

Université du Québec
Institut national de la recherche scientifique
Centre Énergie Matériaux Télécommunications

**NEW FEATURES FOR AUTOMATED AFFECTIVE STATE RECOGNITION
USING MULTIMODAL BRAIN COMPUTER INTERFACES**

By

Andrea Clerico

A thesis submitted in fulfillment of the requirements for the degree of
Master of Science, M.Sc.
in Telecommunications

Evaluation committee

Internal evaluator and committee president:	Prof. Douglas O'Shaughnessy
External evaluator:	Dr. Fabien Lotte Institut national de recherche en informatique et en automatique
Research advisor:	Prof. Tiago H. Falk

Acknowledgements

First of all, I would like to express my gratitude to my supervisor, Dr. Tiago H. Falk, whose expertise added considerably to my graduate experience. Without his support, advice, great ideas, and attention to detail, this work would not have been possible.

I would like to extend my thanks to the members of the evaluation committee, Prof. Douglas O'Shaughnessy and Prof. Fabien Lotte, as well as to the INRS-EMT staff and faculty. This research was made possible by funding from the NSERC CREATE-PERSWADE. I would also like to thank the co-author of our papers: Rishabh Gupta, who have provided constant help and good ideas.

The Multimedia/Multimodal Signal Analysis and Enhancement (MuSAE) Lab is a wonderful environment for research and not only because we have the resources and expertise, but because of the great people that are part of it. I feel honoured to be part of such a team.

Un grand merci à Hubert, Pascal et Alexandre pour le support pendant les derniers deux années et l'aide essentiel avec la partie en français.

A huge thank you goes to my friends, both the ones that are close and far from me. You always represent my acquired family.

Infine, vorrei dedicare questo traguardo alla mia famiglia. Per il loro supporto costante, in ogni luogo e in ogni momento, per il loro incoraggiamento, che non é venuto mai meno. Sono riusciti, involontariamente, ad essere ancora una volta indispensabili. Spero che voi tutti, in un modo o nell'altro, sarete capaci di leggere queste parole.

"Non sarà facile ma sai,
Si muore un po' per poter vivere..."

Abstract

The quantity of multimedia content is rapidly expanding and automated affective recognition can enable the development of intelligent retrieval, music recommendation, video playlist generators, and browsing interfaces to the users' current desires, preferences, or affective states. To achieve this goal, the field of affective computing has emerged, in particular the development of so-called affective brain-computer interfaces (aBCIs), which measure the user's affective state directly from measured brain waves using non-invasive tools, such as electroencephalography (EEG).

Typically, conventional features extracted from the EEG signal have been used, such as frequency subband powers and/or inter-hemispheric power asymmetry indices. More recently, the coupling between EEG and peripheral physiological signals, such as the galvanic skin response (GSR), have also been proposed. Here, we first show the importance of EEG amplitude modulations and propose several new features that measure the amplitude-amplitude cross-frequency coupling per EEG electrode, as well as linear and non-linear connections between multiple electrode pairs. When tested on a publicly available dataset of music video clips tagged with subjective affective ratings, the proposed features were shown to outperform conventional benchmark EEG features by as much as 20%. When combined with EEG-GSR coupling features, further improvements were observed. Overall, decision level fusion with support vector machine classifiers trained on three proposed amplitude modulation feature sets resulted in improvements of 22% over the benchmark features, thus showing the importance of the proposed system to characterize human affective states during music clip watching.

In a second study, we explore the combination of two neuroimaging modalities to further improve affective BCI accuracy. More specifically, we explore the use of functional near-infrared spectroscopy (fNIRS) combined with EEG. Experimental results on a publicly available dataset of human emotions elicited by images showed improvements of up to 8% could be achieved in emotion recognition once fNIRS and EEG were combined, relative to using the individual modalities alone.

Keywords Emotion Classification, Affective Computing, Multimedia Content, Electroencephalography, Physiological signals, Signal Processing, Pattern classification, fNIRS

Contents

Acknowledgements	iii
Abstract	v
Contents	vii
List of Figures	ix
List of Tables	xi
Liste des Figures	xi
Liste des Tableaux	xv
List of Abbreviations	xvii
Sommaire récapitulatif	xix
0.1 Introduction	xix
0.1.1 ICO affectives	xx
0.1.2 Hybrid BCI	xxv
0.1.3 Objectifs de recherche et contribution	xxv
0.1.4 Organisation du document	xxvi
0.2 Échelles subjectives d'émotions	xxvi
0.2.1 Émotions	xxvii
0.2.2 MAE: mannequins d'auto-évaluation	xxix
0.3 Traits caractéristiques	xxx
0.3.1 Intégration multimodale : EEG-SPIR	xxxii
0.3.2 Tâches et bases de données	xxxii
0.4 Résultats expérimentaux et discussion	xxxv
0.4.1 Tâche 1: Résultats expérimentaux	xxxv
0.4.2 Tâche 1: Discussion	xxxvi
0.4.3 Tâche 2: Résultats expérimentaux et Discussion	xxxviii
0.5 Conclusion	xxxviii
1 Introduction	1
1.1 Affective BCIs	3
1.2 Hybrid BCI	15

1.3	Thesis objectives and contributions	17
1.4	Outline	17
2	Subjective scales of emotions	19
2.1	Emotions	20
2.2	SAM: The Self-Assessment Manikin	23
3	Methods and Materials	27
3.1	Features	28
3.1.1	Benchmark EEG Features	28
3.1.2	Proposed EEG Features: Amplitude Modulation	28
3.1.3	Proposed Features: PANS-CNS coupling	33
3.1.4	Multimodal Integration: NIRS-EEG	36
3.2	Tasks and Databases	37
3.2.1	Dataset for Emotional Analysis using EEG and Physiological signals	37
3.2.2	eNTERFACE 06'	38
3.3	Feature Selection	39
3.4	Classification	40
4	Results and Discussion	43
4.1	Task 1: Experimental Results	43
4.2	Task 1: Discussion	48
4.3	Task 2: Experimental Results and Discussion	52
5	Conclusion	55
5.1	Summary of the Research Outcomes	55
5.2	Future Research and Directions	56
	Bibliography	59

List of Figures

1.1	Hemodynamic activation response detected with a 7 channels Functional Near-Infrared Spectroscopy (fNIRS) detector.	8
1.2	Light from the light source is guided to the head by an optode. A photo-detector will collect the light which leaves the head at a distance of some centimetres. The photons follow a banana-shaped path from light source to detector.	9
2.1	First two layers of Parrot’s emotion classification.	21
2.2	Plutchik’s Wheel of Emotions	23
2.3	Circumplex of affects [113]. Discrete emotional terms were organised around the two dimensions of arousal (top to bottom) and valence (left to right). . .	24
2.4	The Self-Assessment Manikin (SAM) used to rate the affective dimensions of valence (top panel), arousal (middle panel), and dominance (bottom panel). . .	25
3.1	Signal processing steps used to compute the EEG amplitude modulation feature sets.	29
3.2	Protocol description of the acquisition of the eNTERFACE 06’ database. . .	39
4.1	Accuracy versus number of features for the baseline, IAMI, and fused feature sets for the (a) arousal, (b) valence, (c) dominance, and (d) liking affective dimensions.	46
4.2	Accuracy versus number of features for the AMF, PAC, and SF feature sets for the (a) arousal, (b) valence, (c) dominance, and (d) liking affective dimensions. . .	47

List of Tables

4.1	Selected 20 best features using minimum Redundancy Maximum Relevance (mRMR) algorithm for the arousal and valence dimensions. In the AMI, AMC and AME columns, the top-20 features corresponded to the log-ratio ones. In the <i>SF</i> column, <i>AI</i> indicates the asymmetry index between the indicated channels.	44
4.2	Selected 20 best features using mRMR algorithm for the dominance and liking dimensions. In the AMI, AMC and AME columns, the top-20 features corresponded to the log-ratio ones. In the <i>SF</i> column, <i>AI</i> indicates the asymmetry index between the indicated channels.	45
4.3	Performance comparison of SVM classifiers for different feature sets and feature-level fusion along with the required number of features needed to achieve such results. Asterisks indicate accuracy significantly higher than chance achieved with a random voting classifier ($p < 0.05$). Values within parentheses indicate the relative improvement, in percentage, over the SF baseline set.	48
4.4	Performance comparison of different decision-level fusion schemes along with the required number of features needed to achieve such results. Asterisks indicate accuracy significantly higher than chance achieved with a random voting classifier ($p < 0.05$). Values within parentheses indicate the relative improvement, in percentage, over the SF baseline set.	48
4.5	Performance comparison of different feature and decision (*) level fusion schemes considering the contribution of fNIRS to three different datasets (EEG, PHY and EEG + PHY) for the two affective states arousal and valence.	53
4.6	Performance comparison of different feature-level fusion schemes considering the contribution of fNIRS to three different datasets (EEG, PHY and EEG + PHY) for the two affective states arousal and valence.	53

Liste des Figures

1.1	Réponse d'activation hémodynamique détectée avec un détecteur 7 canaux de fNIRS.	8
1.2	La lumière provenant de la source lumineuse est guidée vers la tête d'une optode. Une photo-détecteur collectera la lumière qui laisse la tête à une distance de quelques centimètres. Les photons suivent une trajectoire en forme de banane de la source de lumière au détecteur.	9
2.1	Deux premières couches de la classification de l'émotion de Parrott.	21
2.2	La roue des émotions de Plutchik.	23
2.3	Circumplex de l'affects [113]. termes émotionnels discrets ont été organisées autour des deux dimensions de l'excitation (de haut en bas) et Valence (de gauche à droite)	24
2.4	Le mannequin d'auto-évaluation (SAM) utilisé pour évaluer les dimensions affectives de valence (en haut), l'excitation (panneau du milieu), et la domination (panneau du bas).	25
3.1	Étapes de traitement de signal utilisé pour calculer la modulation d'amplitude de le signal EEG.	29
3.2	Description du protocole de l'acquisition de la base de données de eNTERFACE 06'.	39
4.1	Précision par rapport au nombre de composantes du signal pour le baseline, proposé, et les ensembles de composantes pour les dimensions affectives (a) excitation, (b) valence, (c) domination, et (d) aimer.	46
4.2	Précision par rapport au nombre de composantes pour les AMF, PAC, et SF sets pour les dimensions affectives (a) excitation, (b) valence, (c) domination, et (d) aimer.	47

Liste des Tableaux

4.1	20 meilleures composantes du signal en utilisant l'algorithme mRMR pour les dimensions de l'excitation et de valence. Dans les colonnes AMI, AMC et AME, les meilleures-20 composantes correspondent aux les quelles log-ratio. Dans la colonne <i>SF</i> , <i>AI</i> indique l'indice d'asymétrie entre les canaux indiqués.	44
4.2	20 meilleures composantes du signal en utilisant l'algorithme mRMR pour les dimensions de la domination et l'aime. Dans les colonnes AMI, AMC et AME, les meilleures-20 composantes correspondent aux les quelles log-ratio. Dans la colonne <i>SF</i> , <i>AI</i> indique l'indice d'asymétrie entre les canaux indiqués.	45
4.3	Comparaison des performances des classificateurs SVM pour différents ensembles de composants et de leur fusion ainsi que le nombre requis de composants nécessaires pour obtenir de tels résultats. Les astérisques indiquent une précision nettement plus élevé que le hasard réalisé avec un classificateur de vote aléatoire ($p < 0,05$). Les valeurs entre parenthèses indiquent l'amélioration relative, en pourcentage, sur la ligne du baseline SF.	48
4.4	Comparaison des performances des différents régimes de fusion au niveau de la décision ainsi que le nombre requis de composants nécessaires pour obtenir de tels résultats. Les astérisques indiquent une précision nettement plus élevé que le hasard réalisé avec un classificateur de vote aléatoire ($p < 0,05$). Les valeurs entre parenthèses indiquent l'amélioration relative, en pourcentage, sur la ligne du baseline SF.	48
4.5	Comparaison des performances des différents ensamble des composants et fusion ou niveau de la décision (<i>ast</i>) tenu compte de la contribution de NIRS à trois ensembles de données différents (EEG, PHY et EEG + PHY) pour les deux états affectifs éveil et de valence.	53
4.6	Comparaison des performances des différents ensamble des composants tenu compte de la contribution de NIRS à trois ensembles de données différents (EEG, PHY et EEG + PHY) pour les deux états affectifs éveil et de valence.	53

List of Abbreviations

aBCI	Affective Brain-Computer Interface
AEM	Affective Evaluation Methods
AMC	Amplitude Modulated Coherence
AME	Amplitude Modulated Energy
AMI	Amplitude Modulated Interaction
AMF	Amplitude Modulation Features
ANS	Autonomic Nervous System
AM	Autoregressive Model
BCI	Brain-Computer Interface
BP	Blood Pressure
BVP	Blood Volume Pulse
CFC	Cross-Frequency Coherence
CNS	Central Nervous System
ECG	Electrocardiography
ECoG	ElectroCorticoGraphy
EEG	Electroencephalography
EMG	Electromyography
ESC	Envelope-to-Signal Coupling
ERD	Event Related Desynchronization
ERP	Event Related Potential
ERSP	Event Related Spectral Power
EP	Evoked Potential
fMRI	Functional Magnetic Resonance Imaging

fNIRS	Functional Near-Infrared Spectroscopy
GSR	Galvanic Skin Response
hBCI	Hybrid Brain-Computer Interface
HHI	Human-Human Interaction
HR	Heart Rate
HRV	Heart Rate Variability
HMI	Human-Machine Interface
ICA	Independent Component Analysis
IAMI	Inter-Hemispheric Amplitude Modulated Interaction
IAPS	International Affective Picture System
LOSO	Leave-one-sample-out
mRMR	minimum Redundancy Maximum Relevance
MBLL	Modified Beer-Lambert Law
MEG	Magnetoencephalography
ModI	Modulation Index
PANS	Peripheral Autonomic Nervous Systems
PAC	Phase-Amplitude Coupling
PCA	Principal Components Analysis
PNS	Peripheral Nervous System
PPG	Photoplethysmography
RBF	Radial Basis Function
RVM	Relevance Vector Machine
SAM	Self-Assessment Manikins
SF	Spectral Features
ST	Skin Temperature
SMR	Sensorimotor Rhythm
SVM	Support Vector Machine

Sommaire récapitulatif

0.1 Introduction

Les émotions sont généralement conçues comme des réponses physiologiques et physiques, une partie de la communication naturelle entre les êtres humains; elles sont en mesure d’influencer notre intelligence, de façonner nos pensées et régissent nos relations interpersonnelles [26, 81, 83]. Dans le cadre de la recherche fondamentale, elles ont toujours été considérées comme inutiles et difficiles à encoder sous la forme d’un programme informatique. Or, des découvertes récentes en neurosciences, en psychologie et en sciences cognitives ont modifié cette idée et délimitée leur véritable rôle en ce qui concerne le comportement rationnel et intelligent. Les compétences émotionnelles contribuent ainsi massivement aux préférences d’apprentissage et d’adaptation, jouant ainsi le rôle de composante de base de l’intelligence.

À ce jour, les machines n’ont pas été tenues d’avoir des compétences émotionnelles; l’interaction humain-machine (IHM) typiquement impliqué consiste généralement en l’envoi d’information par des humains à des ordinateurs via un clavier et/ou une souris. Cependant, avec les progrès enregistrés récemment dans le domaine de l’informatique affective [102], ce paradigme est sur le point de changer. On s’attend de plus en plus des machines qu’elles possèdent un niveau d’intelligence de base afin de créer une IHM appropriée [108], suivant ainsi les principes d’une interaction humain-humain. En fin de compte, le domaine de l’informatique affective a vu le jour avec l’objectif de réduire la distance entre l’utilisateur final et la machine; il serait alors possible d’obtenir une interaction adaptative dans le cadre duquel la machine comprend les émotions de l’utilisateur et ses états cognitifs, s’ajustant ainsi au contexte. De tels instruments permettent à l’ordinateur de mieux répondre aux besoins humains [91].

La naissance du domaine de l’informatique affective, combinée aux progrès technologiques des dernières années (par exemple, augmentation de la puissance de calcul, baisse du coût du stockage d’informations, amélioration de la vie des batteries, etc.), a abouti à la résurgence des interfaces cerveau-ordinateur (ICO). Traditionnellement, les ICO ont été utilisées pour aider les personnes souffrant de capacité de communiquer ou de mobilité limitées, comme dans le cas du syndrome d’enfermement. Par exemple, les ICO de type *P300-Speller* permettent d’épeler [78], tandis qu’il est possible de conduire des fauteuils roulants électriques avec

d'autres designs [76]. Suivant l'émergence des technologies portables, cependant, de nouvelles applications visant les consommateurs voient le jour, comme des jeux [12], des moniteurs d'attention [58], et des appareils mobiles [132], pour ne nommer que quelques-uns.

Alors qu'une interface ICO 'active' recueille des signaux neuronaux et les traduit en signaux de commande, une interface 'passive' recueille les états mentaux implicites de son utilisateur. Le concept d'ICO affectives (ICOa) est ainsi apparu. En outre, comme les technologies portables et différentes technologies de neuroimagerie deviennent plus populaires et accessibles, les ICO multimodales sont de plus en plus attirantes. Ces ICO multimodales, aussi appelées ICO hybrides, peuvent combiner différentes technologies de neuroimagerie, de neuroimagerie avec surveillance physiologique, ou même de neuroimagerie combinée à des technologies d'interface homme-machine classiques, tels qu'un clavier.

Une récente revue de la littérature [6] a démontré que les ICO hybrides combinant l'électroencéphalographie (EEG) et la spectroscopie proche-infrarouge fonctionnelle (SPIR) sont en plein essor. La majorité des études existantes, cependant, ont mis l'accent sur l'utilisation d'ICO multimodales hybrides actives, tel que défini ci-dessus. Dans ce mémoire, nous mettons l'accent sur les ICO hybrides passives, aussi appelées ICO hybrides affectives. Comme les ICO affectives et hybrides jouent un rôle clé dans ce travail, le reste de ce chapitre se concentre sur ces sujets. Le chapitre se termine par une description détaillée de la contribution de ce mémoire.

0.1.1 ICO affectives

Les ICO affectives découlent naturellement d'une tentative d'appliquer les principes de l'IHM en dehors du domaine médical avec des ICO. Les applications des ICOa concernent typiquement le divertissement, le style de vie, et l'ergonomie [12]; cela peut fournir un moyen de détecter les réactions spontanées et naturelles à des stimuli évoquant des émotions. Les ICOa peuvent être utilisées pour une communication des états émotionnels pour informer les machines sur les états affectifs de leurs utilisateurs [66], transmettant ainsi 'l'état de sensibilisation des utilisateurs' aux systèmes ICO existants. La conscience émotionnelle peut aider à créer plusieurs systèmes pouvant accueillir des utilisateurs et à développer des paradigmes de formation plus efficaces. Typiquement les ICO se répartissent en deux catégories: actives ou passives. Les ICO actives, comme leur nom l'indique, sont des systèmes qui traduisent des patrons d'activations neuronaux en signaux de commande machine. Les ICOs passives [142], d'autre part, permettent de détecter des informations implicites concernant l'état de l'utilisateur, même en l'absence d'une expression émotionnelle visible. En tant que tel, les ICO affectives ont souvent été classées comme ICO passives. Voici donc la terminologie qui sera employée dans ce texte.

Dans la littérature sur les ICO, l'électroencéphalographie (EEG) est une technique de neuroimagerie privilégiée en raison de ses caractéristiques bien connues, son accessibilité, de

même qu'en raison de sa haute résolution temporelle [56]. L'utilisation de l'EEG a commencé dans les années 1970 avec des études spectrales sur l'évaluation du contrôle de l'ordinateur pour des patients internés [23, 46], de même que pour déterminer le niveau de charge mental [55]. Des expériences antérieures ont mis l'accent sur la détection de traits caractéristiques capables de classer les conditions de faible et de forte charge mentale [50, 67]. Au fil du temps, les technologies d'analyse d'EEG se sont améliorées et ont permis une meilleure évaluation de l'activité corticale; la structure temporelle des potentiels électriques cérébraux, de même que la dynamique des réactions neuronales, mettent aujourd'hui en évidence l'activité d'intégration inter-corticale au sein de différents secteurs cérébraux.

La plupart des fonctions cognitives du cerveau sont basées sur les interactions coordonnées d'un grand nombre de neurones qui sont distribuées dans différentes zones spécialisées du cerveau. De plus, l'un des sujets les plus actifs en neurosciences est la synchronisation de paires individuelles de neurones à différents niveaux d'organisation du tissu cérébral, à beaucoup plus grande échelle, dans une zone du cerveau ou entre ses différentes parties [9]. Plus précisément, les opérations normales cognitives nécessitent l'intégration transitoire de nombreux domaines fonctionnels largement répartis dans le cerveau. Les signaux électrochimiques voyagent à travers le cerveau et le crâne; ils sont ensuite enregistrés par des électrodes appliquées directement sur le cuir chevelu. Les électrodes sont placées directement sur le cuir chevelu et sont généralement au nombre de 32, 64, 128 ou 256. Les légères variations de tension dans les axones des neurones sont appelées potentiels gradués. Chaque électrode enregistre la somme des potentiels gradués des nombreux neurones.

Outre l'avantage d'être applicable de manière non invasive, l'EEG est un instrument particulièrement intéressant en raison de sa haute résolution temporelle. Il est sensible aux changements qui surviennent dans l'espace de millisecondes. Ce qui est déterminant dans le choix d'un instrument tel que l'EEG est son coût relativement faible, rendant le dispositif accessible à tous les chercheurs. Ces trois caractéristiques font de l'EEG le dispositif plus largement utilisé. Cependant, l'inconvénient majeur est sa faible résolution spatiale, ce qui signifie qu'il est relativement inefficace pour différencier des régions spécifiques du cerveau.

Il est possible d'extraire des informations complémentaires à celles obtenues avec l'EEG en mesurant d'autres réactions psychologiques. Ces informations proviennent la plupart du temps du réflexe cutané galvanique (RCG) ou de la photopléthysmographie. Normalement utilisés pour des expériences cliniques [79], afin d'évaluer le stress mental et physique d'un sujet dans des situations et des tâches spécifiques, ce n'est que tout récemment que ces mesures ont été incluses dans la détection des émotions. Les signaux physiologiques comprennent également les signaux en provenance du système nerveux périphérique (SNP) [80, 94, 140]. Les expériences dans ce domaine ont pour but d'étudier et reconnaître les émotions et les états émotionnels connexes d'un utilisateur en réponse à des images, des vidéos ou des sons. La préférence émotionnelle est l'un des facteurs des plus importants lors du choix d'un contenu

multimédia et de sa consommation pour l'expérimentation. De plus, le fait de connaître l'émotion du spectateur en regardant les vidéos se révèle très important pour aider les systèmes de recommandation à mieux comprendre les préférences des usagers. L'information sur la réaction émotive de l'utilisateur peut être obtenue simplement par un questionnaire direct, à savoir l'auto-évaluation des émotions ou bien, comme alternative, par la détection des réponses physiologiques centrales. En outre, de récents développements obtenus dans le domaine de l'interaction homme-machine permettent désormais l'analyse émotionnelle en temps réel. Les signaux physiologiques qui sont générés par le corps pendant le fonctionnement de divers systèmes physiologiques sont alors encodés. Ainsi, certains signaux physiologiques contiennent des informations qui peuvent être extraites à partir de ces signaux pour inférer l'état de ces systèmes physiologiques. Le processus d'extraction d'informations peut être très simple: par exemple, le pouls permet de mesurer les battements de cœur. Un processus d'extraction d'information peut être bien plus complexe également, ce qui peut nécessiter une analyse de la structure de tissus à l'aide d'une machine sophistiquée.

Selon le type d'énergie, les signaux physiologiques peuvent être des:

- Signaux bioélectriques, produits par les cellules nerveuses et les cellules musculaires. Les signaux électrocardiographiques en sont un exemple: le changement de potentiel dans de nombreuses cellules permet de générer un champ électrique qui varie, créant du coup un signal bioélectrique.
- Signaux biomécaniques, liés au mouvement, à la pression et au débit d'un système physiologique. Un exemple est la respiration avec le mouvement de la poitrine.
- Signaux bioacoustiques, créés par les systèmes physiologiques impliqués dans la circulation sanguine et la respiration. L'écoulement du sang dans le cœur et l'ouverture et la fermeture des voies respiratoires génèrent tous deux des signaux acoustiques uniques.
- Signaux de bioimpédance: l'impédance de la peau dépend de la composition de la peau, de la distribution du sang et le volume sanguin à travers la peau. La mesure de l'impédance aide à trouver l'état de la peau et le fonctionnement des différents systèmes physiologiques. Un exemple est la réponse galvanique de la peau analysée dans ce rapport.
- Signaux biooptiques, détectés par la spectroscopie proche-infrarouge fonctionnelle (SPIR). Ces signaux sont produits par les variations optiques et le fonctionnement du système physiologique, comme l'oxygénation du sang mesurée en utilisant la lumière transmise et réfléchie passant par les vaisseaux sanguins.

Le cœur est non seulement un muscle qui sert à pomper le sang dans tout le corps, mais il est aussi relié au système nerveux autonome (SNA) et reflète les émotions et les sentiments d'un individu. Ainsi, cela signifie que le cœur humain en bonne santé ne bat pas à un rythme

précis, mais plutôt qu'il est un peu irrégulier en plus d'être constamment influencé par des stimuli externes ou internes. Par exemple, une situation de stress permet un ajustement de la réponse du cœur. La conséquence pourrait conduire à une augmentation ou une diminution de la fréquence cardiaque et la variation est connue sous le nom variabilité du rythme cardiaque (VRC).

La capacité du corps à conduire de l'électricité permet de transmettre l'information d'une partie à l'autre; tous les tissus organiques y participent. Ainsi, la peau conduit l'activité électrique; cette dernière peut être détectée et utilisée pour comprendre les changements dans le système nerveux sympathique, en particulier pour 'mesurer' les émotions par le décodage de stimuli internes ou externes. L'activité électrique de la peau est toujours en variation constante et légère. La fluctuation de la conductivité est appelé RCG et peut être détectée, analysée et même représentée graphiquement afin d'observer la réponse du cœur ou du cerveau. Or, de toute évidence, il est impossible de détecter le type d'émotion à l'œuvre uniquement à partir de l'état physique. Les dispositifs qui mesurent la réponse de la peau sont de faible coût et le signal est alors facilement enregistré. Ces dispositifs sont souvent désignés comme des instruments de rétroaction puisque la réponse ne se mesure pas de manière 'active'. En effet, la RCG mesure des phénomènes générés par le corps de la personne elle-même. Les évaluations rapportées sont la réponse galvanique de la peau mesurée. L'indice RCG a souvent été utilisé pour examiner l'état mental et les émotions. Par exemple, il a été utilisé pour distinguer les différents états de stress et les situations de la charge cognitive.

L'étude de la variabilité des voies respiratoires chez des sujets sains et pathologiques est une méthode intéressante pour analyser le comportement de l'organisme au cours d'activités diverses, ou tout simplement pendant le sommeil. Par exemple, comme il est expliqué dans [116], cette méthode a été utilisée pour détecter des événements d'apnée du sommeil; il était alors utile de distinguer les signaux de la respiration normale, des traces de mouvement et de l'apnée du sommeil.

Le dernier signal physiologique étudié dans ce rapport est la température de la peau (TP). Il décrit les variations de température mesurée sur la surface de la peau. Ce signal a été utilisé dans de nombreuses études soit comme indice physiologique pour surveiller des humains ou des animaux, ou bien pour analyser la réponse émotionnelle à des facteurs extérieurs. Par exemple, dans [66], le signal a été enregistré et analysé pour comprendre la réponse des sujets pendant qu'ils regardaient des vidéoclips musicaux; dans cette situation une augmentation de TP est due à l'embarras. En outre, l'indice TP a l'avantage de pouvoir être évalué sur les deux côtés du corps et, par conséquent, l'asymétrie de la température peut être mesurée. Les variations de l'asymétrie latérale de TP peuvent être liées à des variations individuelles chez les sujets [111].

La technique SPIR consiste à mesurer la demande métabolique locale à la surface du cortex. Cette technique non invasive d'imagerie optique permet de détecter l'absorption de

lumière dans le sang oxygéné et désoxygéné. Elle offre également une alternative à l'imagerie par résonance magnétique fonctionnelle (IRMf), à la magnétoencéphalographie (MEG) ou à l'EEG chez les populations étudiées, de même que pour les populations particulières qui sont plus difficiles à tester. En effet, le cou des nourrissons étant fragiles, ils doivent reposer leur tête sur les genoux d'un parent lorsque que le test d'imagerie a lieu; cette contrainte fait alors de la technique SPIR un examen particulièrement adapté pour eux. En outre, au cours de la mesure de l'activité cérébrale, le sujet effectue simplement une tâche opératoire afin de produire des motifs caractéristiques de leur activité cérébrale; la réponse hémodynamique est alors induite et transmise en tant que sortie. La position des détecteurs est souvent limitée à la région frontale du cerveau comme il ne s'y trouve pas de cheveux obstruant la transmission de la lumière; il est toutefois possible d'imager les autres régions du cuir chevelu. Une différence notable à considérer toutefois pour la sensibilité de la mesure est la couleur de la peau. Il faut spécifier que les signaux de SPIR peuvent perdre de leur qualité sur les peaux foncées.

La Fig. 1.1 présente un sujet portant les détecteurs de SPIR sur son front. De petits systèmes portables détectent des signaux entre l'émetteur et le détecteur (de la forme d'une "banane"), comme le montre la Fig. 1.2, à l'aide de petits émetteurs et de détecteurs opérant grâce à une lumière infrarouge projetée sur le cuir chevelu. Un simple casque maintient ces optodes en place. La transmission en forme de banane pénètre généralement dans le crâne jusqu'à 2-3 cm à l'intérieur du cortex. C'est alors que deux longueurs d'onde différentes sont émises, grâce auxquelles le système peut mesurer séparément l'oxy-hémoglobine et désoxy-hémoglobine (par exemple, l'appareil Hitachi ETG-4000 utilise des longueurs d'onde de 690 nm et 830 nm).

Il y a des avantages considérables reliés à l'utilisation de la SPIR: l'appareil peut être transporté dans divers lieux tels que les foyers des particuliers, des hôpitaux dotés d'unités néonatales, et dans des écoles. Ce système est mobile puisque sa taille est relativement plus petite (que l'IRMf), il est portable, et a recours à la lumière plutôt qu'à des champs magnétiques; de surcroît, il est complètement silencieux, ce qui permet son utilisation dans des contextes naturels. Ainsi, la SPIR possède de nombreux avantages lorsque l'on étudie l'acquisition précoce du langage.

Les concentrations d'hémoglobine oxygénée, désoxygénée et totale peuvent être calculées pour la lumière individuelle sous la forme de signaux dépendant du temps grâce à une sonde émetteur-détecteur, laquelle est souvent conçue pour une région spécifique du cerveau; ces signaux peuvent représenter la réponse hémodynamique moyenne d'un participant aux stimuli sensoriels ou à des réponses motrices. Ces signaux fournissent la base des données d'un système SPIR. Typiquement, la position des paires d'optodes permet de localiser topographiquement les signaux. La loi de Beer-Lambert modifiée 1.1, comme indiqué dans [51], est utilisée pour calculer la concentration d'hémoglobine oxygénée ΔHbO et désoxygéné ΔHbR en fonction de leur absorption de la lumière, où ΔOD est la variation de l'intensité lumineuse,

l est la distance géométrique entre la source et le détecteur, α_{HbO} et α_{HbR} sont les coefficients d'extinction spécifique de l'hémoglobine, et où b est la longueur du trajet des photons entre la source et le détecteur (voir Eq.1.1).

0.1.2 Hybrid BCI

Lorsqu'on combine deux systèmes, avec au moins l'un d'eux comprenant une ICO, une ICO hybride (ICO_h) est alors composée. Selon [101], une de les conditions à remplir pour qu'une ICO soit considérée comme ICO_h est que le dispositif doit recevoir au moins un signal détecté par le cerveau (électrique ou hémodynamique). Par exemple, pour les dispositifs d'ECG, d'EMG ou de mesure de fixation du regard, etc. le signal provenant du cerveau, que l'utilisateur peut moduler, doit apporter une contribution à la ICO. Cette opération volontaire peut être effectuée par un changement du potentiel électrique, des champs magnétiques ou du flot hémodynamique. Une partie des conditions qui sont décrites dans les travaux de [101] sont appliquées aux ICO basées sur l'imagerie motrice et la désynchronisation liée à l'évènement. En fait, on n'a pas véritablement besoin de satisfaire toutes ces conditions en pratique, car l'analyse subséquente vise justement à découvrir de nouvelles mesures pour l'informatique affective. Une ICO_h affective pourrait être formée à partir d'un algorithme qui estime l'état émotionnel en se basant sur de nombreux signaux, combiné avec un autre algorithme qui estime l'intention de l'utilisateur. Par ailleurs, les ICO multimodales affectives peuvent en fait augmenter l'efficacité de la détection d'une émotion lors d'un état affectif de base (c'est-à-dire, mesurée en termes de valence et d'activation physiologique).

0.1.3 Objectifs de recherche et contribution

Au fil des pages suivantes, nous étudions la performance de mesures objectives standards pour la comparaison de la performance d'individus avec des scores d'intelligibilité subjectifs obtenus auprès des utilisateurs qui formulent des évaluations de contenu divers. Nous proposons également une nouvelle mesure en affinant une métrique précédemment proposée utilisée dans l'étude de la maladie d'Alzheimer, sur la base du couplage croisé fréquence amplitude-amplitude entre les sous-bandes de l'EEG [37]. Nous démontrons que les paramètres étudiés peuvent atteindre une performance fiable avec le recours à la base de données DEAP qui est accessible au public. Ces traits caractéristiques surpassent les mesures de référence (puissance spectrale) et l'utilisation d'un système multimodal RCG-EEG démontre une amplitude de la fonctionnalité de couplage d'environ 20%. En outre, la fusion de classifieurs au niveau de la décision a mené à de meilleurs résultats que la fusion au niveau des traits caractéristiques, avec le classifieur Amplitude Modulation Features (AMF) menant à plus de 67% de décisions correctes dans les dimensions de l'activation physiologique et de dominance. Ces résultats suggèrent l'importance des caractéristiques proposées pour la reconnaissance de l'état affec-

tif et signalent l'importance de traits caractéristiques de l'EEG comme la modulation de l'amplitude pour le marquage affectif de vidéoclips et de contenu divers.

Une partie du travail décrit dans ce mémoire a préalablement été publié ou est actuellement à l'étude pour publication pour des actes de conférences et des journaux scientifiques. Dans [17], nous avons évalué la performance des métriques objectives extraites avec le couplage croisé en fréquence entre des paires de canaux d'EEG. Cette expérimentation fut répétée pour chaque paire d'électrodes; un article décrivant l'utilisation du couplage croisé en fréquence pour les signaux d'EEG et signaux physiologiques a été soumis récemment aux *IEEE Transactions on Affective Computing*. En outre, dans les travaux présentés lors de la conférence SPIR 2014, l'effet de la SPIR dans une ICOa a été testé, en étudiant les avantages de la multimodalité ICO pour le domaine de l'informatique affective.

0.1.4 Organisation du document

Ce travail est organisé comme suit. Dans cette section, nous avons présenté le problème de l'amélioration de le ICO pour l'informatique affective et décrit les objectifs de travail et les contributions. Dans le Chapitre 2, nous discutons de l'état des études des émotions. Différents modèles d'émotions sont présentés ainsi que la méthode d'auto-évaluation utilisée dans le présent ouvrage. Dans le Chapitre 3, nous présentons améliorations aux mesures aICO qui visent à la mise au point d'un dispositif automatisé affective. Nous comparons ensuite les performances obtenues avec des repères couramment utilisés dans aICO et de discuter des avantages et des inconvénients. Nous concluons ensuite ce document de thèse avec un sommaire des résultats (Chapitre 4) et par l'ajout de quelques considérations finales au Chapitre 5.

0.2 Échelles subjectives d'émotions

Les méthodes affectives d'évaluation (MAE) sont devenues un domaine d'intérêt pour de nombreux chercheurs en tant que moyen d'étudier l'engagement et de l'expérience d'un sujet émotionnel lors de la réception de différents stimuli. Ceux-ci ont dû affronter deux principaux défis: la précision de la terminologie à utiliser et la reconnaissance des effets du cadre opératoire lors du déclenchement d'une émotion. Tout d'abord, les catégories et les définitions des différents états émotionnels et primitifs proviennent d'une gamme de théories disparates; les chercheurs qui œuvrent dans ce domaine ont questionné la bonne façon de déterminer les émotions et leurs codes, de les communiquer et de les décrire; ce secteur de recherche fut très actif au cours de la dernière décennie. Le deuxième défi concerne le contexte dans lequel les émotions sont provoquées, en particulier l'environnement (réel ou stimulé), car il joue un rôle crucial sur l'expérience émotionnelle [36, 103].

En tant que tel, plusieurs efforts ont été consacré à clarifier les termes tels que ‘affect’ et ‘émotion’. De nombreuses définitions ont été élaborées, mais comme l’ont souligné certains chercheurs [35, 107], il y a un chevauchement complet de ces terminologies entre elles; il n’y pas de distinction claire qui puisse être faite à ce moment. Ce fut seulement en 2006 que la distinction entre les deux termes a été précisée; on a défini ‘affect’ comme le terme plus large, relié aux émotions, humeurs et préférences.

L’autre aspect important qui doit être pris en considération lorsque des chercheurs évaluent des états affectifs est la sensibilité personnelle du sujet [36]. La sélection des participants à partir d’un même groupe d’utilisateurs cibles ne veut pas dire qu’ils réagiront tous de la même façon; leurs réactions peuvent être comparées à des données obtenues à partir d’un MAE [95]. En effet, les expériences provenant de l’enfance ou de l’âge adulte, ainsi que les facteurs culturels sont la base de la formation de la personnalité de chaque être humain. En outre, les mêmes émotions, détectées lors de journées différentes, peuvent être représentées selon une gamme variable de traits caractéristiques physiologiques [84]. Pour surmonter ce problème, une normalisation des traits caractéristiques a été testée [59, 105]; cela a abouti à une amélioration considérable pour l’appréciation correcte de ces variations. Cependant, l’inconvénient qui en résulte provient du fait que la mise en œuvre de ces protocoles n’est pas facile dans une application en temps réel, en raison de la nécessité de recourir à des séances post-opératoires.

0.2.1 Émotions

Afin d’examiner les émotions, deux catégorisations différentes ont été déterminées: les émotions discrètes [107] et la dimension des émotions [118].

- Émotions discrètes

La théorie de l’émotion discrète prétend qu’il existe un ensemble d’émotions de base qui sont biologiquement déterminées, c’est-à-dire des réponses émotionnelles dont l’expression et la reconnaissance sont les mêmes pour tous les individus, invariablement des limites propres aux différences ethniques et à la culture [107]. Une liste de huit émotions de base est développée dans certaines études de référence [57]: l’intérêt, la joie, la surprise, la détresse, le dégoût, la colère, la honte et la peur. Des recherches [133] déclarent que l’expérience émotionnelle en général est en fait constituée d’une combinaison complexe de ces émotions de base. En outre, la présence d’un seul d’entre eux est rare et difficile à évaluer. Diverses catégorisations discrètes d’émotions ont été proposées en somme. Par exemple, d’une part, les six émotions de base introduites par [31] telles que la colère, le dégoût, la peur, le bonheur, la tristesse et la surprise; d’autre part, l’arborescence des émotions avancée par d’autres chercheurs[97] où le premier niveau recouvre six émotions primaires (l’amour, la joie, la surprise, la colère, la tristesse, la

peur) et où le deuxième niveau recouvre les émotions secondaires, dérivées des primaires (sans être toutefois combinées) (Fig. 2.1).

- Dimension des émotions

La catégorisation des dimensions d'émotions propose une alternative à la théorie de l'émotion discrète puisque la structure 'réelle' des émotions n'est pas parfaitement ajustée à une quelconque configuration qui serait suffisamment rigide. Cette méthode propose une conceptualisation des émotions humaines qui tient en compte dans leur définition la situation d'un sujet dans un espace tridimensionnel, c'est-à-dire un sujet qui ressent une émotion subtile alternant entre deux ou trois de ces soi-disantes émotions 'primitives'. Les modèles tridimensionnels pour catégoriser les émotions, contrairement à la théorie des émotions de base, suggèrent qu'un système neurophysiologique interconnecté et commun à tous est responsable de tous les états affectifs en général. L'approche de la dimension originale comprend l'activation physiologique (intensité de l'émotion, exprimé en calme-activation), la valence (l'agrément, l'expérience positive ou négative) et la dominance (contrôle et de la nature dominante de l'émotion) [118]. Plusieurs modèles pour représenter les dimensions de ces émotions ont été proposés, tels que la roue de l'émotion de Plutchik [106] par exemple, ou l'échelle de valence-activation [113]. Toutefois, une limitation de ce modèle consiste dans la possibilité de l'éveil du sujet et dans la détermination des échelles de valence, est un inconvénient qui empêche une caractérisation plus détaillée des émotions. Pour répondre à cette limitation, le 'circomplexe des affects' de Russel est un diagramme qui distribue les émotions discrètes le long de deux axes selon les variables de la valence et de l'activation. De plus, une organisation dans le cercle a été introduite par d'autres chercheurs [106] qui se concentrent sur le diagramme d'une roue des émotions (Fig. 2.2).

En outre, un autre modèle fut conçu [113] lequel est largement utilisé pour déterminer la structure en trois dimensions de l'émotion basée sur les objets, l'expérience et la situation. Ce diagramme est composé de 18 paires d'adjectifs bipolaires [13] qui sont notés avec une échelle de 9 points. Ensuite, la méthode d'analyse des évaluations des sujets de contenu génère alors des scores interprétables sur trois dimensions respectives : la valence, l'activation physiologique et la dominance. Chaque émotion peut être classée dans un plan à l'activation à l'horizontale et de valence sur les axes verticaux (Fig. 2.3). L'activation peut varier d'un état d'indifférence à un état plus actif, dominé par des sentiments comme l'activation et l'action. La valence consiste en la douceur, le sentiment de tristesse, le stress ou la joie et extatique. Une troisième dimension qui peut être ajoutée est la dominance. Ce facteur varie entre un état de soumission à la relance (défense et faible) et le contrôle complet, le sentiment de puissance.

Le modèle utilisé dans le présent ouvrage est donc le 'circomplexe de l'affect' développé par Russel. Il a déjà été utilisé dans les études de comportement des consommateurs où un modèle abrégé a été utilisé, de même que pour des études organisationnelles où les émotions produites suite à l'exposition à un produit commercial sont mesurées de manière spécifique. En outre, dans d'autres cas [8], il a été utilisé dans la construction de personnages d'animation qui affichent des émotions. Pour l'auto-évaluation le long de ces échelles, nous utilisons les mannequins d'auto-évaluation (MAE) comme détaillé ci-dessous.

0.2.2 MAE: mannequins d'auto-évaluation

Alors que le modèle valence-activation physiologique-dominance gagnait du terrain dans différents domaines de recherche, notamment ceux de la publicité [47] de la formulation de stratégies, et de l'imitation à la copie de leur développement [143], il demeurait un obstacle majeur: il fallait rendre ce modèle plus fonctionnel, applicable et utile. Une amélioration était requise: il fallait à parvenir à l'obtention d'un traitement efficace d'une réponse rapide à un stimulus donné afin d'être en mesure de refléter avec précision une gamme complète de sentiments, en plus de le faire sans avoir recours à des stimuli pouvant induire des interférences linguistiques, c'est-à-dire en utilisant uniquement des objets visuels pour l'expérimentation. L'auto-évaluation du mannequin (MAE) [13] représente une solution valable à ces objectifs. Initialement conçu comme un programme informatique interactif, cette méthode a ensuite été appliquée à l'illustration sur papier, tel que représenté sur la Fig. 2.4. Chacune des trois grandes dimensions affectives est généralement représentée par cinq niveaux différents, illustrés par des pictogrammes [120]. La dimension de la valence est représentée par un ensemble de chiffres allant d'un sourire chez l'un au froncement des sourcils chez l'autre. Une gamme qui commence par l'activation à un chiffre décrit les dimensions de l'activation. Enfin, la dimension de dominance est représentée par le même chiffre avec différentes tailles, compte tenu de la plus grande qui correspond à un contrôle complet de la situation. Puisque les pictogrammes sont associés à une échelle de 9 points, les utilisateurs sont autorisés à choisir l'une des images ou à sélectionner entre deux images.

La méthode MAE peut être considérée comme efficace pour l'évaluation directe et rapide des trois principales dimensions affectives généralisées associées à la réaction émotionnelle d'une personne à un événement. Certains soutiennent que ce modèle est capable de comprendre à des expériences émotionnelles avant qu'elles ne soient exprimées, signalées lors du processus cognitifs de l'information traitée dans la partie linguistique du cerveau [53]. Mais surtout, cette méthode simple et non verbale peut être utilisée avec une variété de populations soumises à des limites d'expérimentation. Par exemple: les non-anglophones [13, 29], les enfants [4] et les personnes atteintes de troubles ou syndromes cliniques [134]. De plus, une consigne du protocole spécifie que les participants ont moins de 15 secondes pour rem-

plir le formulaire MAE, ce qui prévient donc le sentiment de fatigue résultant lors d'études beaucoup plus longues.

0.3 Traits caractéristiques

Traits caractéristiques d'EEG traditionnels

Typiquement, les traits caractéristiques spectraux, tels que la puissance dans des sous-bandes de fréquence, sont utilisés pour mesurer les états émotionnels induits par le visionnement de vidéoclips musicaux, d'images, et/ou de clips vidéo (par exemple, [66, 126]), de même que la charge mentale et le stress (par exemple, [50, 67]). En outre, une asymétrie inter-hémisphérique en puissance a été identifiée dans la littérature comme étant liée à ce type de phénomène [25, 60], en particulier dans les régions frontales du cerveau [19]. Dans ce travail, les composantes spectrales (*Spectral Features*, SF) sont utilisées comme point de comparaison pour les nouveaux traits caractéristiques d'EEG proposés. Un total de 128 traits caractéristiques spectraux ($32 \text{ électrodes} \times 4 \text{ sous-bandes}$) et 56 indices d'asymétrie ($14 \text{ paires inter-hémisphériques} \times 4 \text{ sous-bandes}$) ont été calculés à partir des paires d'électrodes suivantes: Fp1-FP2, AF3-AF4, F7-F8, F3-F4, FC5-FC6, FC1-FC2, T7-T8, C3-C4, CP5-CP6, CP1-CP2, P7-P8, P3-P4, PO3-PO4 et O1-O2 (voir Fig. 3.1). Au total, 184 traits caractéristiques SF sont utilisés comme référence.

Traits caractéristiques de modulation d'amplitude, AMF

Trois traits caractéristiques de couplage amplitude-amplitude (*Amplitude Modulation Features*, AMF) sont extraits: l'énergie de la modulation d'amplitude (AME), les interactions en modulation d'amplitude (AMI), et la cohérence de la modulation d'amplitude (AMC), tel que présenté à la Fig. 3.1. Pour calculer ces trois ensembles de traits caractéristiques, les signaux d'EEG sont tout d'abord décomposés en quatre sous-bandes (thêta, alpha, bêta et gamma). L'enveloppe temporelle est ensuite extraite pour chacune des quatre séries temporelles obtenues grâce à la transformée de Hilbert [75]. Afin de mesurer le couplage croisé en fréquence amplitude-amplitude, une seconde décomposition des enveloppes d'amplitude EEG est effectuée dans les mêmes quatre sous-bandes que dans la première décomposition en utilisant un banc de filtres du second ordre. Pour faire la distinction entre les fréquences de modulation et de sous-bandes, les premières sont dénommées m-thêta (4-8 Hz), m-alpha (8-12 Hz), m-bêta (12-30 Hz) et de m-gamma (30-45 Hz). En utilisant la transformée de Hilbert pour extraire l'enveloppe des signaux, les types d'interactions entre les fréquences sont limitées par le théorème de Bedrosian, qui stipule que les signaux d'enveloppe ne peuvent contenir que des fréquences (c'est-à-dire, fréquences modulées) jusqu'à la fréquence maximale de son signal d'origine [11, 125]). En tant que tel, seuls les dix patrons de couplage croisé en fréquence présentés à la figure Fig. 3.1 sont possibles (par électrode), à savoir : theta_m-theta,

alpha_m-theta, alpha_m-alpha, beta_m-theta, beta_m-alpha, beta_m-beta, gamma_m-theta, gamma_m-alpha, gamma_m-beta, and gamma_m-gamma. À partir de ces modèles, les trois ensembles de traits caractéristiques sont calculés, tel qu'indiqué ci-dessous :

Énergie de la modulation d'amplitude (*Amplitude Modulation Energy*, AME) Parmi les dix patrons possibles par électrode, deux mesures de l'énergie sont calculées. La première quantifie le rapport de l'énergie dans une paire de modulation de fréquence donnée sur l'énergie totale dans toutes les paires de sous-bandes possibles. La seconde mesure le logarithme du rapport de l'énergie de modulation au cours d'un vidéoclip musical de 60 secondes à l'énergie de modulation au cours d'une période de repos de 3 secondes.

Interactions en modulation d'amplitude (*Amplitude Modulation Interaction*, AMI) Afin d'intégrer la synchronie de modulation d'amplitude inter-électrodes, les interactions en modulation d'amplitude [17] sont également calculées. Contrairement au travail décrit dans [17], où les interactions sont uniquement calculées par paires interhémisphériques symétriques, ici, nous mesurons les interactions entre les 496 combinaisons possibles de paires d'électrodes pour chacun des dix patrons de couplage inter-fréquence. De plus, une seconde mesure du rapport logarithmique entre les 60 secondes d'un vidéoclip musical et les 3 secondes de repos a été obtenue.

Cohérence de la modulation d'amplitude (*Amplitude Modulation Coherence*, AMC) Tandis que l'AMI capture des interactions non linéaires entre les patrons de couplage amplitude-amplitude inter-électrodes, le coefficient de corrélation de Pearson entre les modèles peut également être utilisée pour quantifier la cohérence, ou des interactions entre les patrons linéaires.

Couplage phase-amplitude PANS-CNS

Les signaux électrophysiologiques émanent de systèmes dynamiques interagissant les uns avec les autres à différentes fréquences. Le couplage phase-amplitude (*Phase-Amplitude Coupling*, PAC) est un des types d'interaction possible et consiste généralement en une modulation de l'amplitude d'oscillateurs à haute fréquence par la phase d'oscillateurs à basse fréquence [Samiee *et al.*]. En règle générale, de telles mesures de couplage phase-amplitude sont calculées à partir des signaux d'EEG seuls [121]. Toutefois, il a récemment été démontré que la phase de l'activité électrodermale couplée à l'amplitude de l'EEG est corrélée aux états affectifs, en particulier pour ce qui est des stimuli de forte valence et très agréables ou désagréables [69, 70]. Dans ce travail, nous testons trois mesures de couplage entre la phase du RCG et l'amplitude de l'EEG.

Couplage enveloppe-sur-signal (*Envelope-to-Signal Coupling*, ESC) Le trait caractéristique de couplage le plus simple peut être calculé à l'aide du coefficient de corréla-

tion de Pearson entre l'enveloppe de l'amplitude du signal d'EEG et le signal d'activité électrodermale.

Cohérence croisée en fréquence (*Cross-Frequency Coherence*, CFC) La cohérence croisée en fréquence évalue le carré de la cohérence entre les signaux du RCG filtrés entre 0-1 Hz et l'enveloppe de l'EEG filtrée entre 4-45 Hz, tel que dans [96].

Indice de modulation (ModI) La dernière mesure de couplage PANS-CNS testé dans ce travail est l'indice de modulation (ModI), qui, tel que récemment démontré [130], permet de caractériser avec précision l'intensité d'un couplage, en particulier pour la reconnaissance des émotions [70].

0.3.1 Intégration multimodale : EEG-SPIR

Tel que mentionné à la Section 1.2, les ICO affectives hybrides peuvent améliorer la reconnaissance des états affectifs, comme une modalité peut pallier les lacunes de l'autre. Ici, nous étudions les avantages d'une ICOa hybride utilisant deux stratégies de fusion multimodale, à savoir la fusion de traits caractéristiques et la fusion de décisions. La fusion de traits caractéristiques [42], comme son nom l'indique, consiste en la concaténation de traits caractéristiques provenant de différentes modalités dans un même vecteur utilisé pour la classification. La malédiction de la dimensionnalité devient vite un problème avec les stratégies de fusion de traits caractéristiques et ainsi ceux-ci font généralement l'objet d'une sélection préalable. La fusion de traits caractéristiques a été largement utilisée pour la détection des états affectifs, puisque le corps exprime ces derniers de différentes manières [66]. La seconde approche repose généralement sur plusieurs classifieurs, chacun entraîné sur une modalité spécifique. La classification finale, cependant, repose sur une combinaison de la décision de chaque classifieur individuel. Typiquement, un vote majoritaire ou une combinaison pondérée est utilisé comme stratégie de fusion de décisions. Ici, les deux stratégies, soit la fusion de traits caractéristiques et la fusion de décisions, sont étudiées pour les configurations suivantes: EEG-SPIR, EEG-PANS, EEG-SPIR-PANS. Afin d'étudier les gains obtenus avec ces stratégies de fusion, seuls les traits caractéristiques traditionnels d'EEG sont utilisés.

0.3.2 Tâches et bases de données

Bases de données

La base de données DEAP, publiquement disponible sur Internet [66], a été utilisée pour évaluer les gains en performance obtenus avec les traits caractéristiques proposées à la Section 3.1.3. Trente-deux sujets sains (nombre égal d'hommes et de femmes, âge moyen de 26,9 ans) ont visionné 40 vidéoclips musicaux, tandis que leurs signaux neurophysiologiques étaient enregistrés. Ces quarante vidéos ont été soigneusement sélectionnés parmi environ 200

vidéos, afin de ne conserver que ceux suscitant les évaluations les plus élevées dans chacun des quatre quadrants du plan valence-activation physiologique (ou *arousal*) [113]. Les participants devaient auto-évaluer leur valence, activation physiologique et dominance perçues (trois primitives émotionnelles), ainsi que d'autres évaluations subjectives telles que l'appréciation (*liking*) et la familiarité, pour chacun des 40 vidéoclips. Les trois primitives émotionnelles ont été mesurées à l'aide d'une échelle d'auto-évaluation "mannequin" à 9 points [13], tel qu'illustré à Figure 2.4. L'échelle d'appréciation a été introduite afin d'évaluer les goûts de l'utilisateur, et non ses sentiments, vis-à-vis le vidéoclip. L'échelle à 9 points avec des symboles de pouces vers le bas ou le haut a ainsi été adoptée. Enfin, le niveau de familiarité a été mesuré suivant une échelle de 5 points allant de "je n'ai jamais entendu parler de ce vidéo avant" à "j'écoute ce vidéo régulièrement".

Plusieurs signaux neurophysiologiques ont été enregistrés pendant l'expérience : 32 canaux d'EEG (BIOSEMI Active II, suivant un placement d'électrodes standard 10-20), la température de la peau, la réponse électrodermale, la respiration, et le photopléthysmogramme. Les signaux bruts ont été enregistrés à une fréquence d'échantillonnage de 512 Hz, puis sous-échantillonnées à 128 Hz. Les signaux d'EEG ont ensuite été filtré par un filtre passe-bande entre 4 et 45 Hz, pré-traités à l'aide d'une analyse en composantes principales pour éliminer les artéfacts oculaires, puis moyennés en utilisant la référence commune. Le lecteur intéressé est invité à consulter [66] pour plus de détails sur cette base de données. La tâche (Tâche 1) visée par l'utilisation de cet ensemble de données est de détecter le niveau d'activation physiologique, de valence, de dominance et d'appréciation en utilisant les traits caractéristiques de couplage d'EEG et de PANS-CNS nouvellement proposés, individuellement ou au sein d'une configuration ICOa hybride.

Pour valider les gains obtenus avec une ICOa hybride, la base de données eNTERFACE 06, elle aussi disponible sur Internet, a été utilisée [117]. Cette base de données est composée d'enregistrements de SPIR, d'EEG, d'activité électrodermale, de signaux de respiration et de photopléthysmographie recueillis auprès de 5 sujets masculins et droitiers âgés de 22 à 38 ans. Plusieurs canaux d'EEG (AF7, AF8, AFZ, FP1, FP2 Fpz, F7 et F6) ont été retirés d'un ensemble de 64 électrodes (BIOSEMI ActiveII, suivant le système 10-20 et avec une fréquence d'échantillonnage de 1024 Hz) en raison de l'occlusion de la zone frontale due à l'installation des capteurs de SPIR. Les participants ont observé 450 images tirées du IAPS (système international d'images affectives) [73], choisies afin de représenter 3 classes différentes (images positives, négatives et neutres). Les sujets ont ensuite évalué leurs niveaux perçus de valence et d'activation physiologique en utilisant une échelle à 5 points. Les images ont été visualisées dans un ordre aléatoire, suivant le schéma illustré à la Figure 3.2.

L'expérience est composée de trois sessions distinctes. Une séance comporte 30 répétitions contenant 150 images chacune, pour un total de 450 images. Les répétitions ont été ordonnées de manière à ne pas avoir la même émotion provoquée pas deux images consécutives. Une

répétition est composée de la visualisation de blocs de 5 images (12,5 secondes) entrecoupés par des écrans noirs de 10 secondes pour permettre aux signaux de SPIR de retourner à leur niveau de base. Plus d'informations au sujet de la base de données eNTERFACE 06 sont disponibles dans [117].

Sélection de traits caractéristiques

Comme mentionné ci-dessus, un grand nombre de traits caractéristiques ont été extraits. Plus précisément, un total de 184 traits caractéristiques SF, 20480 AMF et 1408 PAC ont été extraits. Dans le cas d'une tâche de classification, la dimension élevée de ces données peut conduire à du surapprentissage, particulièrement lorsque la fusion de traits caractéristiques est utilisée. Dans de tels cas, des algorithmes de sélection ou de classement de traits caractéristiques sont généralement utilisés. Récemment, plusieurs algorithmes de sélection de traits caractéristiques ont été comparés sur une tâche de reconnaissance des émotions [60]. Dans nos expériences, l'algorithme mRMR (minimum Redundancy Maximum Relevance) [99] a été utilisé, car, tel que démontré dans [131], il peut mener à une amélioration des performances lorsqu'utilisé avec une machine à vecteurs de support. L'algorithme mRMR est basé sur la mesure de l'information mutuelle et tente d'optimiser deux critères simultanément : le critère du maximum de pertinence (i.e., maximisation de l'information mutuelle moyenne entre chaque trait caractéristique et le vecteur cible) et le critère de minimum de redondance (i.e., minimisation de l'information mutuelle moyenne entre deux traits caractéristiques). Dans ce travail, l'algorithme identifie des traits caractéristiques quasi-optimaux par l'entremise d'une sélection ascendante basée sur les traits caractéristiques maximisant les critères combinés max-min. Le quart des données disponibles a été mis de côté pour le classement des traits caractéristiques. Les trois-quarts restants ont été utilisés pour l'entraînement et le test des classifieurs par validation croisée, comme décrit ci-après.

Classification

Les machines à vecteurs de support (SVM) ont été largement utilisées pour la reconnaissance des états affectifs (par exemple, [131]). Compte tenu de leur utilisation répandue, cette approche ne fait pas l'objet d'une description ici et le lecteur intéressé est appelé à consulter [119] et ses références pour plus d'informations. Des SVM sont entraînés sur quatre problèmes de classification binaire différents, à savoir : la détection de valences faibles/élevées, d'activations physiologiques faibles/élevées, de dominances faibles/élevées et appréciations faibles/élevées. Comme les évaluations subjectives sont recueillies sur une échelle de 9 points, nous définissons un niveau élevé comme correspondant à des valeurs supérieures ou égales à 5; des valeurs plus petites que 5 correspondent ainsi à un niveau faible. Tel que mentionné précédemment, 25% des données disponibles est mis de côté pour le classement des traits caractéristiques. Les 75% restants sont utilisés suivant une validation croisée avec exclusion d'un

échantillon (*Leave-One-Sample-Out*). Dans nos expériences, un noyau gaussien a été utilisé avec la librairie Python scikit-learn [98]. Puisque nous sommes intéressés à évaluer l'efficacité des traits caractéristiques proposés, et non de différents types de classifieurs, nous utilisons les paramètres par défaut pour l'entraînement des SVM (à savoir, $\lambda = 1$ and $\gamma_{RBF} = 0.01$).

Pour la tâche 2, une variante des SVM a été utilisée, à savoir les machines à vecteur de pertinence (Relevance Vector Machine, RVM) [41, 129, 139]. Grâce à un traitement bayésien, les RVM permettent d'obtenir des décisions 'souples', c'est-à-dire probabilistes (par opposition à des décisions dures comme dans le cas des SVM). Comme mentionné ci-haut, deux stratégies de fusion sont explorées afin d'améliorer les performances de classification, à savoir la fusion de traits caractéristiques et la fusion de décisions. Avec la base de données DEAP, dans le cadre de la fusion de traits caractéristiques, nous explorons la combinaison des trois ensembles de traits (SF, PAC, et AMF) suivant la sélection obtenue par l'algorithme mRMR. D'autre part, pour la fusion de décisions, les décisions des trois classifieurs formés avec les traits SF, PAC et AMF ont été fusionnés en utilisant un système de vote majoritaire, avec des poids égaux. Afin de déterminer si les performances atteintes sont significatives, un test t indépendant ($p < 0,05$) est utilisé pour comparer les résultats à ceux d'un classifieur aléatoire, tel que suggéré dans [66]. Avec la base de données eNTERFACE, la fusion de décisions a été testée lors de la combinaison des trois modalités (EEG, SPIR et signaux physiologiques).

0.4 Résultats expérimentaux et discussion

Cette section décrit les résultats obtenus par les tâches 1 et 2 décrites précédemment et fournit une discussion approfondie sur les résultats obtenus. Les résultats présentés ici ont été présentés dans [16–18].

0.4.1 Tâche 1: Résultats expérimentaux

Tables 4.1 et 4.2 montrent les vingt meilleurs traits caractéristiques sélectionnées pour les dimensions de l'activation/valence et dominance/appréciation, respectivement (les noms de fonctions figurant dans les tableaux sont explicites). Les traits "ratio" correspondent aux rapports logarithmiques entre les périodes de vidéo et de base (voir la section 3.1.2). Dans la catégorie SF, les traits caractéristiques 'ai' correspondent à l'indice d'asymétrie entre les canaux indiqués. En raison de contraintes d'espace, les traits caractéristiques IIAM (*IAMI*) ne sont pas inclus dans le tableau de classement.

Les figures 4.1(a)-(d) représentent l'exactitude des classifieurs SVM pour l'activation physiologique, la valence, la dominance, et l'appréciation, respectivement, en fonction du nombre de composantes utilisées. Ces exactitudes sont présentées pour les classifieurs formés avec seulement les traits IIAM proposés (Interaction inter-hémisphérique d'amplitude modulée), la période de référence, ainsi que l'ensemble des traits combinés. Comme on peut le voir,

les traits caractéristiques proposés surpassent ceux de référence pour les quatre dimensions affectives, mais avec des gains plus élevés observés dans les dimensions de la valence et de l’appréciation. Fait intéressant, ce sont les dimensions les plus difficiles à estimer avec des caractéristiques EEG classiques dans [66].

La figure 4.2 (a)-(d), à son tour, rapporte l’exactitude de l’activation, de la valence, de la dominance et de l’appréciation des classifieurs SVM, respectivement, en fonction du nombre de traits caractéristiques utilisés pour les trois ensembles (AMF, PAC, SF). Au sein de ce scénario, les caractéristiques ont été ajoutées une par une sur la base du classement fourni par l’algorithme mRMR. Un nombre maximum de caractéristiques égales à 184 a été utilisé pour coïncider avec le numéro de la référence SF fixée. Comme on peut le voir, l’ensemble AMF surpasse les deux autres ensembles à travers toutes les dimensions affectives.

Le tableau 4.3 indique les résultats les plus élevés de précision obtenus avec des ensembles de classifieurs, ainsi que le nombre de traits nécessaires pour atteindre de tels résultats. Les résultats suivis d’un astérisque indiquent des valeurs nettement plus élevées ($p < 0.05$) que celles obtenues avec un classifieur de vote aléatoire, comme dans [66]. Le nombre entre parenthèses montre l’amélioration relative, en pourcentage, par rapport à la référence SF. Comme on peut le voir, la modulation d’amplitude des ensembles de traits caractéristiques proposés permet d’atteindre des résultats nettement supérieurs à la chance pour les quatre dimensions. Pour les dimensions de valence et de dominance, les caractéristiques de l’AMF proposées provoquent des améliorations de 9% et 14%, respectivement.

La dernière colonne du tableau 4.3 montre les résultats obtenus avec la fusion de traits caractéristiques. Comme on peut le voir, la fusion n’a pas abouti à des performances supérieures à celles obtenues avec l’ensemble AMF proposé. Le tableau 4.4, à son tour, présente les performances obtenues avec deux systèmes de fusion de niveau de décision: AMI/AMC/AME et l’AMF/PAC/SF. Tandis que le niveau de décision de fusion de classifieurs AMF/PAC/SF n’a pas conduit à des gains, la fusion de décision des traits AMC/AME/AMI a amélioré la précision des classifieurs pour la dominance et l’appréciation. Dans l’ensemble, avec la fusion de décision des traits AMI/AMC/AME, les gains de 22%, 9%, 14% et 2% ont été observés pour les dimensions de l’activation, la valence, la dominance et l’appréciation, respectivement.

0.4.2 Tâche 1: Discussion

Une analyse approfondie des cent meilleures traits caractéristiques sélectionnés (tableaux 4.1 a été menée afin de comprendre les régions les plus pertinentes du cerveau et les bandes de fréquences les plus souvent impliquées. Dans le jeu de traits SF, les indices d’asymétrie prévalent sur les caractéristiques spectrales de puissance. En fait, pour chacune des quatre dimensions affectives, les 54 fonctionnalités AI étaient présentes dans les 60 premiers caractéristiques sélectionnés. Par ailleurs, l’analyse des 20 premiers traits révèle une prédominance de la région frontale avec une activité cohérente dans la bande θ . Une expression per-

tinente de l'émotion par les signaux du cerveau à l'aide de l'asymétrie a été décrite dans [24], premier modèle qui concerne lobes frontaux du cerveau des émotions. D'autres études [19, 22, 44] ont également confirmé le concept d'asymétrie précitée.

Des travaux antérieurs [70] sur les PAC, à son tour, ont montré que le couplage entre EEG et RCG est pertinent dans la classification de l'émotion. Dans notre étude, nous avons constaté une prédominance des traits ESC sur les deux autres (CFC et ModI). En effet, les traits ESC sont classées dans les 50 premiers longs métrages pour chaque dimension émotionnelle. En ne considérant que les 20 traits, la région la plus impliquée du cerveau était la région frontale, tel qu'observé avec les caractéristiques d'asymétrie. En outre, en ce qui concerne la CFC, la bande de fréquence qui a abouti à des caractéristiques les plus importantes était thêta.

En ce qui concerne l'ensemble de l'AMF proposée, deux scénarios différents ont surgi. Pour l'activation physiologique et la valence, le trait caractéristique prédominant était l'AMC, correspondant à 58% des traits sélectionnés. Pour les dimensions de l'appréciation et de dominance, à son tour, les traits AMI correspondent à 57% des traits sélectionnés. Les traits AMC sont des mesures de couplage de modulation d'amplitude linéaires entre les différentes régions du cerveau, alors que AMI peut mesurer les effets potentiellement non-linéaires. Ces résultats suggèrent différentes relations fondées variablement sur les émotions primitives. En outre, environ 80% des traits caractéristiques ont été sélectionnés à partir du rapport logarithmique entre la vidéo et les périodes de référence, suggérant ainsi l'importance de la fonction de normalisation.

Comme on le voit à partir des figures 4.1, les traits IIAM sont nettement supérieurs aux traits de référence pour l'ensemble des quatre tâches de classification. La performance des classifieurs, lesquels sont formés soit sur la proposition ou sur l'ensemble des fonctionnalités de base, en général ont eu tendance à se stabiliser autour de 60 traits caractéristiques; toutefois, la fonction proposée fut institué selon une précision beaucoup plus élevé. Cela est particulièrement vrai pour les dimensions de la valence et de l'appréciation, lesquels sont généralement les plus difficiles à classifier avec les traits caractéristiques d'EEG classiques [66]. En fait, d'après le tableau 4.3 pour la dimension de valence, environ un tiers des traits ont été nécessaires pour surpasser les traits de référence (53 vs 146 caractéristiques). Ces résultats mettent en évidence les avantages de la représentation compacte des caractéristiques proposées pour la caractérisation fiable de l'état affectif. Une analyse en profondeur des soixante meilleures traits IIAM a montré que près de la moitié provenait de motifs spectro-temporels alpha-bande. Pour les dimensions de la valence et de l'activation, le patron $A(\alpha, m - \theta)$ était le plus important. Des travaux antérieurs ont lié l'interaction alpha-thêta à la mémoire [15], modulée par la suite par les dimensions de la valence et de l'activation [63], suggérant ainsi que les références internes (c'est-à-dire les affects) ont peut-être joué un rôle clé lors de l'expérience. Les traits proposés semblent être en mesure de caractériser de tels effets reliés

à la mémoire, à la différence des caractéristiques classiques, corroborant ainsi leur complémentarité.

0.4.3 Tâche 2: Résultats expérimentaux et Discussion

Les Tableaux 4.5 et 4.6 montrent la précision et le score F1 obtenus soit avec une seule ou plusieurs modalités (en utilisant la SPIR). En particulier le tableau 4.5 indique les résultats avec les 20 premières composantes sélectionnées. Chaque fois que la SPIR a été ajoutée, on constate une augmentation de la précision. Le maximum a été atteint avec une amélioration de huit pourcent de la valeur initiale lorsque l'on combine les trois modalités ensemble (EEG, signaux physiologiques et SPIR) en utilisant une technique de fusion de décision avec vote majoritaire. La fusion de traits caractéristiques a également été testée. La meilleure amélioration (tableau 4.6) a été atteinte lors de l'ajout de signaux physiologiques et de la SPIR (7%).

La SPIR ajoute des informations importantes pour les tâches de détection affectives, ce qui démontre l'importance de la multimodalité [87]. En outre, il y a une forte contribution du nombre de traits caractéristiques de SPIR lorsque la fusion de traits caractéristiques est utilisée. En effet, lorsque qu'on combine l'EEG et la SPIR, cette dernière compte pour 20% des traits caractéristiques sélectionnés, alors que lorsque on combine les signaux physiologiques et la SPIR, elle compte pour 80%.

0.5 Conclusion

Dans ce travail, de nouveaux traits caractéristiques de couplage amplitude-amplitude en fréquence ont été proposés pour la reconnaissance de l'état affectif lors du visionnement de vidéoclips musicaux. Les résultats expérimentaux avec la base de données accessible au public DEAP ont dévoilé des traits caractéristiques dépassant la performance de traits de référence (puissance spectrale et couplage RCG-EEG) par 20%. La fusion de décision a amené à de meilleurs résultats que la fusion au niveau des traits caractéristiques, avec le classifieur AMF étant responsables de plus de 67% des décisions correctes concernant les dimensions de l'activation physiologique et de dominance. Ces résultats suggèrent l'importance des caractéristiques proposées pour la reconnaissance de l'état affectif et signalent l'importance de la modulation d'amplitude en EEG pour le marquage affectif de vidéoclips et de contenu.

En outre, un nouveau modèle multimodale a été exploré, en introduisant l'utilisation des informations hémodynamiques pour l'informatique affective au cours du visionnement des images. Les résultats expérimentaux avec la base de données eNTERFACE montrent que la SPIR améliore les résultats jusqu'à huit pourcent en précision. La performance suggère que l'utilisation d'informations hémodynamiques pour l'étude des images affectives est bénéfique.

Chapter 1

Introduction

Emotions are usually conceived as physiological and physical responses, are part of natural communication between humans, and are able to influence our intelligence, shape our thoughts and govern our interpersonal relationships [26, 81, 83].

To date, machines have not been required to have emotional skills and Human-Machine Interface (HMI) has typically involved humans providing information to computers via e.g. a keyboard and/or mouse. However, with the advances witnessed recently within the affective computing field [102], this paradigm is on the verge of shifting. Machines are slowly being required to have a basic level of intelligence in order to create a proper human-machine interaction [108], thus following the principles of human-human interaction. Ultimately, the field of affective computing has emerged with the goal of reducing the distance between the end user and the machine by enabling an adaptive interaction where the machine understands the user's emotions and cognitive states, thus adjusting accordingly. This allows instruments to better address human needs [91].

The birth of the affective computing field, combined with the technological advances witnessed in recent years (e.g., increase in computing power, lower cost of storage, improved battery life, etc.), has resulted in the resurgence of the so-called Brain-Computer Interface (BCI). Traditionally, BCIs have been used to help individuals with communication and mobility issues, such as those with locked-in syndrome, to communicate via brain-speller [78] or to drive powered wheelchairs [76]. With wearable technologies emerging, however, new consumers applications are on the rise, such as gaming [12], attention monitors [58], and mobile devices [132], to name a few.

By moving the BCI landscape from an ‘active’ interface, which collects neural signals and translates them into machine control signals, to a more ‘passive’ interface that collects implicit user mental states, the concept of Affective Brain-Computer Interface (aBCI) has appeared. Moreover, as wearable technologies and neuroimaging technologies become more popular and accessible, multimodal BCIs are on the rise. Such multimodal BCIs have been called Hybrid Brain-Computer Interface (hBCI)s, as they can combine multiple neuroimaging technologies, neuroimaging with physiological monitoring, or even neuroimaging with conventional human-machine interface technologies, such as the keyboard [6].

A recent literature review [6] has shown that hBCIs combining Electroencephalography (EEG) and fNIRS are burgeoning. The majority of existing studies, however, have focused on using hybrid multimodal BCIs as active interfaces, as defined above. In this thesis, we place focus on affective hBCIs. Since affective and hybrid BCIs play a key role in this work, the remainder of this chapter will focus on such topics. The chapter concludes with a detailed description of the thesis contributions.

1.1 Affective BCIs

Affective BCI is a natural development of BCIs and an attempt to use HMI outside the medical domain. Applications of aBCI were tested for entertainment, life-style and ergonomic [12] and it may provide a means of detecting spontaneous and natural reactions to emotion-evoking stimuli. Affective BCIs can be used for active communication of emotional states and moods and passive sensing of affect to inform machines about the affective states of their users, thus bringing ‘user state awareness’ to existing BCI systems. This result can be achieved either using neural or physiological data, since emotion is a neurophysiological response [66]. Emotional awareness may help create more user accommodating systems and develop more effective training paradigms. Typically, BCIs fall into two possible categories: active or passive. Active BCIs (as well as Reactive BCIs), as the name suggests, are systems which translate neural patterns into machine control signals. Passive BCIs, on the other hand, detect implicit information regarding the user state and intention, and enable situational interpretation and control in absence of overt emotional expression. As such, passive BCIs have commonly been referred to as affective BCIs and that is the terminology that will be used herein.

Within the BCI literature, EEG has been the dominant neuroimaging modality due to its well-studied features, affordability, portability and high temporal resolution [56]. Most of the brain’s cognitive functions are based on the coordinated interactions of large numbers of neurons that are distributed within and across different specialized brain areas. Synchronization on various levels of organization of brain tissue, from individual pairs of neurons to much larger scales, within one area of the brain or between different parts, is one of the most active topics in current neuroscience [9]. Specifically, normal cognitive operations require the transient integration of numerous functional

areas widely distributed over the brain and in constant interaction with each other. The electrochemical signals travel through the brain and skull, and they are recorded by electrodes applied directly to the scalp. The electrode cap is placed directly on the scalp and it usually consists of 32, 64, 128 or 256 electrodes. The slight changes of voltage in the axons of the neurons are called graded potentials. Each electrode records the sum of the graded potentials of the numerous neurons [56].

The use of EEG-based aBCIs mostly focused on the detection of features able to classify low and high workload conditions [50, 67]. These examples introduced one of the most common technique used with EEG, the Evoked Potential (EP), a combination of a phase-locked external stimulus and the study of synchronous neural activity over a region [85]. The EP is usually very low and hard to detect, but the repetition of the stimulus over several trials and the averaging of all the events in the time-domain illustrates the hidden EP. The type of stimuli (e.g. visual) and the positivity or negativity of the components (P300 potential) define different EPs. The components of the P300 are the peaks in voltage, negative or positive, of the transitory waveform. The number indicates the time period in milliseconds after the stimulus. It has been demonstrated [136] that the evoked waveforms always contain some small differences depending on the proposed stimuli and the main objective of BCI is to discriminate and categorize them.

A further technique that combines BCI with EEG involves changes in Sensorimotor Rhythm (SMR)s, that represent the change in the rythm of the EEG signal due to a movement, or only its imagination [85, 137]. Similar to the EP, SMRs makes use of spectral analysis in order to detect and map the changes, and also build a control system that uses spontaneous brain activity. Hence, main difference between the two techniques relies on the fact that with EP the study focuses on the change in electric potential belonging to a precise stimulus, whereas with SMRs there is no need

for external input and the study can be conducted only making use of imagination. The capability of shaping our thoughts using concentration has been widely tested. For instance,[90, 138] tried to understand the positive or negative answer of the subjects from their imagination. They were able to reach over 95% of accuracy.

Focusing on studies of aBCI, EEG has been widely used. Most of the published works make use of the spectral power features as the most relevant characteristic to discriminate emotions. For instance [100], proposed the use of Independent Component Analysis (ICA) to extract features and discriminate between pictures that are pleasant, unpleasant or neutral. Data from 8 subjects have been collected, using a selection of images from the International Affective Picture System (IAPS) as stimuli. The features used are the Event Related Potential (ERP). Using a one-way ANOVA, they demonstrated that the extracted ERPs were able to discriminate between active/neutral and pleasant/unpleasant images ($p < 0.05$) in the P7 and F8 electrodes.

In the work of [54], the authors try to measure affective state primitives, such as arousal and valence (more on this topic in chapter 2), in response to the IAPS via a Support Vector Machine (SVM) classifier. Ten participants were used for the collection of the EEG data. Each of them rated the images using the Self-Assessment Manikins (SAM) that are going to be described in details in the next chapter. Regarding the features, they extracted frequency band powers together with the cross-correlation between them, the peak frequencies in the alpha band and the Hjorth parameters [52], for a total of 114 features. The mRMR feature selection algorithm was used to select the top-40 features. The SVM classifier was able to discriminate the valence and arousal scales, respectively, with an accuracy of 71% and 81%, after the removal of ambiguous cases located near a neutral emotional state in their picture dataset.

Another common feature used to discriminate affective states is the so called asymmetry index between electrodes. For example, in [66] spectral powers from theta (4-8 Hz), slow alpha (8-10 Hz), alpha (8-12 Hz), beta (12-30 Hz), and gamma (30+ Hz) were computed, and the difference between the spectral power of all the symmetrical pairs of electrodes on the right and left hemisphere were used as features. A total of 216 features from 32 EEG electrodes were extracted from each of the 32 subjects that were watching music video clips. More details on the dataset can be found in Section 3.2.1. A naive Bayes classifier was used since it is simple and generalizable for unbalanced datasets, and accuracies of 62% and 58% were achieved to discriminate between high and low arousal and valence states, respectively. Together with feature fusion the authors also tried to perform decision fusion using different modalities (EEG, multimedia content analysis and physiological signals). The advantage was that the classification was conducted independently for each signal and it was performed an optimal weighting scheme to adjust the relative contribution of each modality. Using an optimal weighting with all the three modalities accuracies of 62% and 65% respectively for arousal and valence were reached.

Some studies use also invasive techniques. For instance, in the work of [Bartsch], it was made use of the ElectroCorticoGraphy (ECoG) invasive technique to discriminate emotions. In particular, the authors want to examine the cortical communication during emotional picture processing in epilepsy patients. The images were selected from the IAPS in order to discriminate between pleasantness and unpleasantness. The study made use of the Event Related Spectral Power (ERSP) features from each electrode and a granger causality analysis for the connectivity. Results demonstrated that the amplitude in the gamma band increases when observing emotional compared to neutral pictures in ventral cortical and deep areas (e.g. amygdala), whereas directional interactions were

observed among ventral visual and deep structures that appeared to use theta oscillations as a carrier frequency.

Apart from the relevant advantage of being non-invasive, with high temporal resolution and reasonable cost in the market, the major drawback of EEG is its poor spatial resolution, meaning that it is relatively ineffective at differentiating specific regions in the brain.

To overcome this limitation, other neuroimaging technologies have been explored, such as Functional Magnetic Resonance Imaging (fMRI), Magnetoencephalography (MEG), and fNIRS. MEG makes use of the neural magnetic fields in order to map brain activity. As an example, the work of [72] sets up an online BCI system trained with MEG data in order to discriminate an imaginary movement of the tongue against a left little finger. They used Autoregressive Model (AM) [48] to capture the spectral characteristics of the 150 channels of MEG. Feature selection is applied before training the SVM classifier. The performance reached a maximum of 92% of accuracy only using the 5% of the extracted features. fMRI, in turn, provides a measure of cerebral blood flow and previous works have used it to explore the cognitive regulation of emotional responses to aversive events [14]. Indeed, highly negative scenes and the analysis of the fMRI activation response have been used to evaluate the reappraisal of negative affects. It determines an increase in activation of the lateral and medial prefrontal cortex, confirming that the prefrontal region is involved in multiple emotion processing systems. MEG and fMRI scanner, however, are very expensive and bulky, thus are not portable. As such, they have seen limited use in BCIs or aBCIs. As emphasized in [6], however, fNIRS can play a key role in overcoming the spatial resolution limitation of the EEG.

fNIRS measures local metabolic demand in the surface of the cortex. This non invasive diffuse optical-imaging technique does so by detecting the absorption of light of oxygenated and deoxy-

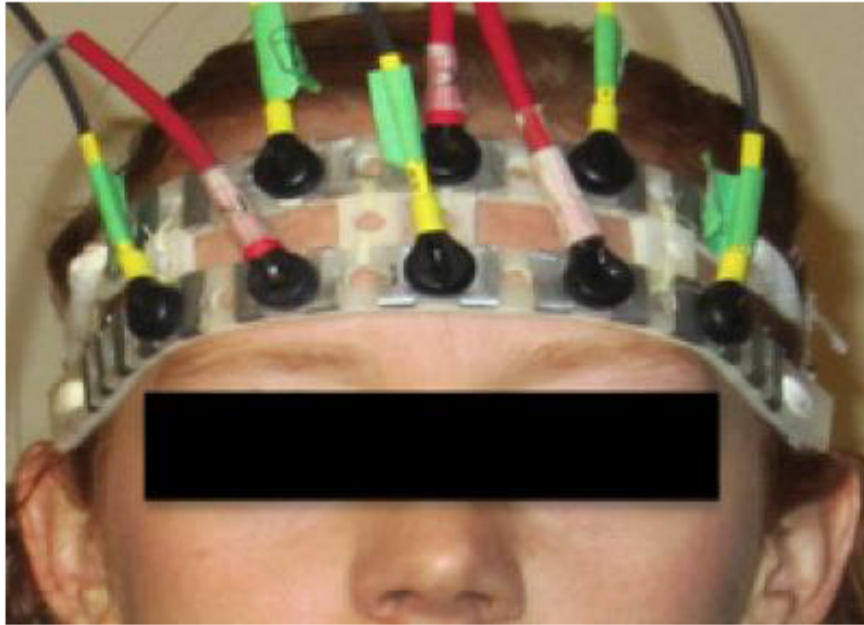


Figure 1.1 – Hemodynamic activation response detected with a 7 channels fNIRS detector.

generated blood. It also provides an easier alternative to using fMRI, MEG and EEG in testing populations, in particular populations that are difficult to test. Indeed, infants being able to rest their head on a parent's lap as imaging takes place, makes fNIRS a particularly suitable exam for infants [6]. Moreover, during the task, the subject simply performs a mental task to produce characteristic patterns in their brain activity, and the hemodynamic response is induced and transmitted as an output. The position of the detectors are suitable for the frontal region of the brain because there is no impedance caused by the hair, but they could be collocated with EEG in every region of the scalp.

Figure 1.1 shows a subject wearing fNIRS optodes (light sources and detectors) on the forehead. Small portable fNIRS systems detect optical signals between the emitter and the detector, which

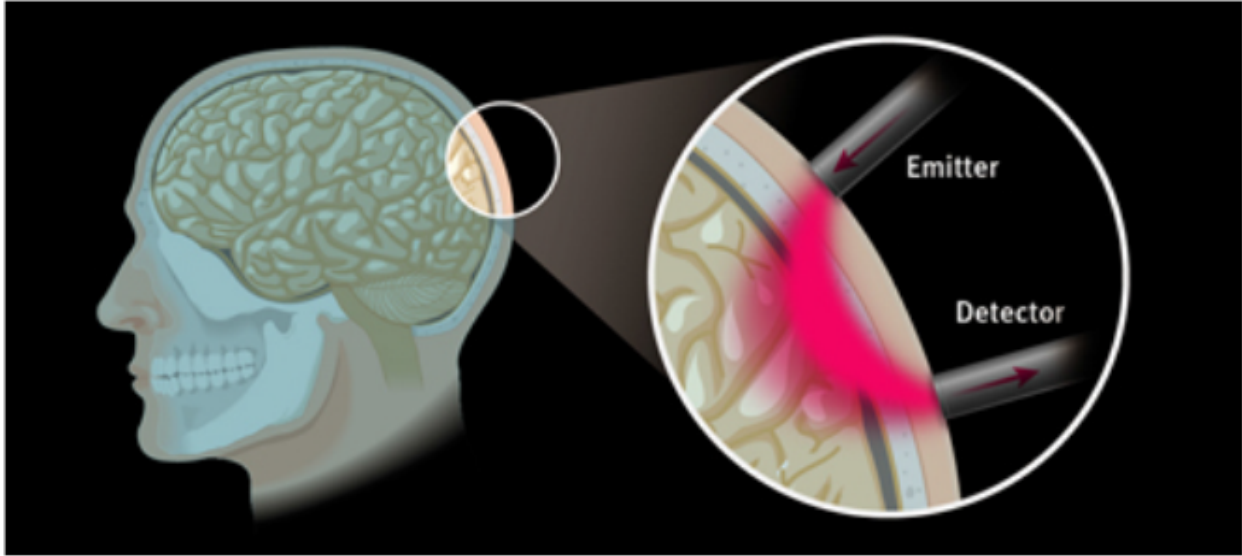


Figure 1.2 – Light from the light source is guided to the head by an optode. A photo-detector will collect the light which leaves the head at a distance of some centimetres. The photons follow a banana-shaped path from light source to detector.

are known to travel in a banana-arch shape through the scalp, skull and cortical tissue, as shown in Fig. 1.2. A simple headset ensures these optodes remain stable. The banana shaped intersection of non-invasive near infrared light emitted and detected by the 2 optodes typically penetrates the skull up to 2-3cm into the cortex. Since two different wavelengths are emitted, the system can measure oxy- and deoxy-hemoglobin separately (e.g., Hitachi ETG-4000 fNIRS system uses 690nm and 830nm wavelengths).

Concentrations of oxygenated and deoxygenated hemoglobin can be calculated for individual light emitter-detector on a probe, which is frequently designed for a specific region of the brain, in the form of time-dependent signals. These concentrations may represent a participant's averaged hemodynamic response to sensory stimuli or motor responses and are the basis of fNIRS data. Typically, the position of optode pairs on a probe enables data to be displayed topographically. The simplified Beer-Labmert law, as detailed in [51], calculates the concentration of oxygenated

hemoglobin ΔHbO and deoxygenated hemoglobin ΔHbR based on their different light absorption.

The concentrations are calculated as:

$$\Delta HbO = \frac{\Delta OD}{bl\alpha_{HbO}} \quad \Delta HbR = \frac{\Delta OD}{bl\alpha_{HbR}} \quad (1.1)$$

where ΔOD is the change in light intensity, l is the geometric distance between the source and the detector, α_{HbO} and α_{HbR} are the specific extinction coefficients of hemoglobin and b is the length of the photon path between source and detector.

Important advantages of the fNIRS system are that it can be transported into participants' homes, hospitals' neonatal units, and school systems due to their much smaller size (when compared to fMRI) and portability, and their use of light as opposed to magnetism makes them completely silent, hence ratifying fNIRS as the most popular modality for the study of hemodynamic response [6]. For those reasons fNIRS will be one modality of interest in this work.

To date, hemodynamic response has not been widely used in aBCIs [49], and the majority of published studies have focused on measuring stress and mental workload [27, 115]. One exception is the work described in [128] where the authors studied the affective reaction of the subjects during interactions with a robot, discriminating between aversion and affinity. In order to study the emotional response the authors analyzed one channel that corresponded to participants' oxygenated hemoglobin concentration changes in the left PreFrontal Cortex. It was demonstrated that there is a connection between the level of concentration of blood in the prefrontal cortex and the aversity with the robot.

In the work of [89], the authors investigated the attachment between mothers and infants using fNIRS and focusing on the orbitofrontal cortex. Twenty six infant-mother pairs were used as participants and for each of them changes in concentrations of oxygenated and deoxygenated hemoglobin were analyzed revealing that oxy-Hb in the vicinity of the orbitofrontal region increases following exposure to rewarding stimuli as can be someones' infant.

Moreover, it has been widely shown that physiological signals include emotional information that can be used for emotion assessment, and they comprise the signals originated from the Peripheral Nervous System (PNS) [80, 94, 140]. Experiments have been used to investigate and recognize emotions and related emotional states of the subject in response to images [74] or videos [66]. Emotional preference is one of the most important factors in multimedia content selection and consumption.

Typical neurophysiological signals that have been used in affective computing include:

- Bioelectrical signals, generated by nerve cells and muscle cells. Electrocardiographic (ECG) signals are an example, and they vary from resting potential to action potential under certain conditions. The change of potential in many cells generate an electric field which fluctuates and in this process it is to emit bioelectric signal.
- Biomechanical Signals, related to motion, pressure and flow of the physiological system i.e. mechanical function. An example is respiration with the movement of the chest.
- Bioacoustics Signals, created by the physiological systems which are dealing with the flow of blood and air. The flow of the blood in the heart, the opening and closing of chest in respiratory system generate unique acoustic signals.

- Bioimpedance Signals: The impedance of the skin depends upon the composition of skin, blood distribution and blood volume through the skin. The measurement of impedance helps in finding the state of skin and functioning of various physiological systems. An example is Galvanic Skin Response (GSR), which measures the amount of sweat.

The heart is not just a muscle that serves to pump blood to the whole body. It is connected with the Autonomic Nervous System (ANS), thus also reflects emotions and feelings of the person. Normally, healthy human heart does not beat at a precise rate, but it is slightly irregular and it is constantly influenced by external or internal stimuli. For example, a stress situation makes an adjustment on the response of the heart. The consequence could lead to an increase or decrease of the Heart Rate (HR) and its variation, leading to the widely-used metric called Heart Rate Variability (HRV). Several emotions have been correlated with a variation in HR. Anger, fear, happiness and sadness are able to increase the average of HR and shape cardiac acceleration [110], whereas in [74] it was demonstrated that HR is significantly correlated with pleasantness. Moreover, negative states such as angry, sadness and fearful have been discriminated from happiness and surprise by [32], combining HR and other physiological signals such as GSR and Electromyography (EMG). Regarding HRV, it has been shown that fear, anger and happiness are able to modify the energy in the high frequency bands [110]. Statistical and frequency features from HR and HRV are the most useful for emotion detection [61].

The Blood Volume Pulse (BVP) signal given by the pletysmograph is a relative measure of the Blood Pressure (BP). Changes in BP are influenced by cardiac output (contractions of the heart), vasoconstriction and vasodilatation (reduction and augmentation of the vessels diameter by contraction and dilatation), and also by mechanisms occurring in other organs [124]. Thus, it is

possible to extract HR and HRV from the Blood Volume Pulse (BVP) curve [66], together with the derived features.

The ability to conduct electricity is the method of the body to relay information from one part to another, and all tissues are included. Thus, also the skin has an electrical activity and it can be detected and used to understand the changes in the sympathetic nervous system, in particular to ‘measure’ emotions due to internal or external stimulus. The electrical activity of the skin is always in constant and slight variation. The fluctuation of the conductivity is called GSR and can be detected, analyzed and graphed on a chart for observation as the heart or brain response, but obviously it is not possible to detect the type of emotion or physical state. The devices which measure the response of the skin are low cost and the signal is easily captured and robust. They are often referred to as feedback instruments because the response is not measured ‘actively’. Indeed, the GSR test is a passive device and measures the current generated by the person’s body itself. The reported feedback is the measured GSR. In many past studies, GSR has been used to detect emotions [66]. Stress and surprise have been connected with an increase of activity in sweat glands [109]. Moreover, previous studies [74] demonstrated that the amplitude of GSR is correlated with the level of arousal using images from IAPS. Discrimination of pleasantness using olfactory stimuli has been conducted by [28] and the authors found that higher GSR amplitude is related to unpleasant odors. Moreover, emotional words can influence the GSR signal, as stated in [123]. Thus, in order to integrate GSR in emotion detection have been used features obtained from amplitude and electrodermal temporal parameters.

In turn, the study of respiration variability in healthy and pathological subjects has been an interesting method to analyze the behavior of the body during activities or simply during the sleep.

For instance, as explained in [116], it has been used to detect sleep apnea events, and the work was useful to distinguish the signals in normal respiration, motion artifact and sleep apnea. Nevertheless, respiration is also influenced by emotional processes [112, 135]. The works in [65, 110] showed that a low breathing rate is affected by relaxation whereas high arousal is expressed with an irregular rhythm, with rapid variation and interruption of respiration. Time and frequency features have been tested with respiration [66]. For instance, statistical features and signal derivatives have been used to obtain instantaneous changes; spectral features have also been explored [66].

Skin Temperature (ST) describes the temperature as measured on the surface of the skin. Since muscles are tense under strain, the blood vessels will be contracted and therefore the temperature will decrease. ST has been used in many studies either as a physiological index to monitor humans or animals, or to analyze the emotional response to external factors [66]. Moreover, ST has the advantage that it can be assessed on both sides of the body and, therefore, asymmetry in temperature can be assessed. ST is usually a slow signal and it is hard to detect a variation in a short time period. However, ST is highly related to vasoconstriction and sweat and it is possible to depict thermal reactions over a few seconds. As an example, [32] demonstrated that the average over an interval of time (10 seconds) was useful to discriminate anger from fear and sadness. In turn, [86] studied the thermal reaction in response to emotional music and the authors found that there is an increase in temperature for calm positive music. Moreover, first derivative and spectral power was tested in [66] to study arousal and valence states.

Recent studies have explored the combination of multiple signals as a new way to discriminate emotions. For instance, in the work of [69] the instantaneous phase of GSR was coupled with the amplitude of EEG while subjects were experiencing music video clips. The authors argued that

physiological processes are dynamic rather than static and physiological signals interact between each other instead of being independent. As a result, it was demonstrated that GSR is more coupled with EEG signals of the temporal lobe. Moreover, the coupling increases in high arousal and high/low valence music clips. This coupling has been reproduced in the present work, as explained in Section 3.1.3.

Affective BCIs that rely on multiple signal modalities at once have been recently termed ‘hybrid’ BCIs. The second section to follow describes hybrid BCIs in more detail.

1.2 Hybrid BCI

A primary limitation of affective computing is the use of monomodal affect recognition systems approach. In natural Human-Human Interaction (HHI), multimodal communication and interaction takes place and oftentimes, the use of only one modality can result in an ambiguous and sometimes incomplete way of sharing information. Developing multimodal affective systems, however, is a challenging task and many open problems and questions exist. For example, which modalities should be combined that are complementary to each other? Which features convey such complementarity more effectively? Which classifiers are better for the task at hand? How do we tackle the different temporal latencies of the measured signals (e.g. order of milliseconds for EEG or seconds for skin temperature). The first attempt on a multimodal approach was made on audio and video data with promising results [6]. Presently, many researchers are exploring which modalities contribute the most for automatic affective recognition and physiological signals have gained grounds for emotion recognition [32, 102, 109]

An hBCI has been defined as a combination of two or more signals, in which at least one of them is a BCI channel [101]. According to [101], one of the conditions to be satisfied in order for a system to become an hBCI is that the device must have at least one signal detected from the brain (either electrical or hemodynamic). The other signal might come from another input (e.g. Electrocardiography (ECG), EMG or Eye Gaze etc.). According to [2], the hBCI can be divided into three categories, *pure*-hBCI, combining two or more conventional BCIs, *physiological*-hBCI with the association of a BCI and another physiological signal, and finally *mixed*-hBCI to indicate the integration of a BCI with a non-physiological input. Since passive BCI has been recognized as a BCI by [136], the work in [6] extends the definition of *physiological*-hBCI stating that a non-neurophysiological modality has to be one of its components.

To date, most of the work on hBCI has been for active BCIs. For instance, in the work of [34], EEG and fNIRS measurements were used simultaneously for Event Related Desynchronization (ERD)-based BCIs. It was demonstrated that the combination of EEG and fNIRS for motor imagery and motor execution tasks improve the classification accuracy compared to the one extracted for the single modality. Concentration of oxygenated hemoglobins and deoxygenated hemoglobins were used as fNIRS features. Affective hybrid BCIs, on the other hand, are relatively new and the majority of the work has been devoted to stress or mental workload monitoring [91, 92].

Nonetheless, it has been demonstrated that emotional responses involve different pathways in the central and autonomic nervous system including cardiovascular, respiratory and electrodermal [68, 77]. As such, the present study has explored the benefits of developing a hybrid affective BCI. The main thesis objective and contributions are described next.

1.3 Thesis objectives and contributions

The main objectives of this work are three-fold. First, we build on a claim by [37] that temporal dynamics information is more valuable for emotion recognition than static features, such as overall spectral powers or asymmetry indices. As such, we have developed a new EEG feature that measures the spectral dynamics of different amplitude modulation envelopes, thus providing new insights into cross-frequency coupling measures. We showed that the proposed metric outperformed several other benchmark measures. A new connectivity metric was also proposed and showed to further improve emotion recognition accuracy. Second, we explore the gains attained by combining fNIRS and/or physiological signals into an affective BCI, thus gauging the benefits of each modality and the complementarity provided for the task at hand. Lastly, we explore different fusion techniques, such as feature- and decision-level fusion, to explore their inherent benefits for emotion recognition. The outcomes of this research have been submitted to the IEEE Transactions on Affective Computing [18], as well as presented at the IEEE Neural Engineering Conference [17] and the fNIRS conference 2014 [16].

1.4 Outline

This work is organized as follows. In this chapter, we introduced the problem of the improvement of BCIs for affective computing and described the work objectives and contributions. In Chapter 2 we discuss the state-of-the-art of the study of emotions. Different models of emotions are presented together with the self-assessment method used in the present work. In Chapter 3 we present refinements to the aBCI metrics that aim at the development of an automated affective device. We then

compare the performance obtained with benchmarks commonly used in aBCI and discuss advantages and shortcomings. We finally conclude the thesis with a summary of the results (Chapter 4) and some final considerations in Chapter 5.

Chapter 2

Subjective scales of emotions

Affective Evaluation Methods (AEM), as a way to study emotional engagement and experience of a subject using different stimuli, became a field of interest for many researchers and they were able to detect two main challenges: the terminology to be used and the effects of context on emotion elicitation. First, possible names and definitions of different emotional states and primitives have come from a range of possible theories and work on how to determine, communicate and depict them effectively has been very active over the past decade. The second challenge relates to the context in which the emotions are being elicited, particularly the environment (real versus acted) as it plays a crucial role on the emotional experience [36, 103].

As such, great effort has been put in order to clarify terms such as ‘affect’ and ‘emotion’. Many definitions were elaborated, but as highlighted by [107] and others [35] ‘there is an extensive overlap and no clear distinctions can be made at this time’. It was only in 2006 that a clear distinction

between the two terms was created and defined ‘affect’ as the broader term encompassing ‘anything to do with emotions, moods, dispositions and preferences’.

The other important aspect that has to be considered when dealing with the study of affective states is the personal training of the subject [36]. Selecting participants from the same target user group does not mean that they start from the same sensibility and their reaction can be compared in data obtained from an AEM [95]. Indeed, experiences from childhood or adulthood as well as cultural factors, are the base of the formation of every human being. Moreover, the same emotions detected on different days might be represented with a relevant diversity in physiological features [84]. In order to overcome the problem, a normalization of the features has been tested [59, 105], and resulted in an improvement handling across-subject variations, but with the drawback of non real-time processing requirements. In order to better define the task at hand, the following section examines human emotions, as well as their two most common categorizations.

2.1 Emotions

In order to examine emotions, two different categorizations have been determined. Discrete emotions [107] and Dimension of Emotions [118]. More specifically:

- Discrete emotions

The discrete emotion theory claims that there is a set of basic emotions that are biologically determined emotional responses whose expression and recognition are the same for all individuals, overstepping the differences in ethnicity and culture [107]. A list of eight basic emotions

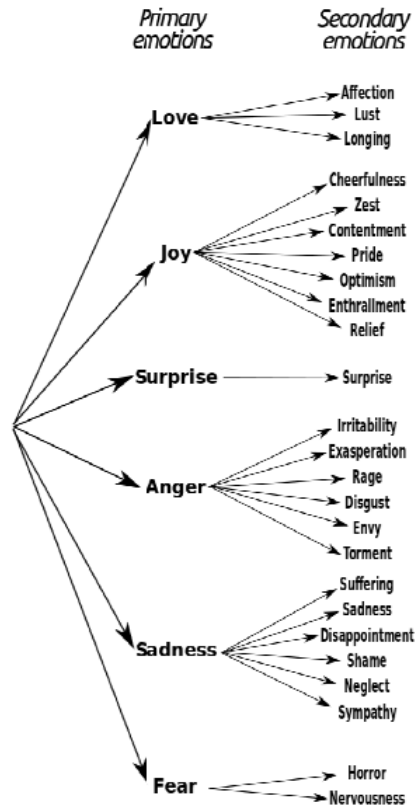


Figure 2.1 – First two layers of Parrot's emotion classification.

is developed in [57]: interest, joy, surprise, distress, disgust, anger, shame and fear. The work in [133] stated that the emotional experience generally consists of a complex combination of those basic emotions, since the presence of only one of them is rare and hard to assess. Various discrete categorizations of emotions have been proposed. For instance the six basic emotions introduced by [31] such as anger, disgust, fear, happiness, sadness and surprise, or the tree structure of emotions advanced by [97] where the first level is composed by six primary emotions (love, joy, surprise, anger, sadness, fear) and the secondary emotions are derivation of the primary ones instead of being their combination (Fig. 2.1).

- Dimensions of Emotions

The dimensional view of emotions proposes an alternative to discrete emotion theory since the 'real' structure of emotions is not perfectly matched by a fixed configuration. The method

proposes a conceptualization of human emotions defining where they lie in a two or three dimensional space of so-called emotional ‘primitives’. Dimensional models of emotion, in contrary to the theory of basic emotions, states that a common interconnected neurophysiological system is responsible for all the affective states. The original dimension approach includes arousal (intensity of the emotion, expressed as calmness-excitement), valence (the pleasantness, positive or negative experience) and dominance (controlling and dominant nature of emotion) [118] primitives. Alternate dimensions of emotions have also been proposed such as Plutchik’s emotion wheel [106] and the valence-arousal scale by [113]. While using only the arousal and valence scales prevents a more detailed characterization of emotions, Russel’s ‘Circumplex of affects’ distributes discrete emotions along the axes of these two dimensions, as depicted by Fig. 2.3. Moreover an organization in circle has been introduced by [106] with the emotion wheel (Fig. 2.2).

The model designed by [113] is widely used to determine the 3-dimensional structure of objects, experiences and situation. It is made of 18 bipolar adjective pairs that are rated with a 9-point scale [13]. Analyses over the ratings generate scores on the dimensions of valence, arousal and dominance. Each emotion can be categorized in a plane with arousal on the horizontal and valence on the vertical axes (Fig. 2.3). Arousal can vary from a status of indifference and tiresome to a more active state dominated by feelings like excitation and action. Valence instead consider the pleasantness, the feeling of sad, stress or joy and ecstatic. One third dimension that can be added is dominance. It ranges between a state of submission to the stimulus (defenseless and feeble) to a complete control, feeling of power.

In this work, we use the Russell’s valence-arousal scale. It has been already used in consumer behavior studies or an abbreviated model has been used in organizational studies where the emotions

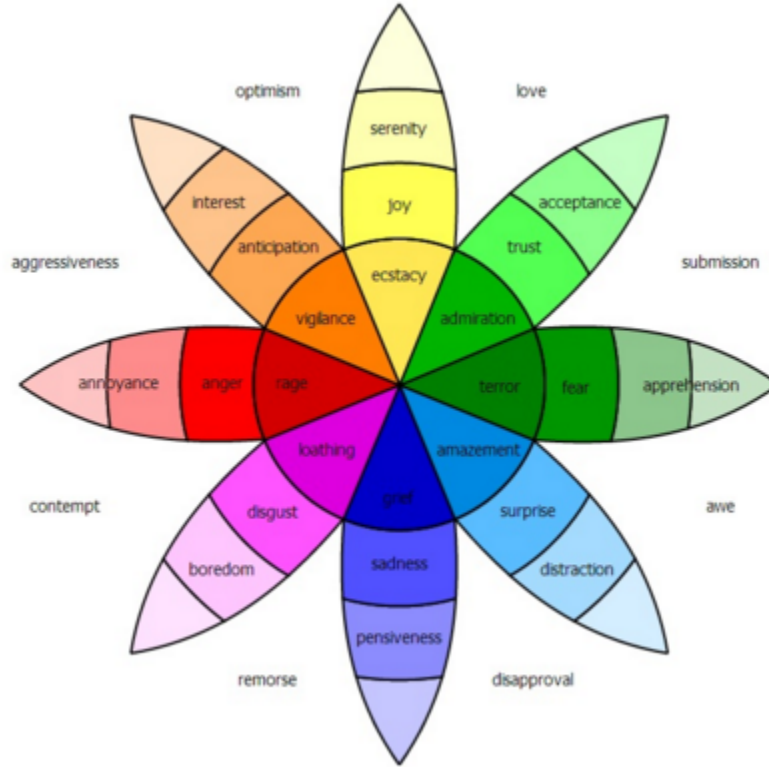


Figure 2.2 – Plutchik’s Wheel of Emotions

towards a specific product marketed by them are measured [30]. Furthermore, in [8] it has been used in the construction of animated agents that exhibit emotions. For self-assessment along these scales, we use the well-known SAM as detailed next.

2.2 SAM: The Self-Assessment Manikin

While the Valence-Arousal-Dominance model was gaining attention in different fields, from advertising to strategy formulation and copy development [47, 143], there was a need to make it more functional, applicable and useful. The impulse for improvement aimed at reaching a quick response to a given stimulus, to be able to precisely reflect a complete range of feelings and to do so without the need for linguistic interferences and using only visual objects. The SAM [13] represents a

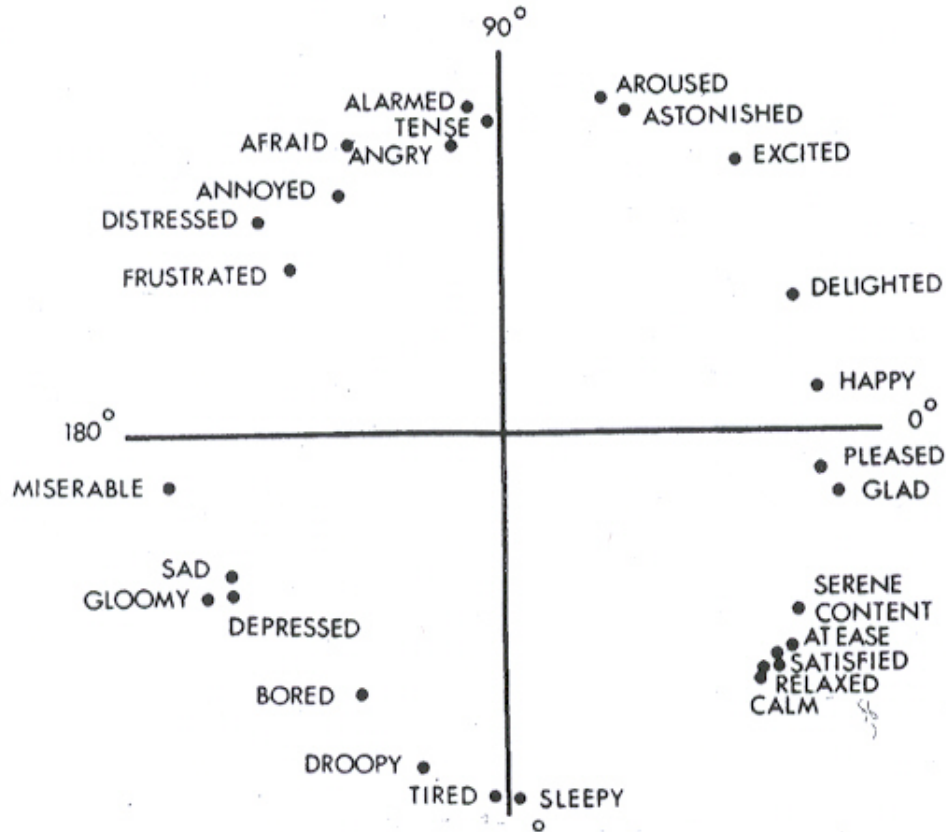


Figure 2.3 – Circumplex of affects [113]. Discrete emotional terms were organised around the two dimensions of arousal (top to bottom) and valence (left to right).

valuable solution to these objectives. Initially designed as an interactive computer program, it was later exported on paper as illustrated in Fig. 2.4. Each of the three major affective dimensions is typically represented by five different levels illustrated by pictograms [120]. The pleasantness dimension is represented by a set of figures going from a smiling one to a frowning, unhappy one. A range that starts with an excited to a relaxed figure describes the arousal dimensions. Finally, the dominance dimension is represented by the same figure in different sizes, considering the biggest one that corresponds to a full control of the situation. Commonly, the pictograms are associated with a 9-point scale, thus users are allowed to choose one of the images or select in between two images.

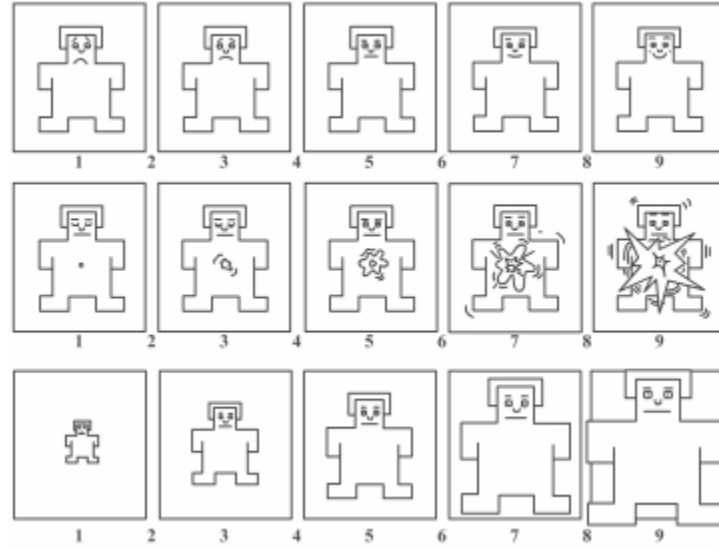


Figure 2.4 – The Self-Assessment Manikin (SAM) used to rate the affective dimensions of valence (top panel), arousal (middle panel), and dominance (bottom panel).

SAM can be seen as an efficient method for quick and direct assessment of three main generalized affective dimensions associated with a person’s emotional reaction to an event. Some argue that it is capable of addressing emotional experiences before cognitive processes signal information to the linguistic part of the brain [53]. Moreover, it is an easy and nonverbal method that can be used with a variety of subject populations, including non-english speakers [13, 29], children [4] and people with disorders or clinical syndromes [134]. Furthermore, participants take less than 15 seconds to fill out the SAM form, thus it prevents exceeding fatigue in longer studies. Given its popularity, the SAM rating scales have been widely used in affective computing research and will be used also in the experiments described herein.

Chapter 3

Methods and Materials

This chapter describes the methods and materials used throughout this study. First, Section 3.1.1 describes conventional state-of-art features used for affective recognition. These features are used here to benchmark the results obtained with the newly proposed features and systems.

Next, Sections 3.1.2 and 3.1.3 describe several innovative features which are proposed and explored, namely EEG amplitude-amplitude cross-coupling features, and coupling features between the Peripheral Autonomic Nervous Systems (PANS) and Central Nervous System (CNS), respectively. Section 3.1.4, in turn, describes a multimodal fNIRS-EEG fusion strategy to improve affective state classification.

Lastly, Sections 3.2, 3.3 and 3.4 describe the two publicly available datasets used in the studies described herein, followed by the feature selection and classification algorithms used, respectively.

3.1 Features

3.1.1 Benchmark EEG Features

Typically, spectral features such as subband spectral powers have been used to measure emotional states elicited from music videos, pictures, and/or movie clips (e.g., [66, 126]), as well as mental workload and stress (e.g., [50, 67]). Moreover, an inter-hemispheric asymmetry in spectral power has been reported in the affective state literature [25, 60], particularly in frontal brain regions [19]. Spectrum subband power features are the most traditional measures used in biomedical signal processing [127]. Within the affective state recognition literature, spectral power in the theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz) and gamma (30-45 Hz) subbands are typically used [60] across different brain regions [5, 122]. In particular, alpha and gamma band inter-hemispheric asymmetry indices have been shown to be correlated with emotional ratings [82, 93]. In this work, Spectral Features (SF) are used as a benchmark. A total of 128 spectral power features ($32 \text{ electrodes} \times 4 \text{ subbands}$) and 56 asymmetry indices ($14 \text{ inter-hemispheric pairs} \times 4 \text{ subbands}$) were computed from the following electrode pairs: Fp1-Fp2, AF3-AF4, F7-F8, F3-F4, FC5-FC6, FC1-FC2, T7-T8, C3-C4, CP5-CP6, CP1-CP2, P7-P8, P3-P4, PO3-PO4, and O1-O2 (see Fig. 3.1 for electrode labels and locations). Overall, a total of 184 SF features are used as benchmark.

3.1.2 Proposed EEG Features: Amplitude Modulation

Cross-frequency amplitude-amplitude coupling in the EEG has been explored in the past as a measure of anxiety and motivation (e.g., [121]), but has been under-explored within the affective

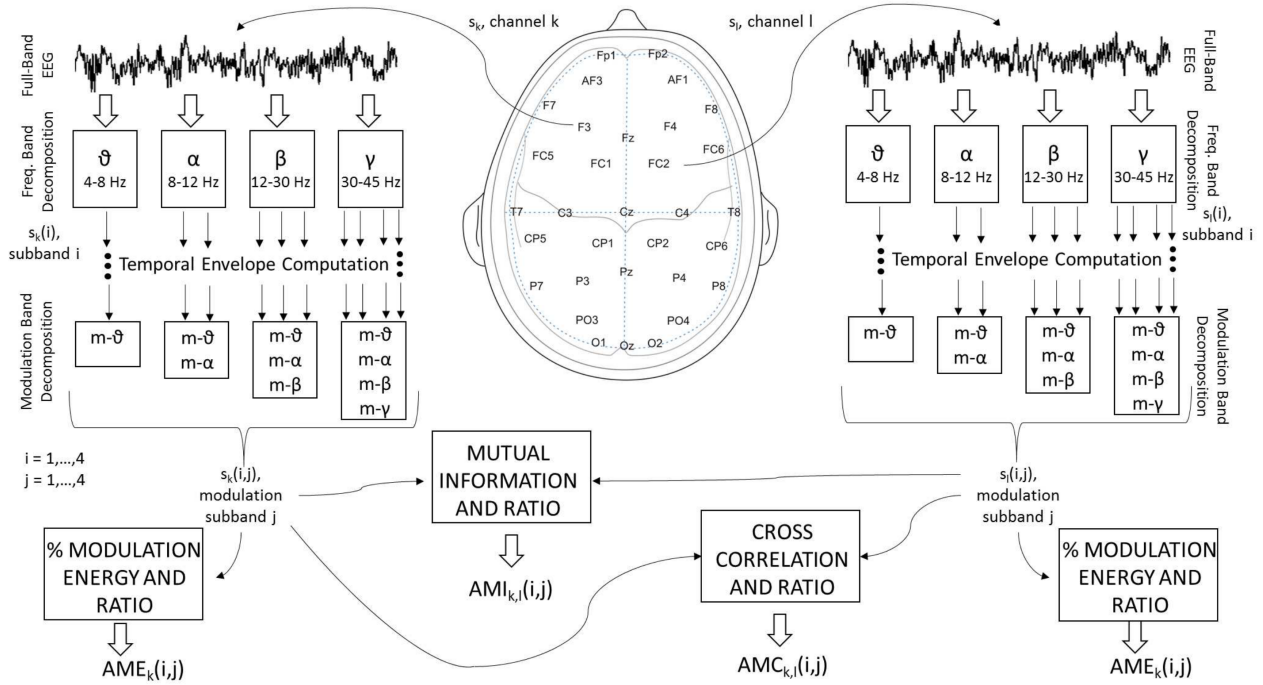


Figure 3.1 – Signal processing steps used to compute the EEG amplitude modulation feature sets.

state recognition community. Recently, beta-theta amplitude-amplitude coupling differences were observed between healthy elderly controls and age-matched Alzheimer’s disease patients; such findings were linked to lack of interest and motivation within the patient population [37]. To explore the benefits of cross-frequency amplitude-amplitude modulations for affective state recognition research, the authors recently showed that non-linear coupling patterns within inter-hemispheric electrode pairs was a reliable indicator of several affective dimensions, but particularly for the valence emotional primitive [17]. Here, we extend this work by extracting a number of other Amplitude Modulation Features (AMF) and show their advantages for affective state recognition.

More specifically, three amplitude-amplitude coupling feature sets are extracted, namely the Amplitude Modulated Energy (AME), Amplitude Modulated Interaction (AMI), and the Amplitude Modulated Coherence (AMC), as depicted by Fig. 3.1. In order to compute these three feature sets,

first the full-band EEG signal s_k for channel ‘k’ (see left side of the figure) is decomposed into the four typical subbands (theta, alpha, beta and gamma). Here, the time-domain index ‘n’ is omitted for brevity, but without loss of generality. For the sake of notation, the decomposed time-domain signal is referred to as $s_k(i), i = 1, \dots, 4$. The temporal envelope is then extracted from each of the four subband time series using the Hilbert transform [75]. In order to measure cross-frequency amplitude-amplitude coupling, a second decomposition of the EEG amplitude envelopes is performed utilizing the same four subbands as in the first decomposition; a second-order time-domain filterbank was used. To distinguish between modulation and frequency subbands, the former are referred to as m-theta (4-8 Hz), m-alpha (8-12 Hz), m-beta (12-30 Hz) and m-gamma (30-45 Hz). For notation, the amplitude-amplitude coupling pattern is termed $s_k(i, j), i, j = 1, \dots, 4$, where ‘i’ indexes spectral subbands and ‘j’ the modulation spectral subbands. By using the Hilbert transform to extract the amplitude envelope, the types of cross-frequency interactions are limited by Bedrosian’s theorem, which states that the envelope signals can only contain frequencies (i.e. modulated frequencies) up to the maximum frequency of its original signal [11, 125]). As such, only the ten cross-frequency patterns shown in Fig. 3.1 are possible (per electrode), namely: theta_m-theta, alpha_m-theta, alpha_m-alpha, beta_m-theta, beta_m-alpha, beta_m-beta, gamma_m-theta, gamma_m-alpha, gamma_m-beta, and gamma_m-gamma. From these patterns, the three feature sets are computed, as detailed below:

Amplitude Modulation Energy, AME

From the ten possible $s_k(i, j)$ patterns per electrode, two energy measures are computed. The first measures the ratio of energy in a given frequency–modulation–frequency pair ($\xi_k(i, j)$) over the total

energy across all possible subband pairs

$$AME_{i,j} = \sum_{i=1}^4 \sum_{j=1}^4 \xi_k(i, j), \quad (3.1)$$

thus resulting in 320 features (32 electrodes $\times$ 10 cross-frequency coupling patterns).

The second measures the logarithm of the ratio of modulation energy during the 60-second music clip to the modulation energy during a 3-second baseline resting period, i.e.,

$$ratioAME_{i,j} = 10 \log \left(\xi_k(i, j)^{video} / \xi_k(i, j)^{baseline} \right), \quad (3.2)$$

thus resulting in an additional 320 features, for a total of 640 $AME_k(i, j)$ features, $k = 1, \dots, 32; i, j = 1, \dots, 4$.

Amplitude Modulation Interaction, AMI

In order to incorporate inter-electrode amplitude modulation (non-linear) synchrony, the amplitude modulation interaction (AMI) features from [17] are also computed. Unlike the work described in [17], where interactions were only computed per symmetric inter-hemispheric pairs, here we measure interactions across all possible 496 electrode pair combinations for each of the ten cross-frequency coupling patterns, thus in the first place extracting 4960 features. The normalized mutual information (MI) is used to measure the interaction as:

$$AMI_{k,l} = \frac{H(s_k) + H(s_l) - H(s_k, s_l)}{\sqrt{H(s_k)H(s_l)}}, \quad (3.3)$$

where the $H(\cdot)$ operator represents marginal entropy and $H(\cdot, \times)$ the joint entropy, and s_k corresponds to $s_k(i, j)$ with the frequency and modulation frequency indices omitted for brevity. Entropy was calculated using the histogram method with 50 discrete bins for each variable. Mutual information has been used widely in affective recognition research (e.g., [20, 45, 64]). Additionally, a second measurement of logarithmic ratio between the 60-second clip and the 3-second baseline has been obtained, thus totalling 9920 AMI features. The inter-hemispheric version of this metric described in [17] was called IAMI (Inter-Hemispheric Amplitude Modulated Interaction).

Amplitude Modulation Coherence, AMC

While the AMI features capture non-linear interactions between inter-electrode amplitude-amplitude coupling patterns, the Pearson correlation coefficient between the patterns can also be used to quantify the coherence, or linear interactions between the patterns. Spectral coherence measures have been widely used in EEG research and were recently shown to also be useful for affective state research (e.g., [62, 141]). Hence, we explore the concept of amplitude modulation coherence, or AMC as a new feature for affective state recognition. The AMC features are computed as:

$$AMC_{k,l} = \frac{\sum_{n=1}^N (s_k(n) - \bar{s}_k)(s_l(n) - \bar{s}_l)}{\sqrt{\sum_{n=1}^N (s_k(n) - \bar{s}_k)^2 \sum_{n=1}^N (s_l(n) - \bar{s}_l)^2}} \quad (3.4)$$

where $s_k(n)$ indicates the n -th sample of the $s_k(i, j)$ time-series (again, the frequency and modulation frequency indices were omitted for brevity), and $\bar{s}_k$ is the average over all samples of such time series. As previously, a total of 9920 AMC features are computed, including the logarithmic ratio with the 3-seconds baseline.

3.1.3 Proposed Features: PANS-CNS coupling

Recent studies, however, have suggested that alternate EEG feature representations may exist that convey more discriminatory information over traditional spectral power and asymmetry indices [60]. Representative features include the recently proposed inter-hemispheric amplitude modulation interaction [17] and EEG graph-theoretic features [40]. In addition to EEG correlates, complementary affective state information has been obtained from physiological signals measured from the PANS [80, 94, 140], particularly the GSR, a measure of the amount of sweat (conductivity) in the skin [10, 104]. More recently, the interaction between the PANS and CNS was measured via a Phase-Amplitude Coupling (PAC) between GSR and EEG signals and promising results were found [70]. Here, several PANS-CNS coupling metrics are explored as correlates of affective states, as detailed below.

Electrophysiological signals are dynamical systems that interact with each other at different frequencies. Phase-Amplitude coupling represents one type of interaction and typically refers to modulation of the amplitude of high-frequency oscillators by the phase of low-frequency ones [130]. Typically, such phase-amplitude coupling measures are computed from EEG signals alone [121], but the concept of electrodermal activity phase coupled to EEG amplitude was recently introduced as a correlate of emotion, particularly for high arousing, very pleasant and very unpleasant stimuli [69, 70]. Here, we test three different GSR-phase and EEG-amplitude coupling measures. For the sake of notation, assume $u(n)$ is the narrowband time-domain GSR signal. Using the Hilbert transform [38], we can extract the signal's instantaneous phase $\phi_u(n)$ as in [70]:

$$\phi_u(n) = \arctan\left(\frac{\text{Hilb}\{u(n)\}}{u(n)}\right), \quad (3.5)$$

where $\text{Hilb}\{\cdot\}$ corresponds to the Hilbert Transform of $u(n)$. For the amplitude envelope of the EEG signal ($A(s_k(n))$), a shape-preserving piecewise cubic interpolation method of neighbouring values is used, as in [70]. Given the GSR signal and phase, as well as the EEG amplitude envelope signals, the following coupling measures were computed.

Envelope-to-Signal Coupling, ESC

The simplest coupling feature can be calculated via the Pearson correlation coefficient between the EEG amplitude envelope signal $A(s_k(n))$ and the raw GSR signal $u(n)$. The Envelope-to-Signal Coupling (ESC) feature can be computed using (3.4) but with $A(s_k(n))$ and $u(n)$ in lieu of $s_k(i, j)$ and $s_l(i, j)$, respectively [3]. ESC has been shown [96] to be particularly useful with noisy data. A total of 32 ESC features were computed.

Cross-Frequency Coherence, CFC

Cross-frequency coherence evaluates the magnitude square coherence between the filtered 0-1 Hz GSR signal $u(n)$ and the filtered envelope of the EEG signal $A(s_k(n))$ between 4-45 Hz, as in [96]. The Cross-Frequency Coherence (CFC) feature is computed as:

$$CFC_k(f) = \frac{|P_{Au}(f)|^2}{P_{AA}(f)P_{uu}(f)}, \quad (3.6)$$

where $|P_{Au}(f)|^2$ is the cross power spectral density of the EEG amplitude $A(s_k(n))$ and GSR signal $u(n)$ at frequency f , and $P_{AA}(f)$ and $P_{uu}(f)$ are the spectral power densities of the two signals, respectively. The CFC feature ranges from 0 (no spectral coherence) to 1 (perfect spectral coherence)

and has been used previously to quantify linear EEG synchrony in different frequency bands and its relationship with emotions [21]. A total of 1344 CFC features were computed corresponding to the coupling of 42 subbands for each of the 32 channels of the EEG with the GSR signal.

Modulation Index, ModI

The last PANS-CNS coupling measure tested is the so-called Modulation Index (ModI), which was recently shown to accurately characterize coupling intensity [130], particularly for emotion recognition [70]. For calculation of the ModI feature, a composite times series is constructed as $[\phi_u(n), A(s_k(n))]$. The phases are then binned and the mean of $A(s_k(n))$ over each phase bin is calculated and denoted by $\langle A_s \rangle \phi_u(m)$, where m indexes phase bin; 18 bins were used in our experiments. Next, the mean amplitude distribution $P(m)$ is normalized by the sum over all bins, i.e.:

$$P(m) = \frac{\langle A_s \rangle \phi_u(m)}{\sum_{m=1}^{18} \langle A_s \rangle \phi_u(m)}. \quad (3.7)$$

The normalized amplitude ‘distribution’ $P(m)$ has similar properties as a probability density function. In fact, in the scenario in which no phase-amplitude coupling exists, $P(n)$ assumes a uniform distribution. Having this said, the ModI feature measures the deviation of $P(m)$ from a uniform distribution. This is achieved by means of a Kullback-Liebler (KL) divergence measure [71] between $P(m)$ and a uniform distribution $Q(m)$, given by:

$$D_{KL}(P, Q) = \sum_{m=1}^{18} U(m) \log \left[\frac{P(m)}{Q(m)} \right]. \quad (3.8)$$

The KL divergence $D_{KL}(P, Q)$ is always greater than zero, and equal to zero only when the two distributions are the same. Finally, the ModI feature is defined as the ratio between the KL divergence and the log of the number of phase bins, i.e.,:

$$ModI = \frac{D_{KL}(P, Q)}{\log(M)}, \quad (3.9)$$

where $M = 18$ in our experiments. A total of 32 ModI features were computed, corresponding to one per electrode.

3.1.4 Multimodal Integration: NIRS-EEG

As mentioned in Section 1.2, hybrid affective BCIs have been shown to improve affective state recognition, as one modality compensates for shortcoming of the other. Here, we investigate the benefits of an hybrid aBCI using two multimodal fusion strategies, namely feature and decision fusion. Feature fusion [42], as the name suggests, concatenates features from different modalities into a composite feature vector to be used for classification. The curse of dimensionality quickly becomes an issue with feature fusion strategies and feature selection is typically used. Feature fusion has been widely used in emotional state detection, since the body express emotional states in different manners [66]. The second approach typically relies on multiple classifiers, each trained on an individual modality. The final classification, however, relies on a combination of each individual classifier's decision. Typically majority voting or weighted combinations are used as decision fusion strategies. Here, feature and decision fusion is investigated for the following multimodal setups: EEG-fNIRS, EEG-PANS, EEG-fNIRS-PANS. In order to investigate the gains obtained with these fusion strategies, only benchmark features are tested.

3.2 Tasks and Databases

3.2.1 Dataset for Emotional Analysis using EEG and Physiological signals

The publicly-available DEAP (Dataset for Emotional Analysis using EEG and Physiological signals) database was used [66] herein to investigate the gains attained with the newly proposed features described in Section 3.1.3. Thirty-two healthy subjects (gender-balanced, average age 26.9 years) were recruited to watch 40 video music clips while their neuro-physiological signals were recorded. The forty videos were carefully selected from a larger set (roughly 200 videos), corresponding to the ones eliciting the 10 highest ratings within each of the four quadrants of the valence-arousal plane [113]. Participants were asked to rate their perceived valence, arousal, and dominance emotional primitives, as well as other subjective ratings such as liking and familiarity for each of the 40 music clips. The three emotional primitives were scored using the 9-point continuous self-assessment manikin scale [13] shown in Figure 2.4. The liking scale was introduced to determine the user's taste, and not their feelings, about the music clip; as such, 9-point scale with thumbs down/up symbols was adopted. Lastly, the familiarity rating was scored using a 5-point scale (from "never heard it before" to "listen to it regularly").

Several neuro-physiological signals were recorded during music clip watching, namely 32-channel EEG (Biosemi Active II, with 10-20 international electrode placement), ST, GSR, respiration, and BVP. The raw signals were recorded at a 512 Hz sample rate and downsampled offline to 128 Hz. The EEG signals were further bandpass filtered from 4-45 Hz, pre-processed using principal component analysis to remove ocular artifacts, and averaged to a common reference. The interested reader is referred to [66] for more details about the database. The task (Task 1) with this dataset

is to detect the level of arousal, valence, dominance and liking using the newly proposed EEG and PANS-CNS coupling features either alone or in a hybrid aBCI setup.

3.2.2 eNTERFACE 06'

To validate the gains obtained with a hybrid affective BCI, the publicly available 'eNTERFACE 06' database was used [117]. The database is comprised of fNIRS, EEG, GSR, respiration, and Photoplethysmography (PPG) signals collected from 5 subjects all male, and right handed, with age ranging from 22 to 38. Several EEG channels (AF7, AF8, AFz, Fp1, Fp2, Fpz, F7, F6) were removed from an EEG set of 64 electrodes (BIOSEMI Active II, with 10-20 international electrode placement with sampling frequency 1024 Hz) because of the occlusion of the frontal area due to the fNIRS installation. The participants watched 450 images from the IAPS [73] chosen to represent 3 different classes (positive, negative and calm). Subjects then rated their perceived levels of valence and arousal using a 5-point scale. The images were visualized in a random order, following the scheme illustrated in Fig. 3.2.

The experiment is composed of three different sessions. One session had 30 trials with 150 images each for a total of 450 images. The trials were ordered not to have the same elicited emotion in repetition. One trial is composed of the visualization of blocks of 5 images (12.5 seconds) interspersed by dark screens of 10 seconds to have the fNIRS levels back to baseline level. More information regarding the eNTERFACE 06 database can be found in [117]. The task (Task 2) with this dataset is to classify valence and arousal levels using multiple modalities, namely EEG-fNIRS, fNIRS-PANS and EEG-fNIRS-PANS.

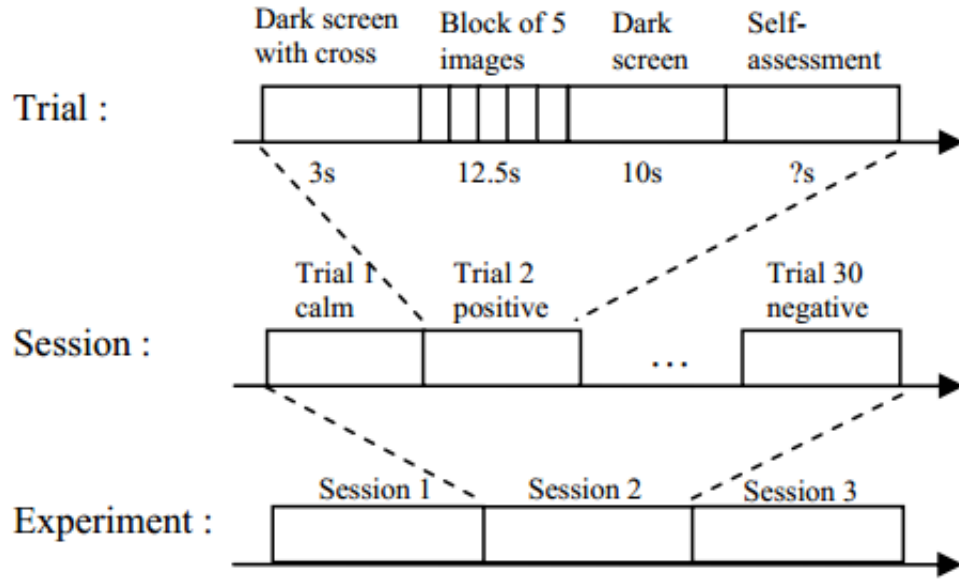


Figure 3.2 – Protocol description of the acquisition of the eNTERFACE 06' database.

3.3 Feature Selection

As mentioned above, a large number of proposed and benchmark features were extracted. More specifically, a total of 184 SF, 20480 AMF, and 1408 PAC features were extracted. For classification purposes, these numbers are large and may lead to classifier overfitting, particularly when feature fusion is performed. In such instances, feature ranking and/or feature selection algorithms are typically used. Recently, several feature selection algorithms were compared on an emotion recognition task [60]. In our experiments, the mRMR algorithm [99] was used, as it showed improved performance when paired with a support vector machine classifier in [131]. The mRMR is a mutual information based algorithm that optimizes two criteria simultaneously: the maximum-relevance criterion (i.e., maximizes the average mutual information between each feature and the target vector) and the minimum-redundancy criterion (i.e., minimizes the average mutual information between two chosen features). In the experiment herein, the algorithm finds near-optimal features using forward selection with the chosen features maximizing the combined max-min criteria. Twenty five percent

of the available data was set aside for feature ranking. The remaining 75% was used for classifier training and testing in a cross-validation scheme, as described next.

3.4 Classification

SVM classifiers have been widely used in bioengineering and in affective state recognition (e.g., [131]). Given their widespread use, a description of the support vector machine approach is not included here and the interested reader is referred to [119] and references therein for more details. SVM classifiers are trained on four different binary classification problems, i.e., detecting low/high valence, low/high arousal, low/high dominance and low/high liking. Since the subjective ratings followed a 9-point scale, we assume values greater or equal to 5 to correspond to high levels; otherwise, a low level was assumed. As mentioned previously, 25% of the available data is set aside for feature ranking. For the remaining 75%, we use a Leave-one-sample-out (LOSO) cross-validation scheme. For our experiments, a Radial Basis Function (RBF) kernel was used and implemented with the Scikit-learn library in Python [98]. Since we are interested in gauging the benefits of the proposed features, and not of the classification schemes, we use the default SVM parameters throughout our experiments (i.e., $\lambda = 1$ and $\gamma_{RBF} = 0.01$).

For Task 2, a variant of the SVM was used, namely the Relevance Vector Machine (RVM) classifier [41, 129, 139] was adopted. The RVM adds an additional Bayesian step that allows for class memberships to be computed, thus resulting in “soft” decisions (as opposed to SVM hard decisions). As already mentioned, in an attempt to improve classification performance, two fusion strategies are explored, namely feature fusion and decision-level fusion. With the DEAP database,

in feature fusion, we explore the combination of the three feature sets (SF, PAC, and AMF) and utilize the top features selected by the mRMR selection algorithm. With classifier decision level fusion, on the other hand, the decisions of the three SVM classifiers trained on the top SF, PAC, and AMF sets were fused using a simple majority voting scheme with equal weights. To test the significance of the attained performances, an independent one-sample t-test against a random voting classifier is used ($p < 0.05$), as suggested in [66]. With the eNTERFACE database, decision fusion was tested when combining the three modalities together (EEG, fNIRS and physiological signals).

Chapter 4

Results and Discussion

This chapter describes the results obtained from Tasks 1 and 2 described previously and provides an in-depth discussion on the obtained findings. The results presented herein have been presented in [16–18].

4.1 Task 1: Experimental Results

Tables 4.1 and 4.2 show the top-20 selected features for the arousal/valence and dominance/liking dimensions, respectively. Feature names listed in the tables are self explanatory. The "ratio" features correspond to the log-ratio ones between the video and baseline periods (see Section 3.1.2). In the SF category, the “AI” features correspond to the asymmetry index between the indicated channels. Due to space constraints the Inter-Hemispheric Amplitude Modulated Interaction (IAMI) features were not included in the ranking table.

Table 4.1 – Selected 20 best features using mRMR algorithm for the arousal and valence dimensions. In the AMI, AMC and AME columns, the top-20 features corresponded to the log-ratio ones. In the SF column, AI indicates the asymmetry index between the indicated channels.

Ranking	Arousal				
	AMI	AMC	AME	PAC	SF
1	ratio_C3_Cz_alpha_m- α	ratio_FC1_CP2_alpha_m- α	ratio_Pz_gamma_m- γ	esc_Fp2	AI_FC5_FC6_theta
2	ratio_T7_O1_theta_m- θ	ratio_T7_P4_alpha_m- α	ratio_AF3_beta_m- β	cfc_AF3_9_Hz	AI_Fp1_Fp2_beta
3	ratio_F4_Cz_beta_m- θ	ratio_F7_Pz_gamma_m- β	ratio_C4_gamma_m- α	esc_C4	AI_P7_P8_alpha
4	ratio_T7_PO3_gamma_m- θ	F3_P3_alpha_m- θ	ratio_FC6_theta_m- θ	esc_CP1	AI_CP1_CP2_beta
5	CP5_P8_theta_m- θ	ratio_C3_P4_alpha_m- α	ratio_T7_beta_m- θ	mod_ind_P3	AI_FC5_FC6_alpha
6	ratio_C3_FC2_beta_m- θ	F7_AF4_beta_m- θ	ratio_O2_alpha_m- θ	esc_PO3	AI_O1_O2_gamma
7	ratio_Fp2_C4_gamma_m- β	ratio_FC1_Pz_alpha_m- θ	ratio_P3_gamma_m- θ	esc_F7	AI_CP5_CP6_theta
8	ratio_CP5_PO4_beta_m- α	ratio_Fp1_T7_gamma_m- γ	ratio_O2_gamma_m- θ	esc_Fz	AI_F3_F4_alpha
9	Oz_Fp2_alpha_m- θ	ratio_F3_P8_gamma_m- α	ratio_T8_gamma_m- γ	cfc_FC2_44_Hz	AI_C3_C4_beta
10	ratio_FC2_Cz_alpha_m- α	Fp1_Cz_alpha_m- θ	ratio_Fp1_beta_m- α	esc_Fp1	Fz_theta
11	ratio_F4_T8_gamma_m- β	ratio_CP2_O2_beta_m- β	ratio_Fz_gamma_m- α	esc_O2	AI_Fp1_Fp2_gamma
12	ratio_FC1_T7_beta_m- α	F7_Fp2_theta_m- θ	ratio_C4_alpha_m- α	esc_P3	AI_FC1_FC2_theta
13	ratio_Fz_O2_beta_m- θ	ratio_F3_PO4_theta_m- θ	ratio_Fz_theta_m- θ	mod_ind_C4	AI_F7_F8_theta
14	ratio_Fp1_FC6_alpha_m- α	ratio_FC1_CP1_alpha_m- α	ratio_FC2_theta_m- θ	esc_Cz	AI_AF3_AF4_beta
15	ratio_T7_Cz_gamma_m- β	ratio_Fp2_P4_gamma_m- θ	ratio_CP5_beta_m- β	cfc_O1_5_Hz	AI_P3_P4_beta
16	ratio_F7_FC1_theta_m- θ	ratio_FC2_C4_alpha_m- α	ratio_Fp1_theta_m- θ	esc_P8	AI_CP1_CP2_gamma
17	ratio_PO3_F4_gamma_m- θ	ratio_F7_F4_gamma_m- α	ratio_CP5_alpha_m- θ	esc_FC5	AI_PO3_PO4_theta
18	ratio_F3_P4_gamma_m- β	AF3_CP6_alpha_m- α	ratio_F4_gamma_m- γ	mod_ind_FC1	AI_T7_T8_gamma
19	FC2_O2_alpha_m- θ	Oz_Cz_theta_m- θ	ratio_F3_gamma_m- α	esc_FC1	AI_O1_O2_beta
20	ratio_CP5_CP2_alpha_m- θ	ratio_AF3_PO3_alpha_m- θ	ratio_CP1_beta_m- θ	esc_AF4	AI_AF3_AF4_gamma
Ranking	Valence				
	AMI	AMC	AME	PAC	SF
1	AF4_C4_theta_m- θ	ratio_CP5_Cz_gamma_m- β	ratio_Fp1_beta_m- β	esc_Fp2	AI_CP1_CP2_theta
2	ratio_FC5_T8_gamma_m- β	ratio_CP5_P3_beta_m- β	ratio_PO4_gamma_m- θ	cfc_AF3_9_Hz	AI_O1_O2_beta
3	F3_F7_alpha_m- α	AF3_FC6_theta_m- θ	ratio_F8_gamma_m- γ	esc_CP2	AI_Fp1_Fp2_gamma
4	ratio_FC6_P4_gamma_m- θ	FC5_F8_theta_m- θ	ratio_T8_alpha_m- α	cfc_FP2_23_Hz	AI_PO3_PO4_beta
5	ratio_Fp2_P8_beta_m- β	CP5_PO4_alpha_m- θ	ratio_AF4_beta_m- θ	esc_T8	O1_theta
6	ratio_F8_T8_alpha_m- θ	ratio_CP6_O2_beta_m- α	ratio_O1_alpha_m- α	mod_ind_C4	AI_FC1_FC2_alpha
7	ratio_AF3_O1_beta_m- α	Fp2_AF4_theta_m- θ	ratio_CP1_theta_m- θ	esc_FC2	AI_P3_P4_alpha
8	ratio_Cz_P4_theta_m- θ	Pz_FC2_beta_m- α	ratio_PO4_alpha_m- θ	esc_PO3	AI_F3_F4_theta
9	ratio_FC1_CP1_beta_m- β	ratio_FC1_T7_gamma_m- γ	ratio_F3_beta_m- θ	esc_F7	AI_P3_P4_theta
10	CP6_O2_beta_m- θ	Cz_CP2_alpha_m- α	ratio_FC2_beta_m- β	cfc_Fp1_4_Hz	AI_T7_T8_beta
11	ratio_CP1_PO3_gamma_m- β	ratio_F7_F4_beta_m- α	ratio_FC6_alpha_m- α	esc_O2	FC2_beta
12	FC2_Cz_gamma_m- β	ratio_C4_O2_theta_m- θ	ratio_P3_beta_m- θ	esc_T7	AI_FC5_FC6_beta
13	ratio_C3_P7_beta_m- θ	Fp1_F8_alpha_m- θ	ratio_FC5_beta_m- θ	esc_C4	AI_F7_F8_alpha
14	F7_CP1_beta_m- β	ratio_T7_CP5_beta_m- α	ratio_C3_alpha_m- θ	mod_ind_Cz	AI_AF3_AF4_theta
15	ratio_F3_Cz_beta_m- α	P3_O1_alpha_m- α	ratio_AF3_beta_m- α	esc_CP1_5_Hz	AI_FC5_FC6_gamma
16	ratio_FC5_CP2_gamma_m- α	ratio_Fp1_FC5_gamma_m- θ	CP2_alpha_m- θ	esc_FC5	AI_Fp1_Fp2_alpha
17	ratio_AF3_F4_gamma_m- β	ratio_O1_FC6_alpha_m- α	ratio_P8_gamma_m- θ	esc_FP1	AI_PO3_PO4_gamma
18	ratio_Fp2_P8_alpha_m- α	AF3_PO4_theta_m- θ	ratio_F4_beta_m- α	cfc_PO4_6_Hz	AI_CP5_CP6_alpha
19	Fz_F8_theta_m- θ	ratio_Fp2_CP6_beta_m- θ	ratio_FP2_theta_m- θ	esc_FC1	AI_F7_F8_theta
20	ratio_Pz_CP2_alpha_m- α	Pz_CP6_alpha_m- θ	ratio_CP6_theta_m- θ	mod_ind_FC1	AI_FC1_FC2_theta

Figures 4.1(a)-(d) depict the accuracy of the SVM classifiers for the arousal, valence, dominance, and liking affective dimensions, respectively, as a function of number of features used. Accuracy plots are shown for classifiers trained with only the proposed IAMI (Inter-Hemispheric Amplitude Modulated Interaction) features, only the baseline, as well as the combined feature set. As can be seen, the proposed features outperform the baseline ones for all four affective dimensions, but

Table 4.2 – Selected 20 best features using mRMR algorithm for the dominance and liking dimensions. In the AMI, AMC and AME columns, the top-20 features corresponded to the log-ratio ones. In the SF column, AI indicates the asymmetry index between the indicated channels.

Ranking	Dominance				
	AMI	AMC	AME	PAC	SF
1	ratio_T7_T8_beta_m- β	ratio_T8_P8_gamma_m- θ	FC6_gamma_m- θ	cfc_F7_27_Hz	AI_P7_P8_alpha
2	ratio_Fp1_Fp2_alpha_m- α	ratio_AF3_PO3_theta_m- θ	ratio_FC5_gamma_m- γ	esc_FC6_9_Hz	AI_FC5_FC6_theta
3	P3_P8_beta_m- θ	AF3_P3_beta_m- α	ratio_CP5_beta_m- α	esc_O2	AI_T7_T8_alpha
4	ratio_C3_Fz_gamma_m- θ	ratio_F8_FC6_alpha_m- α	C4_theta_m- θ	mod_ind_P3	AI_F7_F8_alpha
5	T7_CP5_gamma_m- α	ratio_P3_T8_theta_m- θ	ratio_CP6_beta_m- θ	esc_FP1	AI_FC5_FC6_alpha
6	ratio_Oz_Fp2_beta_m- α	ratio_Fz_F8_gamma_m- β	ratio_F7_alpha_m- θ	esc_FC1	O1_FC1_theta
7	ratio_FC1_CP1_gamma_m- θ	ratio_Pz_PO4_theta_m- θ	ratio_F3_gamma_m- θ	esc_F7	AI_FP1_FP2_beta
8	ratio_F7_F4_theta_m- θ	FC5_FC1_alpha_m- α	ratio_P4_beta_m- β	esc_P3	AI_O1_O2_theta
9	AF3_C4_theta_m- θ	ratio_O1_Pz_gamma_m- θ	ratio_Pz_gamma_m- β	cfc_Fp2_12_Hz	AI_C3_C4_beta
10	ratio_C3_CP6_theta_m- θ	ratio_Fp1_P4_gamma_m- γ	ratio_AF3_gamma_m- α	esc_C4	C4_beta
11	ratio_CP5_T8_beta_m- β	ratio_CP5_Pz_beta_m- α	ratio_F4_theta_m- θ	mod_ind_C4	AI_FC1_FC2_gamma
12	ratio_P7_AF4_beta_m- θ	ratio_CP1_Pz_beta_m- α	ratio_F8_beta_m- β	esc_P7	AI_O1_O2_alpha
13	ratio_P3_FC2_beta_m- θ	ratio_FC1_F4_beta_m- α	ratio_Fp2_theta_m- θ	esc_Oz	AI_CP1_CP2_beta
14	ratio_FC5_CP1_theta_m- θ	FC1_P8_beta_m- β	ratio_P8_gamma_m- α	cfc_O2_4_Hz	AI_P3_P4_gamma
15	Fp2_F4_alpha_m- θ	ratio_Fp1_F8_beta_m- α	ratio_F4_gamma_m- θ	esc_Cz	AI_F3_F4_beta
16	ratio_Oz_FC6_beta_m- α	Pz_T8_theta_m- θ	ratio_P7_beta_m- β	esc_AF4	AI_CP5_CP6_theta
17	ratio_C4_P4_theta_m- θ	ratio_AF4_FC2_theta_m- θ	ratio_T7_theta_m- θ	esc_P8	AI_CP1_CP2_gamma
18	ratio_AF4_P4_alpha_m- α	ratio_Fz_Cz_theta_m- θ	AF3_beta_m- θ	cfc_FC2_8_Hz	AI_AF3_AF4_gamma
19	ratio_AF3_FC2_theta_m- θ	ratio_Fp1_CP6_theta_m- θ	ratio_PO4_alpha_m- α	esc_Fz	AI_FC5_FC6_gamma
20	FC1_CP2_beta_m- β	ratio_C3_P3_alpha_m- θ	ratio_O2_gamma_m- γ	mod_ind_FC5	AI_F7_F8_theta
Liking					
1	CP5_C4_gamma_m- α	ratio_F7_F4_alpha_m- α	ratio_F8_theta_m- θ	cfc_FC1_21_Hz	AI_P7_P8_theta
2	ratio_Oz_Pz_beta_m- α	ratio_AF3_Fz_alpha_m- α	Pz_theta_m- θ	esc_C3	AI_T7_T8_gamma
3	ratio_F4_T8_beta_m- β	F3_O1_theta_m- θ	ratio_F8_gamma_m- γ	esc_F4	F8_theta
4	ratio_CP1_Fp2_alpha_m- α	ratio_P3_T8_gamma_m- β	ratio_F3_beta_m- θ	mod_ind_C4	AI_PO3_PO4_beta
5	ratio_F8_P8_alpha_m- α	ratio_O1_P4_beta_m- α	ratio_O1_theta_m- θ	esc_P8	AI_CP1_CP2_gamma
6	ratio_Pz_CP2_alpha_m- α	Pz_P8_beta_m- β	ratio_AF3_theta_m- θ	cfc_P3_6_Hz	AI_FC1_FC2_alpha
7	FC6_O2_gamma_m- θ	ratio_FC1_F8_alpha_m- θ	ratio_PO4_alpha_m- α	esc_AF3	AI_FC5_FC6_theta
8	ratio_T7_FC2_gamma_m- γ	ratio_C3_Pz_gamma_m- γ	ratio_Fz_beta_m- β	esc_O2	AI_Fp1_Fp2_theta
9	ratio_Oz_CP2_beta_m- α	ratio_F3_FC5_beta_m- β	ratio_C4_gamma_m- γ	esc_P3	AI_O1_O2_alpha
10	ratio_CP1_O1_beta_m- α	T7_T8_theta_m- θ	ratio_Cz_theta_m- θ	cfc_F3_25_Hz	AI_F7_F8_gamma
11	Pz_FC2_alpha_m- θ	ratio_F4_P4_alpha_m- θ	ratio_CP5_gamma_m- α	esc_C4	AI_AF3_AF4_alpha
12	CP5_F8_gamma_m- γ	ratio_FC1_Cz_gamma_m- α	ratio_Fp1_beta_m- β	esc_T7	AI_O1_O2_theta
13	ratio_AF3_F7_gamma_m- β	ratio_F7_Fp2_beta_m- α	ratio_Pz_beta_m- α	mod_ind_FC1	AI_P7_P8_gamma
14	ratio_Pz_O2_beta_m- α	ratio_P7_Pz_gamma_m- β	ratio_AF4_beta_m- θ	esc_CP1	AI_F7_F8_beta
15	O1_Cz_beta_m- β	ratio_P4_O2_gamma_m- γ	ratio_Fp1_alpha_m- θ	cfc_FC5_4_Hz	AI_P3_P4_theta
16	ratio_AF3_F4_beta_m- β	F3_P4_theta_m- θ	F8_alpha_m- α	esc_Fp2	F3_beta
17	ratio_AF3_C4_theta_m- θ	ratio_Oz_Fp2_gamma_m- β	ratio_F4_theta_m- θ	esc_FC2	AI_AF3_AF4_theta
18	ratio_FC1_CP5_theta_m- θ	ratio_T8_P4_alpha_m- θ	ratio_PO3_beta_m- β	mod_ind_O2	AI_P7_P8_beta
19	AF3_PO3_alpha_m- θ	T7_PO3_beta_m- β	ratio_C3_theta_m- θ	esc_F7	AI_CP5_CP6_theta
20	P3_Oz_beta_m- α	ratio_Fp1_F8_theta_m- θ	ratio_AF3_beta_m- β	cfc_Pz_25_Hz	AI_P3_P4_alpha

with higher gains seen in the valence and liking dimensions. Interestingly, these were the hardest dimensions to estimate with conventional EEG features in [66].

The subplots in Figure 4.2 (a)-(d), in turn, report the accuracy of the arousal, valence, dominance and liking SVM classifiers, respectively, as a function of number of features used for the three features sets (AMF, PAC, SF). Within this scenario, features were added one-by-one based on the ranking

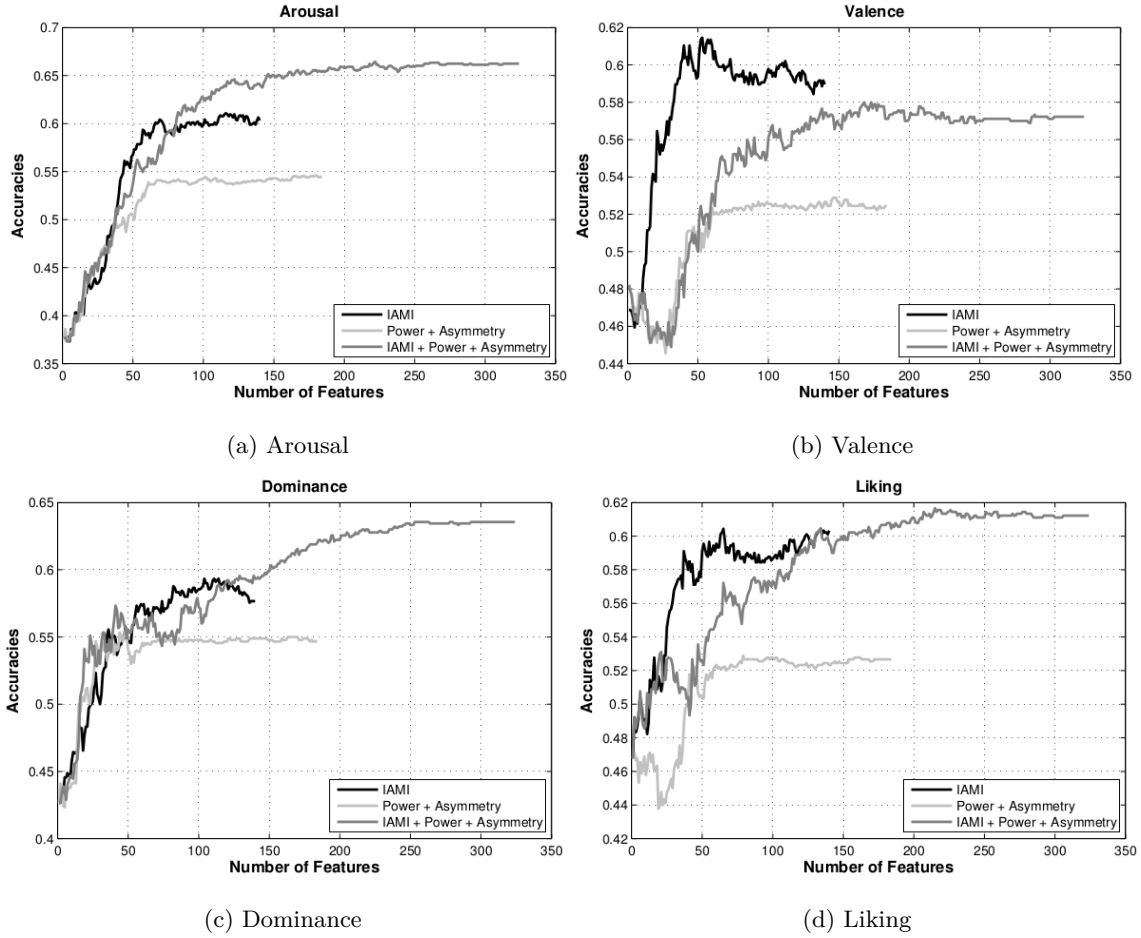


Figure 4.1 – Accuracy versus number of features for the baseline, IAMI, and fused feature sets for the (a) arousal, (b) valence, (c) dominance, and (d) liking affective dimensions.

provided by the mRMR algorithm. A maximum number of features equal to 184 was used to coincide with the number for the SF benchmark feature set. As can be seen, the AMF set outperforms the two other sets across all affective dimensions.

Table 4.3 further reports the highest accuracy results achieved with the individual features sets, as well as the number of features required to achieve such results. Results followed by an asterisk suggest significantly higher values ($p < 0.05$) than those achieved with a random voting classifier, as in [66]. The number within parentheses show the relative improvement, in percentage, relative to the SF baseline set. As can be seen, the proposed amplitude modulation feature sets achieve results

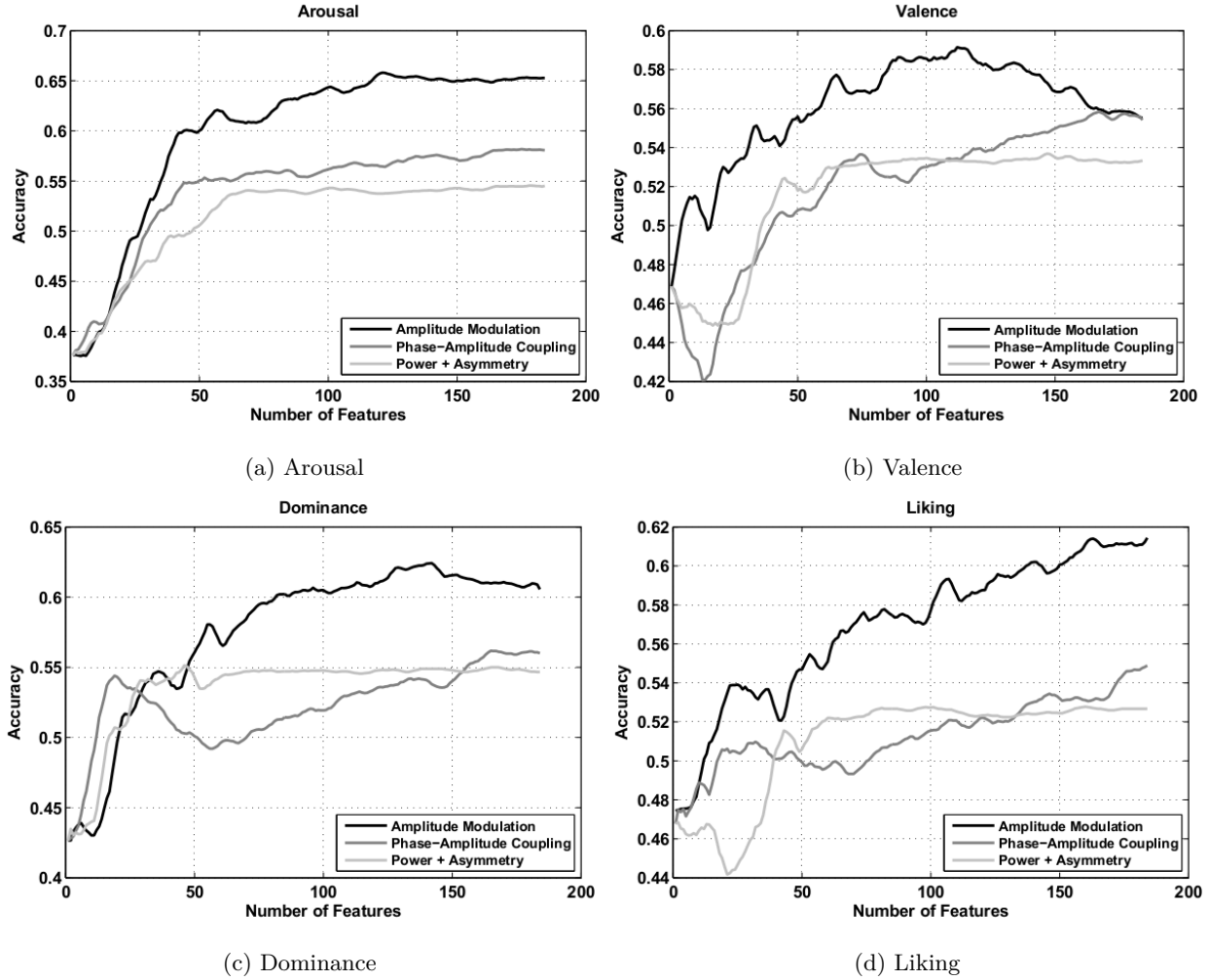


Figure 4.2 – Accuracy versus number of features for the AMF, PAC, and SF feature sets for the (a) arousal, (b) valence, (c) dominance, and (d) liking affective dimensions.

significantly higher than chance for the four dimensions. For the valence and dominance dimensions, the proposed AMF features improve over the baseline by 9% and 14%, respectively.

The last column of Table 4.3 shows the results obtained with feature-level fusion. As can be seen, feature fusion did not result in performances higher than those obtained with the proposed AMF set. Table 4.4, in turn, shows the performances achieved with two decision level fusion schemes: AMI/AMC/AME and AMF/PAC/SF. While decision level fusion of AMF/PAC/SF classifiers did not lead to gains over simple feature fusion, decision level fusion of the AMI/AMC/AME sets

Table 4.3 – Performance comparison of SVM classifiers for different feature sets and feature-level fusion along with the required number of features needed to achieve such results. Asterisks indicate accuracy significantly higher than chance achieved with a random voting classifier ($p < 0.05$). Values within parentheses indicate the relative improvement, in percentage, over the SF baseline set.

Subjective Dimensions	Metric	AMI	AMC	AME	AMF	PAC	SF	IAMI	AMF + SF + PAC
Arousal	Accuracy No. features	0.64* (16) 177	0.67* (22) 125	0.66* (20) 183	0.66* (18) 122	0.58* (5) 168	0.55 180	0.61 (11) 115	0.66* (20) 122
Valence	Accuracy No. features	0.56* (4) 87	0.58* (7) 167	0.59* (9) 102	0.59* (9) 114	0.56* (4) 165	0.54 146	0.61 (11) 53	0.59* (9) 115
Dominance	Accuracy No. features	0.62* (13) 173	0.60* (9) 169	0.61* (13) 90	0.63* (14) 141	0.56 (2) 166	0.55 48	0.59 (7) 82	0.62* (13) 132
Liking	Accuracy No. features	0.60* (2) 167	0.60* (2) 159	0.60* (2) 176	0.62* (5) 163	0.55 177	0.59 79	0.60 (13) 65	0.60* (2) 110

Table 4.4 – Performance comparison of different decision-level fusion schemes along with the required number of features needed to achieve such results. Asterisks indicate accuracy significantly higher than chance achieved with a random voting classifier ($p < 0.05$). Values within parentheses indicate the relative improvement, in percentage, over the SF baseline set.

Subjective Dimensions	Metric	AMI + AMC + AME	AMF + PAC + SF
Arousal	Accuracy No. features	0.67* (22) 125	0.64* (16) 152
Valence	Accuracy No. features	0.59* (9) 40	0.57* (6) 110
Dominance	Accuracy No. features	0.63* (14) 92	0.59* (7) 141
Liking	Accuracy No. features	0.60* (2) 167	0.57* 184

did improve the accuracy of the classifiers for the dominance and liking dimensions. Overall, with AMI/AMC/AME decision level fusion, gains over the SF baseline of 22%, 9%, 14%, and 2% were observed for arousal, valence, dominance, and liking dimensions, respectively.

4.2 Task 1: Discussion

In-depth analysis of the top-100 selected features (Tables 4.1 and 4.2) has been conducted in order to understand the most relevant regions of the brain and the frequency bands mostly involved.

Within the SF feature set, the asymmetry indices prevailed over spectral power features. In fact, for each of the four affective dimensions, all the 54 AI features were present in the first 60 selected features. Moreover, analyzing the first 20 features, they are largely detected from the frontal region with a consistent activity in the theta band. A relevant expression of emotion in brain signals using asymmetry has been described in the literature by [24] who first developed a model that relates frontal lobes of the brain with emotions. Other studies [19, 22, 44] have also confirmed the aforementioned asymmetry concept finding asymmetrical effects in bands including theta.

Previous work [70] on PAC, in turn, showed the coupling between EEG and GSR to be relevant in emotion classification. In our study, we found a predominance of the ESC feature over the other two (CFC and ModI). Indeed ESC features are ranked in the first 50 features for each emotional dimension. Considering only the top 20 features, the most involved region of the brain was the frontal region, as also observed with the asymmetry features. Moreover, regarding the CFC, the frequency band that resulted in the most significant features was theta.

Regarding the proposed AMF set, two different scenarios arose. For the arousal and valence affective states, the predominant feature was the AMC corresponding to 58% of the top-selected features. For the liking and dominance dimensions, in turn, the AMI feature corresponded to 57% of the top-selected features. The AMC feature measures linear amplitude modulation coupling between different brain regions, whereas AMI may measure potentially non-linear effects. These results suggest different relationships based on varying emotional primitives. Moreover, approximately 80% of the features were selected from the log-ratio between the video and baseline periods, thus suggesting the importance of feature normalization.

A more in-depth analysis of the proposed AMC features showed the importance of the $s(\alpha, m - \theta)$ pattern (where channel subscript k was omitted for brevity) for arousal and valence states, thus in line with the work of [63], which related these bands to states of internalized attention and positive emotional experience [1]). Overall, 34% of the top-selected AMC features selected corresponded to alpha band amplitude modulation coupling. Such alpha/theta cross-frequency synchronization has also been previously related to memory usage [15]. To corroborate this hypothesis, the correlation between the proposed features derived from the $s(\alpha, m - \theta)$ patterns and the subjective ‘familiarity’ ratings reported by the participants was computed. The majority of the features showed to be significantly correlated ($\geq 0.3, p < 0.05$) with the familiarity rating, thus suggesting memory may have indeed played an effect on the elicited affective states.

Regarding the AMI features, in turn, a predominance of gamma and beta band (30% and 25%, respectively) coupling was found in the top-selected features, thus inline with results from [60]. It was previously demonstrated that these two bands were able to discriminate between liking and disliking judgements [43]. By analyzing their amplitude modulation cross-frequency coupling via the proposed features, improved results were found, thus showing the importance of EEG amplitude modulation coupling for affective state recognition.

Overall, for the entire AMF set, channels that involved the frontal region provided the most relevant features, thus confirming the importance of the frontal region for affective state recognition [88]. The frontal channels were mostly connected to central and the parietal regions with limited connections with the temporal and occipital regions arising in the top-selected features.

As seen from Figures 4.1, the proposed IAMI features clearly outperform the baseline ones for all four classification tasks. The performance of classifiers trained on either the proposed or baseline

feature sets tended to stabilize at around 60 features, but with the proposed feature set achieving substantially higher accuracy. This was particularly true for the valence and liking dimensions, which are typically the hardest ones to classify using conventional EEG features [66]. In fact, from Table 4.3 for the valence dimension, roughly one third of the features were needed to outperform the baseline features (53 vs 146 features). Such findings highlight the benefits of the compact representation of the proposed features for reliable affective state characterization. An in-depth analysis of the top-60 selected IAMI features showed that roughly half came from alpha-band spectro-temporal patterns. For the valence and arousal dimensions, the $A(\alpha, m - \theta)$ pattern was the most prominent. Previous work has linked alpha-theta coupling to memory [15], which in turn has been shown to be modulated by valence and arousal [63], thus suggesting that internal (affective) references may have played a key role during the experiment. The proposed features seem to be able to characterize such memory effects, unlike conventional features, thus corroborating their complementarity.

With the fused feature set, in turn, performance stabilization occurred once approximately 150 features were used to train the classifiers. Careful analysis of these top-150 features showed that for the four dimensions, roughly 66% of the features belonged to the IAMI set, with the majority of the remaining features belonging to the EEG asymmetry index set. Channel pairs that provided the most relevant features were located in the frontal (Fp1-Fp2 and AF3-AF4) and parietal regions (P7-P8 and P3-P4), thus corroborating the importance of frontal asymmetry for affective state recognition, particularly those involving high arousal states [88]. From Table 4.3, it can be seen that overall the combined feature set resulted in the highest relative gains compared to the baseline, with the exception of the valence affective dimension which saw the IAMI features achieving the best performance. Such findings may be an artifact of the classifier design used here. Since the goal of

the study was to test the effectiveness of the proposed features at classifying four different affective dimensions (or primitives), only default SVM classifier parameters were utilized.

Accuracy plots in Figure 4.2 show that the proposed amplitude modulation feature set outperform both benchmark feature sets (phase-amplitude coupling and spectral features) for classifying users' emotional states elicited by different music video clips. In fact, with feature fusion it was observed that approximately 97 % of the top-selected features came from the AMF set, with a peak of 99% for the arousal and valence dimensions, thus corroborating our hypothesis that cross frequency coupling is a good predictor of affective states. Notwithstanding, fusing the different features led to some improvement, particularly for the arousal dimension where a 20% improvement was seen relative to spectral features when fusing AMF, SF, and PAC features. Decision level fusion with classifiers trained on these three separate modalities, on the other hand, resulted in only 16% relative improvement. Nonetheless, decision level fusion between classifiers trained with the three proposed amplitude modulation feature sets (i.e., AMI, AMC, AME) did result in further improvements, particularly for the arousal dimension, thus suggesting also their complementarity. Decision level fusion was shown previously to be a useful tool for affective state recognition [66].

4.3 Task 2: Experimental Results and Discussion

Tables 4.5 and 4.6 show the level of accuracy and F1 score obtained with single and multi-modality using fNIRS. In particular Table 4.5 shows the results with the first 20 selected features. Every time that fNIRS was added, there was an increase in accuracy. The maximum was reached with an improvement of 8% from the initial value when combining all the three modalities together

Table 4.5 – Performance comparison of different feature-level fusion schemes considering the contribution of fNIRS to three different datasets (EEG, PHY and EEG + PHY) for the two affective states arousal and valence.

	Valence		Arousal	
	Accuracy	F1 Score	Accuracy	F1 Score
EEG	0.53	0.53	0.44	0.43
EEG + NIRS *	0.56	0.56	0.48	0.46
PHY	0.45	0.45	0.44	0.43
PHY + NIRS *	0.51	0.51	0.47	0.46
EEG + PHY	0.53	0.53	0.44	0.43
EEG + PHY + NIRS *	0.57	0.56	0.49	0.47

Table 4.6 – Performance comparison of different feature-level fusion schemes considering the contribution of fNIRS to three different datasets (EEG, PHY and EEG + PHY) for the two affective states arousal and valence.

	Valence			Arousal		
	N of Features Best Performance	Increase in Accuracy	Increase in F1 Score	N of Features Best Performance	Increase in Accuracy	Increase in F1 Score
EEG + NIRS	67	3%	3%	63	6%	6%
PHY + NIRS	9	2%	2%	13	7%	6%
EEG + PHY + NIRS	82	0%	0%	39	6%	6%

(EEG, PHY and fNIRS) using a majority voting decision fusion technique. Feature fusion was also tested calculating the accuracy adding one feature each time. The best improvement (Table 4.6) has been reached when adding PHY and NIRS with a 7% performance gain. fNIRS adds important information for affective computing tasks, demonstrating the importance of multimodality [87]. Moreover, there is a high contribution in number of features coming from fNIRS when using feature fusion. Indeed, when combining EEG and fNIRS, the latter contributes 20% of the top features, whereas when combining PHY and fNIRS, it contributed towards 80% of the top features.

Chapter 5

Conclusion

5.1 Summary of the Research Outcomes

In this work, we investigated the use of several modalities, namely EEG, fNIRS, BVP, Respiration, GSR and ST, for the binary classification of four different emotional states (valence, arousal, dominance and liking). We tested new features and demonstrated the impact of multimodal approaches on the classification performance.

New amplitude-amplitude cross-frequency modulation coupling features for EEG with three different variations (AME, AMC and AMI) were proposed for automated affective state recognition during music video clip watching. Moreover, several PANS-CSN coupling features (GSR-EEG) either alone or in a hybrid aBCI setup have been integrated in the study in order to study common information between different signals.

Experimental results with the publicly-available DEAP database showed the amplitude modulated features outperforming benchmark measures based on spectral power and GSR-EEG phase-amplitude coupling features by as much as 20%. Classifier decision-level fusion was shown to perform better than feature level fusion, with the AMF classifier being responsible for over 67% of the correct decisions in the arousal and dominance dimensions. Such findings suggest the importance of the proposed features for affective state recognition and signal the importance of EEG amplitude modulation for affective tagging of music video clips and content. In turn GSR-EEG phase-amplitude coupling features result in a more compact and accurate representation for affective state recognition compared to the signals treated separately.

Moreover, a new pattern in multimodality has been experimented, introducing the use of hemodynamic information for affective computing during the visualization of images. Experimental results with the 'eNTERFACE 06 database show that including fNIRS with a hybrid aBCI can lead to improvements up to 7% in emotions classification accuracy.

Our proposed feature analysis approach made it possible to delve deeper into the classification results, and shed light on the role of different neurophysiological modalities toward more efficient BCIs for affective state recognition.

5.2 Future Research and Directions

To improve the performance of aBCI and affective hBCIs tasks, supplementary steps could be undertaken. First, the binary classification tasks performed here could be replaced by a regression task where the actual SAM scale ratings could be predicted. Second, classification performance

could be improved by selecting optimal classification models by tuning hyperparameters, comparing different classifiers and explore different fusion strategies (e.g. with statistical models, such as Hidden Markov model) [33]. A third improvement would be to increase the number of participants, as initialized by [39], in order to better validate our implemented system.

Following these steps, an important future research direction could be the realization of real time online BCIs based on the newly implemented feature sets. The restrictions induced by a real-time implementation (setup time computing efficacy, robustness to noise, etc.) will have to be overcome to yield a truly usable brain-computer interface.

Bibliography

- [1] Aftanas L & Golocheikine S (2001). Human anterior and frontal midline theta and lower alpha reflect emotionally positive state and internalized attention: high-resolution EEG investigation of meditation. *Neuroscience letters*, 310(1):57–60.
- [2] Allison BZ, Brunner C, Kaiser V, Müller-Putz GR, Neuper C & Pfurtscheller G (2010). Toward a hybrid brain–computer interface based on imagined movement and visual attention. *Journal of neural engineering*, 7(2):026007.
- [3] Arnulfo G, Hirvonen J, Nobili L, Palva S & Palva JM (2015). Phase and amplitude correlations in resting-state activity in human stereotactical EEG recordings. *NeuroImage*, pages 112:114–27.
- [4] Backs RW, da Silva SP & Han K (2005). A comparison of younger and older adults’ self-assessment manikin ratings of affective pictures. *Experimental aging research*, 31(4):421–440.
- [5] Balconi M & Lucchiari C (2008). Consciousness and arousal effects on emotional face processing as revealed by brain oscillations. A gamma band analysis. *International Journal of Psychophysiology*, 67(1):41–46.
- [6] Banville H & Falk TH (2016). Recent advances and open challenges in hybrid brain-computer interfacing: A technological review. *Brain-Computer Interfaces*. Submitted for publication.
- [Bartsch] Bartsch F. Cortical communication in emotional perception: An analysis of high frequency oscillations and directed connectivity using electrocorticography.
- [8] Becker AE & Hamburg P (1996). Culture, the media, and eating disorders. *Harvard Review of Psychiatry*, 4(3):163–167.
- [9] Bekkedal MY, Rossi J & Panksepp J (2011). Human brain eeg indices of emotions: delineating responses to affective vocalizations by measuring frontal theta event-related synchronization. *Neuroscience & Biobehavioral Reviews*, 35(9):1959–1970.
- [10] Bersak D, McDarby G, Augenblick N, McDarby P, McDonnell D, McDonald B & Karkun R (2001). Intelligent biofeedback using an immersive competitive environment. Paper at the Designing Ubiquitous Computing Games Workshop at UbiComp.
- [11] Boashash B, Lovell B & Kootsookos P (1989). Time-frequency signal analysis and instantaneous frequency estimation: methodology, relationships and implementations. *Circuits and Systems, 1989., IEEE International Symposium on*, IEEE, pages 1237–1242.

- [12] Bos DO, Reuderink B, van de Laar B, Gurkok H, Muhl C, Poel M, Heylen D & Nijholt A (2010). *Human-computer interaction for BCI games: Usability and user experience*. Cyberworlds (CW), 2010 International Conference on, pages 277–281.
- [13] Bradley MM & Lang PJ (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1):49–59.
- [14] Büchel C, Morris J, Dolan RJ & Friston KJ (1998). Brain systems mediating aversive conditioning: an event-related fmri study. *Neuron*, 20(5):947–957.
- [15] Chik D (2013). Theta-alpha cross-frequency synchronization facilitates working memory control—a modeling study. *SpringerPlus*, 2(1):14.
- [16] Clerico A & Falk TH (2015). Towards affective hybrid bci based on fNIRS , EEG and peripheral physiological signals.
- [17] Clerico A, Gupta R & Falk T (2015). Mutual information between inter-hemispheric EEG spectro-temporal patterns: A new feature for automated affect recognition. *IEEE Neural Engineering Conference*, pages 914–917.
- [18] Clerico A, Gupta R & Falk TH (2016). Electroencephalography amplitude modulation analysis for improved automated affective tagging of music video clips.
- [19] Coan JA & Allen JJ (2004). Frontal EEG asymmetry as a moderator and mediator of emotion. *Biological psychology*, 67(1):7–50.
- [20] Cohen I, Sebe N, Garg A, Chen LS & Huang TS (2003). Facial expression recognition from video sequences: temporal and static modeling. *Computer Vision and image understanding*, 91(1):160–187.
- [21] Daly I, Malik A, Hwang F, Roesch E, Weaver J, Kirke A, Williams D, Miranda E & Nasuto SJ (2014). Neural correlates of emotional responses to music: an eeg study. *Neuroscience letters*, 573:52–57.
- [22] Davidson RJ (2004). What does the prefrontal cortex do in affect: perspectives on frontal eeg asymmetry research. *Biological psychology*, 67(1):219–234.
- [23] Davidson RJ, Schwartz GE, Pugash E & Bromfield E (1976). Sex differences in patterns of eeg asymmetry. *Biological Psychology*, 4(2):119–137.
- [24] Davidson RJ, Schwartz GE, Saron C, Bennett J & Goleman DJ (1979). Frontal versus parietal eeg asymmetry during positive and negative affect. *Psychophysiology*, volume 16, pages 202–203.
- [25] Davidson RJ & Tomarken AJ (1989). Laterality and emotion: An electrophysiological approach. *Handbook of neuropsychology*, 3:419–441.
- [26] De Martino B, Kumaran D, Seymour B & Dolan RJ (2006). Frames, biases, and rational decision-making in the human brain. *Science*, 313(5787):684–687.
- [27] De Munck J, Goncalves S, Huijboom L, Kuijer J, Pouwels P, Heethaar R & da Silva FL (2007). The hemodynamic response of the alpha rhythm: an eeg/fmri study. *Neuroimage*, 35(3):1142–1151.

- [28] Delplanque S, Grandjean D, Chrea C, Aymard L, Cayeux I, Le Calve B, Velazco MI, Scherer KR & Sander D (2008). Emotional processing of odors: evidence for a nonlinear relation between pleasantness and familiarity evaluations. *Chemical Senses*, 33(5):469–479.
- [29] Desmet PMA (2003). *Measuring emotion: development and application of an instrument to measure emotional responses to products*. Monk AF, Overbeeke K & Wright PC (éditeurs). Kluwer Academic Publishers, pages 111–123.
- [30] Di Muro F & Murray KB (2012). An arousal regulation explanation of mood effects on consumer choice. *Journal of Consumer Research*, 39(3):574–584.
- [31] Ekman P, Friesen WV, O’Sullivan M, Chan A, Diacoyanni-Tarlatzis I, Heider K, Krause R, LeCompte WA, Pitcairn T, Ricci-Bitti PE *et al.* (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of personality and social psychology*, 53(4):712.
- [32] Ekman P, Levenson RW & Friesen WV (1983). Autonomic nervous system activity distinguishes among emotions. *Science*, 221(4616):1208–1210.
- [33] Falk TH, Guirgis M, Power S & Chau T (2011). Taking nirs-bcis outside the lab: towards achieving robustness against environment noise. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 19(2):136–146.
- [34] Fazli S, Mehnert J, Steinbrink J, Curio G, Villringer A, Müller KR & Blankertz B (2012). Enhanced performance by a hybrid nirs-eeG brain computer interface. *Neuroimage*, 59(1):519–529.
- [35] Forgas JP & Smith CA (2003). Affect and emotion. *The Sage handbook of social psychology*, pages 161–189.
- [36] Fox E (2008). *Emotion science cognitive and neuroscientific approaches to understanding human emotions*. Palgrave Macmillan.
- [37] Fraga FJ, Falk TH, Kanda PA & Anghinah R (2013). Characterizing alzheimers disease severity via resting-awake eeg amplitude modulation analysis. *PloS one*, 8(8):e72240.
- [38] Gabor D (1946). Theory of communication. part 1: The analysis of information. *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, 93(26):429–441.
- [39] Gupta R, Banville HJ & Falk TH (2015). Physyqx: A database for physiological evaluation of synthesised speech quality-of-experience. *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 2015*, IEEE, pages 1–5.
- [40] Gupta R & Falk T (2015). Affective state characterization based on electroencephalography graph-theoretic features. *IEEE Neural Engineering Conference*, pages 577–580.
- [41] Gupta R, Falk TH *et al.* (2016). Relevance vector classifier decision fusion and eeg graph-theoretic features for automatic affective state characterization. *Neurocomputing*, 174:875–884.
- [42] Gürkök H & Nijholt A (2012). Brain-computer interfaces for multimodal interaction: A survey and principles. *International Journal of Human-Computer Interaction*, 28(5):292–307.

- [43] Hadjidimitriou SK & Hadjileontiadis LJ (2012). Toward an eeg-based recognition of music liking using time-frequency analysis. *IEEE Transactions on Biomedical Engineering*, 59(12):3498–3510.
- [44] Hagemann D, Naumann E, Lürken A, Becker G, Maier S & Bartussek D (1999). Eeg asymmetry, dispositional mood and personality. *Personality and Individual Differences*, 27(3):541–568.
- [45] Hamm J, Pinkham A, Gur RC, Verma R & Kohler CG (2014). Dimensional information-theoretic measurement of facial emotion expressions in schizophrenia. *Schizophrenia research and treatment*, 2014:Article ID 243907, 10 pages.
- [46] Harman DW & Ray WJ (1977). Hemispheric activity during affective verbal stimuli: An EEG study. *Neuropsychologia*, 15(3):457–460.
- [47] Havlena WJ & Holbrook MB (1986). The varieties of consumption experience: comparing two typologies of emotion in consumer behavior. *Journal of consumer research*, pages 394–404.
- [48] Haykin S (1999). *Adaptive filters*. volume 6. Citeseer.
- [49] Heger D, Herff C, Putze F, Mutter R & Schultz T (2014). Continuous affective states recognition using functional near infrared spectroscopy. *Brain-Computer Interfaces*, 1(2):113–125.
- [50] Heger D, Putze F & Schultz T (2010). Online workload recognition from eeg data during cognitive tests and human-machine interaction. *KI 2010: Advances in Artificial Intelligence*, Springer, pages 410–417.
- [51] Herff C, Heger D, Putze F, Hennrich J, Fortmann O & Schultz T (2013). Classification of mental tasks in the prefrontal cortex using fnirs. *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 2160–2163.
- [52] Hjorth B (1970). Eeg analysis based on time domain properties. *Electroencephalography and clinical neurophysiology*, 29(3):306–310.
- [53] Höök K, Isbister K & Laaksolahti J (2006). Sensual evaluation instrument. *Proc CHI 06*.
- [54] Horlings R, Datcu D & Rothkrantz LJ (2008). Emotion recognition using brain activity. *Proceedings of the 9th international conference on computer systems and technologies and workshop for PhD students in computing*, ACM, 6 pages.
- [55] Howitt J, Hay A, Shergold G & Ferres H (1978). Workload and fatigue-in-flight EEG changes. *Aviation, space, and environmental medicine*, Vol 49:1197–1202.
- [56] Hwang HJ, Kim S, Choi S & Im CH (2013). EEG-based brain-computer interfaces: a thorough literature survey. *International Journal of Human-Computer Interaction*, 29(12):814–826.
- [57] Izard CE (1992). Basic emotions, relations among emotions, and emotion-cognition relations. *Psychol Rev.*, pages 99(3):561–5.
- [58] Jackson MM & Mappus R (2010). Applications for brain-computer interfaces. *Brain-Computer Interfaces*, Springer, pages 89–103.
- [59] Jatupaiboon N, Pan-ngum S & Israsena P (2013). Real-time eeg-based happiness detection system. *The Scientific World Journal*, 2013.

- [60] Jenke R, Peer A & Buss M (2014). Feature extraction and selection for emotion recognition from EEG. *IEEE Transactions on Affective Computing*, 5:327 – 339.
- [61] Jerritta S, Murugappan M, Nagarajan R & Wan K (2011). Physiological signals based human emotion recognition: a review. *IEEE 7th International Colloquium on Signal Processing and its Applications (CSPA)*, IEEE, pages 410–415.
- [62] Kar R, Konar A, Chakraborty A & Nagar AK (2014). Detection of signaling pathways in human brain during arousal of specific emotion. *International Joint Conference on Neural Networks (IJCNN)*, IEEE, pages 3950–3957.
- [63] Kensinger EA (2004). Remembering emotional experiences: The contribution of valence and arousal. *Reviews in the Neurosciences*, 15(4):241–252.
- [64] Khushaba RN, Greenacre L, Kodagoda S, Louviere J, Burke S & Dissanayake G (2012). Choice modeling and the brain: A study on the electroencephalogram (eeg) of preferences. *Expert Systems with Applications*, 39(16):12378–12388.
- [65] Kim KH, Bang S & Kim S (2004). Emotion recognition system using short-term monitoring of physiological signals. *Medical and biological engineering and computing*, 42(3):419–427.
- [66] Koelstra S *et al.* (2012). DEAP: A database for emotion analysis; using physiological signals. *IEEE Transactions on Affective Computing*, 3(1):18–31.
- [67] Kothe CA & Makeig S (2011). Estimation of task workload from eeg data: new and current tools and perspectives. *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, IEEE, pages 6547–6551.
- [68] Kreibig SD (2010). Autonomic nervous system activity in emotion: A review. *Biological psychology*, 84(3):394–421.
- [69] Kroupi E, Vesin JM & Ebrahimi T (2013). Phase-amplitude coupling between eeg and eda while experiencing multimedia content. *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*, IEEE, pages 865–870.
- [70] Kroupi E, Vesin JM & Ebrahimi T (2014). Implicit affective profiling of subjects based on physiological data coupling. *Brain-Computer Interfaces*, 1(2):85–98.
- [71] Kullback S & Leibler RA (1951). On information and sufficiency. *The annals of mathematical statistics*, pages 79–86.
- [72] Lal TN, Schröder M, Hill NJ, Preissl H, Hinterberger T, Mellinger J, Bogdan M, Rosenstiel W, Hofmann T, Birbaumer N *et al.* (2005). A brain computer interface with online feedback based on magnetoencephalography. *Proceedings of the 22nd international conference on Machine learning*, ACM, pages 465–472.
- [73] Lang PJ, Bradley MM & Cuthbert BN (1997). International affective picture system (iaps): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention*, pages 39–58.
- [74] Lang PJ, Greenwald MK, Bradley MM & Hamm AO (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30(3):261–273.

- [75] Le Van Quyen M, Foucher J, Lachaux JP, Rodriguez E, Lutz A, Martinerie J & Varela FJ (2001). Comparison of hilbert transform and wavelet methods for the analysis of neuronal synchrony. *Journal of neuroscience methods*, 111(2):83–98.
- [76] Leeb R, Friedman D, Slater M & Pfurtscheller G (2012). A tetraplegic patient controls a wheelchair in virtual reality. *BRAINPLAY 07 Brain-Computer Interfaces and Games Workshop at ACE (Advances in Computer Entertainment) 2007*, 37 pages.
- [77] Levenson RW (1992). Autonomic nervous system differences among emotions. *Psychological science*, 3(1):23–27.
- [78] Li H, Li Y & Guan C (2006). An effective bci speller based on semi-supervised learning. pages 1161–1164.
- [79] Lisetti C, Nasoz F, LeRouge C, Ozyer O & Alvarez K (2003). Developing multimodal intelligent affective interfaces for tele-home health care. *International Journal of Human-Computer Studies*, 59(1):245–255.
- [80] Lisetti CL & Nasoz F (2004). Using noninvasive wearable computers to recognize human emotions from physiological signals. *EURASIP journal on applied signal processing*, 2004:1672–1687.
- [81] Loewenstein G & Lerner JS (2003). The role of affect in decision making. *Handbook of affective science*, 619(642):3.
- [82] Mantini D, Perrucci MG, Del Gratta C, Romani GL & Corbetta M (2007). Electrophysiological signatures of resting state networks in the human brain. *Proceedings of the National Academy of Sciences*, 104(32):13170–13175.
- [83] Marg E (1995). Descartes’error: Emotion, reason, and the human brain. *Optometry & Vision Science*, 72(11):847–848.
- [84] McFarland DJ, Sarnacki WA & Wolpaw JR (2011). Should the parameters of a bci translation algorithm be continually adapted? *Journal of neuroscience methods*, 199(1):103–107.
- [85] McFarland DJ & Wolpaw JR (2011). Brain-computer interfaces for communication and control. *Communications of the ACM*, 54(5):60–66.
- [86] McFarland RA (1985). Relationship of skin temperature changes to the emotions accompanying music. *Biofeedback and Self-regulation*, 10(3):255–267.
- [87] Merzagora AC, Izzetoglu M, Polikar R, Weisser V, Onaral B & Schultheis MT (2009). Functional near-infrared spectroscopy and electroencephalography: a multimodal imaging approach. *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*, Springer, pages 417–426.
- [88] Mikutta C, Altorfer A, Strik W & Koenig T (2012). Emotions, arousal, and frontal alpha rhythm asymmetry during beethovens 5th symphony. *Brain topography*, 25(4):423–430.
- [89] Minagawa-Kawai Y, Matsuoka S, Dan I, Naoi N, Nakamura K & Kojima S (2009). Prefrontal activation associated with social attachment: facial-emotion recognition in mothers and infants. *Cerebral Cortex*, 19(2):284–292.

- [90] Miner LA, McFarland DJ & Wolpaw JR (1998). Answering questions with an electroencephalogram-based brain-computer interface. *Archives of physical medicine and rehabilitation*, 79(9):1029–1033.
- [91] Mühl C, Allison B, Nijholt A & Chanel G (2014). A survey of affective brain computer interfaces: principles, state-of-the-art, and challenges. *Brain-Computer Interfaces*, 1(2):66–84.
- [92] Mühl C, Nijholt A, Allison B, Dunne S & Heylen D (2011). Affective brain-computer interfaces (abci 2011). *Affective Computing and Intelligent Interaction*, Springer, pages 435–435.
- [93] Müller MM *et al.* (1999). Processing of affective pictures modulates right-hemispheric gamma band EEG activity. *Clinical Neurophysiology*, 110(11):1913–1920.
- [94] Nasoz F, Lisetti CL, Alvarez K & Finkelstein N (2003). Emotion recognition from physiological signals for user modeling of affect. *Proceedings of the 3rd Workshop on Affective and Attitude User Modelling (Pittsburgh, PA, USA)*.
- [95] Oatley K, Keltner D & Jenkins JM (2006). *Understanding emotions*. Blackwell publishing.
- [96] Onslow AC, Bogacz R & Jones MW (2011). Quantifying phase–amplitude coupling in neuronal network oscillations. *Progress in biophysics and molecular biology*, 105(1):49–57.
- [97] Parrott WG (2001). *Emotions in social psychology: Essential readings*. Psychology Press.
- [98] Pedregosa F *et al.* (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- [99] Peng H, Long F & Ding C (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1226–1238.
- [100] Petersen MK, Stahlhut C, Stopczynski A, Larsen JE & Hansen LK (2011). Smartphones get emotional: mind reading images and reconstructing the neural sources. *Affective computing and intelligent interaction*, Springer, pages 578–587.
- [101] Pfurtscheller G, Allison BZ, Brunner C, Bauernfeind G, Solis-Escalante T, Scherer R, Zander TO, Mueller-Putz G, Neuper C & Birbaumer N (2010). The hybrid bci. *Frontiers in neuroscience*, 4.
- [102] Picard RW (2000). *Affective computing*. MIT press.
- [103] Picard RW & Daily SB (2005). Evaluating affective interactions: Alternatives to asking what users feel. *CHI Workshop on Evaluating Affective Interfaces: Innovative Approaches*, pages 2119–2122.
- [104] Picard RW & Healey J (1997). Affective wearables. *Personal Technologies*, 1(4):231–240.
- [105] Picard RW, Vyzas E & Healey J (2001). Toward machine emotional intelligence: Analysis of affective physiological state. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(10):1175–1191.
- [106] Plutchik R (2001). The nature of emotions human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist*, 89(4):344–350.

- [107] Plutchik R (2003). *Emotions and life: Perspectives from psychology, biology, and evolution*. American Psychological Association.
- [108] Preece J, Rogers Y, Sharp H, Benyon D, Holland S & Carey T (1994). *Human-computer interaction*. Addison-Wesley Longman Ltd.
- [109] Prokasy W (2012). *Electrodermal activity in psychological research*. Elsevier.
- [110] Rainville P, Bechara A, Naqvi N & Damasio AR (2006). Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International journal of psychophysiology*, 61(1):5–18.
- [111] Rimm-Kaufman SE & Kagan J (1996). The psychological significance of changes in skin temperature. *Motivation and Emotion*, 20(1):63–78.
- [112] Ritz T, Dahme B, Dubois AB, Folgering H, Fritz GK, Harver A, Kotses H, Lehrer PM, Ring C, Steptoe A *et al.* (2002). Guidelines for mechanical lung function measurements in psychophysiology. *Psychophysiology*, 39(5):546–567.
- [113] Russell JA (1980). A circumplex model of affect. *Journal of personality and social psychology*, 39(6):1161.
- [Samiee *et al.*] Samiee S, Donoghue T, Tadel F & Baillet S. Phase-amplitude coupling.
- [115] Sammer G, Blecker C, Gebhardt H, Bischoff M, Stark R, Morgen K & Vaitl D (2007). Relationship between regional hemodynamic activity and simultaneously recorded eeg-theta associated with mental arithmetic-induced workload. *Human brain mapping*, 28(8):793–803.
- [116] Sankar AB, Kumar D & Seethalakshmi K (2010). Enhanced method for extracting features of respiratory signals and detection of obstructive sleep apnea using threshold based automatic classification algorithm. *International Journal of Computer Science & Emerging Technologies*, 1(4):38–43.
- [117] Savran¹ A, Ciftci¹ K, Chanel G, Mota JC, Viet LH, Sankur¹ B, Akarun¹ L, Caplier A & Rombaut M (2006). Emotion detection in the loop from brain signals and facial images.
- [118] Scherer KR (2005). What are emotions? and how can they be measured? *Social science information*, 44(4):695–729.
- [119] Schölkopf B & Smola AJ (2002). *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. MIT press.
- [120] Schupp HT, Cuthbert BN, Bradley MM, Cacioppo JT, Ito T & Lang PJ (2000). Affective picture processing: the late positive potential is modulated by motivational relevance. *Psychophysiology*, 37(2):257–261.
- [121] Schutter DJ & Knyazev GG (2012). Cross-frequency coupling of brain oscillations in studying motivation and emotion. *Motivation and emotion*, 36(1):46–54.
- [122] Schutter DJ, Putman P, Hermans E & van Honk J (2001). Parietal electroencephalogram beta asymmetry and selective attention to angry facial expressions in healthy human subjects. *Neuroscience letters*, 314(1):13–16.

- [123] Sequeira H, Hot P, Silvert L & Delplanque S (2009). Electrical autonomic correlates of emotion. *International journal of psychophysiology*, 71(1):50–56.
- [124] Shapiro D, Jamner LD, Lane JD, Light KC, Myrtek M, Sawada Y & Steptoe A (1996). Blood pressure publication guidelines. *Psychophysiology*, 33(1):1–12.
- [125] Smith ZM, Delgutte B & Oxenham AJ (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 416(6876):87–90.
- [126] Soleymani M, Lichtenauer J, Pun T & Pantic M (2012). A multimodal database for affect recognition and implicit tagging. *Affective Computing, IEEE Transactions on*, 3(1):42–55.
- [127] Sörnmo L & Laguna P (2005). *Bioelectrical signal processing in cardiac and neurological applications*. Academic Press.
- [128] Strait M & Scheutz M (2014). Using near infrared spectroscopy to index temporal changes in affect in realistic human-robot interactions. *PhyCS*, pages 385–392.
- [129] Tipping M (2003). *Relevance vector machine*. US Patent 6,633,857.
- [130] Tort AB, Komorowski R, Eichenbaum H & Kopell N (2010). Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *Journal of neurophysiology*, 104(2):1195–1210.
- [131] Wang XW, Nie D & Lu BL (2011a). Eeg-based emotion recognition using frequency domain features and support vector machines. *Neural Information Processing*, Springer, pages 734–743.
- [132] Wang YT, Wang Y & Jung TP (2011b). A cell-phone-based brain–computer interface for communication in daily life. *Journal of neural engineering*, 8(2):025018.
- [133] Watson D (2000). *Mood and temperament*. Guilford Press.
- [134] Wessa M & Flor H (2007). Failure of extinction of fear responses in posttraumatic stress disorder: evidence from second-order conditioning. *American Journal of Psychiatry*.
- [135] Wilhelm FH, Pfaltz MC & Grossman P (2006). Continuous electronic data capture of physiology, behavior and experience in real life: towards ecological momentary assessment of emotion. *Interacting with Computers*, 18(2):171–186.
- [136] Wolpaw J & Wolpaw EW (2012). *Brain-computer interfaces: principles and practice*. OUP USA.
- [137] Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G & Vaughan TM (2002). Brain–computer interfaces for communication and control. *Clinical neurophysiology*, 113(6):767–791.
- [138] Wolpaw JR, Ramoser H, McFarland DJ & Pfurtscheller G (1998). Eeg-based communication: improved accuracy by response verification. *Rehabilitation Engineering, IEEE Transactions on*, 6(3):326–333.
- [139] Wu CH, Tzeng YL, Kuo BC & Tzeng GH (2014). Integration of affective computing techniques and soft computing for developing a human affective recognition system for u-learning systems. *International Journal of Mobile Learning and Organisation 1*, 8(1):50–66.

- [140] Wu D & Parsons TD (2011). Active class selection for arousal classification. *Affective Computing and Intelligent Interaction*, Springer, pages 132–141.
- [141] Xielifuguli K, Fujisawa A, Kusumoto Y, Matsumoto K & Kita K (2014). Pleasant/unpleasant filtering for affective image retrieval based on cross-correlation of eeg features. *Applied Computational Intelligence and Soft Computing*, 2014.
- [142] Zander TO & Kothe C (2011). Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. *Journal of neural engineering*, 8(2):025005.
- [143] Zeitlin DM & Westwood RA (1986). Measuring emotional response. *Journal of Advertising Research*, 26(5):34–44.