

Coming full circle: using interactive data visualization to understand visual information processing in the human brain

Kamen Pavlov

Supervisor: Prof. Jan Aerts
Visual Data Analysis Lab
Co-supervisor: Prof. Cees Van
Leeuwen
Laboratory for Perceptual Dynamics
Mentor: Jansi Thiyagarajan
Visual Data Analysis Lab

Thesis presented in
fulfillment of the requirements
for the degree of Master of Science
in Statistics

Academic year 2015-2016

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 3 |
| 1.1 | Motivation | 3 |
| 1.2 | Domain and Data | 4 |
| 2 | Methodology | 7 |
| 2.1 | Artifacts | 10 |
| 2.1.1 | Proto-Persona | 10 |
| 2.1.2 | Hypothesis Statement | 10 |
| 2.1.3 | Statement of Goals | 11 |
| 2.1.4 | Practical Considerations | 13 |
| 2.2 | Tools | 13 |
| 2.2.1 | Ideation Session | 13 |
| 2.2.2 | Five Design Sheet Technique | 14 |
| 3 | Understanding the User Needs | 15 |
| 3.1 | Diving into the domain research | 15 |
| 3.2 | Summary of requirements | 22 |
| 3.3 | Literature Review | 24 |
| 4 | Data & Design | 27 |
| 4.1 | Data | 27 |

CONTENTS

| | | |
|----------|---|-----------|
| 4.1.1 | Data Sets | 28 |
| 4.1.2 | Data Exploration | 29 |
| 4.1.3 | Data Pre-processing | 34 |
| 4.2 | Design Process | 34 |
| 4.2.1 | Five Design Sheets Technique | 34 |
| 4.2.2 | Tool Development & Implementation | 43 |
| 5 | Discussion | 51 |
| 5.1 | Methodology | 51 |
| 5.2 | Visualization Improvements | 51 |
| 6 | Summary | 53 |
| 7 | Appendix | 55 |

Acknowledgments

Here I want to express my gratitude to my Promoter Jan Aerts and to my Co-Promoter Cees Van Leeuwen for the given opportunity and for the provided facilitation. I am really thankful to my promoter Jansi Thiyagarajan for her big involvement and excellent guiding. At the same time I am really thankful to the domain experts David Alexander and Marcello Giannini for the amazing collaboration and making my life easier by providing the whole information I needed.

Dedicate my work to my parents Dimitar Pavlov, Antoaneta Pavlova, to my brother Emiliyan Pavlov, to my soulmate Danica Dimitrijevic and all other relatives and friends which put up with my ignorance during the writing period. Special thanks to Bushra Monowar Duti, who was in the same time under pressure writing her master thesis and giving me positive example of a positive attitude and big determination.

Chapter 1

Introduction

The data bombarding is happening from every corner of the human activities. This big information flow demands sufficient techniques of analyzing and interpreting the data. In this circumstances the visualization arises as integral part of the data analysis, because of its ability to summarize the data in intuitive for the user manner. The visualization as a means of communication has a big advantage being external perception medium. The internal memory and cognition information processing capacity can be extended substantially by using external perception media such as the visual cognition. The visual perception is superior in comparison to the other external perceptions such as sound, touch and smell, because the brain activity connected to the visual perception is way bigger than in the case of the other external perceptions. It is well known fact that the visual sense of the human being is the most capable sense for perceiving large systems and gives us ability to distinguish patterns very well. In this thesis we will take advantage of this human visual ability by exploring the use of visualization to aid the researchers in the neuroscience Perceptual Dynamics Laboratory of KU Leuven.

1.1 Motivation

Domain experts often do not have time to analyze and interpret the data and they are just using traditional statistical and analytical methods. They need visual analytics medium to make sense of data quickly and focus more on their core research while seeking the help of visualization designers and specialist. For providing superior visualization product the communication between the domain experts and the visualization designers is of major importance. Without putting effort in understanding the needs of the domain experts the visualization designer will have a hard time achieving good assonance between the goals of the domain experts and the visualization. That's why one of the main considerations within the master thesis is understanding good the needs of the domain experts. The main purpose is establishing the data visualization design pipeline by trying different techniques in the design process. These techniques will be explained in detail in the chapter Methodology. As still novice I will try various techniques in order to find out standardized Visualization Design

pipeline which I could use also in the future as a part of eventual work with clients. Ideally I will be happy if some of the results in thesis could help out other people in their work designing a similar products as data visualization for their clients. As a part of the master thesis results we will discuss the extent to which the used techniques proved to be useful.

In general a piece of data is being analyzed in order to answer specific, prior defined questions from the data by using statistical and machine learning methods. As stated in [8] that scenario with concrete and well defined questions the decision making can be left to automated computer program which makes the decision for us on basis of pure mathematical reasoning. If no human judgment is needed, then data visualization design is also not required. In the context of our case study the domain experts are researchers which are dealing with huge amount of data and the decision making cannot be boiled down only to using computer programs. They themselves are trying to make sense out of the data, which includes looking at different visualizations. Thus within my master thesis I have attempted to help them by designing for them useful, two dimensional, novel visualization.

My personal motivation of doing the master thesis arose from my fascination in revealing patterns in complex data structures. In our case study for instance we have complex six dimensional hyper-cube structured data in which by looking at the raw measures across this six dimensions we can't grasp much of information. The "programmed aesthetics" of Manfred Mohr for example is a bright example how patterns can be derived from the multi-dimensionality of the hyper-cube with 100 dimensions boiled down in 2 dimensional representation. This representations of Manfred Mohr have more artistic value than information representative value. Here within the master thesis I will put as main priority the informative power and of the visualization and then will consider improving it's aesthetic look.

I have also big motivation dealing with the particular case study because of my curiosity and fascination about the complicated world of the human brain. In the course of my master thesis I collaborated with Laboratory for Perceptual Dynamics(PDL) at KU Leuven which goal is the characterizing of the dynamics of visual information processing. For that purpose the researchers of the lab are combining psychophysical measurement with computational modeling and mathematical analysis. They investigate the processes underlying awareness, triggered by visual stimulation and activities such as eye movements and exploration of the environment.

1.2 Domain and Data

In the neuroscience field a common way of conducting an experiment is by measuring the subject's brain activity, while the subject is being exposed to some external stimuli. The brain activity is measured locally in the different regions of the brain by using different techniques which are acting on various scales of the brain activity. For instance the Local Field Potential(LFP) focuses on the population activity of several of ten thousands of brain cells, Subddural Electroocortogram(EECoG)-several of millions cells, and the largest scale of Electroencephalogram(EEG) and Magnetoencephalogram(MEG) with whole cortical region as measured focus [4].

The researchers from PDL are conducting experiments by recording the brain signals on the largest scales (EEG, MEG and ECoG) and they are in particular interested in the propagation of the brain activity over the cortex. With this purpose the brain signals are analyzed in the time-space-frequency domain. The brain signals in the experiments they are conducting are measured via different techniques and thus the raw signal can be in different forms. The microscale is not of interest for the domain experts, although in the future there could be possibilities to align their work with the recent scientific results on neuron scale level. The PDL research team investigates the dynamical nature of brain activity introducing space as a relevant variable in their methods. This is in contrast with some established research methods that consider brain activity as made of static sources and therefore analyze independently signals from the different recording channels instead of collating over space. This propagating wave of activity is called Traveling Wave (TW). Within their research they apply mathematical methods to detect TW events. For instance they build TW model by means of Linear Regression. This model in the future we will call TW model. They compare then the model to the measured brain activity to detect propagation events. Other phenomena which can be modeled is the Spiral wave (SW). There are also propagating through time and space, but in contrast to the TWs they have propagation with a form of a spiral. Right now the scientist from PDL are working on building a model which explains the behavior of the SWs. Further explanation of the data will be provided in the Chapter *Data & Design*4.

Chapter 2

Methodology

According to the stated motivation of the master thesis the methodology is a main focus of the thesis. These resources helped in collecting and synthesizing the needed information in order to find out the exact needs of the domain experts and to validate the visualization. Further we will call these resources artifacts and tools. The artifacts are simply pieces of information which are created in the context of use of certain tool. In the course of the communication with the domain experts the content of the artifacts will change so that in the end of the design process they will be complete and robust. All used tools and artifacts are listed in 2.1.

List 2.1: List of Tools and Artifacts being used.

| • Tools | • Artifacts |
|---|--|
| <ul style="list-style-type: none">- Design phase<ul style="list-style-type: none">* Ideation Session* Interviews* Articles- Understanding the user needs<ul style="list-style-type: none">* Five-Design-Sheets Technique | <ul style="list-style-type: none">- Proto-Persona- Problem Statement- Hypothesis Statement- Statement of Goals- Practical Considerations |

This artifact-tool framework arises from the necessity of neat documentation about the progress throughout the master thesis. It is also very intuitive way of documenting where in each action we use a tool in order to change the artifact which is a carrier of the of the information relevant to the design process.

The reason we are having so semantically close artifacts like *Hypothesis Statement* and *Statement of Goals* emerges from the need of rephrasing in order to truly grasp the domain experts problem from different perspectives.

The master thesis is constructed in the framework of the Four Levels of Validation proposed by Tamara Munzner in her book [8]. The four validation levels described in the book are represented in the List 2.2

List 2.2: The Validation levels

1. Domain Situation
2. Data \Task Abstraction
3. Visual Design and Interactivity
4. Algorithm

It should be noted that this four levels are cascading, which means that we can jump from one validation level to another at each time of the design process and that if we have done mistake in one of the levels automatically this mistake will affect the levels below. This framework with cascading levels is very useful in addressing the problems in thematic way, where we can independently validate at each level, if we have done well and have set the artifacts correctly. The mistakes possible mistakes done at each level are generalized in the List 2.3.

List 2.3: Typical mistakes at the given validation levels

- Domain situation: We misunderstood the stakeholders needs.
- Data \Task Abstraction: We are visualizing the wrong thing.
- Visual Design and Interactivity: The presentation of the data is wrong.
- Algorithm: The visualization algorithm is too slow.

If we spot a mistake in some lower level and can't fix it at that level. We have to go upstream at the upper level and try to find the root the problem and tackle it there.

The validation on the level of the domain experts is the first validation be established. This is also the most important level, because communicating and understanding the problem of the domain experts is crucial for the developing of the visualization. If we understand falsely the needs of the stakeholders, the visualization will be useless. That is the reason, why within the master thesis the biggest effort has been payed to understanding the domain situation.

The second validation level is the Data\Task Abstraction. Here having understood well the domain situation we translate the information gathered from the domain experts into **what-why-how** framework. With this framework we are interested in **what** exactly should be visualized, **why** it has to be visualized and **how** to visualize it.

The **why** question converts the information from the domain experts into abstractions *actions* and *targets*. The *targets* are desired structural properties of the data which should be revealed. The *action* what action we should take in order to achieve our visualization *targets*. In the List 7.1 in the Appendix we can see possible abstractions which are taken again from the Tamara Munzner's book [8].

The abstraction translate the domain experts needs in simple, more generic language. They have two main advantages. Firstly we free ourselves from the domain specific jargon and focus on the pure functionality of the visualization. And secondly we avoid going into the pitfall of thinking about visualization design, because at this stage it makes us constrain our thinking in concrete visualization concept with limited functionality. This abstractions encode everything we need for the designing the visualization and will be stored in the artifact *Statement of Goals*.

The third validation level is Visual Design and Interactivity. As the name implies that stage consists of two parts. The first part is the visual design i.e. what the user sees and the interaction part i.e. how the user manipulates the visualized content. in this stage we will use the *Five Design Sheet* tool, which will be discussed in the section *Tools*. With the help of this tool we will build our final visual design concept along with the interactive features.

The fourth validation level is called Algorithm. Here we validate whether visualization algorithm is too computationally complex. Once we fixed the visualization concept, in the course of implementation of the concept it might be the case that it takes too much time to render the visualization or that some interactive features can't be implemented because of some constrains of the visualization engine being used.

In the Figure 2.1 we can see the visualized version of our general methodology structure. Each circle represent one of the validation phases.¹ The fact that they are drawn in one another shows their nested nature. The outer bigger circle shows the domain situation validation. It is not a coincidence that the domain situation circle is the biggest one, since understanding the domain situation determines to biggest extent how successful will be the visualization. Here the size corresponds to the effort put in each validation phase. The black arrow pointing up and down symbolizes the possibility to go upstream and downstream between the levels.

On the right side of the graph we see next to each validation level we have corresponding one or more tools which are used on the particular level. The main purpose of using the tools is acquiring relevant to the design process information. This information is stored in the artifacts. More information about each artifact and tool can be seen in the next two sections.

¹The colors from the graphs will be used in the coming chapters to indicate the certain validation set in which we update the artifacts, whereas each artifact will be shown in a frame with the particular color.

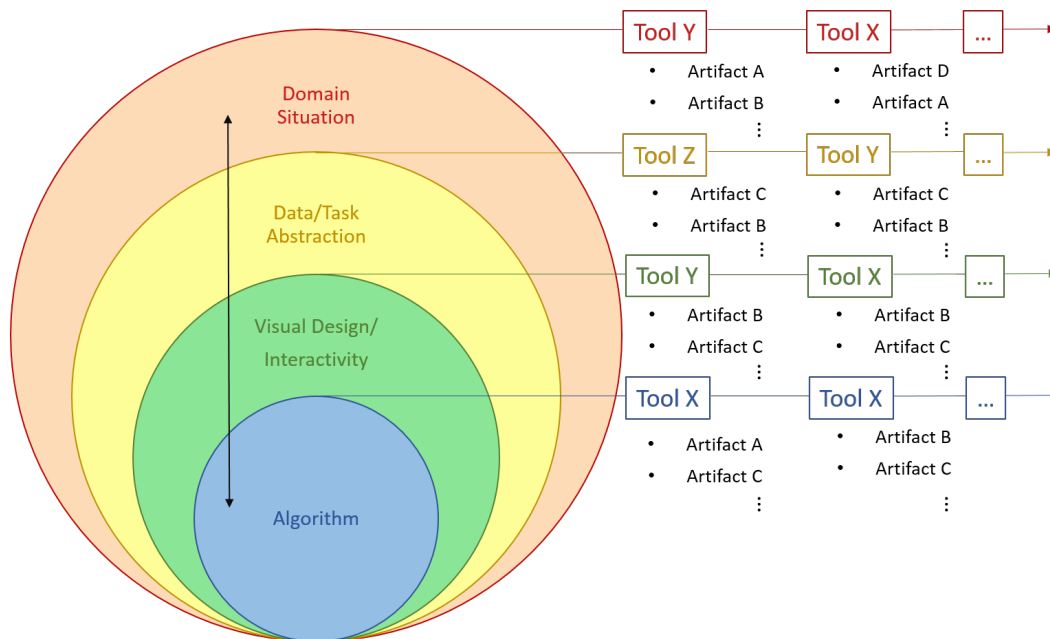


Figure 2.1: Validation phases

2.1 Artifacts

2.1.1 Proto-Persona

The proto-persona artifact represents the relevant characteristics, behavior, activities, needs and means for satisfying these needs. By stating the proto-persona one can get better idea about the profile of the end user and about the reasons why he/she will use the future visualization. The proto-persona here should express the view of the designer of the end users, which in our case study are the domain experts from the Perceptual Dynamics Lab. A template of the proto-persona artifact can be seen in Template 2.1.

Template 2.1: Proto-Persona

| | |
|----------------------|-----------------------------|
| Picture, name, role | Behaviors & Characteristics |
| Needs or Pain points | Possible remedies |

2.1.2 Hypothesis Statement

Another artifact, which has been used as a part of the methodology is *Hypothesis Statement*. The hypothesis statement is one of the basis of each research and project, because it is entangled with the expected outcome. In order to start a project we first build assumptions, which roughly express our view about true nature of the environment of the project. By stating the Hypothesis basically we design our project according to the assumptions we have.

Within the research measures have been conducted and afterwards the measures used to validate or reject the stated hypothesis. In our context we build our hypothesis statement by finding out the underlying hypothesis the researchers are testing within their research. This approach is very useful, because of two reasons. Firstly the measures are in our case the input for the visualization, therefore by discovering the domain experts hypothesis we will find how to use their data for answering their scientific questions. Secondly we will get more insides about the purpose of their research. The hypothesis statement form we use is shown in graph, This hypothesis statement template is taken from the book [5].

Template 2.2: Hypothesis Statement

We believe that [doing this/building this feature/creating this experience] **for** [these people/personas] **will achieve** [this outcome].

We will know this is true when we see [this market feedback, quantitative measure, or qualitative insight].

2.1.3 Statement of Goals

The statement of goals is crucial artifact in the the design process. Before we jump into implementation we have to spend a lot of time reasoning why we want to do the visualization and what purpose it will serve to. The goal of visualization is normally satisfying the need of the user to get certain information, but it can also change the users point of view after she/he uses it. As it is suggested in [6] we have to state the goals in terms of the knowledge, which has to be acquired from the visualization, without referring to a specific content or implementation of design. This corresponds with the idea of the abstraction level, which is explained further.

The reasoning behind not mentioning about any particular implementation lies in the potential danger of getting into the in-the-box thinking. After adding any information of implementation you loose instantly substantial creativity potential, because you've prematurely picked up a design option, which will constrain your thinking in the future. The statement of goals should reveal the desired functionality of the future visualization.

The statement of goals major use will be in the *Data Task/Abstraction* and the *Visual Design/Interactivity* phases, where we will apply it in regard to the **why** and **how** taxonomy abstractions. As already mentioned in the **why** abstraction we will use *actions* to define the domain expert goals and *targets* to highlight the important aspects of data.

The Figure 2.2 shows in which validation phases what part of what artifact we mainly update. The colored artifacts are updated in the according stage. In the phase of finding out the domain situation we update the artifacts *Proto-Persona* and *Hypothesis Statement*. Afterwards, on the Data Abstraction level we translate the already acquired information in *Proto-Persona* and *Hypothesis Statement* into the **what-why-how** abstractions. Further on the vi-

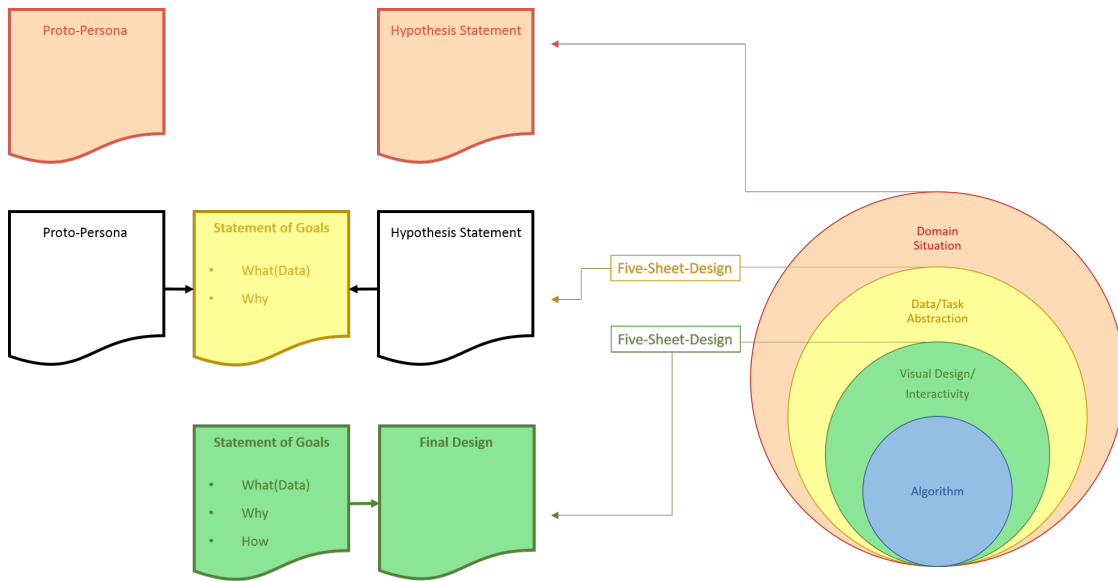


Figure 2.2: Use of the general use of the artifact *Statement of Goals*

visual design level we arrange the **how** abstractions in complete visualization concepts. Out of these design concepts we choose the final design. Note that we update the abstractions and choose the final design in the context of use of the Five-Design-sheets technique which will be explained in the next section.

The Statement of goals has simple form a description of what is being visualized and of a list with the stated goals and the corresponding **why** and **how** abstractions.

Template 2.3: Statement of Goals template

| |
|--|
| <p>What:</p> <p>Goal 1:</p> <p>Why:</p> <p>How:</p> <p>Goal 2:</p> <p>Why:</p> <p>How:</p> <p>...</p> |
|--|

2.1.4 Practical Considerations

The artifact Practical considerations shows the encountered problems throughout the designing process. The practical considerations are associated here with the problems after the visual encoding and the interaction idiom part, because the design is still just on blueprint and once it is implemented in practice you can see its limitations and flaws. In this artifact we list all of this problems with the visualization and we try to work around them and fix them. The columns with information about the practical consideration will represent the key concepts shown in Table 2.4 If we cannot either fix them or work around them in the algorithm part by optimizing the algorithm, we go back in the previous conceptual phase (the visual encoding and interaction idiom) phase and try to redesign the visualization using the artifacts and tools which are available in that phase. In case the practical consideration is solved, then we move on.

Template 2.4: Practical Considerations

Context:

Observation:

How might we...

So that...

Found Solution:

2.2 Tools

The tools used in the course of the master thesis at certain validation stage serve for achieving a particular purpose. Here we will elaborate the purpose of each tool.

2.2.1 Ideation Session

The limits of the possible can only
be defined by going beyond them
into the impossible

Arthur Clarke

The goal of this tools is to stimulate the Domain Expert to discover visualization perspectives he/she has never thought of before. We will call this hypothetical situation Sci-Fi Scenario, because the modern technology and scientific progress are not advanced enough

to make that hypothetical situation reality. Putting the domain experts imagine that situation makes them think without being constraint by the status quo. That way their feed back from the ideation session can reveal some underlying needs for visualizing which are not considered relevant before.

This approach can be considered as a contextmapping technique. The context of use of the product is very important aspect by designing that product. As described in [11] the context has many other properties aside time and space. The actors and the environment are also big determinant for the context in which a product is being used. The dynamics and the action are also big part of the user's experience with the product. The designer always have an idea about the context of the use of the product, but this idea is subjective. By engaging the users in sessions as the one described above, the designer can build empathy towards the user, to discard preset false assumptions about the users needs and to find good insides about the possible application of the product.

Template 2.5: Sci-Fi Ideation session

1. **Preparation:** Preparing the session by choosing Sci-Fi scenario and directions for exploration in the actual session
2. **Sesitization:** Mentally preparing the participants for the actual session
3. **Actual Session:** Conduct the Sci-Fi ideation session with the domain experts
4. **Analysis:** Analyze the gathered information from the session
5. **Adaptation:** Update the artifacts according to the gathered and analyzed information

2.2.2 Five Design Sheet Technique

The *Five Sheet Technique* is a tool which helps us funnel our creativity into coming up with visualization concepts. It is published by the School of Computer Science at Bangor University <http://fds.design/>. In List 7.2 in the Appendix we can see an overview of all activities performed in the first step as it is originally proposed. Since we've already done thorough filtering of ideas via the artifacts *Hypothesis Statement*, *Proto-Persona* and *Statement of Goals* we will skip that step in our methodology. The performed steps by us can be seen also in the Appendix.

Chapter 3

Understanding the User Needs

In that chapter we represent the results of the applied tools for identifying the user needs. Within the use of each tool we update the information in the relevant artifacts, so that the reader has good impression about the evolution of the results in the design process.

3.1 Diving into the domain research

I got my first impressions of the domain situation at my kickoff meeting with the domain experts. At that meeting I had short talk with those members of the PDL team which had potential need of visualization in their work. Eventually I ended up working with two of them - Marcello Giannini and David Alexander. So from now on in each updated artifact information gathered from Marcello and David will be used. The information from our first meeting is summarized in the Proto-Persona artifacts of the two domain experts.

Artifact: Proto-Persona I

Name, Role: Marcello Giannini, PHD Student, Perceptual Dynamics Department

Behaviors & Characteristics: He investigates the occurrence of the traveling waves by working on different methods for their detection. He mostly uses heatmaps as visualization tool.

Needs or Pain points: The heatmaps represent the data just in two domain dimensions(x- and y-axis).

Artifact: Proto-Persona II

Name, Role: David Alexander, Research Scientist, Perceptual Dynamics Department

Behaviors & Characteristics: Investigates the behavior of traveling waves in various experiment conditions. He spends time in implementing by himself visualizations for easier perceiving of the traveling wave behavior.

Firstly in order to get initial idea about the research field of the stakeholders and about the data they are working with I read the papers [3], [1] and [2] published by the domain experts. The Hypothesis Statement can be updated really sufficient from reading the domain researchers papers, because the papers provide you information about the main research goals and the underlying hypothesis the domain experts are testing in their research. The tricky part here is how to adapt the domain experts hypothesis to the *Hypothesis Statement* artifact, which should express the designer's hypothesis about the needs of the domain experts. It should be noted that by stating the hypothesis only on basis of the papers we constrain ourselves only in the space of the already explored in the scientific field. That's why later we will use the Sci-Fi Ideation Session in order to expand our design horizons.

Artifact: Hypothesis Statement

The domain experts believe that the traveling wave along the trials can explain bigger portion of the information about the brain activity than just looking at the averaged signal across trials.[2].

After getting to know the research field of the domain experts through the papers they gave I also made interviews asking them fixed questions prepared beforehand. This kind of interview is chosen, because that way we collect data in highly organized manner, so that the designer can get concrete information from the user more efficiently. The purpose of the interviews is to acquire concrete information, which is non-present in the papers and which will be useful for the visualization design process. The Appendix contains the questions I asked and the summarized answers I received from Marcello and David. I also give the underlying reason I asked each particular question. It should be noted that although I prepared fixed question within the talk some other question have been asked. That's why the information as an answer to all asked question during the interview.

In nutshell from the interview I got already sufficient information about stakeholder's pain points and the nature of their research, so that I am able to formulate problem driven tasks for the solving of this pain points.

Artifact: Proto-Persona I

Name, Role: Marcello Giannini, PHD Student, Perceptual Dynamics Department

Behaviors & Characteristics: He investigates the occurrence of the traveling waves by working on different methods for their detection. He mostly uses heatmaps as visualization tool.

Needs or Pain points: The heatmaps represent the data just in two domain dimensions(x- and y-axis). The heatmaps he's using don't get the data automatically and he has to generate heat map for each chunk of data. He doesn't have much time for implementing visualizations.

Possible remedies: Visualization for confirmatory and exploratory purposes, which has automated way getting the data and which discovers patterns in multidimensional structures. Except of traveling waves other phenomena in the EEG signals could be explorable.

Artifact: Proto-Persona II

Name, Role: David Alexander, Research Scientist, Perceptual Dynamics Department

Behaviors & Characteristics: Investigates the behavior of traveling waves in various experiment conditions. He spends time in implementing by himself visualizations for easier perceiving of the traveling wave behavior.

Needs or Pain points: It is difficult for a lay person to understand the behavior of the traveling waves and spiral waves with conventional visualizations they are using in their research. Need for person outside of the domain field, which can help coming up with more intuitive visualizations of traveling waves.

Possible remedies: Visualization which reveals in intuitive way the characteristics of the traveling wave: direction, phase, positions in time and space.

Artifact: Hypothesis Statement

The domain experts believe that the traveling wave and the spiral wave model along the trials can explain bigger portion of the information about the brain activity than just looking at the averaged signal across trials[2].

They will know this is true, when they see that the correlation between the TW or the SW model and the true brain activity signal is bigger than the correlation between the ERP and the true brain activity signal.

Ideation Session

In the next step we design Ideation session by giving maximal imaginative freedom to the participants, so that they can unleash their creative potential and help us to understand their visualization needs.

With the help of the conventional techniques as interviews and observations the designers can derive explicit knowledge based only on the participants present and past experience. Sometimes however, the needs of the users are strongly based on *tacit* knowledge which is not directly observable, but it lies in the hopes, dreams and passions of the domain experts which may come on the surface by letting the users acting in sci-fi scenario. The Sci-Fi scenario will provoke the domain experts to experience in the provided context, so that their imagination and memories work together. It is important that this collaboration of imagination and memories is not restricted by some constrains in the context. thus by designing the session too much additional information about and constrains by introducing the Sci-Fi scenario to the domain experts should be avoided. The only specific given information about the Sci-Fi scenario will be the following:

"Imagine that there is some almighty device, which is acting on different scales in the brain and which can collect any kind of information. What information would want the device collect for you."

This almighty device manifests the ultimate scientific tool in their domain, which can help the throughout their research. The participants will get automatically excited about the possible existence of such a tool and will become really engaged into the session. The stated ability that this machine can act on different scales gives a hint to the participants that they could explore any possible scale data can be collected from. For example possible data collection scales might be the scale of the human body scale, neuronal scale or more global scale such as human society.

For each participant a toolkit has been prepared so that they can easily transmit their ideas. The tool kit consists sticking notes, pencil and a sheet of paper. On the paper sheet where three example scales are shown one after another according to the size of scale starting from the biggest. The participants should stick the note at that scale field of the note which has the closest scale size with the desired measure written on the note. Afterwards the information of the notes will be discussed with the participants in the Ideation session. The results on the sheet of paper can be seen in the Picture 7.2 in the Appendix.

Analysis

The results of the Ideation session was in form of sticky notes placed on sheet of paper with stripes as a reference indicating some scale on which the almighty device can measure. On the sticky notes the domain experts wrote down the desired measures which the device should provide. This measures are summarized below:

- i *Mechanism of propagation*
 - *volume conduction*
 - *Glia cells*
- ii *Source of Propagation*
- iii *Neural Network that supports TW activity*
- iv *Source localization and its dynamics(causality of the events)*
- v *Measure of diffusion through cells*
- vi *Connections from neuron to neuron(Connectome pathways)*
- vii *Functional role of the single event*
- viii *Measures which provide straight forward comparison between the Traveling Wave process view and other processes which present the brain activity in different light*
- ix *Measures of the impact of the Traveling brain signals to some form of global consciousness communication*

The *volume conduction* measure shows how well the electric signal is collected by the electrode, because the nerve tissue and other matter are as obstacle to the signal and they add noise in the measurement. The noise can be defined also as the amount of unexplained, useless information from the signal. This desired measure of the domain experts shows their wish to have an overview of the noise relevant to the signal. The interesting thing is that the noise in this context is due to some external conditions as the brain and scalp tissue and the distance of the electrode from the source of the signal. From that perspective the noise is dependent of the source of the propagation of the TWs and could be aligned with the *TW Model*. Thus measures which determine the signals noise such as decibel(dB) could be incorporated and then compared via some visualization concept.

The *glia cells* are taking care of the translation of electricity from one cell to another. This is in relation to other two wishes of the domain experts: *measure of diffusion through the cells* and the *anatomical connections from neuron to neuron(pathways)*. Although this wish is not in direct relation to the research of the domain experts, they acknowledge the future possibility to merge the connectome pathways theory with their TW theory. Thus visualization which could be possible bridge between these two theories is also ends up in the bucket list of the stake holders.

Source of propagation and the mechanism of propagation also desired output from the ultimate device. This information is however already part of the acquired information from the other tools so far. In general we included so far in our artifacts the desire of the domain experts of dynamic representation of the traveling, whereas this source of propagation is some part of the dynamics of the traveling wave.

Causality of the events is another property which is desired to be detected by the ultimate device. By *causality of events* the domain experts mean information about the exact cause of the TW activity and its dynamics. The most of the experiments in the PDL have been conducted with visual stimuli. Conclusions about the causality of the events can be made by looking at the *Wave Activity* at the *Experiment conditions* hyper-cube dimension.

The *Functional role of the single event* can be seen as a hypothetical/latent function, which describes the TW activity completely. This function has as an input a particular measures and shows similar output across the trials with the same experimental conditions. The wish of the domain experts that the ultimate device can represent the TW as a function of other variables reveals a underlying need for building connection between the TWs and the available dimensions and measures in the data set. It should be noted that this connections should be with universal structure in our visualization so that also future, still not used data can be Shown afterwards. The samples(time) here plays a big role as a binding unit for establishing these connections. This role of the time as a connection comes from the fact that it is present in each from the measurements and also it is constant component of the commonly used visualizations by the domain experts. One measure, representative of the TW activity is for instance the *Wave fit*. Compare the measures and the dimensions in case of good *Wave fit* to the rest.

The property that the ultimate device can act on different scale levels, made the domain experts to raise some very important aspect for them and namely the view point from which one observes a process. In our case the process is the brain activity. One way of looking at that brain activity is throughout the TWs, which present it like dynamical waves which radiate from the brain, changing trajectory with the time and the different stimuli(events). This traveling waves are at best trackable on certain scale which depends on the wave frequency. Aside the scales on which this processes is observable, the point of view from which we observe the process is also important. For instance, if we observe the waves which are hitting Bondi beach from the shore they are going directly, we will be able only to see the height of some of the highest approaching wave. If we observe the waves from some place on the side of the beach we will be able to get other information about the waves as the number of the approaching waves, their width and the distance between them. In our case the Traveling Wave model is the result of looking at the brain activity process from certain perspective. Other perspective is the averaged signal across trials. In that regard comparison of the goodness of fit(Wave fit, ERP fit) across the dimensions is desired.

Adaptation

Now we incorporate this newly gathered information in our artifacts. This is the last time we update the artifacts *Hypothesis Statement* and *Proto Persona*.

Artifact: Proto-Persona I

Name, Role: Marcello Giannini, PHD Student, Perceptual Dynamics Department

Behaviors & Characteristics: He investigates the occurrence of the traveling waves by working on different methods for their detection. He mostly uses heatmaps as visualization tool.

Needs or Pain points: The heatmaps he's using don't get the data automatically and he has to generate heat map for each chunk of data. The heatmaps represent the data just in two domain dimensions(x- and y-axis). He doesn't have much time for implementing visualizations. He needs a way to discover what plays ultimately a role in determining the TW model. He also wants to investigate to what extent the noise is preserved in the different signal measures, because of some other external factors and because of the TW nature.

Possible remedies: Visualization for confirmatory and exploratory purposes, which has automated way getting the data and which discovers patterns in multidimensional structures. Except of traveling waves other phenomena in the EEG signals could be explorable and then compared between each other. One way of comparison might be between the distribution of the signal measures. By exploring the distribution of this signal measures across the dimensions can also reveal the patterns which determine the TW Activity.

Artifact: Proto-Persona II

Name, Role: David Alexander, Research Scientist, Perceptual Dynamics Department

Behaviors & Characteristics: Investigates the behavior of traveling waves in various experiment conditions. He spends time in implementing by himself visualizations for easier perceiving of the traveling wave behavior.

Needs or Pain points: It is difficult for a lay person to understand the behavior of the traveling waves and spiral waves with conventional visualizations they are using in their research. Need for person outside of the domain field, which can help coming up with more intuitive visualizations of traveling waves. He needs a way of determining the quality of the TW model across the dimensions.

Possible remedies: Visualization which reveals in intuitive way the characteristics of the traveling wave: direction, phase, positions in time and space and which compares brain activity from different points of view across the dimensions by representing their goodness of fit measure.

Artifact: Hypothesis Statement

The domain experts believe that the traveling wave and the spiral wave model along the trials **can explain** bigger portion of the information about the brain activity than just looking at the averaged signal across trials[2].

They will know this is true, when they see that the correlation between the TW model and the true brain activity signal is bigger than the correlation between the ERP and the true brain activity signal.

The domain experts believe that depending on the model(the point of view) different aspects of the pure brain activity are explained.

They will know this is true, when they see that the true signal behaves differently in the parts of the dimensions, where the *Wave Model* has big correlation with the brain activity signal.

It could be possible that the neural anatomy of the brain has an influence of the TW activity.

This could be shown by finding and comparing the source of propagation of the TWs to the neuronal connections.

It could be possible that across the different *Experimental Conditions* patterns about the influence of the events on the TW activity can be found.

This could be shown by showing the discrepancy between the true signal and the *TW Model* along the *Experimental Conditions* domain.

3.2 Summary of requirements

Here we convert the content of our artifacts *Hypothesis Statement* and *Proto-Persona* into a **what-why-how** abstractions. This abstractions are summarized in the body of the *Statement of Goals* artifact. More information about this artifact can be found in the Appendix chapter. In the **what** part we represent each aspect of the data, which has become clear to the designer so far. The **why** abstractions construct the reasons behind certain goal and the **how** abstractions build the solutions for this goals in the visual design context.

Goal 1: The first main aspect which stands out in the artifacts is about the *TW Map* measure and the need to be aligned to the some dimensions and other measures. The *TW Map* represents the brain activity signal from a dynamical point of view where the waves are propagating linearly. For fitting the *TW Map* we use the x, y, z coordinates of electrodes and we are doing it across the trial dimension, for each point in the other five hyper-cube dimensions Electrodes, Subject, Time, Frequency and Experiment Conditions. That means that we can track the *TW Map* in this five dimensions with other signal and model measures. The ERP is present also in this five dimensions and it represents the average over trials. Other candidates for comparison expressed in the artifacts are the Wave fit, the noise, etc. Wave Fit is not present in the electrodes dimension, because it is expressing correlation of the trial average ERP with the TW model. This correlation is calculating by integrating nodes dimensions. In connection to the traveling wave it turns out also to be desirable to explore the data so that we find out what other aspects from the data explain the TW activity.

Goal 2: The second matter summarized from the artifacts concerns the need of the domain experts to present the dynamic behavior of the TWs. The domain experts spent al-

ready quite time of trying to display the TW in intuitive way. They also use some conventional 3D neuroscience visualization software which represent the TWs on the scalp. This software is capable to represent the direction, but it can't reveal the source of propagation, which turned out to be desired property.

Goal 3: The third goal comes into picture as a necessity for the previous two goals and for the hyper-cube multidimensional nature of the data. Since the focus of the domain experts after all are the brain activity measures of the signal and of the TW model which are localized in the hyper-cube dimensions, one primal visual design consideration is look up and locate this measures along dimensions. This is also connected with the need of the domain experts to to get clear and intuitive view of the TW wave activity.

In the **how** abstractions we are answering the question of how the visualize the goals. After doing that our *Statement of Goals* is finished and now we can use it for generating the initial designs in the *Five Sheet Design* tool.

Artifact: Statement of Goals

What: Input: Grid of points with recorded time series brain activity data contained in 6 dimensional hyper-cube space:

- Dimensions
 - subject(nominal);
 - time/samples($\in R$);
 - electrode coordinates: $x, y, z(\in R^3)$;
 - frequency($\in R$);
 - trials(nominal);
- Measures
 - phase(continuous; $\in [0; 2\pi]$);
 - EEG voltage signals(continuous; $\in [10\mu V; 100\mu V]$);
 - experiment conditions(nominal);
 - *Wave fit*($\in [0; 1]$);
 - ERP ($\in [0; 2\pi]$);
 - *Wave Activity*($\in [0; 1]$);
 - *TW Model*($\in [0; 2\pi]$)

Goal 1:

Why: To **discover** and **compare trends** and **patterns** in the signal's measures in along different dimensions which align with the occurrence of TW and other phenomena happening in the brain topography along the dimensions.

How: **Juxtapose** and **map** the signal in relation to the dimensions and other measures by using colors, shapes, angles, bar charts etc. **Express** and **juxtapose** the signal measures **distributions** via box-plots, histograms, heat maps, line charts etc.

Goal 2:

Why: To **present** in easy to grasp way the dynamics of the traveling wave activity across the electrodes and time.

How: **Map** the behavior of the traveling waves in the space and time by using shapes, lines, angles, coordinates, etc.

Goal 3:

Why: To **look up** or **locate** desired measures in the dimensions.

How: **Highlight** the measures according certain **properties**(clusters, value ranges, etc.) by using colors, shapes, thickness of lines, coordinates, etc. **Filter** the observation units along measures dimensions.

■ what *action* idioms; ■ what *target* idioms; ■ how idioms;

3.3 Literature Review

Here we discussed the already established visualization approaches used in the literature of neuroscience. The first widely used approach is the simple *Line Chart*. The *Line Chart* represents the trend of the electrical signal, visualized along the time. One example can be seen in Figure 7.1 in the Appendix. The signal is the Event Related Potential which is time locked between some event indicated at time 0. For example, when the signal has other domain as frequency and we want to compare the signals along the frequency and time as well, the *Line Chart* can be still useful, by overlaying the other signal-time lines for the various frequencies. When we have a big number of frequencies however, the overlaying of too many lines will make the chart unreadable. That weakness of the *Line chart* is illustrated in the paper [10] where also an alternative of the *Line Chart* is proposed. This alternative is called *Lasagna Plot* in the paper, but more generic name for it is *Heat Map*.

An example of the *Heat Map* with domain expert data can be seen in Figure 4.3. This visualization method is used widely by the domain experts in their work. Plenty of examples for it can be seen in the domain experts paper [1]. The advantage of the *Heat Map* in compari-

son to the line chart is that it solves the overlaying problem by representing each line of at certain location on the y-axis so that we have 2D grid representation, whereas each cell of the grid is colored according to the value of the signal. The choice of colors could be either diverging or sequential [7]. The disadvantage of that visual representation is that it is valid only for representing signal in two domains and doesn't allow incorporating more than two dimensions.

Another way of visualizing brain activity is the topographical representation of the brain signal, where the value of the signal is represented as a color on 2D or 3D map of the brain. This way of visualizing is also commonly used by the domain experts and illustration of it can be seen in the Supplementary videos section on <http://www.sciencedirect.com/science/article/pii/S1053811913000633>. With this approach we show how the brain activity dynamically changes with the time in the different regions of the brain. The downside of the approach is restricted in visualizing only in the space domain of the signal and only the magnitude and dynamic characteristics as direction can be incorporated. Again here there is no room left for other domains.

From the paper [12] I got the idea of representing the electrodes recording the information on the scalp as points and combine them with the parallel coordinates.

Chapter 4

Data & Design

In this chapter explanation of the data is provided and then in the context of the visual design validation phase we create the final design of the visualization.

4.1 Data

The data has been provided to me after I already finished the gathering of the information about the domain situation. In that section the scale, the structure and the basic nature of the data will be explored. Understanding the data is crucial prerequisite for making good visualization design, because without knowing the data structure and scale, it could very be that we end up with useless design which doesn't fit the data format and can't handle the size of the data.

The data's nature is complex and highly dimensional with additional time variable which makes its interpretation complicated. The data includes EEG voltage signals from which are extracted phase and power for time and frequency interval of interest. Pure EEG signal is μV is being pre-processed with mathematical and statistical methods such as Independent Component Analysis(ICA) for denoising and Morlet wavelets for decomposing the signal also into the time-frequency domain. The resulting data is hyper-cube with six dimensions. This dimensions are shown below:

- Subjects(S)
- Experimental Conditions(C)
- Trials(L)
- Frequency(F)
- Samples(T)

- Electrodes(E)

Here we will explain each one of the hyper-cube dimensions listed above. Note that from now on we will refer to the hyper-cube dimensions by simply calling them dimensions. The Trials are time intervals with fixed duration during a single experiment, which are time locked around certain event. That event could be some visual stimuli in the context of the conducted experiment. The time zero is when the event occur. The time preceding assumes negative values and the time following time zero assumes positive values. The time is expressed in milliseconds(ms). The *Subjects* are the people participating in the experiment. The Experimental Conditions represent some certain attributes of the data which are recorded in the context of the experiment. This attributes are different from experiment to experiment and could be for example the on set and the onset of the eye movements of the subject. Another Experimental condition can be indicator, when button is being pressed. The Samples represent the time domain and correspond to the points within a trial. The *Frequency* represents the frequency domain where the oscillatory component extracted with wavelet analysis are defined. The *Electrodes* record the brain activity at certain location on the subject's scalp. Each one of the electrodes has particular location with 3 dimensional coordinates(x,y and z).

A conventional way of representing a hyper-cube space is via Star Scheme which can be seen in the Figure 4.1. We can derive the data in six dimensions. The fact table shown in the middle of the Star scheme is showing most of the measures which are used by the domain experts. We already mentioned the *TW model* and the pure EEG signal. After decomposing the signal in the time-frequency domain we get periodical waves localized in time and with certain frequency, phase and amplitude. In [2] it is said that the amplitude is not so reliable measure because of the volume conduction, thus the domain experts look mostly at the phase as a measure. One way of deriving the *TW model* is by fitting regression with the phase as a response and the coordinates of the electrodes as explanatory variables. The fit of the measured phase into the TW model is computed by correlation. This correlation results in the measure called *Wave Activity*, whereas it doesn't have the electrodes dimension as domain anymore. This *Wave activity* is averaged across the trials. On the other hand the *TW model* is averaged across trials. This is called the *Wave ER*. There is another measure, which is called simply ER which stands for Evoked Response and represents the averaged phase across trials. The correlation between the *ER* and the *Wave ER* across the nodes results into other measure called *Wave fit*. As we can notice some of the measures are averaged across trials or correlations across nodes which means that some of the measures don't have all six dimensions as domains.

4.1.1 Data Sets

The domain experts store the data in 3 dimensional hyper-cubes with dimensions trials, samples(time) and nodes. Such files the with 3 dimensional hyper-cube structure are stored for each frequency value, each subject and each experimental condition. The data provided is just for one particular subject and one experimental condition and for 6 frequency levels between 4 and 8 *Hz*. The values of the 3D hyper-cubes I got are the phase measures(PH)

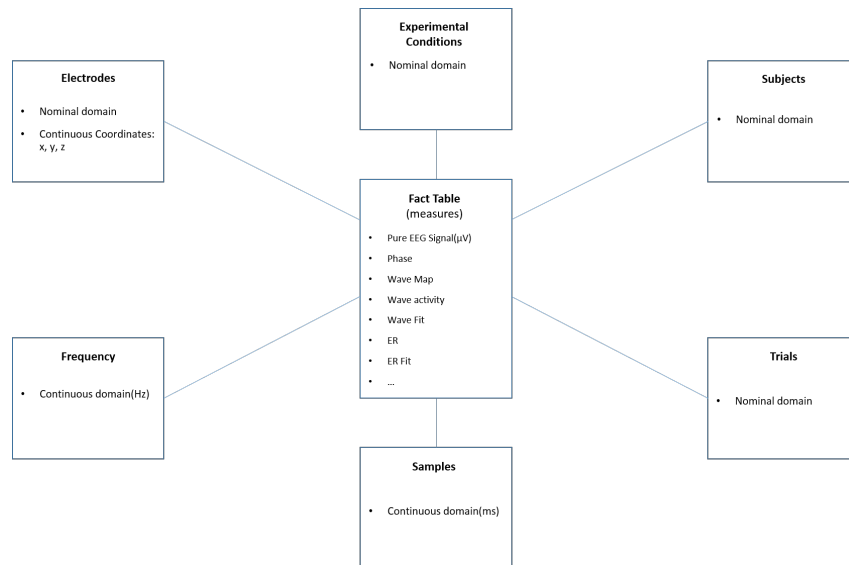


Figure 4.1: Star Scheme representing the dimensions of the data

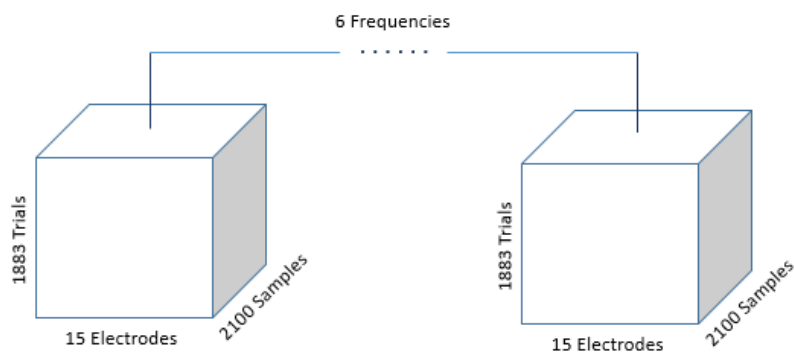


Figure 4.2: The structure of the data given by the domain experts for one subject and without Experiment Conditions

and logarithmic power(LP) of the potential which is the amplitude of the signal. This measures have certain number of units in each hyper-cube dimension. For instance the number of units in the trials dimension are $n_L = 1883$, for the time dimension we have $n_T = 2100$ units for each 2ms time interval and for $n_N = 15$ electrodes. In the picture below we can see the structure of this 3D hyper-cubes.

4.1.2 Data Exploration

In order to explore the data we look at the distributions of the measures for different hyper-cube dimension location, at the stacked bar plots of the measures frequencies along the dimensions, at line charts and heat maps in order to find out which is the best way of summarizing the distribution of measures along the different hyper-cube dimensions. We need

them in order to help the user recognizing patterns of the TW along the dimensions.

First we will have a look at the heat maps similar to these which are used by the domain experts. In Figure 4.3 the heat maps of the phase(top) and the LP(bottom) measures are displayed. This heat maps are slices from the 3D hyper cubes for the first trial and the frequency $4Hz$. At first glance we notice that the bottom values are related to the onset, whereas the peak is slightly before the onset. On the y-axis we map the values across the electrodes dimension. As we can see that this representation gives us good comparison possibility for the case of two hyper-cube dimensions and two measures. when there are more dimensions and more measures to be compared we have to look for other approaches. Furthermore when we want to compare this two measures simultaneously with other measure in narrower time span as the wave fit in Figure 4.4, this way is not sufficient anymore.

When we look at the different trials as units which are coming from the same population we would expect that the measure distribution from trial to trial to be similar. This generalization can be made also for the different subjects, since each person can be seen as a sample of the whole human kind. On the Figures 4.5 and 4.6 present the distributions for the trial 1 and 2 respectively where on each graph also the distributions along all sites(electrodes) are displayed. When we compare the both graphs it is feasible that they are nearly identical which means that the samples from the trials are coming from the same population. This means that, if we average across trials, we will preserve the general trend for all trials. However, that is only the case if we look only from that point of view. When we look at the electrodes dimension for example, we see that the LP and PH distributions among the sites are considerably different from each other, so averaging across nodes won't be good idea, because huge information will be lost.

We want to incorporate in the final visualization some representation of the measures distribution along the different dimensions. For that purpose the stack bar charts are chosen for the designing of the visualization, because they preserve information about the proportions of various value ranges within the dimensions. Now we will verify, whether they are a good choice by generating the stack bar charts for the LP and phase measures along the electrodes dimension and also along the trials dimension. The charts can be seen in Figures 4.7 4.8.

When we look at their stack diagrams for the LP measure, we are not able to distinguish between the electrodes, although we already know that there should be a difference in the distributions of the measures. This suggest one big weakness of the stack bar charts namely that it doesn't show any difference in the proportions, because the differences in the quantities building the proportions are very small percent from the whole. Thus in our case this kind of representation turns out not to be suitable especially for the phase measures, which are periodical measures and they exist in nearly equal quantities along all dimensions. Other opportunities to summarize the distribution of the measure along the dimensions unites are showing the distribution parameters such as standard deviation, median, mean and mode. The discrepancy among the latter three is indicator of the skewness of the distribution([9]) and the former one indicates the variability of the distribution. The first line chart in Figure 4.9 below shows the average of the LP measure for each unit in the trial dimension. Since the main interactive feature is to select specific regions from the Parallel Coordinates representation, the main purpose of this average line charts is to compare the aspects of the

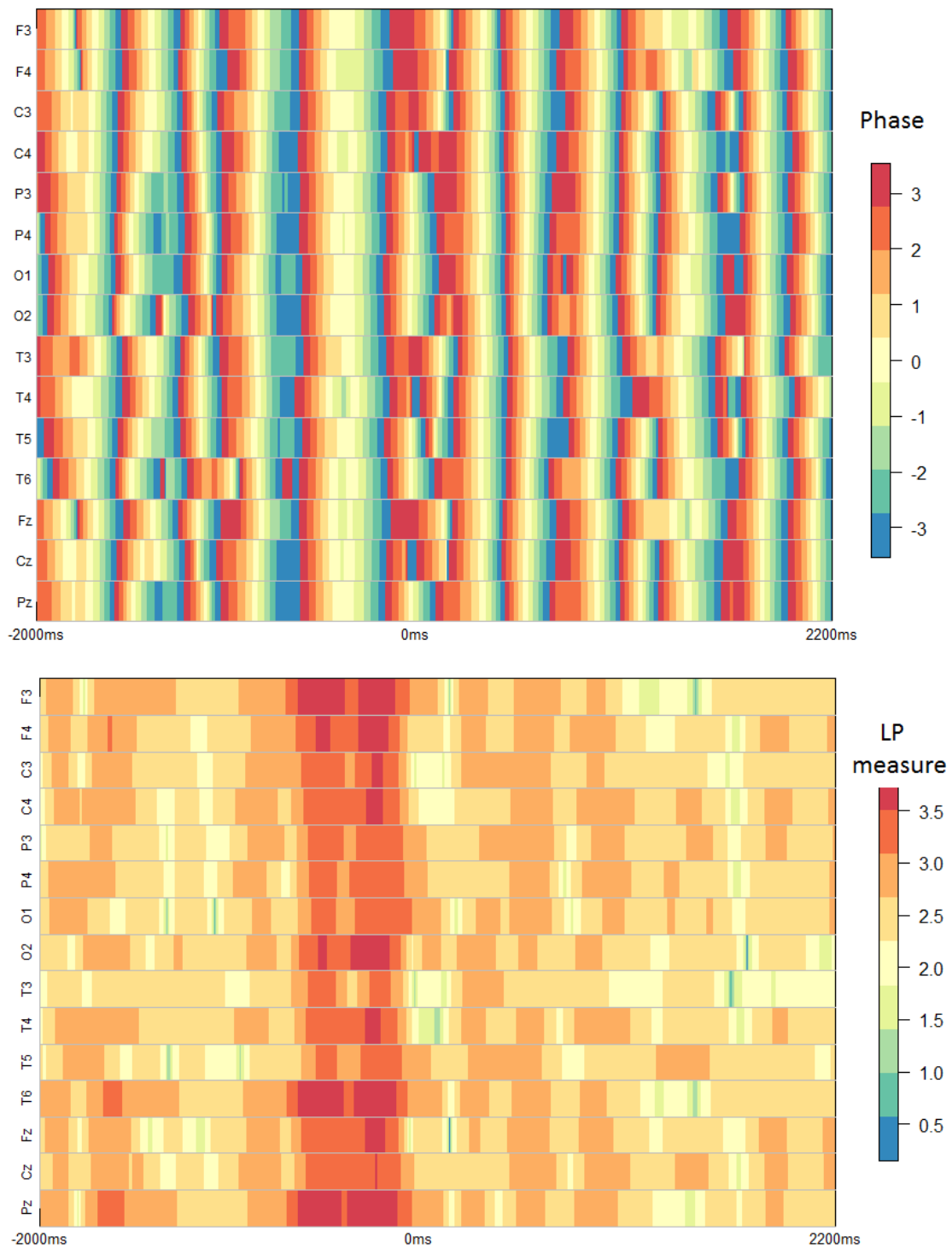


Figure 4.3: Heat maps of the LP(bottom) and phase(top) measure for the first trial at $4Hz$ Frequency, time locked around the onset in the time span from -2000ms to 2200ms.

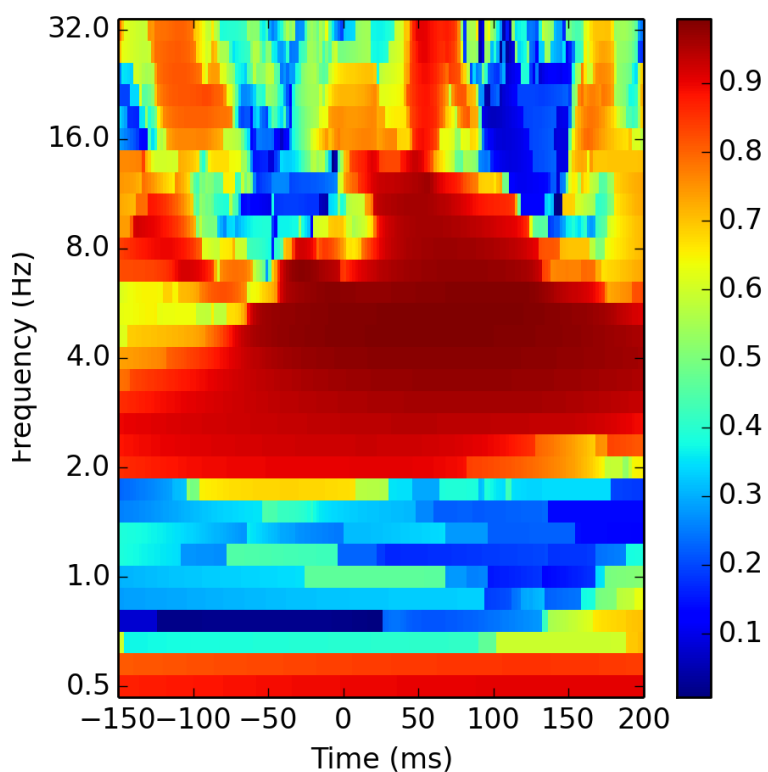


Figure 4.4: Wave fit across the the frequency and time domain

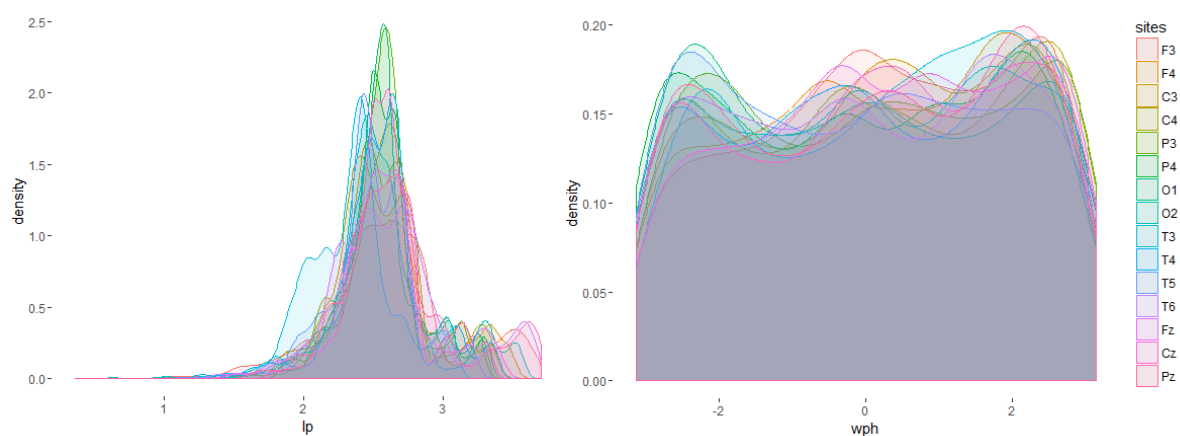


Figure 4.5: The distribution of the LP and PH measures for each of the 15 electrodes at Trial 1

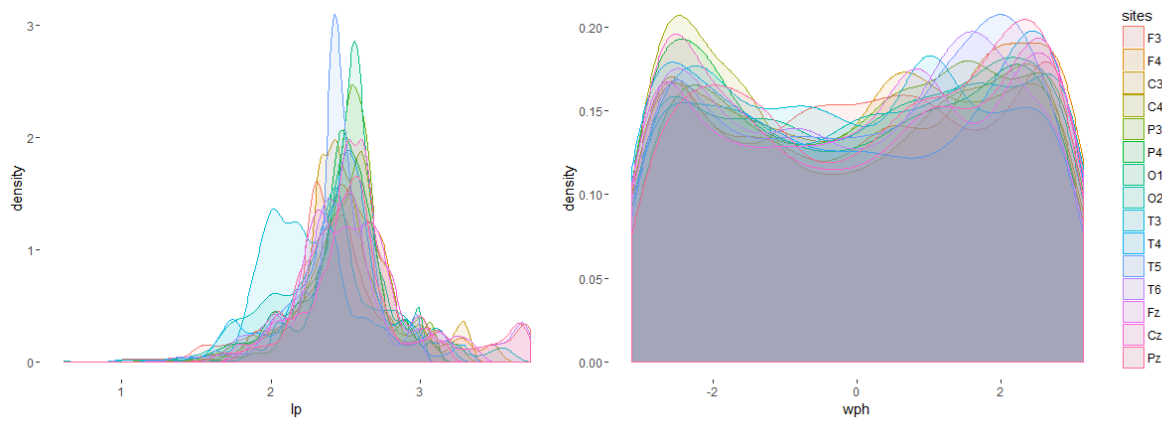


Figure 4.6: The distribution of the LP and PH measures for each of the 15 electrodes at Trial 2

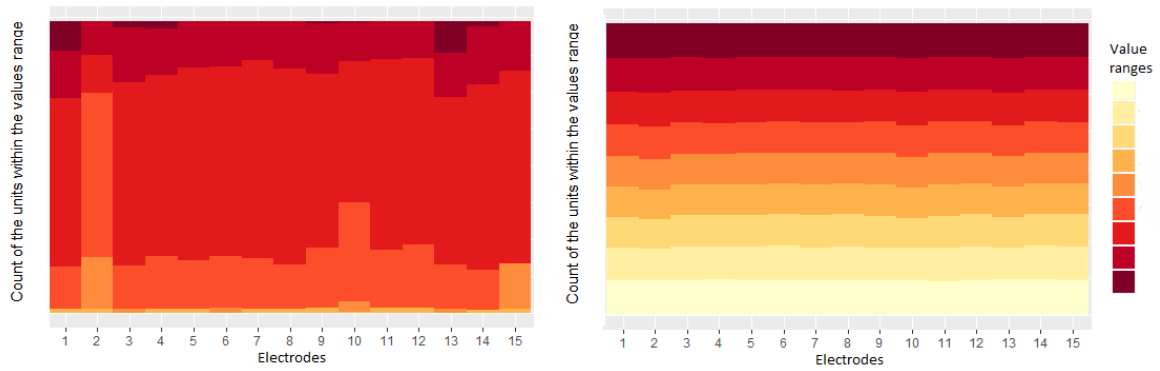


Figure 4.7: Stack bar charts along the electrodes dimension

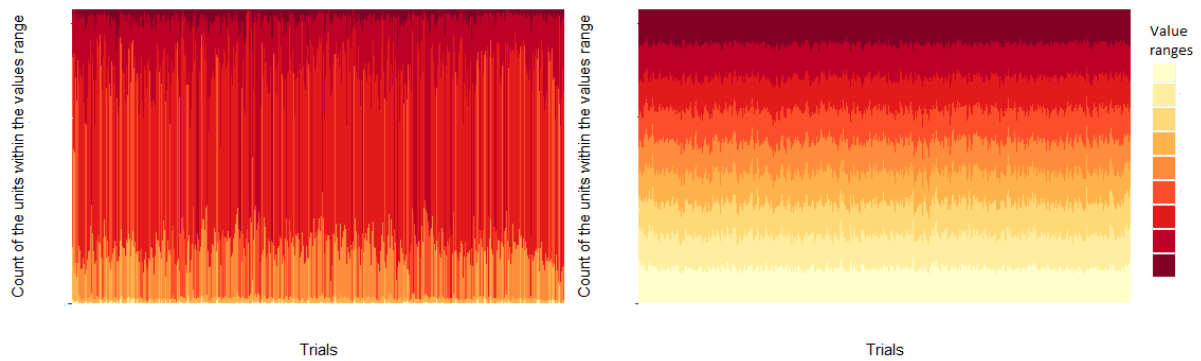


Figure 4.8: Stack bar charts along the trials dimension

measures distribution along the dimension between the selected and discarded parallel coordinates. Thus the line chart below shows the average values along the trials for the selected and discarded units. The selected units are these among all which are contained only in the electrode 10. We can see that although the line for the selected units is in general above the one for the not selected ones, they also overlap a lot which makes this chart not so informative.

The next thing we try is instead of representing the actual averages, to display the differences between the both averages. This is done in the second plot of Figure 4.9. The third plot showing again the differences, but this time as a bar chart where the blue color represents positive difference and the red color represents negative differences. As we can see now even in the trials we can discover some patterns and namely how the signal diverges along the trials between electrode 10 and the grand average of the electrodes. Here we see in particular that the LP averages across trials are bigger in the case of electrode 10 than in the case of the grand average over electrodes.

4.1.3 Data Pre-processing

From the data analysis we come to the conclusion that the statistical measure average of the signal will be used in our visualization as a suitable method for summarizing the distributions of the signal along the dimensions of the hyper-cube. Therefore as data pre-processing techniques before rendering the visualization we will calculate the mean of the signal's measures for each single point along the shown dimensions. We will also use histogram calculations of the signal's measures.

4.2 Design Process

Here we will conduct data analysis and after we got better inside of the structure of the data we will jump into the designing process. Note that in reality I first did the Designing part and then I started to analyze the data which turned out to be mistake since the design concept had to be corrected a couple of times.

4.2.1 Five Design Sheets Technique

Initial Designs

In our case we do slight modification by skipping the brainstorming step, since we've already generated and filtered the design ideas in the *Statement of Goals* artifact.

In the first step we combine the **how** abstractions into complete visualization concepts. We use three initial design sheets, so that each sheet represents some unique design visual-

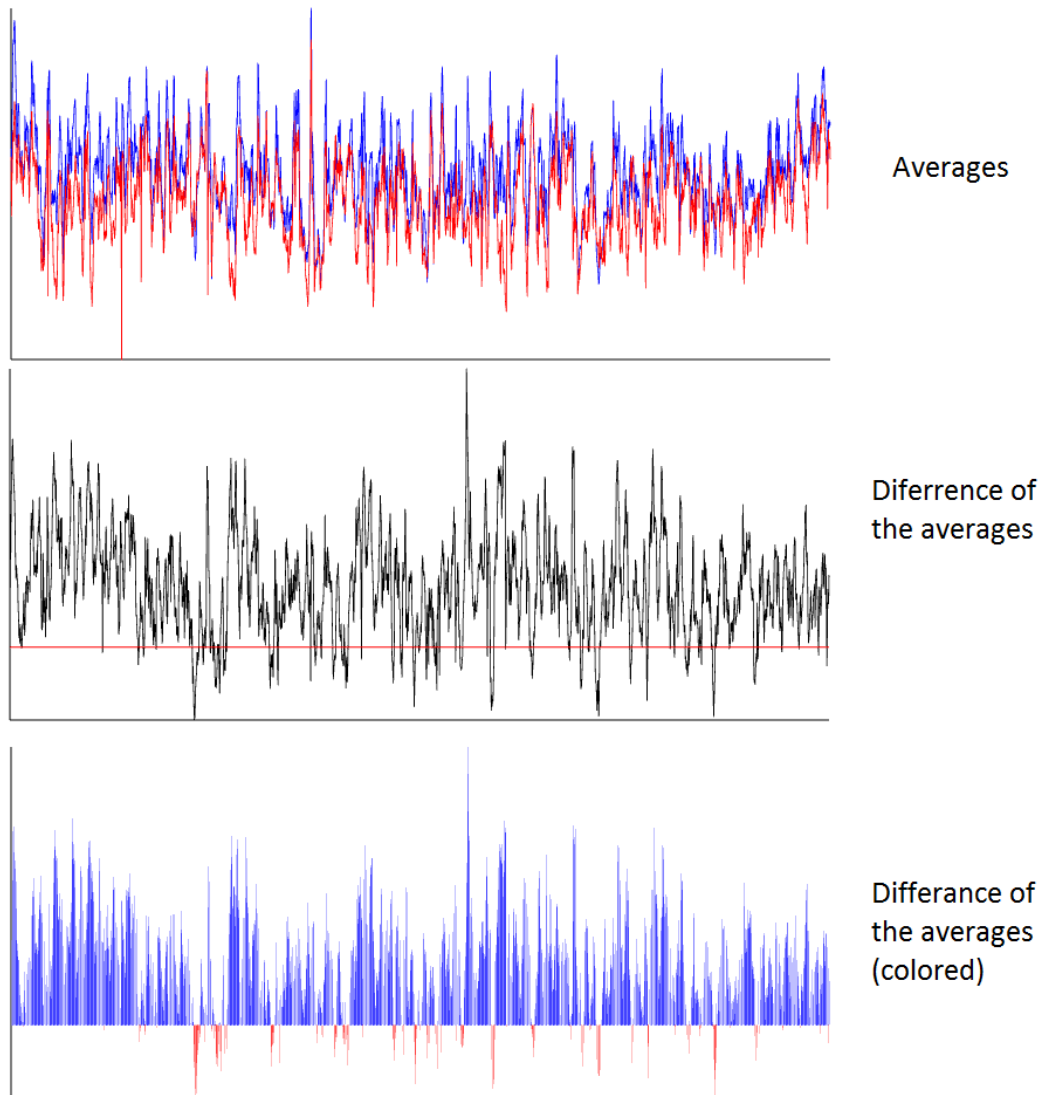


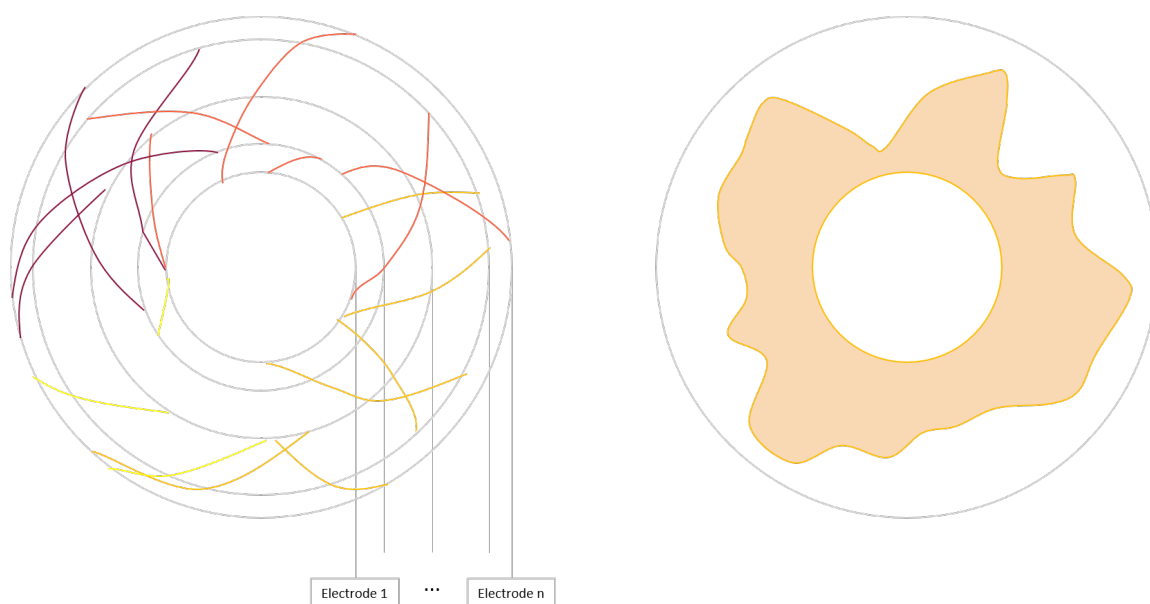
Figure 4.9: Different representation of LP averages across the trials dimension for the cases of the whole hyper-cube and only a slice of the hyper-cube. The First graph is the pure line charts of the two averages. The second one is line chart of the difference between the two averages with a red horizontal line representing the 0 value. And the third plot the difference of the averages in a bar chart.

ization. Each sheet should have the structure represented in 7.3 in the Appendix. Here we design the whole visualization on a sheet, which is shown afterwards to the stakeholders for discussion. The design sheets can be more than three, but the underlying concept is not to occupy too much from the domain experts time by discussing too much design proposals.

In the **Focus** and **Interactivity** section of each initial design is explained in detail which goal is achieved with the particular design and how in particular this goal is achieved.

Artifact: Initial Design 1

Layout:



Focus: The circles on the left graph represent the electrodes and the distance between the circumferences is the relative distance between the electrodes in one of the three space dimensions. Since one of the most reliable part of the wave signal is the phase, because it is at most robust to the volume conduction phenomena, here the position on the circumferences corresponds to the phase angle. The colored lines show the propagation of the TWs in the space and phase or time domains. The colors of the lines can represent the values of some ordinal dimension such as frequency bands or samples. The colored areas on the right graph are the counts of the TWs along the phase or time domain. The stacked bar charts beneath the circled graphs show what part from all units are actually associated with a TW across the dimensions or other measure domains. This representation can achieve partially **Goal 1**, so that the user can see in the stacked bars what is the presence of the TWs along measures and dimension i.e. which parts of the data explain the TW model. The visualization can achieves also partially **Goal 2**, whereas it can be seen how the TW propagate in one chosen spacial dimension and in the phase or the sample(time) domain

Interactivity: The user is able select traveling waves corresponding to some particular

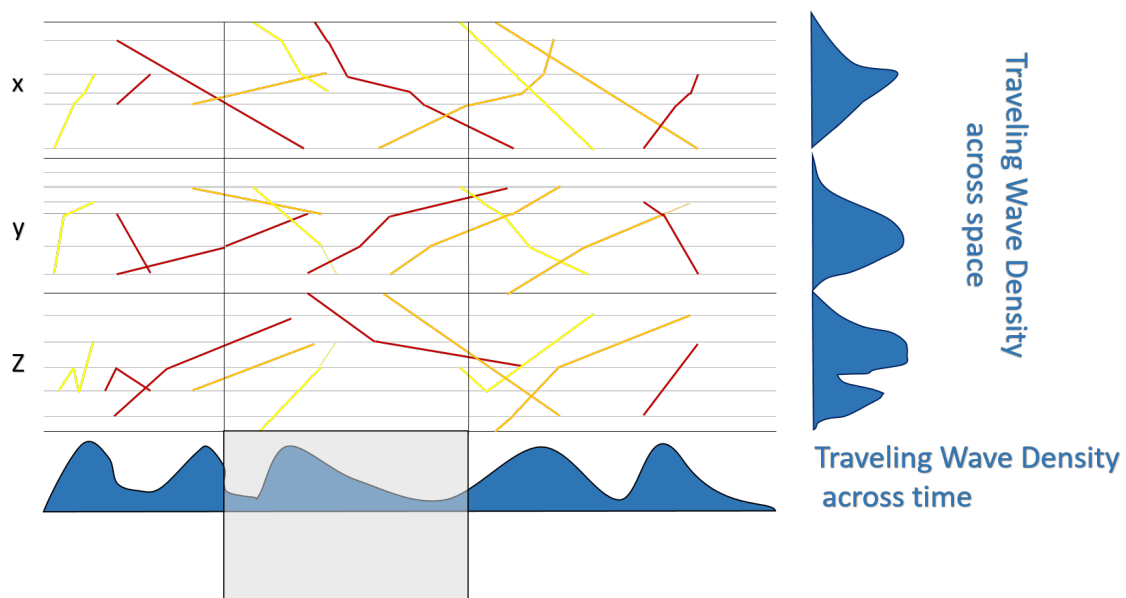
color, which is associated with certain value range of an ordinal dimension such as frequency or time. The densities of the TWs with a particular color is represented on the right graph.

Pros: It has potential use of discovering of pattern in the phase-time-frequency domain of the Traveling Waves and then relating this patterns to some other dimensions with stacked bar charts. It also represents the dynamic nature of the TW but only in one spatial dimension.

Cons: Here we visualize directly the traveling waves which are detected beforehand by looking at certain trial, frequency and time point, we have high correlation between the fitted TW model and the pure signal. Thus we don't represent the dynamics of the pure data, but only of that part of the data which agrees with their theoretical models. This discards to big extent the possibility for exploration. The propagation of the TWs is tilted and the density representation is less accurate because of the circular forms. Also the initial format of the data as the domain experts store it doesn't contain exact indicator which units in the hyper-cube are TWs. Thus we have to see at the *Wave Activity* measure and to indicate the those units as TW which are above certain threshold. This can be very computationally intensive.

Artifact: Initial Design 2

Layout:



Focus: The blue lines are representing the waves propagating in the three dimensions x , y and z and across the time which is increasing from left to right on the horizontal axis (the black lines). The color of the TW can represent the ordinal dimension Frequency. The gray horizontal lines are indicating the electrodes and the distance be-

tween each gray line indicates the relative distance between the electrodes in the particular dimension. The linear propagation behavior of the TW suggests, that the traveling waves should be represented by the straight blue lines. With this feature we achieve the **Goal 1**. Since they could be also other dynamic behavior in the brain signal this visualization could represent also other signal behavior models such as Spiral Waves. The curved blue, horizontal line shows the distribution of the TWs along the time. The gray square selects the time span we are interested in. In the stacked bar charts below again the density of the TWs are shown for each unit of the dimensions or some measure. This density represents purely what is the number of the occurring traveling wave along certain domain or measure. As in *Initial Design 1* with this stacked bar charts **Goal 2** is achieved.

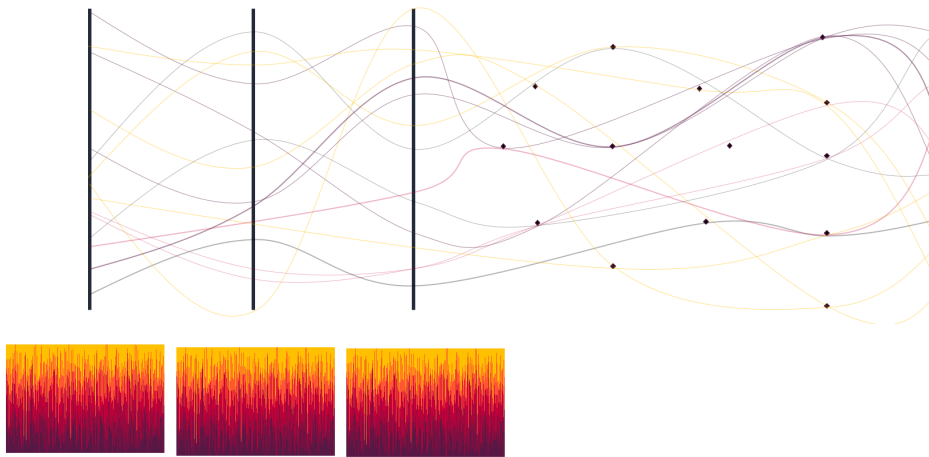
Interactivity: The user has the possibility to select the time span with the show gray square window. The window's width is adjustable by clicking on the left or right edge of the window's black frame and dragging the edge to forwards or backwards along the time. By adjusting the selected time span the histogram of the spiral and Traveling wave is updated, so that only this TWs and SWs are taken into account which are present in the selected time span. the stacked chart is updated in similar manner, whereas only this values of the particular measure are represented along the hyper-space dimension which are present in the selected time span. The user is also able to chose the hyper-dimensions represented on the horizontal axis for the Histogram and stacked bar chart.

Pros: By using this visualization can see how the TWs and SWs are propagating in time and space. we can see the origin of the TWs and SWs and we can compare the distribution of the TWs to the distribution of the distribution of certain measure. This serves for accomplishing the following goals from the statement of goals:

Cons: This design has the same drawback as the Initial Design 1 and namely it doesn't have big exploration power because of lack of representation of more generic and non-modeled data. The initial format of the data as the domain experts store it doesn't contain exact indicator which units in the hyper-cube are TWs. Thus we have to see at the *Wave Activity* measure and to indicate the those units as TW which are above certain threshold. This can be very computationally intensive.

Artifact: Initial Design 3

Layout:



Focus: In that initial design we use parallel coordinates in order to map a particular brain process signals onto the multiple dimensions and measures which domains are represented by the vertical lines. The curves are particular entity taking certain place in the hyper-dimensions and having certain measure value. The color of the curves comes from certain value interval of selected by the user measure. In the stacked bar charts below the corresponding density of the value intervals of the selected measure are show along the domain of the according variable. With the stacked bar charts in contrast to the the previous two designs we don't represent directly the TW density, but the density of some measure selected by the user. The points in between the parallel coordinates on the right are showing the relative position of the electrodes on the $x-y$ plane and along the z dimension respectively, so that the curve with the brain activity information recorded by some electrode passes trough the point with the position of that particular electrode. This is some sort of combining of parallel coordinates with scatter plot. The combination between scatter plot and parallel coordinates is explained in detail in the article [12]

Interactivity: The user can drag and select the range on the horizontal parallel lines he or she is interested in, so that the curves falling in the not selected range are grayed out. This interactive feature makes provides big exploration possibility. After the selection the stacked bar charts below update, so that the corresponding not selected units on the bar chart are also represented in gray. This initial design is suitable for goals 1 and 3. Regarding **Goal 1** the user can drag and select only these *Wave activity* values which are big enough for observing TWs. Then the user can color these *Wave activity* values and compare their behavior along the parallel coordinates to the behavior of other measures. In the same time the user can track what is the presence of the *Wave activity* along the dimensions vie the stacked bar charts. This drag-and-select filtering can be applied for each selected measure, with which we achieve also the **Goal 3**.

Pros: It highlights not only the brain information with high fit to the model, but more generic information, which is good prerequisite for exploration. However it also offers the possibility of verifying and exploring the properties of the TW model along the dif-

ferent measures and hyper-dimensions. It gives the opportunity to compare the behavior of the TW model fit and the averaged signal across trials fit. It also shows the propagation of the of the signals along the space and it locates the signals in the time and frequency domain. It provides flexible measure to measure, measure to hyper-dimension and hyper-dimension to hyper-dimension comparison.

Cons: It can be with long rendering time because of the high amount of curves to be drawn and computationally intensive, because of too much data inter-connections to be calculated. It doesn't show directly the propagation of the TWs.

Final Design

The last step is to describe in much detail the visualization concept that we chose after the discussion of the initial designs. Here we keep the information of the sheet about the Layout, Focus and the Interaction. Instead of the **Pros & Cons** section we put the **Details** section, where we will provide further details important for the implementation of the concept. The final design sheet template can be seen in 7.4 in the Appendix.

After generating the initial designs, it has been discussed with the domain experts, which design from the proposed three to go for. Within the discussions, three main visualization support demanding agendas of the stakeholders emerged on the surface:

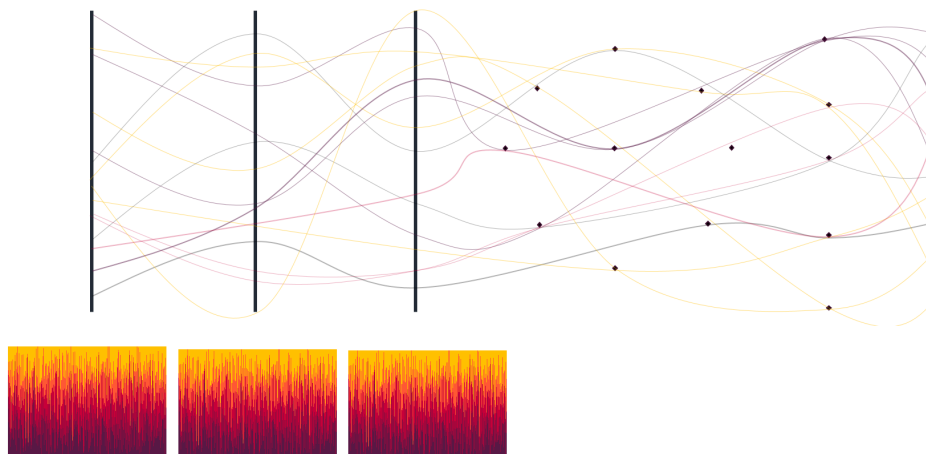
1. **Educational Use** – visualization which makes the lay people better understand the domain expert’s field and in particular the behavior of the Traveling Waves
 - Pluses
 - advantages by incorporating in non-academic environment
 - popularizing the research
 - ease the implementation in the research in non-academical environment
 - Minuses
 - no impact on the present activities of the domain experts
2. **Domain Expert Use** – supports the domain experts in their research by automating their pipeline via interactive visualization interface which connects some golden standard visualizations.
 - Pluses
 - frequent and continuous use
 - time saving throughout automating
 - Minuses
 - none
3. **Scientific Audience Use**– appealing Visualization used for the verification of domain experts hypothesis and revealing the potential of the research. The visualization should be stimulating for the scientist to discover some application or opportunities further developing of the theory.
 - Pluses
 - popularizing the research
 - helps in generating of hypothesis
 - Minuses
 - bigger uncertainty about the impact

Eventually the *Scientific Audience* agenda has been chosen as a priority, because the domain experts haven’t worked much on achieving that agenda and they think that working on visualization design in that direction could be potentially very useful.

Note that so far we already acquired relevant information about this *Scientific Audience* agenda in our artifacts and we don’t have to update them. After setting high priority to that Agenda by the domain experts, we can now also prioritize our goals in the **Statement of Goals** artifact. The goals 1 and 3 turn out to be at most aligned with the chosen agenda. Given the fact that these two goals can be achieved by design 3 and that pros and cons also speak in favor of it, we chose *Initial Design 3* as our final pick. The extended design sheet of that design can be seen below.

Artifact: Final Design

Layout:



Focus: In that initial design we use parallel coordinates in order to map a particular brain process signals onto the multiple dimensions and measures which domains are represented by the vertical lines. The curves are particular entity taking certain place in the hyper-dimensions and having certain measure value. The color of the curves comes from certain value interval of selected by the user measure. In the stacked bar charts below the corresponding density of the value intervals of the selected measure are show along the domain of the according variable. With the stacked bar charts in contrast to the the previous two designs we don't represent directly the TW density, but the density of some measure selected by the user. The points in between the parallel coordinates on the right are showing the relative position of the electrodes on the x - y plane and along the z dimension respectively, so that the curve with the brain activity information recorded by some electrode passes trough the point with the position of that particular electrode. This is some sort of combining of parallel coordinates with scatter plot. The combination between scatter plot and parallel coordinates is explained in detail in the article [12]

Interactivity: The user can drag and select the range on the horizontal parallel lines he or she is interested in, so that the curves falling in the not selected range are grayed out. This interactive feature makes provides big exploration possibility. After the selection the stacked bar charts below update, so that the corresponding not selected units on the bar chart are also represented in gray. This initial design is suitable for goals 1 and 3. Regarding **Goal 1** the user can drag and select only these *Wave activity* values which are big enough for observing TWs. Then the user can color these *Wave activity* values and compare their behavior along the parallel coordinates to the behavior of other measures. In the same time the user can track what is the presence of the *Wave activity* along the dimensions vie the stacked bar charts. This drag-and-select filtering can be applied for each selected measure, with which we achieve also the **Goal 3**.

Algorithms:

- Using polynomial splines by the interpolation of the curves in the parallel coordinates.
- Histogram calculation for each hyper-dimension by using five value bins for the selected measure. Slicing the six dimensional hyper-cube after the user drags and selects the parts of the domain.

Dependencies:

- Using *python* as a programming language.
- Using *Kivy* as a visualization library.

Hardware requirements: Computer with GPU which supports OpenGL 2.0

4.2.2 Tool Development & Implementation

Within the implementation phase various problems have been encountered. This problems are documented in the *Practical Considerations* artifact. The first one is connected to the arrangement of the parallel coordinates.

Artifact: Practical Consideration I

Context: Designing of the visualization.

Observation: Some hyper-cube dimensions are more logical to be next to each other, than others and also the measures and the hyper-dimensions should be arranged so that there time is used as a transition.

How might we position the hyper-dimensions and the measures in the parallel coordinates, **so that** we can most sufficiently compare them and follow the the patterns.

Found solution: The final skeleton of the parallel coordinates is shown in Figure 4.10 We have the time parallel axis as a transition point from signal measures to the other dimensions since it is present domain for all signal measures. On the left of the time parallel we have two other parallel axis for the hyper-cube dimensions. on the right side of time we have parallel axis for the signal measures down and for the TW models measure top. Next the top we have another parallel axis for another model measure. Here the curves between this two parallel lines can be colored id some diverging color palette, so that we compare more efficiently the different measures. Possibilities for comparison here are for instance the *Wave Map* and the *ERP*. On the right of the bottom parallel axis we have another parallel axis so that we can compare two signal measures. On the top right corner the big buttons are represented which are just showing the action we can take by loading the data and selecting role of each data piece.

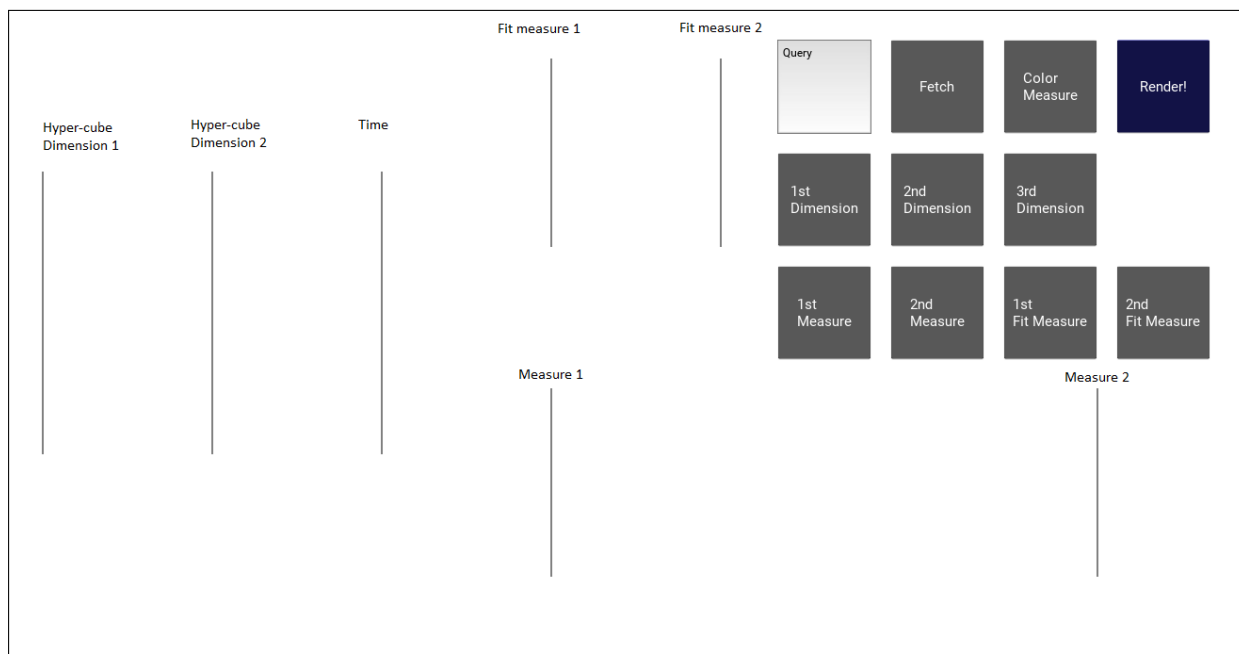


Figure 4.10: Skeleton of the Parallel coordinates, and input data interface

In Figure 4.11 the final arrangement of the parallel coordinates can be seen. This arrangement was suggested by the domain experts during one of the regular meetings with

them. The position of the dimension time is crucial here, because as noted before it is always present on the x-axis of the heat maps the domain experts regularly use. Its integral role makes it key transition element between measures and dimensions. The measures on the top right part of the graph are representing values such as *Wave model* and *Wave Fit* which don't have the dimension electrodes any more, because across this dimension the correlation of the fit with the real signal is built. In the bottom right part we have the measures which have also the electrodes as a hyper-cube dimension. That part offers the opportunity to compare to measures in terms of their values as well as to compare them in terms of the spacial location they are recorded at. Although this part gives more intuitive view about the location of the signal it doesn't provide dynamic representation of the data what we have in the first two initial designs.

Artifact: Practical Consideration II

Context: Designing of the visualization.

Observation: Due to the huge scale of the data we have too many curves to visualize as parallel coordinates. Thus by visualizing the curves for all electrodes and samples ($n_S * n_N = 31500$) we are hitting the limits of that visualization.

How might we modify the parallel coordinate visualization **so that** we make the patterns of the big amount of data more visible.

Found solution: One solution of the By adjusting the transparency of the curves we will make the overlapped curves also visible. The transparency option itself is not enough so we also adjusted the value of the transparency parameter according to the density function of the chosen measure to be colored. That means that the colored measures with values coming from more dense region of the histogram get bigger transparency value, so that they are less visible. Since they are coming from a dense region, by overlaying of the many curves in that region patterns from overlapped curves are build. This is the key to success of the parallel coordinate approach here.

Artifact: Practical Consideration III

Context: Designing of the visualization.

Observation: In order to track the curves of the parallel coordinates we need to make proper color choice.

How might we choose the colors **so that** we make the patterns of the big amount of data more recognizable.

Found solution: For the chosen signal measure we chose some diverging palette and for the fit measures we chose black.[7] We also have to use diverging palette for the curves between the both parallels of the fit measures in Figure 4.10.

After the implementation phase of the visualization we will demonstrate its functional-

ity by showing interactive process in successive screenshots. The first screenshot in Figure 4.11 shows the time on the second parallel coordinate parallel which is used for the transition dimension. The second hyper-cube dimension is the Frequency in our case. Note that the user can select which hyper-cube dimension to put on the parallel coordinate axis. Other choices of dimension can be Experiment Conditions, Subject, Trials. Note that the colored curves are representing phase values averaged over the trials. also for the successive graphs these The upper parallel coordinates indicate the values of the *Wave Fit*. As we can see the *Wave Fit* curves are all in black so that we can distinguish them from the curves of the signal measures. The *Wave Fit* we selected the values between 0.98 and 1, because we are interested only in very high correlation between the *TW Model* and the signal average. Because the provided data has the values for the frequencies between $2Hz$ and $32Hz$, the curves are spreading throughout the whole Frequency domain. In contrast to that the signal measures we we have is only for the Frequency $8Hz$. The histogram below the both measures show the measures distributions. The red histograms represent the distribution of the whole set of values in the hyper-cube, whereas the black histogram is only for the case of the selected measures. In our case we select measures only for the time interval from -1000ms before the onset up till 1000ms after the onset. The time span of the whole hyper-cube time dimension is from -2000ms to 2200ms. The horizontal bar chart in the shows the differences between the averages within certain electrode. The black bars indicate that the difference is positive i.e. the average of the selected phase measures over the time interval of -1000 to 1000 is smaller than the average over the whole available time span. Note that the user is able to select the desired time span with the drag-and-select interactive option which is done by hovering with the computer mouse.

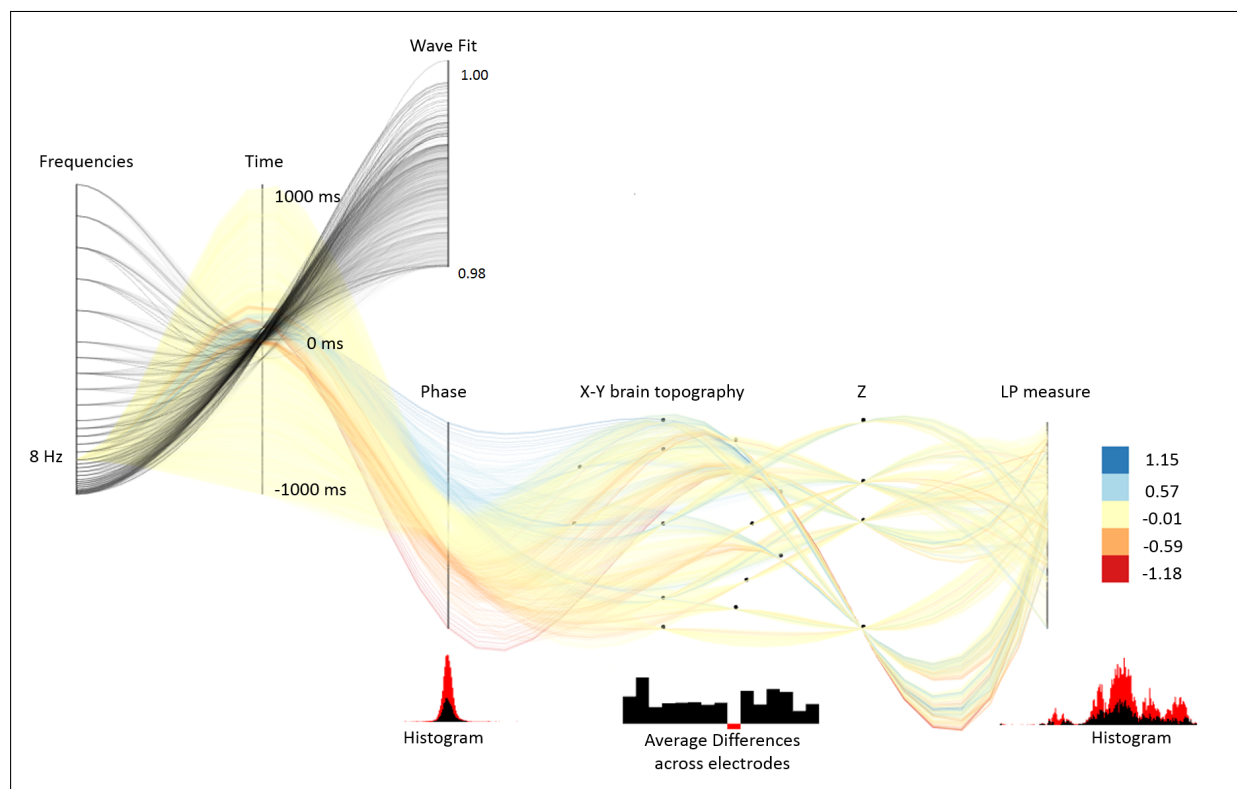


Figure 4.11: Visualization example I with 16500 measure curves

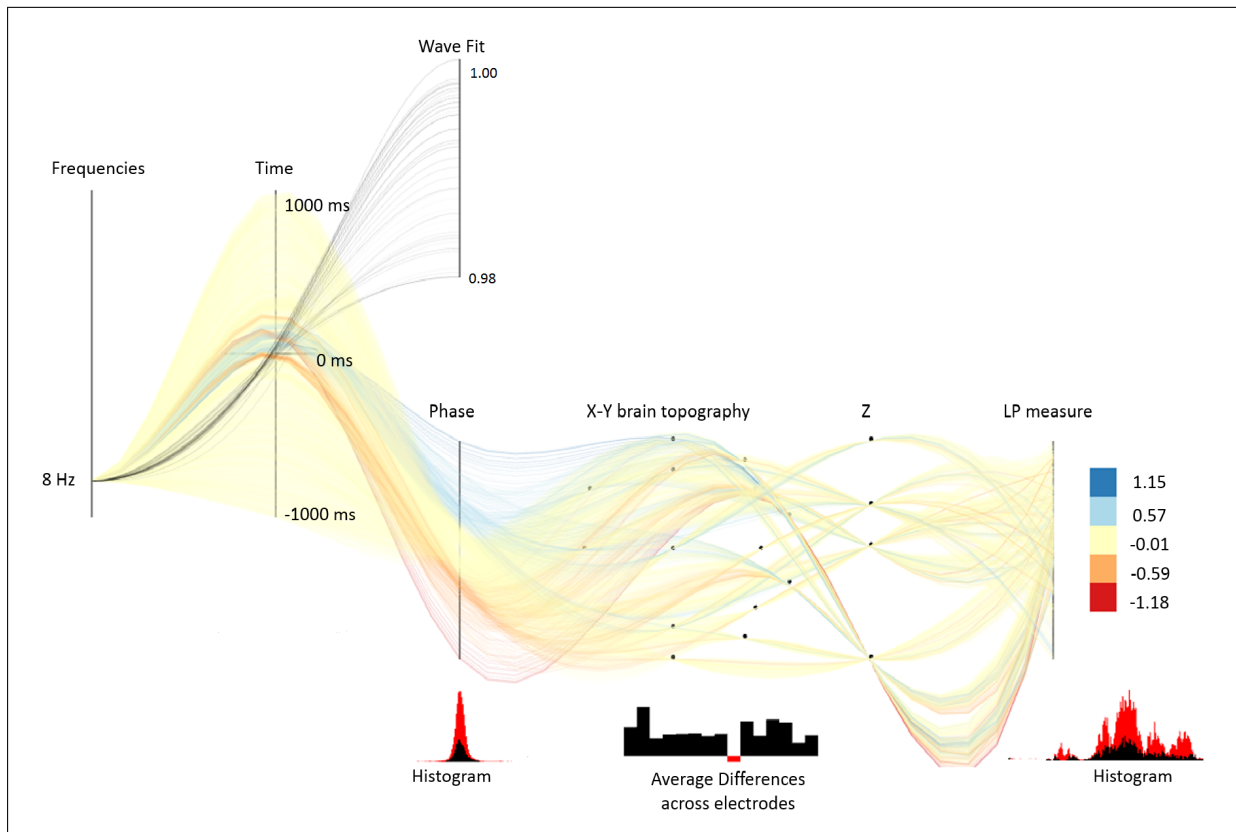


Figure 4.12: Visualization Example II after selecting only Frequency $8Hz$ for the *Wave Fit*

Since we are interested here only in the *Wave Fit* values which go through the Frequency value of $8Hz$ we select just this curves. Note that here the drag-and-select function is not precise enough just to select one value in the parallel coordinate axis of frequency. Thus we have to put the value as an text input.

Next we select only that region of the time parallel coordinate where the signal measures and the TW fit overlapping on the time axis. This is the end of the navigation session. Here the drag-and-select option can be again used by the user. The Figure 4.13 represents the visualization after selecting the time region of overlapping. Here also the histograms and the Average Difference bar charts are updated, since we selected new hyper-cube dimension regions for our signal measures. As we can see the black areas of the histogram are very small, because we've selected very small slice from the hyper-cube. Here we already have filtered a lot of the unnecessarily data and we can see some distinct patterns. For instance on the Z-axis of the brain topography the electrodes which are positioned on the bottom part of the skull have recorded a lot of extreme phase values. Now again with the drag and select feature we select all other nodes which are positioned above the bottom one on the Z-axis. Afterwards the visualization is automatically updated and updated version of it can be seen in Figure 4.14. Now below the time axis another Average Difference bar chart shows up. From that chart we can see that in different time regions the difference between the phase averages of the selected measures and of all measures from the hyper-cube change periodically. They are also connected to the onset time point, whereas after the onset we observe rapid reduce in the average differences. This is due to the periodic

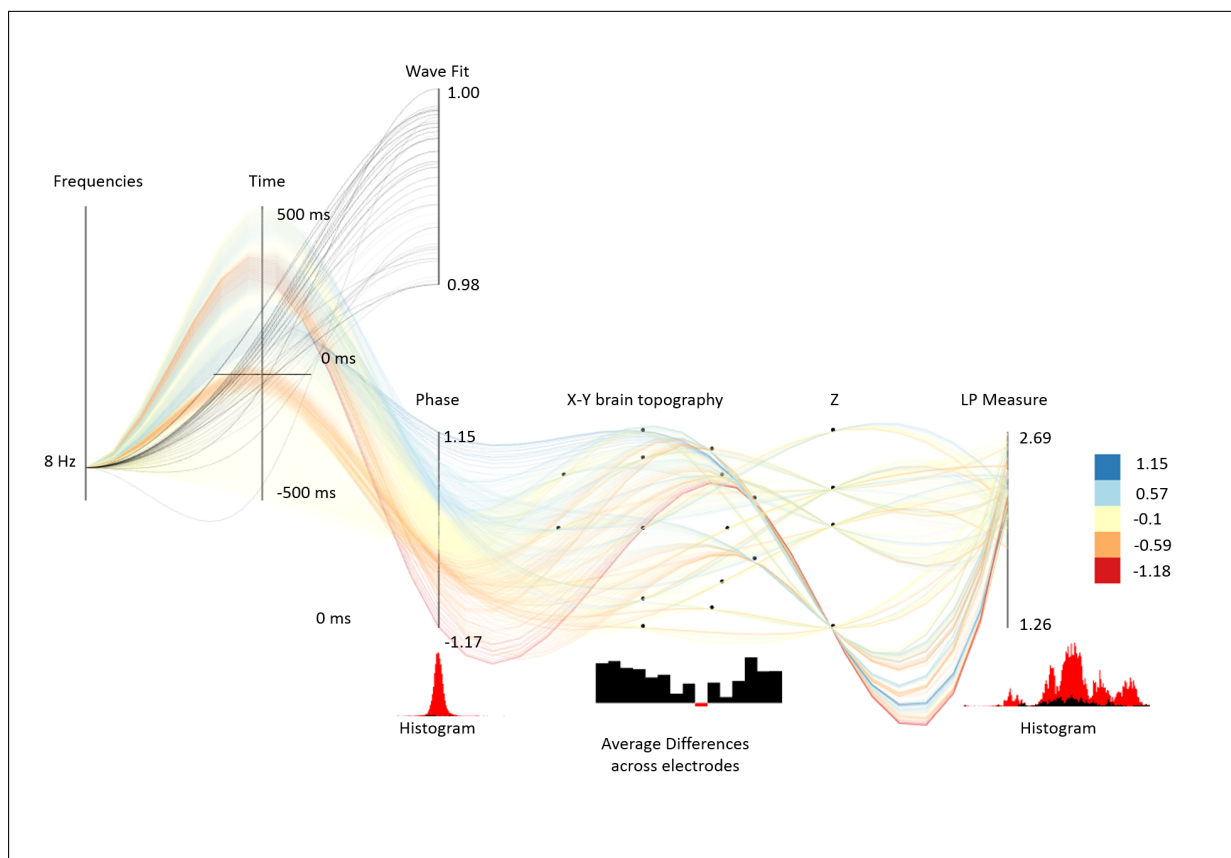


Figure 4.13: visualization Example III after selecting narrower time region

nature of the phase measure. When we select the measure LP to be colored instead of the phase we observe completely different behavior. We can see this behavior of the LP measure in the Figure 4.15. Now we select just the electrodes which are in the bottom of the Z-axis. The change in the visualization we observe in Figure 4.16. This is just one small illustration of the capabilities of the visualization. By looking at different hyper-cube dimensions different the user can explore further pattern.

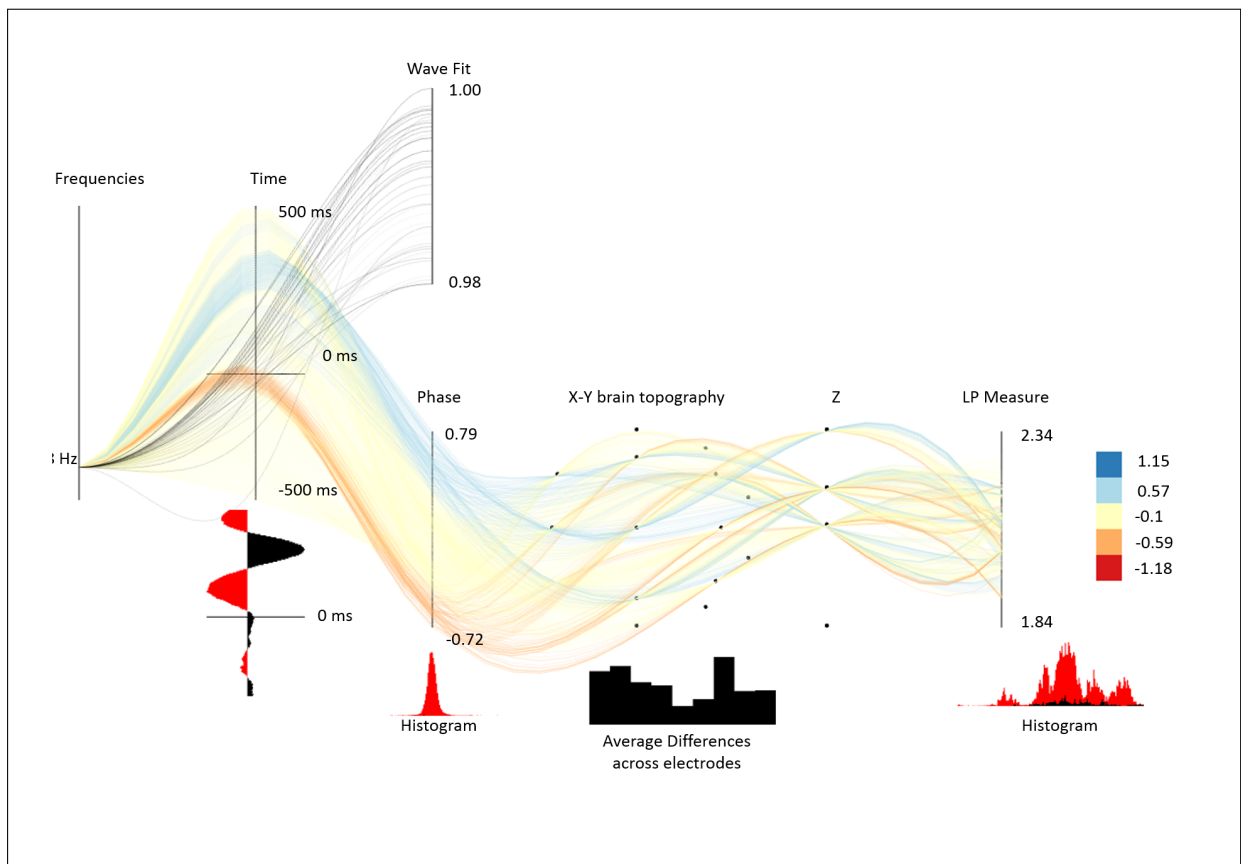


Figure 4.14: Visualization Example IV

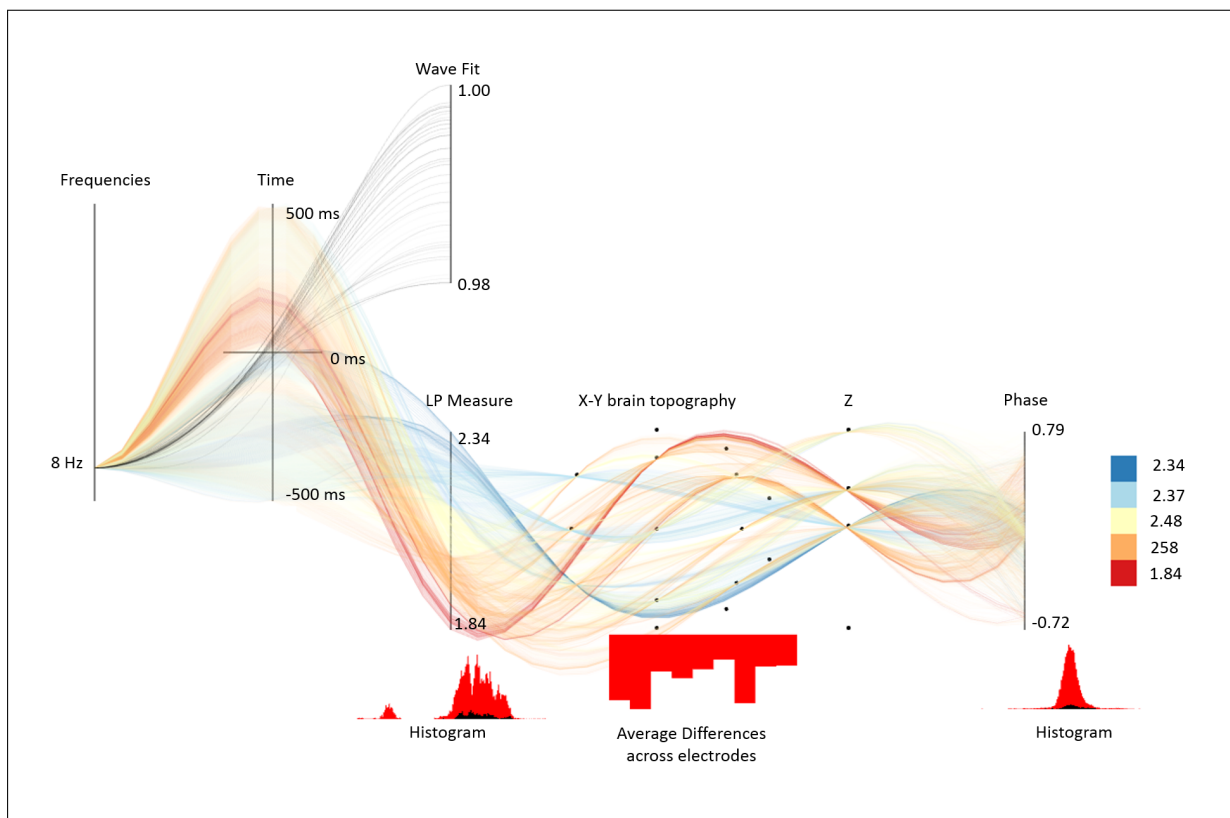


Figure 4.15: Visualization Example VI with the LP as selected measure

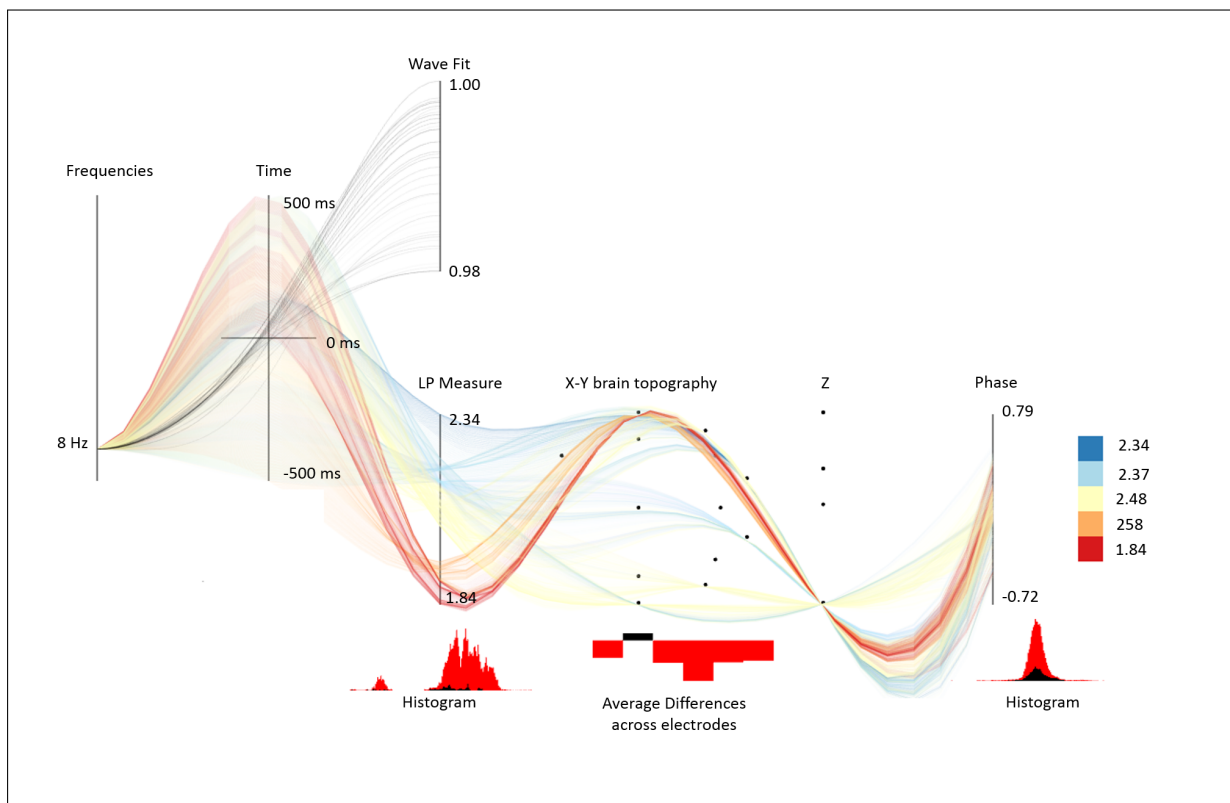


Figure 4.16: Visualization Example VII with the LP as selected measure

Chapter 5

Discussion

The focal points of discussion are the room for improvement of the methodology and of the visualization concept.

5.1 Methodology

In regard to the methodology aside the conventional approaches for understanding the user needs as personal interview and domain literature we conducted *Sci-Fi Ideation Session*. This session really helped discovering some latent needs of the user which weren't feasible from the interviews and the research literature. The mistake which has been made is that after gathering the information about the domain situation I jumped straight into the designing process and I made the *Five Design Sheet Technique* and I've come up with the final design without getting to know the data first. Later when I got the data I did an analysis on it and it turned out that the structure and the scale magnitude of the data demand to do some corrections on the final design. Thus for future reference before jumping into the designing part the designer should be very familiar with the data.

5.2 Visualization Improvements

Regarding the future improvement of the visualization several aspects should be taken into consideration. Firstly how to get better sense of the distribution of current measure for the hyper-cube dimensions. We can include one parallel line which represents the difference between the overall average of the selected measures and the overall average of the whole set of values in the hyper-cube. This line may be as a reference point how the distribution behaves according to the overall measure distribution. Furthermore instead of the mean difference we can represent with the same difference plots the difference of other distribution parameters such as standard deviation, median and the mode. One option is to make

the difference plots horizontal so that we can stack on top of each other two difference plots for two distinct distribution parameter for comparison.

Another improvement to be made is changing of the size proportions of the visualization components, so that for instance, the Average Difference bar charts become bigger and we can put on the x axis the labels of the electrodes. Also the option for sorting the bars according to the position of the bars on one of the axis should be incorporated.

Secondly the visualization can be linked with some more specific visualization. For instance the link might happen via clicking on some certain area of the visualization, so that the other specific visualization pops up in separate window. That other visualization can be functionality complement of our visualization such as one expressing more the dynamic nature of the TWs.

As third more effort can be given in finding out robust solution to our work-around formula of estimating the opacity of the parallel coordinates curves. So far we have to adjust manually the opacity by the change of the distribution of the values. This is very cumbersome for the user, thus some formula robust to the distribution change should be developed, so that the user doesn't have to adjust transparency manually.

The drag-and-select function doesn't work properly in kivy and one option to improve is to switch from *kivy* to *paper.js* which is *javascript* visualization library and has as an advantage that works on layers offers the flexibility to re-render just one visualization layers keeping the other, already drawn layers unchanged. This shift in the languages can result in reduced rendering time. The drawback is, that communication between browser-javascript platform and the python-data-processing platform should be established.

Finally, although we've already tried to improve the algorithm speed, the code is still in experimental condition and it has to be standardized and optimized.

Chapter 6

Summary

We can divide the summary in two parts. One part are the results in regard to the Case Study Visualization and the second kind are the conclusions relating to the Methodology we applied.

Visualization Conclusion

The visualization is useful to the experts firstly as an exploration tool which reveals pattern of the brain signal and the TW model across the dimensions. It can also be used for pairwise comparing of two signal measures. Another very useful application of the visualization is checking which dimensions are influenced by the TW activity. This helps the domain expert to understand what determines the TWs behavior across hyper-cube dimensions. The user is able to check the dependency of the TWs from various dimension by selecting only this region of the parallel coordinates, where there is strong presence of TW activity. After doing so the user can see for each given dimension position what is the difference of the signal averages between the selected TW related signals and the averages of all signals. Other advantage of the visualization is that it is universal and can be used for various experiment conditions which can be represented as parallel coordinates. Furthermore the visualization can be used as summary representation of the complex neuroscience data, so that the lay people can also get intuition about the experiment data.

The disadvantage of the visualization is that it cannot reveal the dynamic nature of the TW. Despite the use of the transparency option, due to the huge scale of the data the visualization has limitations of the number of parallel coordinate curves which can be drawn. Another draw back is that is also very computationally intensive for a huge amount of data.

The interactive part here plays a big role in achieving the purpose of the visualization. By using the drag-and-select we can choose exactly that part of the hyper-cube dimensions the user is interested in. When the user filters out the unnecessarily information, the remaining part becomes better traceable and also the information in form of the average and the portion of the selected measure across dimensions gets actualized. The user can also

specify according to which measure the parallel coordinate curves can be colored. The data can be fetched by setting a query of series of strings, which specifies the measures and the dimensions of the data hyper-cubes to be imported.

Methodology

As already mentioned in the introduction a big part of the purpose of this master thesis is to establish and test a methodology for designing a visualization according to the needs of the clients, so that we can draw conclusions about useful approaches in designing the visualization product. In our case the clients were the domain experts from the KU Leuven Perceptual Dynamics Laboratory. In the course of my work with them I applied some approaches for coming up with visualization designs. As a part of the master thesis pipeline I called this approaches *tools*. In the context of use of this tools I constantly updated some pieces of information called *artifacts*. In the course of the work the *artifact Statement of Goals* shaped in big extent the final visualization design. This artifact was meant to summarize the information into the **what-why-how** framework proposed by Tamara Munzner in her book [8]. It turned out that the paraphrasing of need of the domain experts in different forms such as *Hypothesis Statement* and *Proto-Persona* is very useful. After doing that exercise it was much easier to translate the information into the *Statement of Goals*. From the tools particularly useful was the *Ideation Session*, which really brought on the surface the particular needs of the stake holders. One mistake in the methodology was not to analyze the data before getting started with the designing of the visualization, because that way I didn't have complete understanding of the data structure I had to visualize.

Chapter 7

Appendix

List 7.1: Possible *Targets* and *Actions* used in constructing the **why** abstractions

| • <i>Targets</i> | • <i>Actions</i> |
|------------------|------------------|
| - Outliers | - Discover |
| - Trends | - Present |
| - Features | - Enjoy |
| - Distribution | - Annotate |
| - Extremes | - Record |
| - Dependency | - Derive |
| - Correlation | |
| - Similarity | |
| - Topology | |
| - Shape | |

¹The graph is taken from the public domain https://en.wikipedia.org/wiki/Event-related_potential

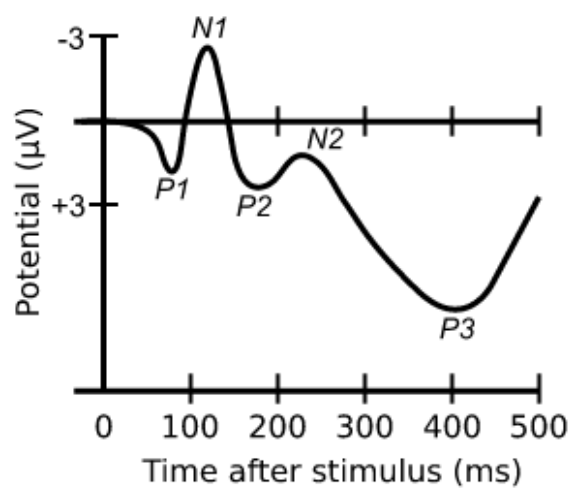


Figure 7.1: Line Chart of the ERP signal.¹

The Five Design Sheet technique

List 7.2: Five Design Sheet original first Step

1. **Generate Ideas:** You can write down the ideas in form of few words association or a quick drawing. We have to write down every relevant idea, which crosses our mind.
2. **Filter the ideas:** See which ideas are overlapping and remove them.
3. **Categorize:** the sketches. Categorize the remaining ideas, so that you combine them easily into visualization concepts
4. **Combine & Refine:** Rearrange the ideas in bigger visualization solutions, which show certain aspect of the data
5. **Question:** Ask questions to verify if the made visualization solutions answer the stakeholders goals.

List 7.3: Five Design Sheet first Step-Initial Designs

1. **Layout:** This is the vision of what the final visualization would look like. Commonly this would appear as a sketched screen-shot of the typical visualization application.
2. **Focus:** Note the functionality of each part visualization.
3. **Interaction:** Note how the user can interact with the visualization
4. **Pros & Cons:** Note the advantages and disadvantages of the visualization

List 7.4: Five Design Sheet second Step-Final Design

1. **Layout:** This is the vision of what the final visualization would look like. Commonly this would appear as a sketched screen shot of the typical visualization application.
2. **Focus:** Note the functionality of each part visualization.
3. **Interaction:** Note how the user can interact with the visualization
4. **Details:** Note additional details of the visualization concept such as:
 - **Algorithms:** Mention crucial math concepts and algorithms used in the visualization
 - **Dependencies:** Mention software libraries for implementing the visualization
 - **Hardware requirements:** Technical requirements, which should be fulfilled

Question 1: I want to find out more about your study, the experiments you are doing and the responses you are measuring.

Answer 1:

Marcello: LP, Phase, ER average, Wave activity, ER average fit, Wave activity fit. Testing. Comparing the TW model fits at single trials.

David: I'm doing using the experiments of other scientists of the lab and using them in direction discovering TW and Spiral wave patterns. Mostly we use the phase as a measure for fitting model and drawing conclusions about, because it is more robust to noise caused by the volume conduction.

Question Agenda 1: To get an input information for updating the artifact.

Question 2: What are the hypothesis you are testing and how do you test them?

Answers 2:

David: We are building models and then test to what extent the models are correlating with the reality at single trial level.

Marcello: Hypothesis that they are some traveling waves which propagate spatially. We are looking of similarities after modeling across trials. Conventional way is to focus on the region of interest in the brain by looking only at the averages across trials at certain electrode. The domain experts however, are fitting model which takes into account the linear propagation of the signal along sites. The predicted values of the model are used for averaging across trials and for building correlations with the true measures and with the averaged measures.

Question Agenda 2: To get input information for updating the artifact.

Question 3: What struggles do you have in your work?

Answers 3:

David: The dynamical representation of the TW in 3D with the time is too messy, and cannot see too much.

Marcello: Lack of automatic way of looking at the data. Too big dimensionality to visualize.

Question Agenda 3: To get input information for updating the artifact.

Question 4: What do you use visualizations for currently (Generating of hypothesis, Exploration, Analytics, Patterns discovery, etc.)?

Answers 4:

David: Patterns discovery, but main focus is on making the data and especially the dynamics of the data more comprehensive, which will help to draw conclusions and would help to the

lay people to understand the underlying mechanism.

Marcello: Visual verification of our traveling wave model. It would be nice to explore the data more.

Question Agenda 4: To get an idea about the function of the visualization used at the moment.

Question 5: Do you rely in big extent on visualizations in your research? Grade the importance you place for visualization?

- Low
- Medium
- High - *David. Marcello*

Question Agenda 5: To see to which extent there is need of visualization. This question was useful to decide whom I will work with within my master thesis.

Question 6: What is the methodology you are using? How do you pre-process and analyze your data(numerical methods)?

Answer 6:

David: Investigating various Brain activity experiments with visual stimuli and various experimental conditions for the occurrence of TWs and Spiral Waves. Building a model, which explains the behavior of Spiral Waves.

Marcello: ICA, Morlet wavelet. Investigating the traveling wave activity by Looking at heat maps along the frequency, time, sites domains. ICA(Noise reduction). Morlet wavelet for localization of the signal in the relevant frequency span.

Question Agenda 6: To get input information for updating the artifact.

Question 7: Is the following terminology important in your research:

- phase-phase cross-frequency coupling
-none
- amplitude-phase cross-frequency coupling
-none
- amplitude-amplitude cross-frequency coupling
-none
- Evoked response measures
-yes

- Spatial scales of brain activity:
 - spikes of single neuron
-**none**
 - synchronized population activity of several tens of thousands nerve cells(local field potential)
-**none**
 - synchronized population activity of several of millions(subddural electrocorticogram)
-**none**
 - synchronized population activity activity of neurons in multiple cortical areas(EEG, MEG)
-**yes**
- Brain activity signal components
 - frequency
-**yes**
 - phase
-**yes**
 - amplitude
-**yes**
- Stimuli
-**yes, the stimuli are time locked events, which are represented by the trials**
- Brain regions of interest(Brain Topography)
-**none**

Question Agenda 7: To orient myself in the information I encounter about the broad neuroscience Domain.

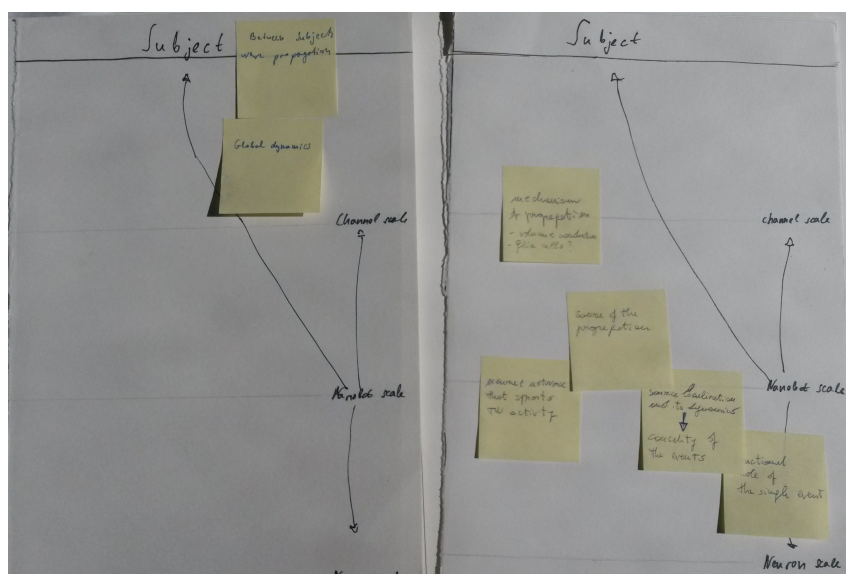


Figure 7.2: The results from the Ideation Session in form of sticky notes filled with content by the domain experts

Bibliography

- [1] David M Alexander, Peter Jurica, Chris Trengove, Andrey R Nikolaev, Sergei Gepshtein, Mikhail Zvyagintsev, Klaus Mathiak, Andreas Schulze-Bonhage, Johanna Ruescher, Tonio Ball, et al. Traveling waves and trial averaging: the nature of single-trial and averaged brain responses in large-scale cortical signals. *Neuroimage*, 73:95–112, 2013.
- [2] David M Alexander, Chris Trengove, and Cees van Leeuwen. Donders is dead: cortical traveling waves and the limits of mental chronometry in cognitive neuroscience. *Cognitive processing*, 16(4):365–375, 2015.
- [3] DM Alexander, C Trengove, JJ Wright, PR Boord, and E Gordon. Measurement of phase gradients in the eeg. *Journal of neuroscience methods*, 156(1):111–128, 2006.
- [4] Ryan T Canolty and Robert T Knight. The functional role of cross-frequency coupling. *Trends in cognitive sciences*, 14(11):506–515, 2010.
- [5] Jeff Gothelf and Josh Seiden. *Lean UX: Applying lean principles to improve user experience*. "O'Reilly Media, Inc.", 2013.
- [6] Noah Iliinsky and Julie Steele. *Designing data visualizations: Representing informational Relationships*. "O'Reilly Media, Inc.", 2011.
- [7] Adam Light and Patrick J Bartlein. The end of the rainbow? color schemes for improved data graphics. *Eos*, 85(40):385–391, 2004.
- [8] Tamara Munzner. *Visualization Analysis and Design*. CRC Press, 2014.
- [9] Michikazu Sato. Some remarks on the mean, median, mode and skewness. *Australian Journal of Statistics*, 39(2):219–224, 1997.
- [10] Bruce J Swihart, Brian Caffo, Bryan D James, Matthew Strand, Brian S Schwartz, and Naresh M Punjabi. Lasagna plots: a saucy alternative to spaghetti plots. *Epidemiology (Cambridge, Mass.)*, 21(5):621, 2010.
- [11] Froukje Sleswijk Visser, Pieter Jan Stappers, Remko Van der Lugt, and Elizabeth BN Sanders. Contextmapping: experiences from practice. *CoDesign*, 1(2):119–149, 2005.
- [12] Xiaoru Yuan, Peihong Guo, He Xiao, Hong Zhou, and Huamin Qu. Scattering points in parallel coordinates. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1001–1008, 2009.

AFDELING
Straat nr bus 0000
3000 LEUVEN, BELGIË
tel. + 32 16 00 00 00
fax + 32 16 00 00 00
www.kuleuven.be

