Data Mining course

Tartu 2011

Abstract

Techniques of getting data from brain are improving constantly. We got a set of intracranial EEG data which was recorded with a novel method by the Department of Epileptology, University of Bonn. All the electrodes are located in the temporal lobe. The data has a very good spatial and temporal resolution.

We are analyzing the data and come up with a novel method for visualizing the EEG data.

Contents

1 Introduction

> *"This is the way science works: Begin with simple, clearly formulated, tractable questions that can pave the way for eventually answering the Big Questions, such as "What are qualia," "What is the self," and even "What is consciousness?""*
> - V.S. Ramachandran

The brain is one of the weirdest objects in the universe. The brain is like a Turing machine which asks about its own capacities or like a Gödel's function which wants to know its own properties.

The main aim for neuroscience is to understand the phenomena of intelligence and subjective consciousness. Although the last decade of the 20th century has been labeled as the Decade of the Brain, not much is known about the brain and the main goals are far from being met. It is not known how we record memories, how our extremely slow hardware[1] can work so efficiently with the incomplete information and how can a subjective consciousness arise at all from a physical processes.

As in every other field of science, one can approach the understanding of the brain in small steps. Neuroscience is getting gradually more complex and theories from various fields, such as mathematics, information technology and physics, are converging to tackle the complicated problems. Although there are many interesting sub-problems in the field of brain studies, this project concentrates on an visualization of intra cranial EEG data.

Chapter 2 describes some properties of the EEG data and describes most of the known methods for visualizing the EEG data. Chapter 3 explains the experimental set up. Chapter 4 describes the data of an epilepsy patient. The data is analyzed with basic methods. Using the understanding of the data and visualization methods used in neuroscience and elsewhere, we develop a novel combination of methods in chapter 5. Some conclusions and futher research questions are analyzed in the chapter 6.

2 Literature Review

Neuroscience and many other fields in different sciences are producing huge quantities of data. Naturally, many methods have been developed for visualizing the data.

[1] Chemical transitions of signal between synapses are hundreds of times slower than our silicon hardware

First, we give a short overview of the EEG data and some basic techniques used in the analysis of the data. Mainly, the paper by M. ten Caat and others is used for examples [1].

Secondly, we describe some novel methods which have been invented recently to improve the techniques.

Thirdly, we describe two methods which are used to analyze other sources of data. First method helps to visualize spike data of multichannel single neuron recordings [2]. Other method, Bubble charts, is used to animate socioeconomic data of the world [3].

Fourthly, we give a short overview of good practices in visualization of biological data by VIZBI - an international conference of biological data visualization [4].

Before the description of visualization methods, some ideas about the signal are given.

2.1 Characteristics of EEG Data

In normal circumstances, EEG measures electrical activity of the brain on the scalp at different locations. In this report, intracranial electrodes are used which give a lot higher quality data. The data is described in the chapter 4, page 24.

Number of electrodes varies greatly but about 64 electrodes is the most common set up. From all the electrodes, electrical signal is recorded at the sampling rates up to 2000 Hz but normally at 1000 Hz.

Normally, an experiment consists of many trials each trial for couple of signals. The data which need to be analyzed, is segmented into trials such that there is up to a second of data before the stimulus and up to a second after the stimulus. Naturally, some studies are different, for example in sleep experiments, EEG signal is measured for hours.

Very good overview on EEG is given in the book "Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields" [5].

2.2 Basic Visualization methods

The most basic method for the visualization of EEG data is the simple time series plot. The time is set out horizontally and voltage vertically. Normally many different electrodes are shown on top of each other.

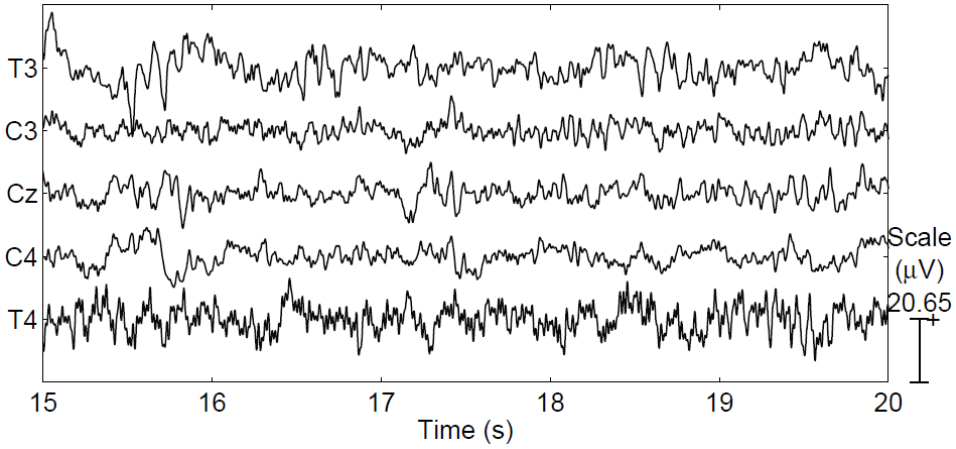


Figure 1: Conventional EEG representation. Measured voltages (in μV) are shown against time (in seconds), for five electrodes (indicated by the labels T3, C3, Cz, C4, and T4). Source: [2]

The plot makes it possible to differentiate between simple phenomena such as sleep vs awake states. Only limited number of electrodes and time intervals can be shown and more complex interactions are normally visible.

In the so called Butterfly plot, all the signals are superimposed on top of each other. Some important features can become visible but single channels cannot be identified any longer.

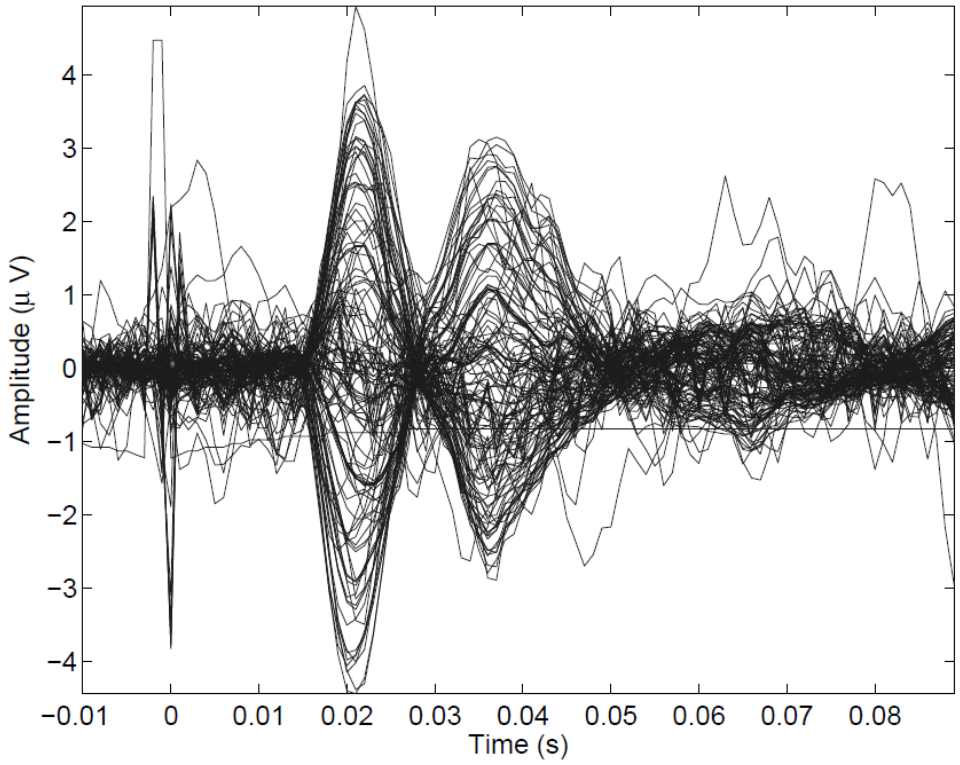


Figure 2: Butterfly plot, with superimposed SEPs for 128 electrodes. The EEG shows local maxima in the variation about 0.021s and 0.035s after the stimulus. (The pattern between approximately 0.015 and 0.060s resembles a butterfly shape.). Source: [2]

Topographic layouts are used to show actual electrode locations on a head.

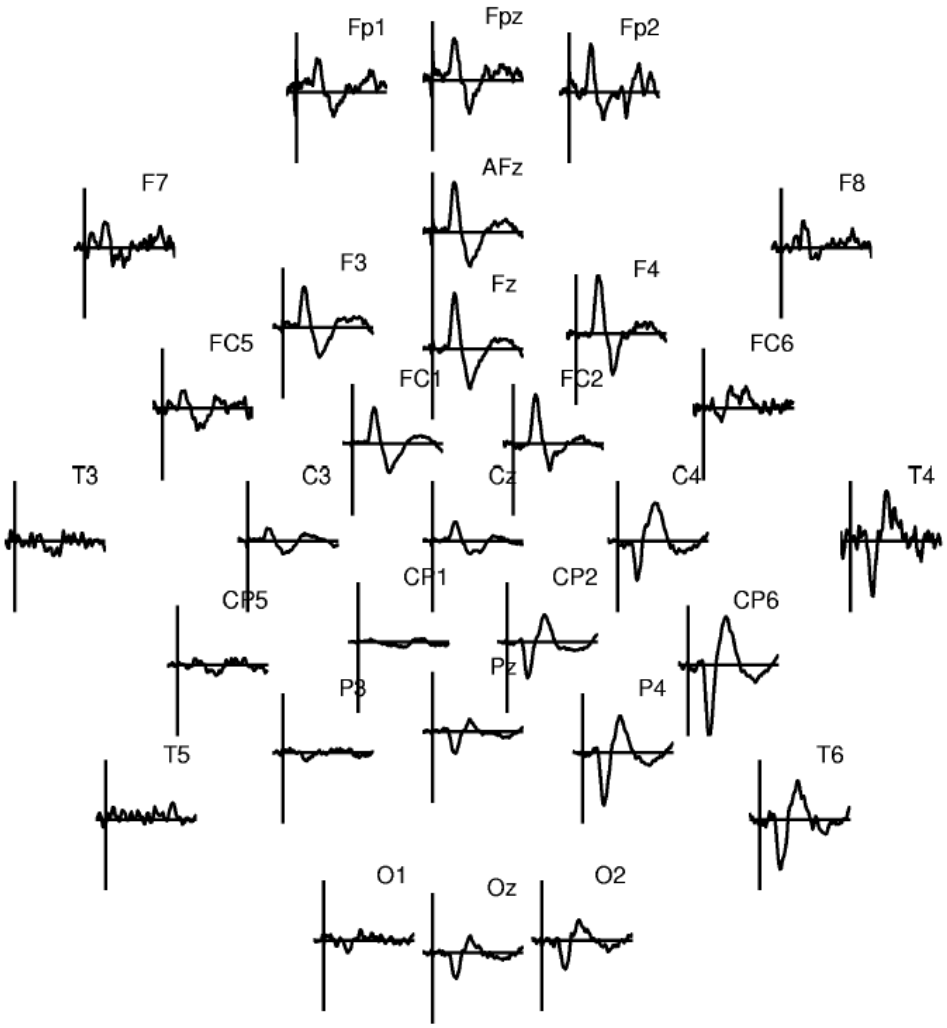


Figure 3: Topographic array, for left median nerve stimulation and thirty electrodes. Top view, nose on top. Source: [2]

Also, different time points can be visualized with topographic maps.

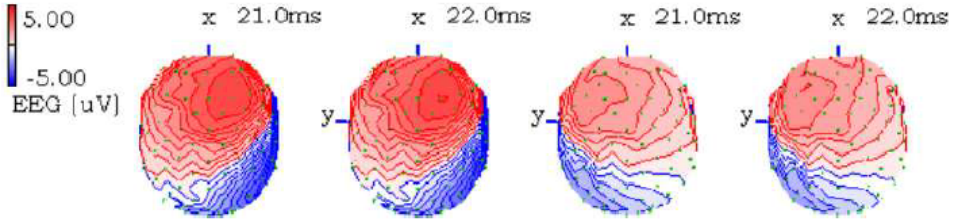


Figure 4: Four topographic maps, including isolines (top view, nose on top). Notice the mirror-symmetry between the two on the left (21 and 22ms after left median nerve stimulation) and the two on the right (21 and 22ms after right median nerve stimulation). Source: [2]

2.2.1 Event Related Potential (ERP)

To get a better signal to noise ration, many trials are quite often averaged up to see the event related potential. ERP analysis can be combined with all the previous plotting methods. For example, figure 5 shows an example of an ERP plot. The activity of all the trials is presented such that the voltage is shown in color code. In the bottom, average ERP over all the trials is shown.

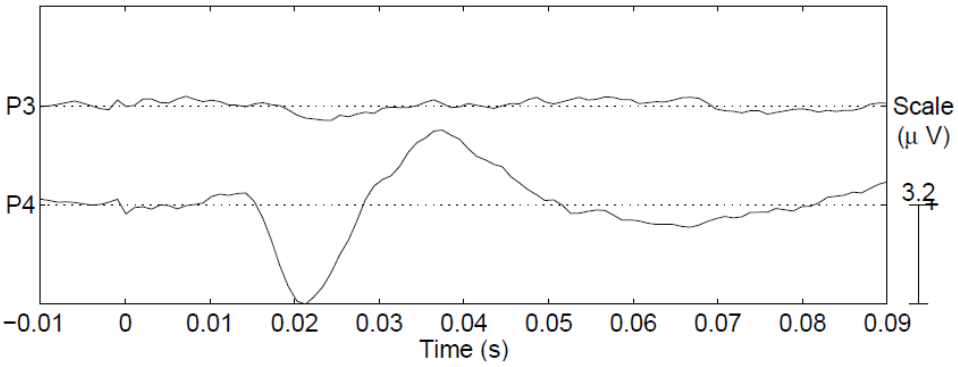


Figure 5: Left-hand stimulation of the median nerve (near the wrist). The average EEG over approximately 500 electrical stimuli is displayed for the electrodes labeled P3 and P4 (attached near the left and the right parietal cortex, respectively). The dotted line indicates the zero-level. The first negative peak for electrode P4, just after 20ms, is referred to as the N20. Source: [2]

2.2.2 Time-Frequency Transformation

The Fourier transform is a mathematical operation that decomposes a function into its constituent frequencies. Each value of the function is a complex number (called complex amplitude) that encodes both a magnitude and phase component[6]. In the current paper, only the amplitude of the Fourier transform is used. The EEG voltage $V_m(t)$ recorded from any electrode pair m is then expressed generally (see figure 6) as a sum over frequency components

$$V_m(t) = \sum_{n=1}^{N} A_{nm} \sin(2\pi f_{nm} t - \phi_{nm}). \tag{1}$$

In neuroscience, it is common to categorize frequency ranges in the following way: delta (1 to 4 Hz), theta (4 to 8 Hz), alpha (8 to 13 Hz) and beta (13 to 30Hz) and very high frequencies (typically 30 to 80 Hz) are referred to as gamma activity. These distinctive labels correspond roughly to frequency bands that often dominate particular human brain states. For example, delta activity with frequencies lower than about 1 or 2 Hz is dominant during deep sleep and in many coma and anesthesia states.

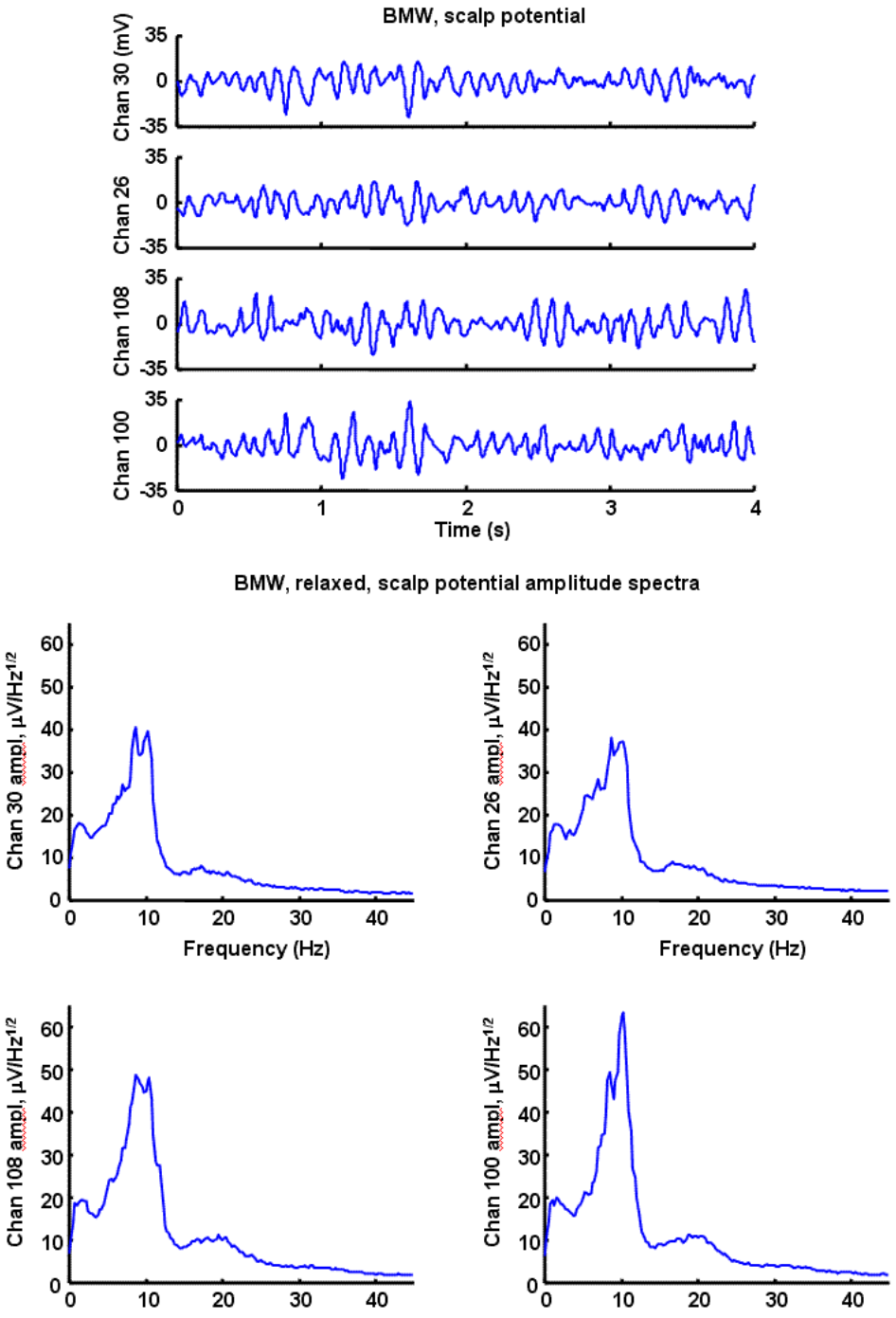


Figure 6: (Upper) Alpha rhythm recorded from a healthy relaxed subject (age 25) with closed eyes using an electrode on the neck as reference. Four seconds of data are shown from four scalp locations. The amplitude is given in microvolts. (Lower) The corresponding amplitude spectra based on the full five minute record reveals dominant activity in the alpha (8-13 Hz) band. Reproduced from Nunez and Srinivasan (2006). Source: [7]

Quite often, all the different frequencies are shown together such that the power band is encoded with the color.

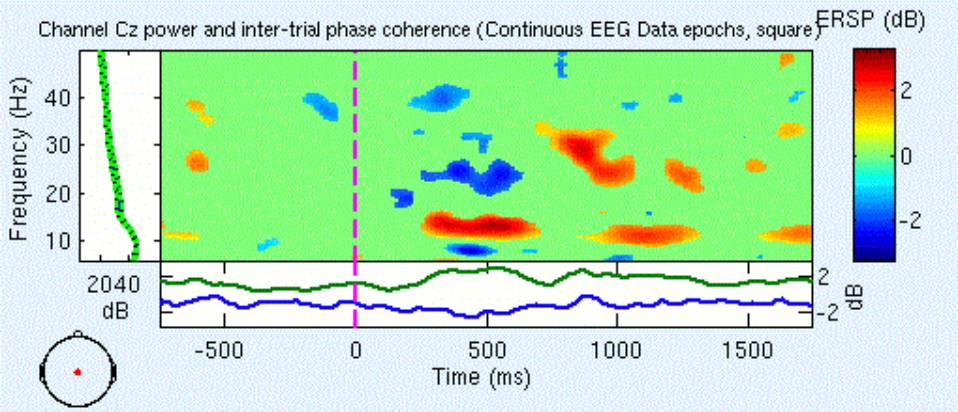


Figure 7: Time-Frequency Transformation. Source: `http://sccn.ucsd.edu/wiki/Chapter_11:_Time/Frequency_decomposition`

2.3 Tiled Parallel Coordinates

Tiled Parallel Coordinates is a new method for visualization of Time-Varying Multichannel EEG Data. The methods is shown in series of figures.

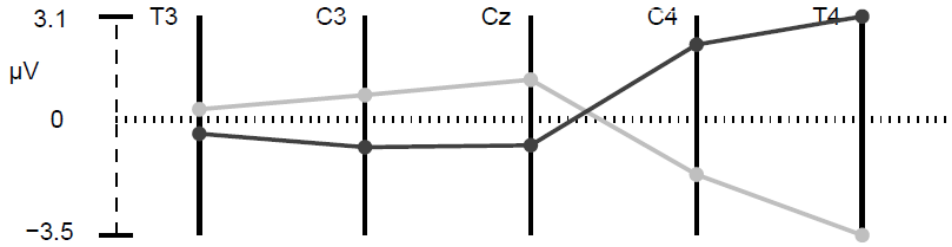


Figure 8: Parallel coordinate representation for two five-dimensional vectors, each of which represents one time step. For each vector, one polyline is drawn. The data have been recorded from five EEG electrodes simultaneously (labeled T3, C3, Cz, C4, and T4). The voltage (μV) is set out vertically. Source:[8]

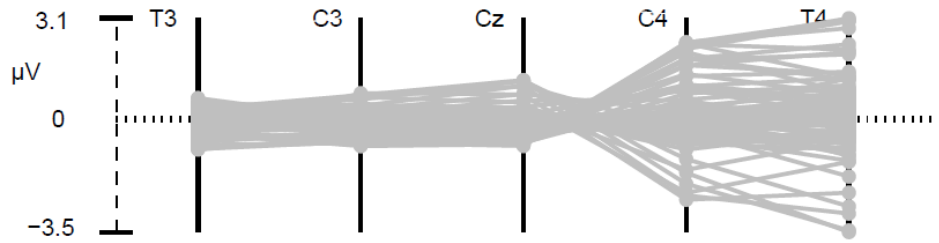


Figure 9: Parallel coordinate representation for 100 time steps and the five electrodes, Source:[8]

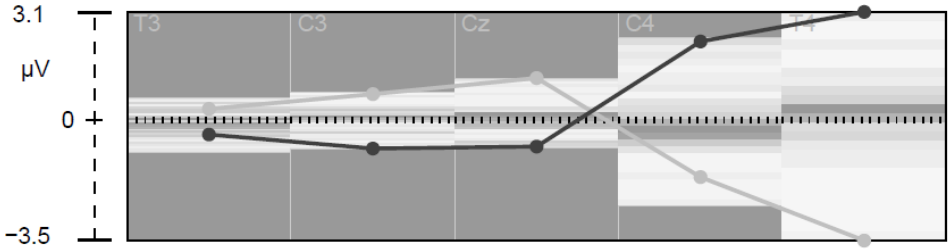


Figure 10: Minmax plot containing five tiles, showing the extreme values for five electrodes. Density map, combined with minmax plot, is reflecting the distribution of the polylines along the vertical axes (dark gray for high, light gray for low densities). Combination of parallel coordinates, the minmax plot, and the density map. Source:[8]

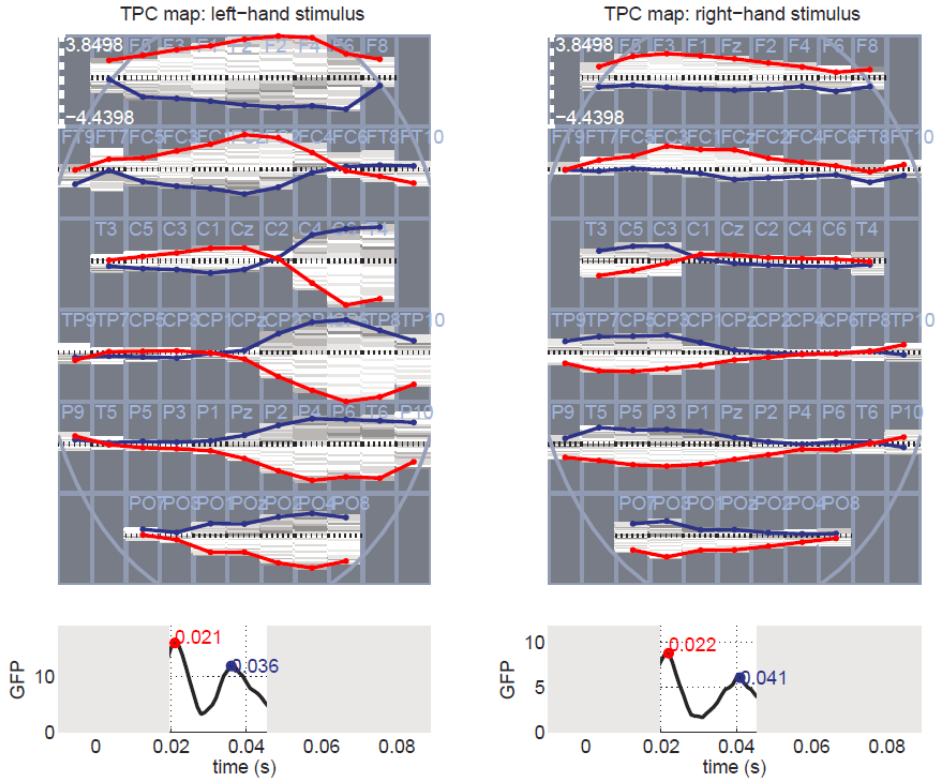


Figure 11: Two TPC maps, both offering a top view of 58 electrodes (nose on top) and showing EEG data for left and right median nerve stimulation, respectively. The red and blue polylines correspond to two time steps. Source:[8]

2.4 Methods from other fields

2.4.1 Spike data

Instead of a continuous EEG signal, multielectrode spike trains give out a discrete signal. This data is difficult to visualize or interpret. Although various methods exist for treating multielectrode datasets quantitatively, there is a lack of techniques that enable a quick visual exploration of the data. A novel method was developed by O. F. Jurjut and others that is using Kohonen self-organizing maps which allows the representation of multiple spike trains through a sequence of color-coded population activity vector. If multiple color sequences (trials) are grouped together one can see an entire dataset visually and extract information about the identity, stimulus-locking and temporal distribution of multi-neuron activity patterns.

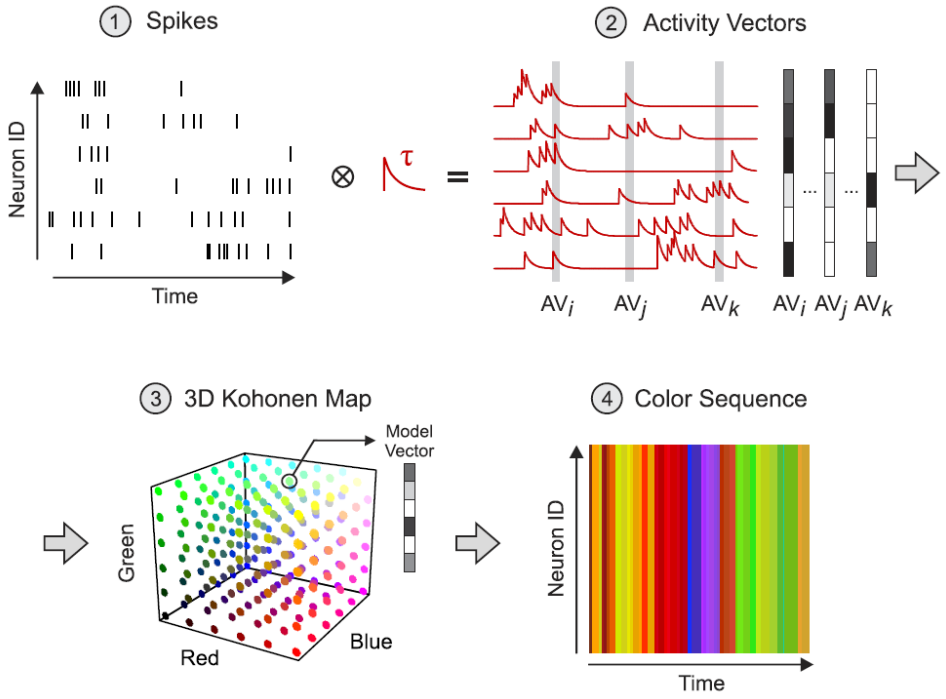


Figure 12: From spikes to color sequences. Steps required to construct a color sequence from spike trains: simultaneously recorded spike trains (1), low-pass filtering by convolution with a decaying exponential function (2), clustering activity vectors with a three-dimensional (3D) Kohonen map (3), assignment of colors to activity vectors based on the position of their corresponding models in the Kohonen map (4). Source: [2]

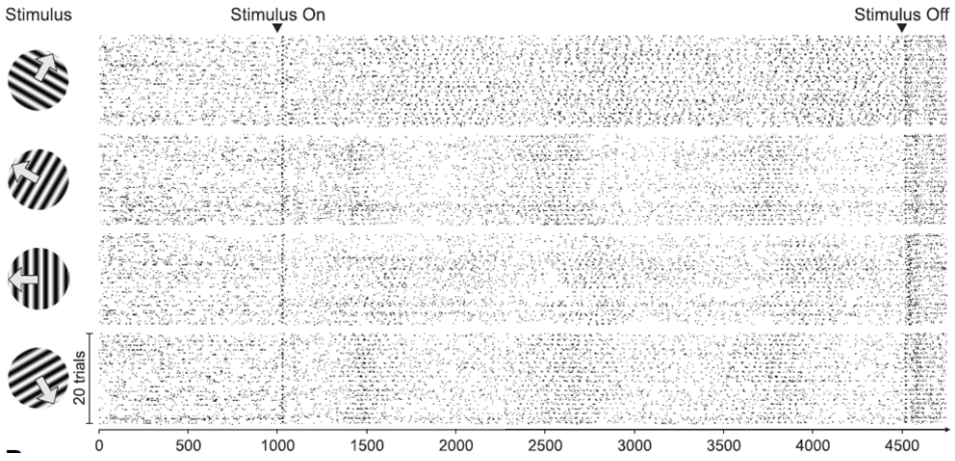


Figure 13: Raw spike data. Four different stimulus, 20 trials with each stimulus and 17 electrodes in each trial. Source: [2]

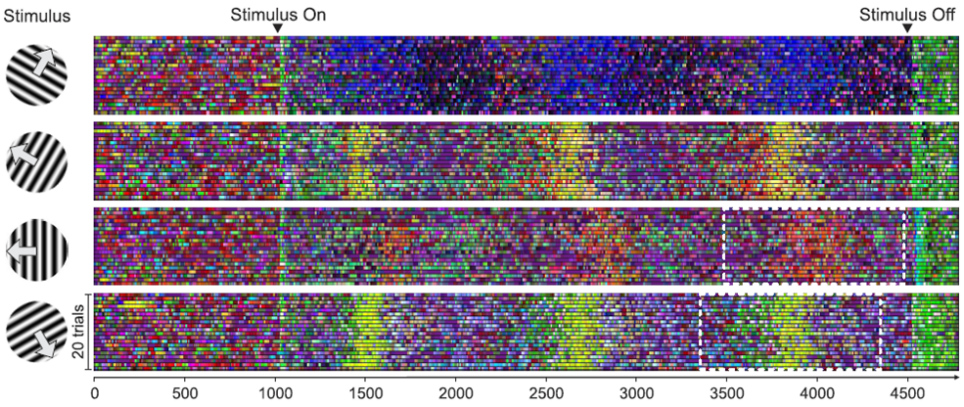


Figure 14: Color code added to each stimulus. Four different stimulus, 20 trials with each stimulus and 17 electrodes in each trial. Source: [2]

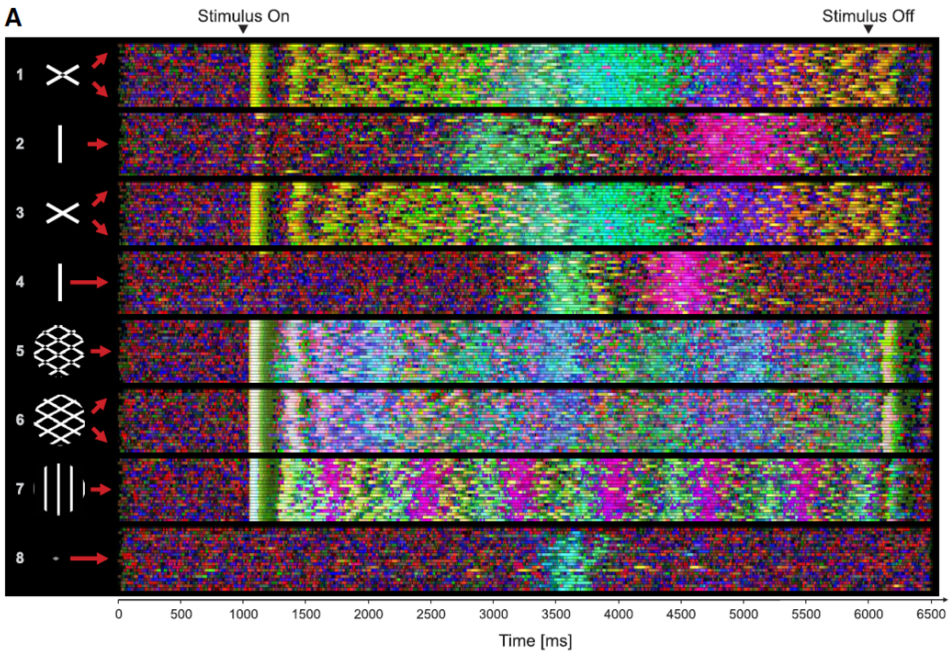


Figure 15: Another set of stimulus. Source: [2]

2.4.2 Self-organizing Maps

As the previous visualization method uses Kohonen's Self organizing maps (SOM), it makes sense to introduce the algorithm. SOM consists of components called nodes which are located on the map i.e. there is a distance

defined between nodes. Each node is a weight vector of the same dimension as the input data vectors i.e. activity vectors in the context of the current work. The usual arrangement of nodes is a regular spacing in a rectangular grid but in the current work, a 3D cube was used. The self-organizing map describes a mapping from a higher dimensional input space to a lower dimensional map space. [9]

To get the map, test data is used with the following algorithm:

1. Randomize the map's nodes' weight vectors
2. Grab an input vector
3. Find the node which has the smallest distance to the input vector
4. Update the nodes in the neighborhood of the node by making them more similar to the input vector.
5. Repeat from 2 while some criteria is met

This produces a map which can be used to map higher dimensional vectors to lower dimensions.

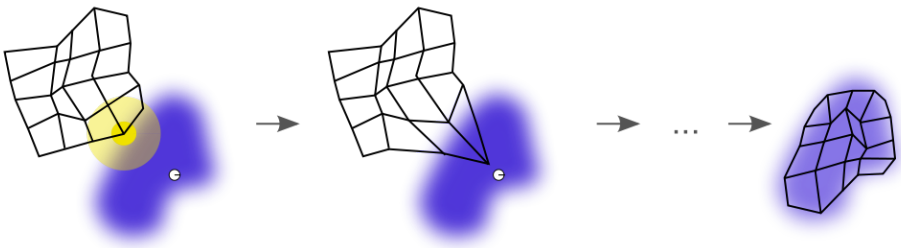


Figure 16: An illustration of the training of a self-organizing map. The blue blob is the distribution of the training data, and the small white disc is the current training sample drawn from that distribution. At first (left) the SOM nodes are arbitrarily positioned in the data space. The node nearest to the training node (highlighted in yellow) is selected, and is moved towards the training datum, as (to a lesser extent) are its neighbours on the grid. After many iterations the grid tends to approximate the data distribution (right). Source: [9]

2.4.3 Bubble charts

The method of Bubble charts was introduced by Swedish non-profit organization Trendalyzer. Their mission is to unveil the beauty of statistical

time series by converting boring numbers into enjoyable, animated and interactive graphics. The web site is called GapMinder. [3]

The basic idea of GapMinder is to convert time series into scatter plot which can display multiple dimensions at once. There are up to 5 dimensions which can be used:

1. x-axis
2. y-axis
3. Color of the dots
4. Size of the dots
5. Time (which could be implemented in terms of a moving image, separate still image frames or with a trail.)

An example of this kind of map can be seen on the figure 17.

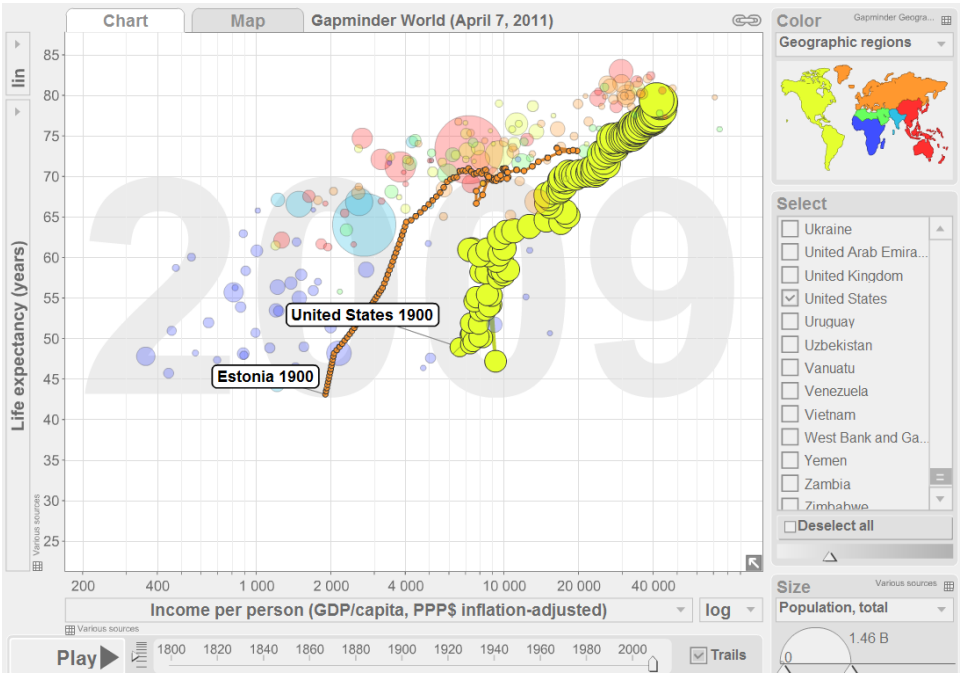


Figure 17: Illustration of the Bubble chart where x-axis is the log scale of income per person, y-axis is the life expectancy, size of the bubble is the size of the country, color shows the continent and the time effect is realizes with trail of two countries: Estonia and United States. Source: [3]

In March 2007, Google acquired Trendalyzer from the Gapminder Foundation and they have made some APIs such that everyone could use the Bubble Charts for free.

2.5 Visualization principles

In the current report we use practical guidelines for visualization from The VIZBI. They organize an international conference series bringing together researchers developing and using computational visualization to address a broad range of biological research areas; the conference also attracts participation from medical illustrators, graphic designers, and graphic artists.

Main source for guidelines is the Tamara Munzner's presentation: Keynote on Visualization Principles - `http://vizbi.org/Videos/26205288`. Figure 18 shows the main summary of the presentation. They give their opinion about different ways of visualizing different types of data.

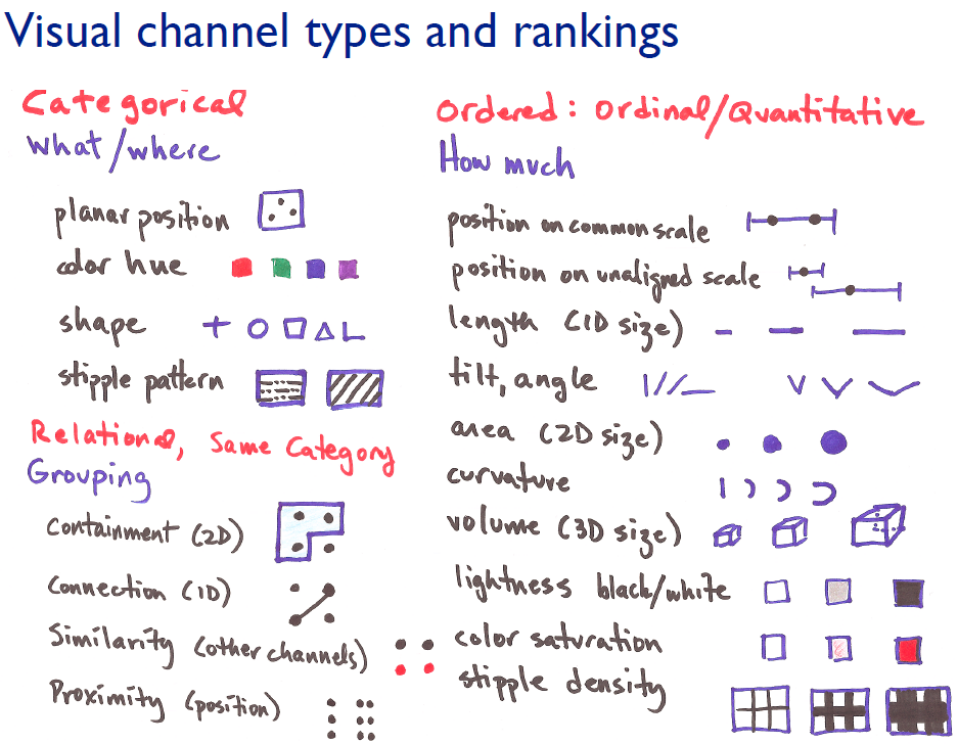


Figure 18: Visual channel types and rankings

3 Methods

3.1 Subjects

Intracranial EEG from a single patient with pharmacoresistant epilepsy is analyzed. Recordings were performed at the Department of Epileptology, University of Bonn, Germany.

3.2 Experimental set up

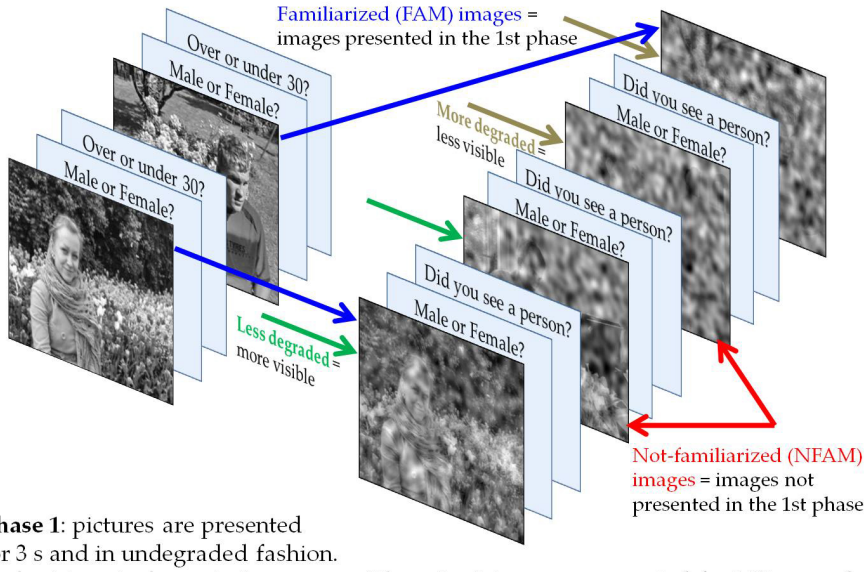
For having stimuli where prior experience could have an impact on sensory processing, relatively complex natural images were used. A set of grey-scale photographs with a variety of backgrounds that contained a single person in the foreground were used.

After each picture in the test phase, the subject had to indicate whether the person on the degraded picture was male or female (objective performance) and whether they had indeed seen a person on the picture or not (subjective performance). Occasionally, subject were also asked whether the picture, now shown in the degraded fashion, had been presented in the familiarization phase of that block. After the questions and before the next picture, the fixation cross was presented randomly for 1200-1400 ms. The response screens appeared 1 second after the onset of the picture. Each block lasted for 3-4 minutes and subjects could take breaks between the blocks. The duration of the whole experiment was 35-45 minutes.

To study conscious recognition, one needs to make sure that the targets are not consciously recognized in every trial. To achieve that in the test phase, the approach of having brief durations (150 ms) and degrading of the target images was used. For that, each target image was edited with random noise, where the amount of noise could be controlled parametrically (see figure 19). For this experiment stimuli were prepared with noise levels from 60% to 90% noise in 5% steps. Before the main experiment started, a threshold experiment was conducted to select for each subject two neighboring degradation levels (e.g. 75% noise and 70% noise) that would provide optimal recognition performance (e.g. not random guessing and not ceiling effects).

The main experiment was divided into 11 experimental blocks. Each experimental block consisted of two phases - the familiarization phase and the test phase (illustrated on figure 19). The goal of the familiarization phase was to create prior knowledge about some of the pictures. In the

first phase 4 pictures were shown clearly for 2 times for 3 seconds each time. The subjects were told to remember these pictures. By the first clear presentation of each picture the subjects were asked whether the person on the picture was male or female and whether they thought he or she was under or over 30 years old. By the second clear presentation of each picture the subjects simply had a task of memorizing the picture.



Phase 1: pictures are presented for 3 s and in undegraded fashion. Each picture is shown twice: once with the task presented above and once in a free viewing condition.

Phase 2: pictures are presented for 150 ms and in degraded fashion. There are also catch trials where no person is present.

Figure 19: Experimental paradigm. Each block is divided into 2 phases. In the first phase prior knowledge is created about some of the images by showing them clearly. In the second phase the effects of prior knowledge and sensory information are tested by showing the pictures briefly and in degraded fashion. Degraded versions of the pictures from phase 1 are presented together with new pictures (manipulation of prior knowledge). In addition, both types of pictures can be shown in two different degradation levels (manipulation of sensory evidence). Source: author

In the test phase of each block these 4 familiarized pictures and 4 new pictures with people on them were presented, belonging to conditions with and without prior knowledge, respectively. Each picture was presented twice during the test phase. Importantly, in the test phase the pictures were presented for 150 ms and in either the lower or the higher degradation

level, as explained above, which constituted the bottom-up factor of sensory evidence. Therefore, each picture with a person on it could meet one of the 4 conditions:

- With prior knowledge and high bottom-up (familiarized and lower degradation)
- With prior knowledge and low bottom-up (familiarized and higher degradation)
- without prior knowledge and low bottom-up (not familiarized and higher degradation)
- high bottom-up (not familiarized and lower degradation).

Additionally, in each test phase 4 catch trials with no persons on them were presented. They served to control for the subjective perception, as in these pictures subjects should not have seen any persons. Indeed, for the subject included in this study, the amount of false perceptions in the catch trials was very low.

3.3 Intracranial recordings

As all the electrodes were implanted for diagnostic purposes, there was no influence on their positioning for this study. The location of electrode contacts was ascertained by MRI in each patient. EEG was referenced using bipolar method such that electrode 1 was referenced to electrode 2 and so one. So that from each of the rod we get 6 signals. Signal is recorded at a sampling rate of 1000 Hz.

Figure 20 shows the electrode rod of the patient whose data is analyzed in the current study.

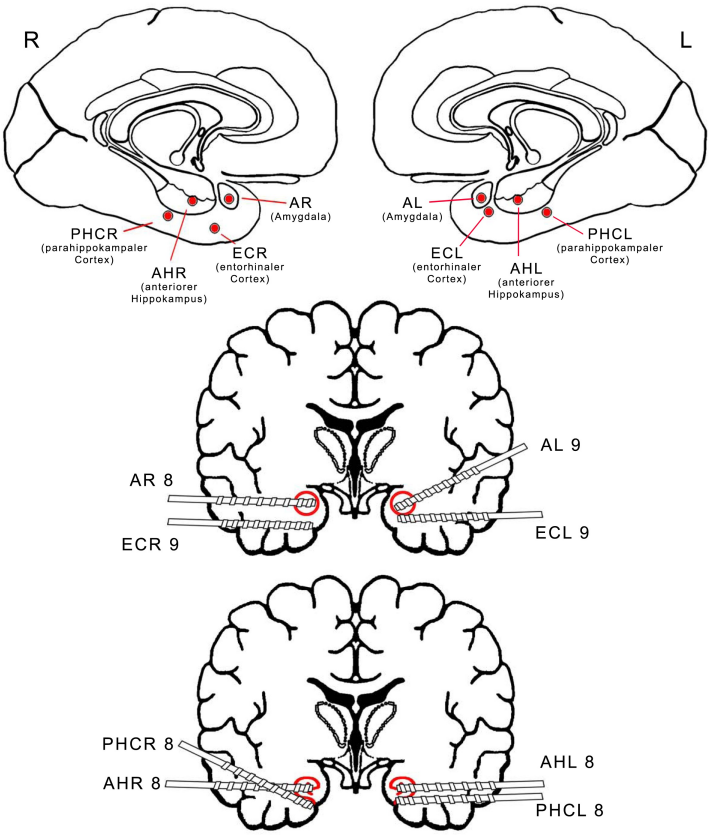


Figure 20: The positions of multi-contact electrode rods and the electrodes used in the current study. Source: Original from the Department of Epileptology, University of Bonn, Germany. Modifications made by us

Intracranial EEG data were filtered (0.5-300 Hz) and segmented into epochs of -1000 ms to 1000 ms around stimulus presentation. Subsequently, trials were visually inspected for artefacts (e.g. epileptiform spikes), and trials containing artefacts were excluded from analysis.

Figure 21 shows a sample data (-500 ms to 500 ms) of a random trial from both channels.

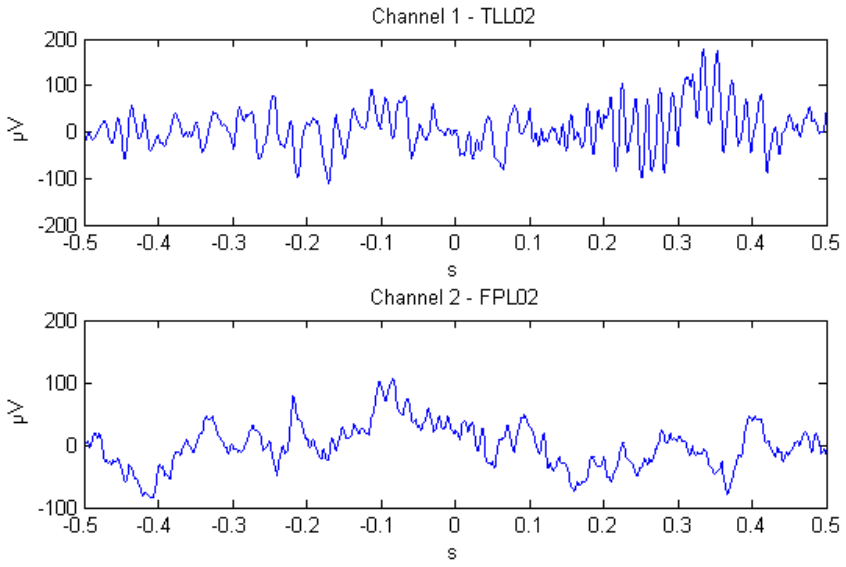


Figure 21: Sample intracranial data from electrodes TLL02 and FPL02. Source: author

3.4 Hypothesis

The hypothesis of the study: according to the theory, post-stimulus high frequency activity should be higher than pre-stimulus activity and there should be some visible effects in the ERP figures.

4 Descriptive Analysis

4.1 3D position of the electrodes

Approximate positions of the electrodes were shown on the figure 20. We got access to the MRI scan of the patient and analyzed the data with the software MRIcron. We were able to locate all the electrodes and extracted the Cartesian coordinates for them. Figure 22 illustrates the data, software and the process of locating the electrodes.

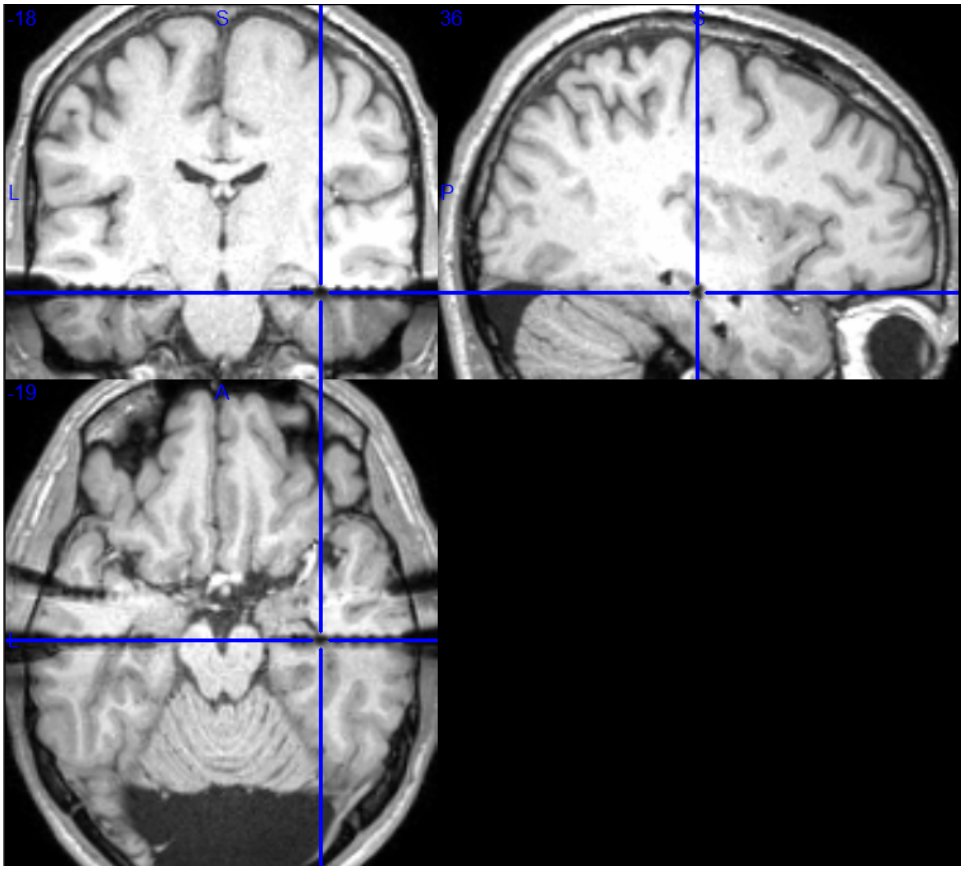


Figure 22: MRI visualization in MRIcron. Third electrode on the rod AHR is in the focus.

Figure 23 shows the coordinates on a 3D plot.

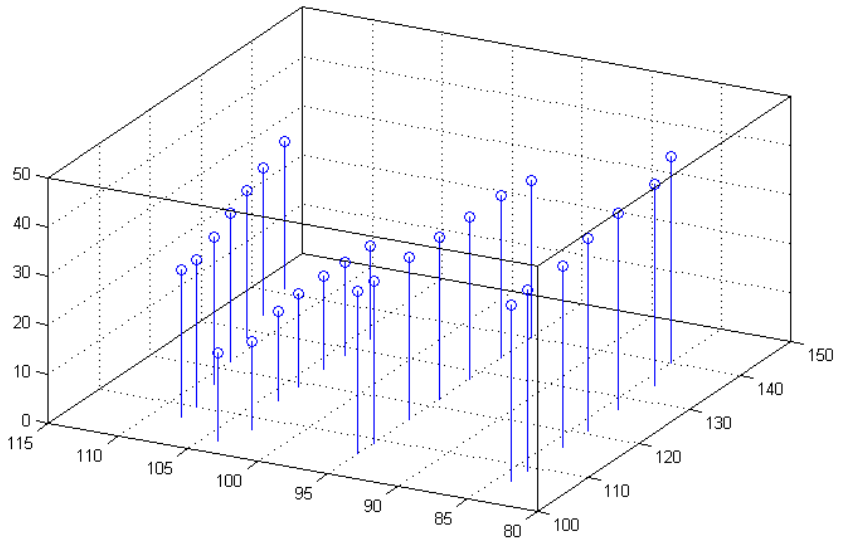


Figure 23: Locations of the electrodes on a 3D plot

4.2 ERP signals

At first we would like to visualize averaged signal strength averaged over all the trials. One of the largest conditions of the experimental set up is the familiarity of the photo - so we keep these to signals separately because there is a change to find some interesting differences in the conditions.

From the figure 24, we can clearly see that there is some event related activity going on.

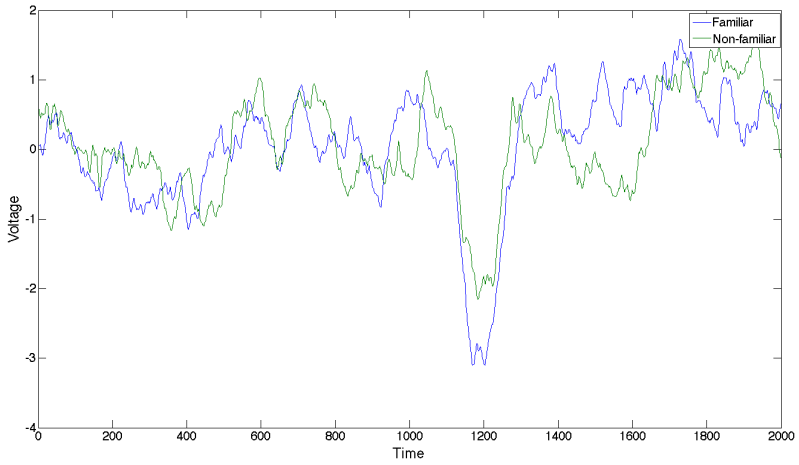


Figure 24: Brain activity after seeing familiar picture differs from the non-familiar picture case.

In the figure 25, all the signals are averaged together. We can see some minor condition specific effects and some major stimulus induced effects.

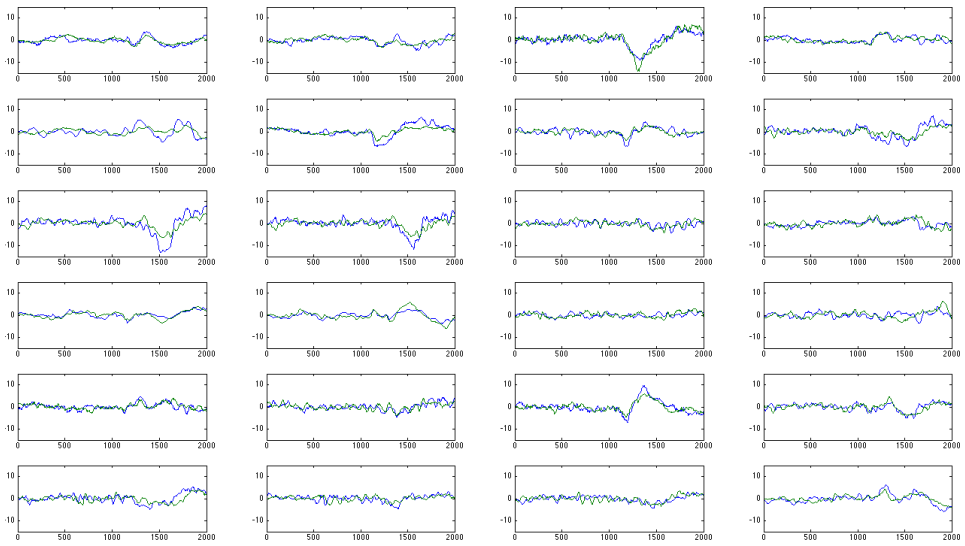


Figure 25: Standardized voltage amplitude over time on every channel.

4.3 Time-Frequency Transformation

On the figure 26 and 27 we can see the power spectrums of the pre-stimulus and post-stimulus signals.

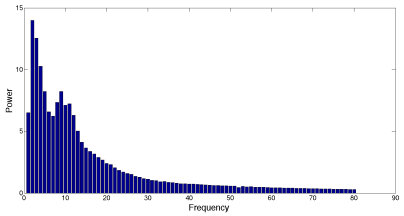


Figure 26: Pre-stimulus power spectrum

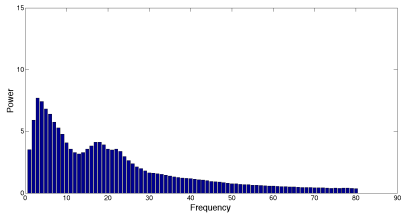


Figure 27: Post-stimulus power spectrum

It is clearly visible that after the the stimulus, power of lower frequencies decrease and higher components increase.

Normally heat maps are used to illustrate and compare the signals. One cell on a heat map represents the power of particular frequency component taken at particular time. Frequency components are obtained with Fast Fourier Transform (FFT). The size of window we perform FFT on depends on the frequency we are analyzing – wider windows for low frequencies and narrower for high.

$$\text{window} = \frac{5}{\text{frequency}}$$

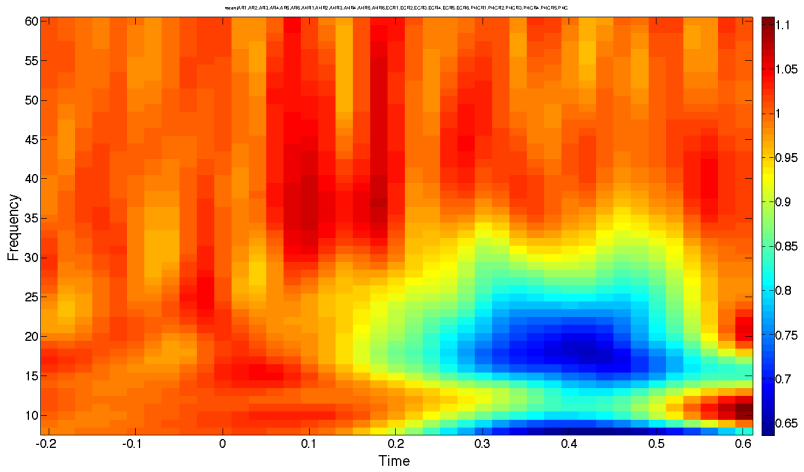


Figure 28: Averaged frequency components over all trials and channels. Spike of activity is clearly visible after the stimulus is introduced.

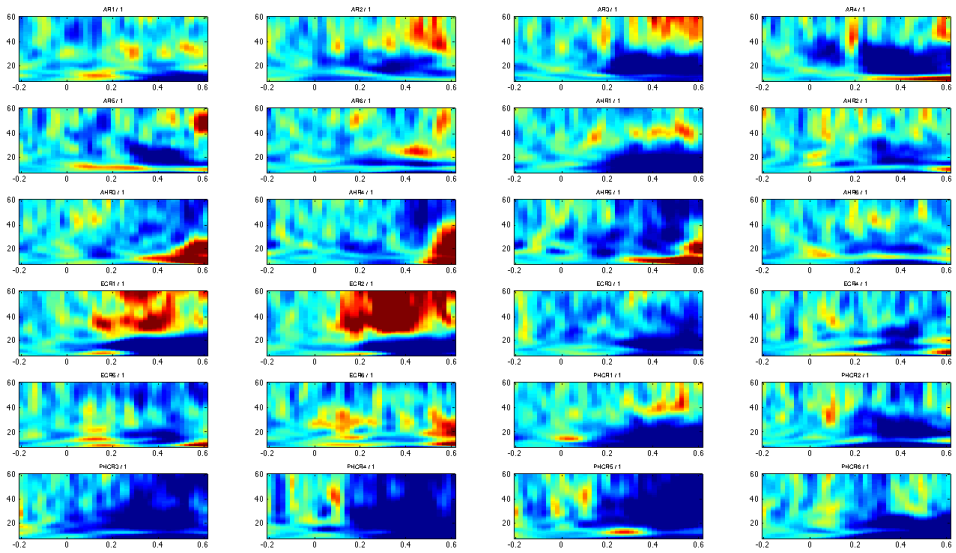


Figure 29: Averaged frequency components over all trials. Here we can see which channels contribute most into the formation of the signal.

4.3.1 Familiar vs. non-familiar

Comparison of post-stimulus signals in familiar and non-familiar cases is shown on the figures 30 and 31.

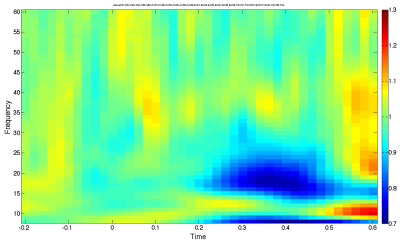


Figure 30: Time-Frequency transformation of the familiar cases

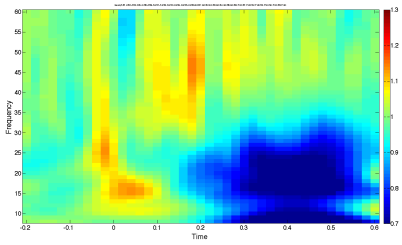


Figure 31: Time-Frequency transformation of the non-familiar cases

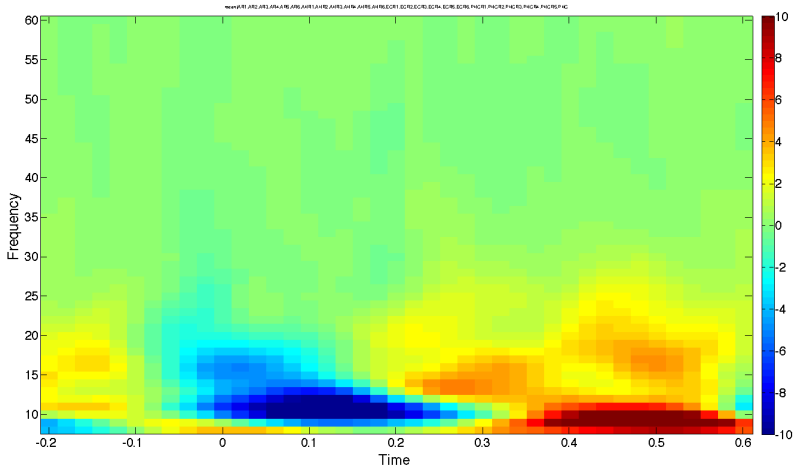


Figure 32: Difference between familiar and non-familiar cases.

Some distinct patterns area clearly visible. In order to study it further we analyzed each channel separately on the figure 33.

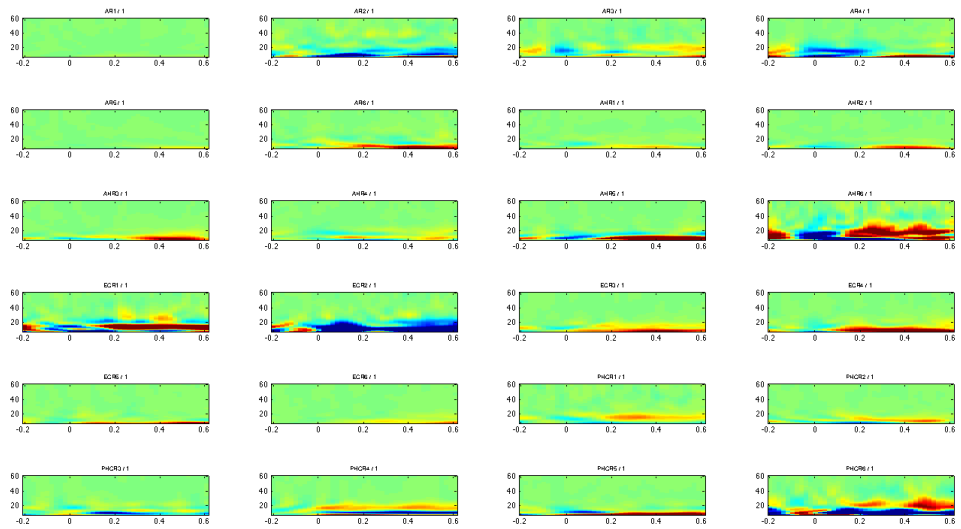


Figure 33: Difference between familiar and non-familiar cases on each channel.

4.4 SOM Implementation

Finally, we used the SOM algorithm which was described on the page 2.4.1. It did not show any significant extra information. Although, from the figure 34, it is clearly visible that there are some effects 300 ms after the stimulus.

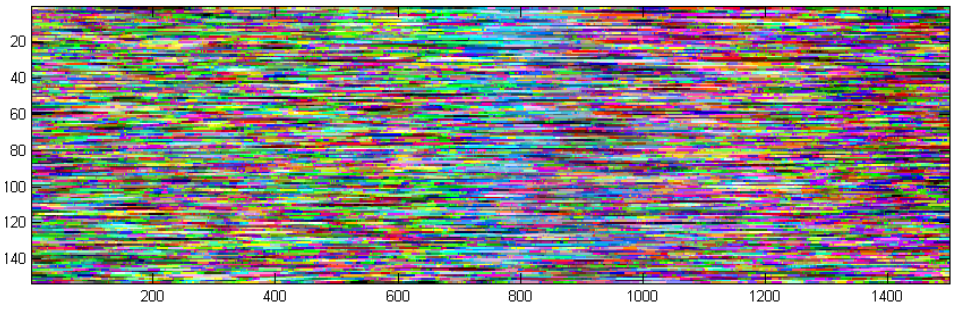


Figure 34: All the 153 trials (77 familiar, 76 non-familiar)

5 Building the Technique for Visualization

From the descriptive analysis we have got an understanding about the information which can be extracted from the data. It also provided us

with some insights of how we should represent the data in our visualization.

Here is the list of measurements we used to characterize the data:

- Voltage – The first, most usual parameter for EEG data, is voltage (see sections 2.2.1 and 4.3).
- Frequency – Another source of valuable information is frequency component analysis (see section 2.2.2). The measurement we use in our set up is the ratio between high frequencies and middle frequencies:

  $$\frac{\text{30 to 70 Hz}}{\text{8 to 20 Hz}}$$

  This ratio grows when brain activity arises and goes down there brain is in more passive.
- Location – Since we have exacted 3D location of each electrode (see section 4.1), we can use this information to observe the behavior of the electrodes which are close to each other. In our set up we use color to denote the deepness of the electrodes.
- Time – Time is addition natural dimension, we observe how our measurement change over the time.
- Correlation – Fifth measurement shows the similar behavior of the electrodes. We calculated the average correlation between the given electrode and all other electrodes.

As described in section 2.4.3 we can plot up to 5 dimensions on a single plot. Here is the mapping of our data to the graphical components of the bubble chart.

- One bubble denotes one electrode
- Y-axis – averaged voltage at given time moment
- X-axis – $\frac{\text{high frequncies}}{\text{low frequencies}}$
- Color of the bubble – how deep the electrode is placed inside the brain
- Time – time
- Size of the bubble – average correlation between current channel and all other channels

The implementation of this setup can be seen on the figures 35, 36, 37 and 38. A moving flash applet can be accessed at `www.ikuz.eu/brain`.

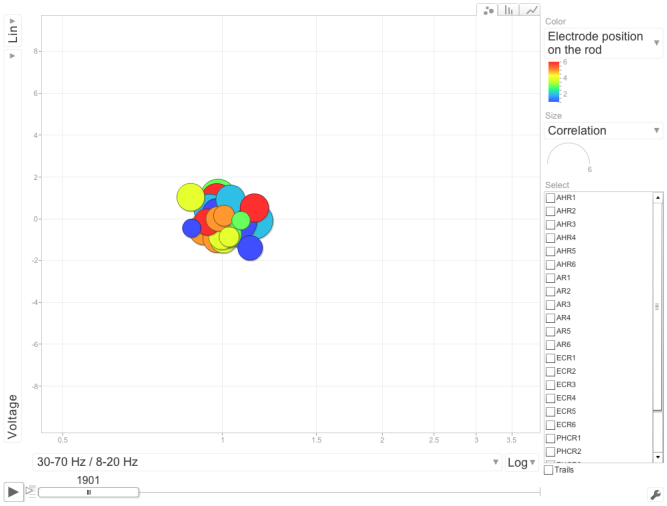


Figure 35: 200 ms before the stimulus: initial position (baseline)

The measurements prior to the stimulus were considered as baseline, thus post-stimulus effects represent changes in the brain activity after the stimulus was introduced.

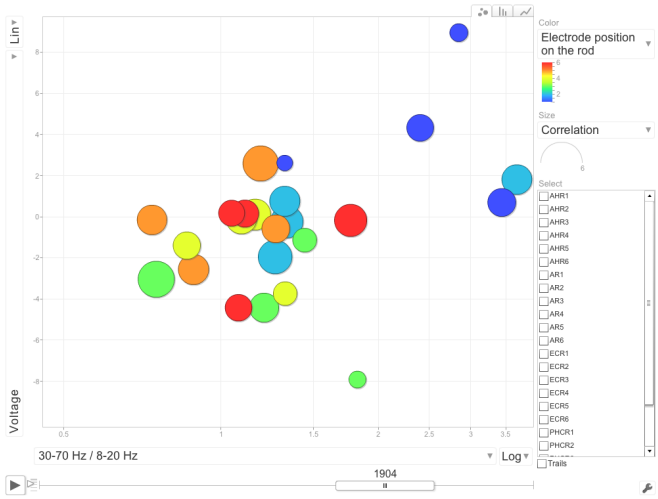


Figure 36: 400 ms after the stimulus: frequency ratio of deepest electrodes grown higher along with voltage

As we can see on Figure 36, post-stimulus powers of high frequency

components grow, which matches with the intuition and Figures 28 and 29. We may also notice that the deepest (blue and light blue) electrodes are most active while superficial (orange and red) stay close the the baseline position, which indicates low activity.

To better observe behavior of the electrodes during the time we add trails to the chart. On the figure 37, we are able to follow the pattern of the electrodes of our interest. Here we can see more clearly that internal electrodes change much more rapidly and seem to have common direction of movement.

Such representation can give quite good intuition for the researcher. He might be able to notice the path which are not worth investigating, find unexpected patterns, observe results of the different analysis methods all at once and in conjunction with each other.

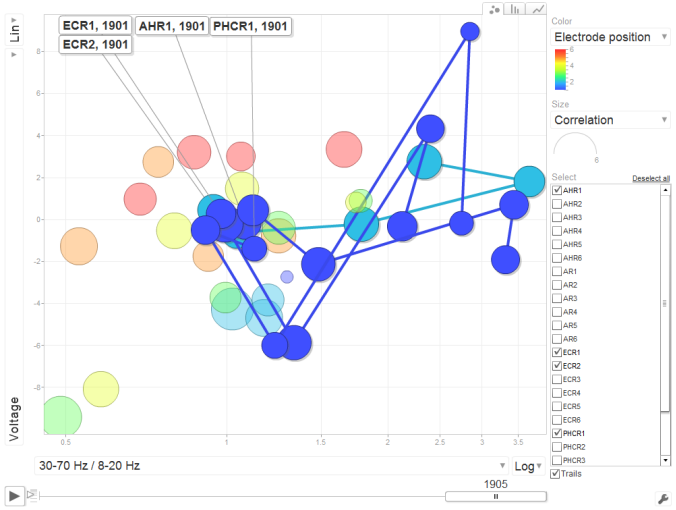


Figure 37: Deepest (blue) electrodes seem to be most active

We used Google Motion Chart API, which is based on the Gapminder software ideas. Data analysis steps were done in Matlab (you can have a look at most important scripts in the Appendix). Python script was used to generate the HTML and JavaScript code necessary for communicating with the API.

One of the great sides of this implementation is the ability to change how measurements are mapped to the graphical objects. On the Figure 38 the same data is represented in different way, which can give some new ideas about the data.

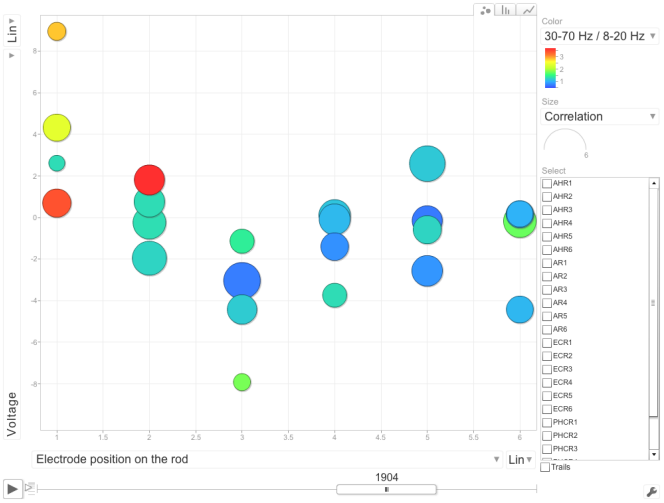


Figure 38: Leftmost bubbles represent internal electrodes, rightmost – superficial ones.

On the X-axis we now plot the 3D position, which means that leftmost bubbles represent the electrodes embedded deeply in the brain and rightmost are the superficial ones. Color now represents the ratio between higher and lower frequencies, the color bar is shown in the right top corner. We can see that after the stimulus (400ms) leftmost electrodes indicate high activity of both voltage and frequency parameters, which confirm the assumption we made before.

We believe this tool can help researches to better understand their data and notice unexpected behavior or surprising relations.

6 Concluding remarks

There are many lots of good methods for visualizing EEG data but normally they are shown separately. We believe that it is possible to combine the data into one graph which would make the analyzes of the data simpler.

Our initial result is already interesting and can be rather valuable tool for a neuroscientist.

In addition, we understand that there are many ways how to improve the result. For example, it is possible to add an extra dimension – it can be done with the shape of the bubble. What is more, it is possible to come up

with lots of new measures which could be useful. Such as phase synchrony and causality measures.

To sum up, we believe that this kind of development is needed in the neuroscience community. If our approach end up being useful we are planning to build a proper toolbox that can be useful for anyone.

7 Appendix

Here are the scripts we used to analyze data (presented FieldTrip format) and create it's representation on the Bubble chart.

Listing 1
Matlab scrtipt: draw time-frequency heat maps for every channel, store frequency decomposition data.

```
fam_nfam = fam;
fam_nfam.trial = [fam_nfam.trial nfam.trial];
fam_nfam.time = [fam_nfam.time nfam.time];
fam_nfam.cfg.trl = [fam_nfam.cfg.trl; nfam.cfg.trl];

freqs_fam_nfam_together = [];
freqs_fam_nfam_datamatrix = [];
for i=1:24
    cfg              = [];
    cfg.channel      = fam_nfam.label(i);
    cfg.output       = 'pow';
    cfg.method       = 'mtmconvol';
    cfg.taper        = 'hanning';
    cfg.foi          = 8:70;
    cfg.t_ftimwin    = 5 ./ cfg.foi;
    cfg.toi          = -0.3:0.02:0.7;
    freq_res_fam_nfam = ft_freqanalysis(cfg, fam_nfam)

    freqs_fam_nfam_together{i} = freq_res_fam_nfam;

    subplot(6,4,i);

    cfg = [];
    cfg.baseline     = [-0.6 -0];
    cfg.baselinetype = 'relative';
    cfg.zlim         = [0.7 1.5];
    cfg.colorbar     = 'no';
    [cfg datamatrix] = ft_singleplotTFR(cfg, freq_res_fam_nfam);
    freqs_fam_nfam_datamatrix{i} = datamatrix;
end
```

Listing 2

Matlab script: collect all necessary data, put it into data matrix, save as .csv file

```
% Bubble Setup 1 -- All trials
%   size - average correlation between current channel and all other channels
%   color - RGB encoded 3D position
%   X-axis - beta / alpha
%   Y-axis - average of absolute value of the voltage over 'ywin' last timeframes
% time - time

% unite fam and nfam
fam_nfam = fam;
fam_nfam.trial = [fam_nfam.trial nfam.trial];
fam_nfam.time = [fam_nfam.time nfam.time];
fam_nfam.cfg.trl = [fam_nfam.cfg.trl; nfam.cfg.trl];

% dataset variables
tstart = 700;
framesize = 20;

% colors
%rgb2hsv;
%colors = zscore(hue);
colors = [1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6 1 2 3 4 5 6];

data = [];

for ch=1:24
    chdata = [];

    % voltage average & baseline for whole channel
    voltage_average = zeros(1,2000);
    for tr = 1:153
        voltage_average = voltage_average + fam_nfam.trial{1,tr}(ch,:);
    end
    voltage_average = voltage_average ./ 153;
    baseline = voltage_average(1:1000);
    voltage_average_std = voltage_average ./ std(baseline);

    % name
    name = ones(5,1) .* ch;

    % time
    time = [800,1000,1200,1400,1600];

    % correlation
    corr = zeros(5,1);
    for t=1:5
        correlation = 0;
        for tr = 1:153
            correlation = correlation + mean(abs(corrcoef(
```

```
                    fam_nfam.trial{1,tr}(:,time(t)-50:time(t)+50)') -
                     eye(24)));
49          end
50          correlation = correlation ./ 153;
51          corr(t) = correlation(ch);
52      end
53
54      % color
55      color = ones(5,1) * colors(ch);
56
57      % frequencies
58      frequencies = zeros(5,1);
59      voltage = zeros(5,1);
60      for t=1:5
61          tm = floor((time(t) - tstart) / 20) + 1;
62
63          % rows -- frequencies 1 to 63 (8Hz to 70Hz)
64          % columns -- time from 1 to 42 (700ms to 1700ms, frame
                20ms)
65          low_freq_avg = ...
66              sum(freqs_fam_nfam_datamatrix{1,ch}(3:12, tm)) / ...
67              numel(freqs_fam_nfam_datamatrix{1,ch}(3:12, tm));
68
69          high_freq_avg = ...
70              sum(freqs_fam_nfam_datamatrix{1,ch}(23:63, tm)) /
                    ...
71              numel(freqs_fam_nfam_datamatrix{1,ch}(23:63, tm));
72
73          frequencies(t) = high_freq_avg / low_freq_avg;
74
75
76      end
77
78      % voltage
79      for t=1:5
80          voltage(t) = mean(voltage_average_std(time(t)-20:time(t)
                +20));
81      end
82
83      time = [1 2 3 4 5];
84      chdata = [name, time', corr, color, frequencies, voltage];
85
86      data = [data; chdata];
87  end
88
89  data(:,3) = zscore(data(:,3)) + 3;
90
91  dlmwrite('../googleViz/bubble_1_all_trials_3_moments.csv',data,'
        delimiter',',');
```

Listing 3
Python script: generate .html file from .csv (Note: this file must be run on a web server)

```
1 #
2 # Generate HTML file for Google BubbleChart from MATLAB matrix
3 #
4
5 filename = 'bubble_1_all_trials_3_moments'
6 electrodes = ('','AR1','AR2','AR3','AR4','AR5','AR6','AHR1','
      AHR2','AHR3','AHR4','AHR5','AHR6','ECR1','ECR2','ECR3','ECR4'
      ,'ECR5','ECR6','PHCR1','PHCR2','PHCR3','PHCR4','PHCR5','PHCR6
      ')
7
8 html = open(filename+'.html', 'w')
9 # head
10 html.write("""
11 <html>
12   <head>
13     <script type="text/javascript" src="http://www.google.com/
         jsapi"></script>
14     <script type="text/javascript">
15       google.load("visualization", "1", {"packages":["
           motionchart"]});
16       google.setOnLoadCallback(drawChart);
17
18       var options = {};
19       options['state'] = '{"dimensions":{"iconDimensions":["dim0
           "]},"orderedByX":false,"time":"1901","iconType":"BUBBLE
           ","xZoomedDataMin":0.48966,"yZoomedDataMin":-9.4157,"
           xAxisOption":"4","yLambda":1,"playDuration":15000,"
           nonSelectedAlpha":0.4,"uniColorForNonSelected":false,"
           yAxisOption":"5","orderedByY":false,"xZoomedIn":false,"
           xZoomedDataMax":3.6768,"yZoomedDataMax":8.9485,"
           yZoomedIn":false,"xLambda":0,"colorOption":"3","
           duration":{"multiplier":1,"timeUnit":"Y"},"
           iconKeySettings":[{"key":{"dim0":"ECR2"},"trailStart
           ":"1901","LabelX":-230,"LabelY":-260},{"key":{"dim0":"
           ECR1"},"trailStart":"1901","LabelX":-241,"LabelY
           ":-294},{"key":{"dim0":"AHR1"},"trailStart":"1901","
           LabelX":-164,"LabelY":-307},{"key":{"dim0":"PHCR1"},"
           trailStart":"1901","LabelX":-67,"LabelY":-344}],"
           sizeOption":"2","showTrails":true}';
20       options['width'] = 1001;
21       options['height'] = 750;
22
23       function drawChart() {
24         var data=new google.visualization.DataTable();
25         data.addColumn("string", "Electrode");
26         data.addColumn("number", "Timeframe");
27         data.addColumn("number", "Correlation");
28         data.addColumn("number", "Electrode position on the rod
             ");
29         data.addColumn("number", "30-70 Hz / 8-20 Hz");
30         data.addColumn("number", "Voltage");
31         data.addRows([""");
32
33 # data
```

```
34 f = open(filename+'.csv')
35 for line in f:
36   (name, time, bsize, color, x, y) = line.split(',')
37   html.write('["'+electrodes[int(name.strip())]+'",'+time.strip
       ()+','+bsize.strip()+','+color.strip()+','+x.strip()+','+y.
       strip()+'],\n');
38 f.close()
39
40 # tail
41 html.write("""
42         ]);
43         var chart=new google.visualization.MotionChart(
44         document.getElementById("chart_div"));
45         chart.draw(data, options);
46       }
47     </script>
48   </head>
49   <body>
50     <div id="chart_div" style="width:1001px; height:750px;"></
         div>
51   </body>
52 </html>""");
53
54 html.close()
```

References

[1] M. ten Caat, N. Maurits, and J. Roerdink, "Tiled parallel coordinates for the visualization of time-varying multichannel eeg data," in *Proc. Eurographics/IEEE VGTC Symposium on Data Visualization (EuroVis)*, p. 61, 2005.

[2] O. Jurjuţ, D. Nikolić, G. Pipa, W. Singer, D. Metzler, and R. Mureşan, "A color-based visualization technique for multielectrode spike trains," *Journal of neurophysiology*, vol. 102, no. 6, p. 3766, 2009.

[3] Gapminder, "Gapminder," November 2011. `http://www.gapminder.org/`.

[4] E. C. Series, "Visualizing biological data," November 2011. `http://vizbi.org/`.

[5] D. Schomer, F. Silva, and F. Silva, *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields.* Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, Lippincott Williams & Wilkins, 2010.

[6] Wikipedia, "Fourier transform — wikipedia, the free encyclopedia," 2011. [Online; accessed 30-November-2011].

[7] P. L. Nunez and R. Srinivasan, "Electroencephalogram," *Scholarpedia*, vol. 2, no. 2, p. 1348, 2007.

[8] M. t. Caat, *Multichannel EEG Visualization*. PhD thesis, University of Groningen, XXXX.

[9] Wikipedia, "Self-organizing map — wikipedia, the free encyclopedia," 2011. [Online; accessed 25-November-2011].