

**FEDERAL UNIVERSITY OF UBERLÂNDIA
FACULTY OF ELECTRICAL ENGINEERING
POST-GRADUATE IN ELECTRICAL ENGINEERING**

**Single channel approach for filtering electroencephalographic
signals strongly contaminated with facial electromyography**

Carlos Magno Medeiros Queiroz

Uberlândia

2022

Carlos Magno Medeiros Queiroz

Single channel approach for filtering electroencephalographic signals strongly contaminated with facial electromyography

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Sciences to the Post-Graduate Program of the Faculty of Electrical Engineering at the Federal University of Uberlândia.

Field of Knowledge: Information Processing

Line of Research: Biomedical Engineering

Prof. Adriano de Oliveira Andrade, PhD

Supervisor

Prof. Luiz Carlos de Freitas, PhD

Coordinator of the Graduate Program in Electrical Engineering

Uberlândia

2022

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

Q3s
2022 Queiroz, Carlos Magno Medeiros, 1971-
Single channel approach for filtering electroencephalographic signals strongly contaminated with facial electromyography [recurso eletrônico] / Carlos Magno Medeiros Queiroz. - 2022.

Orientador: Adriano de Oliveira Andrade.
Tese (Doutorado) - Universidade Federal de Uberlândia, Programa de Pós-Graduação em Engenharia Elétrica.
Modo de acesso: Internet.
Disponível em: <http://doi.org/10.14393/ufu.te.2023.8032>
Inclui bibliografia.
Inclui ilustrações.

1. Engenharia Elétrica. I. Andrade, Adriano de Oliveira, 1975-, (Orient.). II. Universidade Federal de Uberlândia. Programa de Pós-Graduação em Engenharia Elétrica. III. Título.

CDU: 621.3

André Carlos Francisco
Bibliotecário - CRB-6/3408



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
 Coordenação do Programa de Pós-Graduação em Engenharia Elétrica
 Av. João Naves de Ávila, 2121, Bloco 3N - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902
 Telefone: (34) 3239-4707 - www.posgrad.feelt.ufu.br - copel@ufu.br



ATA DE DEFESA - PÓS-GRADUAÇÃO

| | | | | | |
|------------------------------------|---|-----------------|-------|-----------------------|-------|
| Programa de Pós-Graduação em: | Engenharia Elétrica | | | | |
| Defesa de: | Tese de Doutorado, 307, PPGEELT | | | | |
| Data: | Quatro de novembro de dois mil e vinte e dois | Hora de início: | 08:00 | Hora de encerramento: | 11:40 |
| Matrícula do Discente: | 11313EEL003 | | | | |
| Nome do Discente: | Carlos Magno Medeiros Queiroz | | | | |
| Título do Trabalho: | Single channel approach for filtering electroencephalographic signals strongly contaminated with facial electromyography | | | | |
| Área de concentração: | Processamento da Informação | | | | |
| Linha de pesquisa: | Engenharia Biomédica | | | | |
| Projeto de Pesquisa de vinculação: | Coordenador do projeto: Adriano de Oliveira Andrade Título do projeto: CARACTERIZAÇÃO E FILTRAGEM DE ELETROENCEFALOGRAMA CONTAMINADO POR ELETROMIOGRAMA DE MÚSCULOS FACIAIS Agência financiadora: CNPq Número do processo na agência financiadora: 305223/2014-3 Vigência do projeto: 2014 - 2019 | | | | |

Reuniu-se, por meio de videoconferência, a Banca Examinadora designada pelo Colegiado do Programa de Pós-graduação em Engenharia Elétrica, assim composta: Professores Doutores: Pedro Cunha Carneiro - FEELT/UFU; Wellington Maycon Santos Bernardes - FEELT/UFU; Vytutas Abromavicius - VILNIUS TECH; João Loures Salinet Júnior - UFABC; Adriano de Oliveira Andrade - FEELT/UFU, orientador(a) do(a) candidato(a).

Iniciando os trabalhos o(a) presidente da mesa, Dr(a). Adriano de Oliveira Andrade, apresentou a Comissão Examinadora e o candidato(a), agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos(às) examinadores(as), que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

Aprovado.

Esta defesa faz parte dos requisitos necessários à obtenção do título de Doutor.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Adriano de Oliveira Andrade, Professor(a) do Magistério Superior**, em 04/11/2022, às 11:42, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Vytautas Abromavicius, Usuário Externo**, em 04/11/2022, às 11:42, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Wellington Maycon Santos Bernardes, Professor(a) do Magistério Superior**, em 04/11/2022, às 11:43, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **João Loures Salinet Júnior, Usuário Externo**, em 04/11/2022, às 11:43, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Pedro Cunha Carneiro, Professor(a) Substituto(a) do Magistério Superior**, em 04/11/2022, às 11:44, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3971517** e o código CRC **A6203DBC**.

Carlos Magno Medeiros Queiroz

Single channel approach for filtering electroencephalographic signals strongly contaminated with facial electromyography

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Sciences to the Post-Graduate Program of the Faculty of Electrical Engineering at the Federal University of Uberlândia.

Field of Knowledge: Information Processing

Line of Research: Biomedical Engineering

Uberlândia, 04 of November of 2022.

Thesis Committee:

Prof. Adriano de Oliveira Andrade, PhD - Supervisor (UFU)

Prof. João Loures Salinet Junior, PhD - UFABC

Prof. Vytautas Abromavicius, PhD - VILNIUS TECH

Prof. Pedro Cunha Carneiro, PhD - UFU

Prof. Wellington Maycon Santos Bernardes, PhD - UFU

Acknowledgements

Firstly, thank God for granting me the grace to start and finish this work.

To my supervisor, Professor Adriano de Andrade Oliveira, who was present and committed to my learning process from the very beginning until the end. Thank you for your trust and remarkable dedication; your competence has always inspired me.

To my wife, who suffered and rejoiced with me each day of this long journey, she loved me unconditionally.

To my children, who embarked on my dream; your presence has strengthened me. You have been precious throughout.

To my parents, always dreaming my dreams with me, even when they seemed impossible, even when they were far away, they were always very present.

To Luiza Maire, a great friend who walked by my side in this journey of struggles and achievements.

To Gustavo Moreira, a great companion who always fought generously by my side.

To my friends and teachers at NIATS and Biolab, Amanda Rabelo, Fábio Henrique, Samila, Ariana, Sérgio Ricardo, and Alcimar. Thanks to you for your friendship, companionship in the challenges, friendly hand in the works, and moments of relaxation.

To the UFU graduate program in Electrical Engineering.

To the CAPES for the financial support.

"Fear of the Lord is the foundation of true knowledge, but fools despise wisdom and discipline."

Proverbs 1:7 NLT

Abstract

Queiroz, C. M. M. *Single channel approach for filtering electroencephalographic signals strongly contaminated with facial electromyography*. Master's thesis - Federal University of Uberlândia, November 2022.

Eliminating facial electromyographic (EMG) signal from the electroencephalogram (EEG) is crucial for the accuracy of applications such as brain computer interfaces (BCIs) and brain functionality measurement. Facial electromyography typically corrupts the electroencephalogram. Although it is possible to find in the literature a number of multi-channel approaches for filtering corrupted EEG, studies employing single channel approaches are scarce. In this context, this study proposed a single channel method for attenuating facial EMG noise from contaminated EEG. The architecture of the method allows for the evaluation and incorporation of multiple decomposition and adaptive filtering techniques. The decomposition method was responsible for generating EEG or EMG reference signals for the adaptive filtering stage. In this study, the decomposition techniques CiSSA, EMD, EEMD, EMD-PCA, SSA, and Wavelet were evaluated. The adaptive filtering methods RLS, Wiener, LMS, and NLMS were investigated. A time and frequency domain set of features were estimated from experimental signals to evaluate the performance of the single channel method. This set of characteristics permitted the characterization of the contamination of distinct facial muscles, namely Masseter, Frontalis, Zygomatic, Orbicularis Oris, and Orbicularis Oculi. Data were collected from ten healthy subjects executing an experimental protocol that introduced the necessary variability to evaluate the filtering performance. The largest level of contamination was produced by the Masseter muscle, as determined by statistical analysis of the set of features and visualization of topological maps. Regarding the decomposition method, the SSA method allowed for the generation of more suitable

reference signals, whereas the RLS and NLMS methods were more suitable when the reference signal was derived from the EEG. In addition, the LMS and RLS methods were more appropriate when the reference signal was the EMG. This study has a number of practical implications, including the use of filtering techniques to reduce EEG contamination caused by the activation of facial muscles required by distinct types of studies. All the developed code, including examples, is available to facilitate a more accurate reproduction and improvement of the results of this study.

Keywords: EMG; EEG; signal processing; signal decomposition; facial electromyography;

Contents

| | |
|---|------------|
| List of Figures | xix |
| List of Abbreviations | xx |
| 1 Introduction | 1 |
| 1.1 Problem formulation | 1 |
| 1.2 The relevance of the thesis | 1 |
| 1.3 The objective of research | 2 |
| 1.4 The aim of the thesis | 2 |
| 1.5 The objectives of the thesis | 2 |
| 1.6 Research methodology | 4 |
| 1.7 Scientific novelty of the thesis | 5 |
| 1.8 Practical value of the research results | 6 |
| 1.9 Approval of research results | 6 |
| 1.9.1 Article in scientific Journal | 6 |
| 1.9.2 Full papers published in proceedings of conferences | 7 |
| 1.9.3 Participation in research project | 7 |
| 1.9.4 Collaboration work | 8 |
| 1.10 Thesis structure | 10 |
| 2 Background research | 12 |
| 2.1 Electroencephalography and artifacts | 12 |
| 2.2 Filtering methods | 13 |
| 2.3 Proposed approach | 14 |

| | | |
|----------|--|-----------|
| 3 | Materials and methods | 17 |
| 3.1 | Experimental protocol | 17 |
| 3.2 | Data acquisition | 20 |
| 3.2.1 | EEG cap fitting | 21 |
| 3.3 | Signal processing stages for the implementation of a single channel approach to EEG filtering | 22 |
| 3.4 | Estimate of features for filtering assessment | 28 |
| 3.4.1 | Time domain features | 28 |
| 3.4.2 | Frequency domain features | 30 |
| 3.5 | Statistical analysis | 30 |
| 3.5.1 | Characterization of the contamination of the electroencephalogram by distinct facial muscles | 30 |
| 3.5.2 | Comparison of the performance of distinct decomposition methods | 32 |
| 3.5.3 | Comparison of the performance of distinct adaptive filtering methods and experimental conditions | 34 |
| 3.5.3.1 | Evaluation based in time-domain features | 34 |
| 3.5.3.2 | Evaluation based in frequency-domain features | 35 |
| 3.5.4 | Comparison of execution time of decomposition and adaptive filtering methods | 35 |
| 4 | Results | 37 |
| 4.1 | Typical collected signals | 37 |
| 4.2 | Characterization of the contamination of the EEG by EMG signals | 37 |
| 4.3 | Evaluation of decomposition methods to generate reference signals for adaptive filtering | 42 |
| 4.4 | Evaluation of the filtering based on the time-domain features | 42 |
| 4.5 | Evaluation of the filtering based on the frequency domain features | 47 |
| 4.6 | Execution time of decomposition and adaptive filtering methods | 48 |
| 5 | Discussion | 50 |
| 6 | Conclusion | 54 |
| | References | 54 |

| | | |
|----------|---|-----------|
| A | The GL and GH feature vectors | 64 |
| B | Data: contamination of EEG by facial muscles | 69 |
| B.1 | Database | 69 |
| B.2 | Data summary | 71 |
| B.3 | EMG and EEG data | 72 |
| B.3.1 | EMG data | 72 |
| B.3.2 | EEG data | 72 |
| B.4 | Example of a database “.mat” file | 73 |
| B.5 | Access to Database | 74 |
| C | Data: Motor Learning Protocol | 76 |
| C.1 | Motor learning | 76 |
| C.2 | Graphical learning interface | 77 |
| C.2.1 | Timing of events | 78 |
| C.2.2 | Log file | 79 |
| C.2.3 | Configurable features | 80 |
| C.3 | Mouse emulator system | 80 |
| C.4 | Mouse cursor control | 83 |
| C.5 | EMG Signals | 84 |
| C.6 | EEG signals | 85 |
| C.7 | Collection structure | 86 |
| C.8 | EMG and EEG data | 87 |
| C.8.1 | Structure EMG | 88 |
| C.8.2 | Structure EEG | 89 |
| C.9 | Access to Database | 91 |

List of Figures

| | | |
|-----|---|----|
| 1.1 | An infographic that depicts the concept of the thesis proposal. A dataset was made up of EEG and EMG signals from five different facial muscles of 10 volunteers collected during the execution of a specific protocol. A method was devised using features proposed and estimated from the collected data to determine the level of EMG contamination on the EEG and to evaluate the methods of decomposition and adaptive filtering currently indicated in the literature. | 3 |
| 3.1 | Participants were instructed to make a variety of facial expressions by activating muscles whose electrical activity corrupts the electroencephalogram. The facial expressions were performed with both open and closed eyes. In the neutral condition (A) there was no muscular contraction, whereas in the other conditions the following muscles were activated: Frontalis (B) , Masseter (C) , Orbicularis Oculi (D) , Zygomatic (E) and Orbicularis Oris (F) | 18 |
| 3.2 | The five different muscles involved in the protocol: Frontalis (I) , Orbicularis Oculi (II) , Zygomatic (III) , Orbicularis Oris (IV) and Masseter (V) | 19 |
| 3.3 | Each muscle was contracted 15 times in one of three timing patterns: long (3 seconds), medium (1 second), and short (0.5 second). Each pattern of contraction was repeated five times randomly. A two-second neutral period followed each contraction. This protocol was executed with open and closed eyes. | 20 |

| | | |
|-----|--|----|
| 3.4 | (A) Front view of the box housing the acquisition system's RHD2000 USB interface card (Intan, USA); (B) Rear view of the box, showing the input and output signal connectors (analog and digital); (C) Biopotential conditioning card. | 21 |
| 3.5 | "Collect" window created for the data acquisition program, Intan Technologies (Intan, USA), containing items to facilitate the specification of volunteer data collection. | 22 |
| 3.6 | A commercial EEG cap (model EC20 from EASYCAP, Germany) fitted with a connector (Omnetics, USA), containing a 16-channel RHD2216 conditioning card from Intan Technologies (Intan, USA), was already positioned in one subject. | 23 |
| 3.7 | Filtering EEG signals corrupted by EMG via a series of steps. Linear and non-linear trends are eliminated, as well as outliers. The signal is then decomposed using one of the methods outlined, and the resulting components are thresholded. The thresholding process requires the identification of noise periods in the signal, which are provided by a binary signal generated by an EMG burst detector (Figure 3.9). Once the components have been thresholded, the filtered signal is reconstructed, producing a reference signal, i.e., EEG or EMG reference signal, which can be used as a reference for one of the indicated adaptive filters. Various characteristics are estimated to evaluate the filtering process at distinct stages. Note that when the method EMD-PCA is used it is not necessary to execute the soft-thresholding stage. | 24 |
| 3.8 | The raw EMG signal is corrupted by linear and non-linear trends. The estimated non-linear trend and non-trending signals are shown. | 25 |

3.9 Sequence of required steps for detecting EMG bursts. First, the input signal is preprocessed by removing linear and nonlinear trends, and then the resulting signal is decomposed using EMD. The estimated components are soft-thresholded using a priori knowledge of the signal’s noise level. Signal filtration is achieved by summing the thresholded components. The EMG envelope is determined by estimating the signal’s energy, and bursts are detected using a threshold. As a result of this step, a binary signal is generated in which low levels indicate noise and high levels indicate EMG activity. 27

3.10 Set of time (GL , GH , GX_{in} and GX_{out}) and frequency (p^r , p^f , f_{med}^r and f_{med}^f) domain features used to evaluate the performance of distinct methods for adaptive filtering EEG corrupted by facial EMG. The features compare the signals before and after adaptive filtering and consider the regions in which there is the presence and absence of muscle activity. . 29

3.11 Overview of the analysis required for characterisation of contamination caused by different facial muscles, as well as a comparison of the performance of decomposition methods based on the features GH and GL . The analysis takes into account data grouping by participants, muscles, EEG sensors, and filtering methods. For each group, a similarity measure based on the normalised Euclidian distance can be estimated between a pair of vectors representing GH and GL estimates for varying a parameter used in the soft-thresholding of the signal components. The similarity metrics are used to create spatial brain maps that depict the contamination of EMG levels at various areas. Statistical analyses are carried out for GH , GL , and similarity measures. 31

3.12 Examples of typical GL and GH feature vectors obtained for high (Fp2-F8) and low (O1-O2) levels of EMG contamination on EEG signals. By decomposing the raw EEG signals with EMD, the reference signals were obtained. The contamination level of the EEG signal can be captured by the distance between the feature vectors, in the sense that an increased distance is related to a lower signal to noise ratio. The examples demonstrate typical collected signals for the open and closed eyes scenarios. . 33

| | | |
|-----|--|----|
| 4.1 | Typical EMG and EEG signals collected during the experimental trials. The EMG signals from the left and right Frontalis are shown. These signals were filtered to remove linear and non-linear trends. The EMG bursts were detected and then the binary signals oscillating from two levels were generated. Simultaneously collected EEG signals are shown for Fp2-F8 (high contamination) and O1-O2 (low contamination) locations. The binary signals are placed over the EEG signals for the indication of the periods in which there was EMG contamination. | 38 |
| 4.2 | Using the international 10-20 system, the topological maps illustrate how the studied muscles contaminate distinct brain regions. Lighter colours represent the most contaminated locations, whereas darker colours denote the least contaminated areas. The presented results are for subjects from 1 to 5. The colours represent the similarity measure between the <i>GL</i> and <i>GH</i> features. | 39 |
| 4.3 | Topological maps for subjects from 6 to 10. | 40 |
| 4.4 | Typical <i>GL</i> and <i>GH</i> feature vectors estimated using different decomposition techniques for Subject 1. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles. | 41 |

4.5 **(A)** Box plot of GL for distinct decomposition algorithms, regardless of subject and sensor location. The smaller the value of GL , the more appropriate the filtering method. The dashed lines represent the best result obtained with the SSA approach. Statistically significant differences between methods are represented by labels. All possible combinations of two were evaluated. For example, the EEMD method is represented by label “b” and was statistically different from CiSSA (label “a”), EMD (label “c”), EMD-PCA (label “d”) and SSA (label “e”). **(B)** Box plot of GH for distinct decomposition algorithms, regardless of subject and sensor location. The larger the value of GH the more suitable is the method for filtering. **(C)** Box plot of the mean normalized Euclidean distance between GL and GH for each muscle, independent of subjects and EEG sensors. The larger the value of this metric, the more contamination is caused by the muscle. 43

4.6 Box plot of the mean normalized Euclidean distance between GL and GH for each muscle and subject, independent of the EEG sensor. . . . 44

4.7 Evaluation of distinct adaptive filtering methods based on the time-domain features. The assessment is independent of the decomposition method and specific to the type of reference signal. **(A)**, **(B)**, **(C)** and **(D)** show results referent to GH and GL . **(E)** and **(F)** present the results related to GX_{in} and GX_{out} 45

4.8 A typical EEG signal corrupted by facial EMG. The EEG signal is from the Fp2-F8 region because it is most affected by facial electromyography. EMG bursts can be seen on the detrended EEG signal. The EEG signal was used as a reference signal, estimated from EMD, and then filtered using the RLS filter. The residue, which is the difference between the detrended and filtered EEG data, clearly shows the EMG activity that was eliminated from the signal. The inset plots at the top indicate the selection of two EEG regions contaminated by EMG. The filtered signal is shown over the contaminated signal in red. For each region, the detrended EEG, reference signal, filtered EEG, and residue are shown. 46

| | | |
|------|---|----|
| 4.9 | Median frequency and its power of the electroencephalogram (EEG) together with its components for the whole signal (ENTIRE) and the two experimental conditions (OPEN EYES and CLOSED EYES). A contrast between the raw non-filtered signal with the filtered signal is presented. The asterisks show the pair of variables in which the variable associated to the filtered signal was significantly reduced in comparison to the non-filtered signal, i.e., raw signal. | 47 |
| 4.10 | Execution time of distinct decomposition (A) and adaptive filtering (B) methods as function of the number of samples. The box plots show the central trend and dispersion of execution times (C and D). | 49 |
| A.1 | Typical <i>GL</i> and <i>GH</i> feature vectors estimated using different decomposition techniques for Subject 2. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles. | 64 |
| A.2 | Typical <i>GL</i> and <i>GH</i> feature vectors estimated using different decomposition techniques for Subject 3. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles. | 65 |
| A.3 | Typical <i>GL</i> and <i>GH</i> feature vectors estimated using different decomposition techniques for Subject 4. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles. | 65 |
| A.4 | Typical <i>GL</i> and <i>GH</i> feature vectors estimated using different decomposition techniques for Subject 5. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles. | 66 |
| A.5 | Typical <i>GL</i> and <i>GH</i> feature vectors estimated using different decomposition techniques for Subject 6. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles. | 66 |

| | | |
|-----|---|----|
| A.6 | Typical <i>GL</i> and <i>GH</i> feature vectors estimated using different decomposition techniques for Subject 7. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles. | 67 |
| A.7 | Typical <i>GL</i> and <i>GH</i> feature vectors estimated using different decomposition techniques for Subject 8. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles. | 67 |
| A.8 | Typical <i>GL</i> and <i>GH</i> feature vectors estimated using different decomposition techniques for Subject 9. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles. | 68 |
| A.9 | Typical <i>GL</i> and <i>GH</i> feature vectors estimated using different decomposition techniques for Subject 10. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles. | 68 |
| B.1 | Configuration of EEG electrodes in the EEG cap, bipolar transverse montage. Illustration of the connections of sensors Fp2 and F8, channel 08, according to Table B.2. | 73 |
| B.2 | Matlab workspace window displaying variables from a database file on volunteer 1's masseter muscle data. | 74 |
| B.3 | Part of the amplifier_channels structure fields on the Matlab screen. . . | 74 |
| C.1 | Illustration of the position and size of the four buttons according to protocols 1, 2, and 3 (the illustration is not in scale). | 77 |
| C.2 | Timeline of events in the graphical learning interface. | 78 |
| C.3 | Only the initial part of the cursor log file is shown. It displays the cursor movement to click on TARGET_1 and records the moment when TARGET_3 appears. | 79 |
| C.4 | Mouse emulator that acts as a mouse in manual mode, via keyboard at the top of the console, or in automatic mode, receiving external commands. System based on the Arduino Due board. | 82 |

| | | |
|------|--|----|
| C.5 | Temporal diagram of sending a command from an external device to the Mouse Emulator. | 84 |
| C.6 | Finite state machine implemented to control the mouse cursor through the contraction of a single muscle. The cursor control is performed entirely by distinguishing only three types of duration of Frontalis contractions. | 85 |
| C.7 | (A) Conditioning card and acquisition of biological signals RHD2216, Intan, USA, with 16 differential channels. (B) EMG sensors, assembled set with two pairs of PS25255 sensors with high input impedance (typical value of $20\ G\Omega$), dry contact capacitive coupling and bandwidth from $200\ mHz$ to $20\ kHz$ | 86 |
| C.8 | Configuration of the EEG electrodes in the cap, transverse bipolar montage. Illustration of the connections of sensors Fp2 and F8, channel 08 according to Table C.3. | 87 |
| C.9 | Collection structure for a single session totals 60 targets, given that the task is to click on four targets for each repetition. | 88 |
| C.10 | EmgR data structure was created to store the data related to the EMG signal and other signals employed in the control of the myoelectric interface. The data presented are related to volunteer 10, session 5, and protocol 2 and illustrate the data of the five repetitions. The figure shows raw EMG, filtered EMG, Hilbert transform EMG signals, and mouse control signals. | 89 |
| C.11 | It presents some signals of the emgR structure used in the mouse control by the myoelectric interface. The signals illustrated are: the raw EMG (sgn), filtered EMG (FiltEmg), EMG Hilbert transform (hilbert), and the mouse cursor control signals (smouse_cmd). The zoom given in the mouse command signal displays the levels of the mouse control signals in detail. | 90 |

C.12 Data of learning events related to the activity performed by the volunteer 10, session 4, protocol 1, and repetition 1. The “type” column contains the targets (TARGET_ 1,2,3, and 4) and the “CLICK_IN” commands. Considering the indicated timeline, the “pos_time” column indicates when the target appears, for the targets, and when they are clicked, for “CLICK_IN”. The other columns refer to the characteristics of the items indicated in the “type” column. 91

List of Tables

| | | |
|-----|--|----|
| 2.1 | Comparison between the analyzes carried out by Silva (2020) [50] and current thesis. | 16 |
| B.1 | Variables in alphabetical order generated by the Matlab function, <read-_Intan_RHD2000_file.m>, provided by Intan. | 70 |
| B.2 | EEG sensors channel list. | 72 |
| C.1 | Configuration characteristics of the graphical learning interface. | 81 |
| C.2 | Command list supported by the Mouse Emulator control protocol. | 83 |
| C.3 | EEG sensors channel list. | 86 |

List of Symbols

| | |
|--------------|---|
| <i>ANOVA</i> | one-way Analysis of Variance |
| <i>BCIs</i> | Brain-Computer Interfaces |
| <i>BSS</i> | Blind Source Separation |
| <i>CCA</i> | Canonical Correlation Analysis |
| <i>CEP</i> | Human Research Ethics Committee |
| <i>CiSSA</i> | Circulant Singular Spectrum Analysis |
| <i>DWT</i> | Discrete Wavelet Transform |
| <i>EEG</i> | Electroencephalographic |
| <i>EEGr</i> | EEG reference |
| <i>EEMD</i> | Extended Empirical Mode Decomposition |
| <i>EMD</i> | Empirical Mode Decomposition |
| <i>EMG</i> | Electromyographic |
| <i>EMGr</i> | EMG reference |
| <i>EOG</i> | Electrooculogram |
| <i>GLI</i> | Graphical Learning Interface |
| <i>ICA</i> | Independent Component Analysis |
| <i>JADE</i> | Joint Approximation Diagonalization of Eigen-matrices |
| <i>LMS</i> | Least Mean Square |

| | |
|-------------|----------------------------------|
| <i>MNI</i> | Montreal Neurological Institute |
| <i>MWF</i> | Multichannel Wiener Filter |
| <i>NLMS</i> | Normalized Least Mean Square |
| <i>PCA</i> | Principal Component Analysis |
| <i>RLS</i> | Recursive Least Squares |
| <i>RMS</i> | Root-Mean-Square |
| <i>SSA</i> | Singular Spectrum Analysis |
| <i>UFU</i> | Federal University of Uberlândia |
| <i>WT</i> | Wavelet Transform |

Chapter 1

Introduction

1.1 Problem formulation

This thesis focuses on developing a single channel filtering method to reduce facial electromyography in heavily contaminated electroencephalography signals.

1.2 The relevance of the thesis

Electroencephalographic (EEG) signals are widely used in a variety of clinical and commercial applications, including cognitive neuroscience, brain skill quantification, pathological diagnosis, pre-surgical assessment, biometrics, and brain-computer interfaces (BCIs) [1–4].

Unfortunately, noninvasively collected EEG signals have low amplitudes and are often contaminated by artifacts of external and/or physiological origin [5–7]. For intrinsic reasons in the data collection process, in several situations, these contaminating noises can have amplitudes greater than the EEG signal itself [8, 9]. Thus, it is imperative to employ artifact mitigation techniques to enable electroencephalography in these situations.

In this regard, this study aims to create a single channel filtering method capable of accommodating the multiple decomposition and adaptive filtering techniques available in the current literature [10–15].

1.3 The objective of research

The objective of this study was to derive a single channel filtering method for reducing facial electromyography from EEG signals, which must be sufficiently generic and capable of accommodating multiple decomposition and adaptive filtering techniques in a single architecture.

1.4 The aim of the thesis

This thesis aims to create a procedure that allows the evaluation of filtering methods for heavily contaminated EEG signals by employing a protocol that emulates facial muscle artifacts typical in the operation of a human-computer interface. The evaluation encompasses features in the time and frequency domains and can characterize EMG contamination on EEG, compare different decomposition and adaptive filtering approaches, and estimate contamination reduction and EEG signal preservation. Figure 1.1 shows an overview of this study.

1.5 The objectives of the thesis

To achieve the desired results of the thesis, the following objectives were formulated:

- Define an experimental protocol that allows the evaluation of filtering methods for EEG signals heavily contaminated by facial EMG, containing the typical facial artifacts of the operation of a human-computer interface.
- To define the features to be estimated in the time and frequency domains that allow the characterization of EMG contamination on EEG.
- To propose a method for comparing different decomposition and adaptive filtering approaches, evaluating the reduction in the EMG amplitude contamination and EEG signal preservation.
- To compare the execution times of the decomposition and adaptive filtering techniques used in the current study.

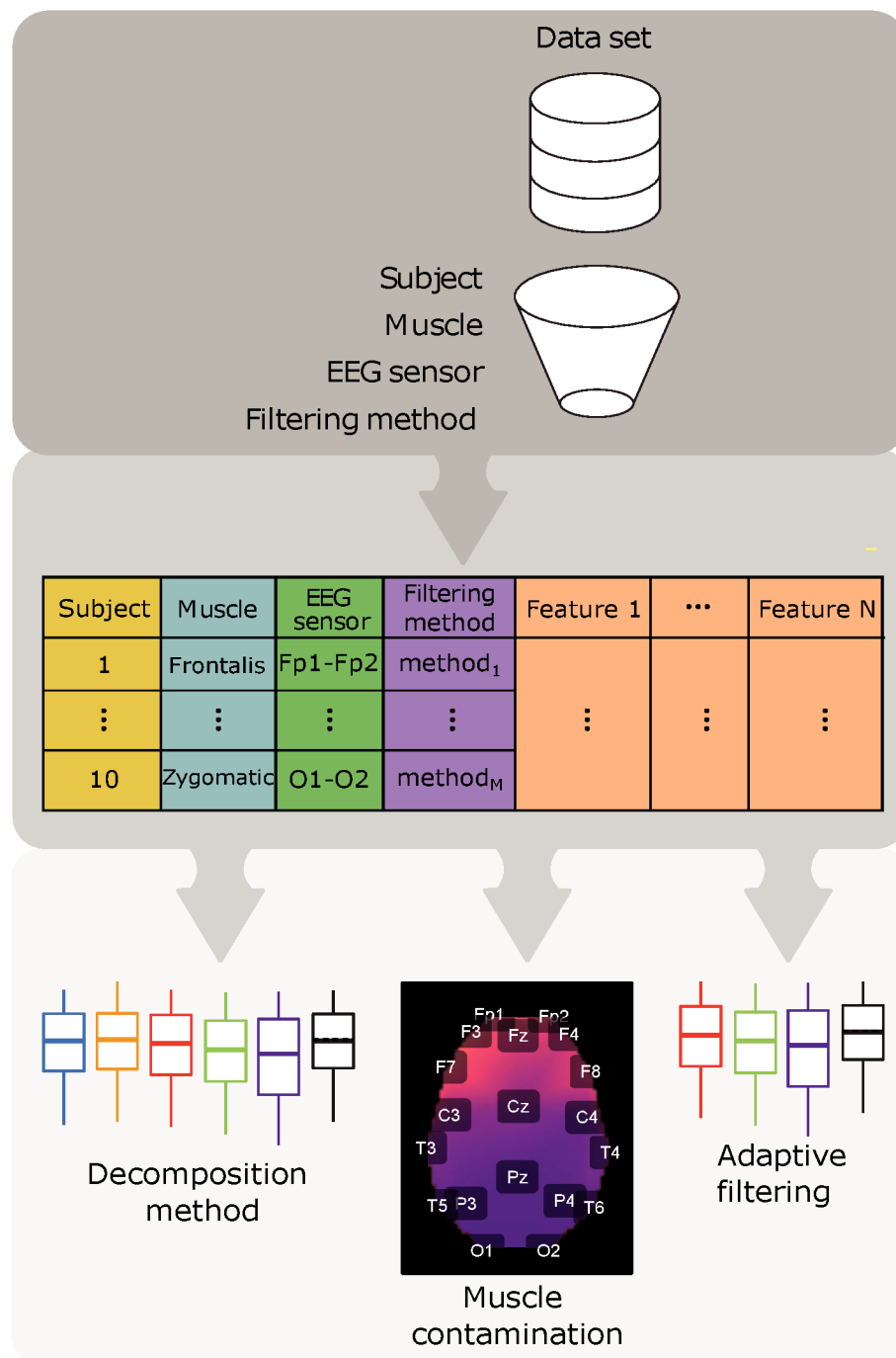


Figure 1.1: An infographic that depicts the concept of the thesis proposal. A dataset was made up of EEG and EMG signals from five different facial muscles of 10 volunteers collected during the execution of a specific protocol. A method was devised using features proposed and estimated from the collected data to determine the level of EMG contamination on the EEG and to evaluate the methods of decomposition and adaptive filtering currently indicated in the literature.

1.6 Research methodology

The research conducted in this thesis was divided into five stages. The first stage refers to the literature survey on the characterization of electromyography, electroencephalography and artifacts. Next, efforts were devoted to studying the main EEG filtering methods available in the current literature.

The second stage involved choosing, assembling, and testing the data-acquisition system. Additions to the customizable graphical interface of the chosen acquisition system were created to facilitate the data collection process. The developed experimental protocol was tested to ensure that defined facial muscle activation generated the desired contamination levels.

In the third stage, two actions were implemented simultaneously. The first action involves the selection of the volunteers, followed by data collection, which consists of the execution of the experimental protocol that comprises the individual activation of the five facial muscles. The second step corresponds to the development and implementation of the decomposition and adaptive filtering code in the Matlab language (MathWorks, USA).

The fourth step involves offline processing of the collected data. The first action in this step consists of viewing all the files recorded during the data collection step, and aims to confirm the levels of the captured signals and the occurrence of the event markers. These actions prevented the occurrence of numerous processing errors. Another important action to speed up the batch processing of the data is to check and correct the names of the collected files according to predefined nomenclature. After the above checks, signal processing must be performed, and the results must be carefully stored to avoid reprocessing and ensure that they are in a format suitable for processing in the next step, that is, statistical analysis.

The final step was statistical analysis of the processed data. The focus of this analysis was the characterization of the contamination of the electroencephalogram by distinct facial muscles, comparison of the performance of different decomposition methods and adaptive filtering under the experimental conditions, and comparison of the execution time required by the evaluated methods.

1.7 Scientific novelty of the thesis

Given that this thesis seeks to introduce a single channel filtering method to reduce contamination of facial EMG on EEG signals, taking into consideration the EMD, EEMD, CiSSA, wavelet, SSA, and EMD-PCA decomposition methods and the RLS, Wiener, LMS, and NLMS adaptive filtering methods, and considering the use of one of the EMG or EEG references, this study postulates the following scientific innovations.

1. The time and frequency domain features proposed in this work allow the characterization of EEG contamination by the facial EMG.
2. The introduction of the normalized Euclidean distance, as a measure of similarity between the proposed GL and GH metrics, allowing visual, qualitative, and quantitative comparisons in the estimated topological maps by volunteers and activated muscles.
3. Among the muscles studied, the masseter had the highest degree of contamination, followed by the frontal and zygomatic muscles, and variability between individuals was identified, indicating that the role of distinct muscles in the contamination of EEG signals varies with regard to the individual.
4. The SSA decomposition method is superior because it successfully preserves the EEG in the uncontaminated regions, while significantly reducing the signal amplitude in the EMG-contaminated regions.
5. The reference based on the EEG signal is preferable because of the lower variability exhibited compared with the EMG reference.
6. The RLS adaptive filtering method is preferred because it produces satisfactory results regardless of the type of the reference signal.
7. The evaluated decomposition and adaptive filtering algorithms have linear computational complexity.
8. Two organized and commented databases were available for employment. One method is suitable for evaluating filtering methods to mitigate facial EMG contamination on EEG. The other is suitable for evaluating motor learning collected

during interaction with a human-computer interface controlled by facial muscles in 11 healthy subjects.

1.8 Practical value of the research results

The results of this research refer to the design of a generic architecture that allows the estimation and evaluation of single channel filtering methods aimed at reducing facial EMG on EEG signals. The use of such an architecture can be of great use in improving the accuracy of processes that employ electroencephalography. The proposed and implemented features in the time and frequency domains allow us to evaluate the reduction of the contamination amplitude and the preservation of the EEG signal, enabling the comparison of the decomposition and adaptive filtering methods, which are essential parts of hybrid filters. Furthermore, using similarity measures between some proposed metrics allows the characterization of contamination through visual, qualitative, and quantitative comparisons of topological maps that can be estimated by volunteers and active muscles. Thus, all Matlab and R scripts are available at <https://doi.org/10.5281/zenodo.6591866>, and duly commented on making this architecture more widespread. In addition, sample data and demonstration scripts were provided to facilitate the understanding and replication of the filtering approach presented and evaluated in this study.

1.9 Approval of research results

In addition to participation in a CNPq Universal Call research project, the results of the research were published in 12 scientific publications, including one peer-reviewed journal article and eleven conference proceedings.

1.9.1 Article in scientific Journal

- **QUEIROZ, C. M. M.** ; da Silva, G. M. ; Walter, S. ; Peres, L. B. ; Luiz, L. M. D. ; Costa, S. C. ; de Faria, K. C. ; Pereira, A. A. ; Vieira, M. F. ; Cabral, A. M. ; ANDRADE, A. O. . “Single channel approach for filtering electroencephalographic

signals strongly contaminated with facial electromyography”. *Frontiers in Computational Neuroscience*, <https://doi.org/10.3389/fncom.2022.822987>. v. 16, p. 1-23, 2022.

1.9.2 Full papers published in proceedings of conferences

- **QUEIROZ, C. M. M.** ; Luiz, L. M. D. ; Silva, G. M. ; NASUTO, S. J. ; Andrade, A. O. . “Remoção de tendências não lineares da eletromiografia em tempo real utilizando filtro de média móvel”. In: XXV Congresso Brasileiro de Engenharia Biomédica, CBEB 2016, Anais do XXV Congresso Brasileiro de Engenharia Biomédica. 2016, Foz do Iguaçu.
- **QUEIROZ, C. M. M.** ; Nasuto, S. J ; Andrade, A. O. . “Development and evaluation of human-computer interface based on facial motor unit activity”. In: 7th International Joint Conference on Biomedical Engineering Systems and Technologies, 2014, Angers, França. 7th International Joint Conference on Biomedical Engineering Systems and Technologies, https://turing.pro.br/anais/ICAART-2014/BIOSTEC/DCBIOSTEC/DCBIOSTEC_2014_13_CR.pdf. 2014. p. 47-53. This paper was presented in the conference area mainly dedicated to doctorate students.
- **QUEIROZ, C. M. M.** ; Silva, B. V. C. ; Nasuto, S. J ; Andrade, A. O. . “Avaliação de diferentes configurações de matrizes de sensores para detecção da atividade muscular facial”. In: XXIV Congresso Brasileiro de Engenharia Biomédica, 2014, Uberlândia. Anais do XXIV Congresso Brasileiro de Engenharia Biomédica. https://www.canal6.com.br/cbeb/2014/artigos/cbeb2014_submission_701.pdf. Bauru: Canal 6, 2014. v. 1. p. 2401-2404.

1.9.3 Participation in research project

1. **Project:** Universal call 14/2013 - Faixa B - CNPq - Processo 477623

Name: Development and Evaluation of Human-Computer Interface Based on Facial Motor Unit Activity

Execution period: 2013 a 2016

Sponsor(es): National Council for Scientific and Technological Development (CNPq) - Financial assistance.

2. **Project:** CNPq - Process: 305223/2014-3

Name: Characterization and filtering of electroencephalogram contaminated by electromyogram of facial muscles

Execution period: 2014 a 2019

Sponsor(es): National Council for Scientific and Technological Development (CNPq) - Scholarship.

1.9.4 Collaboration work

- Silva, G. M. ; Peres, L. B. ; **QUEIROZ, C. M. M.** ; Luiz, L. M. D. ; Vieira, M. F. ; Andrade, A. O. . “Spatial Quantification of Facial Electromyography Artifacts in the Electroencephalogram”. In: XXVI Brazilian Congress on Biomedical Engineering, 2019, Buzios. IFMBE Proceedings. https://doi.org/10.1007/978-981-13-2517-5_68. Singapore: Springer, 2019. v. 70. p. 447-454.
- Silva, G. M. ; Oliveira, F. H. M. ; Peres, L. B. ; **QUEIROZ, C. M. M.** ; Luiz, L. M. D. ; Costa, S. C. ; VIEIRA, M. F. ; Andrade, Adriano O. . “Methodology for Quantification of Frontal Muscle Electromyography Contamination in the Electroencephalogram”. In: World Congress on Medical Physics and Biomedical Engineering, 2018, Praga. doi: 10.1007/978-981-10-9038-7_98. IFMBE Proceedings. Singapore: Springer, 2018. v. 68. p. 535-539.
- Oliveira, M. ; Peres, L. B. ; **QUEIROZ, C. M. M.** ; Silva, J. ; Andrade, A. O. ; Goulart, I. . “Avaliação eletromiográfica da musculatura suprahióidea durante a deglutição de pacientes com hanseníase Virchowiana”. In: XXV Congresso Brasileiro de Engenharia Biomédica, 2016, Foz do Iguaçu. Anais do XXV Congresso Brasileiro de Engenharia Biomédica, 2016.
- Silva, G. M. ; **QUEIROZ, C. M. M.** ; Rabelo, A. G. ; Andrade, A. O. . “Proposta de filtragem do eletroencefalograma da região occipital contaminado por atividade

- muscular facial utilizando decomposição de modos empíricos”. In: XXV Congresso Brasileiro de Engenharia Biomédica, 2016, Foz do Iguaçu. Anais do XXV Congresso Brasileiro de Engenharia Biomédica, 2016.
- Silva, G. M. ; Peres, L. B. ; **QUEIROZ, C. M. M.** ; Luiz, L. M. D. ; Oliveira, I. M. de ; Andrade, A. O. . “Estudo do índice de correlação cruzada entre a atividade eletromiográfica do músculo frontal e o eletroencefalograma”. In: XXV Congresso Brasileiro de Engenharia Biomédica, 2016, Foz do Iguaçu. Anais do XXV Congresso Brasileiro de Engenharia Biomédica, 2016.
 - Peres, L. B. ; Oliveira, M. F. ; **QUEIROZ, C. M. M.** ; Goulart, I. M. B. ; Andrade, A. O. . “Proposta de um protocolo experimental para detecção de atividade eletromiográfica de músculos faciais em indivíduos com hanseníase”. In: VIII Simpósio em Engenharia Biomédica Uberlândia, 2015, Uberlândia. Anais do VIII Simpósio em Engenharia Biomédica. <https://doi.org/10.5281/zenodo.6465503>. Bauru: Canal 6, 2015. v. 1. p. 80-84.
 - Arantes, A. P. B. B. ; Rabelo, A. G. ; **QUEIROZ, C. M. M.** ; Andrade, A. O. . “Proposta de classificação de padrões inerciais e eletromiográficos para aplicação em estratégia de controle de membro superior”. In: Anais do VIII Simpósio em Engenharia Biomédica Uberlândia, 2015, Uberlândia. Anais do VIII Simpósio em Engenharia Biomédica Uberlândia. <https://doi.org/10.5281/zenodo.6465503>. Bauru: Canal 6, 2015. v. 1. p. 282-286.
 - Silva, G. M. ; **QUEIROZ, C. M. M.** ; Lima, G. F. M. ; Barbosa Junior, J. A. F. ; Andrade, A. O. . “Caracterização do EEG contaminado por EMG por meio da decomposição de modos empíricos”. In: VIII Simpósio em Engenharia Biomédica Uberlândia, 2015, Uberlândia. Anais do VIII Simpósio em Engenharia Biomédica Uberlândia. <https://doi.org/10.5281/zenodo.6465503>. Bauru: Canal 6, 2015. v. 1. p. 389-396.
 - Arantes, A. P. B. B. ; **QUEIROZ, C. M. M.** ; Oliveira, E. A. ; Andrade, A. O. . “Avaliação Da Co-Contração Muscular De Um Indivíduo Com Amputação De Membro Superior Durante Contração Isométrica”. In: XXIV Congresso Brasileiro de Engenharia Biomédica, 2014, Uberlândia. Anais do XXIV Congresso Brasileiro

de Engenharia Biomédica. https://www.canal6.com.br/cbeb/2014/artigos/cbeb2014_submission_649.pdf. Bauru: Canal 6, 2014. v. 1. p. 2216-2219.

1.10 Thesis structure

The thesis begins with Chapter 1 which provides a brief overview of the work's motivation, objectives, and structure.

Chapter 2 provides a review of the literature on electroencephalography, artifacts, and current filtering methods, as well as a brief discussion of the filtering proposal in this thesis.

The third Chapter discusses the materials and methods used in this study. It begins by describing the experimental protocol and data acquisition system used. It also describes the signal processing steps, including the implementation of a single channel approach to EEG filtering as well as the resources used in the time and frequency domains to evaluate and compare the performance of different filtering methods, statistical analysis, and the time required for these methods to run.

The findings of this study are detailed in Chapter 4. Typical signals collected are depicted as topological maps that show how the studied muscles contaminate different brain regions. The evaluation results of the decomposition and adaptive filtering methods are presented based on the time and frequency domain characteristics, as well as a comparison of the execution times of the implemented methods.

Finally, Chapters 5 and 6 describe the doctoral work's discussion and conclusions, respectively.

Appendix A presents the behaviour of the GL and GH feature vectors for subjects 2-10. The following appendices describe two databases to facilitate their use. Appendix B presents the database containing EMG and EEG signals collected studying contamination by facial EMG on EEG. Appendix C presents a database created for motor learning studies. This database contains EEG and EMG signals collected from 11 participants who interacted with a human-computer interface. However, before performing the analyses of these data, there was a need to filter the EEG signals that were heavily contaminated due to the use of the electromyographic interface controlled by Frontalis. The search for the most appropriate method resulted in the current thesis.

Therefore, this motor learning database has not yet been analyzed. The analysis of these data should be encouraged in order to verify whether changes in brain activity, detected by changes in rhythmic or macropotential brain waves related to movement, can be used to develop an index capable of estimating motor learning. In this sense, this Appendix C has been added to this thesis precisely to present and disclose the existence of such database for future analysis.

Chapter 2

Background research

2.1 Electroencephalography and artifacts

Electroencephalography is a technique used to record the activity on the scalp of measured cerebral cortex neuronal populations. It is derived from a high temporal resolution, non-invasive macroscopic process and is a low-cost method compared to a functional neuroimaging test [16–18]. The electroencephalogram (EEG) is widely used in a variety of clinical and commercial applications, including cognitive neuroscience, brain-skill quantification, pathological diagnosis, biometrics, and Brain-Computer Interfaces (BCIs) [1–4].

The system for measuring EEG amplifies the tiny disturbances of the electrical potentials of the electrodes positioned on the scalp, which is anatomically separated from the signal-generating sources by the meninges, skull bones, and scalp. Thus, the synaptic potentials which usually have low amplitudes, in the order of millivolts, are strongly attenuated by these anatomical structures, reducing the amplitude of the signals recorded at the scalp [5]. Due to this low amplitude, which typically does not exceed $100 \mu V$, the EEG signal is highly susceptible to artifacts. These artifacts are usually caused by electromagnetic fields generated by nearby electronic devices and the power grid. In addition, artifacts can be produced by other sources of electrophysiological signals, e.g., muscular and heart activity or eye movement [6, 7]. This contamination decreases the performance of applications such as BCI and diagnosis of pathological disfunctions, since the amplitude of the artifact will typically be several orders of magnitude greater than the EEG amplitude [8, 9].

In this context, the characterization and elimination of artifacts is relevant for the correct interpretation and use of EEG. Facial electromyographic (EMG) signals are a primary source of EEG contamination. The main challenge with respect to the elimination of noise generated by the EMG signal lies in the fact that EMG emerges from the anatomically positioned muscles along the skull. Even weak muscular contractions can be detected throughout the scalp due to the phenomenon of conductive volume. In addition, the EMG signal overlaps the spectrum of the EEG signal in virtually all frequency bands [19].

To solve this problem, several EEG filtering methods are described in the literature. However, these methods have some limitations, mainly related to the inability to completely remove noise from the corrupted signal without the introduction of undesired distortions, and the need for a priori noise information for signal filtering. These limitations, associated with several features estimated from the EEG signal to suit the diversity of applications, motivate the search for multiple gold standards for removing EEG artifacts [7, 10–15].

Frequency selector filters, such as a linear Butterworth filter, are one of the main techniques described in the literature for the removal of physiological artifacts from EEG. However, the use of such filter class is only effective when the frequency range of the signal and noise are not overlapped [6].

2.2 Filtering methods

The literature suggests the use of single channel techniques for muscular artifact removal from EEG instead of multichannel techniques, e.g., Independent Component Analysis (ICA) and Canonical Correlation Analysis (CCA). The following methods are commonly employed for this purpose: adaptive filtering [20–22]; Wiener filtering [23, 24], Bayesian filtering [25], Blind Source Separation (BSS) [26, 27], wavelet transform (WT) [28, 29], Empirical Mode Decomposition (EMD) [30, 31], and the combination of these techniques, i.e., hybrid methods [12, 32–35].

An adaptive filter is required when fixed specifications are unknown. The literature describes that the most prevalent family of algorithms for removing EEG artifacts is based on the method of least squares [20, 36]. Adaptive filters vary in time because

their parameters are continuously changing to meet a performance requirement [37].

Wiener filtering is appropriate in situations in which the signal and noise spectrum are overlapping, although it requires an estimated, measured, or reliable reference to operate. Sengupta and Kay (1995) [38] showed that the performance of the multi-channel Wiener filter (MWF) outperformed that of BSS for removal of EEG artifacts of various types, i.e., those that were annotated as unwanted by the user. In addition, Ferdous and Ali (2017) [24] compared Wiener and Kalman filters, and again the Wiener filter was more effective for removing EEG artifacts. However, the Wiener filter was mainly applied to remove ocular artifacts, not including muscular artifacts with low SNR, i.e., lower than -10 dB.

Gao [39] employed an adaptive algorithm to remove ECG from EEG during sleep apnea records by means of Discrete Wavelet Transform (DWT). Iyer [40] compared DWT with an ICA filter for subsequent detection of single-trial evoked potential. Krishnaveni et al. [41] compared the Joint Approximation Diagonalization of Eigen-matrices (JADE) algorithm [42] with DWT for the removal of EOG from EEG.

Empirical Mode Decomposition (EMD) was successfully used for the removal of EEG artifacts in [10, 43] and also in conjunction with BSS methods [43, 44]. A broad review of the application of EMD and its variations on EEG signal processing is given in Sweeney-Reed et al. (2018) [45].

Recent efforts have been focused on the combination of these algorithms for removing artifacts from the EEG. Hybrid methods are, therefore, considered the state of the art in EEG filtering because they use the advantages of different methods in two or more stages and have presented the best results in their applications [6, 7, 12, 14, 35, 44, 46, 47]. The main combinations of algorithms in different filtering stages are: (i) adaptive filtering with BSS-ICA; (ii) EMD with BSS; (iii) wavelet with BSS; (iv) adaptive filtering with EMD [47].

2.3 Proposed approach

Currently, single channel techniques have been shown to be the most effective approach for the removal of facial muscular artifacts from EEG, especially when a reference signal is known [48]. However, the main limitation of this class of noise

removal technique is that its performance is low for signal-to-noise ratios below -10 dB [15, 35, 48], which is typical in EEG contaminated by facial electromyography. To the best of our knowledge, there is lack of studies addressing the removal of facial muscular artifacts from EEG. This is important when there is a need to monitor brain activity during human computer interaction [49].

Recently, in our research group, the thesis Silva (2020) [50] dealt with the development of a new method for removing EMG artifacts in the EEG signal, called EMDRLS. The focus of this work was to evaluate whether such a proposed method was able to mitigate the contamination present in the EEG signal in the context of using a myoelectric interface. This study considered artifacts from the following facial muscles, Masseter and Frontalis.

To contribute to the research on facial EMG removal from EEG, this study presents, a methodology to evaluate EEG filtering for facial EMG removal that is independent of external reference noise and suited for low SNR signals. The evaluation strategy involves determining a single channel reference signal, generated by various decomposition techniques, which are used by adaptive filters to attenuate the facial EMG. Table 2.1 presents a parallel between the Silva (2020) [50] thesis and the current thesis, highlighting the main points addressed and the extent of the study carried out. The reported results of the current research consider the evaluation of various decomposition and adaptive filtering methods, as well as the evaluation of filtering performance in the time and frequency domains. The experimental protocol used in this study is also based on the practical need to assess brain activity for quantification of motor learning during interaction with a myoelectric interface [49].

Table 2.1: Comparison between the analyzes carried out by Silva (2020) [50] and current thesis.

| Thesis Silva, G.M. 2020 [50] | Current thesis |
|---|---|
| <p>It proposes a method</p> <p>Introduces and evaluates a new method for removing EMG artifacts on the EEG signal, called EMDRLS.</p> | <p>It proposes a methodology</p> <p>Single channel filtering evaluation system for reducing facial electromyography on EEG signals, capable of accommodating multiple decomposition and adaptive filtering techniques in a single architecture.</p> |
| <p>Comparisons</p> <p>The performance of the EMDRLS filter is compared with other single channel approaches: wRLS, EMD, Wavelet, and Wiener.</p> | <p>Comparisons</p> <ul style="list-style-type: none"> - Evaluation of the performance of the following decomposition methods: CiSSA, EMD, EEMD, EMD-PCA, SSA, and Wavelet. - Evaluation of the performance of the following adaptive filters: LMS, NLMS, RLS, and Wiener. |
| <p>Experimental protocol</p> <ul style="list-style-type: none"> - 10 healthy subjects, EEG, and EMG (successive and single muscle contractions). - Muscles activated: Masseter and Frontalis. | <p>Experimental protocol</p> <ul style="list-style-type: none"> - 10 healthy subjects, EEG, and EMG (successive and single muscle contractions). - Muscles activated: Masseter, Frontalis, Zygomatic, Orbicularis Oculi, and Orbicularis Oris. |
| <p>Quantitative metrics</p> <ul style="list-style-type: none"> - Signal-to-noise ratio (SNR). - Root mean square error (RMSE). - Mean Power Spectral Density (MPSD). | <p>Quantitative metrics</p> <ul style="list-style-type: none"> - Features in the time domain: GL, GH, GX_{in}, and GX_{out}, calculated based on the median of RMS values estimated in parts of the signal. - Features in the frequency domain: p_r, p_f, f_{med}, and f_{med}^f, calculated based on the median frequency of the power spectral density estimates. |
| <p>Characterization of the contamination</p> <p>The relationship between the normalized average power gain of the contraction and non-contraction windows of the muscles was employed to generate the topographic map.</p> <p>Note: The Zygomatic, Orbicularis Oculi, and Orbicularis Oris muscles were added for contamination characterization.</p> | <p>Characterization of the contamination</p> <p>For each pair of feature vectors (GL and GH) a similarity measure was calculated, based on the normalized Euclidean distance, which is used to generate a topological map.</p> |
| <p>Comparisons made</p> <ol style="list-style-type: none"> 1. Visual. 2. Analysis of filtered signals: metrics, signal-to-noise ratio (SNR). 3. Power spectrum between epochs with uncontaminated and filtered EEG signal: metric, root mean square error of the spectra (RMSE). 4. Spectral preservation of classical EEG rhythms: metrics, mean power spectral density (MPSD). | <p>Comparisons made</p> <ol style="list-style-type: none"> 1. Performance comparison of different decomposition methods: metric, GL and GH. 2. Performance comparison of different adaptive filtering methods: metric, time domain, GL, GH, GX_{in}, and GX_{out}. 3. Spectral preservation of classical EEG rhythms: metrics, in the frequency domain, p^r, p^f, f_{med}^r, and f_{med}^f, according to their fundamental oscillations (Delta, Theta, Alpha, Beta and Gamma). 4. Comparison of runtime of the decomposition and adaptive filtering methods. |

Chapter 3

Materials and methods

3.1 Experimental protocol

Data were collected from ten healthy individuals during the execution of successive facial muscular contractions to characterize the EEG signal contamination by facial muscular activity. This experimental protocol was based on previous published work [49] reporting the implementation of a facial EMG interface and motor learning assessment.

This study followed the Resolution 466/2012 of the National Health Council. The study was conducted at the Centre for Innovation and Technology Assessment in Health of the Federal University of Uberlândia (UFU), Brazil. The experimental protocol was approved by the Human Research Ethics Committee (CEP-UFU), CAAE Number: 43670815.4.0000.5152.

The protocol consisted of two sets of facial muscle contractions, one with the eyes open and one with the eyes closed. The open and closed eyes conditions allowed for the evaluation of the filtering methods considering changes in the EEG amplitude. The 10 participants, 4 women and 5 men (18 - 32 years, mean 26.0, standard deviation 2.5) were instructed to perform a series of facial expressions (Figure 3.1) by contracting five distinct muscles: Frontalis, Masseter, Orbicularis Oculi, Orbicularis Oris and Zygomatic (Figure 3.2).

Each muscle was contracted 15 times following a random timing protocol (Figure 3.3) of three distinct patterns: long (3 seconds), medium (1 second) and short (0.5 second). The onset and duration of the contractions were controlled by an auditory

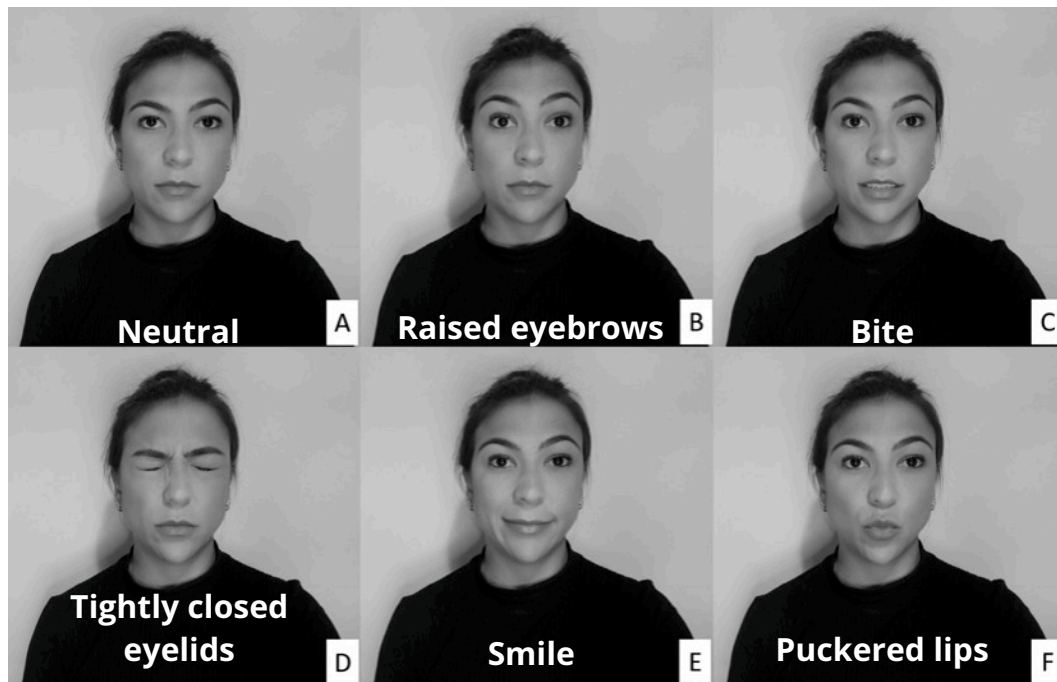
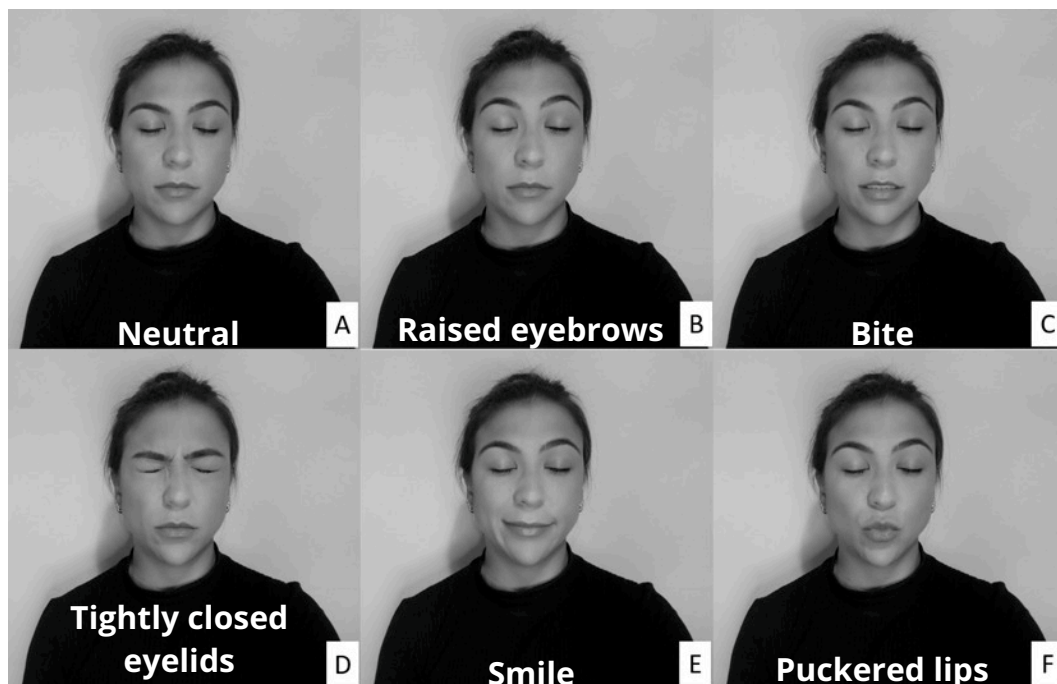
Open eyes**Closed eyes**

Figure 3.1: Participants were instructed to make a variety of facial expressions by activating muscles whose electrical activity corrupts the electroencephalogram. The facial expressions were performed with both open and closed eyes. In the neutral condition (A) there was no muscular contraction, whereas in the other conditions the following muscles were activated: Frontalis (B), Masseter (C), Orbicularis Oculi (D), Zygomatic (E) and Orbicularis Oris (F).



Figure 3.2: The five different muscles involved in the protocol: Frontalis (**I**), Orbicularis Oculi (**II**), Zygomatic (**III**), Orbicularis Oris (**IV**) and Masseter (**V**).

stimulus (beep). The volunteer was asked to maintain the contraction while listening to the beep, and to finish the contraction immediately after the auditory stimulus considering these timing patterns. There were five repetitions of each contraction pattern. Each contraction was followed by a 2-second neutral period. Each participant performed 150 muscle contractions.

Thus, for ten participants, the data set consisted of 500 contractions lasting 3 seconds, 500 contractions lasting 1 second, and 500 contractions lasting 0.5 second, for a total of 2,250 seconds of EEG signals corrupted by facial EMG.

In this study the bipolar, i.e., differential, EEG montage was used to deliberately differentiating potentials between spatially adjacent locations as this may lead to improved signal-to-noise ratio of the collected signal. This type of configuration is also known as longitudinal configuration and widely employed in clinical practice [51]. Although the employed montage was bipolar, the electrodes were positioned by using an EEG cap following the 10-20 International system of EEG electrode placement. EMG signals were detected by using disposable sensors (Meditrace, USA).

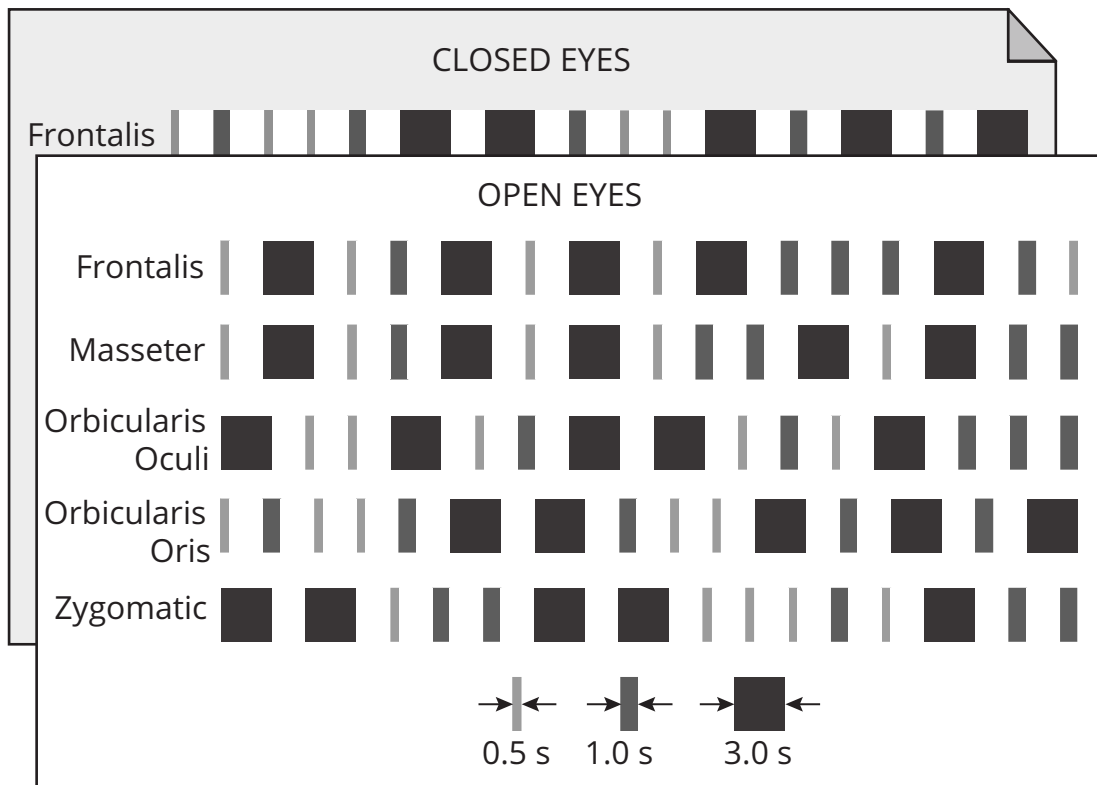


Figure 3.3: Each muscle was contracted 15 times in one of three timing patterns: long (3 seconds), medium (1 second), and short (0.5 second). Each pattern of contraction was repeated five times randomly. A two-second neutral period followed each contraction. This protocol was executed with open and closed eyes.

3.2 Data acquisition

This study compares methods for filtering EMG contamination on EEG signals, so it involves two types of signals with different frequency ranges and amplitudes, and both signals must be collected concurrently. Because the highest frequency of these signals comes from the EMG and considering that the highest frequency of this signal is equal to 1 kHz , one can adopt a sampling frequency of 5 kHz (more than 2 kHz [52]) and a bandpass filter cutoff frequency of 0.1 Hz - 1 kHz .

Among the various commercially available acquisition systems, the acquisition system based on the RHD2000 series of amplifiers from Intan Technologies (Intan, USA) was found to meet the requirements imposed by the experiment [53]. The main features of this acquisition system are biopotential conditioning cards with 16, 32, and 64 bipolar channels, sampling frequencies from 1 kHz to 30 kHz , 16 digital input chan-

nels, and 16 digital output channels synchronized with the biopotential channels and customizable C++/Qt-based multiplatform graphical user interface. Figures 3.4 A and B show the front and back views of the box designed and built to house the USB interface board responsible for forwarding the data, referring to the biopotentials scanned by the two conditioning cards with 16 bipolar channels each (see Figure 3.4 C).

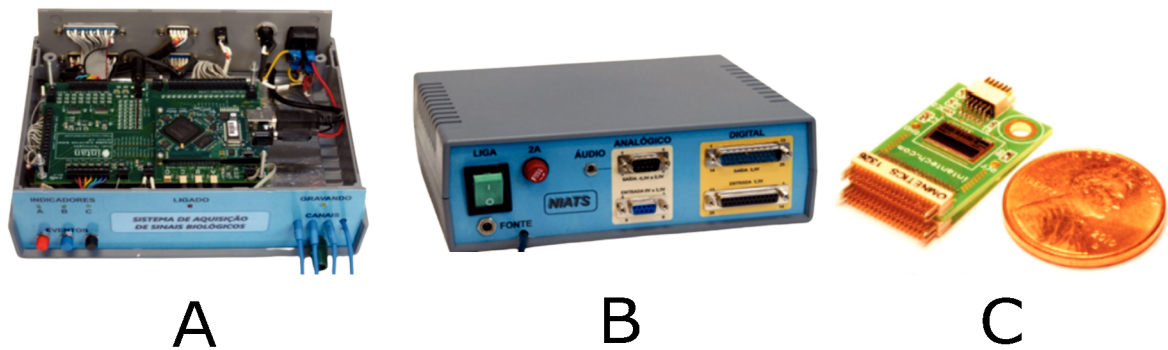


Figure 3.4: **(A)** Front view of the box housing the acquisition system’s RHD2000 USB interface card (Intan, USA); **(B)** Rear view of the box, showing the input and output signal connectors (analog and digital); **(C)** Biopotential conditioning card.

The data acquisition management program, provided for the RHD2000 series from Intan Technologies (Intan, USA), was customized to conduct data collection sessions. A menu item created and named “Collect” allows you to speed up the data entry to specify a given collection. The information “Research”, “Researcher”, “Group”, “Subject”, “Session”, “Protocol”, “Repetition”, “Directory”, “Electrode”, “Cap”, “Cards”, and “Acquisition system” were predefined and easily selected to avoid errors in collection characterization, as illustrated in Figure 3.5.

3.2.1 EEG cap fitting

Electroencephalogram capture follows the international 10/20 positioning standard, employing the transverse bipolar distribution with 16 channels [51, 54]. A commercial EEG cap, model EC20 from EASYCAP, Germany, adapted with the connector (Omnetics, USA) of the 16-channel RHD2216 conditioning card from Intan Technologies (Intan, USA), is used to facilitate the positioning of the electrodes, as illustrated in Figure 3.6.

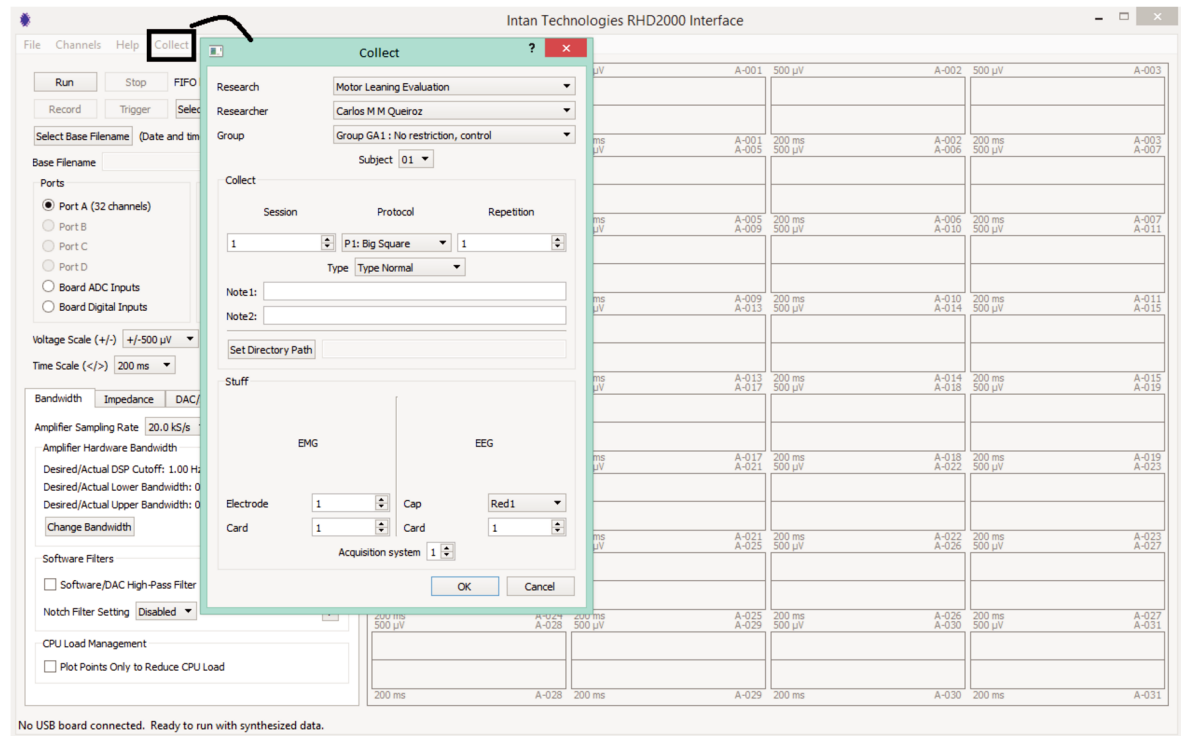


Figure 3.5: “Collect” window created for the data acquisition program, Intan Technologies (Intan, USA), containing items to facilitate the specification of volunteer data collection.

3.3 Signal processing stages for the implementation of a single channel approach to EEG filtering

Figure 3.7 depicts the sequence of steps required to implement a single approach for filtering EEG signals corrupted by facial EMG. The first step is to eliminate any linear and non-linear trends from the collected signals. These tendencies are typically due in part to drift caused by changes in the impedance between the skin and the electrode, as well as cable and skin motion. By fitting a linear model to the time series and then subtracting the resulting straight line from the data, the linear trend is eliminated.

The non-linear trend is estimated by applying a sliding, non-overlapping, rectangular window of 20 milliseconds (100 samples) to the data and then estimating the median of each window. The resultant time-series is interpolated using a Piecewise Cubic Hermite Interpolating Polynomial (pchip) so that it can be re-sampled with the same number of samples as the input time-series. The resultant signal is the non-linear

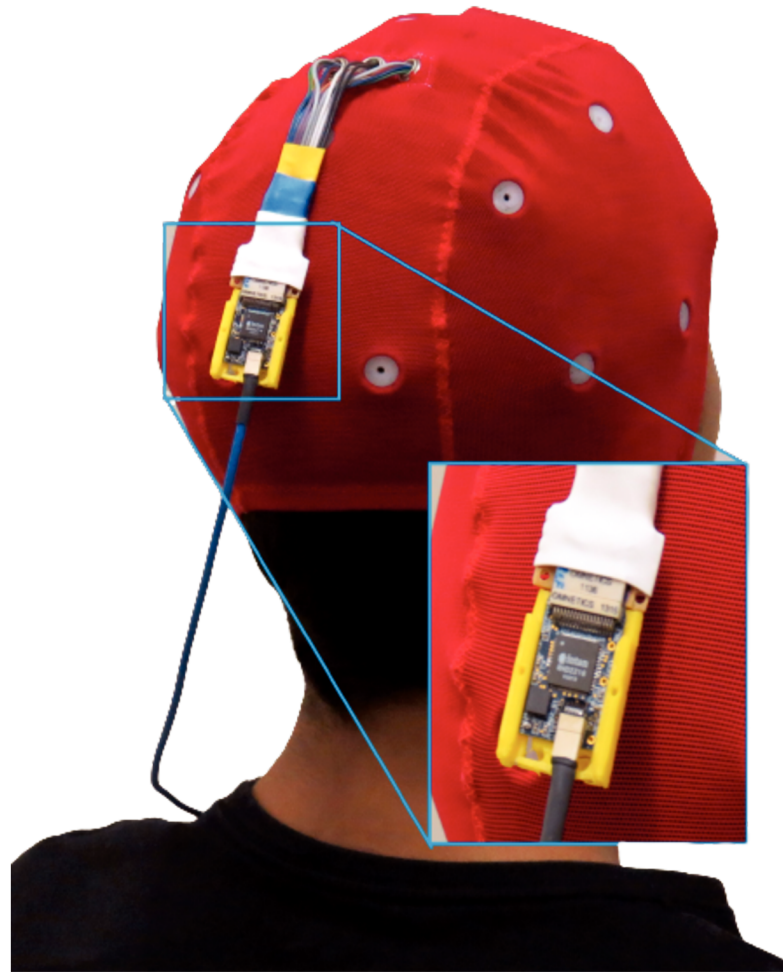


Figure 3.6: A commercial EEG cap (model EC20 from EASYCAP, Germany) fitted with a connector (Omnetics, USA), containing a 16-channel RHD2216 conditioning card from Intan Technologies (Intan, USA), was already positioned in one subject.

trend that should be subtracted from the signal (i.e., the electroencephalogram or electromyogram). Figure 3.8 illustrates an example of the result of this signal processing stage applied to an acquired EMG signal.

Outliers can result from any sudden abnormal changes in data amplitude that exceed or fall below predetermined thresholds. In this study, the upper/lower threshold was established as the mean plus/minus ten times the standard deviation of the data in the EMG-contaminated regions. The outliers were replaced by random scalars drawn from the standard normal distribution.

The pre-processed signal is then decomposed by one of the following decomposition methods: Empirical Mode Decomposition (EMD) [55], Extended Empirical

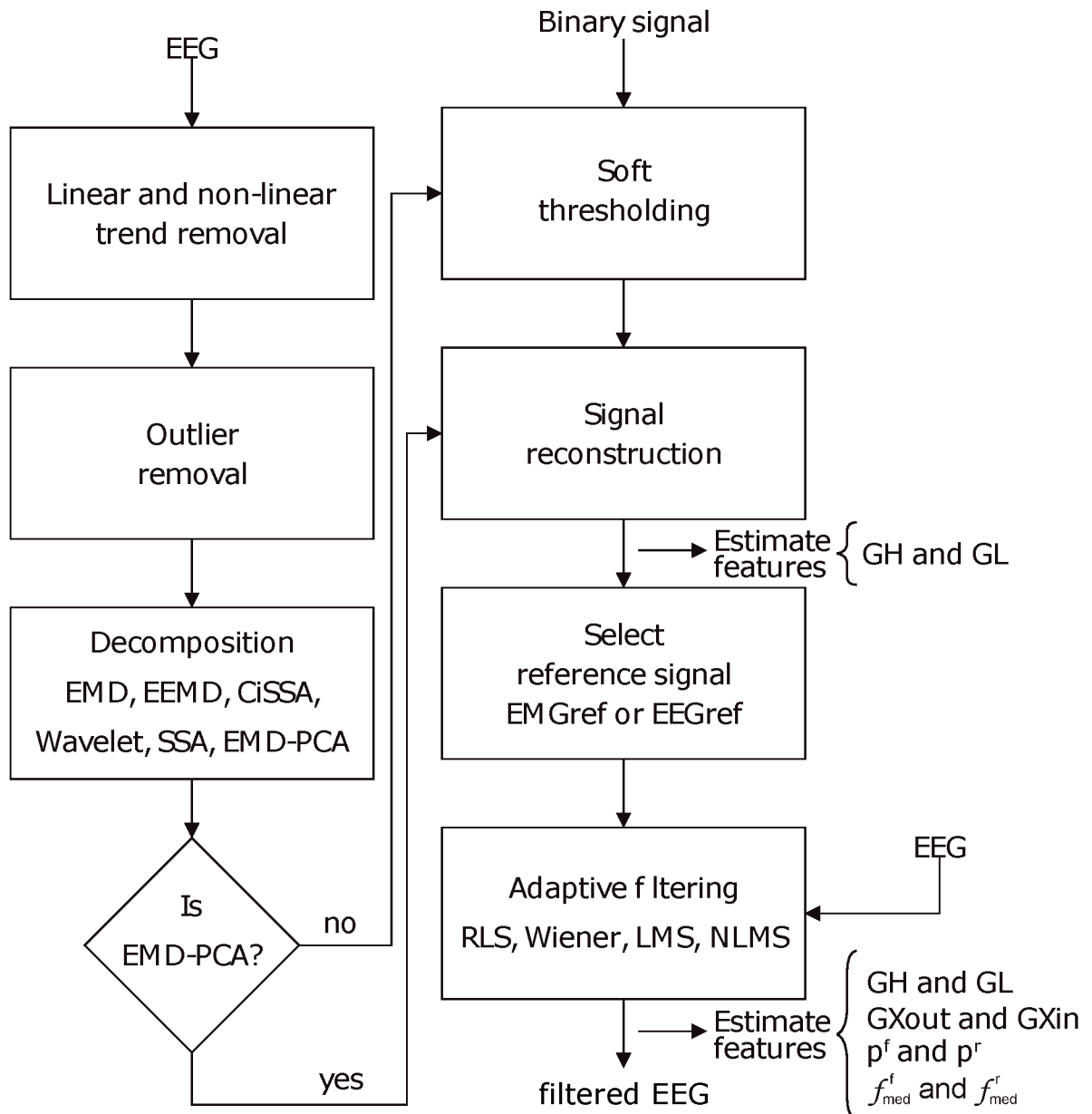


Figure 3.7: Filtering EEG signals corrupted by EMG via a series of steps. Linear and non-linear trends are eliminated, as well as outliers. The signal is then decomposed using one of the methods outlined, and the resulting components are thresholded. The thresholding process requires the identification of noise periods in the signal, which are provided by a binary signal generated by an EMG burst detector (Figure 3.9). Once the components have been thresholded, the filtered signal is reconstructed, producing a reference signal, i.e., EEG or EMG reference signal, which can be used as a reference for one of the indicated adaptive filters. Various characteristics are estimated to evaluate the filtering process at distinct stages. Note that when the method EMD-PCA is used it is not necessary to execute the soft-thresholding stage.

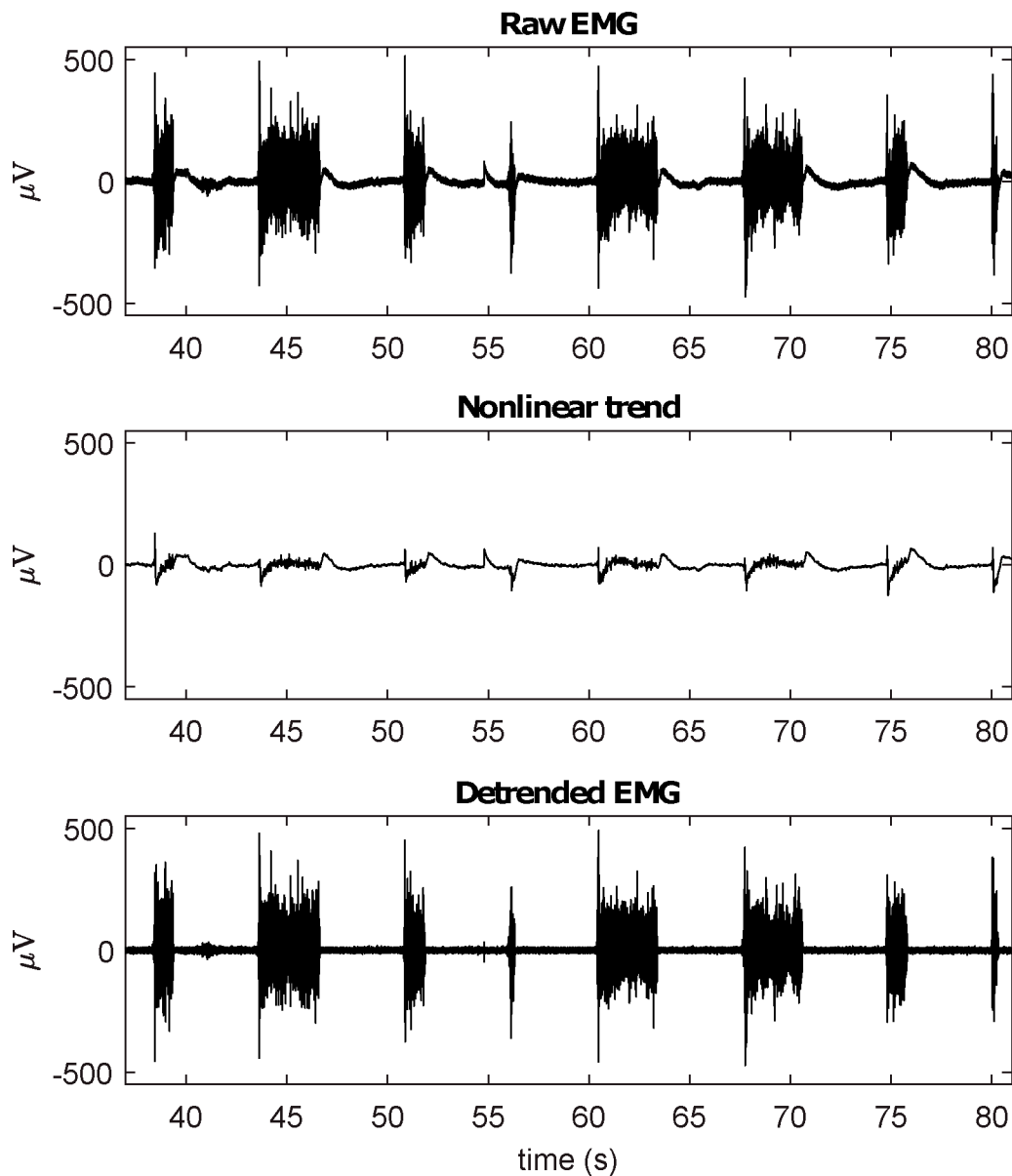


Figure 3.8: The raw EMG signal is corrupted by linear and non-linear trends. The estimated non-linear trend and non-trending signals are shown.

Mode Decomposition (EEMD) [56], Circulant Singular Spectrum Analysis (CiSSA) [57], Wavelet Decomposition [29] or Singular Spectrum Analysis (SSA) [57]. For the methods EMD, EEMD, CiSSA, and Wavelet, the maximum number of components was set to ten. For EMD and EEMD, the pchip interpolation method was utilised. For the EEMD approach, there were 5 ensembles. For Wavelet Decomposition, the mother wavelet

was *coif5*. For the method SSA, the window length was 100 and the proportion of explained variance was 80%.

The obtained components are then soft-thresholded to eliminate noise as explained in a previous work [58]. For each signal component $C = \{c_1, c_2, \dots, c_M\}$, a threshold, $t_m \mid m = \{1, \dots, M\}$, is estimated, and soft-thresholding is applied to individual components as shown in Equation 3.1,

$$tc_m = \text{sign}(c_m)(|c_m| - t_m)_+ \quad (3.1)$$

where tc_m is the de-noised (or thresholded) version of the m th signal component and the function $(x)_+$ is defined as

$$(x)_+ = \begin{cases} 0, & x < 0 \\ x, & x \geq 0. \end{cases} \quad (3.2)$$

The threshold t_m is estimated by using the following strategy: a window of noise is selected from the original signal and then the boundaries of this window are used to extract regions of noise from the signal components. For noise information selection, a binary signal is used. Low-level periods in this binary signal correspond to noise, while high-level periods correspond to EMG regions. Figure 3.9 provides an overview of the required steps for automatic EMG burst detection [58]. Although the EMG signal is the input signal illustrated in Figure 3.9, EMG bursts can be detected directly from EEG that has been corrupted by EMG. Because EMG signals were collected simultaneously with EEG, we decided not to use the EMG-corrupted EEG in this study for EMG burst detection.

The standard deviation of each of those regions is then estimated, multiplied by a constant k , to obtain the required thresholds (t_1, \dots, t_M) . A typical value of k is 1.5 [58]. It is possible to vary k to control the signal filtering.

The EMG reference signal is obtained during the signal reconstruction stage. The adaptive filter may use this signal as a reference. Optionally, an EEG signal may be used as a reference. In this case, the noise-corrupted EEG is subtracted from the EMG reference signal to produce the EEG reference signal. Note that the reference signal is a filtered signal, resulting from the reconstruction of soft-thresholded signal components. In the case of the EMD-PCA method, the signal is reconstructed by selecting

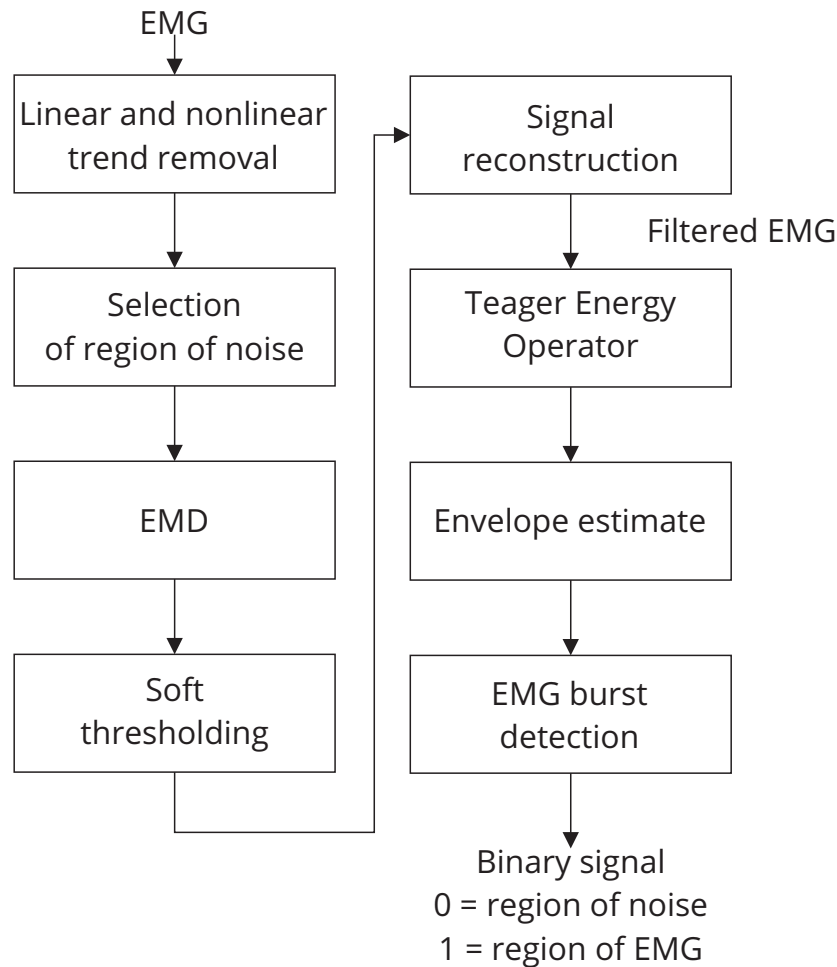


Figure 3.9: Sequence of required steps for detecting EMG bursts. First, the input signal is preprocessed by removing linear and nonlinear trends, and then the resulting signal is decomposed using EMD. The estimated components are soft-thresholded using a priori knowledge of the signal's noise level. Signal filtration is achieved by summing the thresholded components. The EMG envelope is determined by estimating the signal's energy, and bursts are detected using a threshold. As a result of this step, a binary signal is generated in which low levels indicate noise and high levels indicate EMG activity.

the principal components that account for at least 80% of the data variability. For the other methods, the signal is reconstructed based solely in the estimated components.

The reference signal is sent through an adaptive filter, which removes EMG noise from the electroencephalogram. One of the following adaptive filters can be chosen: Recursive Least Squares (RLS), Wiener filter, Least Mean Square (LMS) and Normalized Least Mean Square (NLMS) [59]. Except for the method NLMS, which had an

order of 4, the step-size utilised for all filters was 10^{-7} and the order was 10. Matlab R2022a was used to implement all of the code required for signal processing (MathWorks, USA).

3.4 Estimate of features for filtering assessment

Several features were estimated to enable for the characterisation of EMG contamination on EEG and to compare different approaches used in the investigation. The stages in which the set of features is estimated are depicted in Figure 3.7.

3.4.1 Time domain features

The feature GL assesses the effect of filtering in regions with no EMG activity, i.e., regions with a low binary signal (Figure 3.10). Any filtering method is expected to preserve the amplitude and shape of the signals in this region as much as possible. Equation 3.3 defines GL . The general idea is to apply a sliding, non-overlapping one-second window with 5,000 samples to the data, compute the root-mean-square (RMS) value for each window, and then estimate the median of the RMS values. As the length of the window is one second, the number of samples in each window is equal to the sampling frequency f_s , which is 5 kHz .

$$GL = 20 \log(Xout'_0 / Xin'_0) \quad (3.3)$$

where

$$Xin'_0 = median \left\{ RMS \left(Xin_0 \begin{matrix} i_{2n} \\ i_{1n} \end{matrix} \right) \right\} \quad (3.4)$$

corresponds to the median of the RMS values estimated from the signal Xin_0 , i.e., the non-filtered signal, and

$$Xout'_0 = median \left\{ RMS \left(Xout_0 \begin{matrix} i_{2n} \\ i_{1n} \end{matrix} \right) \right\} \quad (3.5)$$

corresponds to the median of the RMS values estimated from the signal $Xout_0$, i.e., the filtered signal, being $n = \left\{ 1, 2, \dots, \left\lfloor \frac{N}{f_s} \right\rfloor \right\} \in \mathbb{N}$ a set in which each of its values corresponds to a window, $\left\lfloor \frac{N}{f_s} \right\rfloor$ the total number of windows,

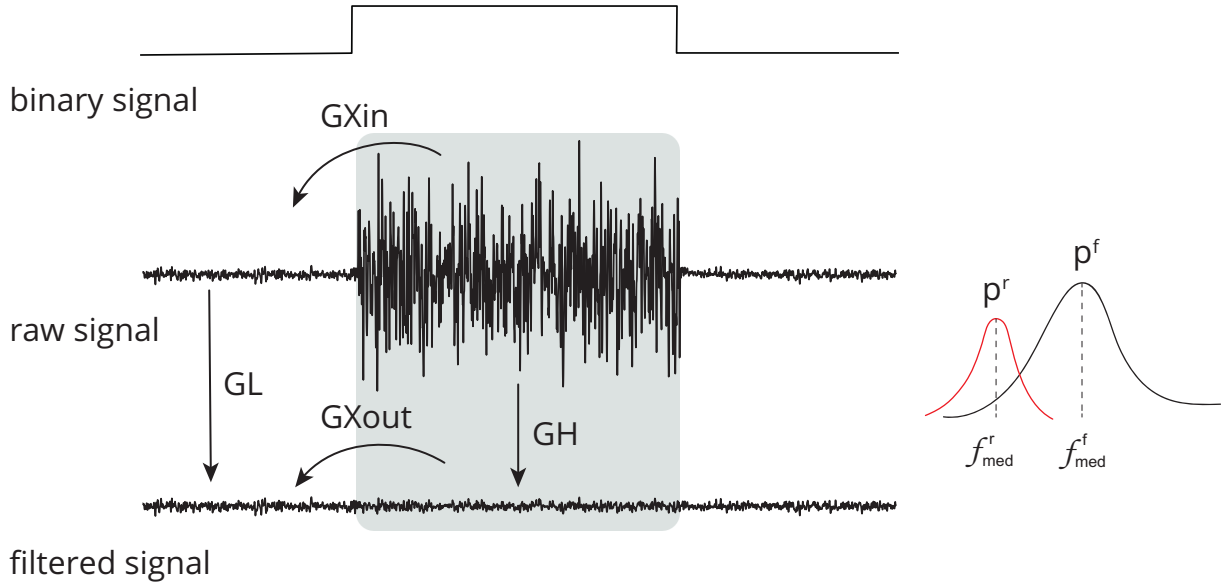


Figure 3.10: Set of time (GL , GH , $GXin$ and $GXout$) and frequency (p^r , p^f , f_{med}^r and f_{med}^f) domain features used to evaluate the performance of distinct methods for adaptive filtering EEG corrupted by facial EMG. The features compare the signals before and after adaptive filtering and consider the regions in which there is the presence and absence of muscle activity.

$i_1 = \left\{ 1, f_s + 1, 2f_s + 1, \dots, \left(\left\lfloor \frac{N}{f_s} \right\rfloor f_s + 1 \right) - f_s \right\} \in \mathbb{N}$ the discrete time in which the window starts, and $i_2 = \left\{ f_s + 1, 2f_s + 1, \dots, \left\lfloor \frac{N}{f_s} \right\rfloor f_s + 1 \right\} \in \mathbb{N}$ the discrete time in which the window ends. N is the number of samples of the signal.

GH (Figure 3.10) is the feature that estimates the ratio between noise-corrupted and filtered signals in an EMG-contaminated region. It is defined in Equation 3.6. It is calculated in a manner similar to that of GL ; hence, equivalent definitions will not be supplied to prevent duplication.

$$GH = 20 \log(Xout'_1 / Xin'_1) \quad (3.6)$$

While the GL and GH features evaluate the ratio of signal amplitudes considering different parts of the binary signal, the $GXin$ and $GXout$ features measure the ratio of signal amplitudes comparing regions with and without noise (Figure 3.10), as given in Equations 3.7 and 3.8. The estimates are similar to that of GL and GH , thus they are not provided.

$$GXin = 20 \log(Xin'_1 / Xin'_0) \quad (3.7)$$

$$GXout = 20 \log(Xout'_1 / Xout'_0) \quad (3.8)$$

3.4.2 Frequency domain features

For the estimate of the frequency domain features (p^r , p^f , f_{med}^r and f_{med}^f), first the power spectral density estimate, pxx , of the discrete-time signal was estimated by using the Yule-Walker method. The signal energy was estimated for the frequency $f = \{0, 0.01, 0.02, \dots, f_s/2\}$ in Hz , considering a model of order 10. The median frequency and its corresponding energy were estimated from pxx for the non-filtered (f_{med}^r and p^r) and filtered (f_{med}^f and p^f) signals.

3.5 Statistical analysis

Statistical analysis was performed using R, which is a language and environment for statistical computing [60]. Considering the studied methods and experimental conditions, the analyses were designed to answer the following research questions: (i) Which facial muscle contributes the most to EEG contamination? (ii) Which decomposition methods are preferable for generating reference signals for adaptive filtering? (iii) Considering its effect on the EEG signal and its components, which adaptive filtering methods are the most desirable?

3.5.1 Characterization of the contamination of the electroencephalogram by distinct facial muscles

Figure 3.11 depicts the main steps employed to characterize the contamination of EEG by facial muscles. The GL and GH features were used to investigate how distinct facial muscles contaminate EEG signals. By varying the parameter k of the soft-thresholding procedure, from 0.1 to 2.0 with a resolution of 0.1, it is possible to obtain, for each decomposition method, a distinct feature vector for GL and GH . Figure 3.12 shows typical vectors for the collected signals. The estimated features, GL and GH , were grouped by subjects, muscles, EEG sensors and filtering methods. For each pair of feature vectors a similarity measure based on the normalized Euclidian distance was computed [61]. These values of the similarity measures were used to generate a

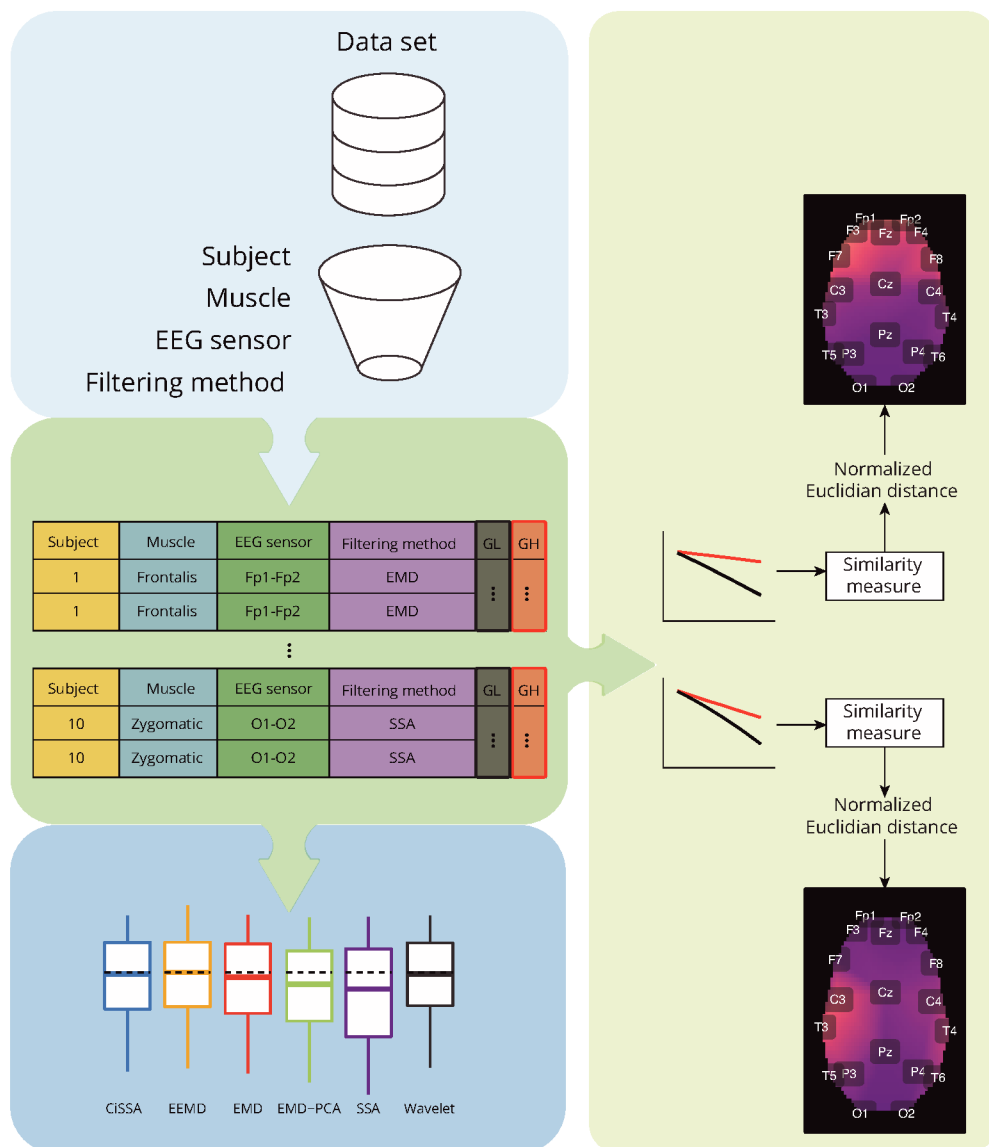


Figure 3.11: Overview of the analysis required for characterisation of contamination caused by different facial muscles, as well as a comparison of the performance of decomposition methods based on the features GH and GL . The analysis takes into account data grouping by participants, muscles, EEG sensors, and filtering methods. For each group, a similarity measure based on the normalised Euclidian distance can be estimated between a pair of vectors representing GH and GL estimates for varying a parameter used in the soft-thresholding of the signal components. The similarity metrics are used to create spatial brain maps that depict the contamination of EMG levels at various areas. Statistical analyses are carried out for GH , GL , and similarity measures.

topological map in which the light colours are associated to the contamination of facial EMG on EEG. A customized function was developed in R to generate the topological maps according to the Montreal Neurological Institute (MNI) coordinates mapped to the International 10-20 System [62]. In addition to the visualization of topological maps, the normalized Euclidean distance between GL and GH were used to quantify the electromyographic contamination produced by distinct muscles. The box plot of the normalized Euclidean distance were computed and the mean of the variables were statistically compared. The statistical analysis was performed considering the scenarios of independence and dependence to the subjects.

To compare the variables, one-way analysis of variance (ANOVA) was used. After fitting the ANOVA model to the data, the model's assumptions were verified, i.e., the evaluation of the homogeneity of variances (Levene's test) and normality of the residuals' distribution (Kolmogorov-Smirnov test). The p-value for all analyses was 0.05. Tukey's honestly significant difference test (Tukey's HSD) was used to examine the significance of differences between sample means. If the variables did not meet the assumptions of ANOVA, the Kruskal-Wallis rank sum test and Dunn's test for multiple comparisons were employed to compare them.

3.5.2 Comparison of the performance of distinct decomposition methods

As seen in Figure 3.7, the decomposition methods in conjunction with the soft-thresholding procedure act as a filtering method that enables the generation of a suitable reference signal for adaptive filtering. In this regard, it is necessary to preserve EMG regions as much as possible so that adaptive filters can attenuate them appropriately.

The GL and GH measures were used to compare the performance of different decomposition approaches. Based on the definitions of GL (Equation 3.3) and GH (Equation 3.6), the most appropriate filtering method is the one that produces the lowest GL and the largest GH , i.e., the method that reduces the signal amplitude in the regions without EMG while preserving the amplitude in the regions of EEG contaminated by EMG as much as possible. For the generation of an appropriate reference

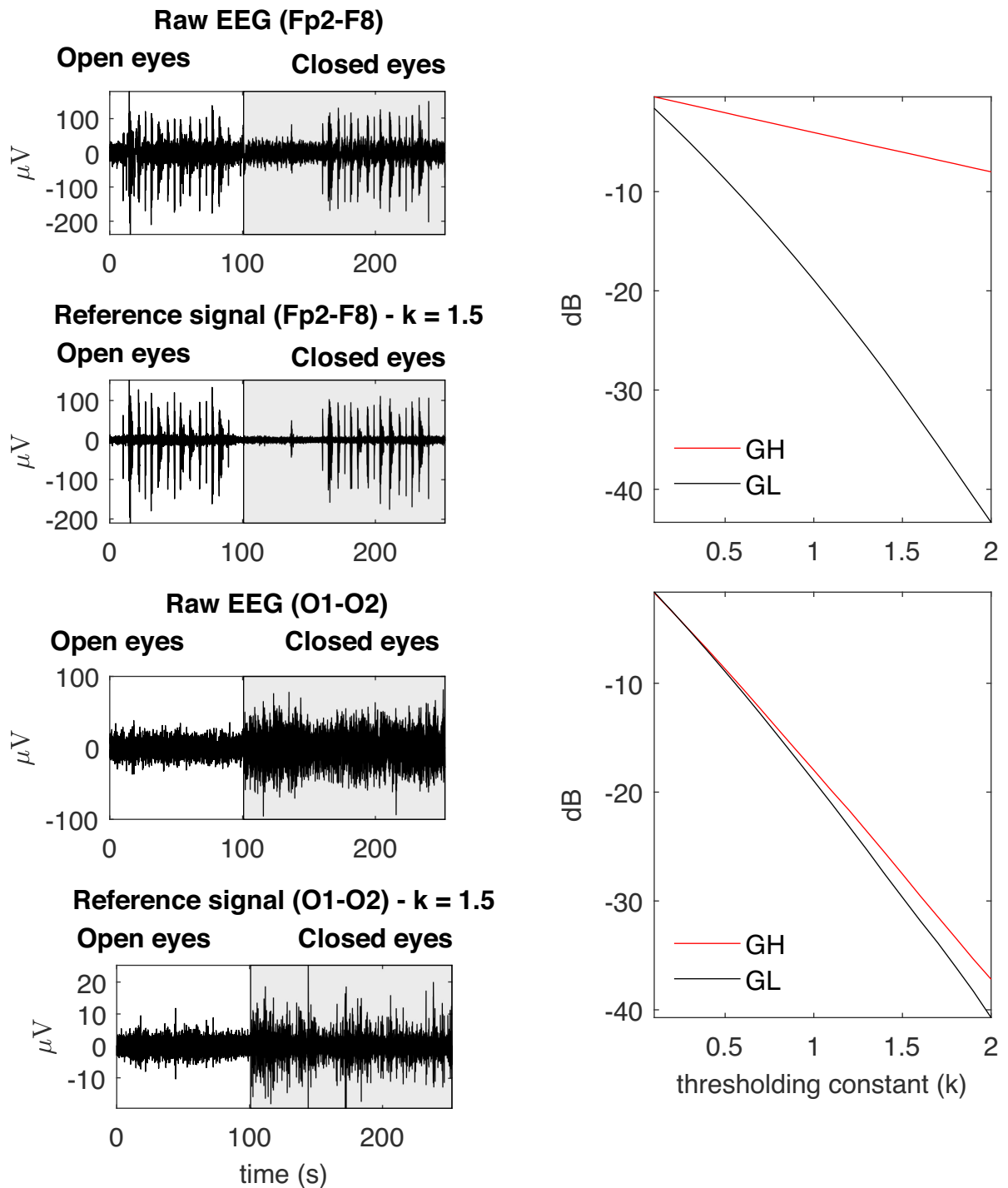


Figure 3.12: Examples of typical *GL* and *GH* feature vectors obtained for high (Fp2-F8) and low (O1-O2) levels of EMG contamination on EEG signals. By decomposing the raw EEG signals with EMD, the reference signals were obtained. The contamination level of the EEG signal can be captured by the distance between the feature vectors, in the sense that an increased distance is related to a lower signal to noise ratio. The examples demonstrate typical collected signals for the open and closed eyes scenarios.

signal, a substantially greater reduction of GL relative to GH is expected.

The box plots of GL and GH were generated, and their respective means were compared statistically. For this purpose, the statistical analysis followed the procedure stated previously, which comprised fitting an ANOVA model to the data, validating the method's assumptions, and employing an alternative non-parametric method if ANOVA was not appropriate.

3.5.3 Comparison of the performance of distinct adaptive filtering methods and experimental conditions

3.5.3.1 Evaluation based in time-domain features

The GL , GH , GX_{in} and GX_{out} features were evaluated for both the EMG ($EMGr$) and EEG ($EEGr$) reference signals. GL and GH were used to evaluate the performance of adaptive filtering for each type of reference signal. In contrast to decomposition methods, the appropriate adaptive filtering method should reduce the signal amplitude in regions contaminated by EMG while preserving signal amplitude in regions of EEG without EMG contamination. Therefore, the appropriate adaptive filtering strategy for reducing the influence of EMG on EEG is the one that produces a GL close to zero (to preserve the EEG signal) and GH less than zero (indicating the reduction of EMG contamination).

The feature GX_{in} and GX_{out} were employed to evaluate and compare the behaviour of the adaptive filtering in the regions of EEG with and without EMG. The lower the GX_{out} compared to GX_{in} , the greater the attenuation of the EEG regions contaminated by EMG. In addition, the closer to zero is GX_{out} ($GX_{out} \rightarrow 0$), the greater the capacity of the adaptive filtering method to preserve the amplitude of the EEG signal in EMG-contaminated regions.

Box plots were used to visually investigate the values of central trend, dispersion and symmetry of the characteristics for each of the scenarios investigated. To verify the differences between the characteristics estimated from different adaptive filtering methods, the means of the variables were compared by ANOVA. If the premises of such a model were not verified, a non-parametric approach was then employed, as previously explained.

3.5.3.2 Evaluation based in frequency-domain features

The non-filtered and filtered EEG signals were decomposed into their fundamental oscillations (Delta, Theta, Alpha, Beta, and Gamma) and box plots for the median frequency and its associated power were calculated. The study took into account the signal regions of the open and closed eye experimental conditions, as well as the overall signal that combines these two regions.

The normality of the variables were verified by the Kolmogorov–Smirnov test ($p > 0.05$), and then if the variables had a normal distribution the non-paired t-student test ($p < 0.05$) was applied to verify whether the mean of the variable related to the filtered signal reduced in relation to the non-filtered signal, i.e., the raw signal. If the distribution of the variables were not normal then the non-parametric Mann-Whitney U test was used ($p < 0.05$). Outliers were removed by eliminating observations that were outside of the following interval $[Q1 - 1.5IQR, Q3 + 1.5IQR]$, in which $Q1$ is the first quartile, $Q3$ the third quartile and IQR the interquartile range.

3.5.4 Comparison of execution time of decomposition and adaptive filtering methods

The execution time of the adaptive decomposition and filtering methods were evaluated by the means of the Matlab *timeit* function. This function performs multiple calls from the routine under analysis and returns the median value of the time measurements. The process for estimating the execution time employed actual and equal data for all methods. The evaluation considers samples with sizes from 25,000 to 1,000,000, with increments of 25,000, i.e. 40 different intervals. For each sample size, eight execution times were estimated to obtain a more representative estimate of the execution time.

The machine that processed the data had the following features: Ryzen 9 5900X 12-core/24-threads @3.7GHz; RAM 2x16GB DDR4 @ 3200MHz; video card Asus RTX 3070 8GB. The Matlab Parallel Computing Toolbox was used to run the applications in parallel, utilising all of the computer's processors and available memory.

The comparison of execution time was performed by the non-parametric Kruskal-Wallis test ($p < 0.05$) as the Shapiro-Wilk test confirmed the distribution of the vari-

ables were not normal ($p < 0.05$). The pairwise comparison between variables was performed by the Nemenyi test ($p < 0.05$).

Chapter 4

Results

4.1 Typical collected signals

Figure 4.1 shows typical EMG and EEG signals simultaneously collected during the experimental trials. The EMG bursts were detected by using the procedure described in Figure 3.9. The number of EMG bursts and their duration are in accordance to the protocol illustrated in Figure 3.3. The binary signals resulting from the EMG burst detection are plotted together with two typical EEG signals, one for the Fp2-F8 region, which is more contaminated, and the other for the O1-O2 region, which is less contaminated. Note that in the region when the eyes are closed (see the raw EEG: O1-O2) it is possible to see an increase in the amplitude of the EEG signal.

4.2 Characterization of the contamination of the EEG by EMG signals

Figures 4.2 and 4.3 depict topological maps for each subject and activated muscle. The visual inspection of the maps allow us to conclude that the Masseter is the muscle which produced the largest level of contamination, followed by the Frontalis and Zygomatic. These maps suggest also that the contamination and its spread over the brain map is dependent on the subject, which may be related to specific anatomical characteristics of the individual.

Figure 4.4 shows typical GL and GH feature vectors estimated for Subject 1. In

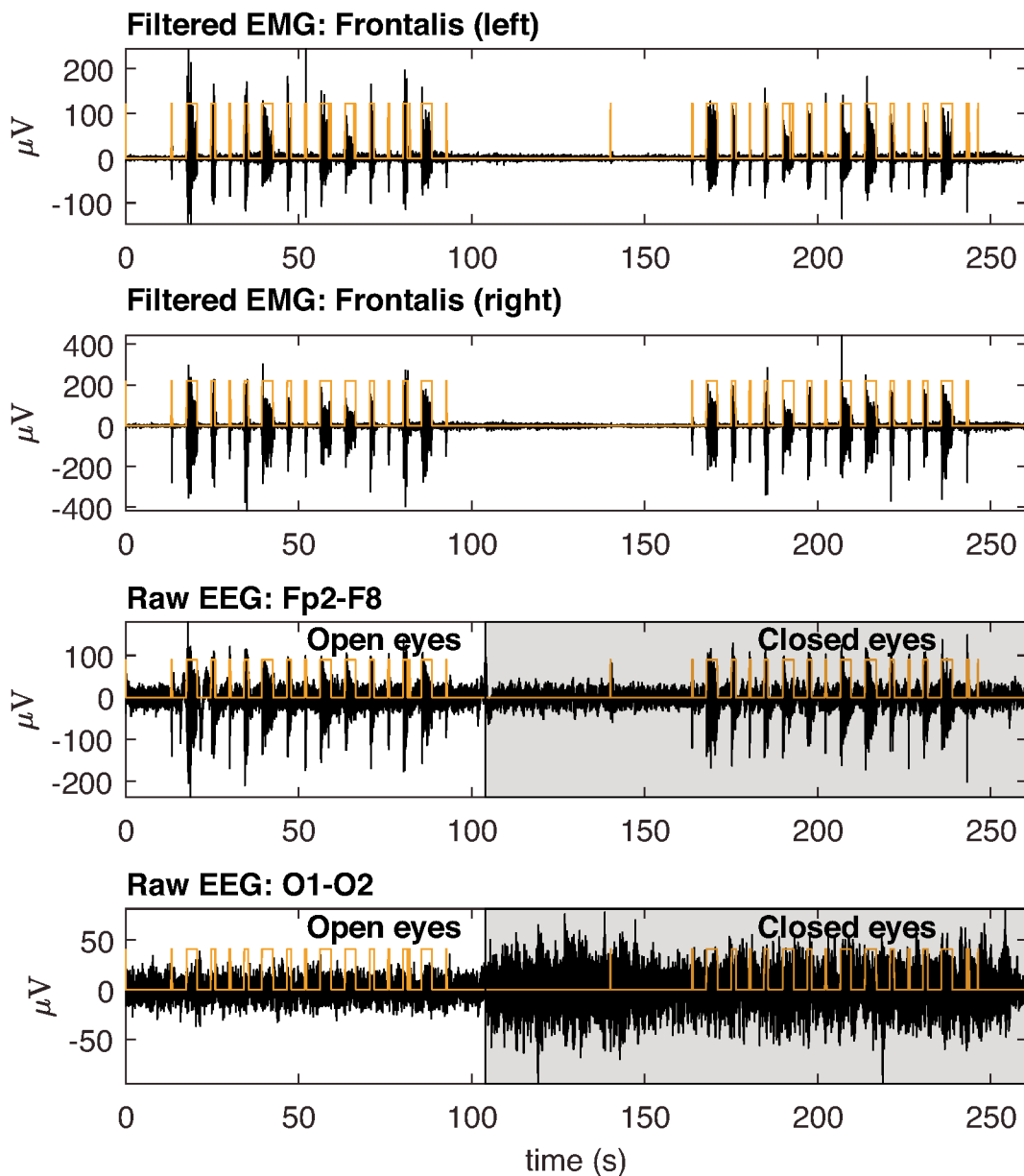


Figure 4.1: Typical EMG and EEG signals collected during the experimental trials. The EMG signals from the left and right Frontalis are shown. These signals were filtered to remove linear and non-linear trends. The EMG bursts were detected and then the binary signals oscillating from two levels were generated. Simultaneously collected EEG signals are shown for Fp2-F8 (high contamination) and O1-O2 (low contamination) locations. The binary signals are placed over the EEG signals for the indication of the periods in which there was EMG contamination.

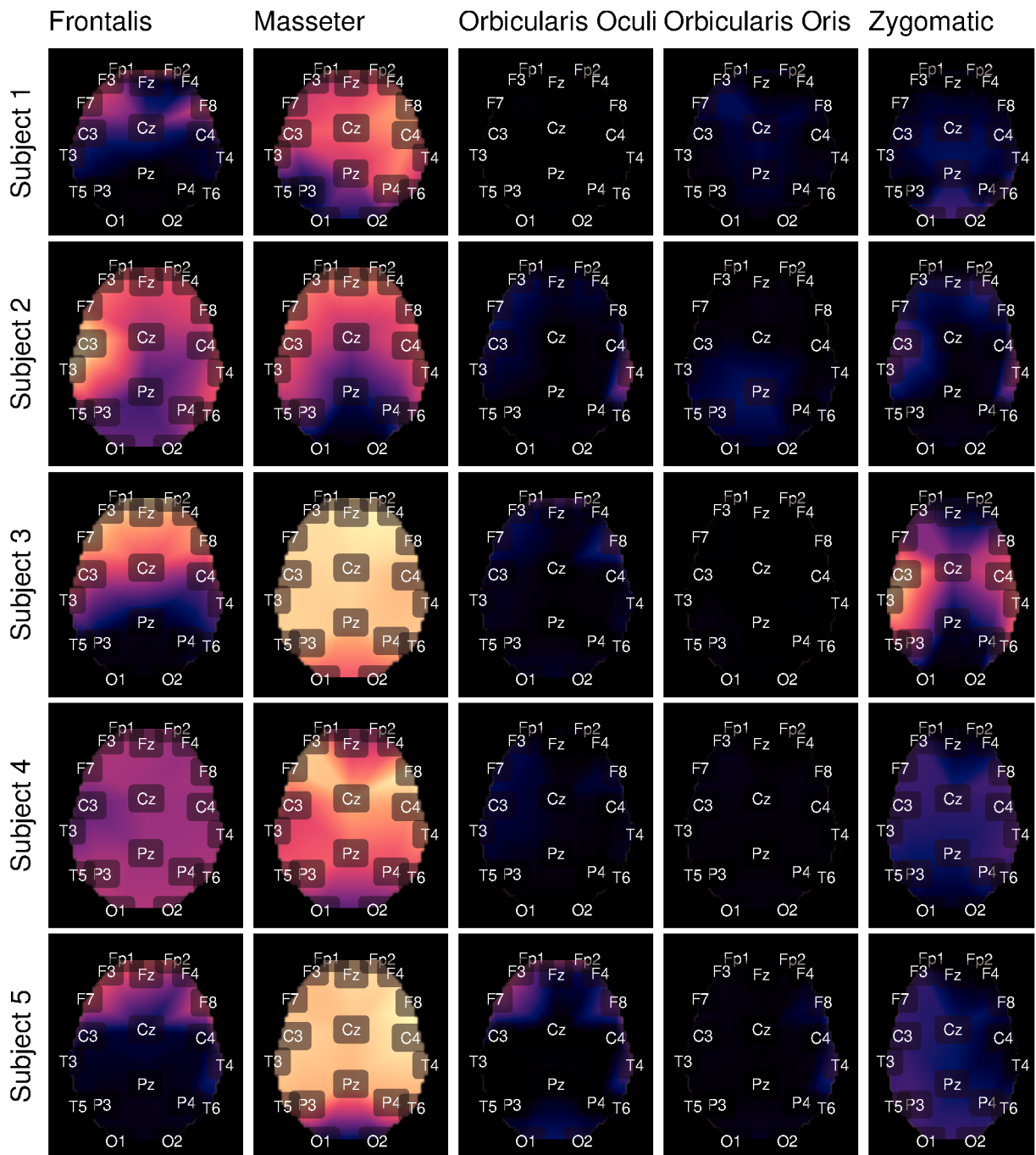


Figure 4.2: Using the international 10-20 system, the topological maps illustrate how the studied muscles contaminate distinct brain regions. Lighter colours represent the most contaminated locations, whereas darker colours denote the least contaminated areas. The presented results are for subjects from 1 to 5. The colours represent the similarity measure between the *GL* and *GH* features.

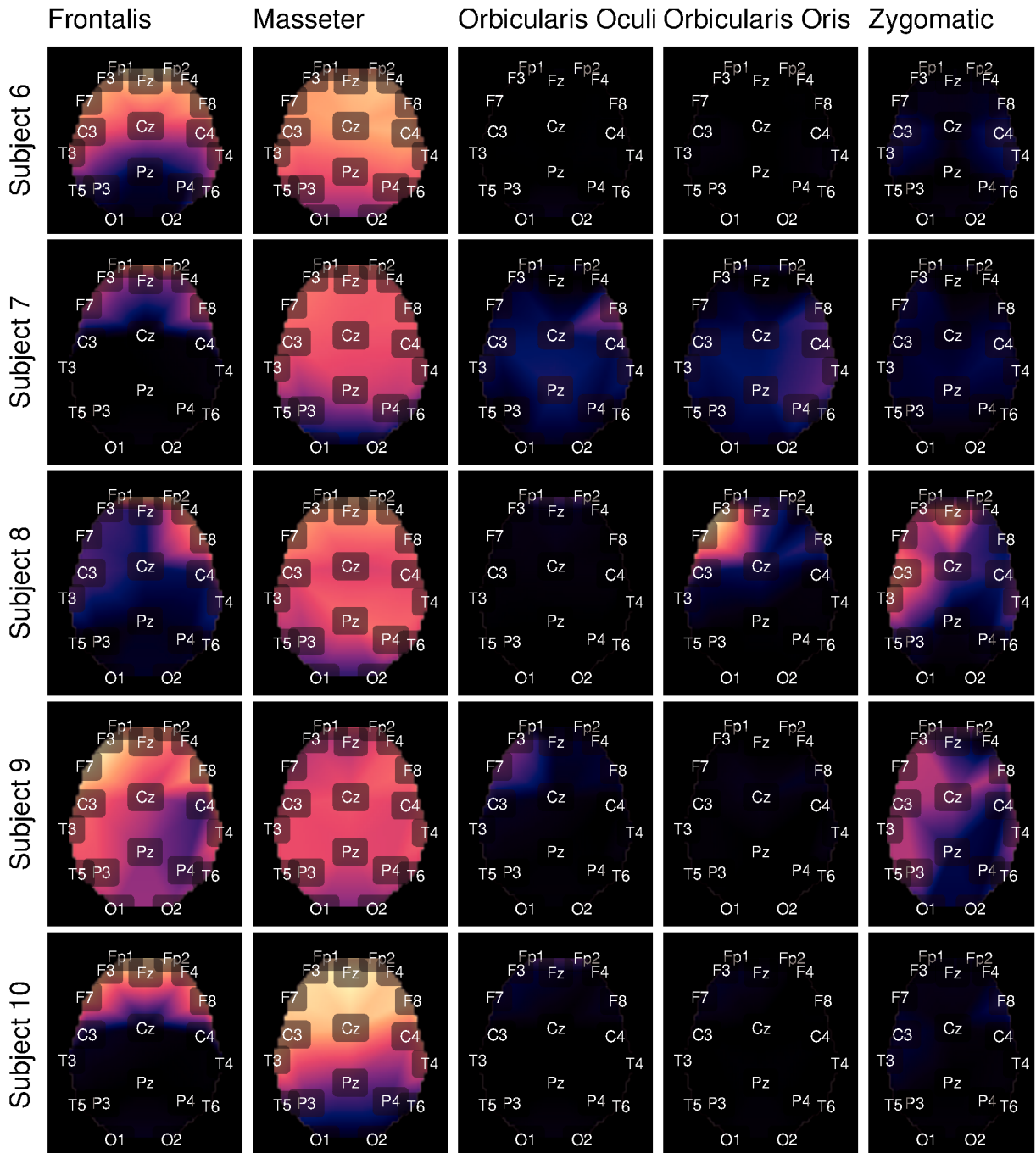


Figure 4.3: Topological maps for subjects from 6 to 10.

each graphic six pairs of feature vectors are presented. Each pair of feature vectors was estimated from a specific decomposition method. The behaviour of the feature vectors are similar for all subjects. The interpretation of the results is straightforward in the sense that the more similar the GL and GH feature vectors, the less contaminated the EEG signal is. For instance, for the occipital region (O1-O2) there is a high similarity (i.e., low distance) between the feature vectors for nearly all muscles, while for the

frontal region (e.g., Fp1-Fp2) the produced contamination is higher for the Frontalis and Masseter. The estimates of feature vectors for the other subjects are available in Appendix A (Figures A1 to A9).

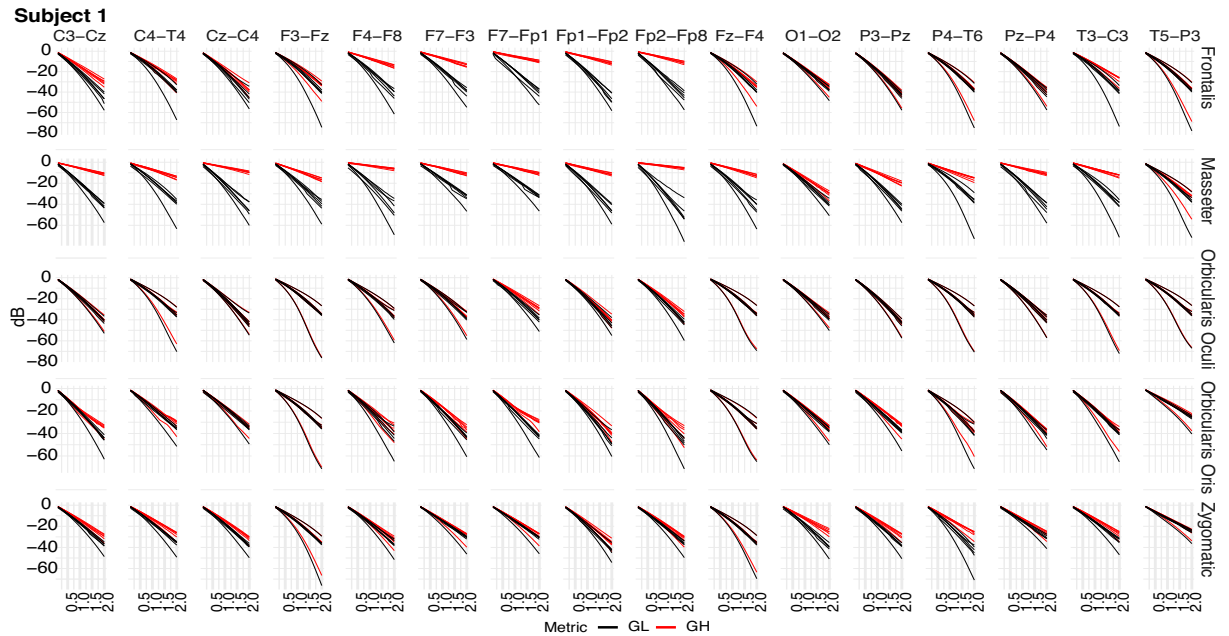


Figure 4.4: Typical GL and GH feature vectors estimated using different decomposition techniques for Subject 1. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles.

Box plots of the mean of the normalised Euclidean distance between GL and GH vectors for each muscle are shown in Figures 4.5 C and 4.6. Figure 4.5 C shows results that are independent of subjects and decomposition methods, whereas Figure 4.6 shows results that are dependent on subjects but independent of decomposition methods.

The results shown in Figures 4.5 C and 4.6 are consistent with those observed in Figures 4.2 and 4.3 (topological maps). In general, the Masseter was the muscle responsible for the highest EEG signal contamination, followed by the Frontalis and Zygomatic. The Frontalis exhibited the highest level of contamination variability, whereas the Orbicularis Oris produced the least. There was no significant difference between the mean normalised distances estimated from the Orbicularis Oculi and Orbicularis Oris muscles (Figure 4.5 C). Figure 4.6 shows that a similar result was found for all subjects.

ANOVA could not be employed for the statistical analysis since its assumptions were violated. To compare the variables, the Kruskal-Wallis rank sum test and Dunn's test for multiple comparisons were used in all analyses.

4.3 Evaluation of decomposition methods to generate reference signals for adaptive filtering

In Figures 4.5 A and B, the box plots of GL and GH are presented for each decomposition method, independently of subject and brain area. From the box plots, it is possible to compare and contrast the distributions of feature vectors for each investigated method. For GL the dashed lines indicate the median of the method which produces the largest amplitude reduction in the regions in which there is no EMG contamination. SSA was the most appropriate method among those considered.

For GH , the dashed lines represent the median of the approach that yields the lowest amplitude reduction in EMG-contaminated locations. Wavelet was the most appropriate method among those studied. In general, when both metrics, GL and GH , are considered, SSA is the most appropriate because it reduces the signal amplitude in regions without EMG contamination the most, while preserving the regions contaminated by EMG in a satisfactory manner, allowing the generation of an appropriate reference signal for adaptive filters. There was no statistically significant difference between the EEMD and Wavelet methods for GL (Figure 4.5 A). There were no significant differences between CiSSA and EMD, and EEMD and EMD-PCA for GH (Figure 4.5 B).

4.4 Evaluation of the filtering based on the time-domain features

Figure 4.7 depicts the overall behaviour of distinct adaptive filtering methods according to the time-domain features and type of reference signals (EEG or EMG). In general, all adaptive filtering methods were capable of filtering EMG contamination from EEG. As can be seen in Figures 4.7 B and D, the medians of GH are less than

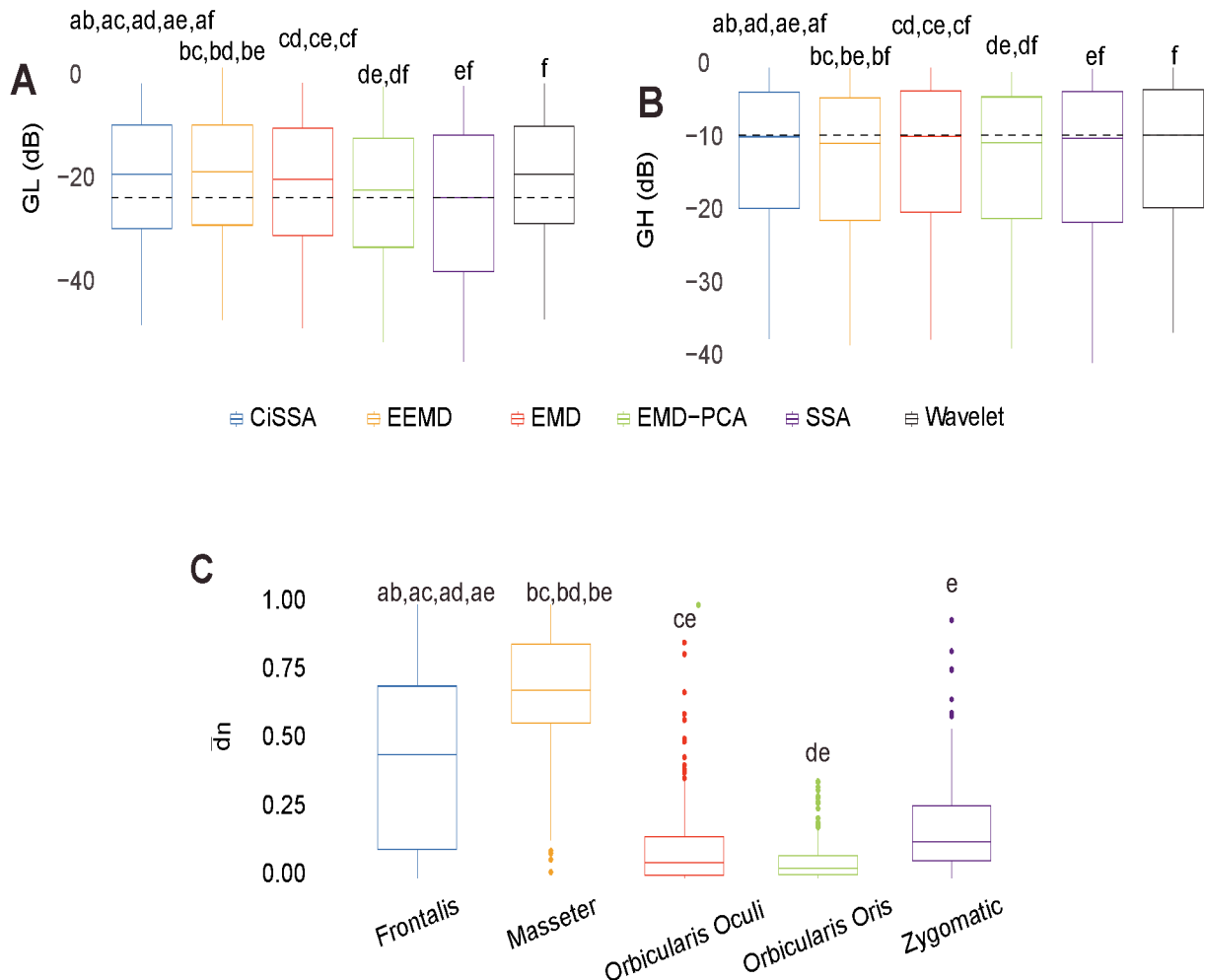


Figure 4.5: **(A)** Box plot of GL for distinct decomposition algorithms, regardless of subject and sensor location. The smaller the value of GL , the more appropriate the filtering method. The dashed lines represent the best result obtained with the SSA approach. Statistically significant differences between methods are represented by labels. All possible combinations of two were evaluated. For example, the EEMD method is represented by label “b” and was statistically different from CiSSA (label “a”), EMD (label “c”), EMD-PCA (label “d”) and SSA (label “e”). **(B)** Box plot of GH for distinct decomposition algorithms, regardless of subject and sensor location. The larger the value of GH the more suitable is the method for filtering. **(C)** Box plot of the mean normalized Euclidean distance between GL and GH for each muscle, independent of subjects and EEG sensors. The larger the value of this metric, the more contamination is caused by the muscle.

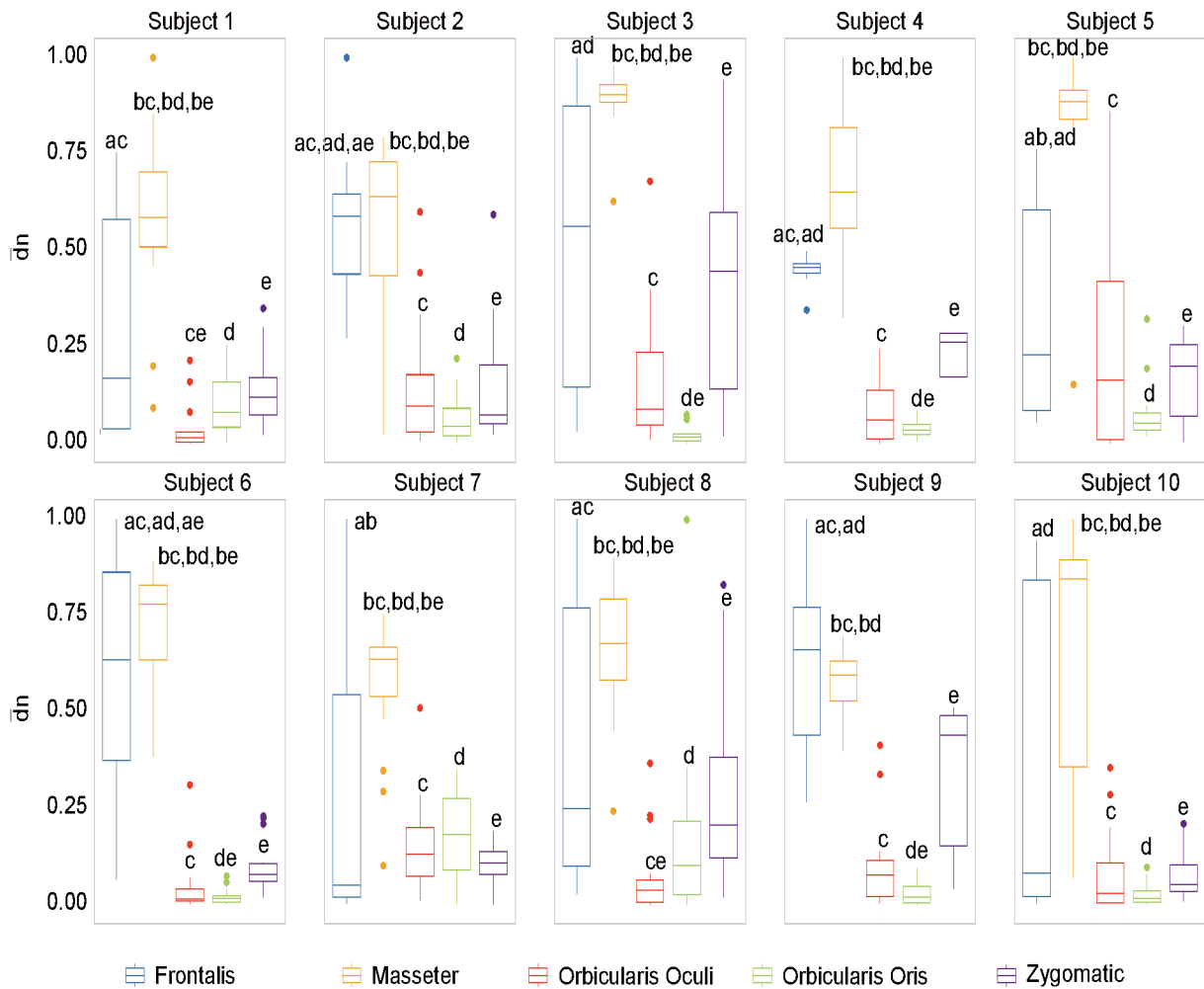


Figure 4.6: Box plot of the mean normalized Euclidean distance between GL and GH for each muscle and subject, independent of the EEG sensor.

zero, confirming, thus, that the EMG contamination was attenuated. On the other hand, when evaluating the preservation of regions of EEG without EMG, the results vary, depending on the type of reference. Ideally the median of GL should be as close to zero as possible. When the reference signal is the EEG, the most appropriate method is the RLS as it causes the lowest changes in the regions of EEG without EMG contamination. When the reference is the EMG signal, the most appropriate methods were the NLMS and Wiener. Figures 4.7 E and F show the behaviour of the variables GX_{in} and GX_{out} for adaptive filtering. For the EEG reference, the obtained results confirm the attenuation of the EMG signal, as all the medians of X_{out} are lower than the median of GX_{in} . On the other hand, for the EMG reference, the results yielded by the Wiener filter were not satisfactory. Considering the time-domain features, the EEG

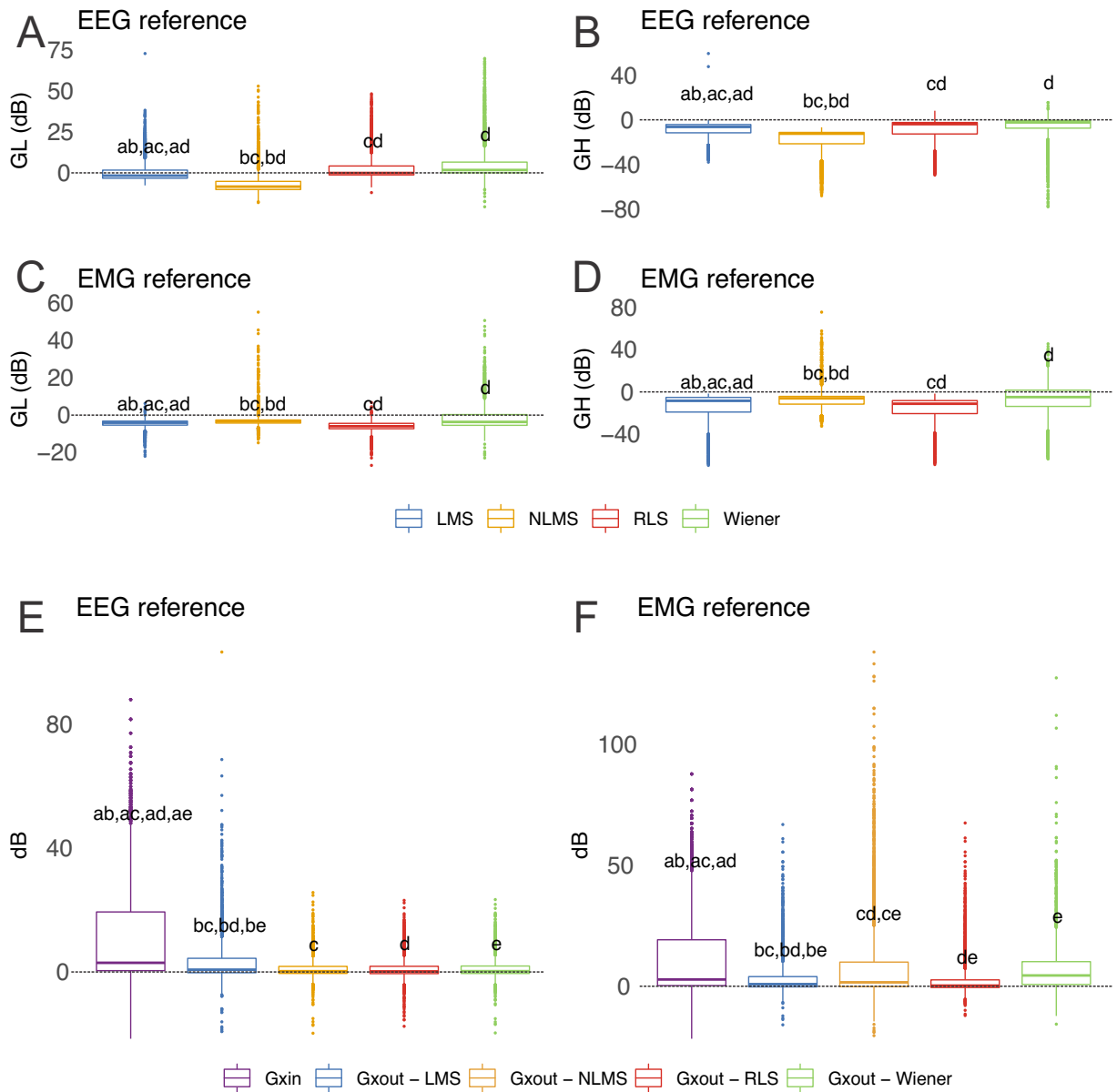


Figure 4.7: Evaluation of distinct adaptive filtering methods based on the time-domain features. The assessment is independent of the decomposition method and specific to the type of reference signal. **(A)**, **(B)**, **(C)** and **(D)** show results referent to GH and GL . **(E)** and **(F)** present the results related to GX_{in} and GX_{out} .

reference signal was more appropriate for the adaptive filtering, as it allowed for the preservation of EEG regions not contaminated by EMG. Furthermore, when evaluating the variables GX_{in} and GX_{out} it is clear that when using the EEG as a reference, the filtered EMG-contaminated region will preserve the EEG activity. Figure 4.8 shows an example of EMG-corrupted EEG and its filtered version. The figure insets show the effect of the adaptive filtering, in which the EMG amplitude is attenuated and the EEG

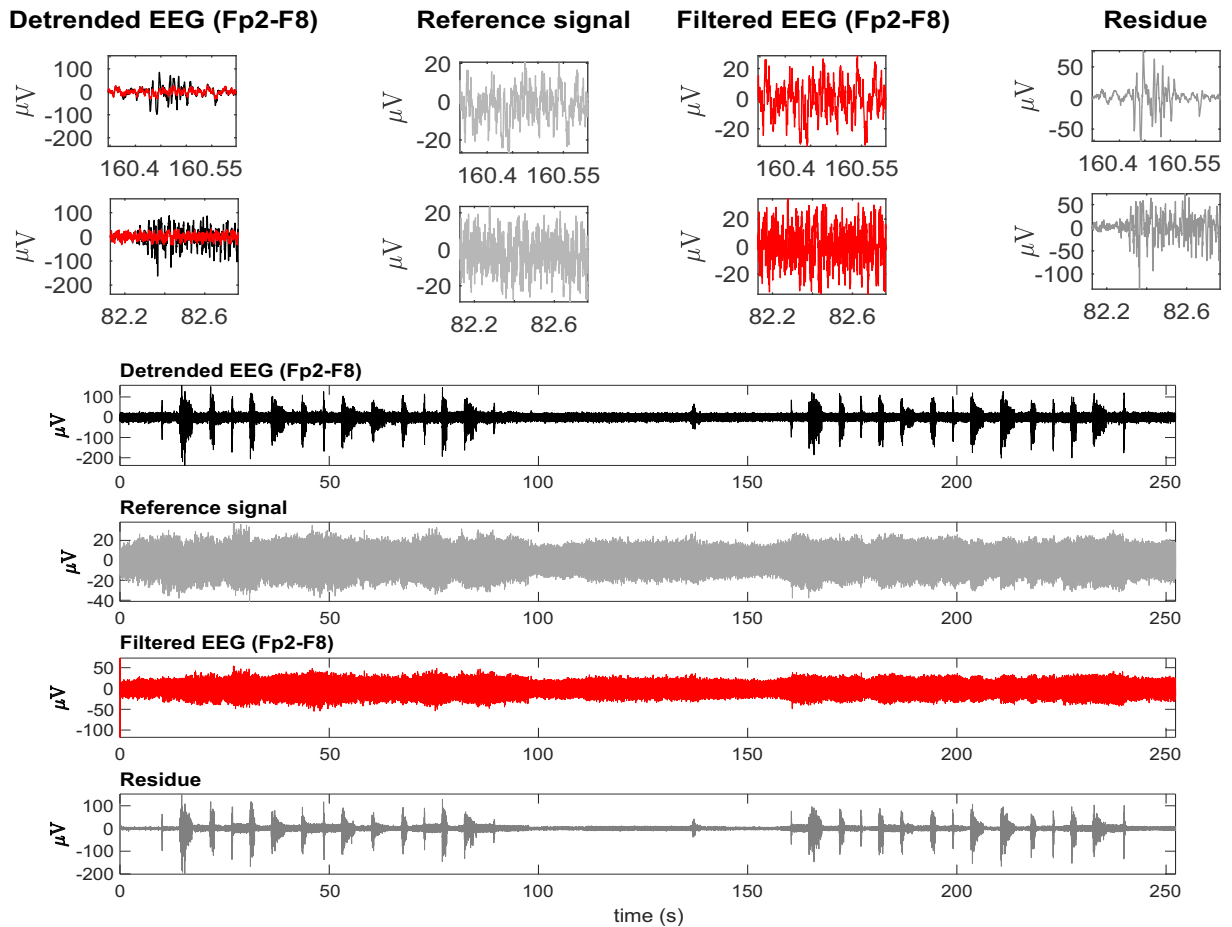


Figure 4.8: A typical EEG signal corrupted by facial EMG. The EEG signal is from the Fp2-F8 region because it is most affected by facial electromyography. EMG bursts can be seen on the detrended EEG signal. The EEG signal was used as a reference signal, estimated from EMD, and then filtered using the RLS filter. The residue, which is the difference between the detrended and filtered EEG data, clearly shows the EMG activity that was eliminated from the signal. The inset plots at the top indicate the selection of two EEG regions contaminated by EMG. The filtered signal is shown over the contaminated signal in red. For each region, the detrended EEG, reference signal, filtered EEG, and residue are shown.

activity in the EMG-contaminated region follows the EEG dynamics of the EEG signal in the neighbour regions in which there are no EMG contamination.

4.5 Evaluation of the filtering based on the frequency domain features

The box plots in Figure 4.9 show the behaviour of the median frequency and its power for the EEG signal together with its components, i.e., Delta (0.5-4 Hz), Theta (4-7 Hz), Alpha (7-13 Hz), Beta (13-30 Hz) and Gamma (30-70 Hz). The results contrast the raw non-filtered signal with the filtered signal. The non-parametric Mann-Whitney U test was used as the distributions of the variables were not normal ($p < 0.05$). There is a

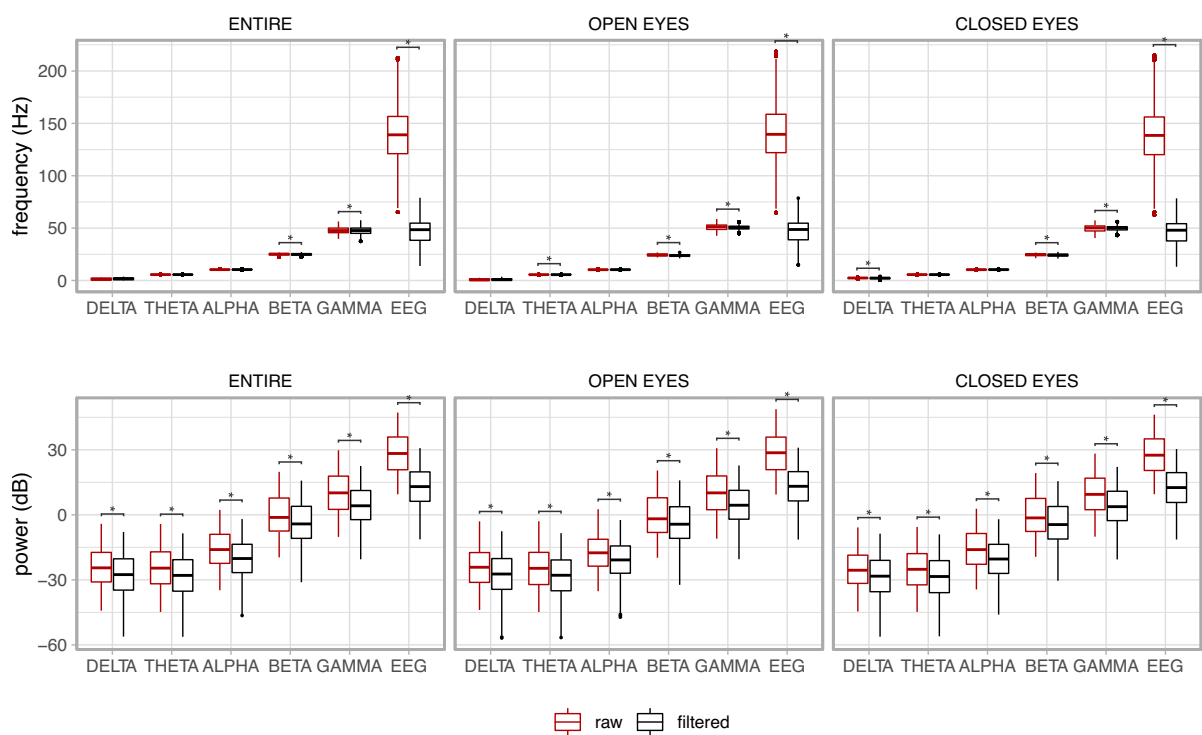


Figure 4.9: Median frequency and its power of the electroencephalogram (EEG) together with its components for the whole signal (ENTIRE) and the two experimental conditions (OPEN EYES and CLOSED EYES). A contrast between the raw non-filtered signal with the filtered signal is presented. The asterisks show the pair of variables in which the variable associated to the filtered signal was significantly reduced in comparison to the non-filtered signal, i.e., raw signal.

clear drop in the frequency and power of the EEG signal resulting from the filtering. As expected, the power of the signal is reduced for all components although their median frequencies are kept in the expected frequency band. The same behaviour is noted for

the experimental conditions of open and closed eyes, and for the whole signal, which considers the joint analysis of the open and closed eye regions.

4.6 Execution time of decomposition and adaptive filtering methods

Figure 4.10 depicts the typical execution times for all methods investigated. The estimates in A and B are based on the mean of the eight execution times for each sample size. The linear relationship between sample size and execution time may indicate that the methods have linear computational complexity, $O(n)$. According to the results in Figure 4.10 C, the execution times for EEMD and SSA are statistically equivalent (Nemenyi Test, $p > 0.05$), just as they are for EMD and CiSSA. Wavelet obtained the shortest execution time. According to the results in 4.10 D, the execution times for LMS and NLMS are statistically equivalent. Furthermore, RLS has a significantly longer execution time than the others.

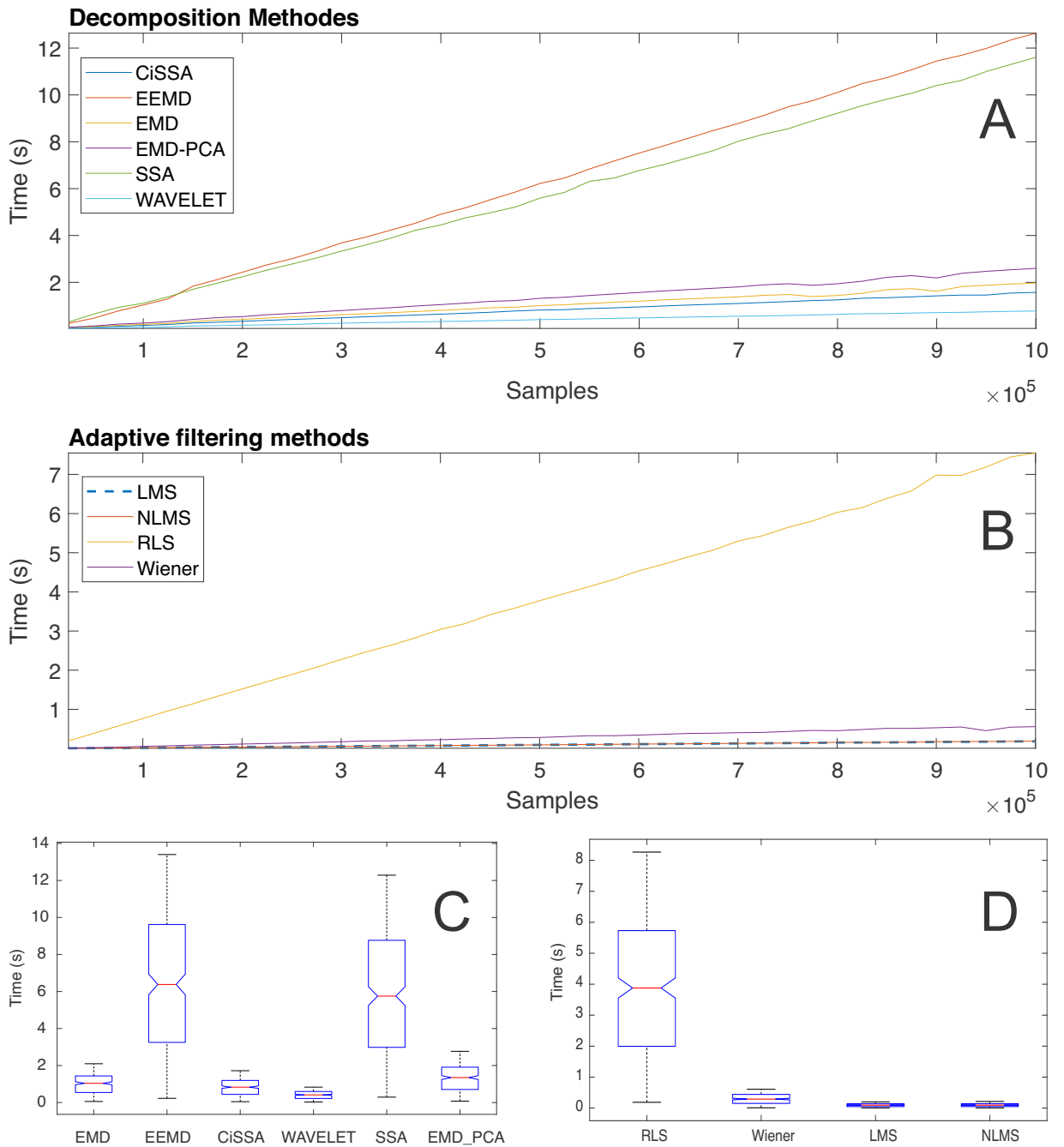


Figure 4.10: Execution time of distinct decomposition (A) and adaptive filtering (B) methods as function of the number of samples. The box plots show the central trend and dispersion of execution times (C and D).

Chapter 5

Discussion

The filtering approach depicted in Figure 3.7 has been proposed for the removal of facial muscular artefacts from EEG signals, which is an important requirement for distinct applications. It is a single channel approach for filtering low SNR EEG signals. In general, the proposed method is based on a hybrid filtering approach, combining adaptive filters with decomposition techniques. Thus, in this research, the performance of several decomposition (EMD, EEMD, CiSSA, Wavelet, SSA and EMD-PCA) and optimal filtering (RLS, Wiener, LMS and NLMS) methods were evaluated.

Although it is possible to find some databases containing EEG signals corrupted by EMG, we could not find any open data set similar to the one that was collected in this research. The relevance of the collected data set is that it considers the influence of the activity of several facial muscles to the contamination of EEG signals. The experimental protocol was carefully designed to take into consideration data collection in practical scenarios, such as the execution of facial expressions commonly used for some human-computer interfaces based on facial EMG [49]. In addition, all EEG signals were collected simultaneously with EMG signals guaranteeing the necessary synchronization between signals and the possibility of annotating the regions in which EEG signals were actually contaminated by EMG. The dataset included a total of 2,250 seconds of EEG signals corrupted by EMG, with the participation of 10 subjects and distinct experimental conditions (e.g., open and closed eyes, EMG bursts of varying durations, and the activation of different facial muscles), allowing for the required variability to test the performance of filtering methods.

Other decomposition and adaptive filtering methods can be added to the filter ar-

chitecture of the EEG single channel filtering approach depicted in Figure 3.7 without altering the entire filtering strategy. This is an interesting feature for the development of computational libraries that can benefit from the use of encapsulated code implementing decomposition and adaptive filtering methods that can be directly plugged into the general steps, i.e., the processing pipe depicted in Figure 3.7. To aid in the diffusion of this architecture, all Matlab and R scripts are available at <https://doi.org/10.5281/zenodo.6591866>. In addition, sample data and demonstration scripts are provided to facilitate comprehension and replication of the filtering approach presented and evaluated in this study.

In this study, the detection of EMG bursts is relevant because these bursts mark the regions contaminated and non-contaminated by electromyography automatically. The identification of these regions are used in the stage of soft-thresholding signal components and for computing the proposed set of features to measure the performance of filtering methods. We decided not to detect the bursts directly from the EEG contaminated signal to guarantee that the noise present in the EEG signal was really from the EMG activity. However, the use of the proposed filtering approach (Figure 3.7) can be applied without the simultaneously collection of EMG signals. If this is the case, it would be necessary to detect EMG bursts directly from the EMG-corrupted EEG.

The time and frequency domain features proposed in this work (Figure 3.10) were advantageous for the characterization of facial EMG contamination on the EEG. Figures 4.2 and 4.3 depict an approach for energy visualisation of topological maps displaying the degree of contamination created by distinct muscles based on this set of characteristics. Utilizing the normalised Euclidean distance as a measure of similarity allowed for the visual, qualitative, and quantitative comparison of topological maps estimated for various subjects and active muscles.

As indicated in Figure 3.7 an important step for adaptive filtering is the generation of reliable reference signals. In the proposed approach the reference signal is generated from the application of soft-thresholding to the signal components. To evaluate the performance of distinct decomposition methods the GL and GH metrics were proposed. These metrics were also employed to characterize the EMG contamination in distinct regions of the scalp (Figures 4.2 and 4.3). An interesting aspect of the contamination is that although there is a general pattern, e.g., muscles such as the Frontalis and Mas-

seter contributed more to the EEG contamination, the way this contamination spread over the scalp is specific to the individual. This fact can be verified by the relatively large variability of the variables presented in Figures 4.5 C and 4.6. These results regarding the higher contamination caused by the Frontalis and Masseter muscles are similar to the results obtained in [50] which estimated contamination by the normalized mean power metric.

According to the findings (Figures 4.2, 4.3, 4.5 and 4.6), the Masseter muscle provided the highest degree of contamination, followed by the Frontalis and Zygomatic. Variability between individuals was an important component of this study; for example, the Orbicularis Oris of subject 8 produced a substantial contamination of the EEG obtained in the frontal region. This may involve anatomical and behavioural characteristics of the individual. This requires that EEG filtering methods be devised to accommodate data variations introduced by anatomical, physiological, and experimental settings. This also justifies the more sophisticated experimental methodology utilised in this study.

In general, all decomposition methods investigated in this study were suitable for generating adequate adaptive filtering reference signals. Nonetheless, we believe that the SSA method is superior because it successfully preserved EEG in non-contaminated regions while lowering the signal amplitude in EMG-contaminated regions significantly (Figures 4.5 A and B). Considering the type of signal reference (EMG or EEG) the results reported in Figure 4.7 suggest that the reference based on the EEG signal is preferable because of the lower variability exhibited (Figure 4.7 E and F) when compared to the EMG reference.

This study also investigated various adaptive filtering algorithms (LMS, NLMS, RLS, and Wiener) to reduce electromyographic activity as much as possible in EEG data heavily contaminated by EMG. When using the EEG as a reference signal, RLS and NLMS were the best methods among those tested. In general, these methods best ensured: (i) attenuation of electromyographic activity in regions of EMG-contaminated EEG (Figure 4.7 B), (ii) preservation of electroencephalographic activity in regions of EEG without EMG (Figure 4.7 A), and (iii) the dynamics of the resulting signal, not mischaracterizing the EEG signal over time (Figure 4.7 and Figure 4.8). However, when the reference signal was the EMG, the LMS and RLS algorithms performed best. Thus,

the RLS method is the most preferred, as it produced satisfactory results regardless of the type of reference signal. The preference for the RLS filtering method is consistent with the result obtained in [50] that points out the EMDRLS method with advantages in both attenuation of EMG noise in the EEG signal and preservation of the EEG signal.

The results shown in Figure 4.9 confirm that the single channel approach proposed in this research was capable of reducing EMG contamination on the EEG signals, while preserving relevant information of the electroencephalogram. For instance, typical values of frequency were found for each EEG component. In addition, there was reduction in the power of the non-filtered EEG signal compared to its filtered version, and this could also be observed for each EEG component.

The results presented in Figure 4.10 reveal that there is a linear relationship between the number of samples, i.e., signal length, and the execution time required to process the data. This suggests that the decomposition and filtering algorithms have a linear computational complexity. In total, ninety hours were spent processing the entire data set of this investigation, taking into account the processing time for ten subjects and all experimental conditions. Methods that need less time to process data are desirable in this regard, even if they do not produce optimal results.

In future research, the proposed method could be utilised to reduce interference produced by electromyography in applications controlled by electroencephalographic activity (such as brain-computer interfaces). Furthermore, while this study only involves healthy people, we encourage the examination and confirmation of the technique created for people with disorders like amyotrophic lateral sclerosis. This would allow for the creation of more robust assistive technology as well as a better understanding of the electroencephalographic activity associated with this type of clinical condition.

Chapter 6

Conclusion

This study introduced a single channel filtering method for reducing facial electromyography from EEG signals. The proposed method is sufficiently general to accommodate multiple decomposition and adaptive filtering techniques within a single architecture.

Using a data set that enabled the generation of EMG-corrupted EEG in experiments involving facial muscular activation, the filtering method was evaluated. The set of time and frequency domain characteristics enabled the visualisation and quantification of facial EMG contamination of the EEG. This set of features allowed for comparative analysis of filtering methods.

The results indicated that the Masseter was the muscle that contaminated the EEG the most; however, individual variation should not be disregarded, as the contraction of other facial muscles in some people may generate significant contamination on EEG signals.

In general, all investigated decomposition and adaptive filtering methods effectively filtered facial EMG-corrupted EEG; however, the decomposition method SSA reduced EMG contamination while preserving the EEG signal more. This method's relative slowness in comparison to other studies is its most significant drawback. In terms of the adaptive filtering method, it was observed that the reference signal (EMG or EEG) affects the method's performance, despite the methods' similarities.

References

- [1] J. J. Foxe and A. C. Snyder, "The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention," *Frontiers in psychology*, vol. 2, p. 154, 2011. <https://doi.org/10.3389/fpsyg.2011.00154>.
- [2] M. Abo-Zahhad, S. M. Ahmed, and S. N. Abbas, "A new eeg acquisition protocol for biometric identification using eye blinking signals," *International Journal of Intelligent Systems and Applications*, vol. 7, no. 6, p. 48, 2015. <https://doi.org/10.5815/ijisa.2015.06.05>.
- [3] V. Mihajlovic, B. Grundlehner, R. Vullers, and J. Penders, "Wearable, wireless EEG solutions in daily life applications: What are we missing?," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, pp. 6–21, jan 2015. <https://doi.org/10.1109/jbhi.2014.2328317>.
- [4] R. A. Ramadan and A. V. Vasilakos, "Brain computer interface: control signals review," *Neurocomputing*, vol. 223, pp. 26–44, feb 2017. <https://doi.org/10.1016/j.neucom.2016.10.024>.
- [5] A. Hero, "Signal processing identity [president's message]," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 4–4, 2006. <https://doi.org/10.1109/msp.2006.1657808>.
- [6] K. T. Sweeney, T. E. Ward, and S. F. McLoone, "Artifact removal in physiological signals—practices and possibilities," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, pp. 488–500, may 2012. <https://doi.org/10.1109/titb.2012.2188536>.
- [7] J. A. Urigüen and B. Garcia-Zapirain, "EEG artifact removal—state-of-the-art and

- guidelines,” *Journal of Neural Engineering*, vol. 12, p. 031001, apr 2015. <https://doi.org/10.1088/1741-2560/12/3/031001>.
- [8] P. L. Nunez and R. Srinivasan, *Electric Fields of the Brain*. Oxford University Press, jan 2006. <https://doi.org/10.1093/acprof:oso/9780195050387.001.0001>.
- [9] W. O. Tatum, B. A. Dworetzky, and D. L. Schomer, “Artifact and recording concepts in EEG,” *Journal of Clinical Neurophysiology*, vol. 28, pp. 252–263, jun 2011. <https://doi.org/10.1097/wnp.0b013e31821c3c93>.
- [10] D. Safieddine, A. Kachenoura, L. Albera, G. Birot, A. Karfoul, A. Pasnicu, A. Biraben, F. Wendling, L. Senhadji, and I. Merlet, “Removal of muscle artifact from EEG data: comparison between stochastic (ICA and CCA) and deterministic (EMD and wavelet-based) approaches,” *EURASIP Journal on Advances in Signal Processing*, vol. 2012, jul 2012. <https://doi.org/10.1186/1687-6180-2012-127>.
- [11] F. Gabsteiger, H. Leutheuser, P. Reis, M. Lochmann, and B. M. Eskofier, “ICA-based reduction of electromyogenic artifacts in EEG data: Comparison with and without EMG data,” in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, aug 2014. <https://doi.org/10.1109/embc.2014.6944466>.
- [12] V. Bono, S. Das, W. Jamal, and K. Maharatna, “Hybrid wavelet and EMD/ICA approach for artifact suppression in pervasive EEG,” *Journal of Neuroscience Methods*, vol. 267, pp. 89–107, jul 2016. <https://doi.org/10.1016/j.jneumeth.2016.04.006>.
- [13] R. Upadhyay, P. Padhy, and P. Kankar, “EEG artifact removal and noise suppression by discrete orthonormal s-transform denoising,” *Computers and Electrical Engineering*, vol. 53, pp. 125–142, jul 2016. <https://doi.org/10.1016/j.compeleceng.2016.05.015>.
- [14] L. Frølich and I. Dowding, “Removal of muscular artifacts in EEG signals: a comparison of linear decomposition methods,” *Brain Informatics*, vol. 5, pp. 13–22, jan 2018. <https://doi.org/10.1007/s40708-017-0074-6>.

- [15] J. A. Mucarquer, P. Prado, M.-J. Escobar, W. El-Deredy, and M. Zanartu, "Improving EEG muscle artifact removal with an EMG array," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, pp. 815–824, mar 2020. <https://doi.org/10.1109/tim.2019.2906967>.
- [16] B. W. McMenamin, A. J. Shackman, L. L. Greischar, and R. J. Davidson, "Electromyogenic artifacts and electroencephalographic inferences revisited," *NeuroImage*, vol. 54, pp. 4–9, jan 2011. <https://doi.org/10.1016/j.neuroimage.2010.07.057>.
- [17] R. T. Pivik, R. J. Broughton, R. Coppola, R. J. Davidson, N. Fox, and M. R. Nuwer, "Guidelines for the recording and quantitative analysis of electroencephalographic activity in research contexts," *Psychophysiology*, vol. 30, no. 6, pp. 547–558, 1993. <https://doi.org/10.1111/j.1469-8986.1993.tb02081.x>.
- [18] M. Mamun, M. Al-Kadi, and M. Marufuzzaman, "Effectiveness of wavelet denoising on electroencephalogram signals," *Journal of Applied Research and Technology*, vol. 11, pp. 156–160, feb 2013. [https://doi.org/10.1016/s1665-6423\(13\)71524-4](https://doi.org/10.1016/s1665-6423(13)71524-4).
- [19] I. Goncharova, D. McFarland, T. Vaughan, and J. Wolpaw, "EMG contamination of EEG: spectral and topographical characteristics," *Clinical Neurophysiology*, vol. 114, pp. 1580–1593, sep 2003. [https://doi.org/10.1016/s1388-2457\(03\)00093-2](https://doi.org/10.1016/s1388-2457(03)00093-2).
- [20] A. G. Correa, E. Laciari, H. D. Patiño, and M. E. Valentinuzzi, "Artifact removal from EEG signals using adaptive filters in cascade," *Journal of Physics: Conference Series*, vol. 90, p. 012081, nov 2007. <https://doi.org/10.1088/1742-6596/90/1/012081>.
- [21] R. Kher and R. Gandhi, "Adaptive filtering based artifact removal from electroencephalogram (EEG) signals," in *2016 International Conference on Communication and Signal Processing (ICCSP)*, IEEE, apr 2016. <https://doi.org/10.1109/iccsp.2016.7754202>.
- [22] P. S. R. Diniz, *Adaptive Filtering*, vol. 4. Springer US, 2008. <https://doi.org/10.1007/978-0-387-68606-6>.

- [23] H. Maki, T. Toda, S. Sakti, G. Neubig, and S. Nakamura, "EEG signal enhancement using multi-channel wiener filter with a spatial correlation prior," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, apr 2015. <https://doi.org/10.1109/icassp.2015.7178449>.
- [24] J. Ferdous and S. Ali, "A comparison of wiener and kalman filters for the artifact suppression from eeg signal," *International Journal of Science and Research*, vol. 6, pp. 2029–2035, Apr. 2017.
- [25] F. Morbidi, A. Garulli, D. Prattichizzo, C. Rizzo, and S. Rossi, "Application of kalman filter to remove TMS-induced artifacts from EEG recordings," *IEEE Transactions on Control Systems Technology*, vol. 16, pp. 1360–1366, nov 2008. <https://doi.org/10.1109/tcst.2008.921814>.
- [26] C. J. James and C. W. Hesse, "Independent component analysis for biomedical signals," *Physiological Measurement*, vol. 26, pp. R15–R39, dec 2004. <https://doi.org/10.1088/0967-3334/26/1/r02>.
- [27] L. Albera, A. Kachenoura, P. Comon, A. Karfoul, F. Wendling, L. Senhadji, and I. Merlet, "ICA-based EEG denoising: a comparative analysis of fifteen methods," *Bulletin of the Polish Academy of Sciences: Technical Sciences*, vol. 60, pp. 407–418, dec 2012. <https://doi.org/10.2478/v10175-012-0052-3>.
- [28] P. P. Ngoc, V. D. Hai, N. C. Bach, and P. V. Binh, "EEG signal analysis and artifact removal by wavelet transform," in *IFMBE Proceedings*, pp. 179–183, Springer International Publishing, 2015. https://doi.org/10.1007/978-3-319-11776-8_44.
- [29] A. Turnip and J. Pardede, "Artefacts removal of EEG signals with wavelet denoising," *MATEC Web of Conferences*, vol. 135, p. 00058, 2017. <https://doi.org/10.1051/matecconf/201713500058>.
- [30] P. Gaur, R. B. Pachori, H. Wang, and G. Prasad, "An empirical mode decomposition based filtering method for classification of motor-imagery EEG signals for enhancing brain-computer interface," in *2015 International Joint Conference on Neural Networks (IJCNN)*, IEEE, jul 2015. <https://doi.org/10.1109/ijcnn.2015.7280754>.

- [31] M. E. Alam and B. Samanta, "Performance evaluation of empirical mode decomposition for EEG artifact removal," in *Volume 4B: Dynamics, Vibration, and Control*, American Society of Mechanical Engineers, nov 2017. <https://doi.org/10.1115/imece2017-71647>.
- [32] X. Chen, C. He, and H. Peng, "Removal of muscle artifacts from single-channel EEG based on ensemble empirical mode decomposition and multiset canonical correlation analysis," *Journal of Applied Mathematics*, vol. 2014, pp. 1–10, 2014. <https://doi.org/10.1155/2014/261347>.
- [33] I. Daly, R. Scherer, M. Billinger, and G. Muller-Putz, "FORCe: Fully online and automated artifact removal for brain-computer interfacing," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, pp. 725–736, sep 2015. <https://doi.org/10.1109/tnsre.2014.2346621>.
- [34] S. Salsabili, S. H. Sardoui, and M. B. Shamsollahi, "Interictal EEG denoising using independent component analysis and empirical mode decomposition," in *2015 38th International Conference on Telecommunications and Signal Processing (TSP)*, IEEE, jul 2015. <https://doi.org/10.1109/tsp.2015.7296475>.
- [35] K. Zeng, D. Chen, G. Ouyang, L. Wang, X. Liu, and X. Li, "An EEMD-ICA approach to enhancing artifact rejection for noisy multivariate neural data," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, pp. 630–638, jun 2016. <https://doi.org/10.1109/tnsre.2015.2496334>.
- [36] M. Kim and S.-P. Kim, "A comparison of artifact rejection methods for a BCI using event related potentials," in *2018 6th International Conference on Brain-Computer Interface (BCI)*, IEEE, jan 2018. <https://doi.org/10.1109/iww-bci.2018.8311530>.
- [37] J. Gerardo, J. C., and J. Velazquez, "Applications of adaptive filtering," in *Adaptive Filtering Applications*, InTech, jun 2011. <https://doi.org/10.5772/16873>.
- [38] S. K. Sengupta and S. M. Kay, "Fundamentals of statistical signal processing: Estimation theory," *Technometrics*, vol. 37, p. 465, nov 1995. <https://doi.org/10.2307/1269750>.

- [39] J. Gao, H. Sultan, J. Hu, and W.-W. Tung, "Denoising nonlinear time series by adaptive filtering and wavelet shrinkage: A comparison," *IEEE Signal Processing Letters*, vol. 17, pp. 237–240, mar 2010. <https://doi.org/10.1109/lsp.2009.2037773>.
- [40] D. Iyer and G. Zouridakis, "Single-trial evoked potential estimation: Comparison between independent component analysis and wavelet denoising," *Clinical Neurophysiology*, vol. 118, pp. 495–504, mar 2007. <https://doi.org/10.1016/j.clinph.2006.10.024>.
- [41] V. Krishnaveni, S. Jayaraman, L. Anitha, and K. Ramadoss, "Removal of ocular artifacts from EEG using adaptive thresholding of wavelet coefficients," *Journal of Neural Engineering*, vol. 3, pp. 338–346, nov 2006. <https://doi.org/10.1088/1741-2560/3/4/011>.
- [42] D. Rutledge and D. J.-R. Bouveresse, "Independent components analysis with the JADE algorithm," *TrAC Trends in Analytical Chemistry*, vol. 50, pp. 22–32, oct 2013. <https://doi.org/10.1016/j.trac.2013.03.013>.
- [43] C. Zhang, J. Yang, Y. Lei, and F. Ye, "Single channel blind source separation by combining slope ensemble empirical mode decomposition and independent component analysis," *Journal of Computational Information Systems*, vol. 8, pp. 3117–3126, Apr. 2012.
- [44] K. Sweeney, *Motion artifact processing techniques for physiological signals*. phdthesis, Faculty of Science and Engineering, Department of Electronic Engineering, National University of Ireland Maynooth, Ireland, Apr. 2013.
- [45] C. M. Sweeney-Reed, S. J. Nasuto, M. F. Vieira, and A. O. Andrade, "Empirical mode decomposition and its extensions applied to EEG analysis: A review," *Advances in Data Science and Adaptive Analysis*, vol. 10, p. 1840001, apr 2018. <https://doi.org/10.1142/s2424922x18400016>.
- [46] N. P. Castellanos and V. A. Makarov, "Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis," *Journal of Neuroscience Methods*, vol. 158, pp. 300–312, dec 2006. <https://doi.org/10.1016/j.jneumeth.2006.05.033>.

- [47] M. M. N. Mannan, M. Y. Jeong, and M. A. Kamran, "Hybrid ICA—regression: Automatic identification and removal of ocular artifacts from electroencephalographic signals," *Frontiers in Human Neuroscience*, vol. 10, may 2016. <https://doi.org/10.3389/fnhum.2016.00193>.
- [48] X. Chen, A. Liu, J. Chiang, Z. J. Wang, M. J. McKeown, and R. K. Ward, "Removing muscle artifacts from EEG data: Multichannel or single-channel techniques?," *IEEE Sensors Journal*, vol. 16, pp. 1986–1997, apr 2016. <https://doi.org/10.1109/jsen.2015.2506982>.
- [49] A. O. Andrade, A. A. Pereira, C. G. P. Jr, and P. J. Kyberd, "Mouse emulation based on facial electromyogram," *Biomedical Signal Processing and Control*, vol. 8, pp. 142–152, mar 2013. <https://doi.org/10.1016/j.bspc.2012.09.001>.
- [50] G. M. d. Silva, *Caracterização e filtragem de eletroencefalograma contaminado por eletromiografia dos músculos faciais*. PhD thesis, UFU, 2020.
- [51] E. Kutluay and G. P. Kalamangalam, "Montages for noninvasive EEG recording," *Journal of Clinical Neurophysiology*, vol. 36, pp. 330–336, sep 2019. <https://doi.org/10.1097/wnp.0000000000000546>.
- [52] B. Moon, "Sampling rates, aliasing, and the analysis of electrophysiological signals," in *Proceedings of the 1996 Fifteenth Southern Biomedical Engineering Conference*, pp. 401–404, IEEE, 1996. <https://doi.org/10.1109/SBEC.1996.493260>.
- [53] R. R. Harrison, "A versatile integrated circuit for the acquisition of biopotentials," in *2007 IEEE Custom Integrated Circuits Conference*, pp. 115–122, IEEE, 2007. <https://doi.org/10.1109/CICC.2007.4405694>.
- [54] A. C. N. Society *et al.*, "A proposal for standard montages to be used in clinical eeg," <http://www.acns.org/pdf/guidelines/Guideline-6.pdf>, 2006.
- [55] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal*

- Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, pp. 903–995, mar 1998. <https://doi.org/10.1098/rspa.1998.0193>.
- [56] Z. Wu and N. E. Huang, “Ensemble empirical mode decomposition: a noise-assisted data analysis method,” *Advances in Adaptive Data Analysis*, vol. 01, pp. 1–41, jan 2009. <https://doi.org/10.1142/s1793536909000047>.
- [57] J. Bógalo, P. Poncela, and E. Senra, “Circulant singular spectrum analysis: A new automated procedure for signal extraction,” *Signal Processing*, vol. 179, p. 107824, feb 2021. <https://doi.org/10.1016/j.sigpro.2020.107824>.
- [58] A. O. Andrade, S. Nasuto, P. Kyberd, C. M. Sweeney-Reed, and F. V. Kaniijn, “EMG signal filtering based on empirical mode decomposition,” *Biomedical Signal Processing and Control*, vol. 1, pp. 44–55, jan 2006. <https://doi.org/10.1016/j.bspc.2006.03.003>.
- [59] B. Farhang-Boroujeny, *Adaptive Filters Theory and Applications*. Wiley, 1999.
- [60] R Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021.
- [61] P. Barrett, “Euclidean distance: raw, normalized, and double-scaled coefficients,” Tech. Rep. 6, 9 2005. <https://www.pbarrett.net/techpapers/euclid.pdf>.
- [62] M. Okamoto, H. Dan, K. Sakamoto, K. Takeo, K. Shimizu, S. Kohno, I. Oda, S. Isobe, T. Suzuki, K. Kohyama, and I. Dan, “Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10–20 system oriented for transcranial functional brain mapping,” *NeuroImage*, vol. 21, pp. 99–111, jan 2004. <https://doi.org/10.1016/j.neuroimage.2003.08.026>.
- [63] D. M. Wolpert, Z. Ghahramani, and M. I. Jordan, “An internal model for sensorimotor integration,” *Science*, vol. 269, no. 5232, pp. 1880–1882, 1995. <https://doi.org/10.1126/science.7569931>.
- [64] C. G. Atkeson *et al.*, “Learning arm kinematics and dynamics,” *Annual review of neuroscience*, vol. 12, no. 1, pp. 157–183, 1989. <https://doi.org/10.1146/annurev.ne.12.030189.001105>.

- [65] B. Preilowski, *Phases of motor-skills acquisition: a neuropsychological approach*. Univ., 1977. <https://doi.org/10.1037/e668292012-002>.
- [66] U. Halsband and R. K. Lange, "Motor learning in man: a review of functional and clinical studies," *Journal of Physiology-Paris*, vol. 99, no. 4-6, pp. 414–424, 2006. <https://doi.org/10.1016/j.jphysparis.2006.03.007>.
- [67] B. Pollok, D. Latz, V. Krause, M. Butz, and A. Schnitzler, "Changes of motor-cortical oscillations associated with motor learning," *Neuroscience*, vol. 275, pp. 47–53, 2014. <https://doi.org/10.1016/j.neuroscience.2014.06.008>.
- [68] K. J. Pope, S. P. Fitzgibbon, T. W. Lewis, E. M. Whitham, and J. O. Willoughby, "Relation of gamma oscillations in scalp recordings to muscular activity," *Brain topography*, vol. 22, no. 1, pp. 13–17, 2009. <https://doi.org/10.1007/s10548-009-0081-x>.
- [69] E. Felton, R. Radwin, J. Wilson, and J. Williams, "Evaluation of a modified fitts law brain–computer interface target acquisition task in able and motor disabled individuals," *Journal of neural engineering*, vol. 6, no. 5, p. 056002, 2009. <https://doi.org/10.1088/1741-2560/6/5/056002>.
- [70] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement.," *Journal of experimental psychology*, vol. 47, no. 6, p. 381, 1954. <https://doi.org/10.1037/h0055392>.

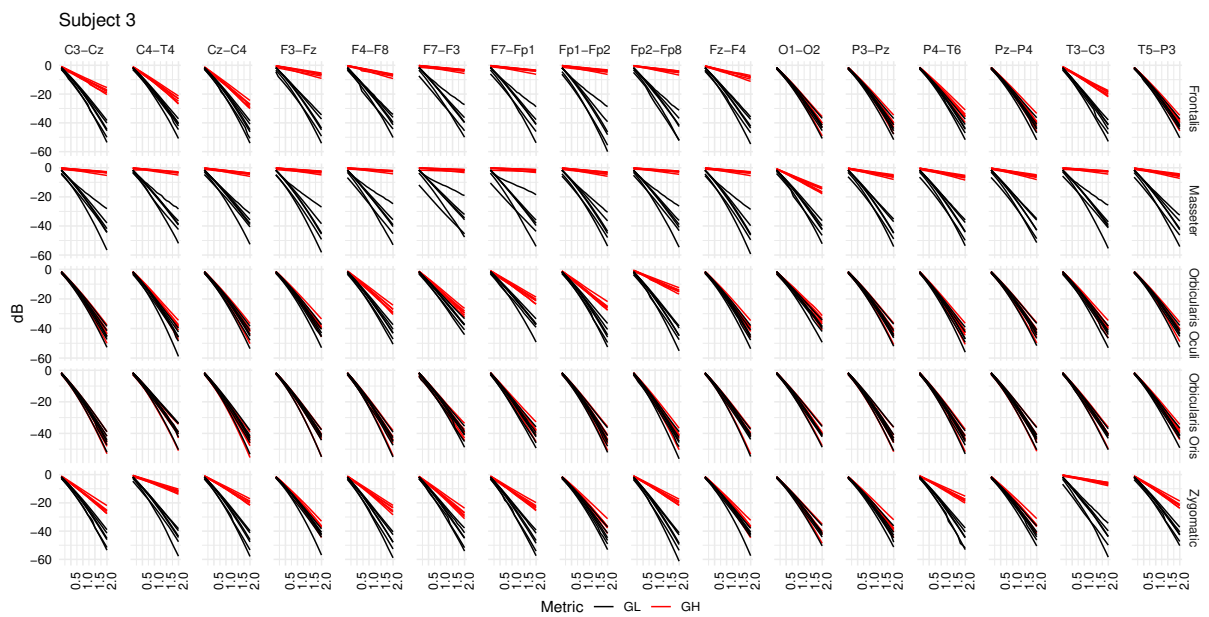


Figure A.2: Typical *GL* and *GH* feature vectors estimated using different decomposition techniques for Subject 3. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles.

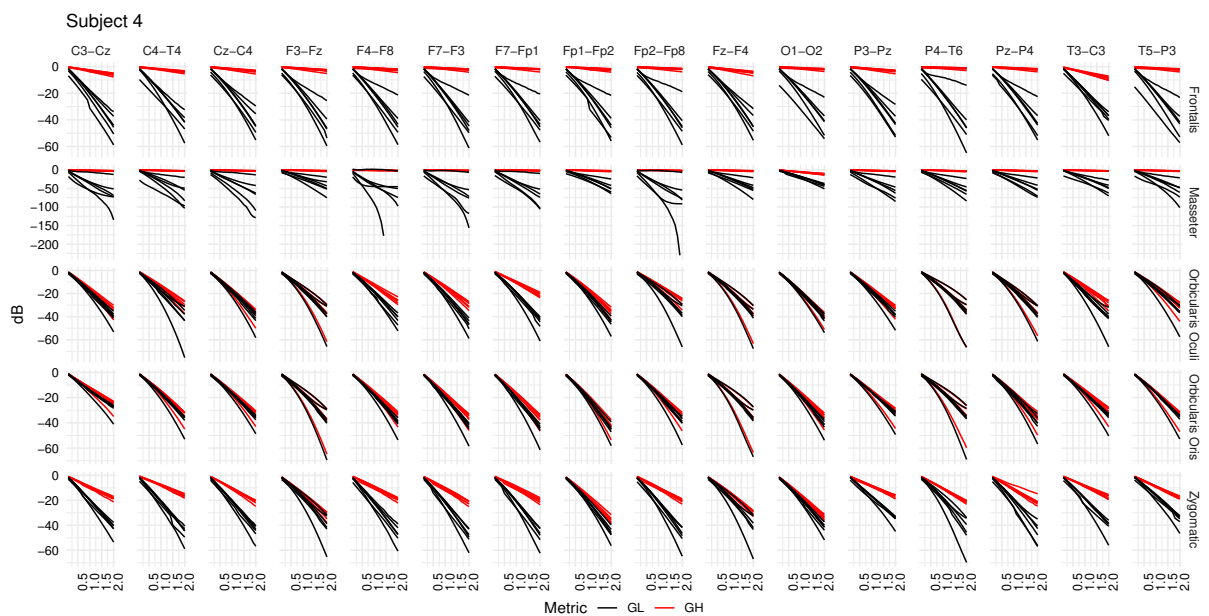


Figure A.3: Typical *GL* and *GH* feature vectors estimated using different decomposition techniques for Subject 4. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles.

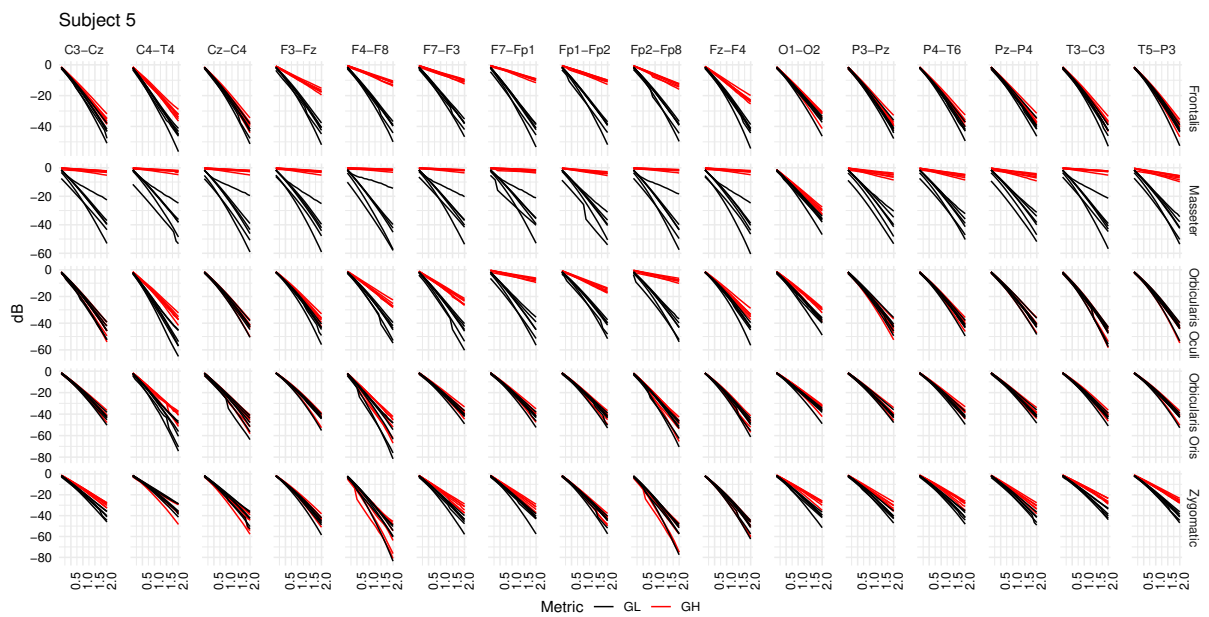


Figure A.4: Typical *GL* and *GH* feature vectors estimated using different decomposition techniques for Subject 5. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles.

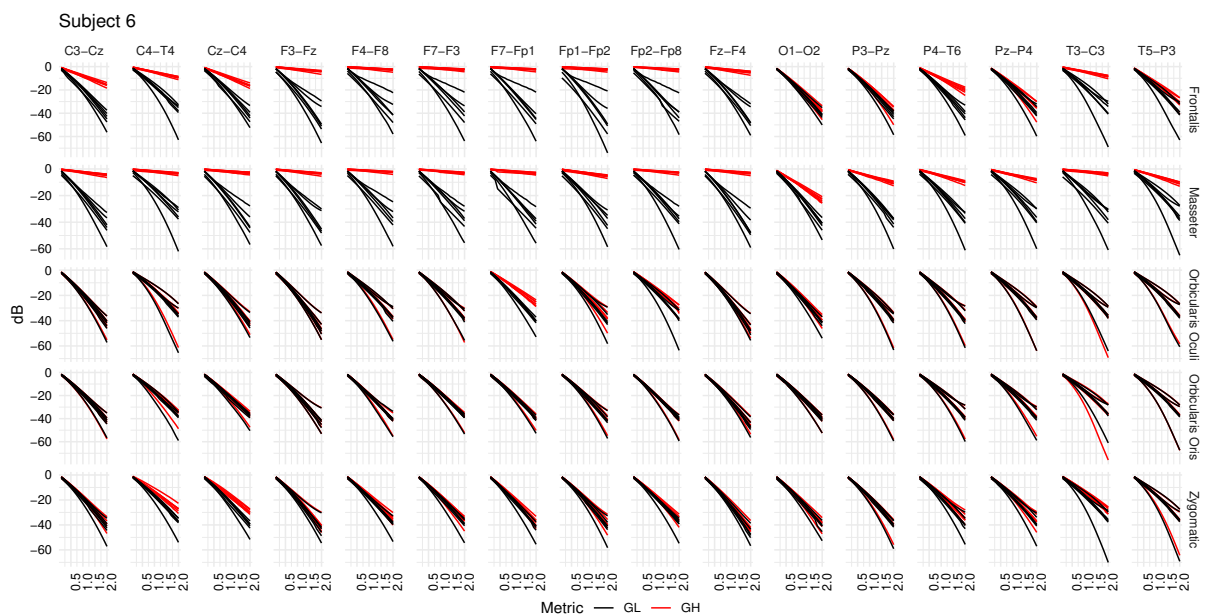


Figure A.5: Typical *GL* and *GH* feature vectors estimated using different decomposition techniques for Subject 6. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles.

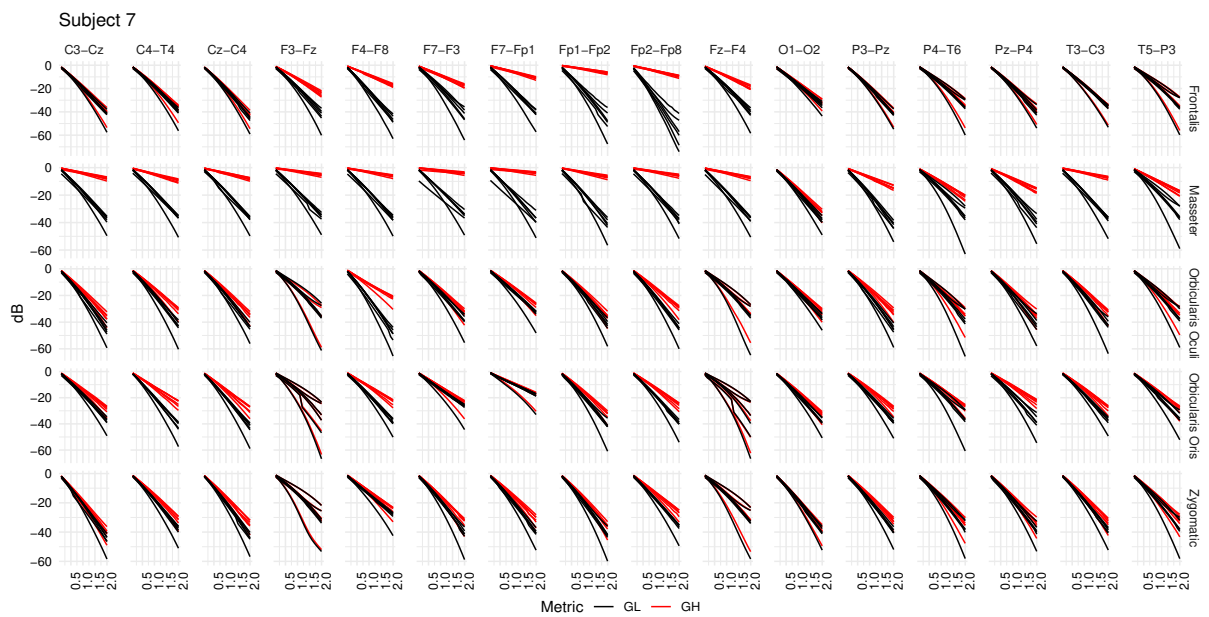


Figure A.6: Typical *GL* and *GH* feature vectors estimated using different decomposition techniques for Subject 7. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles.

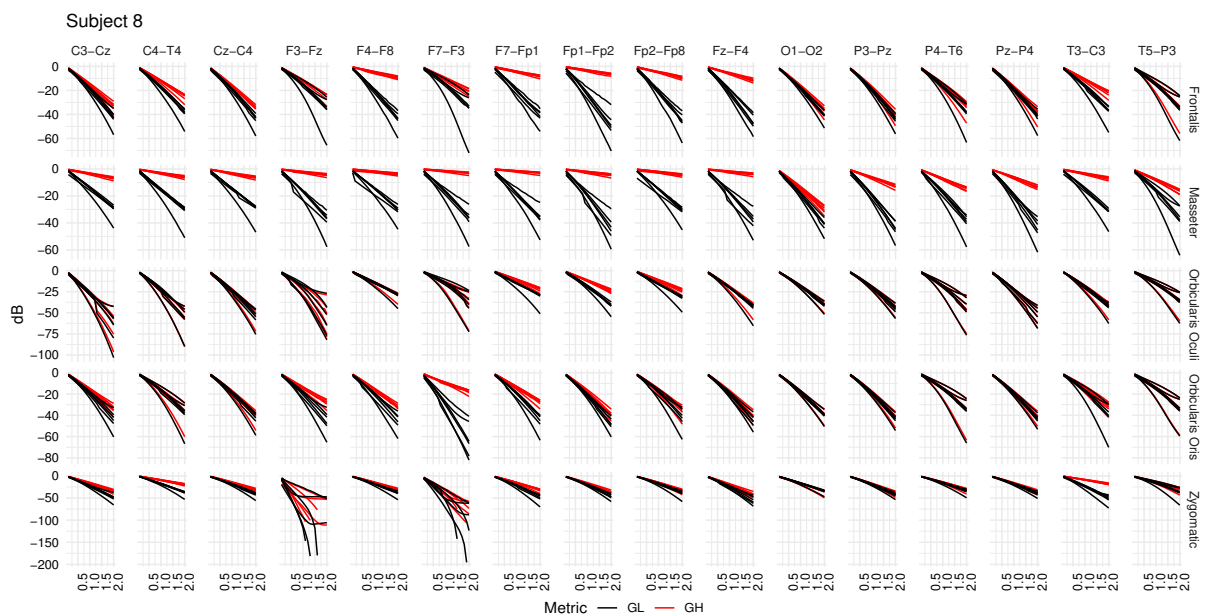


Figure A.7: Typical *GL* and *GH* feature vectors estimated using different decomposition techniques for Subject 8. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles.

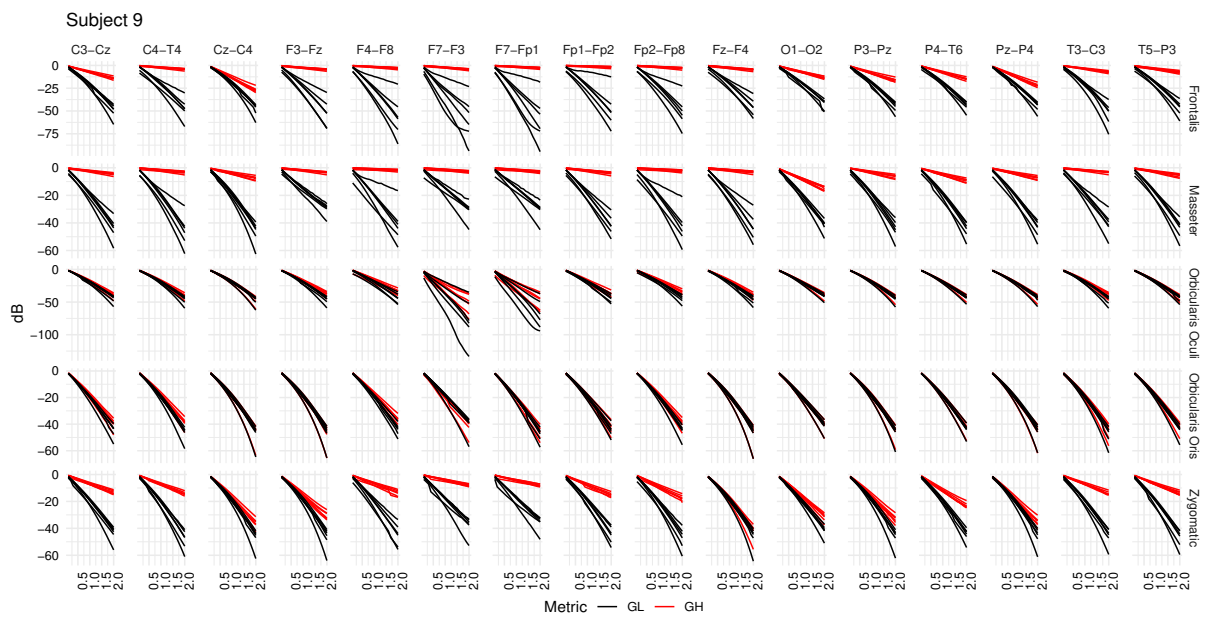


Figure A.8: Typical *GL* and *GH* feature vectors estimated using different decomposition techniques for Subject 9. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles.

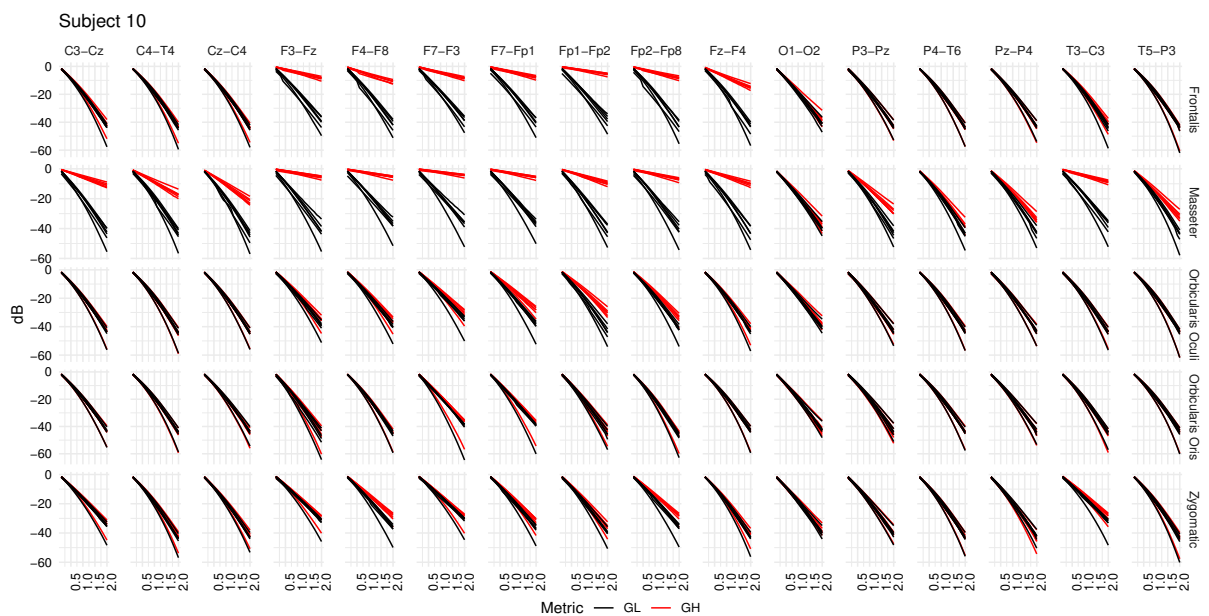


Figure A.9: Typical *GL* and *GH* feature vectors estimated using different decomposition techniques for Subject 10. Each plot consists of six vector pairs, one pair for each method. The outcomes are presented for individual EEG sensors and muscles.

Appendix B

Data: contamination of EEG by facial muscles

This appendix aims to present the dataset that contains the EMG and EEG signals collected for the study of contamination by facial EMG on EEG, i.e., the data used in this thesis. This database were collected under the responsibility of the doctoral student of Electrical Engineering, UFU, Gustavo Moreira Silva, linked to the NIATS laboratory.

B.1 Database

The RHD2000 series acquisition system was used to collect data from Intan Technologies (Intan, USA). The RHD2000 data acquisition system generates a file with the extension “.rhd”. Therefore, every file generated during a data collection will have the extension “.rhd”. The data contained in this extension (“.rhd”) cannot be manipulated directly.

The Matlab function `<read_Intan_RHD2000_file.m>`, provided by Intan, must be used to read the file generated by the RHD2000 acquisition system.

The `<read_Intan_RHD2000_file.m>` function opens a window that allows you to choose the desired “.rhd” file. After selecting the desired file, Matlab will read the data, generating several variables containing the data collected. Table B.1 presents all variables in alphabetical order generated by the `<read_Intan_RHD2000_file.m>` function, the variables in blue contain the data collected by the acquisition system.

The data contained in the database of this thesis are formed by the set of variables

Tabela B.1: Variables in alphabetical order generated by the Matlab function, <read_Intan_RHD2000_file.m>, provided by Intan.

| Variable | Description |
|---|---|
| amplifier_channels | Structure with information about the amplifier_data variable. |
| amplifier_data | Data digitized by acquisition cards. The matrix contains all the biological data collected by EMG (card A) and EEG (card B). |
| aux_input_channels | Not used. |
| aux_input_data | Not used. |
| aux_input_dataCopy | Not used. |
| board_dig_in_channels | Structure with information about the digital input signals refers to the board_dig_in_data variable. |
| board_dig_in_data | Digital data. The vector contains the marker of changes moment of eyes state. A marker can be added by using the buttons on the front panel of the Intam data acquisition system. |
| frequency_parameters | Structure containing information about the filters and sampling frequency used. |
| notes | Not used. |
| spike_triggers | Not used. |
| supply_voltage_channels | Structure with information about the acquisition cards monitoring channels, referring to the supply_voltage_data variable. |
| supply_voltage_data | Two vectors. They have the values of the voltages monitored in both acquisition cards used in the collection. |
| t_amplifier | Time vector, referring to the amplifier_data variable. |
| t_aux_input | Not used. |
| t_dig | Time vector, referring to the board_dig_in_data variable. |
| t_supply_voltage | Time vector, referring to the supply_voltage_data variable. |

listed in Table B.1, grouped by Matlab in the “.mat” format. Therefore, when using data from this database, the `<read_Intan_RHD2000_file.m>` function is unnecessary. Thus, for each volunteer who participated in the research (S1 to S10), there are five “.mat” files, referring to the five muscles activated to generate contamination of facial EMG on EEG.

B.2 Data summary

Summary of the data generated by the Matlab function `<read_Intan_RHD2000_file.m>` presented in Table B.1:

1. **amplifier_data:**

It is a matrix with the EMG (card A) and EEG (card B) signals collected. The `amplifier_channels` structure and the `t_amplifier` vector refer to `amplifier_data` matrix.

2. **board_dig_in_data:**

As expected in the experimental protocol, the collection is performed with the eyes in two states: only open and only closed. The `board_dig_in_data` variable contains digital values (0 or 1). The first moment containing 1s marks the beginning of the eye-open-only state and the second moment indicates the eye-closed-only state. The `board_dig_in_channels` structure and the `t_dig` vector refer to `board_dig_in_data` variable.

3. **supply_voltage_data:**

Data collection is performed by two acquisition cards, A(EMG) and B(EEG). The supply voltage of each card is monitored, and the measured values are stored in this variable. The `supply_voltage_channels` structure and the `t_supply_voltage` vector refer to `supply_voltage_data` variable.

B.3 EMG and EEG data

The data of interest for the collection are in the [amplifier_data](#) variable, a matrix with n rows and m columns, where n is the number of channels (EMG and EEG) and m is the number of data acquired at the sampling frequency for the entire collection time interval. EMG channels are collected by “Port A”, and EEG channels are collected by “Port B”, but all EMG and EEG data are in the same [amplifier_data](#) variable. The name “Port A” or “Port B” comes from the choice of cable used to connect the acquisition cards to the RHD2000 USB interface system. Thus, based on the choice adopted in connecting the EMG and EEG cards in this research, the [amplifier_data](#) variable has the first lines belonging to “Port A” (EMG) and the other lines of the matrix being from “Port B” (EEG). The [amplifier_channels](#) structure indicates which card was used with the name “Port A” and “Port B” in the `port_name` field.

B.3.1 EMG data

The muscle signal data does not have a specific and fixed channel between collections. Thus, the channel number for a muscle must be verified at each data collection. The information for each channel must be consulted in the [amplifier_channels](#) structure at the `custom_channel_name` field.

Tabela B.2: EEG sensors channel list.

| channel x EEG sensor | | | |
|----------------------|--------------|-------------|---------------|
| Ch00: F3-Fz | Ch01: F4-F8 | Ch02: F7-F3 | Ch03: Fp1-Fp2 |
| Ch04: T3-C3 | Ch05: C4-T4 | Ch06: T5-P3 | Ch07: P4-T6 |
| Ch08: Fp2-F8 | Ch09: F7-Fp1 | Ch10: Fz-F | Ch11: Cz-C4 |
| Ch12: O1-O2 | Ch13: P3-Pz | Ch14: Pz-P4 | Ch15: C3-CZ |

B.3.2 EEG data

A bipolar 16-channel acquisition card acquires EEG signals. The channels of each EEG sensor are the same for all collections, thanks to the use of the adapted EEG cap,

and the connection of the EEG signals card is always carried out on the same cable B. Table B.2 shows the relationship of the channels with the EEG sensors on card B. Figure B.1 illustrates the bipolar transverse mounting configuration of the fitted EEG cap.

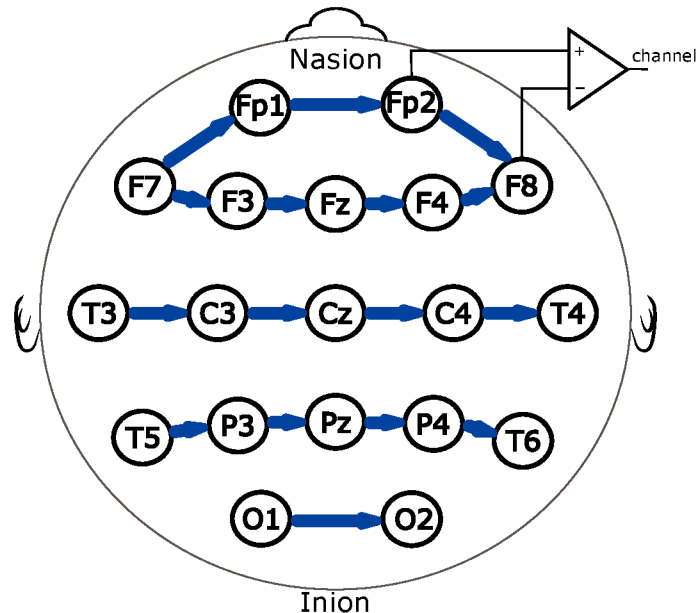
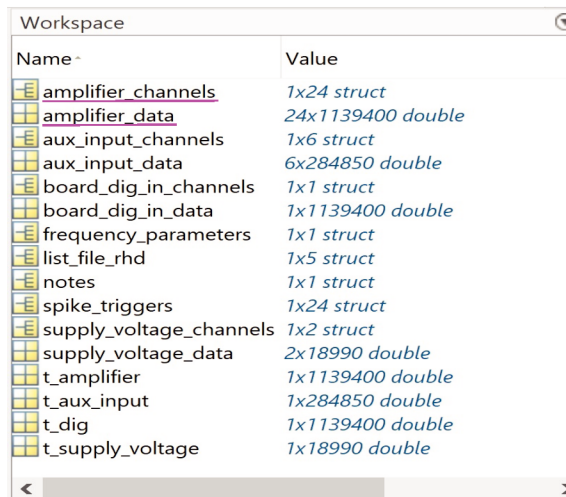


Figure B.1: Configuration of EEG electrodes in the EEG cap, bipolar transverse montage. Illustration of the connections of sensors Fp2 and F8, channel 08, according to Table B.2.

B.4 Example of a database “.mat” file

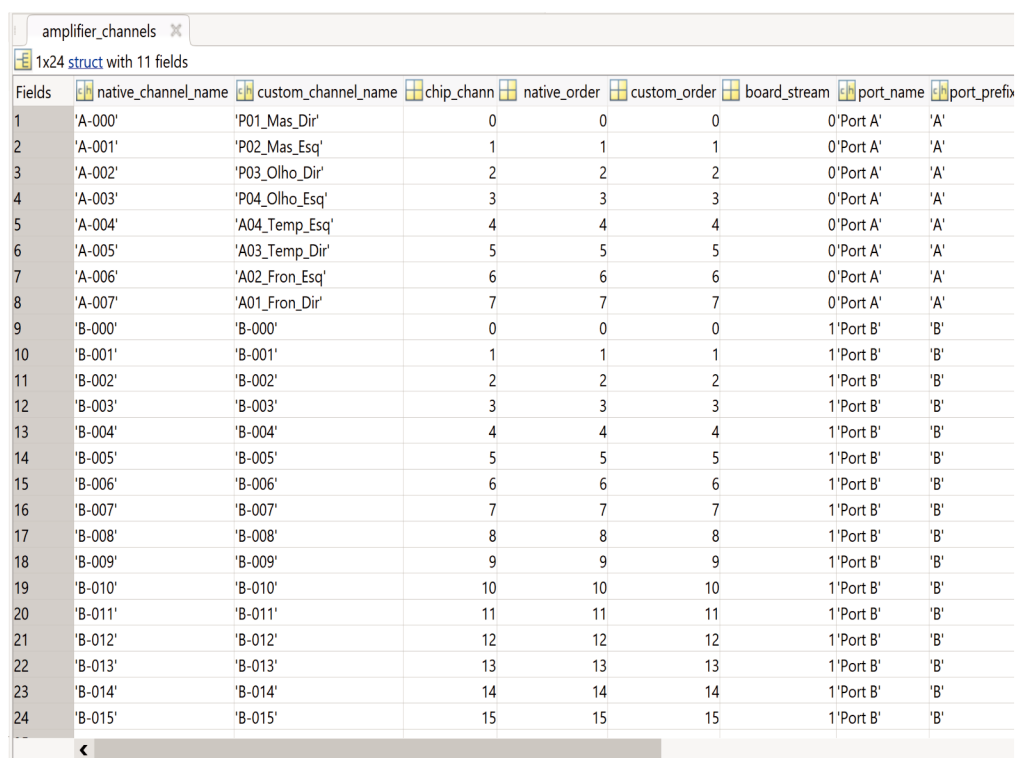
The file of volunteer 01 and masseter muscle is shown to illustrate the use of the database. Figure B.2 illustrates the Matlab workspace when loading the V01_Masseter_160129_111428.mat file. Figure B.2 shows the `amplifier_channels` and `amplifier_data` variables, where we can highlight that the `amplifier_data` variable has data from 24 channels with 1,139,400 points (double).

Figure B.3 illustrates part of the `amplifier_channels` variable fields. In the field called `custom_channel_name`, it is possible to visualize the 24 channels, the first eight being EMG channels, “Port a” in the `port_name` field, and the remaining 16 channels, “Port B” in the `port_name` field, indicating that they belong to the set of EEG cap channels and that follow the distribution shown in Table B.2.



| Name | Value |
|-------------------------|-------------------|
| amplifier_channels | 1x24 struct |
| amplifier_data | 24x1139400 double |
| aux_input_channels | 1x6 struct |
| aux_input_data | 6x284850 double |
| board_dig_in_channels | 1x1 struct |
| board_dig_in_data | 1x1139400 double |
| frequency_parameters | 1x1 struct |
| list_file_rhd | 1x5 struct |
| notes | 1x1 struct |
| spike_triggers | 1x24 struct |
| supply_voltage_channels | 1x2 struct |
| supply_voltage_data | 2x18990 double |
| t_amplifier | 1x1139400 double |
| t_aux_input | 1x284850 double |
| t_dig | 1x1139400 double |
| t_supply_voltage | 1x18990 double |

Figure B.2: Matlab workspace window displaying variables from a database file on volunteer 1's masseter muscle data.



| Fields | native_channel_name | custom_channel_name | chip_chann | native_order | custom_order | board_stream | port_name | port_prefix |
|--------|---------------------|---------------------|------------|--------------|--------------|--------------|-----------|-------------|
| 1 | 'A-000' | 'P01_Mas_Dir' | 0 | 0 | 0 | 0 | 0'Port A' | 'A' |
| 2 | 'A-001' | 'P02_Mas_Esq' | 1 | 1 | 1 | 1 | 0'Port A' | 'A' |
| 3 | 'A-002' | 'P03_Olho_Dir' | 2 | 2 | 2 | 2 | 0'Port A' | 'A' |
| 4 | 'A-003' | 'P04_Olho_Esq' | 3 | 3 | 3 | 3 | 0'Port A' | 'A' |
| 5 | 'A-004' | 'A04_Temp_Esq' | 4 | 4 | 4 | 4 | 0'Port A' | 'A' |
| 6 | 'A-005' | 'A03_Temp_Dir' | 5 | 5 | 5 | 5 | 0'Port A' | 'A' |
| 7 | 'A-006' | 'A02_Fron_Esq' | 6 | 6 | 6 | 6 | 0'Port A' | 'A' |
| 8 | 'A-007' | 'A01_Fron_Dir' | 7 | 7 | 7 | 7 | 0'Port A' | 'A' |
| 9 | 'B-000' | 'B-000' | 0 | 0 | 0 | 0 | 1'Port B' | 'B' |
| 10 | 'B-001' | 'B-001' | 1 | 1 | 1 | 1 | 1'Port B' | 'B' |
| 11 | 'B-002' | 'B-002' | 2 | 2 | 2 | 2 | 1'Port B' | 'B' |
| 12 | 'B-003' | 'B-003' | 3 | 3 | 3 | 3 | 1'Port B' | 'B' |
| 13 | 'B-004' | 'B-004' | 4 | 4 | 4 | 4 | 1'Port B' | 'B' |
| 14 | 'B-005' | 'B-005' | 5 | 5 | 5 | 5 | 1'Port B' | 'B' |
| 15 | 'B-006' | 'B-006' | 6 | 6 | 6 | 6 | 1'Port B' | 'B' |
| 16 | 'B-007' | 'B-007' | 7 | 7 | 7 | 7 | 1'Port B' | 'B' |
| 17 | 'B-008' | 'B-008' | 8 | 8 | 8 | 8 | 1'Port B' | 'B' |
| 18 | 'B-009' | 'B-009' | 9 | 9 | 9 | 9 | 1'Port B' | 'B' |
| 19 | 'B-010' | 'B-010' | 10 | 10 | 10 | 10 | 1'Port B' | 'B' |
| 20 | 'B-011' | 'B-011' | 11 | 11 | 11 | 11 | 1'Port B' | 'B' |
| 21 | 'B-012' | 'B-012' | 12 | 12 | 12 | 12 | 1'Port B' | 'B' |
| 22 | 'B-013' | 'B-013' | 13 | 13 | 13 | 13 | 1'Port B' | 'B' |
| 23 | 'B-014' | 'B-014' | 14 | 14 | 14 | 14 | 1'Port B' | 'B' |
| 24 | 'B-015' | 'B-015' | 15 | 15 | 15 | 15 | 1'Port B' | 'B' |

Figure B.3: Part of the amplifier_channels structure fields on the Matlab screen.

B.5 Access to Database

The database, contamination of EEG by facial muscles, can be accessed via:

Andrade, Adriano de Oliveira, & Queiroz, Carlos Magno. (2022). Single channel

approach for filtering electroencephalographic signals strongly contaminated with facial electromyography [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.7133259>.

Appendix C

Data: Motor Learning Protocol

The motor learning protocol was created to propose an index capable of estimating motor learning in healthy people, acquired during interaction with a human-computer interface, controlled by facial muscles based on the correlation between motor facial and brain activities. This database were collected under the responsibility of the doctoral student of Electrical Engineering, UFU, Carlos Magno Medeiros Queiroz, linked to the NIATS laboratory.

C.1 Motor learning

Motor learning can be conceived as establishing an internal model representing the accurate correspondence between sensory and perceived motor information [63]. During the initial phase of motor learning, the movements are low-skilled, highly dependent on feedback, and require high demands on attention [64]. With practice, the precision and speed of actions increase, and thus, as motor learning progresses, sensory feedback processing becomes less critical [65]. Generally, two forms of motor learning can be recognized, explicit and implicit learning. Explicit learning involves the conscious recall of past experiences. Implicit learning is an unintentional, non-conscious form of learning characterized by improving the motor act. Motor skill progresses from explicit control, in the early stages of learning, to more implicit or automatic control when already well assimilated. Results from studies suggest that motor learning consists of three distinct phases [66]: Early stage - slow motor action under great sensory attention, irregular pattern of movements, significant variation in performance time; In-

intermediate stage - gradual learning of the sensory-motor map, increased performance; Advanced stage - automated speed and movement, skillful performance, isochronous movements, and complete sensory field control. For each of these stages, changes in brain activity can be detected through changes in rhythmic waves [67, 68] or brain macro potentials related to movement. Such changes can be analyzed and correlated in order to estimate the learning stage.

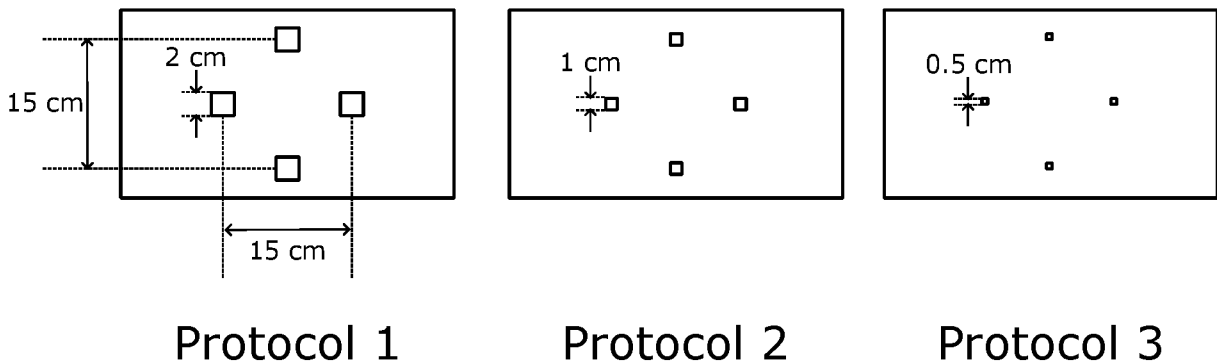


Figure C.1: Illustration of the position and size of the four buttons according to protocols 1, 2, and 3 (the illustration is not in scale).

C.2 Graphical learning interface

The graphical learning interface (GLI) employed is an environment created to offer: specific and controlled tasks, with measurable execution times and virtual, possible to be performed using a computer cursor. The graphical environment of this interface was developed to maximize the user's focus on the task to be performed, which implies a totally black graphical interface capable of presenting objects with configurable graphical characteristics. The conceived task consists of clicking on four buttons with variable sizes, according to the difficulty level, which are presented on the screen consecutively; the next one only appears after the previous one has been clicked and in predefined positions. Figure C.1 presents three screen options, each containing four buttons, offering three difficulty levels. The distance between the horizontal and vertical buttons is always the same and equal to 15 cm. The size of the buttons is variable 2 cm, 1 cm, and 0.5 cm for protocols 1, 2, and 3, respectively, which implies increasing difficulty according to Fitts' law [69, 70]. The GLI was developed in C# language employing

Microsoft's Visual Studio development environment.

The essential aspects of this GLI are described, addressing the timing of events, the log file of the actions performed, and the configurable characteristics.

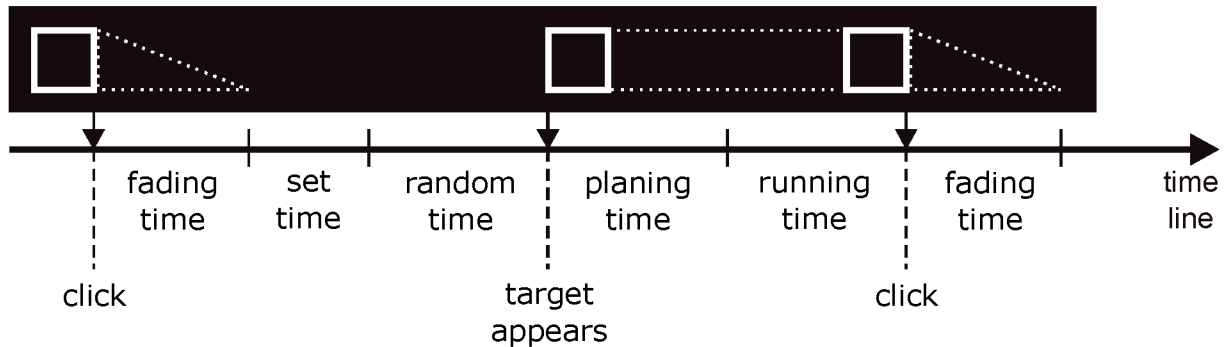


Figure C.2: Timeline of events in the graphical learning interface.

C.2.1 Timing of events

Considering that the objective of this interface is to create an environment to evaluate motor learning through the correlation of biopotentials related to muscle and brain activities, each graphic element has an essential role in constructing this evaluation environment. The timing of the graphical elements, buttons (targets), is important in this context because it allows for creating situations with appropriate and typical conditions when assessing motor learning. In this sense, the timing of events related to buttons (task targets) was conceived, as shown in Figure C.2.

In the timeline, a target, when clicked, starts fading until it disappears, then there is a minimum waiting time of fixed duration, followed by a random waiting time. During the total timeout (set time + random time), the screen remains completely dark (nothing projected). After the waiting time, the next target appears on the screen, and the time elapsed between its appearance and the moment it is clicked consists of the planning time and the running time. The planning time comprises the interval from which the target appears until the moment when the user initiates an action with the cursor. The execution time is considered from the end of the planning interval until the moment when the target is actually clicked correctly. All the events: the instant of target appearance, actions, and cursor position, are registered in a file, enabling later data analysis.

| TIME | ACTION | X_POS | Y_POS | WIDTH | HEIGHT | |
|-------------------------|----------|-------|-------|-------|--------|-----------------------|
| 08/09/2016 10:54:01.280 | TARGET_1 | 24 | 641 | 84 | 84 | — FIRST TARGET |
| 08/09/2016 10:54:01.311 | MOVE | 286 | 209 | 0 | 0 |] MOVING |
| 08/09/2016 10:54:02.608 | MOVE | 309 | 209 | 0 | 0 | |
| 08/09/2016 10:54:02.718 | MOVE | 411 | 238 | 0 | 0 | |
| 08/09/2016 10:54:02.827 | MOVE | 498 | 242 | 0 | 0 | |
| 08/09/2016 10:54:02.936 | MOVE | 529 | 239 | 0 | 0 | |
| 08/09/2016 10:54:03.155 | MOVE | 533 | 236 | 0 | 0 | |
| 08/09/2016 10:54:03.265 | MOVE | 690 | 101 | 0 | 0 | |
| 08/09/2016 10:54:03.374 | MOVE | 717 | 74 | 0 | 0 | |
| 08/09/2016 10:54:03.483 | MOVE | 718 | 74 | 0 | 0 | |
| 08/09/2016 10:54:03.702 | MOVE | 709 | 73 | 0 | 0 | |
| 08/09/2016 10:54:03.811 | MOVE | 676 | 77 | 0 | 0 | |
| 08/09/2016 10:54:03.921 | MOVE | 675 | 77 | 0 | 0 | — CLICK IN THE TARGET |
| 08/09/2016 10:54:04.296 | CLICK_IN | 675 | 77 | 0 | 0 | — NEXT TARGET |
| 08/09/2016 10:54:09.395 | TARGET_3 | 660 | 641 | 84 | 84 |] MOVING |
| 08/09/2016 10:54:10.380 | MOVE | 659 | 660 | 0 | 0 | |
| | | ⋮ | | | | ⋮ |

Figure C.3: Only the initial part of the cursor log file is shown. It displays the cursor movement to click on TARGET_1 and records the moment when TARGET_3 appears.

C.2.2 Log file

The developed GLI registers all the events related to the cursor and the instant when each button appears on the screen in a text file. Each record corresponds to a line containing the following fields: TIME, ACTION, X_POS, Y_POS, WIDTH, and HEIGHT. The cursor's position is registered every 100 *ms* if the current position differs from the previous one. For instance, for a new cursor position, a new line is generated in which the fields: TIME contains the date and time, ACTION contains the word MOVE, the X_POS and Y_POS fields contain the cursor position on the screen, and the WIDTH and HEIGHT fields, which are not used in this case, both contain the value zero. For an interface click event, the record line will contain: the TIME field with date and time; the ACTION field will be equal to CLICK_IN or CLICK_OUT, depending on whether the click was inside the target (button) or not; the X_POS and Y_POS fields contain the position of the click, and the WIDTH and HEIGHT fields are also not used in this case, both contain the value zero. The appearance of a target on the screen is recorded in a new line: the TIME field contains date and time; the ACTION field contains the name of the target (TARGET_1, TARGET_2, TARGET_3, and TARGET_4); the X_POS and

Y_POS fields contain the position of the upper left corner of the target and the WIDTH and HEIGHT fields contain the width and height of the target displayed. The log file was designed to contain all the necessary information to enable synchronism with the EMG and EEG data and to allow reconstitution of the entire cursor movement and actions in the GLI environment. Figure C.3 shows, as an example, the initial part of any log file. The record contains the cursor movement required to click on TARGET_1 and then records the appearance of the next target, TARGET_3.

C.2.3 Configurable features

To make the GLI more flexible and to adapt to the variations needed for a given experiment, several customizable features have been implemented. Table C.1 shows all the configurable features, with their respective acceptable value ranges and the value adopted in the experiment.

C.3 Mouse emulator system

The developed human-computer interface promotes the user's iteration with the computer through the facial muscles. The biopotentials emanating from muscle activation, detected by electrodes, are converted into commands capable of controlling the computer cursor. The program capable of performing such conversion, EMG signals into commands for the computer cursor, was developed and integrated with the program of the graphical interface of the acquisition system. Considering that the conversion program must be multiplatform and aiming to avoid problems in controlling the mouse in different operating systems, it was decided to build a micro-processed device capable of receiving digital commands from the acquisition system and emulating a mouse to control the computer's cursor. This option, a mouse emulator device, was conceived and implemented to avoid future problems of incompatibility and competition with other programs when it is desired to control the cursor in different operating systems.

Figure C.4 illustrates the developed mouse emulator system that employs an Arduino Due board, based on the Atmel SAM3X8E arm Cortex-M3 microcontroller, and its main features are 54 digital inputs/outputs, 4 UARTs, 84 MHz operating frequency.

Tabela C.1: Configuration characteristics of the graphical learning interface.

| Feature | Acceptable range | Adopted value |
|--|-------------------|------------------|
| General | | |
| To emit one beep when the target is reached | False, True | False |
| Display markup to calibrate scale ¹ | False, True | False |
| Value in mm refers to the value in pixels (Scale) | Integer > 0 | 150 ² |
| Distance in pixels from the center of the screen to the center of the target | 1 to 400 | 283 ² |
| To show messages on the screen | False, True | False |
| Number of repetitions for the same protocol | Integer: 0 to 60 | 4 |
| Name of the file of action register | Name of file .txt | Supplied |
| Target Frame | | |
| Target Frame Color (RGB) | 0-255;0-255;0-255 | 255;255;255 |
| Target Frame Thickness in Pixels | 1 - 100 | 1 |
| Frame blink frequency in ms | 1 - 1000 | 0 |
| The color of the blink on the frames | 0-255;0-255;0-255 | 255;255;255 |
| Target area | | |
| Color of the target area) | 0-255;0-255;0-255 | 0;0;0 |
| Blink frequency on the target area in ms | 0- 1000 | 0 |
| Blink color of the target area | 0-255;0-255;0-255 | 255;255;255 |
| Target | | |
| Time in s to fade target after hit | 1 - 100 | 2 |
| Minimum time in s for the appearance of the next target | 1 - 100 | 1 |
| Maximum time in s for the appearance of the next target | 1 - 100 | 3 |
| Size in mm of target - Large protocol | 1 - 4000 | 20 |
| Size in mm of target - Medium protocol | 1 - 4000 | 10 |
| Size in mm of target - Small protocol | 1 - 4000 | 5 |

¹ It was used only to adjust the scale for the monitor in use. It opens a scaling window.

² These values could vary from monitor to monitor.



Figure C.4: Mouse emulator that acts as a mouse in manual mode, via keyboard at the top of the console, or in automatic mode, receiving external commands. System based on the Arduino Due board.

The Arduino Due board has two micro USB connectors, one for programming the Arduino itself and another for a native USB port usually recognized as a mouse or keyboard by a computer. The mouse emulator system allows the user to control the computer cursor utilizing a set of buttons on its console, manual mode, or through external digital signals, automatic mode. The system has also a display by which the user can view the actions performed in manual or automatic mode. The manual mode was created only to enable tests of the emulator itself. The automatic mode was created to operate through the acquisition system, receiving commands generated from muscle activity. For this, a communication protocol was created that includes the four movement commands (up, down, right, and left), right-click, left-click and step size configu-

ration (resolution). But by myoelectric activity, in automatic mode, only the movement and right-click commands are recognized. The command word consists of 6 bits, and there is only one control signal called “enable”. Commands are sent via a DB25 connector present on the side of the mouse emulator. Table C.2 shows the commands that make up the protocol created for communication between the mouse emulator and the human-computer interface.

Tabela C.2: Command list supported by the Mouse Emulator control protocol.

| Binary | Decimal | Function | Description |
|--------------|----------|-------------|--|
| D6-D0 | | | |
| 0000000 | 0 | Unused | For future use. |
| 0000001 | 1 to 120 | Resolution | The value present in this range is understood as the new resolution value of the mouse movement. |
| 1111000 | | | |
| 1111001 | 121 | Move up | Moves the cursor up, only once. |
| 1111010 | 122 | Move down | Moves the cursor down, only once. |
| 1111011 | 123 | Move left | Moves the cursor to the left, only once |
| 1111100 | 124 | Move right | Moves the cursor to the right, only once. |
| 1111101 | 125 | Left click | Enables and disables left click, only once. |
| 1111110 | 126 | Right click | Enables and disables right click, only once. |
| 1111111 | 127 | Menu mode | Changes the Mouse Emulator’s operation mode to the Start Menu. |

To send a command according to the protocol created, the command word must be activated and then keep the “enable” command line active for a minimum of 3 *ms*, as illustrated in Figure C.5.

C.4 Mouse cursor control

Since this is a myoelectric interface, the computer cursor will be controlled by muscle activation, Frontalis. To control the mouse with only one muscle, the finite state machine, presented in Figure C.6, was implemented considering only three distinct types of contraction. As mentioned, the mouse emulator is designed to execute four

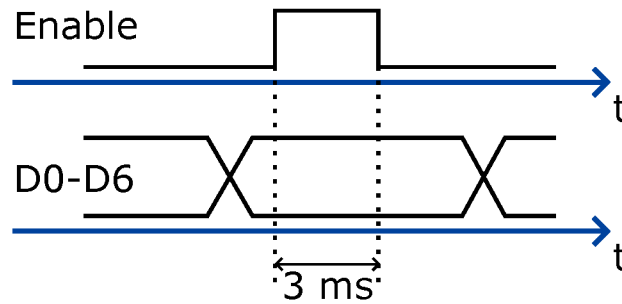


Figure C.5: Temporal diagram of sending a command from an external device to the Mouse Emulator.

move commands (up, down, right, and left) and right click. The cursor has an arrow shape and always starts pointing to the right. It starts on “StandBy” state and remains stationary, pointing to the right, as long as there is no contraction. If a contraction occurs, the mouse state will be changed depending on the duration of this contraction (C1, C2, or C3, see Figure C.6). If there is a contraction that lasts less than 0.4 s, contraction of type C1, the cursor arrow will be rotated 90 degrees clockwise and then the mouse returns to the “StandBy” state. Thus, to point the cursor in the desired direction, it is enough to perform other contractions of type C1. However, if the contraction is of type C2, that is, with a duration equal to or greater than 0.4 s and less than 1 s, the mouse cursor starts moving in the current direction of the arrow and remains in motion until any contraction occurs; if it does, the cursor stops and the mouse returns to the “StandBy” state. The click command is activated for a contraction lasting 1 s or more, and then the mouse returns to the “StandBy” state.

C.5 EMG Signals

The right and left Frontalis muscles were used to control the myoelectric interface. The EMG signals were acquired by the RHD2000 biological data acquisition system from Intan, USA, and the RHD2216 acquisition card was employed. The RHD2216 card contains 16 differential channels with automatic gain, a 16-bit digital converter, and sampling rate from 1 *kHz* to 30 *kHz*. The EMG signals were acquired by two differential channels and sampling rate equal to 1 *kHz*. The EMG sensors used are based on the integrated circuit PS25255 from Plessey Semiconductors, USA, and have

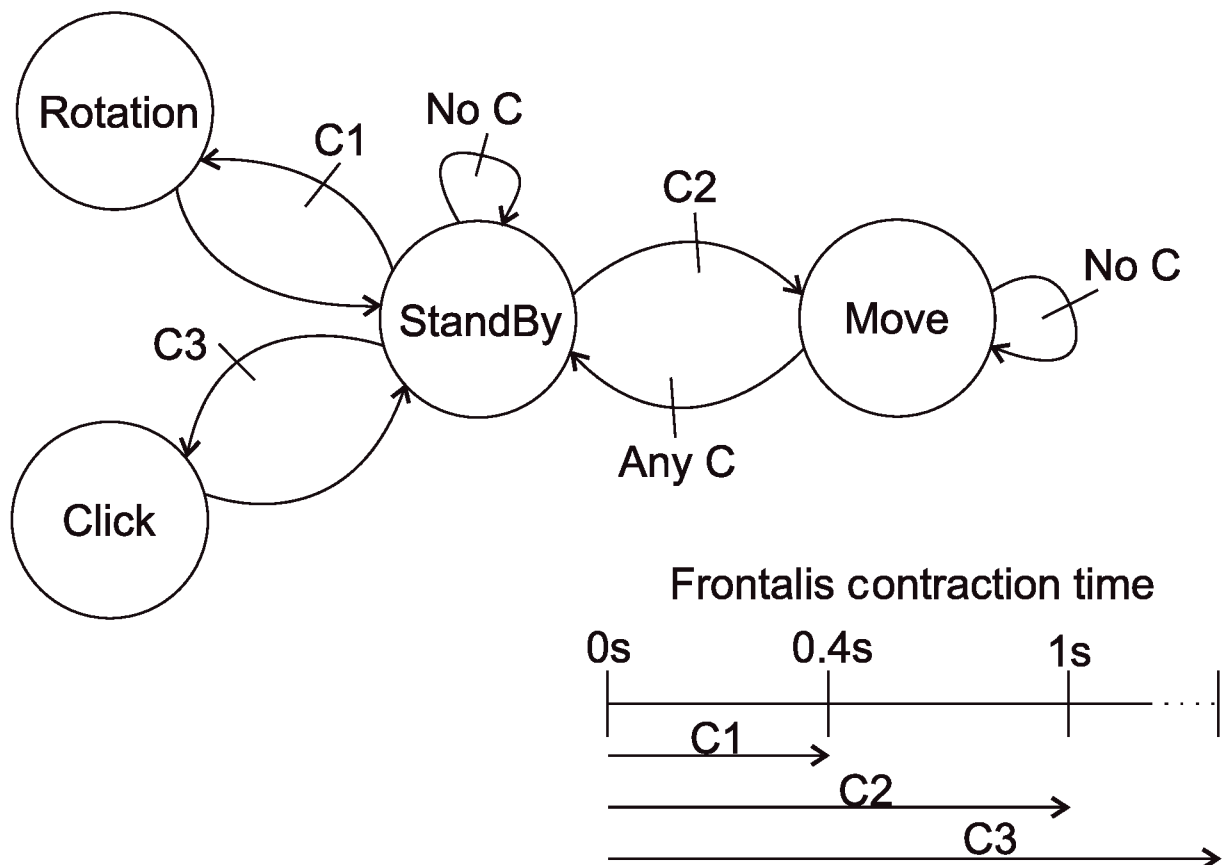


Figure C.6: Finite state machine implemented to control the mouse cursor through the contraction of a single muscle. The cursor control is performed entirely by distinguishing only three types of duration of Frontalis contractions.

high input impedance (typical value of $G\Omega$), dry contact capacitive coupling (does not require the use of gel), bandwidth from 200 mHz to 20 kHz , supply voltage from $\pm 2.4 Vdc$ to $\pm 5.5 Vdc$ and voltage gain of 10 times. Figures C.7 A and B show the RHD2216 acquisition card and EMG sensors, set up with two pairs of PS25255 sensors, respectively.

C.6 EEG signals

The EEG signals were also acquired by a 16-channel bipolar RHD2216 acquisition card. The channels of each EEG sensor are the same for all collections, thanks to the adapted EEG cap (Figure 3.6). Table C.3 presents the relationship of the channels with the EEG sensors. Figure C.8 illustrates the configuration of the transverse bipolar

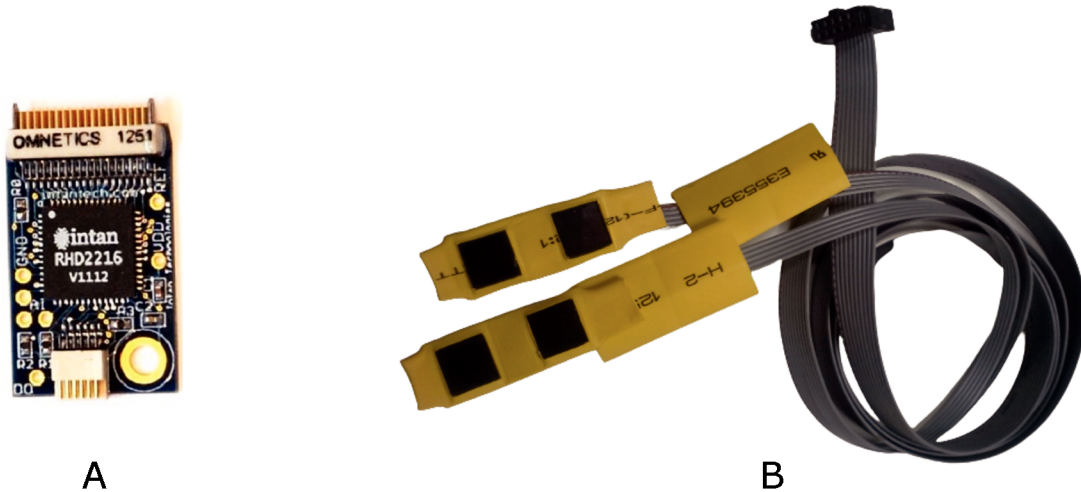


Figure C.7: **(A)** Conditioning card and acquisition of biological signals RHD2216, Intan, USA, with 16 differential channels. **(B)** EMG sensors, assembled set with two pairs of PS25255 sensors with high input impedance (typical value of $20\text{ G}\Omega$), dry contact capacitive coupling and bandwidth from 200 mHz to 20 kHz .

montage used in the adapted EEG cap.

Tabela C.3: EEG sensors channel list.

| channel x EEG sensor | | | |
|----------------------|--------------|-------------|---------------|
| Ch00: F3-Fz | Ch01: F4-F8 | Ch02: F7-F3 | Ch03: Fp1-Fp2 |
| Ch04: T3-C3 | Ch05: C4-T4 | Ch06: T5-P3 | Ch07: P4-T6 |
| Ch08: Fp2-F8 | Ch09: F7-Fp1 | Ch10: Fz-F | Ch11: Cz-C4 |
| Ch12: O1-O2 | Ch13: P3-Pz | Ch14: Pz-P4 | Ch15: C3-CZ |

C.7 Collection structure

The collection was designed in order to evoke in the volunteers participating in the study the three phases of learning; that is, the volunteers must experience during the collection a change from a motor control, initially from the explicit, poorly qualified, to a more precise motor control, and agile, passing through an intermediate stage with a clear increase in task performance [64, 65]. Thus, with this objective in mind,

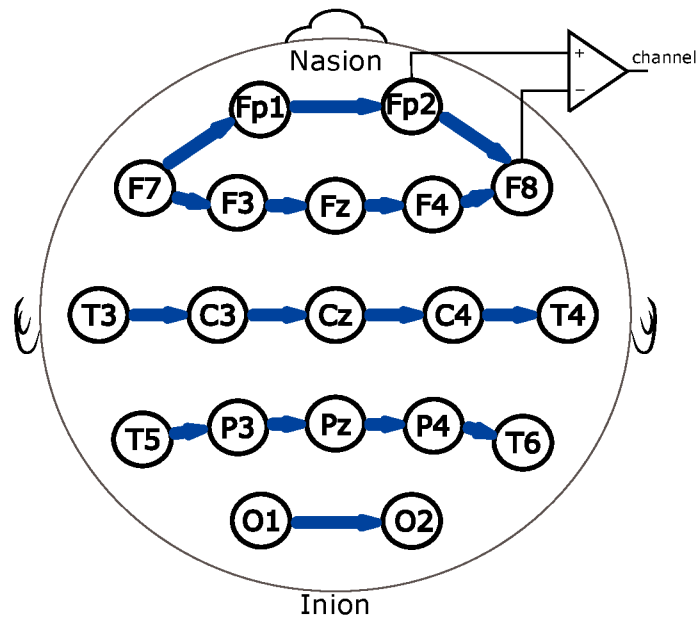


Figure C.8: Configuration of the EEG electrodes in the cap, transverse bipolar montage. Illustration of the connections of sensors Fp2 and F8, channel 08 according to Table C.3.

the collection contains five sessions that must be carried out on different days and, preferably, consecutive ones. The volunteer works with the three protocols in each session, which present different difficulty levels, as illustrated in Figure C.1. For each protocol, the volunteer performs five repetitions of the set of tasks consisting of moving the cursor and clicking on four targets (buttons) that appear in predefined positions but randomly. Figure C.9 illustrates the structure of the complete collection for a single session. Therefore, considering that for a single session, there are 60 targets, which implies 300 targets per volunteer, and considering 11 volunteers, we have a total of 3,300 targets to be analyzed.

C.8 EMG and EEG data

The data of interest in was stored in the variables `emgR` and `eegR`, which are organized structures containing several variables related to the collected data, estimated vectors, and characteristics of interest to the learning evaluation process.

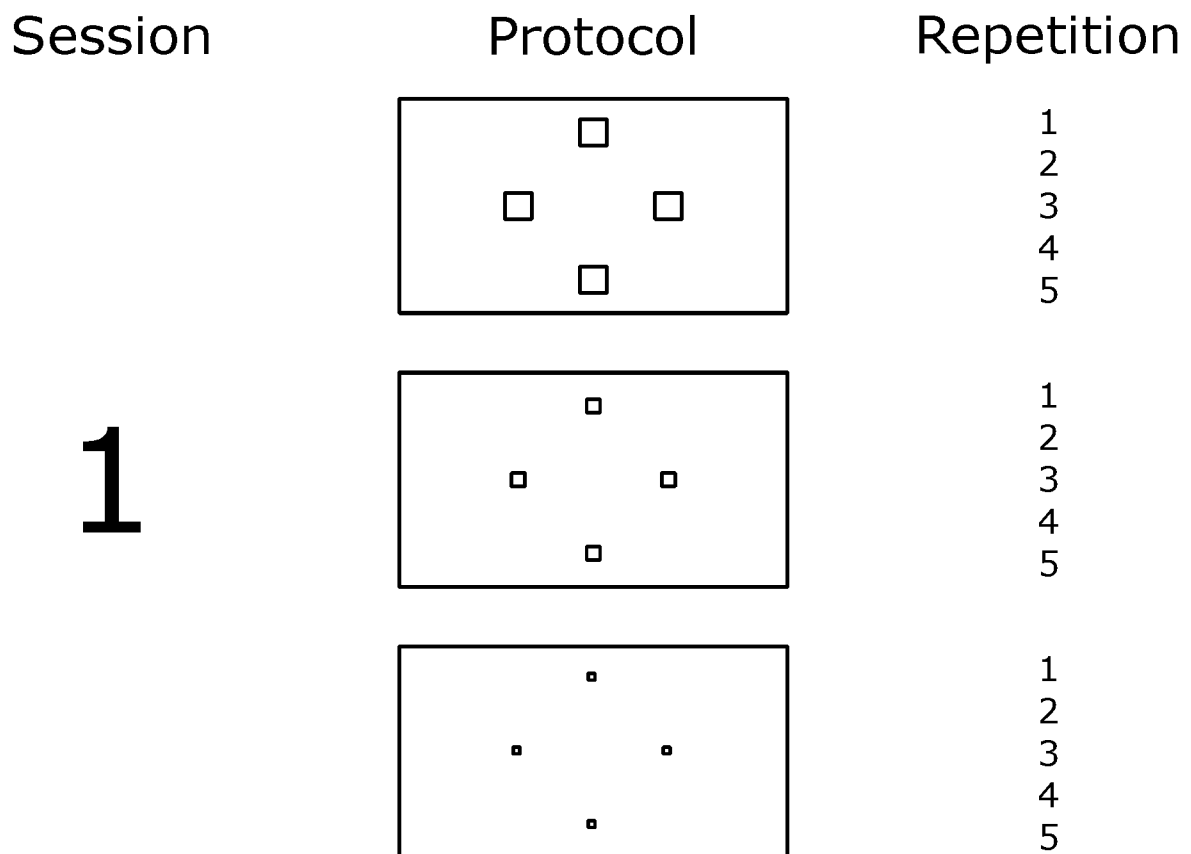


Figure C.9: Collection structure for a single session totals 60 targets, given that the task is to click on four targets for each repetition.

C.8.1 Structure EMG

Figure C.10 shows the structure created to store the data referring to electromyography and control data of the operation of the electromyographic interface used to evaluate motor learning. From this data structure created, it is possible to locate a given collection simply by indicating the volunteer number (1 to 11), the session number (1 to 5), the protocol number (1 to 3), and the repetition number (1 to 5). Figure C.10 shows only part of the data; among them, we can mention: the raw EMG (sgn), filtered EMG (FiltEmg), EMG Hilbert transform (hilbert), and the mouse cursor control signals (smouse_cmd). Figure C.11 shows the raw EMG signals employed to control the mouse cursor during the tasks of clicking on the targets presented by the graphical learning interface. The envelope of the EMG signal was obtained by applying the Hilbert transform to the filtered EMG signal, and through this envelope signal, it is pos-

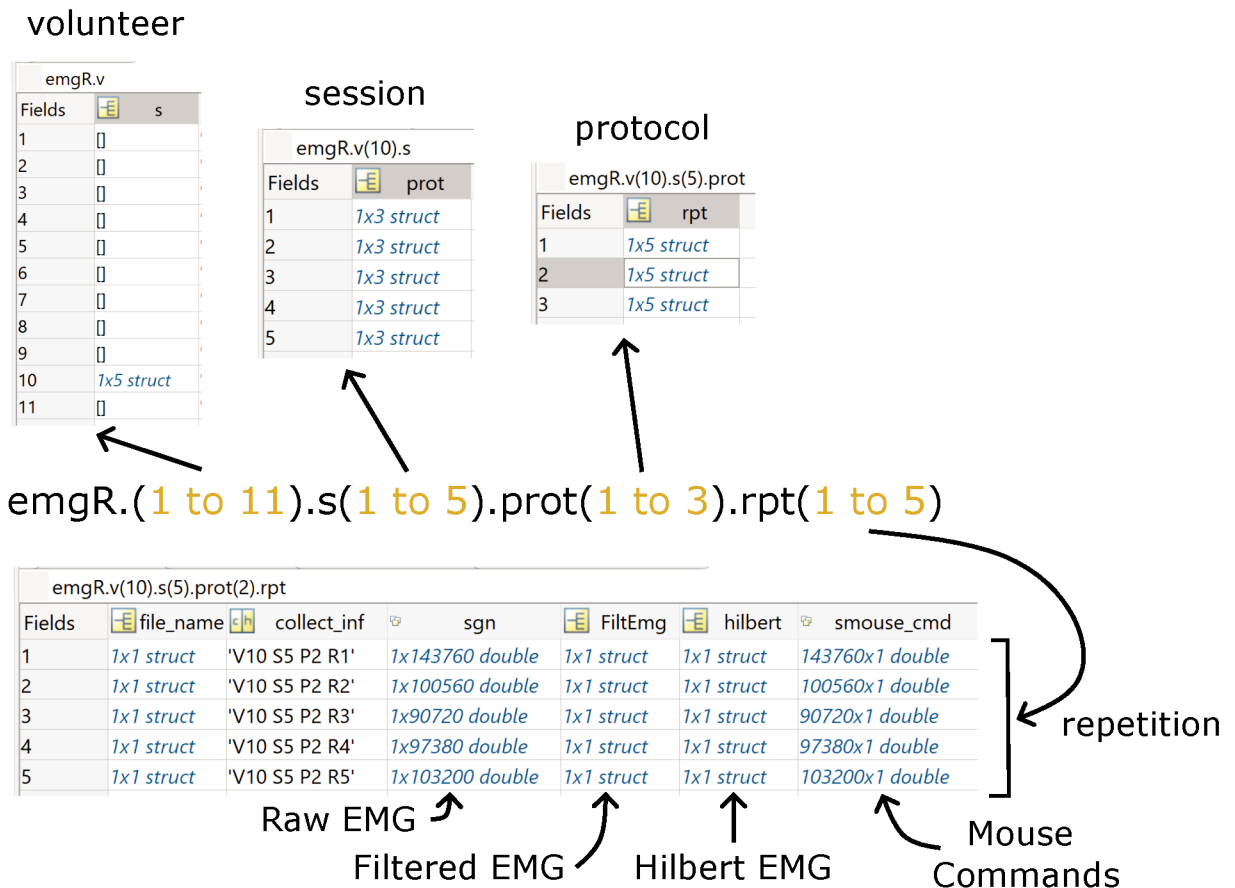


Figure C.10: EmgR data structure was created to store the data related to the EMG signal and other signals employed in the control of the myoelectric interface. The data presented are related to volunteer 10, session 5, and protocol 2 and illustrate the data of the five repetitions. The figure shows raw EMG, filtered EMG, Hilbert transform EMG signals, and mouse control signals.

sible to visualize, more easily, the duration of the contractions (C1, C2, and C3) and understand the application of the finite state machine illustrated in Figure C.6.

C.8.2 Structure EEG

The eegR structure contains EEG signals from the EEG cap and other motor learning data. The eegR data structure is also organized in the same way as the emgR structure. For example, Figure C.12 illustrates the data regarding the events of the tasks performed by volunteer 10, session 4, protocol 1 (2 cm target), and repetition 1. The illustrated data show that target 4 (“TARGET_4”), in field 10 of the table, emerged

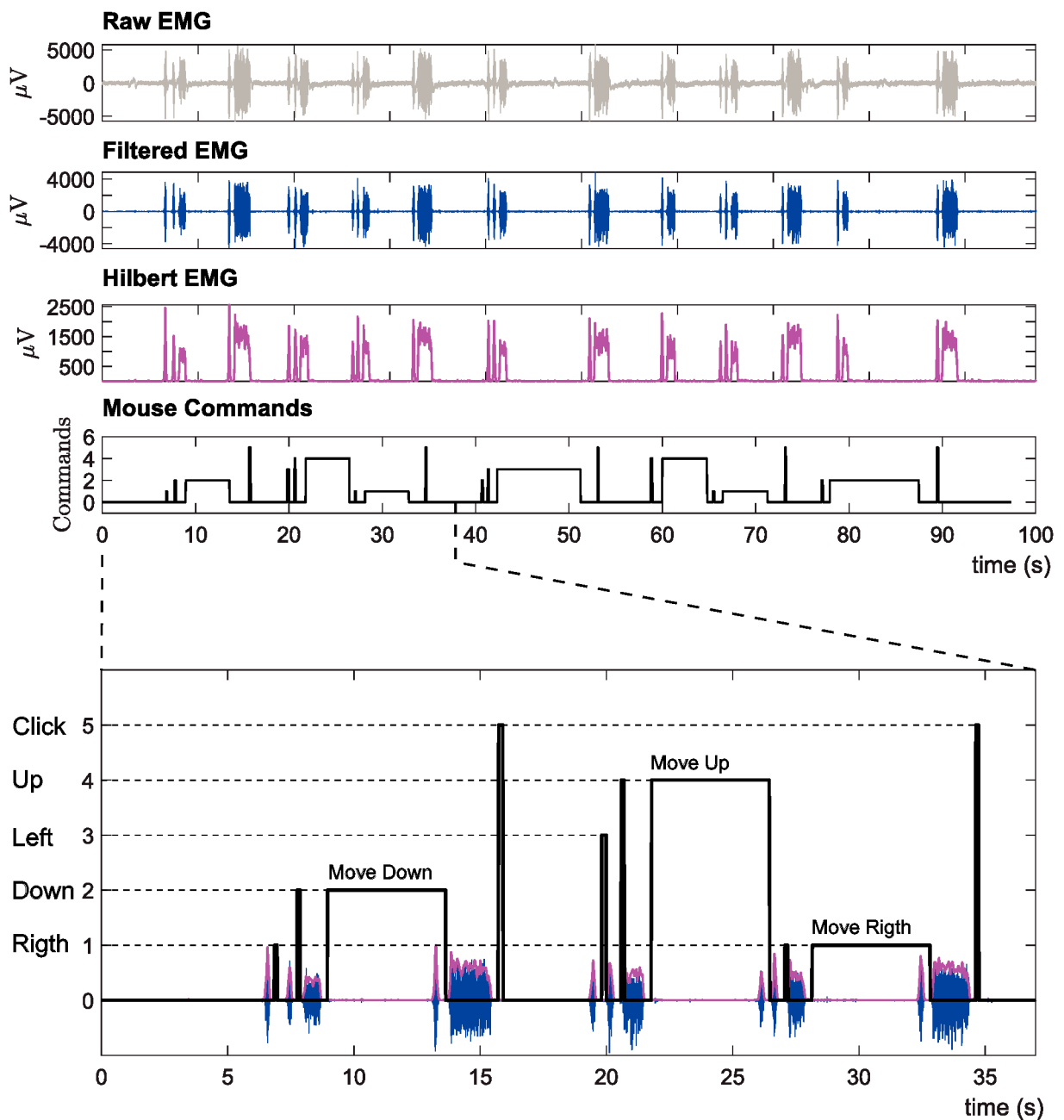


Figure C.11: It presents some signals of the emgR structure used in the mouse control by the myoelectric interface. The signals illustrated are: the raw EMG (sgn), filtered EMG (FiltEmg), EMG Hilbert transform (hilbert), and the mouse cursor control signals (smouse_cmd). The zoom given in the mouse command signal displays the levels of the mouse control signals in detail.

at time 4.396 s, positioned at $x_pos=342$ and $y_pos=323$, with equal width and height 84, and was clicked at time 15.615 s as indicated in field 9 (Figure C.12).

| eegR.v(10).s(4).prot(1).rpt(1).L_events | | | | | | |
|---|----------|------------|-------|-------|-------|--------|
| Fields | pos_time | type | x_pos | y_pos | width | height |
| 1 | 97155 | 'CLICK_IN' | 386 | 375 | 0 | 0 |
| 2 | 82124 | 'TARGET_4' | 342 | 323 | 84 | 84 |
| 3 | 79035 | 'CLICK_IN' | 683 | 47 | 0 | 0 |
| 4 | 65118 | 'TARGET_1' | 24 | 641 | 84 | 84 |
| 5 | 60016 | 'CLICK_IN' | 693 | 694 | 0 | 0 |
| 6 | 44403 | 'TARGET_3' | 660 | 641 | 84 | 84 |
| 7 | 39315 | 'CLICK_IN' | 981 | 375 | 0 | 0 |
| 8 | 19710 | 'TARGET_2' | 342 | 959 | 84 | 84 |
| 9 | 15615 | 'CLICK_IN' | 406 | 407 | 0 | 0 |
| 10 | 4396 | 'TARGET_4' | 342 | 323 | 84 | 84 |

Time Line ↑

Figure C.12: Data of learning events related to the activity performed by the volunteer 10, session 4, protocol 1, and repetition 1. The “type” column contains the targets (TARGET_ 1,2,3, and 4) and the “CLICK_IN” commands. Considering the indicated timeline, the “pos_time” column indicates when the target appears, for the targets, and when they are clicked, for “CLICK_IN”. The other columns refer to the characteristics of the items indicated in the “type” column.

C.9 Access to Database

The database, motor learning protocol, can be accessed via:

Andrade, Adriano de Oliveira, & Queiroz, Carlos Magno. (2022). Single channel approach for filtering electroencephalographic signals strongly contaminated with facial electromyography [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.7133259>.