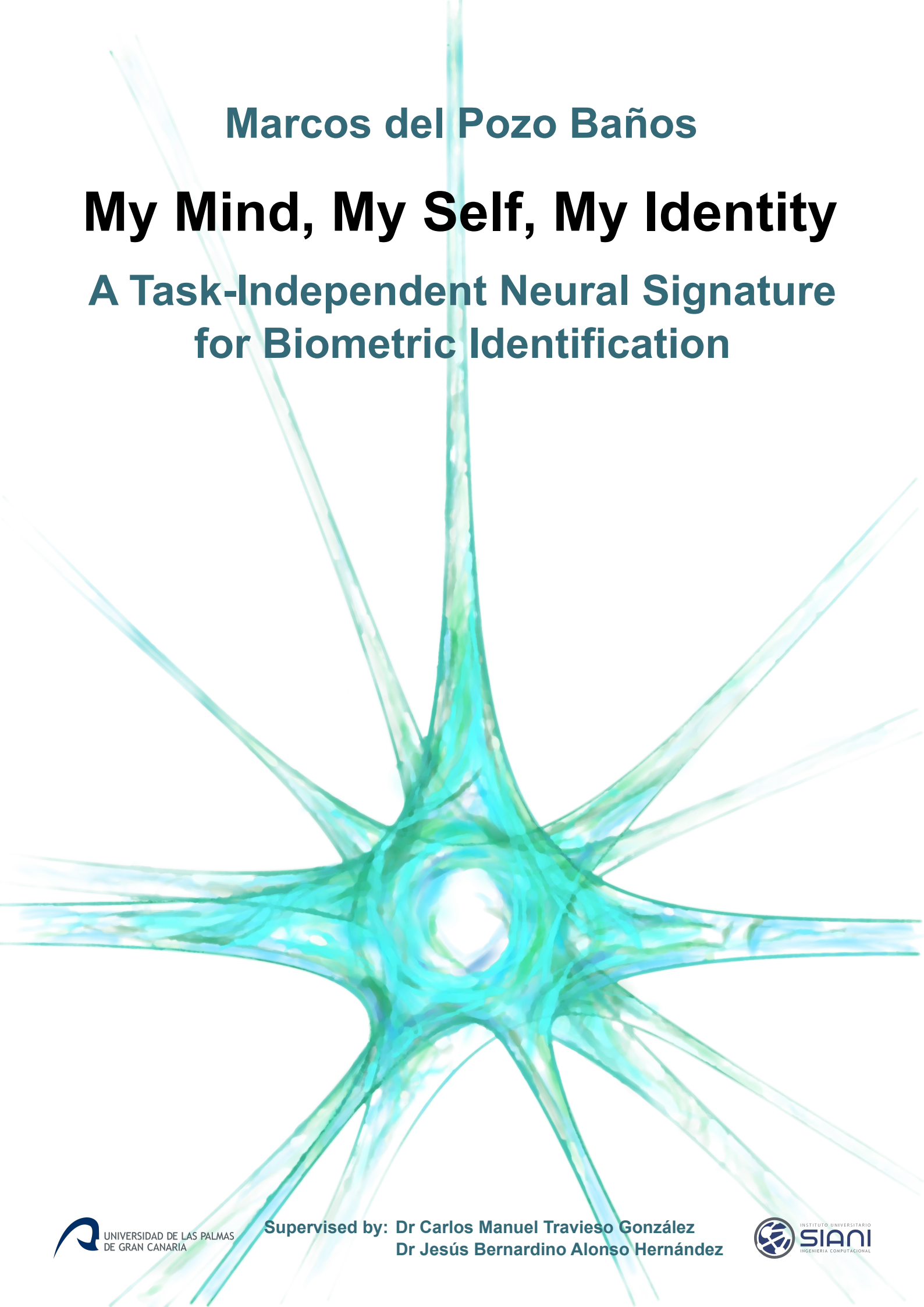
Marcos del Pozo Baños

My Mind, My Self, My Identity

A Task-Independent Neural Signature for Biometric Identification



UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

Supervised by: Dr Carlos Manuel Travieso González
Dr Jesús Bernardino Alonso Hernández

INSTITUTO UNIVERSITARIO SIANI INGENIERIA COMPUTACIONAL

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Instituto Universitario de Sistemas Inteligentes
y Aplicaciones Numéricas en Ingeniería

EDUARDO M. RODRÍGUEZ BARRERA, SECRETARIO DEL INSTITUTO UNIVERSITARIO DE SISTEMAS INTELIGENTES Y APLICACIONES NUMÉRICAS EN INGENIERÍA (SIANI) DE LA UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA,

CERTIFICA

Que el Consejo de Doctores del Instituto Universitario de Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería (SIANI), en su sesión de fecha 17 de septiembre de 2015, tomó el acuerdo de dar el consentimiento para la tramitación de la Tesis Doctoral titulada "My Mind, My Self, My Identity: A Task-Independent Neural Signature for Biometric Identification (Mi Mente, Mi Yo, Mi Identidad: Una Firma Neuronal Independiente de la Tarea Mental aplicada a la Identificación Biométrica)", presentada por el doctorando **D. Marcos del Pozo Baños**, dirigida por el Dr. D. Carlos Manuel Travieso González y el Dr. D. Jesús Bernardino Alonso Hernández, propuesta para acceder a la mención de doctor europeo, a la vista de la idoneidad y calidad de su contenido, interés y relevancia del tema a nivel internacional, y la realización de una estancia de más de un trimestre en un centro de investigación europeo fuera de España.

Para que así conste, y a los efectos oportunos se expide el correspondiente certificado a 18 de septiembre de 2015.

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

INSTITUTO UNIVERSITARIO DE SISTEMA INTELIGENTES Y APLICACIONES NUMÉRICAS EN INGENIERÍA

PROGRAMA DE DOCTORADO

DOCTORADO EN SISTEMA INTELIGENTES
Y APLICACIONES NUMÉRICAS EN INGENIERÍA

TESIS DOCTORAL

My Mind, My Self, My Identity: A Task-Independent Neural Signature for Biometric Identification

Mi Mente, Mi Yo, Mi Identidad: Una Firma Neuronal Independiente de la Tarea Mental aplicada a la Identificación Biométrica

AUTOR: D. Marcos del Pozo Baños

DIRECTOR: Dr. D. Carlos Manuel Travieso González
Dr. D. Jesús Bernardino Alonso Hernández

El Director **El Codirector** **El Doctorando,**

Las Palmas de Gran Canaria a 22 de septiembre de 2015

My Mind, My Self, My Identity: A Task-Independent Neural Signature for Biometric Identification

Marcos Del Pozo Baños

September 23, 2015

Acknowledgements

My mother has always been right to point out that I like to build castles in the air. So, when I was just beginning to conceive this thesis around 2009/2010, I was already thinking that I would not write an acknowledgements chapter. I regarded it too sentimental for a scientific work. Well, here I am now, unable to imagine this document without naming the people that have made it possible.

To my parents: José Manuel del Pozo Cubero and María Isabel Baños Quintana, for being the best providers. They have provided me with their genes (which have certainly played a part or two), with a splendid education, with never ending food, money and accommodation whenever needed without questions, and with their unconditional support for my *"flying castles"*.

To my partner: Dara Almeida Medina, for walking this journey alongside me. For taking care of my mental hygiene while I was in *"thesis-land"*, and making sure that I got at least some of that stuff that normal people have (I am learning now that they call it *"life"*). She also listened to my ir/rationales and pretend-to-listen when I got too boring, both of invaluable help.

To my supervisor: Dr Carlos M. Travieso, for many years of help in my career development. One has to put all this time in context to appreciate the amount of trust that he has put in me since the very beginning. I certainly owe him much more than just this thesis, and I see him as nothing less than my mentor.

To my co-supervisor: Dr Jesús B. Alonso, for his incalculable advice and directions. This work would not have been even remotely close to what it is without him.

To Dr Christoph T. Weidemann, for gifting me with the enriching opportunity of working at the Department of Psychology, Swansea University (Wales, UK). Also, for his crucial guidance during the statistical analysis of the results.

To W. Paul Boyce, for proofreading this document. We would not understand it without his input.

To all the people above I say thank you.

Preface

"Cogito ergo sum."

René Descartes (1637)

Throughout history, many philosophers have argued that thinking is the proof of our own existence. We opened this preface with René Descartes' statement that shares this sentiment and it is arguably the most famous example of the concept. Although a precise English translation will read as *"I am thinking, therefore I exist"*, it is commonly – and very conveniently for this case – translated as *"I think, therefore I am"*. Interpreting *"I am"* as representing the "*self*" as opposed to existence, we can borrow the quote to mean that what we think and who we are, are actually the same thing.

This is not the first study to link the *self* to the mind. One example is J. LeDoux's work Synaptic Self [1]. In it, LeDoux presents a compelling scientific argument to support the view that who we are is unambiguously determined by our brains. Unlike him though, we defend that the *self* is fully defined by the mind/brain, with physical aspects of the individual being mere vehicles.

Our *self* then is linked to our intellectual existence. Being subject to brain plasticity, it is in constant evolution. If we change the way we think, we change who we are. When this change is relatively small, we call it maturation, self-growth or just learning. But if the change dramatically affects the way we interact with the world, we will be perceived, perhaps even self-perceived, as a different person. Hence the expression *"he/she is not the same person since..."*.

In addition, our brain is possibly the pinnacle of our uniqueness. With its synapses shaped not only by nature, but also by nurture, even the brain networks of identical twins are different. Accordingly, as similar as twins can be, they think differently and are distinct individuals. The brain, as a unique entity, is therefore susceptible to being used as a means for identification.

The terms *"self"* and *"identity"* are treated separately in the philosophy literature [2]. In particular, an individual is commonly said to be composed of multiple *identities*, such as their nationality, religion, profession, race and sexual orientation. These identities together form the individual's *self*. However, although this may be convenient for some arguments, we believe it falls far from its practical definition; i.e. *"the state of having unique identifying characteristics held by no other person or thing"* [3], as a single one of these identities cannot define individuality.

Furthermore, the coupling of one's identity with groups of individuals, or categories, seems to contradict the pursuit of equality in societal terms. Conversely, we propose that the identity of an individual is solely defined by the way he/she interacts with the environment (i.e. thinks). Thus, all remaining bits of information are secondary, and most of them are consequences of the former. Those that are not; i.e. those that are out of the subject's control (nationality, race, etc.), should be left out of the definition of *self* or *identity*, as they do not necessarily reflect the way the individual feels or thinks; although they likely have an effect on it. Accordingly, we prefer the statement *"I am myself, who happens to be Spanish and European white"*, to *"I am Spanish and European white, and this is who I am"*.

Taking into account the definitions above, we consider that what we think (my mind), who we are (my self), and our identity (my identity) are concomitant terms, and are virtually the same thing.

Hence, it seems inconsistent to have our legal identity bounded to features of our body (not transcending the physical for the psychological). The ePassport relies on facial, fingerprint and iris recognition to verify the identity of an individual.

By our definition of identity, the previous are just convenient solutions, as far as such characteristics are relatively immutable, non-transferable and non-copyable. We argue then that a more direct and reliable way to measure who a person is, is by means of its brain activity.

Contents

1 Introduction **1**

1.1 Hypothesis . . . 3

1.2 Document structure . . . 3

2 State of the art: comprehensive review **5**

2.1 Review methodology . . . 7

2.2 EEG genetics . . . 10

2.3 Most used databases in the literature . . . 12

2.3.1 Poulos' database . . . 12

2.3.2 Zhang's database . . . 13

2.3.3 Keirn's database . . . 15

2.3.4 IIIaBCI03 database . . . 16

2.3.5 BCI2000 database . . . 16

2.3.6 Tottori database . . . 17

2.3.7 Lanzhou database . . . 18

2.4 Experiments and results . . . 18

2.4.1 Resting states . . . 19

2.4.2 Event related potentials . . . 23

2.4.3 Multiple-tasks . . . 29

2.4.4 Consumer EEG devices . . . 35

2.4.5 Other approaches . . . 37

2.4.6 Advances made in parallel to the present work . . . 38

2.5 Discussion . . . 41

2.5.1 Which are the subject specific traits of the EEG? 42
2.5.2 Where are these traits in terms of location? 42
2.5.3 Where are these traits in terms of frequency? 43
2.5.4 Are they constant across time? 44
2.5.5 Are they constant across cognitive tasks? 44
2.5.6 Which is the best design approach? 45
2.5.7 Medical versus consumer devices 46
2.6 Conclusions 46

3 Materials **49**
3.1 Raw databases 50
3.1.1 BCI2000 database 50
3.1.2 DEAP database 51
3.1.3 Keirn's database 51
3.1.4 P. Ullsperger's database 52
3.1.5 Yeom's database 52
3.1.6 Zhang's database 52
3.2 Preprocessing 53
3.2.1 Filtering 54
3.2.2 Channel rejection 54
3.2.3 Voltage normalization 55
3.2.4 Events rejection 55
3.2.5 Channel interpolation 56
3.2.6 Session rejection 56
3.2.7 Subject rejection 57
3.2.8 Channel selection and sampling frequency 57
3.3 Final databases 57

4 EEG artefact rejection: Localized Component Filtering **61**
4.1 Structure of a BSS EEG artefact rejection method 64
4.2 Localized Component Filtering concept 65
4.2.1 Implemented LCF prototype 68

4.3 Materials for the validation of LCF . . . 71
4.4 Experimentation methodology . . . 73
4.4.1 Evaluated systems . . . 73
4.4.2 Quantitative analysis of synthetic EEG data . . . 75
4.4.3 ERP analysis of real EEG data . . . 77
4.4.4 Time-frequency analysis of real EEG data . . . 78
4.5 Results and discussion . . . 79
4.5.1 Simulated EEG data results . . . 79
4.5.2 Real EEG data results . . . 81
4.5.3 Time-frequency analysis . . . 83
4.6 Conclusions . . . 87

5 EEG time-frequency exploration 91

5.1 Methodology . . . 92
5.1.1 A note on the application of artefact rejection methods . . . 93
5.1.2 Configuration of the PSD . . . 94
5.1.3 Representation of the time and frequency domains . . . 96
5.1.4 Properties of the discriminant information . . . 96
5.1.5 Task-independence of the neural signature . . . 98
5.1.6 Qualitative study . . . 99
5.1.7 Quantitative study . . . 99
5.2 Configuration of the PSD: results, discussion and conclusions . . . 102
5.2.1 Spectrogram width and number of FFT coefficients . . . 103
5.2.2 Window overlap . . . 106
5.2.3 Frequency range . . . 107
5.2.4 Review of spectrogram width with optimal parameters . . . 109
5.2.5 Review of EEG segment length with optimal parameters . . . 111
5.2.6 A comparison across systems . . . 113
5.2.7 Conclusions . . . 115
5.3 Representation of the time and frequency domains: results, discussion and conclusions . . . 115

5.3.1 Montages . . . 116
5.3.2 Spectral normalization . . . 118
5.3.3 Conclusions . . . 121
5.4 Properties of the discriminant information: results, discussion and conclusions . . . 121
5.4.1 Spatial distribution . . . 121
5.4.2 Frequency distribution . . . 123
5.4.3 Uniqueness . . . 125
5.4.4 Permanence . . . 127
5.4.5 Conclusions . . . 129
5.5 Task-independence of the neural signature: results, discussion and conclusions . . . 130
5.5.1 Task-independence . . . 130
5.5.2 Permanence of task-independence . . . 134
5.5.3 Conclusions . . . 137
5.6 Summary and overall conclusions . . . 138

6 Biometric system implementation 141
6.1 Methodology . . . 141
6.2 Baseline design . . . 143
6.3 Feature extraction methods: results, discussion and conclusions . . . 146
6.3.1 Cepstral coefficients . . . 146
6.3.2 Spectral envelope coefficients . . . 148
6.3.3 Other evaluated features . . . 153
6.3.4 Conclusions . . . 154
6.4 Other evaluated designs . . . 154
6.5 Experiments using all the available subjects . . . 157
6.6 Conclusions . . . 164

7 Discussion, advances and future steps 167
7.1 Discussion on the task-independent neural signature . . . 167
7.2 Discussion on the task-independent EEG biometric approach . . . 169

7.3 Generated knowledge 171
7.4 Future research 173

A Supplementary results and statistical tests **175**

B Career development **227**
B.1 Non-thesis specific development 227
B.2 Thesis specific development 228

Acronyms **231**

List of Figures **235**

List of Tables **245**

Bibliography **257**

Chapter 1

Introduction

The search of genetic traits within Electroencephalogram (EEG) has received great attention from the scientific community almost since the first human EEG recordings by Hans Verger in 1924 [4]. Unravelling this genotype-phenotype map will allow us to build advanced and inexpensive tools for understanding many diseases and diagnosing them early, especially those affecting the brain [5,6]. This is mainly because a tool based on the quantitative measure of EEG properties will be closer to gene function than the traditional interpretation of cognitive tests [5]. Crucially, these tools will also allow us to measure the evolution of patients and evaluate the effects of treatments in a quantitative manner.

The understanding of EEG genetic traits will also be applied on the development of biometric security systems. In this era of abundance of information and network communication, biometric systems have slowly become a necessity for controlling access to sensitive places and/or information. These methods are based on unreproducible, non-transferable characteristics of an individuals physiology, such as the fingerprints or facial features.

EEG-based subject identification is a relatively new biometric discipline which finds its origins in the advances of the above mentioned human genetics and clinical neurophysiology studies. A system based on a subject's EEG poses a very attractive design due to its potential for reliability and difficulty in artificial repli-

cation. Passwords will be harder to steal, as users will not need to perform any revealing action. Even if stolen, the system could be tuned to respond not to the passwords' semantic meanings, but to the subjects' specific EEG patterns. These are extremely hard to reproduce by another individual, if at all possible. Furthermore, if a user is forced to enter their 'password', their elevated stress levels, or other changes in baseline EEG activity, could be detected by the system which could then forbid access.

The main disadvantage of this design is the inconvenience of the recording method. EEG devices, although strictly non-invasive, are relatively more invasive than other biometrics. When considering clinical or experimental EEG devices, the preparation of the recording session can be an arduous task, requiring the application of conductive gel to the user's scalp as well as precise trial and error checks on signal clarity. This has been partially solved by the proliferation of consumer EEG devices, which significantly reduce the preparation time and make use of dry sensor technology. Although currently offered consumer devices have lower quality signals than clinical devices, we can expect this to be gradually overcome by future advances in the sensors field.

The current work is a study of the identity traits within EEG, and more specifically, within its time-frequency representation. This is motivated by:

- the importance of understanding the genotype-phenotype map of neural activity, which will lead to advances in medicine and psychology;

- and the potential offered by neural biometric identification, especially in systems that require a robust security method and those that integrate any form of Brain Computer Interface (BCI) and will directly benefit from it, such as remote health monitoring/diagnosing devices or games.

1.1 Hypothesis

Throughout this research, we will provide evidence supporting the following hypothesis:

> *There exists, within the human EEG, a task-independent pattern concomitant to the individual's identity.*

The second part of this hypothesis *"... concomitant to the individual's identity"*, refers to the neural activity, as recorded by the EEG, being defined by the subject's phenotype, and therefore, being unique.

The first part of the hypothesis *"... a task-independent pattern ..."*, refers to the mentioned unique activity being persistent across mind states, cognitive tasks and experimental conditions — from now on, we will refer to these three as simply 'tasks'. Persistent not in a strict sense, but in a way that fluctuations across tasks are comparatively smaller than differences in the neural signatures across individuals.

1.2 Document structure

We have carried out the current research in four main phases, which define the different chapters of this document:

1. **State of the art:** (Chapter 2) Due to this being a relatively new biometric modality, the literature lacked a review work summarizing the most important findings on the subject. In fact, the state of the art was characterized by the absence of a structured methodology. Authors had focussed their efforts on applying new algorithms and architectures, without understanding the fundamental properties of the targeted information, or basing this on genetic and neurophysiologic studies that analysed the data from a different point of view.

2. **Recollection of publicly available EEG databases:** (Chapter 3) In a bid to identify the commonalities as intrinsic properties of the neural signature rather than idiosyncrasies of the data sets, we performed the study presented here on 6 publicly available databases, each of different nature. The set of databases included: resting states, motor and problem-solving tasks, event related potentials and elicited emotions.

3. **EEG time-frequency exploration:** (Chapter 5) To cover the weaknesses identified in the state of the art, the bulk of this research encompasses an in depth analysis of the discriminant information within the time-frequency representation of the EEG. This includes numerous experiments, each designed to evaluate a single or a small set of qualities under controlled conditions.

4. **Biometric system implementation:** (Chapter 6) Finally, to evaluate the advances made during the exploration of the subject traits within the EEG, we applied our findings to the design of a biometric verification system.

In addition to the above points, a parallel research on the rejection of noise from the EEG signal was conducted at the Department of Psychology, College of Human Health and Science, Swansea University (Wales, UK). This resulted in a new artefact rejection approach: **Localized Component Filtering (LCF)**, which was ultimately used for the preprocessing of the data sets used in the EEG time-frequency exploration (chapter 4).

The above phases and research correspond to this text's main structure, each representing an individual chapter. To finalize, we will present a global discussion on the subject, enumerate the produced innovations and outline some potential research avenues for the future (chapter 7).

Chapter 2

State of the art: comprehensive review

This chapter presents a comprehensive review of the state of the art on EEG biometric identification. As well as presenting the evolution of this field (from 1980 to 2015), we provide an overview of related genetic and neurophysiology studies. We also propose a categorization of methods based on the tasks used in the experiments. The most frequently used databases, some of which are public, are identified and named to allow the comparison of results from these and future researches. We show that, although basic questions remain unanswered, EEG has been found to contain subject specific information that can be used for discrimination. Moreover, approaches such as a multi-day-session training, the fusion of signals from different electrodes and frequency bands, and the application of Linear Discriminant Classifiers (LDCs) and Support vector machine (SVM) are recommended to maximize the system's performance. Overall, the goal of discrimination is more challenging than initially expected, as it relies on data extracted from complex, heterogeneous EEG traits. Such traits are the results of elaborated models of inheritance, which makes discrimination very sensitive to variables such as time, frequency, sensor location, recording condition and system design.

In particular, we focussed on finding answers to the following questions:

- What are the subject specific traits of a given EEG signal?
- Where are these traits in terms of location?
- What are the frequencies of these traits?
- Are they constant across time?
- Are they constant across cognitive tasks?
- Which is the best design approach?

Although there exist overviews of the field, each centred their attention on a small set of the algorithms applied over a specific approach for EEG based biometric systems [7–10][1]. In contrast, this is a broad study considering all the publications on the matter, covering all the techniques and strategies as well as their relationships. From this, we will show that most of the questions asked above were unanswered or only partially answered.

The bulk of the analysis presented here deals with articles published before 2014. This is the date when the original review of the state of the art was finalised, the conclusion of which led to the EEG time-frequency exploration (section 1.2). To complete the overview, we have added a review of the studies published during the period spanned by the work presented here: from 2014 to April 2015, in section 2.4.6.

We will first detail the review methodology utilised in this research. Following that, we will present the major findings on the genetic traits of the human EEG signal. The individual main EEG databases used in the considered literature, and the results obtained on each of the studied approaches will then be presented. Subsequently, all the findings presented here will be pieced together in a bid to answer the questions posed above. Finally we will present our conclusions.

[1]Since the publication of this review of the literature, two more complete reviews have been published [11, 12].

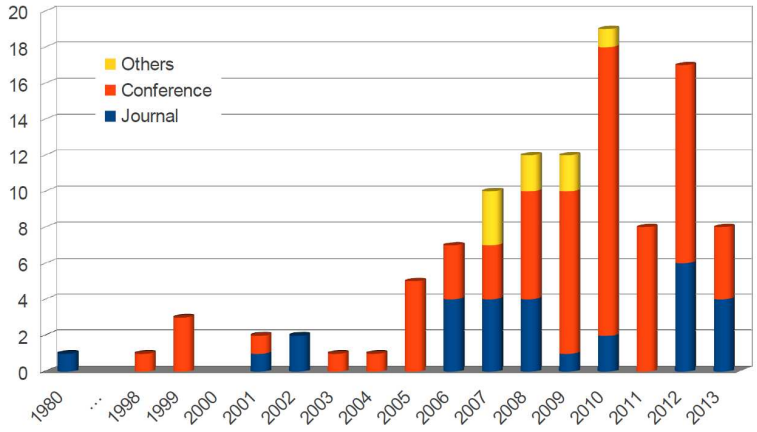


Figure 2.1: Number of publications per year up to 2013. Categories: "conferences", "journals" and "other" are considered separately.

2.1 Review methodology

The research was executed in two phases. Firstly, we gathered literature on genetic and neurophysiological fields regarding the phenotypic features of EEG. This allowed us to establish a scientific foundation for the topic and provided a non-technical point of view of the problem. The book "Genetics and the Electroencephalogram" by F. Vogel [13] is of special relevance here. This book is a comprehensive resource in the genetics field, and provides a compilation of findings on the genetic, clinical and neurophysiological aspects of EEG.

The second phase was comprised of the search of biometric literature. After an in depth exploration, 108 works were finally found between 1998 and 2013 (fig. 2.1). This emphasizes the novelty of the area. Instead of filtering the sources to include only journals, as is typically the case in these studies, all 108 works were reviewed, from which 87 were finally included.

We propose the following classification of the literature based on the recording tasks:

- **Resting states:** Resting with Eyes Closed (REC) and Resting with Eyes Open (REO) have been extensively explored in the literature. These were the first tasks used for identification via EEG.

- **Event Related Potential (ERP):** Brain activity elicited by an external stimulus has also been used for identification via EEG. Specially EEG signals triggered by stimulation of the visual sense or Visual Evoked Potential (VEP).

- **Multiple-tasks:** This includes studies based on EEG recorded during various cognitive tasks. Usually, they focused on the performance difference between recorded tasks and on the fusion of different tasks to boost the system's performance.

- **Indirect:** Some authors have tried to identify users by recognizing a thought password or responses to specific stimulus, rather than subject-specific EEG traits.

- **Others:** This category includes reviews, dissertations, reports and any other work that is related to the subject but does not present any new system architecture or experiment results.

We propose a further differentiation based on the hardware used: **medical** or **consumer** equipment. The latter represents a cheaper alternative which requires no conductive gel and is considerably easier to use. However, they also provide lower signal to noise ratios and sensitivity. Accordingly, it seems advisable to keep in mind which hardware has been used in each case when comparing results. Conveniently, all these categories can also be used to classify genetic and neurophysiological works on the matter.

Figures 2.2 and 2.3 show the distribution of publications across each category. Note that these figures do not represent percentage values, as some works fall into several categories. It can be seen that the vast majority of studies have focussed on REC and VEP (ERP) modalities. This is consistent with genetic and neurophysiological studies. Multiple-tasks studies have also received special attention, as authors attempted to find the most appropriate paradigm for their systems.

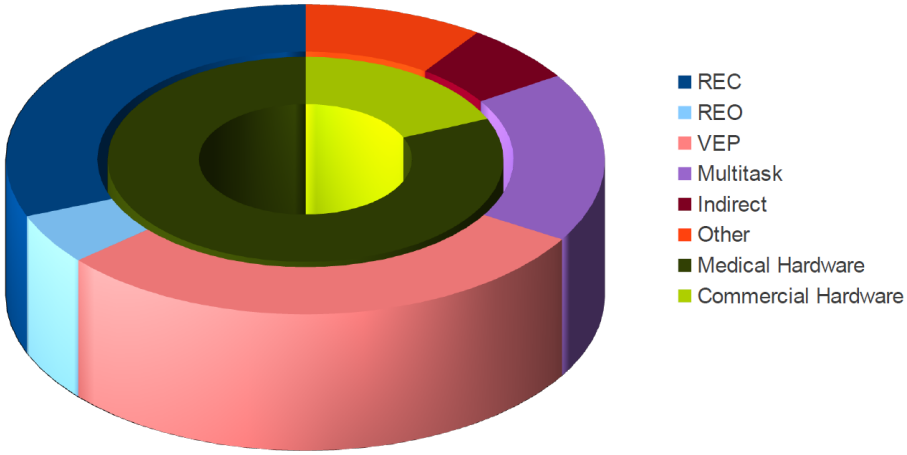


Figure 2.2: Distribution of publications per category up to 2013. A single publication can fall into multiple categories.

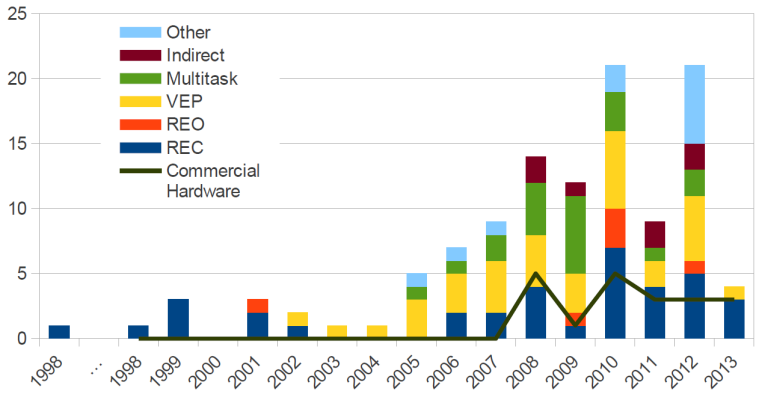


Figure 2.3: Distribution of publications per category and per year up to 2013. Single publication can fall into multiple categories.

In addition, as commercial EEG hardware have relatively recently appeared, the proportion of studies using them is obviously lower than those using medical equipment.

To maximize clarity of the analysis, we tagged and clustered publications into research groups. This grouping convention is based on any overlap of authors across relevant publications. This also helped to identify databases, which had no names and tended to be difficult to track. The identification and naming of databases allowed the comparison of results from different publications.

2.2 EEG genetics

Identifying the genetic traits of EEG has proven to be an arduous task. These are complex heterogeneous traits, as they are the result of elaborate models of inheritance. For example, some genes appear to have different effects in different brain areas and EEG frequencies. In addition, exogenous factors have also been proven to influence human EEG signals and have to be considered when evaluating the results [6].

Twin studies have proven to be immensely beneficial in the understanding of this genotype-phenotype relationship. H. Davis and P. Davis were the first to study the EEG signals of twins [14]. After evaluating a number of EEG traits; mainly based on distinguishing marks of the posterior rhythm, the authors concluded that the resting EEG of Monozygotic (MZ) twins were identical; i.e. as similar as recordings from an individual across time (fig. 2.4). This was not the case for dizygotic (Dizygotic (DZ)) twins, whom had less similarity between their EEG signals, although they were significantly higher than those of unrelated subjects. These results were later confirmed by numerous works across time, regardless of their methods of examination: visual inspection; measures on paper; or computer evaluations [13].

Finally, the first direct connection between the morphology of EEG signals and the genotype of subjects was made by F. Vogel [16]. Some of the effects of time over the EEG signals also proved to be concordant across MZ twins. Interestingly, these results were also found in studies of twins reared apart, excluding the effects of exogenous factors. Similarities were also present in reactions to hyperventilation, photic stimulation and during sleep. The Non-REM sleep EEG patterns have been reported as more heritable than those of resting EEG or ERPs, especially in the range between 8 and 16 Hz [17].

Familial studies reasserted the findings of experiments conducted with twins. EEG traits such as the mean spectral power and frequency values of alpha and beta

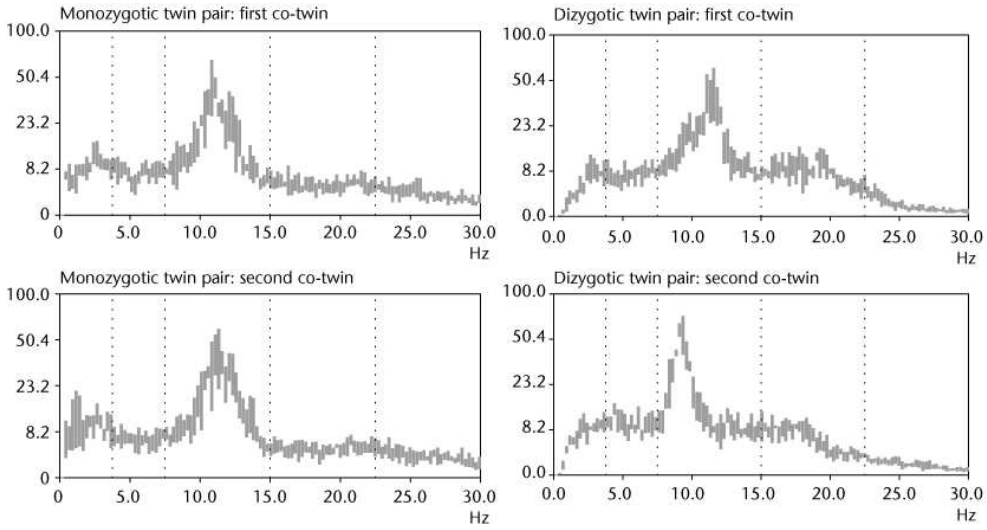


Figure 2.4: Individual EEGs of MZ (left) and DZ (right) twins. Shaded areas represent variation [15].

bands were found highly heritable [13, 18]. Moreover, some of these parameters seem to form a continuous phenotypic range rather than clusters in well defined discrete classes, so that traits of family members are more alike than those of unrelated individuals [19].

Other authors focussed on the brain response to external stimulus (ERPs). VEPs and Auditory Evoked Potentialss (AEPs) were recorded from twins and families, showing similar results to those obtained with the resting EEG recordings [5, 13, 20, 21]. This came as no surprise given the strong relationship between the resting EEG and ERPs described previously by F. Vogel.

Computational tools produced some interesting works. Those of H.H. Stassen et. al. [22–24] are of special relevance. In them, the authors provided quantitative evidences of such comparability between twins and family members by applying similarity measures. Indeed, their first work could be regarded as unintentionally being the inaugural publication of an EEG biometric system [22]. However, they used it as a tool for the evaluation of the inheritance of EEG traits, without much concern about performance as a security system in real scenarios. Other works using computational tools focussed on the classification of Vogel's proposed

EEG spectral variants: alpha and beta variants [16, 25]. These major categories were then divided into sub-classes based on finer details of the activity, such as particularly slow alpha rhythms.

On top of so the so far described complexity of the problem, the EEG signal seems to change with maturation [13, 21]. Furthermore, some of these changes may not be homogeneous across subjects, rhythms, or brain areas. To remove this variable factor across time and subjects, normalizing methods have been proposed [26]. Still, some works claimed to have found a remarkable stability of EEG spectral distribution between sessions recorded more than a year apart [27].

Overall, neurophysiological research suggests that alpha power and peak frequency over occipital regions present the strongest heritability [6, 13, 21]. This laid the foundations of the first works on EEG biometric identification. After all, having detected the phenotype traits of EEG signals, it may be enough for a biometric system to detect, and differentiate between, such features.

2.3 Most used databases in the literature

Some databases have been extensively utilised in the field of EEG based subject identification. However, as most of them had no name, they were difficult to track. In some cases, studies used the same data set but with a different number of subjects or channels, making the tracking even harder. Here, we identified the most used databases in the literature and propose names to them (table 2.1).

2.3.1 Poulos' database

Poulos' database was used in the first series of published works on EEG biometric identification by M. Poulos et. al. back in 1998, and it has been used by this group in many subsequent publications (hence the proposed name). The original purpose of this database is unclear, as it was never specified by the authors. It is unknown whether it was created for the sole purpose of subject identification, or

Table 2.1: Databases used on multiple publications of the relevant literature.

DDBB	**# subjects**	**Categories**
Poulos'	4 users + 75 impostors	REC
Zhang's	125 (48 healthy + 77 alcoholics)	ERP, VEP
Keirn's	7	Multiple-tasks
IIIaBCI03	3	Multiple-tasks
BCI2000	109	Multiple-tasks
Tottori	23	Consumer EEG; REC
Lanzhou	11 users + 11 impostors	Consumer EEG; REC

if, as with many databases used in the literature, it was part of a data set collected with other aims.

This is a small database composed mainly of only 4 registered users (here referred to as *SetR*), from which 45 recordings were taken. Each session consisted on 3 minutes of REC while the voltage difference from electrodes 02 and Cz was recorded at a sample rate of 128 Hz. The database also contains a second subset of data (*SetX*) composed of a single recording session from 75 different subjects used as impostor attacks in experiments.

2.3.2 Zhang's database

X.L. Zhang recorded two databases for the purpose of studying visual memory [28] and the effects of alcoholism on memory [29]. It is unclear whether these studies actually used two independent databases or if the latter is an extension of the former. However, as they were taken under the same exact conditions and their only difference resides in the number of subjects, they could be considered the same for the practical purpose of results comparison. Thus, from here on, they will both be referred to as a single database named Zhang's database, which is publicly available [30].

In this case, subjects were exposed to visual stimulus consisting of black and white images taken from the set of Snodgrass and Vanderwart [31]. These images

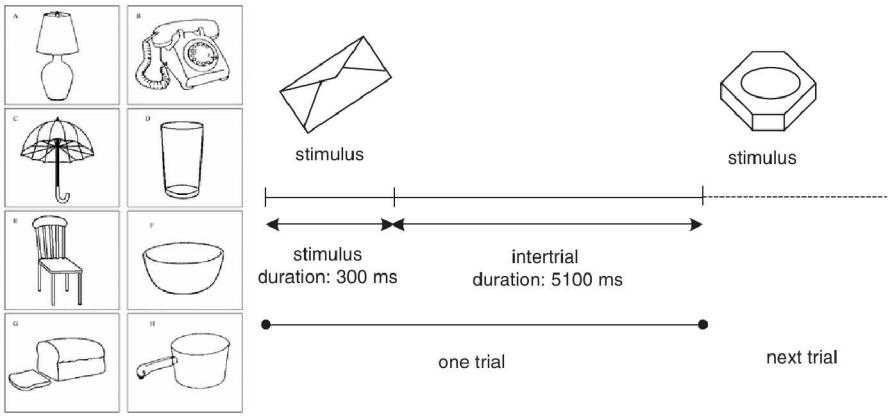


Figure 2.5: Original experimental paradigm of Zhang's database. Snodgrass and Vanderwart image samples (left). Representation of the paradigm (right).

represent easily identifiable objects with defined verbal labels (fig. 2.5). The trials consisted of two stimuli; *S1* and *S2*, each lasting 300 milliseconds and separated by 1.6 seconds. The subjects were asked to determine whether *S1* was the same as *S2*. In some cases, only one stimulus was presented (fig. 2.5).

Forty trials were recorded from each subject, leaving 3.2 seconds between trials. EEG pre-stimulus data was kept as baseline from 190 milliseconds before any stimulus was presented, and 1.44 seconds of EEG post-stimulus data was registered from the moment a stimulus was presented. A set of 61 channels were used, with all channels referenced to Cz. Data was sampled at 256 Hz and hardware filtered between 0.02 and 50 Hz.

This was the first database used for subject identification with a recording paradigm different than REC and REO. With a total of 125 subjects (48 healthy males 25.81 $\pm$ 3.38 years old and 77 alcoholic patients 35.83 $\pm$ 5.33 years old[2] [29]), it also represented an important step forward in the size of the databases used for this problem at the time. Having said that, the use of alcoholic patients on biometric experiments could be problematic due to the effects of alcoholism on EEG recordings (this will be further discussed in section 2.5).

[2]X.L. Zhang reported on [28] 14 males (24.3 + 3.1 years old) and 14 females (23.2 + 1.7 years old), all healthy. This was the only difference found between the databases.

There are a number of works that may have used this database [32–45]. The work published by A. Zuquete et. al. [46] used a database that is similar enough to allow for a direct comparison of results, as the only differences found in the experimentation was the inter-stimulus time (increased to 5.1 seconds) and the length of the EEG recordings (1 second post stimulus onset). In fact, Zuquete's database is also composed of healthy and alcoholic subjects.

2.3.3 Keirn's database

Z.A. Keirn recorded EEG signals from 7 subjects (6 males and 1 female between the ages of 21 and 48) while they were performing different mental tasks. Their aim was to explore new human-machine interaction through the brain [47]. Although this is a small database, its relevance resides on the multiple-tasks recording paradigm. This allowed R. Palaniappan (the only author that had used it so far[3]) to study the effects of different mental tasks on the problem of EEG based identification [48–51].

A total of 5 tasks were recorded. Each repeated 5 times under both REC and REO on every session. Two sessions were taken from each subject in a time span of 2 weeks. The tasks were:

T1 *Baseline measurements.* This task was taken as a baseline for comparison. In this case, subjects were only asked to relax.

T2 *Complex problem solving.* Subjects were asked to mentally solve non-trivial multiplication problems.

T3 *Geometric figure rotation.* Subjects were presented with an image of a 3-dimensional complex object and were asked to mentally rotate it.

T4 *Mental letter composition.* Subjects had to mentally write a letter to a friend or a family member.

[3]As detailed in subsequent sections, we have also made use of Keirn's database.

T5 *Visual counting.* Subjects were asked to visualize numbers being written on a blackboard sequentially, with the previous number being erased before a new number is written.

Only the EEG signals from electrodes 01, 02, P3, P4, C3 and C4 were captured on these experiments. Data was hardware filtered between 0.1 and 100 Hz and sampled at 205 Hz. The EEG signal was recorded during 10 seconds of each task. The database is publicly available [52].

2.3.4 IIIaBCI03 database

IIIaBCI03 database is another multiple-tasks database that has been used on different studies [44,53–57]. It is the data set IIIa of the BCI competition of 2003 [58], provided by the Laboratory of Brain-Computer Interfaces of the Graz University of Technology (Gert Pfurtscheller, Alois Scholögl).

The database is composed of only 3 subjects performing the following 4 tasks:

T1 : Imagining the movement of the left hand.

T2 : Imagining the movement of the right hand.

T3 : Imagining the movement of the foot.

T4 : Imagining the movement of the tongue.

Each subject performed each task 60 times. EEG was recorded by 60 electrodes (+4 references) at a sampling rate of 250 Hz. The signal was hardware filtered between 0.5 and 100 Hz and with a 50 Hz notch filter.

2.3.5 BCI2000 database

The BCI2000 data set consists of 109 subjects performing different motor/imagery tasks [59–61]. A total of 64 EEG electrodes, as per the international 10-10 system,

were recorded. One of the mastoid channels was used as reference. Recordings were taken with a sampling rate of 160 Hz.

Each subject performed 14 experimental blocks: two one-minute baseline blocks (one REO and one REC), and three two-minute blocks of the four following tasks:

T1 A target appears on either the left or the right side of the screen. The subject opens and closes the corresponding fist until the target disappears. Then the subject relaxes.

T2 A target appears on either the left or the right side of the screen. The subject imagines opening and closing the corresponding fist until the target disappears. Then the subject relaxes.

T3 A target appears on either the top or the bottom of the screen. The subject opens and closes either both fists (if the target is on top) or both feet (if the target is on the bottom) until the target disappears. Then the subject relaxes.

T4 A target appears on either the top or the bottom of the screen. The subject imagines opening and closing either both fists (if the target is on top) or both feet (if the target is on the bottom) until the target disappears. Then the subject relaxes.

So far, this database have only been used by three studies [62–64], but it is one of the largest public databases in terms of number of subject.

2.3.6 Tottori database

Tottori database was the first to be recorded with a consumer EEG equipment and used for the problem of subject discrimination. It was collected in the Tottori University (Japan) by C. Miyamoto et. al. and was used in several studies in the practical application of EEG for identification [65–68].

The authors recorded data between 1 and 24 Hz at a sample rate of 128 Hz. The hardware had a minimum and maximum voltage range sensibility of 5 and 80 uVpp respectively and it only contained the frontal electrodes Fp1 and Fp2.

The recording sessions were kept simple, in order to make the EEG identification process more practical. Ten sessions were registered the same day from 23 subjects while they were REC. The last minute of a 3 minute recording was used for the experiments.

2.3.7 Lanzhou database

The Lanzhou University (China) database was also recorded with consumer EEG equipment. In this case, a Nexus-4 machine was used. This is a wireless device connected through bluetooth with a sample rate of 256 Hz. In a bid to simplify the process for practical applications, only electrodes Cz and A2 were collected [69–72].

Again, EEG was recorded while subjects were REC. The database contains 11 registered users (6 males and 5 females between 20 and 24 years old) and 11 intruders. Five recording sessions distributed over a few days were taken for both users and intruders. Around 6 months later, an extra recording session was implemented only for users, not intruders.

2.4 Experiments and results

This section provides a detailed review of the experiments and results in the state of the art on EEG based identification following the proposed categorization. For the sake of clarity and in order to easy the tracking of events, authors have been clustered in *groups*[4].

[4]The name of the main author at the moment of their first publication was used as a convention for naming groups.

2.4.1 Resting states

The first publication available on automatic EEG biometric identification was in 1980 by H.H. Stassen [22]. He experimented with EEG signals recorded by two posterior electrodes from a group of 82 subjects from 4 different psychiatric diagnostic groups. By building a region, delimited by the maximum and minimum Power Spectrum Density (PSD) of each subject, he evaluated the level of similarity between subjects. He finally reported a classification rate of around 90%.

However, Stassen's was an isolated work meant to build a tool for future genetic and neurophysiological studies. The first series of related papers started in September 1998 by M. Poulos et. al. [73]. Their work was based on previous findings about the inheritance of EEG traits, which stated that the posterior rhythm is the most genetically determined trait of the resting EEG. In a bid to facilitate the arousal of such rhythm, a REC database was used. The proposed system relied on the Absolute Power Spectrum (APS)[5] of the alpha rhythm (7-12.5 Hz) recorded by the O2 channel and classified by a Computational Geometry based classifier (CG) algorithm. When tested on 4 subjects of Poulos' database *SetR*, this approach reached a success rate of 95% with regard to classification. The authors also reported a 96.2% impostors rejection rate on the *SetX*.

These results encouraged Poulos' group to keep investigating the usage of EEG as a biometric modality for subject discrimination. In the following years, they published a set of experiments interchanging coefficients of the PSD and Auto-Regressive (AR) models as descriptors and CG and Artificial Neural Network (ANN) as classifiers [74–77]. Systems were always applied over the alpha rhythm and tested using Poulos' database. In these studies, they explored for the first time the differences between sub-bands of the alpha rhythm, concluding that the mid-range sub-bands provided the most discriminative information. In a later work, Poulos et. al. used a spectrum range that included, for the first time, all the main

[5]PSD refers to frequency spectral vector, while APS refers to a single power value within a band.

Table 2.2: Systems used on Poulos' database. Columns specify: the name given to the system for future reference (Sys.), filtered frequency bands (Freq. [Hz]), extracted features (Feat.), applied classifiers (Cls.) and references of papers that used each design (Ref.).

Sys.	**Freq. [Hz]**	**Feat.**	**Cls.**	**Ref.**
Pou1	[7-12.5]; 0.3 Hz wide sub-bands	PSD	CG	[73]
Pou2	[7-10]		LVQ-ANN	[74, 77]
Pou3	[8-11]			
Pou4	[9-12]			
Pou5	[7.5-12.5]	AR		[75]
Pou6	[7.5-12.5]; 1 Hz width sub-bands		CG	[76]
Pou7	[1-30]		LVQ-ANN	[78]
Pou8		Bilinear-AR		

brain waves (i.e. delta, theta, alpha and beta), but failed to evaluate each band individually [78].

Unfortunately, Paulos' team's results cannot be compared directly, as different experimentation procedures were used on each occasion (tables 2.2 and 2.3). However, they seem to suggest that PSD coefficients provide better accuracy than AR models. Both methods have now been widely used by many researchers.

However, the fact that Poulos' database contained only 4 subjects prevented the establishment of any robust conclusions. In 2001, R.B. Paranjape et. al. published a study that tested 40 subjects' REO and REC [79]. The system used AR coefficients as features and a Discriminant Function Analysis for classification. The authors noted that an increase in the order of the AR model was necessary to accommodate the rise in the number of registered users. Using an order of 15, they achieved an 85% success rate. This provided more solid evidence for the possibility of performing subject discrimination via EEG.

Paranjape et. al. suggested that better performance would be expected if more than a single channel was used. In 2006, G. Mohammadi et. al. explored this [80]. AR coefficients and ANN were again used on a 10-subjects REC database. Authors reported accuracies between 80 and 95% (depending on the order of the AR model)

Table 2.3: Results with Poulos' database. Success rate (Succ.) of systems in table 2.2 on classification and verification experiments. The percentage of training samples and/or the Cross-Validation (CV) method used on each case is specified (Tr.% / CV). In this case, training percentage is specified as *SetR*% | *SetX*% (section 2.3.1). For works [74] and [77], only the best results are shown.

Sys.	**Classification**		**Verification**	
	Tr.% / CV	**Succ.**	**Tr.% / CV**	**Succ.**
Pou1	55.56 | 40	95.00%	Not tested	
Pou6	22.23 | 0	47.91%		
Pou2	55.56 | 0	83.75%	55.56 | 33.34	87.50%
Pou3		86.25%		**89.28%**
Pou4		86.25%		87.85%
Pou2	44.45 | 0	91.00%	Not tested	
Pou3		94.00%		
Pou4		**95.00%**		
Pou7		68.00%	44.45 | 40	79.28%
Pou8		78.00%		80.00%
Pou5	Not implemented			79.28%

when one posterior channel was used, and between 85 and 100% by fusing two or more channels. In addition, they noted by visual examination that electrodes from the back of the scalp provided better performances than electrodes from anterior regions. This assertion was later proved empirically by several works [81, 82] and was in line with the fact that the alpha rhythm is more prominent in occipital areas.

In 2009, P. Tangkraingkij et. al. published the first detailed study on the spatial distribution of discriminant information [83]. Their system applied Independent Component Analysis (ICA) directly over the raw EEG for feature extraction and relied on an ANN for classification. When the Joint Approximate Diagonalization of Eigenmatrices (JADE)-ICA algorithm was employed, the system achieved 100% accuracy for a 20-subjects REO database using all available 16 channels. When testing different channels combinations, they managed to retain perfect classifica-

tion using only 3 channels. The authors concluded that the most discriminative channels under the REO are Fp1, P3 and C4.

However, a year later, M.K. Abdullah et. al. reported a significant drop in performance when moving from central to parietal electrodes under REO, but not on a REC scenario [84]. Contrary to P. Tangkraingkij et. al., they concluded that the P3 electrode should be dropped on REO. The authors also compared the performance between left and right hemispheres, which were found not to be statistically significantly different. In addition, REC outperformed REO in general terms. Nevertheless, on a second work with different processing techniques, their results failed to replicate some of these findings, specially those regarding left versus right hemispheres and REC versus REO [85].

In 2011, P. Campisi et. al. also examined the contribution of different brain areas and different frequencies [81]. Their results showed again a decrease in performance when moving from posterior to anterior areas. This decrease was more pronounced when frequency rhythms above 33.33 Hz were removed from the EEG signals. The best results were obtained with the temporal triplet T7-Cz-T8 and a cut-off frequency of 33.33 Hz (96.08% accuracy over a 48-subjects database). Results also showed an increase in performance between 4 and 9 percentage points between the regular AR coefficients and their Reflection Coefficients (RC) counterpart.

In a later work, Campisi's team bolstered their conclusion on the spatial distribution of discriminant information [82]. The results also showed that performance was maximized using the alpha and delta rhythms, followed by the theta rhythm and finally the gamma rhythm. When the selective filter was set to include more than a single rhythm, the performance of the system increased. In particular, the best outcome was obtained when all 4 rhythms were included and electrodes O1-POz-O2 were used, reaching almost 99% accuracy.

2.4.2 Event related potentials

In 2002, R. Palaniappan et. al. used for the first time VEPs for biometric identification [32]. They aimed to extract discriminant information from higher brain functions like perception and memory. Their system extracted the APS of the gamma band from 61 electrodes. An ANN was applied for classification. This achieved, on average, 90.95% accuracy on 10 subjects of Zhang's database (tables 2.4 to 2.8).

In the following years, Palaniappan's group focussed on perfecting their VEPs based design. They experimented with Principal Component Analysis (PCA), both as a preprocessing step for noise reduction (PCA-NR) [33] and as a feature reduction technique [36]. They tested different band widths such as that of the late gamma band (between 30 and 50 Hz [38]), and numerous filters [35]. They also tried to increase the system's performance normalizing the features [34] and applying different classification techniques. Two of their works offer the results of an interesting set of experiments combining and directly comparing several of these methods [37, 39].

In a bid to reduce the volume of data processed by the classifier, in 2007, Palaniappan's group applied the Multiple Signal Classification (MUSIC) algorithm to select features from the PSD coefficients of the dominant frequencies between 25 and 56 Hz [86]. An ANN was used for classification. This approach achieved 97.61% accuracy for 102 subjects of Zhang's database. However, this pool of users included both healthy and alcoholic subjects (see section 2.3.2). Hence, the previous result, and for that matter the results of all experiments with this characteristic, should be carefully interpreted (see section 2.5 for a discussion).

In 2006, landmarks of VEPs, APS and AR coefficients were directly compared by A.J. Power et. al. [87]. A LDC and a 13-subjects database were used. Results showed that relative amplitudes between two channels of P100 and N75 performed around 15 percentage points above N135[6]. When fused, P100 and N75 between Oz

[6]The terms P100, N75 and N135 refer to peaks on an ERP, so that P100 is a positive peak

Table 2.4: Systems used on Zhang's database. Along with those presented on table 2.5. Refer to table 2.2 for details of the columns' nomenclature.

Sys.	**Freq. [Hz]**	**Feat.**	**Cls.**	**Ref.**
Zha01	[32-48]	APS	FA-ANN	[32]
Zha02			RB-ENN	[86]
Zha03			Manh. kNN	
Zha04	[30-50] Butterworth	APS	SFA-ANN	[37, 39]
Zha05			Eucl. kNN	[39]
Zha06			Manh. kNN	[37, 39]
Zha07			LDC	[39]
Zha08		Norm.APS	BP-ANN	[34, 36, 40]
Zha09			SFA-ANN	[39]
Zha10			RB-ENN	[86]
Zha11			Eucl. kNN	[39]
Zha12			Manh. kNN	[39, 86]
Zha13			LDC	[39]
Zha14		PCA-NR; APS	SFA-ANN	[33, 39]
Zha15			Eucl.kNN	[39]
Zha16			Manh. kNN	
Zha17			LDC	
Zha18		PCA-NR; Norm.APS	BP-ANN	[38]
Zha19			SFA-ANN	[37–39]
Zha20			Eucl. kNN	[39]
Zha21			Manh. kNN	[37, 39]
Zha22			LDC	[39]

and O2 obtained a success rate of 82%. The APS ratio of the beta band reached 59.6% accuracy, outperforming that of the alpha band on more than 40 percentage points. The 4-order AR coefficients produced a 63.5% success rate. In all cases, the best results were obtained with occipital channels.

A year later, G.K. Singhal and P. RamKumar proposed a system based on

100 msecs post-stimulus and N75 and N135 are negative peaks 75 and 135 msecs post-stimulus respectively.

Table 2.5: Systems used on Zhang's database. Along with those presented on table 2.4. Refer to table 2.2 for details of the columns' nomenclature.

Sys.	**Freq. [Hz]**	**Feat.**	**Cls.**	**Ref.**
Zha23	[30-50] Elliptic	Norm.APS	BP-ENN	[35]
Zha24			LDC	[41]
Zha25		Norm. APS; 13 ch. by GA		
Zha26		Norm. APS; 23 ch. by GA		
Zha27		Norm. APS; 40 ch. by GA		
Zha28	[20-50] Elliptic	AvgMnt; Norm. APS; 1 ch. by DBI	RB-ENN	[40]
Zha29		AvgMnt; Norm.APS; 35 ch. by DBI		
Zha30		AvgMnt; Norm. APS; 50 ch. by DBI		
Zha31	[30-70] Elliptic	Norm. and whitened APS; PCA		[36]
Zha32	[26-56] SD-FIR	AvgMnt; APS of dominant frequency by MUSIC		[86]
Zha33			Manh. kNN	
Zha34	[30-50]	AR model of order 14 + Peak value of PSD; LDC	Eucl. kNN	[42]
Zha35	Full range	AR model of order 4		[43]
Zha36			lSVM	
Zha37		From channles C3, Cz, C4, P3, Pz, P4, O1, O2: MFCC + spectral features + APS + pitch + zero crossing rate + probability of voicing + jitter and shimmer and their statistics; Correlation-based floating forward feature selection		[44]

Table 2.6: Results with Zhang's database. Success rate of systems in tables 2.4 and 2.5 evaluated on classification with 20 subjects. For systems with configurable parameters accuracy ranges are given. Refer to table 2.3 for details of the columns' nomenclature.

Sys.	**Tr. % / CV**	**Succ.**	**Min-Max Succ.**
Using 20 subjects of Zhang's database			
Zha08	50%	99.06%	98.98-99.15
Zha14		94.18%	93.50-97.75
Zha37*	66.67%	92.80%	-
Zha04	Leave-One-Out	71.14%	70.50-73.00
Zha05		66.12%	63.13-67.63
Zha06		70.88%	67.38-72.38
Zha07		84.00%	-
Zha09		66.26%	65.38-69.88
Zha11		62.48%	59.88-63.63
Zha12		65.64%	63.25-67.00
Zha13		85.75%	-
Zha14		91.93%	91.63-93.25
Zha15		91.54%	90.75-92.25
Zha16		94.18%	93.75-95.25
Zha17		84.25%	-
Zha19		92.84%	92.25-95.25
Zha20		89.32%	87.88-90.38
Zha21		92.04%	89.50-93.13
Zha22		96.50%	-
Zha34		**100%**	-

Zhan37*: In this case, even though only 20 subjects of the Zhang's database were used, the pool contained 10 healthy subjects and 10 alcoholics.

Table 2.7: Results with Zhang's database. Success rate of systems in tables 2.4 and 2.5 evaluated on classification with 40 subjects. For systems with configurable parameters, accuracy ranges are given. Refer to table 2.3 for details of the columns' nomenclature.

Sys.	Tr. samples	Succ.	Min-Max Succ.
Using 40 subjects from Zhang's database			
Zha08	50%	95.69%	95.37-96.13
Zha18		95.40%	89.83-97.33
Zha19		82.44%	81.54-85.59
Zha23		95.42%	95.00-96.63
Zha31		**99.08%**	98.75-99.62
Zha08	Leave-One-Out	93.08%	91.38-94.00
Zha28		13.63%	-
Zha29		98.29%	98.06-98.56
Zha30		**99.00%**	-
Zha24	50% for GA; 50% for Leave-One-Out	72.25%	55.00-.90.00
Zha25		44.65%	35.00-65.00
Zha26		73.95%	53.00-86.00
Zha27		**82.00%**	75.00-95.00

VEPs measured by a single occipital electrode (Oz) [88]. The algorithm was again based on VEP's landmarks (phases, amplitudes and latencies) and relied on a similarity measure with a 2 dimensional Gaussian kernel classifier. Their method achieved 78% classification accuracy on a 10-subjects database.

In 2009, K. Das et. al. studied the spatio-temporal pattern responsible for encoding subject discriminant data [89]. They computed what they called the "*fisherbrains*": coefficients of the Linear Discriminant Analysis (LDA) (fig. 2.6). In the temporal domain, they concluded that the most informative signal arises between 120 and 200 msecs after the stimulus is presented, while pre-stimulus signal led to poor results (chance level). Presumably, these 120 msecs time-scales are required in order to process the visual information. In terms of space, occipital regions proved again to be more critical for subject identification. When a Linear SVM (lSVM) was used for classification, the system achieved performances

Table 2.8: Results with Zhang's database. Success rate of systems in tables 2.4 and 2.5 evaluated on classification with 10, 102, 120 and 122 subjects. For systems with configurable parameters accuracy ranges are given. Refer to table 2.3 for details of the columns' nomenclature.

Sys.	Tr. samples	Succ.	Min-Max Succ.
Using 10 subjects from Zhang's database			
Zha01	50%	90.95%	88.00-95.00
Using 102 subjects from Zhang's database			
Zha02	10 k-folds	96.43%	95.09-96.64
Zha03		90.70%	89.75-92.87
Zha10		96.01%	94.63-96.58
Zha12		89.85%	88.48-91.94
Zha32		**97.61%**	96.77-98.12
Zha33		95.85%	95.00-96.13
Using 120 subjects from Zhang's database			
Zha35	4 k-folds	95.02%	93.24-96.34
Zha36		98.86%	-
Using 122 subjects from Zhang's database			
Zha37	66.67%	61.70%	-

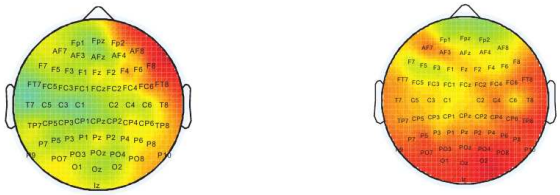


Figure 2.6: Fisherbrains. K. Das' fisherbrains from two subjects from a 170 msecs post-stimulus VEP [89].

between 90% and 95% accuracy on a 20-subjects database.

A. Ferreira et. al. proposed in 2010 a system based solely on an SVM [90, 91]. This was the first time that an SVM with a non-lineal kernel, in particular an Radial Basis Function (RBF) kernel, was utilised for EEG based identification. They achieved an 84.33% success rate on a 13-subjects database. Although worse than those of the state of the art, this result was remarkably high considering that

the system relied on raw EEG voltages.

Finally, in 2012, P. Nguyen et. al. borrowed techniques from the voice-processing field [44]. In particular, they built the following feature vector:Mel-Frequency Cepstral Coefficients (MFCC), spectral features, energy and pitch measurements, zero crossing rate, probability of voicing, jitter, shimmer, and their associated statistics. A correlation-based feature selection, with sequential floating forward selection, was then applied to reduce the dimensionality of the data. An lSVM was applied during classification. The model achieved 92.80% and 61.70% classification accuracies on 20 and 122 subjects of Zhang's database respectively (in both cases, healthy and alcoholic users where included).

2.4.3 Multiple-tasks

In a bid to explore the possibilities of new recording tasks, at the end of 2005 R. Palaniappan published the first study to compare the performance of different mental tasks [48]. As a starting point, he used AR coefficients and an LDC, a configuration extensively used in the field. Four subjects of Keirn's database were used for testing. The lowest accuracy; 92.45%, was obtained by tasks 1 and 4, while the best result; 97.40%, was obtained by task 2. Recalling from section 2.3.3, tasks 1 and 4 correspond to relaxed and letter composition tasks respectively, while task 2 corresponds to solving complex mathematical problems.

Palaniappan also experimented with the fusion of features from two tasks. This resulted in an increase in performance. The best results were obtained fusing tasks 2 and 4, with 99.05% of success, and tasks 2 and 4 with 98.95%. These combinations included both the best and one of the worst performing tasks when they were evaluated individually.

Palaniappan's team produced three more papers using 5 subjects of Keirn's database, reasserting that the fusion of different tasks increases performance [49–51]. However, they obtained contradictory results regarding the differences between paradigms. For example, some experiments reported tasks 2 and 4 being

Table 2.9: Systems used on Keirn's database. All these systems share a common high-pass pre-filtering stage at 0.2 Hz using an elliptic filter. In addition, the EEG signal separated into 20 segments of 0.5 seconds windows. Refer to table 2.2 for details of the columns' nomenclature.

Sys.	**Feat.**	**Cls.**	**Ref.**
Kei1	6 order AR	LDC	[48]
Kei2	Kei1's features + APS and IH-APS of [8-13]Hz, [14-20]Hz and [21-50]Hz bands; PCA	LDC	[49]
Kei3	Kei2's features + inter-hemispheric channel linear complexity + non-linear complexity	2 stage Manh. distance OCC	[51]
Kei4	Kei3's features; PCA		

the best and worst respectively, reversing the conclusions of Palaniappan's first work. When groups of 3 were evaluated, tasks 1, 2 and 4 appeared again in the best combinations (tables 2.9, 2.10 and 2.11).

In 2007, Marcel and Millan published the first work that actually studied the effects of time over an EEG based identification system [92]. They did so using a data set including the imagined tasks of moving the left hand, moving the right hand and producing words. Their system was based on Normalized (Norm.) PSD features and a Gaussian Mixture Model (GMM) classifier. Results showed a drop in performance when training and testing sets got apart in time. They also noted that this drop can be circumvented by using a multi-day-session training approach. They obtained a best Equal Error Rate (EER) of 6.60% with the left hand movement task.

At the end of that year, S. Sun proposed a Multi Task Learning (MTL) approach where an ANN was trained with features extracted from multiple tasks [93]. The 9-subjects database used contained only two tasks: imagine left and right hand movements. An increase of around 4 percentage points was observed when the MTL approach was used, obtaining 95.60% and 94.81% classification on each task respectively.

Table 2.10: Results with Keirn's database. Success rate and EER of systems in table 2.9 evaluated on classification and verification experiments with a maximum of 5 subjects. In both cases, a CV method using 50% of the database for training and 50% for testing was applied. Results of individual and combined tasks are presented. Refer to table 2.3 for details of the columns' nomenclature.

	Classification		**Verification**	
Sys.:	Kei1*	Kei2	Kei3	Kei4
Task/s	**Succ.**		**EER**	
Task 1	92.45%	97.77%	0.15%	0.20%
Task 2	97.40%	96.88%	0.07%	0.07%
Task 3	94.30%	97.84%	0.00%	0.00%
Task 4	92.45%	98.64%	0.10%	0.12%
Task 5	95.30%	98.40%	0.00%	0.02%
Tasks 1 & 2	98.98%	99.60%		
Tasks 1 & 3	98.15%	99.52%		
Tasks 1 & 4	98.50%	99.60%		
Tasks 1 & 5	98.50%	99.60%		
Tasks 2 & 3	98.85%	99.24%		
Tasks 2 & 4	99.05%	99.76%		
Tasks 2 & 5	98.90%	99.32%		
Tasks 3 & 4	98.60%	99.56%		
Tasks 3 & 5	98.35%	99.40%		
Tasks 4 & 5	98.75%	99.60%		

Kei1*: Only 4 of the 5 subjects were used.

The previous scenario of imagined movements of the left hand outperforming the right was also studied by other researchers. For example, in 2009, X. Bao's et. al. found no significant differences between these tasks on their experiments on IIIaBCI03 database [53, 55]. Their results also showed that imagining tongue movement produced better results compared to hand tasks; around 10 percentage points better, and that frequencies between 20 and 30 Hz provided more discriminant information than lower ones. Nevertheless, these results, and any result

Table 2.11: Results with Keirn's database. Success rate of systems in table 2.9 evaluated on classification with a maximum of 5 subjects. A CV method using 50% of the database for training and 50% for testing was applied. Refer to table 2.3 for details of the columns' nomenclature.

Sys. Kei2	
Task/s	**Succ.**
Tasks 1, 2 & 3	99.84%
Tasks 1, 2 & 4	99.88%
Tasks 1, 2 & 5	99.68%
Tasks 1, 3 & 4	99.84%
Tasks 1, 3 & 5	99.60%
Tasks 1, 4 & 5	99.82%
Tasks 2, 3 & 4	99.98%
Tasks 2, 3 & 5	99.74%
Tasks 2, 4 & 5	99.78%
Tasks 3, 4 & 5	99.60%
Tasks 1, 2, 3 & 4	100%
Tasks 1, 2, 3 & 5	99.80%
Tasks 1, 2, 4 & 5	99.80%
Tasks 1, 3, 4 & 5	99.98%
Tasks 2, 3, 4 & 5	99.78%
Tasks 1, 2, 3, 4 & 5	100%

obtained with IIIaBCI03 database (tables 2.12 to 2.14), must be approached with caution, as the database only contains 3 subjects. In fact, in a later work from the same group, their results did show a superiority of an imagined left hand movement task over the right hand equivalent [57].

In 2012, S. Yang and F. Deravi compared motor imagery tasks and actual movements [62] using the *BCI2000* database. However, they obtained inconclusive results. One task gave better results when the actual movement was carried out while the other performed better under the imagery condition. When the authors studied the spatial distribution, they concluded that parietal and occipital channels

Table 2.12: Systems used on IIIaBCI03 database. Systems 1 to 6 used only channels C3, C4, P3, P4, O1 and O2. Refer to table 2.2 for details of the columns' nomenclature.

Sys.	**Freq. [Hz]**	**Feat.**	**Cls.**	**Ref.**
IIIa1	[0.5-100]	AR + Linear Complexity + PSD + Phase Locking Value	BP-ANN	[53]
IIIa2	[8-13]			
IIIa3	[14-20]			
IIIa4	[21-30]			
IIIa5	[2-40]	IIIa1's features + Energy Entropy + Mutual Information + Cross-correlation		[54]
IIIa6	[2-40]	ARMA model		[55]
IIIa7	[8-30]	DWT		[56]
IIIa8	[8-13]			
IIIa9	[14-20]			
IIIa10	[21-30]			
IIIa11	Full	SCBI-NR; 10 highest coefficients of LDC		[57]
IIIa12	Same as Zha37 on table 2.5			[44]

Table 2.13: Results with IIIaBCI03 database. Success rate and EER of systems in table 2.12 evaluated on classification and verification experiments. A CV method using 50% of the database for training and 50% for testing was applied.

Sys.	**Task1**	**Task2**	**Task3**	**Task4**
Classification				
IIIa1	81.20%	82.10%	82.80%	90.60%
IIIa2	∼61%	∼62%	∼68%	∼68%
IIIa3	∼56%	∼57%	∼54%	∼55%
IIIa4	∼81%	∼72%	∼76%	∼88%
IIIa6	76.70%	77.90%	80.90%	92.20%
IIIa11	82.40%	79.10%	81.70%	88.10%
Verification				
IIIa6	80.80%	80.50%	81.30%	92.80%

Table 2.14: Results with IIIaBCI03 database. Success rate (Succ.) rate and EER of systems in table 2.12 evaluated on classification and verification experiments. A CV method using 50% of the database for training and 50% for testing was applied. These were MTL experiments; i.e. systems used EEG from all the recorded paradigms to determine the user's identity.

Sys.	**Succ.**
Classification Experiment	
IIIa7	84.07%
IIIa8	∼74%
IIIa9	∼65%
IIIa10	∼85%
IIIa12	99.00%
Verification Experiment	
Sys.	**EER**
IIIa5	∼13%

were the most discriminant.

Finally, some works that used multiple-tasks databases focussed merely on achieving high performances and did not report individual task results. For example, C. He et. al. applied Multivariate AR (mAR) coefficients and Naive Bayes classification [94, 95]. When tested on verification mode using a 4-subjects data set with 5 motion related tasks and 16 electrodes distributed around the scalp, the system obtained a Half Total Error Rate (HTER) of 6.70%. Soon after, He's group improved the design applying ICA as a spatial filtering tool and using a 7 order AR model [96]. Such architecture obtained a HTER of 2.20% on an expanded version of the database which contained 7 subjects.

Nguyen's et. al. study, introduced in section 2.4.2, also tested several multiple-tasks databases [44]. Their system obtained a 99.00% classification on the *IIIaBCI03* data set, a 46.24% on a 9-subjects database similar to the previous and an 80.80% on another 9 subjects database with left and right hand imagery movement tasks.

2.4.4 Consumer EEG devices

Once the viability of identifying subjects using their EEG was established, researchers started to realize its possibilities as a high level security system [97]. However, for this to become a reality, they first needed to overcome the hardware issues. Medical and research EEG apparatuses are expensive and need qualified staff, conductive gel and at least 15 minutes of preparation. Some authors already tried to simplify this process by using less electrodes[7]. However, thanks to the advances made in the field of sensors, easy to use electrodes and commercial EEG caps started to be used for subject identification in 2008.

A. Riera et. al. presented the first true viability test of a real EEG verification system [98]. In doing so, they used a database composed of 51 registered users and 36 intruders, with REC sessions recorded in a time frame of 34 ± 74 days for each subject, thus including the effect of time on their experiments. The system used only two dry electrodes: FP1 and FP2. They applied a total of 28 combinations of features and classifiers. These included AR and PSD coefficients, plus 3 statistical measurements, and 4 variants of the LDC. For each registered user, the 5 best combinations were selected. Using partially overlapping training and test sets, authors reported a classification accuracy of 98.10% on the aforementioned database. Following this, the best 15 classifiers were fused into a single verification system, achieving an EER of 3.4%.

Around the same time, C. Miyamoto et. al. published two papers using Tottori database [65,66] (table 2.15). In a bid to build a practical system, the authors used simple designs based on spectral features and similarity measures. These models achieved errors of around 20% in verification. A year later, they managed to lower this to 11.00% using the concavity and convexity of the spectral distribution from the alpha band [67].

F. Kennet provided a statistical study of multiple features [99]. His 12-subjects

[7]The first work with REC, while not approaching the research from this perspective, used a single occipital channel

Table 2.15: Systems and results with Tottori database. A CV method using 50% of the database for training and 50% for testing was applied. Refer to table 2.2 for details of the columns' nomenclature.

Sys.	**Freq. [Hz]**	**Feat.**	**Cls.**	**EER**	**Ref.**
Tot1	[7.8-13.3]	Non-dominant region of the power spectrum	Similarity measures	30%	[65]
Tot2		Variance of spectral power		30%	
Tot3		Fusion of Tot1 and Tot2		21%	
Tot4	Full range	N maximum spectral values and their frequencies		30%	[66]
Tot5		Sum of frequency values over a threshold		31%	
Tot6	[8-13]	Concavity of spectral distribution, variance, convexity of spectral distribution		11%	[67]
Tot7	Score fusion of Tot4 and Tot5			28%	[66]

database, recorded with a ThinkGear device, contained multiple tasks: relaxation; counting; reading a text; thinking on a colour; thinking on rotating a mouse; thinking on a password; thinking on music; and thinking on words starting with the letter "M". Several features were computed. Results showed that the power of high frequencies, between 20 to 50 Hz, were the most promising based on their $std/mean$ rate, while the mean of the phase angle produced the worst score. In addition, features from the mid-range frequencies were the most normal, i.e. the more stable. No relation between features and tasks was found on a Manhattan metric (Manh.) based analysis.

In 2010, F. Su et. al. presented a work on the effects that diet and/or circadian rhythms have on these systems [100]. They recorded a 40-users REC database with the HXD-I EEG consumer device. The database included sessions captured before and after the intake of coffee and at different times of the day. Using AR coefficients

and PSD from 5 to 32 Hz as features, the system achieved 97.5% of accuracy when classifying with a combination of LDC and k-Nearest Neighbour (kNN), 81.90% when classifying with an ANN, and 79.60% when classifying with an SVM. The results also showed a decrease in accuracy of around 10 percentage points when coffee was drunk and a rise when the length of the EEG signal increased; peaking at around 3 minutes long. A later study demonstrated that these differences on performance were in fact statistically significant, and that diet and coffee affected the system reliability [101].

Finally, Q. Zhao et. al. published a number of studies on Lanzhou database [69–72] (tables 2.16 and 2.17). The latest of these works was the first to study the influence of time over long periods (around 6 months) when using a commercial EEG device [72]. The results reiterated the conclusions previously found by Marcel and Millan [92]. Moreover, they implemented a fully functional system which could record EEG from a mobile device and send it to a server for processing. Other fully implemented systems have also been published [102, 103].

2.4.5 Other approaches

All the works introduced so far aimed to identify or verify the users' identity directly from their EEG activity. On the other hand, some works applied an indirect approach based on information other than the subject's unique EEG features that would ultimately help to identify them.

Palaniappan's group have been especially interested in such a strategy. After studying the P300 wave under target and non-target visual stimulus [104, 105], they proposed a system that allows to introduce a password via automatic stimulus response. Such a system flashed letters [106] or numbers [107] to users on a screen. When the target (the character that matched the subsequent password digit) was flashed, the subjects' brain produced a specific P300 response. This was detected by the system, which then assembled the password. They managed to achieve a perfect accuracy when 3 users performed 5 attempts of login.

Table 2.16: Systems used on Lanzhou database. Refer to table 2.2 for details of the columns' nomenclature.

Sys.	Freq. [Hz]	Feat.	Cls.	Ref.
Lan1	Full range	6 order AR coefficients	kNN	[70]
Lan2		ApEn + C_0-complexity + correlation dimension D_2 + largest Laypunov exponent		
Lan3	[0-40]	Fast-ICA; central freq. + max. power + average peak-to-peak value + power ratio		[69]
Lan4	[4-7]			[70]
Lan5	[8-13]			
Lan6	[12-15]			
Lan7	DWT-NR	From an adaptive AR model: Lan7's features + activity, mobility and complexity Hjorth parameters		[72]
Lan8	DWT-NR; [8-13], [14-30] and [4-7]Hz	from AR model: maximum PSD and its frequency + APS of each rhythm	Naive Bayes Classifier	[71]

S.K. Yeom et. al. [108, 109] used the differences between VEPs elicited by self and non-self images to verify users (fig. 2.7). When tested on a verification experiment, the model achieved an 86.10% success rate on a 10-subjects database (Yeom's database will be described in section 3.1.5). In addition, the authors noted that frontal, central and parietal regions provided the most discriminant data.

2.4.6 Advances made in parallel to the present work

So far, we have analysed studies published before 2014. Between 2014 and April 2015, and during the execution of the remaining of the present work, a total of 22 extra publications have been identified, including conference and journal articles, technical reports and a patent [110]. From these, the following are of special

Table 2.17: Results with Lanzhou's database. Success rate of systems in table 2.16 evaluated on classification and verification experiments. Refer to table 2.3 for details of the columns' nomenclature.

Sys.	Notes	Succ.
Classification Experiments		
Lan3	3-kfolds over 3 registered users	99.23%
	3-kfolds over 4 registered users	∼98.50%
	3-kfolds over 5 registered users	∼97.75%
	3-kfolds over 6 registered users	∼97.60%
	3-kfolds over 7 registered users	∼97.25%
	3-kfolds over 8 registered users	∼97.20%
	3-kfolds over 9 registered users	∼96.90%
	3-kfolds over 10 registered users	96.77%
Lan1		83.78%
Lan2		<25%
Lan4		97.29%
Lan5		94.59%
Lan6		94.59%
Lan8	Using 4 secs of EEG	∼65%
	Using 60 secs of EEG	∼100%
Verification Experiments		
Lan7	11 registered users + 11 intruders	∼66%
	Same day test	94.60% TAR
	1 week latter test	83.64% TAR
	6 months latter test	78.20% TAR

interest.

Two published papers executed their experiments on the 108-subjects BCI2000 data set. M. Fraschini et. al. proposed a system based on the eigenanalysis of a functional brain map of the Phase Lock Index measured on 6 bands between 0.5 and 50 Hz [63]. When evaluated with an Euclidean metric (Eucl.) distance classifier on a verification paradigm, their architecture achieved an EER of 4.4% and 6.5% on REO and REC respectively. Unfortunately, the authors gave no

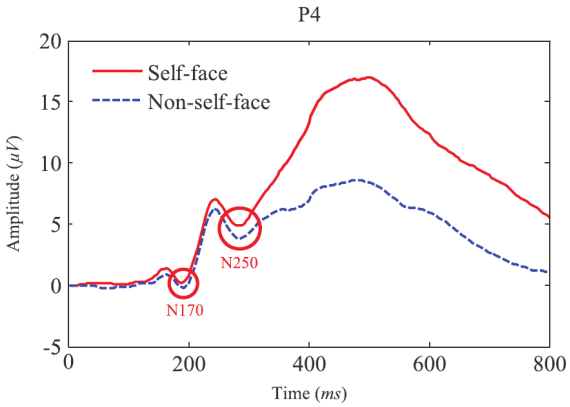


Figure 2.7: VEPs elicited by self and non-self face images [109].

details of the experimentation methodology.

Campisi's team ran a classification test on the BCI2000 database [64]. They reported a perfect classification accuracy under leave-one-out CV, on both REC and REO. Their system characterized the EEG signals by the Spectral Coherence (COH) of sensor pairs. Notably, to obtain the signal's spectral representation, they applied the Welch's periodogram method with a 50%-overlapped 1-second Hanning window. Finally, they applied a Manh. distance based classifier. Their results showed that PSD coefficients of single channels performed, on average, 15 percentage points better than COH of single pairs. However, when the authors fused (at score level) the responses of multiple channels/pairs, COH was the only one to reach perfect classification. This was achieved for both REC and REO when fusing pairs of frontal channels. Finally, they also reported that the PSD of posterior sensors outperformed central and frontal regions under REC, but not under REO .

In a later work, Campisi's team used a 9-subjects database with two REC and REO sessions recorded 1 or 2 weeks apart [111]. In this case, they used AR coefficients extracted individually from delta, theta, alpha, beta and gamma rhythms. LDC was applied for classification. After system optimization, they decided to use a 10-order AR model on 1 second segments and to reject the gamma band ([30, 40] Hz). When applying a CV protocol that trained and tested the

system with different sessions, they achieved perfect classification using parietal trios on REC.

Other proposals included the use of muscle and blink artefacts recorded by EEG sensors for the discrimination of subjects [112–114], practical implementations of EEG biometric identification systems [115,116], and two multimodal systems: one fusing EEG and iris biometries [117] and another one exploiting the gender, age and task information contained in the EEG signal to boost the overall system's performance [118].

2.5 Discussion

At this point, it is clear that the problem of EEG subject identification is sensitive to several factors, such as: the spatial location of electrodes, frequencies, time spans, recording tasks, subject conditions and, of course, the system's architecture. In many cases, this was accentuated by the fact that systems were highly tuned for specific information. An example of this can be seen in the results obtained by K. Das et. al. and S.K. Yeom et. al. using VEPs. While the former reported that the most discriminant information was localized on the occipital region between 120 and 200 msecs after stimulus onset [89], the latter found the most discriminant information on anterior brain regions and from 250 ms post-stimulus onset onwards [108,109]. These differences in time and location are probably due to the important detail that, even though both studies used a VEP paradigm, each was based on different information within the ERPs.

Before assessing each of these factors, let us discuss the usage of EEG activity recorded from unhealthy subjects for the evaluation of EEG based identification systems. Some of the studies found in the literature used data from healthy and alcoholic subjects [43, 44, 46, 86, 119, 120]. The problem with this practice is that alcohol has been shown to affect various aspects of EEG activity [5, 6], including the alpha rhythm [13]. Therefore, it is likely that such data sets were biased,

overrating the systems' performances. To clarify this, we encourage future works including unhealthy individuals to provide a statistical study probing that the computed parameters are not affected by the relevant disorder, i.e. assuring these are uncorrelated.

2.5.1 Which are the subject specific traits of the EEG?

There is enough evidence to confirm the existence of subject identity within EEG. Which these traits are is, at the moment, less clear. Genetic and neurophysiological studies affirm that the power and peak frequency of the alpha band present the strongest heritability relationship, followed by the beta band [18,21]. Nevertheless, these bands by themselves proved to be insufficient to obtain high accuracy rates on biometric systems, forcing researchers to explore the usage of different features. In general, results suggest that it is the overall shape of the spectrum that codes the subject's identity. However, without an exhaustive study, we are prevented from making definitive assertions.

2.5.2 Where are these traits in terms of location?

Several studies reported the optimum sources of discriminant information as being occipital, temporal and parietal during REC [64, 80–82]. This should come as no surprise given that the posterior rhythm was described as highly genetically determined. The opposite scenario was observed under REO, where anterior brain areas provided optimum discriminant information [64, 83–85, 121].

The balance seems to tilt back to occipital regions on VEP experiments [89]. This was attributed to the increase in activation of these areas during VEPs. However, Yeom's et. al. showed that when the experimentation procedure implies personality and self-representation thinking, as in VEPs elicited by self images, frontal brain areas are again the most discriminant [108, 109]. This could still be associated with the relative activation of brain regions, given that the visual

pathway describes a flow of information/activation from posterior to anterior areas as the brain interprets the scene [122].

From the multiple-tasks studies, only Yang's et. al. considered the spatial factor [62]. However, they did so with a small database (only 3 subjects). Their results show parietal areas near the temporal lobes carrying the most discriminant information.

Differences between hemispheres are even less clear. Several studies, including some neurophysiological, concluded that there is no significant hemisphere effects [83, 84, 121, 123]. However, others have reported that such difference do exist [40, 57, 85]. In the latter case, they all described a scenario where the right hemisphere outperformed the left.

The lack of homogeneity in this regard hinders the extraction of any strong conclusion. Some of these variations can be explained by properties of the functioning of the brain and system design. For example, systems analysing ERPs perform better on occipital regions probably due to ERPs being more easily detected in those areas. However, it could also be the case that the overall differences in performance across locations are due to idiosyncrasies of the data, such as set-up artefacts.

2.5.3 Where are these traits in terms of frequency?

The frequency factor has received much less attention. This is largely due to authors focusing on localized information on a specific part of the spectrum, based on previous studies or on their own experience.

On the REC paradigm, neurophysiological as well as biometric studies concluded that alpha and delta rhythms are the most discriminant, followed by the theta and the beta rhythms [18, 21, 82, 123]. Within the alpha rhythm, the middle section has outperformed the first and last portions in some experiments [74, 77]. However, this is not always the case, and the theta rhythm has also been reported to be the most discriminant one [70].

On the remaining tasks, it has usually been the case that higher frequencies produce better classification rates [53, 56, 63, 68, 99].

2.5.4 Are they constant across time?

Genetic and neurophysiological studies have described changes in human EEG across maturation, some of which are related to heritable traits [13]. Even though these must be considered in genetic studies, one can decide not to contemplate them in an adult security system, as they are slow changing effects past developmental years[8].

Genetic and neurophysiological works have observed high stability of EEG PSD between sessions recorded even one year apart [27]. However, apart from I. Nakanishi et. al. visual inspection [67], all biometric studies have contradicted this. S. Marcel and J.D.R. Milla were the first to observe a drop in performance when the training and test sessions grow further apart in time [92]. In addition, studies on diet and circadian effects revealed how sensitive a system could be to changes of the EEG activity elicited by common events, such as the intake of coffee or daily physiological changes [100, 101, 124]. Having said that, authors also noted that this can be circumvented by a multi-day-session training approach [43, 72, 125].

2.5.5 Are they constant across cognitive tasks?

Performance differences on multiple-tasks databases have been reported since the very beginning of this research field, except for Kennet's statistical work [99]. The first Palaniappan's results suggested that the most cognitively demanding tasks provide the best discriminant information [48]. However, this was soon undermined by follow-up work of his, where the findings were inverted [49].

Such a high variability, even between studies performed by the same research

[8]EEG activity has been identified to change dramatically during developmental years until approximately 19-20 years old [13].

group, hinders the extraction of a consistent narrative. Even in a comparison between left and right hand movement imagery tasks, there are several works reporting that the left outperforms the right [57, 92, 93, 119] and several stating something else [53, 55].

What seems to be clear is that the systems extracted some task-dependant discriminant information. This was exhibited by Palaniappan's results, where the best combination of tasks turned out to include the best and the worst performing ones individually. In addition, this hypothesis is in line with S. Sun's results on MTL [93].

2.5.6 Which is the best design approach?

Explored features can be divided in two large groups: "broad" and "specific". The majority of the publications fall in the former group, which refers to features that describe EEG signals extensively. Hence, the discriminant information is not directly presented, but somewhat hidden within the larger picture. The more prominent examples are coefficients of AR models and PSD. In these cases, it is the classifier or a feature selection technique which must find/access the right data.

Specific features include peak APS and its corresponding frequency and ERP characteristics. In these cases, the discriminant information is more directly exposed to the classifier, facilitating its task. These have been the *de facto* features used on consumer EEG databases, mainly due to the pursuit of low processing complexity and high speeds. Although specific features have achieved lower accuracies, they have boosted the performance of broad features when concatenated. For example, R. Palaniappan improved his system on a multiple-tasks database by including APS measurements from each sub-band to the AR coefficients [51]. Even when the AR coefficients already contained such information; in fact APS values were computed from the AR coefficients.

In addition, the use of data from multiple electrodes, rhythms and tasks in-

creases the system's performance, as well as the application of NR techniques. Tools such as LDA for feature reduction and LDC and SVM for classification seem to give better results.

Regarding the indirect methods studied by Palaniappan's work, these do not measure any EEG subject specific traits. They represent BCIs to enter passwords and, as a consequence, they lack most of the robustness of EEG biometric systems.

2.5.7 Medical versus consumer devices

Judging from the relatively few works using consumer EEG devices, they seem to provide worse results than medical equipment. This could be attributed to the quality of the recorded signal. Nevertheless, a specific experiment using the same subjects and tasks recorded with both types of equipment is necessary to draw a firm conclusion on this matter.

2.6 Conclusions

We have presented an extensive research of the state of the art on EEG biometric identification systems. We conclude that the EEG, and in particular its spectral distribution, contains subject specific traits suitable for a biometric identification system. Moreover, results suggest that this is specially true for data within the alpha rhythm.

Perhaps, one of the most striking facts of this analysis is the low level of agreement between studies. This may come as a consequence of the complexity of the heritable model underlying the EEG traits, along with the amount of degrees of freedom in the problem. On top of that, some of the considered variables seem to be dependant on one another. Hence, we conclude that the state of the art only scratches the surface of the problem, and further research is needed to draw firmer conclusions.

Having said that, we can extract from the presented results that: the combination of multiple electrodes, rhythms and tasks increases the system's performance; and a multi-day-session training approach and/or a continuous learning method can counteract the detrimental effects of time. Other features are less clear. Consumer EEG devices may give worst results than medical equipment, presumably because of the quality of the recorded signal. On the other hand, if the traits used are robust enough to noise, this may not be the case. In addition, tools such as PCA-NR, LDA for feature reduction and LDC and SVM seem to provide higher performances. Although this depends, again, on all the other variables of the problem.

Chapter 3

Materials

Almost all of the studies presented in chapter 2 focused on the analyses of isolated databases. Even when multiple data sets were used [27, 44, 126], authors tested systems with different architectures, but never analysed base properties of the EEG traits. In addition, databases were not always publicly available. These two facts hinder the interpretation and reproduction of results. In chapter 5, we will present a study of the basic properties of the EEG discriminant information, performed on 6 publicly available databases with different characteristics. Two of these databases: Keirn's and Zhang's, have been extensively used in the literature. Two more: BCI2000 and Yeom's, have also been used in the literature, but less frequently. The remaining 2: DEAP [127,128] and P. Ullsperger's [129] databases, are new within this context (EEG biometrics).

In this chapter, we will describe in detail the 6 data sets used throughout the current work. First, we will introduced the original properties of the databases and its initial preparation. Subsequently, we will detail the preprocessing applied to each of these databases in a bid to reduce noise and uninteresting differences between them; such as sampling frequencies and montages. We will close with a description of the ten final data sets extracted from the original 6 databases.

3.1 Raw databases

Data sets were initially prepared with the aim of maximizing the available data and testing conditions. However, in some cases, it was necessary to exclude data. For example, we only included subjects identified as healthy/normal to avoid any bias on the obtained results. We also rejected subjects with not enough trials per task, to ensure reasonably balanced data sets. In other cases, the original EEG segments were fragmented maximizing the inter-segment time to increase the number of samples.

3.1.1 BCI2000 database

BCI2000 database has been previously described (section 2.3.5). It is mainly composed of two resting states and 4 tasks, each divided in three conditions. Here, we considered all the possibilities within each individual task. That is: a relaxing stage, left/right for T1, left/right for T2, and fists/feet for T3 and fists/feet for T4. As a result, a total of 11 tasks (4x2 tasks + 1 common relaxing stage + 2 baselines) were differentiated.

We extracted a total of three different sets: BCI2000-Baseline, BCI2000-Tasks and BCI2000-Full. The first contains samples from the REO and REC tasks, each divided in 5 10-second segments, non-overlapped and as temporally spaced as possible. The BCI2000-Tasks set contains the 9 labels (4x2 tasks + 1 common relaxing stage) from the different tasks. Non-overlapped 4-second segments were extracted, again as temporally spaced as possible. The last data set, BCI2000-Full, contains 4-second segments of all of the 11 tasks.

As it is not specified on the documentation, it was assumed that all blocks were recorded in a single session.

3.1.2 DEAP database

The DEAP database (Dataset for Emotion Analysis using EEG, Physiological and Video Signals) was originally collected to study emotional responses [127, 128]. A single session was recorded from 32 participants (50% males, aged between 19 and 37, mean age 26.9) while they visualized 40 60-second music videos which elicited different emotions. These music videos were presented on a 17-inch screen placed at 1 meter from the participant. The resolution was set to fit approximately 2/3 of the screen in a bid to reduce aye movements. The sound was played through stereo speakers. 32 AgCl EEG electrodes and peripheral physiological signals were collected by a Biosemi ActiveTwo system at a sampling rate of 512 Hz. Recordings were made without a reference channel, as is usually the case with the Biosemi hardware.

Before each session, a 2-minute baseline was recorded with subjects relaxed. We extracted three separate sets: DEAP-Baseline, DEAP-Playback and DEAP-Full. The first is composed of 5 20-second segments per subject of the recorded baseline, non-overlapped and as temporally spaced as possible. The playback dataset contains 20-second EEG segments from each of the played videos, non-overlapped and as temporally spaced as possible. Videos were rated in terms of arousal (boring versus stimulating), valence (positive/happy versus negative/sad), like versus dislike, dominance (submissive versus dominant) and familiarity. For the playback set, 4 states corresponding to the quadrants of the arousal-score representation of emotions were identified. The number of states per subject varied due to their personal evaluation of the videos. Finally, the DEAP-Full data set contains a mix of the two previous sets.

3.1.3 Keirn's database

Keirn's database was described in section 2.3.3. To increase the number of samples, we divided the 10-second trials in two 4-second segments separated by 2 seconds.

3.1.4 P. Ullsperger's database

P. Ullsperger et. al. recorded AEPs from 5 subjects [129]. Auditory stimuli (words) were presented to the participants, who had to classify each of the stimulus as synonyms or non-synonyms of a given target. The number of trials varied across subjects. Inter-stimulus time between trait and test stimulus was set to 1 second.

The EEG was recorded from 61 electrodes referenced to PCz and CP1 with a sampling rate of 200 Hz. A notch filter between 47 and 53 Hz was applied to remove line noise. From each trial, segments were extracted from 2 seconds before to 2 seconds after presentation of the second (testing) word.

3.1.5 Yeom's database

This database was used by S.K. Yeom et. al. for an indirect EEG based subject verification system [108, 109] (section 2.4.5). It contains 11 male participants (10 users and 1 impostor); including 1 pair of monozygotic twins, with ages between 20 and 29 years old (mean 26.67). Self and non-self images were presented to the subjects in 2 different-day sessions. Each session consisted of 20,000 trials divided into 2 runs (with a short break in between), 50 blocks per run, and 20 trials per block (10 self and 10 non-self stimuli).

A Neuroscan SynAmps2 system was used to record the EEG signal from 18 electrodes (International 1020 System). Note that the available database is a post-processed version and therefore some of the parameters may differ from the original. The sampling rate was set to 300 Hz and a band-pass filtered between 0.1 and 100 Hz as well as a 50 Hz notch filter were applied. EEG was extracted from 200 ms pre- to 800 ms post-stimulus. The data set consisted of 1-second EEG segments starting at stimulus onset.

3.1.6 Zhang's database

Zhang's database was used as described in section 2.3.2.

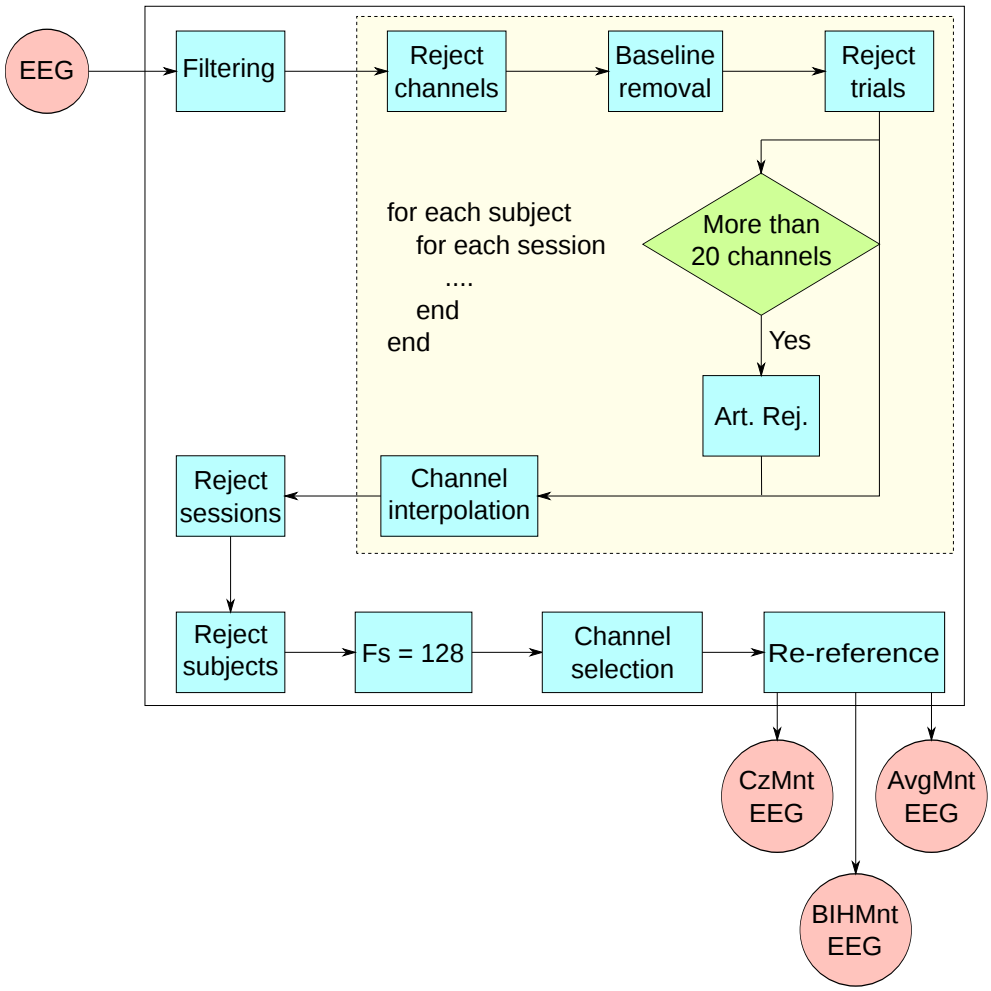


Figure 3.1: Preprocessing steps applied to all databases. Block "Art. Rej." represents an artefact rejection method (chapter 4).

3.2 Preprocessing

As a group, the databases just described comprise a wide range of configuration values. While some were of interest for the current study (tasks, signal length, number of subjects, etc.), others were not (sampling frequency, number of channels, filtered range, etc). In a bid to remove the latter and normalize the databases, we applied a common preprocessing stage (fig. 3.1).

3.2.1 Filtering

All databases were filtered using a High-Pass (HP) Finite Impulse Response (FIR) filter with a low edge frequency of 1 Hz. On data sets with 50 (DEAP) or 60 Hz (BCI2000, Keirn's database) line noise, we applied a FIR notch filter as appropriate. Additionally, we low-pass filtered the EEG activity from Zhang's Database (DDBB) at 40 Hz to homogenize channels, as some of them came filtered at around that frequency.

FIR filters were selected for their linear phase response, as we aimed to distort the signal of interest as little as possible. Relatively high filter orders (table 3.1) had no negative connotations, as we applied the preprocessing offline for computational optimization reasons.

Table 3.1: Filtering details for each database. The table shows: the sampling frequency (Fs) at which each database was processed, the high-pass filter order (HP order), and the frequency at which the notch filter was applied (Notch fcut) and its order (Notch order).

DDBB	**Fs [Hz]**	**HP order**	**Notch fcut [Hz]**	**Notch order**
BCI2000	128	424	60	424
DEAP	256	846	50	424
Keirn's	250	826	60	826
P.Ullsperger's	200	660	50*	-
Yeom's	300	990	60*	-
Zhang's	256	8448	-	-

*: Originally filtered.

3.2.2 Channel rejection

Next, extremely noisy EEG channels were automatically detected and rejected to avoid any interference in subsequent computations. Specifically, the time kurtosis

$$K_c = rn(|kurt(X_c) - 3|) \text{ for } c = 1, 2, ..., C, \tag{3.1}$$

with X_c the EEG of the c-th channel, was used as a noise indicator feature, together with the weighted correlation between channels

$$R_c = rn(median(\Phi_c,\ \forall j \neq c)) \text{ for } i = 1, 2, ..., C \tag{3.2}$$

$$\text{and } \Phi_c = \{1 - |corr(X_{ci}, X_{cj})|e^{-d_{ci,cj}}, \forall j \neq c\}, \tag{3.3}$$

where $d_{ci,cj}$ is the arc distance between channels ci and cj in a sphere with radius one. In both cases, rn is the robust normalization function

$$rn(v) = \frac{20(v - median(v))}{27 \cdot iqr(v)}, \tag{3.4}$$

where v is a feature vector and $median$ and iqr are the median and interquartile functions respectively. Note that $(27/20) \cdot iqr(v_n) \simeq \theta_n$ for a normal distribution v_n with standard deviation (std) θ_n.

Channels with any of these features exceeding a value of 3 were rejected. If the proportion of such channels fell below 10%, they were excluded, otherwise we rejected the 10% of the channels with the largest scores $\sqrt{K_c^2 + R_c^2}$.

3.2.3 Voltage normalization

Apart from neural activity, EEG voltage ranges depend on scull thickness and exogenous factors such as the hardware used and the cap preparation. In a bid to isolate the brain signals, we normalized EEG voltage ranges by equation 3.4. This was preferred over z-score normalization due to the level of artefacts in some of the databases.

3.2.4 Events rejection

When there were a sufficient number of available trials, the noisiest were discarded. Again, the criterion applied to determine the number of retained trials relied on the desire to have a balanced database across subjects, tasks and sessions.

Hence, noise in events was characterized by the voltage range

$$A_e = |rn(\langle p^{(95)}_{(T)}(X_e) - p^{(5)}_{(T)}(X_e)\rangle_{(C)})| \text{ for } e = 1, 2, ..., E \tag{3.5}$$

and the voltage variance

$$V_e = |rn(\langle var_{(T)}(X_e)\rangle_{(C)})| \text{ for } e = 1, 2, ..., E, \tag{3.6}$$

where X_e is the EEG of the e-th event, $p^{(n)}_{(T)}$ and $var_{(T)}(\cdot)$ are the $n\%$ percentile and variance across time (T) and $\langle\cdot\rangle_{(C)}$ represents the average across channels. Events were then sorted based on their score $\sqrt{A_e^2 + V_e^2}$ and those with the highest values were rejected until the desired amount was reached. This amount varied across databases (table 3.2).

3.2.5 Channel interpolation

With artefactual signals removed, rejected channels were interpolated back into the data set. A spherical spline interpolation method was applied for this purpose [130].

3.2.6 Session rejection

Some subjects of Keirn's data set contained 3 recorded sessions. To normalize the number of sessions to 2, the one with the worst signal quality was discarded. The rejection was based on the score

$$\frac{\langle\sqrt{A^2 + V^2}\rangle_{(E)}}{2}, \tag{3.7}$$

where A and V are the features defined in (3.5) and (3.6), and $\langle\cdot\rangle_{(E)}$ is the average along the events dimension.

3.2.7 Subject rejection

When the number of subjects available allowed it, those with the worst signal quality were discarded. The rejection was based on the score

$$\frac{\langle\sqrt{A^2+V^2}\rangle_{(E)}}{2}+30\frac{\hat{C}+1}{C}, \tag{3.8}$$

where A and V are the features defined in (3.5) and (3.6), $\langle\cdot\rangle_{(E)}$ is the average along the events dimension, C is the number of channels and $\hat{C}$ is the number of rejected channels. The number of rejected subjects varied across databases (table 3.2).

3.2.8 Channel selection and sampling frequency

Finally, in order to reduce the volume of data, only a set of 16 channels distributed around the scalp were kept (figure 3.2), except for Keirn's and Yeom's databases, from which all available channels were kept. In addition, databases were down-sampled to 128 Hz.

3.3 Final databases

After the preparation of the 6 publicly available databases described (section 3.1) and preprocessing detailed above (section 3.2), we obtained a total of 10 data sets (table 3.2). Together, they represent a rich variety of tasks (mind states, cognitive tasks and experimental conditions). In particular:

1. BCI2000-Baseline and DEAP-Baseline databases represent resting states,
2. BCI2000-Tasks includes real and imagery motor tasks (it also contains a relaxation task),
3. Keirn's data set provides multiple problem based tasks,
4. DEAP-Playback assesses the effects of emotions,

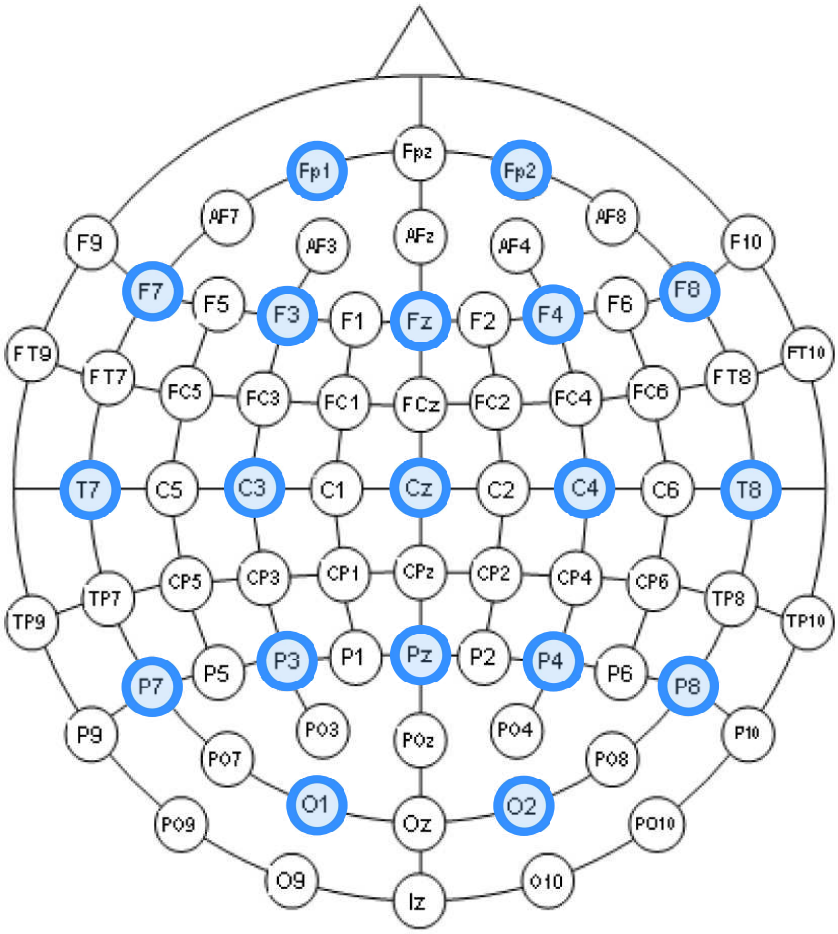


Figure 3.2: Selected EEG channels. Diagram of the 10-20 International System for EEG channel locations. Highlighted channels were kept for experimentation.

5. Yeom's and Zhang's databases contain two extensively studied VEPs, and

6. P.Ullsperger's data introduces the use of AEP.

The above variety helped us to identify properties characteristic of the neural signature and discard those due to idiosyncrasies of the set-up.

Table 3.2: Normalized databases used for experimentation. Columns show the name of the database (Database), the number of subjects (Subj.), the number of tasks (Tasks), the number of sessions (Sess.), the number of trials (Trials), the length of the EEG segments in seconds (Len.) and the descriptive keywords (Keywords).

Database	**Subj.**	**Tasks**	**Sess.**	**Trials**	**Len.**	**Keywords**
BCI2000-Baseline (BB)	100	2	1	5	10	REO; REC
BCI2000-Tasks (BT)	100	9	1	[9, 10]	4	Motor real/imagery
BCI2000-Full (BF)	100	11	1	[9, 10]	4	REO; REC; Motor real/imagery
DEAP-Baseline (DB)	20	1	1	5	20	REO
DEAP-Playback (DP)	20	4	1	[5, 10]	20	Elicited emotions
DEAP-Full (DF)	18	5	1	[5, 10]	20	REO; Elicited emotions
Keirn's (K)	5	5	2	[8, 10]	4	Problem-solving tasks
P.Ullsperger's (P)	5	2	1	180	4.1	AEP
Yeom's (Y)	10	2	2	900	1	VEP (Self-representation)
Zhang's (Z)	30	3	1	[15, 20]	1	VEP

Chapter 4

EEG artefact rejection: Localized Component Filtering

Being a non-invasive technique and a relatively inexpensive way of recording neural activity, EEG is extensively used in research and application development in all brain-related areas. One of the main difficulties faced by EEG devices is the coexistence of interesting and contaminating electrical fields. The former are generated by spiking neurons, while the latter may come from various sources, such as muscle activity or interference from electric activity in the vicinity of the recordings. These contaminating sources are in many cases greater in magnitude than the interesting ones, resulting in a low signal to noise ratio. Advances in recording technology have increased this ratio, but contamination by noise is still a major concern [131, 132].

The most direct and simple way to solve this problem is to reject the artefactual portions of the signal. However, this may not always be possible due to a limited amount of data or constraints of the recording conditions. In these cases, it is desirable to dissociate and remove the noise from the neural activity, i.e. clean the signal. The most common approaches to separate signal from noise are based on the application of Blind Source Separation (BSS) techniques, and in particular ICA [133]. These outperform other methods in rejecting high amplitude noise such

as eye movements [134].

Activity recorded at each EEG sensor represents a combination of multiple sources, some of which are based on brain activity (signals) and some of which are not (noise). BSS algorithms treat the EEG as high dimensional data, with each sensor depicting a different dimension. This space (called the original space) is then transformed into a new one (BSS space) with the aim to have each dimension (component) correspond to an individual source. To the extent that this separation of sources is successful and that artefactual sources can be identified, eliminating the corresponding dimensions and projecting the remaining components back into the original space will produce a clean signal [135].

As opposed to other problems, the presence of noise in the signal can play a positive as well as a negative role in EEG subject identification, as artefacts can arise from an individual's characteristic behavioural patterns. We described in chapter 2 the existence of identification systems based on muscle and blink artefacts [112–114]. As stated, we aimed to study the discriminant information contained in brain activity, and therefore it was necessary to isolate this information from artefacts.

Hence, we began by applying an existing algorithm: Automatic EEG artifact Detection based on the Joint Use of Spatial and Temporal features (ADJUST) [136], to the databases listed in chapter 3. However, the success of ADJUST, and of any rejection method, relies mainly on the performance of the BSS. The better the dissociation of noise and neural activity is, the better the result. S. Romero et. al. offered a comparative study of several of these techniques [137]. In general, they all provide a non-perfect separation of noise and neural activity, and BSS artefactual components do contain neural activity that gets subsequently rejected (neural leakage) [138,139]. This could be circumvented by applying more restrictive conditions during the identification of noisy and clean components, keeping those with mixed signals. The most obvious drawback of this approach is that some artefacts will now be retained.

Therefore, we devised a method to reduce the neural leakage while maintaining the same level of artefact rejection[1]. In particular, we propose a novel methodology that focuses the processing of BSS components: Localized Component Filtering (LCF). The presented algorithm localizes time segments within components contaminated by artefacts, and directs the processing to these segments, keeping the remaining parts of the component in their original form. This allows to lower the identification threshold of artefactual components, as the a priori noisy components undergo further scrutiny, reducing the probability of neural leakage.

Furthermore, we have designed LCF to be easily integrated within any existing artefact rejection system based on BSS. The LCF component can be directly embedded after the BSS^{-1} block, without any modification of the original algorithms (central panel of fig. 4.1). In addition, LCF can also be used by itself, without an artefact rejection method (right panel of fig. 4.1).

LCF is not a closed tool, but a general approach to the problem of neural leakage. Although we will later describe a prototype, this is just one possible design, and any of its components can be modified and improved, while maintaining the underlying concept. In fact, the prototype was kept as simple as possible to assess the validity of the concept.

Below, we will first define the common structure of a BSS EEG artefact rejection method. We will continue by detailing the proposed LCF methodology and describing the particularities of the prototype implemented for experimentation. The two databases (one synthesized and one real) used for testing will then be detailed, followed by the introduction of the experimentation methodology and the obtained results and discussion. We will end with the conclusions extracted from this study.

[1]This research was conducted at the Department of Psychology, College of Human Health and Science, Swansea University (Wales, UK).

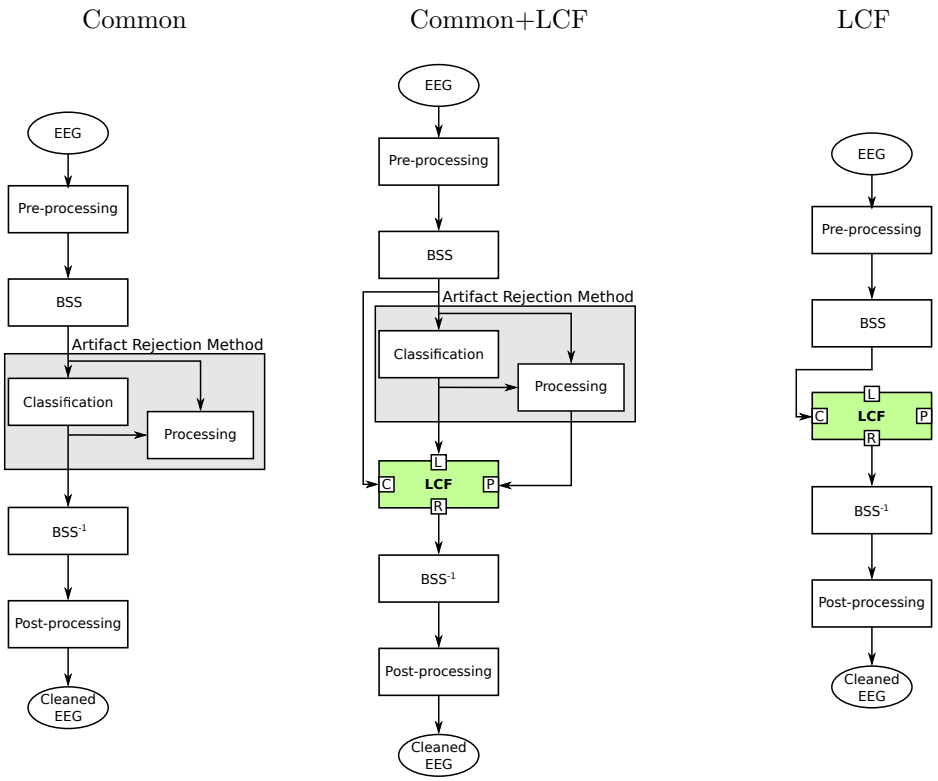


Figure 4.1: Diagrams of feature rejection systems based on BSS. (Left) Common architecture of existing systems. The actual BSS component can be seamlessly interchanged and therefore it has been left out of the "Artefact Rejection Method" box, which defines how the output of BSS is processed. (Centre) Combination of the common architecture and the proposed LCF method. The only difference with the previous diagram is the embedded LCF component. All the other blocks are exactly the same as before. Hence, the simplicity of its integration within any existing system. (Right) LCF can also be used individually, without any artefact rejection method.

4.1 Structure of a BSS EEG artefact rejection method

Six steps may be differentiated in the BSS approach (left panel of fig. 4.1). (*Pre-processing*) The data is initially prepared for further analysis. This generally includes filtering the signal and rejecting highly artefactual channels and/or segments, which may interfere with the BSS algorithm. (*BSS*) The data is then pro-

jected into the BSS space and (*Classification*) extracted components are classified as either neural or artefactual. (*Processing*) Following that, artefactual components are processed to remove their noise. In many cases this is implemented as a complete rejection (zeroing) of the components. (BSS^{-1}) The EEG signal is subsequently reconstructed using the clean components. (*Post-processing*) Finally, processing functions that work better under clean conditions are applied.

The literature contains a variety of architectures following the described procedure. For example, the classification of components into clean and artefactual can be manual or automatic. The latter, although algorithmically more complex, is preferred to avoid the introduction of subjective factors, which will hinder the reproduction of the analysis. Some supervised pattern recognition algorithms represent a very powerful option [140]. However, their requisite for labelled data sets of clean and artefactual components make them cumbersome to adopt. Most of the systems are therefore based on unsupervised techniques, which tend to be less accurate but do not require labelled data. In some cases, systems entirely lack a classification stage, treating each of the components equally [139]. In between approaches can also be found, where clean and artefactual components are processed differently [141].

Most of the available designs differentiate themselves in the way they describe the components. Topological templates of artefacts [142, 143] and statistical properties of their temporal and frequency representations [144, 145] have been extensively used to characterize noisy components. Ultimately, it is the combination of some or all of these features which provide better results [136, 146, 147].

4.2 Localized Component Filtering concept

Although current BSS techniques are unable to perfectly dissociate noise and neural activity, they do provide a representation of the data where artefacts are more easily identifiable. Hence, the principle behind the proposed LCF is to use such

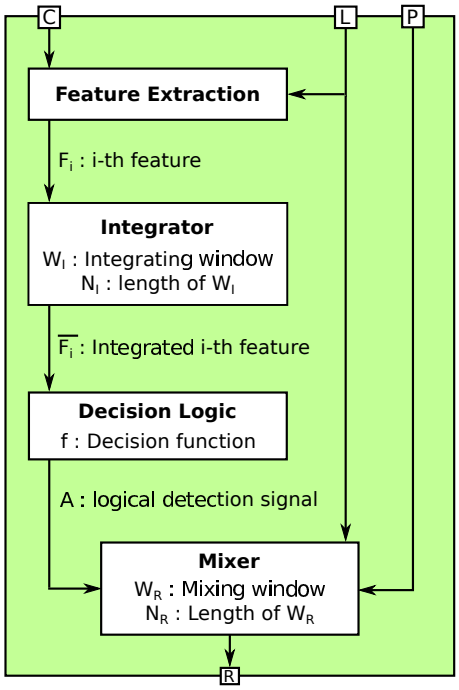


Figure 4.2: Diagram of the LCF step. The inputs and output are (C) the original BSS components, (L) the control signal pointing to the components that will be mixed, (P) the processed or alternative components, and (R) the resulting mixed components.

enhancement to apply a targeted noise treatment. By doing so, we aim to reduce the leakage of neural activity while rejecting the unwanted signal.

Crucially, LCF represents a single block that can be easily integrated within any noise rejection system. Within this block, the following four sub-blocks can be identified (fig. 4.2):

1. *Features Extraction*: Instantaneous characteristics (measures defined for each time instant n) of noisy activity are extracted from the original component. Most of the features used in the literature are statistics measured across time, channels or events, and their adaptation to instantaneous measurements is not always obvious, if at all possible. Features such as the instantaneous voltage and its time derivative can be easily used here. Spectral measurements can also be used if a time-frequency analysis is applied. Other features, such as variance or kurtosis, may be computed within a sliding window centred at

each instant. This module outputs a list of features, with $F_i[c, n]$ representing the i-th feature measured from the c-component at instant n.

2. *Integrator*: Because instantaneous features are usually noisy within the surroundings of an artefactual event, a more stable version is computed by applying an integrating window around each time instant (fig. 4.3). The integrated features $\overline{F_i[c, n]}$ are defined as

$$\overline{F_i[c, n]} = \frac{F_i[c, n] * W_I}{\sum_{k \in K} W_I(k)}, \tag{4.1}$$

where W_I is a window of length N_I, $*$ is the convolution operator, and K is the integration range defined as

$$K = [max(0, N_I/2 - n), min(N_I - 1, N_I/2 + N_C - n)], \tag{4.2}$$

with N_C the length of the component. The range K covers the entire integrating window except when it reaches the beginning or end of the component. The denominator of equation 4.1 is a normalizing factor that effectively transforms the integration into a weighted average around each time instant. It also counteracts the boundary effects of the convolution operation, so that artefacts at the beginning and end of the signal segment can be correctly detected.

3. *Decision Logic*: Once features are extracted and integrated, we identify the presence of noise within each component c at each time instant n. As with the *feature measurement* block, the current module can be designed in many different ways: from using a fixed threshold, to applying more sophisticated classification algorithms. We compute a single logical vector $A[c, n] = f(\{F_i[c, n], \forall i\})$, where f is the decision logic function.

4. *Mixer*: Finally, the detection vector is used as a control signal to mix the original components $C[c, n]$ and the processed (cleaned) ones $P[c, n]$ given by the artefact rejection system (in the absence of a separate artefact rejection

system, $P[c, n]$ can be simply set to zeros). Let M be a mixing signal defined as

$$M[c, n] = A[c, n] * W_M, \tag{4.3}$$

where W_M is a window of length N_M and $\sum W_M[n] = 1$, used to round the edges of detected areas, which in turn smooths the transitions in the mix. The resulting component is defined as

$$Q[c, n] = P[c, n] \cdot (1 - M[c, n]) + C[c, n] \cdot M[c, n]. \tag{4.4}$$

Therefore, Q is a mixed vector where the amount P (alternative/processed component) and C (original/raw component) is determined by the detection of artefacts. If a segment is identified as clean, only the original component will be included in the mix and no signal will be lost. If a segment is identified as artefactual, only the processed component will be used. Borders between P-only and C-only segments contain a mix between the two signals, smoothing the transitions from one to the other, avoiding the introduction of discontinuities (fig. 4.4).

The resulting components can then be back-projected to the original space. Since only artefactual areas are processed, the system is now better protected against the leakage of neural activity. Hence, the result is a robuster version of whatever the original BSS-based rejection method was.

4.2.1 Implemented LCF prototype

We implemented a prototype of the LCF methodology to evaluate the concept. Again, this is just one possible implementation, and many changes can be applied to each of the blocks just described.

1. *Feature Extraction*: As stated before, even though BSS algorithms are incapable of a perfect dissociation between neural and artefactual sources, they

do provide a space where noise in artefactual dimensions stands out from neural activity. As a result, the voltage of these components can be used as an indicator of the presence of noise. Equally, high absolute values of the first time derivative indicate the position of high frequency noise. These two measurements are instantaneous and easy to obtain, and therefore they were included in the current implementation.

In addition, both features were normalized by the equation

$$\Theta(X) = \frac{X - mean[\widehat{X}]}{std[\widehat{X}]}, \tag{4.5}$$

where $mean(\cdot)$ and $std(\cdot)$ are the global mean and std and $\widehat{X}$ is a trimmed version of the vector X built by excluding samples that deviate more than three times the std. Thus, features are formally defined as

$$F_1 = \Theta(|C|) \tag{4.6}$$

and

$$F_2 = \Theta(|dC/dt|) \tag{4.7}$$

2. *Integrator*: We configured the integrator block to have W_I equal to a Hamming window of 0.2 seconds (fig. 4.3).

3. *Decision Logic*: We used a simple threshold (τ) to determine whether a given feature indicated the presence of noise (which corresponds to a value of 1 in the decision vector A), so that

$$V_i[c,n] = \begin{cases} 1 & \text{if } F_i[c,n] > \tau \\ 0 & \text{otherwise.} \end{cases} \tag{4.8}$$

A threshold of $\tau = 1$ was set empirically (note that F_i are z-scored vectors).

However, the response of V_i was still unstable within artefactual segments. To overcome this, detected zones were dilated by an all ones window W_D of

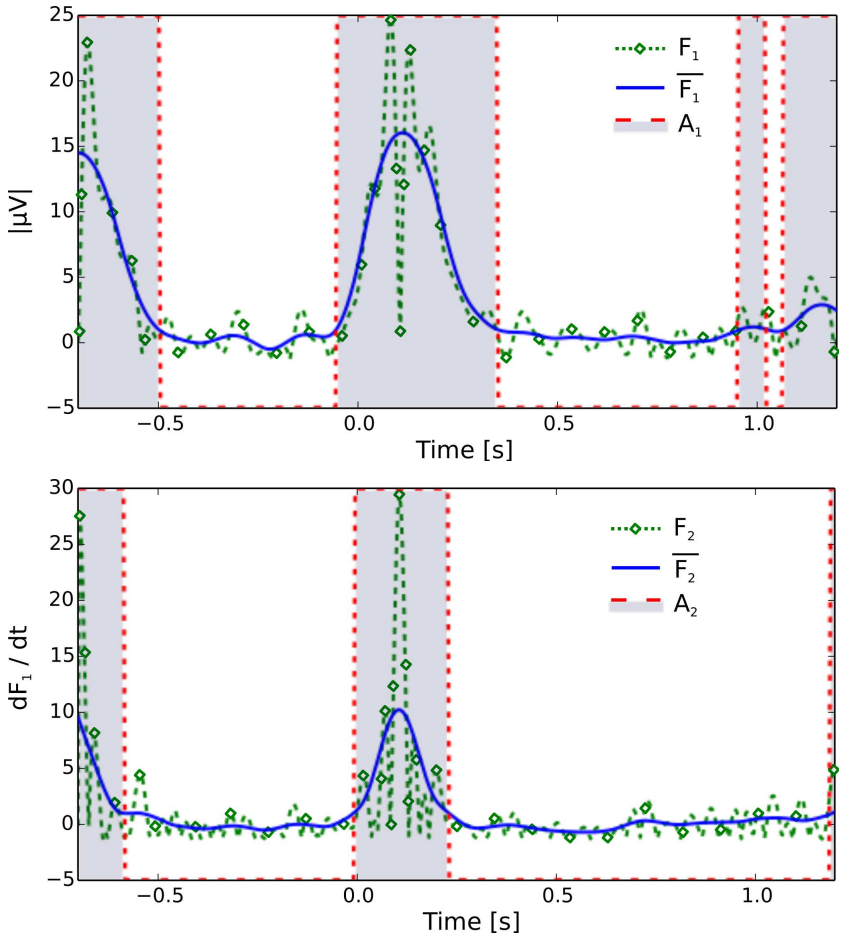


Figure 4.3: Examples of F_1 (top) and F_2 (bottom) defined in the implemented LCF block. Their integrated version $\overline{F_i}$ is smoother, allowing for a more robust detection of A_i (in the figure we expanded the amplitude of A_i to the length of the y axis).

length $N_D = 0.1$ seconds, which produced continuous noise regions that were well separated from adjacent noise regions. The result was then binarized to maintain the mask-like nature of the vector, resulting in

$$\overline{V_i[c,n]} = bn(V_i[c,n] * W_D), \tag{4.9}$$

where $bn(\cdot)$ is a binarizing operator which sets all values different than 0 equal to 1. We then combined the responses from each feature in a single logical detection signal $A[c,n] = \overline{V_1[c,n]} \vee \overline{V_2[c,n]}$, with $\vee$ the logical "or" operator, which pointed the location in time of artefacts within each component.

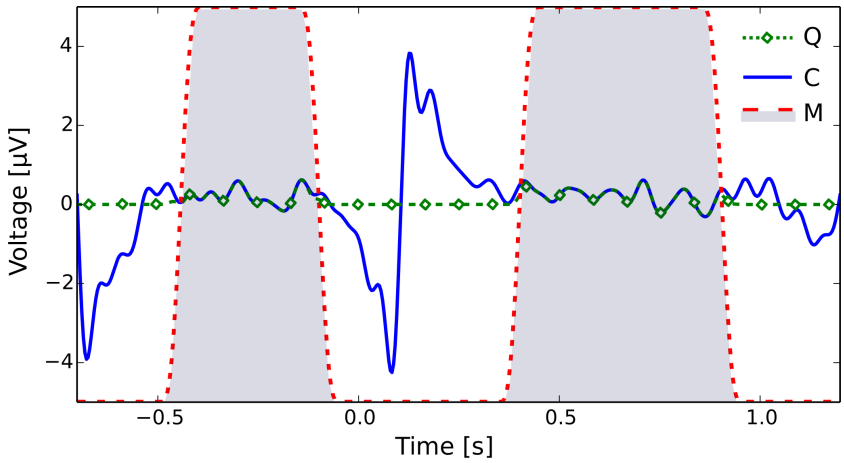


Figure 4.4: Mixing of the component C whose features are depicted in figure 4.3 when its processed version P are all zeros. Transitions between *on* and *off* are smoothed to avoid discontinuities. The amplitude of M was modified to match the y axis.

Two more extra logical operations were finally applied to account for highly noisy events and components. Firstly, if more than 75% of a trial was labelled as noise, the whole trial was rejected. Secondly, if more than 75% of a component was labelled as noise, the whole component was rejected.

4. *Mixer*: We set W_R (from the mixer block defined in the previous section) equal to a Hamming window of 0.1 seconds (fig. 4.4).

4.3 Materials for the validation of LCF

As this study was developed at Swansea University (Wales, UK) we used a different set of databases to evaluate the proposed method. In particular, we utilized a sample of the simulated data set used to assess Fully Automated Statistical Thresholding for EEG artifact Rejection (FASTER) in its original work [146,148], and a real data set recorded by us at Swansea University.

The simulated database has its neural and artefactual activity perfectly dissociated and defined, which allows us to quantify the effect of LCF by measuring how much of the injected noise is rejected and how much of the neural activity is

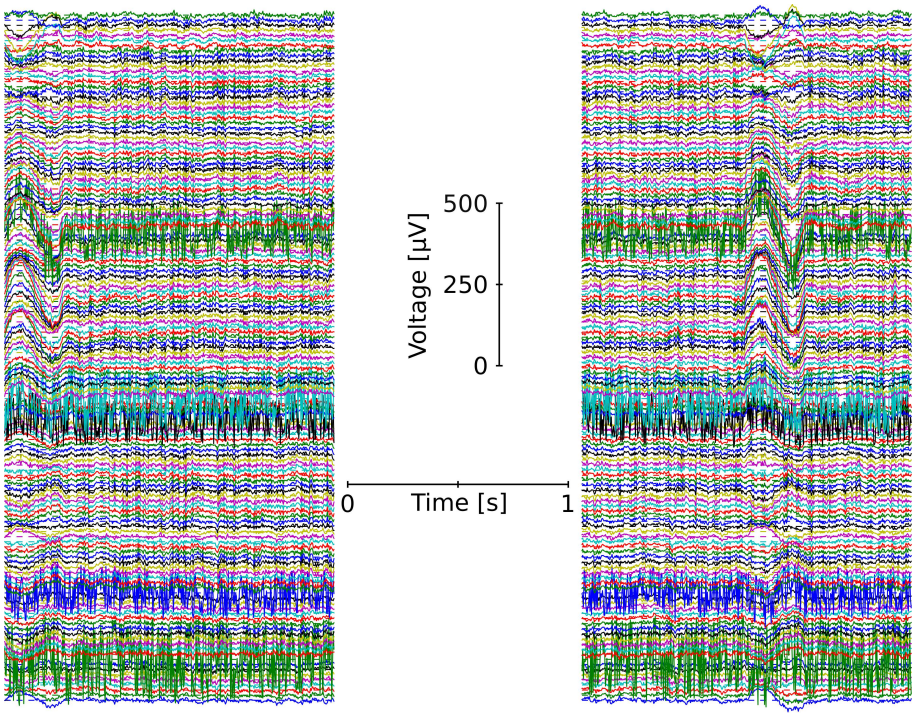


Figure 4.5: Examples of the simulated data set used on FASTER's original work. Contaminated channels and blinks are particularly prominent.

lost (fig. 4.5). Five files with 200-epochs each were available, simulated with the BESA Dipole Simulator program [149]. Artefacts were randomly added following the procedure described in [145]. In addition, a random number of channels were contaminated with white noise, and some of the epochs were also corrupted by a high-amplitude (30–150 μV) low frequency (1–3 Hz) wave.

However, this database represents a simpler version of the problem. To evaluate the system under real conditions, we applied it to real data (fig. 4.6). This contained the EEG recordings of 18 participants while they performed a memory based task. A total of 133 electrodes (128 scalp + 2 mastoids + 3 EOG) from a BIOSEMI system were recorded with a sampling rate of 500 Hz. From each participant, 576 trials were extracted from 0.5 seconds pre-stimulus onset to 1 second post-stimulus onset.

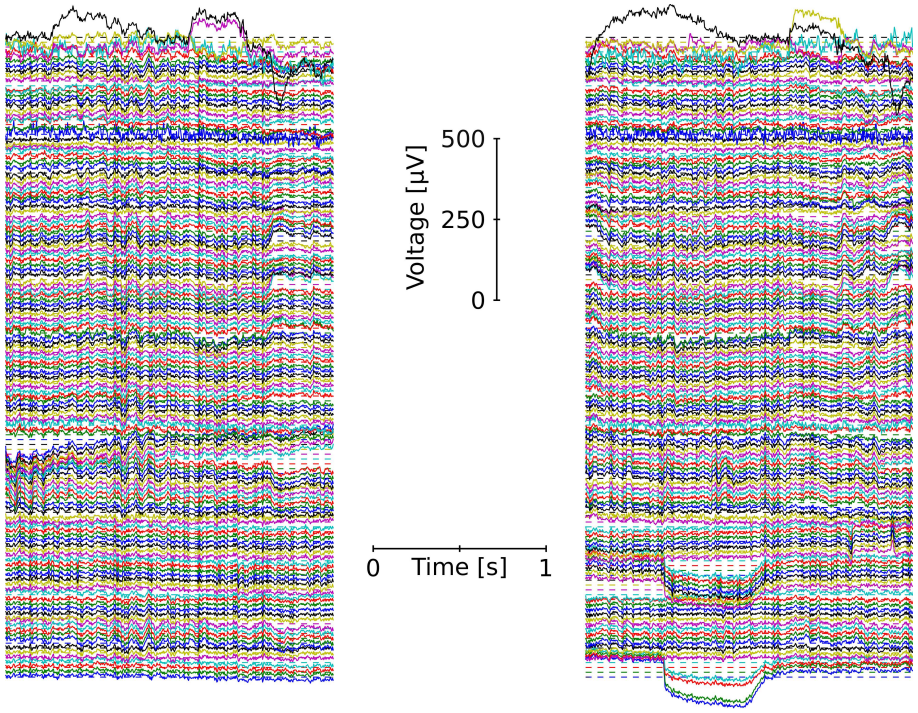


Figure 4.6: Examples of the real data set recorded from a memory based task. Artefacts are substantially more complex than in the synthesized EEG.

4.4 Experimentation methodology

The evaluation of any artefact rejection system is a complex process, especially since it is still impossible to identify noise and neural activity with full certainty. In general, visual judgements and numerical measurements rely on assumptions based on known characteristics of EEG activity. Any deviation from the expected behaviour is therefore labelled as artefactual.

In this case, to assess the effectiveness of the proposed method, we performed the following studies over a set of systems with and without LCF.

4.4.1 Evaluated systems

To generalize the benefits of the proposed LCF methodology, we assessed the performance of three different systems based on BSS algorithms. In order to isolate the effects of LCF in combination with each of the evaluated methods, we

designed the systems to share all the blocks described in section 4.1, except the two enclosed within the "Artefact Rejection Method" label (fig. 4.1).

Thus, the EEG data was first high-pass filtered at 0.5 Hz with a FIR filter of order 99 and Hamming window, and low-pass filtered at 40 Hz with a similar filter of order 100. The baseline was then removed from the signal and highly artefactual channels and events were automatically rejected based on the z-score of several statistics (variance, correlation and Hurst exponent for channel rejection; voltage range, variance and channel deviation for epoch rejection) following the same preprocessing described in [136].

Next, the INFOMAX ICA algorithm was used as a BSS technique [150]. Once the Independent Components (ICs) were computed, each of the following artefact rejection methods were individually applied:

- ADJUST [136]: This method characterizes artefactual ICs by both temporal and spatial features: kurtosis, variance and the spatial distribution of ICs activations. These are then automatically classified into clean or artefactual ICs by an Expectation Maximization algorithm, which maximizes the likelihood of the distribution of classes by an iterative process [151]. The detected artefactual ICs are then zeroed.

- FASTER [146]: In this case, artefactual ICs are described mainly by temporal measurements. In particular, descriptors are: temporal correlation with EOG channels, spectral and voltage gradient, Hurst exponent, and spatial kurtosis. Outliers, as per their z-score, are then labelled as artefactual. The detected artefactual ICs are then zeroed.

- Wavelet enhanced ICA (ICAW) [139]: This technique processes each IC by thresholding their Discrete Wavelet Transform (DWT) coefficients. A threshold equal to the 99.5 percentile of the absolute wavelet coefficients was set empirically.

We post-processed the cleaned reconstructed EEG data by correcting the baseline and interpolating the rejected channels using the spherical spline technique [130]. Finally, we re-referenced the EEG signal to common average and locally interpolate channels showing noisy activity within epochs. This was again implemented as described in [136].

Two versions of the above systems, with and without the LCF component described above, were therefore evaluated. In addition, to asses the absolute effect of LCF, we also tested a system without any artefact rejection method. In this case, the LCF block was fed with an all-zeros alternative signal and configured to process all the extracted ICs (*None+LCF*).

4.4.2 Quantitative analysis of synthetic EEG data

The simulated dataset has the advantage of offering fully controlled testing conditions. Hence, noise and neural activity can be completely dissociated by $Z = Y - X$, where X is the original EEG signal, Y is its contaminated version and Z is the injected artefacts. Similarly, the rejected signal is defined as $R = Y - ar(Y)$, with $ar(\cdot)$ any artefact rejection method. Therefore, under the assumption that X is pure neural activity, any deviation from $X = ar(Y)$, and therefore $Z = R$, is considered misses of the applied process.

In particular, we defined the following measurements:

1. True Rejection Power Rate (TRPR): This is the proportion of power from the artefact signal that has been successfully removed. It corresponds to the overlapped area under Z and R, and is defined as

$$TRPR = \frac{\sum\limits_{n \in \Omega} min(|Z(n)|, |R(n)|)^2}{\sum\limits_{\forall n} Z(n)^2}, \tag{4.10}$$

   where $\Omega = \{n \mid sign(Z(n)) == sign(R(n))\}$.

2. False Rejection Power Rate (FRPR): This accounts for the leakage of neural activity, either by an overcompensation of the artefact or by miss-classification

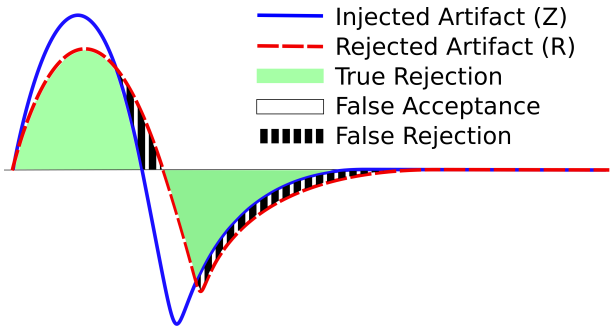


Figure 4.7: Graphical representation of the true and false rejection and false acceptance concepts.

of a clean signal as noise. It corresponds to the area under R and above Z. It is defined as

$$FRPR = \frac{\sum\limits_{n\in(\Omega\cap\Psi)} (R(n) - Z(n))^2 + \sum\limits_{n\ni\Omega} R(n)^2}{\sum\limits_{\forall n} X(n)^2}. \tag{4.11}$$

with $\Psi = \{n \mid |Z(n)| < |R(n)|\}$.

3. True Acceptance Power Rate (TAPR) and False Acceptance Power Rate (FAPR): Once the previous values were computed, TAPR and FAPR were derived from them as $1 - FRPR$ and $1 - TRPR$ respectively.

Note that the worst cases are whenever R and Z have different polarities, as there is no TRPR and the area under Z and R corresponds to FAPR and FRPR respectively (fig. 4.7). Based on the above definitions, we seek to obtain high TAPRs, i.e. retain as much of the original EEG as possible, and low FAPRs, i.e. reject as much of the injected artefacts as possible.

To isolate the influence of LCF, the original X and noisy Y versions of the signal were extracted at the exit of the pre-processing stage, while the cleaned version $ar(Y)$ was sampled at the output of the BSS^{-}1 block (fig. 4.1).

4.4.3 ERP analysis of real EEG data

A common way to analyze EEG data is to average EEG activity across events to produce an ERP. We partitioned the 128 EEG channels into 9 regions of interest (within which we averaged the sensors) as shown in Figure 4.8. For each participant we then computed the ERP for each region and then averaged those ERPs across participants.

In real data sets without labeled artifacts the quantities we used to assess performance on the simulated data set are not readily available. Because the baseline period of each event in the real data set lacked external stimuli that were synchronized across events, we expect the average ERP to be fairly constant and close to 0 μV for clean EEG data. However a small proportion of artifacts (such as those produced by motion) can cause significant deviations from that mean due to their large amplitude.

Our events were locked to the presentation of a test stimulus and it is common to see ERP deflections reflecting the processing of such stimuli. Noise, however, can obscure these synchronous effects of stimulus processing in the EEG and attenuate the resulting ERP. On the other hand, removal of neural activity will also reduce the signal to noise ratio and attenuate the ERP.

As an extra tool for the analysis of the results, we studied the rejected ERPs. These were computed by subtracting the cleaned ERP (by whatever method applied) from the original one. Hence, we expected them to show all the pre-stimulus noise – or any other noise – of the original ERP and no actual ERP wave after the stimulus onset, meaning that all the noise has been removed while all the ERP power has been retained.

Finally, to directly assess the effects of embedding LCF into an existing artifact rejection method, we studied the difference between the ERPs of the original methods and those with LCF integrated. According with the described effects of artifacts, we expected the resulting signals to show activity only within the post-stimulus ERP, meaning that LCF retrieved the lost neural signal while rejecting

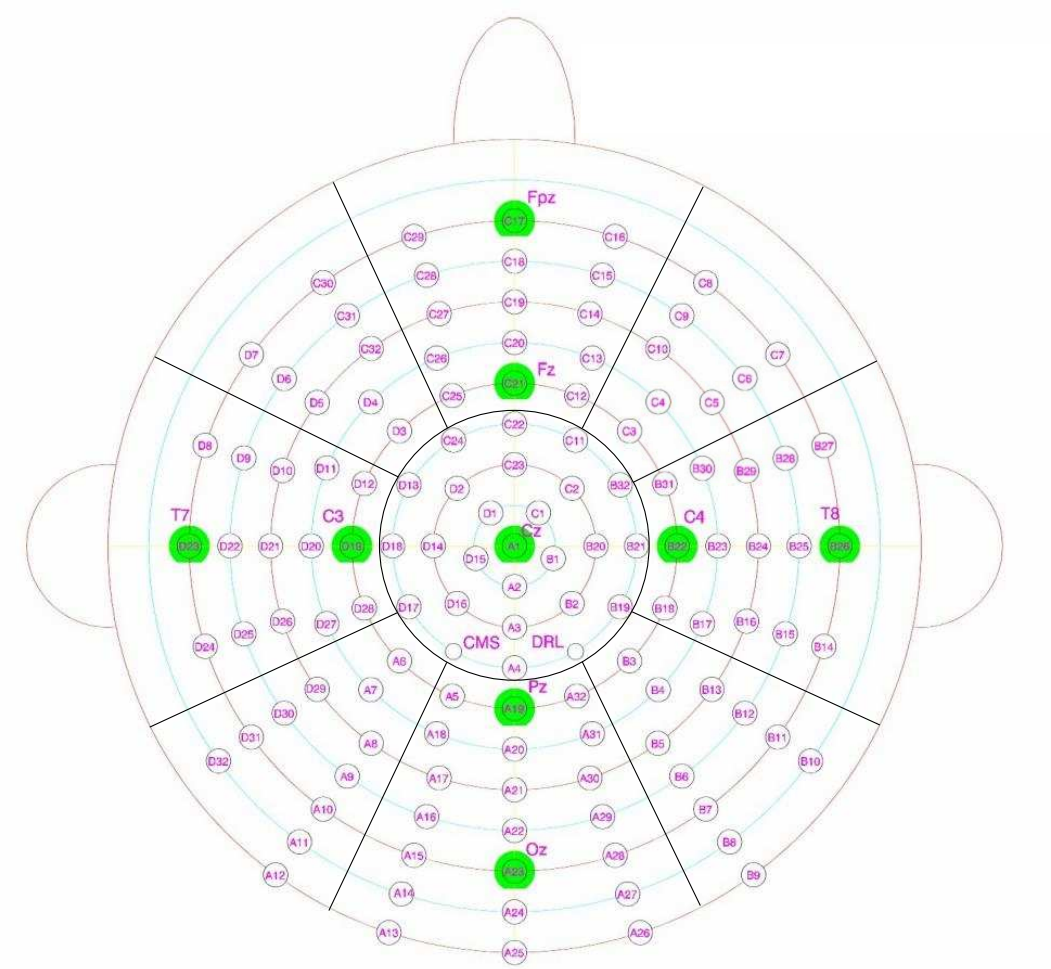


Figure 4.8: Computed regions by averaging BIOSEMI sensors within each area. Channels falling on top of the division lines are included in the cardinal regions, i.e. those with angles 0 deg, 90 deg, 180 deg and 270 deg as measured from the sagittal plane.

the noise.

4.4.4 Time-frequency analysis of real EEG data

Another common way to analyze EEG activity is to perform a time-frequency decomposition of the signal. Because these analyses often reveal interesting effects even for frequencies beyond the 40 Hz low-pass filter that we applied for the ERP analyses (section 4.4.1), we repeated the processing steps without applying a low-pass filter for the time-frequency analyses. We calculated power for 64 frequencies

Table 4.1: Numerical results obtained with the synthetic database as detailed in section 4.4.2. Mean and std is provided for the percentage of processed ICs (% ICs), TAPR and FAPR.

	% ICs	**TAPR**	**FAPR**
None+LCF	All	45.70% ± 1.78	14.64% ± 1.17
ADJUST	2.05% ± 1.40	97.78% ± 1.72	62.77% ± 44.09
ADJUST+LCF		99.33% ± 0.51	66.47% ± 40.59
FASTER	14.67% ± 2.88	61.98% ± 10.12	11.72% ± 1.90
FASTER+LCF		81.51% ± 5.14	12.01% ± 1.68
ICAW	All	51.97% ± 3.78	26.97% ± 3.73
ICAW+LCF		61.74% ± 2.44	27.00% ± 4.11

(logarithmically spaced between 2 and 100 Hz) using Morlet's wavelets with 5 cycles. We averaged power across events and z-transformed the resulting power separately for each frequency on the basis of the respective mean and standard deviation in the 0.5 s pre-stimulus baseline. For simplicity, we will refer to this representation as Event Related Spectrogram (ERS).

4.5 Results and discussion

Results for both quantitative and qualitative analyses will be provided in this section. Note that quantitative analyses are computed from the simulated dataset, which represents a version of the problem with ideal conditions. Hence, results are only an optimistic estimation of the properties of the evaluated systems, and should be considered as such.

4.5.1 Simulated EEG data results

In general, the application of LCF reduced the leakage of neural activity (increased TAPR) without heavily penalizing the rejection of noise (maintained FAPR). Nevertheless, each system showed its own characteristic behaviour (table 4.1).

ADJUST proved to be the most conservative method during classification of ICs, with only 2% of the ICs rejected on average (in one case, no IC was rejected). This left LCF with a reduced scope to improve the TAPR – it obtained less than 2 percentage points of improvement. At the same time, it led to extremely high an variable FAPRs.

On the opposite end of the spectrum sat ICAW, which processed all the available ICs. The FAPR improved with respect to ADJUST by more than 35 percentage points, but at the expense of a 45 percentage point drop on the TAPR. Here, LCF increased the TAPR by almost 10 percentage points while not affecting the FAPR level, i.e. reduced the neural leakage without penalizing the rejection of noise.

Halfway between ADJUST and ICAW lay FASTER, selecting between 11% and 18% of the ICs for rejection. It also provided the best overall results, with 62% of TAPR and 12% of FAPR. When combined with the proposed LCF, the TAPR increased to 81%, again with virtually no deterioration of the FAPR. In addition, the stability of the TAPR was also improved, reducing its std from 10 to 5.

By itself, LCF obtained relatively good results. In particular, it achieved a TAPR of 46% (comparable to ICAW) and an FAPR of 15% (comparable to FASTER).

Finally, looking at the individual results (prior to averaging across sets), there were occasions where LCF improved the detection of artefacts (reduced the FAPR). This may seem strange at first given that the method reduces the amount of signal rejected. However, removing non-artefactual parts of an IC may disrupt the reconstruction of EEG data, resulting in introduced noise. Depending on the polarity of this noise, and given the way we defined the performance measurements, this can result in an increase in FRPR or in FAPR. Thus, its correction can results in an improvement of FAPR.

4.5.2 Real EEG data results

The results for the real EEG data set mirror those from the simulated data (fig. 4.9 to 4.12). ADJUST only tagged few ICs as artifactual, limiting the scope for improvement that could be gained by LCF. FASTER, on the other hand, rejects more ICs which results in more rejected noise and larger levels of neural leakage, which LCF reduces. ICAW processes all ICs, which maximizes the potential effectiveness of LCF.

Because this was a real case scenario, as oposed to the simulated data, there were situations where the BSS algorithm found an especially bad solution for the dissociation of neural and artefactual activity, which resulted in the benefits of LCF being magnified (fig. 4.9). When averaging the ERPs across subjects, differences between methods, although still present, are less obvious (fig. 4.10). By improving the ERP of those especially bad cases, LCF allows to retain cases that may be discarded otherwise, increasing the volume of data available for the EEG analysis.

Looking at the ERPs (fig. 4.10), whereas ADJUST was not affected much by the introduction of LCF, the amplitude of the ERPs increased markedly with LCF when FASTER was used as an artifact rejection method. Notably, this increase in amplitude was confined to the ERPs in response to stimulus onset – no such increases in amplitude are discernible in the baseline periods or more than about 500ms past stimulus onset. A similar pattern of results is evident for ICAW. Finally, LCF by itself (*None+LCF*), performed better than we anticipated, despite the simplicity of the current implementation. The LCF only implementation rejected a good amount of noise (as evident in the reduced baseline amplitude) while retaining most of the ERP.

A complementary way to visualize the results is to examine the rejected ERPs (i.e., ERPs based on the difference between the processed and the original signal) (fig. 4.11). To the extent that artifactual activity is not time-locked to stimulus onset, rejected ERPs should be around zero and significant deflections indicate the removal of time-locked activity. It is evident that LCF tended to reduce the

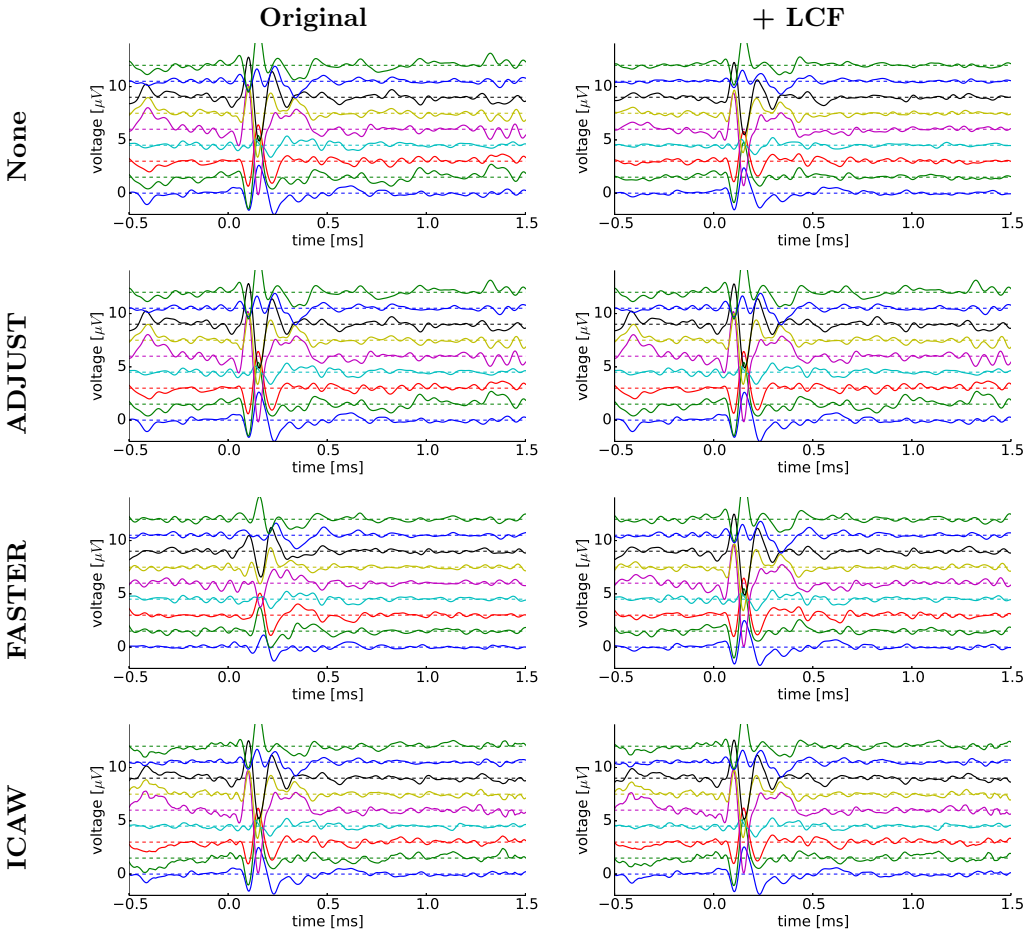


Figure 4.9: ERPs for one participant after applying each of the systems defined in section 4.4.1. Systems are applied both without (Original) and with a subsequent LCF step (+ LCF). The top left image corresponds to the ERPs of the original signal. Lines correspond to the brain regions specified in figure 4.8, so that the top line is RC and subsequent ones are R0D, R45D, R90D, ..., R315D.

rejection of time-locked activity and that FASTER was particularly effective at avoiding the removal of time-locked activity.

Finally, the effects of integrating LCF within the original rejection methods can also be assesed by their contrast (i.e. substracting the LCF integrated system with the original, no LCF design) (fig. 4.12). With the exception of ICAW, it is evident that differences between the two are mainly confined within the stimulus locked ERP, with LCF having little effect outside this area. This adds to the evidence

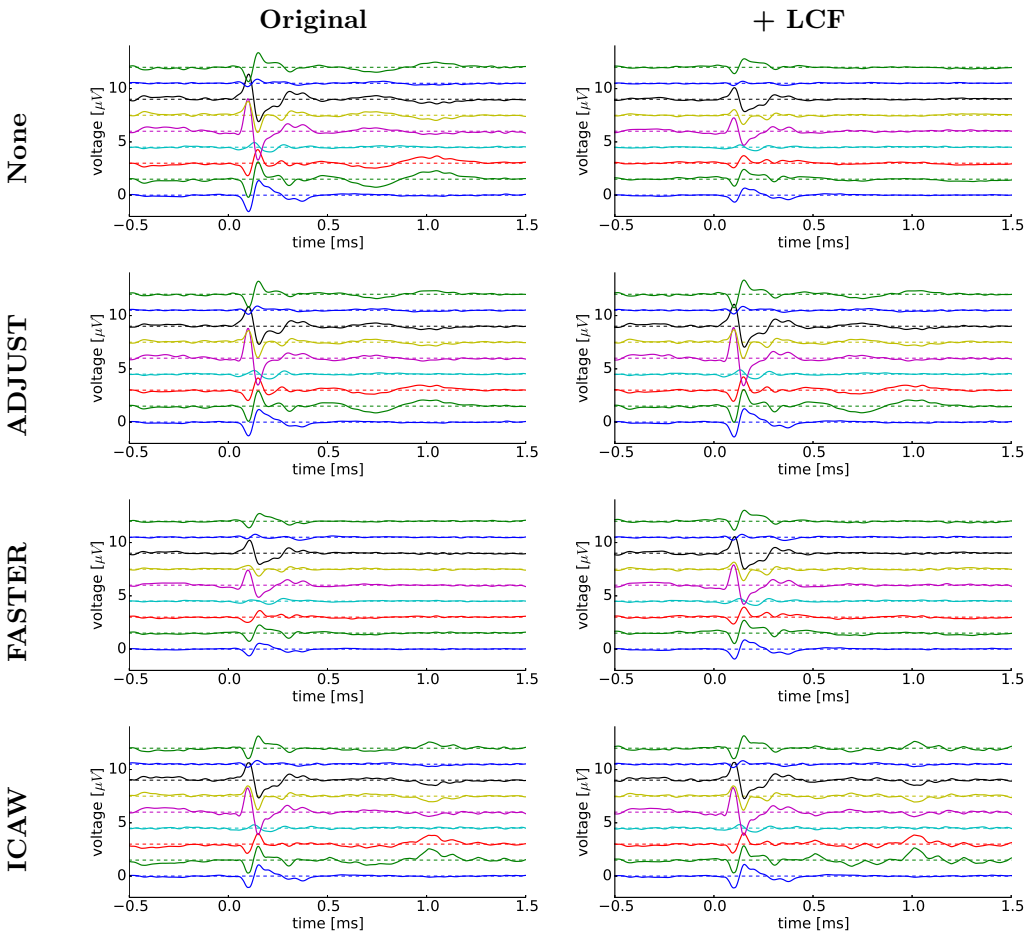


Figure 4.10: ERPs averaged across participants after applying each of the systems defined in section 4.4.1. Systems are applied both without (Original) and with a subsequent LCF step (+ LCF). The top left image corresponds to the ERPs of the original signal. Lines correspond to the brain regions specified in figure 4.8, so that the top line is RC and subsequent ones are R0D, R45D, R90D, ..., R315D.

that LCF regains part of the leaked neural signal, without heavily penalizing in the noise removal.

4.5.3 Time-frequency analysis

The same set of systems, but without a 40 Hz low-pass filter within the preprocessing step, were applied again to both data sets. Overall, the same relationship

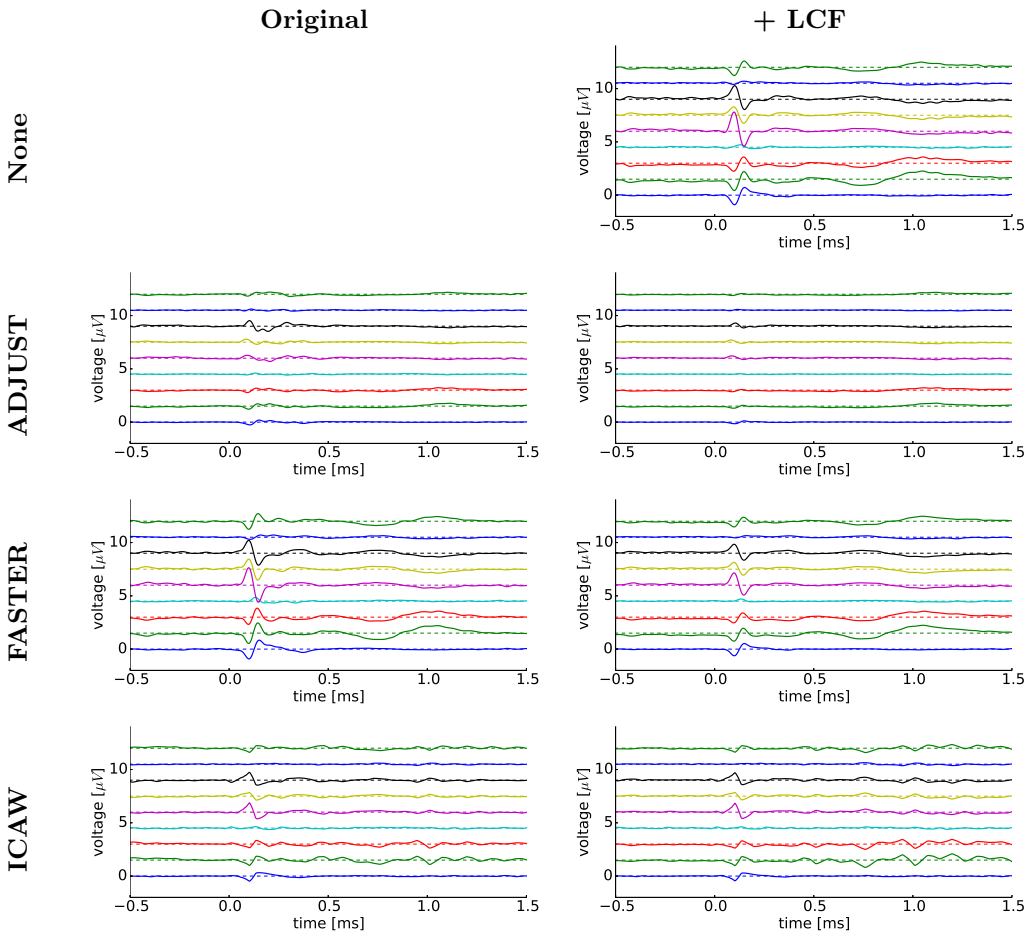


Figure 4.11: Rejected ERPs averaged across participants after applying each of the systems defined in section 4.4.1. Systems are applied both without (Original) and with a subsequent LCF step (+ LCF). Lines correspond to the ERPs of each brain region specified in figure 4.8, so that the top line is RC and subsequent ones are R0D, R45D, R90D, ..., R315D.

between original and LCF systems was observed, with the latter reducing neural leakage.

Interestingly, results of the synthetic data show an improvement in all the systems except those based on FASTER (table 4.2). This was especially true for ADJUST, which lowered the std of its FAPR from 62 to 3. Crucially, LCF discriminated better between noise and neural activity, dropping the FAPR level from 15% to 11%. Looking at the ERP's amplitude of the real EEG data set (fig. 4.13 and 4.14), we can see a similar overall improvement.

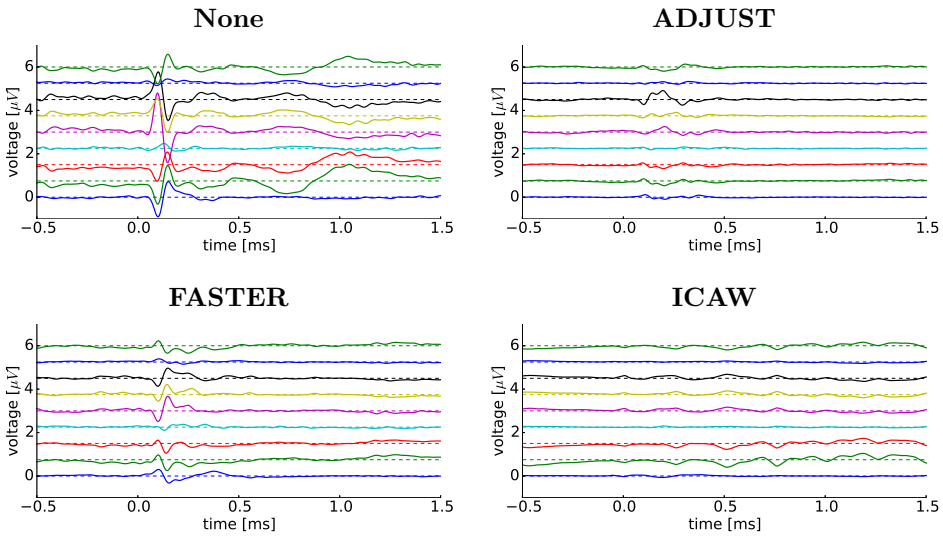


Figure 4.12: Difference between the ERPs of the systems defined in section 4.4.1 without and with LCF. Lines correspond to the ERP difference of each brain region specified in figure 4.8, so that the top line is RC and subsequent ones are R0D, R45D, R90D, ..., R315D.

Table 4.2: Numerical results obtained with the synthetic database and no 40 Hz low-pass filter. Mean and std is provided for the percentage of processed ICs (% ICs), TAPR and FAPR. Refer to table 4.1 to compare these with the results of the original systems (with 40 Hz low-pass filter).

	% ICs	**TAPR**	**FAPR**
None+LCF	All	81.96% ± 1.85	11.07% ± 1.17
ADJUST	2.96% ± 1.96	96.26% ± 2.49	11.55% ± 3.09
ADJUST+LCF		98.80% ± 1.03	17.09% ± 4.29
FASTER	11.93% ± 1.16	75.57% ± 11.50	18.40% ± 6.07
FASTER+LCF		97.83% ± 0.79	20.61% ± 5.23
ICAW	All	73.28% ± 2.56	30.99% ± 1.34
ICAW+LCF		85.88% ± 2.12	32.12% ± 1.52

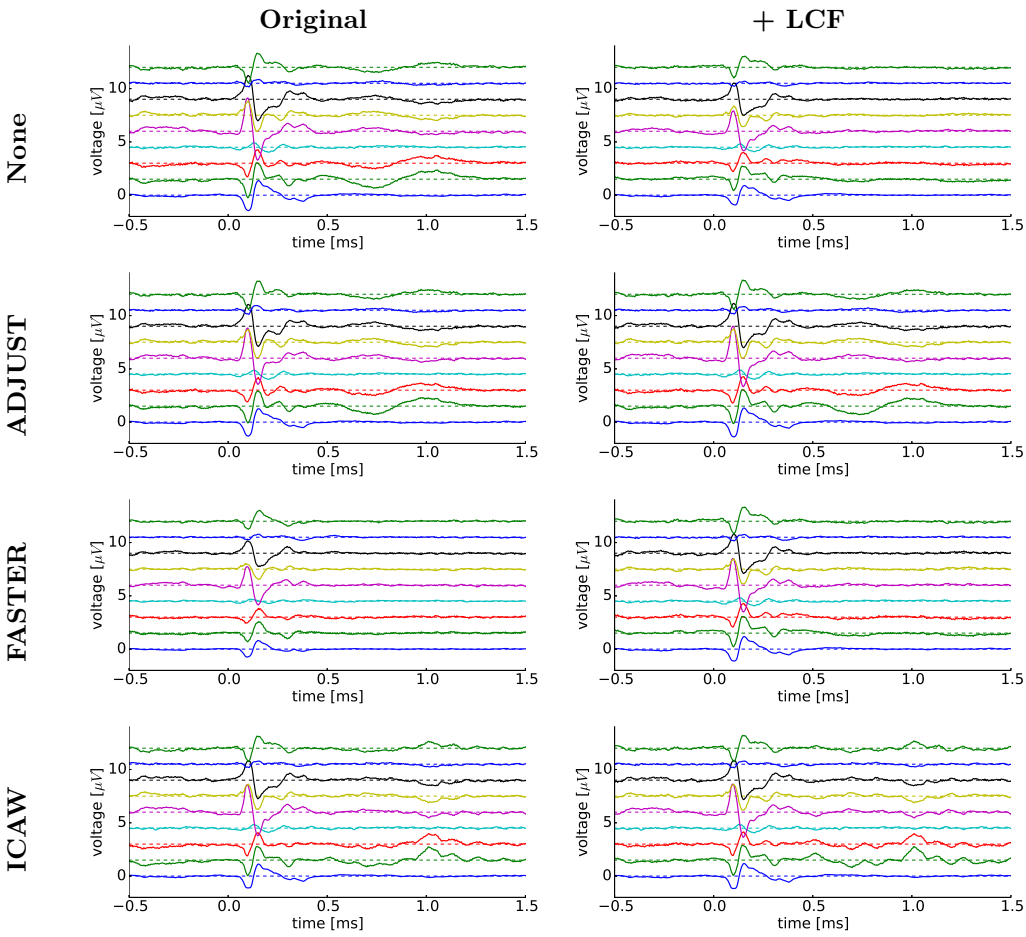


Figure 4.13: ERPs averaged across participants after applying each of the systems defined in section 4.4.1 without low-pass filter. Systems are applied both without (Original) and with a subsequent LCF step (+ LCF). The top left image corresponds to the ERPs of the original signal. Lines correspond to the brain regions specified in figure 4.8, so that the top line is RC and subsequent ones are R0D, R45D, R90D, ..., R315D.

The interpretation of the ERSs is more complicated than that of the ERPs. There are less conceptions about how the ERSs should look like, as opposed to ERPs, which, for example, we know that they should be flat prior to the stimulus on-set. From the time-frequency representation of the real EEG data, the effects of LCF seemed to be mainly localized between 2 and 40 Hz (fig. 4.15). The selection of artifactual ICs by ADJUST, FASTER and ICAW focused these effects on the lowest frequencies (between 2 and 10 Hz). Overall, the usage of any of

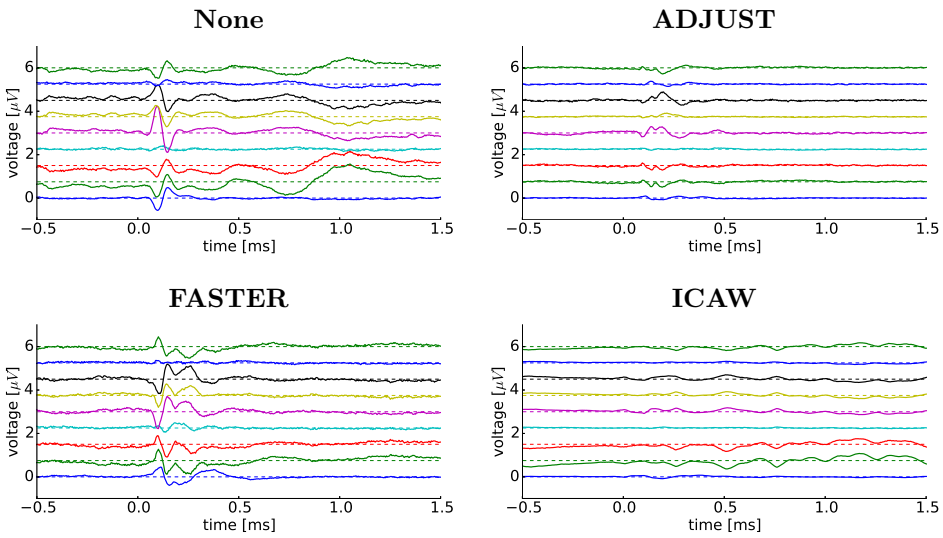


Figure 4.14: Difference between the ERPs of the systems defined in section 4.4.1 without low-pass filter without and with LCF. Lines correspond to the ERP difference of each brain region specified in figure 4.8, so that the top line is RC and subsequent ones are R0D, R45D, R90D, ..., R315D.

these methods with or without LCF had no apparent disruptive effects on the spectrogram, and can, therefore, be used on a time-frequency analysis.

4.6 Conclusions

In this chapter, we have proposed LCF as a novel methodology to boost the performance of existing EEG artefact rejection systems that are based on BSS techniques. The method takes the original and processed components as inputs, and mixes them so that the processed (cleaned) components replace the original ones only when an artefact is detected. We have performed quantitative and qualitative analyses on simulated and real datasets, demonstrating the benefits of LCF, especially in the reduction of neural leakage.

LCF was designed to be compatible with any of the existing systems based on BSS. Moreover, its integration is straight forward and requires no big changes on the system itself, facilitating its adoption. On the other hand, LCF's ability

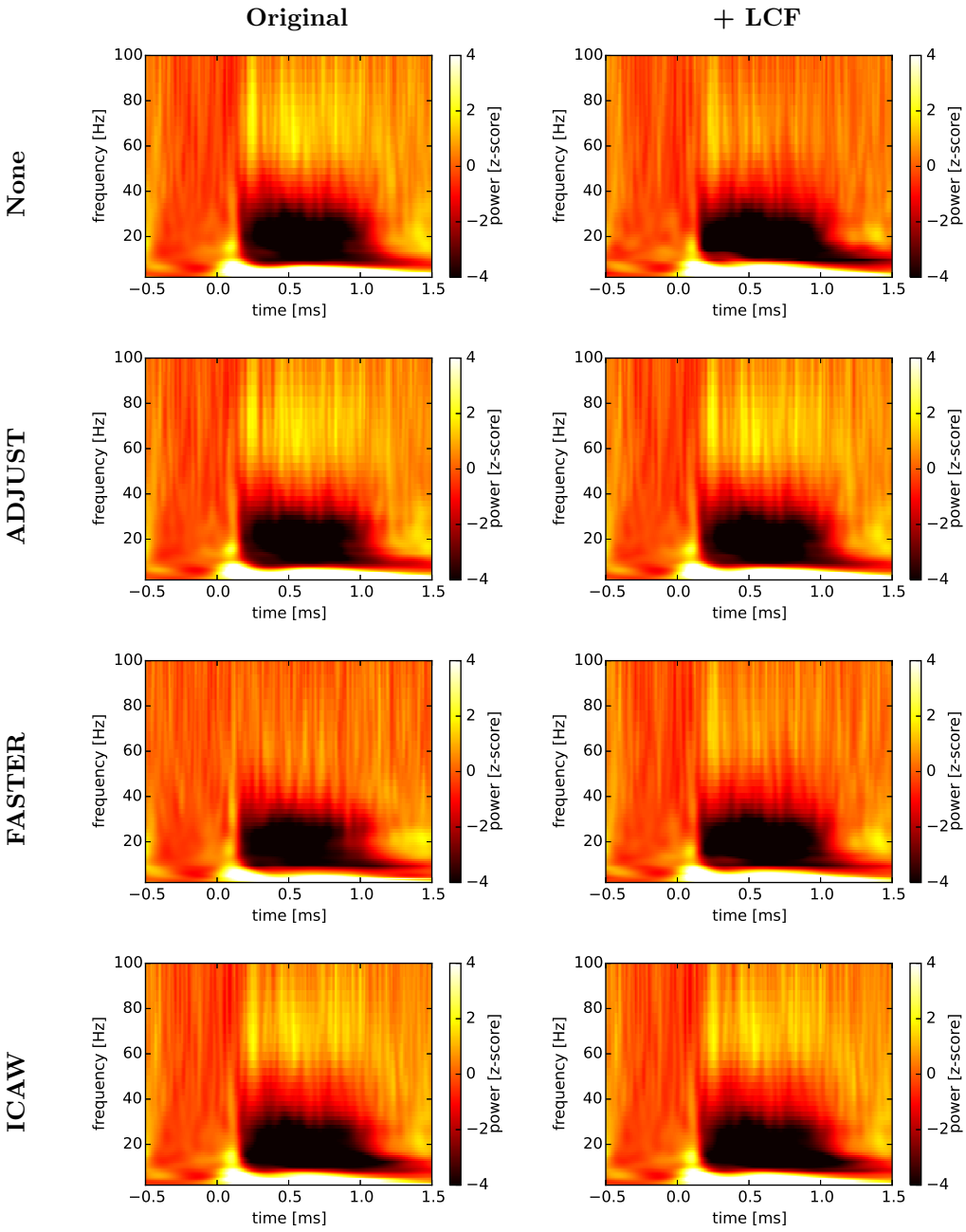


Figure 4.15: ERSs of the central channels cluster (RC) as defined in figure 4.8. Mean across participants, after applying each of the systems defined in the text without low-pass filter, without (Original) and with LCF (+ LCF). The top left image corresponds to the ERS of the original signal.

to retain neural activity allows to further increase the restrictiveness of the system's classification component, without heavily penalizing on signal loss. This is specially interesting in cases such as the one observed with ADJUST, where the system performs over-conservatively.

Furthermore, the implementation of LCF used here as a prove of concept is a very simplistic one, based on the voltage amplitude and its speed of change. Hence, better results are to be expected with more sophisticated artifact localization features. These in turn will allow to increase the restrictiveness of the artifact rejection systems even more.

To conclude, it is important to stress that the benefits of LCF are inversely proportional to the quality of the BSS algorithm. If the algorithm is able to perfectly dissociate neural and noise activity, the inclusion of an LCF block is then unnecessary. However, until such a BSS technique is available, LCF will always serve as an extra tool to ensure that the interesting signal is retained.

Chapter 5

EEG time-frequency exploration

In chapter 2, we identified the lack of a comprehensive study of the spectral discriminant information in biometric EEG identification research. All the biometric studies published to date focused their efforts mainly on improving the accuracy rates through the application of new algorithms. Meanwhile, the exact nature and properties of the processed information has yet to be fully described from a biometric point of view. In this chapter, we attempt to fill this gap with a series of carefully designed experiments. In particular:

1. we provide a set of recommendations to maximize the quality of the discriminant spectral information;

2. we present visual evidence of the existence of such information and associated properties by means of a stacked representation of the PSD;

3. we use 6 publicly available databases – ultimately divided into 10 data sets–, which allow us to distinguish general properties of the EEG signature from idiosyncrasies in individual data sets;

4. we use, for the first time, an AEP database for EEG biometric subject identification; and

5. ***we provide evidence of the existence of a task-independent neural signature***.

Contrary to the state of the art, the study presented in this chapter at no point aimed to obtain high accuracy rates during classification. Instead, it intended to describe in some detail the subject discriminant information within the time-frequency representation of EEG activity, and to characterize the effects of each of the parameters in the classification problem.

5.1 Methodology

The experimentation process was divided into four blocks, each composed of several tests focussed on a common goal (fig. 5.1 and tables 5.1 and 5.2). Configuration details missing in this section will be given within the corresponding results section, as their rationale depends on the results of previously evaluated steps. Complementary qualitative and quantitative studies were carried out during each experiment. In both cases, a Short Time Fourier Transform (STFT) spectrogram, computed with a Hamming window, was used to extract the time-frequency representation of the EEG signal. All the experiments were ran on the preprocessed databases listed in table 3.2. From each data set, except Keirn's and Yeom's, 3 versions were evaluated:

- Raw-EEG: The preprocessed data sets as defined in chapter 3.
- ADJUST-EEG: A cleaned version of the databases applying ADJUST artefact rejection to the raw version (fig. 3.1).
- LCF-EEG: A cleaned version of the databases applying a combination of ADJUST, FASTER and LCF artefact rejection to the raw version. This combines the detection of artefacts from ADJUST and FASTER and uses LCF to compute the cleaned components (fig. 3.1 and chapter 4).

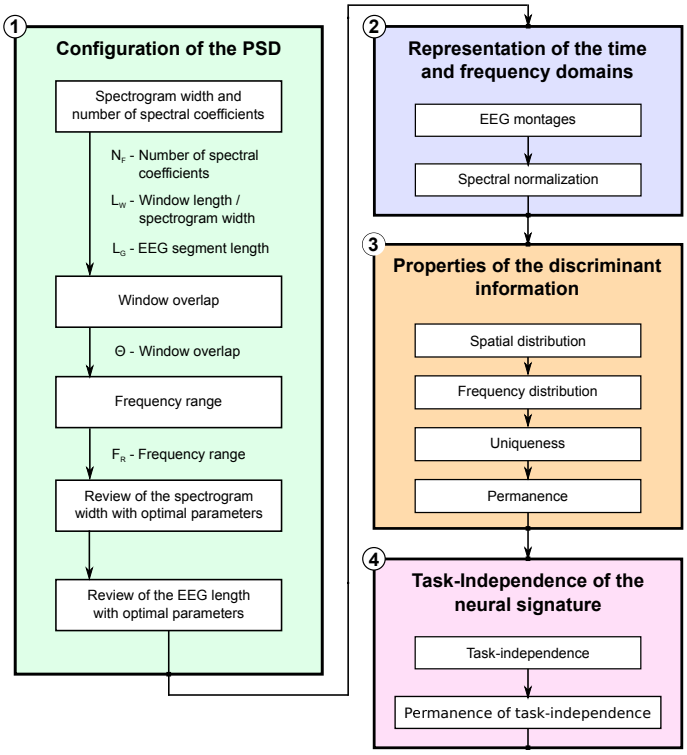


Figure 5.1: Diagram of the experimentation methodology. (1) We identified the configuration of the PSD that maximizes the discriminant information. (2) We find the representation of the time and frequency domains that maximizes the discriminant information. (3) We characterized the properties of the EEG discriminant information. (4) We characterize the task-independence property of the neural signature. This figure complements the information presented in tables 5.1 and 5.2.

5.1.1 A note on the application of artefact rejection methods

When interpreting the results, it is important to keep in mind that those obtained with the cleaned data sets may be subject to optimistic bias. It is possible that, when there are errors present during artefact rejection, the EEG data is modified in unexpected ways other than noise removal (e.g. neural leakage). As we applied the artefact rejection techniques within each subject, these alterations would be specific to that subject, therefore biasing the results.

The issue becomes more complicated when one assumes that the general ac-

Table 5.1: Phases 1 and 2 of the experimentation methodology. This table complements the information presented in figure 5.1.

1. Configuration of the PSD	
Spectrogram width and number of spectral coefficients	Analysis of number of spectral coefficients (N_F), spectrogram window length (L_W) and EEG length (L_G)
Window overlap	Analysis of spectrogram window overlap (Θ)
Frequency range	Analysis of maximum (F_{max}) and minimum (F_{min}) cut-off frequencies
Review of spectrogram width with optimal parameters	Analysis of spectrogram window length (L_W), with specific number of coefficients (N_F), window overlap (Θ) and EEG length (L_G)
Review of EEG length with optimal parameters	Analysis of the EEG length (L_G), with specific number of coefficients (N_F), spectrogram width (L_W) and window overlap (Θ)
2. Representation of the time and frequency domains	
EEG montages	Analysis of multiple EEG montages
Spectral normalization	Analysis of different normalization methods

curacy of the system will improve with the removal of artefacts. We will see that this improvement indeed occurs. However, it is difficult (if at all possible) to asses the extend to which performance differences between raw and cleaned data sets are due to the removal of noise or due to erroneous alterations of the EEG.

Given the analysis introduced in chapter 4, we expected the application of LCF to reduce this bias. However, it remains imperative to bear in mind the above consideration.

5.1.2 Configuration of the PSD

During the first experimentation stage, we defined the configuration of the STFT that maximizes the discriminant information. Specifically, we explored the number of Fast Fourier Transform (FFT) coefficients (N_F), the window length or spectral

Table 5.2: Phases 3 and 4 of the experimentation methodology. This table complements the information presented in figure 5.1.

3. Properties of the discriminant information	
Spatial distribution	Analysis of the spatial distribution of the EEG discriminant information
Frequency distribution	Analysis of the frequency distribution of the EEG discriminant information
Uniqueness	Analysis of the EEG discrimination power with increasing number of subjects
Permanence	Analysis of time stability of the EEG discrimination information
4. Task-independence of the neural signature	
Task-independence	Analysis of the task-independence of the EEG discriminant information
Permanence of task-independence	Analysis of time stability of the task-independence property of the EEG discriminant information

width (L_W), the window overlap percentage (Θ), the EEG segment length (L_G) and the frequency range (F_R). With this aim, we ran the following experiments:

1. *Spectrogram width and number of spectral coefficients*: Hamming window lengths $L_W \in [0.1, 60]$ seconds and the number of FFT spectral coefficients $N_F \in [32, 1024]$ were evaluated on a grid-like set of experiments where the window overlap Θ was fixed to 0.

2. *Window overlap*: Overlap percentage values $\Theta = 0, 25, 50$ and 75 were tested on the best performing configurations (L_W and N_F) of the previous step.

3. *Frequency range*: The width of the spectral band used was modified by varying the maximum F_{max} and minimum F_{min} frequencies between 10 and 60 Hz on individual experiments.

4. *Review of the spectrogram width with optimal parameters*: To clarify the effects of the STFT window length on the system's performance, optimal con-

figurations were represented against L_W. That is, $L_W \in [0.1, 60]$ was tested with the remaining parameters (N_F, L_G and Θ) set based on conclusions from previous experiments.

5. *Review of the EEG length with optimal parameters*: To clarify the effects of the EEG length on the system's performance, optimal configurations were represented against L_G. That is, $L_G \in [L_W, L]$, with L the available length of signal, was tested with the remaining parameters (N_F, L_W and Θ) set based on conclusions from previous experiments.

5.1.3 Representation of the time and frequency domains

Once the STFT was fully configured, we evaluated different time and frequency representations of the data. In particular, we ran the following experiments:

1. *Montages*: Transformations of the signal in the time domain were assessed by EEG montages Common Global Average Reference Montage (AvgMnt), Bipolar Inter-Hemispheric Reference Montage (BIHMnt) and Common Cz Reference Montage (CzMnt).

2. *Spectral normalization*: In a bid to find the optimal representation of the spectral information for the problem, the normalization functions described in table 5.3 were applied to the PSD prior to the classification.

5.1.4 Properties of the discriminant information

This experimentation block ran a number of tests to describe some of the properties of the discriminant PSD information. In particular:

1. *Spatial distribution*: Each EEG channel was evaluated individually to assess the performance of neural activity from different sensor locations (which vary in sensitivity to activity from different brain areas).

Table 5.3: PSD normalization functions. P is a PSD matrix with dimensions N_FxN_T, with N_T the number of time points, and $p^{n\%}_{(F)}$ is the n-% percentile applied along the frequency dimension.

Name	Equation	Name	Equation
powNorm	$\frac{P}{\sum_{\forall f} P}$	*prcNorm*	$\frac{P - p^{5\%}_{(F)}(P)}{p^{95\%}_{(F)}(P) - p^{5\%}_{(F)}(P)}$
normNorm	$\frac{P}{\sqrt{\sum_{\forall f} P^2}}$	*iqrNorm*	$\frac{P - p^{25\%}_{(F)}(P)}{p^{75\%}_{(F)}(P) - p^{25\%}_{(F)}(P)}$
zNorm	$\frac{P - \langle P \rangle_{(F)}}{std_{(F)}(P)}$	*rNorm*	$\frac{P - median_{(F)}(P)}{iqr_{(F)}(P)}$

2. *Frequency distribution*: Each frequency was evaluated individually to asses the performance of different EEG rhythms.

3. *Uniqueness of patterns*: To evaluate the uniqueness of individual spectral patterns, experiments varying the number of users N_S in the system were run, so that the stability of the performance with increased N_S could be assessed.

4. *Permanence of patterns*: This step aimed to judge the independence of the subject characteristic patterns with respect to time. To do so, Keirn's and Yeom's data sets were cross-validated considering each recording session as an indivisible unit. That is, for each subject, training and testing sets contained samples from different sessions. For simplicity, this CV methodology will be referred to as *session-CV*.

5.1.5 Task-independence of the neural signature

The neurophysiological, genetic and biometric avenues of research introduced in section 2 have been carried out under numerous cognitive tasks, including relaxed states and stimulus-elicited activity or ERPs, but always considering these tasks individually. The only multiple-task studies come from the biometric side [48,49,93], but even these have focussed on finding the design or cognitive task that maximizes the system's performance. This should come as no surprise as, historically, functional brain research has primarily studied differences across task-related or condition-related activity.

Here we present a completely different approach, as we hypothesize the existence of a neural pattern homogeneous across tasks and concomitant to the subject's identity. To test this idea, we ran the following two experiments in the final block:

1. *Task-CV*: The equivalent of experiment 4 in section 5.1.4 was applied to test independence with respect to tasks. Hence, the system was trained and tested with disjointed sets of cognitive tasks. Crucially, we executed this segmentation individually for each subject. As a result, in a given iteration, a task may have been used to train some subjects and to test others. For this experiment, K (from K-Folds) was set equal to the number of tasks in the database, truncated to five when exceeded. .

2. *Task-Sess-CV*: In the final step, Keirn's and Yeom's data sets were tested with a combined *Task-CV* and *Sess-CV*. The system was trained with a set of cognitive tasks from a single session and tested with the remaining tasks of the other session. The segmentation was executed individually for each subject, and therefore we have a similar situation to the previous experiment (*Task-CV*) regarding the distribution of tasks between subjects. For this experiment, K (from K-Folds) was set equal to 2.

5.1.6 Qualitative study

To simplify inspection of the EEG PSD, the spectrograms from selected trials, sessions, conditions and channels were stacked along their time axis, resulting in a single, piecewise-continuous (in time) spectrogram (fig. 5.2). This representation of the data helped to visually confirm the existence of different spectral neural signatures across individuals. It also helped to understand the effects that the evaluated parameters have on the signature. Hence, we performed these qualitative analyses on all experiments, to complement quantitative tests.

5.1.7 Quantitative study

To quantify the observations made during visual inspection of the spectrogram, we performed a series of classification experiments. Because the aim of such experiments was not to obtain high accuracy rates, we applied a simple design based on a Bayes classifier.

To this end, we z-normalized PSD coefficients across samples and used them as input to the classifier. The normalizing factors were computed only with the training set, and uninteresting filtered frequencies were removed from the analysis.

Different final responses of the Bayes classifier were computed from the sum-score fusion of N_G neighbouring windows, with $N_G = 1, 2, ..., N_W$, where N_W is the number of windows extracted from a single trial. Hence, for each N_G value, a total of $N_W - N_G + 1$ responses where computed by shifting the fusing scope a single PSD window to the right each time. This process was done to differentiate between the effects of the length of the Hamming window L_W and the length of the EEG segment

$$L_G = L_W[N_G - (1 - N_G)\frac{\Theta}{100}], \tag{5.1}$$

with Θ the window overlap percentage, used to compute a response. When representing results against L_G, intermediate values (not multiples of L_W) were interpolated whenever possible for ease of interpretation.

Table 5.4: Fusion modes used during experimentation.

Mode	Description
No-Fusion	The system is evaluated using a **single PSD coefficient**, individually for **each EEG channel**.
Freq-Fusion	The system is evaluated using **all PSD coefficients** simultaneously (from the selected frequency range), but individually for **each EEG channel**.
Ch-Fusion	The system is evaluated using a **single PSD coefficient**, but combining **all EEG channels**.
Full-Fusion	The system is evaluated using **all PSD coefficients** (from the selected frequency range) from **all EEG channels** simultaneously.

Four fusion modes were differentiated based on the level of feature fusion (table 5.4):

- *Full-Fusion*: Coefficients from all frequencies and channels are fused into a single vector.
- *Freq-Fusion*: Coefficients from all frequencies are fused into a single vector and each channel is evaluated individually.
- *Ch-Fusion*: Coefficients from all channels are fused into a single vector, and each frequency is evaluated individually.
- *No-Fusion*: The coefficient from each channel and frequency is evaluated individually.

For CV, we combined stratified K-Folds and Monte Carlo (MC) techniques in order to benefit from the stability (lack of bias) of the former and the low-variance of the latter [152]. To this end, we repeated 10 times a K-Fold process with 5 folds.

For databases where the number of available subjects exceeded the number used for experimentation (N_S), we used a different subset of subjects on each of the MC repetitions. Such subsets were built in a balanced way, i.e. by trying

to use each subject the same number of times across the whole experimentation process. Unless otherwise specified, we used a maximum of $N_S = 20$ subjects as a compromise between the number of subjects and the number of available databases able to accommodate N_S. The latter was important to avoid, as much as possible, unexpected interactions between the tested factor and N_S.

In addition, to ensure the validity of the results, during the segmentation of the databases all windows extracted from the same trial were kept in the same fold. This prevented the inclusion of EEG segments on both training and testing sets when the window overlap parameter was greater than 0%.

Micro-accuracies were computed for each of the MC repetitions as follows. Let $M_{(i,j)}$ be the N_SxN_S confusion matrix of the j-th K-Fold iteration of the i-th MC repetition, so that $M_i = \sum_{\forall j} M_{(i,j)}$ is the aggregated confusion matrix for the i-th MC repetition. Let A_i be the corresponding mean accuracy rate. We converted accuracy rates to Percentage Reduction of Error (PRE) values for the evaluation of the results [153]. Specifically, we compared A_i with a random process, such that

$$PRE_i = \frac{A_i - \frac{1}{N_S}}{1 - \frac{1}{N_S}}. \tag{5.2}$$

The mean (μ_{PRE}) and 95% Confidence Interval (CI) of the above PRE values were then computed and reported.

We use PREs instead of the more common absolute accuracy values for practical reasons. First, we aimed to shed light on the properties of the discriminant information within the EEG data, relegating absolute performance to a secondary role (as long as it is above chance level). This is especially true given that results were analysed by comparison rather than inspecting isolated values. Second, PREs as measured here will be 0 if the system performs at chance levels and > 0 otherwise, regardless of the number of subjects included in the analysis. In other words, the range of values of the results have a common meaning across databases. This allows for a better representation and more direct interpretation and comparison of the results. Having said that, if required, performance values can be

easily extracted from the given PREs by inverting equation 5.2.

Finally, we differentiated between the following four experimental modes based on the level of focus (table 5.5):

- *1Task-1Sess*: Experiments were run individually for each combination of session and task.
- *1Task-AllSess*: Experiments were run individually for each task, ignoring session labels during the CV partitioning. Mixing sessions on training and testing sets is not the standard procedure from a biometric point of view. However, in the current scenario, there are random factors linked with the set-up process, such as the exact position of sensors or their contact quality, which affect the recorded EEG. One way to alleviate the influence of such factors is to average across sessions. Note that we also executed experiments cross-validating sessions to keep the study relevant within the biometric field (section 5.1.4).
- *AllTask-1Sess*: Experiments were run individually for each session, ignoring task labels during the CV partitioning. As in the previous case, we also executed experiments cross-validating tasks (section 5.1.5).
- *AllTask-AllSess*: Experiments were run over all the available data, ignoring task and session labels during the CV partitioning.

5.2 Configuration of the PSD: results, discussion and conclusions

Within the first study block, we analysed the effects of some basic parameters on the quality of the extracted EEG discriminant information. Specifically, we considered the number of FFT coefficients (N_F), the length of the spectral window (L_W) and its overlap (Θ), and the length of the EEG signal used to compute the

Table 5.5: Focus modes used during experimentation.

Mode	Description
1Task-1Sess	The system is evaluated using data from a **single cognitive task/condition** and a **single recorded session**.
1Task-AllSess	The system is evaluated using data from a **single cognitive task/condition** but **all recorded sessions**.
AllTask-1Sess	The system is evaluated using data from **all cognitive tasks/conditions** but a **single recorded session**.
AllTask-AllSess	The system is evaluated using data from **all cognitive tasks/conditions** and **all recorded sessions**.

final identification response (L_G). We executed *full-fusion* experiments with all four focus modes (see tables 5.4 and 5.5).

5.2.1 Spectrogram width and number of FFT coefficients

First, we ran a number of experiments varying the number of FFT coefficients (N_F), the length of the spectral window (L_W) and the length of the EEG signal used to compute the final identification response (L_G) (Θ was set to 0%). As we suspected a triple interaction between these parameters, we executed all possible combinations in a cube like methodology with the following ranges $N_F = [16, 2048]$, $L_W = [0.1, max(20, L_A)]$ and $L_G = [L_W, L_A]$; with L_A the length of the available signal.

Results

Overall, a common behaviour was observed across databases. Visually, the PSD became more stable with longer windows (above 1 second). At the same time, some EEG traits and details are not apparent until N_F is higher than 64 (fig. 5.2). In quantitative terms, representing the μ_{PRE} surface of L_W against N_F corroborated the qualitative observations (fig. 5.3 and A.1). Overall, the system reached quasi-optimal performance at the diagonal $N_F \approx L_W * F_s$, with F_s the

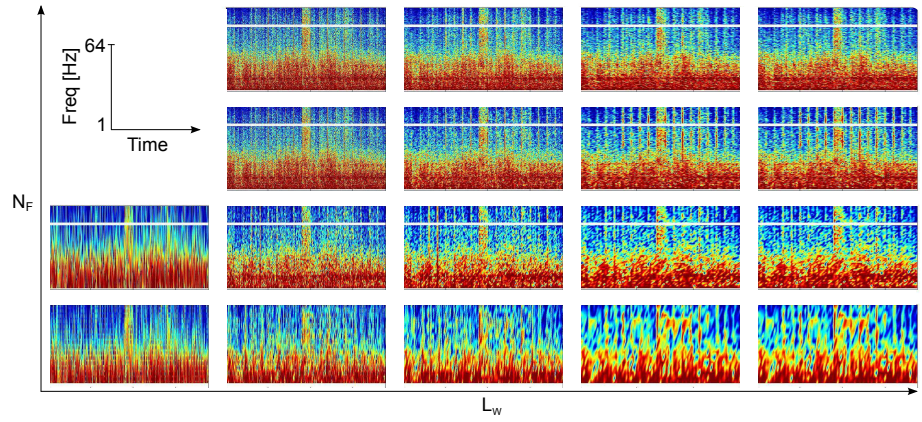


Figure 5.2: PSD with different STFT window length (L_W) and number of FFT coefficients (N_F). The PSD, in dB scale, corresponds to one of the subjects of BCI2000-Baseline database. Within each PSD, the missing frequency band around 50 Hz (vertical axis) was filtered in the pre-processing and manually removed from the analysis. The horizontal axis (time) is piecewise-continuous and contains all the samples from the subject.

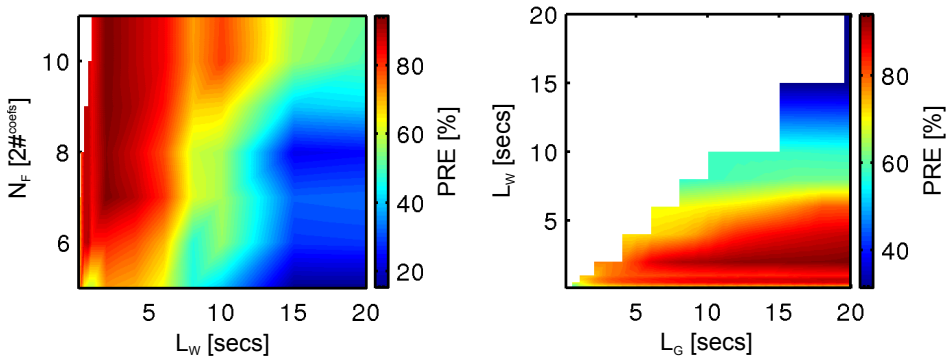


Figure 5.3: Quantitative analysis of the number of spectral coefficients (N_F), the SFTF window length (L_W) and the length of the EEG signal (L_G). Mean PRE results obtained with DEAP-Baseline database on an L_W vs N_F grid with $L_G = 10$ seconds (left) and an L_G vs L_W grid with $N_F = 128$ coefficients (right). In both cases, Θ was fixed to 0% and the results were obtained during the *full-fusion AllTask-AllSess* experiments. A maximum of 20 subjects was used in each experimental iteration.

sampling frequency (table A.1).

Although less prominent and less homogeneous across databases, an increase in L_G was generally followed by a gain in μ_{PRE} values (fig. 5.3 and A.2). In some instances, such as DEAP and Keirn's databases, this leap was as high as 10

percentage points. This could be expected, as increasing L_G means more information being fed to the system to produce the final response. Overall, we found the performance peak at $L_W < L_G$, i.e. when the available signal was divided into individually processed segments instead of in a unique long chunk. The only exceptions to this were DEAP-Playback results of only *AllTask-AllSess* experiments, and Yeom's and Zhang's databases. In the latter pair, the effect might still be there, but we are unable to see it due to the limited EEG available (only 1 second).

In terms of the stability of the results, databases with fewer than 20 subjects (i.e., all subjects were used in each iteration of MC) showed, as expected, lower levels of θ_{PRE}. The interaction between the std and the parameters considered here were not that clear. In general, minimum values of θ_{PRE} were located within the optimal area, i.e. above the N_F-L_W diagonal, while L_G had little or no effect in this regard.

When comparing results from each of the experimentation modes, the use of multiple tasks or sessions had no effect other than the expected averaging of μ_{PRE} and the increase of θ_{PRE}. The application of ADJUST for artefact rejection increased the amount of discriminant information, which was reflected in both mean and std PRE values. Absolute performance dropped back to the levels of raw EEG with the LCF artefact rejection method. In both cases, the effects of N_F, L_W and L_G remained the same.

Discussion

In terms of the parameters N_F and L_W, the optimal diagonal $N_F \approx L_W * F_s$ suggests that it is desirable to retain the maximum frequency resolution. This conflicts with architectures such as [73,77], which focused on the power of relatively wide spectral bands. At the same time, the relationship between performance and frequency resolution is in line with the findings of [79] regarding the need for higher orders of the AR model; and hence better frequency resolutions, to counteract the

rise in the number of users. According to this, wide spectral bands should only be used on neurophysiological studies – where brain rhythms are well described – or as supplementary features into a more complex feature vector.

With no obvious relationship between L_G and the system's performance, a "maximum" approach could be recommended, i.e. use all available EEG data. Perhaps, the most interesting effect of L_G is its ability to dilate the *"optimal configuration area"*, i.e. the set of parameters that yield results similar to the observed best performance. It is this property which allows to reduce L_W and subsequently N_F, maintaining the level of discrimination. A more detailed study of L_W and L_G will be presented in sections 5.2.4 and 5.2.5.

5.2.2 Window overlap

Subsequently, the overlap percentage parameter Θ was optimized by testing configurations of Θ, L_W and N_F, which accounted for the interaction between these three parameters. In this case, based on the previous experiment, the range of N_F was shrunk to $[64, 256]$ in a bid to reduce the volume of experimentation.

Results

Results were less uniform across data sets and system's configurations than in the previous experiment. In general, higher values of overlap yielded similar or better results (fig. 5.4 and A.3). The most obvious effect of Θ was its ability to widen the optimal configuration area (as described for L_G). The effects of Θ remained relatively stable across focus modes (*1Task-1Sess*, *1Task-AllSess*, *AllTask-1Sess* and *AllTask-AllSess*) and after the application of the artefact rejection methods. The most prominent exception was, once again, obtained with the DEAP-Playback data set during *AllTask-AllSess* experiments – the behaviour just explained was observed in *1Task-AllSess* experiments.

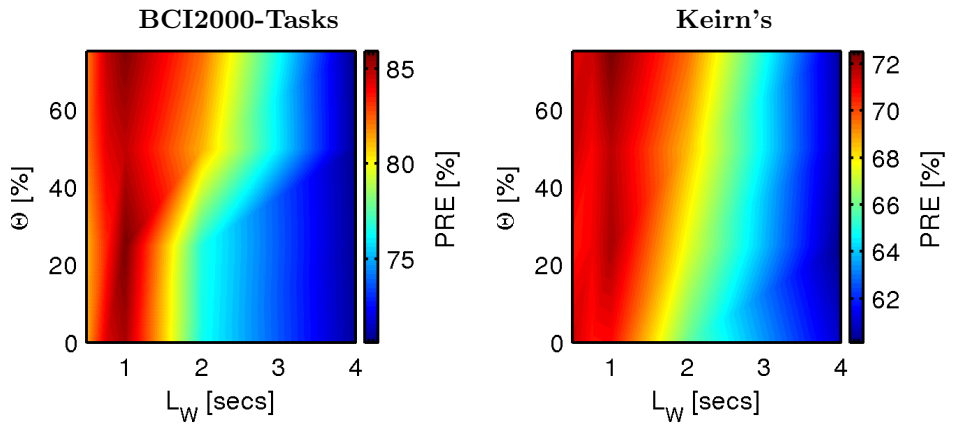


Figure 5.4: Quantitative analysis of the SFTF window overlap (Θ). Mean PRE results on a L_W vs Θ grid with $N_F = 128$ and L_G equal to the maximum EEG available. Results correspond to BCI2000-Tasks (left) and Keirn's (right) data sets, during *full-fusion AllTask-AllSess* experiments. A maximum of 20 subjects was used in each experimental iteration.

Discussion

A priori, it seems reasonable to use some degree of window overlap, as higher Θ translates into more information being fed to the system to produce the final response – similar to L_G. However, although results supported this to some degree, the observed effects are not strong enough to make a final statement in this regard. With this in mind, we decided to use an overlap of 75% in subsequent experiments. This is, however, a somehow arbitrary decision, as we have no strong evidence that this overlap value works better than any other. Therefore, lower overlaps could be used if, for example, processing time or data volume are a concern in the design of the system.

5.2.3 Frequency range

Visual inspection of EEG PSD suggests that spectral neural signatures stabilize for frequencies above 30-40 Hz (fig. 5.2). To explore this further, we systematically limited the maximum F_{max} and minimum F_{min} frequencies in separate experiments. Once more, to reduce the volume of experimentation, only two con-

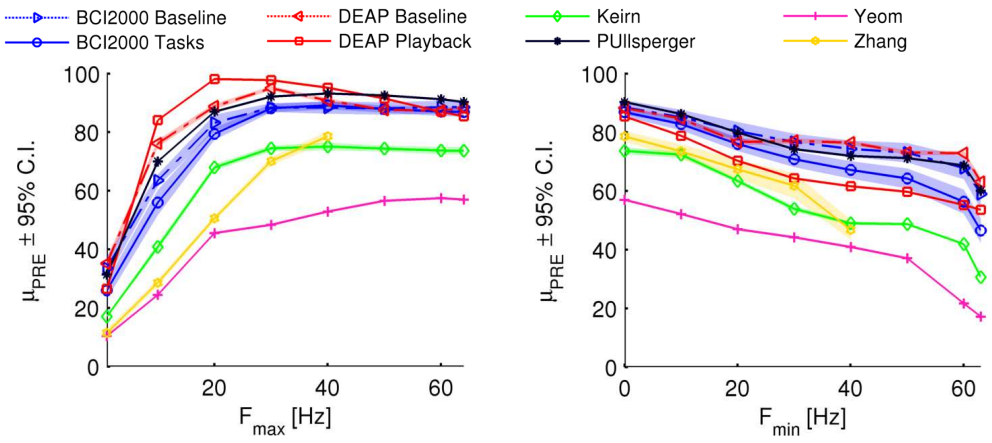


Figure 5.5: Quantitative analysis of the maximum (F_{max}) and minimum (F_{min}) cut-off frequencies. Mean and 95% CI of PRE results obtained in *full-fusion AllTask-AllSess* experiments with different F_{max} (left) and F_{min} (right) values, corresponding to *Conf-HalfLen* system. A maximum of 20 subjects was used in each experimental iteration.

figurations were tested, named *Conf-HalfLen* and *Conf-FullLen*. The latter uses L_W equal to the available EEG signal and hence $\Theta = 0$, while *Conf-HalfLen* utilizes L_W equal to half the available signal and $\Theta = 75$. In both cases, N_F was set to the power of two closest from the right to $L_W * F_s$.

Results

As expected from the observation of PSD representations, the system reached maximum performance at F_{max} equal 30 or 40 Hz in all databases with the exception of Yeom's data set, which reaches it at 50 Hz (fig. 5.5 left). On DEAP databases, there was a decrease of less than 10 percentage points in PRE when raising F_{max} from 30 to 40 Hz (tables A.2 and A.3). Experiments varying F_{min} resulted in a μ_{PRE} curve with two more pronounced increases, one when the first frequencies are added (between 50 and 60 Hz) and the second one with the last frequencies – frequencies below 20 or 30 Hz (fig. 5.5 right and tables A.2 and A.3). This behaviour was also observed after rejecting artefacts with ADJUST and LCF (fig. A.4 and tables A.4 to A.7).

Discussion

In line with [81], a high frequency cut-off of 30 or 40 Hz can be established based on the results. Indeed, most of the systems in the literature have used a maximum frequency below 40 Hz (chapter 2). The described results highlight that architectures using only high frequencies [37, 39] are not necessarily the optimal approach.

The outlier behaviour in Yeom's database may be a result of the classifier capitalizing on differences between users' artefacts – muscle artefacts have been identified to overlap with frequencies above 20 Hz [131]. A post-hoc examination of the spectra from Yeom's database corroborated the existence of noise at high frequencies.

Indeed, EEG artefacts have been shown to be subject discriminant and used successfully for identification [114]. Hence, the performance of systems using frequencies above 20 Hz cannot be considered to be solely based on the users' neural signature without a proper artefact analysis. In our case, the exact same behaviour was obtained with artefact free databases, as processed by ADJUST and LCF. In fact, techniques based on ICA have been described as the *"more promising approaches that have been used for attenuating muscle artefact* [131]. Hence, although it is not possible at the moment to be certain about the source of the performance above 20 Hz, it seems unlikely that it comes from muscle artefacts. Unfortunately, due to an insufficient number of EEG sensors, we lacked an artefact free version of Yeom's database (which we noted had high frequency artefacts) to further test this.

5.2.4 Review of spectrogram width with optimal parameters

Although previous experiments have shed light onto the optimal value of the spectral window length (L_W), they did not address the issue in a definitive way due

to the relationship between L_W and the other parameters. To circumvent this, we re-tested a range of window lengths, setting the remaining parameters to their defined optimal values. That is, N_F was set to the maximum power of 2 closest to $L_W * F_s$, L_G was set to the available signal length, Θ was set to 75% whenever possible (i.e. when L_W is small enough to allow multiple windows with 75% overlap on the available signal length) and 0% otherwise, and F_{max} was set to 40 Hz. Results obtained with Yeom's and Zhang's will be presented but not considered in the analysis, as they only contain 1 second of EEG.

Results

Across databases, maximum performance was reached at L_W between 1 and 2 seconds (fig. 5.6 and table A.8). After this point, performances either stabilised within the 95% CI or decreased. The main exception encountered was on DEAP-Playback data set, for which the maximum was reached between 6 and 8 seconds.

The rejection of artefacts reduced the differences between window lengths below 2 seconds, boosting the performance of the shortest windows. On DEAP-Playback data set, cleaning the EEG shifted the optimal point to earlier in L_W, reaching it between 4 and 6 seconds on ADJUST processed data, and between 2 and 4 seconds on LCF processed data.

Discussion

Similar behaviour was observed across databases. Overall, we obtained optimal performances with window lengths 1 or 2 seconds, with the exception of DEAP-Playback. In addition, the flattening observed on the performance curves for windows sorter than 2 seconds suggests that these configurations are more sensitive to noise.

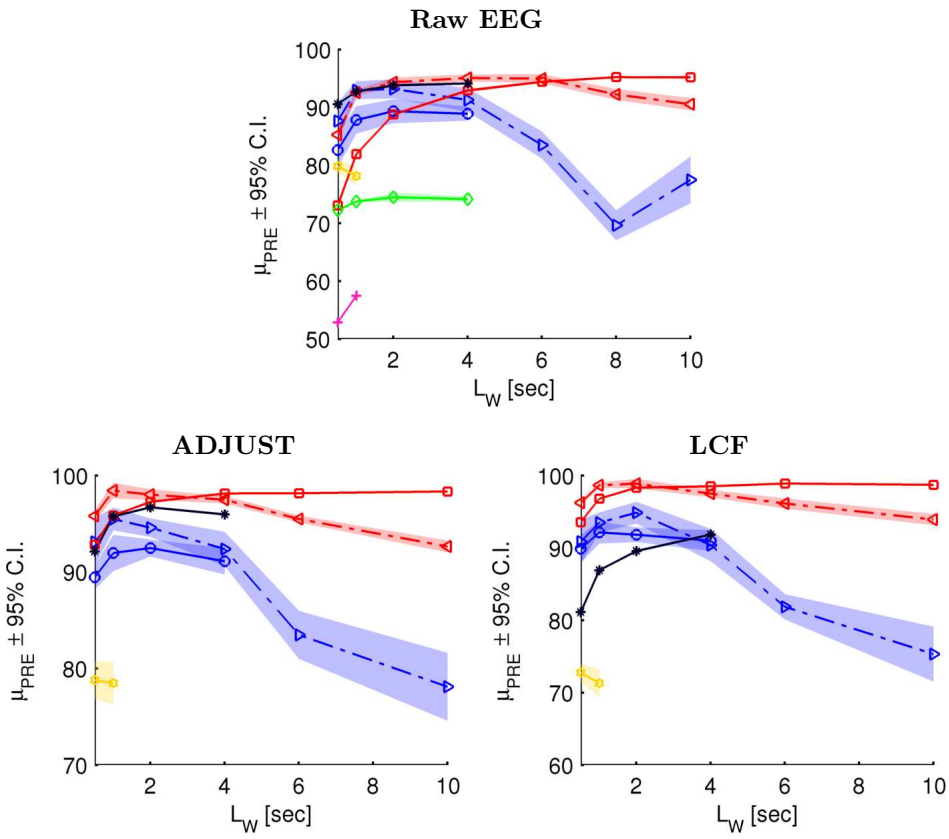


Figure 5.6: Quantitative analysis of the STFT window length (L_W). Mean PRE and 95% CI (shaded area) obtained with different L_W. The remaining parameters were set to their optimal values according to previous experiments. Data correspond to *full-fusion* set-up in *AllTask-AllSess* experiments with raw databases (top) and after the application of ADJUST (bottom-left) and LCF (bottom-right) processing. A maximum of 20 subjects was used in each experimental iteration. Refer to figure 5.5 for details on the legend.

5.2.5 Review of EEG segment length with optimal parameters

A review similar to the previous was run for the length of the EEG signal used to compute the final response (L_G). We defined here configurations *Conf-1s* and *Conf-2s*. These had $\Theta = 75\%$ and a L_W of 1 and 2 seconds respectively, with N_F and L_G set to their optimal values (section 5.2.7). Note that, since the duration of events in Yeom's and Zhang's databases was 1 second, only *Conf-HalfLen* and

Conf-FullLen were possible (a 75% overlap as used in *Conf-1s* in not achievable in this case). Additionally, in BCI2000-Tasks and Keirn's databases, configurations *Conf-HalfLen* and *Conf-2s* are equivalent.

Results

In this case, results are remarkably similar across databases, with PRE curves showing an asymptotic behaviour with increasing L_G (fig. 5.7 and A.5). On databases with EEG signals too sort to show the asymptote itself, the PRE curves still behave in a way similar to databases with longer signals, clearly suggesting the presence of such a limit. Overall, the maximum performance is obtained using all the available EEG.

To study the described behaviour, we fitted the rational model

$$PRE = \frac{a + bL_G}{c + dL_G}, \tag{5.3}$$

with *a*-*d* fitted factors, and used $L_W = 60$ seconds as the maximum PRE point. Such model can be seen to accommodate well the dynamics of the data. Exceptions were found on the curves of P. Ullsperger's database after rejection of artefacts, and on DEAP-Baseline with *Conf-HalfLen* system after the application of ADJUST (fig. A.5). For all data sets (raw and artefact free) and all systems, we found that [92.5%-95%] of the maximum performance (asymptote) was obtained at L_G between 4 and 6 seconds.

Discussion

An increase in performance with longer EEG recordings is to be expected, as this translates in more information being fed into the system. Once the system has enough information to overcome the distortion of noisy segments, the performance levels out. This asymptotic behaviour has been studied here by means of a 1st-degree/1st-degree rational model. For the sorter databases (less than 4 seconds), the amount of data available to fit the model was relatively small, and results

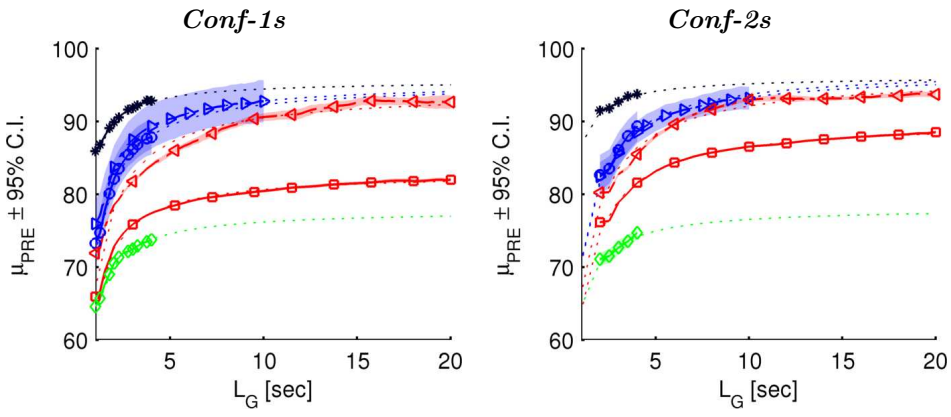


Figure 5.7: Quantitative analysis of the EEG segment length (L_G). Mean PRE and 95% CI (shaded area) obtained with different L_G. Data correspond to *full-fusion* set-up in *AllTask-AllSess* experiments using *Conf-1s* (left) and *Conf-2s* (right) systems. A maximum of 20 subjects was used in each experimental iteration. Refer to figure 5.5 for details on the legend.

should be considered accordingly. Having said that, the homogeneity across all (sort and long) data sets is obvious from the represented curves, which suggest that, although more experimentation is needed to make a final statement, the described behaviour may be a good approximation of the true interaction between performance and L_G.

5.2.6 A comparison across systems

To finalize the analysis of the results obtained within this experimentation block, we compare here the performance of the systems defined in sections 5.2.3 and 5.2.5, i.e. *Conf-HalfLen*, *Conf-FullLen*, *Conf-1s* and *Conf-2s* (tables 5.6, A.9 and A.10). Overall, the fragmentation of the EEG signal into shorter overlapping windows (between 1 and 2 seconds) outperformed the use of all the EEG signal at once. DEAP-Playback data set was the main exception, but only in *AllTask-AllSess* experiments. Raw and LCF processed P. Ullsperger's database also performed better with *Conf-FullLen*, but by less than 1 and 3 percentage points respectively. Yeom's and Zhang's databases should be considered aside, as they only allow

Table 5.6: Quantitative results for different configurations. Mean PRE and 95% CI obtained with different configurations. Data correspond to raw databases tested during *full-fusion AllTask-AllSess* experiments. Within data sets, performances statistically different than the maximum are pointed by * (single tail t-tests with BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names.

Conf	**BB**	**BT**	**DB**	**DP**
Conf-FullLen	68.79*	88.88	64.11*	**95.82**
	[64.52, 73.06]	[87.67, 90.10]	[62.58, 65.63]	**[95.61, 96.04]**
Conf-1s	92.74	87.80	92.63*	81.99*
	[89.85, 95.62]	[85.43, 90.17]	[91.69, 93.57]	[81.70, 82.27]
Conf-2s	**93.16**	**89.32**	**93.68**	88.51*
	[91.60, 94.72]	**[87.25, 91.40]**	**[93.33, 94.04]**	[88.26, 88.75]

Conf	**K**	**P**	**Y**	**Z**
Conf-HalfLen	-	-	52.89*	**78.82**
	-	-	[52.86, 52.92]	**[75.26, 82.37]**
Conf-FullLen	74.42	**94.22**	**57.47**	77.88
	[73.68, 75.17]	**[94.15, 94.30]**	**[57.45, 57.50]**	[76.05, 79.71]
Conf-1s	73.74*	92.83*	-	-
	[73.57, 73.92]	[92.71, 92.94]	-	-
Conf-2s	**74.70**	93.72*	-	-
	[73.81, 75.59]	[93.61, 93.84]	-	-

configurations *Conf-HalfLen* and *Conf-FullLen* due to the limited amount of EEG signal available (1 second).

Such behaviour may be explained by the way *Conf-1s* and *Conf-2s* process the data. By breaking the available data into segments, we isolate localized noise into individual segments, or more precisely, into a set of them as we allow 75% overlap between windows. At the same time, segments with good quality signal are also obtained. While noisy segments may yield random outputs from the system, clean segments will presumably result in accurate responses with high confidences (scores). When averaging outputs across segments, we expect the response from

clean segments to be selected.

5.2.7 Conclusions

Throughout these experiments, we have defined the optimal configuration of STFT to maximize subject-discriminant information. In particular, fragmentation of the EEG signal into shorter overlapping windows (between 1 and 2 seconds) has been identified as the best overall approach. The spectral domain of each segment should then be computed without compromising in frequency resolution ($N_F \approx L_W * F_s$), while also keeping in mind that discriminative information was found primarily under 40 Hz.

With regard to the amount of EEG needed, it is a matter of maximizing the information fed to the system. Having said that, from our analysis we conclude that $L_G = [4, 6]$ seconds results in performances between 92.5% and 95% of the maximum achievable by the data, although we can expect this to vary depending on the quality of the data.

5.3 Representation of the time and frequency domains: results, discussion and conclusions

Once the optimal configuration of N_F, L_W, Θ and L_G was found, we assessed the effects of different time and frequency representations. We compared the discriminant information of AvgMnt, BIHMnt and CzMnt montages, as well as the effects of several PSD normalization methods. In the following experiments, we executed the four focus modes (table 5.5) in *full-fusion* with the four defined systems (*Conf-HalfLen*, *Conf-FullLen*, *Conf-1s* and *Conf-2s*).

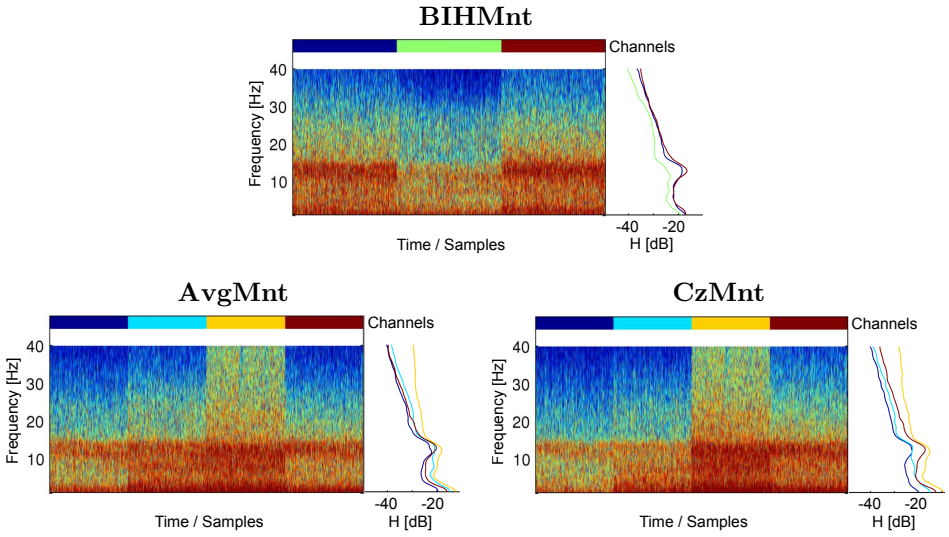


Figure 5.8: Qualitative analysis of EEG montages. PSD (in dB) of a subject from Yeom's database computed with BIHMnt (top), AvgMnt (bottom-left) and CzMnt (bottom-right). The colour bar-code above each PSD represents changes in channels, in the piecewise-continuous time axis. Each channel's spectrum (H) is added to the right of the spectrogram using the same colour scheme as the channel bar-code.

5.3.1 Montages

The above experiments have all been executed with AvgMnt reference. Here, we evaluate the effects of BIHMnt and CzMnt montages on the EEG discriminant information compared to AvgMnt.

Results

Qualitatively, differences across montages were quite subtle (fig. 5.8). Overall, AvgMnt seemed to have slightly more stable spectral patterns. Occasionally, BIHMnt showed a pronounced loss of information, probably due to highly correlated inter-hemispheric channels[1].

Quantitative results between and within databases were quite heterogeneous.

[1]In cases where all rejected sensors cover a common area, e.g left frontal lobe, the result of the interpolation of the sensors closer to the saggital midline may be extremely correlated with that of its hemispheric pair, resulting in an almost flat response of the BIHMnt.

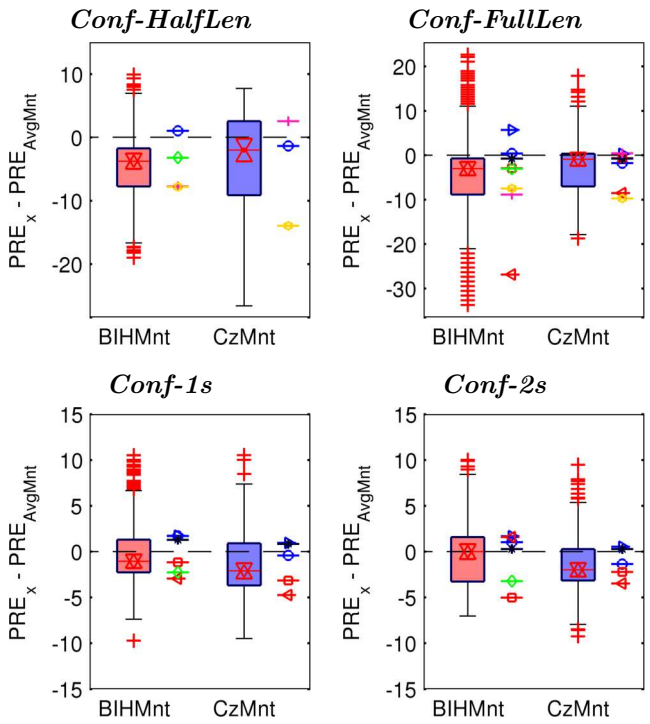


Figure 5.9: Quantitative analysis of EEG montages with raw data sets. Relative PRE values between *BIHMnt-AvgMnt* and between *CzMnt-AvgMnt*. Boxes show results stacked across databases. Box limits are 25 and 95 percentiles, while black bars shows maximum and minimum values after excluding outliers (red crosses). The red line within each box and triangle markers show median values and their 95% CI. To the right of each box, corresponding mean PRE values from each database are shown following the legend of figure 5.5. A maximum of 20 subjects was used in each experimental iteration.

In general, mean PRE values between montages were within 5 percentage points in *Conf-1s* and *Conf-2s* systems, and within 10 percentage points in *Conf-HalfLen* and *Conf-FullLen* (fig. 5.9 and tables A.11 to A.13). The application of artefact rejection had virtually no effect on the described overall relationship between montages, other than defining their differences by reducing the dispersion of the results (fig. A.6 and A.7, and tables A.11 to A.14)).

Discussion

The inhomogeneity of the results should come as no surprise. Optimal montages have been described before as highly dependent on the recording paradigm and level of noise [154]. Interestingly, BIHMnt, although reducing the volume of data by almost half, maintained a remarkably similar performance (within 5 or 10 percentage points), especially with *Conf-1s* and *Conf-2s* systems. This makes it an attractive option in scenarios where the volume of data is a concern. Across databases, AvgMnt is the safest option among the tested montages.

5.3.2 Spectral normalization

Next, we normalized the PSD of each window segment with the methods described in table 5.3. Following previous results, we again used AvgMnt.

Results

In this case, the effect of each normalization varied greatly across system's configurations. Overall, each had a homogenizing effect on the spectral pattern of the EEG, especially for frequencies above ∼20 Hz (fig. 5.10). This effect was particularly strong for *prcNorm*, *iqrNorm* and *rNorm* methods.

Quantitatively, results where again quite heterogeneous (fig. 5.11 and tables A.15 to A.18). While Keirn's database always benefited substantially from normalization – with 20 percentage points of μ_{PPRE} improvement on average –, the others showed great variability across systems. Overall, *normNorm*, *prcNorm* and *zNorm* were the worst performing normalizations. *rNorm* and, to a lesser extent, *iqrNorm* gave equal or better performance than the raw PSD in almost all cases, especially in *Conf-1s* and *Conf-2s* systems.

The rejection of artefacts by ADJUST (fig. A.8 and tables A.19 to A.22) and LCF (fig. A.9 and tables A.23 to A.26) had a great impact on the relationship between normalization methods and the raw PSD. In particular, the latter gained

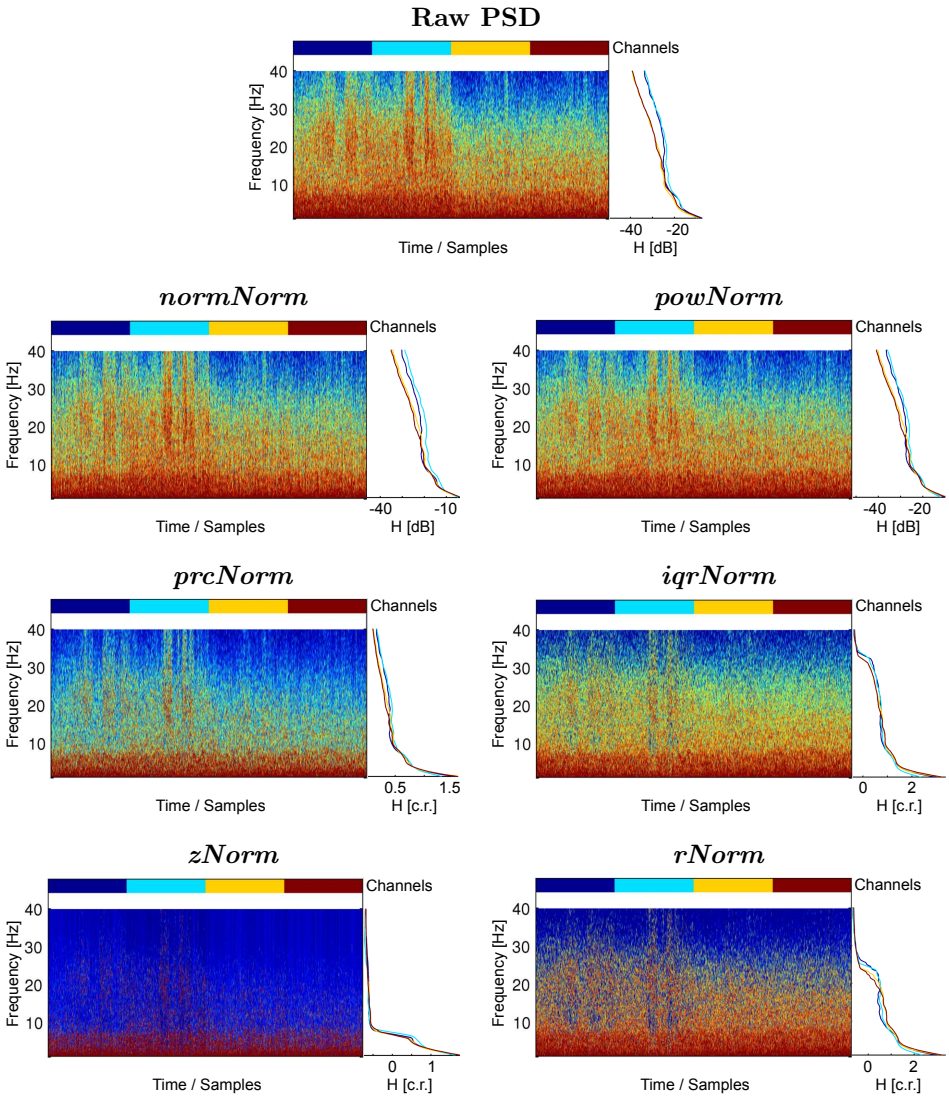


Figure 5.10: Qualitative analysis of PSD normalization. *Conf-2s* PSD, in dB, of a subject from P. Ullsperger's database. *prcNorm*, *iqrNorm*, *zNorm* and *rNorm* are cubic root scaled (c.r.) instead, as they contained negative values.

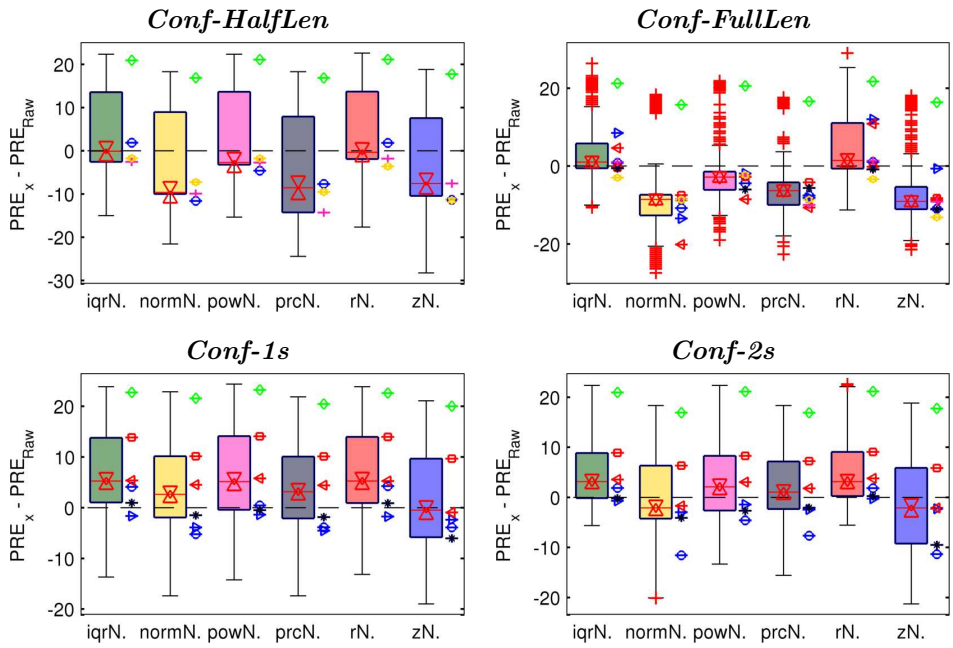


Figure 5.11: Quantitative analysis of PSD normalization. Relative PRE values between the PSD normalized by each of the methods in table 5.3 and the raw PSD. Results are stacked across databases. A maximum of 20 subjects was used in each experimental iteration. Refer to caption of figure 5.9 for details about the meaning of symbols within the image.

the most from the cleaning, which translated in normalization methods being less beneficial, or even disadvantageous.

Discussion

In line with [34, 39] results, normalization approaches based on magnitudes sensitive to outliers had a negative effect under some configurations. On the other hand, those based on more robust measurements, such as *rNorm* and *iqrNorm*, boosted the performance in almost all cases, especially with optimal configurations (*Conf-1s* and *Conf-2s*). This, together with the modulating behaviour of the artefact rejection methods, suggests that:

1. the discriminant information is coded within the relationship between spectral coefficients (i.e. the global spectral shape), rather than in the instant

absolute power of each frequency; and

2. the enhancement observed via robust normalizations is, at least in part, due to the counteraction of effects from artefacts.

5.3.3 Conclusions

In this experimentation block, we have assessed for the first time the effects of EEG montages (time representation) on the EEG subject discriminant information. Although AvgMnt turned out to be the best performing montage on average across system's configurations, the performance of BIHMnt was strikingly similar, especially when considering the highly reduced data volume of BIHMnt. In addition, normalizing the spectral coefficients using measurements robust against outliers reduced the effect of noise and boosted the quality of the data in almost all cases. Hence, it may be applied as an alternative to artefact rejection methods.

5.4 Properties of the discriminant information: results, discussion and conclusions

Next, we described some of the properties of the discriminant information identified in the previous experimentation blocks. Specifically, we studied the spatial and frequency distribution of the discriminant information, as well as its uniqueness across individuals and permanence along time.

We evaluated systems *Conf-HalfLen*, *Conf-FullLen*, *Conf-1s* and *Conf-2s* (sections 5.2.3 and 5.3) with the raw PSD coefficients and with their *rNorm* version.

5.4.1 Spatial distribution

The discriminant information of different sensor locations was evaluated through *AllTask-AllSess* and *AllTask-1Sess* experiments in *freq-fusion* architectures, i.e. running isolated experiments within sensors.

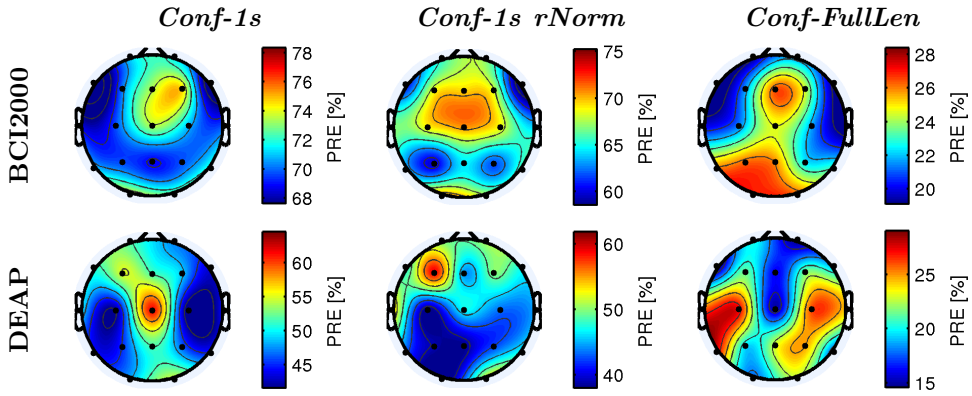


Figure 5.12: Quantitative analysis of the spatial distribution of the discriminant information. Mean PRE values obtained at each location with the REO condition of BCI2000 (top) and DEAP (bottom) data sets when applying the systems *Conf-1s* (left), *Conf-1s rNorm* (centre) and *Conf-FullLen* (right). Results correspond to *freq-fusion ch-focus* experiments, with a maximum of 20 subjects used in each experimental iteration. Refer to fig. A.10 for further related results.

Results

Looking at the PSD representations, although there were instances of patterns with large variability across sensors, there were also numerous cases where the pattern remained relatively equal across channels (fig. 5.8). Quantitative results showed no clear pattern in the spatial distribution of the performance across data sets, tasks or systems. In fact, this distribution varied even within conditions. For example, focusing on the REO condition of BCI2000 and DEAP databases, we observed a great variability across system configurations as well as between data sets (fig. 5.12 and A.10). The removal of artefacts brought the performance between channels closer by boosting those with the worst μ_{PRE} values.

Discussion

The lack of uniformity across and within databases, tasks and systems suggests that there is no obvious 'most-discriminative' region. Rather, the homogenizing effect of removing the artefacts hints that the performance depends more on the

strengths of the applied system and on the idiosyncrasies of each session's set-up, which affects the quality of the signal individually on each channel. This would explain the lack of consensus in the literature regarding the performance of sensor locations during subject identification. Looking at research with similar results, such as in [80–82] or [83–85, 121], they all relied on EEG signals recorded from a single task and from the same database[2].

On the other hand, one has to consider that the systems applied here rely on the general spectral shape of the EEG. Systems that focus on specific characteristics of the EEG during a particular task, such as the power of the alpha rhythm during REC or the P300 amplitude during VEP, can be expected to have a more defined spatial distribution of the discrimination power.

5.4.2 Frequency distribution

To study the effect of each frequency component, we ran *AllTask-AllSess* and *AllTask-1Sess* experiments with *ch-fusion* and *no-fusion* architectures.

Results

In general, quantitative experiments showed noisy μ_{PRE} curves idiosyncratic to each database (fig. 5.13). However, within this variability, we observed some common characteristics across data sets, the most distinctive being a peak within the alpha rhythm. After this peak, the PRE raised again passed 15 or 20 Hz, until it reached the global maximum – in some cases, curves of Keirn's database decreased after reaching this maximum. With *Conf-FullLen* systems, the PRE also increased to the left of the alpha peak (towards lower frequencies) and reached a local maximum at 1 Hz. This behaviour was also reproduced in *no-fusion* experiments within each channel.

[2]Because databases are not generally labelled, it is difficult to be certain whether the same database was used for different analyses. However, the descriptions of the databases suggest that they used the same database.

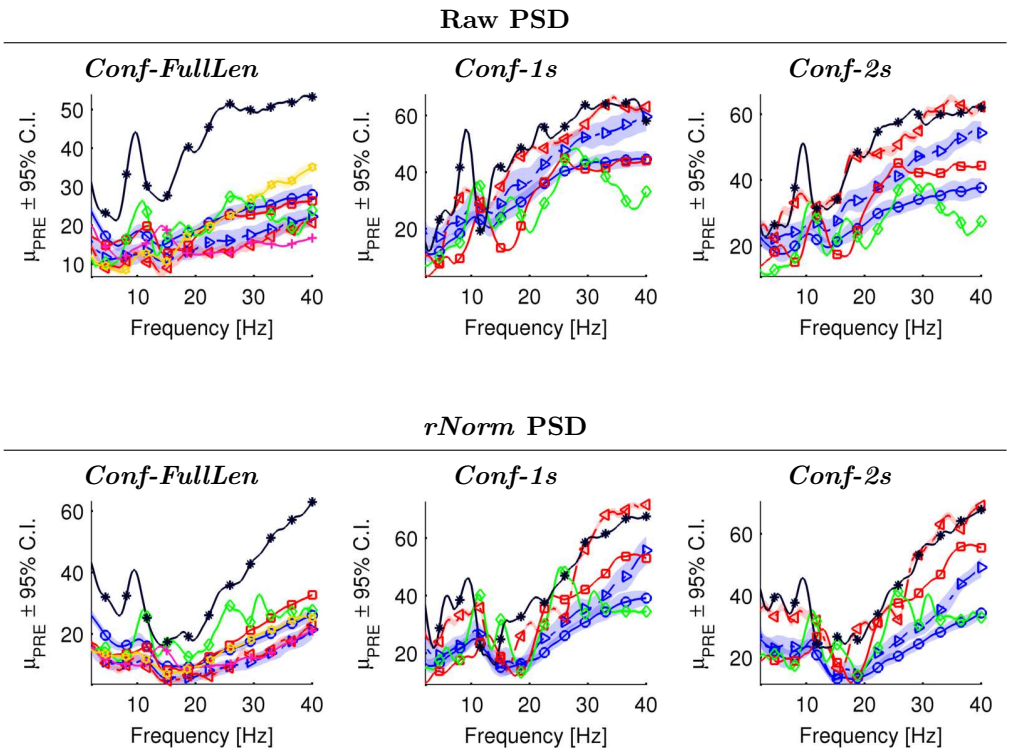


Figure 5.13: Quantitative analysis of the discrimination power frequency distribution. Mean PRE and 95% CI (shaded area) obtained with each frequency (*ch-fusion* experiments) of raw and rNorm PSD. Curves were smoothed by local regression, using weighted linear least squares and a first degree polynomial model with a 3 Hz span. A maximum of 20 subjects was used in each experimental iteration. Refer to figure 5.5 for details on the legend.

The rejection of artefacts had no major effects on the described behaviour (fig. A.11). The performance of frequencies above the alpha rhythm remained virtually the same. With *Conf-1s* and *Conf-2s* systems, the discrimination power of frequencies below the alpha peak raised. With ADJUST processed databases, the gain was as large as 30 percentage points. In some cases, this increase created another local peak between 1 and 8 Hz.

Discussion

The observed behaviour is in line with those presented in section 5.2.3. They also agree with other genetic [18, 21] and biometric [77, 82] studies with regard to the

amount of information within the delta and alpha rhythms.

In addition, our results suggest that frequencies corresponding to the beta rhythm and up to 40 Hz carry as much discriminant information as the delta and alpha rhythms. Furthermore, as reported in [87, 99], the high-beta and gamma band (up to 40 Hz) reached performance levels, on occasion, above those of lower bands.

5.4.3 Uniqueness

The current experimental step aims to asses the uniqueness of the neural signature. If this signature is comprised of a finite number of categorical characteristics of EEG activity, it will only be able to classify subjects into said categories. Hence, individual discrimination beyond these categories would not be possible. This will be reflected as a performance slump when the number of subjects in the system increases.

In order for the system to unequivocally classify among a large number of individuals, the neural signature must be a continuous characteristic. Thus, we tested the uniqueness of the neural signature by evaluating the systems with an increased number of subjects and analysing the dynamics of the obtained PRE curves.

Results

Differences across subjects in the shape of their EEG spectrum can be easily seen in the generated visual representations (fig. 5.14).

From quantitative experiments, systems *Conf-1s* and *Conf-2s* were the most robust against the rise in the number of subjects. In particular, their μ_{PRE} dropped less than 10% when increasing the number of users from 5 to 20 (fig. 5.15). With BCI2000 databases, their performance fell less than 20% between 5 and 100 subjects. The application of ADJUST and LCF artefact rejection attenuated such decay by as much as 6 percentage points (fig. A.12).

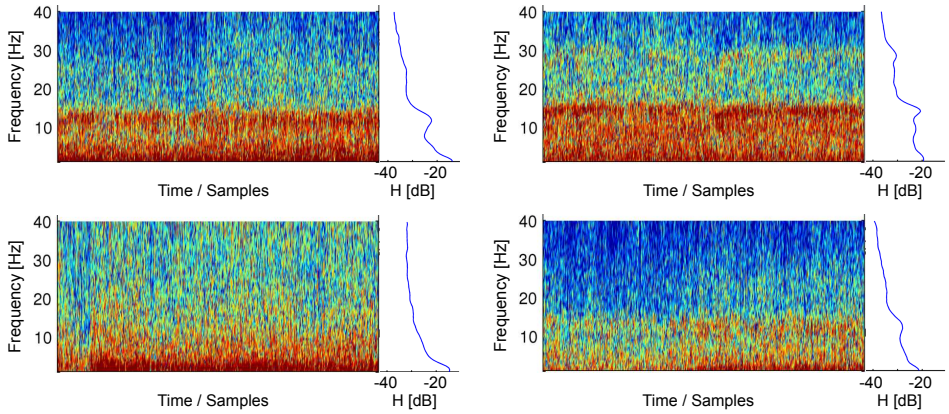


Figure 5.14: Qualitative analysis of uniqueness. PSD, in dB, of 4 subjects from Yeom's database, corresponding to the EEG sensor P3 and a single task (self-representation).

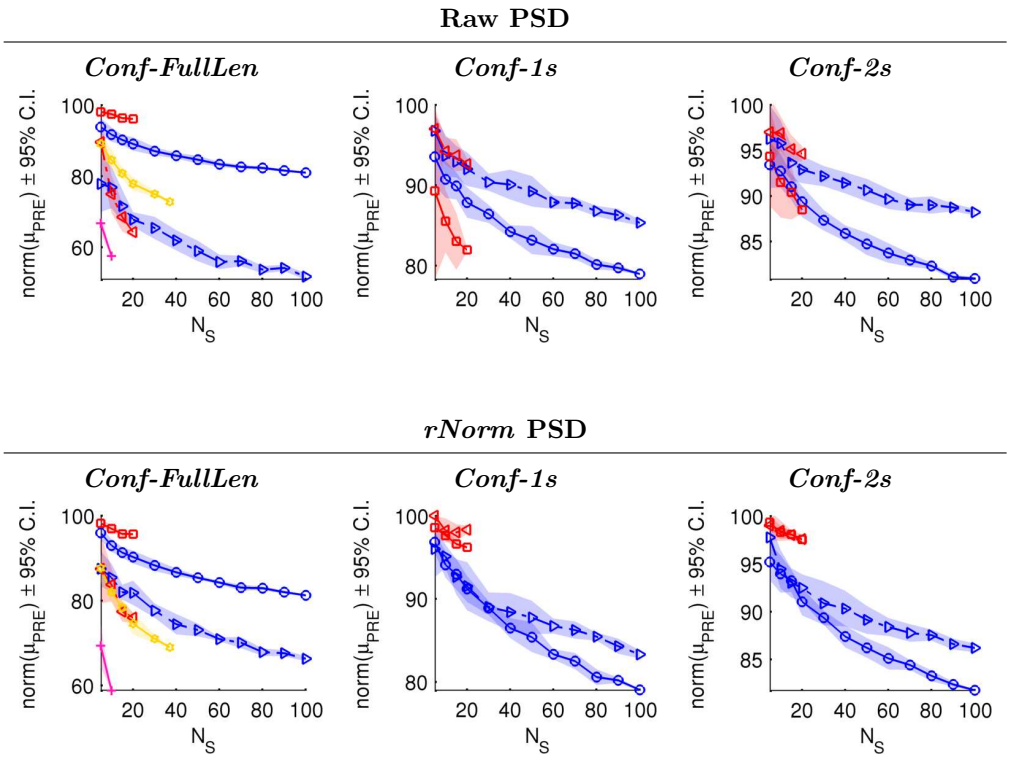


Figure 5.15: Quantitative analysis of uniqueness. Mean PRE and 95% CI (shaded area) for different numbers of subjects (N_S) in the system and each database. Results correspond to the *rNorm* system. Refer to figure 5.5 for details on the legend.

Discussion

Results show that EEG spectral patterns are distinctive enough to discriminate between 100+ subjects. A PRE of 80%, the lowest observed with *Conf-1s* and *Conf-2s* systems for 100 subjects, represents an accuracy rate of almost 80%. This is an encouraging result, especially considering that the systems used here are fairly simple ones. Overall, in line with [19], our results suggest that the neural signature is comprised of continuous features. In this regard, larger databases are necessary to find the true potential of the EEG discrimination power.

5.4.4 Permanence

In a bid to evaluate the stability of the EEG subject traits, we ran *session-CV* experiments. To allow a direct comparison of results, we also ran extra tests using normal CV (no *Sess-CV*) with a K (from k-folds) equal to the number of sessions.

Results

From the representation of the spectrograms, time seemed to have a moderate to large effect on the neural signature (fig. 5.16). Quantitatively, median PRE values decreased less than 5 percentage points on Keirn's database and between 17 and 22 percentage points on Yeom's data set when comparing normal and *Sess-CV* (fig. 5.17 and table A.27). The use of *rNorm* had opposite effects on both databases, increasing the difference by ∼10 percentage points in the former and reducing it by a similar amount in the latter.

Discussion

Genetic and neurophysiological studies have described changes in human EEG activity across maturation [13]. However, these are long term effects that are disproportionally relevant to younger ages (until approximately 19-20 years old). In the short term, the PSD appears to be relatively stable, comparable with other

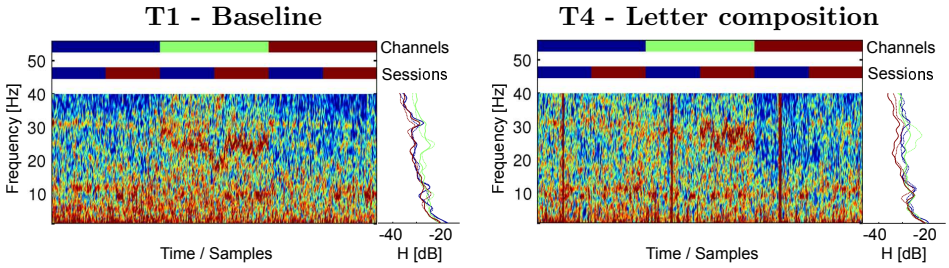


Figure 5.16: Qualitative analysis of permanence. PSD in dB of tasks T1 and T2 of one subject from Keirn's database. The colour bar-codes above each spectrogram represent changes in channels and sessions. PSD for each channel and session is attached to the right of the spectrograms, using the same colour scheme as the channel bar-codes.

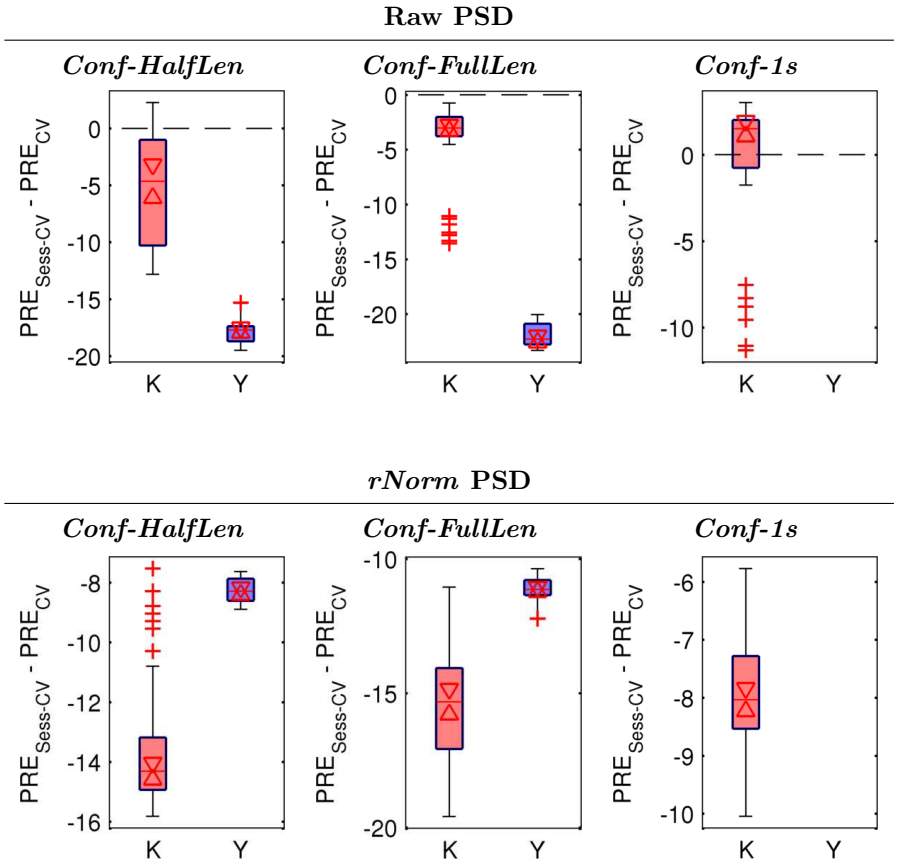


Figure 5.17: Quantitative analysis of permanence. Relative PRE values between normal and *Sess-CV* experiments, obtained with the raw (top) and the *rNorm* (bottom) PSD of Keirn's (K) and Yeom's (Y) databases. Note that for Keirn's data set, *Conf-HalfLen* is equivalent to *Conf-2s*. Refer to the caption of figure 5.9 for details on box markings within the image.

biometric modalities.

The above results are compatible with those obtained by analogous experiments [43, 72, 92, 111, 125]. For these studies, including the present one, the low number of subjects and sessions available is an important limiting factor. Having said that, although a drop in performance is anticipated when train and test sets drift further apart in time:

1. this can be circumvented with a multi-session or a continuous training approach,

2. results suggest that the decline may pose a problem no worse than the one observed on other biometric modalities and

3. this drop would be expected to be lower than the results presented here – with a single training session, the system is unable to model variations across time, leaving it more vulnerable to time effects.

5.4.5 Conclusions

Along this experimentation block, we have defined various properties of the discriminant information within the EEG PSD, some of which were obscured by the lack of consensus across previous research. Specifically, our results suggest that there is no clear location outperforming others systematically across systems, databases and cognitive tasks. Hence, we conclude that the relative performance of individual sensor locations appears to be largely driven by idiosyncrasies of the recordings set-up and by the characteristics of the system.

Looking at the frequency distribution of the discriminant information, results hint at the existence of a performance peak within the alpha rhythm. In addition, the beta rhythm, up to 40 Hz, seemed as much or more discriminant than lower frequencies. Finally, it should be noted that an extra performance peak may arise below 5 Hz.

Vitally, our results suggest that this subject-specific information is 'unique enough' to discriminate between a high number of users (> 100) and is relatively constant along short periods of time. Nevertheless, additional experimentation with larger databases recorded on multiple sessions – with greater temporal distances between each session – are needed to make stronger assertions in this regard.

5.5 Task-independence of the neural signature: results, discussion and conclusions

Finally, we executed a number of experiments specifically designed to test the main hypothesis of this thesis: the existence of a task-independent neural signature. This was indeed hinted by S. Sun results on the proposed MTL approach [93] and by F. Kennet results when trying to identify tasks with the same features used for subject identification [99]. Here, we ran *Task-CV* and *Task-Sess-CV full-fusion* experiments (section 5.1.5) to directly test the hypothesis. The number of MC iterations was increased from 10 to 20. To maximize the number of cognitive tasks and states in each experiment, we replaced the BCI2000-Baseline, BCI2000-Tasks, DEAP-Baseline and DEAP-Playback data sets with their full counterparts BCI2000-Full and DEAP-Full.

5.5.1 Task-independence

We searched for evidence of a task-independent neural signature by means of a *Task-CV* configuration. We set the number of K (from K-Folds CV) equal to the number of tasks in the database, truncated to five when exceeded.

We ran an extra experiment using the *1Task-AllSess* focus mode with a similar K to allow for a better comparison of the results. However, comparisons between *Task-CV* and *1Task-AllSess* experiments should be considered as hints as to the nature of their relationship, rather than as absolute values. Even though

1Task-AllSess confusion matrices were added across cognitive tasks, resulting in a total number of training and testing samples equal to that of the *Task-CV* experiments, they come from as many systems as the number of tasks in the database (one system for each task), as opposed to the single system of *Task-CV* experiments. Although the final computation of the results considers all samples in the database, systems are trained and tested with a different number of samples: *1Task-AllSess* experiments consider samples from each task at a time, while *Task-CV* experiments use all samples simultaneously. Such unbalance will certainly have unexpected effects on the results.

Results

Through the qualitative analysis, we observed that the general shape of the spectrogram, although dynamic, remains relatively stable across cognitive tasks and states, while substantially different across individuals (fig. 5.18). Shape details such as the exact position, height and width of the spectral peaks and valleys are still representative of the subject's identity, but they seem to be more sensitive to task variations.

Quantitative results highlighted a similar scenario (fig. 5.19 and 5.20). PRE values decreased between *Task-CV* and regular *CV* by less than 5 percentage points on average (tables A.28 and A.31). This meant that performance with *Task-CV* was well above chance levels in all cases (tables A.30 and A.31). The use of *rNorm* PSD had no major effect on this relationship. Neither did the application of artefact rejection techniques.

Overall, the difference between *Task-CV* and *1Task-AllSess* PREs showed a big variability across databases and configurations (fig. 5.21 and 5.22). The mean drop in performance was lower than 10 percentage points. As before, *rNorm* and artefact rejection had no major impact on this difference. We observed an exception in Keirn's database, for which *rNorm* reduced the difference by as much as $\sim$10 percentage points.

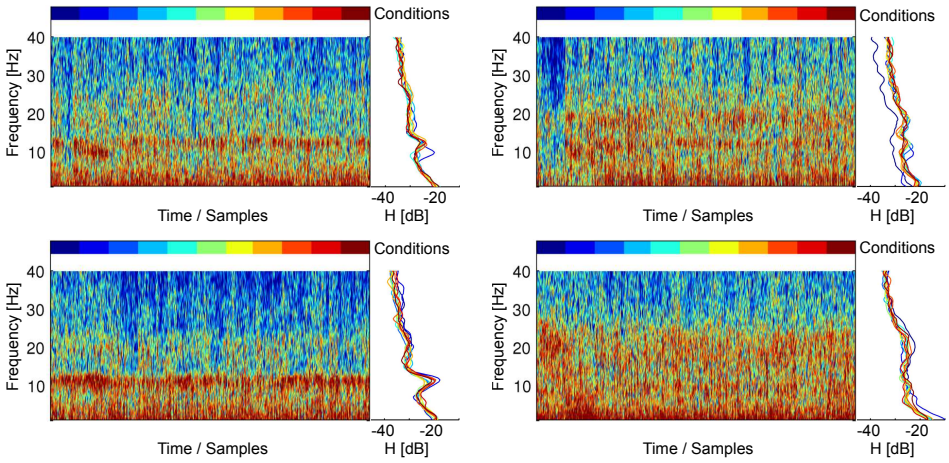


Figure 5.18: Qualitative analysis of task-independence. Spectral representation, in dB, of four subjects from BCI2000-Full. The signal corresponds to the C3 sensor, and was obtained with *Conf-1s* system. The top colour bar-code represent changes in tasks. Each task's PSD is attached to the right of the spectrograms, using the same colour scheme as the task bar-code.

Discussion

If differences in EEG neural signatures were driven predominantly by cognitive tasks rather than the individual's identity, the performance of the biometric system should have experienced a dramatic drop (close to chance levels) in the *Task-CV* experiments. Conversely, the obtained stability supports the hypothesis of a task-independent neural signature. This is especially true when considering that, in the *Task-CV* tests, some subjects were evaluated on the training tasks of others.

This does not suggest that the EEG activity is completely homogeneous across tasks. From the examples provided, task-dependent activity can be easily identified. However, these variations (A) coexist with task-independent features and/or (B) are of a smaller magnitude than differences across individuals. Moreover, the presented quantitative results point to the existence of subject-specific task-specific information, although of lower discrimination power than the task-independent one.

In addition to the above experiments, we also tested the BCI2000-Baseline,

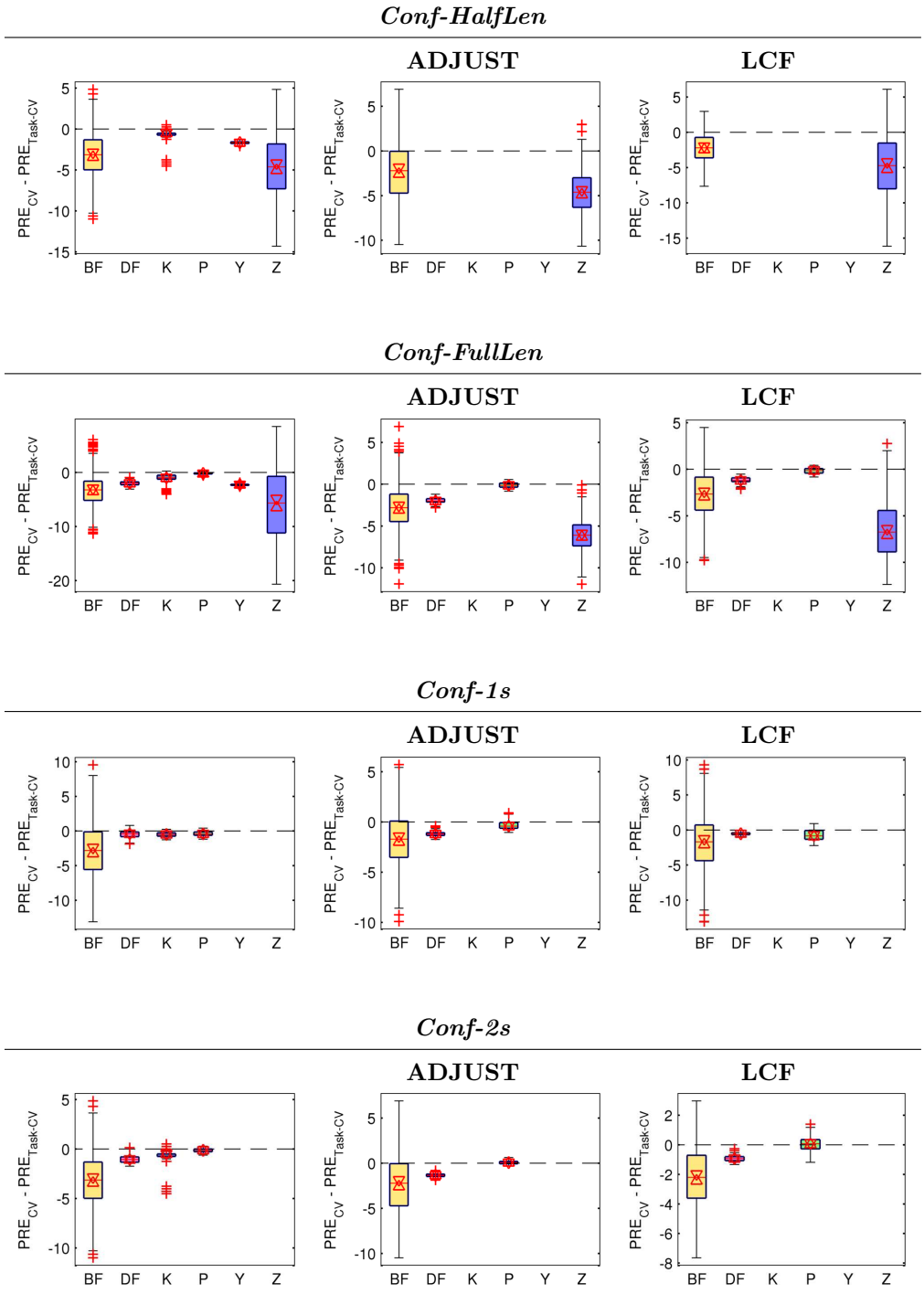


Figure 5.19: Quantitative analysis of the task-independence property. Difference between PRE values of *Task-CV* and regular CV *full-fusion* experiments. A maximum of 20 subjects was used in each experimental iteration. Refer to caption of figure 5.9 for details on box markings within the image.

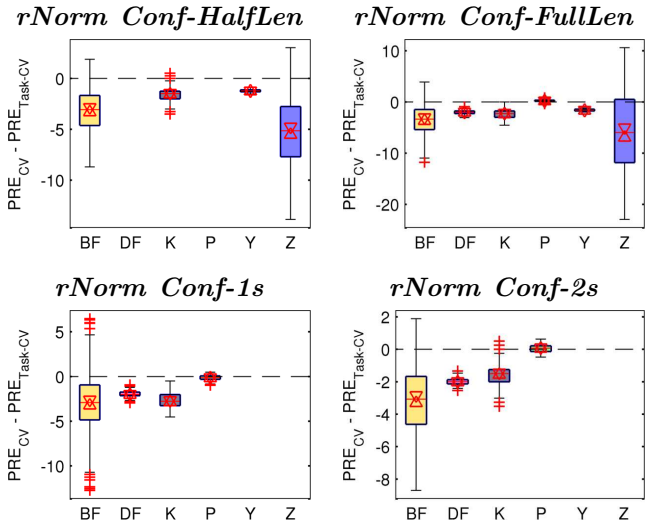


Figure 5.20: Quantitative analysis of the task-independence property with *rNorm* systems. Difference between PRE values of *Task-CV* and regular CV *full-fusion* experiments. A maximum of 20 subjects was used in each experimental iteration. Refer to caption of figure 5.9 for details on box markings within the image.

BCI2000-Tasks and DEAP-Playback datasets[3]. Specifically, from BCI2000-Baseline results we noticed a remarkably high difference in PRE values between *Task-CV* and normal CV or *1Task-AllSess* experiments. We concluded that, with a single cognitive state for training (BCI2000-Baseline has only two states: REO and REC), the system is unable to identify the range of discriminative features that are task-independent. As a result, the system suffers from over fitting and underestimates the true potential of the data, similar to what happened with Keirn's and Yeom's database in *Sess-CV* experiments (section 5.4.4).

5.5.2 Permanence of task-independence

Finally, we ran a double CV experiment on cognitive tasks and sessions to complete the analysis of the EEG traits. As in the previous case, we ran some extra experiments combining focus modes with CV methods.

[3]DEAP-Baseline contained a single state and therefore was inappropriate for these experiments.

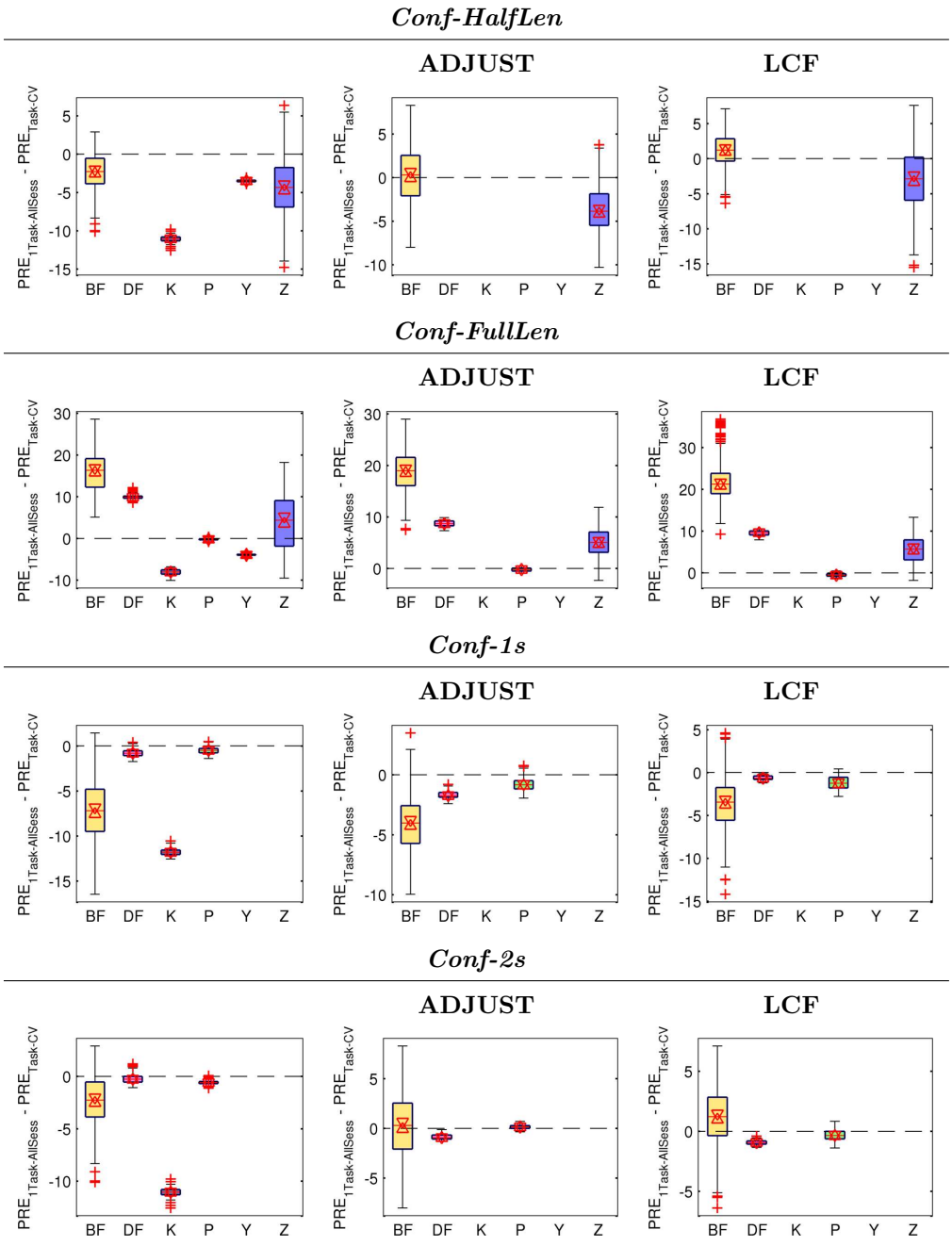


Figure 5.21: *Task-CV* and *1Task-AllSess* comparison. Difference in PRE values between *Task-CV* and *1Task-AllSess full-fusion* experiments. *1Task-AllSess* confusion matrices were added across cognitive tasks. A maximum of 20 subjects was used in each experimental iteration. Refer to caption of figure 5.9 for details on box markings within the image.

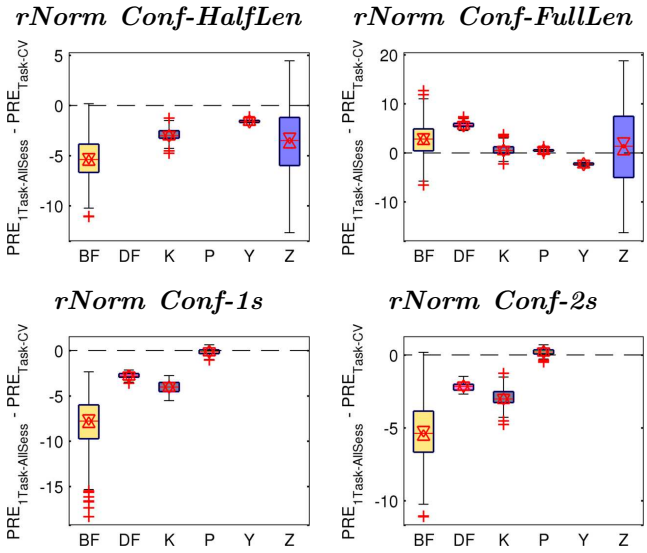


Figure 5.22: *Task-CV* and *1Task-AllSess* comparison with *rNorm* systems. Difference in PRE values between *Task-CV* and *1Task-AllSess full-fusion* experiments. *1Task-AllSess* confusion matrices were added across cognitive tasks. A maximum of 20 subjects was used in each experimental iteration. Refer to caption of figure 5.9 for details on box markings within the image.

Results

From qualitative analysis, time has the same effect as described in section 5.4.4 (fig. 5.23). Qualitatively, the effect of time on PRE values remains as described in the previous experiment, with a drop in performance between *Task-Sess-CV* and *Sess-CV* experiments. In this case, the drop was not greater than 10 percentage points for Keirn's database and not greater than 1 percentage point for Yeom's data set, except with *rNorm Conf-FullLen* system (table A.32). Still, the system's performance remains higher than chance levels (table 5.7).

Discussion

The fact that the effects of time under *Task-Sess-CV* remained as in normal *Sess-CV* tests, supports the existence of a task-independent neural signature. If the performance of the previous experiment (section 5.5.1) had been due to peculiar-

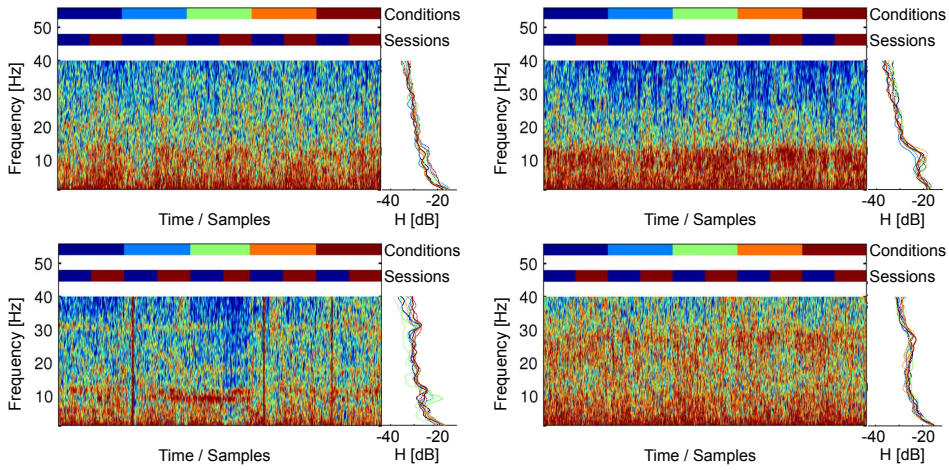


Figure 5.23: Qualitative analysis of the permanence of task-independence. Spectral representation, in dB, of four subjects from Keirn's database. The signal corresponds to the C3 sensor, and was obtained with *Conf-1s* system. The colour bar-codes above the PSDs represent changes in tasks and sessions. PSD for each task and session is attached to the right of the spectrograms, using the same colour scheme as the task bar-code.

ities of the set-up rather than to real subject specific features, PRE values would have slumped to chance levels under *Task-Sess-CV*.

5.5.3 Conclusions

These results provide evidence of the existence of a task-independent neural signature. They suggest that the brain activity, as recorded by an EEG equipment, is more defined by the individual's identity than it is by the performed task or experimental condition. Indeed, the EEG spectrogram reveals that part of this activity is relatively stable across tasks. Specifically, we observed this signature across motor tasks and resting states (BCI2000-Full), emotional states and REO (DEAP-Full), problem-solving tasks (Keirn's), synonyms and non-synonyms AEPs (P. Ulssperger's), self and non-self VEP (Yeom's) and target and non-target VEP (Zhang's).

Our findings also serve to expose the difficulties faced by the neuroscience disciplines, which mainly focus on finding commonalities of brain functioning across

Table 5.7: Quantitative analysis of the permanence of task-independence. Mean PRE and 95% CI obtained with *Sess-CV* experiments and combining *Sess-CV* with *1Task-AllSess* and *Task-CV*. Note that for Keirn's database, *Conf-HalfLen* is equivalent to *Conf-2s*. Refer to table 3.2 for details on databases' code names.

	Sess-CV +		
Dat.	**AllTask-AllSess**	**1Task-AllSess**	**Task-CV**
Conf-HalfLen			
K	73.29	68.52	69.07
	[72.66, 73.92]	[67.72, 69.32]	[65.82, 72.31]
Y	34.36	35.01	33.11
	[33.73, 34.98]	[34.52, 35.51]	[32.31, 33.91]
Conf-FullLen			
K	70.76	56.26	60.15
	[70.30, 71.22]	[55.32, 57.21]	[57.33, 62.97]
Y	35.69	36.36	34.46
	[35.21, 36.16]	[35.98, 36.74]	[33.78, 35.13]
Conf-1s			
K	73.33	71.98	68.81
	[72.50, 74.17]	[71.37, 72.58]	[66.36, 71.26]

individuals. An example of this are BCIs, whose ideal system is one that could be trained offline to differentiate between thought orders and used seamlessly by any person without further tuning.

5.6 Summary and overall conclusions

In this chapter, we have presented the results of an extensive study of the individual's discriminant information within the time-frequency representation of EEG signals. In doing so, we have used 6 databases with different recorded cognitive tasks and states. This, together with the performed complementary qualitative and quantitative analyses, allowed us to distinguish inherent characteristics of the EEG neural signature from idiosyncrasies of individual data sets. Specifically, we

Table 5.8: Quantitative analysis of the permanence of task-independence with *rNorm* systems. Mean PRE and 95% CI obtained with *Sess-CV* experiments and combining *Sess-CV* with *1Task-AllSess* and *Task-CV*. Note that for Keirn's database, *Conf-HalfLen* is equivalent to *Conf-2s*. Refer to table 3.2 for details on databases' code names.

	***Sess-CV* +**		
Dat.	**AllTask-AllSess**	**1Task-AllSess**	**Task-CV**
Conf-HalfLen			
K	82.18	78.14	73.58
	[81.34, 83.02]	[77.41, 78.87]	[70.20, 76.97]
Y	42.60	42.97	41.76
	[42.46, 42.74]	[42.78, 43.16]	[41.38, 42.13]
Conf-FullLen			
K	76.95	65.51	66.53
	[75.96, 77.93]	[64.95, 66.08]	[63.58, 69.47]
Y	47.73	47.59	46.12
	[47.49, 47.97]	[47.44, 47.74]	[45.64, 46.59]
Conf-1s			
K	87.01	83.02	78.19
	[86.64, 87.38]	[82.20, 83.84]	[75.19, 81.19]

ran four experimentation blocks, each with a specific goal, which resulted in the following synthesized recommendations and conclusions:

1. Configuration of the PSD: recommendations

   (a) Record, at least, 5 seconds of EEG to perform the identification.

   (b) Divide the EEG into segments between 1 or 2 seconds long.

   (c) If data volume and computational speed is not an issue, use some degree of window overlap.

   (d) Compute the spectral representation for each window, using a number of spectral coefficients similar to the number of samples within the window to maintain the spectral resolution.

(e) Retain a bandwidth from the lowest frequencies to 30 or 40 Hz.

(f) Perform classifications for each window individually and generate a single response by fusing scores.

2. Representation of the time and frequency domains: recommendations

(a) Use *AvgMnt* as the default montage.

(b) Consider using *BIHMnt* in cases where processing time or data volume is a concern.

(c) As a substitute for complex artefact rejection methods, you may normalize the spectral coefficients with a method robust to outliers.

3. Properties of the discriminant information: conclusions

(a) There seems to be no best performing sensor location found across systems, databases and/or cognitive tasks.

(b) In terms of frequency distribution, there is a performance peak within the alpha rhythm. Frequencies within the beta rhythm (up to 40 Hz) are much or more discriminant than lower bands. Frequencies below 8 Hz may also contain an important amount of discriminant information.

(c) Subject traits within the EEG activity are 'unique enough' to discriminate 100+ subjects when an appropriate system configuration is used.

(d) Subject-specific EEG spectral patterns seem to be 'permanent enough' to use them as a biometric modality.

(e) ***From the full discriminant information within EEG spectrum, a significant part of it is independent from the ongoing task.***

Chapter 6

Biometric system implementation

For the final step of the current work, we applied the knowledge gained on the properties and characteristics of the neural signature to the design of an EEG biometric verification system. In this chapter, we will describe the architecture of all the evaluated systems, with special emphasis on those that yielded relevant results. We will show that the real cepstrums coefficients and the Lineal Prediction Coefficients (LPC) can code most of the discriminant information within the PSD of the EEG. We will begin describing the experimentation methodology followed in this study.

6.1 Methodology

All the experiments ran in this phase implemented a verification paradigm[1]. This is arguably more suitable for the biometry in hand than that of classification[2], as in an EEG biometric system, the user's consent and collaboration will always be required to proceed. More so than in other biometries, since users need to wear an EEG device during the process.

[1]In a verification paradigm, the system is presented with the user's biometry along with his or her identity, and it has to decide whether the two pieces of information match or not.

[2]In a classification paradigm, the system is only presented with the user's biometry and it has to decide which registered user it corresponds to, if any.

As before, experiments combined stratified K-Folds and MC CV methodologies. In this case, a subject was randomly selected as the positive class for each MC iteration. The remaining subjects were therefore treated as impostors. In addition, we applied an open-set segmentation, where different sets of impostors were used for training and testing. Hence, impostors were segmented at the subjects level, using all data from an impostor (all tasks and sessions) either for training or testing – as opposed to the segmentation of the registered user data, which was done at the sample level.

To keep these experiments in line with our proposed hypothesis of a task-independent neural signature, experiments were run on a *task-CV* basis if a single session was available, and on a *task-sess-CV* basis for Keirn's and Yeom's data sets, as described in section 5.1.5. The number of folds (from K-Folds) was set to 2, thus setting half of the data for training and half for testing[3], and the number of MC iteration was fixed to 20 — bare in mind that in each MC iteration a different subject is used as the positive class. In databases where the number of subjects was smaller than 20, we repeated the whole K-Fold MC process N times until 20 experiments were executed. For example, for Keirn's database (5 subjects) we repeated the whole process 4 times (4 x 5 = 20).

For the computation of results, we proceed as in the previous chapter, i.e. we aggregated the confusion matrices from each K-Fold and computed average and 95% CI across MC iterations. In this chapter, we will report accuracy results (Acc.), defined as the average between Genuine Acceptance Rate (GAR) or sensitivity (percentage of positive samples correctly classified) and specificity (percentage of negative samples correctly classified). Receiver Operating Characteristic (ROC) curves, showing False Acceptance Rate (FAR) against GAR for

[3]For Keirn's and Yeom's data sets, where a *task-sess-CV* is applied. That is, for the positive class (registered user), only half of the data from one session is used for training, and the opposing half from the other session for testing. This is done to maintain cognitive tasks crossed between sets.

different decision thresholds, will also be provided together with optimal performance points.

6.2 Baseline design

Our baseline design was derived entirely from the findings of chapter 5. The EEG signal was first divided into 75%-overlapped windows of length 2 seconds – except for Yeom's and Zhang's databases, for which 0.5-second windows were used. We then applied a Hamming window and computed the PSD coefficients of each of these segments. Next, we built a feature vector for each window by concatenating the 1-to-40 Hz spectral coefficients from all the EEG channels. An LDC was used for classification. For each sample, the final score was obtained by averaging the scores of each of the overlapped windows.

We relied on the normalization of the spectral coefficients by *rNorm* (table 5.3) as an alternative to artefact rejection algorithms (section 5.3.2). The application of any of the methods described in section 4 to a practical system is not straight forward. Firstly, the computation of the BSS model requires large volumes of data and is computationally expensive. Secondly, BSS model need to be trained for each recording session, i.e. each time the user put on the EEG cap, which makes the previous point even more troublesome. Hence, although possible to some degree, the adaptation of these techniques to be suitable for the current problem is out of the scope of this work.

Results

As it was the case in the experiments of section 5.3.2, the application of *rNorm* had mixed effects on the system's performance (table 6.1). On average, *rNorm* had a neutral or positive effect on performance and/or stability. However, this effect was not statistically significant (t-test; $df = 38$; $p < 0.05$), except on DEAP-Full dataset, for which *rNorm* boosted performance by 12 percentage points and

Table 6.1: Accuracy results of the baseline system. Mean and 95% CI accuracy values obtained with the baseline design using the raw and *rNorm* PSD coefficients. Within data sets, performances statistically lower than the maximum are pointed by * (single tail t-tests with $df = 38$ and BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names.

System	**BF**	**DF**	**K**
Raw	80.52	82.79*	**66.47**
	[77.82, 83.23]	[78.54, 87.04]	**[61.65, 71.30]**
rNorm	**82.75**	**94.70**	62.99
	[79.22, 86.27]	**[92.05, 97.36]**	[60.95, 65.03]
System	**P**	**Y**	**Z**
Raw	81.99	68.72	**77.58**
	[77.83, 86.14]	[64.41, 73.03]	**[72.18, 82.98]**
rNorm	**84.61**	**72.81**	71.13
	[80.36, 88.86]	**[68.85, 76.77]**	[66.87, 75.39]

reduced *CI* by 3 points.

The benefits of *rNorm* were more easily seen in the obtained ROC curves, especially for lower values of FAR (fig. 6.1 and 6.2). With BCI2000-Full, DEAP-Full and P. Ullsperger's data sets, *rNorm* systems reported higher GAR for FARs below 10%. The opposite was true for Keirn's database, whose ROC curve was closer to chance levels with the *rNorm* system. Results of Yeom's data set remained virtually unaltered by *rNorm*. Finally, we observed mixed effects in Zhang's database case, where the *rNorm* system gave worst results for FARs above 1% and better below that point.

Discussion

These results reinforced the conclusions of section 5.3.2 regarding *rNorm*, which may have a positive effect on the system's performance. After inspecting the

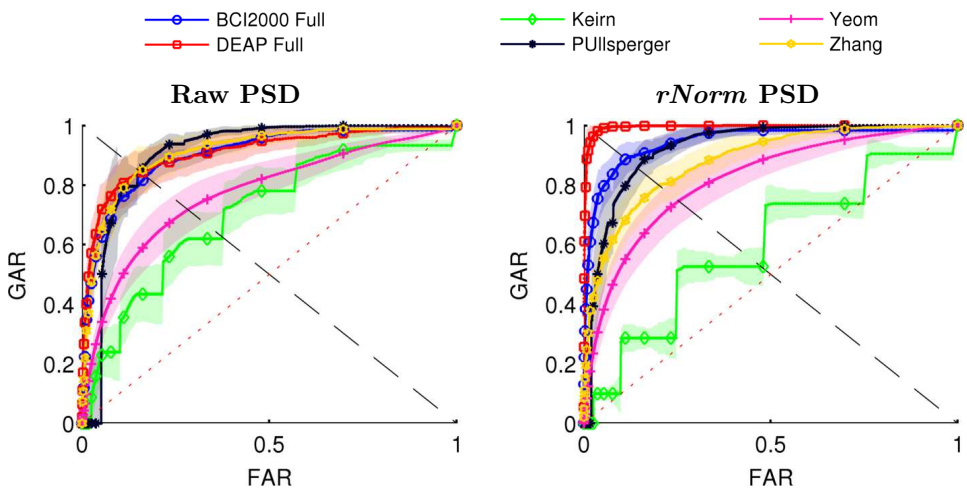


Figure 6.1: ROC curves of the baseline design. Mean ROC curves and std (shaded area) obtained with the raw (left) and *rNorm* (right) PSD coefficients. A maximum of 20 subjects was used in each experimental iteration. The dotted red diagonal line represents chance accuracy levels, while the black dashed diagonal line represents EER points.

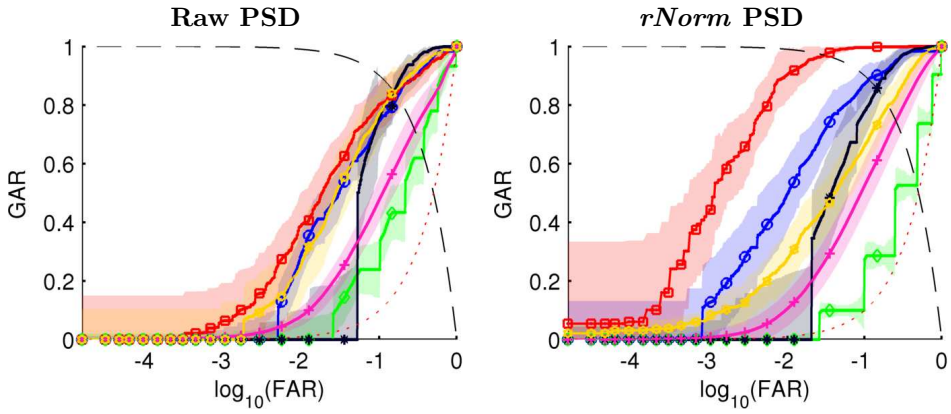


Figure 6.2: ROC curves of the baseline design with logarithmic FAR axis. Mean ROC curves and std (shaded area) obtained with the raw (left) and *rNorm* (right) PSD coefficients. A maximum of 20 subjects was used in each experimental iteration. Refer to fig. 6.1 for legend details.

validation results[4] of Zhang's database, we concluded that the system may have suffered from over fitting with the application of *rNorm*. To maximize the potential of subsequent systems, we continued to consider both versions of the PSD (raw and *rNorm*).

6.3 Feature extraction methods: results, discussion and conclusions

Once the baseline had been set, we explored the use of different descriptors to characterize the discriminant information. In chapter 5, we identified the overall shape of the EEG's spectrum as the source of the discriminant information. Hence, we considered the application of cepstral analysis and spectral envelope coefficients.

6.3.1 Cepstral coefficients

Cepstral coefficients have been extensively used on signal processing problems [155]. The Real Cepstrums (RCeps) (usually just called cepstrums) are defined as

$$C(q) = |FFT^{-1}(log(|FFT(X(t))|^2))|^2, \tag{6.1}$$

where $X(t)$ is a signal. Here, we delimited the cepstral analysis to the frequencies in [1, 40] Hz. In addition, following our previous results, we performed FFT^{-1} to the raw and *rNorm* PSD. Finally, we configured FFT^{-1} to compute the same number of coefficients as in the FFT step.

The cepstral space codes the broad shape of the spectrum in the lower quefrencies (first coefficients), and its details and periodicity in higher quefrencies. As we have previously concluded that the discriminant information is coded in the overall shape of the spectrum, we expected to find the bulk of the discriminative

[4]Unless otherwise specified, the term "validation results" refer to the results obtained with the training set. As we are not directly dealing with the optimization of parameters during training, there is no need to have a disjoint validation set.

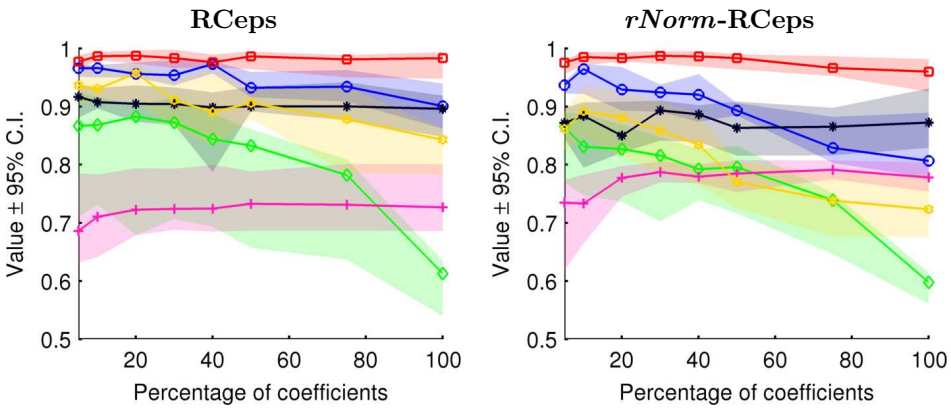


Figure 6.3: Performance of RCeps$_{P\%}$ system with increasing number of coefficients. Mean accuracy values and 95% CI (shaded area) obtained with the first P% of the real raw (left) and *rNorm* (right) cepstral coefficients. A maximum of 20 subjects was used in each experimental iteration. Refer to fig. 6.1 for legend details.

power in the first few coefficients. To test this, we designed a system that used the first P% of the computed coefficients (RCeps$_{P\%}$), and configured it with different values of P.

Results

Across databases and normalization conditions, maximum performance was achieved at different values of P (fig. 6.3). Having said that, the majority of the configurations gave performances close the maximum (tables A.33 to A.35). On BCI2000-Full, Keirn's and Zhang's databases, using too many cepstral coefficients translated on a lost of performance. This was especially acute on Keirn's and Zhang's data sets, whose accuracy dropped ~ 20 percentage points.

Looking at the ROC curves of the RCeps$_{P\%}$ from raw PSD, we noticed a steady improvement of the GAR for lower FARs peaking at $P = 20\%$ (fig. 6.4). After that point, the GARs oscillated, regaining the maximum values in some cases but not in others. The *rNorm*-RCeps showed a poorer performance for lower FAR points than the raw RCeps (A.13).

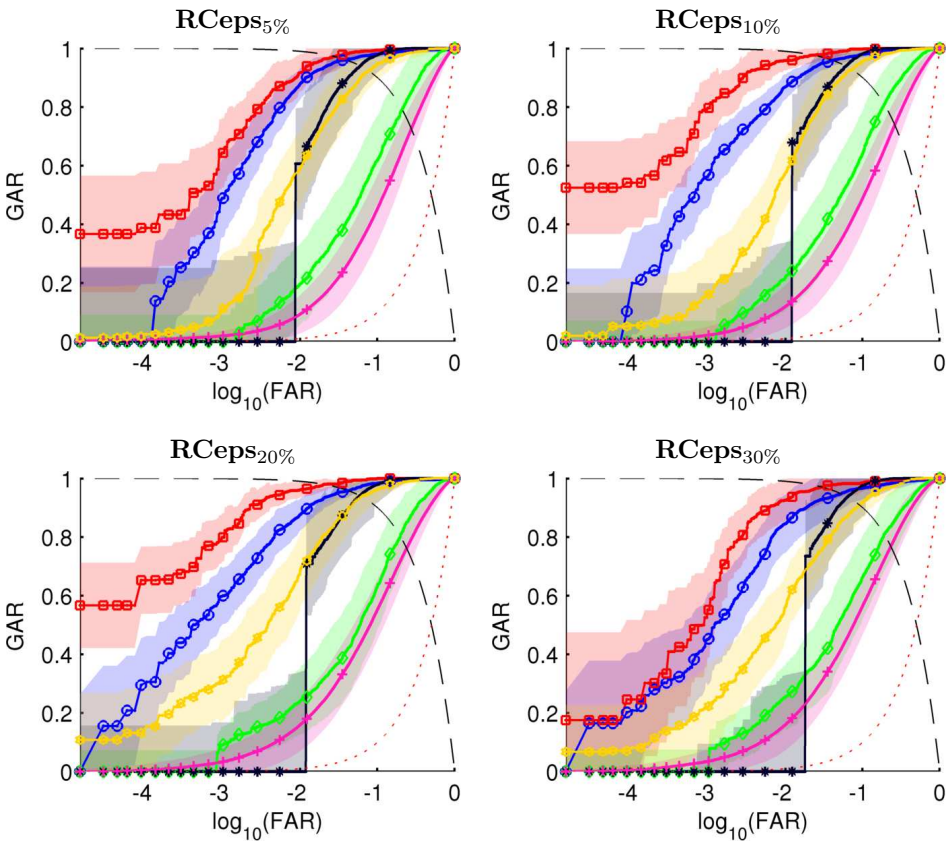


Figure 6.4: ROC curves of RCeps$_{P\%}$ systems coefficients with logarithmic FAR axis. Mean ROC curves and std (shaded area) obtained with different P values. A maximum of 20 subjects was used in each experimental iteration. Refer to fig. 6.1 for legend details.

Discussion

As with the baseline experiment, results were not fully homogeneous across databases and systems, hindering the selection of the optimal configuration point. We chose to retain 20% of the cepstrums for the remaining experiments, and rejected the application of *rNorm*.

6.3.2 Spectral envelope coefficients

The coefficients of an AR model (or LPC) have been a popular choice for subject characterization within the EEG biometric identification literature (chapter 2).

This model predicts samples of a time series as a function of the past p observations, where p is the order of the system. This is typically defined as

$$X[t] = c + \sum_{i=1}^{p} \varphi_i X[t-i] + \epsilon_t, \tag{6.2}$$

with c a constant, φ_i the parameters of the model and ϵ_t white noise. The above can also be seen as the output of an all-pole Infinite Impulse Response (IIR) system with noise presented at its input. Therefore, the LPC describe the spectral shape of the modelled signal. The higher the order of the model, the more detailed the description.

As with the cepstrums, we would like to capture the overall shape of the spectrogram and discard irrelevant details from each sample. Hence, we expected to find a peak in accuracy with increased order N (LPC_N). Once an optimal order was fixed, we explored the use of other representation forms of the LPC. Specifically: RC and Line Spectral Pairs (LSP), which have been identified as robust against noise. In addition, we also tested the performance of the model's fitting error ϵ.

Results

Overall, results show a behaviour similar to that of the cepstrums, with high variation of the optimal point (fig. 6.5) and great number of configurations performing similar to the optimal (tables A.37 and A.38). An accuracy within 1.5 percentage points to the maximum is reached at order 8. Orders above this had no effect on the performance of the system, except on Zhang's database, whose accuracy decreased abruptly passed order 25.

The tendency was less clear when looking at GAR for lower FARs, with large variation across databases. Having said that, a steady increase in GAR was observed in all cases up to order 8, followed by oscillating performances. In some cases, the maximum GARs were obtained with higher orders; e.g. LPC_{20} for BCI2000-Full database and LPC_{40} for DEAP-Full data set.

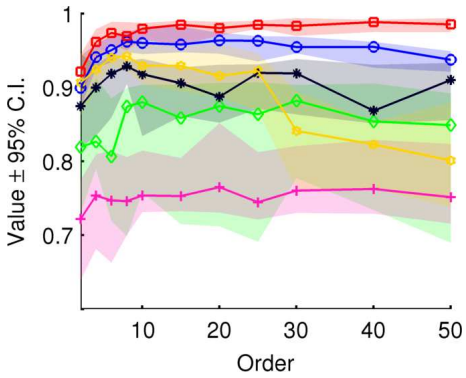


Figure 6.5: Performance of LPC$_N$ systems. Mean and 95% CI (shaded area) verification accuracy obtained with different values of N. A maximum of 20 subjects was used in each experimental iteration. Refer to fig. 6.1 for legend details.

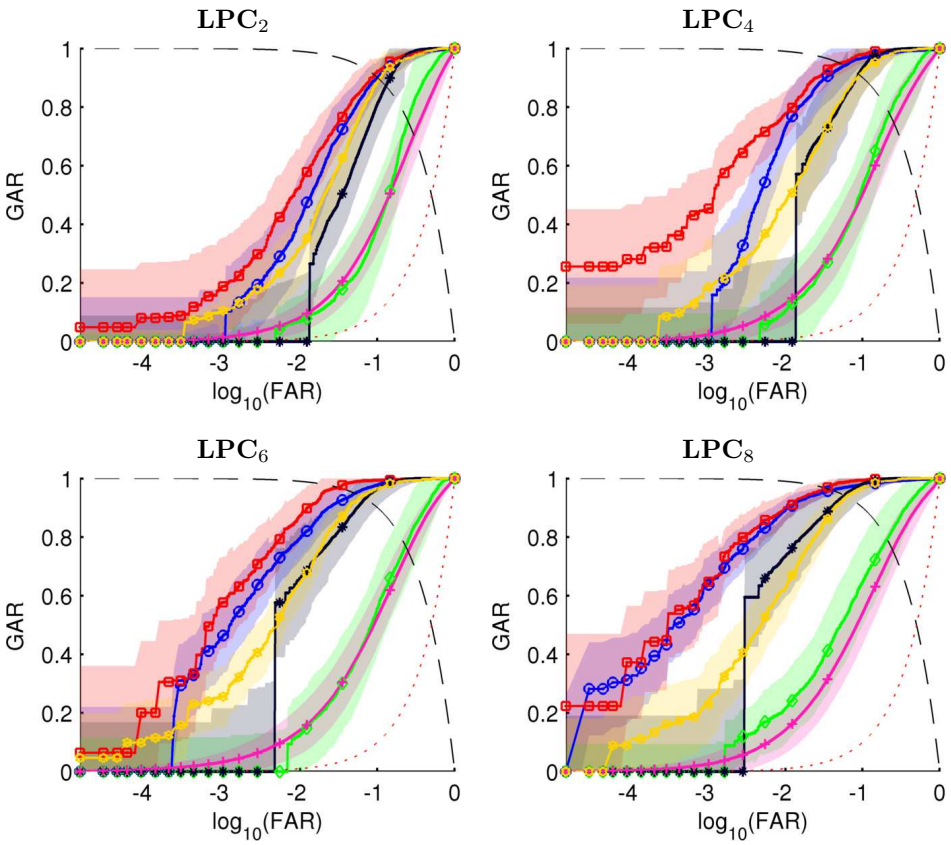


Figure 6.6: ROC curves of LPC$_N$ systems coefficients with logarithmic FAR axis. Mean ROC curves and std (shaded area) obtained with different N. A maximum of 20 subjects was used in each experimental iteration. Refer to fig. 6.1 for legend details.

Table 6.2: Results of spectral envelope coefficients. Mean and 95% CI verification accuracy of LPC_8, RC_8, LSP_8 and the ϵ_8. Within data sets, performances statistically lower than the maximum are pointed by * (single tail t-tests with $df = 38$ and BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names.

System	BF	DF	K
LPC_8	**96.01**	96.82	**79.10**
	[95.04, 96.99]	[95.66, 97.97]	**[72.03, 86.18]**
RC_8	95.72	96.52	77.29
	[94.66, 96.77]	[95.00, 98.04]	[70.94, 83.63]
LSP_8	95.27	**97.91**	78.77
	[93.31, 97.24]	**[97.18, 98.63]**	[69.77, 87.78]
ϵ_8	81.86*	73.29*	64.27*
	[78.44, 85.27]	[68.51, 78.08]	[59.08, 69.46]
System	**P**	**Y**	**Z**
LPC_8	**89.88**	74.73	93.93
	[85.49, 94.26]	[69.57, 79.88]	[92.19, 95.66]
RC_8	85.47	73.50	93.14
	[79.73, 91.21]	[67.98, 79.03]	[90.51, 95.77]
LSP_8	88.35	**74.79**	**94.03**
	[83.25, 93.45]	**[69.02, 80.57]**	**[92.31, 95.75]**
ϵ_8	80.97*	62.02*	81.15*
	[77.01, 84.92]	[57.69, 66.35]	[76.29, 86.01]

Comparing the accuracy of 8-order LPC, RC, LSP and ϵ, only ϵ_8 performed substantially worse than the rest (table 6.2). LPC_8, RC_8 and LSP_8 resulted in virtually equivalent performances. Inspecting the ROC curves at lower FAR values, LPC_8 produced, on average, best GAR (fig. 6.7).

Discussion

From the above results, we conclude that an order of 8 is a good compromise between accuracy and feature dimensionality, and is sufficient to characterize the

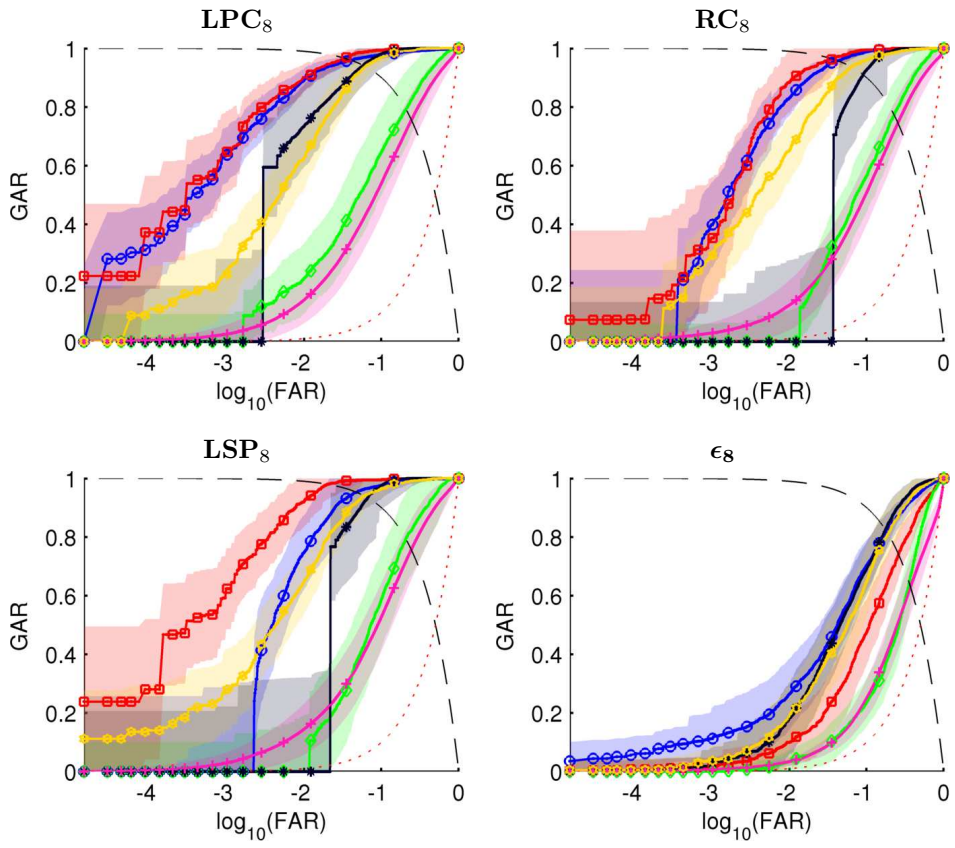


Figure 6.7: ROC curves for different spectral envelope coefficients. Mean and std (shaded area) ROC curves obtained with LPC_8, RC_8, LSP_8 and ϵ_8. A maximum of 20 subjects was used in each experimental iteration. Refer to fig. 6.1 for legend details.

discriminant spectral shape of the EEG activity. In addition, LPC_8 will be used in future designs, as there was no clear benefit from the application of the more computationally expensive RC_8 and LSP_8.

These results contradict some of the conclusions in the literature. R.B. Paranjape et. al. concluded that an increase in the order of the AR model is necessary to bear with the rise in the number of users [79]. Even with only 5 subjects, they reported an increase in classification accuracy of 7 percentage points when moving from order 9 to 15. In addition, Campisi's team found RC to outperform LPC [81, 111]. These discrepancies may be due to differences in the experimentation methodology: classification versus verification experiments, differences in the

systems' architectures, and/or idiosyncrasies of the databases.

6.3.3 Other evaluated features

In addition to the described features, we also tested the following ones[5]:

- Time statistics of feature vectors: We computed several statistical measurements across time (across windows of the STFT) and used the results as inputs to the LDC. Specifically, we computed the mean, std, kurtosis, skewness and the percentiles 5, 25, 50 (median), 75 and 95. Mean and median statistics performed at a level similar to, or worse than, the original system, while the remaining measurements performed relatively poor. In addition, ROC curves showed a poorer behaviour away from the EER point. Having said that, as with the BIHMnt in section 5.3.1, this method has the advantage of reducing the number of vectors to be evaluated, as the N_W windows are converted into a single vector. Thus, it might be considered in cases where the volume of data or the processing speed is a concern.

- Feature statistics of feature vectors: We computed the statistics just described, but across features within each sample instead of across time. Hence, the number of windows remained the same and the length of the feature vector was reduced to one (the static value). In this case, results were substantially worse than that of the original systems.

- APS of bands: We divided the PSD into bands and computed the powers within each band. Results were significantly worse than that of the baseline system; i.e. based on the full PSD vector.

[5]Since these features yielded worse or equivalent performances with a more complex design, their results are not reported in detail here.

6.3.4 Conclusions

Overall, the first design step showed some encouraging results (table 6.3). Performance of the baseline design were already above 80% for all databases except Yeom's. RCeps and LPC clearly outperformed the PSD coefficients in all cases and the *rNorm*-PSD in 3 out of 6 cases. This was particularly interesting considering the dimensionality reduction of the feature vector of the former two systems compared to the PSD and *rNorm*-PSD. Based on the results, $\text{RCeps}_{20\%}$ and LPC_8 features were considered on subsequent design.

6.4 Other evaluated designs

A number of different architectures were evaluated in addition to the ones described so far. However, as they all yielded worse or equal performances with a more complex design, their results are not reported in detail here. In particular, the following designs were tested:

- Feature fusion: We fused $\text{RCeps}_{20\%}$ and LPC_8 in a single vector and fed it to the LDC. This fusion performed similarly to the individual features $\text{RCeps}_{20\%}$ and LPC_8, which evidenced the high level of correlation between the information extracted by both methods.

- Fusion of statistical measurements: We combined the statistical measurements described in section 6.3.3 taken from $\text{RCeps}_{20\%}$ or LPC_8 in a single feature vector. The fusion of time-statistics performed similar to, or worse than, the individual mean and median vectors. On the other hand, the concatenation of the feature-statistics, i.e. taken within each vector instead of across time, produced a remarkable increase in accuracy compared to individual statistics. In some cases there were more than 10 percentage points of improvement. Therefore, we considered the fusion of feature-statistics to the original $\text{RCeps}_{20\%}$ or LPC_8 vectors. Based on R. Palaniappan's results [51]

Table 6.3: Results of selected configurations. Mean and 95% CI results of the selected configurations of PSD, RCeps and LPC systems. A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names.

System	BF	DF	K
PSD	80.52*	82.79*	66.47*
	[77.82, 83.23]	[78.54, 87.04]	[61.65, 71.30]
rNorm-PSD	82.75*	94.70	62.99*
	[79.22, 86.27]	[92.05, 97.36]	[60.95, 65.03]
$RCeps_{20\%}$	95.73	**97.59**	80.06
	[94.56, 96.90]	**[96.12, 99.05]**	[73.97, 86.16]
rNorm-$RCeps_{20\%}$	93.04*	97.42	**80.11**
	[90.99, 95.09]	[95.67, 99.17]	**[75.66, 84.55]**
LPC_8	**96.01**	96.82	79.10
	[95.04, 96.99]	[95.66, 97.97]	[72.03, 86.18]
System	**P**	**Y**	**Z**
PSD	81.99*	68.72*	77.58*
	[77.83, 86.14]	[64.41, 73.03]	[72.18, 82.98]
rNorm-PSD	84.61	72.81	71.13*
	[80.36, 88.86]	[68.85, 76.77]	[66.87, 75.39]
$RCeps_{20\%}$	89.19	73.94	93.18
	[85.55, 92.83]	[69.04, 78.84]	[90.54, 95.82]
rNorm-$RCeps_{20\%}$	85.12	**75.37**	86.92*
	[80.18, 90.07]	**[71.41, 79.33]**	[82.93, 90.91]
LPC_8	**89.88**	74.73	**93.93**
	[85.49, 94.26]	[69.57, 79.88]	**[92.19, 95.66]**

we were expecting this to improve the system's accuracy. Nevertheless, once more, results were similar to or worse than those of the original systems. The enhancement observed by Palaniappan may therefore be due to a suboptimal extraction of the AR coefficients (he used EEG segments of 0.5 seconds).

- Score fusion of $RCeps_{20\%}$ and LPC_8 individual systems: This performed equal to, or worse than, the feature fusion version.

- Multi-window length: Throughout chapter 2 we used systems with different STFT window lengths (L_W). We considered the possibility that these configurations extracted uncorrelated discriminant information, and created a design based on multiple sub-systems with different L_W fused at score level. Results were equal to or worse than those obtained with the original systems.

- Projection methods: We applied PCA, LDA and ICA to all the described architectures. This was not a bid to reduce the data dimensionality, but rather an attempt to present the discriminant information to the classifier in a more suitable way, in the hope of boosting the performance and/or improve the stability of the system. We applied such projection techniques to the features of each sensor individually (as sensor experts) and to the vector containing the features from all sensors. In both cases, results were similar or worst than those of the original systems.

- Score fusion of sensor experts: This architecture contained as many subsystems as EEG sensors. The LDC scores obtained for all sensors were then averaged to build the final response. We evaluated this with all the described systems. In all cases, results were similar to those of the original systems.

- Non-lineal classifiers: In the classification phase, we evaluated the performance of RBF-SVM and ANN. The kernel and cost parameters of the former were optimized applying a CV procedure within the training data. To avoid any over fitting, this inner CV followed the same principle as the one used for testing, meaning different subjects were used as impostors during training and validation. Similarly, multiple configurations of the hidden layers were evaluated for the ANN. In both cases, we only managed to equal the results of LDC. Interestingly, the optimization of the RBF-SVM parameters showed a clear tendency to create a linear model, rather than a non-linear one.

Table 6.4: Results obtained when testing with all the available subjects. Mean and 95% CI results of $RCeps_{20\%}$ and LPC_8 systems. Although results of DEAP-Full, Keirn's, P. Ullsperger's and Yeom's data sets remained the same as before – as they contain less or equal to 20 subjects –, they are replicated here for ease of inspection. Refer to table 3.2 for details on databases' code names.

System	BF	DF	K
RCEPS	95.40 [94.70, 96.10]	**97.59 [96.12, 99.05]**	**80.06 [73.97, 86.16]**
LPC	**95.42 [94.79, 96.05]**	96.82 [95.66, 97.97]	79.10 [72.03, 86.18]
System	**P**	**Y**	**Z**
RCEPS	89.19 [85.55, 92.83]	73.94 [69.04, 78.84]	91.58 [89.30, 93.87]
LPC	**89.88 [85.49, 94.26]**	**74.73 [69.57, 79.88]**	**93.24 [91.53, 94.94]**

6.5 Experiments using all the available subjects

Finally, we re-evaluated the best performing systems – i.e. $RCeps_{20\%}$ and LPC_8 with an LDC classifier – with all the subjects of BCI2000-Full and Zhang's data sets. That is, 100 subjects of BCI2000-Full database and 30 subjects for Zhang's data set.

Results

Surprisingly, results remained virtually equal despite the increase in the number of subjects: from 20 to 100 in the case of BCI2000-Full data set, and from 20 to 30 in the case of Zhang's database (table 6.4).

After inspecting the ROC curves (tab. 6.5, fig. 6.8 and 6.9), we concluded that both $RCeps_{20\%}$ and LPC_8 systems maximized their potential; i.e. obtained accuracy values close to their EER. The only notable exception was P. Ullsperger's data set, which underperformed on both systems by $\sim$5 percentage points. The

Table 6.5: Optimal ROC points obtained when testing with all the available subjects. Refer to table 3.2 for details on databases' code names.

Feature	BF	DF	K	P	Y	Z
$RCeps_{20\%}$	96.85	97.99	80.71	94.42	76.19	93.21
LPC_8	96.06	96.89	79.36	94.46	75.61	94.02

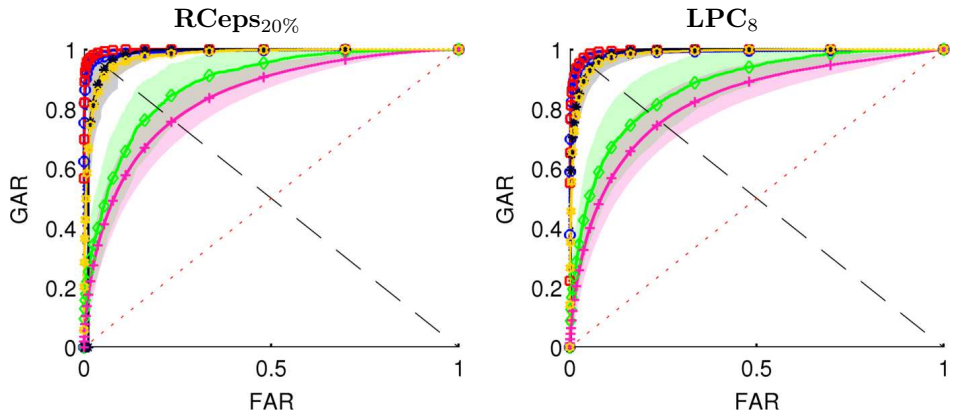


Figure 6.8: ROC curves obtained when testing with all the available subjects. Mean ROC curves and std (shaded area) of $RCeps_{20\%}$ and LPC_8 systems. Refer to fig. 6.1 for legend details.

obtained (reported) accuracy corresponds to a sensitivity of 99% and a specificity of 78% and 79% for $RCeps_{20\%}$ and LPC_8 respectively. Therefore, the system underestimated the threshold, resulting in a sub-optimal configuration situated to the right of the EER point.

On BCI2000-Full and DEAP-Full databases, $RCeps_{20\%}$ and LPC_8 systems showed a strong performance, reflected in their ROC curves. GAR was above 88% for an FAR of 1% (table 6.6), and above 82% for an FAR of 0.5% (table 6.7). Experiments with Zhang's database resulted on a GAR above 80% for 1% FAR, and 55% for 0.5% FAR. The remaining data sets had relatively poor performances at these FAR values.

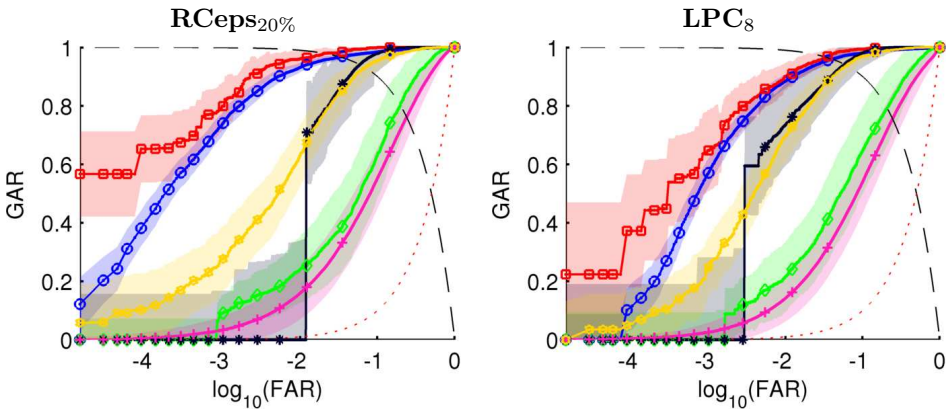


Figure 6.9: ROC curves with logarithmic FAR axis obtained when testing with all the available subjects. Mean ROC curves and std (shaded area) of $\text{RCeps}_{20\%}$ and LPC_8 systems. Refer to fig. 6.1 for legend details.

Table 6.6: GAR for an FAR of 1% obtained when testing with all the available subjects. Mean and std of GAR obtained with $\text{RCeps}_{20\%}$ and LPC_8 systems. Results computed with all the available subjects for each database. Refer to table 3.2 for details on databases' code names.

System	BF	DF	K
$\text{RCeps}_{20\%}$	**93.17** [**91.54, 94.80**]	**96.11** [**93.44, 98.77**]	**24.05** [**12.82, 35.28**]
LPC_8	88.49 [86.17, 90.81]	90.04 [85.37, 94.72]	21.61 [9.70, 33.52]

System	P	Y	Z
$\text{RCeps}_{20\%}$	0.00 [0.00, 34.03]	**16.01** [**7.06, 24.95**]	63.18 [52.79, 73.58]
LPC_8	**73.54** [**59.48, 87.60**]	14.27 [6.26, 22.28]	**69.65** [**60.74, 78.57**]

Table 6.7: GAR for an FAR of 0.5% obtained when testing with all the available subjects. Mean and std of GAR obtained with $RCeps_{20\%}$ and LPC_8 systems. Results computed with all the available subjects for each database. Refer to table 3.2 for details on databases' code names.

System	BF	DF	K
$RCeps_{20\%}$	**89.83**	**94.32**	**17.92**
	[87.43, 92.24]	**[90.51, 98.13]**	**[8.75, 27.09]**
LPC_8	82.08	85.53	16.16
	[79.05, 85.11]	[78.68, 92.37]	[5.64, 26.68]
System	**P**	**Y**	**Z**
$RCeps_{20\%}$	0.00	**10.33**	49.44
	[0.00, 29.37]	**[2.70, 17.97]**	[38.29, 60.58]
LPC_8	**64.89**	8.83	**54.86**
	[48.05, 81.73]	[2.33, 15.32]	**[45.17, 64.56]**

Discussion

Overall, we obtained relatively good and stable accuracy results across systems and databases. Although BCI2000, Keirn's, Yeom's and Zhang's databases have also been used by other authors, a direct comparison between our results and those within the state of the art is impossible due numerous reasons (tables 6.8 to 6.11). In addition to differences in the number of subjects and tasks used and in the applied experimentation methodology, the following two factors should be considered:

Firstly, the design proposed here is a generalised one. Unlike most of the systems presented by other authors, we have designed the system using multiple databases and selected the collective optimal configurations – these configurations were not necessarily optimal for individual data sets. Hence, our system is not tuned to maximize the performance within a single data set, but to work well across multiple scenarios. In fact, the performance of the three largest databases are within 6 percentage points within each other, suggesting that this level of

Table 6.8: Comparison of results obtained with BCI2000 database. Columns correspond to the publication reference (Ref.), the number of subjects used (# Subj.), the task used (Task), the CV method applied (CV method), and the success rate (Succ.). In addition, NIt refers to the number of MC iterations. For completeness, classification results from table A.30 are also shown, these correspond to the analysis of chapter 5.

Ref.	**# Subj.**	**Task**	**CV method**	**Succ.**
Classification experiments				
[64]	109	REO and REC	K-Fold; K = 5	100%
[62]	18	T4 (section 2.3.5)	K-Folds; K = 3	96.5%
This	20	All	Open-set; *Task-CV*; K-Fold + MC; $K = 5 + NIt = 20$	84.88%
Verification experiments				
[63]	109	REO	Unspecified	95.6% (from EER)
This	100	All	Open-set; *Task-CV*; K-Fold + MC; $K = 2 + NIt = 20$	95.42%

discrimination is indeed a property of the neural signature and not due to idiosyncrasies of the data. This is also supported by the observed stability of the results when increasing the number of subjects (from 20 to 100 for BCI2000-Full).

Secondly, our proposed approach on a task-independent neural signature represents a more complex problem than that of task-specific identification via EEG. As highlighted in chapter 5, task-specific neural activity carries extra discriminant information, which systems from other architectures may potentially exploit, hence, obtaining higher overall performances. We will present in chapter 7 an argument in favour of our approach, regardless of its inherent increase in difficulty. In addition, on Keirn's and Yeom's data sets, sessions were also crossed between training and testing sets (*task-sess-CV*), a practice that was not generally followed in the literature.

The relatively lower performance observed in some cases when compared with other works, and the pronounced decline on GAR for lower values of FAR in all but BCI2000-Full and DEAP-Full data sets, may be due to several reasons:

Table 6.9: Comparison of results obtained with Keirn's database. For more results on this database, refer to tables 2.10 and 2.11. Refer to table 6.8 for a description of the table's columns. For completeness, classification results from table 5.8 are also shown, these correspond to the analysis of chapter 5.

Ref.	# Subj.	Task	CV method	Succ.
Classification experiments				
[49]	5	All on MTL	K-Folds; K = 2	100%
This	5	All	Open-set; *Task-Sess-CV*; K-Fold + MC; $K = 5$ + $NIt = 20$	88.55%
Verification experiments				
[48, 51]	5	Task 3	K-Fold; $K = 2$	100% (from EER)
[51]	5	Task 1	K-Fold; $K = 2$	80% (from EER)
This	5	All	Open-set; *Task-Sess-CV*; K-Fold + MC; $K = 2$ + $NIt = 20$	80.06%

Table 6.10: Comparison of results obtained with Yeoms's database. Refer to table 6.8 for a description of the table's columns. For completeness, classification results from table 5.8 are also shown, these correspond to the analysis of chapter 5.

Ref.	# Subj.	Task	CV method	Succ.
Classification experiments				
This	10	All	Open-set; *Task-Sess-CV*; K-Fold + MC; $K = 2$ + $NIt = 20$	52.96%
Verification experiments				
[109]	10	All	K-Fold; K = 10	86.10%
This	10	All	Open-set; *Task-Sess-CV*; K-Fold + MC; $K = 2$ + $NIt = 20$	74.73%

Table 6.11: Comparison of results obtained with Zhang's database. For more results on these database, refer to tables tables 2.6 to 2.8. Refer to table 6.8 for a description of the table's columns. For completeness, classification results from table A.30 are also shown, these correspond to the analysis of chapter 5.

Ref.	# Subj.	Task	CV method	Succ.
Classification experiments				
[37, 39]	20	All (ignored)	Leave-One-Out	100%
[34, 36, 40]	20	All (ignored)	K-Folds; K = 2	99.06%
[44]*	20	All (ignored)	K-Folds; K = 3	92.80%
This	20	All	*Task-CV*; K-Fold + MC; $K = 3$ + $NIt = 20$	75.90%
Verification experiments				
This	20	All	Open-set; *Task-CV*; K-Fold + MC; $K = 2$ + $NIt = 20$	93.24%

*: The pool of users contained 10 healthy subjects and 10 alcoholics.

Keirn's and Yeom's databases were tested under the *task-sess-CV* paradigm, which is arguably more difficult than the *task-CV* one. This is especially true when considering that a single session is used for training which in turn leaves the system prone to detecting the idiosyncrasies of that specific recording session. We can expect higher accuracy values when more sessions are available for training, allowing the system to reject session idiosyncrasies [43, 72, 92, 125].

Performance on Keirn's and P. Ullsperger's databases was certainly affected by the reduced number of available subjects (only 5 subjects). Following the described experimentation methodology, 1 subject is designated as the registered user and only 2 subjects are used as impostors for training, with the other 2 utilized for testing. With such a small number of impostors, we can expect the system's performance to be more sensitive to the idiosyncrasies of each CV partition.

The above should also be considered, to a lesser degree, for Yeom's database, which only has 10 subjects. In addition, this dataset contains EEG from two MZ

twins. How these subjects are distributed among the registered and impostors sets should be presumed to have a major impact on the test result.

Finally, in the case of Yeom's and Zhang's databases, the systems' performance was certainly compromised by the length of the EEG segments (only 1 second). This forced us to use a window length L_W of 0.5 seconds, as opposed to the optimal 1 or 2 seconds found in chapter 5.

6.6 Conclusions

In this chapter, we have evaluated multiple methods and system architectures in a bid to exploit the identity traits within the time-frequency representation of EEG activity. Overall, we have found that $\text{RCeps}_{20\%}$ and LPC_8 are the best performing characteristics across databases from those tested. Adding any extra feature and fusing $\text{RCeps}_{20\%}$ and LPC_8 in various ways had a neutral or negative effect on the system's accuracy, as did the application of non-linear classification algorithms such as ANN and RBF-SVM.

Specifically, we conclude that:

1. The information extracted by RCeps and LPC is highly correlated, as evidenced by the fact that fusing them had no effect on the system's performance.

2. The problem of verification based on RCeps or LPC features is a linear problem, as evidenced by the fact that the evaluated non-linear algorithms only equalled the performance of LDC.

3. RCeps and LPC can encode most of the discrimination power of the EEG's spectrum, as evidenced by the fact that all the systems tested performed worse than, or similarly to, those based solely on RCeps or LPC – including systems that fused different measurements of the spectral data.

The information coded in $\text{RCeps}_{20\%}$ and LPC_8 is remarkably discriminant and robust. Hence, to improve the design of the overall system, it may be more productive to search for the required extra discrimination power in sources other than the spectral shape of individual sensors, or to apply a more sophisticated preprocessing stage.

We may also conclude that the observed errors are probably not due to weaknesses in these features, but mainly due to intrinsic properties of the data, such as the number of subjects available or perturbations of the EEG signal. We suggest that this should be solved at the recording stage in a live system, for example, requiring 5 seconds of clean EEG signal to perform the verification.

Chapter 7

Discussion, advances and future steps

> "It's much more interesting to live not knowing than to have answers which might be wrong."
>
> Richard Feynman

In this final chapter, we will discuss the obtained results and give arguments in favour of the proposed approach to the problem of EEG biometric identification. We will also highlight the advances in the field driven by this work. Finally, we will provide some possible future research steps in line with the obtained results.

7.1 Discussion on the task-independent neural signature

Throughout this research, we have provided evidence supporting the hypothesis presented in section 1.1, i.e.:

There exists, within human EEG activity, a task-independent pattern concomitant to the individual's identity.

In addition, we have used this property to successfully build a system capable of differentiating individuals via their EEG activity, irrespective of cognitive state. Such a system would not have succeed without the existence of task-independent, discriminant information.

Note that this research is fundamentally different to previously published neurophysiological and biometric studies. Let us consider iris and gait recognition systems as analogous examples. While the former uses properties 'inherent'[1] to the individual, the latter focuses on how the individual performs a given task. Hence, we could argue that task-specific studies are closer in nature to describing idiosyncratic activity during cognitive processing. In contrast, our research on task-independent characteristics, tries to describe identity in and of itself. Having said that, it is very probable that previous studies have also been processing, inadvertently, part of the task-independent signature.

The origin of this neural signature is unclear. Given that it is independent of the recorded condition, we may associate it with unconscious processes working uninterruptedly in the background, similar to M.E. Raichle's et. al. concept of a *'default mode of brain function'* [156–158]. Such concept hypotheses the existence of an intrinsic activity which *"instantiates the maintenance of information for interpreting, responding to and even predicting environmental demands"*. The default mode is not completely suppressed when the subject engages in a task, but rather attenuated: *"Its functions are spontaneous and virtually continuous, being attenuated only when we engage in goal-directed actions"* [157]. This is in line with the observation that the presented neural signature is a dynamic pattern across conditions. The task-independence property arises from the fact that such

[1]Strictly speaking, irises are not inherent parts of an individual, as a subject can still be without them.

fluctuations have a smaller magnitude than differences across subjects. Hence, we can interpret these fluctuations as task specific activity superimposing the default mode. Interestingly, when talking about the continuity of the default mode, M.E. Raichle et. al. noticed its relationship with the individual's 'self': *"This is consistent with the continuity of a stable, unified perspective of the organism relative to its environment (a 'self')"* [157].

Another possible explanation is that the signature is purely due to the structure of the neural networks, and has nothing to do with their underlying cognitive processes. Given the nature of the electrical fields and their propagation through the skull, two networks with identical functionality but different organization will produce different EEG signals. Accordingly, the EEG neural signature would be broadly defined by the disposition of networks within the brain, with cognitive processes playing a modulating role.

As is typically the case, the solution is probably neither of the above propositions, but a combination of both.

7.2 Discussion on the task-independent EEG biometric approach

To date, the approach followed in the literature has been based on the analysis of EEG from isolated conditions. Even when the system was fed with signals from multiple tasks, they were usually labelled with the task itself, so that systems could differentiate amongst them and exploit task-specific information (MTL).

In a recent publication, Campisi's team expanded on the idea of an acquisition protocol, where users were asked to perform a particular task while their EEG was recorded for identification/verification [11]. Specifically, they have focussed their efforts mainly on REC and REO conditions [64, 81, 82, 111]:

Within this paradigm subjects are typically seated in a comfortable

> *chair with both arms resting, in a dimly lit or completely dark room. Generally, external sounds and noise are minimized to favour the relaxed state of the subjects. Participants are asked to perform a few minutes of resting state with eyes closed or eyes open, avoiding any focusing or concentration, but staying awake and alert.* ([11])

T. Pham's et. al. went a step further and proposed a system which assigns specific tasks to groups of subjects [118]. By identifying the performed task during verification, they can effectively reduce the problem's complexity by a factor of N, where N is the number of considered tasks.

It is undeniable that performing these tasks during the verification of a user's identity is, in many real scenarios, cumbersome. For example, if this modality is to be integrated with the biometric passport, performing *"a simple 'resting state' protocol"* (as described by Campisi's team [11]) is completely impractical. In addition, this biometry will almost certainly find an application within other BCIs, which originally intended for different purposes other than subject identification. For example, a recent patent issued by Google embeds the identification of the user within a multi-sensor diagnostic system [110].

To overcome these difficulties, we have proposed a completely new approach where subjects are not asked to perform any specific task. Instead, the system is tuned to extract the subject's task-independent neural signature. This will leave the subject free to perform any other operations. For example, in an airport border control, the EEG activity could be collected while the users present their passport, introduce any required information, and/or provide any other biometry. In a general purpose BCI, the verification of the user's identity could take place in the background invisibly. As a result, the security procedure will not interfere with the user experience of the system. Moreover, verification checks could happen continuously or periodically, again, without interfering in the operation of the device.

7.3 Generated knowledge

In this work, we have researched the state of the art in EEG identification biometry and, to a lesser degree, equivalent genetic and neurophysiological studies (chapter 2). When published in May 2014 [159], this research was the first of its kind [2]. In it, we identified a general lack of consensus on several fundamental questions, such as the frequency and space distribution of the discriminant information, and the conditions to optimise its extraction.

In a bid to shed light on those questions, we conducted an in depth analysis of the time-frequency representation of the EEG (chapter 5). Such study presents a number of innovations with respect to previously published research:

- We proposed a visual representation of the data, based on the concatenation of EEG spectrograms, and provided visual evidence of the existence of subject traits within the EEG PSD.
- Through such representation, we ran qualitative analyses complementing the quantitative ones. This approach proved to be vital for the interpretation of the results.
- We performed the experiments on 6 databases, ultimately divided into 10 different data sets of varying characteristics. These included resting stages, motor real and imagery tasks, problem solving tasks, VEPs, and AEPs. Such variability allowed us to differentiate between idiosyncratic and common EEG behaviours and to extract robust conclusions on the general properties of the neural signature.
- In addition, one of the databases used was composed of AEPs. On previous biometric studies, VEPs were the only ERPs used.
- We executed the analysis on a carefully designed incremental methodology. We assessed each parameter individually to understand its effects on the

[2] Campisi's team published a similar work around the same date [11]

problem. Whenever possible, we used three versions of each data set: a raw version and two artefact free versions processed with automatic tools.

- We identified the optimal conditions to extract the subject-specific information within the time-frequency representation of the EEG data through the STFT.

Again, when published in 2015, this was the first comprehensive analysis of the properties of the subject traits within the time-frequency representation of the EEG [160].

Next, we evaluated the proposed hypothesis through two dedicated experiments (section 5.5). As a result, ***we have provided, for the first time, evidence of the existence of a task-independent neural signature***. This suggests that neural activity, as recorded via EEG, is primarily determined by the subjects identity rather than by the performed tasks. We anticipate this finding will open a new set of experimental possibilities within brain research fields. Genetic and neurophysiological studies could use this neural signature to further the understanding of the EEG inheritance model, differentiating between task-independent and task-dependent activity. Through targeted experiments, it could also be used in the study of neuroplasticity, neurodegenerative diseases or the default mode of brain function. Overall, it represents a step forward in the understanding of the individual differences in brain activity, which will, in turn, help in the understanding of the commonalities.

Finally, we applied our findings to the design of a task-independent EEG biometric verification system (chapter 6). We identified $\text{RCeps}_{20\%}$ and LPC_8 as two very powerful descriptors capable of coding most of the discriminant information within the spectral shape of the EEG signal. Although LPC had been a popular choice within the literature, this is the first work to use RCeps[3]. Moreover, thanks to our previous analysis of the EEG data, we were also able to explain in detail

[3]P. Nguyen et. al. used mel-cepstrums instead of real-cepstrums [44].

why these two features perform so well and to understand its behaviour across different configurations.

In addition to the above innovations, we have also presented here a new methodology for the rejection of EEG artifacts: LCF (chapter 4). This was developed at the Psychology Department of the College of Human Health and Science, Swansea University (Wales, UK). This methodology was applied to obtain an artefact free version of the data sets for subsequent experiments.

7.4 Future research

The main weakness of the presented research and, for that matter, the related studies in the literature, is the size and nature of the databases used for experimentation. In addition, this research has also been limited by the number of conditions within each database. Specifically, we observed the task-independent signature across motor tasks, relaxation and resting states (BCI2000-Full), emotional states and REO (DEAP-Full), problem-solving tasks and relaxation (Keirn's), synonyms and non-synonyms AEPs (P. Ulssperger's), self and non-self VEP (Yeom's) and target and non-target VEP (Zhang's). We suggest that focus should be placed on collecting a database with: 100+ subjects, 20+ sessions during a minimum time span of 1 year, and 20+ recording conditions – including relaxation states, motor and problem-solving tasks and various ERPs. This will, without doubt, represent a major step forward in the understanding of the neural signature.

At the same time, different sources of discriminant information within EEG activity should be explored and combined with the spectral shape information described here. Some of these sources have already been proposed in the literature, such as blink signals [161] or inter-sensor spectral coherence [64]. However, they will benefit from a description to the level of that provided in the current study.

A multidisciplinary study of the genetics and neurophysiology of the presented task-independent signature, carried out on a database composed of families as

well as MZ and DZ twins, would be valuable to determine the extent to which this neural signature is defined by nature and nurture. This will also help in understanding the origin of such a pattern. From a biometric point of view, it will be vital for the discrimination of twins, to identify non-genetic traits.

For studies that focus on the identification of EEG activity commonalities across clinical patients, the reported evidence of a neural signature may suggest new ways of analysing and/or interpreting their data. For example, a condition may have effects idiosyncratic for each individual, that may be more easily identifiable as changes or disruptions in the subject's neural signature.

Overall, we expect the task-independence property of the reported neural signature to be key for the study, design and implementation of future BCIs.

Appendix A

Supplementary results and statistical tests

To avoid cluttering the text with figures and tables, we only inserted the most representative examples. Here, we provide supplementary results and statistical tests. A description of the figures and tables below, beyond that of their captions, can be found in the corresponding section of the main text.

On cases were statistical tests were performed, we always ran targeted independent t-tests. To adjust the p-value for the typical increase of Type I error in multiple testing, we applied the Benjamini-Hochberg False Discovery Rate (BHFDR) method across all the obtained values. This is, within an experiment (e.g. EEG montage evaluation), we simultaneously adjusted all the p-values obtained with all the considered databases, systems and conditions. This approach was preferred over analysis of variance (ANOVA) due to the high number of variables involved in the experiments and the strong n-way interactions exhibited – especially between the variables "database", "database mode" (raw or artefact free) and "system" –, which hindered the interpretation of the results.

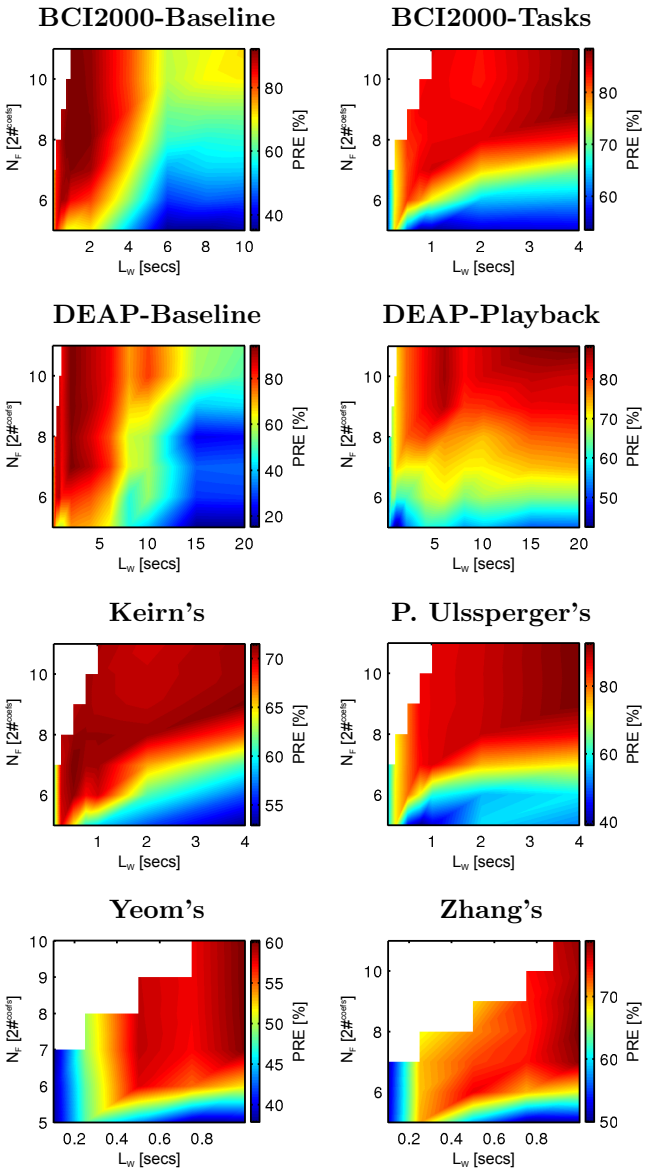


Figure A.1: Mean PRE results obtained with each database on an L_W vs N_F grid. L_G was set to the length of the available EEG on each database (table 3.2) and $\Theta = 0\%$. Results correspond to *full-fusion no-focus* experiments. A maximum of 20 subjects was used in each experimental iteration. Refer to section 5.2.1 for a description of the results.

Table A.1: Independent t-test analysis of the $N_F \approx L_W * F_s$ diagonal. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) > \mu_{PRE}(B)$. With A and B representing the conditions above and below the diagonal $N_F \approx L_W * F_s$ respectively. Refer to table 3.2 for details on databases' code names, and to section 5.2.1 for a description of the results.

(a) Results with raw data sets.

Dat.	**t**	**df**	**p-value**	**SE**	**r**
BB	20.65	598	**< 0.001**	1.05	1.00
BT	13.39	388	**< 0.001**	1.06	1.00
DB	18.93	738	**< 0.001**	1.27	1.00
DP	2.07	738	**< 0.05**	0.86	1.00
K	18.13	388	**< 0.001**	0.41	1.00
P	17.16	388	**< 0.001**	1.24	1.00
Y	6.66	238	**< 0.001**	1.19	1.00
Z	7.80	248	**< 0.001**	1.44	1.00

(b) Results with ADJUST processed data sets.

Dat.	**t**	**df**	**p-value**	**SE**	**r**
BB	17.75	598	**< 0.001**	1.07	1.00
BT	10.52	388	**< 0.001**	0.71	1.00
DB	16.80	738	**< 0.001**	1.35	1.00
DP	7.33	738	**< 0.001**	0.60	1.00
P	12.23	388	**< 0.001**	0.83	1.00
Z	8.95	248	**< 0.001**	1.39	1.00

(c) Results with LCF processed data sets.

Dat.	**t**	**df**	**p-value**	**SE**	**r**
BB	16.86	598	**< 0.001**	1.15	1.00
BT	11.24	388	**< 0.001**	0.83	1.00
DB	18.73	738	**< 0.001**	1.39	1.00
DP	7.15	738	**< 0.001**	0.55	1.00
P	9.11	388	**< 0.001**	0.90	1.00
Z	9.11	248	**< 0.001**	1.48	1.00

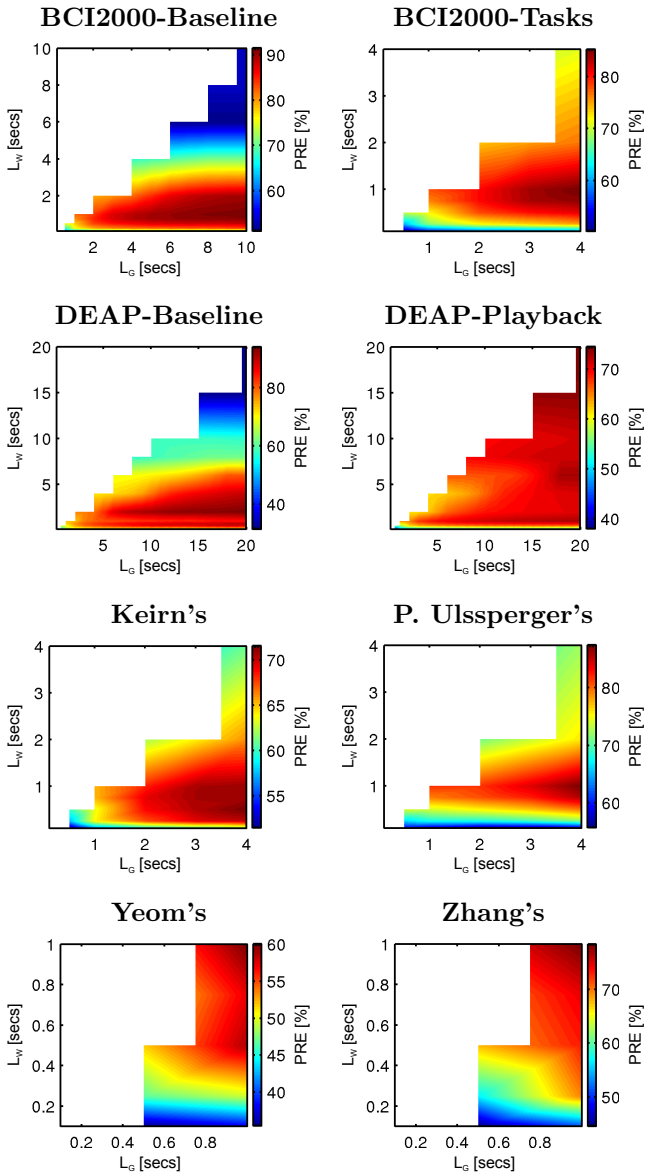


Figure A.2: Mean PRE results obtained with each database on an L_G vs L_W grid. In all cases, $N_F = 128$ coefficients and $\Theta = 0\%$. Results correspond to the *full-fusion no-focus* experiments. A maximum of 20 subjects was used in each experimental iteration. Refer to section 5.2.1 for a description of the results.

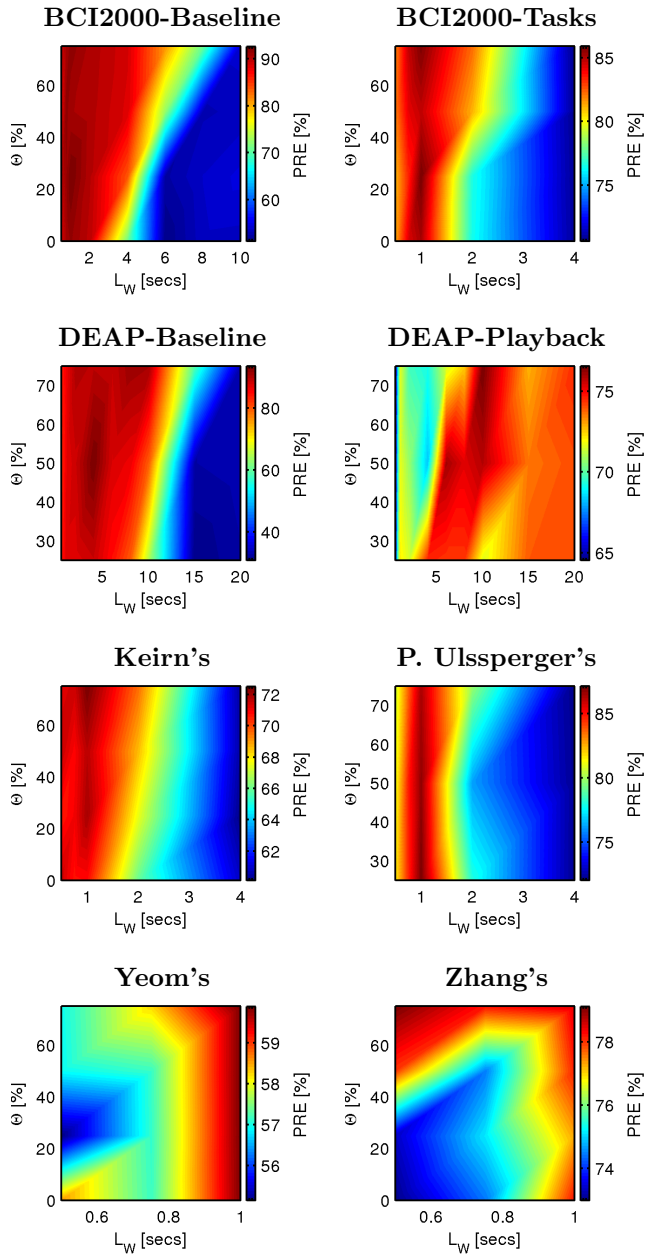


Figure A.3: μPRE results obtained with each database on an L_W vs Θ grid. L_G was set to the length of the available EEG on each database (table 3.2) and $N_F = 128$ coefficients. Results correspond to *full-fusion no-focus* experiments. A maximum of 20 subjects was used in each experimental iteration. Refer to section 5.2.1 for a description of the results.

Table A.2: Independent t-test analysis of F_{max} and F_{min} with *Conf-HalfLen* system. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B conditions specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.2.3 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	$(F_{max} = 30Hz) - (F_{max} = 20Hz)$					$(F_{max} = 40Hz) - (F_{max} = 30Hz)$				
BB	3.75	18	**< 0.01**	1.39	0.64	-0.05	18	0.97	1.14	-0.01
BT	3.99	18	**< 0.01**	1.53	0.67	0.42	18	0.71	1.28	0.09
DB	12.02	18	**< 0.001**	0.52	0.94	-8.66	18	**< 0.001**	0.49	-0.89
DP	-4.85	18	**< 0.001**	0.08	-0.74	-22.14	18	**< 0.001**	0.12	-0.98
K	10.31	18	**< 0.001**	0.64	0.92	0.84	18	0.44	0.72	0.18
P	77.62	18	**< 0.001**	0.07	1.00	15.74	18	**< 0.001**	0.07	0.96
Y	123.21	18	**< 0.001**	0.02	1.00	232.71	18	**< 0.001**	0.02	1.00
Z	26.22	18	**< 0.001**	0.74	0.99	8.44	18	**< 0.001**	1.01	0.88
	$(F_{min} = 10Hz) - (F_{min} = 20Hz)$									
BB	2.63	18	**< 0.05**	1.78	0.51					
BT	2.79	18	**< 0.05**	1.95	0.53					
DB	14.14	18	**< 0.001**	0.57	0.95					
DP	53.59	18	**< 0.001**	0.16	1.00					
K	15.17	18	**< 0.001**	0.59	0.96					
P	48.32	18	**< 0.001**	0.13	1.00					
Y	191.48	18	**< 0.001**	0.03	1.00					
Z	3.33	18	**< 0.01**	1.83	0.60					

Table A.3: Independent t-test analysis of F_{max} and F_{min} with *Conf-FullLen* system. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B conditions specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.2.3 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	$(F_{max} = 30Hz) - (F_{max} = 20Hz)$					$(F_{max} = 40Hz) - (F_{max} = 30Hz)$				
BB	3.78	18	**< 0.01**	2.31	0.65	-0.04	18	0.97	2.59	-0.01
BT	2.57	18	**< 0.05**	2.36	0.50	1.11	18	0.31	2.53	0.24
DB	9.42	18	**< 0.001**	0.95	0.90	-1.50	18	0.17	1.12	-0.32
DP	11.21	18	**< 0.001**	0.11	0.93	-14.68	18	**< 0.001**	0.11	-0.96
K	21.00	18	**< 0.001**	0.28	0.98	2.38	18	**< 0.05**	0.33	0.47
P	33.49	18	**< 0.001**	0.09	0.99	4.49	18	**< 0.001**	0.06	0.71
Y	-8.85	18	**< 0.001**	0.02	-0.89	146.21	18	**< 0.001**	0.02	1.00
Z	6.11	18	**< 0.001**	2.51	0.81	3.84	18	**< 0.01**	2.21	0.65
	$(F_{min} = 10Hz) - (F_{min} = 20Hz)$									
BB	4.71	18	**< 0.001**	1.20	0.73					
BT	6.14	18	**< 0.001**	1.49	0.81					
DB	26.76	18	**< 0.001**	0.99	0.99					
DP	62.45	18	**< 0.001**	0.14	1.00					
K	19.48	18	**< 0.001**	0.42	0.97					
P	44.72	18	**< 0.001**	0.14	1.00					
Y	232.57	18	**< 0.001**	0.03	1.00					
Z	3.41	18	**< 0.01**	2.03	0.61					

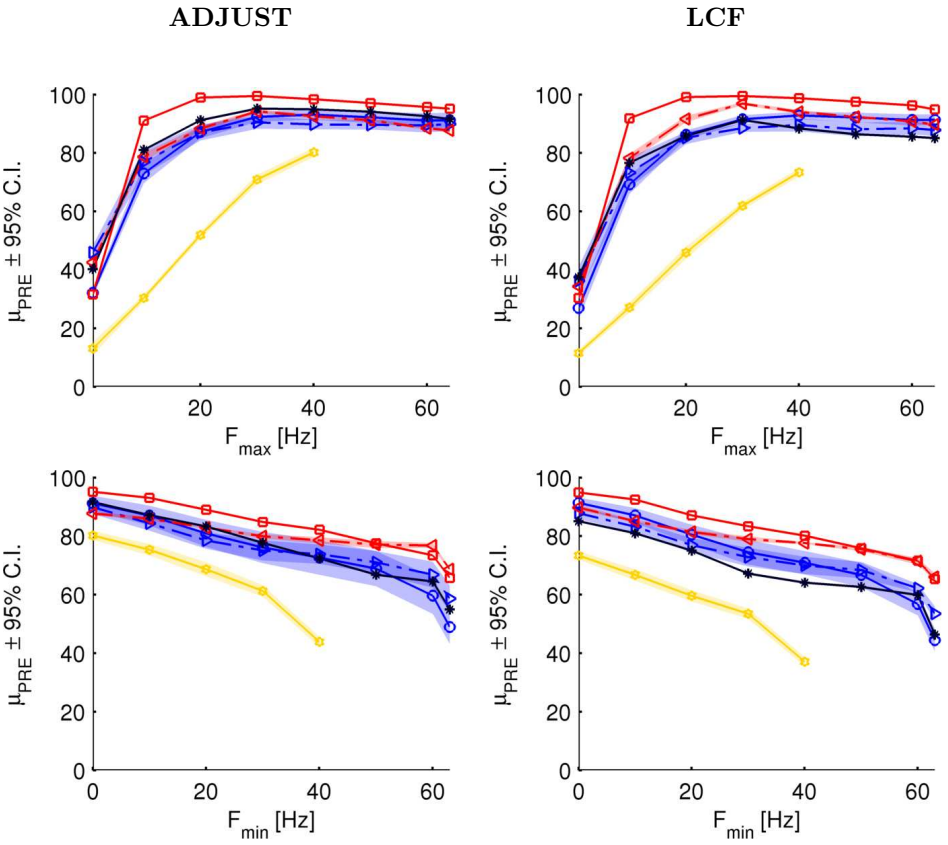


Figure A.4: Quantitative analysis of the maximum (F_{max}) and minimum (F_{min}) cut-off frequencies with artefact free data sets. Mean and 95% CI of PRE results obtained in *full-fusion no-focus* experiments with different F_{max} (top) and F_{min} (bottom) values. Results correspond to ADJUST (left) and LCF (right) pre-processed databases, tested with the *Conf-HalfLen* system. A maximum of 20 subjects was used in each experimental iteration. Refer to figure 5.5 for details on the legend, and to section 5.2.3 for a description of the results.

Table A.4: Independent t-test analysis of F_{max} and F_{min} with ADJUST processed data sets. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B conditions specified on top of each cell as $(A) - (B)$. Refer to section 5.2.3 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	$(F_{max} = 30Hz) - (F_{max} = 20Hz)$					$(F_{max} = 40Hz) - (F_{max} = 30Hz)$				
BB	2.23	18	$<$ **0.05**	1.55	0.45	-0.56	18	0.61	1.23	-0.12
BT	3.65	18	$<$ **0.01**	1.50	0.63	1.08	18	0.32	1.17	0.24
DB	12.74	18	$<$ **0.001**	0.46	0.94	-3.87	18	$<$ **0.01**	0.44	-0.65
DP	8.61	18	$<$ **0.001**	0.06	0.89	-23.23	18	$<$ **0.001**	0.05	-0.98
P	70.43	18	$<$ **0.001**	0.06	1.00	-4.47	18	$<$ **0.001**	0.06	-0.71
Z	20.87	18	$<$ **0.001**	0.91	0.98	9.14	18	$<$ **0.001**	1.02	0.90
	$(F_{min} = 10Hz) - (F_{min} = 20Hz)$									
BB	3.48	18	$<$ **0.01**	1.68	0.61					
BT	2.73	18	$<$ **0.05**	2.08	0.52					
DB	9.44	18	$<$ **0.001**	0.32	0.90					
DP	32.97	18	$<$ **0.001**	0.12	0.99					
P	29.35	18	$<$ **0.001**	0.13	0.99					
Z	4.82	18	$<$ **0.001**	1.38	0.73					

Table A.5: Independent t-test analysis of F_{max} and F_{min} with *Conf-FullLen* system and ADJUST processed data sets. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B conditions specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.2.3 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	$(F_{max} = 30Hz) - (F_{max} = 20Hz)$					$(F_{max} = 40Hz) - (F_{max} = 30Hz)$				
BB	2.58	18	**< 0.05**	2.63	0.50	0.04	18	0.97	2.85	0.01
BT	4.23	18	**< 0.001**	2.16	0.69	0.77	18	0.48	2.46	0.17
DB	12.31	18	**< 0.001**	1.00	0.94	3.31	18	**< 0.01**	1.05	0.60
DP	11.92	18	**< 0.001**	0.08	0.94	-26.73	18	**< 0.001**	0.04	-0.99
P	41.59	18	**< 0.001**	0.07	0.99	-18.49	18	**< 0.001**	0.06	-0.97
Z	11.71	18	**< 0.001**	1.29	0.93	6.96	18	**< 0.001**	1.17	0.84
	$(F_{min} = 10Hz) - (F_{min} = 20Hz)$									
BB	3.07	18	**< 0.01**	2.28	0.57					
BT	2.56	18	**< 0.05**	2.57	0.50					
DB	15.76	18	**< 0.001**	0.89	0.96					
DP	27.13	18	**< 0.001**	0.13	0.99					
P	46.62	18	**< 0.001**	0.08	1.00					
Z	8.56	18	**< 0.001**	0.93	0.89					

Table A.6: Independent t-test analysis of F_{max} and F_{min} with *Conf-HalfLen* system and LCF processed data sets. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B conditions specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.2.3 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	$(F_{max} = 30Hz) - (F_{max} = 20Hz)$					$(F_{max} = 40Hz) - (F_{max} = 30Hz)$				
BB	2.96	18	**< 0.05**	1.12	0.55	0.93	18	0.39	1.13	0.20
BT	4.72	18	**< 0.001**	1.07	0.73	1.16	18	0.29	1.08	0.25
DB	9.68	18	**< 0.001**	0.54	0.91	-7.36	18	**< 0.001**	0.39	-0.85
DP	7.33	18	**< 0.001**	0.05	0.85	-11.95	18	**< 0.001**	0.06	-0.94
P	63.38	18	**< 0.001**	0.09	1.00	-37.98	18	**< 0.001**	0.08	-0.99
Z	14.37	18	**< 0.001**	1.11	0.95	13.26	18	**< 0.001**	0.86	0.95
	$(F_{min} = 10Hz) - (F_{min} = 20Hz)$									
BB	4.54	18	**< 0.001**	1.39	0.71					
BT	2.55	18	**< 0.05**	1.90	0.49					
DB	9.71	18	**< 0.001**	0.38	0.91					
DP	77.13	18	**< 0.001**	0.07	1.00					
P	35.04	18	**< 0.001**	0.17	0.99					
Z	5.85	18	**< 0.001**	1.22	0.79					

Table A.7: Independent t-test analysis of F_{max} and F_{min} with *Conf-FullLen* system and LCF processed data sets. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B conditions specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.2.3 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	$(F_{max} = 30Hz) - (F_{max} = 20Hz)$					$(F_{max} = 40Hz) - (F_{max} = 30Hz)$				
BB	3.30	18	$<$ **0.01**	2.12	0.59	-0.05	18	0.97	2.08	-0.01
BT	2.94	18	$<$ **0.05**	2.86	0.55	0.59	18	0.59	2.69	0.13
DB	8.95	18	$<$ **0.001**	1.12	0.89	5.26	18	$<$ **0.001**	0.88	0.76
DP	10.67	18	$<$ **0.001**	0.07	0.92	-4.97	18	$<$ **0.001**	0.05	-0.74
P	48.59	18	$<$ **0.001**	0.08	1.00	-18.14	18	$<$ **0.001**	0.08	-0.97
Z	9.44	18	$<$ **0.001**	1.35	0.90	7.02	18	$<$ **0.001**	1.25	0.84
	$(F_{min} = 10Hz) - (F_{min} = 20Hz)$									
BB	2.69	18	$<$ **0.05**	2.66	0.52					
BT	4.77	18	$<$ **0.001**	2.00	0.73					
DB	10.85	18	$<$ **0.001**	0.77	0.92					
DP	34.39	18	$<$ **0.001**	0.12	0.99					
P	49.77	18	$<$ **0.001**	0.11	1.00					
Z	5.66	18	$<$ **0.001**	1.23	0.78					

Table A.8: Independent t-test analysis of L_W. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B conditions specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.2.4 for a description of the results.

(a) Results with raw data sets.

Dat.	t	df	p-value	SE	r	t	df	p-value	SE	r
	$(L_W = 1s) - (L_W = 0.5s)$					$(L_W = 2s) - (L_W = 1s)$				
BB	3.36	18	**< 0.01**	1.58	0.60	0.21	18	0.84	1.01	0.05
BT	3.09	18	**< 0.05**	1.68	0.57	1.09	18	0.35	1.39	0.24
DB	14.32	18	**< 0.001**	0.51	0.95	4.12	18	**< 0.01**	0.43	0.68
DP	54.50	18	**< 0.001**	0.16	1.00	40.22	18	**< 0.001**	0.17	0.99
K	14.80	18	**< 0.001**	0.10	0.96	2.15	18	0.08	0.34	0.43
P	31.59	18	**< 0.001**	0.07	0.99	16.51	18	**< 0.001**	0.06	0.97
Y	198.11	18	**< 0.001**	0.02	1.00	-	-	-	-	-
Z	-2.00	18	0.09	0.85	-0.41	-	-	-	-	-

(b) Results with ADJUST processed data sets.

Dat.	t	df	p-value	SE	r	t	df	p-value	SE	r
	$(L_W = 1s) - (L_W = 0.5s)$					$(L_W = 2s) - (L_W = 1s)$				
BB	1.96	18	0.09	1.21	0.40	-1.33	18	0.25	0.67	-0.29
BT	2.53	18	**< 0.05**	1.00	0.49	0.56	18	0.64	0.92	0.12
DB	5.84	18	**< 0.001**	0.45	0.79	-0.97	18	0.40	0.44	-0.21
DP	39.95	18	**< 0.001**	0.08	0.99	16.18	18	**< 0.001**	0.09	0.96
P	21.10	18	**< 0.001**	0.17	0.98	12.40	18	**< 0.001**	0.08	0.94
Z	-0.24	18	0.83	1.31	-0.05	-	-	-	-	-

(c) Results with LCF processed data sets.

Dat.	t	df	p-value	SE	r	t	df	p-value	SE	r
	$(L_W = 1s) - (L_W = 0.5s)$					$(L_W = 2s) - (L_W = 1s)$				
BB	1.76	18	0.13	1.49	0.37	1.50	18	0.20	0.91	0.32
BT	2.08	18	0.08	1.09	0.42	-0.40	18	0.74	0.81	-0.09
DB	10.29	18	**< 0.001**	0.24	0.92	0.63	18	0.60	0.33	0.14
DP	43.08	18	**< 0.001**	0.08	0.99	32.00	18	**< 0.001**	0.05	0.99
P	36.37	18	**< 0.001**	0.16	0.99	22.71	18	**< 0.001**	0.12	0.98
Z	-1.43	18	0.22	1.01	-0.31	-	-	-	-	-

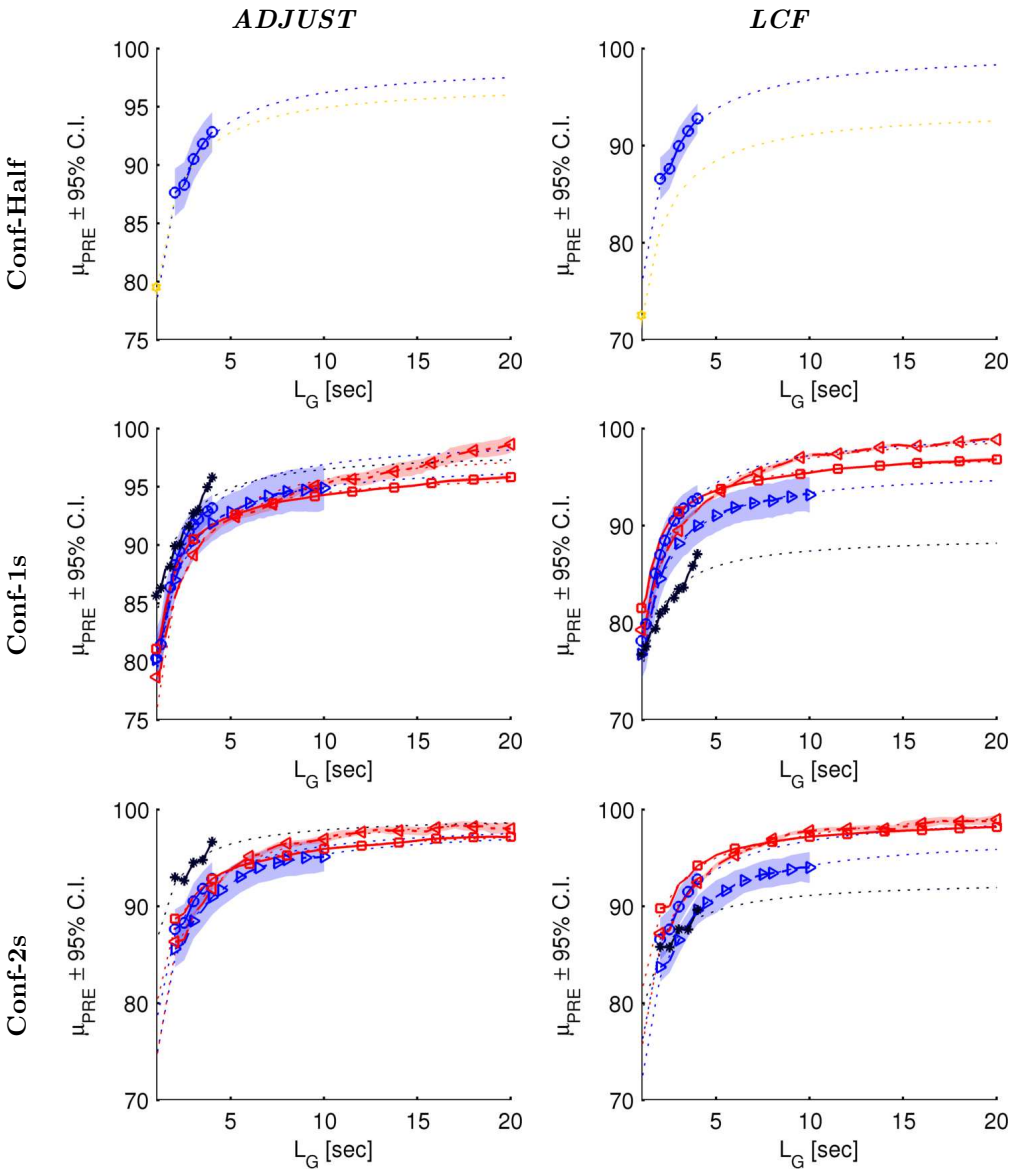


Figure A.5: Quantitative analysis of the EEG segment length (L_G) with artefact free data sets. Mean PRE and 95% CI (shaded area) obtained with different L_G. Data corresponds to *full-fusion* set-up on *no-focus* experiments using *Conf-1s* (left) and *Conf-2s* (right) systems. A maximum of 20 subjects was used in each experimental iteration. Refer to figure 5.7 for details on the legend, and to section 5.2.5 for a description of the results.

Table A.9: Quantitative results for different configurations with ADJUST processed data sets. Mean PRE and 95% CI obtained with different configurations. Data corresponds *full-fusion no-focus* experiments. Within data sets, performances statistically different than the maximum are pointed by * (single tail t-tests with BHFDR adjusted $p > 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 5.2.6 for a description of the results.

Conf	**BB**	**BT**	**DB**	**DP**
Conf-FullLen	70.74*	68.26*	70.42*	**98.06**
	[67.62, 73.85]	[63.24, 73.29]	[68.28, 72.56]	**[97.96, 98.16]**
Conf-1s	94.89	**94.37**	**98.63**	95.83*
	[93.02, 96.77]	**[93.16, 95.57]**	**[97.92, 99.35]**	[95.66, 96.01]
Conf-2s	**95.11**	89.32*	98.00	97.16*
	[93.62, 96.59]	[86.69, 91.94]	[97.34, 98.66]	[97.05, 97.27]

Conf	**K**	**P**	**Y**	**Z**
Conf-HalfLen	-	-	-	**79.51**
	-	-	-	**[78.31, 80.71]**
Conf-FullLen	-	96.01*	-	78.30
	-	[95.90, 96.13]	-	[76.93, 79.67]
Conf-1s	-	95.79*	-	-
	-	[95.63, 95.95]	-	-
Conf-2s	-	**96.65**	-	-
	-	**[96.61, 96.69]**	-	-

Table A.10: Quantitative results for different configurations with LCF processed data sets. Mean PRE and 95% CI obtained with different configurations. Data corresponds to *full-fusion no-focus* experiments. Within data sets, performances statistically different than the maximum are pointed by * (single tail t-tests with BHFDR adjusted $p > 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 5.2.6 for a description of the results.

Conf	**BB**	**BT**	**DB**	**DP**
Conf-FullLen	67.26*	65.74*	68.42*	**98.70**
	[63.23, 71.29]	[59.82, 71.65]	[67.30, 69.54]	**[98.55, 98.85]**
Conf-1s	93.16	**93.47**	98.84	96.80*
	[91.32, 94.99]	**[92.66, 94.29]**	[98.60, 99.08]	[96.58, 97.03]
Conf-2s	**94.00**	89.79*	**98.95**	98.18*
	[92.40, 95.60]	[87.73, 91.85]	**[98.45, 99.45]**	[98.10, 98.26]

Conf	**K**	**P**	**Y**	**Z**
Conf-HalfLen	-	-	-	**72.53**
	-	-	-	**[71.42, 73.64]**
Conf-FullLen	-	**92.24**	-	71.94
	-	**[92.13, 92.35]**	-	[70.54, 73.33]
Conf-1s	-	87.06*	-	-
	-	[86.88, 87.24]	-	-
Conf-2s	-	89.63*	-	-
	-	[89.54, 89.71]	-	-

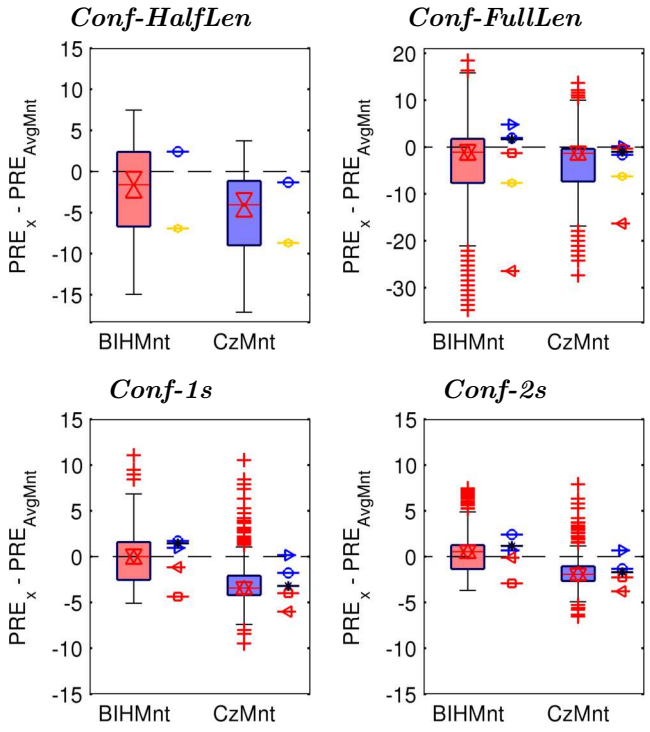


Figure A.6: Quantitative analysis of EEG montages with ADJUST processed data sets. Relative PRE values between *BIHMnt-AvgMnt* and between *CzMnt-AvgMnt*. Results are stacked across databases. Refer to caption of figure 5.9 for details about the meaning of symbols within the image, and to section 5.3.1 for a description of the results.

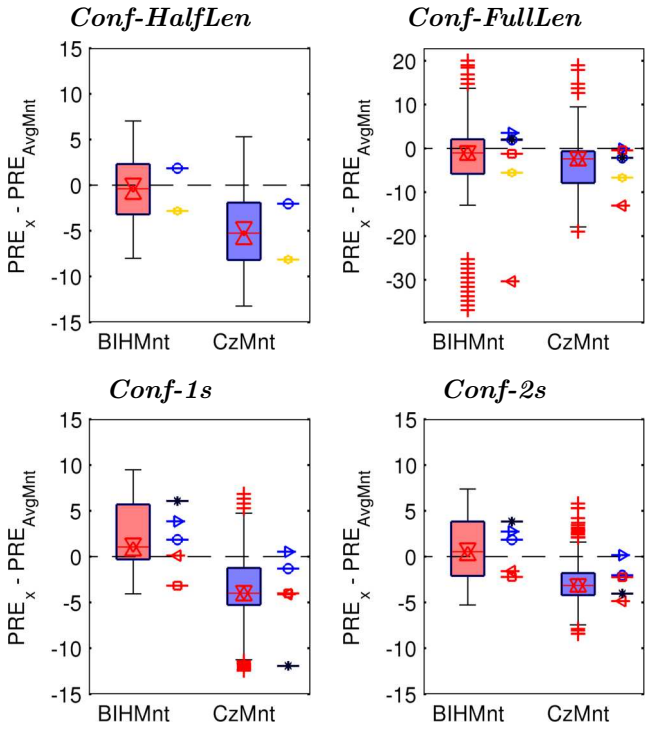


Figure A.7: Quantitative analysis of EEG montages with LCF processed data sets. Relative PRE values between *BIHMnt-AvgMnt* and between *CzMnt-AvgMnt*. Results are stacked across databases. Refer to caption of figure 5.9 for details about the meaning of symbols within the image, and to section 5.3.1 for a description of the results.

Table A.11: Independent t-test analysis of EEG montages with *Conf-HalfLen* system. H_0: $\mu_{PRE}(A) - \mu_{PRE}(B) \geq 10$. H_1: $\mu_{PRE}(A) - \mu_{PRE}(B) < 10$. With A and B montages specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.1 for a description of the results.

(a) Results with raw data sets.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(BIHMnt) - (AvgMnt)					(CzMnt) - (AvgMnt)				
BT	5.94	18	**< 0.001**	1.92	0.80	4.51	18	**< 0.001**	1.99	0.71
K	13.09	18	**< 0.001**	0.52	0.95	-	-	-	-	-
Y	137.44	18	**< 0.001**	0.02	1.00	703.62	18	**< 0.001**	0.02	1.00
Z	1.14	18	0.15	1.91	0.25	-2.10	18	1.00	1.89	-0.42

(b) Results with ADJUST processed data sets.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(BIHMnt) - (AvgMnt)					(CzMnt) - (AvgMnt)				
BT	8.59	18	**< 0.001**	1.40	0.89	5.81	18	**< 0.001**	1.45	0.79
Z	2.69	18	**< 0.01**	1.15	0.51	0.93	18	0.20	1.41	0.20

(c) Results with LCF processed data sets.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(BIHMnt) - (AvgMnt)					(CzMnt) - (AvgMnt)				
BT	8.21	18	**< 0.001**	1.41	0.88	6.67	18	**< 0.001**	1.09	0.83
Z	8.16	18	**< 0.001**	0.88	0.88	2.21	18	**< 0.05**	0.84	0.44

Table A.12: Independent t-test analysis of EEG montages with *Conf-FullLen* system. H_0: $\mu_{PRE}(A) - \mu_{PRE}(B) \geq 10$. H_1: $\mu_{PRE}(A) - \mu_{PRE}(B) < 10$. With A and B montages specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.1 for a description of the results.

(a) Results with raw data sets.

Dat.	t	df	p-value	SE	r	t	df	p-value	SE	r
	(BIHMnt) - (AvgMnt)					(CzMnt) - (AvgMnt)				
BB	6.64	18	**< 0.001**	2.36	0.83	4.12	18	**< 0.001**	2.50	0.68
BT	4.38	18	**< 0.001**	1.79	0.70	3.41	18	**< 0.01**	2.32	0.61
DB	-16.43	18	1.00	1.02	-0.96	1.58	18	0.07	0.93	0.33
DP	50.58	18	**< 0.001**	0.14	1.00	70.95	18	**< 0.001**	0.13	1.00
K	17.27	18	**< 0.001**	0.40	0.97	-	-	-	-	-
P	144.91	18	**< 0.001**	0.06	1.00	165.98	18	**< 0.001**	0.06	1.00
Y	59.51	18	**< 0.001**	0.02	1.00	493.04	18	**< 0.001**	0.02	1.00
Z	1.68	18	0.06	1.50	0.35	0.22	18	0.45	1.28	0.05

(b) Results with ADJUST processed data sets.

Dat.	t	df	p-value	SE	r	t	df	p-value	SE	r
	(BIHMnt) - (AvgMnt)					(CzMnt) - (AvgMnt)				
BB	6.16	18	**< 0.001**	2.40	0.81	4.62	18	**< 0.001**	2.20	0.72
BT	4.81	18	**< 0.001**	2.77	0.73	3.28	18	**< 0.01**	2.49	0.59
DB	-14.48	18	1.00	1.13	-0.96	-4.48	18	1.00	1.41	-0.71
DP	142.71	18	**< 0.001**	0.06	1.00	138.50	18	**< 0.001**	0.07	1.00
P	194.77	18	**< 0.001**	0.06	1.00	108.34	18	**< 0.001**	0.08	1.00
Z	2.84	18	**< 0.01**	0.82	0.54	2.87	18	**< 0.01**	1.31	0.54

(c) Results with LCF processed data sets.

Dat.	t	df	p-value	SE	r	t	df	p-value	SE	r
	(BIHMnt) - (AvgMnt)					(CzMnt) - (AvgMnt)				
BB	6.62	18	**< 0.001**	2.04	0.83	4.33	18	**< 0.001**	2.30	0.70
BT	4.34	18	**< 0.001**	3.00	0.70	2.00	18	**< 0.05**	3.08	0.41
DB	-25.49	18	1.00	0.80	-0.98	-3.48	18	1.00	0.88	-0.61
DP	115.38	18	**< 0.001**	0.08	1.00	114.60	18	**< 0.001**	0.08	1.00
P	214.05	18	**< 0.001**	0.06	1.00	61.42	18	**< 0.001**	0.13	1.00
Z	4.31	18	**< 0.001**	1.03	0.69	2.68	18	**< 0.01**	1.24	0.51

Table A.13: Independent t-test analysis of EEG montages with *Conf-1s* system. H_0: $\mu_{PRE}(A) - \mu_{PRE}(B) \geq 5$. H_1: $\mu_{PRE}(A) - \mu_{PRE}(B) < 5$. With A and B montages specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.1 for a description of the results.

(a) Results with raw data sets.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(BIHMnt) - (AvgMnt)					(CzMnt) - (AvgMnt)				
BB	4.64	18	**< 0.001**	1.45	0.72	3.89	18	**< 0.001**	1.53	0.66
BT	3.95	18	**< 0.001**	1.29	0.66	3.29	18	**< 0.01**	1.32	0.59
DB	3.40	18	**< 0.01**	0.60	0.61	0.39	18	0.38	0.67	0.09
DP	21.96	18	**< 0.001**	0.17	0.98	11.44	18	**< 0.001**	0.16	0.93
K	25.27	18	**< 0.001**	0.11	0.98	-	-	-	-	-
P	77.75	18	**< 0.001**	0.08	1.00	96.16	18	**< 0.001**	0.06	1.00

(b) Results with ADJUST processed data sets.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(BIHMnt) - (AvgMnt)					(CzMnt) - (AvgMnt)				
BB	5.48	18	**< 0.001**	1.09	0.77	4.84	18	**< 0.001**	1.06	0.73
BT	7.05	18	**< 0.001**	0.83	0.84	3.17	18	**< 0.01**	1.06	0.58
DB	9.79	18	**< 0.001**	0.39	0.91	-1.42	18	0.97	0.70	-0.30
DP	6.42	18	**< 0.001**	0.10	0.82	10.87	18	**< 0.001**	0.09	0.92
P	83.94	18	**< 0.001**	0.08	1.00	16.99	18	**< 0.001**	0.11	0.97

(c) Results with LCF processed data sets.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(BIHMnt) - (AvgMnt)					(CzMnt) - (AvgMnt)				
BB	8.88	18	**< 0.001**	1.00	0.89	4.06	18	**< 0.001**	1.36	0.67
BT	12.21	18	**< 0.001**	0.56	0.94	5.20	18	**< 0.001**	0.89	0.76
DB	48.50	18	**< 0.001**	0.11	1.00	3.07	18	**< 0.01**	0.29	0.57
DP	14.19	18	**< 0.001**	0.13	0.95	8.17	18	**< 0.001**	0.12	0.88
P	108.47	18	**< 0.001**	0.10	1.00	-70.03	18	1.00	0.10	-1.00

Table A.14: Independent t-test analysis of EEG montages with *Conf-2s* system. H_0: $\mu_{PRE}(A) - \mu_{PRE}(B) \geq 5$. H_1: $\mu_{PRE}(A) - \mu_{PRE}(B) < 5$. With A and B montages specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.1 for a description of the results.

(a) Results with raw data sets.

Dat.	t	df	p-value	SE	r	t	df	p-value	SE	r
	(BIHMnt) - (AvgMnt)					(CzMnt) - (AvgMnt)				
BB	5.62	18	**< 0.001**	1.19	0.78	5.22	18	**< 0.001**	1.06	0.76
BT	3.33	18	**< 0.01**	1.92	0.60	2.00	18	**< 0.05**	1.99	0.41
DB	20.34	18	**< 0.001**	0.32	0.98	4.83	18	**< 0.001**	0.32	0.73
DP	-0.04	18	0.55	0.14	-0.01	17.33	18	**< 0.001**	0.16	0.97
K	3.45	18	**< 0.01**	0.52	0.61	-	-	-	-	-
P	62.72	18	**< 0.001**	0.08	1.00	86.75	18	**< 0.001**	0.06	1.00

(b) Results with ADJUST processed data sets.

Dat.	t	df	p-value	SE	r	t	df	p-value	SE	r
	(BIHMnt) - (AvgMnt)					(CzMnt) - (AvgMnt)				
BB	7.10	18	**< 0.001**	0.80	0.85	6.81	18	**< 0.001**	0.83	0.84
BT	5.01	18	**< 0.001**	1.40	0.75	2.36	18	**< 0.05**	1.45	0.47
DB	16.79	18	**< 0.001**	0.29	0.97	2.62	18	**< 0.05**	0.46	0.50
DP	28.49	18	**< 0.001**	0.07	0.99	32.24	18	**< 0.001**	0.08	0.99
P	121.33	18	**< 0.001**	0.05	1.00	62.62	18	**< 0.001**	0.05	1.00

(c) Results with LCF processed data sets.

Dat.	t	df	p-value	SE	r	t	df	p-value	SE	r
	(BIHMnt) - (AvgMnt)					(CzMnt) - (AvgMnt)				
BB	7.77	18	**< 0.001**	1.00	0.87	4.81	18	**< 0.001**	1.07	0.73
BT	4.66	18	**< 0.001**	1.41	0.72	2.08	18	**< 0.05**	1.09	0.42
DB	10.58	18	**< 0.001**	0.32	0.92	0.40	18	0.38	0.39	0.09
DP	28.21	18	**< 0.001**	0.10	0.99	25.21	18	**< 0.001**	0.11	0.98
P	163.33	18	**< 0.001**	0.05	1.00	8.35	18	**< 0.001**	0.12	0.88

Table A.15: Independent t-test analysis of PSD normalization with *Conf-HalfLen* system. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(iqrNorm) - (Raw)					(normNorm) - (Raw)				
BT	1.74	18	0.12	1.72	0.36	-6.13	18	$<$ **0.001**	2.14	-0.81
K	54.42	18	$<$ **0.001**	0.38	1.00	42.65	18	$<$ **0.001**	0.40	0.99
Y	-128.03	18	$<$ **0.001**	0.02	-1.00	-527.09	18	$<$ **0.001**	0.02	-1.00
Z	-0.78	18	0.47	2.17	-0.17	-3.17	18	$<$ **0.01**	2.28	-0.58
	(powNorm) - (Raw)					(prcNorm) - (Raw)				
BT	-2.18	18	0.05	2.00	-0.44	-3.46	18	$<$ **0.01**	2.19	-0.61
K	57.44	18	$<$ **0.001**	0.37	1.00	41.62	18	$<$ **0.001**	0.41	0.99
Y	-160.27	18	$<$ **0.001**	0.02	-1.00	-621.81	18	$<$ **0.001**	0.02	-1.00
Z	-0.81	18	0.45	2.16	-0.18	-4.07	18	$<$ **0.01**	2.34	-0.67
	(rNorm) - (Raw)					(zNorm) - (Raw)				
BT	2.40	18	$<$ **0.05**	1.71	0.47	-5.41	18	$<$ **0.001**	2.06	-0.77
K	55.04	18	$<$ **0.001**	0.38	1.00	48.90	18	$<$ **0.001**	0.36	1.00
Y	-106.26	18	$<$ **0.001**	0.02	-1.00	-460.96	18	$<$ **0.001**	0.02	-1.00
Z	-1.56	18	0.15	2.31	-0.33	-4.32	18	$<$ **0.001**	2.65	-0.69

Table A.16: Independent t-test analysis of PSD normalization with *Conf-FullLen* system. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to section 5.3.2 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(iqrNorm) - (Raw)					(normNorm) - (Raw)				
BB	3.47	18	**< 0.01**	2.44	0.61	-6.53	18	**< 0.001**	2.05	-0.83
BT	4.44	18	**< 0.001**	2.10	0.70	-6.19	18	**< 0.001**	2.39	-0.81
DB	4.70	18	**< 0.001**	0.99	0.72	-22.92	18	**< 0.001**	0.88	-0.98
DP	-4.09	18	**< 0.01**	0.11	-0.67	-57.68	18	**< 0.001**	0.13	-1.00
K	109.22	18	**< 0.001**	0.19	1.00	49.84	18	**< 0.001**	0.32	1.00
P	-10.61	18	**< 0.001**	0.05	-0.92	-132.35	18	**< 0.001**	0.07	-1.00
Y	74.08	18	**< 0.001**	0.01	1.00	-536.16	18	**< 0.001**	0.02	-1.00
Z	-2.24	18	**< 0.05**	1.33	-0.45	-7.47	18	**< 0.001**	1.16	-0.86
	(powNorm) - (Raw)					(prcNorm) - (Raw)				
BB	-0.78	18	0.47	2.44	-0.17	-3.56	18	**< 0.01**	2.28	-0.62
BT	-2.00	18	0.07	2.34	-0.41	-4.94	18	**< 0.001**	2.34	-0.74
DB	-10.41	18	**< 0.001**	0.82	-0.92	-12.12	18	**< 0.001**	0.88	-0.94
DP	-23.69	18	**< 0.001**	0.12	-0.98	-37.80	18	**< 0.001**	0.11	-0.99
K	103.00	18	**< 0.001**	0.20	1.00	84.94	18	**< 0.001**	0.20	1.00
P	-81.67	18	**< 0.001**	0.07	-1.00	-94.12	18	**< 0.001**	0.06	-1.00
Y	-105.72	18	**< 0.001**	0.02	-1.00	-378.87	18	**< 0.001**	0.03	-1.00
Z	-1.82	18	0.10	1.24	-0.38	-7.51	18	**< 0.001**	1.15	-0.86
	(rNorm) - (Raw)					(zNorm) - (Raw)				
BB	5.21	18	**< 0.001**	2.30	0.76	-0.29	18	0.79	2.20	-0.06
BT	5.62	18	**< 0.001**	2.07	0.78	-1.95	18	0.08	2.24	-0.40
DB	10.76	18	**< 0.001**	1.01	0.92	-10.49	18	**< 0.001**	0.82	-0.92
DP	-1.59	18	0.15	0.11	-0.34	-48.75	18	**< 0.001**	0.17	-1.00
K	113.23	18	**< 0.001**	0.19	1.00	56.26	18	**< 0.001**	0.29	1.00
P	-19.76	18	**< 0.001**	0.04	-0.98	-153.96	18	**< 0.001**	0.07	-1.00
Y	56.61	18	**< 0.001**	0.02	1.00	-440.68	18	**< 0.001**	0.02	-1.00
Z	-2.48	18	**< 0.05**	1.35	-0.48	-8.79	18	**< 0.001**	1.49	-0.89

Table A.17: Independent t-test analysis of PSD normalization with *Conf-1s* system. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(iqrNorm) - (Raw)					(normNorm) - (Raw)				
BB	-1.02	18	0.35	1.59	-0.22	-2.01	18	0.07	1.94	-0.41
BT	1.36	18	0.21	1.45	0.29	-4.44	18	**< 0.001**	1.45	-0.70
DB	11.77	18	**< 0.001**	0.46	0.93	7.70	18	**< 0.001**	0.59	0.86
DP	87.60	18	**< 0.001**	0.16	1.00	57.31	18	**< 0.001**	0.18	1.00
K	59.29	18	**< 0.001**	0.38	1.00	54.63	18	**< 0.001**	0.39	1.00
P	15.49	18	**< 0.001**	0.06	0.96	-16.68	18	**< 0.001**	0.09	-0.97
	(powNorm) - (Raw)					(prcNorm) - (Raw)				
BB	-0.80	18	0.45	1.70	-0.18	-2.43	18	**< 0.05**	1.86	-0.48
BT	-0.95	18	0.38	1.49	-0.21	-2.93	18	**< 0.05**	1.43	-0.55
DB	12.85	18	**< 0.001**	0.45	0.94	8.57	18	**< 0.001**	0.52	0.89
DP	106.52	18	**< 0.001**	0.13	1.00	73.02	18	**< 0.001**	0.14	1.00
K	60.89	18	**< 0.001**	0.38	1.00	49.64	18	**< 0.001**	0.41	1.00
P	-6.99	18	**< 0.001**	0.07	-0.84	-21.10	18	**< 0.001**	0.09	-0.98
	(rNorm) - (Raw)					(zNorm) - (Raw)				
BB	-1.06	18	0.33	1.64	-0.23	-1.16	18	0.29	2.05	-0.25
BT	2.23	18	**< 0.05**	1.29	0.45	-3.33	18	**< 0.01**	1.80	-0.60
DB	10.61	18	**< 0.001**	0.50	0.92	-1.65	18	0.13	0.58	-0.35
DP	103.14	18	**< 0.001**	0.14	1.00	66.51	18	**< 0.001**	0.15	1.00
K	54.98	18	**< 0.001**	0.41	1.00	48.76	18	**< 0.001**	0.41	1.00
P	13.20	18	**< 0.001**	0.07	0.95	-81.67	18	**< 0.001**	0.07	-1.00

Table A.18: Independent t-test analysis of PSD normalization with *Conf-2s* system. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(iqrNorm) - (Raw)					(normNorm) - (Raw)				
BB	-0.73	18	0.49	0.93	-0.16	-2.51	18	**< 0.05**	1.17	-0.49
BT	1.74	18	0.12	1.72	0.36	-6.13	18	**< 0.001**	2.14	-0.81
DB	11.13	18	**< 0.001**	0.32	0.93	-4.31	18	**< 0.001**	0.39	-0.69
DP	79.21	18	**< 0.001**	0.11	1.00	49.56	18	**< 0.001**	0.13	1.00
K	54.42	18	**< 0.001**	0.38	1.00	42.65	18	**< 0.001**	0.40	0.99
P	-2.83	18	**< 0.05**	0.07	-0.53	-57.83	18	**< 0.001**	0.07	-1.00
	(powNorm) - (Raw)					(prcNorm) - (Raw)				
BB	-1.18	18	0.28	1.21	-0.25	-2.12	18	0.06	1.14	-0.43
BT	-2.18	18	0.05	2.00	-0.44	-3.46	18	**< 0.01**	2.19	-0.61
DB	10.47	18	**< 0.001**	0.29	0.92	5.67	18	**< 0.001**	0.32	0.78
DP	70.37	18	**< 0.001**	0.12	1.00	55.76	18	**< 0.001**	0.13	1.00
K	57.44	18	**< 0.001**	0.37	1.00	41.62	18	**< 0.001**	0.41	0.99
P	-38.51	18	**< 0.001**	0.07	-0.99	-25.39	18	**< 0.001**	0.08	-0.98
	(rNorm) - (Raw)					(zNorm) - (Raw)				
BB	-0.28	18	0.79	0.94	-0.06	-1.98	18	0.08	1.15	-0.40
BT	2.40	18	**< 0.05**	1.71	0.47	-5.41	18	**< 0.001**	2.06	-0.77
DB	13.50	18	**< 0.001**	0.28	0.95	-5.48	18	**< 0.001**	0.38	-0.77
DP	76.17	18	**< 0.001**	0.12	1.00	47.22	18	**< 0.001**	0.12	1.00
K	55.04	18	**< 0.001**	0.38	1.00	48.90	18	**< 0.001**	0.36	1.00
P	3.63	18	**< 0.01**	0.07	0.63	-117.75	18	**< 0.001**	0.08	-1.00

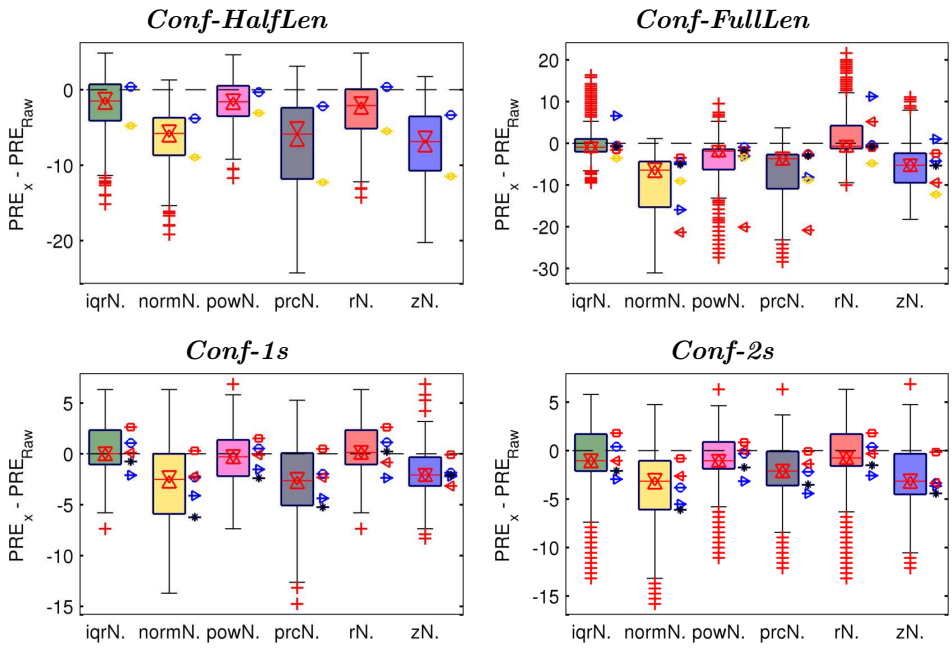


Figure A.8: Quantitative analysis of PSD normalization with ADJUST processed data sets. Relative PRE values between the PSD normalized by each of the methods in table 5.3 and the raw PSD. Results are stacked across databases. A maximum of 20 subjects was used in each experimental iteration. Refer to caption of figure 5.9 for details about the meaning of symbols within the image, and to section 5.3.2 for a description of the results.

Table A.19: Independent t-test analysis of PSD normalization with *Conf-HalfLen* system and ADJUST processed data sets. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(iqrNorm) - (Raw)					(normNorm) - (Raw)				
BT	0.27	18	0.80	1.37	0.06	-5.47	18	**< 0.001**	1.29	-0.77
Z	-3.46	18	**< 0.01**	1.38	-0.61	-6.70	18	**< 0.001**	1.34	-0.83
	(powNorm) - (Raw)					(prcNorm) - (Raw)				
BT	-1.02	18	0.35	1.39	-0.22	-3.05	18	**< 0.01**	1.31	-0.56
Z	-2.87	18	**< 0.05**	1.08	-0.54	-7.67	18	**< 0.001**	1.60	-0.86
	(rNorm) - (Raw)					(zNorm) - (Raw)				
BT	1.01	18	0.35	1.25	0.22	-2.84	18	**< 0.05**	1.56	-0.54
Z	-4.79	18	**< 0.001**	1.15	-0.73	-9.09	18	**< 0.001**	1.26	-0.90

Table A.20: Independent t-test analysis of PSD normalization with *Conf-FullLen* system and ADJUST processed data sets. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(iqrNorm) - (Raw)					(normNorm) - (Raw)				
BB	3.54	18	**< 0.01**	1.86	0.62	-7.36	18	**< 0.001**	2.17	-0.85
BT	5.04	18	**< 0.001**	2.41	0.75	-4.28	18	**< 0.001**	2.51	-0.69
DB	-0.69	18	0.52	1.07	-0.15	-18.95	18	**< 0.001**	1.13	-0.97
DP	-20.92	18	**< 0.001**	0.07	-0.98	-39.00	18	**< 0.001**	0.09	-0.99
P	-12.40	18	**< 0.001**	0.06	-0.94	-56.39	18	**< 0.001**	0.09	-1.00
Z	-4.69	18	**< 0.001**	0.77	-0.72	-10.62	18	**< 0.001**	0.85	-0.92
	(powNorm) - (Raw)					(prcNorm) - (Raw)				
BB	-1.50	18	0.17	2.03	-0.32	-3.75	18	**< 0.01**	2.17	-0.64
BT	1.44	18	0.19	2.34	0.31	-2.02	18	0.07	2.46	-0.41
DB	-18.84	18	**< 0.001**	1.07	-0.97	-20.05	18	**< 0.001**	1.04	-0.98
DP	-21.97	18	**< 0.001**	0.08	-0.98	-32.30	18	**< 0.001**	0.08	-0.99
P	-29.46	18	**< 0.001**	0.06	-0.99	-37.19	18	**< 0.001**	0.08	-0.99
Z	-3.67	18	**< 0.01**	0.92	-0.63	-10.24	18	**< 0.001**	0.86	-0.92
	(rNorm) - (Raw)					(zNorm) - (Raw)				
BB	6.04	18	**< 0.001**	1.86	0.80	0.50	18	0.64	2.01	0.11
BT	5.30	18	**< 0.001**	2.48	0.76	1.29	18	0.24	2.57	0.28
DB	4.13	18	**< 0.001**	1.25	0.68	-9.22	18	**< 0.001**	1.03	-0.90
DP	-11.79	18	**< 0.001**	0.09	-0.93	-31.06	18	**< 0.001**	0.08	-0.99
P	-11.80	18	**< 0.001**	0.07	-0.94	-61.42	18	**< 0.001**	0.09	-1.00
Z	-6.33	18	**< 0.001**	0.76	-0.82	-13.28	18	**< 0.001**	0.92	-0.95

Table A.21: Independent t-test analysis of PSD normalization with *Conf-1s* system and ADJUST processed data sets. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(iqrNorm) - (Raw)					(normNorm) - (Raw)				
BB	-2.24	18	$<$ **0.05**	0.94	-0.45	-3.09	18	$<$ **0.01**	1.33	-0.57
BT	-1.77	18	0.11	0.77	-0.37	-5.23	18	$<$ **0.001**	0.82	-0.76
DB	0.30	18	0.78	0.35	0.07	-4.90	18	$<$ **0.001**	0.45	-0.74
DP	27.89	18	$<$ **0.001**	0.09	0.99	2.52	18	$<$ **0.05**	0.12	0.49
P	-8.26	18	$<$ **0.001**	0.09	-0.88	-60.10	18	$<$ **0.001**	0.10	-1.00
	(powNorm) - (Raw)					(prcNorm) - (Raw)				
BB	-1.56	18	0.15	0.98	-0.33	-3.24	18	$<$ **0.01**	1.35	-0.59
BT	-1.94	18	0.08	0.71	-0.40	-3.75	18	$<$ **0.01**	0.86	-0.64
DB	-0.27	18	0.80	0.39	-0.06	-5.46	18	$<$ **0.001**	0.42	-0.77
DP	17.57	18	$<$ **0.001**	0.09	0.97	5.66	18	$<$ **0.001**	0.09	0.78
P	-27.35	18	$<$ **0.001**	0.09	-0.99	-56.75	18	$<$ **0.001**	0.09	-1.00
	(rNorm) - (Raw)					(zNorm) - (Raw)				
BB	-2.57	18	$<$ **0.05**	0.92	-0.50	-2.23	18	$<$ **0.05**	0.99	-0.45
BT	-1.69	18	0.13	0.66	-0.35	-3.08	18	$<$ **0.01**	0.90	-0.57
DB	-2.53	18	$<$ **0.05**	0.33	-0.49	-7.07	18	$<$ **0.001**	0.45	-0.85
DP	30.71	18	$<$ **0.001**	0.09	0.99	-0.90	18	0.41	0.11	-0.20
P	2.74	18	$<$ **0.05**	0.08	0.52	-24.44	18	$<$ **0.001**	0.09	-0.98

Table A.22: Independent t-test analysis of PSD normalization with *Conf-2s* system and ADJUST processed data sets. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A)-(B)$. Refer to section 5.3.2 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(iqrNorm) - (Raw)					(normNorm) - (Raw)				
BB	-2.28	18	**< 0.05**	1.29	-0.45	-3.71	18	**< 0.01**	1.49	-0.64
BT	0.27	18	0.80	1.37	0.06	-5.47	18	**< 0.001**	1.29	-0.77
DB	-3.03	18	**< 0.01**	0.35	-0.56	-6.50	18	**< 0.001**	0.40	-0.82
DP	33.11	18	**< 0.001**	0.05	0.99	-10.53	18	**< 0.001**	0.08	-0.92
P	-52.40	18	**< 0.001**	0.04	-1.00	-119.74	18	**< 0.001**	0.05	-1.00
	(powNorm) - (Raw)					(prcNorm) - (Raw)				
BB	-2.58	18	**< 0.05**	1.23	-0.50	-3.28	18	**< 0.01**	1.35	-0.59
BT	-1.02	18	0.35	1.39	-0.22	-3.05	18	**< 0.01**	1.31	-0.56
DB	0.00	18	1.00	NaN	0.00	-3.24	18	**< 0.01**	0.42	-0.59
DP	12.77	18	**< 0.001**	0.07	0.94	-1.17	18	0.28	0.07	-0.25
P	-33.15	18	**< 0.001**	0.05	-0.99	-81.89	18	**< 0.001**	0.04	-1.00
	(rNorm) - (Raw)					(zNorm) - (Raw)				
BB	-1.90	18	0.09	1.36	-0.39	-2.54	18	**< 0.05**	1.43	-0.49
BT	1.01	18	0.35	1.25	0.22	-2.84	18	**< 0.05**	1.56	-0.54
DB	-0.80	18	0.45	0.39	-0.18	-7.64	18	**< 0.001**	0.44	-0.86
DP	32.45	18	**< 0.001**	0.06	0.99	-2.18	18	0.05	0.07	-0.44
P	-36.44	18	**< 0.001**	0.04	-0.99	-113.20	18	**< 0.001**	0.04	-1.00

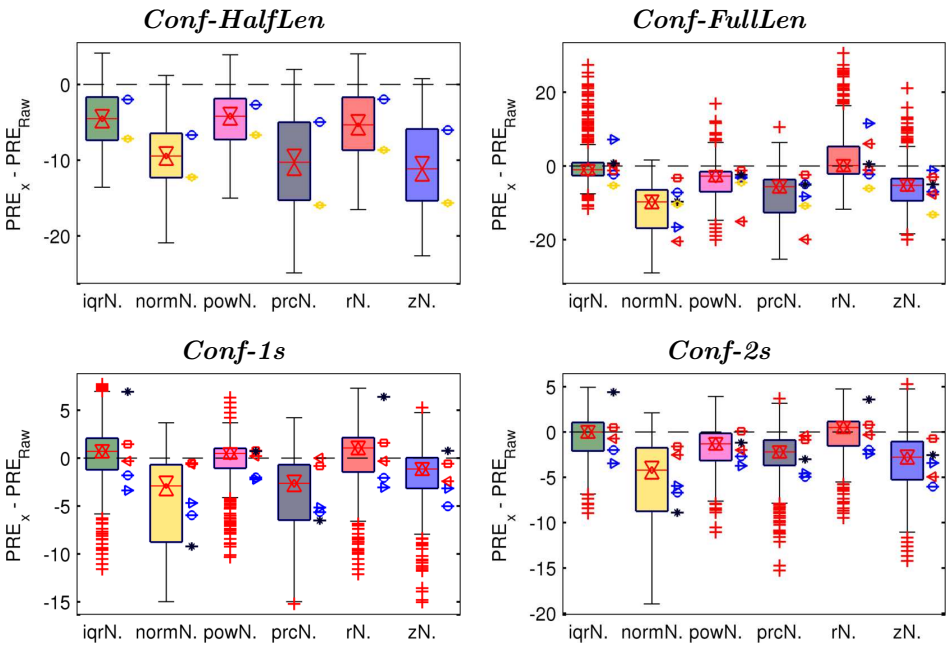


Figure A.9: Quantitative analysis of PSD normalization with LCF processed data sets. Relative PRE values between the PSD normalized by each of the methods in table 5.3 and the raw PSD. Results are stacked across databases. A maximum of 20 subjects was used in each experimental iteration. Refer to caption of figure 5.9 for details about the meaning of symbols within the image, and to section 5.3.2 for a description of the results.

Table A.23: Independent t-test analysis of PSD normalization with *Conf-HalfLen* system and LCF processed data sets. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(iqrNorm) - (Raw)					(normNorm) - (Raw)				
BT	-3.76	18	**< 0.01**	1.09	-0.64	-8.53	18	**< 0.001**	1.20	-0.89
Z	-8.21	18	**< 0.001**	0.88	-0.88	-10.71	18	**< 0.001**	1.14	-0.92
	(powNorm) - (Raw)					(prcNorm) - (Raw)				
BT	-4.73	18	**< 0.001**	1.14	-0.73	-7.40	18	**< 0.001**	1.16	-0.86
Z	-5.52	18	**< 0.001**	1.21	-0.78	-14.54	18	**< 0.001**	1.10	-0.96
	(rNorm) - (Raw)					(zNorm) - (Raw)				
BT	-3.22	18	**< 0.01**	1.05	-0.58	-5.46	18	**< 0.001**	1.10	-0.77
Z	-9.38	18	**< 0.001**	0.92	-0.90	-17.32	18	**< 0.001**	0.90	-0.97

Table A.24: Independent t-test analysis of PSD normalization with *Conf-FullLen* system and LCF processed data sets. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(iqrNorm) - (Raw)					(normNorm) - (Raw)				
BB	2.71	18	**< 0.05**	2.64	0.52	-7.52	18	**< 0.001**	2.20	-0.86
BT	2.60	18	**< 0.05**	3.07	0.50	-3.33	18	**< 0.01**	3.25	-0.60
DB	0.29	18	0.79	0.73	0.06	-29.70	18	**< 0.001**	0.69	-0.99
DP	-14.90	18	**< 0.001**	0.08	-0.96	-24.12	18	**< 0.001**	0.14	-0.98
P	9.18	18	**< 0.001**	0.07	0.90	-84.10	18	**< 0.001**	0.12	-1.00
Z	-6.35	18	**< 0.001**	0.84	-0.82	-12.43	18	**< 0.001**	0.84	-0.94
	(powNorm) - (Raw)					(prcNorm) - (Raw)				
BB	-1.47	18	0.18	2.19	-0.31	-3.89	18	**< 0.01**	2.12	-0.66
BT	0.64	18	0.55	3.06	0.14	-2.49	18	**< 0.05**	2.89	-0.49
DB	-23.23	18	**< 0.001**	0.65	-0.98	-25.64	18	**< 0.001**	0.78	-0.99
DP	-13.96	18	**< 0.001**	0.09	-0.95	-31.48	18	**< 0.001**	0.08	-0.99
P	-26.60	18	**< 0.001**	0.09	-0.99	-61.34	18	**< 0.001**	0.08	-1.00
Z	-4.99	18	**< 0.001**	0.91	-0.74	-11.87	18	**< 0.001**	0.91	-0.94
	(rNorm) - (Raw)					(zNorm) - (Raw)				
BB	4.81	18	**< 0.001**	2.40	0.73	-0.54	18	0.61	2.22	-0.12
BT	3.88	18	**< 0.01**	3.03	0.65	0.64	18	0.55	3.21	0.14
DB	8.76	18	**< 0.001**	0.68	0.89	-9.25	18	**< 0.001**	0.83	-0.90
DP	-14.46	18	**< 0.001**	0.07	-0.96	-23.99	18	**< 0.001**	0.13	-0.98
P	7.12	18	**< 0.001**	0.07	0.85	-70.31	18	**< 0.001**	0.07	-1.00
Z	-6.95	18	**< 0.001**	0.88	-0.84	-13.27	18	**< 0.001**	0.99	-0.95

Table A.25: Independent t-test analysis of PSD normalization with *Conf-1s* system and LCF processed data sets. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(iqrNorm) - (Raw)					(normNorm) - (Raw)				
BB	-2.21	18	**< 0.05**	1.52	-0.44	-4.08	18	**< 0.01**	1.15	-0.67
BT	-1.70	18	0.12	1.14	-0.36	-4.86	18	**< 0.001**	1.09	-0.74
DB	-1.57	18	0.15	0.20	-0.33	-2.55	18	**< 0.05**	0.25	-0.49
DP	13.01	18	**< 0.001**	0.11	0.95	-4.73	18	**< 0.001**	0.11	-0.73
P	62.57	18	**< 0.001**	0.11	1.00	-83.13	18	**< 0.001**	0.11	-1.00
	(powNorm) - (Raw)					(prcNorm) - (Raw)				
BB	-1.70	18	0.12	1.30	-0.36	-3.90	18	**< 0.01**	1.32	-0.66
BT	-2.13	18	0.06	0.96	-0.43	-4.17	18	**< 0.001**	1.11	-0.68
DB	0.68	18	0.52	0.31	0.15	0.00	18	1.00	0.27	0.00
DP	6.97	18	**< 0.001**	0.11	0.84	-7.83	18	**< 0.001**	0.11	-0.87
P	6.56	18	**< 0.001**	0.11	0.83	-45.83	18	**< 0.001**	0.14	-1.00
	(rNorm) - (Raw)					(zNorm) - (Raw)				
BB	-1.92	18	0.08	1.59	-0.40	-2.77	18	**< 0.05**	1.14	-0.53
BT	-2.07	18	0.06	0.96	-0.42	-3.59	18	**< 0.01**	1.04	-0.63
DB	-0.93	18	0.39	0.34	-0.20	-6.49	18	**< 0.001**	0.37	-0.82
DP	14.92	18	**< 0.001**	0.11	0.96	-5.49	18	**< 0.001**	0.11	-0.78
P	57.93	18	**< 0.001**	0.11	1.00	8.09	18	**< 0.001**	0.09	0.88

Table A.26: Independent t-test analysis of PSD normalization with *Conf-2s* system and LCF processed data sets. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	(iqrNorm) - (Raw)					(normNorm) - (Raw)				
BB	-3.66	18	**< 0.01**	0.95	-0.63	-3.94	18	**< 0.01**	1.51	-0.66
BT	-3.76	18	**< 0.01**	1.09	-0.64	-8.53	18	**< 0.001**	1.20	-0.89
DB	-1.77	18	0.11	0.42	-0.37	-5.62	18	**< 0.001**	0.45	-0.78
DP	6.65	18	**< 0.001**	0.07	0.83	-27.17	18	**< 0.001**	0.06	-0.99
P	52.26	18	**< 0.001**	0.08	1.00	-123.42	18	**< 0.001**	0.07	-1.00
	(powNorm) - (Raw)					(prcNorm) - (Raw)				
BB	-4.09	18	**< 0.01**	0.91	-0.68	-3.35	18	**< 0.01**	1.36	-0.60
BT	-4.73	18	**< 0.001**	1.14	-0.73	-7.40	18	**< 0.001**	1.16	-0.86
DB	-6.04	18	**< 0.001**	0.33	-0.80	-1.31	18	0.23	0.32	-0.28
DP	1.42	18	0.19	0.06	0.30	-16.25	18	**< 0.001**	0.05	-0.96
P	-12.34	18	**< 0.001**	0.10	-0.94	-29.09	18	**< 0.001**	0.10	-0.99
	(rNorm) - (Raw)					(zNorm) - (Raw)				
BB	-2.25	18	**< 0.05**	1.08	-0.45	-2.54	18	**< 0.05**	1.35	-0.49
BT	-3.22	18	**< 0.01**	1.05	-0.58	-5.46	18	**< 0.001**	1.10	-0.77
DB	-1.00	18	0.35	0.32	-0.22	-11.87	18	**< 0.001**	0.42	-0.94
DP	15.67	18	**< 0.001**	0.05	0.96	-10.61	18	**< 0.001**	0.07	-0.92
P	51.86	18	**< 0.001**	0.07	1.00	-44.47	18	**< 0.001**	0.06	-0.99

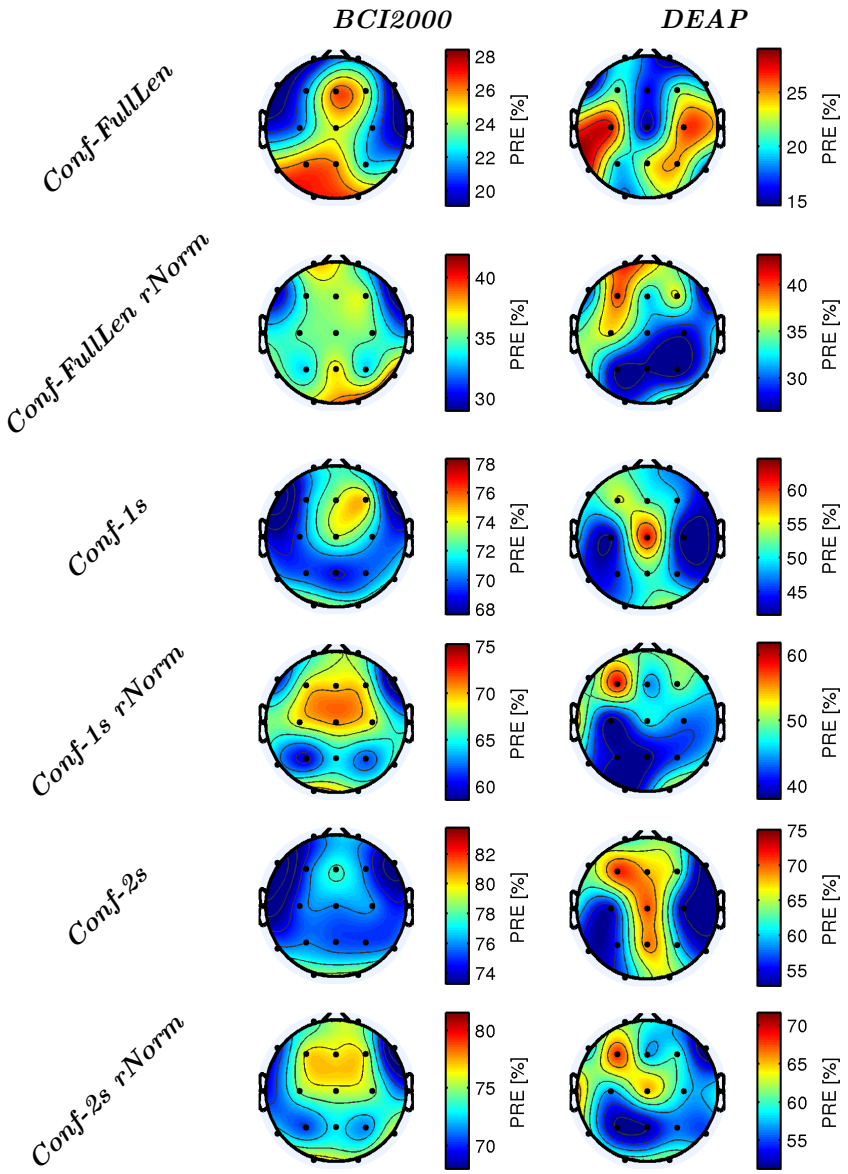


Figure A.10: Quantitative analysis of the spatial distribution of the discriminant information. Mean PRE values obtained at each location with the REO condition of BCI2000 (left) and DEAP (right) data sets when applying each system with and without *rNorm*. Results correspond to *freq-fusion ch-focus* experiments, with a maximum of 20 subjects used in each experimental iteration. Refer to section 5.4.1 for a description of the results.

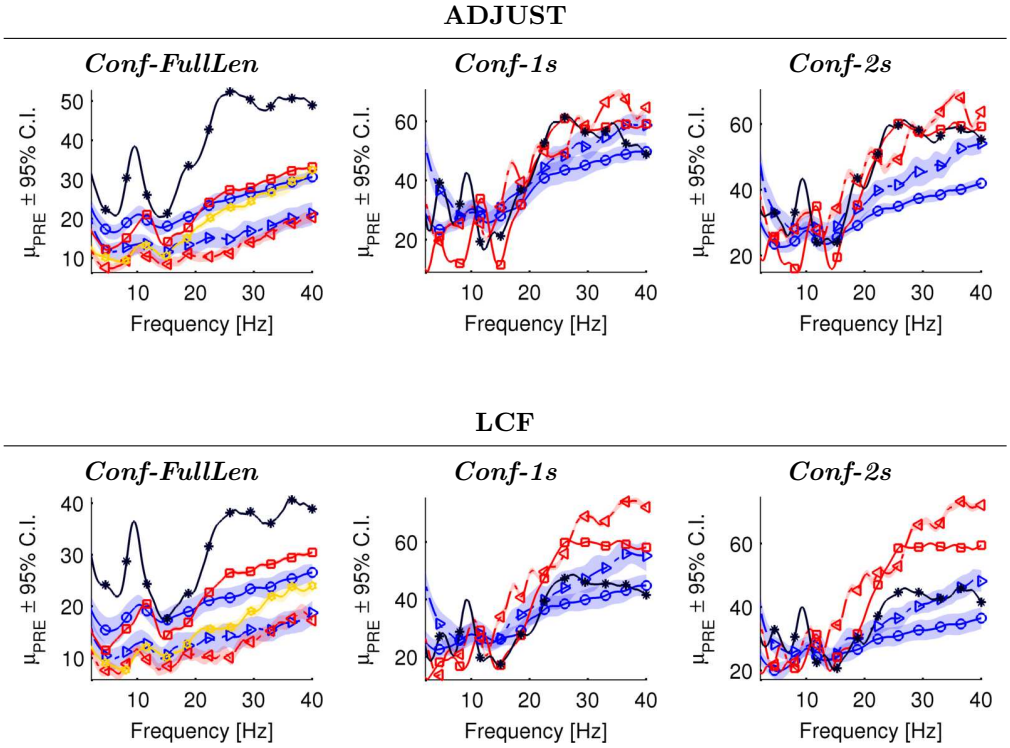


Figure A.11: Quantitative analysis of the discrimination power frequency distribution with artefact free data sets. Mean PRE and 95% CI (shaded area) obtained with each frequency (*ch-fusion* experiments). Curves where smoothed by local regression, using weighted linear least squares and a first degree polynomial model with a 3 Hz span. Graphs correspond to results obtained with the raw PSD of ADJUST (top) and LCF (bottom) processed databases. A maximum of 20 subjects was used in each experimental iteration. Refer to figure 5.5 for details on the legend, and to section 5.4.2 for a description of the results.

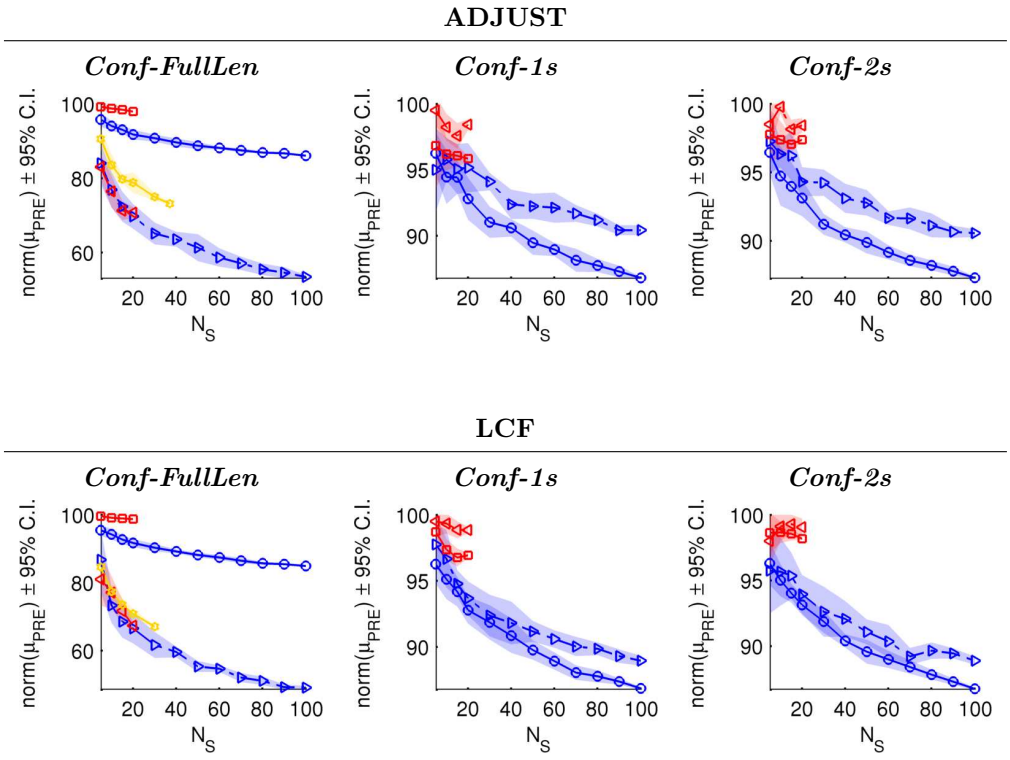


Figure A.12: Quantitative analysis of uniqueness with artefact free data sets. Mean PRE and 95% CI (shaded area) for different numbers of subjects (N_S) in the system and each database (refer to figure 5.5 for details on the legend). Graphs correspond to results obtained with the raw PSD of ADJUST (top) and LCF (bottom) processed databases. Refer to figure 5.5 for details on the legend, and to section 5.4.3 for a description of the results.

Table A.27: Independent t-test of permanence with raw data sets. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) > \mu_{PRE}(B)$. With A the regular CV procedure (no *sess-CV*) and B the *sess-CV* procedure. Refer to table 3.2 for details on databases' code names, and to section 5.4.4 for a description of the results.

(a) Results of *Conf-HalfLen* system – equivalent to *Conf-2s* for Keirn's data set.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	Raw PSD					*rNorm* PSD				
K	2.99	18	**< 0.01**	1.69	0.56	20.13	18	**< 0.001**	0.68	0.98
Y	45.01	18	**< 0.001**	0.39	1.00	66.52	18	**< 0.001**	0.12	1.00

(b) Results of *Conf-FullLen* system.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	Raw PSD					*rNorm* PSD				
K	3.70	18	**< 0.01**	1.01	0.64	24.94	18	**< 0.001**	0.62	0.98
Y	59.98	18	**< 0.001**	0.36	1.00	72.95	18	**< 0.001**	0.15	1.00

(c) Results of *Conf-1s* system.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	Raw PSD					*rNorm* PSD				
K	-0.09	18	0.54	1.09	-0.02	27.00	18	**< 0.001**	0.30	0.99

Table A.28: Independent t-test of task-independence with raw data sets. H_0: $\mu_{PRE}(A) - \mu_{PRE}(B) \geq 5$. H_1: $\mu_{PRE}(A) - \mu_{PRE}(B) < 5$. A and B are specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.5.1 for a description of the results.

(a) Results with raw systems.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	Conf-HalfLen					*Conf-FullLen*				
BF	-3.06	38	**< 0.01**	0.62	-0.44	-2.64	38	**< 0.01**	0.70	-0.38
DF	-	-	-	-	-	-36.35	38	**< 0.001**	0.08	-0.99
K	-21.96	38	**< 0.001**	0.19	-0.96	-24.12	38	**< 0.001**	0.16	-0.97
P	-	-	-	-	-	-120.84	38	**< 0.001**	0.04	-1.00
Y	-93.57	38	**< 0.001**	0.04	-1.00	-78.54	38	**< 0.001**	0.03	-1.00
Z	-0.47	38	0.36	0.87	-0.07	0.44	38	0.71	1.63	0.07
	Conf-1s					*Conf-2s*				
BF	-2.56	38	**< 0.01**	0.90	-0.38	-3.06	38	**< 0.01**	0.62	-0.44
DF	-41.50	38	**< 0.001**	0.11	-0.99	-50.18	38	**< 0.001**	0.08	-0.99
K	-69.19	38	**< 0.001**	0.06	-1.00	-21.96	38	**< 0.001**	0.19	-0.96
P	-57.49	38	**< 0.001**	0.08	-0.99	-124.02	38	**< 0.001**	0.04	-1.00

(b) Results with *rNorm* systems.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	Conf-HalfLen					*Conf-FullLen*				
BF	-3.97	38	**< 0.001**	0.48	-0.53	-2.50	38	**< 0.01**	0.64	-0.37
DF	-	-	-	-	-	-34.94	38	**< 0.001**	0.08	-0.98
K	-23.45	38	**< 0.001**	0.14	-0.97	-14.08	38	**< 0.001**	0.19	-0.91
P	-	-	-	-	-	-103.16	38	**< 0.001**	0.05	-1.00
Y	-112.36	38	**< 0.001**	0.03	-1.00	-73.08	38	**< 0.001**	0.05	-1.00
Z	0.30	38	0.66	0.81	0.05	0.50	38	0.71	1.94	0.08
	Conf-1s					*Conf-2s*				
BF	-2.83	38	**< 0.01**	0.76	-0.41	-3.97	38	**< 0.001**	0.48	-0.53
DF	-40.82	38	**< 0.001**	0.07	-0.99	-62.60	38	**< 0.001**	0.05	-0.99
K	-13.36	38	**< 0.001**	0.18	-0.90	-23.45	38	**< 0.001**	0.14	-0.97
P	-70.83	38	**< 0.001**	0.07	-1.00	-90.54	38	**< 0.001**	0.06	-1.00

Table A.29: Independent t-test of task-independence with artefact free data sets. H_0: $\mu_{PRE}(A) - \mu_{PRE}(B) \geq 5$. H_1: $\mu_{PRE}(A) - \mu_{PRE}(B) < 5$. A and B are specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.5.1 for a description of the results.

(a) Results with ADJUST processed databases.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	Conf-HalfLen					*Conf-FullLen*				
BF	-3.65	38	**< 0.001**	0.76	-0.50	-3.46	38	**< 0.001**	0.64	-0.48
DF	-	-	-	-	-	-46.46	38	**< 0.001**	0.07	-0.99
P	-	-	-	-	-	-64.94	38	**< 0.001**	0.07	-1.00
Z	-0.81	38	0.24	0.56	-0.13	2.43	38	0.99	0.45	0.36
	Conf-1s					*Conf-2s*				
BF	-5.22	38	**< 0.001**	0.61	-0.64	-3.65	38	**< 0.001**	0.76	-0.50
DF	-78.65	38	**< 0.001**	0.05	-1.00	-84.59	38	**< 0.001**	0.04	-1.00
P	-50.75	38	**< 0.001**	0.09	-0.99	-109.08	38	**< 0.001**	0.05	-1.00

(b) Results with LCF processed databases.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	Conf-HalfLen					*Conf-FullLen*				
BF	-6.13	38	**< 0.001**	0.45	-0.70	-3.97	38	**< 0.001**	0.59	-0.53
DF	-	-	-	-	-	-57.61	38	**< 0.001**	0.07	-0.99
P	-	-	-	-	-	-72.74	38	**< 0.001**	0.07	-1.00
Z	-0.18	38	0.47	1.03	-0.03	2.25	38	0.99	0.69	0.34
	Conf-1s					*Conf-2s*				
BF	-3.54	38	**< 0.001**	0.90	-0.49	-6.13	38	**< 0.001**	0.45	-0.70
DF	-117.53	38	**< 0.001**	0.04	-1.00	-93.63	38	**< 0.001**	0.04	-1.00
P	-24.66	38	**< 0.001**	0.17	-0.97	-49.02	38	**< 0.001**	0.10	-0.99

Table A.30: Quantitative results of task-independence with raw data sets. Mean PRE and 95% CI obtained in *task-CV* experiments. A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 5.5.1 for a description of the results.

(a) Results with raw systems.

Dat.	***Conf-HalfLen***	***Conf-FullLen***	***Conf-1s***	***Conf-2s***
BF	83.07	82.21	81.85	83.07
	[82.13, 84.01]	[81.10, 83.33]	[80.50, 83.20]	[82.13, 84.01]
DF	-	94.89	94.34	97.03
	-	[94.78, 95.00]	[94.18, 94.50]	[96.92, 97.14]
K	73.49	73.10	73.23	73.49
	[73.40, 73.59]	[72.96, 73.25]	[73.13, 73.33]	[73.40, 73.59]
P	-	94.19	92.64	93.70
	-	[94.14, 94.24]	[92.55, 92.73]	[93.65, 93.75]
Y	51.25	55.17	-	-
	[51.18, 51.31]	[55.10, 55.23]	-	-
Z	74.63	70.93	-	-
	[73.28, 75.98]	[68.11, 73.74]	-	-

(b) Results with *rNorm* systems.

Dat.	***Conf-HalfLen***	***Conf-FullLen***	***Conf-1s***	***Conf-2s***
BF	83.91	82.94	84.08	83.91
	[83.13, 84.70]	[82.00, 83.88]	[82.94, 85.22]	[83.13, 84.70]
DF	-	92.81	95.17	96.26
	-	[92.73, 92.90]	[95.06, 95.28]	[96.19, 96.32]
K	93.31	92.14	93.59	93.31
	[93.10, 93.52]	[91.86, 92.42]	[93.32, 93.85]	[93.10, 93.52]
P	-	93.73	93.75	94.16
	-	[93.69, 93.77]	[93.66, 93.84]	[94.11, 94.21]
Y	49.79	57.23	-	-
	[49.73, 49.86]	[57.14, 57.33]	-	-
Z	68.74	67.95	-	-
	[67.55, 69.93]	[64.70, 71.20]	-	-

Table A.31: Quantitative results of task-independence with artefact free data sets. Mean PRE and 95% CI obtained in *task-CV* experiments. A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 5.5.1 for a description of the results.

(a) Results with ADJUST processed data sets.

Dat.	***Conf-HalfLen***	***Conf-FullLen***	***Conf-1s***	***Conf-2s***
BF	87.77	86.31	88.09	87.77
	[86.66, 88.89]	[85.27, 87.35]	[87.18, 89.00]	[86.66, 88.89]
DF	-	95.49	96.84	97.97
	-	[95.42, 95.56]	[96.77, 96.92]	[97.90, 98.03]
P	-	95.82	95.39	96.84
	-	[95.69, 95.94]	[95.27, 95.50]	[96.78, 96.90]
Z	74.32	71.46	-	-
	[73.45, 75.20]	[70.76, 72.16]	-	-

(b) Results with ADJUST processed data sets.

Dat.	***Conf-HalfLen***	***Conf-FullLen***	***Conf-1s***	***Conf-2s***
BF	87.78	86.06	88.19	87.78
	[87.07, 88.49]	[85.14, 86.98]	[86.79, 89.59]	[87.07, 88.49]
DF	-	96.51	98.22	98.32
	-	[96.43, 96.59]	[98.19, 98.26]	[98.25, 98.38]
P	-	92.15	86.43	89.95
	-	[92.10, 92.20]	[86.10, 86.77]	[89.80, 90.11]
Z	67.85	63.74	-	-
	[66.39, 69.31]	[62.55, 64.93]	-	-

Table A.32: Independent t-test of the permanence of the task-independence property with raw data sets. H_0: $\mu_{PRE}(A) - \mu_{PRE}(B) \leq C$. H_1: $\mu_{PRE}(A) - \mu_{PRE}(B) > C$. With A and B corresponding to conditions *Task-Sess-CV* and *Task-CV* respectively. The threshold C is set to 10 for Keirn's database (K) and to 1 for Yeom's database (Y). Refer to section 5.5.2 for a description of the results.

(a) Results with raw systems.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	Conf-HalfLen					*Conf-FullLen*				
K	-3.65	38	1.00	1.58	-0.50	0.44	38	0.46	1.37	0.07
Y	0.51	38	0.45	0.49	0.08	0.58	38	0.44	0.40	0.09
	Conf-1s					*Conf-2s*				
K	-4.43	38	1.00	1.24	-0.57	-3.65	38	1.00	1.58	-0.50

(b) Results with *rNorm* systems.

Dat.	**t**	**df**	**p-value**	**SE**	**r**	**t**	**df**	**p-value**	**SE**	**r**
	Conf-HalfLen					*Conf-FullLen*				
K	-0.84	38	0.89	1.66	-0.13	0.28	38	0.52	1.48	0.04
Y	-0.82	38	0.89	0.19	-0.13	2.43	38	**< 0.05**	0.25	0.36
	Conf-1s					*Conf-2s*				
K	-0.82	38	0.89	1.44	-0.13	-0.84	38	0.89	1.66	-0.13

Table A.33: Accuracy results of RCeps$_{P\%}$ systems and raw PSD coefficients (Part A). Mean and 95% CI accuracy values obtained with different values of P. Within data sets, performances statistically lower than the maximum are pointed by * (single tail t-tests with $df = 38$ and BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 6.3.1 for a description of the results

System	BF	DF	K
RCeps$_{5\%}$	96.02	97.17	79.14
	[95.30, 96.75]	[95.88, 98.45]	[71.17, 87.10]
RCeps$_{10\%}$	95.89	**97.75**	**80.81**
	[94.85, 96.93]	**[96.39, 99.10]**	**[74.63, 87.00]**
RCeps$_{20\%}$	95.73	97.59	80.06
	[94.56, 96.90]	[96.12, 99.05]	[73.97, 86.16]
RCeps$_{30\%}$	94.67	96.80	79.64
	[92.20, 97.15]	[94.42, 99.19]	[72.78, 86.51]
RCeps$_{40\%}$	**96.15**	96.69	79.85
	[94.77, 97.53]	[94.01, 99.37]	[74.43, 85.26]
RCeps$_{50\%}$	93.43*	97.29	76.75
	[91.75, 95.11]	[95.62, 98.96]	[71.07, 82.44]
RCeps$_{75\%}$	91.46*	94.70	73.05
	[88.02, 94.90]	[90.85, 98.55]	[67.71, 78.39]
RCeps$_{100\%}$	88.73*	95.83	59.06*
	[84.94, 92.52]	[93.20, 98.45]	[55.80, 62.32]

Table A.34: Accuracy results of RCeps$_{P\%}$ systems and raw PSD coefficients (Part B). Mean and 95% CI accuracy values obtained with different values of P. Within data sets, performances statistically lower than the maximum are pointed by * (single tail t-tests with $df = 38$ and BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 6.3.1 for a description of the results

System	P	Y	Z
RCeps$_{5\%}$	88.80	70.92	92.42
	[84.04, 93.57]	[65.78, 76.06]	[89.84, 95.01]
RCeps$_{10\%}$	89.16	71.93	92.78
	[84.90, 93.43]	[66.77, 77.08]	[90.58, 94.97]
RCeps$_{20\%}$	89.19	73.94	**93.18**
	[85.55, 92.83]	[69.04, 78.84]	**[90.54, 95.82]**
RCeps$_{30\%}$	89.17	74.46	90.01
	[85.75, 92.58]	[69.64, 79.29]	[86.89, 93.14]
RCeps$_{40\%}$	87.58	74.53	87.91
	[83.65, 91.51]	[69.93, 79.13]	[83.10, 92.73]
RCeps$_{50\%}$	89.27	74.59	89.57
	[85.81, 92.72]	[69.85, 79.33]	[85.85, 93.30]
RCeps$_{75\%}$	**89.57**	**74.94**	84.48*
	[86.57, 92.56]	**[70.23, 79.66]**	[79.10, 89.87]
RCeps$_{100\%}$	88.07	74.62	82.81*
	[84.36, 91.79]	[69.69, 79.55]	[79.10, 86.52]

Table A.35: Accuracy results of RCeps$_{P\%}$ systems and *rNorm* PSD coefficients (Part A). Mean and 95% CI accuracy values obtained with different values of P. Within data sets, performances statistically lower than the maximum are pointed by * (single tail t-tests with $df = 38$ and BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 6.3.1 for a description of the results

System	BF	DF	K
RCeps$_{5\%}$	92.95	96.16	79.05
	[91.34, 94.57]	[94.38, 97.93]	[71.17, 86.94]
RCeps$_{10\%}$	**94.95**	97.40	78.75
	[93.27, 96.62]	[95.78, 99.03]	[72.35, 85.15]
RCeps$_{20\%}$	93.04	**97.42**	**80.11**
	[90.99, 95.09]	**[95.67, 99.17]**	**[75.66, 84.55]**
RCeps$_{30\%}$	92.05*	97.11	77.80
	[90.18, 93.91]	[95.21, 99.00]	[73.29, 82.31]
RCeps$_{40\%}$	91.53*	95.72	75.87
	[89.08, 93.97]	[92.49, 98.95]	[71.02, 80.71]
RCeps$_{50\%}$	89.23*	96.72	75.54
	[87.11, 91.34]	[94.42, 99.02]	[70.64, 80.44]
RCeps$_{75\%}$	83.94*	95.26	69.97*
	[80.47, 87.42]	[91.83, 98.69]	[66.33, 73.60]
RCeps$_{100\%}$	82.45*	93.87	58.41*
	[78.65, 86.26]	[90.25, 97.50]	[56.25, 60.58]

Table A.36: Accuracy results of RCeps$_{P\%}$ systems and *rNorm* PSD coefficients (Part B). Mean and 95% CI accuracy values obtained with different values of P. Within data sets, performances statistically lower than the maximum are pointed by * (single tail t-tests with $df = 38$ and BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 6.3.1 for a description of the results

System	P	Y	Z
RCeps$_{5\%}$	86.38 [82.11, 90.66]	70.36 [65.69, 75.04]	85.85 [82.75, 88.94]
RCeps$_{10\%}$	86.69 [82.58, 90.81]	72.15 [67.69, 76.62]	**87.81 [83.52, 92.09]**
RCeps$_{20\%}$	85.12 [80.18, 90.07]	75.37 [71.41, 79.33]	86.92 [82.93, 90.91]
RCeps$_{30\%}$	**88.15 [84.82, 91.48]**	**77.08 [72.83, 81.34]**	84.78 [80.38, 89.17]
RCeps$_{40\%}$	86.29 [81.83, 90.75]	76.46 [72.02, 80.90]	83.02 [78.67, 87.37]
RCeps$_{50\%}$	86.56 [82.82, 90.30]	76.57 [72.04, 81.10]	77.15* [71.99, 82.32]
RCeps$_{75\%}$	86.55 [83.07, 90.03]	76.50 [71.64, 81.36]	74.42* [68.41, 80.43]
RCeps$_{100\%}$	85.70 [81.54, 89.86]	76.52 [71.96, 81.07]	73.35* [68.00, 78.69]

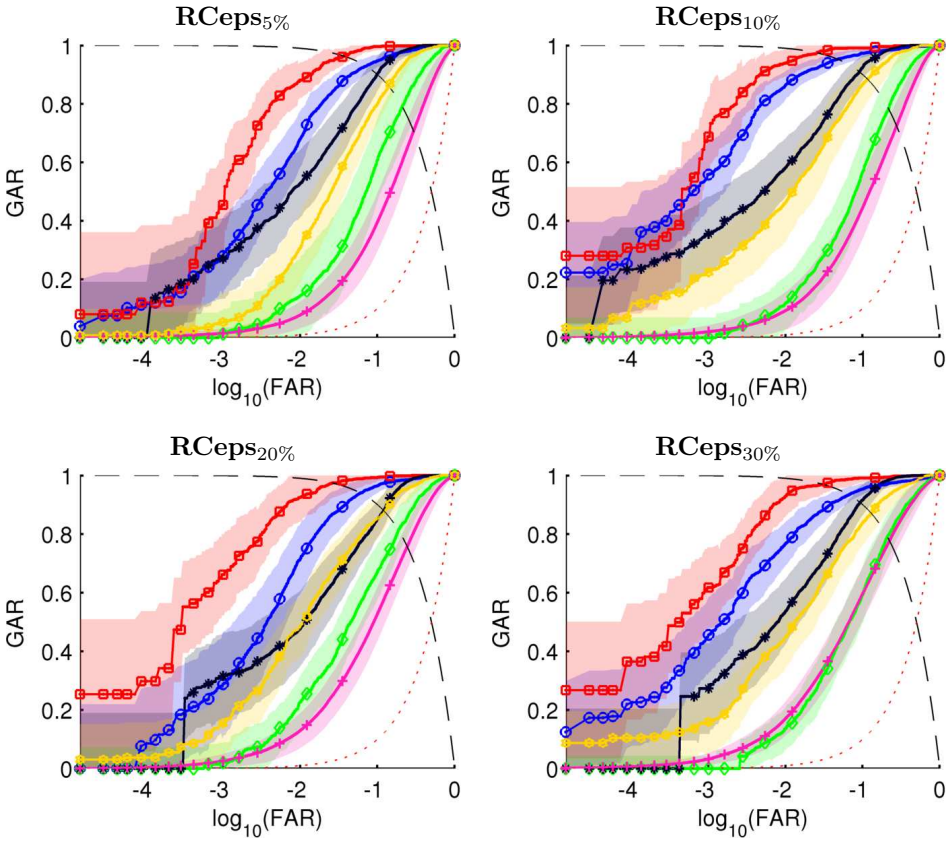


Figure A.13: ROC curves of RCeps$_{P\%}$ systems coefficients with logarithmic FAR axis. Mean ROC curves and std (shaded area) obtained with different configurations of RCeps$_{P\%}$. A maximum of 20 subjects was used in each experimental iteration. Refer to fig. 6.1 for legend details.

Table A.37: Accuracy results of LPC_N systems (Part A). Mean and 95% CI accuracy values obtained with different values of N. Within data sets, performances statistically lower than the maximum are pointed by * (single tail t-tests with $df = 38$ and BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 6.3.2 for a description of the results

System	BF	DF	K
AR_2	90.09*	91.02*	75.90
	[87.37, 92.81]	[88.64, 93.39]	[68.12, 83.68]
AR_4	94.09	94.76*	77.60
	[92.61, 95.58]	[92.61, 96.92]	[68.77, 86.43]
AR_6	94.93	97.12	76.94
	[93.39, 96.47]	[96.12, 98.12]	[68.98, 84.90]
AR_8	96.01	96.82	79.10
	[95.04, 96.99]	[95.66, 97.97]	[72.03, 86.18]
AR_{10}	95.67	96.48	80.31
	[94.26, 97.07]	[94.73, 98.23]	[73.33, 87.29]
AR_{15}	95.83	97.43	79.75
	[94.36, 97.30]	[96.25, 98.62]	[72.66, 86.85]
AR_{20}	**96.10**	96.78	79.97
	[95.05, 97.15]	[94.79, 98.76]	[72.81, 87.12]
AR_{25}	95.34	97.95	79.31
	[93.99, 96.68]	[96.77, 99.13]	[72.21, 86.41]
AR_{30}	95.64	97.51	**80.45**
	[94.58, 96.71]	[95.94, 99.08]	**[73.53, 87.38]**
AR_{40}	94.44	**98.04**	79.60
	[93.08, 95.80]	**[97.03, 99.05]**	[73.18, 86.01]
AR_{50}	93.17*	97.82	77.97
	[91.50, 94.85]	[96.63, 99.01]	[71.10, 84.84]

Table A.38: Accuracy results of LPC_N systems (Part B). Mean and 95% CI accuracy values obtained with different values of N. Within data sets, performances statistically lower than the maximum are pointed by * (single tail t-tests with $df = 38$ and BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 6.3.2 for a description of the results

System	P	Y	Z
AR_2	85.87	69.88	89.98*
	[81.20, 90.55]	[63.99, 75.77]	[87.34, 92.62]
AR_4	87.51	73.69	91.19
	[83.54, 91.48]	[68.22, 79.16]	[88.00, 94.38]
AR_6	87.94	74.17	93.92
	[83.45, 92.43]	[69.49, 78.85]	[92.49, 95.35]
AR_8	89.88	74.73	**93.93**
	[85.49, 94.26]	[69.57, 79.88]	**[92.19, 95.66]**
AR_{10}	88.41	75.17	92.03
	[84.33, 92.49]	[70.21, 80.13]	[89.35, 94.71]
AR_{15}	**89.95**	75.78	91.93
	[86.95, 92.95]	[70.76, 80.80]	[89.53, 94.34]
AR_{20}	88.88	**76.28**	90.08
	[85.87, 91.89]	**[71.28, 81.28]**	[86.50, 93.66]
AR_{25}	88.41	75.22	90.05
	[84.18, 92.64]	[70.35, 80.10]	[86.52, 93.58]
AR_{30}	89.51	76.05	83.07*
	[86.26, 92.76]	[71.39, 80.71]	[78.77, 87.36]
AR_{40}	88.54	75.66	80.48*
	[85.16, 91.92]	[70.71, 80.60]	[76.06, 84.90]
AR_{50}	89.18	75.69	80.30*
	[85.74, 92.61]	[70.64, 80.73]	[75.02, 85.57]

Appendix B

Career development

Since at least half of my training has taken place outside the subject of this thesis, but during the time span of my Ph.D., I would like to go through those that were part of my doctoral training. Note that, although strictly outside the scope of this thesis, they are research related exercises and therefore had a direct influence on the thesis. Finally, I will also enumerate the work directly related to this thesis.

B.1 Non-thesis specific development

In September 2008, I enrolled in the Division of Digital Signal Processing at University Institute for Technological Development and Innovation in Communications (IDeTIC), University of Las Palmas de Gran Canaria (ULPGC) (Spain) as a researcher trainee for the development of a face recognition biometric system applied to videos, under the supervision of Dr Carlos M. Travieso. This led in January 2010 to my B.S. final year project, titled "Study, design and implementation of a biometric face recognition system for TV videos. BioSFaRV (Biometric System for Face Recognition from Videos)", which in turn resulted in my first publication [162].

I was then fortunate to continue my training with Travieso, now supervisor of my Ph.D. thesis. I undertook my my M.Sc. on Intelligent Systems and Numerical

Applications in Engineering at IUSIANI, ULPGC (Spain) between September 2009 and December 2010. Having used an existing face detection tool with BioSFaRV, I had the opportunity to develop a tool of my own as a master thesis titled "Study, design and implementation of a face detection system based on color contours", under the supervision of Dr. Modesto Castrillón Santana.

At the same time, I continued working on the problem of face biometrics with Travieso, mainly as part of the research project "Research of Face Recognition Biometric Systems Applied to the Hotel Industry", financed under the "Programa Innova Canarias 2020". This culminated in several publications in international conferences and journals as well as book chapters [163–173].

My training continued with the supervision of several students during their final year projects and during their trainee periods at IDeTIC. In addition, I participated in applications for research funding, as well as on the elaboration of subsequent progress reports, and I was part of the management and submission-review teams of several international scientific events and journals.

I was also able to work on some of the other biometric modalities studied in the group, which helped extend my knowledge and insight of signal processing techniques. Most notably, I had the opportunity to join the project "Laboratorio de Sistemas y Autómatas Inteligentes en Biodiversidad", financed by the Spanish Agency for International Development Cooperation (AECID), Minister of Foreign Affairs and Cooperation (MAEC), Government of Spain, focussed on the application of biometric technologies for the conservation of the biodiversity. Again, a number of publications resulted from these collaborations, including two patents [174–187].

B.2 Thesis specific development

All the training just described was vital for the execution of this thesis. The processing of EEG data posed a completely new challenge, especially considering

that the group had no direct experience with it. While exploring the required concepts and techniques, my understanding of what meticulous research entailed has evolved, and I have learned to be extremely critical with my own work.

In January 2014 I joined the Psychology Department of the College of Human Health and Science, Swansea University (Wales, UK). Under the supervision of Dr. Christoph T. Weidemann, I recorded and processed EEG data and supervised other graduate and post-graduate students. This proved to be an important step forward in my training, especially considering that this group's expertise is from the other end of the spectrum within the neuroscience disciplines. There, I further developed my understanding of the processing and analysis of EEG data, and I gained new insights into research methodologies adopted by behavioural sciences such as experimental psychology.

From the research undertaken in this thesis, the following articles have arisen:

- "Electroencephalogram subject identification: A review", published in the Journal of Expert Systems with Applications on May 2014 [159].
- "EEG Biometric Identification: A Thorough Exploration of the Time-Frequency Domain", published in the Journal of Neural Engineering on 2015 [160].
- "Real Cepstrums on Electroencephalogram Biometric Identification", published in the IC3-2015 International Conference on Contemporary Computing on August 2015 [160].
- "A Task-Independent Neural Signature: Our Brain's Fingerprint", work in progress.
- "Localized Component Filtering to Boost Electroencephalogram Artifact Rejection", work in progress.
- "My Mind, My Identity: a Task-Independent Neural Signature for Biometric Verification", work in progress.

Acronyms

ADJUST Automatic EEG artifact Detection based on the Joint Use of Spatial and Temporal features

AEP Auditory Evoked Potentials

ANN Artificial Neural Network

ANOVA analysis of variance

ApEn Approximate Entropy

APS Absolute Power Spectrum

AR Auto-Regressive

ARMA AR Moving Average

AvgMnt Common Global Average Reference Montage

BCI Brain Computer Interface

BHFDR Benjamini-Hochberg False Discovery Rate

BIHMnt Bipolar Inter-Hemispheric Reference Montage

BP Back Propagation algorithm of an ANN

BSS Blind Source Separation

CG Computational Geometry based classifier

CI Confidence Interval

COH Spectral Coherence

CV Cross-Validation

CzMnt Common Cz Reference Montage

DDBB Database

DBI Davies Bouldin Index

DWT Discrete Wavelet Transform

DZ Dizygotic

EEG Electroencephalogram

EER Equal Error Rate

ENN Elman ANN

ERP Event Related Potential

ERS Event Related Spectrogram

Eucl. Euclidean metric

FA-ANN Fuzy ARTMAP ANN

FAR False Acceptance Rate

FAPR False Acceptance Power Rate

FASTER Fully Automated Statistical Thresholding for EEG artifact Rejection

FFT Fast Fourier Transform

FIR Finite Impulse Response

FRPR False Rejection Power Rate

GA Genetic Algorithms

GAR Genuine Acceptance Rate

GMM Gaussian Mixture Model

HP High-Pass

HTER Half Total Error Rate

IC Independent Component

ICA Independent Component Analysis

ICAW Wavelet enhanced ICA

IH Inter-Hemispheric

IIR Infinite Impulse Response

JADE Joint Approximate Diagonalization of Eigenmatrices

kNN k-Nearest Neighbour

LCF Localized Component Filtering

LDA Linear Discriminant Analysis

LDC Linear Discriminant Classifier

lSVM Linear SVM

LPC Lineal Prediction Coefficients

LSP Line Spectral Pairs

LVQ-ANN Linear Vector Quantization ANN

Manh. Manhattan metric

mAR Multivariate AR

MC Monte Carlo

MFCC Mel-Frequency Cepstral Coefficients

MTL Multi Task Learning

MUSIC Multiple Signal Classification

MZ Monozygotic

Norm. Normalized

PCA Principal Component Analysis

Manh. Manhattan metric

mAR Multivariate AR

MFCC Mel-Frequency Cepstral Coefficients

MTL Multi Task Learning

MUSIC Multiple Signal Classification

MZ Monozygotic

Norm. Normalized

NR Applied for noise reduction

OCC One Class Classifier

PCA Principal Component Analysis

PRE Percentage Reduction of Error

PSD Power Spectrum Density

RC Reflection Coefficients

RCeps Real Cepstrums

REC Resting with Eyes Closed

REO Resting with Eyes Open

RB-ENN Resilient BP-ENN

RBF Radial Basis Function

ROC Receiver Operating Characteristic

SCBI Second Order Blind Separation

SD-FIR Sum Difference FIR filter

SFA-ANN Simplified FA-NN

std standard deviation

STFT Short Time Fourier Transform

Succ. Success rate

SVM Support vector machine

TAPR True Acceptance Power Rate

TRPR True Rejection Power Rate

VEP Visual Evoked Potential

List of Figures

2.1 **Number of publications per year up to 2013.** Categories: "conferences", "journals" and "other" are considered separately. 7

2.2 **Distribution of publications per category up to 2013.** A single publication can fall into multiple categories. 9

2.3 **Distribution of publications per category and per year up to 2013.** Single publication can fall into multiple categories. 9

2.4 **Individual EEGs of MZ (left) and DZ (right) twins.** Shaded areas represent variation [15]. 11

2.5 **Original experimental paradigm of Zhang's database**. Snodgrass and Vanderwart image samples (left). Representation of the paradigm (right). 14

2.6 **Fisherbrains.** K. Das' fisherbrains from two subjects from a 170 msecs post-stimulus VEP [89]. 28

2.7 **VEPs elicited by self and non-self face images [109].** 40

3.1 **Preprocessing steps applied to all databases.** Block "Art. Rej." represents an artefact rejection method (chapter 4). 53

3.2 **Selected EEG channels.** Diagram of the 10-20 International System for EEG channel locations. Highlighted channels were kept for experimentation. 58

4.1 **Diagrams of feature rejection systems based on BSS.** (Left) Common architecture of existing systems. The actual BSS component can be seamlessly interchanged and therefore it has been left out of the "Artefact Rejection Method" box, which defines how the output of BSS is processed. (Centre) Combination of the common architecture and the proposed LCF method. The only difference with the previous diagram is the embedded LCF component. All the other blocks are exactly the same as before. Hence, the simplicity of its integration within any existing system. (Right) LCF can also be used individually, without any artefact rejection method. 64

4.2 **Diagram of the LCF step.** The inputs and output are (C) the original BSS components, (L) the control signal pointing to the components that will be mixed, (P) the processed or alternative components, and (R) the resulting mixed components. 66

4.3 **Examples of F_1 (top) and F_2 (bottom) defined in the implemented LCF block.** Their integrated version $\overline{F_i}$ is smoother, allowing for a more robust detection of A_i (in the figure we expanded the amplitude of A_i to the length of the y axis). 70

4.4 **Mixing of the component C whose features are depicted in figure 4.3 when its processed version P are all zeros.** Transitions between *on* and *off* are smoothed to avoid discontinuities. The amplitude of M was modified to match the y axis. 71

4.5 **Examples of the simulated data set used on FASTER's original work.** Contaminated channels and blinks are particularly prominent. 72

4.6 **Examples of the real data set recorded from a memory based task.** Artefacts are substantially more complex than in the synthesized EEG. 73

4.7 **Graphical representation of the true and false rejection and false acceptance concepts.** 76

4.8 **Computed regions by averaging BIOSEMI sensors within each area.** Channels falling on top of the division lines are included in the cardinal regions, i.e. those with angles 0 deg, 90 deg, 180 deg and 270 deg as measured from the sagittal plane. 78

4.9 **ERPs for one participant after applying each of the systems defined in section 4.4.1.** Systems are applied both without (Original) and with a subsequent LCF step (+ LCF). The top left image corresponds to the ERPs of the original signal. Lines correspond to the brain regions specified in figure 4.8, so that the top line is RC and subsequent ones are R0D, R45D, R90D, ..., R315D. 82

4.10 **ERPs averaged across participants after applying each of the systems defined in section 4.4.1.** Systems are applied both without (Original) and with a subsequent LCF step (+ LCF). The top left image corresponds to the ERPs of the original signal. Lines correspond to the brain regions specified in figure 4.8, so that the top line is RC and subsequent ones are R0D, R45D, R90D, ..., R315D. 83

4.11 **Rejected ERPs averaged across participants after applying each of the systems defined in section 4.4.1.** Systems are applied both without (Original) and with a subsequent LCF step (+ LCF). Lines correspond to the ERPs of each brain region specified in figure 4.8, so that the top line is RC and subsequent ones are R0D, R45D, R90D, ..., R315D. 84

4.12 **Difference between the ERPs of the systems defined in section 4.4.1 without and with LCF.** Lines correspond to the ERP difference of each brain region specified in figure 4.8, so that the top line is RC and subsequent ones are R0D, R45D, R90D, ..., R315D. 85

4.13 **ERPs averaged across participants after applying each of the systems defined in section 4.4.1 without low-pass filter.** Systems are applied both without (Original) and with a subsequent LCF step (+ LCF). The top left image corresponds to the ERPs of the original signal. Lines correspond to the brain regions specified in figure 4.8, so that the top line is RC and subsequent ones are R0D, R45D, R90D, ..., R315D. 86

4.14 **Difference between the ERPs of the systems defined in section 4.4.1 without low-pass filter without and with LCF.** Lines correspond to the ERP difference of each brain region specified in figure 4.8, so that the top line is RC and subsequent ones are R0D, R45D, R90D, ..., R315D. 87

4.15 **ERSs of the central channels cluster (RC) as defined in figure 4.8.** Mean across participants, after applying each of the systems defined in the text without low-pass filter, without (Original) and with LCF (+ LCF). The top left image corresponds to the ERS of the original signal. 88

5.1 **Diagram of the experimentation methodology.** (1) We identified the configuration of the PSD that maximizes the discriminant information. (2) We find the representation of the time and frequency domains that maximizes the discriminant information. (3) We characterized the properties of the EEG discriminant information. (4) We characterize the task-independence property of the neural signature. This figure complements the information presented in tables 5.1 and 5.2. 93

5.2 **PSD with different STFT window length (L_W) and number of FFT coefficients (N_F).** The PSD, in dB scale, corresponds to one of the subjects of BCI2000-Baseline database. Within each PSD, the missing frequency band around 50 Hz (vertical axis) was filtered in the pre-processing and manually removed from the analysis. The horizontal axis (time) is piecewise-continuous and contains all the samples from the subject. 104

5.3 **Quantitative analysis of the number of spectral coefficients (N_F), the SFTF window length (L_W) and the length of the EEG signal (L_G).** Mean PRE results obtained with DEAP-Baseline database on an L_W vs N_F grid with $L_G = 10$ seconds (left) and an L_G vs L_W grid with $N_F = 128$ coefficients (right). In both cases, Θ was fixed to 0% and the results were obtained during the *full-fusion AllTask-AllSess* experiments. A maximum of 20 subjects was used in each experimental iteration. 104

5.4 **Quantitative analysis of the SFTF window overlap (Θ).** Mean PRE results on a L_W vs Θ grid with $N_F = 128$ and L_G equal to the maximum EEG available. Results correspond to BCI2000-Tasks (left) and Keirn's (right) data sets, during *full-fusion AllTask-AllSess* experiments. A maximum of 20 subjects was used in each experimental iteration. 107

5.5 **Quantitative analysis of the maximum (F_{max}) and minimum (F_{min}) cut-off frequencies.** Mean and 95% CI of PRE results obtained in *full-fusion AllTask-AllSess* experiments with different F_{max} (left) and F_{min} (right) values, corresponding to *Conf-HalfLen* system. A maximum of 20 subjects was used in each experimental iteration. 108

5.6 **Quantitative analysis of the STFT window length (L_W).** Mean PRE and 95% CI (shaded area) obtained with different L_W. The remaining parameters were set to their optimal values according to previous experiments. Data correspond to *full-fusion* set-up in *AllTask-AllSess* experiments with raw databases (top) and after the application of ADJUST (bottom-left) and LCF (bottom-right) processing. A maximum of 20 subjects was used in each experimental iteration. Refer to figure 5.5 for details on the legend. 111

5.7 **Quantitative analysis of the EEG segment length (L_G).** Mean PRE and 95% CI (shaded area) obtained with different L_G. Data correspond to *full-fusion* set-up in *AllTask-AllSess* experiments using *Conf-1s* (left) and *Conf-2s* (right) systems. A maximum of 20 subjects was used in each experimental iteration. Refer to figure 5.5 for details on the legend. 113

5.8 **Qualitative analysis of EEG montages.** PSD (in dB) of a subject from Yeom's database computed with BIHMnt (top), AvgMnt (bottom-left) and CzMnt (bottom-right). The colour bar-code above each PSD represents changes in channels, in the piecewise-continuous time axis. Each channel's spectrum (H) is added to the right of the spectrogram using the same colour scheme as the channel bar-code. 116

5.9 **Quantitative analysis of EEG montages with raw data sets.** Relative PRE values between *BIHMnt-AvgMnt* and between *CzMnt-AvgMnt*. Boxes show results stacked across databases. Box limits are 25 and 95 percentiles, while black bars shows maximum and minimum values after excluding outliers (red crosses). The red line within each box and triangle markers show median values and their 95% CI. To the right of each box, corresponding mean PRE values from each database are shown following the legend of figure 5.5. A maximum of 20 subjects was used in each experimental iteration. 117

5.10 **Qualitative analysis of PSD normalization.** *Conf-2s* PSD, in dB, of a subject from P. Ullsperger's database. *prcNorm*, *iqrNorm*, *zNorm* and *rNorm* are cubic root scaled (c.r.) instead, as they contained negative values. 119

5.11 **Quantitative analysis of PSD normalization.** Relative PRE values between the PSD normalized by each of the methods in table 5.3 and the raw PSD. Results are stacked across databases. A maximum of 20 subjects was used in each experimental iteration. Refer to caption of figure 5.9 for details about the meaning of symbols within the image. 120

5.12 **Quantitative analysis of the spatial distribution of the discriminant information.** Mean PRE values obtained at each location with the REO condition of BCI2000 (top) and DEAP (bottom) data sets when applying the systems *Conf-1s* (left), *Conf-1s rNorm* (centre) and *Conf-FullLen* (right). Results correspond to *freq-fusion ch-focus* experiments, with a maximum of 20 subjects used in each experimental iteration. Refer to fig. A.10 for further related results. 122

5.13 **Quantitative analysis of the discrimination power frequency distribution.** Mean PRE and 95% CI (shaded area) obtained with each frequency (*ch-fusion* experiments) of raw and rNorm PSD. Curves were smoothed by local regression, using weighted linear least squares and a first degree polynomial model with a 3 Hz span. A maximum of 20 subjects was used in each experimental iteration. Refer to figure 5.5 for details on the legend. 124

5.14 **Qualitative analysis of uniqueness.** PSD, in dB, of 4 subjects from Yeom's database, corresponding to the EEG sensor P3 and a single task (self-representation). 126

5.15 **Quantitative analysis of uniqueness.** Mean PRE and 95% CI (shaded area) for different numbers of subjects (N_S) in the system and each database. Results correspond to the *rNorm* system. Refer to figure 5.5 for details on the legend. . 126

5.16 **Qualitative analysis of permanence.** PSD in dB of tasks T1 and T2 of one subject from Keirn's database. The colour bar-codes above each spectrogram represent changes in channels and sessions. PSD for each channel and session is attached to the right of the spectrograms, using the same colour scheme as the channel bar-codes. 128

5.17 **Quantitative analysis of permanence.** Relative PRE values between normal and *Sess-CV* experiments, obtained with the raw (top) and the *rNorm* (bottom) PSD of Keirn's (K) and Yeom's (Y) databases. Note that for Keirn's data set, *Conf-HalfLen* is equivalent to *Conf-2s*. Refer to the caption of figure 5.9 for details on box markings within the image. 128

5.18 **Qualitative analysis of task-independence.** Spectral representation, in dB, of four subjects from BCI2000-Full. The signal corresponds to the C3 sensor, and was obtained with *Conf-1s* system. The top colour bar-code represent changes in tasks. Each task's PSD is attached to the right of the spectrograms, using the same colour scheme as the task bar-code. 132

5.19 **Quantitative analysis of the task-independence property.** Difference between PRE values of *Task-CV* and regular CV *full-fusion* experiments. A maximum of 20 subjects was used in each experimental iteration. Refer to caption of figure 5.9 for details on box markings within the image. 133

5.20 **Quantitative analysis of the task-independence property with *rNorm* systems.** Difference between PRE values of *Task-CV* and regular CV *full-fusion* experiments. A maximum of 20 subjects was used in each experimental iteration. Refer to caption of figure 5.9 for details on box markings within the image. . . . 134

5.21 ***Task-CV* and *1Task-AllSess* comparison.** Difference in PRE values between *Task-CV* and *1Task-AllSess full-fusion* experiments. *1Task-AllSess* confusion matrices were added across cognitive tasks. A maximum of 20 subjects was used in each experimental iteration. Refer to caption of figure 5.9 for details on box markings within the image. 135

5.22 ***Task-CV* and *1Task-AllSess* comparison with *rNorm* systems.** Difference in PRE values between *Task-CV* and *1Task-AllSess full-fusion* experiments. *1Task-AllSess* confusion matrices were added across cognitive tasks. A maximum of 20 subjects was used in each experimental iteration. Refer to caption of figure 5.9 for details on box markings within the image. 136

5.23 **Qualitative analysis of the permanence of task-independence.** Spectral representation, in dB, of four subjects from Keirn's database. The signal corresponds to the C3 sensor, and was obtained with *Conf-1s* system. The colour bar-codes above the PSDs represent changes in tasks and sessions. PSD for each task and session is attached to the right of the spectrograms, using the same colour scheme as the task bar-code. 137

6.1 **ROC curves of the baseline design.** Mean ROC curves and std (shaded area) obtained with the raw (left) and *rNorm* (right) PSD coefficients. A maximum of 20 subjects was used in each experimental iteration. The dotted red diagonal line represents chance accuracy levels, while the black dashed diagonal line represents EER points. 145

6.2 **ROC curves of the baseline design with logarithmic FAR axis.** Mean ROC curves and std (shaded area) obtained with the raw (left) and *rNorm* (right) PSD coefficients. A maximum of 20 subjects was used in each experimental iteration. Refer to fig. 6.1 for legend details. 145

6.3 **Performance of RCeps$_{P\%}$ system with increasing number of coefficients.** Mean accuracy values and 95% CI (shaded area) obtained with the first P% of the real raw (left) and *rNorm* (right) cepstral coefficients. A maximum of 20 subjects was used in each experimental iteration. Refer to fig. 6.1 for legend details. . . . 147

6.4 **ROC curves of RCeps$_{P\%}$ systems coefficients with logarithmic FAR axis.** Mean ROC curves and std (shaded area) obtained with different P values. A maximum of 20 subjects was used in each experimental iteration. Refer to fig. 6.1 for legend details. . . . 148

6.5 **Performance of LPC$_N$ systems.** Mean and 95% CI (shaded area) verification accuracy obtained with different values of N. A maximum of 20 subjects was used in each experimental iteration. Refer to fig. 6.1 for legend details. . . . 150

6.6 **ROC curves of LPC$_N$ systems coefficients with logarithmic FAR axis.** Mean ROC curves and std (shaded area) obtained with different N. A maximum of 20 subjects was used in each experimental iteration. Refer to fig. 6.1 for legend details. . . . 150

6.7 **ROC curves for different spectral envelope coefficients.** Mean and std (shaded area) ROC curves obtained with LPC$_8$, RC$_8$, LSP$_8$ and ϵ_8. A maximum of 20 subjects was used in each experimental iteration. Refer to fig. 6.1 for legend details. . . . 152

6.8 **ROC curves obtained when testing with all the available subjects.** Mean ROC curves and std (shaded area) of RCeps$_{20\%}$ and LPC$_8$ systems. Refer to fig. 6.1 for legend details. . . . 158

6.9 **ROC curves with logarithmic FAR axis obtained when testing with all the available subjects.** Mean ROC curves and std (shaded area) of RCeps$_{20\%}$ and LPC$_8$ systems. Refer to fig. 6.1 for legend details. . . . 159

A.1 **Mean PRE results obtained with each database on an L_W vs N_F grid.** L_G was set to the length of the available EEG on each database (table 3.2) and $\Theta = 0\%$. Results correspond to *full-fusion no-focus* experiments. A maximum of 20 subjects was used in each experimental iteration. Refer to section 5.2.1 for a description of the results. . . . 176

A.2 **Mean PRE results obtained with each database on an L_G vs L_W grid.** In all cases, $N_F = 128$ coefficients and $\Theta = 0\%$. Results correspond to the *full-fusion no-focus* experiments. A maximum of 20 subjects was used in each experimental iteration. Refer to section 5.2.1 for a description of the results. . . . 178

A.3 **μPRE results obtained with each database on an L_W vs Θ grid.** L_G was set to the length of the available EEG on each database (table 3.2) and $N_F = 128$ coefficients. Results correspond to *full-fusion no-focus* experiments. A maximum of 20 subjects was used in each experimental iteration. Refer to section 5.2.1 for a description of the results. 179

A.4 **Quantitative analysis of the maximum (F_{max}) and minimum (F_{min}) cut-off frequencies with artefact free data sets.** Mean and 95% CI of PRE results obtained in *full-fusion no-focus* experiments with different F_{max} (top) and F_{min} (bottom) values. Results correspond to ADJUST (left) and LCF (right) pre-processed databases, tested with the *Conf-HalfLen* system. A maximum of 20 subjects was used in each experimental iteration. Refer to figure 5.5 for details on the legend, and to section 5.2.3 for a description of the results. 182

A.5 **Quantitative analysis of the EEG segment length (L_G) with artefact free data sets.** Mean PRE and 95% CI (shaded area) obtained with different L_G. Data corresponds to *full-fusion* set-up on *no-focus* experiments using *Conf-1s* (left) and *Conf-2s* (right) systems. A maximum of 20 subjects was used in each experimental iteration. Refer to figure 5.7 for details on the legend, and to section 5.2.5 for a description of the results. 188

A.6 **Quantitative analysis of EEG montages with ADJUST processed data sets.** Relative PRE values between *BIHMnt-AvgMnt* and between *CzMnt-AvgMnt*. Results are stacked across databases. Refer to caption of figure 5.9 for details about the meaning of symbols within the image, and to section 5.3.1 for a description of the results. 191

A.7 **Quantitative analysis of EEG montages with LCF processed data sets.** Relative PRE values between *BIHMnt-AvgMnt* and between *CzMnt-AvgMnt*. Results are stacked across databases. Refer to caption of figure 5.9 for details about the meaning of symbols within the image, and to section 5.3.1 for a description of the results. 192

A.8 **Quantitative analysis of PSD normalization with ADJUST processed data sets.** Relative PRE values between the PSD normalized by each of the methods in table 5.3 and the raw PSD. Results are stacked across databases. A maximum of 20 subjects was used in each experimental iteration. Refer to caption of figure 5.9 for details about the meaning of symbols within the image, and to section 5.3.2 for a description of the results. 201

A.9 **Quantitative analysis of PSD normalization with LCF processed data sets.** Relative PRE values between the PSD normalized by each of the methods in table 5.3 and the raw PSD. Results are stacked across databases. A maximum of 20 subjects was used in each experimental iteration. Refer to caption of figure 5.9 for details about the meaning of symbols within the image, and to section 5.3.2 for a description of the results. 206

A.10 **Quantitative analysis of the spatial distribution of the discriminant information.** Mean PRE values obtained at each location with the REO condition of BCI2000 (left) and DEAP (right) data sets when applying each system with and without *rNorm*. Results correspond to *freq-fusion ch-focus* experiments, with a maximum of 20 subjects used in each experimental iteration. Refer to section 5.4.1 for a description of the results. 211

A.11 **Quantitative analysis of the discrimination power frequency distribution with artefact free data sets.** Mean PRE and 95% CI (shaded area) obtained with each frequency (*ch-fusion* experiments). Curves where smoothed by local regression, using weighted linear least squares and a first degree polynomial model with a 3 Hz span. Graphs correspond to results obtained with the raw PSD of ADJUST (top) and LCF (bottom) processed databases. A maximum of 20 subjects was used in each experimental iteration. Refer to figure 5.5 for details on the legend, and to section 5.4.2 for a description of the results. 212

A.12 **Quantitative analysis of uniqueness with artefact free data sets.** Mean PRE and 95% CI (shaded area) for different numbers of subjects (N_S) in the system and each database (refer to figure 5.5 for details on the legend). Graphs correspond to results obtained with the raw PSD of ADJUST (top) and LCF (bottom) processed databases. Refer to figure 5.5 for details on the legend, and to section 5.4.3 for a description of the results. 213

A.13 **ROC curves of RCeps$_{P\%}$ systems coefficients with logarithmic FAR axis.** Mean ROC curves and std (shaded area) obtained with different configurations of RCeps$_{P\%}$. A maximum of 20 subjects was used in each experimental iteration. Refer to fig. 6.1 for legend details. 224

List of Tables

2.1 **Databases used on multiple publications of the relevant literature.** . . 13

2.2 **Systems used on Poulos' database.** Columns specify: the name given to the system for future reference (Sys.), filtered frequency bands (Freq. [Hz]), extracted features (Feat.), applied classifiers (Cls.) and references of papers that used each design (Ref.). 20

2.3 **Results with Poulos' database.** Success rate (Succ.) of systems in table 2.2 on classification and verification experiments. The percentage of training samples and/or the Cross-Validation (CV) method used on each case is specified (Tr.% / CV). In this case, training percentage is specified as $SetR\%$ | $SetX\%$ (section 2.3.1). For works [74] and [77], only the best results are shown. 21

2.4 **Systems used on Zhang's database.** Along with those presented on table 2.5. Refer to table 2.2 for details of the columns' nomenclature. 24

2.5 **Systems used on Zhang's database.** Along with those presented on table 2.4. Refer to table 2.2 for details of the columns' nomenclature. 25

2.6 **Results with Zhang's database.** Success rate of systems in tables 2.4 and 2.5 evaluated on classification with 20 subjects. For systems with configurable parameters accuracy ranges are given. Refer to table 2.3 for details of the columns' nomenclature. 26

2.7 **Results with Zhang's database.** Success rate of systems in tables 2.4 and 2.5 evaluated on classification with 40 subjects. For systems with configurable parameters, accuracy ranges are given. Refer to table 2.3 for details of the columns' nomenclature. 27

2.8 **Results with Zhang's database.** Success rate of systems in tables 2.4 and 2.5 evaluated on classification with 10, 102, 120 and 122 subjects. For systems with configurable parameters accuracy ranges are given. Refer to table 2.3 for details of the columns' nomenclature. 28

2.9 **Systems used on Keirn's database.** All these systems share a common high-pass pre-filtering stage at 0.2 Hz using an elliptic filter. In addition, the EEG signal separated into 20 segments of 0.5 seconds windows. Refer to table 2.2 for details of the columns' nomenclature. 30

2.10 **Results with Keirn's database.** Success rate and EER of systems in table 2.9 evaluated on classification and verification experiments with a maximum of 5 subjects. In both cases, a CV method using 50% of the database for training and 50% for testing was applied. Results of individual and combined tasks are presented. Refer to table 2.3 for details of the columns' nomenclature. 31

2.11 **Results with Keirn's database.** Success rate of systems in table 2.9 evaluated on classification with a maximum of 5 subjects. A CV method using 50% of the database for training and 50% for testing was applied. Refer to table 2.3 for details of the columns' nomenclature. 32

2.12 **Systems used on IIIaBCI03 database.** Systems 1 to 6 used only channels C3, C4, P3, P4, O1 and O2. Refer to table 2.2 for details of the columns' nomenclature. 33

2.13 **Results with IIIaBCI03 database.** Success rate and EER of systems in table 2.12 evaluated on classification and verification experiments. A CV method using 50% of the database for training and 50% for testing was applied. 33

2.14 **Results with IIIaBCI03 database.** Success rate (Succ.) rate and EER of systems in table 2.12 evaluated on classification and verification experiments. A CV method using 50% of the database for training and 50% for testing was applied. These were MTL experiments; i.e. systems used EEG from all the recorded paradigms to determine the user's identity. 34

2.15 **Systems and results with Tottori database.** A CV method using 50% of the database for training and 50% for testing was applied. Refer to table 2.2 for details of the columns' nomenclature. 36

2.16 **Systems used on Lanzhou database.** Refer to table 2.2 for details of the columns' nomenclature. 38

2.17 **Results with Lanzhou's database.** Success rate of systems in table 2.16 evaluated on classification and verification experiments. Refer to table 2.3 for details of the columns' nomenclature. 39

3.1 **Filtering details for each database.** The table shows: the sampling frequency (Fs) at which each database was processed, the high-pass filter order (HP order), and the frequency at which the notch filter was applied (Notch fcut) and its order (Notch order). 54

3.2 **Normalized databases used for experimentation.** Columns show the name of the database (Database), the number of subjects (Subj.), the number of tasks (Tasks), the number of sessions (Sess.), the number of trials (Trials), the length of the EEG segments in seconds (Len.) and the descriptive keywords (Keywords). 59

4.1 **Numerical results obtained with the synthetic database as detailed in section 4.4.2.** Mean and std is provided for the percentage of processed ICs (% ICs), TAPR and FAPR. 79

4.2 **Numerical results obtained with the synthetic database and no 40 Hz low-pass filter.** Mean and std is provided for the percentage of processed ICs (% ICs), TAPR and FAPR. Refer to table 4.1 to compare these with the results of the original systems (with 40 Hz low-pass filter). 85

5.1 **Phases 1 and 2 of the experimentation methodology.** This table complements the information presented in figure 5.1. 94

5.2 **Phases 3 and 4 of the experimentation methodology.** This table complements the information presented in figure 5.1. 95

5.3 **PSD normalization functions.** P is a PSD matrix with dimensions N_FxN_T, with N_T the number of time points, and $p^{n\%}_{(F)}$ is the n-% percentile applied along the frequency dimension. 97

5.4 **Fusion modes used during experimentation.** 100

5.5 **Focus modes used during experimentation.** 103

5.6 **Quantitative results for different configurations.** Mean PRE and 95% CI obtained with different configurations. Data correspond to raw databases tested during *full-fusion AllTask-AllSess* experiments. Within data sets, performances statistically different than the maximum are pointed by * (single tail t-tests with BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names. 114

5.7 **Quantitative analysis of the permanence of task-independence.** Mean PRE and 95% CI obtained with *Sess-CV* experiments and combining *Sess-CV* with *1Task-AllSess* and *Task-CV*. Note that for Keirn's database, *Conf-HalfLen* is equivalent to *Conf-2s*. Refer to table 3.2 for details on databases' code names. 138

5.8 **Quantitative analysis of the permanence of task-independence with *rNorm* systems.** Mean PRE and 95% CI obtained with *Sess-CV* experiments and combining *Sess-CV* with *1Task-AllSess* and *Task-CV*. Note that for Keirn's database, *Conf-HalfLen* is equivalent to *Conf-2s*. Refer to table 3.2 for details on databases' code names. 139

6.1 **Accuracy results of the baseline system.** Mean and 95% CI accuracy values obtained with the baseline design using the raw and *rNorm* PSD coefficients. Within data sets, performances statistically lower than the maximum are pointed by * (single tail t-tests with $df = 38$ and BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names. 144

6.2 **Results of spectral envelope coefficients.** Mean and 95% CI verification accuracy of LPC_8, RC_8, LSP_8 and the ϵ_8. Within data sets, performances statistically lower than the maximum are pointed by * (single tail t-tests with $df = 38$ and BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names. 151

6.3 **Results of selected configurations.** Mean and 95% CI results of the selected configurations of PSD, RCeps and LPC systems. A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names. 155

6.4 **Results obtained when testing with all the available subjects.** Mean and 95% CI results of $RCeps_{20\%}$ and LPC_8 systems. Although results of DEAP-Full, Keirn's, P. Ullsperger's and Yeom's data sets remained the same as before – as they contain less or equal to 20 subjects –, they are replicated here for ease of inspection. Refer to table 3.2 for details on databases' code names. 157

6.5 **Optimal ROC points obtained when testing with all the available subjects.** Refer to table 3.2 for details on databases' code names. 158

6.6 **GAR for an FAR of 1% obtained when testing with all the available subjects.** Mean and std of GAR obtained with $RCeps_{20\%}$ and LPC_8 systems. Results computed with all the available subjects for each database. Refer to table 3.2 for details on databases' code names. 159

6.7 **GAR for an FAR of 0.5% obtained when testing with all the available subjects.** Mean and std of GAR obtained with $\text{RCeps}_{20\%}$ and LPC_8 systems. Results computed with all the available subjects for each database. Refer to table 3.2 for details on databases' code names. 160

6.8 **Comparison of results obtained with BCI2000 database.** Columns correspond to the publication reference (Ref.), the number of subjects used (# Subj.), the task used (Task), the CV method applied (CV method), and the success rate (Succ.). In addition, NIt refers to the number of MC iterations. For completeness, classification results from table A.30 are also shown, these correspond to the analysis of chapter 5. 161

6.9 **Comparison of results obtained with Keirn's database.** For more results on this database, refer to tables 2.10 and 2.11. Refer to table 6.8 for a description of the table's columns. For completeness, classification results from table 5.8 are also shown, these correspond to the analysis of chapter 5. 162

6.10 **Comparison of results obtained with Yeoms's database.** Refer to table 6.8 for a description of the table's columns. For completeness, classification results from table 5.8 are also shown, these correspond to the analysis of chapter 5. . . 162

6.11 **Comparison of results obtained with Zhang's database.** For more results on these database, refer to tables tables 2.6 to 2.8. Refer to table 6.8 for a description of the table's columns. For completeness, classification results from table A.30 are also shown, these correspond to the analysis of chapter 5. 163

A.1 **Independent t-test analysis of the $N_F \approx L_W * F_s$ diagonal**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) > \mu_{PRE}(B)$. With A and B representing the conditions above and below the diagonal $N_F \approx L_W * F_s$ respectively. Refer to table 3.2 for details on databases' code names, and to section 5.2.1 for a description of the results. 177

A.2 **Independent t-test analysis of F_{max} and F_{min} with *Conf-HalfLen* system**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B conditions specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.2.3 for a description of the results. . 180

A.3 **Independent t-test analysis of F_{max} and F_{min} with *Conf-FullLen* system**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B conditions specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.2.3 for a description of the results. . 181

A.4 **Independent t-test analysis of F_{max} and F_{min} with ADJUST processed data sets**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B conditions specified on top of each cell as $(A) - (B)$. Refer to section 5.2.3 for a description of the results. 183

A.5 **Independent t-test analysis of F_{max} and F_{min} with *Conf-FullLen* system and ADJUST processed data sets**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B conditions specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.2.3 for a description of the results. 184

A.6 **Independent t-test analysis of F_{max} and F_{min} with *Conf-HalfLen* system and LCF processed data sets**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B conditions specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.2.3 for a description of the results. 185

A.7 **Independent t-test analysis of F_{max} and F_{min} with *Conf-FullLen* system and LCF processed data sets**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B conditions specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.2.3 for a description of the results. 186

A.8 **Independent t-test analysis of L_W**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B conditions specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.2.4 for a description of the results. 187

A.9 **Quantitative results for different configurations with ADJUST processed data sets.** Mean PRE and 95% CI obtained with different configurations. Data corresponds *full-fusion no-focus* experiments. Within data sets, performances statistically different than the maximum are pointed by * (single tail t-tests with BHFDR adjusted $p > 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 5.2.6 for a description of the results. 189

A.10 **Quantitative results for different configurations with LCF processed data sets.** Mean PRE and 95% CI obtained with different configurations. Data corresponds to *full-fusion no-focus* experiments. Within data sets, performances statistically different than the maximum are pointed by * (single tail t-tests with BHFDR adjusted $p > 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 5.2.6 for a description of the results. 190

A.11 **Independent t-test analysis of EEG montages with *Conf-HalfLen* system**. H_0: $\mu_{PRE}(A) - \mu_{PRE}(B) \geq 10$. H_1: $\mu_{PRE}(A) - \mu_{PRE}(B) < 10$. With A and B montages specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.1 for a description of the results. 193

A.12 **Independent t-test analysis of EEG montages with *Conf-FullLen* system**. H_0: $\mu_{PRE}(A) - \mu_{PRE}(B) \geq 10$. H_1: $\mu_{PRE}(A) - \mu_{PRE}(B) < 10$. With A and B montages specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.1 for a description of the results. 194

A.13 **Independent t-test analysis of EEG montages with *Conf-1s* system**. H_0: $\mu_{PRE}(A) - \mu_{PRE}(B) \geq 5$. H_1: $\mu_{PRE}(A) - \mu_{PRE}(B) < 5$. With A and B montages specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.1 for a description of the results. . 195

A.14 **Independent t-test analysis of EEG montages with *Conf-2s* system**. H_0: $\mu_{PRE}(A) - \mu_{PRE}(B) \geq 5$. H_1: $\mu_{PRE}(A) - \mu_{PRE}(B) < 5$. With A and B montages specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.1 for a description of the results. . 196

A.15 **Independent t-test analysis of PSD normalization with *Conf-HalfLen* system**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results. 197

A.16 **Independent t-test analysis of PSD normalization with *Conf-FullLen* system**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to section 5.3.2 for a description of the results. 198

A.17 **Independent t-test analysis of PSD normalization with *Conf-1s* system**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results. 199

A.18 **Independent t-test analysis of PSD normalization with *Conf-2s* system**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results. 200

A.19 **Independent t-test analysis of PSD normalization with *Conf-HalfLen* system and ADJUST processed data sets**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results. 202

A.20 **Independent t-test analysis of PSD normalization with *Conf-FullLen* system and ADJUST processed data sets**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results. 203

A.21 **Independent t-test analysis of PSD normalization with *Conf-1s* system and ADJUST processed data sets**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results. 204

A.22 **Independent t-test analysis of PSD normalization with *Conf-2s* system and ADJUST processed data sets**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to section 5.3.2 for a description of the results. 205

A.23 **Independent t-test analysis of PSD normalization with *Conf-HalfLen* system and LCF processed data sets**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results. 207

A.24 **Independent t-test analysis of PSD normalization with *Conf-FullLen* system and LCF processed data sets**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results. 208

A.25 **Independent t-test analysis of PSD normalization with *Conf-1s* system and LCF processed data sets**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results. 209

A.26 **Independent t-test analysis of PSD normalization with *Conf-2s* system and LCF processed data sets**. H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) \neq \mu_{PRE}(B)$. With A and B normalization methods (table 5.3) specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.3.2 for a description of the results. 210

A.27 **Independent t-test of permanence with raw data sets.** H_0: $\mu_{PRE}(A) = \mu_{PRE}(B)$. H_1: $\mu_{PRE}(A) > \mu_{PRE}(B)$. With A the regular CV procedure (no *sess-CV*) and B the *sess-CV* procedure. Refer to table 3.2 for details on databases' code names, and to section 5.4.4 for a description of the results. 214

A.28 **Independent t-test of task-independence with raw data sets.** H_0: $\mu_{PRE}(A) - \mu_{PRE}(B) \geq 5$. H_1: $\mu_{PRE}(A) - \mu_{PRE}(B) < 5$. A and B are specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.5.1 for a description of the results. 215

A.29 **Independent t-test of task-independence with artefact free data sets.** H_0: $\mu_{PRE}(A) - \mu_{PRE}(B) \geq 5$. H_1: $\mu_{PRE}(A) - \mu_{PRE}(B) < 5$. A and B are specified on top of each cell as $(A) - (B)$. Refer to table 3.2 for details on databases' code names, and to section 5.5.1 for a description of the results. 216

A.30 **Quantitative results of task-independence with raw data sets.** Mean PRE and 95% CI obtained in *task-CV* experiments. A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 5.5.1 for a description of the results. 217

A.31 **Quantitative results of task-independence with artefact free data sets.** Mean PRE and 95% CI obtained in *task-CV* experiments. A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 5.5.1 for a description of the results. . . . 218

A.32 **Independent t-test of the permanence of the task-independence property with raw data sets.** H_0: $\mu_{PRE}(A) - \mu_{PRE}(B) \leq C$. H_1: $\mu_{PRE}(A) - \mu_{PRE}(B) > C$. With A and B corresponding to conditions *Task-Sess-CV* and *Task-CV* respectively. The threshold C is set to 10 for Keirn's database (K) and to 1 for Yeom's database (Y). Refer to section 5.5.2 for a description of the results. 219

A.33 **Accuracy results of RCeps$_{P\%}$ systems and raw PSD coefficients (Part A).** Mean and 95% CI accuracy values obtained with different values of P. Within data sets, performances statistically lower than the maximum are pointed by * (single tail t-tests with $df = 38$ and BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 6.3.1 for a description of the results . . 220

A.34 **Accuracy results of RCeps$_{P\%}$ systems and raw PSD coefficients (Part B).** Mean and 95% CI accuracy values obtained with different values of P. Within data sets, performances statistically lower than the maximum are pointed by * (single tail t-tests with $df = 38$ and BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 6.3.1 for a description of the results . . 221

A.35 **Accuracy results of RCeps$_{P\%}$ systems and *rNorm* PSD coefficients (Part A).** Mean and 95% CI accuracy values obtained with different values of P. Within data sets, performances statistically lower than the maximum are pointed by * (single tail t-tests with $df = 38$ and BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 6.3.1 for a description of the results 222

A.36 **Accuracy results of RCeps$_{P\%}$ systems and *rNorm* PSD coefficients (Part B).** Mean and 95% CI accuracy values obtained with different values of P. Within data sets, performances statistically lower than the maximum are pointed by * (single tail t-tests with $df = 38$ and BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 6.3.1 for a description of the results 223

A.37 **Accuracy results of LPC$_N$ systems (Part A).** Mean and 95% CI accuracy values obtained with different values of N. Within data sets, performances statistically lower than the maximum are pointed by * (single tail t-tests with $df = 38$ and BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 6.3.2 for a description of the results 225

A.38 **Accuracy results of LPC$_N$ systems (Part B).** Mean and 95% CI accuracy values obtained with different values of N. Within data sets, performances statistically lower than the maximum are pointed by * (single tail t-tests with $df = 38$ and BHFDR adjusted $p < 0.05$). A maximum of 20 subjects was used in each experimental iteration. Refer to table 3.2 for details on databases' code names, and to section 6.3.2 for a description of the results 226

Bibliography

[1] J. LeDoux, *Synaptic Self: How Our Brains Become Who We Are*. Penguin Publishing Group, 2003.

[2] D. Oyserman, K. Elmore, and G. Smith, *Handbook of self and identity*, ch. Self, Self-Concept, and Identity, pp. 69–104. New York: Guilford Press., second ed., 2012.

[3] *Collins Concise English Dictionary*, ch. definition of "identity".

[4] T. Collura, "History and evolution of electroencephalographic instruments and techniques," *Journal of Clinical Neurophysiology*, vol. 10, no. 4, pp. 476–504, 1993.

[5] H. Begleiter and B. Porjesz, "Genetics of human brain oscillations," *International Journal of Psychophysiology*, vol. 60, pp. 162–171, May 2006.

[6] B. Zietsch, J. Hansen, N. Hansell, G. Geffen, N. Martin, and M. Wright, "Common and specific genetic influences on eeg power bands delta, theta, alpha, and beta," *Biological Psychology*, vol. 75, pp. 154–164, January 2007.

[7] *Cognitive Biometrics: Challenges for the Future*, vol. 92, 2010.

[8] K. Revett, "Cognitive biometrics: a novel approach to person authentication," *International Journal of Cognitive Biometrics*, vol. 1, no. 1, pp. 1–9, 2012.

[9] W. Khalifa, A. Salem, M. Roushdy, and K. Revett, "A survey of eeg based user authentication schemes," in *8th International Conference on Informatics and Systems (INFOS)*, May 2012.

[10] Y. Singh, S. Singh, and A. Ray, "Bioelectrical signals as emerging biometrics: Issues and challenges," *ISRN Signal Processing*, vol. 45, no. 7, p. 13, 2012.

[11] P. Campisi and D. La Rocca, "Brain waves for automatic biometric based user recognition," 2014.

[12] M. Abo-Zahhad, S. Ahmed, and S. Abbas, "State-of-the-art methods and future perspectives for personal recognition based on electroencephalogram signals," *IET Biometrics*, vol. 4, no. 3, pp. 179–190, 2015.

[13] F. Vogel, *Genetics and the Electroencephalogram*. Springer Verlag, 2000.

[14] H. Davis and P. Davis, "Action potentials of the brain: In normal persons and in normal states of cerebral activity," *Archives of Neurology And Psychiatry*, vol. 36, no. 6, pp. 1214–1224, 1936.

[15] H. Stassen, "Electroencephalogram (eeg) and evoked potentials." `http://www.els.net`, July 2006.

[16] F. Vogel, "The genetic basis of the normal human electroencephalogram (eeg)," *Human Genetic*, vol. 10, no. 2, pp. 91–114, 1970.

[17] L. De Gennaro, C. Marzano, F. Fratello, F. Moroni, M. Pellicciari, F. Ferlazzo, S. Costa, A. Couyoumdjian, G. Curcio, E. Sforza, A. Malafosse, L. Finelli, P. Pasqualetti, M. Ferrara, M. Bertini, and P. Rossini, "The electroencephalographic fingerprint of sleep is genetically determined: a twin study.," *Annals of Neurology*, vol. 64, pp. 455–60, October 2008.

[18] S. Eischen, J. Luckritz, and J. Polich, "Spectral analysis of eeg from families," *Biological Psychology*, vol. 41, no. 1, pp. 61–8, 1995.

[19] H. Stassen, G. Bomben, and D. Hell, "Familial brain wave patterns: study of a 12-sib family," *Psychiatric Genetics*, vol. 8, no. 3, pp. 141–154, 1998.

[20] J. Yung, M. Lader, and G. Fenton, "A twin study of the genetic influences on the electroencephalogram," *Journal of Medical Genetics*, vol. 9, pp. 13–16, March 1972.

[21] C. van Beijsterveldt and G. van Baal, "Twin and family studies of the human electroencephalogram: a review and a meta-analysis," *Biological Psychology*, vol. 61, pp. 111–138, October 2002.

[22] H. Stassen, "Computerized recognition of persons by {EEG} spectral patterns," *Electroencephalography and Clinical Neurophysiology*, vol. 49, no. 12, pp. 190–194, 1980.

[23] H. Stassen, G. Bomben, and P. Propping, "Genetic aspects of the eeg: An investigation into the within-pair similarity of monozygotic and dizygotic twins with a new method of analysis," *Electroencephalography and Clinical Neurophysiology*, vol. 66, pp. 489–501, 1995.

[24] H. Stassen, D. Lykken, P. Propping, and G. Bomben, "Genetic determination of the human eeg. survey of recent results on twins reared together and apart," *Human Genetics*, vol. 80, no. 2, pp. 165–76, 1988.

[25] J. Varner, R. Potter, and J. Rohrbaugh, "A procedure for automatic classification of eeg genetic variants," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 13, pp. 451–452, November 1991.

[26] M. Doppelmayr, W. Klimesch, T. Pachinger, and B. Ripper, "Individual differences in brain dynamics: important implications for the calculation of event-related band power.," *Biological Cybernetics*, vol. 79, no. 1, pp. 49–57, 1998.

[27] M. Näpflin, M. Wildi, and J. Sarnthein, "Testretest reliability of resting eeg spectra validates a statistical signature of persons," *Clinical Neurophysiology*, vol. 118, pp. 2519–2524, November 2007.

[28] X. Zhang, H. Begleiter, B. Porjesz, W. Wang, and A. Litke, "Event related potentials during object recognition tasks," *Brain Research Bulletin*, vol. 38, no. 6, pp. 531–8, 1995.

[29] X. Zhang, H. Begleiter, B. Porjesz, and A. Litke, "Electrophysiological evidence of memory impairment in alcoholic patients," *Biol Psychiatry*, vol. 42, no. 12, pp. 1157–71, 1997.

[30] X. Zhang, "Public database." `http://archive.ics.uci.edu/ml/datasets/EEG+Database`, 1995.

[31] J. Snodgrass and M. Vanderwart, "A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity," *Journal of experimental psychology Human learning and memory*, vol. 6, no. 2, pp. 174–215, 1980.

[32] R. Palaniappan and P. Raveendran, "Individual identification technique using visual evoked potential signals," *Electronics Letters*, vol. 38, no. 25, pp. 1634–1635, 2002.

[33] R. Palaniappan and K. Ravi, "A new method to identify individuals using signals from the brain," in *Proceedings of the 2003 Joint Conference of the Fourth International Conference on Information, Communications and Signal Processing, 2003 and Fourth Pacific Rim Conference on Multimedia*, vol. 3, pp. 1442–1445, 2003.

[34] R. Palaniappan, "Method of identifying individuals using VEP signals and neural network," in *IEEE Proceedings on Science, Measurement and Technology*, vol. 151, pp. 16–20, 2004.

[35] K. Ravi and R. Palaniappan, "Recognising individuals using their brain patterns," in *Third International Conference on Information Technology and Applications, 2005. ICITA 2005*, vol. 2, pp. 520–523, 2005.

[36] R. Palaniappan and D. Mandic, "Energy of brain potentials evoked during visual stimulus: a new biometric?," in *Proceedings of the 15th International Conference on Artificial Neural*

Networks: formal models and their applications, ICANN'05, vol. 2, pp. 735–740, Springer-Verlag, 2005.

[37] K. Ravi and R. Palaniappan, "Leave-one-out authentication of persons using 40 hz eeg oscillations," in *The International Conference on Computer as a Tool, EUROCON'05*, vol. 2, pp. 1386–1389, 2005.

[38] R. Ravi and R. Palaniappan, "Neural network classification of late gamma band electroencephalogram features," *Soft Computing*, vol. 10, no. 2, pp. 163–169, 2006.

[39] R. Palaniappan and K. Ravi, "Improving visual evoked potential feature classification for person recognition using pca and normalization," *Pattern Recognition Letters*, vol. 27, no. 7, pp. 726–733, 2006.

[40] R. Palaniappan and D. Mandic, "Eeg based biometric framework for automatic identity verification," *The Journal of VLSI Signal Processing*, vol. 49, no. 2, pp. 243–250, 2007.

[41] K. Ravi and R. Palaniappan, "A minimal channel set for individual identification with eeg biometric using genetic algorithm," in *International Conference on Conference on Computational Intelligence and Multimedia Applications*, vol. 2, pp. 328–332, 2007.

[42] A. Yazdani, A. Roodaki, S. Rezatofighi, K. Misaghian, and S. Setarehdan, "Fisher linear discriminant based person identification using visual evoked potentials," in *9th International Conference on Signal Processing, ICSP'08*, pp. 1677–1680, 2008.

[43] K. Brigham and B. Kumar, "Subject identification from electroencephalogram (EEG) signals during imagined speech," in *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)*, pp. 1–8, IEEE, 2010.

[44] *A Proposed Feature Extraction Method for EEG-based Person Identification*, 2012.

[45] S. Altahat, X. Huang, D. Tran, and D. Sharma, "People identification with rms-based spatial pattern of eeg signal," in *12th International Conference on Algorithms and Architectures for Parallel Processing, ICA3PP-12*, pp. 310–318, 2012.

[46] *Biometric Authentication using Brain Responses to Visual Stimuli*, vol. 6406, INSTICC Press, 2010.

[47] Z. Keirn and J. Aunon, "A new mode of communication between man and his surroundings," *Biomedical Engineering*, vol. 37, no. 12, pp. 1209–1214, 1990.

[48] R. Palaniappan, "Identifying individuality using mental task based brain computer interface," in *Third International Conference on Intelligent Sensing and Information Processing, ICISIP'05*, pp. 238–242, 2005.

[49] R. Palaniappan, "Electroencephalogram signals from imagined activities: a novel biometric identifier for a small population," in *Proceedings of the 7th international conference on Intelligent Data Engineering and Automated Learning, IDEAL'06*, pp. 604–611, 2006.

[50] R. Palaniappan and L. Patnaik, "Identity verification using resting state brain signals," in *Encyclopedia of Information Ethics and Security* (M. Quigley, ed.), pp. 335–341, IGI Global, 2007.

[51] R. Palaniappan, "Two-stage biometric authentication method using thought activity brain waves," *International Journal of Neural Systems*, vol. 18, no. 1, pp. 59–66, 2008.

[52] Z. Keirn, "Public database." `http://www.cs.colostate.edu/eeg/main/data/1989_Keirn_and_Aunon`, 1990.

[53] X. Bao, J. Wang, and J. Hu, "Method of individual identification based on electroencephalogram analysis," in *International Conference on New Trends in Information and Service Science, NISS'09*, pp. 390–393, 2009.

[54] J. Hu, "Multifeature biometric system based on eeg signals," in *Proceedings of the 2nd International Conference on Interaction Sciences: Information Technology, Culture and Human, ICIS'09*, pp. 1341–1345, ACM, 2009.

[55] J. Hu, "New biometric approach based on motor imagery eeg signals," in *International Conference on Future BioMedical Information Engineering, FBIE'09*, pp. 94–97, 2009.

[56] D. Jiang and J. Hu, "Research of computing in eeg password based on wavelet," in *, FBIE'09*, pp. 98–101, 2009.

[57] D. Xiao and J. Hu, "Identification of motor imagery eeg signal," in *International Conference on Biomedical Engineering and Computer Science. ICBECS*, pp. 1–4, 2010.

[58] B. C. III, "Public database." `http://www.bbci.de/competition/iii/`, 2003.

[59] BCI2000, "Public database." `www.bci2000.org`, 2000.

[60] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. Stanley, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[61] G. Schalk, D. Mcfarl, T. Hinterberger, N. Birbaumer, and J. Wolpaw, "Bci2000: A general-purpose brain-computer interface (bci) system," *IEEE Transactions on Biomedical Engineering*, vol. 51, p. 2004, 2004.

[62] S. Yang and F. Deravi, "On the effectiveness of eeg signals as a source of biometric information," in *Third International Conference on Emerging Security Technologies, EST'12*, pp. 49–52, 2012.

[63] M. Fraschini, A. Hillebrand, M. Demuru, L. Didaci, and G. Marcialis, "An eeg-based biometric system using eigenvector centrality in resting state brain networks," *Signal Processing Letters, IEEE*, vol. 22, no. 6, pp. 666–670, 2015.

[64] D. La Rocca, P. Campisi, B. Vegso, P. Cserti, G. Kozmann, F. Babiloni, and F. De Vico Fallani, "Human brain distinctiveness based on eeg spectral coherence connectivity," *Biomedical Engineering, IEEE Transactions on*, vol. 61, no. 9, pp. 2406–2412, 2014.

[65] C. Miyamoto, S. Baba, and I. Nakanishi, "Biometric person authentication using new spectral features of electroencephalogram (eeg)," in *International Symposium on Intelligent Signal Processing and Communications Systems, ISPACS'08*, pp. 1–4, 2008.

[66] *Personal Authentication Using New Feature Vector of Brain Wave*, 2008.

[67] I. Nakanishi, S. Baba, and C. Miyamoto, "Eeg based biometric authentication using new spectral features," in *International Symposium on Intelligent Signal Processing and Communication Systems, ISPACS'09*, pp. 651–654, 2009.

[68] I. Nakanishi, K. Ozaki, and S. Li, "Evaluation of the brain wave as biometrics in a simulated driving environment," in *Proceedings of the International Conference of the Biometrics Special Interest Group, BIOSIG'12*, pp. 1–5, 2012.

[69] Q. Zhao, H. Peng, B. Hu, Q. Liu, L. Liu, Y. Qi, and L. Li, "Improving individual identification in security check with an eeg based biometric solution," in *Proceedings of the International Conference on Brain Informatics*, BI, pp. 145–155, 2010.

[70] Q. Zhao, H. Peng, B. Hu, L. Li, Y. Qi, Q. Liu, and L. Liu, "Towards an efficient and accurate eeg data analysis in eeg-based individual identification," in *Ubiquitous Intelligence and Computing*, vol. 6406 of *Lecture Notes in Computer Science*, pp. 534–547, Springer Berlin / Heidelberg, 2010.

[71] B. Hu, C. Mao, W. Campbell, P. Moore, L. Liu, and G. Zhao, "A pervasive eeg-based biometric system," in *International workshop on Ubiquitous affective awareness and intelligent interaction, UAAII '11*, pp. 17–24, 2011.

[72] B. Hu, Q. Liu, Q. Zhao, Y. Qi, and H. Peng, "A real-time electroencephalogram (eeg) based individual identification interface for mobile security in ubiquitous environment," in *IEEE Asia-Pacific Services Computing Conference, APSCC*, pp. 436–441, 2011.

[73] M. Poulos, M. Rangoussi, and E. Kafetzopoulos, "Person identification via the eeg using computational geometry algorithms," in *Proceedings of the Ninth European Signal Processing, EUSIPCO'98*, pp. 2125–2128, 1998.

[74] M. Poulos, M. Rangoussi, and N. Alexandris, "Neural network based person identification using eeg features," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP'99*, vol. 2, pp. 1117–1120, 1999.

[75] M. Poulos, M. Rangoussi, V. Chrissikopoulos, and A. Evangelou, "Person identification based on parametric processing of the eeg," in *Proceedings of the 6th IEEE International Conference on Electronics, Circuits and Systems, ICECS'99*, vol. 1, pp. 283–286, 1999.

[76] M. Poulos, M. Rangoussi, V. Chrissikopoulos, and A. Evangelou, "Parametric person identification from the eeg using computational geometry," in *Proceedings of the 6th IEEE International Conference on Electronics, Circuits and Systems, ICECS'99*, vol. 2, pp. 1005–1008, 1999.

[77] M. Poulos, M. Rangoussi, N. Alexandris, and A. Evangelou, "On the use of eeg features towards person identification via neural networks," *Medical Informatics and the Internet in Medicine*, vol. 26, no. 1, pp. 35–48, 2001.

[78] M. Poulos, M. Rangoussi, N. Alexandris, and A. Evangelou, "Person identification from the eeg using nonlinear signal classification," *Methods of Information in Medicine*, vol. 41, no. 1, pp. 64–75, 2002.

[79] R. Paranjape, J. Mahovsky, L. Benedicenti, and Z. Koles, "The electroencephalogram as a biometric," in *Proceedings on Canadian Conference on Electrical and Computer Engineering*, vol. 2, pp. 1363–1366, 2001.

[80] G. Mohammadi, P. Shoushtari, B. Molaee Ardekani, and M. B. Shamsollahi, "Person identification by using ar model for eeg signals," in *Proceeding of World Academy of Science, Engineering and Technology*, vol. 11, pp. 281–285, 2006.

[81] P. Campisi, G. Scarano, F. Babiloni, F. DeVico Fallani, S. Colonnese, E. Maiorana, and L. Forastiere, "Brain waves based user recognition using the "eyes closed resting conditions" protocol," in *IEEE International Workshop on Information Forensics and Security, WIFS'11*, pp. 1–6, IEEE, 2011.

[82] D. La Rocca, P. Campisi, and G. Scarano, "Eeg biometrics for individual recognition in resting state with closed eyes," in *Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG)*, pp. 1–12, 2012.

[83] P. Tangkraingkij, C. Lursinsap, S. Sanguansintukul, and T. Desudchit, "Selecting relevant eeg signal locations for personal identification problem using ica and neural network," in *Eighth IEEE/ACIS International Conference on Computer and Information Science, ICIS'09*, pp. 616–621, 2009.

[84] M. Abdullah, K. Subari, J. Loong, and N. Ahmad, "Analysis of the EEG signal for a practical biometric system," *World Academy of Science, Engineering and Technology*, vol. 4, no. 8, pp. 1123–1127, 2010.

[85] M. Abdullah, K. Subari, J. Loong, and N. Ahmad, "Analysis of effective channel placement for an eeg-based biometric system," in *IEEE EMBS Conference on Biomedical Engineering and Sciences, IECBES*, pp. 303–306, IEEE, 2010.

[86] R. Palaniappan and D. Mandic, "Biometrics from brain electrical activity: A machine learning approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 738–742, 2007.

[87] A. Power, E. Lalor, and R. Reilly, "Can visual evoked potentials be used in biometric identification?," in *28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS'06*, pp. 5575–5578, 2006.

[88] G. Singhal and P. RamKumar, "Person identification using evoked potentials and peak matching," in *Biometrics Symposium*, pp. 1–6, 2007.

[89] K. Das, S. Zhang, B. Giesbrecht, and P. Eckstein, "Using rapid visually evoked eeg activity for person identification," in *International Conference on Engineering in Medicine and Biology Society*, pp. 2490–2493, 2009.

[90] A. Ferreira, C. Almeida, P. Georgieva, A. Tomé, and F. Silva, "Advances in eeg-based biometry," in *Proceedings of the 7th international conference on Image Analysis and Recognition - Volume Part II*, ICIAR'10, pp. 287–295, 2010.

[91] A. Ferreira, C. Almeida, P. Georgieva, and A. Tome, "Person identification using vep signals and svm classifiers," in *International Joint Conference on Neural Networks. IJCNN*, pp. 1–8, 2010.

[92] S. Marcel and J. Millan, "Person authentication using brainwaves (eeg) and maximum a posteriori model adaptation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 743–752, 2007.

[93] S. Sun, "Multi-task learning for eeg-based biometrics," in *19th International Conference on Pattern Recognition, ICPR'08*, pp. 1–4, 2008.

[94] C. He, X. Lv, and Z. Wang, "Hashing the mar coefficients from eeg data for person authentication," in *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'09*, pp. 1445–1448, 2009.

[95] C. He, "Person authentication using eeg brainwaves signals," Master's thesis, The university of British Columbia, The Faculty of Graduate Studies, 2010.

[96] C. He and J. Wang, "An independent component analysis (ica) based approach for eeg person authentication," in *3rd International Conference on Bioinformatics and Biomedical Engineering, ICBBE'09*, pp. 1–4, 2009.

[97] A. S. J. Thorpe, P. C. van Oorschot, "Pass-thoughts: authenticating with our minds," in *Proceedings of the 2005 workshop on New security paradigms, NSPW'05*, pp. 45–56, 2005.

[98] A. Riera, A. Soria-Frisch, M. Caparrini, C. Grau, and G. Ruffini, "Unobtrusive biometric system based on electroencephalogram analysis," *EURASIP Journal on Advances in Signal Processing*, no. 18, 2008.

[99] F. Kennet, "Brain wave based authentication," Master's thesis, Gjøvik University College, Department of Computer Science and Media Technology, June 2008.

[100] F. Su, L. Xia, A. Cai, Y. Wu, and J. Ma, "Eeg-based personal identification: from proof-of-concept to a practical system," in *20th International Conference on Pattern Recognition, ICPR'10*, pp. 3728–3731, 2010.

[101] F. Su, L. Xia, A. Cai, and J. Ma, "Evaluation of recording factors in eeg-based personal identification: A vital step in real implementations," in *IEEE International Conference on Systems Man and Cybernetics, SMC*, pp. 3861–3866, 2010.

[102] J. Klonovs, C. Petersen, H. Olesen, and A. Hammershoj, "Id proof on the go: Development of a mobile eeg-based biometric authentication system," *IEEE Vehicular Technology Magazine*, vol. 8, no. 1, pp. 81–89, 2013.

[103] K. Mohanchandra, G. Lingaraju, P. Kambli, and V. Krishnamurthy, "Using brain waves as new biometric feature for authenticating a computer user in real-time," *International Journal of Biometrics and Bioinformatic*, vol. 7, pp. 49–57, 2013.

[104] C. Gupta, R. Palaniappan, and S. Swaminathan, "Novel analysis technique for a brain biometric system," *International Journal of Medical Engineering and Informatics (IJMEI)*, vol. 1, pp. 266–273, 2008.

[105] C. Gupta, R. Palaniappan, and R. Paramesran, "Exploiting the p300 paradigm for cognitive biometrics," *International Journal of Cognitive Biometrics*, vol. 1, no. 1/2012, pp. 26–38, 2012.

[106] C. Gupta, Y. Khan, R. Palaniappan, and F. Sepulveda, "Wavelet framework for improved target detection in oddball paradigms using p300 and gamma band analysis," *Int. Journal of Biomedical Soft Computing and Human Sciences*, vol. 14, no. 2, pp. 61–67, 2009.

[107] R. Palaniappan, J. Gosalia, K. Revett, and A. Samraj, "PIN generation using single channel eeg biometric," in *Advances in Computing and Communication*, vol. 193 of *Communications in Computer and Information Science*, (Tiergartenstraße 17, 69121 Heidelberg, Germany), pp. 378–385, Springer, 2011.

[108] S. Yeom, H. Suk, and S. Lee, "Eeg-based person authentication using face-specific self representation," vol. 38 of *Proceedings of the Korea Computer Congress*, pp. 379–382, 2011.

[109] S. Yeom, H. Suk, and S. Lee, "Person authentication from neural activity of face-specific visual self-representation," *Pattern Recognition*, vol. 46, no. 4, pp. 1159–1169, 2013.

[110] H. Kayyali, "Medical device and method with improved biometric verification," 2014. US Patent 8,679,012.

[111] D. La Rocca, P. Campisi, and G. Scarano, "Stable eeg features for biometric recognition in resting state conditions," in *Biomedical Engineering Systems and Technologies*, vol. 452 of *Communications in Computer and Information Science*, pp. 313–330, Springer, November 2014.

[112] T. Pham, W. Ma, D. Tran, P. Nguyen, and D. Phung, "Eeg-based user authentication using artifacts," in *International Joint Conference SOCO14-CISIS14-ICEUTE14*, pp. 343–353, Springer, 2014.

[113] Y. Zhang and M. Juhola, "Biometric verification of a user based on eye movements," *International Journal of Biometrics*, vol. 6, no. 2, pp. 106–124, 2014.

[114] M. Abo-Zahhad, S. M. Ahmed, and S. N. Abbas, "A novel biometric approach for human identification and verification using eye blinking signal," *IEEE SIGNAL PROCESSING LETTERS*, vol. 22, no. 7, pp. 876–880, 2015.

[115] J. Hu and Z. Mu, "Authentication system using eeg biometric for smart home," *Applied Mechanics and Materials*, vol. 457, pp. 1228–1231, 2014.

[116] J. Hu and Z. Mu, "Authentication system for biometric applications using mobile devices," *Applied Mechanics and Materials*, vol. 457, pp. 1224–1227, 2014.

[117] B. Sabarigiri and D. Suganyadevi, "An efficient multimodal biometric authentication based on iris and electroencephalogram (eeg)," in *Fifth International Conference on Control, Communication and Power Engineering–CCPE*, 2014.

[118] T. Pham, W. Ma, D. Tran, P. Nguyen, and D. Phung, "Multi-factor eeg-based user authentication," in *Neural Networks (IJCNN), 2014 International Joint Conference on*, pp. 4029–4034, IEEE, 2014.

[119] A. Zǫuete, B. Quintela, and J. Silva Cunha, "Biometric authentication using electroencephalograms: a practical study using visual evoked potentials," *Electrónica e Telecomunicazões*, vol. 5, no. 2, pp. 185–194, 2010.

[120] H. X., S. Altahat, D. Tran, and D. Sharma, "Human identification with electroencephalogram (eeg) signal processing," in *International Symposium on Communications and Information Technologies (ISCIT)*, pp. 1021–1026, 2012.

[121] P. Tangkraingkij, C. Lursinsap, S. Sanguansintukul, and T. Desudchit, "Personal identification by eeg using ica and neural network," in *International Conference on Computational Science and Its Applications - Part III*, vol. 6018 of *Lecture Notes in Computer Science*, pp. 419–430, Springer-Verlag, March 2010.

[122] M. Bear, B. Connors, and M. Paradiso, *Neuroscience: Exploring the Brain*. Lippincott Williams & Wilkins, third ed., 2006.

[123] C. van Beijsterveldt, P. Molenaar, E. de Geus, and D. Boomsma, "Genetic and environmental influences on eeg coherence," *Behavior Genetics*, vol. 28, pp. 443–453, November 1998.

[124] F. Su, H. Zhou, Z. Feng, and J. Ma, "A biometric-based covert warning system using eeg," in *5th IAPR International Conference on Biometrics (ICB)*, pp. 342–347, 2012.

[125] M. Kostílek and J. Šťastný, "Eeg biometric identification: Repeatability and influence of movement-related eeg," in *International Conference on Applied Electronics (AE)*, pp. 147–150, 2012.

[126] X. Huang, S. Altahat, D. Tran, and L. Shutao, "Human identification with electroencephalogram (eeg) for the future network security," in *7th International Conference on Network and System Security*, vol. 7873 of *Lecture Notes in Computer Science*, pp. 575–581, 2013.

[127] DEAP, "Public database." `http://www.eecs.qmul.ac.uk/mmv/datasets/deap/`, 2012.

[128] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "Deap: A database for emotion analysis ;using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.

[129] P. Ullsperger, "Public database." `http://sccn.ucsd.edu/wiki/Chapter_02:_STUDY_Creation`, 2014.

[130] F. Perrin, J. Pernier, O. Bertrand, and J. Echallier, "Spherical splines for scalp potential and current density mapping," *Electroencephalography and clinical Neurophysiology*, vol. 72, no. 2, pp. 184–187, 1989.

[131] S. D. Muthukumaraswamy, "High-frequency brain activity and muscle artifacts in MEG/EEG: A review and recommendations," *Frontiers in human neuroscience*, vol. 7, no. 128, pp. 1–11, 2013.

[132] M. Fatourechi, A. Bashashati, R. K. Ward, and G. E. Birch, "Emg and eog artifacts in brain computer interface systems: A survey," *Clinical neurophysiology*, vol. 118, no. 3, pp. 480–494, 2007.

[133] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent component analysis*, vol. 46. John Wiley & Sons, 2004.

[134] I. Daly, N. Nicolaou, S. Nasuto, and K. Warwick, "Automated artifact removal from the electroencephalogram: A comparative study," *Clinical EEG and Neuroscience*, vol. 44, no. 4, pp. 291–306, 2013.

[135] S. Makeig, A. Bell, T. Jung, and T. Sejnowski, "Independent component analysis of electroencephalographic data," in *in Advances in Neural Information Processing Systems*, pp. 145–151, MIT Press, 1996.

[136] A. Mognon, J. Jovicich, L. Bruzzone, and M. Buiatti, "ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features," *Psychophysiology*, vol. 48, no. 2, pp. 229–240, 2010.

[137] S. Romero, M. A. Mañanas, and M. J. Barbanoj, "A comparative study of automatic techniques for ocular artifact reduction in spontaneous EEG signals based on clinical target variables: a simulation case," *Computers in biology and medicine*, vol. 38, no. 3, pp. 348–360, 2008.

[138] C. A. Joyce, I. F. Gorodnitsky, and M. Kutas, "Automatic removal of eye movement and blink artifacts from EEG data using blind component separation," *Psychophysiology*, vol. 41, no. 2, pp. 313–325, 2004.

[139] N. P. P. Castellanos and V. A. A. Makarov, "Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis.," *Journal of Neuroscience Methods*, no. 158, pp. 300–312, 2006.

[140] S.-Y. Shao, K.-Q. Shen, C. J. Ong, E. Wilder-Smith, and X.-P. Li, "Automatic EEG artifact removal: a weighted support vector machine approach with error correction," *Biomedical Engineering, IEEE Transactions on*, vol. 56, no. 2, pp. 336–344, 2009.

[141] S. Vorobyov and A. Cichocki, "Blind noise reduction for multisensory signals using ICA and subspace filtering, with application to EEG analysis," *Biological Cybernetics*, vol. 86, no. 4, pp. 293–303, 2002.

[142] Y. Li, Z. Ma, W. Lu, and Y. Li, "Automatic removal of the eye blink artifact from EEG using an ica-based template matching approach," *Physiological measurement*, vol. 27, no. 4, p. 425, 2006.

[143] F. C. Viola, J. Thorne, B. Edmonds, T. Schneider, T. Eichele, and S. Debener, "Semi-automatic identification of independent components representing EEG artifact," *Clinical Neurophysiology*, vol. 120, no. 5, pp. 868–877, 2009.

[144] A. Greco, N. Mammone, F. Morabito, and M. Versaci, "Semi-automatic artifact rejection procedure based on kurtosis, Renyi's entropy and independent component scalp maps," *International Journal of Medical, Health, Biomedical and Pharmaceutical Engeneering*, vol. 1, pp. 466–470, 2007.

[145] A. Delorme, T. Sejnowski, and S. Makeig, "Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis," *Neuroimage*, vol. 34, no. 4, pp. 1443–1449, 2007.

[146] H. Nolan, R. Whelan, and R. Reilly, "FASTER: Fully automated statistical thresholding for EEG artifact rejection," *Journal of Neuroscience Methods*, vol. 192, no. 1, pp. 152–162, 2010.

[147] I. Winkler, S. Haufe, and M. Tangermann, "Automatic classification of artifactual ica-components for artifact removal in EEG signals," *Behavioral and Brain Functions*, vol. 7, no. 1, 2011.

[148] H. Nolan, R. Whelan, and R. Reilly, "Sample of the simulated database used to evaluate FASTER."

[149] BESA, "EEG signal simulator."

[150] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural computation*, vol. 7, no. 6, pp. 1129–1159, 1995.

[151] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the royal statistical society. Series B (methodological)*, pp. 1–38, 1977.

[152] R. Kohavi, *Wrappers for Performance Enhancement and Oblivious Decision Graphs*. PhD thesis, Department of Computer Science, Stanford University, 1995.

[153] C. M. Judd, G. H. McClelland, and C. S. Ryan, *Data analysis: A model comparison approach*. 711 Third Avenue, New York, NY 10017 USA: Routledge, 2011.

[154] M. H. Libenson, *Practical approach to electroencephalography*. 1600 John F. Kennedy Blvd., Ste 1800, Philadelphia, PA 19103-2899, USA: Saunders Elsevier, 2012.

[155] J. B. Alonso, *La evaluación acústica del sistema fonador*. Universidad de Las Palmas de Gran Canaria, Vicerrectorado de Calidad e Innovación Educativa, 2008.

[156] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman, "A default mode of brain function," *Proceedings of the National Academy of Sciences*, vol. 98, no. 2, pp. 676–682, 2001.

[157] D. A. Gusnard and M. E. Raichle, "Searching for a baseline: functional imaging and the resting human brain," *Nature Reviews Neuroscience*, vol. 2, no. 10, pp. 685–694, 2001.

[158] M. E. Raichle and A. Z. Snyder, "A default mode of brain function: a brief history of an evolving idea," *Neuroimage*, vol. 37, no. 4, pp. 1083–1090, 2007.

[159] M. del Pozo-Baños, J. Alonso, J. Ticay-Rivas, and C. Travieso, "Electroencephalogram subject identification: A review," *Expert Systems with Applications*, vol. 41, pp. 6537–6554, May 2014.

[160] M. del Pozo-Baños, C. Travieso, C. Weidemann, and J. Alonso, "EEG biometric identification: A thorough exploration of the time-frequency domain," *Journal of Neural Engineering*, vol. 0, p. 0, September 2015.

[161] M. Abo-Zahhad, S. M. Ahmed, and S. N. Abbas, "A new biometric modality for human authentication using eye blinking," in *Biomedical Engineering Conference (CIBEC), 2014 Cairo International*, pp. 174–177, IEEE, 2014.

[162] M. del Pozo-Baños, C. Travieso, J. Alonso, and M. Ferrer, "Face identification based on tv videos," in *43rd Annual 2009 International Carnahan Conference on Security Technology*, pp. 119–125, October 2009.

[163] M. del Pozo-Baños, C. Travieso, M. Ferrer, and J. Alonso, "A gender classifier problem," in *1er Workshop de Tecnologas Multibiomtricas para la Identificacin de Personas (WTM-IP)*, vol. 1, pp. 25–29, July 2010.

[164] C. Travieso, M. del Pozo-Baños, M. Ferrer, and J. Alonso, "Reducing features using discriminative common vectors," *Cognitive Computation*, vol. 2, pp. 160–164, September 2010.

[165] M. del Pozo-Baños, C. Travieso, J. Alonso, and M. Ferrer, "Discriminative multi-projection vectors: Modifying the discriminative common vectors approach for face verification," in *2010 IEEE International Carnahan Conference on Security Technology (ICCST)*, pp. 190–197, October 2010.

[166] J. Solé-Casals, C. Travieso-González, M. del Pozo-Baños, and J. Alonso, "Face detection algorithm using emd - application to biometric recognition system.," in *BIOSIGNALS 2011 - Proceedings of the International Conference on Bio-inspired Systems and Signal Processing*, pp. 565–569, January 2011.

[167] M. del Pozo-Baños, C. Travieso, J. Alonso, and M. Ferrer, "Gender verification system based on jade-ica - application to biometric identification system," in *BIOSIGNALS 2011 - Proceedings of the International Conference on Bio-inspired Systems and Signal Processing*, pp. 570–576, January 2011.

[168] M. del Pozo-Baños, C. Travieso, J. Ticay-Rivas, and J. Alonso, *Advanced Biometric Technologies*, ch. A Gender Detection Approach, pp. 225–245. InTech, August 2011.

[169] C. Travieso, M. del Pozo-Baños, and J. Alonso, *Biometric Systems, Design and Applications*, ch. Facial Indentification Based on Transform Domains for Images and Videos, pp. 57–76. InTech, October 2011.

[170] J. Ticay-Rivas, M. del Pozo-Baños, C. Travieso, J. Alonso, and M. Ferrer, "Multi-spectral information regarding face recognition," in *2011 IEEE International Carnahan Conference on Security Technology (ICCST)*, pp. 1–9, October 2011.

[171] C. Travieso, M. del Pozo-Baños, and J. Alonso, "Comparison between thermal and visible facial features on a verification approach," in *Biomedical Engineering Systems and Technologies*, pp. 298–308, Springer Berlin Heidelberg, 2013.

[172] C. Travieso, M. Pozo-Baños, and J. Alonso, *Comparison between Thermal and Visible Facial Features on a Verification Approach*, vol. 357 of *Communications in Computer and Information Science*, pp. 298–308. Springer Berlin Heidelberg, 2013.

[173] J. Hernández-Travieso, C. Travieso, M. del Pozo-Baños, and J. Alonso, "Expression detector system based on facial images," in *6th International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS)*, pp. 411–418, February 2013.

[174] J. Ticay-Rivas, M. del Pozo-Baños, W. Eberhard, J. Alonso, and C. Travieso, "Spider recognition by biometric web analysis," in *Proceedings of the 4th international conference on Interplay between natural and artificial computation: new challenges on bioinspired applications - Volume Part II*, IWINAC'11, pp. 409–417, 2011.

[175] C. Travieso, M. del Pozo-Baños, and J. Ticay-Rivas, "Análisis de los óvalos de las vocales a y o bajo técnicas grafométricas," in *XXVI Simposio de la Unin Cientfica Internacional de Radio 2011 (URSI)*, p. 83, September 2011.

[176] J. Ticay-Rivas, M. del Pozo-Baños, C. Travieso, J. Arroyo-Hernández, S. Pérez, J. Alonso, and F. Mora-Mora, "Pollen classification based on geometrical, descriptors and colour features using decorrelation stretching method," in *Artificial Intelligence Applications and Innovations - 12th INNS EANN-SIG International Conference, EANN 2011 and 7th IFIP WG 12.5 International Conference, AIAI 2011, Part II*, vol. 364 of *IFIP Advances in Information and Communication Technology*, pp. 342–349, September 2011.

[177] C. Travieso, J. Alonso, J. Ticay-Rivas, and M. del Pozo-Baños, "Apnea detection based on hidden markov model kernel," in *Proceedings of the 5th international conference on Advances in nonlinear speech processing*, NOLISP'11, pp. 71–79, 2011.

[178] C. Travieso-González, J. Ticay-Rivas, M. del Pozo-Baños, W. Eberhard, and J. Alonso, "Improving spider recognition based on biometric web analysis," in *17th Iberoamerican Congress, Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, CIARP*, vol. 7441 of *Lecture Notes in Computer Science*, pp. 438–446, September 2012.

[179] M. del Pozo-Baños, J. Ticay-Rivas, J. Cabrera-Falcón, J. Arroyo, C. Travieso-González, L. Sánchez-Chavez, S. Pérez, J. Alonso, and M. Ramírez-Bogantes, *Biodiversity Enrichment in a Diverse World*, ch. Image Processing for Pollen Classification. InTech, August 2012.

[180] J. Ticay-Rivas, M. del Pozo-Baños, M. Gutiérrez-Ramos, W. Eberhard, C. Travieso-González, J. Alonso, and M. Ramírez-Bogantes, *Biodiversity Conservation and Utilization in a Diverse World*, ch. Image Processing for Spider Classification. InTech, August 2012.

[181] C. Travieso, J. Alonso, M. del Pozo-Baños, J. Ticay-Rivas, and K. Lopez-de Ipiña, "Automatic apnea identification by transformation of the cepstral domain," *Cognitive Computation*, vol. 4, pp. 1–8, September 2012.

[182] C. Travieso González, M. del Pozo Baños, J. Ticay Rivas, and J. Alonso Hernndez, "Método para la identificación y clasificación automática de especies de granos de polen."

[183] J. Ticay-Rivas, C. Travieso, M. del Pozo-Baños, and W. Eberhard, *Study of the Discriminate Information on Spider Webs*, pp. 95–114. Nova Publisher, June 2013.

[184] J. Ticay-Rivas, M. del Pozo-Baños, W. Eberhard, J. Alonso, and C. Travieso, "Spider specie identification and verification based on pattern recognition of it cobweb," *Expert Systems with Applications*, vol. 40, pp. 4213–4225, August 2013.

[185] T. C.M., J. Ticay-Rivas, J. Briceño, M. del Pozo-Baños, and J. Alonso, "Hand shape identification on multirange images," *Information Sciences*, vol. 275, no. 0, pp. 45–56, 2014.

[186] C. Travieso González, J. Ticay Rivas, M. del Pozo Baños, and J. Alonso Hernández, "Método para la identificación y clasificación automática de especies arácnidas a través de sus telas de araña."

[187] M. del Pozo-Baños, J. Ticay-Rivas, J. Alonso, and C. Travieso, "Features extraction techniques for pollen grain classification," *Neurocomputing*, vol. 150, Part B, pp. 377–391, February 2015.

Mi Mente, Mi Yo, Mi Identidad: Una Firma Neuronal Independiente de la Tarea Mental aplicada a la Identificación Biométrica (Resumen)

Marcos Del Pozo Baños

23 de septiembre de 2015

Índice general

1. Introducción **1**

1.1. Hipótesis 3

1.2. Estructura del documento 4

2. Estado del arte: revisión completa **7**

2.1. Uso de individuos no-sanos para la evaluación de sistemas 8

2.2. ¿Cuáles son los rasgos del sujeto localizados en el EEG? 9

2.3. ¿Dónde se localizan estos rasgos en el cerebro? 9

2.4. ¿Dónde se localizan estos rasgos en la frecuencia? 11

2.5. ¿Son constantes a lo largo del tiempo? 11

2.6. ¿Son constantes a lo largo de las diferentes tareas cognitivas? 12

2.6.1. ¿Cuáles son las mejores estrategias de diseño? 13

2.7. Conclusiones 14

3. Materiales **17**

3.1. Preprocesado 17

3.2. Bases de datos finales 18

4. Eliminación de artefactos del EEG: Filtrado Localizado de Componentes **19**

4.1. Conclusiones 20

5. Exploración tiempo-frecuencia del EEG **23**

5.1. Configuración de la Densidad Espectral de Potencia (PSD) 24

5.2. Representación de los dominios de tiempo y frecuencia 25

5.3. Propiedades de la información discriminatoria 26

5.4. Firma neuronal independiente del estado cognitivo 27

5.5. Conclusiones 28

6. Implementación de un sistema biométrico **31**

6.1. Diseño base 31

6.2. Coeficientes cepstrales reales 32

6.2.1. Coeficientes de envolvente espectral 32

6.3. Otros sistemas evaluados 33

6.4. Comparación con otros sistemas del estado del arte. 34

6.5. Conclusiones 35

7. Exposición, avances y lineas futuras **37**

7.1. Reflexión sobre la firma neuronal independiente del estado cognitivo 37

7.2. Reflexión sobre la estrategia de una biometría del EEG independiente del estado cognitivo 39

7.3. Conocimiento generado 41

Bibliografía **45**

Capítulo 1

Introducción

El presente documento es un resumen del original, titulado "My Mind, My Self, My Identity: A Task-Independent Neural Signature for Biometric Identification". Así pues, se centra en enumerar las conclusiones principales de cada uno de los capítulos, sin detallar los procedimientos de experimentación ni los resultados obtenidos. A su vez, no se presentarán imágenes ni tablas, sino que se hará referencia a las mismas del documento original.

La búsqueda de rasgos genéticos en el electroencefalograma (EEG) ha recibido especial atención por parte de la comunidad científica casi desde las primeras grabaciones del EEG humano, tomadas por Hans Verger en 1924 [1]. El descifrado de este mapa genotipo-fenotipo permitirá el desarrollo de herramientas para el estudio, el diagnóstico y el diagnóstico temprano de muchas enfermedades, especialmente aquellas que afectan al cerebro [2, 3]. Dichas herramientas, basadas

en medidas cuantitativas del EEG, estarán más próximas a la función genética que la tradicional interpretación de pruebas cognitivas [2]. A su vez, permitirán el seguimiento de pacientes y la evaluación de tratamientos también de manera cuantitativa.

La comprensión de los rasgos genéticos del EEG abrirá las puertas al desarrollo de sistemas biométricos de seguridad. En esta era de abundancia de la información y redes de comunicación, la biometría se ha erigido como una de las mejores soluciones, si no la mejor, para controlar el acceso a información y/o zonas restringidas. Estos sistemas se basan en características irreproducibles e intransferibles de la fisiología humana, tales como la huellas dactilares o los rasgos faciales.

La identificación de sujetos basada en el EEG es una disciplina relativamente joven, que encuentra sus orígenes en avances de los estudios genéticos y neurofisiológicos nombrados al inicio. Un sistema biométrico del EEG es especialmente atractivo debido a su potencial robustez y a la dificultad de falseamiento de la biometría. Las contraseñas serán más difíciles de robar dado que los usuarios no tendrán que realizar ninguna acción reveladora, como la introducción de la misma a través de un teclado. Incluso en el caso de robo el sistema puede estar programado para responder, no al significado semántico de la contraseña, sino a rasgos del EEG específicos del individuo. Estos rasgos son imposibles o extremadamente difíciles de reproducir por otro individuo. Además, si a un usuario se le fuerza a introducir su contraseña, el sistema puede monitorizar y responder a los niveles elevados de estrés u otros cambios en la actividad normal del EEG y bloquear el acceso.

Actualmente, la principal desventaja de la biometría del EEG es la inconveniencia del método de captura de datos. Los aparatos de EEG, aunque estrictamente no invasivos, son relativamente más invasivos que otras biometrías. En el caso de equipos clínicos o de investigación, la preparación de las sesiones de grabación

puede ser una ardua tarea que requiere de la aplicación de gel conductivo sobre el cuero cabelludo del usuario y de múltiples comprobaciones y ajustes de la calidad de la señal. Dicho esto, la proliferación de dispositivos EEG de consumo que reducen substancialmente los tiempos de preparación y hacen uso de tecnologías de sensores secos, han facilitado en gran medida la tarea de captura. Aunque estos dispositivos actualmente sufren de una menor calidad de la señal, cabe esperar que avances futuros en el campo de los sensores solventen gradualmente esta desventaja.

El presente trabajo es un estudio de los rasgos individuales del EEG, y más concretamente de aquellos en la representación tiempo-frecuencia del EEG. Dicho estudio está motivado por:

- La importancia de comprender el mapa genotipo-fenotipo de la actividad neuronal, el cual llevará a importantes avances en los campos de la medicina y la psicología.

- El potencial de la identificación biométrica del EEG, especialmente para sistemas que requieren métodos de seguridad robustos y para aquellos que ya integran cualquier forma de Interfaz Cerebro Ordenador (BIC), tales como dispositivos de monitorización y/o diagnóstico remotos y juegos.

1.1. Hipótesis

En este trabajo se presentan evidencias respaldando la siguiente hipótesis:

> *Existe en el electroencefalograma humano, un patrón concomitante a la identidad del individuo e independiente del estado cognitivo.*

La primera parte de esta hipótesis: "... concomitante a la identidad del indivi-

duo...", se refiere a actividad neuronal definida por el fenotipo del sujeto, y por lo tanto, única entre individuos.

La segunda parte: "... independiente del estado cognitivo", se refiere a la persistencia de dicha actividad en los distintos estados mentales y condiciones de grabación. Persistente no en el sentido estricto, sino de forma que las variaciones entre tareas sean de menor magnitud que las diferencias entre firmas neuronales de individuos.

1.2. Estructura del documento

Este trabajo ha sido llevado a cabo en cuatro etapas bien diferenciadas que definen los distintos capítulos del presente documento:

1. **Estado del arte:** (Capítulo 2) Dado que se trata ésta de una modalidad biométrica relativamente joven, la literatura no contaba con una recopilación de los avances más importantes en el área. Además, el estado del arte se caracterizaba por la ausencia de una metodología estructurada, lo cual dificultó la extracción de un argumento lógico. Los autores habían centrado sus esfuerzos en la aplicación de nuevos algoritmos y arquitecturas, sin antes comprender las propiedades básicas de la información procesada, o basándose en estudios genéticos y neurofisiológicos que analizan el EEG desde un punto de vista distinto al de la biometría.

2. **Recolección de bases de datos de EEG públicas:** (Capítulo 3) En un intento por identificar propiedades comunes intrínsecas de la firma neuronal y no idiosincrásicas de los datos, se hizo uso de 6 bases de datos públicas de distinta naturaleza. El conjunto de bases de datos usadas incluye: estados de relajación, tareas motoras y de resolución de problemas, potenciales evocados (ERPs) y emociones evocadas.

3. **Exploración del EEG en tiempo-frecuencia:** (Capítulo 5) Con el fin de suplir los puntos débiles detectados en el estado del arte, una parte importante de esta investigación se dirigió al análisis en profundidad de la información discriminatoria contenida en el espacio tiempo-frecuencia del EEG. Para ello, se llevaron a cabo numerosos experimentos, cada uno diseñado para evaluar una única propiedad o un pequeño conjunto de ellas, bajo condiciones controladas.

4. **Implementación de un sistema biométrico:** (Capítulo 6) Finalmente, los avances logrados durante la exploración de los rasgos individuales del EEG, se aplicaron en el diseño de un sistema de verificación biométrico basado en el EEG.

En paralelo a las etapas descritas, se realizó un estudio de las técnicas de corrección de artefactos del EEG durante una estancia en el Departamento de Psicología, Facultad de Ciencia y Salud Humana, Universidad de Swansea (Gales, Reino Unido). Dicha investigación concluyó en la propuesta de una nueva técnica de eliminación de artefactos: **Filtrado Localizado de Componentes (LCF)** (capítulo 4). Esta técnica se aplicó finalmente en el preprocesado de las bases de datos.

El documento cierra con disertaciones sobre las posibles fuentes fisiológicas de la firma neuronal independiente del estado cognitivo y su aplicación en el campo de la biometría, así como con una recopilación de las innovaciones producidas en el presente trabajo y una serie de propuestas de potenciales líneas de investigación futuras (capítulo 7).

Capítulo 2

Estado del arte: revisión completa

En este capítulo se recoge una revisión completa del estado del arte en identificación biométrica del EEG. Además de la evolución de dicha modalidad desde 1980 al 2015, también se presentan los principales avances de estudios genéticos y neurofisiológicos relacionados. Así mismo, se propone una categorización de los métodos aplicados basada en los estados cognitivos empleados durante la experimentación. Se identifican y nombran las bases de datos más usadas en la literatura, algunas de las cuales son de dominio público, con el fin de facilitar la comparación de resultados entre éstas y futuras investigaciones. Se muestra que, aunque muchas preguntas básicas siguen sin una respuesta clara, se ha identificado la existencia de información discriminatoria de la identidad en el EEG, la cual puede ser usada como biometría. Finalmente, se recomienda la aplicación de estrategias tales como el entrenamiento multi-sesión, la fusión de señales de distintos electrodos y frecuencias,y el uso de Clasificadores Lineales Discriminadores (LDC) y Máquinas de Vectores Soporte (SVM).

En general, el proceso de identificación es más complicado de lo anticipado puesto que se apoya sobre rasgos del EEG complejos y heterogéneos. Éstos son, a su vez, el resultado de elaborados modelos de herencia que hacen el proceso de

identificación extremadamente sensible a variables como el tiempo, la frecuencia, la localización de sensores, las condiciones de grabación y la arquitectura del sistema.

En particular, este capítulo se centra en encontrar respuestas a las siguientes preguntas:

- ¿Cuáles son los rasgos del sujeto localizados en el EEG?
- ¿Dónde se localizan estos rasgos en el cerebro?
- ¿Dónde se localizan estos rasgos en la frecuencia?
- ¿Son constantes a lo largo del tiempo?
- ¿Son constantes a lo largo de las diferentes tareas cognitivas?
- ¿Cuáles son las mejores estrategias de diseño?

2.1. Uso de individuos no-sanos para la evaluación de sistemas

Algunos estudios en la literatura hacen uso de bases de datos que incluyen individuos no-sanos además de sanos. Por ejemplo, los trabajos [4–9] hacen uso de pacientes alcohólicos. El problema de esta práctica es que el alcoholismo afecta a diversos aspectos de la actividad cerebral [2, 3], como las ondas alfa [10]. De esta manera, el problema de clasificación queda dividido de manera efectiva en dos: identificación entre sujetos sanos e identificación entre sujetos no-sanos; y los resultados sufren de sobre-estimación del rendimiento del sistema.

Así pues, para despejar cualquier duda, trabajos de biometría futuros que hagan uso de bases de datos con sujetos sanos y no-sanos, deben incluir un análisis específico que corrobore que los parámetros de identificación no están afectados por la condición del sujeto, es decir, que no existen correlación.

2.2. ¿Cuáles son los rasgos del sujeto localizados en el EEG?

Existen suficientes evidencias que confirman la existencia de información referente a la identidad del individuo contenida en el EEG. Qué rasgos son estos, por el momento, está menos claro. Estudios genéticos y neurofisiológicos afirman que la potencia y la frecuencia de pico de las ondas alfa presentan la mayor heredabilidad, seguidas por las ondas beta [11, 12].

Sin embargo, estas características por sí mismas fueron insuficientes para obtener buenas tasas de identificación en los estudios de biometría, lo que forzó la exploración y uso de nuevos descriptores. En general, los resultados sugieren que es en la forma global del espectro donde está codificada la identidad del sujeto.

2.3. ¿Dónde se localizan estos rasgos en el cerebro?

Múltiples estudios concluyeron que las fuentes óptimas de información discriminatoria se localizan en las regiones occipitales, temporales y parietales durante estados relajados con ojos cerrados (REC) [13–16]. Esto concuerda con los resultados de estudios genéticos, en los que se describen las ondas alfa como altamente determinadas genéticamente. Bajo estados de relajación con ojos abiertos (REO), se concluye que son las regiones anteriores del cerebro las que proporcionan la máxima información discriminatoria [16–20].

La balanza se inclina de nuevo hacia regiones occipitales en experimentos con Potenciales Visuales Evocados (VEPs) [21]. Esto puede atribuirse al aumento de la activación en estas áreas durante VEPs. Yeom's et. al. mostraron que cuando el procedimiento de experimentación hace uso de la personalidad y de imágenes del

propio usuario, las zonas frontales del cerebro vuelven a ser las más discriminatorias [22, 23]. Lo anterior sigue concordando con la relativa activación de las regiones cerebrales, dado que el "camino visual" describe el flujo de información/activación desde áreas posteriores hacia anteriores a medida que el cerebro interpreta la escena [24].

De los estudios con múltiples estados cognitivos, tan solo Yang et. al. consideraron el factor espacial [25]. Sus resultados muestran que áreas parietales cercanas a los lóbulos temporales contienen la información más discriminatoria. A pesar de ello, sus resultados fueron obtenidos con una base de datos extremadamente pequeña (tan sólo 3 usuarios).

La diferencia de rendimiento entre hemisferios está aún menos clara. Existen estudios tanto biométricos como neurofisiológicos que concluyen que no existen diferencias significativas entre ambos hemisferios [17–19, 26]. Al mismo tiempo, otros trabajos concluyeron que dicha diferencia sí existe, y que es el hemisferio derecho el que supera al izquierdo [20, 27, 28].

La falta de unanimidad en este sentido dificulta la extracción de conclusiones. Algunas de estas variaciones pueden ser explicadas por la propiedades del funcionamiento del cerebro humano y por los algoritmos aplicados. Por ejemplo, los sistemas especializados en ERPs funcionan mejor en regiones occipitales probablemente debido a que el complejo ERP es más fácilmente extraído de estas áreas. Así mismo, la discrepancia de resultados y conclusiones también puede ser debida a idiosincrasias de las bases de datos usadas en cada estudio.

2.4. ¿Dónde se localizan estos rasgos en la frecuencia?

El factor frecuencia ha recibido mucha menos atención que el factor espacio. Esto se debe, en gran medida, a que los autores se centraron en información ya localizada en una parte específica del espectro, basándose en estudios previos o en su propia experiencia.

En experimentos con REC, estudios neurofisiológicos y biométricos concluyeron que las ondas alfa y delta son las mas discriminatorias, seguidas por las ondas teta y beta [11,12,15,26]. Dentro del rango de las ondas alfa, la sección media ha superado las partes primera y última en algunos experimentos [29, 30]. En ocasiones, las ondas teta también han sido descritas como las más discriminatorias [31].

En el resto de experimentos, las altas frecuencias han producido mejores tasas de clasificación en la mayoría de los casos [32–36].

2.5. ¿Son constantes a lo largo del tiempo?

Estudios genéticos y neurofisiológicos han descrito cambios en el EEG humano durante la maduración, algunos de los cuales están relacionados con rasgos hereditarios [10]. Dichos cambios pueden ser ignorados en sistemas biométricos para adultos, puesto que se ralentizan pasados los primeros 19 o 20 años de desarrollo.

En cuanto a periodos de tiempo más cortos, hasta 1 año, estudios genéticos y neurofisiológicos han descrito la Densidad Espectral de Potencia (PSD) del EEG como altamente estable [37]. Sin embargo, a excepción del análisis cualitativo de I. Nakanishi et. al. [38], todos los estudios biométricos han contradicho esto en mayor o menor medida. S. Marcel y J.D.R. Milla fueron los primeros en observar una caída en el rendimiento del sistema a medida que las sesiones de entrenamiento

y test se distanciaban en el tiempo [39]. Estudios sobre los efectos de la dieta y los ritmos circadianos han revelado cuan sensibles pueden ser estos sistemas a cambios en el EEG producidos por situaciones cotidianas, como la ingesta de café o cambios fisiológicos diarios [40–42].

Algunos autores proponen la aplicación de estrategias de entrenamiento multi-sesión para contrarrestar dicha caída [7, 43, 44].

2.6. ¿Son constantes a lo largo de las diferentes tareas cognitivas?

A excepción del análisis estadístico de Kennet [32], todos los estudios biométricos han observado diferencias en el rendimiento de los sistemas cuando éstos son alimentados con EEG de distintos estados cognitivos. Los primeros resultados de Palaniappan sugirieron que los estados cognitivos más demandantes producían un mejor rendimiento [45]. Sin embargo, posteriores trabajos del mismo autor resultaron en conclusiones totalmente opuestas [46].

Esta disparidad, incluso entre trabajos llevados a cabo por el mismo grupo de investigación, dificulta la extracción de una narrativa consistente. Incluso comparando entre el movimiento de las manos derecha e izquierda, se encuentran estudios cuyas conclusiones abarcan todos los posibles resultados [6, 28, 33, 39, 47, 48].

Por otro lado, parece claro que los sistemas extraen algún tipo de información específica a la tarea realizada. Los resultados de Palaniappan mostraron que la fusión de información proveniente de las tareas con mayor y menor información discriminante de manera individual, producía los mejores resultados. Esta hipótesis está en línea con los resultados de S. Sun sobre el Entrenamiento Multi-Tarea (MTL) [47].

2.6.1. ¿Cuáles son las mejores estrategias de diseño?

Los descriptores explorados en la literatura pueden dividirse en dos grandes grupos: "generales" y "específicas". La mayoría de las publicaciones usan características del primer grupo, las cuales describen el EEG de manera general. Así pues, la información discriminatoria no se presenta de manera directa, sino que se encuentra "escondida" entre mucha otra información no relevante. Los ejemplos más importantes son los coeficientes de modelos Auto-Regresivos (AR) y de la PSD. En estos casos, es el clasificador o el selector de características quien debe encontrar y extraer la información correcta.

Descriptores "específicos" incluyen la potencia de pico y su frecuencia, así como características de los complejos ERP. En estos casos, la información discriminatoria se presenta de manera más directa al clasificador, facilitando así su labor. Este grupo de características han sido las más utilizadas en estudios con dispositivos EEG de consumo, principalmente debido a su baja complejidad y velocidad de procesado. Aunque los descriptores específicos han obtenido peores resultados que los generales, éstos han aumentado el rendimiento de los segundos en fusión. Por ejemplo, R. Palaniappan mejoró su diseño al incluir medidas de potencia de diferentes sub-bandas de frecuencia a los coeficientes AR, a pesar de que dichas medidas fueron extraídas directamente de estos coeficientes AR [49].

Así mismo, el uso de información proveniente de múltiples electrodos, ondas y estados cognitivos aumenta el rendimiento del sistema. También lo hacen la aplicación de técnicas de reducción de ruido y de herramientas de clasificación tales como LDC y SVM.

2.7. Conclusiones

A partir del extenso estudio del estado del arte realizado, se concluye que el EEG, y en particular su distribución espectral, contiene rasgos característicos de la identidad del sujeto. Estos rasgos pueden ser usados en un sistema de identificación biométrica, en especial aquellos extraídos de las ondas alfa.

Quizás, uno de los hechos más sorprendentes es el grado de disparidad en las conclusiones entre estudios. Esto puede ser consecuencia de la complejidad del modelo hereditario que gobierna los rasgos del EEG, y de la cantidad de variables que afectan al rendimiento del sistema. Así pues, el problema de identificación a partir del EEG es sensible a numerosos factores, tales como: la localización espacial de electrodos, las bandas de frecuencia, el tiempo, los estados cognitivos, la condición de los sujetos y, por supuesto, la arquitectura del sistema aplicado. En muchos casos, la alta especialización de los sistemas, diseñados para un problema muy concreto, ha acentuado dicha sensibilidad. Un ejemplo de ésto puede encontrarse en los resultados obtenidos por K. Das et. al. y S.K. Yeom et. al. con VEPs. El primero concluyó que la información discriminatoria se localiza en el lóbulo occipital del cerebro, entre 120 y 200 milisegundos después de la presentación del estímulo [21]. Por otro lado, S.K. Yeom et. al. concluyeron que la información más discriminatoria se encontraba en regiones anteriores del cerebro, a partir de 250 ms después de la presentación del estímulo [22,23]. Estas diferencias en tiempo y localización son, probablemente, debidas a que, a pesar de que ambos estudios usaron VEPs, cada uno de ellos se centraba en distinta información dentro de los mismos.

De los resultados considerados aquí se desprende que: la combinación de múltiples electrodos, ondas y estados cognitivos aumenta el rendimiento del sistema, y las estrategias de entrenamiento multi-sesión y de aprendizaje continuo pueden contrarrestar los efectos negativos del tiempo. El resto de propiedades están menos claras. Herramientas como los algoritmos de reducción de artefactos, LDC y SVM

parecen mejorar los resultados, pero éstos dependen también del resto de variables.

Por tanto, se concluye que el estado del arte tan sólo alcanza a arañar la superficie del problema, así como que se necesita más investigación para extraer conclusiones más firmes que las aquí presentadas.

Capítulo 3

Materiales

La mayoría de los estudios considerados en el capítulo 2 se centraron en el análisis de una única base de datos. Aún cuando se hizo uso de múltiples bases de datos [8,37,50], el objetivo de los autores era el de evaluar el rendimiento de varios sistemas, nunca el de analizar las propiedades básicas de los rasgos individuales del EEG. Además, las bases de datos usadas no siempre eran de dominio público, lo que dificulta la reproducción de los resultados.

Este capítulo describe en detalle las bases de datos usadas en el presente trabajo. Se introducen las propiedades originales de estas bases de datos y su preparación inicial. Se detalla el preprocesado aplicado a cada una de las bases de datos en un intento por reducir el ruido y las diferencias no relevantes entre ellas, tales como la frecuencia de muestreo o el montaje de electrodos. El capítulo finaliza con una descripción de las 10 bases de datos finales extraídas de las 6 originales.

3.1. Preprocesado

En conjunto, las bases de datos usadas abarcan un amplio rango de valores de configuración. Aunque algunos de éstos eran de interés para el estudio realizado

(tarea cognitiva, longitud de la señal, número de sujetos, etc), otros eran irrelevantes (frecuencia de muestreo, número de canales, rango de filtrado, etc). En un intento por eliminar estos últimos, se aplicó un preprocesado común a todas las bases de datos para normalizar estos factores (fig. 3.1).

3.2. Bases de datos finales

Tras la preparación y el preprocesado de las 6 bases de datos de dominio público, se obtuvieron un total de 10 conjuntos de datos (tabla 3.2). Éstos representan una rica variedad de estados cognitivos:

1. BCI2000-Baseline y DEAP-Baseline contienen estados de relajación.
2. BCI2000-Tasks incluye tareas motoras reales e imaginadas, además de tareas de relajación.
3. La base de datos de Keirn proporciona múltiples tareas de resolución de problemas.
4. DEAP-Playback contiene distintas emociones.
5. Las bases de datos de Yeom y Zhang contienen VEPs.
6. La base de datos de P. Ullsperger contiene AEPs.

Capítulo 4

Eliminación de artefactos del EEG: Filtrado Localizado de Componentes

Siendo el EEG una técnica de captura de la actividad neuronal no-invasiva, éste ha sido extensamente utilizado en la investigación y el desarrollo de aplicaciones en todas las áreas relacionadas con el cerebro. Una de las principales dificultades de los dispositivos EEG es la coexistencia de campos eléctricos de interés y contaminantes. Mientras los primeros son generados por el disparo de neuronas, los segundos pueden provenir de varias fuentes, tales como la actividad muscular o interferencias de dispositivos eléctricos, incluido el propio dispositivo EEG. Estas fuentes contaminantes son, en muchos casos, de mayor magnitud que las señales de interés, lo que resulta en una mala relación señal a ruido. Avances en los campos de la electrónica y los sensores han aumentado esta relación, pero la contaminación del EEG por ruido es, aún hoy en día, la mayor preocupación a la hora de usar estas señales [51, 52].

En este capítulo se presenta una nueva técnica para la eliminación de artefac-

tos, la cual reduce la pérdida de señal inherente a los métodos basados en métodos de Separación de Fuente Ciega (BSS). Dicha técnica localiza los segmentos contaminados dentro del espacio BSS y centra el procesado de componentes únicamente en esos segmentos: Filtrado Localizado de Componentes (LCF). Además, se diseña LCF de manera que pueda ser implementada en cualquier técnica de procesado existente basada en métodos BSS. Su integración es directa, y no requiere grandes cambios en los sistemas originales.

4.1. Conclusiones

De los análisis cualitativos y cuantitativos en bases de datos reales y simuladas, se desprende que la integración LCF en los sistemas existentes ofrece claras ventajas en la reducción de la pérdida de señal. Por otro lado, la habilidad de LCF para mantener la información relevante permite aumentar la agresividad del preprocesado a la hora de identificar cuales de los componentes BSS contiene ruido. Ésto resulta especialmente interesante en situaciones donde el sistema original funciona de manera muy conservadora.

Además, el prototipo de LCF usado aquí no es más que una prueba del concepto con un diseño sencillo, basado en la amplitud del voltage y en su velocidad de cambio. Se pueden esperar mejores resultados con descriptores del ruido más sofisticados. Ésto, a su vez, permitirá aumentar aún más la agresividad de la identificación de componentes contaminados. En el caso límite, LCF podrá ser utilizado de manera individual, procesando cada uno de los componentes BSS.

Es importante resaltar que los beneficios de LCF son inversamente proporcionales a la calidad del algoritmo BSS. Si el algoritmo es capaz de disociar las fuentes de actividad neuronal del resto, el componente de LCF sería innecesario. Hasta que los avances en dichos algoritmos logren una perfecta disociación, LCF

servirá como una herramienta extra para la reducción de pérdida de señal

Capítulo 5

Exploración tiempo-frecuencia del EEG

En el capítulo 2, se identifica en el estado del arte la falta de un estudio exhaustivo de la información espectral discriminatoria. Todos los estudios biométricos publicados hasta la fecha se centraron, principalmente, en mejorar las tasas de acierto mediante la aplicación de nuevos algoritmos. Mientras tanto, la naturaleza y las propiedades de la información procesada aún no habían sido descritas desde el punto de vista de la biometría.

Contrariamente al estado del arte, el estudio presentado en este capítulo no pretende, en ningún momento, obtener altas tasas de acierto durante la clasificación. En su lugar, pretende describir con el mayor detalle posible la información discriminatoria del sujeto contenida en el espacio tiempo-frecuencia de la actividad EEG, así como caracterizar cada uno de los parámetros del problema de clasificación. De hecho, algunas de las decisiones tomadas en el diseño y análisis de los experimentos están fundamentados en este objetivo. Se intenta así cubrir el vacío encontrado en el estado del arte.

En particular, a lo largo de los 4 cuatro bloques de experimentación:

1. Se obtiene una serie de recomendaciones para maximizar la calidad de la información espectral discriminatoria.

2. Se presentan evidencias visuales de la existencia de dicha información y sus propiedades mediante una representación en batería del espectrograma de múltiples muestras.

3. Se utilizan 6 bases de datos públicas, lo cual ha permitido distinguir entre propiedades generales de la información y aquellas idiosincrásicas de los datos.

4. Se hace uso, por primera vez, de Potenciales Auditorios Evocados (AEPs) para la identificación biométrica de sujetos.

5. ***Se presentan evidencias de la existencia de una firma neuronal independiente del estado cognitivo***.

5.1. Configuración de la Densidad Espectral de Potencia (PSD)

En el primer bloque de experimentación, se analizan los efectos en la calidad de la información de algunos parámetros básicos. En concreto, se consideran el número de coeficientes de la Transformada Rápida de Fourier (FFT) N_F, la longitud de la ventana espectral L_W y su solapamiento Θ, y la longitud de la señal de EEG usada para general la respuesta final L_G.

En estos experimentos, se define la configuración óptima de la Transformada de Fourier de Tiempo Corto (STFT) para maximizar la extracción de información discriminatoria. En concreto, la fragmentación de la señal del EEG en ventanas cortas solapadas, de entre 1 y 2 segundos, se ha identificado como la mejor estrate-

gia. La representación espectral de cada segmento debe ser calculada sin escatimar en la resolución en frecuencia ($N_F \approx L_W * F_s$), teniendo en cuenta que la información discriminatoria se encontró localizada principalmente por debajo de los 40 Hz.

En cuanto a la cantidad de señal EEG necesaria para realizar la identificación, es una cuestión de maximizar la información proporcionada al sistema. Dicho ésto, tras el análisis se concluye que $L_G = [4, 6]$ segundos resulta en un rendimiento entre el 92.5 % y el 95 % del máximo alcanzable por los datos disponibles. Sin embargo, es de esperar que esto varíe en función de la calidad de la señal.

5.2. Representación de los dominios de tiempo y frecuencia

Una vez encontrada la configuración óptima de los parámetros N_F, L_W, Θ y L_G, se han evaluado los efectos de diferentes representaciones de los dominios de tiempo y frecuencia. Se ha comparado la información discriminatoria de los montajes *AvgMnt*, *BIHMnt* y *CzMnt*, y los efectos de varios métodos de normalización de la PSD.

Aunque *AvgMnt* resultó ser el montaje con mejor rendimiento medio, los resultados de *BIHMnt* fueron sorprendentemente similares, en especial considerando la significante reducción del volumen de datos introducida por *BIHMnt*. Además, la normalización espectral con medidas robustas frente a valores atípicos, redujo los efectos del ruido y mejoró la calidad de la información extraída en la mayoría de casos. Así pues, éstas técnicas pueden ser aplicadas como alternativa a métodos de eliminación de artefactos.

5.3. Propiedades de la información discriminatoria

A continuación, se describen algunas de las propiedades de la información discriminatoria identificada en los bloques de experimentación anteriores. En particular, se estudia la distribución espacial y en frecuencia de la información, además de la naturaleza única entre individuos y permanencia a lo largo del tiempo.

Los resultados sugieren que no hay una clara localización que supere a las demás de manera sistemática entre sistemas, bases de datos y estados cognitivos. Por tanto, se concluye que el rendimiento relativo de cada sensor parece estar principalmente determinado por idiosincrasias de los datos y del montaje y por las características del sistema.

En cuanto a la distribución en frecuencia, los resultados apuntan a la existencia de un pico de rendimiento dentro de las ondas alfa. Además, las ondas beta, hasta 40 Hz, parecen contener tanta o más información que frecuencias más bajas. Finalmente, en ocasiones, otro pico de rendimiento puede verse por debajo de los 5 Hz.

Los resultados también sugieren que esta información discriminatoria es suficientemente única para discriminar un gran número de usuarios (> 100) y relativamente constante en cortos periodos de tiempo. Sin embargo, se necesita más experimentación con bases de datos mayores, grabadas en múltiples sesiones distanciadas en el tiempo, para extraer conclusiones más firmes en este aspecto.

5.4. Firma neuronal independiente del estado cognitivo

Los resultados de S. Sun en MTL [47] y de F. Kennet cuando trataba de identificar estados cognitivos usando las mismas características que para la identificación de usuarios [32] ya sugerían la existencia de una firma neuronal independiente del estado cognitivo. En esta última fase de experimentación, se ejecutan una serie de experimentos especialmente diseñados para evaluar la hipótesis principal de esta tesis.

Si las diferencias entre las firmas neuronales estuvieran principalmente determinadas por el estado cognitivo, la tasa de acierto del sistema biométrico habría sufrido una caída dramática (cerca de niveles aleatorios) en los experimentos de validación cruzada de estados. Por el contrario, la estabilidad en la tasa de acierto observada respalda la hipótesis planteada. En especial, considerando que en los experimentos de validación cruzada de estados algunos sujetos fueron testeados con EEG de estados cognitivos usados para entrenar a otros sujetos.

Ésto no sugiere que la actividad del EEG sea completamente homogénea entre estados cognitivos. En los ejemplos mostrados, actividad específica al estado cognitivo es fácilmente identificable. Sin embargo, estas variaciones (A) coexisten con características independientes del estado y/o (B) tienen una magnitud menor que las diferencias entre individuos. De hecho, los resultados cuantitativos presentados señalan la existencia de información discriminatoria específica del estado, pero con una capacidad discriminatoria menor que la de la información independiente del estado.

El hecho de que los efectos del tiempo en experimentos con validación cruzada de estados cognitivos se mantuvieran como en experimentos sin validación cruzada de estados, respalda aún más la existencia de una firma neuronal independiente. Si

el rendimiento de los experimentos previos se debiera a peculiaridades del montaje en lugar de a características reales de los sujetos, la tasa de acierto se hubiera desplomado al aplicar validación cruzada en las sesiones.

Los resultados sugieren que la actividad cerebral capturada por el EEG, está más definida por la identidad del individuo que por el estado cognitivo. De hecho, el espectrograma del EEG revela que parte de esta actividad es relativamente estable entre estados. En particular, se observa esta firma neuronal independiente entre tareas motoras y de relajación (BCI2000-Full), estados emocionales y REO (DEAP-Full), resolución de problemas (Keirn's), AEPs de sinónimos y no-sinónimos (P. Ulssperger's), VEPs de representaciones propias y de otras personas (Yeom's) y VEP de objetivos y no-objetivos (Zhang's).

5.5. Conclusiones

En este capítulo, se presentan los resultados de un estudio exhaustivo de la información discriminatoria en el espacio tiempo-frecuencia del EEG. Para ello, se han empleado 6 bases de datos con distintos estados cognitivos. Esto, en conjunto con los análisis cualitativos y cuantitativos, han permitido distinguir entre las características inherentes a la firma neuronal y las idiosincrasias de las bases de datos. Se han ejecutado cuatro bloques de experimentación, cada uno con un objetivo específico, los cuales resultaron en las siguientes recomendaciones y conclusiones:

1. Configuración de la PSD: recomendaciones.

    a) Grabar, al menos, 5 segundos del EEG para llevar a cabo la identificación.

    b) Dividir el EEG en segmentos de entre 1 y 2 segundos de longitud.

c) Si el volumen de datos y la velocidad de procesado no es un problema, aplicar solapamiento entre los segmentos.

d) Calcular la representación espectral de cada ventana, manteniendo la máxima resolución en frecuencia posible.

e) Mantener un ancho de banda desde la frecuencia más baja posible hasta 30 o 40 Hz.

f) Llevar a cabo la clasificación para cada una de las ventanas individualmente y generar la respuesta final combinando el resultado de cada una de ellas.

2. Representación de los dominios temporal y de la frecuencia: recomendaciones.

   a) Usar *AvgMnt* como el montaje por defecto.

   b) Considerar *BIHMnt* en casos donde la velocidad de procesado o el volumen de datos sea un problema.

   c) Como alternativa a métodos complejos de eliminación de artefactos, normalizar los coeficientes espectrales con métodos robustos frente a muestras atípicas.

3. Propiedades de la información discriminatoria: conclusiones

   a) Parece no haber una localización que proporcione mejores resultados que el resto para todos los sistemas, bases de datos y/o estados cognitivos.

   b) Existe un pico de rendimiento dentro de las ondas alfa. Frecuencias en las ondas beta (hasta 40 Hz) son tan o más discriminatorias que otras frecuencias más bajas. Frecuencias por debajo de 5 Hz también pueden contener una importante cantidad de información discriminatoria.

c) Los rasgos individuales en el EEG son suficientemente únicos como para discriminar 100+ sujetos cuando se aplica el sistema apropiado.

d) La firma neuronal parece ser lo suficientemente permanente como para ser usada como una biometría.

e) ***Del total de la información discriminatoria localizada en el espectro del EEG, una parte significativa de la misma es independiente del estado cognitivo.***

Capítulo 6

Implementación de un sistema biométrico

Para finalizar esta investigación, se han aplicado los conocimientos sobre las propiedades y características de la firma neuronal adquiridos en el capítulo anterior, en el diseño de un sistema de verificación biométrica.

6.1. Diseño base

Se parte de un diseño base derivado del usado en el capítulo 5. La señal de EEG se dividió en ventanas de 2 segundos solapadas un 75 %. Se aplica una ventana de Hamming y se calculan los coeficientes PSD para cada uno de los segmentos. Se concatenan los coeficientes de entre 1 y 40 Hz de todos los sensores EEG para formar el vector de características. Se realiza la clasificación mediante LDC. Para cada muestra, se agregan los resultados de cada una de las ventanas/segmentos para obtener la respuesta final.

6.2. Coeficientes cepstrales reales

Los coeficientes cepstrales han sido extensamente aplicados a problemas de procesado de señal [53]. El espacio cepstral codifica la forma general del espectro en sus bajas cuefrencias (primeros coeficientes cepstrales), y los detalles y periodicidad en cuefrencias superiores. Así pues, tras una serie de experimentos con vectores de características $\text{RCeps}_{P\,\%}$, con distintos valores de P, donde P es el porcentaje de los primeros coeficientes usados (el resto fueron desechados), se encuentra que es suficiente con mantener un 20 % de los mismos para obtener un rendimiento óptimo.

6.2.1. Coeficientes de envolvente espectral

Los coeficientes de un modelo AR (o Coeficientes Predictores Lineales, LPC) han sido una opción popular dentro del campo de la biometría del EEG (capítulo 2). Un modelo AR predice muestras de una serie temporal en función de p muestras pasadas, donde p es el orden del modelo. Ésto puede verse como la salida de un sistema todo-polos de Respuesta al Impulso Infinita (IIR) cuando se le presenta ruido a su entrada. Así pues, los LPC describen la forma espectral de la señal modelada. Entre mayor sea el orden del modelo, más detallada será esta descripción.

Al igual que ocurrió con los cepstrums reales, se encuentra que es suficiente con modelar la señal con orden 8 (LPC_8). Órdenes superiores resultaron en rendimientos similares o incluso peores. Además, se testeó el uso de otras representaciones de los LPC, tales como los Coeficientes de Reflexión (RC), las Líneas de Pares Espectrales (LSP) y el error del modelo AR (ϵ). Mientras ϵ obtuvo resultados significativamente peores que LPC, tanto RC como LSP obtuvieron resultados similares. Sin embargo, hay que señalar que el cálculo de estos coeficientes es computacionalmente más costoso que en el caso de los LPC.

Estos resultados contradicen algunas de las conclusiones encontradas en la literatura. R.B. Paranjape et. al. concluían que un incremento en el orden del modelo AR era necesario para mantener el rendimiento del sistema frente a un incremento en el número de usuarios [54]. Con tan sólo 5 sujetos, los autores encontraron un incremento en la tasa de clasificación de 7 puntos porcentuales al aumentar el orden de 9 a 15. Así mismo, el equipo de Campisi encontró que los RC superaban en rendimiento a los LPC [14, 55]. Estas discrepancias pueden ser debidas a diferencias en la metodología de experimentación: experimentos de clasificación frente a verificación, arquitecturas del sistema y/o idiosincrasias de las bases de datos.

6.3. Otros sistemas evaluados

En esta fase, se evaluaron muchos más diseños además de los descritos. Sin embargo, todos ellos resultaron en tasas de acierto peores o similares; éstos últimos con arquitecturas más complejas. Así pues, se decidió no presentar sus resultados de manera detallada, pero sí una descripción de los mismos. En particular, los siguientes diseños fueron testeados:

- Vectores de características basados en medidas estadísticas de la PSD o de los coeficientes RCeps o LPC.
- Fusión a nivel de datos de todos los vectores de características, considerando todas las combinaciones posibles.
- Fusión a nivel de resultados de todos los vectores de características, considerando todas las combinaciones posibles.
- Fusión a nivel de resultados de sistemas con segmentaciones del EEG de distinta longitud.

- Aplicación de algoritmos de reducción de la dimensionalidad, tales como el Análisis Linear Discriminante (LDA), el Análisis de Componentes Principales (PCA) y el Análisis de Componentes Independientes (ICA).
- Fusión a nivel de resultados de sistemas expertos de sensores.
- Aplicación de clasificadores no lineales como SVM y Redes Neuronales Artificiales (ANN).

6.4. Comparación con otros sistemas del estado del arte.

A pesar de que las bases de datos BCI2000, de Keirn, de Yeom y de Zhang han sido utilizadas por otros autores para la evaluación de sistemas biométricos, no se puede realizar una comparación directa entre sus resultados y los obtenidos en este trabajo. Además de las diferencias en el número de sujetos y estados cognitivos, y en la metodología de experimentación empleada, los siguientes dos factores han de tenerse en cuenta:

En primer lugar, el diseño propuesto aquí es un diseño generalizado. Al contrario de la mayoría de sistemas presentados por otro autores, en el presente trabajo se ha diseñado el sistema haciendo uso de múltiples bases de datos, de manera que las configuraciones seleccionadas son óptimas de manera general, pero sub-óptimas de manera particular para cada base de datos. Por tanto, este sistema no esta optimizado para maximizar el rendimiento frente a una única base de datos, sino para obtener buenos resultados en cualquier escenario. De hecho, los resultados de las tres bases de datos mayores se encuentran distanciados por tan sólo 6 puntos porcentuales, lo que sugiere que dicho grado de discriminación es propio del sistema y de la firma neuronal, y no debido a idiosincrasias de los datos.

En segundo lugar, la estrategia propuesta basada en el uso de una firma neuronal independiente del estado cognitivo representa un problema indudablemente más complejo que el de la identificación vía EEG de un único estado cognitivo. Tal y como se señaló en el capítulo 5, la actividad neuronal de un estado específico contiene información discriminatoria, la cual es explotada por los sistemas del estado del arte. En el capítulo 7 se presentan una serie de argumentos a favor de esta estrategia, a pesar del adquirido aumento en complejidad.

6.5. Conclusiones

En este capítulo, se han evaluado múltiples métodos y arquitecturas de sistema en un intento por explotar los rasgos de identidad contenidos en el espacio tiempo-frecuencia del EEG. En general, se observa que $\text{RCeps}_{20\,\%}$ y LPC_8 son, de aquellas testeadas, las dos características con mejores resultados. Cualquier descriptor extra añadido al vector $\text{RCeps}_{20\,\%}$ o LPC_8 tuvo siempre un efecto neutral o negativo en el rendimiento del sistema, al igual que lo tuvo la aplicación de algoritmos de clasificación no lineales tales como ANN y SVM.

De los resultados presentados se concluye que:

1. La información extraída por RCeps y LPC está altamente correlacionadas, evidenciada por el hecho de que la fusión de los mismos no tiene ningún efecto en la tasa de acierto del sistema.

2. El problema de verificación basado en características RCeps o LPC es un problema lineal, evidenciado por el hecho de que los algoritmos no lineales aplicados, SVM y ANN, tan sólo igualaron la tasa de acierto de LDC.

3. RCeps y LPC pueden codificar la mayoría de la información discriminatoria del espectro del EEG, evidenciado por el hecho de que todos los demás sis-

temas testeados obtuvieron tasas de acierto peores o similares a los sistemas basados únicamente en RCeps o LPC, incluyendo sistemas que fusionaban múltiples características.

Así pues, la información codificada por $\text{RCeps}_{20\,\%}$ y LPC_8 es sorprendentemente discriminatoria y robusta. Para mejorar el diseño del sistema sería más productivo explorar información extra de fuentes distintas al espectro del EEG de sensores individuales, o aplicar etapas de preprocesado más sofisticadas.

Capítulo 7

Exposición, avances y lineas futuras

En este último capítulo, se reflexiona sobre la fuente de la firma neuronal independiente del estado cognitivo, se aportan argumentos a favor de la estrategia propuesta para el problema de identificación biométrica del EEG y se resaltan los avances en los campos de la biometría y la neurociencia logrados por el presente trabajo.

7.1. Reflexión sobre la firma neuronal independiente del estado cognitivo

A lo largo de esta investigación, se han proporcionado evidencias que respaldan la hipótesis presentada en la sección 1.1:

> *Existe en el electroencefalograma humano, un patrón concomitante a la identidad del individuo e independiente del estado cognitivo.*

Además, se ha usado esta propiedad para construir satisfactoriamente un sistema capaz de diferenciar individuos en función de su actividad EEG, independientemente del estado cognitivo. Dicho sistema no habría sido posible sin la existencia de información discriminatoria independiente del estado.

Dicha investigación es fundamentalmente diferente a estudios neurofisiológicos y biométricos publicados anteriormente. Sean los sistemas de reconocimiento del iris y de la marcha como ejemplos análogos. El primero de estos sistemas hace uso de propiedades inherentes al individuo, mientras que el segundo se centra en la manera en la que el individuo ejecuta una acción (caminar). Por tanto, se puede argumentar que los estudios biométricos centrados en el EEG de un estado cognitivo concreto están más cerca, en naturaleza, de describir actividad idiosincrásica del individuo durante el procesado cognitivo. En contraposición, esta investigación, centrada en características independientes de la tarea, intenta describir la identidad del individuo en sí misma. Dicho ésto, es muy probable que estudios previos hayan usado, de manera inadvertida, parte de la firma neuronal independiente de la tarea.

El origen de esta firma neuronal sigue sin estar claro. Puesto que es independiente del estado cognitivo, se podría asociar con procesos inconscientes que trabajan de manera ininterrumpida, de forma similar al concepto de *'modo estándar del funcionamiento cerebral'* introducido por M.E. Raichle et. al. [56–58]. Dicho concepto propone la existencia de actividad intrínseca que se encarga del mantenimiento de información para la interpretación, respuesta e incluso la predicción de las demandas del entorno. El modo estándar no desaparece completamente cuando el sujeto comienza una tarea, si no que se atenúa [57]. Ésto concuerda con nuestra observación sobre el dinamismo de la firma neuronal. La propiedad de independencia de la tarea surge del hecho de que dichas fluctuaciones tienen una magnitud inferior que las diferencias entre individuos. Por tanto, pueden ser interpretadas

como actividad específica del estado superpuestas al modo estándar del cerebro. A modo de anécdota, resaltar que M.E. Raichle et. al., en su disertación sobre la continuidad del modo estándar, notaron su relación con el 'yo' del individuo [57].

También existe la posibilidad de que la firma neuronal sea puramente debida a la estructura de las redes neuronales subyacentes, y que no tenga nada que ver con los procesos cognitivos que ejecutan. Dada la naturaleza de los campos eléctricos y de su propagación a través del cráneo, dos redes con funcionalidad idéntica pero distinta organización producirían distintas señales EEG. Así, la firma neuronal del EEG estaría principalmente definida por la disposición de redes neuronales dentro del cerebro, y los procesos cognitivos jugarían un papel modulador de la misma.

Como suele ocurrir en estos casos, es probable que la solución no sea ninguna de las anteriores, sino una combinación de ambas.

7.2. Reflexión sobre la estrategia de una biometría del EEG independiente del estado cognitivo

Hasta la fecha, la estrategia seguida en la literatura se ha basado en el análisis del EEG para condiciones cognitivas únicas. Incluso cuando un sistema era alimentado con señales de múltiples estados, éstos eran etiquetados de manera que el sistema pudiera diferenciar entre ellos y explotar información especifica a cada uno de los mismos (MTL).

En una publicación reciente, el equipo de Campisi reflexionó sobre la idea de definir un protocolo de adquisición de EEG, donde se pedía a los usuarios realizar una tarea particular mientras se registraba su EEG para la identificación o verificación [59]. En concreto, Campisi et. al. centraron sus esfuerzos en las condiciones de reposo REC y REO [14–16, 55]:

En este paradigma, se le pide a los sujetos que se sienten en una silla confor-

table con ambos brazos en reposo, en una sala con luz tenue o totalmente oscura. Generalmente, los sonidos externos se minimizan para favorecer la relajación de los sujetos. Se le pide a los participantes que se mantengan unos minutos en estado de relajación con sus ojos cerrados o abiertos, evitando cualquier tipo de atención o concentración, pero manteniéndose despiertos y alerta [59].

T. Pham's et. al. fueron más allá y propusieron un sistema que asignaba tareas específicas a cada uno de los usuarios [60]. Así, al identificar la tarea realizada, la complejidad del problema se reduce en un factor N, donde N es el número de tareas consideradas.

No se puede negar que la ejecución de cualquier tarea específica durante la verificación de la identidad de un usuario es, en muchos casos reales, problemática. Por ejemplo, si esta modalidad fuera integrada en el pasaporte biométrico, llevar a cabo un *simple protocolo de estado de relajación* (en palabras del equipo de Campisi [59]) es completamente inviable. Además, este tipo de biometría encontraría aplicación, principalmente, en otros BCIs, los cuales fueron originalmente creados con otros propósitos. Por ejemplo, una patente reciente de Google integra la identificación del usuario en un sistema de diagnóstico multi-sensor [61].

Como solución a estos problemas, se propone una estrategia novedosa donde no se le pide al usuario que ejecute ninguna acción en particular. En su lugar, el sistema extrae información discriminatoria independiente del estado cognitivo. A su vez, esto deja al usuario libre para ejecutar cualquier otra acción. Por ejemplo, en la cabina de pasaporte biométrico, el EEG del usuario puede ser grabado mientras este presenta su pasaporte, introduce cualquier información requerida y/o proporciona otras biometrías. En una BCI de propósito general, la verificación de la identidad del usuario podría tener lugar de fondo y de manera invisible. Así, el procedimiento de seguridad no interferiría con la experiencia del usuario. Además, se podrían ejecutar chequeos de seguridad de manea continua o periódica, una vez

más, sin interferir con el funcionamiento del dispositivo.

7.3. Conocimiento generado

En este trabajo, se ha realizado un estudio del estado del arte en la identificación biométrica del EEG y, en menor medida, de estudios genéticos y neurofisiológicos equivalentes (capítulo 2). En el momento de su publicación en mayo del 2014 [62], fue el primer artículo de este tipo. En él, se identifica una falta general de consenso en varias preguntas fundamentales, tales como la distribución en frecuencia y espacio de la información discriminatoria, y las condiciones para optimizar su extracción.

En un intento por arrojar luz sobre estas cuestiones, se llevó a cabo un análisis exhaustivo del espacio tiempo-frecuencia del EEG (capítulo 5). Dicho estudio representa las siguientes innovaciones con respecto a investigaciones previamente publicadas:

- Se propuso una representación visual de la información, basada en la concatenación de espectrogramas del EEG, que permite aportar evidencias visuales de la existencia de rasgos individuales en el PSD del EEG.

- A través de dicha representación, se llevaron a cabo análisis cualitativos que complementaron los cuantitativos. Esta estrategia fue vital para la correcta interpretación de los resultados.

- Se ejecutaron los experimentos en 6 bases de datos, divididas en última estancia en 10 conjuntos de datos con distintas características. Éstas incluían estados de reposo, tareas motoras reales e imaginadas, tareas de resolución de problemas, VEPs y AEPs. Tal variabilidad permitió diferenciar entre resultados característicos e idiosincrásicos, y extraer conclusiones robustas sobre

las propiedades generales de la firma neuronal.

- Además, una de las bases de datos usadas estaba compuesta por AEPs. En estudios biométricos previos, VEPs eran los únicos ERPs usados.

- Se ejecutó el análisis siguiendo una metodología incremental cuidadosamente diseñada. Se abordaron cada uno de los parámetros de manera individual para entender su interacción con el problema de identificación. Siempre que fue posible, se usaron tres versiones de las bases de datos: una en crudo y dos libres de artefactos.

- Se identificaron las condiciones óptimas para la extracción de los rasgos individuales en el espacio tiempo-frecuencia del EEG a través de la STFT.

Una vez más, cuando estos avances fueron publicados en el 2015, representaron el primer análisis exhaustivo de las propiedades de los rasgos individuales dentro del espacio tiempo-frecuencia del EEG [63].

A continuación, se evaluó la hipótesis propuesta con dos experimentos dedicados (sección 5.4). Como resultado, ***proporcionamos, por primera vez, evidencias de la existencia de una firma neuronal independiente del estado cognitivo***. Esto sugiere que la actividad neuronal, grabada por el EEG, está principalmente determinada por la identidad del sujeto y no por la tarea realizada. De ser así, esto podría tener implicaciones importantes para el estudio del cerebro humano y de su funcionamiento.

Finalmente, se aplicaron estos avances para el diseño de un sistema de verificación biométrica del EEG independiente del estado cognitivo (capítulo 6). Se identifican los $\mathrm{RCeps}_{20\,\%}$ y los LCF_{8} como dos potentes descriptores capaces de codificar la mayor parte de la información discriminatoria contenida en la forma del espectro del EEG. Aunque LCF había sido una opción popular en la literatu-

ra, este es el primer trabajo en usar RCeps. Gracias al análisis previo, fue posible explicar el funcionamiento de estas dos características.

Además de las innovaciones detalladas, también se ha presentado una nueva metodología para la eliminación de artefactos en el EEG: LCF (capítulo 4). Ésta fue desarrollada en el Departamento de Psicología de la Facultad de Salud y Ciencia Humana, Universidad de Swansea (Gales, Reino Unido). Dicha metodología se aplicó para obtener una de las versiones libres de artefactos de las bases de datos.

Bibliografía

[1] T. Collura, "History and evolution of electroencephalographic instruments and techniques," *Journal of Clinical Neurophysiology*, vol. 10, no. 4, pp. 476–504, 1993.

[2] H. Begleiter and B. Porjesz, "Genetics of human brain oscillations," *International Journal of Psychophysiology*, vol. 60, pp. 162–171, May 2006.

[3] B. Zietsch, J. Hansen, N. Hansell, G. Geffen, N. Martin, and M. Wright, "Common and specific genetic influences on eeg power bands delta, theta, alpha, and beta," *Biological Psychology*, vol. 75, pp. 154–164, January 2007.

[4] R. Palaniappan and D. Mandic, "Biometrics from brain electrical activity: A machine learning approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 738–742, 2007.

[5] *Biometric Authentication using Brain Responses to Visual Stimuli*, vol. 6406, INSTICC Press, 2010.

[6] A. Zźquete, B. Quintela, and J. Silva Cunha, "Biometric authentication using electroencephalograms: a practical study using visual evoked potentials," *Electrónica e Telecomunicazões*, vol. 5, no. 2, pp. 185–194, 2010.

[7] K. Brigham and B. Kumar, "Subject identification from electroencephalogram (EEG) signals during imagined speech," in *Fourth IEEE International Con-*

ference on Biometrics: Theory Applications and Systems (BTAS), pp. 1–8, IEEE, 2010.

[8] *A Proposed Feature Extraction Method for EEG-based Person Identification*, 2012.

[9] H. X., S. Altahat, D. Tran, and D. Sharma, "Human identification with electroencephalogram (eeg) signal processing," in *International Symposium on Communications and Information Technologies (ISCIT)*, pp. 1021–1026, 2012.

[10] F. Vogel, *Genetics and the Electroencephalogram*. Springer Verlag, 2000.

[11] C. van Beijsterveldt and G. van Baal, "Twin and family studies of the human electroencephalogram: a review and a meta-analysis," *Biological Psychology*, vol. 61, pp. 111–138, October 2002.

[12] S. Eischen, J. Luckritz, and J. Polich, "Spectral analysis of eeg from families," *Biological Psychology*, vol. 41, no. 1, pp. 61–8, 1995.

[13] G. Mohammadi, P. Shoushtari, B. Molaee Ardekani, and M. B. Shamsollahi, "Person identification by using ar model for eeg signals," in *Proceeding of World Academy of Science, Engineering and Technology*, vol. 11, pp. 281–285, 2006.

[14] P. Campisi, G. Scarano, F. Babiloni, F. DeVico Fallani, S. Colonnese, E. Maiorana, and L. Forastiere, "Brain waves based user recognition using the "eyes closed resting conditions" protocol," in *IEEE International Workshop on Information Forensics and Security, WIFS'11*, pp. 1–6, IEEE, 2011.

[15] D. La Rocca, P. Campisi, and G. Scarano, "Eeg biometrics for individual recognition in resting state with closed eyes," in *Proceedings of the Internatio-*

nal Conference of the Biometrics Special Interest Group (BIOSIG), pp. 1–12, 2012.

[16] D. La Rocca, P. Campisi, B. Vegso, P. Cserti, G. Kozmann, F. Babiloni, and F. De Vico Fallani, "Human brain distinctiveness based on eeg spectral coherence connectivity," *Biomedical Engineering, IEEE Transactions on*, vol. 61, no. 9, pp. 2406–2412, 2014.

[17] P. Tangkraingkij, C. Lursinsap, S. Sanguansintukul, and T. Desudchit, "Selecting relevant eeg signal locations for personal identification problem using ica and neural network," in *Eighth IEEE/ACIS International Conference on Computer and Information Science, ICIS'09*, pp. 616–621, 2009.

[18] P. Tangkraingkij, C. Lursinsap, S. Sanguansintukul, and T. Desudchit, "Personal identification by eeg using ica and neural network," in *International Conference on Computational Science and Its Applications - Part III*, vol. 6018 of *Lecture Notes in Computer Science*, pp. 419–430, Springer-Verlag, March 2010.

[19] M. Abdullah, K. Subari, J. Loong, and N. Ahmad, "Analysis of the EEG signal for a practical biometric system," *World Academy of Science, Engineering and Technology*, vol. 4, no. 8, pp. 1123–1127, 2010.

[20] M. Abdullah, K. Subari, J. Loong, and N. Ahmad, "Analysis of effective channel placement for an eeg-based biometric system," in *IEEE EMBS Conference on Biomedical Engineering and Sciences, IECBES*, pp. 303–306, IEEE, 2010.

[21] K. Das, S. Zhang, B. Giesbrecht, and P. Eckstein, "Using rapid visually evoked eeg activity for person identification," in *International Conference on Engineering in Medicine and Biology Society*, pp. 2490–2493, 2009.

[22] S. Yeom, H. Suk, and S. Lee, "Eeg-based person authentication using face-specific self representation," vol. 38 of *Proceedings of the Korea Computer Congress*, pp. 379–382, 2011.

[23] S. Yeom, H. Suk, and S. Lee, "Person authentication from neural activity of face-specific visual self-representation," *Pattern Recognition*, vol. 46, no. 4, pp. 1159–1169, 2013.

[24] M. Bear, B. Connors, and M. Paradiso, *Neuroscience: Exploring the Brain*. Lippincott Williams & Wilkins, third ed., 2006.

[25] S. Yang and F. Deravi, "On the effectiveness of eeg signals as a source of biometric information," in *Third International Conference on Emerging Security Technologies, EST'12*, pp. 49–52, 2012.

[26] C. van Beijsterveldt, P. Molenaar, E. de Geus, and D. Boomsma, "Genetic and environmental influences on eeg coherence," *Behavior Genetics*, vol. 28, pp. 443–453, November 1998.

[27] R. Palaniappan and D. Mandic, "Eeg based biometric framework for automatic identity verification," *The Journal of VLSI Signal Processing*, vol. 49, no. 2, pp. 243–250, 2007.

[28] D. Xiao and J. Hu, "Identification of motor imagery eeg signal," in *International Conference on Biomedical Engineering and Computer Science. ICBECS*, pp. 1–4, 2010.

[29] M. Poulos, M. Rangoussi, and N. Alexandris, "Neural network based person identification using eeg features," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP'99*, vol. 2, pp. 1117–1120, 1999.

[30] M. Poulos, M. Rangoussi, N. Alexandris, and A. Evangelou, "On the use of eeg features towards person identification via neural networks," *Medical Informatics and the Internet in Medicine*, vol. 26, no. 1, pp. 35–48, 2001.

[31] Q. Zhao, H. Peng, B. Hu, L. Li, Y. Qi, Q. Liu, and L. Liu, "Towards an efficient and accurate eeg data analysis in eeg-based individual identification," in *Ubiquitous Intelligence and Computing*, vol. 6406 of *Lecture Notes in Computer Science*, pp. 534–547, Springer Berlin / Heidelberg, 2010.

[32] F. Kennet, "Brain wave based authentication," Master's thesis, Gjøvik University College, Department of Computer Science and Media Technology, June 2008.

[33] X. Bao, J. Wang, and J. Hu, "Method of individual identification based on electroencephalogram analysis," in *International Conference on New Trends in Information and Service Science, NISS'09*, pp. 390–393, 2009.

[34] D. Jiang and J. Hu, "Research of computing in eeg password based on wavelet," in *, FBIE'09*, pp. 98–101, 2009.

[35] I. Nakanishi, K. Ozaki, and S. Li, "Evaluation of the brain wave as biometrics in a simulated driving environment," in *Proceedings of the International Conference of the Biometrics Special Interest Group, BIOSIG'12*, pp. 1–5, 2012.

[36] M. Fraschini, A. Hillebrand, M. Demuru, L. Didaci, and G. Marcialis, "An eeg-based biometric system using eigenvector centrality in resting state brain networks," *Signal Processing Letters, IEEE*, vol. 22, no. 6, pp. 666–670, 2015.

[37] M. Näpflin, M. Wildi, and J. Sarnthein, "Test–retest reliability of resting eeg spectra validates a statistical signature of persons," *Clinical Neurophysiology*, vol. 118, pp. 2519–2524, November 2007.

[38] I. Nakanishi, S. Baba, and C. Miyamoto, "Eeg based biometric authentication using new spectral features," in *International Symposium on Intelligent Signal Processing and Communication Systems, ISPACS'09*, pp. 651–654, 2009.

[39] S. Marcel and J. Millan, "Person authentication using brainwaves (eeg) and maximum a posteriori model adaptation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 743–752, 2007.

[40] F. Su, L. Xia, A. Cai, Y. Wu, and J. Ma, "Eeg-based personal identification: from proof-of-concept to a practical system," in *20th International Conference on Pattern Recognition, ICPR'10*, pp. 3728–3731, 2010.

[41] F. Su, L. Xia, A. Cai, and J. Ma, "Evaluation of recording factors in eeg-based personal identification: A vital step in real implementations," in *IEEE International Conference on Systems Man and Cybernetics, SMC*, pp. 3861–3866, 2010.

[42] F. Su, H. Zhou, Z. Feng, and J. Ma, "A biometric-based covert warning system using eeg," in *5th IAPR International Conference on Biometrics (ICB)*, pp. 342–347, 2012.

[43] B. Hu, Q. Liu, Q. Zhao, Y. Qi, and H. Peng, "A real-time electroencephalogram (eeg) based individual identification interface for mobile security in ubiquitous environment," in *IEEE Asia-Pacific Services Computing Conference, APSCC*, pp. 436–441, 2011.

[44] M. Kostílek and J. Šťastný, "Eeg biometric identification: Repeatability and influence of movement-related eeg," in *International Conference on Applied Electronics (AE)*, pp. 147–150, 2012.

[45] R. Palaniappan, "Identifying individuality using mental task based brain computer interface," in *Third International Conference on Intelligent Sensing and Information Processing, ICISIP'05*, pp. 238–242, 2005.

[46] R. Palaniappan, "Electroencephalogram signals from imagined activities: a novel biometric identifier for a small population," in *Proceedings of the 7th international conference on Intelligent Data Engineering and Automated Learning, IDEAL'06*, pp. 604–611, 2006.

[47] S. Sun, "Multi-task learning for eeg-based biometrics," in *19th International Conference on Pattern Recognition, ICPR'08*, pp. 1–4, 2008.

[48] J. Hu, "New biometric approach based on motor imagery eeg signals," in *International Conference on Future BioMedical Information Engineering, FBIE'09*, pp. 94–97, 2009.

[49] R. Palaniappan, "Two-stage biometric authentication method using thought activity brain waves," *International Journal of Neural Systems*, vol. 18, no. 1, pp. 59–66, 2008.

[50] X. Huang, S. Altahat, D. Tran, and L. Shutao, "Human identification with electroencephalogram (eeg) for the future network security," in *7th International Conference on Network and System Security*, vol. 7873 of *Lecture Notes in Computer Science*, pp. 575–581, 2013.

[51] S. D. Muthukumaraswamy, "High-frequency brain activity and muscle artifacts in MEG/EEG: A review and recommendations," *Frontiers in human neuroscience*, vol. 7, no. 128, pp. 1–11, 2013.

[52] M. Fatourechi, A. Bashashati, R. K. Ward, and G. E. Birch, "Emg and eog artifacts in brain computer interface systems: A survey," *Clinical neurophysiology*, vol. 118, no. 3, pp. 480–494, 2007.

[53] J. B. Alonso, *La evaluación acústica del sistema fonador*. Universidad de Las Palmas de Gran Canaria, Vicerrectorado de Calidad e Innovación Educativa, 2008.

[54] R. Paranjape, J. Mahovsky, L. Benedicenti, and Z. Koles, "The electroencephalogram as a biometric," in *Proceedings on Canadian Conference on Electrical and Computer Engineering*, vol. 2, pp. 1363–1366, 2001.

[55] D. La Rocca, P. Campisi, and G. Scarano, "Stable eeg features for biometric recognition in resting state conditions," in *Biomedical Engineering Systems and Technologies*, vol. 452 of *Communications in Computer and Information Science*, pp. 313–330, Springer, November 2014.

[56] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman, "A default mode of brain function," *Proceedings of the National Academy of Sciences*, vol. 98, no. 2, pp. 676–682, 2001.

[57] D. A. Gusnard and M. E. Raichle, "Searching for a baseline: functional imaging and the resting human brain," *Nature Reviews Neuroscience*, vol. 2, no. 10, pp. 685–694, 2001.

[58] M. E. Raichle and A. Z. Snyder, "A default mode of brain function: a brief history of an evolving idea," *Neuroimage*, vol. 37, no. 4, pp. 1083–1090, 2007.

[59] P. Campisi and D. La Rocca, "Brain waves for automatic biometric based user recognition," 2014.

[60] T. Pham, W. Ma, D. Tran, P. Nguyen, and D. Phung, "Multi-factor eeg-based user authentication," in *Neural Networks (IJCNN), 2014 International Joint Conference on*, pp. 4029–4034, IEEE, 2014.

[61] H. Kayyali, "Medical device and method with improved biometric verification," 2014. US Patent 8,679,012.

[62] M. del Pozo-Baños, J. Alonso, J. Ticay-Rivas, and C. Travieso, "Electroencephalogram subject identification: A review," *Expert Systems with Applications*, vol. 41, pp. 6537–6554, May 2014.

[63] M. del Pozo-Baños, C. Travieso, C. Weidemann, and J. Alonso, "EEG biometric identification: A thorough exploration of the time-frequency domain," *Journal of Neural Engineering*, vol. 0, p. 0, September 2015.

Ph. D. Thesis

As early as 1936 -- only 12 years after Hans Verger recorded the first human Electroencephalogram (EEG) -- twin EEG research by H. and P. Davis evidence the existence of a brain activity inheritance model. In fact, the fields of biometrics has successfully used this property to build systems capable of identifying users from their EEG.

Here we present evidence supporting the hypothesis: "There exists, within the human EEG, a task-independent pattern concomitant to the individual's identity". In doing so we have: (1) researched the state of the art in EEG identification biometry and, to a lesser degree, equivalent genetic and neurophisiological studies; (2) conducted an in depth analysis (both qualitatively and quantitatively) of the time-frequency representation of the EEG from a biometrical point of view; and (3) designed and implemented a task-independent EEG biometric verification system. These research stages resulted in: (1) the identification of a lack of consensus in the literature on several fundamental questions regarding the subject traits within the EEG; (2) a description of properties of the discriminant information within the time-frequency space of the EEG and a set of guidelines for its optimal extraction -- shedding light in most of the unanswered questions identified in (1) --; and (3) a task-independent EEG biometric verification system with success rates between 73.94% and 97.59% in 6 publicly available databases. In addition, during a research stay at the Psychology Department of the College of Human Health and Science, Swansea University (Wales, UK), we developed a new methodology for the enhancement of EEG artefact rejection systems based on blind source separation methods.

by Marcos del Pozo Baños

Supervised by: Dr Carlos Manuel Travieso González
Dr Jesús Bernardino Alonso Hernández

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

INSTITUTO UNIVERSITARIO SIANI INGENIERIA COMPUTACIONAL