KU LEUVEN

An Information Theoretical Approach to EEG Source-Reconstructed Connectivity

Axel Faes

Thesis submitted for the degree of Master of Science in Artificial Intelligence, option Engineering and Computer Science

Thesis supervisor:
Prof. dr. ir. Marc Van Hulle

Assessor:
Mansoureh Fahimi
Prof. dr. Daniele Marinazzo

Mentor:
Mansoureh Fahimi

Academic year 2017 – 2018

© Copyright KU Leuven

Without written permission of the thesis supervisor and the author it is forbidden to reproduce or adapt in any form or by any means any part of this publication. Requests for obtaining the right to reproduce or utilize parts of this publication should be addressed to the Departement Computerwetenschappen, Celestijnenlaan 200A bus 2402, B-3001 Heverlee, +32-16-327700 or by email info@cs.kuleuven.be.

A written permission of the thesis supervisor is also required to use the methods, products, schematics and programmes described in this work for industrial or commercial use, and for submitting this publication in scientific contests.

Acknowledgements

This thesis has been quite a long journey. I was able to get invested into an interesting field of research and learn a lot of new things. Both about research, and about myself.

I would like to thank everybody who kept me busy and supported me the last year. I would also like to thank the jury for reading the text.

Axel Faes

Contents

Acknowledgements i

Abstract iii

List of Figures iv

1 Introduction 1

2 Source Reconstruction 3

3 Brain Connectivity 5

3.1 Volume Conduction . . . 5

4 Information Theory 7

4.1 Entropy . . . 7

4.2 Joint Entropy . . . 8

4.3 Mutual Information . . . 10

4.4 Multi-Variate Information Theory . . . 11

4.5 Continuous Data . . . 11

4.6 Coding Theory . . . 11

4.7 Cybernetics . . . 11

5 EEG Experiment 13

6 Evaluation 15

6.1 Calculating Bin Sizes . . . 15

6.2 Comparing Common Regions . . . 15

6.3 Per Subject Comparison . . . 15

7 Implementation 21

7.1 Language . . . 21

7.2 Data Conversion . . . 21

8 Related Work 23

9 Future Work 25

9.1 Directed Information . . . 25

9.2 Open Source Connectivity Package . . . 25

9.3 Data Flow Connectivity . . . 25

10 Conclusion 27

Bibliography 29

Abstract

List of Figures

4.1 Venn diagram for information measures. 9

6.1 Entropy comparison between bin sizes 15
6.2 Experiment 16
6.3 Comparison 17
6.4 Entropy comparison between bin sizes 18
6.5 Entropy comparison between bin sizes 19
6.6 Entropy comparison between bin sizes 20

Chapter 1

Introduction

Chapter 2

Source Reconstruction

Chapter 3

Brain Connectivity

3.1 Volume Conduction

Chapter 4

Information Theory

Information theory provides us the tools to study the information processing capabilities of different systems. These systems include computers, artificial intelligence and also the brain. The story of Information Theory begins with Shannon who provided the tools to estimate information [1]. The most fundamental aspect of information theory is the concept of the 'bit'.

As a standard, information theory deals with 'bits'. One bit of information represents a choice between two equally probable options. A perfectly balanced coin toss contains one bit of information. It has 50% chance of landing on head and 50% chance of landing on tails. A bit is a measure of information and a measure of uncertainty. Thus uncertainty and information are tightly intertwined. If we are completely certain about a certain event, there is no information to be gained.

This chapter will cover the most important aspects of information theory. First and formost, this means discussing entropy. Once we have an understanding of entropy, extensions can be discussed. These include joint entropy and conditional entropy. With these tools, we can go into the area of mutual information, which plays an important role in this thesis. Another important aspect is the generalisation into multivariate systems and dealing with continuous (as opposed to discrete) systems. Finally, some explanation is provided of coding theory and cybernetics.

4.1 Entropy

The next step is entropy. There are different kinds of entropy, such as thermodynamic entropy. Here, we are talking about information entropy. Information entropy is the average uncertainty associated with a random variable. In other words, information entropy is the average rate of information from a random or stochastic variable.

In order to calculate the entropy, we require some discrete random variable X, with $x_1 ... x_n$ the different values from X. $P(x)$ is the probability mass function. The entropy $H(X)$ can be calculated as:

$$H(X) = -\sum_{i=1}^{n} P(x_i) log_2(P(x_i)) \tag{4.1}$$

Interesting to note is that $log_2(P(x_i))$ is the information about value x_i. This formula can be extended into the conditional entropy of two events X and Y. The entropy $H(X|Y)$ is the entropy of random variable X given that the outcome of Y is known.

$$H(X|Y) = \sum_{i,j} P(x_i, y_j) log_2(\frac{P(y_j)}{P(x_i, y_j)}) = -\sum_{i,j} P(x_i, y_j) log_2(\frac{P(x_i, y_j)}{P(y_j)}) \tag{4.2}$$

If random variables X and Y are independent of eachother, then $H(X|Y) = H(X)$. If the variables are completely independent, knowing anything about Y, will not change anything we know about X. Similarly, if $H(X|Y) = 0$, then X is completely determined by Y.

The rule of Bayes is also applicable to conditional entropy:

$$H(X|Y) = H(Y|X) + H(Y) - H(X) \tag{4.3}$$

4.2 Joint Entropy

$H(X)$ and $H(X|Y)$ are basic notions of information measures. Figure 4.1 visualises the notion of entropy. This figure also contains two information measures that are not yet described, $I(X, Y)$ and $H(X, Y)$. Respectively, they are the mutual information and the joint entropy.

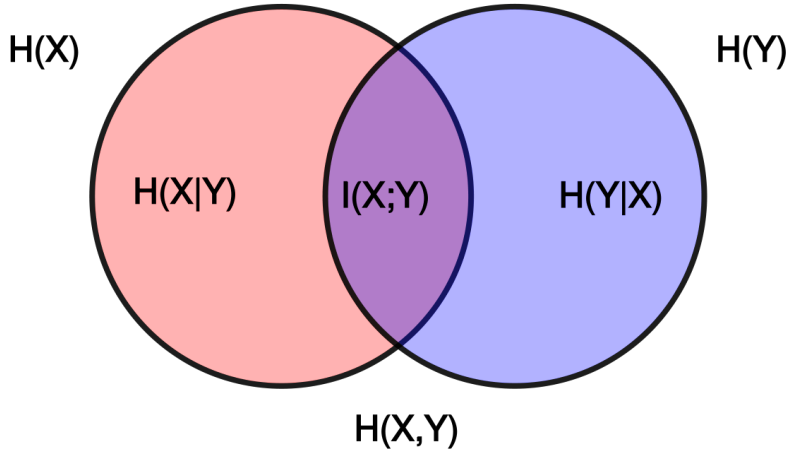
In the figure, $H(X, Y)$ is the complete information content of X and Y. Using the following equation, we can calculate $H(X, Y)$:

$$H(X, Y) = -\sum_{i=1}^{n} \sum_{j=1}^{m} P(x_i, y_j) log_2(P(x_i, y_j)) \tag{4.4}$$

Using figure 4.1, we can see some interesting relations. $H(X, Y)$ is related to $H(X|Y)$ and $H(X)$:

$$H(X, Y) = H(X|Y) + H(Y) \tag{4.5}$$

Figure 4.1: Venn diagram for information measures.



This can be validated:

$$
\begin{aligned}
H(X,Y) &= -\sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i,y_j)log_2(P(x_i,y_j)) \\
&= -\sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i,y_j)log_2(P(y_i)P(x_i|y_j)) \\
&= -\sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i,y_j)log_2(P(y_i)) - \sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i,y_j)log_2(P(x_i|y_j)) \\
&= -\sum_{j=1}^{m} P(y_j)log_2(P(y_i)) - \sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i,y_j)log_2(P(x_i|y_j)) \\
&= H(Y) + H(X|Y)
\end{aligned}
$$

This relation is useful for the actual implementation of information theoretical algorithms. An important note about entropy, conditional entropy and joint entropy is that they are non-negative. It would not make sense for an random variable to have a negative information content. Joint entropy is also always greater or equal to individual entropy and joint entropy is smaller or equal to the sum of individual entropies.

$$H(X) \geq 0 \tag{4.6}$$

$$H(X|Y) \geq 0 \tag{4.7}$$

$$H(X,Y) \geq 0 \tag{4.8}$$

$$H(X,Y) \geq H(X) \tag{4.9}$$

$$H(X,Y) \leq H(X) + H(Y) \tag{4.10}$$

4.3 Mutual Information

Mutual information is the amount of information that is common between two variables. Figure 4.1 shows the mutual information as the intersection between $H(X)$ and $H(Y)$. The figure also shows how mutual information can be computed. The individual entropies are summed and the joint entropy is subtracted. Another equal interpretation of mutual information is that mutual information measures the reduction in uncertainty or information when we observe one of the variables.

$$I(X,Y) = H(X) + H(Y) - H(X,Y) \tag{4.11}$$

$$I(X,Y) = H(X) - H(X|Y) \tag{4.12}$$

Mutual information can also be calculated using the probability mass functions directly. In this case we get:

$$I(X,Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i,y_j) log_2(\frac{P(x_i,y_j)}{P(x_i)P(y_j)}) \tag{4.13}$$

While mutual information does show whether there is a relationship or correlation between two variables, it does not give information about the "shape" of the relationship. Making the analogy with figure 4.1, mutual information shows how much overlap there is, but it does not explain anything else. Mutual information makes no assumptions of what the distribution of the variables X and Y is like.

Mutual information, like entropy, is non-negative. Additionally, just like there is conditional entropy, there is also conditional mutual information. Conditional mutual information indicates the mutual information given another variable is given:

$$I(X,Y|Z) = \sum_{k=1}^{o}\sum_{i=1}^{n}\sum_{j=1}^{m} P(x_i,y_j,z_k) log_2(\frac{P(x_i,y_jz_k)P(z_k)}{P(x_i,z_k)P(y_jz_k)}) \tag{4.14}$$

4.4 Multi-Variate Information Theory

In the previous sections, information theory has been observed through a bivariate lens. These methods can be generalised to multivariate variations. When comparing different brain regions, we do not want to restrict ourselves to a bivariate case.

Joint entropy can be easily extended to a multivariate case. In the bivariate case, the joint probability mass function was used and a summation over both random variables was done. The multivariate case simply generalises this equation:

$$H(X_1, ..., X_n) = -\sum_{x_1} ... \sum_{x_n} P(x_1, ..., x_n) log_2(P(x_1, ..., x_n)) \tag{4.15}$$

Equation 4.5 can also be extended into a multivariate case. In this case, we get:

$$H(X_1, ..., X_n) = -\sum_{k=1}^{n} H(X_k | X_{k-1}, ..., X_1) \tag{4.16}$$

The multivariate case of conditional entropy becomes:

$$H(Y|X_1, ..., X_n) = H(Y, X_1, ..., X_n) - H(X_1, ..., X_n) \tag{4.17}$$

Having extended the different forms of entropy into a multivariate case, we can make a multivariate method for mutual information. The multivariate mutual information becomes a recursive function:

add mutual information + conditional

4.5 Continuous Data

explain that we have continuous data, not discrete

4.6 Coding Theory

4.7 Cybernetics

Chapter 5

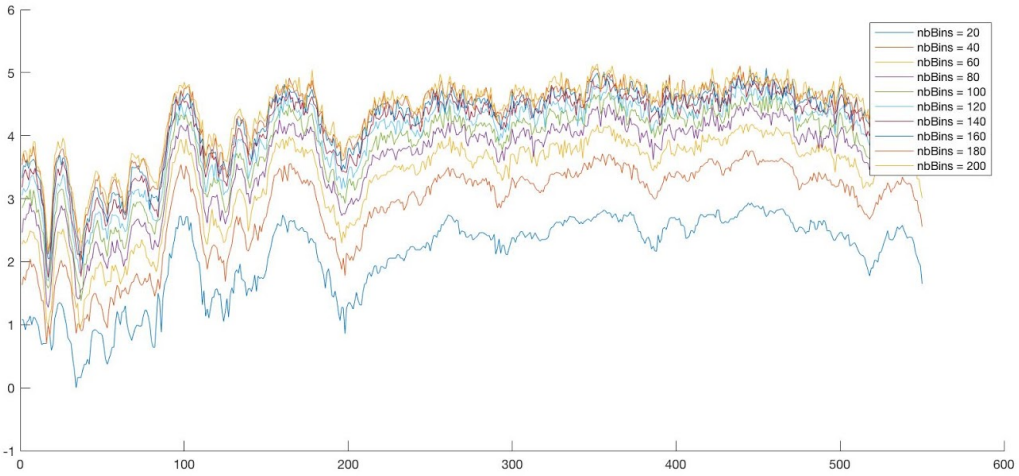
EEG Experiment

Chapter 6

Evaluation

6.1 Calculating Bin Sizes

Figure 6.1: Entropy comparison between bin sizes



6.2 Comparing Common Regions

6.3 Per Subject Comparison

6.3.1 Effect of Trials Used

Figure 6.2: Experiment

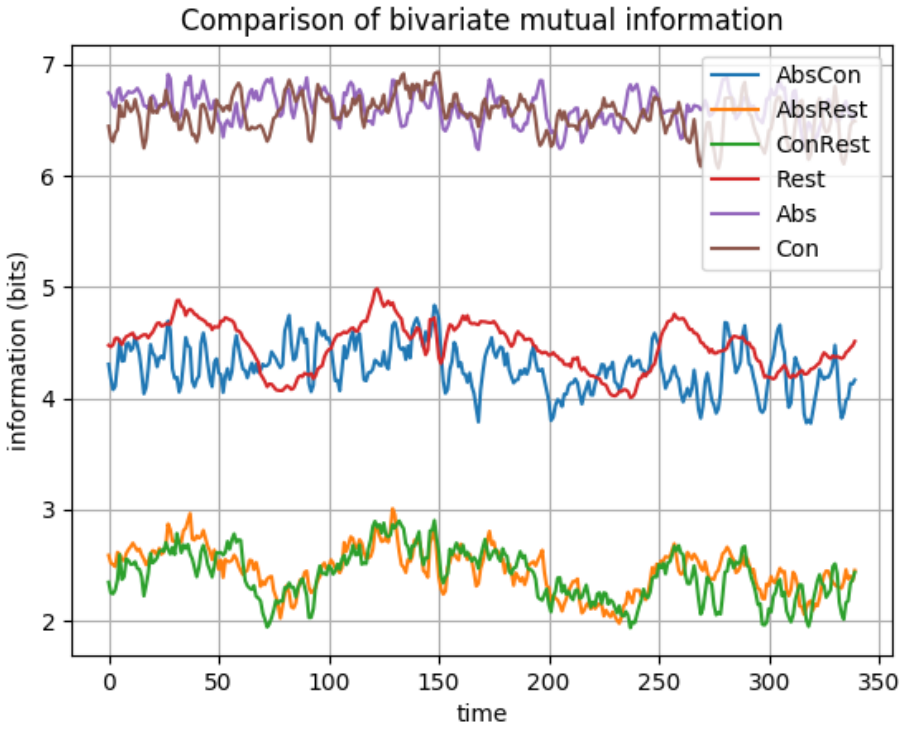


Figure 6.3: Comparison

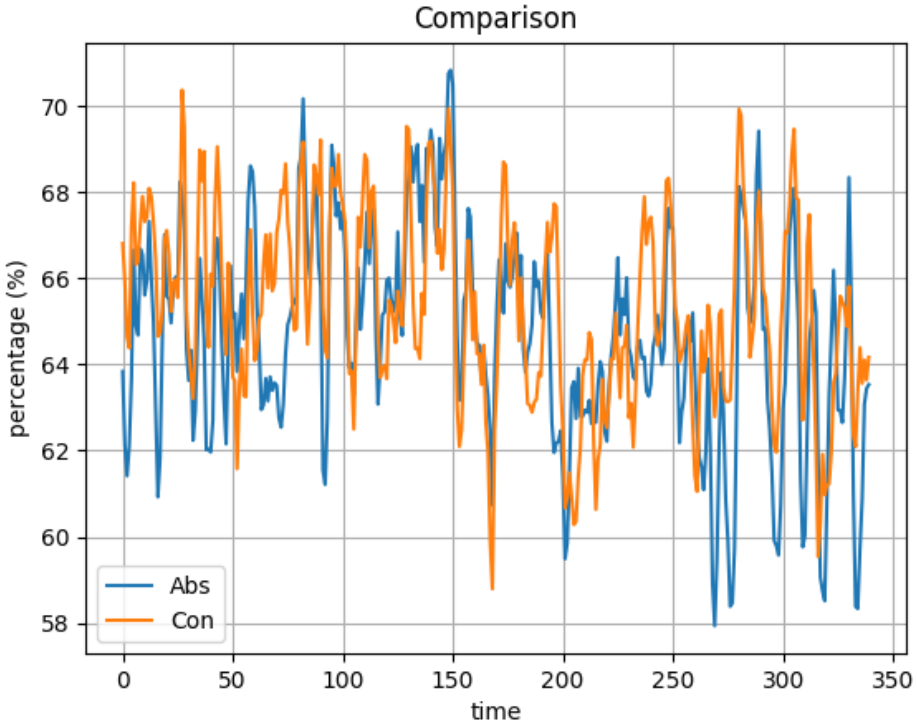


Figure 6.4: Entropy comparison between bin sizes

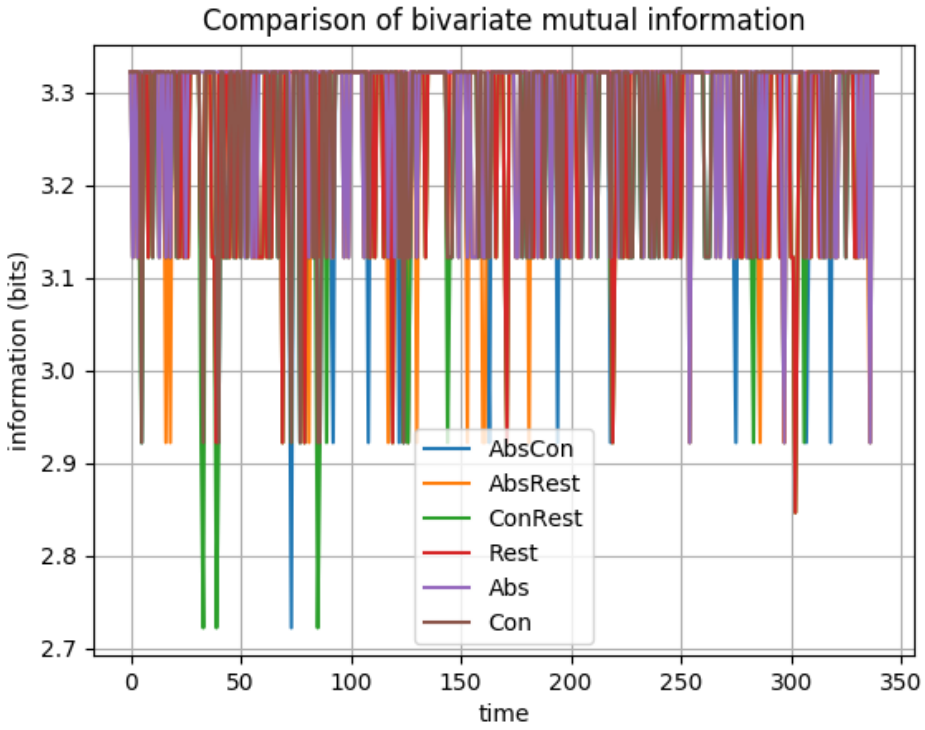


Figure 6.5: Entropy comparison between bin sizes

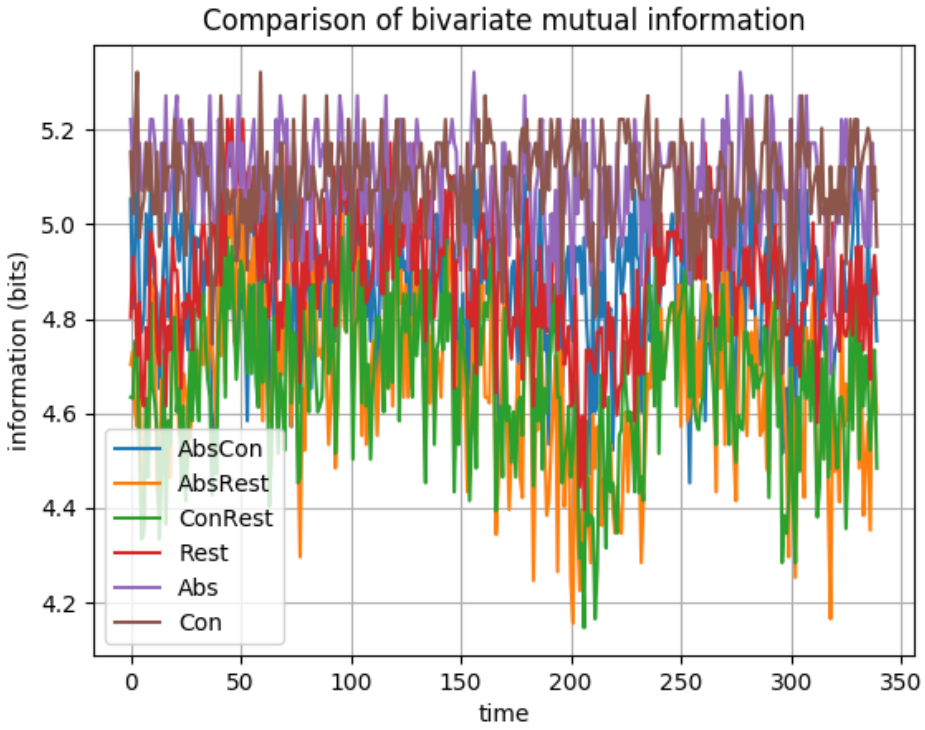
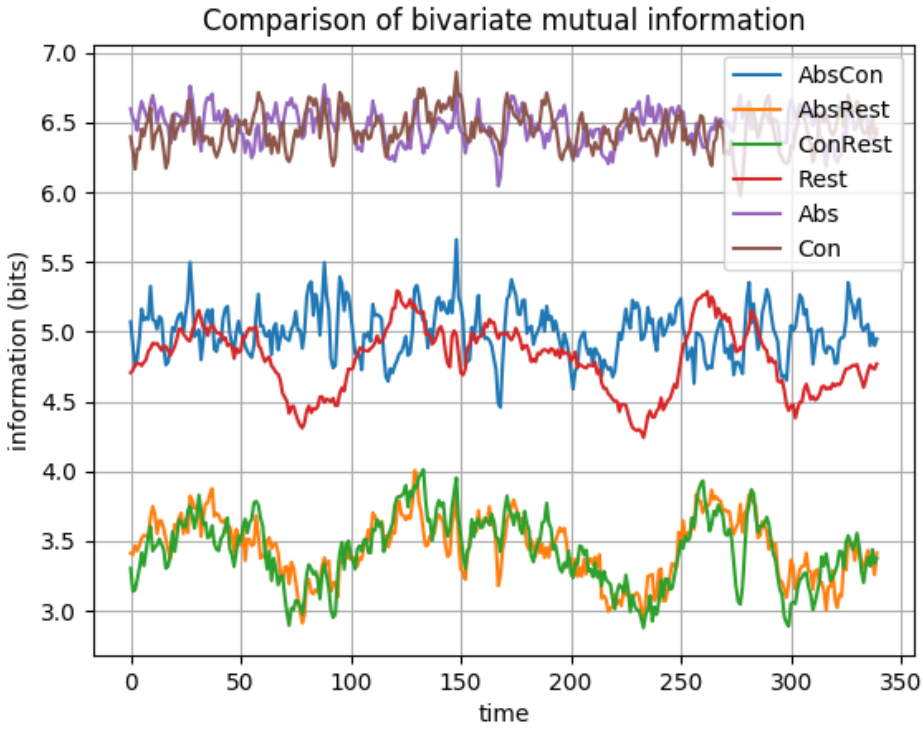


Figure 6.6: Entropy comparison between bin sizes



Chapter 7

Implementation

7.1 Language

7.2 Data Conversion

Chapter 8

Related Work

Chapter 9

Future Work

9.1 Directed Information

9.2 Open Source Connectivity Package

9.3 Data Flow Connectivity

Chapter 10

Conclusion

Bibliography

[1] J. W. Shin and S. J. Kim. A mathematical theory of communication. *University of Ulinois Ress*, 1949.

Master's thesis filing card

Student: Axel Faes

Title: An Information Theoretical Approach to EEG Source-Reconstructed Connectivity

UDC: 681.3*I20

Abstract:

Thesis submitted for the degree of Master of Science in Artificial Intelligence, option Engineering and Computer Science

Thesis supervisor: Prof. dr. ir. Marc Van Hulle

Assessor: Mansoureh Fahimi
Prof. dr. Daniele Marinazzo

Mentor: Mansoureh Fahimi