



HAL
open science

Analyse de la dynamique neuronale pour les Interfaces Cerveau-Machines : un retour aux sources.

Michel Besserve

► **To cite this version:**

Michel Besserve. Analyse de la dynamique neuronale pour les Interfaces Cerveau-Machines : un retour aux sources.. Interface homme-machine [cs.HC]. Université Paris Sud - Paris XI, 2007. Français. NNT: . tel-00559128

HAL Id: tel-00559128

<https://theses.hal.science/tel-00559128>

Submitted on 24 Jan 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE DE DOCTORAT

SPÉCIALITÉ : PHYSIQUE

*École Doctorale « Sciences et Technologies de l'Information,
des Télécommunications et des Systèmes »*

Présentée par

Michel BESSERVE

Sujet :

**Analyse de la dynamique neuronale pour les Interfaces
Cerveau-Machines : un retour aux sources.**

Soutenue le 22 novembre 2007 devant les membres du jury :

M Ali MOHAMMAD-DJAFARI	Président du jury
M Olivier BERTRAND	Rapporteur
M François CABESTAING	Examineur
Mme Line GANERO	Directeur de thèse
M Christian JUTTEN	Rapporteur
M Jacques MARTINERIE	Directeur de thèse

Résumé

Les Interfaces Cerveau-Machine sont des dispositifs permettant d'instaurer un canal de communication entre le cerveau humain et le monde extérieur sans utiliser les voies usuelles nerveuses et musculaires. Le développement de tels systèmes se situe à l'interface entre le traitement du signal, l'apprentissage statistique et la neurophysiologie.

Dans cette thèse, nous avons réalisé et étudié un dispositif d'Interface Cerveau-Machine non invasif asynchrone, c'est-à-dire capable d'identifier des actions mentales associées à des tâches motrices ou cognitives imaginées sans synchronisation sur un événement contrôlé par un système externe. Celui-ci est basé sur l'analyse en temps réel de signaux électro-encéphalographiques (EEG) issus d'électrodes disposées à la surface de la tête d'un sujet humain.

Du point de vue méthodologique, nous avons implémenté plusieurs techniques de prétraitement de ces signaux et comparé leur influence sur les performances du système. Ces techniques comprennent : 1) l'utilisation directe du signal issu des capteurs EEG, 2) l'exploitation de méthodes de séparation de sources qui permettent de résumer les signaux EEG par un faible nombre de composantes spatiales et 3) la reconstruction de l'activité des sources de courant corticales par résolution du problème inverse en EEG. De plus, plusieurs mesures permettant de quantifier l'activité cérébrale sont exploitées et comparées : la puissance spectrale, la cohérence et la synchronie de phase. Nos résultats montrent que la reconstruction préalable de l'activité corticale par problème inverse, ainsi que l'utilisation de mesures d'interaction à distance permettent d'améliorer les performances du dispositif.

Mots clés : Interface cerveau-machine, synchronie de phase, cohérence, classification, électro-encéphalographie, problème inverse, séparation de sources, séparateur à vaste marge.

Cette thèse a été réalisée au laboratoire de Neurosciences Cognitives et Imagerie Cérébrale, CNRS UPR 640-LENA, 47, boul. de l'Hôpital, 75651 Paris Cedex 13.

Page web : <http://cogimage.dsi.cnrs.fr/>

Analysis of neural dynamics for Brain-Machine Interfaces : back to the sources.

Abstract

Brain-Machine Interfaces are aimed at providing a new communication channel to link the human brain to external world without using the usual nervous and muscular pathways. The achievement of such devices is at the conjunction of signal processing, machine learning and neurophysiology.

In this thesis, we achieved and studied a non invasive Brain Machine Interface. This system is asynchronous : the user communicates through self-paced imagined motor or cognitive tasks which are not triggered by an external stimulation system. It is based on the real time analysis of electroencephalographic recordings from scalp electrodes in a human subject.

From a methodological point of view, several preprocessing of these signals have been implemented and their impact on the efficiency of communication system have been compared. These methods include : 1) direct processing of the signal from EEG sensors, 2) blind source separation algorithms, 3) reconstruction of cortical sources' activities with inverse problem solving. Moreover, several measures are exploited and compared to quantify cerebral activity : spectral power, coherence and phase synchrony.

Our results show that a using a previous source reconstruction of cortical activity through inverse problem and computing long distance interaction measures improve the performance of the device.

Keywords : Brain Machine Interface, phase synchrony, coherence, classification, electro-encephalography, inverse problem, blind source separation, support vector machine.

This thesis has been completed at the Cognitive Neuroscience & Brain Imaging Lab in Paris, France.

Address : Laboratoire de Neurosciences Cognitives et Imagerie Cérébrale, CNRS UPR 640-LENA, 47, boul. de l'Hôpital, 75651 Paris Cedex 13, FRANCE.

Webpage : <http://cogimage.dsi.cnrs.fr/>

Remerciements

Lorsqu'il y a presque quatre ans, j'ai décidé de m'engager dans une thèse, ce n'est pas tout à fait par hasard que j'ai rendu visite à Jacques Martinerie et Line Garnero au LENA. J'avais en effet déjà réalisé un stage au sein de l'équipe Analyse de la Dynamique Neuronale dirigée par Jacques. À cette occasion, j'avais pu constater combien les signaux d'imagerie cérébrale pouvaient être un champ d'application fertile pour le traitement du signal et les méthodes statistiques. Mais au-delà de cette perspective scientifique, je savais aussi que je pourrais trouver ici un environnement de travail exceptionnellement humain pour réaliser une thèse. Je me rappelle encore quelques paroles de Jacques à cette époque : "Il y a de mauvais moments à passer dans une thèse et c'est important de bien les gérer !" Bien sûr, il y en a eu : comment ne pas subir des déceptions lorsque l'on désire réaliser des projets ambitieux ? C'est en particulier dans ces moments-là que les conseils et les encouragements de Line et Jacques ont été déterminants. Il y a beaucoup d'autres choses pour lesquelles je voudrais les remercier : leur curiosité scientifique qui fut un véritable moteur pour moi, leur grande disponibilité (pardon d'en avoir abusé...) et la grande liberté d'initiative qu'ils m'ont laissée dans mes choix des scientifiques. Si j'ai décidé de persévérer dans la voie de la recherche, c'est en grande partie grâce à vous.

Bernard Renault fut la première personne à me recevoir au laboratoire, en tant que directeur de l'unité à l'époque. Je tiens à cette occasion à le remercier de l'intérêt qu'il a manifesté pour ma candidature et, par la suite, pour mon travail.

Au cours de cette thèse, plusieurs personnes m'ont donné la chance de travailler avec eux. Ce fut l'occasion d'échanger beaucoup d'idées et de découvrir leurs problématiques, par exemple en appliquant mes méthodes de classification aux données d'imagerie cérébrale qu'ils avaient acquises. Je remercie en particulier Mathieu Phillippe et Geneviève Florence de l'IMASSA¹, ainsi que Mario Chavez, Karim Jerbi et Sylvain Baillet au laboratoire. Je remercie aussi Florence Bouchet pour son aide cruciale lors des expériences d'enregistrement EEG, et aussi pour sa franchise et sa bonne humeur !

Merci à l'ensemble des membres de mon équipe pour leurs utiles conseils et leur soutien lors des nombreuses discussions sur mes présentations et mon travail en général. Je garde encore en mémoire ces nombreuses réunions de l'équipe ADN qui furent le siège de débats très animés et enrichissants. Merci à nouveau à Mario Chavez qui depuis de nombreuses années me prodigue régulièrement conseils, publications intéressantes et

1. Institut de Médecine Aérospatiale et du Service de Santé des Armées

aussi vin et parmesan en grande quantité, muchas gracias! Je salue aussi Fred, ma voisine de bureau pendant ces quatre années, avec toi j'ai partagé beaucoup de choses, merci pour ta présence. Merci aussi à Laurent Pezard, Sébastien, François, Miguel, Mario Valderama et Benoît Crépon pour leur soutien de tous les jours et les moments agréables passés ensemble.

Merci aussi à Françoise Pougeol, Catherine Mouly, Jean-Claude Bourzeix, Denis Schwartz, Antoine Ducorps, Karim N'Diaye, Laurent Hugueville et Manik Bhattacharjee pour leur aide dans mes déboires informatiques, matériels et administratifs. Votre disponibilité et votre expérience m'ont beaucoup aidé.

Je salue tous les étudiants et anciens étudiants du LENA avec qui j'ai partagé de très bons moments autant à l'intérieur qu'à l'extérieur du labo. Merci à Benoît Cottureau, aux deux Karim, à Julien, Florence, Anael, Joan, Manik, Laurence, Agnès, Guillaume, Max, Yohan, J-D, Yann, Zoé, Nathalie, Shasha, Valentin, et pardon pour tous ceux que j'oublie... Vous contribuez tous à rendre ce labo vivant et agréable.

Cette (longue) liste serait incomplète sans les personnes extérieures au laboratoire, qui m'ont progressivement amené dans la voie de l'enseignement et de la recherche. Il y a parmi eux mon grand-père et mon oncle ainsi que beaucoup d'enseignants qui m'ont transmis leur passion pour la science et la connaissance depuis l'école primaire jusqu'à l'ENS Cachan.

Merci à Christophe, Guillaume, Jean-Mich, Natacha, Nico et Slimane pour leur amitié et leur sincérité. J'exprime ma profonde reconnaissance à Isabelle pour sa grande patience, son soutien quotidien et son amour, ainsi qu'à mes parents et à l'ensemble de ma famille pour leurs encouragements.

Je tiens enfin à exprimer ma gratitude à Olivier Bertrand et Christian Jutten, qui m'ont fait l'honneur d'accorder de leur temps pour rapporter mon travail de thèse. Je joins à ces remerciements François Cabestaing et Ali Mohammad Djafari qui ont examiné attentivement ce travail et ont pris part à ce jury malgré le contexte mémorable des grèves de transports du mois de novembre 2007.

Notations

$a_i^f(t)$	signal analytique du capteur i autour de la bande de fréquence f (16)
\mathcal{H}	transformée de Hilbert (16)
$\varphi_i^f(t)$	phase instantanée du signal analytique du capteur i à la fréquence f (17)
f^*	fonction discriminante associée à un classifieur (23)
\mathbf{x}	vecteur de quantification de l'activité cérébrale (23)
y	variable de classe (26)
\mathbf{J}_p	densité volumique de courant primaire (36)
$\mathbf{q}(\mathbf{r})$	moment dipolaire du dipôle de courant à la position \mathbf{r} (38)
\mathbf{q}	vecteur des composantes de l'ensemble des moments dipolaires corticaux (38)
Λ	matrice de champ de sensibilité (38)
\mathbf{m}	vecteur des mesures EEG (capteurs \times temps) (42)
\mathbf{W}	matrice de problème inverse (42)
$\Gamma_{\mathcal{W}}$	matrice de densité spectrale estimée par la méthode de Welch (62)

Table des matières

Résumé	iii
English abstract	v
Remerciements	vii
Notations	ix
Introduction	1
Définition et buts d'une ICM	1
Schéma général	5
Objectif de la thèse	8
Lignes directrices	9
1 Fondements neurophysiologiques des ICM	11
1.1 Quelques notions sur le cortex	11
1.2 Les corrélats neurophysiologiques des états mentaux	13
1.2.1 Les mesures d'activité locale	13
1.2.2 Les interactions à longue distance	16
1.3 Tâches cognitives et ICM asynchrones	18
1.3.1 L'imagination : aspects psychologiques et physiologiques	18
1.3.2 L'imagerie motrice	19
1.3.3 L'imagerie visuo-spatiale	19
1.3.4 L'imagerie auditive	21
1.3.5 Le calcul mental	21
2 Du signal à l'information transmise	23
2.1 Les méthodes de classification	23
2.1.1 Principe	23
2.1.2 Quelques algorithmes de classification	23
2.2 Évaluer la performance des classifieurs	27
2.2.1 Les mesures de performance	27
2.2.2 Généralisation	29

2.2.3	Complexité	31
2.3	Les prétraitements	31
2.3.1	Sélection de variables	32
2.3.2	Filtrage spatial	33
2.3.3	Trouver la bonne quantification	34
3	Utiliser le problème inverse en ICM	35
3.1	Le problème direct en MEG et EEG	35
3.1.1	Origine physiologique de la MEG et de l'EEG	35
3.1.2	Modélisation par dipôle de courant	36
3.2	Le problème inverse	39
3.2.1	État de l'art	40
3.2.2	Solutions linéaires	42
3.2.3	Imagerie corticale des phénomènes de synchronisation	43
3.2.4	Le problème inverse dans les ICM	44
4	Expérience d'ICM asynchrone	47
4.1	Objectifs et contraintes	47
4.1.1	Généralités	47
4.1.2	Reconstruire les sources corticales	48
4.1.3	Un système temps réel	48
4.1.4	Choix du paradigme	48
4.1.5	Ma contribution	49
4.2	Protocole expérimental	49
4.2.1	Tâches	49
4.2.2	Entraînement	50
4.2.3	Feedback	51
4.2.4	Données enregistrées	53
4.3	Enregistrements	53
4.3.1	Système d'acquisition	54
4.3.2	Protocole de stimulation	55
4.3.3	Acquisition des données anatomiques	55
4.3.4	Architecture du traitement temps réel	56
4.4	Observations	57
4.4.1	Sessions d'apprentissage	57
4.4.2	Sessions de feedback	58
5	Analyse au niveau des capteurs	59
5.1	Méthodes	59
5.1.1	Prétraitements	59
5.1.2	Mesures utilisées	60
5.1.3	Normalisation	62
5.1.4	Classification	62
5.1.5	Sélection de variables	63

5.1.6	Validation des performances	64
5.2	Résultats	65
5.2.1	Choix des bandes de fréquences	65
5.2.2	Sélection de variables et quantification	65
5.2.3	Influence des états mentaux choisis	68
5.3	Discussion	69
5.3.1	Choix des bandes de fréquence	69
5.3.2	Sélection de variables	69
5.3.3	Quantification	70
5.3.4	Influence des tâches mentales	71
6	Analyse par séparation de sources	75
6.1	Cadre de l'étude	75
6.1.1	Hypothèses	75
6.1.2	Principe de la méthode SOBI	76
6.2	Application à des données MEG	77
6.2.1	Article publié dans International Congress Series	78
6.3	CSDA : Un algorithme hybride de séparation de sources pour la classification	83
6.3.1	Présentation	83
6.3.2	Article présenté à la troisième conférence EMBS-IEEE Neural Engineering	84
7	Analyse au niveau des sources corticales.	89
7.1	Méthodes	89
7.1.1	Reconstruction de l'activité corticale	89
7.1.2	Quantification	91
7.1.3	Classification	92
7.2	Résultats : Classification à partir de la puissance	94
7.2.1	Sélection de variables sur les sources corticales	94
7.2.2	Résultats en validation croisée inter-sessions	95
7.3	Résultats : Classification à partir de la cohérence	99
7.3.1	Taux de classification	99
7.3.2	Variables discriminantes	99
7.4	Discussion	100
7.4.1	Résultats de classification	100
7.4.2	Variables discriminantes	100
8	Discussion générale	115
8.1	Analyse de l'ensemble des résultats	115
8.1.1	Filtrage spatial et sélection de variables	115
8.1.2	Les tâches mentales	116
8.2	Perspectives	117
8.2.1	Améliorations méthodologiques	117

8.2.2 Études expérimentales futures	118
Annexe A : Travaux sur la fatigue cognitive	119
Annexe B : Classification d'une tâche de contrôle visuomoteur en MEG	133
Annexe C : OOpci, une librairie orientée objet pour les ICM	165
Index	169
Bibliographie	171

Introduction

Définition et buts d'une ICM

Le concept d'Interface Cerveau Machine (ICM : en anglais BCI, Brain Computer Interface) est apparu il y a une quinzaine d'années pour définir un système qui instaure un canal de communication entre le cerveau et le monde extérieur sans utiliser les voies usuelles nerveuses et musculaires. Les dispositifs d'ICM sont ainsi destinés en particulier à la palliation de handicaps et peuvent être divisés en plusieurs catégories ; on distingue ainsi :

- les ICM *entrée* versus *sortie* : cette distinction sépare respectivement les interfaces qui transmettent une information issue de l'environnement vers le cerveau, comme les prothèses auditives ou oculaires (entrée), et les interfaces qui communiquent un signal de commande du cerveau vers un dispositif externe (sortie). Nous nous restreignons ici au cas des ICM sortie.
- les ICM *directes* versus *indirectes* : Cette notion se réfère à la technique d'acquisition de l'activité cérébrale (Donoghue, 2002). Une ICM directe exploite le signal neuronal par le biais d'enregistrements invasifs (voir section et (Lebedev and Nicolelis, 2006)). Au contraire les ICM indirectes, qui sont les plus répandues, exploitent des phénomènes qui sont indirectement liés à l'activité neuronale. C'est le cas de l'Imagerie par Résonance Magnétique fonctionnelle (IRMf), qui mesure la consommation d'oxygène des populations de neurones, et dans une moindre mesure de l'EEG et de la MEG qui enregistrent sur le scalp le champ électromagnétique engendré par l'activation de certains types de neurones (voir section 3.1).
- les ICM *synchrones* versus *asynchrones* : Les paradigmes les plus utilisés en ICM exploitent les changements d'activités cérébrales provoqués par un stimulus. On qualifie ces systèmes de synchrones car l'instant d'apparition du stimulus par rapport auquel est effectué l'analyse du signal doit être imposé par le protocole de stimulation et contraignent le protocole de communication. Au contraire un paradigme asynchrone ne dépend pas d'un stimulus particulier : le changement d'activité cérébrale est décidé par l'utilisateur. Techniquement, le paradigme asynchrone nécessite un traitement continu sur des fenêtres glissantes des signaux enregistrés (Scherer et al., 2004), alors qu'un système synchrone réalise des traitements à des instants précis du protocole de stimulation. Certaines publications ont une définition plus stricte des ICM asynchrones, où l'on demande aussi que le système

puisse être allumé par l’initiative de l’utilisateur (et en n’utilisant que son activité cérébrale)(Borisoff et al., 2004).

- les ICM *dépendantes* versus *indépendantes* : les ICM indépendantes utilisent exclusivement les signaux du système nerveux central, tandis que les ICM dépendantes peuvent exploiter des signaux auxiliaires comme le clignement des yeux.

Notre propos dans cette thèse se restreint aux *ICM indirectes asynchrones indépendantes* qui présentent des avantages sur les autres dispositifs en terme de simplicité de réalisation ainsi qu’en terme d’applications envisageables. En effet, les dispositifs non-invasifs peuvent être utilisés chez des sujets sains et ne présentent pas de risque additionnel chez un sujet déjà touché par un handicap. De plus, les approches asynchrones donnent plus de liberté au sujet et pourraient permettre, à terme, d’effectuer des actions complexes comme le déplacement 2D d’un curseur. Pour finir, la notion d’ICM indépendante, plus contraignante, permet l’utilisation du système par des patients n’ayant aucun moyen de communication avec l’extérieur. Cette dernière population fait en effet l’objet de l’application prioritairement visée par ces dispositifs : la restauration de la communication vers le monde extérieur chez un patient atteint de handicap moteur très sévère.

ICM et handicap moteur

Pour entrer de manière plus précise dans les mécanismes liés au handicap, on représente Fig 1 l’anatomie des voies motrices pyramidales qui commandent les mouvements volontaires fins des bras et des jambes. La commande motrice prend sa source au sein du cortex moteur primaire. Les fibres nerveuses se projettent alors à travers la substance blanche dans le tronc cérébral où a lieu la *décussation* : les fibres de l’hémisphère droit changent de côté pour se projeter sur les membres du côté gauche, et inversement. Les muscles des membres et du tronc poursuivent alors leur chemin dans la moelle épinière où ils se connectent à d’autres neurones moteurs, se projetant à leur tour dans les fibres musculaires. La commande des muscles du cou et du visage est similaire, à l’exception qu’elle ne passe pas par la moelle épinière : les fibres nerveuses se projettent sur des neurones moteurs du tronc cérébral, appelés nerfs crâniens, qui sont directement reliés aux muscles. Bien qu’il existe aussi des circuits de régulation — notamment de la posture et du tonus musculaire — qui entrent en jeu dans l’exécution des mouvements et font appel à d’autres structures nerveuses, les paralysies les plus sévères sont provoquées par les atteintes de ce système. Les personnes les plus susceptibles d’être aidées par une ICM sont celles atteintes d’une paralysie totale ou quasi totale, non seulement des membres supérieurs et inférieurs, mais qui empêche aussi la communication orale. L’ensemble de ces symptômes est observé dans le *locked-in syndrome* (LIS) ou *syndrome d’enfermement* qui est la conséquence d’une atteinte du tronc cérébral faisant généralement suite à un accident vasculaire cérébral. Cet état tient son nom du fait que la conscience du patient reste parfaitement intacte et active, alors que son corps est totalement immobilisé (les seuls mouvements volontaires possibles sont des mouvements verticaux des yeux et des clignements). Ce syndrome touche actuellement environ 400 personnes en

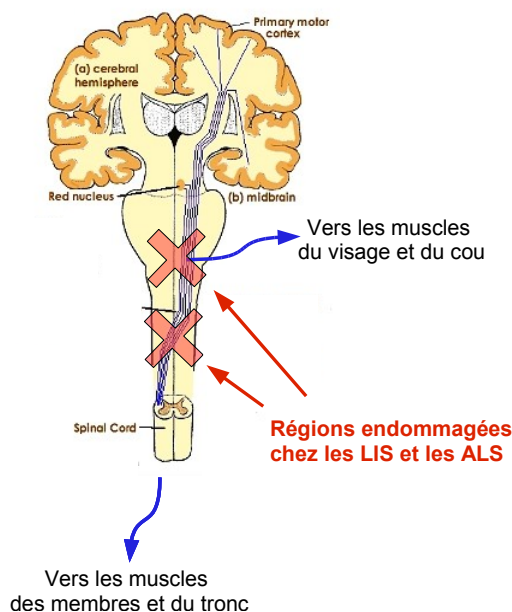


FIGURE 1 – Représentation schématique de la voie motrice pyramidale (image reprise de <http://freednerd.wordpress.com/category/the-brain/motor-tract/>)

France². Il existe aussi une pathologie qui évolue vers un état similaire au LIS. La *sclérose latérale amyotrophique* (en anglais ALS : amyotrophic lateral sclerosis) est une dégénérescence des neurones moteurs de la moelle épinière et du tronc cérébral d'origine encore inconnue. Le malade perd progressivement ses capacités motrices : d'abord les bras, puis les jambes, puis l'usage de la parole, et enfin les capacités de déglutition et l'usage des muscles faciaux. Cependant, là aussi, les capacités cognitives et la conscience restent intactes tout au long de la maladie. Cette maladie touche plus de 8000 personnes en France³. Il est important de préciser que, malgré leur handicap sévère, les patients ne semblent pas forcément dans la dépression et évaluent leur état de "bien-être" plus positivement que ce que pensent leurs proches (en moyenne). L'envie de communiquer — et plus généralement d'interagir avec le monde extérieur — est donc encore présente. Dans ce cadre, les dispositifs de communication tels que les ICM visent à améliorer la qualité de vie des patients et de leur entourage.

Il existe déjà des techniques de communication palliative pour les LIS :

- la communication par le pH de la bouche : l'imagerie d'aliments provoque des variations du pH de la bouche qui peuvent être modulées par les LIS (Wilhelm et al., 2006). Ce dispositif de communication, bien que sommaire, a l'avantage de

2. source : http://www.attaquerebrale.org/fileadmin/user_upload/Docs_Soignants/Mise_au_point/Locked-In_Syndrome.pdf

3. source : http://www.doctissimo.fr/html/dossiers/sclerose_laterale_amyotrophique.htm

demander peu de matériel.

- Lorsque les mouvements des yeux ou les clignements des paupières sont possibles (ce n'est pas toujours le cas), ils peuvent être à la base d'un dispositif de communication plus ou moins automatisé (Hori et al., 2006). Par exemple, on peut établir un protocole de communication de type : {un clignement = OUI, 2 clignements =NON}. Ce dispositif peut être exploité par des dispositifs de clavier virtuel : les lettres défilent sur un écran jusqu'à ce que l'utilisateur sélectionne la lettre par un clignement. On peut citer l'exemple célèbre de Jean-Dominique Bauby, atteint d'un LIS et auteur du roman *Le scaphandre et le papillon* écrit entièrement à l'aide de sa paupière gauche.

Par rapport à ces techniques, les ICM se distinguent par le fait qu'elles vont chercher l'information à transmettre à sa source au sein du cortex. Cela présuppose bien entendu que les régions corticales nécessaires à la transmission de l'information soient saines. Ceci est souvent vérifié, car la région causant le handicap est le plus souvent le tronc cérébral. L'avantage d'une ICM par rapport à un système de communication par mouvement des yeux est qu'elle ne nécessite l'activation d'aucun muscle (certains patients ne peuvent pas ou presque pas bouger les yeux). Par rapport à un dispositif de communication par le pH, un bénéfice est attendu en terme de vitesse de transmission de l'information (voir section 2.2.1). On suppose en particulier qu'il est possible de modifier plus rapidement son activité cérébrale que son activité salivaire.

Autres applications

À terme, les ICM sortie peuvent aussi être envisagées comme un nouveau canal de communication pour des utilisateurs non handicapés, notamment pour des applications ludiques. On peut citer notamment l'adaptation du jeu Pong par l'équipe du Berlin Brain Computer Interface⁴. En plus des ICM — qui communiquent des informations sur des changements volontaires d'activité cérébrale d'un sujet — les systèmes de monitoring sont aussi un champ d'application possible des techniques d'analyse temps réel de l'activité cérébrale. Ces systèmes ont pour but de détecter automatiquement des modifications neurophysiologiques involontaires chez un sujet. En particulier, les dispositifs destinés à détecter les baisses de vigilance des conducteurs ou des pilotes d'avion pourraient permettre d'anticiper de graves accidents. Un exemple d'application de nos techniques à ce domaine est présenté dans l'annexe A. Finalement, le monitoring de l'activité cérébrale est utilisé aussi à des fins thérapeutiques dans un domaine connexe à l'ICM : le neurofeedback. Celui-ci consiste en une tâche de contrôle en temps réel de certaines caractéristiques de l'activité cérébrale du patient. Ce type de protocole a été très utilisé pour soigner le trouble du déficit de l'attention et hyperactivité (Fuchs et al., 2003), et son usage s'est étendu notamment à la restauration de fonctions cérébrales après une attaque (Nelson, 2007). L'amélioration des techniques de classification de l'activité cérébrale utilisées en ICM pourrait aussi contribuer à ouvrir de nouvelles voies dans le domaine du neurofeedback.

4. <http://www.youtube.com/watch?v=qCSSBEXBCbY>

Schéma général

Un système d'ICM est une chaîne de traitement temps réel de l'activité cérébrale. Les signaux enregistrés par un dispositif d'acquisition subissent plusieurs traitements numériques : des prétraitements améliorant la qualité du signal, une quantification extrayant les caractéristiques utiles (en anglais "features") du signal, puis une classification permettant d'estimer l'*état mental* du sujet à partir de ces caractéristiques. Le terme état mental désignera tout au long de ce manuscrit un ensemble de caractéristiques physiologiques, comportementales et/ou psychologiques, que l'on attribue à un sujet à un instant donné. Il pourra ainsi correspondre par exemple à la nature de la tâche mentale effectuée par le sujet (dans les chapitres 5,6 et 7, ainsi que dans l'annexe B), ou bien au niveau de performance atteint par un sujet sur une tâche (performance élevée ou faible dans l'annexe A). En ce sens, un état mental du sujet est défini par des caractéristiques objectives externes au sujet (type de stimulation, mesure comportementale de la performance, consigne donnée au sujet), mais l'on supposera qu'il correspond aussi à une signature spécifique de l'activité cérébrale du sujet qui puisse être estimée à partir des signaux acquis. L'information sur l'état mental ainsi obtenue peut être utilisée comme un signal d'entrée pour un programme (par exemple un clavier virtuel) ou pour une commande (exemple d'un fauteuil roulant électrique). En général, cette information est communiquée à l'utilisateur par une rétro-action (feedback) qui peut être visuelle (mouvement d'un curseur sur un écran), auditive, ou tactile. Ce dispositif permet ainsi des séances de feedback ayant pour but d'améliorer le contrôle du système par l'utilisateur. Cette structure, représentée schématiquement Fig 2, apparaît comme une chaîne dans laquelle chaque maillon occupe un rôle fondamental dans le fonctionnement global. Les différents éléments de cette chaîne sont décrits dans les paragraphes suivants.

Système d'acquisition

Le fonctionnement d'un tel système repose sur un dispositif de mesure de l'activité cérébrale. Parmi les techniques non invasives qui ont été utilisées en ICM on trouve d'abord l'EEG (Électro-encéphalographie). Depuis peu commencent à apparaître des systèmes basés sur l'IRMf (Imagerie par Résonance Magnétique fonctionnelle), la MEG (Magnéto-encéphalographie), ainsi que la tomographie optique. Cette tendance ne témoigne pas forcément d'une évolution vers la technique d'imagerie fonctionnelle cérébrale la plus adaptée, mais plutôt d'une exploration de toutes les possibilités techniquement accessibles de réalisation d'une ICM. Ainsi, les expériences qui étaient initialement pratiquées sur des appareils peu coûteux ont été adaptées à des appareils plus volumineux pour lesquels les protocoles étaient plus contraignants et les traitements temps réel plus difficiles à mettre en oeuvre⁵. En parallèle se développent les ICM invasives,

5. Il faut préciser ici que l'utilisation des ICM n'a pas qu'un but applicatif : pouvoir discriminer différents états mentaux en essai unique est un argument fort pour valider les performances d'une technique d'imagerie et permet de tester des hypothèses nouvelles sur codage de l'information dans le cerveau.

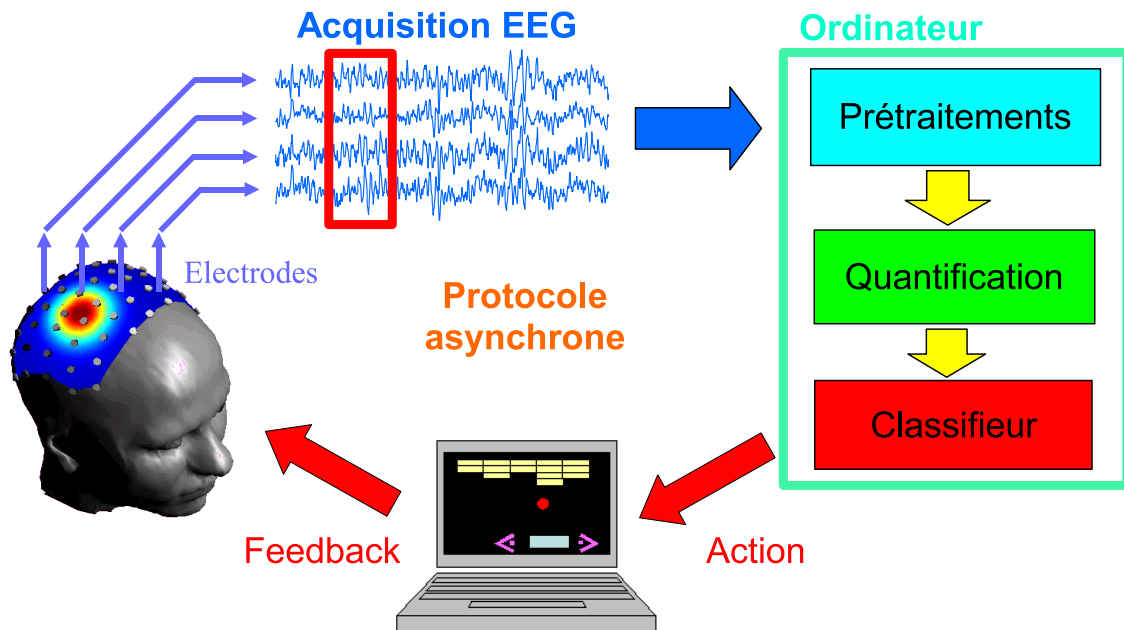


FIGURE 2 – Schéma général d'une ICM

utilisant chez l'homme les grilles d'électrodes sous-durales et les électrodes EEG stéréotaxiques de patients épileptiques implantés (en vue d'une exploration pré-chirurgicale), tandis qu'en parallèle se développent chez l'animal des interfaces basées sur des enregistrements unitaires ou multi-unitaires (Lebedev and Nicolelis, 2006; Wessberg et al., 2000; Quiroga et al., 2006).

Du point de vue des performances, les systèmes non invasifs d'ICM héritent des qualités et défauts de leur système d'imagerie : une très bonne résolution temporelle pour l'EEG et la MEG, et une très bonne résolution spatiale pour l'IRMf. Ainsi, bien que nous manquions encore de recul sur les capacités des différents systèmes, on peut supposer que l'IRMf est à préférer pour séparer des états mentaux très proches : par exemple la perception de différents stimuli visuels peut être différenciée par leurs activations respectives sur les aires rétinotopiques (Thirion et al., 2006). À l'inverse, pour des états mentaux très différenciés, un système devant réagir à l'échelle de la seconde sera préférentiellement basé sur l'EEG ou la MEG. Enfin, à l'aide des techniques invasives, il est possible de cibler une population de neurones ayant une fonction très précise, avec une qualité de signal bien supérieure.

Si départager les systèmes du point de vue des performances est difficile, la situation est plus claire en ce qui concerne la mise en oeuvre des applications. L'EEG et les implants corticaux sont pour l'instant les seuls dispositifs ambulatoires, permettant à l'ICM de sortir du laboratoire. Il n'est cependant pas exclu que certaines évolutions technologiques de l'IRMf et de la MEG puissent aussi conduire à des dispositifs ambu-

latoires à très long terme (cela suppose notamment la miniaturisation ou la suppression de leurs dispositifs de refroidissement). À plus court terme, la tomographie optique, une technique d'imagerie non invasive plus récente (Arridge, 1999), pourrait être utilisée de façon ambulatoire. Dans le cadre de cette thèse, c'est l'électro-encéphalographie qui a été choisie pour ses aspects ambulatoire et non-invasif, et pour sa bonne résolution temporelle. Des signaux MEG ont aussi été exploités car, bien qu'issus d'un dispositif non ambulatoire, ils reflètent les mêmes aspects biophysiques de l'activité neuronale (voir chapitre 3).

Prétraitements

Le prétraitement regroupe un ensemble de techniques qui permettent d'améliorer la qualité du signal cérébral. Parmi ces techniques on trouve les méthodes d'élimination d'artefacts, qui permettent de ne conserver que les signaux liés à l'activité cérébrale. Les artefacts à éliminer des signaux EEG et MEG sont principalement les activités cardiaques, oculaires et musculaires. Les techniques usuelles pour les éliminer sont des filtres spatiaux dérivés des techniques de ACP ou d'ACI (voir section 2.3.2) ainsi que des méthodes de filtrage temporel.

Un deuxième type de prétraitement est destiné à rehausser les caractéristiques informatives du signal. Une des premières méthodes utilisées en ICM fut le changement de référence du signal EEG. Ainsi, l'application d'un opérateur de référence moyenne commune ou d'un Laplacien spatial des données EEG permettent de rehausser les oscillations du cortex moteur par rapport au bruit (McFarland et al., 1997). D'autres traitements font appel à des techniques de filtrage spatial ou spatio-temporel adaptées au problème. Celles-ci mettent en valeur les composantes du signal qui différencient au mieux deux tâches (Muller-Gerking et al., 1999). La plupart de ces filtres adaptés se basent sur des critères statistiques pour traiter les signaux. On commence cependant à s'intéresser aux techniques de problème inverse qui exploitent des informations anatomiques sur les sujets. Elles permettent d'estimer à chaque instant l'activité des assemblées neuronales du cortex à partir des enregistrements EEG ou MEG. Ce type de prétraitement permet ainsi de quantifier l'activité des aires corticales plutôt que celle des capteurs.

Quantification

Il faut ensuite extraire des signaux prétraités les caractéristiques donnant des informations sur l'état mental du sujet. Celles-ci mettent en évidence des phénomènes neurophysiologiques spécifiques à chaque état mental. Le type d'ICM, *synchrone* ou *asynchrone* est un facteur déterminant de cette quantification, car elle autorise ou non la quantification de phénomènes liés à une stimulation. Les caractéristiques extraites dépendent fortement du système d'acquisition et de la tâche étudiée, mais aussi du sujet qui est souvent une source importante de variabilité. On fait généralement la distinction entre deux types de mesures : les mesures d'activité locale caractérisant l'activité au

sein d'une région corticale, et les mesures d'interaction à distance quantifiant le couplage entre différentes régions (voir section 1.2).

Classification

Une fois que l'activité est quantifiée, il faut être capable d'estimer l'état mental à partir des variables de quantification obtenues. Ce type de problème est appelé classification ou reconnaissance de formes ; il appartient au domaine de l'apprentissage statistique. En effet, il faut qu'un algorithme "apprenne", à partir d'exemples, à utiliser les informations contenues dans l'espace de quantification afin de différencier les états. Des méthodes paramétriques classiques telles que, l'Analyse linéaire discriminante sont utilisées depuis plusieurs dizaines d'années dans ce domaine. Cependant, la dernière décennie a vu émerger des outils non-paramétriques capables d'atteindre de très bonnes performances avec un faible nombre de données telles que les Séparateurs à Vaste Marge (SVM) (voir chapitre 2). Une fois qu'un algorithme d'apprentissage a été implémenté, une estimation de l'état mental doit pouvoir être associée en temps réel aux données acquises. Cette estimation peut être discrète (un entier donnant la classe estimée) ou continue (assimilable à une probabilité d'appartenance à une classe). Lors d'une séance de feedback, c'est cette quantité qui est communiquée au sujet par l'intermédiaire d'un retour sensoriel (visuel, auditif ou tactile).

Protocole expérimental

Nous englobons dans le terme "protocole expérimental" tous les aspects qui concernent les séances d'enregistrements de l'activité cérébrale. Cela comprend le choix entre ICM asynchrone versus synchrone, le choix des états mentaux à discriminer, les protocoles de stimulation et de feedback. Certains paramétrages de la chaîne de traitement, comme le réglage du nombre de classes ou la largeur des fenêtres d'analyse sur lesquelles l'état mental est estimé, sont aussi des choix importants.

Objectif de la thèse

L'objectif principal de la thèse est de développer et tester de nouveaux paradigmes expérimentaux et de nouvelles méthodes de traitement pour les ICM asynchrones indépendantes basées sur l'EEG.

Plus particulièrement, nous apportons notre contribution à 3 niveaux :

- Les *méthodes de prétraitement* : nous appliquons des méthodes de prétraitement par filtrage linéaire basées sur : 1) des informations statistiques qui guident l'optimisation des filtres spatiaux (ces travaux s'inspirent des méthodes de séparation aveugle de sources) et 2) des informations anatomiques : utilisation des méthodes de reconstruction de sources en EEG. Ces techniques ont jusqu'ici été très peu utilisées dans le domaine des ICM et jamais à notre connaissance dans le cas d'une ICM asynchrone.

- La *quantification* : nous utilisons de manière conjointe des mesures d’activité locale (puissance spectrale) et d’interaction à distance (cohérence et synchronie de phase) et quantifions leur apport.
- La réalisation et le traitement d’une expérience d’*ICM asynchrone* : nous avons réalisé des enregistrements EEG haute densité dans un paradigme d’ICM asynchrone exploitant des tâches mentales d’imagination motrice et non motrice. Ce protocole expérimental est géré par une interface temps réel qui permet de traiter les signaux acquis et de convertir le résultat de la classification en un retour visuel pour le sujet.

Lignes directrices

Ce manuscrit a un contenu fortement pluridisciplinaire dont la structure est résumée figure 3. Le sujet se situe à l’interface entre le traitement du signal (au sens large) et les neurosciences. Ainsi les deux premiers chapitres sont destinés au positionnement de la thématique ICM dans ces deux grands champs disciplinaires. Le chapitre 1 fait la revue des aspects de la MEG et de l’EEG exploitables dans le cadre d’une ICM. Les notions de synchronisation locale et à distance sont introduites avec les méthodes de mesure qui leur sont associées (désynchronisation liée à un événement, synchronie de phase, cohérence). La fin du chapitre fait la revue des données d’imagerie cérébrale fonctionnelle permettant de caractériser certaines tâches mentales qui seront exploitées dans les expériences d’ICM. Le chapitre 2 introduit les méthodes de classification ainsi que deux aspects susceptibles d’améliorer les performances sur des signaux spatiaux-temporels tels que la MEG et l’EEG : la sélection de variables et les techniques de filtrage spatial. Les méthodes de problème inverse, qui peuvent s’apparenter à un filtrage spatial, font l’objet d’un chapitre particulier car elles font appel à d’importantes notions biophysiques et mathématiques. Le protocole d’ICM asynchrone réalisé au laboratoire est présenté chapitre 8, ainsi que les aspects matériels et logiciels de l’expérience. Nos contributions méthodologiques exploitent trois niveaux d’analyse successifs (capteurs, composantes spatiales, et sources corticales) et sont décrites respectivement aux chapitres 5,6 et 7. L’ensemble des résultats est finalement interprété dans une discussion générale chapitre 8. Les méthodes présentées ont été appliquées sur d’autres jeux de données que ceux obtenus lors de notre expérience d’ICM. Ces travaux ont donné lieu à des articles que pour des raisons de cohérence nous présentons en annexe. Ainsi, en annexe A, les méthodes sont appliquées à la classification de la performance des sujets lors d’une tâche cognitive. En annexe B sont présentés des résultats de classification appliqués à des enregistrements MEG d’une tâche de contrôle visuo-moteur. Enfin, l’annexe C rend compte des travaux informatiques effectués, avec la réalisation d’une boîte à outils orientée objet sous Matlab pour la mise en oeuvre des expériences et le traitement des données d’ICM.

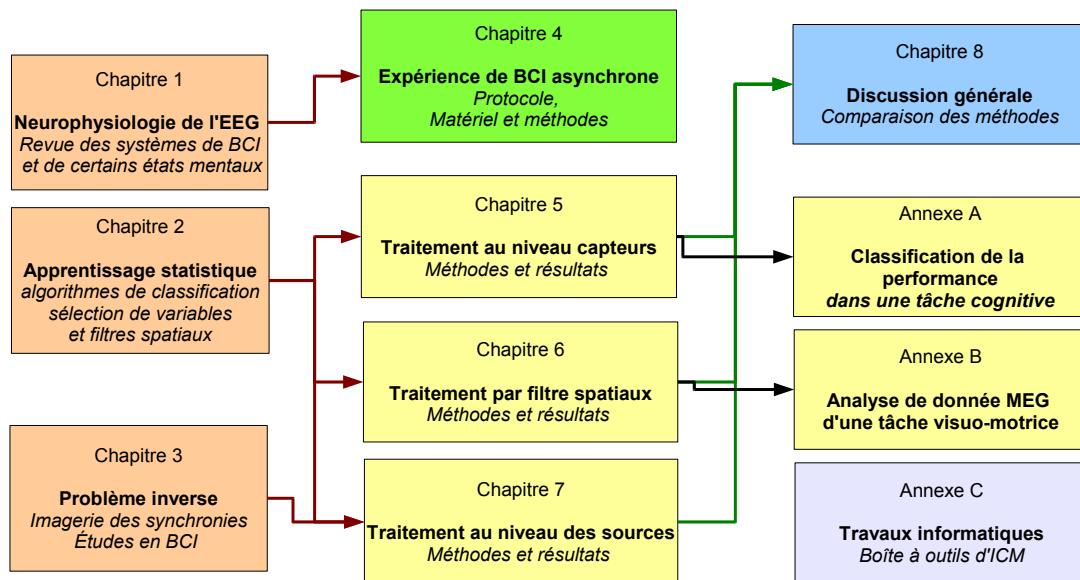


FIGURE 3 – Organisation du manuscrit : les premier chapitres de revue introduisent les méthodes appliquées aux données d'une expérience d'ICM asynchrone réalisée au laboratoire et utilisées dans chaque chapitre de résultats. L'ensemble des résultats et discuté au chapitre 8.

Chapitre 1

Fondements neurophysiologiques des ICM

Dans ce chapitre, les aspects fondamentaux qui permettent de lier l'exécution d'une tâche mentale à des modifications l'activité cérébrale mesurables à la surface de la tête d'un sujet. Nous nous restreindrons aux principes de mesure de l'activité cérébrale qui s'appliquent aux signaux EEG et MEG qui seront exploités dans nos travaux. Précisons cependant que une grande partie du domaine de l'imagerie cérébrale utilise des mesures de l'activité hémodynamique du cerveau qui peut être mesurée notamment en IRM fonctionnelle et en tomographie optique.

1.1 Quelques notions sur le cortex

La partie supérieure du système nerveux central, représentée figure 1.1 est appelée cortex, c'est une couche de matière grise d'environ 5 mm d'épaisseur divisée en deux hémisphères. Chaque hémisphère est lui-même composé de quatre lobes séparés par de grands "plis" : les sillons. On attribue au cortex une organisation anatomofonctionnelle où certaines aires localisées correspondent à des modules spécialisés (voir [Kandel et al. \(2000\)](#) et aussi section 1.3) :

- la partie avant du cortex (lobe frontal) supervise tous les processus qui soutiennent l'action (décision planification, exécution),
- les régions postérieures traitent les informations sensorielles : la vision dans l'aire occipitale, l'audition dans l'aire temporale et le toucher dans l'aire somatosensorielle,
- ces informations sont coordonnées par des aires dites "associatives" comme celles du cortex pariétal (voir figure 1.1).

La matière du cortex est composée de neurones connectés entre eux par de nombreuses fibres nerveuses. En particulier, l'enveloppe du cortex, dite substance grise, est composée de différentes couches de neurones. Elle se différencie par son aspect de la substance blanche, à l'intérieur de cette enveloppe, qui n'est composée que de fibres nerveuses qui connectent entre eux les neurones des deux hémisphères cérébraux (gauche et droit). Au

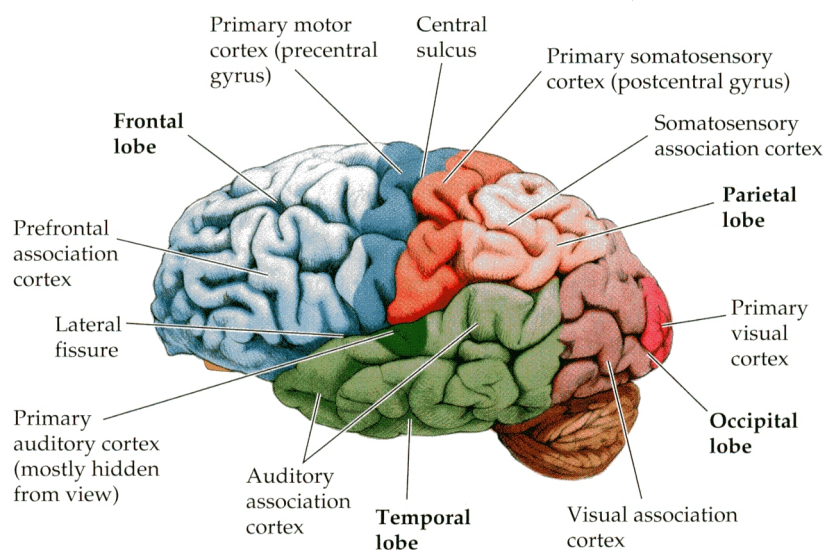


FIGURE 1.1 – Organisation anatomofonctionnelle du cortex humain sur une vue latérale de l'hémisphère gauche du cortex : les parties antérieures (à gauche sur l'image) sont attribuées à la préparation de l'action et à son exécution par le cortex moteur. Les parties postérieures et temporales rassemblent les aires qui traitent les informations perceptuelles (image reprise de http://www.neuropsychopathologie.fr/downloads/structure_cerveau004.jpg).

sein de ces assemblées neuronales, les informations se propagent grâce à des phénomènes de diffusion et de transfert ionique. D'un point de vue électromagnétique, l'activation d'un neurone se traduit par des mouvements de charge générant un dipôle de courant microscopique (voir chapitre 3). C'est l'activation simultanée de plusieurs millions de neurones voisins qui engendre un dipôle macroscopique dont le potentiel électrique et le champ magnétique rayonné sont détectables à la surface de la tête par l'EEG et la MEG respectivement (Nunez, 1981; Hämäläinen et al., 1993).

Pour quantifier l'activité corticale liée à la tâche issue des enregistrements EEG ou MEG, le décours de ceux-ci est en général segmenté en fenêtres temporelles sur lesquelles différents traitements sont effectués. La plupart des caractéristiques du signal qui sont estimées peuvent être divisées en deux familles, suivant l'extension spatiale du phénomène physique que l'on veut mesurer : les activités locales et les interactions à distance. Les activités locales sont mesurées au niveau d'un détecteur (une électrode dans le cas de l'EEG). On suppose que cette mesure reflète la modulation de l'activité neuronale dans une aire corticale focale localisée près du détecteur. À l'inverse, les mesures d'interactions à longue distance quantifient le couplage entre les signaux enregistrés au niveau de deux détecteurs distincts. Ce couplage peut, bien que pas nécessairement,

0. repris de http://www.neuropsychopathologie.fr/downloads/structure_cerveau004.jpg

révéler un transfert d'information entre deux assemblées neuronales distinctes. L'interprétation physiologique d'un tel phénomène est l'objet d'un grand nombre de travaux utilisant des techniques d'enregistrement invasives et non invasives, chez l'homme et chez l'animal (Rodriguez et al., 1999; Varela et al., 2001; Brovelli et al., 2004; Jerbi et al., 2007; Lachaux et al., 1999; Gray and Singer, 1989).

Sachant que l'activité électromagnétique enregistrée au niveau du scalp est une signature indirecte et atténuée des processus neuronaux sous-jacents¹. L'amplitude des signaux fonctionnels mesurés dépend probablement de paramètres variés incluant le rapport signal à bruit du système d'acquisition, la taille des populations impliquées, le paradigme expérimental et éventuellement l'état d'alerte et la motivation du sujet. Pour favoriser la détection de ces signaux, beaucoup de techniques de traitement du signal ont été utilisées dans les systèmes d'ICM non invasifs (Bashashati et al., 2007). Cependant, les phénomènes neurophysiologiques exploités sont en nombre réduit. Ils sont présentés dans les parties suivantes, ainsi que les mesures qui permettent de les quantifier.

1.2 Les corrélats neurophysiologiques des états mentaux

1.2.1 Les mesures d'activité locale

Étant donné que les paradigmes expérimentaux calés à un stimulus ont été largement utilisés à la fois pour les études cognitives fondamentales et pour les applications en ICM, un grand nombre d'outils d'analyse ont été testés et développés pour quantifier les activités dites *liées à un événement* (event-related en anglais). Le phénomène de ce type le plus anciennement connu est le *potentiel évoqué* (PE) qui est une onde apparaissant lorsque l'on moyenne sur un grand nombre d'essais des signaux à des instants référencés par rapport à un stimulus. Même si, à l'heure actuelle, les processus physiologiques sous-jacents à la génération des PE sont encore débattus, plusieurs types de PE induits par des stimuli spécifiques sont observés de manière robuste et peuvent être exploités. Par exemple, un clavier virtuel, appelé P300 speller, a été développé en se basant sur un PE spécifique faisant suite à l'apparition d'un stimulus attendu (Barret, 1996). À cause des fluctuations de ligne de base, les PE ne peuvent être détectés qu'en réalisant une moyenne des signaux sur plusieurs essais, réalisés dans les mêmes conditions expérimentales, pour chaque instant référencé par rapport au début de la stimulation. Bien qu'en théorie cette exigence rende le PE inutilisable pour la classification d'essais uniques, ce problème a été contourné dans le P300 speller en accumulant de multiples essais avant la classification (Farwell and Donchin, 1988). Dans ce système, une grille de lettres est présentée à l'utilisateur à qui l'on demande de compter le nombre de fois

1. la principale différence entre les enregistrements M/EEG et les enregistrements invasifs est donc que les effets liés à la tâche doivent être suffisamment forts pour être observés de manière non invasive à la surface de la tête

ou la lettre “cible” qu’il veut voir sélectionnée apparaît en surbrillance. En répétant un nombre de fois suffisant l’illumination de l’ensemble des lignes et des colonnes du système, il est possible de déduire quelle lettre (c’est-à-dire quelle ligne et quelle colonne) provoque une onde P300 lorsqu’elle apparaît en surbrillance. La P300 peut être détectée par exemple par son amplitude, son aire, par analyse temps-fréquence ou par corrélation avec un signal “patron” (Bashashati et al., 2007). Il existe aussi une onde lente exploitée en ICM, en particulier dans l’équipe de Berlin (Krepki, 2006), qui est le potentiel de préparation motrice (Bereitschaftspotential), variation négative précédant un mouvement (Deecke and Kornhuber, 2003).

Une autre quantification des modifications locales d’activités est donnée par des mesures appelées *Désynchronisation liée à un événement : DLE* (en anglais *Event Related Desynchronization : ERD*) ou *Synchronisation liée à un événement : SLE* (en anglais *Event Related Synchronization : ERS*) (Pfurtscheller and Aranibar, 1979; Pfurtscheller and Lopez Da Silva, 1999). La DLE et la SLE représentent respectivement une diminution ou une augmentation de la puissance des oscillations dans une bande de fréquence et à une localisation spécifique en comparaison de la puissance de la ligne de base estimée durant une période de référence (c’est-à-dire une fenêtre temporelle précédant l’apparition du stimulus ou alternativement une période d’enregistrement durant une condition de contrôle). On distingue plusieurs bandes de fréquences où ces phénomènes ont des rôles fonctionnels spécifiques. Des ondes thêta (4-8Hz) sont associées entre autres au fonctionnement de la mémoire à long terme (Klimesch, 1999). Dans la bande alpha 8-12Hz, les rythmes des régions occipitales sont associés à un fonctionnement “ au repos ” du système visuel : ils diminuent à l’ouverture des yeux et lors de l’apparition d’un stimulus visuel, cependant ils caractérisent aussi l’état de vigilance d’un sujet (Cantero et al., 2002)(voir aussi l’annexe A). Dans une bande de fréquence similaire, les rythmes mu (13Hz) sont présents dans le cortex moteur à l’état de repos et disparaissent pendant l’exécution d’un mouvement (Pfurtscheller and Aranibar, 1979; Salmelin and Hari, 1994). Dans les bandes de fréquences supérieures, il a été montré que les rythmes bêta disparaissent durant le mouvement et montrent une forte augmentation environ une à deux secondes après la fin du mouvement (ce phénomène est nommé rebond bêta (Cassim et al., 2001)). Finalement, les modifications de puissance dans la bande gamma (30-90Hz) semblent jouer un rôle important dans des processus cognitifs de haut niveau (Tallon-Baudry and Bertrand, 1999) comme la formation de la représentation cohérente d’un objet visuel à partir des éléments qui le constituent (le problème du “ binding ” visuel). Si ces phénomènes de D/SLE peuvent être vus dans les signaux en essai unique lorsqu’il s’agit des bandes thêta ou alpha, les phénomènes dans les bandes de plus haute fréquence sont le plus souvent mis en évidence par des analyses temps fréquences normalisées par rapport à une ligne de base et moyennées sur plusieurs essais. Historiquement, la présence d’ERS et de DLE dans des bandes de fréquences spécifiques est interprétée comme l’activation ou la désactivation d’une assemblée neuronale, mais cette théorie est remise en question par des études récentes : les phénomènes de DLE et SLE sont hautement dépendants de la tâche et de l’aire corticale impliquée (Lopes Da Silva, 2006) et un rythme dans une bande de fréquence donnée

peuvent suivant les cas correspondre à la désactivation ou à l'activation du réseau sous-jacent (Pfurtscheller, 2006). La DLE dans les bandes alpha et bêta apparaissant dans les aires motrices a été très utilisée pour la classification de différentes tâches d'imagerie motrice (Pfurtscheller et al., 1997) et dans plusieurs systèmes d'ICM comme celui de Berlin (Krepki, 2006) et celui de Graz (Neuper et al., 2003).

Comme les phénomènes de PE et de D/SLE sont liés à un événement, leur détection est toujours envisageable dans une ICM synchrone. Par contre dans les ICM asynchrones, l'instant de début d'une tâche mentale est soit inutilisé soit non disponible, ce qui rend le PE trop difficile à détecter. De plus, la nature transitoire du potentiel évoqué n'est pas compatible avec la caractérisation d'états continus. Au contraire, bien que les D/SLE soient usuellement référencées par rapport à une ligne de base ce phénomène peut aussi bien être considéré comme une (dé)activation liée à un événement ou bien à une tâche. Techniquement, l'alternative au calcul de D/SLE pour la quantification des états continus est de découper les enregistrements en fenêtres glissantes et d'appliquer des méthodes classiques de mesure de puissance spectrale dans différentes bandes de fréquences. Les outils usuels pour extraire l'information fréquentielle des signaux M/EEG sont l'analyse de Fourier et le filtrage passe-bande. Dans ce cas, les variables considérées sont des puissances spectrales dans différentes bandes de fréquences f obtenues en calculant la puissance moyenne du signal filtré durant chaque fenêtre temporelle de durée T : Ces outils peuvent être vus comme des méthodes d'analyse spectrale non paramétrique, car aucune hypothèse n'est faite sur le spectre des signaux. Un système d'ICM asynchrone contrôlant un robot a ainsi été réalisé par l'équipe de Martigny (Millan del R. et al., 2004). Une autre technique qui s'est révélée efficace pour la classification d'EEG asynchrone dans des études offline est la modélisation autorégressive (Anderson et al., 1998; Keirn and Aunon, 1990). C'est une approche paramétrique, car elle suppose que le spectre de puissance prenne la forme d'une fraction rationnelle décrite par un faible nombre de coefficients (typiquement 3 à 10).

Les mesures spectrales sont aussi adaptées pour détecter les Potentiels évoqués visuels en régime permanent (en anglais Steady State Visual Evoked Potential : SSVEP) induits dans le cortex visuel (occipital) par des stimulations visuelles répétitives et qui sont exploitées dans de récents paradigmes d'ICM (Middendorf et al., 2000).

Au lieu de mettre en évidence et d'exploiter des activités fonctionnelles qui avaient été mises en évidence par les neurophysiologistes en association avec des tâches cognitives, certaines équipes se sont intéressées depuis longtemps à la possibilité de réguler des activités cérébrales sans avoir recours à des protocoles cognitifs, mais seulement par feedback. Ainsi, l'équipe de Tübingen entraîne ses sujets à contrôler des potentiels corticaux lents apparaissant dans les régions centrales (Birbaumer et al., 1999). Par ailleurs, l'équipe du Wadsworth Center entraîne ses sujets à réguler leur rythme mu (Wolpaw et al., 1991). Contrairement aux systèmes basés sur la détection des corrélats neurophysiologiques d'une tâche mentale, qui nécessitent une optimisation et un ajustement des méthodes de traitement à l'utilisateur, les méthodes basées sur le feedback transfèrent tout l'effort d'adaptation à l'utilisateur et requièrent donc un apprentissage très long (Curran et al., 2003). De plus, les sujets ont finalement recours naturellement

à des tâches d'imagerie motrice pour apprendre à contrôler le dispositif (Curran et al., 2003; Wolpaw et al., 1991; Birbaumer et al., 1999). Ainsi, la tendance actuelle est à l'utilisation de tâches mentales prédéfinies, qui permettent de transférer la phase d'apprentissage du côté système plutôt que du côté utilisateur. Pour réaliser cet objectif, il est nécessaire d'utiliser des méthodes de classification de plus en plus élaborées. Il est aussi possible d'améliorer la quantification des tâches mentales en incorporant des variables supplémentaires comme les mesures d'interactions à distance.

1.2.2 Les interactions à longue distance

Il est raisonnable de supposer qu'un mécanisme d'intégration de l'information au sein du cerveau (ou intégration fonctionnelle) doit impliquer des interactions entre les différents réseaux de neurones participants (Varela et al., 2001). La quantification de ces interactions a été abordée avec de multiples techniques qui mesurent l'information commune présente dans les signaux issus d'aires corticales différentes. On peut résumer le principe de ces mesures comme la recherche d'une relation fonctionnelle entre les différents signaux; c'est pourquoi ce type de mesures est appelé dans la littérature connectivité fonctionnelle (en contraste avec la connectivité anatomique qui évalue les connexions nerveuses entre les aires)(Friston, 1994). La connectivité fonctionnelle étudie la structure du réseau cérébral à un instant particulier et caractérise la coopération entre plusieurs régions spécialisées. Les mesures d'interaction à longue distance les plus utilisées sont la *cohérence* et la *synchronie de phase* (Lachaux et al., 1999).

La cohérence mesure les interactions entre deux signaux $z_1(t)$ et $z_2(t)$ à une fréquence particulière en calculant la corrélation entre leurs transformées de Fourier $Z_1(f)$ et $Z_2(f)$ à cette fréquence² :

$$C_{12}(f) = \frac{|\mathbb{E} [Z_1(f)Z_2(f)^*]|}{\sqrt{\mathbb{E} [|Z_1(f)|^2] \mathbb{E} [|Z_2(f)|^2]}} \quad (1.1)$$

La mesure quantifiant la valeur des interactions à distance est en général la valeur absolue de la cohérence. Cependant, il est à noter que certains auteurs préfèrent utiliser la partie imaginaire plutôt que la valeur absolue afin de quantifier les interactions à distance en évitant le biais introduit par les effets de conduction de volume qui se manifestent dans la partie réelle (car la conduction n'introduit pas de déphasage entre les signaux).

Au contraire, la synchronie de phase mesure seulement l'information de phase commune aux deux signaux. Elle est calculée en convertissant les signaux à bande étroite obtenus par filtrage passe-bande de chaque capteur i autour d'une fréquence f , $z_i^f(t)$, en signaux analytiques $a_i^f(t)$ en utilisant la transformée de Hilbert \mathcal{H} par la formule

$$a_i^f(t) = z_i^f(t) + \mathbf{i}\mathcal{H} \left[z_i^f \right] \quad (1.2)$$

2. Dans l'équation 1.1, l'étoile indique l'opérateur de conjugaison.

Cela permet de définir la phase instantanée du signal filtré (Pikovski et al., 2001) par une décomposition en module et argument $a_i^f(t) = |a_i^f(t)|e^{i\varphi_i^f(t)}$ où i est le nombre complexe de module 1 et d'argument $\frac{\pi}{2}$. La valeur de verrouillage de phase ou *phase locking value* (PLV) entre les canaux 1 et 2 dans une bande de fréquence donnée est alors

$$S_{12}(f) = \frac{1}{T} \left| \sum_{t=1}^T e^{i(\varphi_2^f(t) - \varphi_1^f(t))} \right| \quad (1.3)$$

Plusieurs travaux dans la littérature ont mis en avant l'intérêt de cette mesure pour l'étude des signaux EEG. Dans une étude précédente (Rodriguez et al., 1999) réalisée au laboratoire, la synchronie de phase a été estimée sur des signaux EEG enregistrés durant une tâche de stimulation visuelle par des images de visages. Les modifications de synchronie de phase liées à la tâche entre des électrodes distantes ont montré des signatures significativement différentes entre les conditions " perception d'un visage " et " pas de perception de visage ". De plus, sur des enregistrements intracérébraux chez des patients épileptiques (Le Van Quyen et al., 2005), il a été montré que l'emploi de la mesure de synchronie de phase permettait d'anticiper les crises d'épilepsie.

Un autre argument en faveur de l'utilisation conjuguée de ces deux types de mesures est qu'elles peuvent être considérées comme deux aspects du même phénomène dans le cadre de la théorie des assemblées cellulaires résonnantes (Varela, 1995). La synchronie neuronale est aujourd'hui devenue un concept prometteur pour améliorer la caractérisation et la compréhension des aspects globaux de la dynamique cérébrale. Les neurosciences ont en effet fourni des preuves abondantes de ce phénomène à tous les niveaux d'organisation du système nerveux, allant des paires individuelles de neurones à des échelles macroscopiques : des assemblées de neurones au sein d'une aire corticale ou dans deux aires différentes. Généralement les oscillations sont supposées refléter la synchronisation locale de ces assemblées, le principe sous-jacent étant que ce sont les activités de grandes populations de neurones synchronisées qui permettent d'obtenir des oscillations mesurables à la surface du scalp. Au contraire la synchronisation de phase à longue distance entre deux électrodes est supposée être la signature d'un verrouillage de phase entre des régions corticales distinctes. Ainsi, ces deux mesures sont complémentaires pour caractériser l'activité cérébrale et essentielles pour appuyer l'hypothèse d'intégration globale de l'information au sein du système nerveux central (Tononi, 2005).

Enfin, des résultats récents dans le domaine des ICM suggèrent que combiner des mesures de synchronie de phase avec des mesures de puissance spectrale améliore les taux de bonne classification des signaux EEG (Wei et al., 2007; Gysels et al., 2005; Gysels and Celka, 2004).

1.3 Tâches cognitives et ICM asynchrones

À notre connaissance, parmi les études qui reportent l'utilisation de tâches cognitives dans le cadre d'ICM asynchrone, la plupart exploitent un paradigme d'imagerie motrice des bras, gauche ou droit, et des pieds (Leeb et al., 2007; Scherer et al., 2004). Dans un système de contrôle asynchrone du mouvement d'un robot mobile (Millan del R. et al., 2004), les auteurs donnent le choix entre plusieurs tâches cognitives au sujet, finalement ce sont les tâches motrices qui sont choisies et une tâche de rotation mentale d'un cube et il n'y a pas d'étude quantitative sur les tâches utilisées. Finalement le reste de la littérature sur l'utilisation de tâches cognitives autres que les tâches motrices proviennent d'études off-line des expériences de Keirn et Aunon (Keirn and Aunon, 1990; Anderson et al., 1998) qui, sous des hypothèses de spécialisation hémisphériques, ont choisi un jeu de cinq tâches mentales à différencier parmi lesquelles figuraient des tâches mentales de calcul, de rotation de figure géométrique, de visualisation de nombres et de composition d'un texte. La comparaison des résultats de classification sur ces tâches ne permet pas de mettre en évidence des différences entre le pouvoir discriminant des différentes tâches, vraisemblablement à cause de la méthode de validation du classifieur utilisée (validation leave one out) qui ne prend pas en compte les difficultés dues à la variabilité des signaux EEG entre différentes sessions. Afin de réaliser une étude quantitative de l'utilisation de différentes tâches d'imagerie motrices pour un système d'ICM asynchrone, nous nous proposons de faire une revue des données d'imagerie fonctionnelle sur les tâches mentales qui seront utilisées dans nos expériences.

1.3.1 L'imagination : aspects psychologiques et physiologiques

Considérant que nous voulons réaliser une ICM asynchrone, les tâches choisies ne doivent pas dépendre d'un stimulus externe. Une alternative est d'envisager que le sujet puisse " se stimuler " lui-même de façon interne. Il est donc naturel de considérer des tâches mentales qui permettent d'activer sans stimulus les aires sensorielles du cortex. C'est précisément la spécificité de tâches d'*imageries mentales*³ Kosslyn et al. (2001) (imageries visuelles, motrices et auditives). En effet, l'existence de l'imagerie mentale, définie comme un accès à une information perceptuelle mémorisée est maintenant étayée par des expériences en imagerie fonctionnelle cérébrale. Dans ces études, il est montré que l'imagination et la perception d'un stimulus activent des aires communes dans le cortex sensoriel de la modalité considérée. Ces résultats sur l'imagerie sont aussi à relier aux travaux sur la mémoire qui se focalisent sur l'accès à l'information mémorisée. Durant ce processus, les régions sensorielles semblent faire l'objet d'un contrôle " top-down " ⁴ à partir des régions frontales via l'hippocampe (Donoghue, 2002). Du point de

3. si étymologiquement imagerie fait référence à la modalité visuelle, ce terme a été adopté dans la littérature pour décrire des phénomènes impliquant les autres sens : ainsi, une imagerie auditive est un son entendu à l'intérieur de sa tête

4. Le contrôle top-down s'oppose aux traitements " bottom up " caractérisés par le cheminement progressif de l'information bas-niveau issue des organes sensoriels à travers des modules de traitements successifs qui en extraient des caractéristiques de haut niveau.

vue des ICM, l'imagerie motrice a été étudiée de manière prépondérante. Cependant, quelques travaux encourageants en ICM utilisent d'autres modalités (Curran et al., 2003). Nous allons ici revoir les résultats neurophysiologiques relatifs aux imageries motrices, auditives et visuo-spatiales. À celles-ci nous ajoutons les résultats concernant les tâches de calcul mental, qui ne sont pas du domaine des imageries mentales, mais peuvent y faire appel (pour verbaliser ou visualiser un nombre par exemple) et peuvent au même titre que ces dernières être effectuées de manière autonome par le sujet.

1.3.2 L'imagerie motrice

Ce sont les similarités entre les temps d'exécution d'un mouvement imaginé et d'un mouvement réel qui ont porté à croire que l'imagerie pouvait avoir une composante motrice (Decety and Michel, 1989). Depuis les recherches ont fait état d'une activation des aires de la planification et de la préparation du mouvement (Jeannerod and Decety, 1995) dans les mouvements imaginés (en particulier l'aire prémotrice numéro 6 et l'aire motrice supplémentaire). Cependant, l'activation de l'aire motrice primaire n'est pas systématiquement trouvée dans les études d'imagerie et a fait l'objet de controverses. Ce sont les études en EEG qui ont rendu compte de celle-ci par la mise en évidence de potentiels corticaux lents (Beisteiner et al., 1995) et de désynchronisation du rythme mu (Neuper and Pfurtscheller, 1999) en provenance de la région motrice primaire contralatérale. Il est aussi important de préciser que des aires non motrices entrent aussi en jeu, notamment dans l'imagination d'un mouvement : le lobule inférieur pariétal (ayant un rôle dans la représentation spatiale) et le cortex frontal dorsolatéral (impliqué dans la planification du mouvement).

Les ICM basées sur l'imagerie motrice exploitent principalement l'activité du cortex moteur primaire pour détecter quel mouvement est imaginé parmi un choix de différents membres (main gauche, main droite, pied). Ce paradigme est fondé sur l'hypothèse que l'imagerie activera la région du cortex moteur primaire spécifique du membre imaginé, réparti selon l'homoncule de Penfield représenté figure 1.2. Bien entendu, en utilisant des techniques non invasives, les aires spatialement distantes de cette région (par exemple celles activées par des mouvements des pieds et des doigts) seront plus facilement distinguées que des aires proches (main et doigt). Pour que ces signaux soient observés, la stratégie d'imagerie du mouvement doit être kinesthésique (c'est-à-dire focalisée sur la sensation que le mouvement réel produirait) plutôt que visuelle (basée sur l'observation du mouvement) (Stinear et al., 2006). Ce peut être une limitation du système, car la capacité à effectuer une stratégie donnée dépend du sujet.

1.3.3 L'imagerie visuo-spatiale

L'imagerie visuelle est la plus étudiée, car peut-être la plus communément expérimentée. D'après les études en IRMf, il existe un fort recouvrement entre les aires activées pendant l'imagerie visuelle et celles activées pendant la vision (Kosslyn et al., 2001).

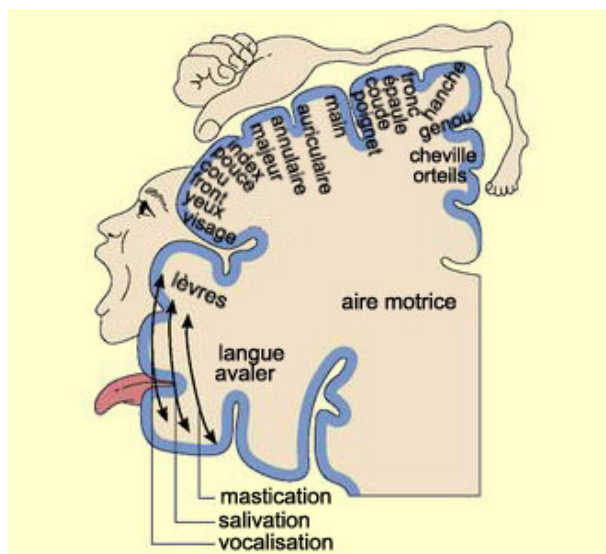


FIGURE 1.2 – Homuncule moteur de Penfield : coupe des régions du cortex moteur en correspondance avec les régions du corps qu'elles contrôlent. On remarque que la taille de la région correspondant à un membre n'est pas proportionnelle à la taille du membre, mais plutôt à la complexité des mouvements qui peuvent être effectués (image reprise de http://lecerveau.mcgill.ca/flash/i/i_06/i_06_cr/i_06_cr_mou/i_06_cr_mou.html).

Une étude MEG montre en particulier une suppression du rythme alpha dans le sulcus pariéto-occipital (Salenius et al., 1995). Certaines régions réalisant des traitements de percepts visuels spécifiques (l'aire fusiforme des visages par exemple) semblent être aussi activées dans les tâches d'imagerie. De même, l'imagerie de représentations spatiales semble faire participer les aires pariétales de la voie dorsale (spécialisées dans les traitements spatiaux). Cette thématique a surtout été abordée en étudiant le fonctionnement de la mémoire spatiale dans des tâches de navigation dans un environnement virtuel. Il semble que l'hippocampe ait un rôle prépondérant qui se manifeste par des activités dans la bande thêta observées sur des électrodes EEG intracorticales, celles-ci permettraient de relayer les informations mémorisées vers le cortex postérieur pariétal qui générerait les changements de point de vue dans la représentation spatiale⁵. Ces régions pariétales se retrouvent aussi activées dans les tâches de rotation mentale où le sujet doit simuler les transformations visuelles produites par un objet en mouvement (Kosslyn et al., 2001) comme le confirme une analyse en champs évoqués MEG, (Kawamichi et al., 2007).

5. En particulier, cette aire générerait les passages d'un point de vue allocentrique, où le sujet voit une scène de l'extérieur, à un point de vue égocentrique où le sujet se représente la scène comme s'il était dedans.

1.3.4 L'imagerie auditive

C'est une modalité d'imagerie qui a encore été très peu étudiée. Les travaux concernent essentiellement l'imagerie musicale. Une étude TEP montre que l'imagerie auditive active le gyrus temporal supérieur, le gyrus supramarginal et les cortex médians et inférieurs (Zatorre et al., 1996). D'autres études montrent que les aires auditives associatives du gyrus temporal supérieur ont un rôle dans l'imagination auditive. Une étude MEG en champs évoqués a mis en évidence cette activation au cours de l'imagerie auditive. Cependant, l'activation de régions visuelles et motrices a aussi été relevée dans la littérature. Cette association d'autres modalités n'est pas surprenante, en particulier chez des musiciens qui peuvent s'imaginer jouer la musique sur un instrument pendant qu'ils se la rappellent.

Les études EEG sur l'imagerie ne sont pas très développées peut-être parce que celle-ci n'est pas observable par un comportement visible (en anglais, on utilise le mot *covert* pour qualifier ces phénomènes cachés à l'observateur). Du point de vue des applications en ICM, ces états sont intéressants car ils ne nécessitent pas de stimulation externe (utilisation possible en asynchrone). Contrairement au choix fait dans les ICM basées sur l'imagerie motrice, utiliser des imageries dans des modalités variées peut permettre de contraster des aires sensorielles distinctes. Une limitation de ce choix étant l'interaction possible entre modalités d'imagerie. Ainsi, une imagerie motrice peut avoir une composante visuelle et une imagerie auditive une composante motrice, ce qui peut conduire à un recouvrement des aires activées. La littérature considère ce phénomène comme dépendant de la stratégie d'imagerie employée, mais il est difficile de savoir à quel point ces stratégies peuvent être modifiées par le sujet pour obtenir des configurations d'activités cérébrales différenciées. Il faut aussi préciser que l'imagerie n'est pas le seul moyen de modifier son état mental de manière interne : une autre possibilité est la modulation de l'attention à un stimulus externe.

1.3.5 Le calcul mental

Le calcul mental semble impliquer certaines parties du lobe frontal associées à des représentations linguistiques (Dehaene, 1999) et des régions pariétales qui pourraient correspondre à des traitements visuo-spatiaux (Zago et al., 2001). En EEG et en MEG, une augmentation des oscillations frontales dans la bande thêta a été mise en évidence (Sasaki et al., 1996).

Chapitre 2

Du signal à l'information transmise

2.1 Les méthodes de classification

2.1.1 Principe

L'application de la procédure de quantification à plusieurs essais, ou plusieurs fenêtres temporelles, permet d'obtenir un ensemble de points distribués dans l'espace des variables de quantification. Les ensembles de points qui correspondent aux différents états mentaux forment des nuages de points représentant chaque classe. L'étape suivante de classification consiste dans un premier temps à exploiter un ensemble d'essais appelé *ensemble d'apprentissage* pour partitionner l'espace de quantification d'une façon qui permet de séparer de manière optimale les classes les unes des autres. Cette phase est ainsi appelée phase d'apprentissage. La plupart des techniques usuelles (décrites ci-dessous) reviennent à ajuster une surface de séparation entre les nuages de chaque classe, appelée *frontière de décision* (en anglais *decision boundary*), entre les domaines associés à chaque classe. Une fois le classifieur entraîné (c'est-à-dire quand la phase d'apprentissage est terminée), les nouveaux essais peuvent être classés comme appartenant à une des classes suivant leur appartenance à un des domaines de l'espace de quantification. Dans cette étude, nous allons nous restreindre au cas de la classification binaire où seulement deux classes sont distinguées. Dans ce cas, la surface de décision est en général déterminée en utilisant une *fonction discriminante* (Duda et al., 2001) $f^*(\mathbf{x})$ avec la règle suivante : si $f^*(\mathbf{x}) > 0$ alors l'essai associé au vecteur de quantification \mathbf{x} est classé dans la classe 1, sinon l'essai est attribué à la classe 2. Ce principe de la classification binaire est résumé figure 2.1. Nous allons brièvement présenter quelques méthodes de classification ainsi que les problèmes usuellement rencontrés avec ces approches.

2.1.2 Quelques algorithmes de classification

- Deux considérations principales caractérisent un algorithme de classification :
- la frontière de décision peut être linéaire ou non linéaire

Espace de quantification

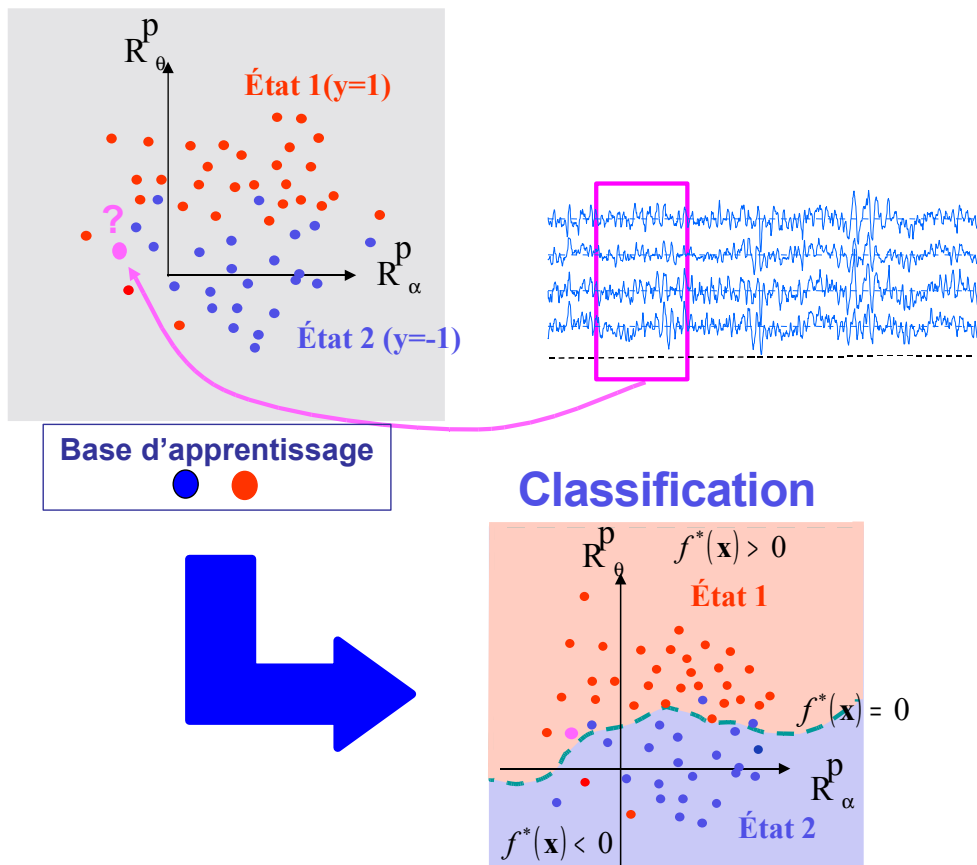


FIGURE 2.1 – Principe de la classification binaire : une base d'apprentissage est constituée à partir des points de quantification appartenant à chaque classe qui forment deux nuages. Une fonction discriminante f^* est ajustée à ces données de manière à partitionner l'espace en deux domaines distincts suivant le signe de f^* . La frontière entre les deux domaines est appelée frontière de décision.

- la fonction discriminante à ajuster peut résulter d’un modèle probabiliste (algorithmes génératifs ou basés sur un modèle) ou bien ils peuvent être exprimés directement en utilisant des points de données issus de l’ensemble d’apprentissage (algorithmes discriminants ou basés sur les données).

Nous allons présenter des algorithmes illustrant ces caractéristiques fondamentales.

Analyse linéaire discriminante

Les algorithmes de classification linéaire, qui représentent une grande partie des techniques disponibles, sont basés sur l’ajustement d’une fonction discriminante linéaire aux données. Cette fonction est de la forme $f(\mathbf{x}) = \boldsymbol{\omega}\mathbf{x} + b$ où b est le biais et $\boldsymbol{\omega}$ un vecteur orthogonal à la frontière de décision $f(\mathbf{x}) = 0$. Une approche pour ajuster cette fonction linéaire est l’*Analyse linéaire discriminante* (ALD), dont l’idée originale se trouve dans Fisher (1936). Cette méthode revient à ajuster aux données un modèle probabiliste gaussien (Duda et al., 2001) aux données afin de déterminer la classe la plus probable. Dans sa formulation la plus simple, en supposant :

- que le vecteur de quantification \mathbf{x} a une distribution gaussienne avec des moyennes conditionnelles à la classe différentes $\boldsymbol{\mu}_1$ et $\boldsymbol{\mu}_2$ pour les classes 1 et 2 respectivement et des matrices de covariances conditionnelles identiques $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$
- que les classes 1 et 2 ont des probabilités a priori égales $\pi_1 = \pi_2$

, la fonction de décision optimale dans un cadre bayésien est alors de la forme :

$$f_{ALD}^*(\mathbf{x}) = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}^{-1} \left(\mathbf{x} - \frac{\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2}{2} \right) = \boldsymbol{\omega}_{ALD} \mathbf{x} + b_{ALD} \quad (2.1)$$

Une illustration de cette technique sur des données jouets, dans la figure 2.2(a), montre la frontière de décision entre les deux nuages de points de l’ensemble d’apprentissage. L’ALD est un algorithme intéressant pour plusieurs raisons : il est rapide, simple à implémenter et à comprendre. En conséquence, il a été largement utilisé dans la communauté des ICM (Garrett et al., 2003; Bostanov, 2004; Scherer et al., 2004).

Séparateur à vaste marge linéaire

Contrairement à l’ALD, qui construit un modèle probabiliste de chaque classe en utilisant tous les points de données, le principe sous-jacent du *Séparateur à Vaste Marge linéaire* (SVM, en anglais Support Vector Machine) est de minimiser une erreur de séparation linéaire en se concentrant sur les points voisins de la frontière de décision (Vapnik, 1998). Les points bien classés qui sont positionnés suffisamment loin de la surface de séparation sont donc ignorés durant le processus d’apprentissage. Plus précisément, la fonction discriminante a la même forme que dans l’ALD : $f(\mathbf{x}) = \boldsymbol{\omega}\mathbf{x} + b$. Une marge est définie comme le domaine situé entre deux hyperplans encadrant la surface de décision, et dont les équations sont données par $f(\mathbf{x}) = 1$ et $f(\mathbf{x}) = -1$. Le SVM tend à maximiser la largeur de cette marge (d’une valeur de $\frac{2}{\|\boldsymbol{\omega}\|}$), ce qui revient à minimiser la norme euclidienne du vecteur $\boldsymbol{\omega}$ tout en maintenant les points de données à l’extérieur de la marge. Si l’on désigne par \mathbf{x}_i le i -ème vecteur de l’ensemble d’apprentissage et par

y_i la classe qui lui correspond ($y_i = 1$ pour la classe 1 et $y_i = -1$ pour la classe 2), alors fonction discriminante optimale $f^*(\mathbf{x}) = \boldsymbol{\omega}^* \mathbf{x} + b^*$ est la solution :

$$(\boldsymbol{\omega}^*, b^*) = \arg \min_{(\boldsymbol{\omega}, b)} C \sum_{\forall i} \epsilon_i + \frac{\|\boldsymbol{\omega}\|^2}{2} \quad (2.2)$$

$$\text{soumis à} \quad \begin{cases} y_i(\langle \boldsymbol{\omega}, \mathbf{x}_i \rangle + b) \geq 1 - \epsilon_i, \forall i \\ \epsilon_i \geq 0, \forall i \end{cases} \quad (2.3)$$

où $\sum_{\forall i} \epsilon_i$ est le terme d'erreur de séparation des classes, $\frac{\|\boldsymbol{\omega}\|^2}{2}$ est le terme de régularisation et C est un paramètre de régularisation défini par l'utilisateur (Muller et al., 2001). Dans la figure 2.2(b), qui illustre l'utilisation de ce classifieur sur des données jouets, certains points au voisinage de la surface séparatrice sont appelés vecteurs de support et définissent les limites de la marge de chaque côté de la surface. Le SVM dispose de "bonnes" propriétés statistiques détaillées dans Vapnik (1998). De plus, cet algorithme a l'avantage d'être une méthode à noyau : il est ainsi possible de rechercher la fonction discriminante sous la forme :

$$f_{\boldsymbol{\alpha}, b}^*(\mathbf{x}) = \sum_i \alpha_i y_i K(\mathbf{x}, \mathbf{x}_i) + b$$

où K est un noyau défini positif. Dans le cas linéaire, K correspond au produit scalaire usuel, mais il est aussi possible d'utiliser d'autres noyaux comme le noyau RBF¹ ou un noyau polynomial² qui permettent de construire des fonctions discriminantes non linéaires. Dans tous les cas, la solution optimale est donnée par la résolution du problème dual

$$\arg \max_{\boldsymbol{\alpha}} \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j (\mathbf{K})_{ij}$$

sous les contraintes

$$0 \leq \alpha_i \leq C \quad (i = 1, \dots, n), \quad \sum \alpha_i y_i = 0 \quad (2.4)$$

où \mathbf{K} est la matrice de Gram des données d'apprentissage. Celle-ci est définie par $(\mathbf{K})_{ij} = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$ dans le cas linéaire et par $(\mathbf{K})_{ij} = K(\mathbf{x}_i, \mathbf{x}_j)$ dans le cas général. La résolution de ce problème fait appel à un algorithme de programmation quadratique (Altman and Gondzio, 1998).

K plus proches voisins

Une autre méthode guidée par les données est l'algorithme des *K plus proches voisins* (KPPV, en anglais K Nearest Neighbor) (Hastie et al., 2001). Le principe de cette méthode est le suivant : l'étiquette de classe d'un point est la classe prédominante

-
1. $K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma^2}\right)$
 2. $K(\mathbf{x}_i, \mathbf{x}_j) = (\langle \mathbf{x}_i, \mathbf{x}_j \rangle + 1)^d$

dans l'ensemble des K plus proches voisins du point considéré appartenant à l'ensemble d'apprentissage (K est défini par l'expérimentateur). Les KPPV ne reposent sur aucun modèle, mais sont dépendants de la métrique utilisée pour mesurer la proximité (ou similarité) de deux paires de points (usuellement la distance euclidienne). La fonction discriminante associée à cet algorithme est de la forme :

$$f_{KPPV}^*(\mathbf{x}) = \sum_{n \in N_K(\mathbf{x})} y_n \quad (2.5)$$

où $N_K(\mathbf{x})$ est l'index des K plus proches voisins de \mathbf{x} dans l'ensemble d'apprentissage. D'un point de vue algorithmique, les KPPV ne nécessitent pas d'apprentissage car il n'y a aucun paramètre à ajuster ; cependant, l'ensemble d'apprentissage doit être conservé en mémoire dans sa totalité pour classer de nouvelles données. En comparaison avec d'autres méthodes linéaires, la surface de décision $f_{KPPV}^*(\mathbf{x})$ peut être fortement non linéaire comme illustré dans l'exemple figure 2.2(c).

Une large variété d'autres méthodes ont été utilisées en ICM, comme les réseaux de neurones (Hiraiwa et al., 1990; Anderson et al., 1998; Haselsteiner and Pfurtscheller, 2000) ou les modèles de Markov cachés (Obermeier et al., 2001). Du point de vue des applications en ICM, un avantage des réseaux de neurones comme le perceptron multicouche est la capacité de pouvoir entraîner le classifieur sur de nouveaux points de données sans recommencer la phase d'apprentissage en entier. Ceci peut être utilisé pour améliorer en temps réel le transfert d'information dans une session de feedback. Cependant, la phase initiale d'apprentissage de ces méthodes a un coût de calcul relativement élevé et qui augmente vite avec le nombre de variables. Une revue du domaine de la classification pourra être trouvée dans (Duda et al., 2001) ou (Hastie et al., 2001), et une revue exhaustive des algorithmes utilisés en ICM pour la classification des signaux EEG est dressée dans (Lotte et al., 2007a).

2.2 Évaluer la performance des classifieurs

2.2.1 Les mesures de performance

L'apprentissage d'un classifieur a pour but de minimiser l'erreur de classification généralement mesurée par le *taux de bien classés* ou *précision de classification* (PC, en anglais classification accuracy). Afin de l'interpréter, il faut prendre pour référence le pourcentage atteint par une classification aléatoire. Lorsque l'on essaie de classer un essai en choisissant aléatoirement entre deux classes, on atteint un taux moyen de biens classés de 50%. De même, lorsque l'on traite d'un problème à 3 classes, le taux de biens classés par chance est de 33.3%. Pour quantifier le niveau de performance atteint, le coefficient kappa quantifie la distance entre le taux de biens classés obtenu et le niveau de la chance (Townsend et al., 2006).

Une autre manière de mesurer la performance est l'*aire sous la courbe de ROC* (AUC). La *courbe de ROC* (Receiver Operating Characteristic) représente l'évolution

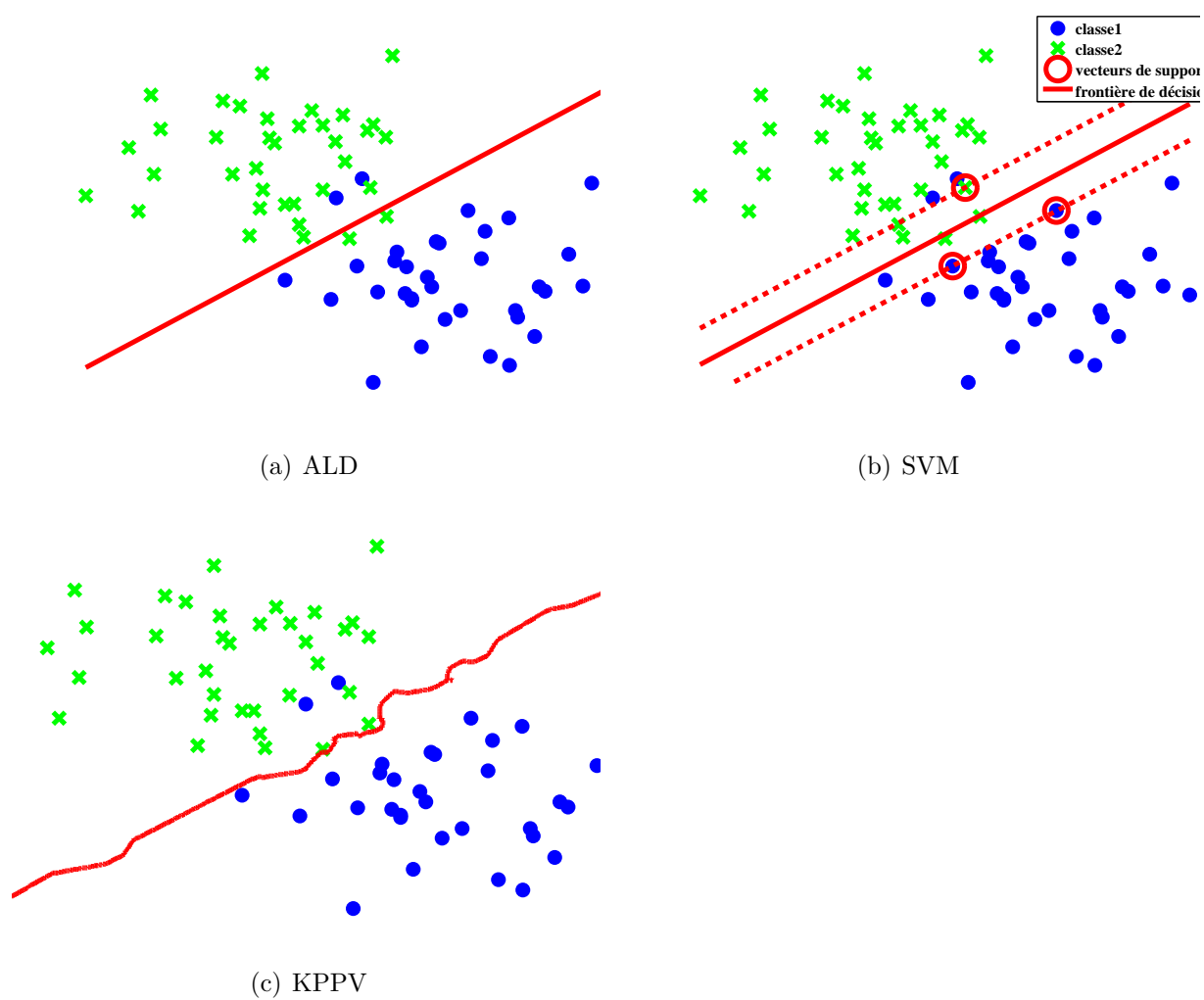


FIGURE 2.2 – Frontières de décision de trois méthodes de classification calculées à partir d'un ensemble d'apprentissage en dimension 2.

du taux de faux positifs en fonction du taux de vrais positifs. Ces deux variables sont obtenues par un seuillage de la fonction discriminante à différentes valeurs (Green and Swets, 1974; Duda et al., 2001). Une valeur de la fonction discriminante au dessus du seuil prédit la classe 1 et une valeur en dessous du seuil prédit la classe 2; donc à une valeur extrême de la courbe tous les essais sont attribués à la classe 1 tandis qu'à l'autre extrémité ils sont tous dans la classe 2. En pratique, cette mesure permet un contrôle du niveau acceptable de faux négatifs qui garantit la fiabilité d'un diagnostic. En résumé, l'aire sous la courbe de ROC est une mesure globale du pouvoir discriminant de la fonction discriminante, quel que soit le seuil choisi. Si l'aire sous cette courbe atteint 1, la discrimination est considérée comme parfaite.

Finalement pour les applications en ICM, il est évident que la performance d'un classifieur doit être interprétée en fonction de la longueur de la fenêtre temporelle de l'essai à classer. Par exemple, si l'on choisit deux classifieurs avec des taux de bien classés semblables, mais que le premier effectue la classification à partir d'une fenêtre d'une seconde seulement tandis que l'autre utilise deux secondes d'enregistrement EEG, alors le premier transfère deux fois plus d'informations sur l'état mental que le deuxième par unité de temps. De la même manière, à taux de bien classés égaux, un classifieur à quatre classes envoie deux fois plus d'informations binaires qu'un classifieur à deux classes pour le même nombre d'essais. Pour permettre la comparaison des taux de classification indépendamment des paramètres expérimentaux choisis (durée des essais et nombre de classes), la notion de taux de transfert d'information (ITR) a été empruntée au domaine de la théorie des communications numériques. Elle peut être calculée en utilisant le taux de biens classés p , le nombre de classes N et la durée d'un essai T d'après la formule :

$$ITR(\text{bits}/\text{minute}) = \frac{1}{T}(\log_2 N + p \log_2 p + (1 - p) \log_2 \frac{1 - p}{N - 1})$$

Cette équation permet la comparaison de toutes sortes de paradigmes d'ICM : d'après la littérature, la limite accessible actuellement par les ICM non invasives est d'environ 25 bits/minute (Krepki, 2006).

2.2.2 Généralisation

La notion de *généralisation* est cruciale pour vraiment quantifier la performance d'un classifieur : elle correspond à la capacité d'un classifieur, optimisé sur des données d'apprentissage, à classer fidèlement de nouvelles données (qui n'étaient pas dans l'ensemble d'apprentissage). Afin d'illustrer ce point, considérons un ensemble d'algorithmes d'apprentissages dotés d'un paramètre qui permet de régler leur fidélité aux données d'apprentissage. En augmentant la fidélité à ces données, l'erreur de classification diminue progressivement sur les données d'apprentissage³. Mais l'erreur de classification sur des données non apprises (données de validation) s'arrêtera de décroître et recommencera

3. Remarquons qu'il est tout à fait possible d'obtenir une erreur de classification nulle sur les données d'apprentissage. Par exemple, il suffit pour cela d'utiliser l'algorithme des 1PPV : le plus proche voisin de chaque point d'apprentissage étant lui-même, sa classe lui sera attribuée.

même à croître lorsque les caractéristiques des données d'apprentissage qui sont prises en compte par l'algorithme ne représentent plus des propriétés robustes des données ; c'est ce qu'on appelle le *sur apprentissage*. Ce scénario est illustré dans la figure 2.3 où l'erreur de validation E_{val} apparaît comme un outil intéressant pour choisir le meilleur réglage de la fidélité aux données. À cause de ce problème, la performance du classifieur doit toujours être calculée sur des *données de validation* ou *test* pour mesurer la capacité de généralisation. Les méthodes de mesure empirique de la capacité de généralisation

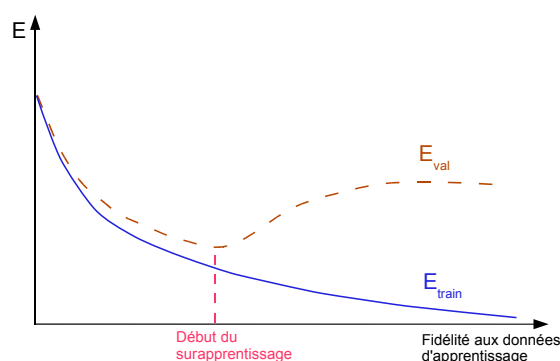


FIGURE 2.3 – Illustration du phénomène de surapprentissage : lorsque que l'on augmente la fidélité aux données d'apprentissage de l'algorithme, l'erreur sur les données d'apprentissage diminue progressivement, cependant, l'erreur sur des données de validation (différentes des données d'apprentissage) atteint un minimum. Au-delà de ce minimum l'augmentation de la fidélité aux données d'apprentissage détériore les capacités de généralisation de l'algorithme.

sont basées sur le principe de la *validation croisée* : priver l'ensemble d'apprentissage d'une partie des données afin de l'utiliser comme ensemble de validation. Cette partition est réalisée en général de manière répétitive pour différents échantillonnages de l'ensemble de données. Pour chaque échantillonnage le taux de biens classés est évalué sur l'ensemble de validation après entraînement sur l'ensemble d'apprentissage réduit, puis on calcule les taux moyens de biens classés à travers tous les échantillonnages qui ont été faits. Suivant la quantité de données qui sont retenues pour la validation, on distingue différents types de validation croisée (Hastie et al., 2001) :

- *validation croisée "leave one out"* : un élément de l'ensemble d'apprentissage est retiré à chaque échantillonnage
- *validation croisée "k fois"* : les essais de l'ensemble d'apprentissage sont partitionnés en k sous-ensembles de tailles égales, puis à chaque échantillonnage on retire (avec remise) un de ces sous-ensembles.

2.2.3 Complexité

La problématique de la généralisation a conduit les chercheurs à se pencher sur le problème suivant : comment réaliser un algorithme avec de bonnes capacités de généralisation ? C'est de là qu'est apparue la notion de complexité. En apprentissage statistique, elle quantifie la capacité d'une famille de fonctions discriminantes à séparer avec une fidélité parfaite les nuages de points de différentes classes lorsqu'ils sont très intriqués. Cette propriété est quantifiée par la dimension de Vapnik-Chervonenkis (Vapnik and Chervonenkis, 1971) d'un ensemble de fonctions \mathcal{F} ; celle-ci correspond au nombre maximal de points d'apprentissage qu'il est possible de séparer parfaitement par une fonction discriminante appartenant à \mathcal{F} , quelque soit la distribution des points et leur classe. En accord avec cette définition simplifiée, les fonctions discriminantes des KPPV ont une plus grande complexité que les fonctions affines de l'ALD. Plus les fonctions discriminantes sont complexes, mieux elles peuvent s'ajuster aux données d'apprentissage. Il a été démontré que le taux d'erreur sur un ensemble test d'un classifieur peut être majoré par une borne dont la valeur augmente avec la complexité de \mathcal{F} (Blumer et al., 1989). Ainsi, une trop grande complexité peut conduire au phénomène néfaste du surapprentissage. Pour les classifieurs linéaires, la complexité peut augmenter en fonction de la dimension de l'espace de quantification. C'est pourquoi en classification des données M/EEG, l'utilisation d'un grand nombre de paramètres de quantification peut poser des problèmes. Ce problème est aussi connu de façon générale sous le nom de *malédiction de la dimension* (Bellman, 1957), résumant le fait que l'estimation de paramètres dans un espace de grande dimension est très difficile. En effet, le nombre de données nécessaires pour mener à bien l'estimation croît très vite, afin de compenser le nombre de degrés de liberté supplémentaires qui sont ajoutés par chaque dimension supplémentaire. Afin de limiter l'effet de la haute dimension, plusieurs techniques sont présentées dans la partie suivante. Il faut cependant préciser qu'en plus de la complexité de l'ensemble de fonctions discriminantes \mathcal{F} du classifieur, la distribution des données d'apprentissage entre aussi en compte dans le phénomène de surapprentissage, il est ainsi possible d'affiner la notion de complexité en prenant en compte la structure empirique des données d'apprentissages, c'est le cas avec la mesure de complexité de Rademacher (Bartlett and Mendelson, 2002).

2.3 Les prétraitements

Pour simplifier l'espace de quantification, deux approches peuvent être considérées :
1) sélectionner un sous-ensemble de dimensions, les plus utiles pour la classification ou
2) trouver des directions privilégiées dans l'espace de quantification qui permettent de résumer l'ensemble des informations contenues dans les données. Ces deux approches sont présentées dans les sections suivantes.

2.3.1 Sélection de variables

La sélection de variables consiste à retenir un sous-ensemble de variables, optimal pour réaliser la classification (Guyon and Elisseeff, 2003). Parmi celles-ci, les plus élémentaires sont les méthodes de sélection univariées dites de “tri de variables” (en anglais *variable ranking*). Celles-ci quantifient le pouvoir discriminant de chaque variable une à une par un score. Le score le plus souvent utilisé dans le cas de deux classes étant le *score de Fisher*. Celui-ci correspond au quotient de la variance interclasses sur la variance intraclasse d’un nuage de points. Dans le cas d’une classification binaire univariée, soient (\bar{X}_1, \bar{X}_2) les moyennes empiriques et $(\hat{\sigma}_1^2, \hat{\sigma}_2^2)$ les variances empiriques conditionnelles à chaque classe pour la variable de quantification X , alors pour des classes de même effectif, le score de Fisher vaut :

$$F_{12} = \frac{(\bar{X}_1 - \bar{X}_2)^2}{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}$$

Les variables possédant le pouvoir discriminant le plus élevé sont alors conservées pour la classification. Ces méthodes ont deux inconvénients :

1. lorsque le nombre de variables est important, certaines variables peuvent obtenir un score élevé par un effet purement aléatoire. Il est possible de déterminer un seuil de significativité qui prenne en compte ce phénomène dû à des comparaisons multiples (par correction de Bonferoni ou par test de permutation (Pantazis et al., 2003)). Cependant, plus le nombre de variables augmente, plus le pouvoir discriminant d’une variable doit être important pour le distinguer d’un effet aléatoire.
2. la corrélation entre les variables n’est pas prise en compte, la méthode peut ainsi sélectionner des variables très redondantes alors que certaines, moins discriminantes, auraient apporté des informations complémentaires (Guyon and Elisseeff, 2003).

Pour sélectionner directement un sous-ensemble de variables ayant un pouvoir discriminant optimal, des méthodes “ gloutonnes ” d’accumulation ou d’élimination récursives de variables ont été proposées comme la *stepwise LDA* (Jenrich, 1960) ou le *SVM-RFE* (Rakotomamonjy, 2003). Il s’agit soit d’accroître un sous-ensemble de variables en choisissant la variable supplémentaire la plus informative, ou à l’inverse de réduire le sous-ensemble en excluant la variable qui par son absence provoque le moins de pertes d’information. Ces méthodes ont un apport certain par rapport aux méthodes de sélection univariées dans le sens qu’elles sont moins susceptibles de sélectionner des variables redondantes. Cependant, elles nécessitent de tester l’apport d’informations discriminantes de l’ensemble des variables disponibles à chaque itération et sont donc coûteuses en temps de calcul.

Une dernière possibilité qui a été élaborée est d’inclure la sélection de variables directement dans le calcul du classifieur (embedded methods en anglais) en introduisant un terme pénalisant les fonctions discriminantes utilisant un nombre de variables élevé. Par exemple, pour l’algorithme l0SVM (Weston et al., 2003) le terme de norme L2 dans la fonction objectif du SVM est remplacé par un terme approchant une mesure

de norme 0. Ce choix favorise ainsi les fonctions de décisions ayant un faible nombre de coefficients non nuls dans le vecteur directeur ω . Ces méthodes ont ainsi l'avantage de minimiser directement un coût de classification, tout en évitant le temps de calcul dû au parcours systématique de l'espace des variables (Guyon and Elisseeff, 2003).

2.3.2 Filtrage spatial

Au lieu de réduire le nombre de dimensions dans la base canonique définie par les variables de quantification, il peut être envisagé de décomposer le signal EEG dans une base adaptée aux données, de manière à extraire les caractéristiques informatives du signal original sur un faible nombre de composantes. Soit M la matrice des mesures EEG (nombre de capteurs \times temps), alors le décours temporel des composantes \mathbf{c} obtenues par la matrice de filtrage spatial θ est $\mathbf{C} = \theta M$.⁴ La méthode de filtrage spatial la plus répandue est l'*analyse en composantes principales* (ACP). Celle-ci est une transformation unitaire qui permet de concentrer l'énergie du signal sur un faible nombre de composantes. Une autre technique de filtrage spatial est l'*analyse en composantes indépendantes* (ACI), introduite dans Hérault et al. (1985) et formalisée dans Comon (1994). Elle consiste à optimiser la matrice de filtrage spatial de façon à maximiser un critère d'indépendance mutuelle des composantes de sortie. Plusieurs critères d'indépendance peuvent être utilisés, par exemple FastIca (Hyvärinen and Oja, 1997) exploite un critère de non gaussianité des signaux : la négentropie, tandis que JADE (Cardoso and A. Souloumiac, 1993) exploite les cumulants d'ordre supérieur et Infomax (Bell and Sejnowsky, 1997) utilise un critère d'information mutuelle. Ces méthodes permettent de résoudre, sous certaines conditions, un problème de séparation de sources qui consiste à retrouver des signaux "sources" originaux qui ont été mélangés linéairement. De façon générale, l'ACI exploite des statistiques d'ordre supérieur à deux pour réaliser cette séparation. Cependant, il est aussi possible d'exploiter les propriétés de corrélation temporelle de signaux colorés pour inverser un mélange de sources si leurs spectres sont différents (linéairement indépendants). C'est le cas de la séparation aveugle du second ordre (en anglais Second Order Blind Identification : SOBI) (Belouchrani et al., 1997). Cette possibilité est particulièrement intéressante si l'on considère qu'une tâche mentale est caractérisée par des oscillations apparaissant dans des bandes de fréquences et des régions spécifiques. Chaque oscillation pourrait ainsi être considérée comme un signal "source" qui pourrait être isolé de l'activité EEG de fond par un algorithme tel que SOBI, pouvant ainsi permettre une meilleure quantification de la tâche mentale. C'est cette approche que nous envisageons au chapitre 6. La méthode de filtrage spatial la plus répandue dans le domaine des ICM est celle des *patterns spatiaux communs* (en anglais *Common Spatial Patterns* : CSP) (Muller-Gerking et al., 1999). Le principe est de trouver des filtres spatiaux qui maximisent la puissance obtenue en sortie par une classe

4. Cette transformation linéaire peut être appliquée aussi dans l'espace de quantification. Cependant, le signal récolté sur les capteurs EEG étant un mélange linéaire des activités des sources corticales (voir chapitre 3), il est plus adapté d'effectuer le changement de base sur le signal brut avant la quantification.

(correspondant à un état mental) par rapport à celle obtenue dans l'autre classe. Soient \mathbf{C}_1 et \mathbf{C}_2 les matrices de covariance des signaux EEG conditionnés à la classe. Pour des classes équilibrées (de moyenne nulle), la covariance totale est $\mathbf{C}_t = \frac{1}{2}(\mathbf{C}_1 + \mathbf{C}_2)$. Le filtre CSP est obtenu à partir de ces quantités en deux étapes.

1. blanchiment : soit la décomposition en valeurs et vecteurs propres $\mathbf{C}_t = \mathbf{U}_t \mathbf{D}_t \mathbf{U}_t^T$, on calcule la matrice de blanchiment $\mathbf{C}_t^{-1/2} = \mathbf{U}_t (\mathbf{D}_t)^{-1/2} \mathbf{U}_t^T$. En appliquant cette matrice de blanchiment aux données, on obtient des données blanchies de matrice de covariance totale $\tilde{\mathbf{C}}_t = \mathbb{I}$ et de matrices de covariances conditionnelles $\tilde{\mathbf{C}}_1 = \mathbf{C}_t^{-1/2} \mathbf{C}_1 \mathbf{C}_t^{-1/2}$ et $\tilde{\mathbf{C}}_2 = \mathbf{C}_t^{-1/2} \mathbf{C}_2 \mathbf{C}_t^{-1/2}$.
2. Diagonalisation conjointe : Comme les nouvelles covariances vérifient $\frac{1}{2}(\tilde{\mathbf{C}}_1 + \tilde{\mathbf{C}}_2) = \mathbb{I}$, alors elles sont diagonalisables dans la même base \mathbf{V} vérifiant $\tilde{\mathbf{C}}_1 = \mathbf{V}(\tilde{\mathbf{D}}_1)\mathbf{V}^T$ et $\tilde{\mathbf{C}}_2 = \mathbf{V}(\tilde{\mathbf{D}}_2)\mathbf{V}^T$. Pour un même vecteur propre, la valeur propre (qui correspond à la puissance du signal dans la direction de ce vecteur propre) λ^2 de $\tilde{\mathbf{C}}_2$ vaut $2 - \lambda^1$, où λ^1 est la valeur propre de $\tilde{\mathbf{C}}_1$. Ainsi, lorsqu'une valeur propre est maximale pour $\tilde{\mathbf{C}}_1$ elle est minimale pour $\tilde{\mathbf{C}}_2$ et inversement. Les filtres spatiaux choisis pour quantifier les activités sont donc de la forme $\mathbf{v}_k^T \mathbf{C}_t^{-1/2}$ où \mathbf{v}_k est un vecteur propre de $\tilde{\mathbf{C}}_1$ de valeur propre extrême.

Différents prétraitements par filtrage spatial linéaire ont été comparés dans [Naeem et al. \(2006\)](#) dans le cadre des ICM. Les résultats obtenus montrent une supériorité du CSP sur les autres techniques telles que FastICA ou SOBI. Cependant, ces résultats sont biaisés par le fait que, contrairement aux autres filtres spatiaux, le CSP est une méthode *supervisée*, c'est-à-dire qu'elle utilise non seulement les signaux EEG enregistrés, mais aussi l'information sur la classe de chacun de ces signaux pour séparer au mieux les données.

2.3.3 Trouver la bonne quantification

Tout comme la recherche de caractéristiques fonctionnelles (puissance, synchronie, cohérence) dans les signaux EEG, la sélection de variables et le filtrage spatial participent à la démarche générale de trouver la meilleure quantification aux données EEG et MEG. Dans cette lignée, incorporer un maximum d'informations a priori sur le problème, comme par exemple un modèle biophysique génératif pour les signaux EEG, est une autre voie pour accroître encore le taux de biens classés. C'est l'approche qui est introduite au chapitre suivant.

Chapitre 3

Utiliser le problème inverse en ICM

Contrairement aux signaux d'IRMf, les signaux MEG et EEG récoltés à la surface du scalp ne permettent pas de localiser directement les sources corticales qui génèrent ces activités. Il faut pour cela recourir à des méthodes de reconstruction d'activité des sources corticales. Ces techniques procèdent en deux étapes :

- il s'agit d'abord de modéliser le champ magnétique ou le potentiel généré par l'activité de chaque source corticale. La question posée est alors : quel est le signal obtenu sur les capteurs MEG et EEG lorsqu'une source corticale est active ? La solution de ce *problème direct* utilise une modélisation physique et physiologique des tissus de la tête pour déduire une relation linéaire entre l'activité des sources corticales J et les signaux EEG et MEG engendrés m .
- réciproquement le problème inverse pose la question : quelle peut être l'activité des sources corticales sachant que les mesures MEG ou EEG valent m ? Contrairement au problème direct, le problème inverse ne peut admettre de solution unique, car il est largement sous-déterminé (cela est dû au fait que le nombre de capteurs est très inférieur au nombre de sources corticales). Des contraintes supplémentaires sont donc imposées au problème pour choisir une réponse particulière dans l'espace des solutions possibles.

Nous allons rappeler le principe de la modélisation du problème direct ainsi que les différentes solutions de problème inverse existantes.

3.1 Le problème direct en MEG et EEG

3.1.1 Origine physiologique de la MEG et de l'EEG

Les signaux électromagnétiques enregistrés à l'extérieur de la tête sont générés par des déplacements d'ions qui sont à la base de l'électrophysiologie. Plus précisément, les ions sont échangés au moyen de canaux ioniques à travers la paroi de cellules appelées neurones, provoquant localement des différences de concentration entre l'intérieur des neurones et l'extérieur (appelé milieu extracellulaire). Un neurone est constitué du *corps cellulaire*, ou soma, prolongé par un long canal nerveux : l'*axone*. Le neurone reçoit

des informations d'axones d'autres neurones par l'intermédiaire de synapses, reliées au corps par une arborescence de dendrites. L'essentiel des mouvements de charge a lieu pendant l'excitation du neurone : lorsque des informations arrivent sur les synapses de l'arbre dendritique, celles-ci libèrent des neurotransmetteurs provoquant un potentiel postsynaptique excitateur qui abaisse localement le potentiel par rapport au reste de la cellule. Cette différence de potentiel provoque des mouvements de charges du soma vers l'arbre dendritique à l'intérieur de la cellule, que l'on appelle courant primaire. Ce dernier est compensé par un deuxième mouvement de charges à l'extérieur de la cellule appelé courant secondaire (Gloor, 1985). Compte tenu du fort rapport d'échelle entre la distance du neurone à la surface du cortex et la longueur caractéristique sur laquelle s'étend le courant primaire (la distance entre l'arbre dendritique et le soma), le phénomène biophysique peut être modélisé de manière satisfaisante par un *dipôle de courant*. Ce modèle permet d'évaluer le champ électromagnétique généré par les mouvements de charges à la surface du cuir chevelu ; cependant, le champ théoriquement prédit par ce modèle pour un seul neurone est si faible qu'il ne serait pas mesurable sans la contribution d'un réseau plus étendu. En effet, le champ et le potentiel observés en MEG et EEG résultent principalement de l'activité de certaines cellules corticales, les *neurones pyramidaux* (illustrés figure 3.1), à cause de leur structure particulière (Nunez, 1981) :

- les neurones pyramidaux dans une portion de cortex de quelques millimètres reçoivent statistiquement les mêmes signaux afférents, ce qui fait que leurs potentiels post-synaptiques sont synchronisés temporellement. L'excitation de cette portion de cortex donne naissance à des courants primaires et secondaires simultanés de dizaines de millions de neurones.
- l'axe arbre dendritique-axone des neurones pyramidaux est toujours orienté perpendiculairement à la surface du cortex (Nunez and Silberstein, 2000) ; les courants primaires de ces cellules ont donc tous la même orientation dans une portion de cortex.

Si cette cohérence spatiale et temporelle des courants des cellules pyramidales permet d'observer en surface l'activité du cortex, il semble que d'autres organes du système nerveux central puissent aussi contribuer à l'activité MEG et EEG. C'est le cas de l'hippocampe (Tesche and Karhu, 2000), du thalamus (Butz et al., 2006; Gross et al., 2002; Timmermann et al., 2002) et du cervelet (Butz et al., 2006; Gross et al., 2002; Tesche and Karhu, 1997; Timmermann et al., 2002).

3.1.2 Modélisation par dipôle de courant

La mise en équation du problème direct fait appel aux équations de Maxwell dans leur forme quasi statique. Soit \mathbf{J}_p la densité volumique de courant primaire et $\mathbf{J}_s = \sigma \mathbf{E}$ la densité de courant secondaire due aux porteurs de charge dans le milieu extracellulaire de conductivité σ , alors l'équation de Maxwell pour le champ magnétique s'écrit

$$\mathbf{rot} \mathbf{B} = \mu_0 (\mathbf{J}_p + \sigma \mathbf{E})$$

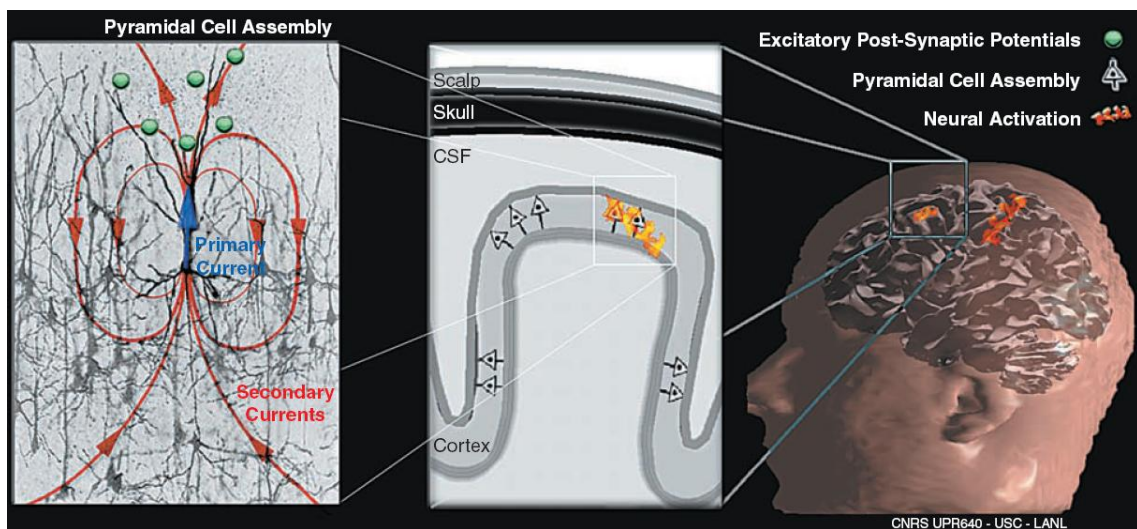


FIGURE 3.1 – Illustration de l'origine physiologique de l'EEG et de la MEG reprise de [Baillet et al. \(2001\)](#). Les techniques de résolution du problème inverse ont pour but d'estimer l'amplitude des courants primaires engendrés par les potentiels post-synaptiques au sommet de l'arbre dendritique des neurones pyramidaux. Ces courants primaires sont orientés perpendiculairement à la surface du cortex, ce qui rend l'activité des assemblées neuronales localement synchronisées mesurable à l'extérieur de la tête par l'EEG et la MEG.

En introduisant le potentiel vecteur \mathbf{A} et le potentiel scalaire V tels que

$$\mathbf{E} = -\mathbf{grad}V \quad (3.1)$$

$$\mathbf{B} = -\mathbf{rot}\mathbf{A} \quad (3.2)$$

$$\mathit{div}\mathbf{A} = 0 \quad (3.3)$$

on obtient les équations de Poisson

$$\Delta\mathbf{A} = \mu_0(\mathbf{J}_p + \sigma\mathbf{E}) \quad (3.4)$$

$$\Delta V = \frac{\mathit{div}\mathbf{J}_p}{\sigma} \quad (3.5)$$

Par application du principe de superposition, le potentiel électrique sur la i -ème électrode v_i et la composante du champ magnétique enregistrée sur le i -ème capteur MEG b_i dépendent linéairement de l'amplitude des courants primaires et peuvent être exprimés sous forme d'intégrales volumiques :

$$v_i = \int_{\mathcal{C}} \lambda_i^V(\mathbf{r}) \mathbf{J}_p(\mathbf{r}) dv(\mathbf{r}) \quad (3.6)$$

$$b_i = \int_{\mathcal{C}} \lambda_i^B(\mathbf{r}) \mathbf{J}_p(\mathbf{r}) dv(\mathbf{r}) \quad (3.7)$$

où \mathcal{C} est l'ensemble du volume occupé par le cortex. λ_i^B , ou respectivement λ_i^V , représentent le *champ sensibilité* (en anglais *lead field*) du magnétomètre (respectivement de l'électrode) i vis-à-vis de la distribution de courant sous-jacente. En général, la distribution de courant volumique est résumée par un nombre fini de dipôles de courants (des diracs de courant volumique) répartis de façon homogène sur le cortex. Cet artifice permet de représenter la distribution par un nombre fini de paramètres qui pourront être estimés. Il faut insister sur le fait que les dipôles ainsi constitués sont *macroscopiques* et représentent la superposition de dizaines de millions de *dipôles microscopiques* associés aux neurones pyramidaux présents dans les quelques millimètres cubes de cortex environnants. Si l'on note $\mathbf{q}(\mathbf{r}_i)$ le moment dipolaire (vecteur à trois composantes) à la i -ème position \mathbf{r}_i sur le cortex alors la relation du problème direct fait intervenir le champ de sensibilité sous forme matricielle :

$$v_i = \sum_{i=1}^N \lambda_i^V(\mathbf{r}_i) \mathbf{q}(\mathbf{r}_i) = \mathbf{\Lambda} \mathbf{q} \quad (3.8)$$

où l'on note abusivement par \mathbf{q} le vecteur contenant les paramètres des moments dipolaires macroscopiques à toutes les positions. Si N est le nombre de dipôles et M le nombre de capteurs alors la matrice de champ de sensibilité $\mathbf{\Lambda}$ est de dimension $(M \times 3N)$. Cependant, dans le cas où l'orientation des dipôles est contrainte a priori à être perpendiculaire à la surface corticale du cortex, seul un paramètre d'amplitude est à estimer pour chaque dipôle et $\mathbf{\Lambda}$ est une matrice $(M \times N)$. On peut obtenir l'expression

de \mathbf{A} sous forme analytique dans le cas d'un milieu conducteur homogène, cependant cette hypothèse n'est pas vérifiée ici, car les différents milieux traversés par le champ (cerveau, os et peau) ont des conductivités moyennes différentes. De plus, même en négligeant ce fait en première approximation, la conductivité des tissus humains est nettement plus élevée que l'air ambiant, qui est un isolant. Ces problèmes expliquent les méthodes de complexité croissante qui ont été élaborées dans le domaine :

- **Modèles de têtes sphériques** : il se trouve que les équations du problème direct ont une solution analytique dans le cas de milieux conducteurs séparés par des sphères concentriques. Les solutions les plus simples du problème direct utilisent donc des modèles de têtes sphériques que l'on ajuste à l'anatomie du sujet. On peut choisir d'utiliser une enveloppe représentant la tête, ou bien trois enveloppes pour prendre en compte les changements de conductivités aux interfaces cerveau/os et os/peau.
- **Modèles réalistes** : l'approximation tête sphérique est en général peu réaliste et peut conduire à de grosses erreurs de localisation des activités corticales. Il est possible d'exploiter les données anatomiques du sujet, fournies par un examen IRM ou un scanneur à rayons X, pour segmenter les différents tissus et obtenir un modèle réaliste de la tête. Cette stratégie est à la base de deux méthodes de résolution numérique du problème direct : la méthode des *intégrales de frontières* (Boundary Element Method, BEM) (Meijs et al., 1989; Yvert et al., 1995) qui utilise un maillage surfacique des différentes enveloppes, et la méthode par *éléments finis* (Finite Element Method, FEM) (Haueisen et al., 1995; Thevenet et al., 1991) basée sur un maillage volumique des tissus. Comme dans le modèle sphérique, un nombre plus ou moins important de tissus peut être modélisé ; cependant, le calcul de certains maillages à partir des données anatomiques est difficile à estimer (par exemple l'interface cerveau/os).
- Les modèles réalistes sont plus précis, mais sont plus coûteux en temps de calcul. Certains modèles proposent un choix intermédiaire. C'est le cas du **modèle à sphères multiples** (Ermer et al., 2001) qui au lieu de considérer un seul modèle de tête sphérique, ajuste une sphère différente pour chaque capteur, tangente à la surface du scalp au voisinage de celui-ci. De cette manière, un modèle réaliste est obtenu sans recourir à la résolution numérique des équations.

3.2 Le problème inverse

Le problème inverse dans le cadre du modèle dipolaire du problème direct, consiste à estimer la position l'orientation et l'amplitude des sources de courant primaire qui génèrent les enregistrements EEG et MEG sur un laps de temps donné. Les techniques utilisées se divisent en deux familles qui font appel à des choix de modélisation différents :

- les *méthodes dipolaires* ajustent un petit nombre de dipôles aux données. Il est généralement nécessaire d'estimer au préalable le nombre optimal de dipôles, après quoi la position, l'orientation et l'amplitude de ceux-ci doivent être ajustés par

des méthodes non linéaires. Dès lors que l'on ajuste plusieurs dipôles, la résolution du problème fait appel à des méthodes d'optimisation complexes et coûteuses en temps de calcul.

- les modèles de *sources distribuées* considèrent un grand nombre de dipôles répartis régulièrement dans une partie ou dans la totalité du volume cérébral. Ces modèles recherchent des solutions d'imagerie de l'activité des sources corticales où les régions actives sont localisées implicitement là où l'activité estimée des dipôles distribués a la plus forte amplitude. Afin de limiter le nombre de paramètres à estimer, il est possible de restreindre la position des dipôles au cortex et de contraindre leur orientation normalement à la surface corticale (il reste alors un seul paramètre scalaire d'amplitude à estimer au lieu de trois).

Cependant, une troisième famille de techniques, les méthodes de *balayage*, se place dans la continuité de l'approche dipolaire tout en donnant des résultats proches des méthodes distribuées. Nous allons tenter de résumer de manière cohérente l'ensemble des méthodes proposées qui sont d'une grande variété.

3.2.1 État de l'art

Les dipôles de courant équivalents

L'ajustement de dipôles équivalents à un vecteur de données mesurées \mathbf{m} est fondé sur le principe d'une régression non linéaire. Soit d le nombre de dipôles, il s'agit de trouver les paramètres de localisation $\mathbf{r}_0 = (r_1, \dots, r_d)$ et les moments des dipôles $\mathbf{q}_0 = (q_1, \dots, q_d)$ de courant qui minimisent l'écart entre les données reconstruites par le problème direct et les données observées \mathbf{m} , en utilisant un critère des moindres carrés :

$$(\mathbf{r}_0, \mathbf{q}_0) = \arg \min_{(\mathbf{r}, \mathbf{q})} \|\mathbf{m} - \mathbf{\Lambda}(\mathbf{r})\mathbf{q}\| \quad (3.9)$$

Dans cette équation, la matrice $\mathbf{\Lambda}(\mathbf{r})$ est le champ de sensibilité associé aux positions \mathbf{r} . Compte tenu de la dépendance non linéaire de $\mathbf{\Lambda}(\mathbf{r})$ vis-à-vis de la position des dipôles, il n'y a pas de solution analytique et il est nécessaire de recourir à des méthodes d'optimisation qui permettent de parcourir l'espace des d-uplets de positions possibles (simplexe, Gauss-Newton, ou Newton-Raphson). Bien entendu, la difficulté du problème augmente fortement avec le nombre de dipôles à ajuster. Le problème se complique aussi lorsque le modèle doit être adapté à des données qui dépendent du temps : on choisit en général un modèle où la position des dipôles reste fixe tandis que l'orientation évolue en fonction du temps (cette hypothèse est compatible avec le fait que l'orientation de la surface corticale varie beaucoup sur de faibles distances, et, dans le cas du champ magnétique, modifie grandement la répartition du champ à l'extérieur de la tête) (Scherg and Cramon, 1985). Ce modèle est appelé aussi modèle de dipôles spatio-temporels. Une autre extension des dipôles de courant équivalents prend en compte la statistique du bruit de mesure pour estimer les paramètres de manière plus robuste, c'est l'approche par moindres carrés généralisés qui revient à introduire une matrice de pondération des termes d'erreurs de l'équation 3.9. De façon générale, les modèles de dipôles équivalents

ont l'inconvénient de nécessiter la connaissance a priori du nombre de dipôles à modéliser et constituent un modèle très simplifié de l'activité corticale, peu susceptible de la représenter de manière réaliste en essai unique.

Les méthodes de balayage linéaires

Pour obtenir une modélisation plus réaliste et ne pas déterminer a priori le nombre de dipôles, des méthodes de balayage ont été élaborées. Il s'agit de déterminer la contribution des dipôles, placés en chaque voxel du volume cérébral, à l'activité enregistrée à la surface en les considérant un par un¹. Parmi ces techniques, on trouve les beamformers et la méthode MUSIC (Moshier et al., 1992). Les beamformers associent à chaque dipôle considéré un filtre spatial qui a pour but d'estimer son niveau d'activation à partir des données enregistrées en surface. Le beamformer le plus simple est linéairement contraint à variance minimale (en anglais Linearly Constrained Minimum Variance : LCMV)(Van Veen et al., 1997) : il s'agit de déterminer un vecteur de filtrage spatial des données qui assure un gain unitaire à l'activité du dipôle considéré tout en ayant une énergie de sortie minimale, ce qui a pour but de minimiser la contribution des autres sources corticales. Pour chaque dipôle i , le filtre linéaire $\mathbf{w}^*(\mathbf{r}_i)$ est donc obtenu par la minimisation sous contrainte suivante (où $\mathbf{\Lambda}(\mathbf{r}_i)$ est le champ de sensibilité à la position \mathbf{r}_i) :

$$\mathbf{w}^*(\mathbf{r}_i) = \arg \min_{\mathbf{w}} \|\mathbf{w}\mathbf{m}\mathbf{m}^T\mathbf{w}^T\| \quad \text{sous la contrainte} \quad \mathbf{w}\mathbf{\Lambda}(\mathbf{r}_i) = 1 \quad (3.10)$$

En appliquant ces filtres spatiaux aux données, on obtient la contribution de chaque dipôle aux signaux et les dipôles qui obtiennent le meilleur score sont sélectionnés pour modéliser les données. Des techniques non linéaires comme la magnétométrie à ouverture synthétique (Synthetic Aperture Magnetometry : SAM) ont aussi été envisagées dans la littérature. Enfin, la méthode MUSIC (Multiple signal classification), utilise au préalable une décomposition des données en un sous-espace signal et un sous-espace bruit (par décomposition en éléments simples) avant de balayer l'espace des dipôles corticaux en recherchant ceux qui contribuent le plus à l'espace signal.

Le défaut majeur que l'on attribue à ces modèles est leur incapacité à estimer l'activité de sources corrélées. Ce défaut est une conséquence du principe fondamental de la méthode qui est de considérer chaque dipôle indépendamment des autres. Les modèles de sources distribués corrigent ce problème en estimant globalement l'activité de toutes les sources.

Les modèles de sources distribuées

Au lieu de chercher un faible nombre de dipôles actifs ou de les balayer un par un, les modèles distribués fournissent une image en 3 dimensions de l'activité cérébrale².

1. on évite ainsi de parcourir l'ensemble des d-uplets de dipôles
2. on parle aussi de tomographie de l'activité cérébrale

Les principales méthodes qui ont été employées consistent toutes à résoudre un problème de régression avec un terme de régularisation quadratique qui permet d'obtenir une solution qui dépend linéairement des mesures \mathbf{m} (voir section suivante). L'inconvénient des méthodes distribuées est l'obtention d'activations dans un grand nombre de régions du cortex. Des méthodes statistiques de seuillage permettent de trouver parmi l'ensemble des régions celles qui ont des activations significatives (Pantazis et al., 2003). Les méthodes linéaires produisent en général des solutions de faible résolution spatiale et biaisées. Pour les améliorer, il existe des méthodes itératives qui raffinent progressivement la résolution spatiale de la solution (Cottareau et al., 2007; Gavit et al., 2001).

Dans le cadre des ICM, le problème inverse doit être appliqué aux données en temps réel. Cela impose l'utilisation de méthodes dont le temps de calcul de la distribution de courants à partir des signaux des capteurs est très rapide. C'est essentiellement le cas des méthodes distribuées linéaires. Nous allons détailler dans le paragraphe suivant ces méthodes linéaires de résolution du problème inverse.

3.2.2 Solutions linéaires

Les solutions linéaires des modèles distribués sont optimales pour un problème inverse régularisé dont la fonction objectif est quadratique (Baillet et al., 2001). Ils prennent la forme :

$$\hat{\mathbf{q}} = \arg \min_{\mathbf{q} \in \mathbb{R}^N} \|\mathbf{m} - \mathbf{\Lambda} \mathbf{q}\|^2 + \alpha \|\mathbf{F} \mathbf{q}\|^2 \quad (3.11)$$

où F est une matrice. Le premier terme de 3.11 est présent pour que la solution soit fidèle aux données tandis que le second terme tend à renforcer certaines propriétés souhaitables de la solution qui résument des connaissances *a priori* sur la solution. Le compromis entre les deux termes est réglé par le coefficient de régularisation α qui tend à renforcer la fidélité aux données quand il est petit ou à renforcer l'*a priori* quand il est grand. Les différents choix qui ont été faits dans la littérature pour le terme de régularisation sont :

- $\mathbf{F} = \mathbb{I}_N$ pour l'estimation à *contrainte de norme minimale* (en anglais Minimum Norm Estimate : MNE)
- $\mathbf{F} = \mathbf{\Delta}$ qui est une formulation discrète de l'opérateur Laplacien, imposant ainsi une contrainte de régularité spatiale à la solution.
- $\mathbf{F} = \text{diag}(\|\mathbf{\Lambda}_i\|)^{-p}$ pour le minimum norme pondéré par la norme des colonnes de la matrice de champ de sensibilité. Cette technique permet de rehausser l'amplitude estimée des dipôles de courant dont l'activité observée à la surface de la tête est fortement atténuée (Fuchs et al., 1999).

La solution générale de ce problème est de la forme :

$$\hat{\mathbf{q}} = (\mathbf{F}^T \mathbf{F})^{-1} \mathbf{\Lambda}^T (\mathbf{\Lambda} (\mathbf{F}^T \mathbf{F})^{-1} \mathbf{\Lambda}^T + \alpha \mathbb{I}_n)^{-1} \mathbf{M} = \mathbf{W} \mathbf{m} \quad (3.12)$$

La particularité de telles solutions est qu'elles ne nécessitent qu'une multiplication matricielle des données capteurs avec une matrice de problème inverse \mathbf{W} qui ne

varie pas en fonction du temps³. Cette propriété est très intéressante pour réaliser des calculs temps réel des activités corticales. À ces techniques s'ajoutent des méthodes de pondération a posteriori de l'amplitude des dipôles de courants, ces techniques ont pour but de diminuer l'erreur de localisation des activités que l'on mesure par la position du dipôle qui atteint le maximum d'amplitude (Pascual-Marqui, 2002; Dale et al., 2000).

3.2.3 Imagerie corticale des phénomènes de synchronisation

Le paradigme ICM asynchrone que nous mettons en oeuvre impose de travailler en essai unique sur les aspects oscillatoires du signal EEG ou MEG. Les diverses méthodes présentées ci-dessus ont en général été conçues pour traiter des données de potentiels évoqués (donc moyennées) et peuvent être, dans certains cas, adaptées à l'imagerie des activités oscillatoires.

Localisation d'oscillations

Dans le cas de l'analyse des synchronisations locales, les principales modifications rapportées dans la littérature passent par un filtrage dans une bande de fréquence préalablement au problème inverse.

- dans Salmelin and Hari (1994) les auteurs ajustent au cours du temps la position d'un dipôle aux signaux non moyennés filtrés dans la bande alpha. Ils obtiennent ainsi un ensemble de positions de dipôles qui se focalisent autour de la région corticale d'où proviennent les oscillations.
- dans Gross et al. (2001) une nouvelle méthode est introduite, baptisée DICS (Dynamic Imagery of Cortical Oscillations). L'approche est inspirée du LCMV, à la différence que l'optimisation est réalisée à une fréquence particulière à partir d'une matrice de densité spectrale, qui représente la matrice de covariance des signaux à cette fréquence.
- des solutions de type minimum norme ont aussi été appliquées à des représentations temps fréquence des signaux.

Localisation d'interactions

Si l'on s'intéresse aux phénomènes d'interactions à distance, on est conduit le plus souvent à travailler autour d'une fréquence donnée en mesurant la cohérence ou la synchronie de phase (voir section 1.2.2). Bien que l'on puisse s'inspirer des méthodes d'imagerie des synchronisations locales, il faut considérer dans ce cas l'ensemble des couples de sources corticales possibles. Par exemple, pour 10000 dipôles répartis sur le cortex on dénombre environ $(10000)^2/2$ couples de dipôles. Cette approche est trop coûteuse en temps de calcul surtout pour des applications temps réel. Dans Gross et al. (2001), pour calculer la cohérence les auteurs proposent trois solutions pour résoudre ce problème : 1) calculer seulement l'interaction des sources avec une référence externe

3. Car elle ne dépend que de l'anatomie du sujet

(par exemple un signal électromyographique), 2) calculer l'interaction des sources avec la référence du maximum de puissance dans la bande considérée, 3) faire une estimation préalable au niveau des capteurs, sélectionner les sources au voisinage des capteurs d'intérêt. Ce type de méthode a été adapté aux problèmes inverses distribués dans [Jerbi et al. \(2007\)](#). Une autre possibilité proposée par [David and Garnero \(2002\)](#) est l'utilisation d'un algorithme itératif pour réduire le nombre de régions d'intérêt entre lesquelles la synchronie est calculée. Plus récemment [Grasman et al. \(2004\)](#) proposent l'estimation conjointe d'un modèle à plusieurs dipôles et des interactions entre eux. Enfin, [Dossevi et al. \(2008\)](#) proposent d'appréhender le réseau d'interactions par les premiers vecteurs propres de la matrice de corrélation (ou de cohérence). Parmi toutes ces approches, les solutions de [Gross et al. \(2001\)](#) et [Dossevi et al. \(2008\)](#) sont les moins coûteuses en temps de calcul, au prix d'une vision plus restreinte du réseau d'interactions.

3.2.4 Le problème inverse dans les ICM

L'application du problème inverse dans le cadre des ICM a pour l'instant fait l'objet de quelques travaux. Dans [Qin et al. \(2004\)](#) les auteurs réalisent un positionnement de dipôles corticaux dans l'aire de la main des cortex moteurs droit et gauche puis reconstituent l'activité de ces deux dipôles correspondant à des tâches d'imagerie. D'autres études ([Lotte et al., 2007b](#); [Congedo et al., 2006](#)) utilisent la méthode distribuée sLORETA ([Pascual-Marqui, 2002](#)) pour estimer l'activité oscillatoire locale des sources impliquées dans un protocole d'anticipation de mouvement et trouvent des résultats compétitifs par rapport aux vainqueurs de la compétition BCI 2000 sur le même jeu de données. Le problème du grand nombre de sources à estimer est résolu soit par filtrage spatial de type CSP, soit par une méthode originale de localisation de région d'intérêt floue dans le domaine espace-fréquence des activités corticales.

Dans [Grave de Peralta Menendez et al. \(2005\)](#), le problème inverse ELECTRA est utilisé pour estimer des potentiels de champs locaux et différencier deux tâches de tapotement, l'une avec la main droite et l'autre avec la main gauche. D'après les auteurs, l'utilisation du problème inverse conduit à une amélioration du taux de classification de 5% chez deux sujets. Enfin, dans [Cincotti et al. \(2008\)](#), les auteurs appliquent un problème inverse de norme minimale et localisent les activités discriminantes entre les imageries main droite et main gauche sur le cortex. Pour l'instant, l'ensemble des publications fait état de peu de comparaisons quantitatives entre les ICM par problème inverse et les ICM traditionnelles (approche par filtrage spatial ou capteurs). En particulier, ces méthodes n'ont été appliquées que sur des protocoles synchrones d'imagerie motrice et de mouvement réel pour lesquels on a de forts a priori de localisation qui permettent d'utiliser des méthodes dipolaires. Il serait donc intéressant de quantifier l'apport généré par un problème inverse distribué dans des tâches mentales asynchrones et impliquant aussi de plus vastes aires corticales.

De façon générale, il est naturel de s'interroger sur l'intérêt d'un prétraitement par problème inverse linéaire par rapport aux méthodes déjà existantes de filtrage spatial

de type analyse en composantes indépendantes. Nous avançons ici des arguments en faveur de l'exploration de cette voie :

- bien que le problème inverse réalise un filtrage spatial linéaire, le résultat obtenu peut ensuite être traité de manière non linéaire. Il est par exemple possible de s'intéresser à la localisation spatiale des activités ou de quantifier des interactions à distance. Ces deux aspects ne sont pas envisageables de manière directe lorsqu'on décompose les données en différents modes spatio-temporels⁴.
- la décomposition du problème inverse est basée sur des a priori physiques et anatomiques indépendants du signal enregistré, tandis que les méthodes de séparation de sources utilisent des estimations des propriétés statistiques des données (matrice de covariance) qui sont sensibles au bruit. Ceci est renforcé par le fait que l'on dispose en général de peu d'enregistrements en comparaison du nombre de paramètres statistiques à estimer. Les approches statistiques peuvent donc être bien loin de leurs performances asymptotiques.
- la comparaison des performances en classification obtenues par différentes méthodes de problème inverse peut permettre d'améliorer la résolution spatiale de ces méthodes. De plus, les localisations discriminantes obtenues peuvent être interprétées d'un point de vue neurophysiologique.

Ces justifications montrent qu'il est donc raisonnable d'envisager les techniques de reconstruction de sources comme une alternative potentielle aux méthodes déjà utilisées.

4. En particulier, les signaux obtenus par PCA ou ICA sont supposés spatialement indépendants, il est donc mal justifié de quantifier les interactions entre leurs composantes

Chapitre 4

Expérience d'ICM asynchrone

Dans cette partie, nous présenterons l'expérience d'ICM qui a été développée dans cette thèse. En particulier, le protocole expérimental ainsi que l'architecture matérielle et logicielle du dispositif seront détaillées.

4.1 Objectifs et contraintes

4.1.1 Généralités

Les expériences d'ICM constituent une partie non négligeable du temps qui a été consacré à cette thèse, d'une part pour l'enregistrement des sujets et d'autre part pour la réalisation des protocoles de stimulation et des algorithmes de traitement des données qui sont implémentés en temps réel.

Des données d'enregistrements d'ICM sont cependant disponibles sur internet, en particulier les jeux de données des compétitions BCI2000 et BCI2003¹ qui ont permis de comparer différentes méthodes de classification des signaux EEG. Dans le cadre de notre travail, il a été choisi d'acquérir notre propre jeu de données expérimentales pour les raisons suivantes :

- Il existe à notre connaissance un seul jeu de données publiques concernant les ICM asynchrones et qui ne fait intervenir que des tâches d'imagerie motrice. De plus, celui-ci n'a pour l'instant fait l'objet d'aucun résultat, car trop peu d'équipes ont tenté de relever la compétition sur ce jeu de données.
- Nous nous proposons d'étudier l'apport de techniques de reconstruction de sources à partir du problème inverse en EEG. Pour bénéficier au mieux de ces techniques, il faut connaître l'anatomie du sujet enregistré (maillage surfacique ou volumique de la tête et du cortex) qui n'est pas disponible dans les bases de données actuelles.
- La quantité de données disponibles est en général faible : il existe un ensemble d'apprentissage et un ensemble test d'une centaine d'essais maximum. Nous pensons que la validation de méthodes de classification doit être faite sur plusieurs sujets et plusieurs sessions d'enregistrements, voire plusieurs jours d'utilisation.

1. http://ida.first.fraunhofer.de/projects/bci/competition_iii/

Cela permet d'éviter autant que possible le biais qui découle de l'essai de différentes méthodes sur le même jeu de données². De plus, l'influence de la méthode de validation et la stabilité du classifieur à travers le temps peuvent être étudiées (à l'intérieur d'une même session, entre sessions différentes ou entre différents jours d'enregistrements).

- La réalisation expérimentale d'une ICM permet d'appréhender tous les facteurs qui entrent en jeu dans la performance du dispositif. Il est ainsi possible d'étudier l'influence des expériences de feedback sur le système.

Pour résumer, nos expériences ont pour but de réaliser une ICM asynchrone qui autorise en temps réel une reconstruction de l'activité corticale des sujets et un retour (feedback) à l'utilisateur de son activité cérébrale.

4.1.2 Reconstruire les sources corticales

Pour reconstruire l'activité corticale aussi fidèlement que possible, il est nécessaire d'obtenir des informations individuelles par une technique d'imagerie anatomique sur chaque sujet. Cela contraint donc à choisir des sujets pour lesquels ces images sont disponibles. De plus, la position tridimensionnelle des électrodes sur la tête du sujet doit être relevée pendant l'expérience EEG. À ces contraintes expérimentales s'ajoutent des contraintes algorithmiques qui imposent d'utiliser des techniques de reconstruction de sources suffisamment rapides pour être utilisées en temps réel (voir chapitre 3).

4.1.3 Un système temps réel

En plus des méthodes de reconstruction de sources, le temps réel contraint l'ensemble de la chaîne de traitement des données : les méthodes de classification et de quantification (analyses fréquentielles et calcul des interactions) doivent avoir une complexité limitée. De plus, le système d'acquisition doit être choisi pour permettre le transfert en temps réel des données enregistrées vers la chaîne de traitement.

4.1.4 Choix du paradigme

Nos expériences ont pour but de tester un système d'ICM peu contraignant pour l'utilisateur. Dans ce cadre, un paradigme asynchrone autorise le sujet à changer son état mental sans attendre la stimulation suivante. De plus, un apprentissage principalement du côté machine semble s'imposer pour limiter l'apprentissage côté utilisateur par feedback, qui est en général un processus long et fastidieux. Le système doit donc être capable de différencier des états mentaux du sujet en minimisant l'adaptation de celui-ci. Pour ceci, nous avons choisi de tester un panel de six états mentaux différents.

2. L'essai successif de différentes méthodes sur un petit ensemble d'apprentissage peut conduire à trouver de bons résultats qui ne sont pas reproductibles sur un jeu de données plus important. Ce phénomène peut être interprété comme un cas particulier de non-correction pour comparaisons multiples.

Chacun d'entre eux est adapté au paradigme asynchrone en ce sens que le sujet peut déclencher ces états sans stimulation externe. Comme il a été développé dans la section 1.3, les états d'imagerie mentale sont particulièrement adaptés à cette situation. De plus, l'imagerie peut être sollicitée dans des modalités diverses, dont la plus utilisée en ICM est l'imagerie motrice. Il peut être utile d'envisager d'autres modalités (visuelles ou auditives) pour obtenir des signatures EEG plus contrastées à cause de la diversité des régions cérébrales impliquées (voir 1.3). Il est aussi intéressant d'utiliser d'autres tâches cognitives, comme le calcul mental, facilement réalisables par le sujet et pouvant solliciter l'activation d'aires cérébrales associées à l'imagerie (visuelle par exemple) et plus généralement à la mémoire de travail. L'ensemble de ces considérations nous a amené à élaborer le protocole expérimental décrit ci-dessous.

4.1.5 Ma contribution

Le système qui va être présenté dans les parties suivantes découle de l'ensemble des considérations ci-dessus. À mon arrivée au laboratoire le matériel nécessaire pour réaliser les enregistrements EEG était déjà disponible. De plus, la solution pour transmettre en temps réel les données du système d'acquisition vers la chaîne de traitement, élaborée sous Matlab, avait été trouvée par Jacques Martinerie. Hormis ces deux points, l'ensemble des aspects décrits ci-dessous ont été réalisés par moi. Ceci inclut la mise au point du protocole expérimental, la réalisation des sessions d'enregistrement (avec l'aide de Jacques Martinerie et Florence Bouchet), l'implémentation de la chaîne de traitement temps réel ainsi que les routines de stimulation.

4.2 Protocole expérimental

Les sujets étaient assis dans un fauteuil à environ 1 m d'un écran de stimulation, dans une salle d'enregistrement isolée des expérimentateurs. Durant les expériences, ils avaient pour consigne d'effectuer des tâches mentales en suivant les instructions affichées sur l'écran tout en évitant les clignements et les contractions musculaires en dehors des périodes de repos.

4.2.1 Tâches

La littérature évoquée à la section 1.3 fait état de l'activation des régions sensorielles du cortex lors des tâches d'imagerie. Afin de permettre la différenciation maximale de ces tâches à partir des signaux EEG, nous avons utilisé des tâches susceptibles d'activer des régions différentes en utilisant d'une part différents types d'imagerie motrice, mais aussi en les contrastant avec des tâches d'imagerie non motrices. La différenciation de ces tâches se fait principalement sur un axe antéro-postérieur représenté figure 4.1, qui contraste avec l'approche plus classique en ICM reposant sur la différenciation des activités des hémisphères gauche et droit. Ainsi, six tâches mentales ont été réalisées par les sujets :

- trois tâches d'imagerie motrice :
 - imagination du mouvement de l'*index* droit
 - imagination de la *préhension* d'un objet avec le bras droit
 - imagination du mouvement de la *langue*
- trois tâches non motrices :
 - imagerie auditive d'un morceau de *musique* instrumentale
 - imagerie *visuo-spatiale* de la navigation dans un environnement familier
 - *calcul* mental : soustractions répétées du nombre sept à partir d'un nombre initial choisi arbitrairement.

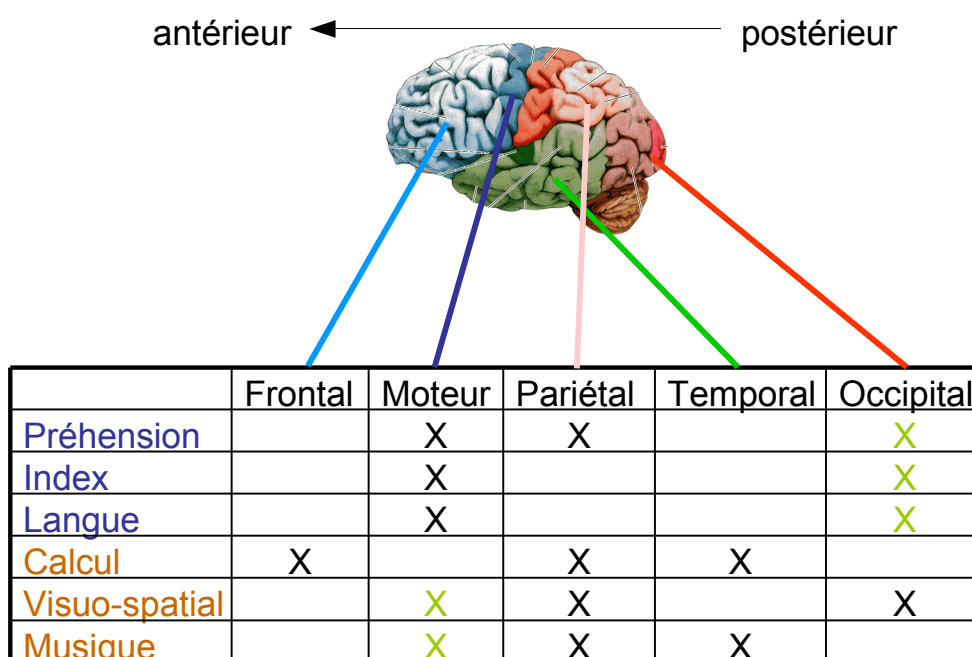


FIGURE 4.1 – Aires susceptibles d'être activées par chaque tâche d'imagerie, ordonnées suivant l'axe antéro-postérieur. Les croix noires reportent les régions associées à ce type d'imagerie, les croix vertes reportent les régions qui peuvent être activées si la stratégie d'imagerie du sujet fait appel à des modalités secondaires (par exemple l'imagerie motrice peut faire appel à une modalité visuelle : voir section 1.3).

4.2.2 Entraînement

Les sessions d'entraînement permettent de former une base d'apprentissage pour la classification des différents états mentaux. Ce sont sur ces données que sont testées et comparées différentes méthodes de traitement des données ainsi que les différents couples de tâches à séparer.

Chaque session d'entraînement dure environ 7 minutes, où se succèdent des blocs de 20s durant lesquels le sujet doit effectuer la tâche affichée sur l'écran en continu. Ces blocs sont entrecoupés de périodes de repos de durée comprise entre 2 et 3 s durant lesquelles est affiché le mot PAUSE sur l'écran. Ces pauses sont l'occasion pour le sujet de bouger, se détendre et cligner des yeux. Le protocole de stimulation pour l'entraînement est résumé figure 4.2.

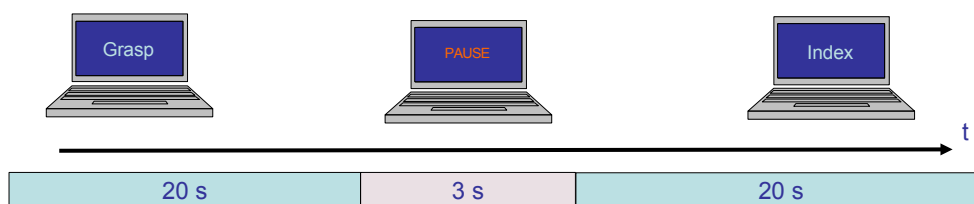


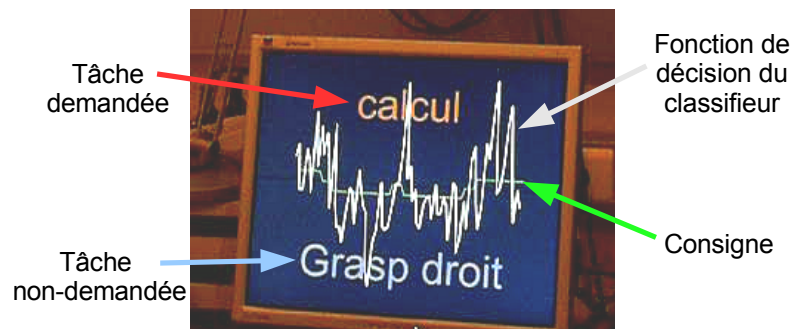
FIGURE 4.2 – Schéma du protocole d'entraînement. Ici, la tâche d'imagerie motrice de préhension (GRASP) est effectuée pendant 20s. Après une pause de 3s durant laquelle le sujet peut se détendre, il commence un nouveau bloc de 20s durant lequel il effectue une imagerie motrice de l'index.

4.2.3 Feedback

Des sessions de feedback sont réalisées sur certains couples de tâches qui sont évalués comme les plus discriminants lors de l'entraînement. L'expérimentateur choisit un couple de tâches ainsi que les paramètres d'un classifieur qui a préalablement été entraîné à séparer ces deux tâches. Une séance de feedback dure environ 6 minutes durant lesquelles le sujet cherche à contrôler la hauteur d'une courbe représentée figure 4.3. Cette courbe (en blanc) représente l'évolution de la fonction discriminante du classifieur, qui est affichée en temps réel à partir des variables de quantification (voir chapitre 2) calculées sur des fenêtres de deux secondes. Les fenêtres de quantification étant choisies avec un recouvrement d'une seconde, l'affichage de la courbe est aussi mis à jour toutes les secondes. Les deux tâches à différencier sont affichées en haut et en bas de l'écran durant toute la durée de la session. La consigne est d'effectuer la tâche qui apparaît en surbrillance (en orange) sur l'écran (si la tâche est affichée en haut, une bonne classification doit correspondre à une position haute de la courbe). De manière similaire aux sessions d'apprentissage, la tâche à effectuer reste constante pendant la durée des blocs de 20s. Ceux-ci sont entrecoupés de pauses d'environ 3s durant lesquelles le sujet peut cligner des yeux et contracter ses muscles. Des photographies des différentes étapes d'un enregistrement de feedback sont présentées figure 4.3(a).



(a)



(b)

FIGURE 4.3 – Protocole de feedback lors de l'enregistrement d'un sujet. En (a) les photographies successives montrent une vue de l'expérimentateur avec une image vidéo infrarouge du sujet et la duplication sur un écran de contrôle du retour visuel donné au sujet. En (b), description du retour visuel : le décours temporel de la fonction discriminante du classifieur (voir chapitre 2) est renvoyé au sujet en temps réel (courbe blanche), la tâche illuminée en orange indique la consigne et la direction dans laquelle cette courbe doit évoluer (vers le haut ou le bas), l'évolution de la consigne (courbe verte en escalier) permet de quantifier la performance du système.

4.2.4 Données enregistrées

À la suite de cette expérience, on dispose pour cinq sujets de trois jours d'enregistrement durant lesquels sont disponibles 4 à 8 sessions d'entraînement (par jour). Pour les expériences de feedback, on dispose de moins d'enregistrements (en moyenne une ou deux sessions sur chacun des deux derniers jours) et qui concernent des couples de tâches choisis en fonction des sujets. Comme il est schématisé dans la figure 4.2.4, certaines sessions de feedback sont réalisées à partir d'un classifieur ayant appris sur les sessions d'entraînement du jour précédent, tandis que d'autres sessions utilisent un classifieur ayant appris sur une session de feedback précédente du même jour. Il a été observé empiriquement une différence entre ces deux types de session (voir section 4.4.2).

La variabilité des tâches utilisées et des ensembles d'apprentissage exploités pour les sessions de feedback rend leur analyse systématique plus difficile. Dans les chapitres suivants, les analyses sont donc réalisées sur les sessions d'entraînement pour les 15 couples de tâches possibles pour une classification binaire.

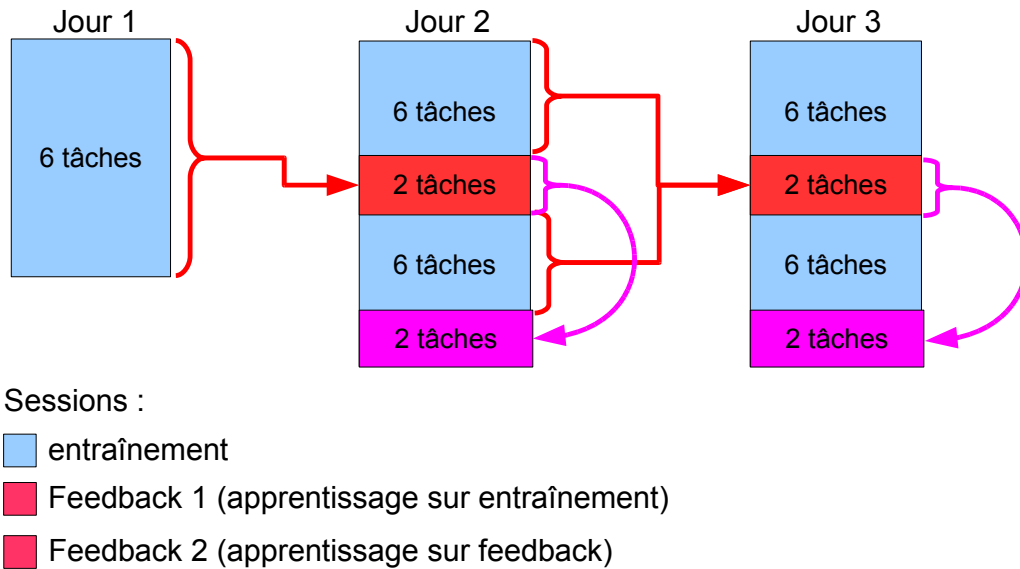


FIGURE 4.4 – Schéma des sessions réalisées par les sujets pendant les trois jours d'enregistrement. Les sessions de feedback sont réalisées avec un classifieur entraîné soit sur les sessions d'entraînement du jour précédent, soit sur une session de feedback du jour même.

4.3 Enregistrements

Le protocole ainsi défini nécessite la mise en place de nombreux éléments lors de l'expérimentation. La structure matérielle de l'ensemble du système expérimental est

représentée figure 4.5 et ses différents constituants sont décrits dans les parties suivantes.

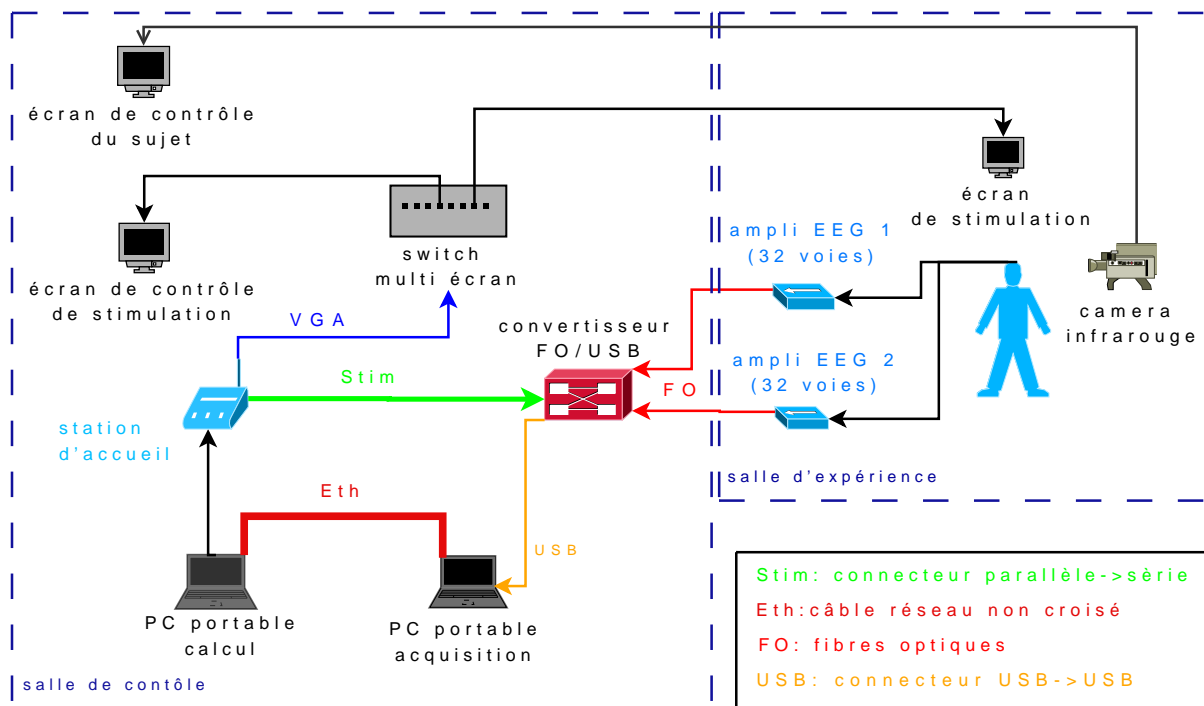


FIGURE 4.5 – Schéma de câblage de l'expérience

4.3.1 Système d'acquisition

Les enregistrements sont réalisés à partir d'un casque EEG Braincap de 62 électrodes pour les sujets 1 à 3, et 61 électrodes pour les sujets 4 et 5. Les signaux EEG ainsi que ceux issus des électrodes périphériques (une ou deux³ électrodes d'électrooculogramme : EOG, et une électrode d'électromyogramme : EMG) sont envoyés vers deux amplificateurs 32 voies BrainAmp. Ils sont alors amplifiés, puis filtrés entre .1 et 250Hz, puis échantillonnés à 500Hz et numérisés avec une résolution de $.5 \mu V$. Les signaux numériques obtenus sont transmis par fibre optique vers l'extérieur de la salle d'expérience vers un *convertisseur Fibre optique -> USB*, qui reçoit aussi les marqueurs de stimulation issus du PC portable de calcul (voir section suivante). L'ensemble de ces signaux est transmis par cordon USB au *PC portable d'acquisition* qui enregistre les données via le logiciel d'acquisition *Vision Recorder*⁴.

3. suivant le casque utilisé : une électrode pour les 3 premiers sujets, deux électrodes pour les deux autres

4. http://www.brainproducts.com/downloads/down_recorder_attention.html

4.3.2 Protocole de stimulation

Lors des expériences de feedback, la stimulation doit incorporer des informations issues du traitement temps réel des signaux acquis. Il a donc été choisi de gérer l'envoi des stimulations via le programme Matlab qui réalise la quantification et la classification des signaux EEG et fonctionnant sur le *PC portable de calcul*. À chaque changement de l'écran de stimulation, l'affichage est géré via les fonctions de la boîte à outils *Cogent Graphics*⁵ qui permet d'avoir un contrôle temps réel sur le signal vidéo. Les commandes d'affichage sont envoyées vers un écran externe grâce à la sortie VGA du PC, alors que l'interface avec les programmes Matlab reste accessible à l'expérimentateur sur l'écran LCD en configurant un mode double affichage sur la carte vidéo. L'écran de stimulation est aussi reproduit sur un écran de contrôle qui permet à l'expérimentateur de surveiller le bon fonctionnement de la séance.

Pour synchroniser la stimulation avec le système d'acquisition, à chaque changement de l'écran de stimulation un signal binaire TTL est écrit sur le port parallèle de la station d'accueil du PC de calcul⁶. Ce port parallèle est connecté au convertisseur *Fibre optique/USB*. L'accès au port parallèle à partir de Matlab est géré par les fonctions de la librairie *NTport*⁷.

4.3.3 Acquisition des données anatomiques

L'estimation de l'activité corticale des sujets nécessite l'obtention de données supplémentaires associées à l'anatomie du sujet :

- la position tridimensionnelle des électrodes sur le cuir chevelu du sujet, obtenue grâce au système de localisation hyperfréquence *Polhemus*. En même temps que les électrodes, les positions de trois repères anatomiques sont relevées sur le nez, l'oreille droite et l'oreille gauche.
- l'IRM anatomique pondérée en T_1 de chaque sujet (voir encadré suivant) a été segmentée à l'aide du logiciel *Brainvisa* afin d'extraire un maillage surfacique du cortex⁸ et de la tête du sujet. Les trois repères anatomiques acquis sous *Polhemus* sont alors localisés sur la surface de la tête dans le logiciel *Brainstorm* et permettent de positionner les différents maillages dans le même référentiel que les électrodes (le système de coordonnées CTF associé au sujet⁹).

5. <http://www.vislab.ucl.ac.uk/CogentGraphics/>

6. la station d'accueil permet de disposer d'un port parallèle, qui est absent sur les PC portables récents

7. <http://www.zealsoftstudio.com/ntport/download.html>

8. on utilise l'interface entre la matière grise et la matière blanche

9. voir http://neuroimage.usc.edu/brainstorm/GUI_CoordinateSystem.htm

Quelques notions sur l'Imagerie par Résonance Magnétique :

Cette méthode d'imagerie est basée sur un principe physique : lorsque ils sont placés dans un champ magnétique constant \mathbf{B}_0 , les protons de l'hydrogène de l'eau acquièrent un moment magnétique \mathbf{M}_0 orientés en moyenne le long de l'axe de \mathbf{B}_0 . L'émission d'une impulsion radio à la fréquence de résonance dite *fréquence de Larmor*, perpendiculairement à l'axe de \mathbf{B}_0 , permet de faire basculer le moment magnétique perpendiculairement à l'axe de \mathbf{B}_0 : c'est le phénomène de *résonance magnétique nucléaire*.

À la fin de cette impulsion, le moment magnétique retourne à sa valeur initiale par un mouvement de précession, émettant une onde électromagnétique à la fréquence de Larmor qui permettra la mesure. Ce phénomène de relaxation est régi par plusieurs temps de relaxation :

- T_1 pour le retour de la composante longitudinale (parallèle à \mathbf{B}_0) du moment magnétique à sa valeur initiale \mathbf{M}_0 .
- T_2 est le temps de relaxation théorique pour le retour à 0 de la composante transverse.
- T_2^* est le temps de relaxation observé en pratique pour la composante transverse. À cause des hétérogénéités du champ magnétique au sein du tissu biologique, il est inférieur à T_2 .

Des techniques de mesure des temps de relaxation en tout point de la tête d'un sujet permettent de contraster différents phénomènes au sein des tissus biologiques. Chaque temps de relaxation est adapté à un problème particulier :

- l'IRM dite "pondérée en T_1 " permet de bien contraster la substance grise et la substance blanche du système nerveux central, et donc d'en visualiser les organes.
- la pondération en T_2 met en évidence les tissus pathologiques (oedèmes, les tumeurs, etc...), qui ont un temps de relaxation supérieur aux tissus sains.
- la pondération en T_2^* fait apparaître les fluctuations de la concentration d'oxygène dans le sang. Le signal BOLD (Blood Oxygen Level Dependent) ainsi enregistré est lié à l'activité neuronale. Il est à la base de l'IRM fonctionnelle.

4.3.4 Architecture du traitement temps réel

Le protocole de feedback nécessite un traitement en temps réel des signaux issus du système d'acquisition. Ce traitement repose sur la transmission des signaux acquis depuis le PC d'acquisition vers le PC de calcul par protocole TCP-IP. Le principe de la transmission, résumé figure 4.6, est le suivant :

- toutes les 40 ms, les informations acquises sont envoyées sur le port 51234 du PC d'acquisition par le logiciel Brain Vision Recorder¹⁰.
- un driver Matlab lancé sur le PC de calcul vient régulièrement lire ce port pour récupérer ces données¹¹, les concatène pour constituer un bloc qui correspond à la moitié d'une fenêtre d'analyse (de durée d'une seconde¹²). Une fois complet, ce bloc est transmis, via le port 5432 du PC de calcul, à la routine Matlab qui gère la stimulation et la classification .

10. http://www.brainproducts.com/products/recorder/index_recorder.html

11. La gestion des communication TCP/IP est réalisée sous Matlab en utilisant la boîte à outils *TCP/UDP/IP Toolbox 2.0.5* disponible à l'adresse <http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=345>

12. pour permettre un rafraîchissement d'une seconde lors du feedback, les fenêtres d'analyse, qui durent 2s, sont concaténées à partir de deux demi-fenêtres successives d'1s dans la routine Matlab de stimulation

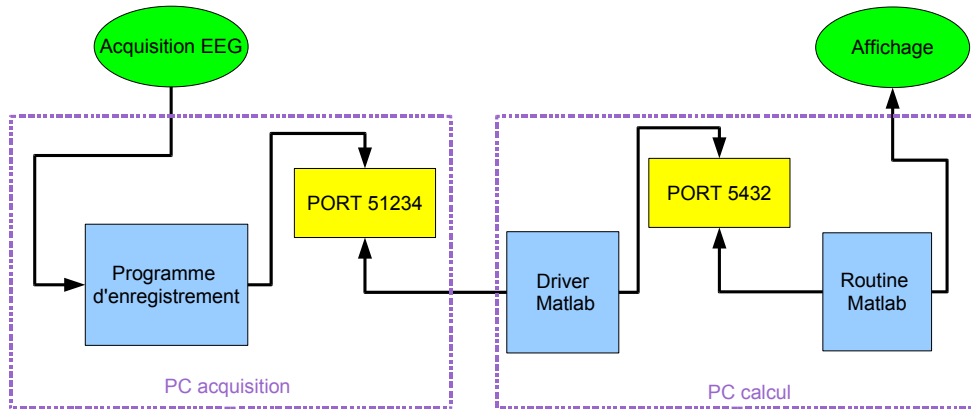


FIGURE 4.6 – Schéma du traitement temps réel des données acquises.

4.4 Observations

Nous répertorions ici les remarques intéressantes sur l'expérience, qui ont été observées par l'ensemble des sujets.

4.4.1 Sessions d'apprentissage

L'exécution des tâches n'a pas posé de difficultés de compréhension à l'ensemble des sujets. Hormis pour la tâche de calcul mental, les consignes données aux sujets étaient peu contraignantes pour s'adapter aux capacités d'imagerie de chaque sujet. Par exemple, pour la tâche d'imagerie de l'index droit certains sujets imaginaient simplement lever leur doigt répétitivement tandis que d'autres l'utilisaient pour dessiner une forme imaginaire dans l'espace. Ces faibles contraintes pouvaient être un inconvénient si l'objectif était d'obtenir des résultats reproductibles à travers les sujets, cependant l'ensemble de la littérature reporte des résultats qui dépendent fortement du sujet.

Par contre, il était vérifié oralement tout au long des jours d'enregistrements que le sujet maintenait les mêmes stratégies pour chaque état mental. Ainsi les premières sessions de chaque jour ont été retirées des analyses, car elles correspondaient à des périodes où le sujet devait se rappeler la stratégie mentale qu'il avait employée pour chaque tâche. De manière générale, les sujets ont trouvé plus laborieuses les tâches d'imagerie motrice à cause de leur caractère répétitif et ont eu parfois des difficultés à imaginer certaines d'entre elles comme le mouvement de l'index. De plus, la succession des tâches, toutes les 20s, a été considérée comme confortable, contrairement à une expérience précédente où des périodes d'une minute étaient utilisées et où certains sujets se distraient ou s'endormaient.

4.4.2 Sessions de feedback

Après plusieurs sessions d'apprentissage, un classifieur optimal pouvait être calculé pour différencier chaque couple de tâches mentales. En fonction de ces résultats, le sujet effectuait une session de feedback sur le couple de classes le plus discriminant. À la plupart des premières sessions de feedback, les sujets étaient fortement perturbés par le fait que le comportement de la courbe qu'ils voyaient évoluer à l'écran n'était pas ce qu'ils attendaient. Lorsqu'ils envisageaient de changer de stratégie pour tenter d'améliorer les résultats, nous les autorisions à le faire tant qu'ils respectaient la consigne associée à la tâche. Les difficultés de contrôle de la courbe diminuaient fortement lorsque l'apprentissage du classifieur était réalisé sur des sessions précédentes de feedback. Il semble donc que les sessions de feedback, bien que basées sur les mêmes tâches que les sessions d'apprentissage, conduisent à des signatures EEG différentes, ce qui pourrait être dû soit à l'attention portée par le sujet sur l'écran qui devient plus soutenue, soit à l'objectif du feedback qui incite le sujet, non pas à penser seulement à la tâche demandée, mais aussi à contrôler la courbe qui s'affiche à l'écran. Ce phénomène pourrait être étudié en détail dans de futurs travaux.

Chapitre 5

Analyse au niveau des capteurs

5.1 Méthodes

Cette partie constitue une étude de référence pour les chapitres suivants. Il s'agit ici de dresser les premiers résultats concernant les enregistrements d'ICM en utilisant le niveau d'analyse le plus simple à réaliser : celui des capteurs EEG. Il s'agit donc d'exploiter les mesures de synchronie locale sur chaque capteur et à distance pour chaque couple de capteurs pour quantifier l'activité cérébrale et estimer l'état mental.

5.1.1 Prétraitements

Fenêtres glissantes

La base du protocole asynchrone est l'analyse des enregistrements sur des fenêtres glissantes. La durée des fenêtres utilisées influence directement la performance du système en terme de taux de classification et en terme de temps de réponse. Ainsi, plus la largeur des fenêtres d'analyse augmente, plus il est a priori possible d'identifier l'état mental associé à un signal car plus d'informations sont disponibles. Cependant une plus grande durée de fenêtre détériore aussi certaines caractéristiques du système :

- l'état mental doit être maintenu durant toute la durée de la fenêtre, ce qui limite le débit d'information du système
- lors des séances de feedback, l'information discriminante extraite d'une fenêtre est restituée au sujet avec un délai qui est la somme de : 1) la durée nécessaire pour enregistrer les signaux de la fenêtre d'analyse, 2) la durée de traitement de ces signaux (quantification et classification). Un délai trop important rend le feedback inutile car le sujet ne peut plus percevoir l'influence qu'il a sur le dispositif.

Afin d'exploiter une durée de fenêtre maximale tout en conservant un délai de feedback acceptable pour l'utilisateur, il a été choisi de faire les analyses sur des fenêtres de 2s.

Élimination d'artefacts

Un algorithme de correction d'artefacts est appliqué aux données afin d'éliminer la contribution des signaux EOG dans les enregistrements. Soit \mathbf{M}_b la matrice des mesures EEG "brutes" issues d'une fenêtre d'analyse, on réalise une décomposition en valeur singulière de la matrice étendue $\begin{pmatrix} \mathbf{M}_b \\ \mathbf{EOG} \end{pmatrix}$ où l'on rajoute une ligne contenant le signal EOG¹. La décomposition obtenue prend la forme \mathbf{UDV}^T où \mathbf{U} et \mathbf{V} regroupent les vecteur colonnes orthonormés spatiaux et temporels respectivement et \mathbf{D} est la matrice diagonale des valeurs singulières. Le coefficient de corrélation de chaque vecteur temporel de \mathbf{V} avec l'EOG est évalué par $\mathbf{V}^T \mathbf{EOG}^T / \|\mathbf{EOG}\|$. Les valeurs singulières correspondant à des vecteurs temporels \mathbf{V} qui présentent un coefficient de corrélation supérieur à .9 sont alors annulés dans la matrice \mathbf{D} pour former la nouvelle matrice $\tilde{\mathbf{D}}$ (les autres valeurs singulières sont conservées). Celle-ci permet de reconstituer le signal corrigé qui prend la forme $\mathbf{U}\tilde{\mathbf{D}}\mathbf{V}^T$ (Wallstrom et al., 2004).

Référence moyenne commune

Il a été montré empiriquement que l'application d'une référence moyenne commune permet de rehausser les oscillations du signal EEG (McFarland et al., 1997). L'opération de référence commune consiste simplement à retirer à chaque instant sur chaque voie la moyenne des signaux à travers l'ensemble des électrodes de scalp. Elle est réalisée après l'élimination d'artefacts EOG par l'application sur le vecteur de mesures EEG de la matrice de filtrage spatial :

$$\boldsymbol{\theta}_{CAR} = \begin{pmatrix} 1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & 1 \\ 0 & \dots & 0 \end{pmatrix} - \frac{1}{N_{elec} + 1} \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \\ 1 & \dots & 1 \end{pmatrix}$$

Dans cette matrice, il a été rajouté une ligne supplémentaire qui correspond à l'activité de l'électrode de référence. En effet, on peut considérer que la référence est une électrode comme les autres qui enregistre un signal nul. Il est donc possible d'inclure cette électrode dans le calcul de la référence moyenne commune, permettant ainsi de combler un vide dans la répartition des capteurs sur le scalp en associant un signal non nul à l'électrode de référence.

5.1.2 Mesures utilisées

Mesures de synchronie locale

Pour chaque électrode, la puissance spectrale dans une bande de fréquence est estimée par un périodogramme de Welch, qui consiste à moyenner les périodogrammes

1. Pour simplifier les signaux \mathbf{M}_b et \mathbf{EOG} seront supposés centrés

calculés à partir de plusieurs fenêtrages temporels du signal. Son expression pour l'électrode n est :

$$P_W(n, f_k) = \frac{1}{I} \sum_{i \in 1}^I P_i(n, f_k)$$

où $P_i(n, f_k) = \frac{\alpha}{L} (\tilde{\mathbf{M}}_i \tilde{\mathbf{M}}_i^H)_{nn}$ ² est le périodogramme calculé à partir de la transformée de Fourier à temps discret $\tilde{\mathbf{M}}_i$ du signal $\tilde{\mathbf{m}}$ résultant de la multiplication du signal mesuré m par la i -ème fonction de fenêtrage w_i (L désigne le nombre de points temporel de la fenêtre et α est un coefficient de normalisation qui dépend de la forme de la fenêtre).

L'intérêt de cette mesure est de réduire la variance de l'estimateur de la densité spectrale en chaque point fréquentiel, en moyennant les estimations issues de différents fenêtrages. Dans notre cas, la répartition des fonctions de fenêtrage est décrite figure 5.1.2. Nous avons choisi des fenêtres de Hamming (qui minimisent les oscillations du spectre estimé), de durée 300ms réparties sur la totalité de la fenêtre d'analyse avec un recouvrement de 150ms entre fenêtres voisines. Cette durée (300 ms) permet d'assurer une résolution fréquentielle raisonnable (de l'ordre de 3Hz) vis à vis des spectres couramment observés en EEG tout en assurant une diminution de la variance de l'estimateur par moyennage d'une dizaine de fenêtres. Les fréquences f_k sont choisies de manière à échantillonner uniformément la bande $[0, 250Hz]$ avec 2048 points. L'activité dans une bande de fréquence $[f_l, f_h]$ est alors quantifiée par la somme des valeurs du périodogramme dans la bande concernée : $\sum_{f_k \in [f_l, f_h]} P_W(f_k)$

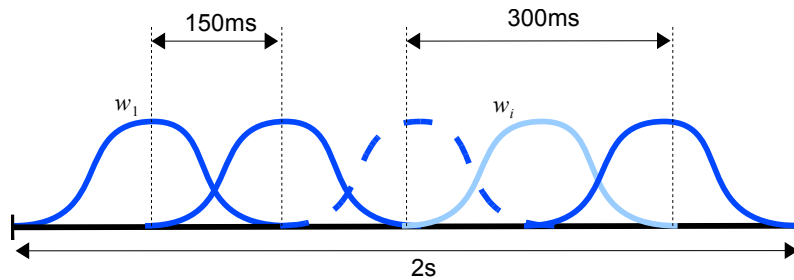


FIGURE 5.1 – Répartition des fonctions de fenêtrage du périodogramme de Welch sur une fenêtre d'analyse.

Mesures d'interaction à distance

Parmi les méthodes possibles pour quantifier les interactions dans chaque bande de fréquences, nous en avons retenu et comparé deux :

2. où H est l'opérateur de conjugaison Hermitienne

- Cohérence : la cohérence à été calculée de manière similaire au périodogramme de Welch, en calculant les coefficients de la matrice de densité spectrale (Brillinger, 2001) estimée $\Gamma_{\mathcal{W}}(f_k)$ à partir plusieurs fonctions de fenêtrage de Hamming

$$\Gamma_{\mathcal{W}}(f_k) = \frac{1}{I} \sum_{i \in 1}^I \Gamma_i(f_k) \quad \text{où} \quad \Gamma_i(f_k) = \frac{\alpha}{L} \tilde{M}_i \tilde{M}_i^H$$

La cohérence entre l'électrode n et l'électrode m dans la bande $[f_l, f_h]$ est alors donnée par

$$C_{nm} = \frac{|\sum_{f_k \in [f_l, f_h]} (\Gamma_{\mathcal{W}}(f_k))_{nm}|}{\sqrt{|\sum_{f_k \in [f_l, f_h]} (\Gamma_{\mathcal{W}}(f_k))_{nn}| |\sum_{f_k \in [f_l, f_h]} (\Gamma_{\mathcal{W}}(f_k))_{mm}|}}$$

- Synchronie de phase : Comme précisé dans le chapitre 1 la synchronie de phase nécessite l'utilisation d'un filtrage passe bande (Chavez et al., 2006) puis d'un filtrage de Hilbert qui permet de calculer le signal analytique associé au signal de chaque électrode. En utilisant la transformée de Fourier à temps discret, l'ensemble de ces opérations peut être effectué en une seule opération de filtrage qui élimine les fréquences négatives du signal et sélectionne les fréquence positives dans une bande de fréquence prédéfinie. Ce filtrage est réalisé dans le domaine fréquentiel par multiplication de la transformée de Fourier du signal par une fonction de transfert trapézoïdale³ représentée dans la figure 5.2. En sortie de ce filtre, un signal analytique est obtenu qui permet de calculer la phase instantanée de chaque signal et la mesure de synchronie de phase dans la bande $[f_l, f_h]$:

$$S_{nm} = \frac{1}{N} \left| \sum_{k=1}^n e^{i(\phi_n(t_k) - \phi_m(t_k))} \right|$$

5.1.3 Normalisation

Pour assurer un fonctionnement optimal de l'algorithme de classification, toutes les variables de quantification subissent une normalisation préalable à l'apprentissage du classifieur. Elle consiste à soustraire à chaque variable sa moyenne et à diviser le résultat par son écart type (ces deux grandeurs statistiques sont évalués sur l'ensemble d'apprentissage). Les données test sont normalisées à l'aide de ces mêmes valeurs avant la classification.

5.1.4 Classification

Pour l'ensemble des résultats de ce chapitre il a été choisi d'utiliser un SVM linéaire. Ce choix est motivé par deux arguments :

³. afin de ne pas biaiser l'estimation de la phase instantanée, la fonction de transfert doit aussi être réelle positive pour ne pas introduire de déphasage.

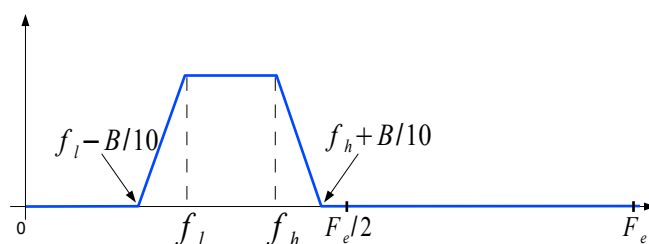


FIGURE 5.2 – Forme de la fonction de transfert du filtre permettant de calculer le signal analytique dans la bande $[f_l, f_h]$, on note $B = f_h - f_l$.

- compte tenu du nombre de variables déjà élevé, il n'est pas nécessaire d'utiliser un classifieur non-linéaire qui risquerait d'avoir pour effet d'accentuer le phénomène de surapprentissage.
- on dispose d'un faible nombre de segments EEG d'apprentissage par rapport au nombre de dimensions (au mieux ces quantités sont égales si l'on réalise une quantification par synchronie locale). Comme illustré dans l'annexe B sur des données MEG, une méthode paramétrique comme l'analyse discriminante linéaire risque donc d'avoir de mauvaises performances, dues à la mauvaise estimation des paramètres à partir d'un faible nombre d'échantillons.

Le paramètre de régularisation C doit en théorie être réglé par validation croisée. Cependant, le coût d'une telle validation nécessite un temps de calcul prohibitif. Compte tenu du nombre de variables utilisées et de leur normalisation, nous avons utilisé $C = 1$ qui est une valeur suffisamment grande pour contraindre le terme d'erreur de l'équation 2.3 à être nul⁴. La solution du SVM est alors celle du cas séparable où aucun point n'est autorisé à pénétrer dans la marge séparant les deux classes (Vapnik, 1998). Bien qu'une valeur de C moins élevée puisse améliorer les résultats, celle qui a été choisie permet d'assurer une bonne convergence de l'algorithme dans tous les cas (en effet, en utilisant un coefficient C trop faible, nous avons été confronté à des problèmes de convergence de l'algorithme de programmation quadratique). Pour l'optimisation rapide du SVM, nous utilisons la boîte à outil *QPC*⁵ de programmation quadratique et en particulier la méthode du point intérieur implémentée par la fonction *qpip*.

5.1.5 Sélection de variables

Le nombre de variables qui sont exploitées pour la classification est important. En particulier, le nombre de mesures d'interactions à distance issues des données fait intervenir le produit du carré du nombre d'électrodes et le nombre de bandes de fréquence.

4. Dans le problème dual équation 2.4, cela se traduit par des coefficients α_i qui sont tous strictement inférieurs à C .

5. <http://www.sigpromu.org/quadprog/index.html>

On obtient ainsi plusieurs milliers de mesures de synchronie pour une même fenêtre d'analyse. Il est intéressant de diminuer ce nombre important pour 3 raisons :

- permettre aux étapes de quantification et de classification d'être calculées en temps réel
- réduire l'occupation mémoire
- limiter le surapprentissage du SVM

Il s'agit donc de conserver parmi l'ensemble des variables celles qui favorisent la discrimination entre les deux tâches considérées. Pour cela, les variables sont rangées en fonction de leur pouvoir discriminant grâce au score de Fisher évalué sur l'ensemble d'apprentissage (voir section 2.3.1). Seul un nombre prédéfini de variables les plus discriminantes est alors utilisé pour la classification.

La reproductibilité des variables sélectionnées en fonction des sujets et des sessions d'enregistrements n'a pas été quantifiée. Cependant, on peut s'attendre à une instabilité de la sélection, compte tenu du nombre élevé d'électrodes considérées⁶, du faible nombre d'échantillons (qui entraîne une variance considérable de la statistique) et de la variabilité de l'anatomie des sujets.

La robustesse des variables sélectionnées est bien entendu un facteur qui influe directement sur la performance du système : si les variables sélectionnées lors d'une séance ne sont plus discriminantes lors de la séance suivante, une augmentation du taux d'erreur est à prévoir. Pour ne pas surestimer la performance du système, la sélection de variables est réalisée en excluant l'ensemble test.

5.1.6 Validation des performances

Les résultats de classification sont quantifiés par deux types de validations croisées :

- *validation croisée inter-sessions* : pour un sujet lors d'un jour d'enregistrement, une session est retirée de l'ensemble d'apprentissage journalier pour servir d'ensemble test et le classifieur est construit à partir de l'ensemble des sessions restantes. Ceci est répété pour chaque session et le taux de classification sur l'ensemble test est moyenné à travers les sessions.
- *validation croisée inter-jours* : pour un sujet, l'apprentissage est réalisé sur un jour et le test sur les deux autres jours d'enregistrement. Ceci est répété pour chaque jours d'enregistrement et le résultat est moyenné à travers la totalité des ensembles test.

Ces deux mesures permettent de quantifier la stabilité dans le temps de la quantification.

6. deux électrodes voisines peuvent être à proximité d'une même région, avoir un pouvoir discriminant similaire et être l'une ou l'autre sélectionnées

TABLE 5.1 – Bandes de fréquences choisies pour chaque sujet en Hz

sujet	bande 1	bande 2	bande 3	bande 4	bande 5
1	1-6	6-11	11-18	18-27	27-39
2	1-6	6-15	15-24	24-32	32-39
3	3-9	9-16	16-22	22-27	27-38
4	1-6	6-10	10-15	15-26	26-38
5	4-10	10-16	16-21	21-27	27-39

5.2 Résultats

5.2.1 Choix des bandes de fréquences

Des études précédentes ont été réalisées dans des bandes de fréquences qui n'étaient pas spécifiques aux différents sujets. Après avoir constaté une grande variabilité dans les spectres de ces sujets, nous avons décidé de réaliser une segmentation spécifique pour chacun d'entre eux. Les segmentations sont basées sur les valeurs du score de Fisher calculées sur les périodogrammes de chaque capteur EEG. Plus précisément, soit

$$F_{i,j}(f) = \frac{1}{N} \sum_n \frac{|\mu_i(n, f) - \mu_j(n, f)|}{\sqrt{(\sigma_i^2(n, f) + \sigma_j^2(n, f))}} \quad (5.1)$$

le racine carrée du score de Fisher (qui correspond à la statistique du T de Student) entre les classes i et j , moyennée sur l'ensemble des N électrodes à la fréquence f . Afin d'extraire les bandes de fréquences qui permettent de réaliser au mieux la discrimination de chaque classe par rapport aux autres, il est possible de tracer l'évolution fréquentielle des scores moyens de chaque classe par rapport aux autres. Ceux-ci sont tracés figure 5.3 pour le sujet 1, chaque axe rassemblant les courbes de score d'une classe par rapport à toutes les autres. À partir des courbes relatives à chaque sujet, 5 bandes de fréquences ont été segmentées manuellement en extrayant les maxima locaux les plus marqués de ces courbes et en définissant des bandes jointives centrées sur ces maxima. Les bandes de fréquences choisies par cette méthode sont répertoriées table 5.1 et seront utilisées dans les quantifications réalisées dans ce chapitre ainsi qu'au chapitre 7.

5.2.2 Sélection de variables et quantification

Pour les différents types de quantification dans les bandes de fréquence données table 5.1, le taux de classification en validation croisée inter-sessions a été évalué pour l'ensemble des discriminations possibles chez les cinq sujets. Les taux moyens de bien classés en validation croisée inter session, obtenus chez les cinq sujets, sont représentés figures 5.4(a) à 5.4(d) en fonction du nombre de variables sélectionnées par score de Fisher. Les différentes quantifications qui ont été testées sont : 1) l'utilisation des synchronies locales mesurées par la puissance spectrale (figure 5.4(a)), 2) l'utilisation

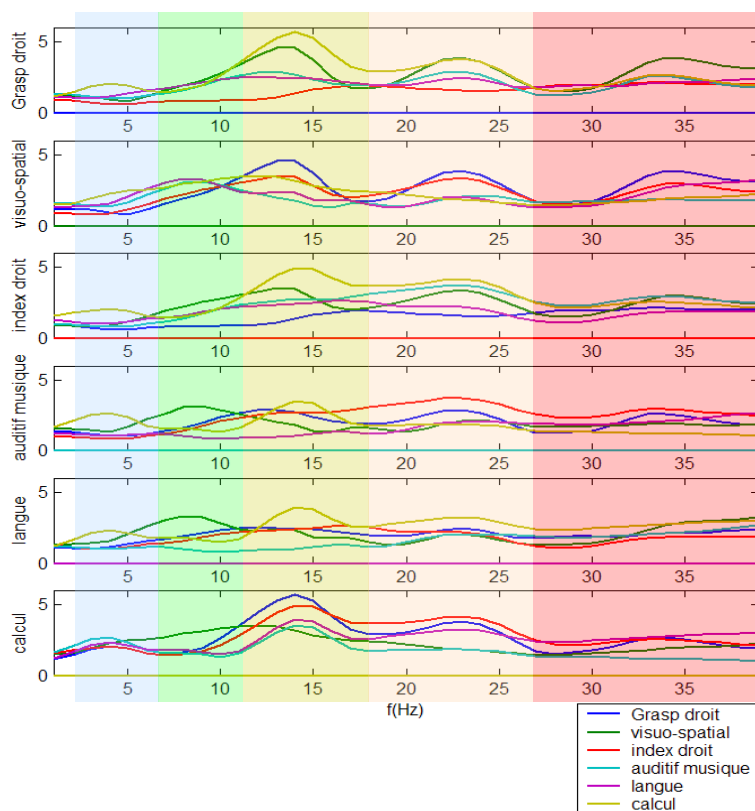
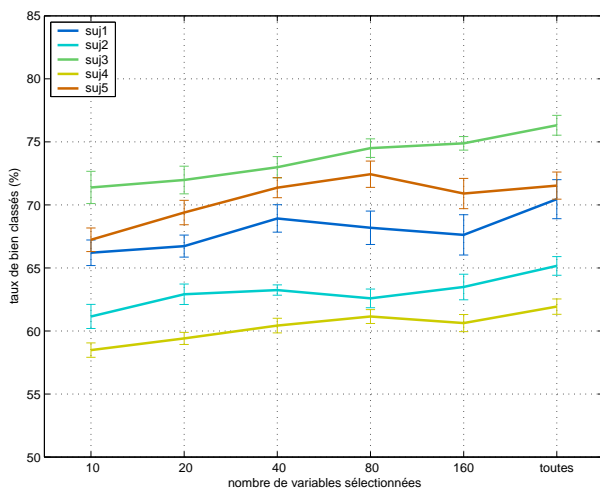


FIGURE 5.3 – Tests t en fonction de la fréquence pour l'ensemble des couples de classes possibles pour le sujet 1

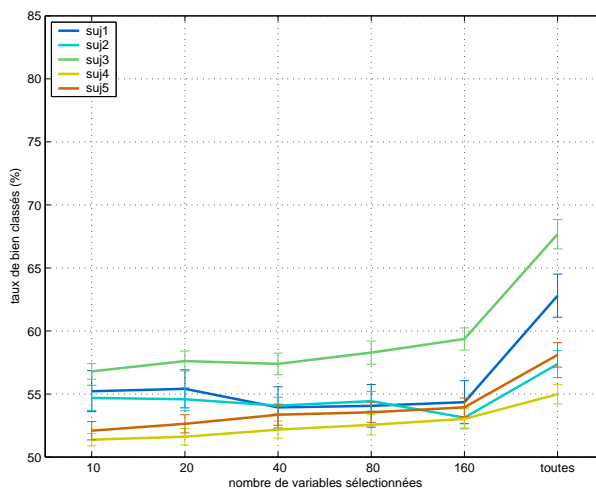
des interactions à distance quantifiées par la synchronie de phase sur tous les couples d'électrodes (figure 5.4(b)), 3) l'utilisation d'interactions à distance quantifiées par la cohérence sur tous les couples d'électrodes (figure 5.4(c)), 4) la combinaison synchronie locale (quantifiée par puissance spectrale) et synchronie à distance (quantifiée par la cohérence) (figure 5.4(d)).

L'ensemble de ces figures montre généralement une augmentation du taux de classification lorsque le nombre de variables augmente. En particulier, le taux de bien classés est ainsi maximal lorsque toutes les variables sont sélectionnées. Pour la mesure de puissance, la sélection d'un sous ensemble de 40 ou 80 variables conduit à des résultats proches de ceux sans sélection, c'est à dire lorsque toutes les variables sont considérées (moins de 3% de différence). Par contre, il existe un contraste net entre les résultats sans sélection et avec sélection sur les mesures d'interaction, en particulier sur la cohérence (de 4 à 13% d'amélioration). Ce contraste se retrouve aussi lorsque les mesures de puissance et de cohérence sont combinées dans le classifieur.

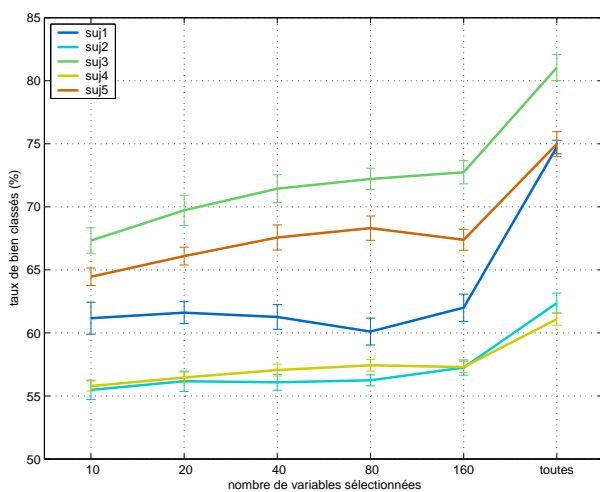
S'agissant de la comparaison des différentes quantifications, il y a une nette diffé-



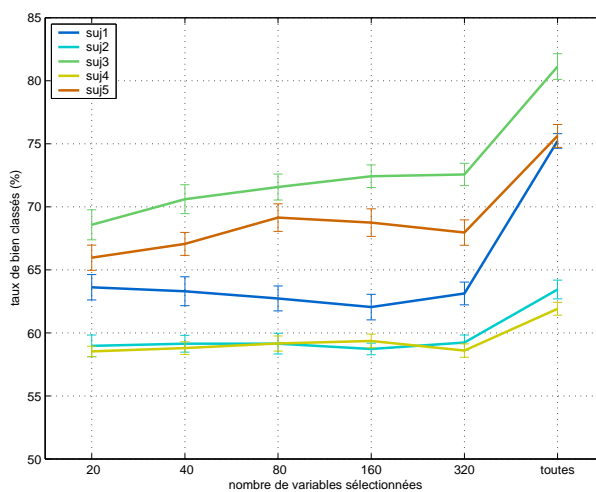
(a) mesures de puissance



(b) mesures de synchronie de phase



(c) mesures de cohérence



(d) mesures de puissance et de cohérence

FIGURE 5.4 – Taux de bien classés en fonction du nombre de variables sélectionnées en validation croisée inter-session pour différentes variables de quantification. Les résultats de classification sont moyennés sur l'ensemble des couples de tâches possibles. Lorsque toutes les variables sont utilisées par le classifieur, cela correspond à 300 variables pour les mesures de puissance (60 électrodes \times 5 bandes de fréquences) et à 9000 variables pour les mesures de cohérence ou de synchronie de phase (1800 couples d'électrodes \times 5 bandes de fréquences).

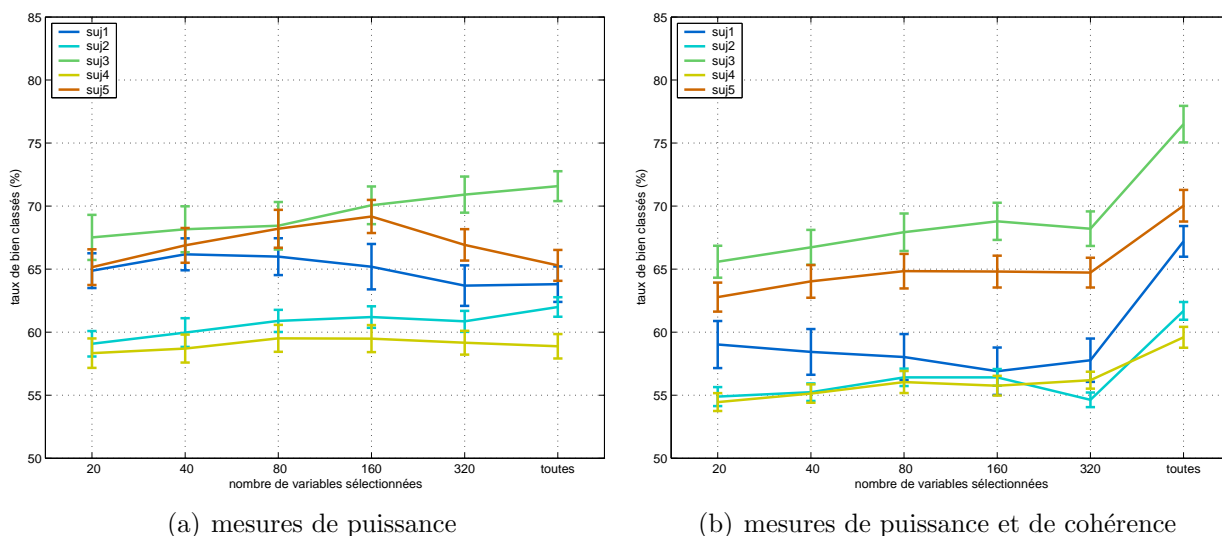


FIGURE 5.5 – Taux de bien classés en fonction du nombre de variables sélectionnées en validation croisée *inter-jours* pour différentes variables de quantification. Les résultats de classification sont moyennés sur l’ensemble des couples de tâches possibles.

rence entre les deux mesures d’interactions à distance (synchronie de phase et cohérence) la deuxième donnant des résultats nettement supérieurs à la deuxième (au moins 5% d’amélioration). La comparaison entre cohérence et puissance est moins tranchée ; en utilisant le maximum de variables disponibles la cohérence donne des résultats supérieurs sur les meilleurs sujets (1, 3 et 5) et des résultats comparables ou légèrement inférieurs sur les deux autres. Par contre, la sélection d’un faible sous ensemble de variables est nettement à l’avantage de la mesure de puissance. En combinant la meilleure mesure d’interaction (la cohérence) avec la mesure de synchronie locale (la puissance), les résultats obtenus approchent le maximum des performances obtenues sur les deux mesures séparément.

Pour quantifier la stabilité à long terme du classifieur, les résultats de validation croisée *inter-jours* sont aussi représentés figures 5.5(a) et 5.5(b) en utilisant uniquement les variables de puissance ou la réunion des variables puissance et cohérence (respectivement). Par rapport aux résultats obtenus en validation croisée *inter session*, on observe une perte de 5% pour la performance maximale obtenue en général en utilisant toutes les variables de quantification. De plus, l’utilisation de la cohérence combinée à la puissance apparaît encore comme plus discriminante que la puissance seule.

5.2.3 Influence des états mentaux choisis

Après avoir étudié globalement l’influence de la quantification sur les performances du classifieur, il est intéressant de savoir si ces résultats sont spécifiques à certaines tâches mentales. Afin d’étudier les performances en discrimination des différents couples

d'états mentaux, les résultats en validation croisée inter-session par classe et par sujet basés sur la *meilleure quantification* (puissance et cohérence) sont représentés figure 5.6. Ces résultats sont classés par performance croissante et regroupés suivant le type de couple d'état mental : couple de tâches motrices, couple de tâches non-motrices et finalement couple mixte (tâche motrice versus non-motrice). À l'intérieur du groupe "moteur", la tâche la plus contrastée est le couple imagination de la langue versus imagination de la préhension. Au sein du groupe "non-moteur", les couples faisant intervenir le calcul sont les meilleurs. Enfin dans le groupe mixte, le contraste calcul versus une autre tâche motrice donne les meilleurs résultats. Ces résultats sont observés chez tous les sujets à l'exception du sujet 2, qui donne les moins bons résultats.

Pour évaluer l'effet de la quantification en fonction de l'état mental étudié, les taux de bien classés pour la quantification par la puissance et par la puissance combinée à la cohérence sont comparés en fonction du couple d'états mentaux (figure 5.7). Il apparaît ainsi que l'apport de la mesure de cohérence permet d'améliorer les performances du classifieur pour les couples "non moteurs" ou "mixtes", en particulier dans les tâches impliquant le calcul.

5.3 Discussion

5.3.1 Choix des bandes de fréquence

Les bandes choisies sont reproductibles à travers les sujets et ce notamment dans les hautes fréquences (bandes bêta). Cependant, le découpage des basses fréquences est assez variable. On remarque de façon générale que le découpage au voisinage de 12Hz correspond plutôt à l'activité mu (13-14Hz) du cortex moteur et non à l'activité alpha traditionnellement observée dans le cortex occipital (8-12Hz). Cela peut traduire le fait que l'activité du cortex moteur est plus discriminante pour les tâches considérées que le cortex occipital. En basses fréquences (bandes 1 et 2) le découpage est plus variable. Il semble impliquer la bande delta (1-4Hz), qui a été observée en MEG dans des tâches visuo-motrices (Jerbi et al., 2007) et la bande thêta (4-8Hz), qui peut être impliquée dans le calcul mental. La variabilité peut s'expliquer par les différences inter-individuelles associées à ces bandes qui ont été reportées dans le cas du thêta (Klimesch, 1999).

5.3.2 Sélection de variables

Contrairement à l'a priori que nous avons posé au départ, la sélection de variables semble nuire à la classification. Bien que pour la puissance la détérioration reste minimale, dans le cas des interactions à distance la sélection d'une centaine de variables les plus discriminantes contient bien moins d'informations que l'ensemble des variables. Cette observation peut refléter la difficulté rencontrée par la sélection de variables univariées à trouver des variables discriminantes dans un espace de grande dimension, où des variables de bruit peuvent être sélectionnées accidentellement par le jeu des comparaisons

multiples. Une autre interprétation possible qui ouvre des perspectives sur les chapitres suivants est que l'information contenue dans les mesures d'interactions à distance puisse être répartie sur un grand nombre de variables. Ce phénomène pourrait être dû à la diffusion des activités électriques corticales, la mesure de cohérence entre deux capteurs ne fournissant qu'une information partielle sur l'interaction entre deux aires corticales situées dans leur voisinage. Dans ce cas, une analyse des données au niveau des sources corticales devrait améliorer la quantification des interactions.

5.3.3 Quantification

Les résultats montrent l'intérêt d'utiliser les mesures d'interaction à distance dans le cadre des ICM asynchrones. Concernant les deux mesures d'interaction à distance proposées, la mesure de cohérence apparaît ici plus performante que la mesure de synchronie de phase. Nous pensons que cela est dû aux méthodes de calcul respectives de ces mesures. En effet l'utilisation de multiples fenêtres pour calculer la cohérence à la manière du périodogramme de Welch confère à la mesure de cohérence de bonnes propriétés asymptotiques (Rosenblatt, 1985) qui s'expliquent intuitivement par le moyennage de mesure indépendantes qui réduit la variance de l'estimateur⁷ par la loi des grands nombres. Le calcul de la synchronie de phase est quand à lui réalisé sur la fenêtre complète en utilisant un simple filtrage dans le domaine de Fourier. Dans les paradigmes usuels, autres que la BCI, ce n'est pas un problème pour ce type de mesure car elles sont moyennées à travers les essais. Une étude plus poussée pourrait conduire à une amélioration de la méthode d'estimation de la synchronie, la rendant accessible à l'utilisation en essai unique requise par les ICM (en utilisant éventuellement les travaux de (Lachaux et al., 2000)).

La mesure de cohérence apparaît ainsi comme une mesure compétitive en terme de quantité d'informations par rapport à la mesure de synchronie locale qu'est la puissance spectrale. L'amélioration procurée par celle-ci chez plusieurs sujets permet d'affirmer que la cohérence apporte des informations discriminantes supplémentaires (ou complémentaires) par rapport à la puissance spectrale.

Enfin, on observe que les mesures utilisées restent stables au cours des différents jours d'enregistrement, car l'écart entre les taux de classification inter-jours et inter-sessions est peu élevé (perte de 5%). C'est écart peut rendre compte des fluctuations de ligne de base qui ont lieu entre les jours d'enregistrements successifs. Il pourrait être envisagé de le corriger en réalisant une calibration du classifieur à chaque nouveau jour d'enregistrement. On remarque aussi que la sélection de variables n'apporte pas, là encore, plus de robustesse à la classification.

7. En pratique, comme il existe un recouvrement des fenêtres temporelles utilisées, les mesures sur chaque fenêtres ne sont pas tout à fait indépendantes. La réduction de la variance l'estimateur est donc a priori moins importante

5.3.4 Influence des tâches mentales

L'analyse par couple d'états mentaux des résultats de classification met en évidence l'intérêt de contraster une tâche motrice avec une tâche non motrice. En effet, chez tous les sujets, les meilleurs résultats sont obtenus en utilisant un couple de tâches de ce type. En particulier, le calcul apparaît bien contraster avec les imageries motrices. Dans le groupe des couples d'imagerie motrice, c'est sans surprise que l'on trouve une meilleure discrimination pour les tâches *préhension versus langue* qui correspondent aux aires du cortex moteur primaire les plus éloignées. Cependant il est possible que l'activation d'aires pariétales traitant l'aspect spatial de la tâche de préhension contribue aussi à favoriser la discrimination (en comparaison avec la tâche d'imagerie de l'index). Finalement, au sein du groupe des couples non-moteurs, c'est encore le calcul qui semble se différencier le plus facilement des autres tâches.

Il peut être raisonnable d'interpréter ces bonnes performances de la tâche de calcul par le fait qu'elle implique un engagement plus important de la mémoire de travail que les autres tâches non motrices. De plus son meilleur contraste avec les tâches d'imagerie motrice peut s'expliquer par un très faible engagement des aires motrices, en comparaison des tâches de navigation et d'imagerie qui peuvent inciter le sujet à se représenter un mouvement pour se déplacer ou pour battre le rythme de la musique.

Finalement, l'apport spécifique des mesures de cohérence en comparaison avec les puissances spectrales bénéficie aux couples qui font intervenir une tâche non motrice, ce qui est cohérent avec l'hypothèse que ces tâches sont caractérisées par des réseaux corticaux largement distribués qu'il est possible de contraster entre eux ou avec des imageries motrices. Pour le couple moteur *index versus langue*, il n'y a pas d'amélioration, ce qui soutiendrait l'idée que les réseaux corticaux à large échelle sont similaires et se différencient principalement par l'aire du cortex moteur spécifique à la tâche.

Les variables exploitées par le classifieur pour la discrimination ne sont pas étudiées dans ce chapitre. Les aires corticales et les bandes de fréquences impliquées dans ces tâches sont étudiées au chapitre 7.

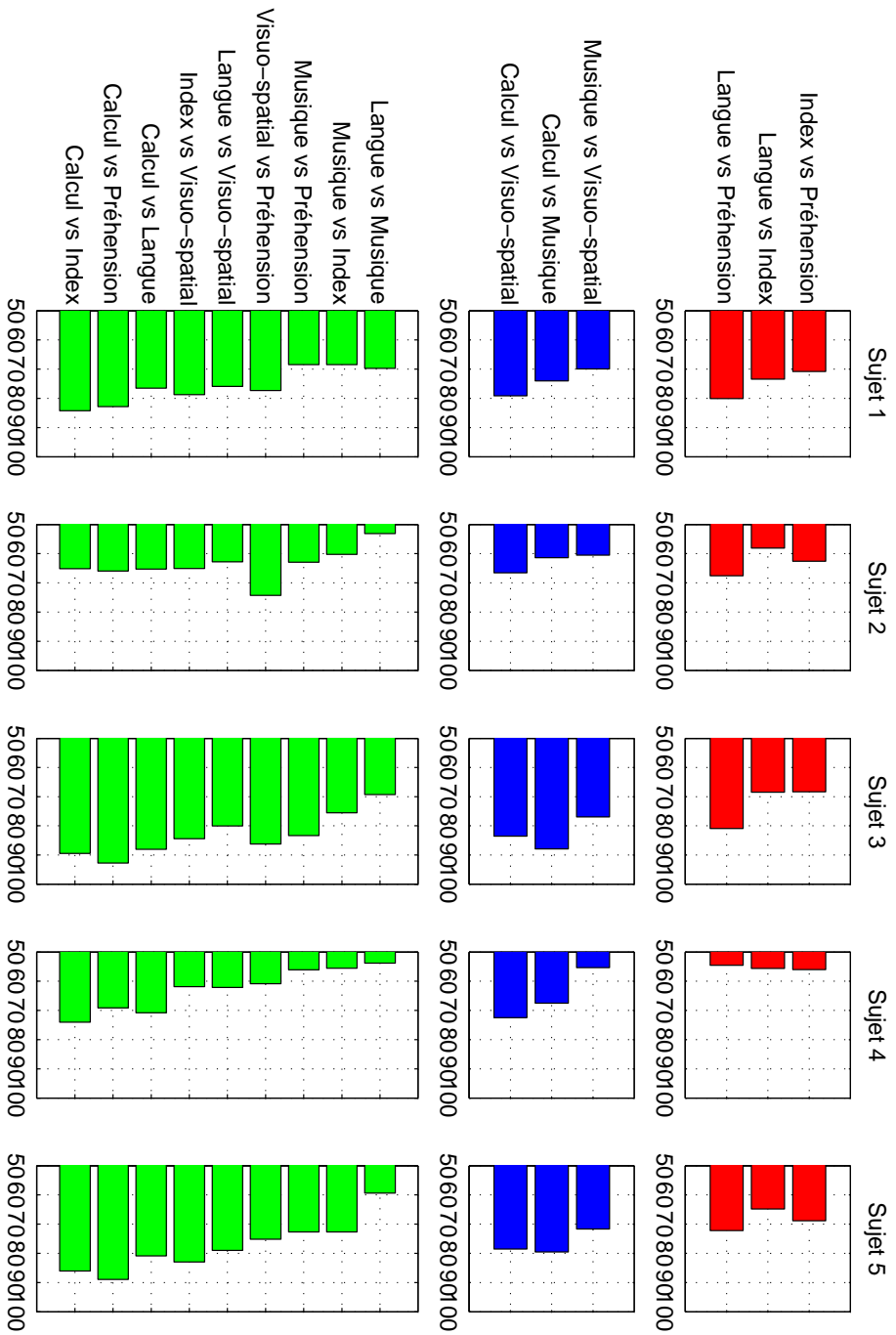


FIGURE 5.6 – Comparaison du taux de bien classés en validation croisée inter-sessions entre les différents groupes de discriminations possibles : en rouge tâche motrice versus tâche motrice, en bleu tâche non motrice versus non motrice et en vert tâche motrice versus non motrice.

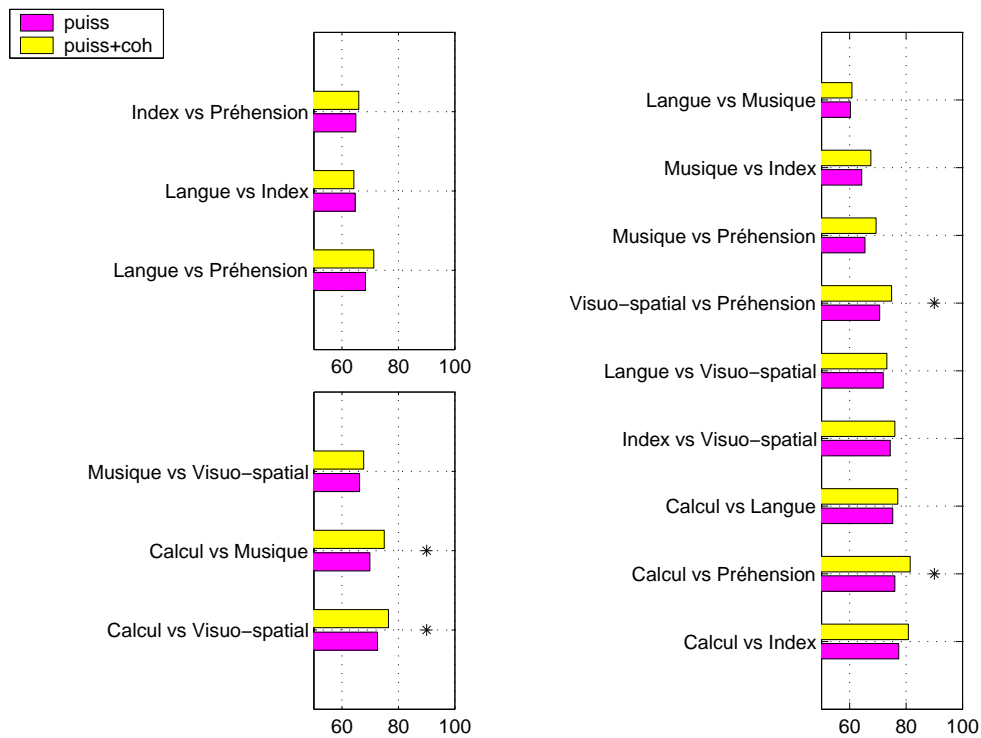


FIGURE 5.7 – Comparaison des taux de classification en validation croisée intersession pour les quantifications par la puissance seulement et par la puissance combinée à la cohérence. Les résultats sont donnés pour chaque classe, moyennés à travers les sujets. Les étoiles indiquent les couples pour lesquels la performance du classifieur est significativement supérieure lorsque les mesures de puissances sont combinées aux mesures de synchronies ($p_{cor} < .05$). Le test utilisé est un test Fisherien apparié.

Chapitre 6

Analyse par séparation de sources

6.1 Cadre de l'étude

Dans ce chapitre, nous incluons dans la chaîne de traitement des signaux EEG ou MEG un prétraitement par filtre spatial optimisé en fonction des propriétés statistiques du signal EEG. Les méthodes utilisées pour déterminer ce filtre spatial sont inspirées de travaux en séparation de sources (Belouchrani et al., 1997; Cardoso and Souloumiac, 1996).

6.1.1 Hypothèses

Nous supposons que les mesures EEG ou MEG sont une combinaison linéaire d'activités indépendantes deux à deux qui reflètent l'activation ou la désactivation de réseaux corticaux sous-jacents spécifiques à certaines tâches mentales¹. Les signaux mesurés $\mathbf{m}(t)$ s'expriment donc en fonction d'un vecteur de composantes $\mathbf{c}(t)$ tel que :

$$\mathbf{m}(t) = \mathbf{N}\mathbf{c}(t)$$

où \mathbf{N} est la matrice de mélange des différentes composantes. Comme des oscillations apparaissant en EEG et en MEG ont été associées à des traitements cognitifs spécifiques (voir chapitre 1), extraire les activités de ces ondes, sur différentes composantes, et mesurer leur densité spectrale de puissance peut permettre de réaliser une quantification optimale des signaux EEG ou MEG en vue de la différenciation de tâches cognitives. Pour estimer l'activité des composantes \mathbf{c} , il faut déterminer une matrice de filtrage spatial $\boldsymbol{\theta}$ qui permet d'inverser au mieux le mélange des composantes enregistrées sur les capteurs :

$$\hat{\mathbf{c}}(t) = \boldsymbol{\theta}\mathbf{m}(t) \approx \mathbf{c}(t)$$

1. L'indépendance correspond à une hypothèse de spécialisation fonctionnelle selon laquelle des régions corticales spécialisées réalisent des traitements de manière indépendante. Bien qu'elle ne soit pas vérifiée a priori, cette hypothèse simplificatrice permet de négliger les interactions à distance en aires corticales.

À cause de la nature des hypothèses faites sur les composantes, celles-ci ne peuvent en général être retrouvées qu'à une permutation et un facteur d'échelle près (Comon, 1994).

Pour déterminer $\boldsymbol{\theta}$, plusieurs méthodes de séparation de sources peuvent être envisagées suivant les hypothèses statistiques faites sur les sources. Dans le cas présenté ici, une hypothèse particulièrement adaptée est de supposer que toutes les composantes sont indépendantes et qu'elles ont des spectres différents (linéairement indépendants). Dans ce cas, il est possible d'utiliser l'algorithme SOBI (Second Order Blind Identification) (Belouchrani et al., 1997).

6.1.2 Principe de la méthode SOBI

De façon similaire à d'autres algorithmes de séparation de sources, l'hypothèse d'indépendance implique une hypothèse de décorrélation des composantes qui permet d'appliquer une matrice de blanchiment \mathbf{B} aux signaux EEG de façon à imposer une matrice de corrélation unité aux signaux blanchis \mathbf{m}_b ².

À l'issue du blanchiment, $\hat{\mathbf{c}}$ est lié à \mathbf{m}_b par la relation $\hat{\mathbf{c}} = \mathbf{G}\mathbf{m}_b$ où \mathbf{G} est à rechercher parmi les matrices orthogonales (Hyvärinen and Oja, 1997)³.

Or l'hypothèse d'indépendance des composantes permet de supposer que les matrices de covariance retardées du vecteur \mathbf{c} sont diagonales, quel que soit le retard τ .

$$R_{\mathbf{c}}(\tau) = \mathbb{E} [\mathbf{c}(t)\mathbf{c}(t - \tau)^T] = \mathbf{D}(\tau)$$

donc les matrices de covariance retardées du vecteur de mesures blanchies \mathbf{m}_b sont diagonalisables dans une base orthogonale commune.

$$R_{\mathbf{m}_b}(\tau) = \mathbf{G}\mathbb{E} [\mathbf{c}(t)\mathbf{c}(t - \tau)^T] \mathbf{G}^T = \mathbf{G}\mathbf{D}(\tau)\mathbf{G}^T$$

Pour déterminer \mathbf{G} il est alors possible de réaliser une diagonalisation conjointe de plusieurs matrices de covariances retardées pour différentes valeurs de retards en utilisant la méthode de Jacobi étendue (Cardoso and Souloumiac, 1996). Celle-ci permet d'ajuster une matrice de changement de base orthogonale (ou unitaire dans le cas complexe) de façon à minimiser itérativement la somme des carrés des termes hors diagonaux de l'ensemble des matrices considérées. Le succès de cette méthode repose sur le fait que les différentes composantes doivent avoir des fonctions d'autocorrélations (les éléments diagonaux des matrices $\mathbf{D}(\tau)$) linéairement indépendantes (Belouchrani et al., 1997). Les composantes recherchées doivent donc nécessairement avoir des densités spectrales de puissance linéairement indépendantes.

2. Soit $\hat{\mathbf{R}}_{\mathbf{m}}(0) = \mathbf{V}\mathbf{D}\mathbf{V}^T$ la diagonalisation de la matrice de covariance empirique des signaux EEG et alors la matrice de blanchiment s'écrit : $\mathbf{B} = \mathbf{D}^{-1/2}\mathbf{V}^T$, où $\mathbf{D}^{-1/2}$ représente la matrice \mathbf{D} et où chaque élément diagonal a été remplacé par l'inverse de sa racine carrée. De plus, pour garantir une solution stable, les valeurs propres de faible amplitude et les vecteurs propres correspondants sont supprimés

3. Dans le cas complexe la matrice doit être unitaire, et les transpositions dans les formules suivantes doivent être remplacées par des conjugaisons Hermitiennes

Finalement, lorsque ces conditions sont vérifiées, le filtre spatial obtenu est $\boldsymbol{\theta} = \mathbf{G}^T \mathbf{B}$. Une estimation de la matrice de mélange \mathbf{N} est alors donnée par $\mathbf{N} = \boldsymbol{\theta}^\dagger$ qui désigne la pseudo inverse du filtre spatial. La i -ème colonne de \mathbf{N} correspond à la projection sur les capteurs de l'activité de la i -ème composante, il est ainsi possible de visualiser la distribution spatiale de chaque composante pour interpréter son origine physiologique.

La matrice de filtrage spatial ainsi réalisée peut être utilisée comme un prétraitement destiné à rehausser les signaux oscillatoires du signal EEG. Après le prétraitement, l'activité des composantes peut être quantifiée par leurs puissances spectrales dans différentes bandes de fréquences. Ces variables peuvent ensuite être utilisées par un classifieur (par exemple un SVM). C'est ce qui est réalisé dans l'article suivant.

6.2 Application à des données MEG

Présentation

Dans l'article suivant, cette méthode est appliquée sur des données MEG de contrôle visuo moteur enregistrées par Karim Jerbi au laboratoire ([Jerbi et al., 2007](#)). L'objectif était de différencier les périodes de repos des périodes où le sujet effectuait la tâche visuo motrice. Ces signaux MEG étaient quantifiés sur des fenêtres temporelles d'une seconde en utilisant des mesures de puissance spectrale dans six bandes de fréquence (2-4 Hz, 5-7 Hz, 8-12 Hz, 13-30 Hz, 30-60 Hz, 60-90 Hz). Une adaptation de l'algorithme SOBI original a été utilisée afin de prendre en compte l'information donnée par la tâche mentale dans la séparation de sources : après blanchiment, les matrices de covariances retardées sont estimées séparément pour les deux conditions expérimentales donnant deux sous-ensembles de matrices $R_{\mathbf{m}_b}^1(\tau_i)$ et $R_{\mathbf{m}_b}^2(\tau_i)$ pour différents retards τ_i . L'algorithme de diagonalisation conjointe est appliqué alors à l'ensemble de ces matrices pour trouver la matrice solution $\boldsymbol{\theta}$. Cette modification permet de détecter des composantes qui ont des variations de spectre différentes en fonction de la tâche mentale effectuée⁴.

Les résultats (figure 1 page 207) montrent un meilleur taux de bien classés lorsqu'une quantification sur les composantes issues de l'algorithme SOBI est utilisée, par rapport à une quantification similaire sur les électrodes. Cependant, ces meilleurs résultats sont obtenus en sélectionnant un faible nombre de variables (et donc de composantes) par le score de Fisher. Lorsque le nombre de variables sélectionnées augmente, le taux de classification diminue, probablement à cause de la sélection de composantes de bruit. La distribution sur les capteurs et la densité spectrale de puissance des composantes les plus discriminantes sont représentées figure 2 page 207. Elles montrent une composante latéralisée susceptible de correspondre au cortex moteur primaire et une composante

4. Deux composantes associées à des tâches différentes peuvent avoir des propriétés spectrales identiques en moyenne. C'est le cas si l'on considère des DLE dans la bande mu dans le cortex moteur, associées aux tâches de mouvement du bras gauche et au bras droit : en moyenne sur les deux conditions, les activités dans la bande mu peuvent être les mêmes, par contre leurs densités spectrales de puissance dans chaque conditions varient de manière opposée.

centrale qui pourrait être associée à l'aire motrice supplémentaire. De plus, conformément aux hypothèses de la méthode, ces deux composantes ont des propriétés spectrales différentes : la première composante est caractérisée par une désynchronisation (DLE : voir chapitre 1) dans la bande mu lors de l'exécution de la tâche visuo-motrice, tandis que pour la deuxième composante la DLE a lieu principalement dans la bande bêta.

6.2.1 Article publié dans International Congress Series

Prediction of cognitive states using MEG and Blind Source Separation

M. Besserve^{a,b,*}, K. Jerbi^{a,1}, L. Garnero^{a,b}, J. Martinerie^{a,b}

^a CNRS UPR 640-LENA Cognitive Neuroscience and Brain Imaging Lab, Paris, France

^b Pierre and Marie Curie University, Paris, France

Abstract. The present study investigates the predictability of a subject's state based on the classification of the underlying brain activity recorded via magnetoencephalography (MEG). We use Second Order Blind Identification (SOBI) to reduce the high dimensionality of MEG sensors into a smaller number of task-related components. A classification of distinct cognitive states is then achieved by feeding the spectral power of these components into a Support Vector Machine (SVM). We tested this approach on data from one subject during a visuomotor control experiment and found that our method outperforms classification based on the spectral powers computed directly from the MEG sensor array. Our findings suggest that combining SOBI and SVM may provide a reliable classifier for the prediction of cognitive states in MEG. © 2007 Elsevier B.V. All rights reserved.

Keywords: Magnetoencephalography (MEG); Blind Source Separation (BSS); Second Order Blind Identification (SOBI); Support Vector Machine (SVM); Visuomotor control

1. Introduction

The prediction of cognitive states is crucial to applications such as Brain Computer Interfaces (BCI). We investigated this problem using magnetoencephalography (MEG) for the classification of continuous cognitive tasks.

Computing the power spectral densities of MEG sensors in selected frequency bands is a straightforward method to extract task relevant parameters for classification, but the high number of sensors leads to a high number of features that impairs classification accuracy. We

* Corresponding author. Laboratoire de Neurosciences Cognitives and Imagerie Cérébrale, CNRS UPR640-LENA; 47, boul. de l'hôpital; 75651 Paris Cedex 13 - France. Tel.: +33 1 42 16 11 72; fax: +33 1 45 86 25 37.

E-mail address: michel.besserve@chups.jussieu.fr (M. Besserve).

¹ Currently in post-doc at Inserm U280, Mental Processes and Brain Activation Lab, Lyon, France.

solved this problem by applying a Blind Source Separation (BSS) technique to represent MEG sensor data as an instantaneous linear mixing of a restricted number of task relevant components. Second Order Blind Identification (SOBI) is a BSS technique particularly suitable to our problem since it allows retrieving uncorrelated components relying on their spectral properties [1,2]. Once a reduced number of task-related components had been determined using SOBI, the next step was to estimate the power spectral density of each one of these relevant components. The power values averaged over specific frequency bands were then processed by a classifier in order to differentiate between two cognitive tasks based on MEG recordings of the ongoing neural activity. To achieve good classification accuracy, we used the linear Support Vector Machine (SVM) classifier combined with variable selection.

2. Materials and methods

2.1. SOBI algorithm

We assumed that the vector of observed MEG signals $\mathbf{x}(t)$ (which is supposed centered to simplify notations) is an instantaneous linear mixing of a vector of uncorrelated components with unit variance $\mathbf{s}(t)$ such that $\mathbf{x}(t) = \mathbf{M}\mathbf{s}(t)$, where \mathbf{M} is the mixing matrix. SOBI only uses MEG signals to compute both an estimated unmixing matrix $\mathbf{U} \approx \mathbf{M}^{-1}$ and the estimated time course of the components, $\hat{\mathbf{s}}(t) = \mathbf{U}\mathbf{x}(t)$. This estimate is achieved using information on temporal coherence of the wide-band (0–100 Hz in our case) MEG signals contained in the empirical delayed correlation matrices $\mathbf{R}_x(\tau_i)$ defined at many time delays $(\tau_i)_{i=1..D}$ such that:

$$\mathbf{R}_x(\tau_i) = \frac{1}{N-1} \sum_{n=0}^{N-1} \mathbf{x}(n)\mathbf{x}^T(n-\tau_i).$$

The algorithm we used proceeds in two steps:

- Whitening: the original MEG signals are transformed into whitened signals $\mathbf{z}(t)$ (uncorrelated components of unit variance) which are the first principal components of $\mathbf{x}(t)$ representing 99% of the total variance. This is achieved with a whitening matrix \mathbf{Q} such that $\mathbf{z} = \mathbf{Q}\mathbf{x}$.
- Approximate joint diagonalization of the delayed correlation matrices $\mathbf{R}_z(\tau_i)$ that aims at finding an orthogonal matrix \mathbf{W} that jointly minimizes the magnitude of the off-diagonal terms of all the matrices $\mathbf{W}\mathbf{R}_z(\tau_i)\mathbf{W}^T$. For our specific classification purposes we computed separately delayed correlation matrices for the signals corresponding to the cognitive state 1, $\mathbf{R}_z^1(\tau_i)$, and the cognitive state 2, $\mathbf{R}_z^2(\tau_i)$ (τ_i ranging from 0 to 300 ms by steps of 2 ms), and included both matrix sets in the joint diagonalization.

2.2. Classification

The resulting estimated unmixing matrix $\mathbf{U} = \mathbf{W}\mathbf{Q}$ was used to estimate the time course of the SOBI components and the distribution of each component on the scalp was given by each column of the estimated mixing matrix $\hat{\mathbf{M}} \approx \mathbf{U}^+$ (the Moore–Penrose pseudo-inverse). The power spectra of these signals were then computed using the Welch periodogram on each

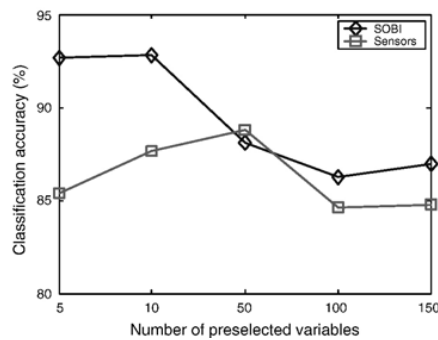


Fig. 1. Classification accuracy obtained with SOBI quantification and sensor quantification for various numbers of selected variables.

successive 1-s time windows of the MEG recordings. The average power obtained for each one of six distinct frequency bands (δ : 2–4 Hz, θ : 5–7 Hz, α : 8–12 Hz, β : 13–30 Hz, γ_1 : 30–60 Hz, γ_2 : 60–90 Hz) provided the final quantification associated with each time window. A learning set made up of quantified time windows corresponding to each of the two cognitive states was used to train a SVM classifier. Although restricting the analysis to the SOBI components reduced the number of variables, the quantification described here still yielded a high dimensional feature space (six times the number of components) and severely impaired the classification accuracy due to overfitting of the learning set. We therefore further reduced the dimensionality via variable selection: i.e. only the most discriminant features (according to a Student's *t*-test) among the spectral powers were fed into the SVM.

3. Results

We applied this novel approach to MEG data recorded with a whole scalp (151 channels) CTF neuromagnetometer during a continuous visuomotor coordination experiment which required subjects to continuously manipulate a track-ball to compensate the random rotations of a cube projected on a display screen [3]. The MEG recording sessions were split into 1-s time windows and separated in two classes: the first corresponding to rest periods (Rest) while

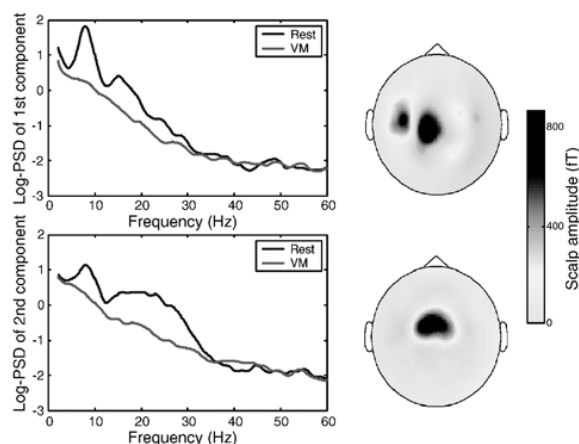


Fig. 2. The two SOBI components for which spectral power density is the most discriminant. Left column: log power spectral density for each class. Right column: scalp topography in femto tesla (fT).

the second contained the epochs of sustained visuomotor control (VM). Each session contained approximately 2 min of rest periods and 4 min of VM activity. We computed classification accuracy with a cross-validation across two experimental sessions of one subject with two quantification methods using the multi-frequency signal power of either a) the SOBI components or b) the whole sensor data directly. The resulting classification accuracies of the two methods are plotted in Fig. 1 for various numbers of preselected variables. Moreover, Fig. 2 shows the scalp topographies (see Section 2.2) and the power spectra of the two SOBI components for which power spectral density features were the most discriminant for classification.

4. Discussion

The classification results show that the use of SOBI allows for a better prediction of the MEG states than the one achieved with sensor quantification, and with fewer selected variables. This highlights the ability of SOBI to extract task relevant components giving more information on the cognitive state than direct MEG sensor measurements. The physiological relevance of MEG components is also supported by the topographies of the two most discriminant SOBI components (Fig. 2). These spatial maps of the scalp amplitude of each component correspond most likely to primary sensorimotor cortex (S1–M1) and the supplementary motor area (SMA) both known to be involved in sensorimotor behavior. In accordance with the literature, the power spectra of these components reveal event-related desynchronization (ERD) in α and β bands during VM activity compared to the resting state. Moreover, the classifier accuracy reached 93% reflecting the high separability of short periods of VM activity and rest using MEG.

5. Conclusion

Our study shows that SOBI can be successfully applied to ongoing MEG signals to extract oscillatory activities involved in a continuous cognitive task. The SOBI-based methodology described in this paper transforms the MEG sensor space to a lower dimensional space that provides a highly interpretable representation of MEG signal and thereby improves the performance of the classifier. More generally, our results also indicate that MEG imaging might be particularly useful to classify mental states. Further work will include other subjects and task specificity of the SOBI components will be studied under more control conditions (e.g. pure motor activity).

References

- [1] A. Belchourani, et al., A blind source separation technique using second order statistics, *IEEE Trans. Signal Process.* 42 (2) (Feb 1997) 434–444.
- [2] A.C. Tang, M.T. Sutherland, C.J. McKinney, Validation of SOBI components from high-density EEG, *Neuroimage* 25 (2) (2005) 539–553.
- [3] K. Jerbi, et al., Imaging cortical oscillations during sustained visuomotor coordination in MEG, *Proc. IEEE ISBI* (2004) 380–383.

6.3 CSDA : Un algorithme hybride de séparation de sources pour la classification

6.3.1 Présentation

La méthode utilisée dans l'article précédent a deux inconvénients :

- le choix des retards à prendre en compte dans le processus de diagonalisation conjointe contient une part d'arbitraire
- la quantification de l'activité de composantes nécessite le calcul de leur puissance spectrale dans des bandes qui ne sont pas connues a priori
- il est nécessaire d'entraîner un classifieur sur les variables de quantifications extraites de chaque composante

L'objectif du papier suivant est de rassembler les étapes de filtrage spatial, de quantification et de classification dans un même algorithme. Pour accomplir cet objectif les modifications suivantes sont apportées à l'algorithme SOBI

1. le problème d'optimisation de la matrice de filtrage spatial est formulé dans le domaine fréquentiel plutôt que dans le domaine temporel : les matrices de covariances retardées sont remplacées par des *matrices de densité spectrales* dans une gamme de fréquence susceptibles de contenir des oscillations spécifiques à la tâche.
2. les variables de quantification sont des densités spectrales de puissances qui correspondent aux éléments diagonaux des matrices de densité spectrales
3. le classifieur utilisé est du type analyse linéaire discriminante (ALD), avec un nombre de paramètres réduits pour éviter le surapprentissage
4. l'optimisation de la matrice de filtre spatial a pour objectif de maximiser un critère de performance du classifieur.

L'ensemble de ces considérations permet de montrer qu'une matrice de filtrage spatial peut être ajustée par la méthode de Jacobi pour maximiser le pouvoir discriminant des densités spectrales de puissances des composantes obtenues par le filtre et ce en réalisant une diagonalisation conjointe de la différence entre les matrices de densité spectrales associées à chaque tâche à différencier. Cette méthode est développée dans l'article suivant. Elle est comparée à la méthode usuelle de patterns spatiaux communs (CSP : voir section 2.3.2) sur des données simulées et sur nos enregistrements EEG d'ICM asynchrone. Les résultats (table 1 et figure 4 dans l'article) montrent que la méthode permet d'analyser le signal EEG sans sélectionner une bande de fréquence d'intérêt, alors que cette étape est nécessaire pour le CSP. Comme il est précisé dans les perspectives à la fin de l'article, l'algorithme gagnerait à être amélioré en introduisant une sélection automatique de bandes de fréquences d'intérêt. Cela permettrait de limiter le nombre de paramètres du classifieur en ce concentrant sur les fréquences les plus discriminantes. Une méthode de sélection récursive des fréquences est évoquée dans l'article. Cependant, une autre piste pourrait être envisagée en exploitant les principes de l'analyse en composantes parcimonieuses (Cardoso and Pham, 2003; Gribonval

and Nielsen, 2003). Cette approche permet de contraindre un problème de séparation de sources de manière à favoriser les solutions dont la distribution de puissance est concentrée dans une région fréquentielle de faible étendue.

6.3.2 Article présenté à la troisième conférence EMBS-IEEE Neural Engineering

Cross-Spectral Discriminant Analysis (CSDA) for the classification of Brain Computer Interfaces.

Michel Besserve, Line Garnero and Jacques Martinerie

Abstract—Mining discriminant informations through the spatial and frequency domains is a challenging problem in Brain Computer Interfaces (BCI), especially when the experimental paradigm is not as well-known as motor imagery. We propose a new classification algorithm that extracts relevant oscillatory activities from ongoing electroencephalographic (EEG) signals without prior knowledge of the discriminant frequency bands. In this approach, a spatial filtering matrix transforms EEG signals into several components. In order to minimize the generalization error of a Bayesian classifier, the spatial filter is tuned using Spectral Density Matrices. The algorithmic implementation of this method uses explicitly the cross-spectral properties of EEG data. Results on both realistic simulations of EEG data and real asynchronous BCI experiments show the interest of the method and its superiority over the Common Spatial Pattern (CSP) algorithm.

I. INTRODUCTION

Powerful classification methods of EEG rhythms are crucial for BCI applications, providing more and more efficient communication devices for impaired patients. In particular Common Spatial Pattern (CSP) has become a popular spatial filtering technique for the preprocessing of EEG in the classification of Event Related Desynchronization (ERD) occurring during motor imagery [1]. This technique is currently used as a semi-supervised preprocessing step: EEG signals are filtered in a given frequency band (e.g. the 8-12 Hz alpha band); CSP is then applied to obtain many spatial filters aimed at discriminating between different motor imageries. The operator then selects the useful spatial filters according to their scalp topography (generally those located on the motor cortex). Finally, a classifier uses power measurements of the resulting filtered components to predict the motor imagery performed by the subject. From the above description, one can observe that the use of CSP strongly depends on a priori knowledge of the EEG correlates of the experimental paradigm.

This prior knowledge is often missing in the field of asynchronous BCI, as continuous mental tasks can not be studied in the well established ERD/ERS framework (no stimulus onset is available). In particular, if we want to explore various mental tasks, for example non motor tasks, the task-related topographic patterns and frequency contents of EEG recordings can be unknown and largely subject specific. We propose here a spatial filtering technique that can process the whole frequency range at the same time. This approach is based on the properties of Spectral Density

Matrices (SDM). For each frequency point, the diagonal of SDM provides the Power Spectral Density (PSD) of each channel whereas off-diagonal terms are called cross-spectra and provide information on the correlation between each channel at the considered frequency.

To ensure an automated selection of the appropriate component, the proposed approach includes directly the computation of the spatial filter as an optimization step of the classifier. This classifier is based on a Gaussian model, like Linear Discriminant Analysis (LDA).

The layout of the paper is as follows: In Sec. II we describe in details the CSDA algorithm. Sec. III presents the simulated and experimental data used for the validation of the algorithm. Results are provided in Sec. IV. and Sec. VI draws the conclusion of this work.

II. CSDA ALGORITHM

A. Mathematical framework

We assume the EEG is analysed using sliding time-windows, each time-window can correspond to a mental state (1 or 2) of the subject. Let $e_i(t)$ be the EEG signal recorded on the i -th time-window, it is a matrix of size (N, M) for a signal of N channels and M time points per window. Using e_i we want to infer the mental state $y_i \in \{1, 2\}$ of the subject. We use a spatial filtering matrix θ to transform e_i into $z_i = \theta e_i$ and denote by $Z_i(\nu)$ the vector of its discrete Fourier transform at frequency ν . Then $S_i^{zz}(\nu) = Z_i(\nu)Z_i(\nu)^T$ is the estimate of the SDM at frequency ν for the i th time-window. Our approach defines the appropriate spatial filtering θ such that the Power Spectral Densities (PSD) estimates of z_i (i.e. the diagonal terms of $S_i^{zz}(\nu)$) are the most useful for the discrimination of the two mental states.

B. Bayesian classifier

At each time window i all the diagonal elements of all spectral density matrices $S_i^{zz}(\nu_k)$, $k = 1, \dots, K$ (k denotes the number of frequency samples) can be used to construct a single feature vector x_i . We then use Bayesian decision theory to compute an optimal classifier to predict the mental state y from a new vector x . Basic assumptions of gaussianity and statistical independence of the features lead to the following discriminant function, assuming the two classes have equal prior probabilities:

$$f(x) = \log\left(\frac{P(x|y=1)}{P(x|y=2)}\right) \quad (1)$$

$$= \left(x - \frac{\mu_1 + \mu_2}{2}\right)^T \Sigma^{-1} (\mu_1 - \mu_2), \quad (2)$$

with

$$\Sigma = \text{diag}([\sigma_{1,1}^2, \dots, \sigma_{N,K}^2]),$$

and

$$\boldsymbol{\mu} \mathbf{j} = [\mu_{j_{1,1}}, \dots, \mu_{j_{N,K}}]^T,$$

where $\sigma_{n,k}^2$ is the estimated variance of the n -th diagonal element of $\mathbf{S}_i^{zz}(\nu_k)$ and $\mu_{j_{n,k}}$ its estimated class conditional mean for class j . When $f(\mathbf{x}) \geq 0$, the classifier predicts class 1, otherwise it predicts class 2.

C. Bounding the classification error

If we assume that the data actually respects the previous Gaussian model, one can estimate the Bayes error of the dataset, which is the lowest expectation of error rate any classifier can achieve on this data. This error rate can be reached asymptotically by the above mentioned Bayesian classifier and is defined by $\varepsilon_{\text{Bayes}} = P(\{f(\mathbf{x}) \geq 0\}, \{y = 2\}) + P(\{f(\mathbf{x}) < 0\}, \{y = 1\})$. Unfortunately, Bayes error can not be computed analytically in our case. Nevertheless, some bounds of this error rate are well-known and one of the simplest is the Bhattacharyya bound [2] that gives the following majoration under our assumptions:

$$\varepsilon_{\text{Bayes}} \leq \frac{1}{2} \exp(-\rho)$$

where

$$\rho = \frac{1}{8} (\boldsymbol{\mu} \mathbf{1} - \boldsymbol{\mu} \mathbf{2})^T \Sigma^{-1} (\boldsymbol{\mu} \mathbf{1} - \boldsymbol{\mu} \mathbf{2})$$

We assume that the number of time points enables us to use asymptotic properties on the periodogram estimate of PSD for stationary signals. Then at each frequency point the n th diagonal term of the spectral density matrix verifies [3]:

$$\begin{aligned} \sigma_{n,k}^2 &= \text{var}(\mathbf{S}^{zz}(\nu_k))_{nn} = \text{mean}(\mathbf{S}^{zz}(\nu_k))_{nn}^2 \\ &= \frac{1}{2} (\mu_{1_{n,k}}^2 + \mu_{2_{n,k}}^2) \end{aligned}$$

The exponential term of the error bound thus becomes

$$\rho = \frac{1}{8} \sum_{n,k} \frac{(\mu_{1_{n,k}} - \mu_{2_{n,k}})^2}{1/2(\mu_{1_{n,k}}^2 + \mu_{2_{n,k}}^2)}$$

One can notice at this stage that all the quantities in this expression are features of the periodogram of $\mathbf{z} = \boldsymbol{\theta} \mathbf{e}$ and thus depend on $\boldsymbol{\theta}$. To minimize the error bound, $\rho(\boldsymbol{\theta})$ has to be maximized. A less tight bound leads to the maximization of the more tractable quantity $\rho^*(\boldsymbol{\theta})$ such that:

$$\rho^*(\boldsymbol{\theta}) = \frac{1}{4} \sum_{n,k} \frac{(\mu_{1_{n,k}} - \mu_{2_{n,k}})^2}{(\mu_{1_{n,k}} + \mu_{2_{n,k}})^2} < \rho \quad (3)$$

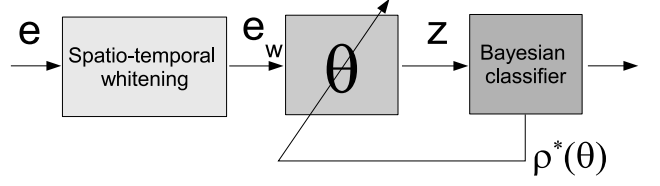


Fig. 1. Scheme of the CSDA method.

D. Optimizing the spatial filter

Minimizing the error bound relatively to the spatial filtering matrix $\boldsymbol{\theta}$ is equivalent in our case to solve the problem:

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} \rho^*(\boldsymbol{\theta}) \quad (4)$$

The variables involved in the expression of ρ^* can be expressed directly from the Spectral Density Matrices of the original EEG signals. Let $\langle \mathbf{S}^{ee}(\nu_k) \rangle_1$ and $\langle \mathbf{S}^{ee}(\nu_k) \rangle_2$ be the empirical mean spectral density matrices of the EEG signal \mathbf{e} for each class. Then, by definition, the terms at the numerators of (3) are the squared diagonal terms of

$$\langle \mathbf{S}^{zz}(\nu_k) \rangle_1 - \langle \mathbf{S}^{zz}(\nu_k) \rangle_2 = \boldsymbol{\theta} (\langle \mathbf{S}^{ee}(\nu_k) \rangle_1 - \langle \mathbf{S}^{ee}(\nu_k) \rangle_2) \boldsymbol{\theta}^T,$$

while the denominators of (3) are the squared diagonal terms of

$$\langle \mathbf{S}^{zz}(\nu_k) \rangle_1 + \langle \mathbf{S}^{zz}(\nu_k) \rangle_2 = \boldsymbol{\theta} (\langle \mathbf{S}^{ee}(\nu_k) \rangle_1 + \langle \mathbf{S}^{ee}(\nu_k) \rangle_2) \boldsymbol{\theta}^T.$$

The exact resolution of this problem has not been tackled yet. We propose here to apply a blind source separation techniques to compute an approximate solution of (4). This falls in two steps:

1) Perform a "spatio-temporal whitening" of EEG data, i.e. to transform \mathbf{e} into a signal \mathbf{e}_w such that $(\langle \mathbf{S}^{e_w e_w}(\nu_k) \rangle_1 + \langle \mathbf{S}^{e_w e_w}(\nu_k) \rangle_2)$ is close to the identity matrix for all the frequencies ν_k . In our work, we approximate the solution by successively whitening in the frequency domain and then computing a usual spatial whitening matrix (see APPENDIX for the details).

2) Once whitened, the matrices involved in the denominator of (3) are close to the identity matrix. If $\boldsymbol{\theta}$ is set to an Hermitian matrix, one can simply maximize the numerator while leaving the denominator unchanged. Under this constraint, one can maximize $\rho^*(\boldsymbol{\theta})$ with the Extended Jacobi method currently used in the Blind Source Separation community [4]. This amounts to perform the joint approximate diagonalization of the matrices:

$$A^k(\boldsymbol{\theta}) = \boldsymbol{\theta} (\langle \mathbf{S}^{e_w e_w}(\nu_k) \rangle_1 - \langle \mathbf{S}^{e_w e_w}(\nu_k) \rangle_2) \boldsymbol{\theta}^T.$$

This iterative procedure minimizes successively the non-diagonal elements of the set of matrices. Once the optimal matrix $\hat{\boldsymbol{\theta}}$ is obtained, the optimal classifier is given by (3) and can be applied on new data. The whole method is summarized in Fig. 1.

E. Comparison with CSP

CSDA has been designed to perform a good classification without a precise knowledge of the task related frequency band. To validate this, we evaluate the influence of the broadness of the chosen frequency band of analysis for CSDA and compare it to an equivalent CSP based classification method. To allow for comparison between CSDA and CSP, the same model of Bayesian classifier defined by (2) is applied to the power measurements of the band-pass filtered components produced by CSP. So, in the case of CSDA, all the values of the Power Spectral Density of each component in the considered frequency band are used for classification, whereas only one power feature by component is used for CSP-based classification.

III. DATA

A. Simulated data

In this experiment, to simulate EEG activities during two mental tasks, we used a realistic head model to compute EEG activities generated by four dipolar sources respectively located on 1)left primary motor cortex, 2)right primary motor cortex, 3)SMA and 4)left occipital area. The time courses of the sources were generated by auto regressive filtering of a white noise at a sampling rate of 500 Hz (two second-order band-pass filters simulate activities in the alpha and beta bands). As illustrated by the task-conditional PSD of each source (Fig.2), only the two first sources have task related activities in the alpha and beta bands. As the two task-dependent activities are negatively correlated (for source 1 beta increases from state 1 to state 2 while alpha is decreasing) a bad choice of the frequency band for the filtering of the data can severely impair the performance of CSP as it takes into account only the total power in the considered frequency band. Once generated, the source activities are projected on the cortex using EEG forward modeling [5] and a gaussian white noise is added on each EEG channel achieving a Signal to Noise Ratio of 20dB.

B. Real data

Methodology is also applied offline to an asynchronous BCI experiment performed in our lab on 3 subjects. The subjects were asked to perform continuously a mental task indicated on a screen during consecutive 20s epochs. Successive tasks were separated by a resting period of 3 seconds. There were six different mental tasks, including three motor imagery tasks (grasping an object with the right hand: GRASP, moving the right finger, moving the tongue) and three non-motor tasks (visuo spatial navigation, imagined music: MUSIC, calculus). These tasks appeared successively in random order on the screen. For each subject, we recorded 4 to 6 sessions of 6 min each per day during 3 days. The EEG data were recorded using a 60 electrodes BrainCap™. Data were amplified using a BrainAmps™ (Brain Products, Inc) 64 channels system sampled to 500Hz.

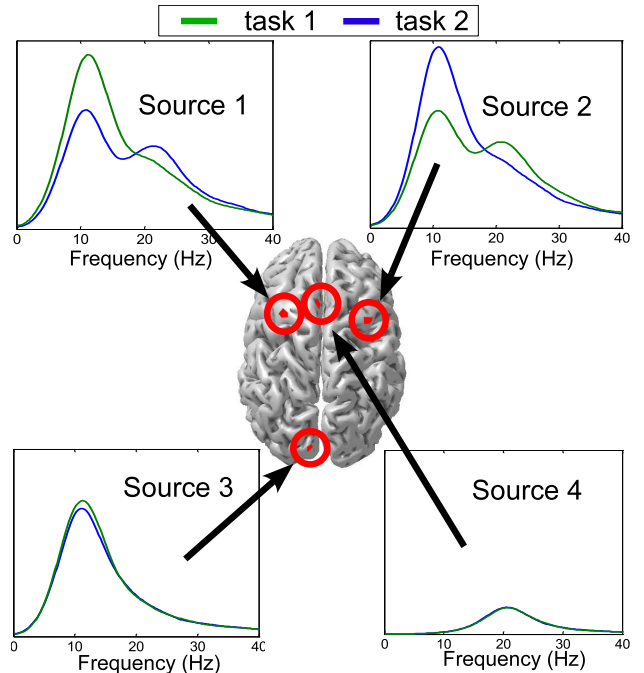


Fig. 2. Power Spectral Density of the four current dipoles simulated on a realistic cortex for the two tasks.

TABLE I
CLASSIFICATION ERROR FOR CSDA AND CSP FOR NARROW OR WIDE BAND SIGNALS (SIMULATIONS).

Frequency band	CSP	CSDA
8-12Hz	21.40% \pm 1.0	21.57% \pm 1.1
8-20Hz	34.83% \pm 1.0	6.48% \pm 0.6

IV. RESULTS

A. Simulated data

To estimate classification error of CSDA and CSP, 100 simulated 2s time-windows were generated for each task and used for CSDA and CSP learning using either a narrow (8-12Hz) or a large (8-20Hz) frequency band of analysis, then 100 other time windows were used for test. This procedure was iterated 20 times for each algorithm. The results are reported on Table 1. While performance is equivalent on narrow band, we observe significant differences on wide band signals as the growing of the band tends to improve CSDA performance whereas CSP performance is decreasing. As our minimization of the bound on Bayes error was only approximate (in particular due to the spatio-temporal whitening step), it is interesting to test if our algorithm really maximizes the objective ρ^* . The evolution of ρ^* and of the coefficient of the Bhattacharyya bound ρ during the iterations of the Jacobi Method is plotted on Fig. 3. We can observe the two quantities clearly increase during the optimization procedure. Moreover ρ is very well approximated by $2\rho^*$: this can justify that even if we don't maximize directly ρ it is nearly maximized too.

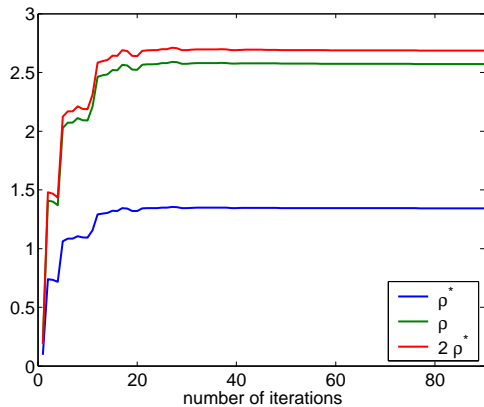


Fig. 3. Evolution of the parameters of the bounds of classification error during the iterative optimization of the extended Jacobi algorithm.

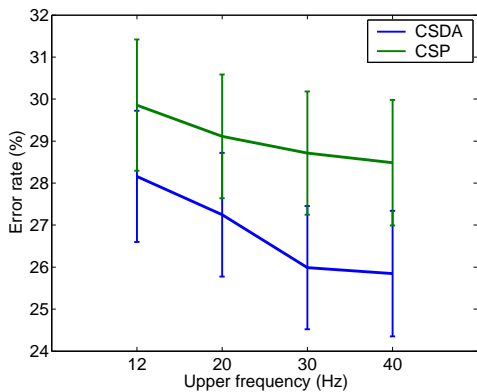


Fig. 4. Classification accuracy of CSDA and CSP for a lower frequency of 8Hz and an increasing upper frequency (error bars indicate standard errors)

B. Real data

In this paper we focus on the discrimination of the two following tasks: GRASP versus MUSIC. A complete study over all the tasks will be reported elsewhere. For classification purposes the signals were cut into successive 2s time-windows with 1s overlap. For one subject, we performed a cross validation across the experimental sessions of the same day for the three days of recordings. We tested different frequency bands of analysis, ranging from 8Hz to 12, 20, 30 or 40 Hz. The resulting average error rate is reported Fig. 4. We observed a slightly different result from simulated data: whereas the CSP error slowly decreases with a broader frequency range, CSDA error decreases more rapidly. This illustrates again the better capacity of the CSDA classifier to gather informations from a broad frequency range.

V. CONCLUSIONS AND FUTURE WORK

A. Conclusions

The results obtained in simulated and real data support the interest of CSDA especially if the discriminant frequency bands are unknown. While CSP can loose useful information, CSDA gathers more discriminant informations as the

frequency range increases. Moreover, this work is putting forward that particular model based classifiers can be elaborated that can benefit from algorithm such as the Jacobi method, usually applied in the blind source separation community. As the optimization is a joint diagonalization process of Spectral Density Matrices, one can interpret CSDA as spatial filter that concentrates the discriminant informations initially contained in cross-spectral terms into the diagonal avoiding the classifier to manage directly the high dimensional space of the whole of the coefficients of SDMs. This example shows that classification and blind source separation can possibly be linked in a new procedure.

B. Future Works

The spatio-temporal whitening step is the weak point of this method as it is currently done only in a suboptimal way. More work has to be done to find an optimal solution to this problem. Moreover, as the frequency range increases, this model based method has a growing number of parameters to estimate and thus can lead to overfitting. We will exploit Recursive Feature Elimination (RFE) methods to automatically select relevant frequency bands and improve classifier accuracy.

VI. APPENDIX

The spatio-temporal whitening is performed in two steps:

Step 1: Temporal whitening

Temporal whitening amounts to turn the signal of each EEG channel into a white noise leading to the signal e_t , i.e. a process with constant spectrum (equal to 1 in our case). This whitening is done by filtering separately each channel in the frequency domain and leads to the new coefficients of the Spectral Density Matrices:

$$(\mathbf{S}_i^{e_t e_t}(\nu_k))_{n,m} = \frac{(\mathbf{S}_i^{zz}(\nu_k))_{n,m}}{\sqrt{(\mathbf{S}_i^{zz}(\nu_k))_{n,n} (\mathbf{S}_i^{zz}(\nu_k))_{m,m}}}$$

Step 2: Spatial whitening

Spatial whitening is then applied on the signal e_t resulting from the previous step using a whitening matrix \mathbf{W} . This matrix is computed from the eigenvalue decomposition of the covariance matrix of e_t : $\mathbf{R}_{e_t e_t} = \mathbf{V} \mathbf{D} \mathbf{V}^T$. The choice of

$$\mathbf{W} = \mathbf{R}_{e_t e_t}^{-1/2} = \mathbf{V} \mathbf{D}^{-1/2} \mathbf{V}^T$$

yields a signal $e_w = \mathbf{W} e_t$ with uncorrelated components of unit variance.

REFERENCES

- [1] J. Muller-Gerking, G. Pfurtscheller, and H. Flyvbjerg, "Designing optimal filters for single-trial EEG classification in a movement task," *Clin. Neurophysiol.*, vol. 110, pp. 787–798, 1999.
- [2] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. Wiley-Interscience, 2001.
- [3] M. Rosenblatt, *Stationnary sequences and random fields*. Birkhauser, 1985.
- [4] J.-F. Cardoso and A. Souloumiac, "Jacobi angles for simultaneous diagonalization," *SIAM Journal on Matrix Analysis and Applications*, vol. 17, no. 1, pp. 161–164, 1996.
- [5] J. J. Ermer, J. Mosher, S. Baillet, and R. Leahy, "Rapidly recomputable EEG forward models for realistic head shapes," *Phys Med Biol*, vol. 46, no. 4, pp. 1265–1281, Apr 2001.

Chapitre 7

Analyse au niveau des sources corticales.

Au chapitre 5 nous avons étudié la classification des tâches mentales de notre ICM en nous basant sur la quantification des activités oscillatoires issues des capteurs EEG. Ces activités étant un reflet atténué et mélangé des activités des macro-colonnes corticales (voir chapitre 3), on peut envisager la reconstruction de celles-ci afin d'améliorer le transfert d'information du système global. Ainsi, dans ce chapitre, nous étudions la possibilité d'inclure l'estimation du problème inverse dans la chaîne de traitement du dispositif d'ICM. Bien que les méthodes implémentées, en aval du problème inverse dans cette chaîne, soient préservées dans leur globalité, certains ajustements des méthodes de quantification et de classification ont été réalisés pour prendre en compte la très haute dimensionnalité de l'espace de quantification des sources corticales.

7.1 Méthodes

7.1.1 Reconstruction de l'activité corticale

Problème direct

La modélisation du potentiel électrique généré au niveau des capteurs en fonction de l'activité des dipôles de courants situés en tous points du cortex, aussi connue sous le nom de problème direct, est ici réalisée à partir d'un modèle réaliste de l'anatomie corticale. Celui-ci est basé sur deux enveloppes : un maillage de la surface de la tête et un maillage de la surface du cortex de chaque sujet¹ qui sont représentés tous les deux figure 7.1 pour un sujet. Ces maillages sont extraits de l'IRM anatomique de chaque sujet à l'aide du logiciel Brainvisa. En chaque sommet du maillage du cortex est disposé un dipôle de courant virtuel orienté perpendiculairement à la surface (Dale and Sereno, 1993); ce maillage est décimé de façon à contenir 10000 dipôles². À partir de

1. en réalité la surface utilisée correspond à l'interface substance grise/substance blanche du cortex

2. cette opération est réalisée avec la boîte à outils Brainstorm

ces informations, le modèle direct le plus élémentaire est un modèle de tête sphérique pour lequel la distribution de potentiel due à chaque source corticale distribuée au sein du cortex peut être connue analytiquement. Cependant, la forme de la tête (qui est loin d'être sphérique pour la plupart d'entre nous) est une information importante qui modifie grandement la répartition du potentiel à la surface. Afin d'exploiter un modèle plus réaliste, nous avons utilisé une méthode à sphères multiples, implémentée dans Brainstorm, qui ajuste à chaque électrode une sphère tangente à la surface de la tête pour calculer plus précisément le potentiel engendré sur cette électrode par un dipôle dans son voisinage. Le résultat est stocké dans la matrice de sensibilité des capteurs $\mathbf{\Lambda}$ (voir chapitre 3).

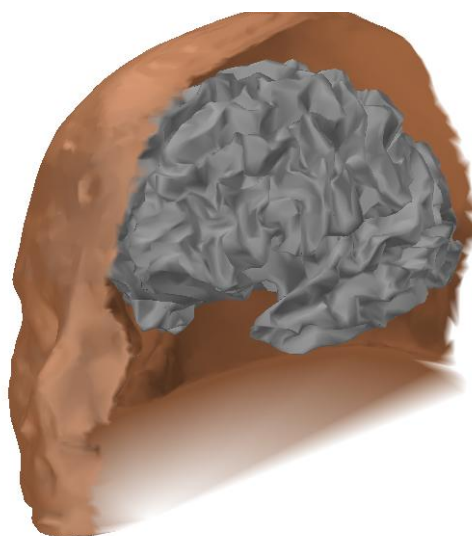


FIGURE 7.1 – Enveloppes générées à partir de l'IRM anatomique du sujet, utilisées pour le calcul du problème direct. La surface externe correspond à la surface de la tête et la surface interne est l'interface entre la matière grise et la matière blanche du cortex.

Problème inverse

Afin de minimiser le temps de calcul nécessaire pour l'estimation des sources corticales et aussi de prendre en compte la complexité des signaux en essai unique, nous avons choisi d'utiliser le problème inverse avec contrainte de norme minimale défini au chapitre 3. On rappelle que $\mathbf{\Lambda}$ est la matrice de champ de sensibilité issue du problème direct, liant l'activité de N sources corticales au potentiel mesuré sur n électrodes. L'estimation $\hat{\mathbf{q}}$ de l'activité des sources est déduite des signaux mesurés \mathbf{m} sur les électrodes à l'aide de la formule matricielle :

$$\hat{\mathbf{q}} = (\mathbf{\Lambda}^T \mathbf{\Lambda} + \alpha \mathbb{I}_N)^{-1} \mathbf{\Lambda}^T \mathbf{m}$$

Le coefficient de régularisation α est choisi grâce à une heuristique : $\alpha = .1 \lambda_1$ ou λ_1 est la plus grande valeur propre de la matrice $\Lambda\Lambda^T$. Compte tenu du faible nombre de lignes (environ 60) de la matrice Λ en comparaison au nombre de colonnes (environ 10000), la matrice de problème inverse peut être exprimée d'une manière qui facilite grandement son calcul (Dale and Sereno, 1993) :

$$\hat{\mathbf{q}} = \Lambda^T(\Lambda\Lambda^T + \alpha\mathbb{I}_n)^{-1}\mathbf{m} = \mathbf{W}\mathbf{m}$$

ce qui permet d'avoir à inverser une matrice (60×60) à la place d'une matrice (10000×10000). L'ensemble des mesures de quantification de l'activité des sources corticales pour un sujet est réalisé de manière rapide à partir de la matrice de problème inverse \mathbf{W} .

7.1.2 Quantification

De manière similaire à l'analyse au niveau des capteurs, l'activité corticale est quantifiée par une mesure de synchronie locale, la puissance spectrale dans chaque bande de fréquence, et par une mesure d'interaction à distance, la cohérence³. Ces deux quantités sont évaluées par la méthode de Welch (voir chapitre 5) et peuvent être calculées à partir de la matrice de densité spectrale des électrodes $\Gamma_{\mathcal{W}}(f_k)$ (estimée elle aussi avec la méthode de Welch). Ainsi, par application de la matrice de problème inverse, l'expression de la puissance spectrale dans la bande $[f_l, f_h]$ sur la source n est :

$$P_W(n, [f_l, f_h]) = \sum_{f_k \in [f_l, f_h]} \mathbf{W}_n \Gamma_{\mathcal{W}}(f_k) \mathbf{W}_n^T$$

où \mathbf{W}_n est la n-ième ligne de la matrice de problème inverse (correspondant à la n-ième source corticale). De même la cohérence entre la source m et la source n est exprimée par :

$$C_{nm}([f_l, f_h]) = \frac{|\sum_{f_k \in [f_l, f_h]} \mathbf{W}_n \Gamma_{\mathcal{W}}(f_k) \mathbf{W}_m^T|}{\sqrt{|\sum_{f_k \in [f_l, f_h]} \mathbf{W}_n \Gamma_{\mathcal{W}}(f_k) \mathbf{W}_n^T| |\sum_{f_k \in [f_l, f_h]} \mathbf{W}_m \Gamma_{\mathcal{W}}(f_k) \mathbf{W}_m^T|}}$$

Cependant la cohérence ne peut être calculée entre chaque couple de sources (il y aurait environ 50 millions de couples à considérer), un certain nombre de sources de référence ont donc été présélectionnées. Pour cela la source la plus discriminante dans chaque bande de fréquence pour chaque couple de tâche possible a été choisie à l'aide du score de Fisher. Un ensemble de 15 sources de référence par bande de fréquence est ainsi obtenu. Afin de réduire le nombre de sources de référence à considérer (pour minimiser le nombre de variables de quantification), un algorithme d'élimination récursif est implémenté de la façon suivante :

3. qui est avantageuse par rapport à la synchronie de phase en terme de temps de calcul et en terme de taux de classification, voir chapitre 5

1. Dans chaque bande de fréquence, on choisit sur le cortex les deux sources n et m les plus proches selon la distance euclidienne (parmi les 15 présélectionnées). Ces sources ont les puissances les plus discriminantes pour le couple de classes i et le couple de classes j respectivement.
2. soit $F_i(n)$ la valeur du score de Fisher de la source n pour le couple de classe i . On élimine n de l'ensemble des sources de référence si $F_i(m) > .6F_i(n)$. C'est-à-dire si le remplacement de la source n par la source m entraîne une perte de pouvoir discriminant inférieure à 40%⁴ pour la classe i . A l'inverse, si l'inégalité est valable en permutant n , m , i et j , alors la source m est éliminée. La valeur de 40% a été sélectionnée empiriquement de façon à ne pas saturer la mémoire vive de l'ordinateur pendant le calcul de la quantification : elle permet d'obtenir un nombre maximum de 6 sources de références.
3. l'opération 2 est répétée sur les couples de sources les plus proches qui n'ont pas été choisis auparavant, jusqu'à convergence.

Cette élimination récursive permet de sélectionner dans chaque bande de fréquence 3 à 7 sources de référence, dont la puissance est discriminante pour l'ensemble des 15 couples de tâches. Les mesures de cohérence dans une bande sont calculées sur l'ensemble des couples de sources corticales qui contiennent une des sources de référence. Une illustration du résultat de la méthode de sélection de sources de référence est donnée figure 7.2. On remarque dans ce cas que quatre sources ont été sélectionnées, réparties de manière relativement homogène sur le cortex.

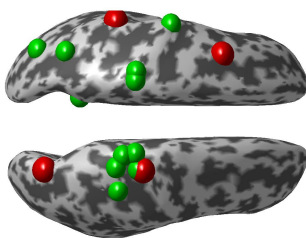


FIGURE 7.2 – Sources de référence calculées dans la bande à partir des maxima du score de Fisher pour chaque couple de classes. À partir de 15 sources de référence, une procédure d'élimination permet de sélectionner les quatre sources qui apparaissent en rouge. Les sources non sélectionnées apparaissent en vert.

7.1.3 Classification

Concernant l'application temps réel de ce système, l'utilisation d'une quantification sur l'ensemble des sources corticales pose des problèmes. En effet, le calcul de l'activité

⁴ cette valeur de 40% a été choisie de façon à obtenir un nombre suffisamment restreint de sources de référence

corticale de chacune des 10000 sources nécessite un temps supérieur au délai requis par le temps réel. Il est donc intéressant de sélectionner des variables afin de réduire le nombre d'activités corticales à calculer. L'outil de base reste dans ce cadre le SVM, qui est ici très avantageux car son temps de calcul ne dépend pas du nombre de variables et est moins sensible à la haute dimension de l'espace de quantification. Une modification du SVM adaptée à la sélection de variables est donc utilisée et a fait l'objet d'un article au colloque GRETSI 2007 (Besserve et al., 2007a). Cependant, les résultats du chapitre 5 montrent que la sélection de variables n'est pas toujours avantageuse pour la performance du classifieur, en particulier lorsqu'il s'agit de prendre en compte les mesures d'interaction à distance. Pour vérifier cette propriété, les méthodes de classification par SVM sans sélection de variables ou en utilisant le score de Fisher seront aussi envisagées, de façon similaire au chapitre 5.

Algorithme L2-AROM

Afin de réduire le nombre de variables, plusieurs méthodes de sélection de variables peuvent être envisagées (Guyon and Elisseeff, 2003). Ce type d'algorithme calcule à partir de la base d'apprentissage un sous-ensemble de variables, optimal pour la classification. Nous utilisons ici une technique de sélection de variables multivariée qui réalise simultanément le calcul du classifieur optimal.

Le classifieur utilisé est un Séparateur à Vaste Marge (SVM) linéaire (Vapnik, 1998) pour lequel on remplace la norme 2 par une norme 0 dans la fonction objectif. La méthode détermine donc un hyperplan séparateur des données d'apprentissage d'équation $\langle \boldsymbol{\omega}^*, \mathbf{x} \rangle + b^* = 0$, comme solution du problème suivant (Weston et al., 2003) :

$$(\boldsymbol{\omega}^*, b^*) = \arg \min_{\boldsymbol{\omega}, b} \|\boldsymbol{\omega}\|_0$$

sous la contrainte

$$\forall i, y_i(\langle \boldsymbol{\omega}, \mathbf{x}_i \rangle + b) \geq 1$$

La norme 0 utilisée est égale au nombre de coefficients non nuls dans le vecteur $\boldsymbol{\omega}$. On recherche donc à calculer une séparatrice affine des données dont l'équation possède un nombre de variables minimales. Une telle méthode de sélection de variables présente l'intérêt de ne pas sélectionner les variables redondantes comme pourrait le faire une méthode univariée telle qu'un test T de Fisher. Ce problème est complexe à résoudre de façon exacte. Cependant, l'algorithme récursif L2-AROM présenté ci-dessous permet de le résoudre de façon approchée (Weston et al., 2003) :

1. On pose $\mathbf{z} = [1, 1, \dots, 1]^T$
2. Résoudre

$$\hat{\boldsymbol{\alpha}} = \arg \max_{\boldsymbol{\alpha}} \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \langle \mathbf{z} * \mathbf{x}_i, \mathbf{z} * \mathbf{x}_j \rangle$$

sous les contraintes

$$\sum \alpha_i y_i = 0 \quad \text{et} \quad \forall i, \alpha_i \geq 0$$

3. Soit $\hat{\omega} = \sum y_i \hat{\alpha}_i(\mathbf{x}_i * \mathbf{z})$ issu de la solution de 2. On pose $\mathbf{z} \leftarrow |\mathbf{z} * \hat{\omega}|$
4. Revenir à 2 jusqu'à convergence de \mathbf{z} . La solution est alors $\omega^* = \hat{\omega}$

Cet algorithme revient donc à implémenter récursivement un Séparateur à Vaste Marge (SVM) classique (Vapnik, 1998) sur les données $\mathbf{z} * \mathbf{x}$. Le vecteur ω^* ainsi obtenu contient seulement un très faible nombre de coefficients non nuls, qui correspondent aux variables sélectionnées. Cette technique a la particularité exceptionnelle d'avoir un temps de calcul qui ne dépend pas du nombre total de variables car le problème est formulé dans un espace dual. Cette caractéristique est importante, car le problème inverse utilisé génère un nombre de variables très important.

7.2 Résultats : Classification à partir de la puissance

7.2.1 Sélection de variables sur les sources corticales

Dans de premiers résultats (Besserve et al., 2007a), l'algorithme L2-AROM a été appliqué aux données de 3 sujets quantifiées par la puissance des sources corticales dans quatre bandes de fréquence (10-12,15-20,20-30 et 30-40Hz), calculée sur des fenêtres glissantes de 2s.

Pour comparer le taux de classification (évalué en validation croisée inter-jours) obtenu sur les sources avec celui obtenu sur les électrodes, le même classifieur (L2-AROM) a été appliqué en utilisant directement comme quantification les puissances spectrales des électrodes EEG dans les mêmes bandes de fréquences. On représente table. 7.1 le taux de bien classés moyenné à travers les 15 couples de discrimination binaires possibles entre les 6 tâches pour les deux approches. Le taux moyen de bonne classification de la méthode basée sur les sources corticales atteint 75.01% pour le meilleur des trois sujets et apparaît significativement supérieur à la performance obtenue par la méthode basée sur les électrodes EEG ($p < .05$) par un test non paramétrique réalisé sur l'ensemble des trois sujets. Les résultats sont encourageants compte tenu qui s'agit de résultats

TABLE 7.1 – Taux de bonne classification de l'algorithme L0-SVM à partir des activités mesurées sur les électrodes et à partir des estimations de l'activité corticale.

Mesures	sources	électrodes
sujet 1	65.91% \pm 1.2	63.12% \pm 1.5
sujet 2	60.67% \pm 1.3	61.84% \pm 1.0
sujet 3	75.01% \pm 2.0	72.76% \pm 1.9

de prédiction d'un jour à l'autre sur des fenêtres temporelles de 2s. Dans la figure 7.3 nous illustrons les propriétés de convergence de l'algorithme en traçant l'évolution du vecteur \mathbf{z} associé à chaque variable de quantification. On observe au fil des itérations

qu'un nombre de variables de plus en plus faible se voit attribuer des coefficients importants relativement aux autres. Une vingtaine de variables sont finalement sélectionnées en 34 itérations, ce qui illustre la convergence rapide de l'algorithme malgré le grand nombre de variables initial. La répartition topographique des variables choisies est illus-

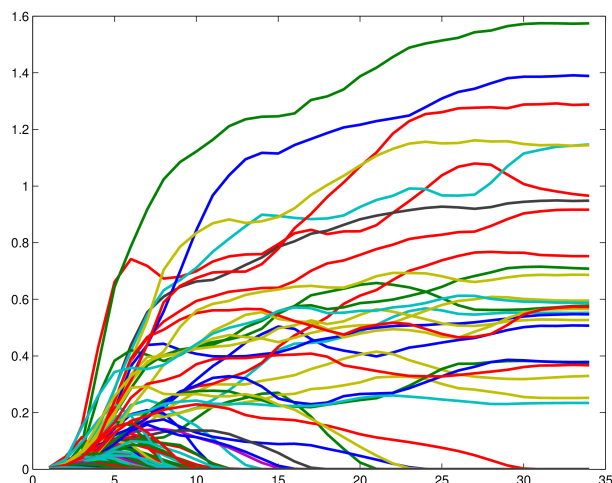


FIGURE 7.3 – Exemple d'évolution de la valeur absolue des coefficients du vecteur \mathbf{z} correspondant à chaque variable en fonction des itérations successives de l'algorithme L2-AROM. Seules 25 variables parmi 40000 atteignent une valeur significative lorsque l'algorithme a convergé et sont donc sélectionnées.

trée Fig. 7.4 pour la discrimination entre les tâches "index droit" et "visuo-spatial". Les activations significativement différentes (mesurées par un test T de Student univarié sur chaque source, corrigé pour comparaison multiples par test de permutation) entre les deux tâches dans la bande 15-20 Hz sont représentées sur un cortex gonflé où les régions sombres représentent le fond des sillons corticaux. On observe une activation moins importante du cortex frontal gauche, controlatéral au mouvement, lorsque le sujet imagine bouger l'index droit. La navigation visuo-spatiale diminue quant à elle la puissance des régions occipitales et pariétales droites. Ces localisations correspondent bien à des régions impliquées dans les tâches demandées et confirment que le sujet les réalise correctement. L'algorithme de sélection de variables a sélectionné une variable dans chacune de ces régions ainsi qu'une variable supplémentaire n'apparaissant pas significative par test univarié. L'algorithme semble se comporter ici de manière optimale en sélectionnant des variables non redondantes pour extraire le maximum d'information des activités corticales en vue de la classification.

7.2.2 Résultats en validation croisée inter-sessions

Influence de la sélection de variables

Les résultats de la section précédente (présentés au colloque GRETSI 2007) étaient donnés pour une validation croisée à travers les jours d'enregistrement. Pour comparer

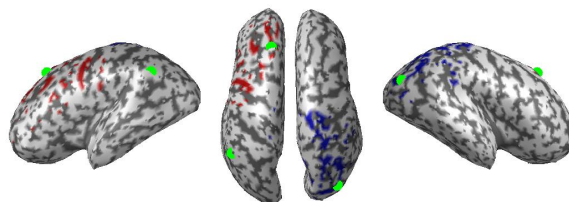


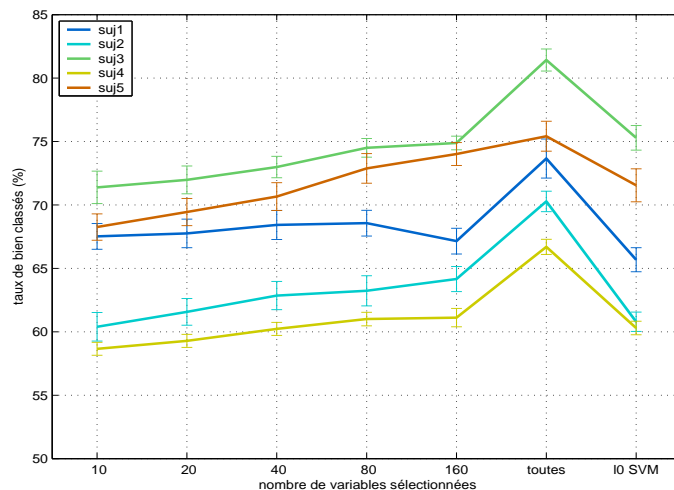
FIGURE 7.4 – Topographie sur un cortex gonflé des activations corticales présentant des différences significatives dans la bande 15-20Hz entre les tâches "index droit" et "navigation visuo-spatiale" pour un sujet (rouge : "index droit" < "navigation visuo-spatiale", bleu pour le contraire). Les points verts correspondent aux variables sélectionnées par l'algorithme L2-AROM

nos résultats avec ceux du chapitre 5 et tester les performances à court terme de la classification sur les sources, nous présentons ici des résultats en validation croisée à travers les sessions du même jour d'enregistrement, avec des quantifications réalisées dans 5 bandes de fréquences (voir chapitre 5) moyennés sur les trois jours et sur l'ensemble des couples de tâches. La performance en validation croisée de la quantification par la puissance spectrale des sources est présentée pour chaque sujet sur la figure 7.5(a). Elle montre que la norme zéro donne une sélection de variable équivalente à la sélection de 80 sources bien qu'il n'en sélectionne qu'environ 30. Le maximum de performance est clairement atteint en utilisant pas de sélection de variables. Le score obtenu dépasse de 5% celui atteint par la quantification par la puissance sur les électrodes (voir chapitre 5).

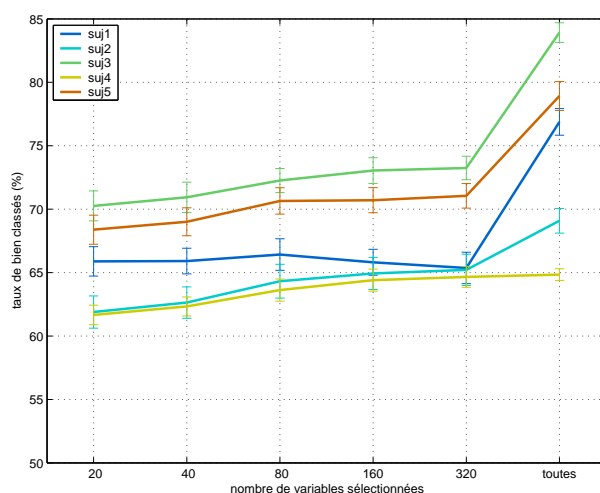
Analyse des variables de puissance discriminantes

Afin de connaître les régions impliquées dans la classification des différentes tâches, nous avons appliqué à l'ensemble des variables de puissance pour chaque sujet et chaque jour d'expérimentation un test de permutation basé sur la statistique du maximum du test T de Student. Celui-ci indique pour chaque bande de fréquence si la puissance de chaque source dans cette bande est significativement différente entre deux situations expérimentales (qui correspondent ici à des tâches mentales), à partir d'un seuillage de la statistique (Pantazis et al., 2003). Afin de caractériser des variables associées à chaque tâche plutôt qu'à chaque couple de tâches, nous avons cumulé le résultat de ce test à travers les jours et les couples impliquant une même tâche. Ainsi, soit $t_{ab}(n, j)$ la valeur de la statistique du test t calculé sur la n-ième variable le jour j entre les tâches a et b et $T_{a,b,j}$ le seuil calculé par les permutations, alors un indice du pouvoir discriminant de la classe a, par rapport aux autres classes, est calculé par la formule

$$I(n) = \sum_j \sum_b \mathbb{1}_{t_{ab}(n,j) > T_{a,b,j}} - \mathbb{1}_{t_{ab}(n,j) < -T_{a,b,j}}$$



(a) Classification à partir de la puissance spectrale sur les sources corticales



(b) Classification à partir de la puissance spectrale et de la cohérence sur les sources corticales

FIGURE 7.5 – Taux de bien classés en fonction du nombre de variables sélectionnées pour la quantification par mesures de *puissance* ou *puissance et cohérence* sur les sources corticales. Lorsque toutes les variables sont utilisées par le classifieur, cela correspond à environ 50000 variables pour la puissance spectrale (10000 sources \times 5 bandes de fréquences) auxquelles s'ajoutent environ 250000 variables dans le cas où les cohérences sont aussi utilisées (environ 5 sources de référence \times 10000 sources corticales \times 5 bandes de fréquences). Ces résultats sont à comparer à ceux obtenus avec une quantification équivalente sur les électrodes (figures 5.4(a) et 5.4(d)).

Ainsi $I(n)$ est très positif lorsqu'il existe plusieurs jours d'expérience pour lesquelles cette variable a une moyenne significativement supérieure dans la classe a, par rapport aux autres classes. Inversement, si la moyenne conditionnelle à la classe a est plus souvent inférieure (significativement) aux autres classes, alors l'indice $I(n)$ est négatif. Le tracé de l'indice $I(n)$ pour chaque classe sur l'ensemble des variables de puissance est donné pour tous les sujets sur les figures 7.6 à 7.10. Concernant les tâches motrices, la préhension fait apparaître des baisses d'activités significatives dans les régions centrales chez la plupart des sujets (sujets 1, 2, 3 et 5). Cette baisse d'activité est latéralisée du côté gauche (controlatéral au mouvement) chez les sujets 1, 2 et 3, mais seulement pour les bandes de fréquence les plus hautes (18 à 40 Hz), pour les sujets 1 et 3. Lorsque cette latéralisation est présente, elle implique des régions antérieures au sillon central, mais on observe cependant chez les quatre sujets concernés une baisse d'activité dans les régions pariétales supérieures, en particulier dans les fréquences basses (4-16Hz). En comparaison, la tâche d'imagerie de l'index est associée à des baisses d'activité centrale chez les mêmes sujets, à la différence que la latéralisation est plus claire et que les régions pariétales sont beaucoup moins impliquées. En particulier, pour les sujets 3 et 5, les régions pariétales sont au contraire associées à une augmentation d'activité par rapport aux autres classes. L'imagerie de la langue, quant à elle, présente des résultats en opposition aux deux autres tâches motrices : les régions centrales sont en moyenne plus actives que dans les autres classes chez tous les sujets. En particulier, pour les sujets 3, 4 et 5, un maximum apparaît dans les deux plus hautes bandes de fréquences, très localisé sur la région dorsale centrale (à la jonction entre les deux hémisphères). La hausse d'activité est cependant largement distribuée en particulier chez les sujets 2, 3 et 5, où les régions pariétales et temporales sont activées. Parmi les tâches non-motrices, le calcul est associé à des variables discriminantes reproductibles à travers les sujets :

1. une hausse d'activité par rapport aux autres classes dans les régions centrales est visible chez les sujets 1, 2, 3 et 5.
2. une baisse d'activité dans les régions temporales et le lobe insulaire est observée chez tous les sujets, en particulier à droite, bien que cette baisse ne se produise pas systématiquement dans les mêmes bandes de fréquences pour des sujets différents (chez le sujet 5 cette baisse d'activité dans les bandes 4-10Hz et 16-21Hz est partiellement masquée par la région de hausse d'activité centrale qui est très étendue).

Dans le cas de l'imagerie visuo-spatiale, chez trois sujets (sujets 1, 3, et 5) on observe une baisse d'activité dans les régions pariétales tandis qu'une augmentation d'activité apparaît dans les régions centrales (on remarque que pour le sujet 5 ces deux phénomènes se recouvrent). Un comportement très différent est observé chez les deux autres sujets : une hausse de la puissance dans les régions occipitales autour de 10Hz. Finalement, la tâche d'imagerie auditive fait apparaître des activations semblables à l'imagerie visuo-spatiale : une augmentation dans les régions centrales visible chez quatre sujets dans la bande 20-30Hz, et une diminution focalisée dans la région pariétale dorsale, en particulier autour de 10Hz pour les sujets 1, 2 et 5. À cela s'ajoute chez les sujets 1, 2 et 5 une baisse d'activité dans les régions temporales. Contrairement à l'imagerie visuo-spatiale,

dans le cas de l'imagerie auditive c'est la hausse d'activité des régions centrales qui est le phénomène le plus étendu spatialement.

Fonction discriminante du classifieur

Au-delà de l'analyse statistique du pouvoir discriminant de chaque variable, il est intéressant de savoir quelles variables le classifieur utilise pour réaliser la discrimination. Les données étant normalisées, l'amplitude des coefficients du vecteur ω de la fonction de décision du SVM donne un indice de l'importance accordée à chaque variable pour la discrimination. La distribution des coefficients de la fonction de décision est représentée figure 7.12 pour les couples de tâches présentant le meilleur taux de bien classés (index versus calcul) et pour les trois sujets obtenant les meilleurs résultats. Il apparaît clairement que la région du cortex moteur primaire controlatéral est associée à des coefficients élevés du classifieur, ce qui est cohérent avec la littérature sur les rythmes sensorimoteurs.

7.3 Résultats : Classification à partir de la cohérence

7.3.1 Taux de classification

Le taux de classification moyenné à travers toutes les discriminations possibles, à partir d'une quantification sur la cohérence combinée aux mesures de puissance est représenté figure 7.5(b). L'utilisation de la cohérence en plus de la puissance sur les sources permet encore une amélioration des résultats par rapport à une quantification uniquement sur la puissance chez les trois meilleurs sujets (voir figure 7.5(a)). Le taux de bien classés atteint par le meilleur sujet est de 84% en moyenne sur l'ensemble des couples de tâches, ce qui est très satisfaisant.

Pour comparer l'ensemble des méthodes de quantification, les résultats de classification moyennés pour l'ensemble des sujets sont représentés figure 7.3.1 pour chaque couple de tâches. On observe une augmentation graduelle de la performance en utilisant des quantifications de plus en plus riches.

7.3.2 Variables discriminantes

Les indices $I(n)$ associés aux mesures de cohérence sont représentés figures 7.14 à 7.16 pour chaque classe, pour les trois sujets ayant les meilleurs résultats de classification et pour différentes bandes de fréquences. Dans ces figures sont affichées seulement les cohérences mesurées par rapport à la source de référence donnant la valeur de T la plus élevée dans la bande considérée. L'analyse générale de la position de ces sources de référence (représentée par une sphère rouge) montre que la plupart sont situées dans les régions centrales et pariétales. Toujours de façon générale, la zone où la cohérence est discriminante est répartie autour de la source de référence avec une extension spatiale

variable. Si l'on compare la position et le signe des zones discriminantes avec les distributions obtenues pour la puissance, les résultats sont assez proches en terme de signe et d'amplitude, mais l'extension spatiale peut être réduite comme dans le cas du sujet 1 dans la bande 11-18Hz pour les classes index, préhension et visuo-spatial. La source de référence est souvent positionnée en un point où un maximum de pouvoir discriminant pour la puissance est atteint et le signe de l'indice I est en général constant pour la classe et la bande considérée. Cependant, on observe dans certains cas des configurations plus complexes. Par exemple, pour la tâche de calcul du sujet 5 dans la bande 16-21Hz, alors que les variables discriminantes en puissance dans la région pariétale reflètent une augmentation, la distribution de la cohérence dans cette région par rapport à la source de référence contient deux zones : une zone centrale où la cohérence augmente, et une zone pariétale droite où la cohérence diminue. Finalement, l'ensemble des mesures de cohérence montre des valeurs de significativité (quantifiées par l'amplitude de l'indice I) comparables à celles de la puissance.

7.4 Discussion

7.4.1 Résultats de classification

L'ensemble des résultats obtenus dans ce chapitre et le chapitre 5 met en avant le fait que la sélection de variables nuit au taux de classification des tâches mentales. De plus, l'utilisation de quantifications dans des espaces de très haute dimension ne semble pas perturber la méthode de classification : au contraire, l'apport des mesures de synchronisation à distance, puis de la quantification sur les sources corticales contribue à améliorer les résultats. L'apport des mesures de cohérences sur les sources corticales n'est bénéfique que pour les trois meilleurs sujets. Cependant les sources de référence pour la cohérence ont été calculées en fonction de leur pouvoir discriminant en puissance. Il est donc normal que l'apport des mesures de cohérence par rapport à la puissance soit limité. L'utilisation de méthodes de sélection de variables et en particulier de la méthode L2-AROM peut cependant permettre de réaliser un compromis entre l'amélioration du taux de classification et le coût en temps de calcul de la quantification sur un grand nombre de variables.

7.4.2 Variables discriminantes

Puissance

Dans le cas des tâches motrices, la baisse d'activité observée dans les bandes mu et bêta est cohérente avec les résultats de la littérature sur la désynchronisation liée à un événement ayant lieu au sein des régions motrices et pré-motrices. Le fait que le mouvement d'index imaginé apparaisse plus latéralisé que la préhension avec le bras est aussi cohérente avec les localisations des aires motrices du bras et de la main, l'aire de la main étant plus latérale. La localisation pariétale de la baisse d'activité dans la bande

alpha dans le cas de la préhension pourrait refléter un traitement par le lobule pariétal supérieur pour gérer la coordination du mouvement de préhension. Réciproquement, l'augmentation d'activité dans la région pariétale chez certains sujets pour l'imagerie de l'index peut traduire le fait que cette région réalise moins de traitements que dans les autres tâches. Concernant l'imagination du mouvement de la langue, deux explications sont possibles pour les distributions observées. Tout d'abord, la méthode utilisée pour mettre en évidence les variables discriminantes de chaque classe montre le fait que les régions du cortex moteur associées au bras et à l'index ne subissent pas de baisse d'activité durant les mouvements imaginés de la langue, ainsi ces régions apparaissent comme ayant en moyenne une activité supérieure pendant l'imagination de la langue par rapport aux autres classes. Une autre interprétation possible concerne un aspect physiologique du fonctionnement du cortex moteur : si l'imagerie de la langue active la région motrice qui lui est associée, alors les autres régions du cortex moteur sont inhibées, ce qui se traduit effectivement par des hausses d'activités mu et bêta dans ces régions.

Les corrélats EEG des tâches non motrices ont quant à eux été peu étudiés dans la littérature. Cependant, il est possible de dresser quelques interprétations des topographies observées. Tout d'abord, la désynchronisation dans les régions temporales et pariétales inférieures pendant le calcul pourrait impliquer les aires associatives visuelles (visualisation mentale des chiffres) et du langage (aires du langage de Broca et de Wernicke). Bien que l'on puisse s'attendre dans ce cas à une baisse d'activité plus prononcée dans l'hémisphère gauche (où sont localisées les aires du langage pour la majorité de la population), il est possible que les oscillations dues à l'inhibition de cortex moteur gauche lors de l'absence d'imagerie motrice masquent une baisse de puissance qui résulterait de l'activation des aires du langage. En effet, chez les sujets présentant une forte baisse d'activité pendant les tâches motrices d'index et de préhension, une hausse d'activité apparaît clairement sur le cortex moteur controlatéral au mouvement lors de la tâche de calcul, qui n'active aucune région motrice. Dans le cas de l'imagerie visuo-spatiale, la forte activation du cortex pariétal semble refléter fidèlement le traitement des informations spatiales tandis que les régions motrices présentent une hausse d'activité dans les bandes mu et bêta qui reflète leur inhibition dans cette tâche. La forte baisse de puissance occipitale dans la bande alpha chez deux sujets est difficile à interpréter, elle pourrait cependant refléter l'inhibition des aires visuelles primaires permettant de bloquer l'arrivée des stimuli visuels externes pour favoriser le traitement de scènes visuelles mémorisées par les régions visuelles associatives. Finalement, l'imagerie auditive présente principalement une hausse de puissance dans les régions centrales, incluant le cortex moteur mais aussi une partie du lobe pariétal. Cela peut être associé au faible contenu moteur et visuo-spatial dans cette imagerie par rapport aux autres tâches. Cependant la baisse d'activité dans les régions temporales, en particulier dans l'hémisphère droit chez certains sujets peut refléter le traitement des sons musicaux par les aires auditives associatives. Ce phénomène est cependant de faible ampleur.

Pour résumer, les variables de puissance des sources corticales mettent en évidence des contrastes entre les tâches principalement au niveau des régions motrices, pré-

motrices et pariétales. Il y a une régularité de la distribution de variables à travers les bandes de fréquence, impliquant les régions pariétales principalement de 4 à 30Hz et les régions motrices et pré-motrices de 10 à 40Hz. Un lissage moins important du spectre fréquentiel pourrait être envisagé pour rechercher des variations plus fines de la puissance spectrale, cependant cette augmentation de résolution fréquentielle serait au détriment du rapport signal à bruit et l'amélioration du pouvoir discriminant n'est donc pas garantie. Par ailleurs, il est possible, compte tenu des résultats de classification obtenus, de s'interroger sur l'apport des mesures de cohérence par rapport aux mesures de puissance et en particulier d'analyser s'ils mettent en évidence des phénomènes différents.

Cohérence

D'après la comparaison entre les topographies pour la cohérence et pour la puissance, la cohérence fournit à première vue une information qui conforte celle des puissances des sources corticales. Ceci est raisonnable compte tenu de la capacité des assemblées de neurones corticaux à se synchroniser localement et du fait que le problème inverse produit des solutions qui sont corrélées localement. Ainsi, plus une source est puissante, plus son signal est diffusé sur une large région corticale. Cependant, une analyse plus fine montre que l'information contenue dans la cohérence n'est pas complètement redondante avec celle de puissance spectrale par deux aspects. Tout d'abord l'extension spatiale d'une région cohérente discriminante ne correspond pas forcément à l'extension spatiale de la puissance discriminante dans cette région. La cohérence rend donc compte de la portée d'une oscillation corticale. Il est ainsi possible de différencier une vaste aire qui se synchronise de deux aires voisines qui oscillent sous l'effet de générateurs indépendants. De plus, nous avons mis en évidence que la cohérence par rapport à une source de référence peut diminuer pour une région et augmenter pour une autre. Cette mesure permet donc de quantifier l'appartenance d'une même source à deux réseaux différents, alors que la puissance de cette source peut rester constante. Ces deux propriétés peuvent expliquer pourquoi la cohérence apporte de l'information par rapport aux sources corticales. Cependant, ces phénomènes "de second ordre" ne sont pas les plus évidents à identifier par rapport aux phénomènes qui s'expliquent par une simple augmentation de puissance. Cela est peut être la raison pour laquelle l'accumulation d'un grand nombre de variables de cohérence est nécessaire pour améliorer les performances en classification de l'ICM.

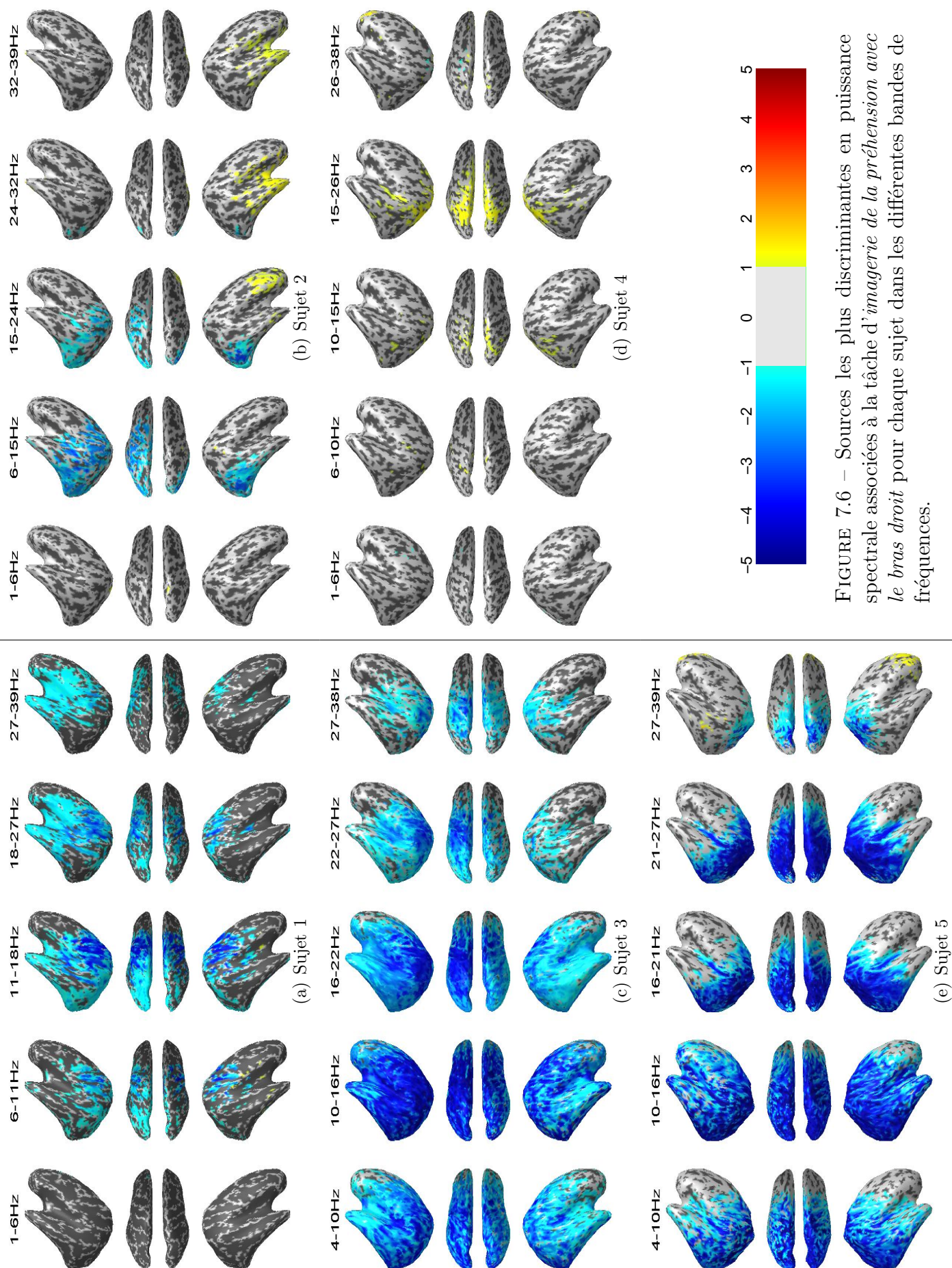


FIGURE 7.6 – Sources les plus discriminantes en puissance spectrale associées à la tâche d’*imagerie de la préhension avec le bras droit* pour chaque sujet dans les différentes bandes de fréquences.

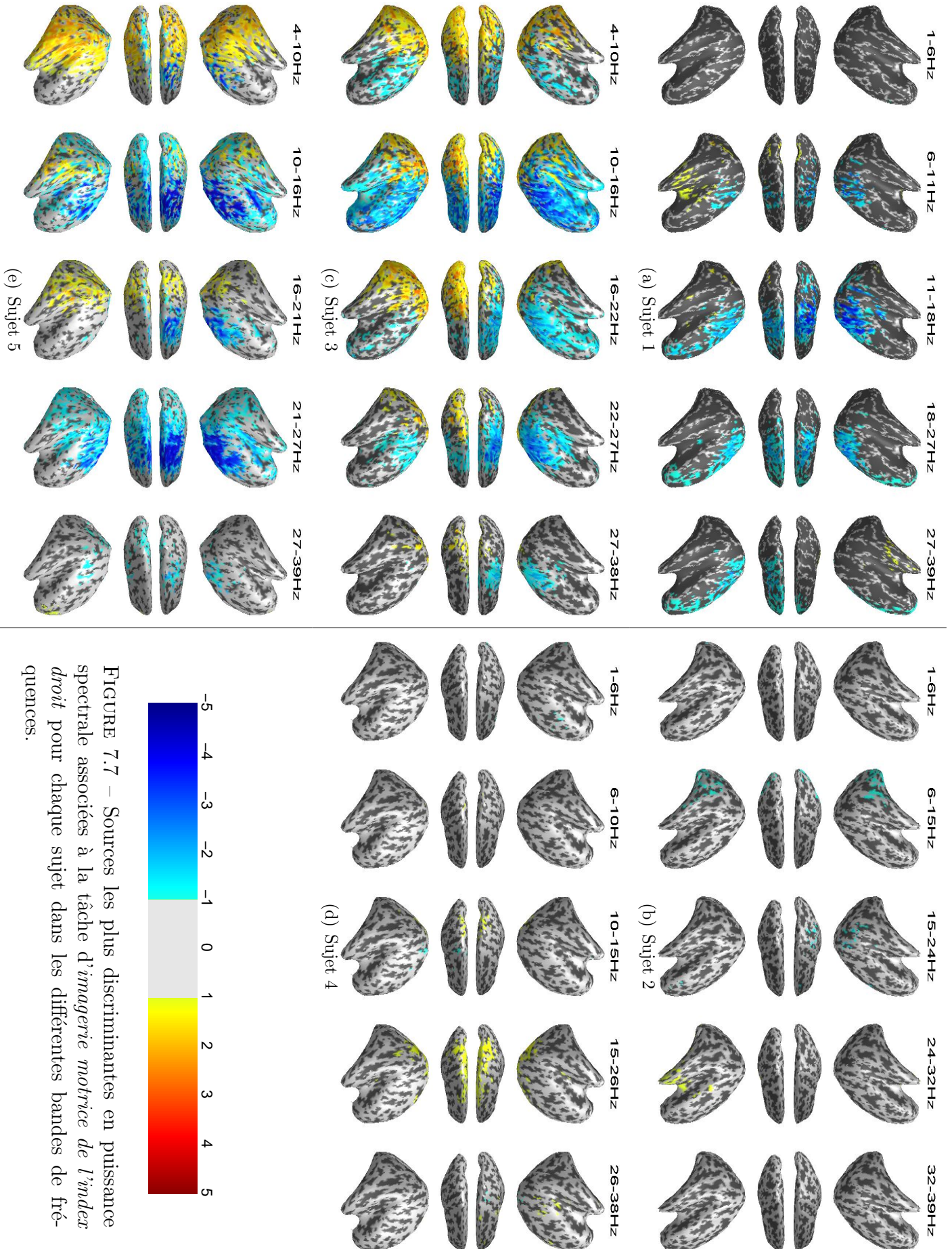


FIGURE 7.7 – Sources les plus discriminantes en puissance spectrale associées à la tâche d'*imagerie motrice de l'index droit* pour chaque sujet dans les différentes bandes de fréquences.

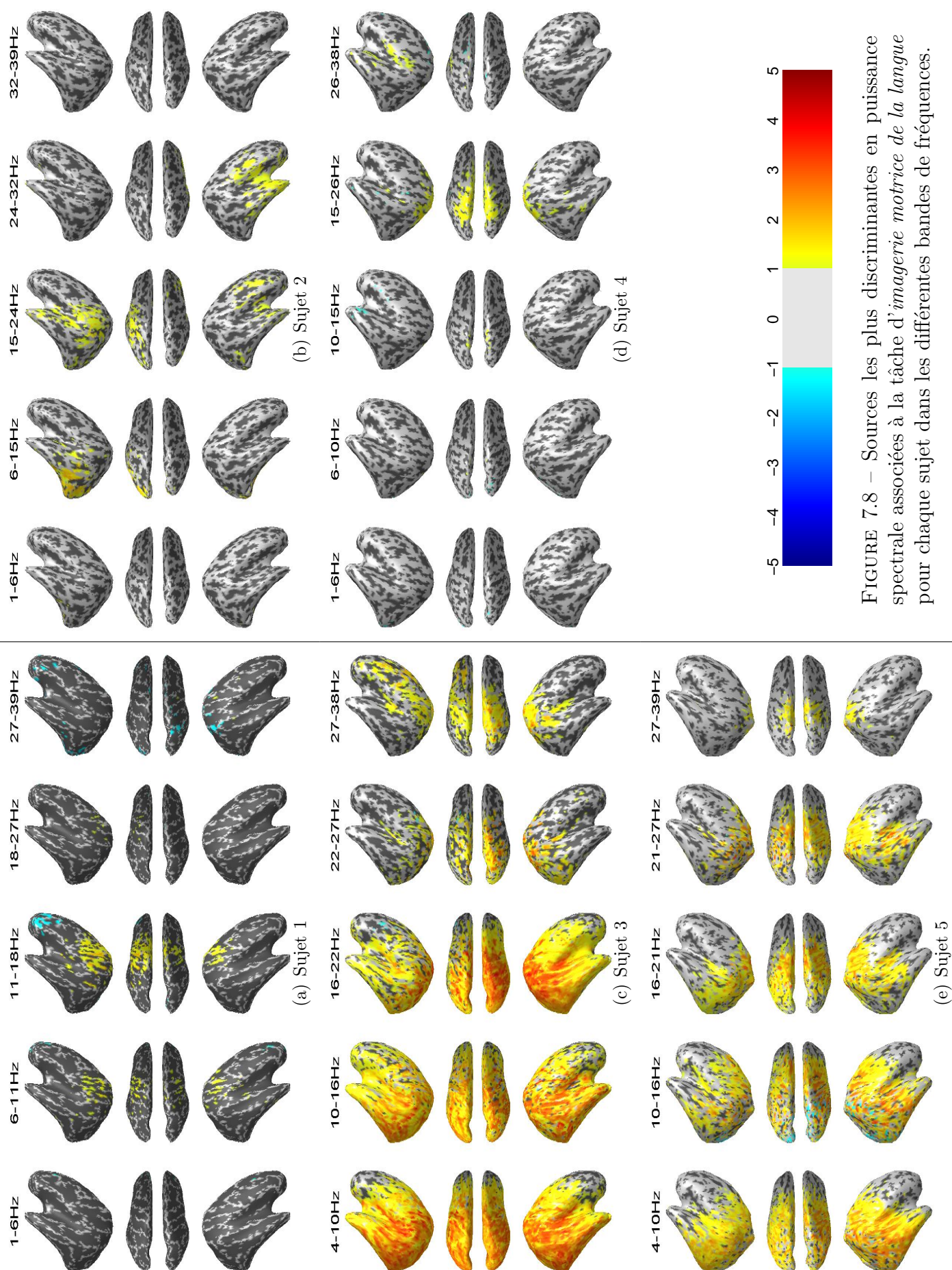


FIGURE 7.8 – Sources les plus discriminantes en puissance spectrale associées à la tâche d’imagerie motrice de la langue pour chaque sujet dans les différentes bandes de fréquences.

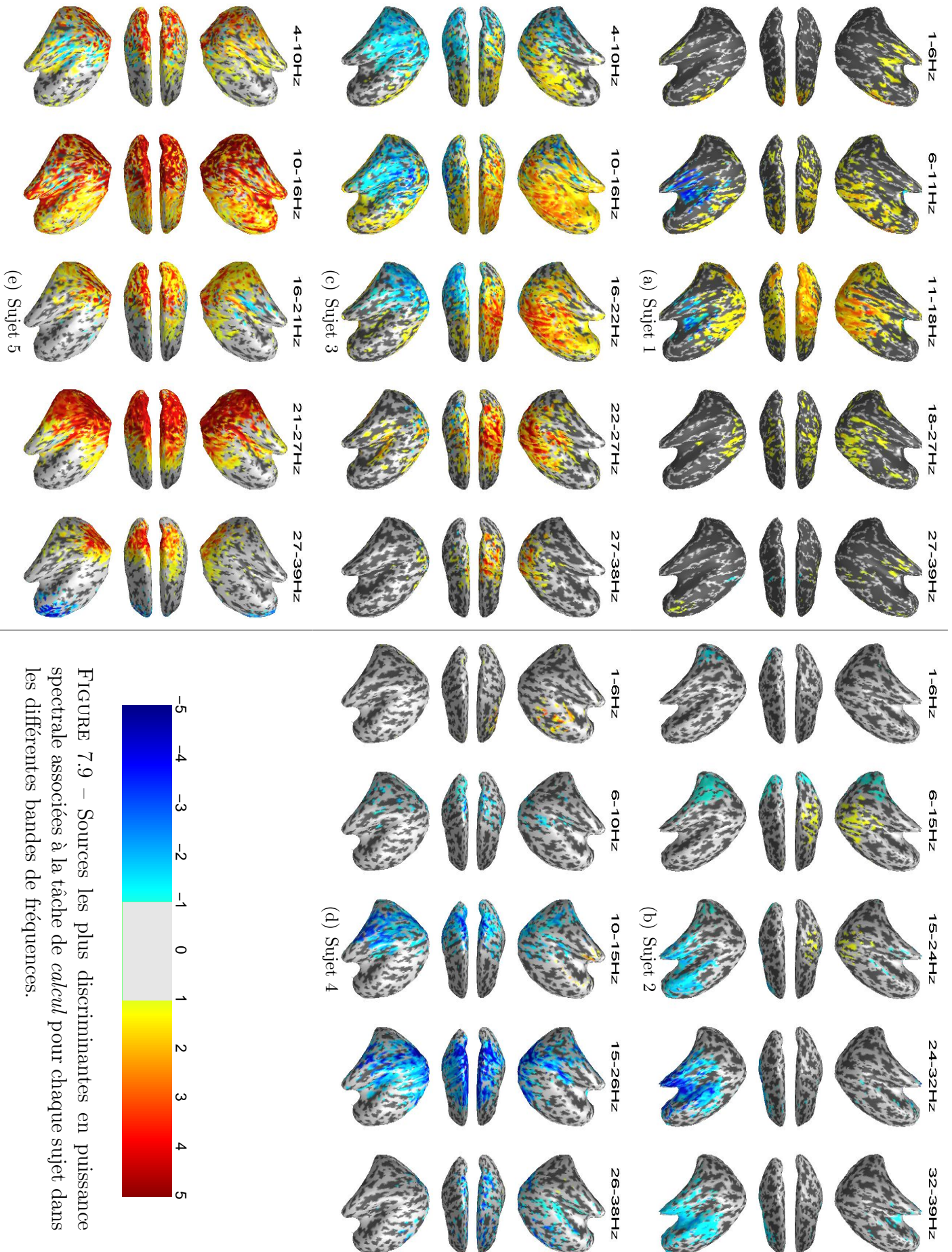


FIGURE 7.9 – Sources les plus discriminantes en puissance spectrale associées à la tâche de *calcul* pour chaque sujet dans les différentes bandes de fréquences.

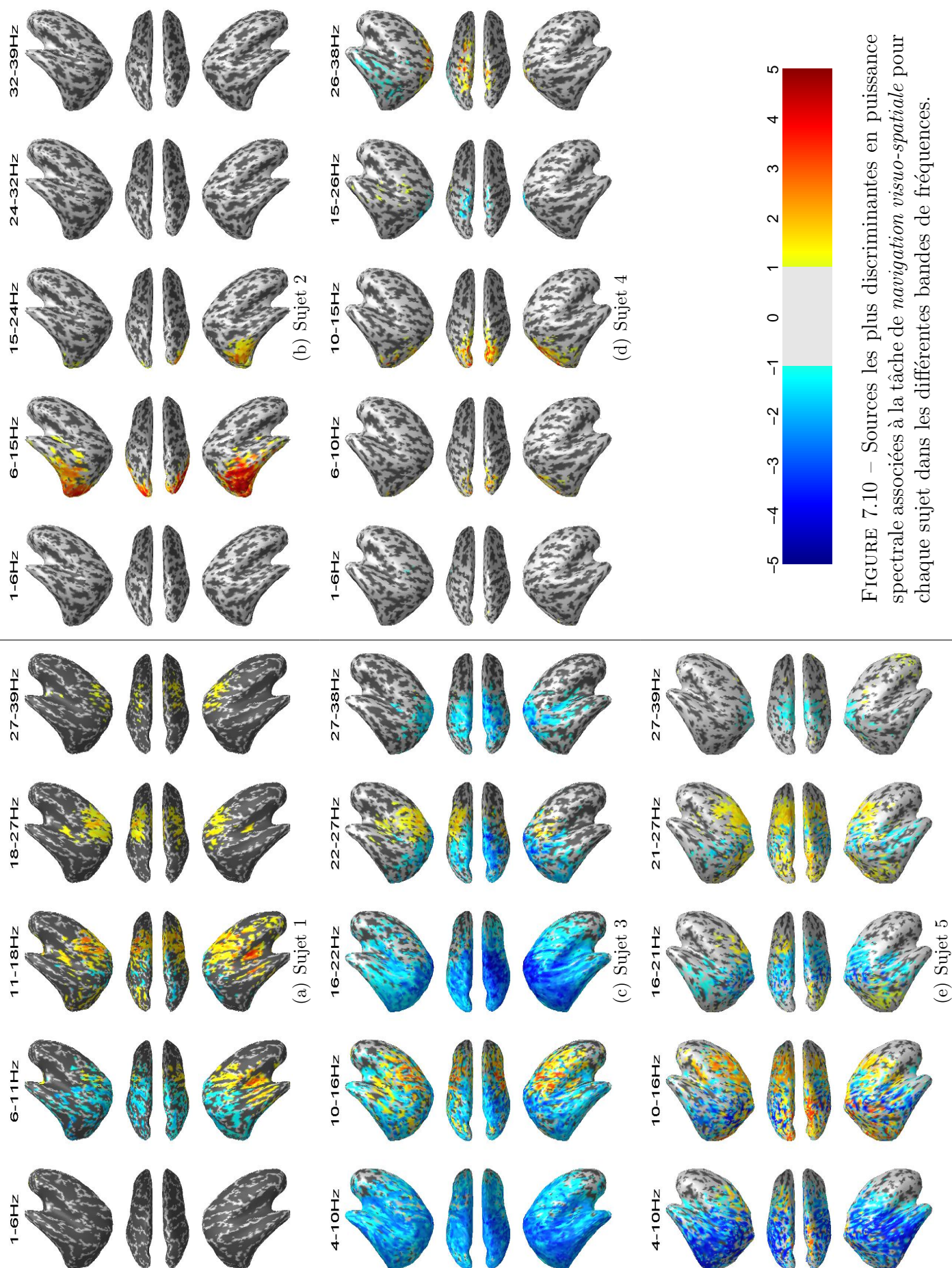


FIGURE 7.10 – Sources les plus discriminantes en puissance spectrale associées à la tâche de *navigation visuo-spatiale* pour chaque sujet dans les différentes bandes de fréquences.

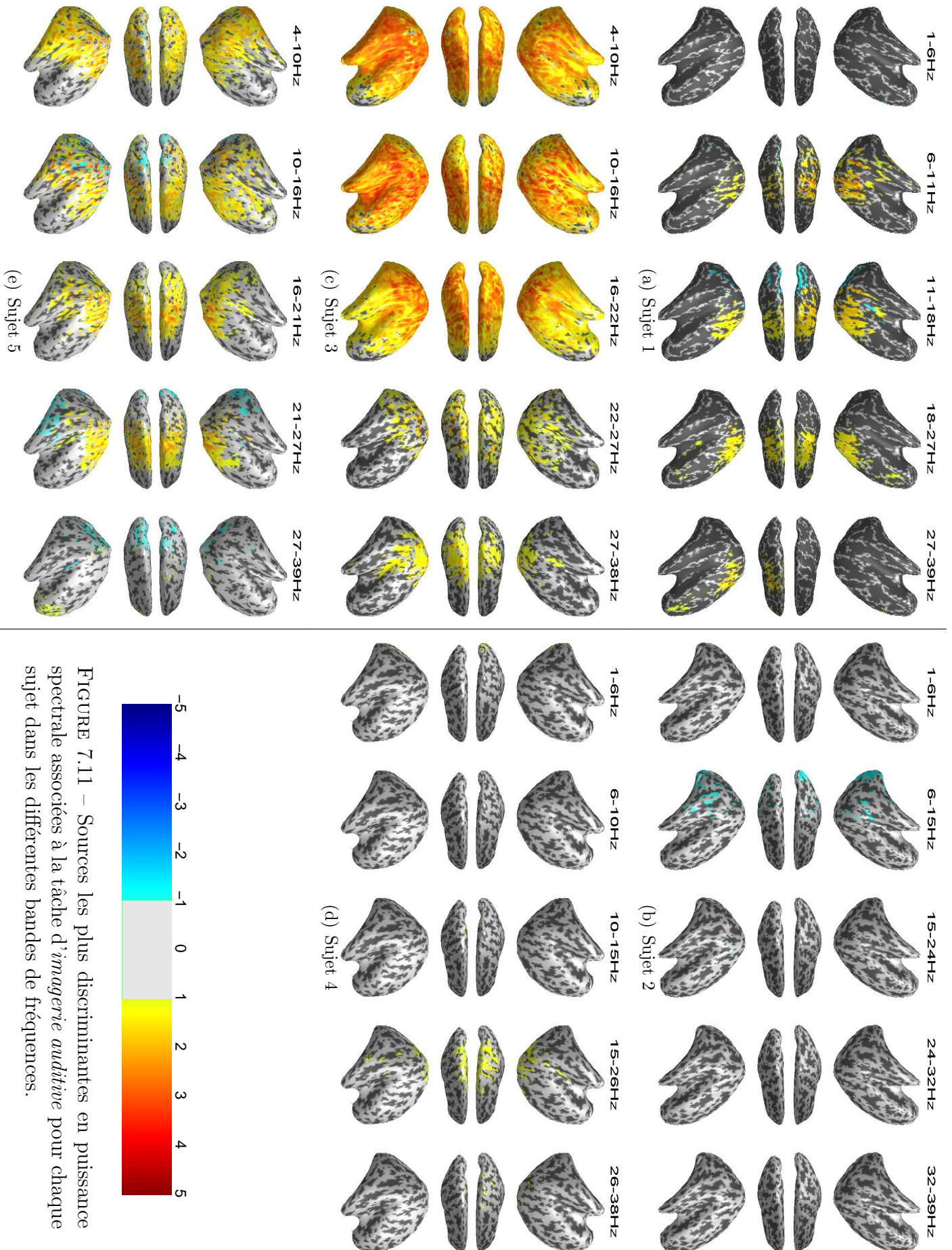


FIGURE 7.11 – Sources les plus discriminantes en puissance spectrale associées à la tâche d'*imagerie auditive* pour chaque sujet dans les différentes bandes de fréquences.

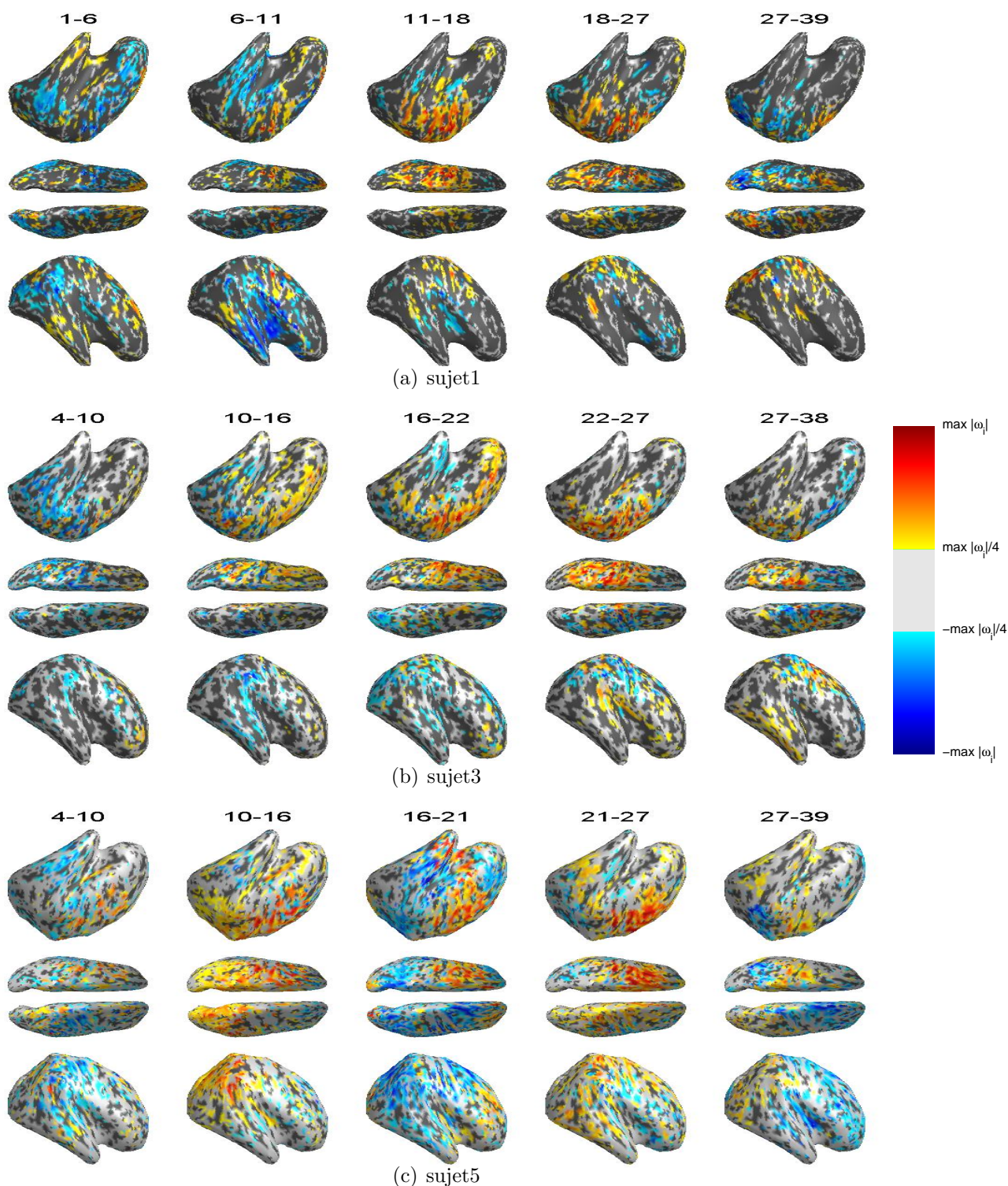


FIGURE 7.12 – Distribution des coefficients ω_i de la fonction discriminante du SVM, associés à chaque variable de quantification pour la discrimination *calcul versus index* pour les trois meilleurs sujets. L'apprentissage du SVM est réalisé sur l'ensemble des variables de puissances et sur les données de chaque jour d'enregistrement. Les coefficients résultants sont moyennés à travers les jours. La tâche de calcul est associée au label de classe $y = 1$ et l'index à $y = -1$: un coefficient positif caractérise donc une variable dont la puissance augmente dans la classe calcul par rapport à la classe index.

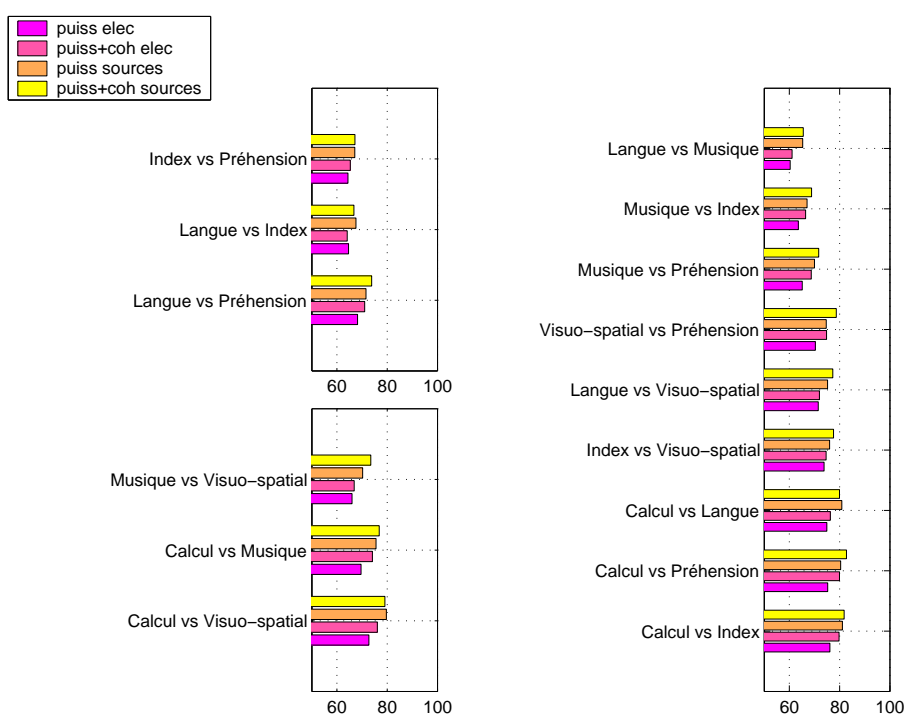
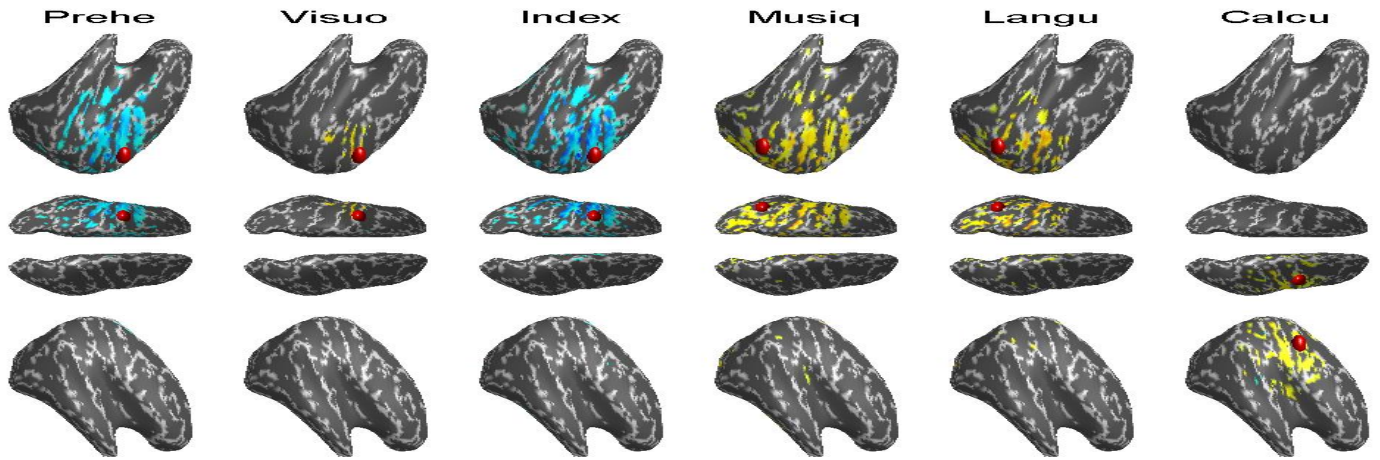
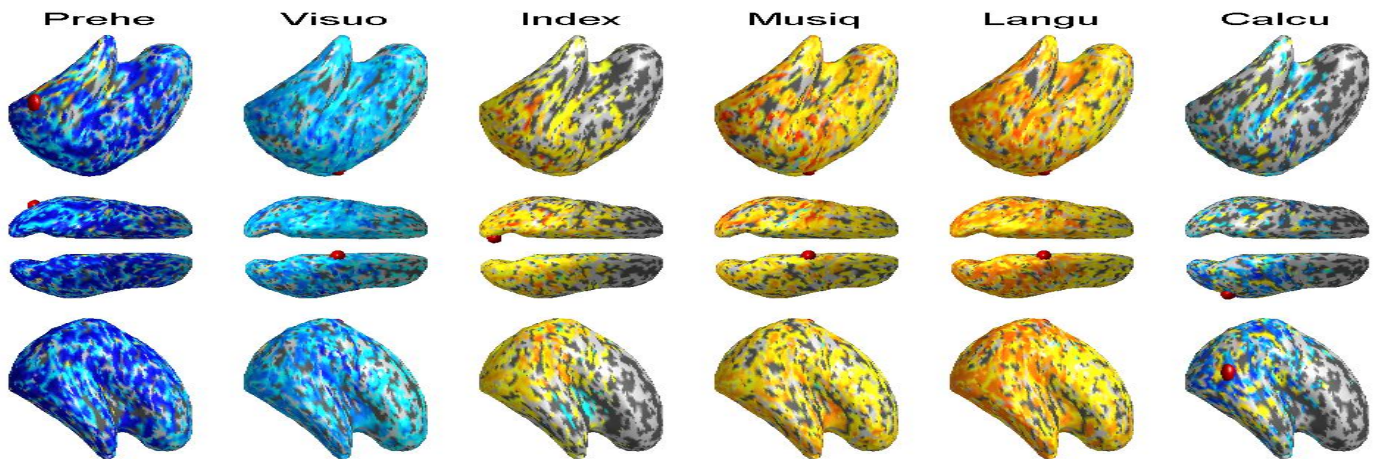


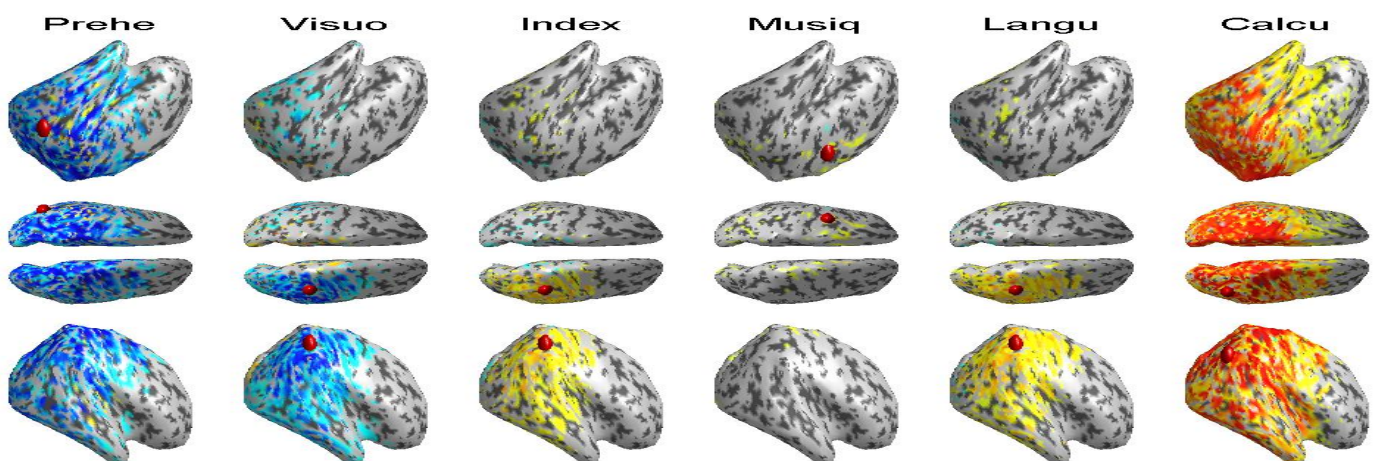
FIGURE 7.13 – Comparaison des taux de classification pour les quatre types de quantification étudiés aux chapitres 5 et 7 : puissance sur les électrodes, puissance sur les sources puissance et cohérence sur les électrodes, puissance et cohérence sur les sources.



(a) sujet1 : 6-11Hz

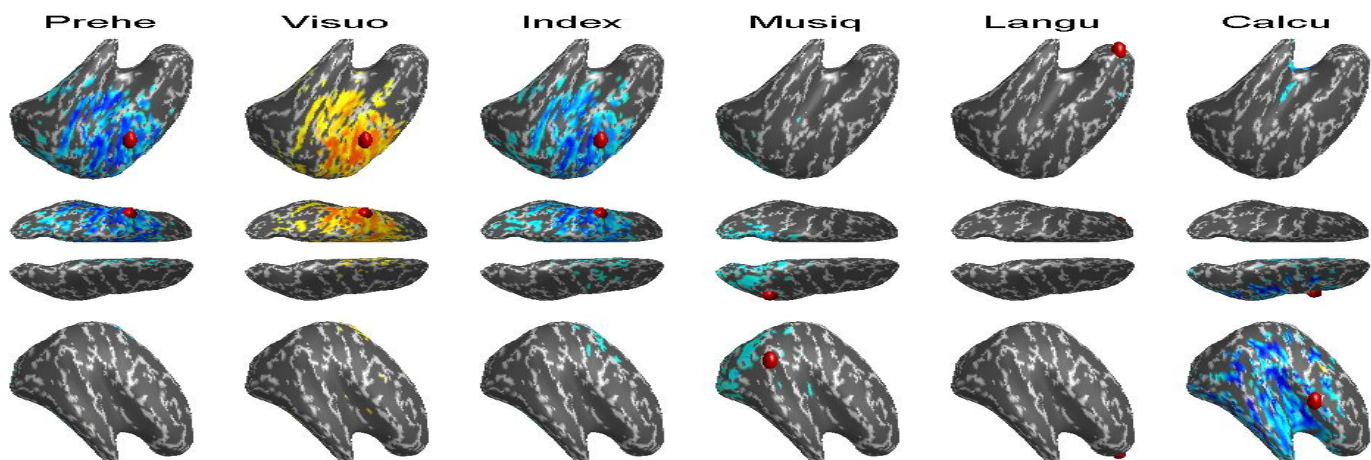


(b) sujet3 : 4-10Hz

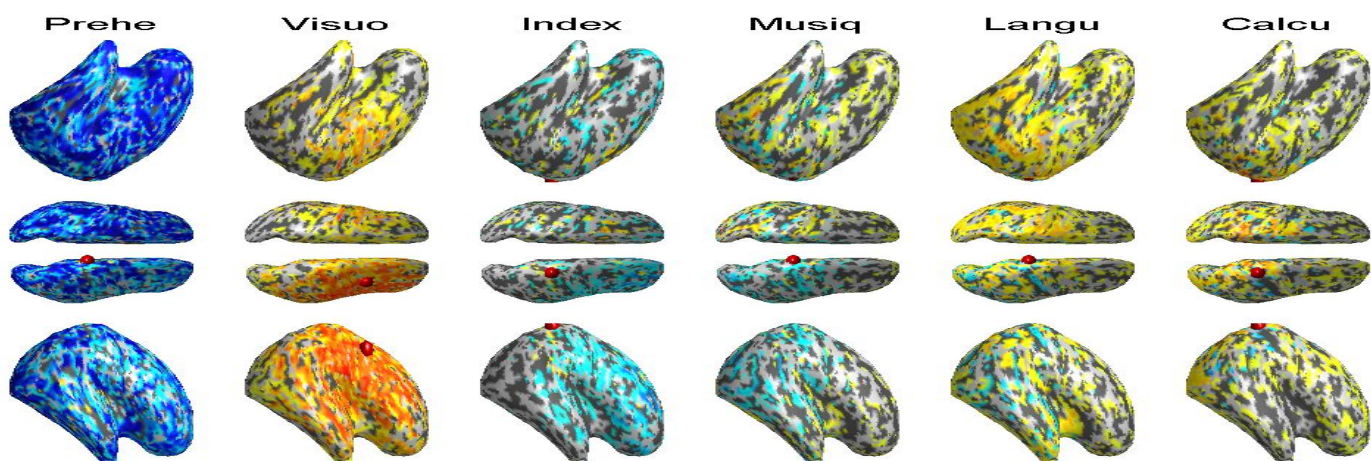


(c) sujet5 : 4-10Hz

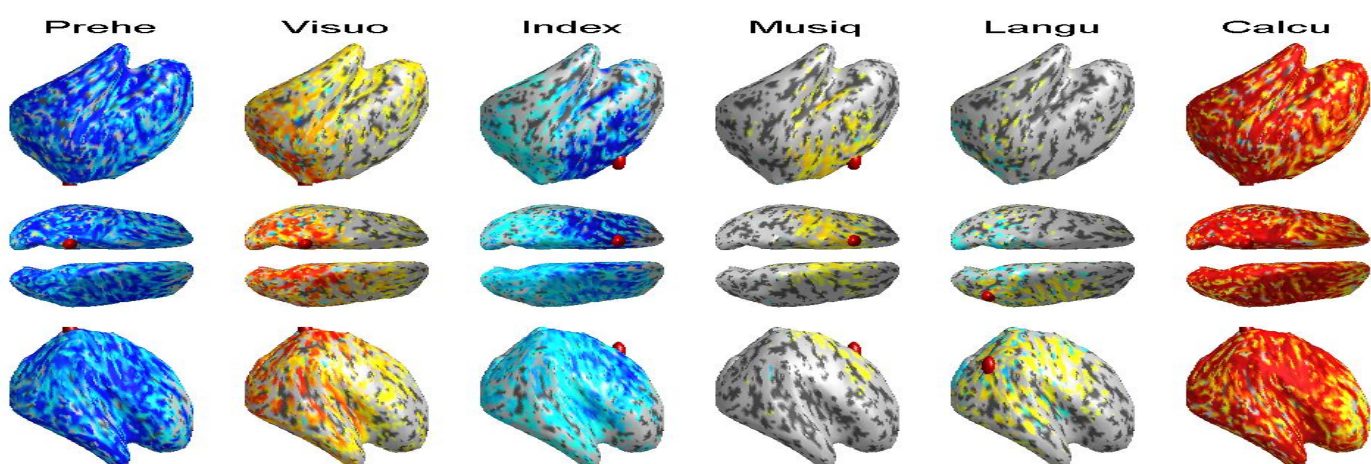
FIGURE 7.14 – Variables de cohérence significativement discriminantes pour les 3 meilleurs sujets dans leur bande 'thêta' pour la source de référence donnant les meilleurs résultats.



(a) sujet1 : 11-18Hz



(b) sujet3 : 10-16Hz



(c) sujet5 : 10-16Hz

FIGURE 7.15 – Variables de cohérence significativement discriminantes pour les 3 meilleurs sujets dans leur bande 'thêta' pour la source de référence donnant les meilleurs résultats.

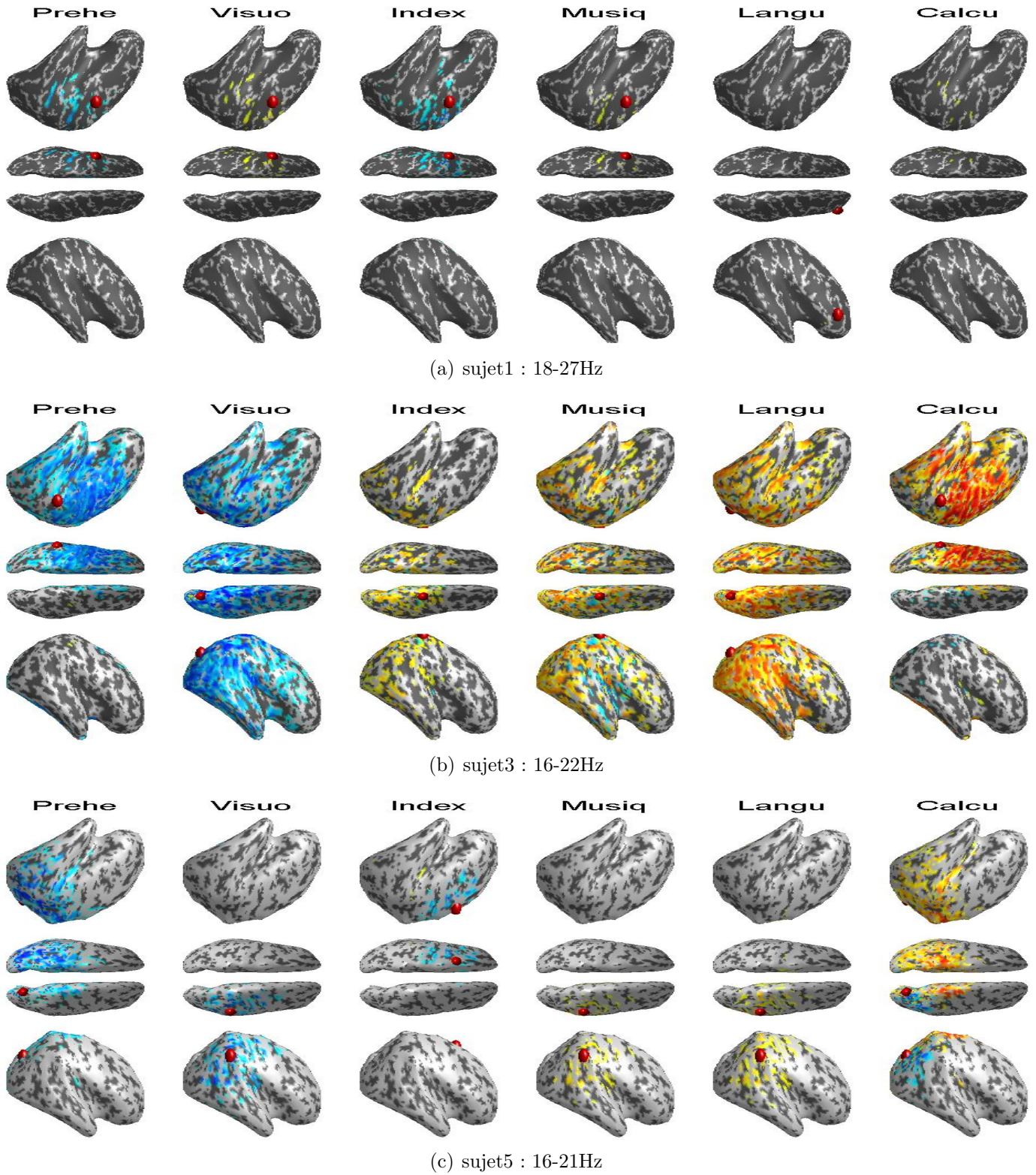


FIGURE 7.16 – Variables de cohérence significativement discriminantes pour les 3 meilleurs sujets dans leur bande 'thêta' pour la source de référence donnant les meilleurs résultats.

Chapitre 8

Discussion générale

8.1 Analyse de l'ensemble des résultats

8.1.1 Filtrage spatial et sélection de variables

Différents types de filtres spatiaux ont été utilisés dans nos travaux. Le filtre spatial le plus élémentaire est la référence moyenne commune qui a été appliquée dans le chapitre 5 et dans les travaux cités en annexe. Dans le chapitre 6 et dans (Besserve et al., 2007b), les filtres spatiaux ont été calculés à partir de méthodes de séparation de sources. Enfin, dans le chapitre 7, le filtre spatial correspond à la matrice de problème inverse appliquée aux mesures EEG pour estimer l'activité des sources corticales. Comme résumé dans le chapitre 2 le principe du filtre spatial est de rehausser les signaux spécifiques aux tâches mentales, par rapport au bruit, pour améliorer leur pouvoir discriminant. Une autre technique qui est aussi présentée est la sélection de variables permettant d'éliminer les variables non informatives. D'après, les résultats de classification obtenus dans nos travaux, nous pouvons observer différentes propriétés de ces techniques.

Par rapport au filtre spatial référence moyenne commune, l'utilisation de méthodes de séparation de sources permet d'améliorer la classification ; cependant, les résultats optimaux ne sont obtenus qu'en sélectionnant un nombre très faible de variables. En effet, dans (Besserve et al., 2007b) lorsque le nombre de variables sélectionnées augmente (chaque variable correspond à la puissance d'une composante du filtre spatial dans une bande de fréquence), le pourcentage de biens classés diminue pour atteindre celui obtenu sur les électrodes. Ce résultat peut s'interpréter de la façon suivante : la séparation de sources permet d'extraire du signal EEG un faible nombre de composantes liées à la tâche, cependant les autres composantes extraites par le filtre sont des composantes de bruit, qui, si elles sont incorporées dans le classifieur (lorsque trop de variables sont sélectionnées), tendent à diminuer le taux de biens classés. Il faut donc être très attentif à la sélection de variables lors de l'utilisation d'une technique de séparation de sources.

À l'inverse de la séparation de sources, le problème inverse fournit un ensemble de filtres spatiaux très redondants. En effet, à chacune des 10000 sources corticales correspond un filtre spatial alors que l'espace de départ de ces filtres ne comporte que 60

dimensions (environ). Il a cependant été démontré au chapitre 7 que l'utilisation de la quantification sur les sources corticales conduisait à une amélioration des résultats lorsque toutes les variables de quantification des sources corticales étaient utilisées par le classifieur. Ce résultat est l'opposé du résultat obtenu avec les méthodes de séparation de sources. Bien qu'il soit difficile d'expliquer rigoureusement ce résultat, la forte redondance des filtres spatiaux utilisés peut permettre au classifieur d'atténuer le bruit engendré par les sources qui ne portent pas d'information discriminante sur la tâche. Ces résultats, qui sont obtenus avec un classifieur SVM, ne pourraient bien sûr pas être reproduits en utilisant l'ALD qui ne peut traiter des problèmes en si haute dimension. Ainsi, peut-être ce résultat est-il précisément dû à la haute dimension de l'espace de quantification, qui, dans ce cas précis, comme dans un certain nombre d'autres¹ cas (Donoho, 2000), peut constituer une bénédiction plutôt qu'une malédiction.

8.1.2 Les tâches mentales

Notre étude de tâches mentales se place dans la continuité de Keirn and Aunon (1990), qui avaient choisi ces tâches (verbales, visuelles, spatiales) selon des critères de spécialisation hémisphérique. Il n'y avait cependant pas eu de résultats permettant de différencier clairement l'intérêt d'une tâche par rapport à une autre (peut-être à cause de la méthode de validation croisée utilisée). De plus, l'étude de ces tâches montrait qu'une même aire pouvait s'activer pour deux tâches différentes (Curran et al., 2003).

Dans cette étude, nous confirmons le recouvrement partiel des aires corticales impliquant différents types d'imagerie mentale. En particulier, les tâches de préhension, de calcul et de navigation visuo-spatiale font toutes intervenir les régions pariétales. Il est cependant probable qu'au sein de cette même aire, des régions plus spécifiques de chaque tâche soient activées dans chaque cas, ce qui expliquerait pourquoi il est quand même possible de différencier ces tâches avec des taux de classification atteignant 80%. Nous observons de plus que la latéralisation des tâches mentales étudiées n'est pas toujours nette et que les tâches se différencient plus par la distribution de leurs activations sur l'axe antéro-postérieur. C'est en particulier le cas du contraste entre tâches motrices et non motrices qui fait intervenir des variations d'activités opposées dans les régions antérieures et postérieures. Si l'on établit un classement des tâches mentales, dans les meilleures tâches on trouve le calcul, l'imagerie de l'index et de la préhension, qui ont toutes trois des distributions spatiales très marquées. En ce qui concerne la préhension, il est surprenant d'observer pour une tâche motrice que les variations d'activités les plus importantes sont localisées au niveau de la région pariétale en position médiane et pas dans le cortex moteur primaire. Cela pourrait correspondre à une activation de la *parietal reaching region* mise en évidence en IRMf (Connolly et al., 2003) et en EEG récemment (Naranjo et al., 2007). Ce phénomène met aussi en évidence le fait que même pour une imagerie motrice, il peut exister d'autres régions que le cortex moteur

1. on peut citer notamment le phénomène de concentration de la mesure qui montre, sous certaines conditions, qu'une variable aléatoire dans un espace de haute dimension prend ses valeurs dans des domaines restreints

qui peuvent s'activer de manière spécifique.

Parmi les tâches les moins spécifiques, on trouve d'abord l'imagerie auditive dans les tâches non motrices. Cela peut s'expliquer par une mauvaise couverture des régions temporales qui sont susceptibles d'être les plus impliquées dans cette tâche. D'après les résultats, on peut considérer qu'il s'agit d'une tâche à éliminer de notre protocole expérimental. Une autre tâche peu spécifique est l'imagerie de la langue, car la topographie de ces activations, qui correspondent à une augmentation bilatérale de l'activation du cortex moteur, est celle d'une tâche non motrice. Là encore, la région associée au mouvement de la langue se situe au voisinage du lobe temporal, et est donc susceptible de ne pas être bien couverte par le casque EEG. Malgré cette non-spécificité, l'imagerie de la langue contraste bien avec les autres tâches d'imagerie motrice et peut donc être utilisée en ICM.

L'ensemble des résultats reporté dans les chapitres précédents fait aussi état d'une variabilité des résultats en fonction des sujets. En effet, deux sujets en particulier ont des performances beaucoup moins bonnes que les trois autres. Ces différences se retrouvent aussi dans l'analyse de la topographie de leurs activités corticales discriminantes qui sont très différentes des autres sujets. L'interprétation de ces résultats est difficile, mais d'après l'analyse des résultats de classification et des topographies associées, les classes qui se différencient par rapport aux autres sujets sont surtout les tâches d'imagination motrice de l'index et de la préhension. La cause des mauvaises performances peut donc être l'utilisation de stratégies mentales visuelles qui font peu intervenir le cortex moteur et se différencient donc peu des tâches de navigation visuo-spatiale ou de calcul. Ce résultat a déjà été reporté dans [Stinear et al. \(2006\)](#).

8.2 Perspectives

8.2.1 Améliorations méthodologiques

Plusieurs directions peuvent être envisagées pour améliorer nos algorithmes. Tout d'abord, l'amélioration de la résolution spatiale pourrait jouer un rôle important. Le problème inverse qui a été utilisé possède une faible résolution spatiale. L'utilisation d'une méthode plus avancée, comme l'utilisation d'une norme L1 au lieu d'une norme L2 pour la régularisation peut être envisagée. Cela suppose cependant de trouver des algorithmes suffisamment performants pour être utilisés en temps réel, car la solution du problème inverse en norme L1 ne dépend pas linéairement des données EEG mesurées. Si un problème inverse mieux résolu est utilisé, il deviendra peut-être aussi possible de localiser les régions corticales actives plutôt que de quantifier l'activité des sources corticales en tous points de la surface corticale.

Les mesures de quantification (puissance et interactions à distance) actuellement utilisées sont moyennées sur des fenêtres temporelles de deux secondes. Une quantification temporelle plus fine peut être envisagée, basée sur l'hypothèse que les phénomènes neurophysiologiques associés à une tâche mentale peuvent être décomposés en une succession de microétats ([Wackermann et al., 1993](#)). En effet, certaines tâches mentales

comme le calcul sont composées de tâches cognitives élémentaires et nécessitent de faire appel successivement à plusieurs régions cérébrales. L'accès à cette dynamique temporelle peut permettre à terme d'améliorer la différenciation des tâches mentales.

8.2.2 Études expérimentales futures

Les études off-line présentées dans ce manuscrit sont une première étape de caractérisation des tâches mentales utilisées. Les premières expériences de feedback qui ont été réalisées ont montré des résultats encourageants. Elles seront prolongées par de nouvelles expériences incluant une quantification de l'activité des sources corticales. Afin d'accélérer la procédure d'apprentissage, une évolution de l'implémentation temps réel du protocole de feedback sera envisagée pour permettre un apprentissage en ligne du classifieur.

Annexe A : Travaux sur la fatigue cognitive

Présentation

Un autre application potentielle des techniques de classification aux données d'imagerie fonctionnelle cérébrale est le monitoring de la fatigue cognitive. Par exemple, la détection des baisses de vigilance chez les pilotes pourrait permettre d'éviter de graves accidents. Dans l'article suivant, nous exploitons des enregistrements EEG réalisés par Matthieu Philippe et Geneviève Florence à l'Institut de Médecine Aérospatiale et du Service de Santé des Armées(IMASSA). La classification des signaux EEG est réalisée pour différencier les périodes de temps de réaction court et long des sujets dans une tâche de compatibilité spatiale. La méthode implémentée réalise un vote pondéré entre six classifieurs SVM exploitant chacun des mesures de puissances spectrales ou de synchronie de phase dans une des trois bandes de fréquences suivantes : thêta, alpha ou bêta (voir dans l'article Figure 1 page 22). Chacun des six classifieurs apporte sa contribution à la classification en fonction du poids qui quantifie son pouvoir discriminant. Ceci permet de mettre en valeur les sous ensembles de variables les plus discriminants.

Principaux résultats

Les signaux ont été analysés sur des fenêtres de 20s. Le taux de bien classés moyenné à travers les sujets est de 71% (voir Table 1 page 29) pour la méthode de vote pondérée dite de "stacking"(Wolpert, 1992). Ceci correspond à un taux de transfert d'information de .5 bits par minute, ce qui rend envisageable la détection des baisses de performances avec un délai de l'ordre d'une ou deux minutes. Les variables de quantification les plus discriminantes sont les mesures de puissance dans la bande thêta qui diminuent lorsque le temps de réponse du sujet est court (voir Figure 5 page 26). Ceci pourrait refléter une inhibition plus forte des circuits de la mémoire à long terme lorsque le temps de réponse est faible. Dans ces conditions, on note aussi une hausse de synchronie de phase dans la bande alpha dans un réseau antéro postérieur (voir Figure 6 page 27) qui pourrait être liée à un traitement attentionnel visuel plus performant.

Perspectives

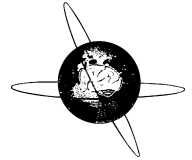
Ces travaux sur la fatigue cognitive sont prolongés dans la thèse de François Laurent, qui va appliquer des méthodes de classification sur des enregistrements EEG en simulateur de vol réalisés dans le cadre d'une collaboration avec l'IMASSA. Ces travaux ont à terme pour but d'anticiper les erreurs de pilotage dues à la fatigue cognitive ou mentale (Laurent et al., 2007).

Article (accepté pour publication dans *Clinical Neurophysiology*)



ELSEVIER

Clinical Neurophysiology xxx (2008) xxx–xxx



www.elsevier.com/locate/clinph

Prediction of performance level during a cognitive task from ongoing EEG oscillatory activities

Michel Besserve^{a,b}, Matthieu Philippe^c, Geneviève Florence^{a,c},
François Laurent^{a,b}, Line Garnero^{a,b}, Jacques Martinerie^{a,b,*}

^a Laboratoire Neurosciences Cognitives et Imagerie Cérébrale, CNRS UPR 640-LENA, 75013 Paris Cedex 13, France

^b UPMC Univ Paris 06, F-75005, Paris, France

^c Institut de Médecine Aéronautique du Service de Santé des Armées, B.P. 73 - 91223 - Brétigny-sur-Orge Cédex, France

Accepted 2 December 2007

Abstract

Objective: Tracking the level of performance in cognitive tasks may be useful in environments, such as aircraft, in which the awareness of the pilots is critical for security. In this paper, the usefulness of EEG for the prediction of performance is investigated.

Methods: We present a new methodology that combines various ongoing EEG measurements to predict performance level during a cognitive task. We propose a voting approach that combines the outputs of elementary support vector machine (SVM) classifiers derived from various sets of EEG parameters in different frequency bands. The spectral power and phase synchrony of the oscillatory activities are used to classify the periods of rapid reaction time (RT) versus the slow RT responses of each subject.

Results: The voting algorithm significantly outperforms classical SVM and gives a good average classification accuracy across 12 subjects (71%) and an average information transfer rate (ITR) of 0.49 bit/min. The main discriminating activities are laterally distributed theta power and antero–posterior alpha synchronies, possibly reflecting the role of a visual-attentional network in performance.

Conclusions: Power and synchrony measurements enable the discrimination between periods of high average reaction time versus periods of low average reaction time in a same subject. Moreover, the proposed approach is easy to interpret as it combines various types of measurements for classification, emphasizing the most informative.

Significance: Ongoing EEG recordings can predict the level of performance during a cognitive task. This can lead to real-time EEG monitoring devices for the anticipation of human mistakes.

© 2007 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.

Keywords: Classification; Support vector machine; Phase synchrony; Phase locking value

1. Introduction

The prediction of mental states by classifying ongoing EEG signals is a challenging problem. Resolving this can lead to various applications, such as Brain–Computer Interfaces (Wolpaw et al., 2002). One part of the problem is the large number of relevant neurophysiological parameters that can be extracted from the EEG. For example, the amplitudes of various oscillations can be measured on each electrode and the long-distance interactions between the underlying

brain areas can be quantified using several methods (coherence, phase synchrony, etc.). The strategy employed by experts in clinical EEG that uses the qualitative detection of oscillatory activities in well-known frequency bands has led us to propose combining the outputs of “expert classifiers” derived from various sets of EEG parameters via a weighted vote of the classifiers. This voting approach is based on the statistical stacking framework (Wolpert, 1992). Here, we apply this method to classify the level of performance of subjects during a cognitive task.

Estimating this level of performance may allow serious human mistakes in working environments, such as aircraft, to be anticipated. Cognitive task performance is related to

* Corresponding author. Tel.: +33 142161171; fax: +33 145862537.
E-mail address: jacques.martinerie@chups.jussieu.fr (J. Martinerie).

many psychological variables, such as attention, arousal or workload, and there is much research in studying EEG correlates of these variables. For example, some studies have used Event Related Potentials to quantify the changes in information processing due to high workload or cognitive fatigue (Hohnsbein et al., 1995; Boksem et al., 2005). However, most work focuses on fluctuations in the power of EEG signals in the theta (4–7 Hz), alpha (7–13 Hz) and beta (13–18 Hz) bands (Smith et al., 2001, 2004; Klimesch, 1999; Grandjean, 1988).

Brain activity can also be quantified by assessing the interactions between different parts of the cortex by measuring, for example, phase synchrony. This has already proved useful in various studies of cognitive and pathological brain activity. We have shown that these measurements from surface EEGs can discriminate between “face perception” and “no face perception” during a visual task (Rodriguez et al., 1999) and, with intra-cerebral recordings, can be used to anticipate epileptic seizures (Le Van Quyen et al., 2005). Recently, research in the domain of Brain–Computer Interfaces has shown that combining spectral power measurements with phase synchrony improves accuracy when classifying EEG signals (Gysels and Celka, 2004; Gysels et al., 2005), and thus should prove useful for predicting the level of performance. Moreover, those two measures can be used within the same biological framework of “Resonant Cell Assemblies” (Varela, 1995), with phase synchrony revealing synchronized activities of neural populations at large spatial scales (Varela et al., 2001), and spectral power possibly revealing the same phenomenon on smaller scales (Gray and Singer, 1989; Lopes da Silva, 1991; Tallon-Baudry et al., 1999; Whittington et al., 2000).

Machine learning applications using EEG measurements have already been developed in similar research domains. For example, good results have been obtained for discriminating tasks having different workloads, under either simulated or real working conditions (Wilson and Fisher, 1998; Wilson and Russell, 2003). However, the experimental framework used in these studies could not detect fluctuations in the performance of a subject undertaking a unique task with constant workload. Neural networks have also been used to monitor the attentional state, using power features (Kiyimik et al., 2004; Jung et al., 1997), correlation and coherence (Makeig et al., 1996) in an auditory cue detection task, revealing that these features are correlated with arousal. More recently, EEG power spectra have been used to give an objective index of driving performance (Lin et al., 2005).

Continuing these studies, here, we have investigated how to classify slight fluctuations in performance level (measured by reaction time) during a cognitive task. We have used information from spectral power and phase synchrony features in three frequency bands from each electrode and electrode pair, respectively. The Support Vector Machine (SVM) was chosen as the elementary classification algorithm because of its good generalization ability (Burges, 1998; Cristianini and Shawe-Taylor, 2000).

However, some of the several sets of measurements chosen may be irrelevant for our classification problem and drop-off classifier accuracy. This is due to the small number of data points available, comparatively to the high number of features, which makes classification highly sensitive to noise. One simple solution is to use cross-validation to choose the best set of features, discarding the rest, although some informative features may be lost. The stacking framework (Wolpert, 1992) is a generalized cross-validation, which performs better than the “choose best” strategy (LeBlanc and Tibshirani, 1996). It combines various estimators according to their respective performance in the cross-validation. We used stacking to combine predictions of SVM classifiers trained with different sets of features. The resulting combination rule is weighted so that the highest weights are allocated to the most accurate classifiers.

Thus, in this study, we propose using a stacking approach to classify the level of performance from power and synchrony measurements computed during a cognitive task. Section 2.1 describes the general methodology. The layout of the paper is as follows. Section 2.2 describes the features extracted from EEG data. Section 2.3 reviews the SVM classifier and the different steps in the stacking approach. Section 2.4 explains how this classifier is applied to EEG data. Section 2.5 describes the experimental data and explains how a performance index is extracted from a reaction time. Sections 3 and 4, respectively, present and discuss the results.

2. Methods

2.1. Scheme of the method

Using EEG data, our approach aims to classify periods of high performance levels versus periods of low performance levels, which are defined from the reaction times (see Section 2.5.3). We used the EEG data from a sliding time-window to compute spectral power (for all electrodes) and phase synchrony (for all couples of electrodes) in three frequency bands (θ , α and β : see below for details). This gives six sets of features (θ , α and β spectral powers, θ , α and β phase synchronies). Each set is fed into an SVM classifier, providing its own estimate of the performance level of the subject (“high” or “low”) associated with each time-window. The six SVM estimates are then combined with the stacking algorithm to give an optimal estimate of performance level (Fig. 1). The different parts of this method are described further in the following sections.

2.2. Feature extraction

2.2.1. Frequency bands

In this study, EEG signals are quantified in the following frequency bands:

- θ band (3–7 Hz)
- α band (7–13 Hz)
- β band (13–18 Hz)

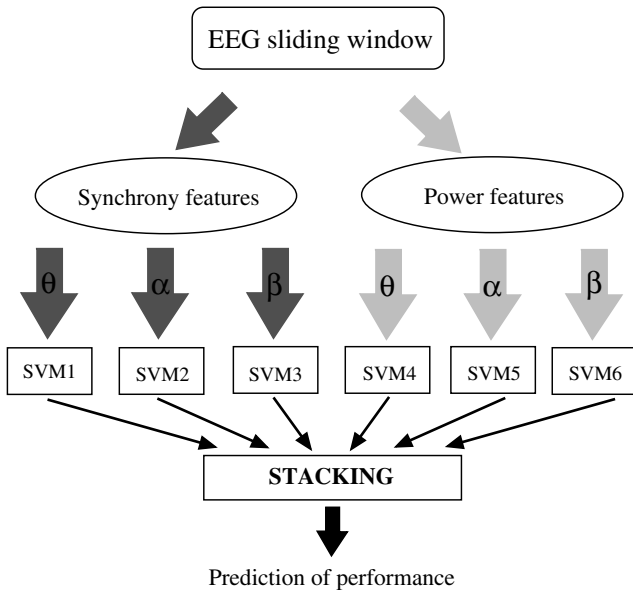


Fig. 1. EEG classification approach: power and synchrony features in three frequency bands are used to train six SVM “experts”. Experts’ opinions is then combined by a stacking algorithm to provide a prediction of subjects’ performance.

Higher frequency bands were not considered since they were found too sensitive to muscle artifacts to be used in single trial analysis. After multiplication by a Hamming window to avoid side effects, the vectors of sampled EEG signals $z(t), t = 1, \dots, T$ of the k th time window W_k were band-pass filtered in each frequency band using zero-phase FIR filters (computed by frequency sampling (Orfanidis, 1996)). Brain activities were quantified, via the filtered signals, by two indexes: local oscillatory activities under each electrode by spectral power, and long-distance interactions by phase synchrony of each electrode pair.

2.2.2. Power values

Spectral power was determined for each frequency band f by calculating the average power of the filtered signal $z_i^f(t)$ of channel i during each time-window W_k containing T time samples:

$$P_i^f(k) = \frac{1}{T} \sum_{t=1}^T |z_i^f(t)|^2$$

2.2.3. Phase synchrony

The phase locking value has been previously described (Lachaux et al., 1999). It is widely used to measure the non-linear relationships between two signals. The phase locking value is calculated by converting the filtered data of each channel i in the frequency band f , $z_i^f(t)$, into an analytical signal using the Hilbert transform \mathbf{H} . This allows us to define properly an instantaneous phase of the filtered signal (Pikovski et al., 2001): $a_i^f(t) = z_i^f + \mathbf{iH}(z_i^f) = A_i^f(t)e^{i\phi_i^f(t)}$, where \mathbf{i} is the complex number of modulus 1 and argument $\frac{\pi}{2}$. The phase locking value between channels i and j in a given frequency band is thus:

$$S_{ij}^f(k) = \frac{1}{T} \left| \sum_{t=1}^T e^{i(\phi_i^f(t) - \phi_j^f(t))} \right|$$

2.2.4. Normalization

For all channels and all electrode pairs, we independently normalized the power and phase synchrony values by subtracting the mean and dividing by the standard deviation over all time-windows. The different features, $P_i^f(k)$ and $S_{ij}^f(k)$, can be used to construct a single vector \mathbf{x}_k associated to the k th time window. The resulting vector has a length $N_f \times (\frac{N_e \times (N_e - 1)}{2} + N_e)$ for N_f frequency bands and N_e electrodes.

2.3. Classification methodology

2.3.1. Background

Let y be the performance class ($y = +1$ for a good performance corresponding to rapid reaction times (RT), and $y = -1$ for a bad performance corresponding to slow RT). The aim of a classifier is to achieve a reliable prediction \hat{y} of y from the feature vector \mathbf{x} . This is solved by calculating a discriminant function $\hat{y} = d^L(\mathbf{x})$ from a learning set $L = \{(\mathbf{x}_k, y_k), k \in I\}$, where I is the set of time-windows taken as learning examples.

2.3.2. The linear support vector machine

The linear SVM is a classification algorithm that separates two classes by calculating a separating hyper-plane in the learning data space and by penalizing badly classified data points (Vapnik, 1998; Burges, 1998; Muller et al., 2001). The resulting discriminant function is:

$$d^L(\mathbf{x}) = \text{sign}(\boldsymbol{\omega}^* \mathbf{x} + b^*)$$

where $\boldsymbol{\omega}^*$ is a normal vector of the separating hyper plane and b^* is a bias constant that can possibly be modified to control the proportion of classification errors in a particular class (false negatives for example). The coefficients of the classifier are computed according to:

$$(\boldsymbol{\omega}^*, b^*) = \arg \min_{\boldsymbol{w}, b} \frac{1}{2} \|\boldsymbol{w}\|^2 + C \sum \xi_k$$

under the constraints

$$y_k(\langle \boldsymbol{w}, \mathbf{x}_k \rangle + b) \geq 1 - \xi_k, \quad \forall k \quad \text{and} \quad \xi_k \geq 0, \quad \forall k$$

For the implementation of the algorithm, the parameter C is to be chosen by the user, a larger C corresponding to assigning a higher penalty to errors. Using preliminary results, we set $C = 1$ for all the experiments described in this paper.

2.3.3. Stacking algorithm

Stacking has been described previously in a very general framework (Wolpert, 1992). In the present study, the aim of this method is to combine the decision rules of N classifiers $(d_n^L(\mathbf{x}))_{n=1, \dots, N}$ to improve the ability of classification to predict the class y if the feature vector \mathbf{x} is known.

The combinations considered here are weighted votes of the form:

$$\hat{y} = \text{sign} \left(\sum_n \beta_n d_n^L(\mathbf{x}) \right)$$

where the optimal weights, β_n^* , can be found by solving:

$$\beta^* = \arg \min_{\beta} \sum_k \left(\left(\sum_n \beta_n d_n^L(\mathbf{x}_k) \right) - y_k \right)^2$$

This is a least square regression fit between the results of the decision rules on training data and the real class y . Unfortunately, this method produces poor results due to overfitting of the training data. Thus, certain classifiers can be given a greater weight despite performing badly on new data. Stacking solves this problem by evaluating the decision rules on new data. This can be done by cross-validation: the learning set is split into J subsets thus $\{L^j\}_{j=1, \dots, J}$. One subset, L^j , is kept as a test set while the remaining elements (denoted as $L \setminus L^j$) are used for learning. This procedure is carried out iteratively on each subset, and the decision rules $d_n^{L \setminus L^j}$ are tested on unlearned data and used to fit a new regression:

$$\beta^* = \arg \min_{\beta} \sum_{j=1}^J \sum_{(\mathbf{x}_k, y_k) \in L^j} \left(\left(\sum_n \beta_n d_n^{L \setminus L^j}(\mathbf{x}_k) \right) - y_k \right)^2 \quad (1)$$

Performance can be further improved by imposing positivity constraints and summing the weights to one.

$$\begin{cases} \sum \beta_n = 1 \\ \beta_n \geq 0, \quad \forall n \end{cases} \quad (2)$$

Moreover, with these constraints, weights can be interpreted as posterior probabilities in the framework of model averaging (Hastie et al., 2001). Thus, Eq. (1) can be solved with a quadratic programming algorithm (see Appendix). The different steps of the stacking algorithm are summarized in Algorithm 1.

Algorithm 1. Stacking procedure

```

 $(L^1, \dots, L^J) \leftarrow$  Split the learning set  $L$  into  $J$ -folds
for  $j \rightarrow 1$  to  $J$  do
  for  $n \rightarrow 1$  to  $N$  do
     $d_n^{L \setminus L^j} \leftarrow$  train classifier  $n$  on  $L \setminus L^j$ 
     $d_n^{L \setminus L^j}(\mathbf{x}_{k \in L^j}) \leftarrow$  test classifier  $n$  on  $L^j$ 
  end for
end for
 $\beta^* \leftarrow$  solve Eq. (1) under the constraints of Eq. (2)

```

2.4. Application to EEG data

2.4.1. Overview

As in most of the EEG classification studies, the classifier was trained for each subject separately to avoid the numer-

ous sources of inter-subject variability to impair the performance of the method. For each subject, using the previously defined features, we compared two approaches for classifying scalp EEGs recorded with 60 electrodes.

- (1) Stacking algorithm (STACK): The stacking algorithm is implemented with a 10-fold split ($J = 10$) and six classifiers dedicated to extracting relevant information from each of the six following feature subsets (see Fig. 1):
 - θ synchrony: 1770 features (corresponding to the number of electrode pairs $\frac{60 \times (60-1)}{2}$)
 - α synchrony: 1770 features
 - β synchrony: 1770 features
 - θ power: 60 features (one by electrode)
 - α power: 60 features
 - β power: 60 features
 We then compare the stacking classifier output, $\hat{y} = \text{sign}(\sum_n \beta_n d_n(\mathbf{x}))$, to the real performance class of the subject obtained from reaction time measurements: $y = +1$ for “high performance” and $y = -1$ for “low performance” (see Section 2.5.3 for explanations).
- (2) Simple SVM algorithm (SVM): For comparison, we applied the classical SVM algorithm to the whole $3 \times 1770 + 3 \times 60 = 5490$ features vector, \mathbf{x} , giving a decision function that mixes the various power and synchrony measurements.

These two methods were implemented on each subject separately using a specific validation procedure explained in the following part.

2.4.2. Validation

To quantify the efficacy of our approach, we computed classification accuracy, which is the rate of good classified points in the test set. As estimating classification accuracy using simple cross-validation had a high variance across the subjects (probably due to the size of the dataset being small and the number of variables used for classification being high), we used a bootstrap algorithm to better estimate the accuracy of the classifiers on the dataset. We quantified the accuracy of the STACK algorithm by randomly splitting the temporally ordered time-windows into ten subsets of equal size using a circular permutation with a random shift, at each bootstrap iteration. Weights and SVM’s discriminant functions were then estimated by applying Algorithm 1 on nine of the ten subsets and the classification accuracy was calculated by testing the resulting stacking classifier on the remaining subset (taking care to discard the time-windows overlapping with windows from other subsets). We generated 100 estimates of classification accuracy by carrying out this procedure at each bootstrap iteration and leaving out one time in each subset. The mean of the 100 estimates was the final estimate of classifier accuracy. This procedure is summarized in Algorithm 2. We used the same type of bootstrap algorithm to estimate the classification accuracy of the SVM algorithm.

At each iteration of this procedure, the learning set had approximately 80 elements and the test set had 10 elements. From classification accuracy, information transfer rate (ITR) can also be computed. This concept, first introduced in communication theory (Shannon and Weaver, 1964), measures the quantity of information per units of time carried by a random process. For a two class problem, if p is the classification accuracy and T the time interval between two decisions of the classifier, ITR is given by $ITR = \frac{1}{T} (1 + p \log p + (1 - p) \log(1 - p))$. It is currently used in BCI to allow comparison between classification methods which are based on different experimental paradigms (Krepki et al., 2006).

Algorithm 2. Bootstrap for the estimation of classification accuracy for STACK algorithm

```

for  $i \rightarrow 1$  to 10
   $D_i \leftarrow$  apply a random permutation to dataset  $D$ 

```

```

   $(D_i^1, \dots, D_i^j) \leftarrow$  split the dataset into 10-folds of equal size.

```

```

  for  $j \rightarrow 1$  to 10

```

```

     $(\beta^*, d_n^L) \leftarrow$  apply Algorithm 1 on the learning set  $D_i \setminus D_i^j$ 

```

```

     $p(i, j) \leftarrow$  compute good classified percentage of stacking classifier with parameters  $(\beta^*, d_n^L)$  on  $D_i^j$ 

```

```

  end for

```

```

end for

```

```

Return mean ( $p$ )

```

We insist on the fact that at each iteration of this bootstrap algorithm, *the classifier is applied to a test set that was not in the learning set of the stacking algorithm*. Indeed, the 10-fold cross-validation procedure of stacking takes a total learning set composed of 9 out of 10 subsets of the bootstrap procedure, and the final test is done on the discarded subset (see Fig. 2).

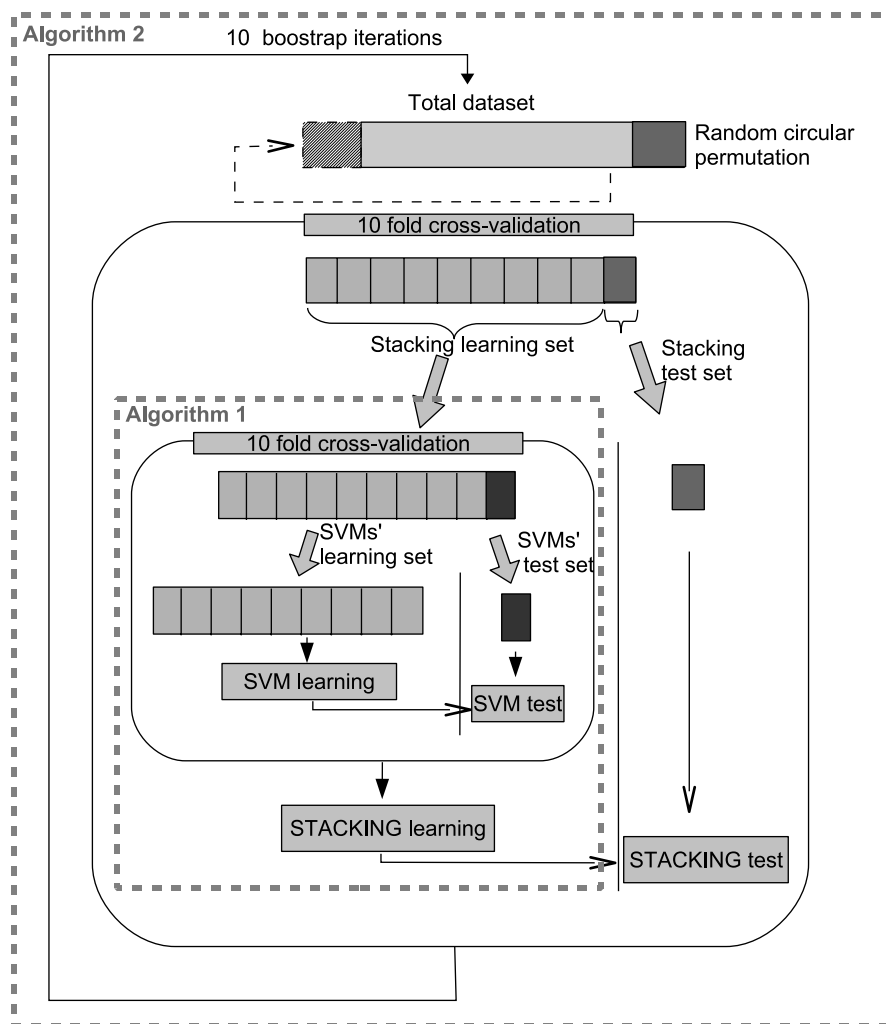


Fig. 2. Identification of learning and test set at each step of the bootstrap estimation of classification accuracy of the stacking algorithm. It is noteworthy that, using 10-fold validation procedures, each classifier (the six SVMs and the stacking) is tested on data that do not belong to the learning set. Moreover, to provide more stable results, the validation process is embedded in a bootstrap procedure, implementing at each step a circular permutation of the data points.

2.5. Data acquisition

2.5.1. Experiment

The experimental protocol was designed to study mental fatigue induced by cognitive tasks. Each subject performed three experimental cognitive tasks successively during one day: a Stroop color word test (St), a Sternberg-derived test (Sb), and a spatial stimulus–response compatibility task (SCT) (Ragot and Renault, 1981). The experiment was organized in two recording periods: morning and afternoon, which were themselves divided into three experimental sessions (one per task). Each St and SCT session consisted of two successive 10 min of recording with a short resting period of one minute between the two. Each Sternberg-like task consisted of one 20-min session. Each subject was seated in a chair one meter from the screen. They had been previously trained to carry out the tasks. In this paper, we describe the results from SCT recordings. During this task, an arrow appeared about every second for 700 ms on the left or right side of the screen. The subjects were asked to respond as fast as they could to the direction of the arrow (pointing left or right) on the screen ignoring the location of the arrow on the screen. Reaction times and errors were recorded during the whole experiment to give quantitative information about the ongoing level of performance.

2.5.2. EEG recordings and pre-processing

The EEG data were recorded in 15 subjects using a 60-electrode ears-referenced BrainCap™. Electrodes were placed according to the standard 10-20 system with AFz used as ground. Data were amplified using a BrainAmps™

(Brain Products, Inc) 64-channel system sampled at 1000 Hz (including trigger signal) and band-pass filtered between 0.3 and 300 Hz. Eye blinks were recorded using a differential montage involving electrode Fp1 and an electrode below the left eye. A common average reference was used off-line to enhance the oscillatory patterns (McFarland et al., 1997). We visually inspected data for contamination by muscular artifacts, with these segments being manually rejected. We did not use the data from three subjects due to high muscular activity. All algorithmic steps were implemented in the Matlab® (The MathsWorks, Inc.) environment. Ocular artifacts were corrected using principal component analysis (PCA) (Wallstrom et al., 2004) successively on each 2-s segments of EEG data. We eliminated principal components giving a correlation coefficient $R > 0.9$ with EOG activity. The corrected data were then rebuilt from the remaining principal components, low-pass filtered at 25 Hz and down-sampled to 200 Hz.

2.5.3. Data segmentation according to performance

The subjects made few mistakes, so we used the reaction times (RT) to measure the ongoing level of performance. First, each epoch was divided into 20-s long sliding windows (corresponding to about 20 successive stimuli of the SCT test) with a 10-s overlap. The first recording of the morning session was discarded to avoid learning effects. The stimulus response compatibility is known to affect reaction time: compatible stimuli (the position of the arrow matches the direction of the arrow) give faster reaction times than incompatible stimuli (the position of the arrow is opposite to the direction) (Ragot and Renault, 1981). To eliminate this bias, the reaction times of compatible stimuli

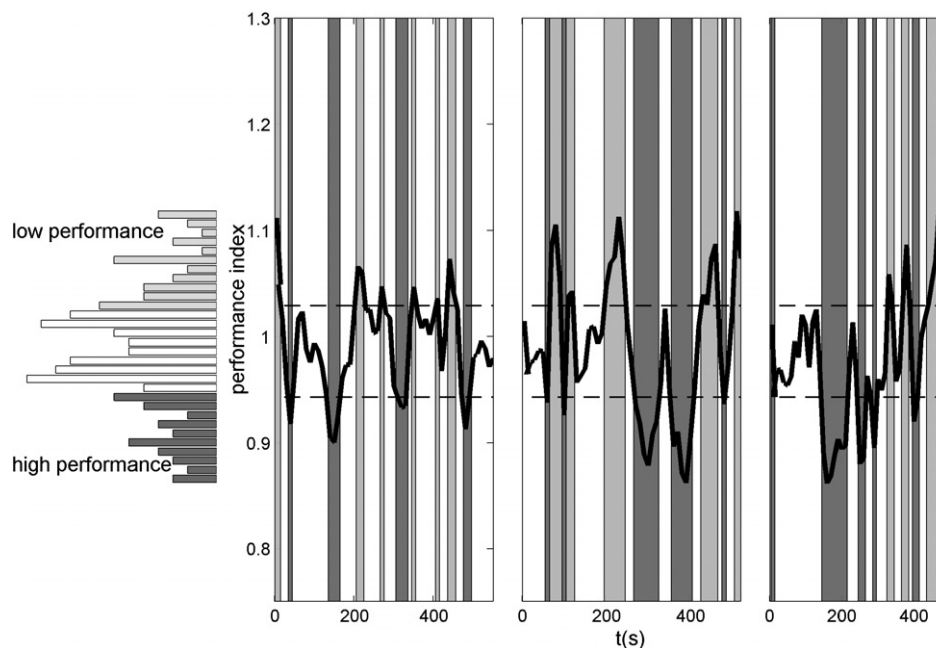


Fig. 3. Performance index (extracted from the reaction time) of one subject during three recordings. The index was calculated using 20-s-long sliding blocks with a 10-s overlap. On the left, histogram of performance index through the sessions with the upper quartile in light gray and the lower quartile in dark gray. On the right, normalized performance index associated with each sliding window along the three 10-min recordings of one subject in the SCT task. The time-windows in dark gray are labeled “high performance” and the windows in light gray are labeled “low performance”.

and incompatible stimuli were normalized by dividing them by their respective median (more robust with outliers than the mean) over the three remaining 10 min recordings (one in the morning and two in the afternoon). We then calculated the overall average normalized reaction time (across compatible and incompatible stimuli at the same time) over each 20-s time-window. This average was used as an index of performance, p_k , for the k th time-window. The histogram of this performance index was split into four equal quartiles and the two classification classes were defined by the first and last quartile. The blocks belonging to the lowest quartile were labeled as a good performance, $y_k = 1$, and those from the highest quartile were labeled as a bad performance, $y_k = -1$, whereas the remaining blocks were discarded from the classification study. This segmentation process is illustrated in Fig. 3 for one subject, the fluctuation of performance across the session can thus be observed. It is noteworthy that performance does not always decrease during each session. Finally, the whole process results in a final classification dataset of approximately 90 blocks per subjects.

3. Results

The rates of correct classification for the SCT task are shown in Fig. 4. The STACK classification was significantly more accurate than the SVM classification according to a non-parametric test ($p < 0.05$) across the 12 subjects. This was also supported by quantitative analysis of average classification accuracies and ITRs reported in Table 1, with the STACK classification being 5% more accurate than the SVM classification. Moreover, we calculated the recognition rates for the two algorithms ten times using the same data with randomized labels (mixing low and high performances) and found that the mean classification accuracy

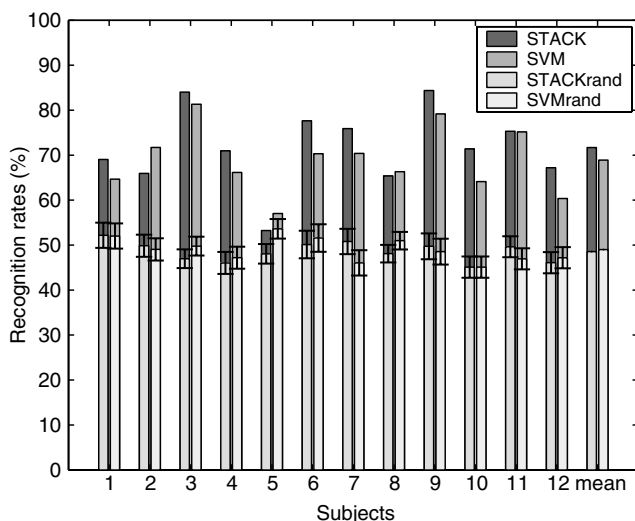


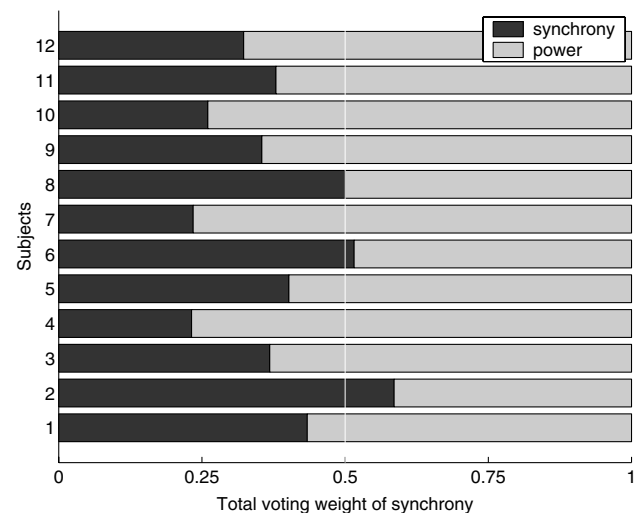
Fig. 4. Recognition rates for each algorithm for each subject performing the SCT task. SVMrand and STACKrand are the mean recognition rates for randomized data. Error bars are standard error around the means.

Table 1

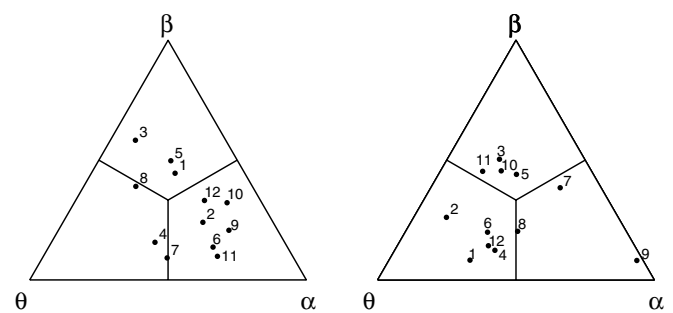
Mean classification accuracies and information transfer rates (ITR) across the subjects for each algorithm

	STACK	SVM	STACKrand	SVMrand
Accuracy (%)	71.70 ± 2.49	68.89 ± 2.10	48.56 ± 2.64	49.00 ± 2.18
ITR (bit/min)	0.495 ± 0.100	0.367 ± 0.078	0.006 ± 0.002	0.006 ± 0.002

for the randomized data was about 50% for the two algorithms (see Fig. 4). We validated by a Wilcoxon signed-rank test for matched samples that original classification results are significantly above this chance level ($p < .01$). Therefore, classification with original data is better than random classification, showing that EEG holds information about the performance level of the subject. The aver-



(a) Sum of the synchronization weights (in black) versus sum of the power weights (in gray) for each subject (one per line).



(b) Barycentric representation of the weights of synchronization features.

(c) Barycentric representation of the weights of power features.

Fig. 5. Weight distribution for algorithm STACK across the subjects: for each subject the repartition was computed using the mean of the weights of the stacking algorithm across bootstrap and cross validation iterations. Graph (a) depicts the proportions of synchronization weights and power weights for each subject. Graphs (b) and (c) are barycentric representations of the weights of frequency bands for, respectively, synchrony and power features: each weight corresponding to a frequency band is affected to a vertex of the triangle and the resulting barycenter of the stacking weights is plotted as a dot for each subject.

age confusion matrix for algorithm STACK is also reported in Table 2; the error rate is nearly the same for the two classes. The weight distributions (through frequencies and measurement types) of the stacking algorithm (Fig. 5) show that power features have greater weights than

synchronization features, although the contribution due to the synchronization weights cannot be neglected (they account for at least 30%). For the synchrony features, alpha activity dominates, whereas for spectral power, theta activity gives the biggest weights.

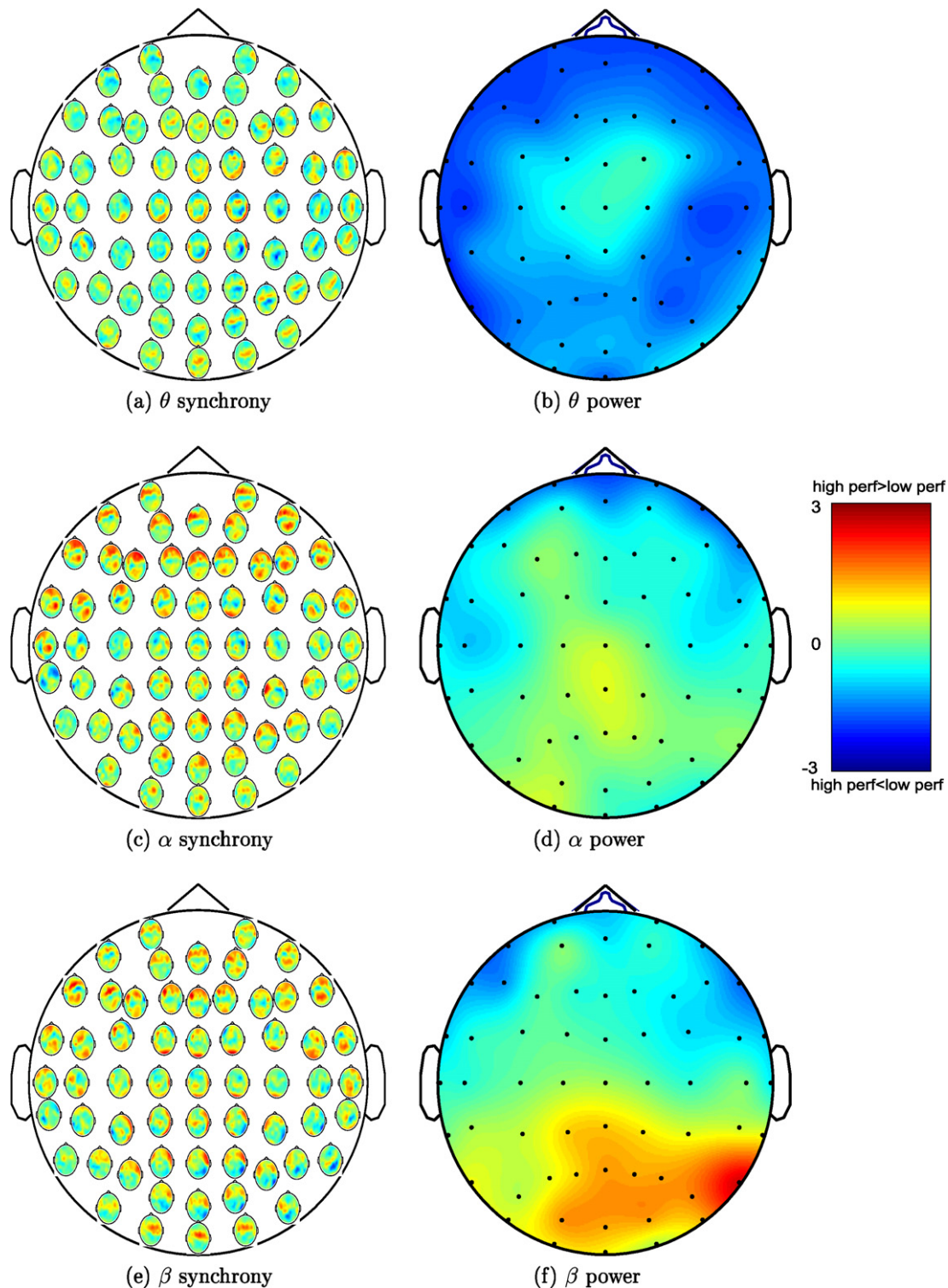


Fig. 6. Distribution of the feature's mean t -value across the seven best subjects (see 3). Scale: hot colors when high performance is greater than low performance, cold colors for the contrary (on the right column the map at each electrode represents synchrony features between this electrode and the rest of the scalp, following (Alba et al., 2007)).

Table 2
Confusion matrix of the STACK algorithm (in percents of the actual class)

Predicted \ Actual	High performance	Low performance
High performance	68.84% ± 2.7	31.16% ± 2.7
Low performance	27.36% ± 2.8	72.64% ± 2.8

We assessed the electrode position of the discriminating variables associated with performance using the Student's *t*-test on each feature to determine the ability of each position to discriminate between “low performance” and “high performance” conditions. The test was carried out on the seven best subjects. The scalp distribution of the mean *t*-value across subjects is shown in Fig. 6 for each features subset. The distribution of the feature sets shows that the discriminating theta power features are widespread on the scalp except on the central region, whereas the discriminating alpha and beta power features are almost parieto-occipital. We noticed signals from peripheral electrodes in the beta power distribution, which may have been due to residual muscle artifacts. The synchrony maps show, apart from local synchrony, discriminating long-distance interactions between fronto-central regions and the rest of the scalp for theta synchrony, whereas the most discriminating alpha and beta synchronies are between parieto-occipital and frontal areas (except for some possibly artifactual short distance interactions located on the temporal lobes). According to the spatial maps presented here, using only a subset of scalp electrodes could be possible, keeping for example one frontal, one parietal and one occipital electrode. To determine the common features of the classifiers, we assessed the reproducibility across the subjects of the electrode labels that have a high weighting coefficient in the vector ω of the SVM decision function. To this end for each subject and each subset of features, we listed the coefficients that belong to the fourth quartile of the coefficients distribution. The labels of the five most reproducible coefficients across the seven best subjects are depicted in Table 3. According to this table, some coefficients have a good reproducibility across the subjects (see the number of subjects in parentheses). This is especially the case for synchrony measures for which some of them are always high across six subjects over seven. Curiously, the coefficients of power features have shown less reproducibility across

the subjects, probably because these coefficients are largely distributed over the scalp. As already shown in Fig. 6, the synchrony features of Table 3 exhibit short-range interactions (FC1–FC3) that can be due to volume conduction, but also long range synchronies reflecting antero–posterior interactions in the alpha and beta bands (P7–AF7). For the power characteristics, the most reproducible coefficients are mainly situated in the whole frontal lobe and also on the left posterior regions.

The same methodology was applied for the Stroop task, recorded on the same experiment, leading to 65% good classification accuracy across the subjects. The lower accuracy can be explained by the difficulty to estimate performance from the reaction time: the inter-stimuli interval was longer (only 10 responses are available to compute the performance index on a 20 s time-window). Thus, for such tasks, it will be useful to consider longer time-windows to compute the classification or adopt another strategy to label periods of low and high performance.

4. Discussion

A general analysis of the classification results reveals a good discrimination between low and high performance states during a SCT task, with the accuracy of classification being greater than 75% for four subjects. However, these results are not as good as those obtained in vigilance state classification (Kiyimik et al., 2004), which gives almost 95%, and motor imagery classification (Pfurtscheller et al., 1997; Muller-Gerking et al., 1999), which gives between 80% and 90%. This is probably because the data used in the vigilance state classification are segmented by an expert, meaning that they contain recognizable EEG oscillations. By contrast, our approach tends to classify according to an objective measure of performance even if there is no visible EEG signature. Moreover, unlike the motor imagery classification used in brain–computer interfaces, our classification approach is asynchronous in that the quantifications do not require knowledge of the external stimuli and their onset. It should also be noticed that these classification results are obtained with few training data (80 time-windows in the training set). Classification accuracy could then be increased with more recordings from each subject. Moreover, the obtained average accuracy (71.7%) has also to be interpreted in terms of informa-

Table 3
Common features across subjects in the classification decision function

θ Synchrony	α Synchrony	β Synchrony	θ Power	α Power	β Power
FC1–FC3(6)	FT8–FC5(6)	C3–Fp2(6)	F8(5)	Fp1(5)	F8(4)
T8–C1(6)	C2–FT8(6)	T8–AF7(6)	FC1(5)	P7(4)	FPZ(4)
PO7–CPz(6)	PO3–C5(6)	P7–AF7(6)	P8(5)	O1(4)	T7(4)
F7–AF7(5)	C4–C2(6)	TP7–F5(6)	PZ(4)	AF4(3)	C5(4)
FC3–AF7(5)	CPz–CPI(6)	CPz–FC3(6)	AF3(3)	AF8(3)	FC6(3)

Only the coefficients of the SVM decision function that belong to the fourth quartile of the coefficients distribution for each subject have been selected. For each one, its recurrence across the seven best subjects is computed, and the five most recurrent labels are listed. The results are given for each of the six feature subsets. For each label selected, the corresponding number of subjects is indicated in parentheses.

tion transfer rate. Our present results correspond to an average ITR of 0.49 bit/min. To give an order of magnitude, it is equivalent to the one of a classifier on 2 min-long time windows giving 99.8% of classification accuracy. A reliable classifier can thus potentially be achieved with a detection delay of a few minutes, which may be fast enough to prevent serious mistakes.

For future applications, it could be useful to distinguish classification errors regarding the high performance (false positives) states and low performance states (false negatives). The balance between false positive rate and false negative rate observed in Table 2 results from the SVM cost function that put an equal weight on the errors in each class. In future studies, focusing on real world applications may require the modification of this balance to allow a maximal limit on the rate of false negatives. This could also be done by an appropriate tuning of the bias constant b of each SVM classifier after the training.

Our results also show that combining multiple SVM by stacking outperforms a single SVM algorithm. There may be two reasons for that. First, the split of the set of features in six subsets reduces the dimension of the features' space, possibly improving the learning process of each SVM. Second, the voting algorithm limits the influence of noisy or uninformative feature subsets that impair SVM performance. Thus the stacking method, initially developed to combine different learning algorithms, also seems to be useful for combining different measurement types. Moreover, the use of weighting in the stacking allows a retroactive analysis of the weights determining the relative importance of each feature subset for the classification.

We focused our present study on a spatial compatibility task. This paradigm has been used in the literature to study the interference effect caused by stimulus-response incompatibility, sometimes called the Simon effect (Peterson et al., 2002; Fan et al., 2002; Liu et al., 2004). In this paper, we have exploited this task to ask a different question: whether we can detect the level of performance of subjects on such a cognitive task from single trial EEG analysis. The predominance of theta in spectral power measurements is consistent with previous studies in related fields because theta oscillations give the clearest results for mental workload (Smith et al., 2001) and cognitive fatigue (Smit et al., 2004). The fact that theta power is lower during high performance compared to low performance can be interpreted as a stronger inhibition of long-term memory networks (Klimesch et al., 2006) which may improve the processing of external stimuli. Interestingly, the lateral distribution of these activities could reflect a hippocampal origin which plays a central role in long-term memory. Moreover, the parietal distribution of alpha power is consistent with results on drowsiness and cognitive fatigue (Tatum et al., 2006; Trejo, 2005). The superiority of spectral power weights over phase synchrony weights is a predictable result, as the large number of electrode pairs used for phase synchrony makes SVM learning more difficult.

Regarding the long-distance synchrony maps, alpha and beta maps are very similar. They both account for an increase of antero-posterior synchrony with better performance. This possibly reflects the activation of a network characterizing visual-attentional control. Such a network has already been reported in the beta band (Gross et al., 2004) and its potential impact on performance in visual tasks is straightforward. This is a paradigmatic example of the added value of synchrony analysis over spectral power analysis, as this phenomenon cannot be observed with the latter approach. Whether this network is task-independent remains an open question; but, if so, this could enable the detection of decreases in visual attention, disregarding its origin (distraction, sleepiness or cognitive fatigue). The predominance of alpha phase synchrony over other frequency bands is difficult to explain, but may be due to synchrony being calculated over long time-windows, thereby discarding transient β synchronies that appear over short time scales.

Our study only investigated the classification of extreme states. This is because intermediate performance levels, which obviously occur between the high and low performance levels, require specific classification or regression methods that have to be validated on a larger dataset. Predicting performance from any EEG time-window requires that these methods be implemented in future work. This can be achieved from longer experiments (several hours of recordings for each subject instead of only 40 min).

This paper presents first results showing that EEG analysis and classification can be applied to performance prediction. As our approach to EEG quantification and classification can be extended to real time algorithms, an on-line EEG-based estimation of performance may be achieved and could provide reliable information with a time constant of a few minutes. Such a device can be useful in various monitoring applications ranging from nuclear plant operators to car and aircraft drivers. Several improvements can be made such as using other behavioral measurement than reaction time. Indeed simulator and virtual reality environments allow quantifying more precisely the performance of a subject, for example by measuring the quality of a pilot's trajectory.

Apart from the potential methodological improvements discussed above, the feasibility of "real world" EEG-based monitoring devices also strongly depends on the future technological improvements of EEG electrodes, already initiated with, for example, the achievement of movement insensitive electrodes. Moreover, it is possible to reduce the number of electrodes. The automatic choice of these latter, using for example variable selection algorithms (Garrett et al., 2003), is out of the topic of this paper but should be the subject of future work.

5. Conclusion

Signal processing and machine learning can be used to discriminate and predict good and bad performance levels

from EEGs, in which there is no prior knowledge about dissociable brain activity patterns. The stacking framework is an intuitive and interpretable approach that can combine many measurements from EEGs to achieve this prediction. It removes the requirement for an EEG expert to analyze the data before classification. We have shown encouraging classification results that may be the first steps toward real-time prediction of cognitive performance states.

Acknowledgments

The authors thank Dr. Frederique Amor and Dr. Mario Chavez for many helpful comments on the original manuscript.

Appendix. Stacking optimization algorithm

The stacking optimization problem of Eq. 1 is of the form

$$\beta^* = \arg \min_{\beta} \sum_{k \in L} \left(\left(\sum_n u_{k,n} \beta_n \right) - y_k \right)^2$$

under the constraints

$$\sum \beta_n = 1 \quad \forall n, \beta_n \geq 0$$

Let us define the matrix $U = (u_{k,n})$. Then the optimization problem can take the matrix form

$$\beta^* = \arg \min_{\beta} (U\beta - y)^T (U\beta - y)$$

Developing the matrix product leads to the final form of a quadratic programming problem

$$\beta^* = \arg \min_{\beta} \beta^T (U^T U) \beta - 2y^T U \beta$$

under the same constraints. This kind of problem can be solved using various algorithms. In this work we used the *quadsolve* algorithm from *The Spider toolbox* (Weston et al., 2003). This is a primal dual method with a predictor–corrector approach (Altman and Gondzio, 1998) which is also used to implement the support vector machine.

References

- Alba A, Marroquin JL, Pena J, Harmony T, Gonzalez-Frankenberger B. Exploration of event-induced EEG phase synchronization patterns in cognitive tasks using a time–frequency–topography visualization system. *J Neurosci Meth* 2007;161:166–82.
- Altman A, Gondzio J. Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization. *Optim Methods Softw* 1999;11:275–302.
- Boksem MA, Meijman TF, Lorist MM. Effects of mental fatigue on attention: an ERP study. *Brain Res Cogn Brain Res* 2005;25(1):107–16.
- Burges CJC. A tutorial on support vector machines for pattern recognition. *Data Min Knowl Discov* 1998;2(2):121–67, [ISSN 1384–5810].
- Cristianini N, Shawe-Taylor J. An introduction to support vector machines. Cambridge University Press; 2000.
- Fan J, Flombaum J, McCandliss BD, Thomas KM, Posner MI. Cognitive and brain consequences of conflict. *Neuroimage* 2002;18:42–57.
- Garrett D, Peterson DA, Anderson CW, Thaut MH. Comparison of linear, non-linear, and feature selection methods for EEG signal classification. *IEEE Trans Neural Syst Rehabil Eng* 2003;11:141–4.
- Grandjean E. *Fitting the task to the man*. Taylor and Francis; 1988.
- Gray C, Singer W. Stimulus specific neuronal oscillations in orientation columns of cat visual cortex. *Proc Natl Acad Sci USA* 1989;86(5):1698–702.
- Gross J, Schmitz F, Schnitzler I, Kessler K, Shapiro K, Hommel B, et al. Modulation of long-range neural synchrony reflects temporal limitations of visual attention in humans. *Proc Natl Acad Sci USA* 2004;101:13050–5.
- Gysels E, Celka P. Phase synchronisation for the recognition of mental tasks in a brain–computer interface. *IEEE Trans Neural Syst Rehabil Eng* 2004;12(4):206–415.
- Gysels E, Renevey P, Celka P. SVM-based recursive feature elimination to compare phase synchronization computed from broadband and narrowband EEG signals in brain–computer interfaces. *Signal Process* 2005;85(11):2178–89.
- Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning*. Springer; 2001.
- Hohnsbein J, Falkenstein M, Hoormann J. Effects of attention and time–pressure on P300 subcomponents and implications for mental workload research. *Biol Psychol* 1995;40(1–2):73–81.
- Jung TP, Makeig S, Stensmo M, Sejnowski TJ. Estimating alertness from the EEG power spectrum. *IEEE Trans Biomed Eng* 1997;44(1):60–9.
- Kiyimik MK, Akin M, Subasi A. Automatic recognition of alertness level by using wavelet transform and artificial neural network. *J Neurosci Meth* 2004;139(2):231–40.
- Klimesch W. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res Brain Res Rev* 1999;29(2–3):169–95.
- Klimesch W, Doppelmayr M, Hanslmayr S. Upper alpha ERD and absolute power: their meaning for memory performance. In: *Event-related dynamics of brain oscillations*, Elsevier; 2006. p 151–163 [chapter 10].
- Krepki R, Curio G, Blankertz B, Muller K. Berlin brain–computer interface. the HCI communication channel discovery. *Int J Hum Comp Studies* 2006;65:460–77.
- Lachaux J, Rodriguez E, Martinerie J, Varela F. Measuring phase synchrony in brain signals. *Hum Brain Mapp* 1999;8:194–208.
- Le Van Quyen M, Soss J, Navarro V, Robertson R, Chavez M, Baulac M, et al. Preictal state identification by synchronization changes in long-term intracranial EEG recordings. *Clin Neurophysiol* 2005;116:559–68.
- LeBlanc M, Tibshirani R. Combining estimates in regression and classification. *J Am Stat Assoc* 1996;91(436):1641–50.
- Lin CT, Wu RC, Jung TP, Liang SF, Huang TY. Estimating driving performance based on EEG spectrum analysis. *EURASIP J Appl Sig P* 2005;19:3165–74.
- Liu X, Banish MT, Jacobson BL, Tanabe JL. Common and distinct neural substrates of attentional control in integrated Simon and spatial Stroop task as assessed by event related fMRI. *Neuroimage* 2004;22:1097–106.
- Lopes da Silva F. Neural mechanisms underlying brain waves: from neural membranes to networks. *Electroencephalogr Clin Neurophysiol* 1991;79(2):81–93.
- Makeig S, Jung TP, Sejnowski TJ. Using feedforward neural networks to monitor alertness from changes in EEG correlation and coherence. *Adv Neur Info Proc Sys* 1996;8:931–7.
- McFarland D, McCane L, Davi S, Wolpaw J. Spatial filter selection for EEG-based communication. *Electroencephalogr Clin Neurophysiol* 1997;103:386–94.
- Muller-Gerking J, Pfurtscheller G, Flyvbjerg H. Designing optimal filters for single-trial EEG classification in a movement task. *Clin Neurophysiol* 1999;110:787–98.

- Muller K, Mika S, Ratsch G, Tsuda K, Scholkopf B. An introduction to kernel-based learning algorithms. *IEEE Tran Neur Netw* 2001;12:181–202.
- Orfanidis SJ. Introduction to signal processing. Prentice-Hall; 1996.
- Peterson BS, Kane MJ, Alexander GM, Lacadie C, Skudlarski P, Hoi-Chung L, et al. An event-related functional MRI study comparing interferences effects in the Simon and Stroop tasks. *Cog Brain Res* 2002;13:427–40.
- Pfurtscheller G, Neuper C, Flotzinger D, Prezenger M. EEG-based discrimination between imagination of right and left hand movement. *Electroencephalogr Clin Neurophysiol* 1997;103:642–51.
- Pikovski A, Rosenblum M, Kurths J. Synchronization: a universal concept in nonlinear science. Cambridge University Press; 2001.
- Ragot R, Renault B. P300 as a function of SR compatibility and motor programming. *Biol Psychol* 1981;13:289–94.
- Rodriguez E, George N, Lachaux J, Martinerie J, Renault B. Perception shadow: long distance gamma band synchronisation and desynchronisation on the human scalp. *Nature* 1999;397:430–3.
- Shannon CE, Weaver W. Mathematical theory of communication. Champaign, IL: University Illinois Press; 1964.
- Smit AS, Eling PA, Coenen AM. Mental effort affects vigilance enduringly: after-effects in EEG and behavior. *Int J Psychophysiol* 2004;53(3):239–43.
- Smith ME, Gevins A, Brown H, Karnik A, Du R. Monitoring task loading with multivariate EEG measures during complex forms of human–computer interaction. *Hum Factors* 2001;43(3):366–80.
- Tallon-Baudry C, Kreiter A, Bertrand O. Sustained and transient oscillatory responses in the gamma and beta bands in a visual short-term memory task in humans. *Vis Neurosci* 1999;16(3):449–59.
- Tatum WO, Husain AM, Benbadis SR, Kaplan PW. Normal adult EEG and patterns of uncertain significance. *J Clin Neurophysiol* 2006;23(3):194–207.
- Trejo LJ. EEG-based estimation of cognitive fatigue. In: Proceedings of SPIE: Bio-monitoring for physiological and cognitive performance during military operations, vol. 5797. 2005.
- Vapnik V. Statistical learning theory. New York: Wiley-Interscience; 1998.
- Varela F. Resonant cell assemblies: a new approach to cognitive function and neuronal synchrony. *Biol Res* 1995;28:81–95.
- Varela F, Lachaux J, Rodriguez E, Martinerie J. The Brainweb: phase synchronisation and large-scale integration. *Nature Rev Neurosci* 2001;2:229–39.
- Wallstrom GL, Kass RE, Miller A, Cohn JF, Fox NA. Automatic correction of ocular artifacts in the EEG: a comparison of regression-based and component-based methods. *Int J Psychophysiol* 2004;53(2):105–19.
- Weston J, Elisseff A, BakIr G, Sinz F. The spider 2003.
- Whittington MA, Traub RD, Kopell N, Ermentrout B, Buhl EH. Inhibition-based rhythms: experimental and mathematical observations on network dynamics. *Int J Psychophysiol* 2000;38(3):315–36.
- Wilson G, Fisher F. Cognitive task classification based on topographic EEG data. *IEEE Trans Biomed Eng* 1998;45(3):277–86.
- Wilson G, Russell C. Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. *Hum Factors* 2003;45(4):635–43.
- Wolpaw J, Birbaumer N, McFarland D, Pfurtscheller G, Vaughan T. Brain–computer interfaces for communication and control. *Clin Neurophysiol* 2002;113:767–91.
- Wolpert DH. Stacked generalization. *Neural Netw* 1992;5(2):241–59.

Annexe B : Classification d'une tâche de contrôle visuo-moteur en MEG

Présentation

L'article suivant comporte une partie revue qui résume ce qui a été dit au chapitres 1 et 2 de ce manuscrit, et une partie résultats où nous avons appliqué nos méthodes de classification sur des enregistrements MEG réalisés au laboratoire par Karim Jerbi (Jerbi et al., 2007). L'objectif est de différencier deux conditions expérimentales : une tâche de repos yeux ouverts et une tâche de contrôle visuo-moteur. Il s'agit en particulier de caractériser l'influence sur le taux de classification de différents paramètres comme les choix des variables de quantification ou la méthode de validation.

Principaux résultats

La quantification et la classification sont réalisées sur des fenêtres glissantes d'une seconde par des mesures de puissance ou de synchronie de phase dans cinq bandes de fréquences. Une comparaison de performance de trois classifieurs (ALD, SVM linéaire et KPPV) met d'abord en évidence les difficultés rencontrées par l'ALD lorsque le nombre de données d'apprentissage est trop faible devant le nombre de variables. L'ensemble des résultats de classification suivants est réalisé avec un SVM linéaire. De manière générale, le taux de bien classés atteint 97% (voir Figure 7 page 28) en utilisant l'ensemble des variables de quantification. Ce résultat est cependant atteint pour une cross-validation à l'intérieur de la même session (validation croisée "10" fold intra-session). Lorsque la validation utilise un ensemble test issu d'une session différente de l'apprentissage (validation croisée inter-session), alors le taux bien classés diminue (voir Figure 5 page 26). Une analyse de la performance est menée séparément pour chaque type de variable dans chaque bande de fréquence considérée, elle montre que puissance et synchronie de phase permettent de bons résultats de classification dans un large domaine de fréquences mais seulement en validation intra-session. En validation inter-session, les taux de classification associés aux variables de synchronies chutent fortement, ainsi que ceux des variables de puissance dans les basses et hautes fréquences correspondant aux bandes

delta et gamma 2.

Il semble donc que certaines mesures soient prédictives à court-terme, mais subissent des fluctuations entre sessions différentes, ce qui pourrait s'expliquer par des variations de la position de la tête du sujet dans le dispositif MEG. Pour la bande delta, cette diminution est peut être aussi due à une difficulté d'estimation de la puissance en basse fréquence. Les oscillations dans la bande gamma sont quand à elles réputées très instables et difficiles à mettre en évidence en essai unique.

Finalement, l'analyse des variables utilisées par le classifieur semble mettre en évidence des activités spécifiques à la tâche dans le cortex moteur primaire, l'aire motrice supplémentaire et les aires visuelles, ainsi que des interactions à distance entre ses régions (voir figure 8 page 29).

Perspectives

Les résultats présentés constituent une étude préliminaire sur les potentialités des méthodes de classification pour les signaux MEG. Le problème de classification traité dans cet article peut être jugé relativement facile puisque'il s'agit de classer une tâche motrice (non imaginée) contre une tâche de repos². Dans une étude plus approfondie nous aborderons des problèmes de discrimination plus difficiles sur cette expérience, comme par exemple la différenciation de la tâche de contrôle visuo-moteur par rapport une tâche motrice équivalente sans retour visuel. Le but recherché sera alors de mettre en évidence l'apport des techniques de classification dans l'étude des corrélats neurophysiologiques de tâches cognitives de haut niveau par rapport aux techniques statistiques usuelles (comme le test t).

Article (soumis à Biological Research)

2. les performances atteintes sur une fenêtre d'une seconde sont cependant remarquables.

Classification methods for ongoing EEG and MEG signals.

MICHEL BESSERVE, KARIM JERBI*, FRANCOIS LAURENT, SYLVAIN

BAILLET, JACQUES MARTINERIE and LINE GARNERO

CNRS UPR 640-LENA, universit  Pierre & Marie Curie-Paris6

and Centre MEG, Paris, FRANCE

ABSTRACT

Classification algorithms help predict the qualitative properties of a subject's mental state by extracting useful information from the highly multivariate non-invasive recordings of its brain activity. In particular, applying them to Magneto-encephalography (MEG) and electro-encephalography (EEG) is a challenging and promising task with prominent practical applications to e.g. Brain Computer Interface (BCI). In this paper, we first review the principles of the major classification techniques and discuss their application to MEG and EEG data classification. Next, we investigate the behavior of classification methods using real data recorded during a MEG visuomotor experiment. In particular, we study the influence of the classification algorithm, of the quantitative functional variables used in this classifier, and of the validation method. In addition, our findings suggest that by investigating the distribution of classifier coefficients, it is possible to infer knowledge and construct functional interpretations of the underlying neural mechanisms of the performed tasks. Finally, the promising results reported here (up to 97% classification accuracy on 1-second time windows) reflect the considerable potential of MEG for the continuous classification of mental states.

Key terms: brain computer interface, electroencephalography, magnetoencephalography, visuomotor control, Support Vector Machine.

*Currently in post-doctoral position at INSERM U821, Lyon, France and LPPA, coll ge de France, Paris, France.

1. Introduction

The non-invasive detection of task-related neurophysiological changes occurring within the human brain is a significant challenge in biomedical engineering. There is a growing interest in using classification techniques to estimate the mental state of a subject, related to a performed task, from multivariate brain functional imaging signals such as functional magnetic resonance imaging (fMRI) (Carlson et al., 2003; Laconte et al., 2006; Haynes and Rees, 2006) or magneto- or electroencephalography (MEG/EEG) (Kubler et al., 2001; Wolpaw et al., 2002; Lal et al., 2005; Lotte et al., 2007) from ongoing, single-trial epochs. In principle, the related algorithms are statistical tools that can be trained to estimate a qualitative variable, the *class label*, from a set of quantitative variables. Therefore, classification techniques aim at predicting the qualitative mental state of the human brain and come in addition to the estimation of functional brain responses per se.

There are two major reasons why these methods should be considered: first, classification tools let us evaluate the predictive power of functional signals and thus quantify the information they convey about the tasks being performed. Classification methods come in complement to massively univariate statistical tools like Statistical Parametric Mapping (Frackowiak et al., 1997) by mining the information contents of multidimensional signals. This approach is more powerful than the univariate statistical tests since the former seeks information in many variables at the same time whereas the latter can only process each variable independently (Carlson et al., 2003). The second reason why classification tools are pertinent to neuroscience is their potential applications to real-time prediction of brain states. Their most popular application is undoubtedly Brain Computer Interfaces (BCI) which has witnessed considerable progress during the past decade (Kubler et al., 2001, Wolpaw et al. 2002; Lebedev and Nicolelis, 2006). A further promising real-time application is *neurofeedback* for the treatment of motor or cognitive dysfunctions such as Attention-Deficit/Hyperactivity Disorders (ADHD) (Fuchs 2003, Strehl 2006).

For some of these applications, exploiting directly the invasive recordings of neural populations can be more efficient than using fMRI, EEG or MEG. In particular, invasive neural interfaces have known considerable progress (Taylor et al., 2002; Wessberg and Nicolelis, 2004); but a lot of work remains to be done to achieve a safe and stable invasive BCI (Lebedev and Nicolelis, 2006). Thus, even if they provide a slower information transfer, devices using non-invasive imaging modalities are still competitive for this application. Among these imaging techniques, Magnetoencephalography (MEG) and electroencephalography (EEG) provide the highest temporal resolution making these techniques ideal for the aforementioned real-time applications. Several parameters, however, need to be carefully taken into account when using classification methods with magneto-electroencephalographic (MEEG) data. First, the choice and design of the experimental paradigm is an important factor that will partially determine subsequent data processing procedures. The most commonly used paradigms involve stimulus-locked subject responses where the spatial and temporal properties of brain responses linked to a stimulus are investigated and referenced in direct relationship to the stimulus onset. In the context of BCI, a synchronous set-up is directed toward evoked brain responses triggered by an external stimulus. By contrast, asynchronous BCI set-ups do not require any external stimulation to infer mental states. Therefore, asynchronous approaches are challenging because they are expected to yield a less tedious and more natural and efficient communication device. In general, such systems operate with data acquired while subjects alternate between two (or

more) sustained mental tasks. A major difference between continuous data acquisition paradigms and classical stimulus-locked experiments is that the task is maintained on a longer time span (typically tens of seconds). In the following, we will refer to this type of paradigm as a *continuous* experiment.

The two types of experimental paradigms, stimulus-locked and continuous, involve different types of data processing and investigate different quantitative measures estimated from the signals. Nevertheless, it is noteworthy that while some techniques like, the analysis of evoked potentials, are only useful for synchronous BCI applications (Farwell and Donchin, 1988; Wolpaw et al., 2002), other task-related phenomena, like sustained oscillatory activities, can be used both with synchronous and asynchronous paradigms (Keirn et al., 1990; Anderson et al., 1998; Pfurtscheller et al. 1997; Borisoff et al., 2004; Milan et al., 2004; Scherer et al., 2004). To elucidate these task-related brain rhythms, spectral analysis has become a standard procedure with two measurements: (a) power spectral density, (which provides an estimation of signal amplitude at different frequencies or within various frequency bands) and (b) long-range signal correlation in the frequency domain (which can be measured via coherence or phase synchrony). Whether they are performed at the sensor or at the cortical level, power estimations are assumed to reflect the synchronization of local neural populations (Singer et al. 1999) while the long-range signal coupling represents long-distance interactions between two signals recorded in distinct brain regions.

The primary objective of this paper is to provide an overview of the major principles of classification approaches applied to EEG and MEG data, and secondly to investigate the practical behavior of such techniques with real MEG data acquired during a continuous visuomotor task. The results shed some light on important issues such as the selection of a classifier and the evaluation of its performances, as well as the identification of physiologically relevant discrimination features for continuous sensorimotor MEEG data.

The layout of the paper is as follows: In Section 2, we review the general principles of classification methods by considering different classification schemes and various measures generally extracted from MEEG recordings for the purpose of classification. In Section 3, we report on the application of classification tools to MEG data recorded during a continuous visuomotor task and we conclude by discussing the utility, interpretation and limitations of the results obtained.

2. Using classification methods for MEEG

2.1 The different steps

Applying statistical analysis tools to MEEG signals usually requires temporal segmentation of the recordings according to precise events of the experimental paradigm. For stimulus-locked paradigms, segmentation bounds are set according to the stimulus onset. In particular, a reference time window preceding the stimulus onset is usually taken as a reference epoch (or baseline) to which the temporal evolution of the cortical signals after stimulus onset is compared. Obviously, this scheme can be temporally reversed if the reference point is the subject's response instead of a stimulus (as in movement preparation or anticipation paradigms). Conversely, if continuous paradigms may contain specific events (for example the beginning of the task), they do not systematically impose a temporal ordering of the whole recording. In this case, if we assume that signal properties are largely invariant over the duration of a given continuous behavioral condition, a long-lasting epoch corresponding to the continuous task can be segmented into successive time windows. This segmentation

provides a large number of time windows which can then be processed as multiple trials of the same experimental condition (although all reference-free).

To characterize brain activities from recordings, several features are then computed from the segmented data (details in next section). These features allow representing each segment as a point in a normed vector space, whose entries are feature values. Therefore, by applying this quantification procedure to multiple trials, we obtain a set of distributed points that may form several ‘clouds’ in the multi-dimensional feature space (Fig. 1). Each cloud of points corresponding to the data of a specific behavioral condition represents a *class*.

The next step of the classification procedure consists in exploiting a first dataset as a “training set” to learn to partition the feature space in a way that optimally distinguishes the various classes from each other. Most of the widely used techniques (described below) fit a separation surface, called the *decision boundary*, between domains corresponding to each class. Once the classifier is “trained”, new trials can then be classified as belonging to one class or another according to their position in the feature space. In this study, we will restrict the discussion to *binary classification* where only two classes have to be distinguished. In this case, the decision boundary is usually determined using a *discriminant function* (Duda and Hart, 2000) $f^*(x)$ with the following rule: if $f^*(x) > 0$ then the trial associated with feature vector x is classified in class 1, else the trial is attributed to class 2. We will briefly present a few classification methods in addition to common problems that arise from these approaches. A schematic overview of the MEEG classification process in the binary classification case of a continuous paradigm is depicted in Figure 1.

2.2. Characterizing brain activity from MEEG measurements

To quantify task-related cortical activations from recordings, the ongoing signals are usually segmented into time windows on which various computations are performed. Most of the estimated parameters can be divided into families depending on the spatial extent of the physiological phenomenon under investigation: local or long range. Local measurements generally provide a measure of task-related activity picked up at a single sensor or electrode. This measure is assumed to reflect the modulation of neural activity at a focal brain area located next to the detector. By contrast, measurements of long distance interactions quantify the coupling between signals detected at two distinct sensors, possibly, although not necessarily, revealing an information transfer between two distant neural ensembles. The physiological interpretation of such long-range phenomena is the focus of a large body of literature ranging from realistic models of neural networks and experimental evidence obtained using both invasive and non-invasive recordings, both in human and non-human primates (Rodriguez et al., 1999; Varela et al., 2001; Brovelli et al., 2004.; Jerbi et al., 2007; Lachaux et al., 1999).

Given that the electromagnetic activity recorded on the scalp is an indirect and attenuated signature of the underlying neural processes, the main difference between MEEG and invasive recording techniques, is that the task-related effects should be sufficiently strong if they are to be observed non-invasively on the scalp surface. The strength of these signals presumably depends on various parameters including the Signal-to-Noise ratio of the system, the size of the neural populations involved, the experimental paradigm and possibly the alertness or motivation of the subject. To help detecting these signals, a lot of signal processing techniques have been exploited in non-invasive BCI systems (Bashashati et al., 2007); but most of them exploit the same physiological phenomena and similar measures to those that will be reviewed in the following sections.

Local activity measurements

Given that stimulus locked experimental paradigms have been widely used both for basic cognitive studies and for BCI applications a large range of analysis tools have been tested and developed over the years to specifically analyze event related activities. The best-studied family of event related activities is Event Related Potentials (ERP). Even if, to date, the physiological processes underlying ERP generation remain debated, many types of ERP induced by specific stimulation set-ups are robustly observed and can be exploited. For instance, a virtual keyboard, called the *P300 speller*, has been developed based on a specific ERP arising after the onset of an attended stimulus (Barett, 1996). Due to baseline fluctuations, ERP can only be detected by averaging data obtained over multiple trials. Although, in theory, this requirement makes ERP unusable for classification of single trial MEG activities, this problem has been solved in the *P300 speller* paradigm by cumulating multiple trials before classification (Farwell and Donchin, 1988).

A further quantification of local modulations is given by measures known as Event Related Desynchronization (ERD) and Synchronization (ERS) (Pfurtscheller and Lopes da Silva, 1999). ERD (or ERS) represent a decrease (or an increase) of oscillatory power at specific frequencies and in specific locations as compared to the baseline power estimated during a reference period (i.e. a time window prior to stimulus onset or alternatively an epoch during a control condition). Many studies report the presence of task specific rhythms in MEG and EEG recordings; In particular, in the 8-12Hz band occipital alpha rhythms are implied in vision whereas mu rhythms (~11 Hz) are present in the motor cortex in resting states and are suppressed during the execution of a movement (Salmelin and Hari, 1994). In higher frequency bands, beta rhythms (15-30 Hz) have been shown to be suppressed during movement and to display a strong increase roughly 1-2 seconds after the movement is terminated (a phenomenon known as *beta rebound*). Finally, gamma band power modulations (30-90 Hz) seem to play an important role in wide range of complex cognitive processes (Tallon-Baudry, 1999). Historically, the presence of ERS and ERD in specific frequency bands are interpreted as an activation or deactivation of the neural assemblies, but this theory is questioned according to recent studies: the ERD/ERS phenomena are highly dependent of the task and the brain area (Lopes Da Silva, 2006) and rhythms in the same frequency band can both correspond to deactivation or activation of the underlying network (Pfurtscheller, 2006). ERD in the alpha and beta bands appearing in the motor areas have been extensively used to classify motor imagery tasks (Pfurtscheller, 1997). The usual way to detect ERD/ERS is to make a baseline normalization and then to compute inter trial averages (Graumann and Pfurtscheller, 2006); however studies in BCI protocol have demonstrated their measurability in single trial (Pfurtscheller, 97).

For continuous paradigms, the starting time of the mental process is either not used or not available, which makes ERP too difficult to detect. Moreover the transient nature of ERP is not compatible with the characterization of continuous states. Besides, even though ERD/ERS is usually referenced to a baseline, this phenomenon can be equally considered as induced by an event or as task related activation. Hence a continuous task can also be characterized by ERD/ERS measurements. Technically speaking, the alternative to stimulus-locked ERD/ERS when processing continuous brain activities is to cut original recordings into sliding windows and perform the spectral power measurements on these windows in multiple frequency bands.

The usual tools to extract frequency information from EEG and MEG signals are Fourier transform and band pass filtering. In this case, the feature considered is the spectral power in different frequency bands f obtained by calculating the average power of the filtered signal $z_i^f(t)$ of channel i during each time window of length T :

$$P_i(f) = \frac{1}{T} \sum_{t=1}^T |z_i^f(t)|^2 \quad (1)$$

These tools can be seen as non parametric spectral analysis methods, as no assumption is done on the power spectrum of the signals. Another method, which has been proved to be efficient for EEG classification, is autoregressive modeling (Anderson, 1998). This is a parametric approach as it assumes that the power spectrum of the signals is described by a few coefficients (typically 3 to 10).

Spectral measurements are also a convenient tool to detect Steady State Visual Evoked Potentials (SSVEP) induced by repetitive visual stimulations in new BCI paradigms (Middendorf et al. 2000).

Long distance interactions

It is reasonable to assume that any mechanism for brain integration must involve interactions between the participating local neural ensembles (Varela et al., 2001). Quantifying these interactions has been investigated via multiple techniques that quantify the presence of common information in the signals that originate from different brain areas. Assessing this functional relationship between distant signals has been named “functional connectivity” by contrast to anatomical connectivity. Functional connectivity investigates the structure of the brain network at a particular time sample and is supposed to characterize the cooperation between many specialized regions. The most widely used measures for long-range interactions quantification are coherence and phase synchrony (Lachaux et al. 1999).

Coherence measures the interaction of two signals in a particular frequency band by computing the correlation between their Fourier transforms $Z_1(f)$ and $Z_2(f)$ of two signals at this frequency.

$$Coh_{12}(f) = \frac{Z_1(f)Z_2(f)^*}{|Z_1(f)||Z_2(f)|}$$

Thus, the resulting quantity can both reflect the correlation between spectral power of the signal (i.e. amplitude) and a temporal synchronization (phase relationship) of the signals.

By contrast, phase synchrony concentrates on assessing the common phase information between two signals. It is calculated by converting the filtered data of each channel i in the frequency band f , $z_i^f(t)$, into an analytical signal $a_i^f(t)$ by the Hilbert transform H with $a_i^f(t) = z_i^f(t) + H[z_i^f(t)]$. This allows to define the instantaneous phase $\varphi_i^f(t)$ of the filtered signal (Pikovsky et al., 2001) by the modulus-argument decomposition $a_i^f(t) = |a_i^f(t)|e^{i(\varphi_i^f(t))}$, where \mathbf{i} is the complex number of modulus 1 and argument $\frac{\pi}{2}$. The phase locking value between channels i and j in a given frequency band f is thus:

$$S_{ij}(f) = \frac{1}{T} \left| \sum_{t=1}^T e^{i(\varphi_j^f(t) - \varphi_i^f(t))} \right|, \quad (2)$$

In a previous study (Rodriguez et al., 1999), we estimated phase synchrony on surface EEG recordings performed during a visual task. Task-related phase synchrony modulations between widely separated electrodes showed significantly different patterns between "face perception" and "no face perception". Additionally, in an intra-cerebral study (Le Van Quyen et al., 2005) we provided evidence for the use of phase synchrony to anticipate epileptic seizures. Moreover, recent findings in the field of Brain-Computer Interfaces suggest that

combining spectral power measurements with phase synchrony improves accuracy when classifying EEG signals (Gysels et al., 2004,2005; Wei et al., 2007).

In the present study, we also investigate the combination of power and phase synchrony measures for classification purposes applied to MEG visuomotor data (details in Section 3). An argument for using both these tools is that they can be considered within the same biological framework of "Resonant Cell Assemblies" (Varela, 1995). In recent years, it has become evident that the concept of neuronal synchrony can improve our characterization and our understanding of the global aspect of brain dynamics. Indeed neuroscience has provided abundant evidence of synchronization at all levels of the nervous system (Varela et al. 2001), ranging from individual pairs of neurons to larger scales within or in between different local neural assemblies. Generally, oscillations are thought to reflect local synchronization of such assemblies, the underlying principle being that it is the phase locked neural activity that predominantly gives rise to measurable scalp oscillations. By contrast, the long range phase synchronization between two electrodes is assumed to be a signature of the phase locked activity between distinct brain regions. Therefore these two measures are complementary in characterizing brain activity and essential for the global integration hypothesis.

2.3. Classification algorithms

Data classification is an issue of primary importance in the field of data mining: it consists in estimating a qualitative variable the *class label* using a set of other variables. Numerous algorithms and methods have been proposed in order to achieve data classification and to improve its efficiency. Most standard classification methods can be broadly described as a two-step procedure which consists first of a learning phase followed by actual estimation of unknown class labels. During the learning phase, a discriminant function is fitted to a portion of the data generally called the training data set and then, in the second phase, the trained model (achieving optimized separation on the training set) is used to discriminate between the classes from new data sets. Two main considerations can distinguish classification algorithms:

- Their decision boundary can be either linear or non-linear
- The fitted discriminant function can result from a probabilistic model (model-based or generative algorithms) or be expressed directly using the data points from the training set (data-based or discriminative algorithms).

We will present three algorithms illustrating these fundamental characteristics.

Linear Discriminant Analysis (LDA)

Linear classification algorithms, which represent a large portion of the available techniques, are based on fitting a linear discriminant function to the data. This linear decision is of the form $f(x) = wx + b$ where b is the bias and w the normal vector to the decision boundary $f(x) = 0$. One approach to fit this linear function to the data is Linear Discriminant Analysis which amounts to fitting a Gaussian probabilistic model to the data (Duda and Hart, 2001). In its simplest formulation, assuming the feature vector x has a Gaussian distribution with different class conditional means μ_1 and μ_2 for class 1 and 2 respectively and the same covariance matrix $\Sigma_1 = \Sigma_2 = \Sigma$ in the two cases, the optimal decision function in a Bayesian framework is of the form:

$$f_{LDA}^*(\mathbf{x}) = (\mu_1 - \mu_2)^T \Sigma^{-1} \left(\mathbf{x} - \frac{\mu_1 + \mu_2}{2} \right) = w_{LDA} x + b_{LDA}$$

An illustration of this technique on toy data is provided Fig. 2 a), showing the affine decision boundary between the two clouds of points of the training set. LDA is an interesting algorithm for many reasons: it is fast, simple to implement and to understand. As a consequence it has been widely used in the BCI community (Garett 2003, Bostanov 2004, Scherer 2004).

Linear Support Vector Machine (SVM)

Contrary to LDA, that constructs a probabilistic model for each class using all data points, the underlying principle of linear Support Vector Machines (SVM) is to minimize a linear separation error focusing on neighboring points of the linear separation surface (Vapnik, 1998). Points that lie sufficiently far from the separation surface are thus ignored in the learning process. More precisely, the linear discriminant function has the same form as in LDA : $f(x) = wx + b$; moreover, we denote by (x_i, y_i) the i -th data point of the training set of feature vector x_i and of class label y_i ($y_i = +1$ if the point belongs to state or class 1, $y_i = -1$ if the point belongs to class 2). A margin is defined as the domain between two hyperplanes around the decision boundary of equations $f(x) = +1$ and $f(x) = -1$. The SVM algorithm tends to maximize the width of that margin (of value $\frac{2}{\|w\|}$), which amounts to minimizing the Euclidian norm of the vector w , while maintaining most of the data points outside the margin. The optimal discriminant function $f^*(x) = w^* x + b^*$ is thus the solution of:

$$(w^*, b^*) = \arg \min_{w,b} C \sum_{\forall i} \varepsilon_i + \frac{\|w\|^2}{2} \text{ subjected to } \begin{cases} y_i (< w, x_i > + b) \geq 1 - \varepsilon_i, \forall i \\ \varepsilon_i \geq 0, \forall i \end{cases}$$

where $\sum_{\forall i} \varepsilon_i$ is the separation error term, $\frac{\|w\|^2}{2}$ is the regularization term and C is a user-defined regularization parameter (Muller et al., 2001). As represented Fig. 2 b) some points close to the surface are called support vectors and define the boundaries of the margin on both sides of the surface.

K-Nearest Neighbor

A further data driven method is the k-nearest-neighbors (KNN) algorithm (Hastie et al. 2001). The principle of this third method is as follows: The class label of an unlabelled point is attributed to the predominant class within the k nearest labeled points belonging to the training set (k is defined by the experimenter). The KNN algorithm does not rely on any model but rather on the metric used to assess the similarity of any pair of points (usually the Euclidian distance). Using the aforementioned notation (x_i, y_i) for the points in the learning set the discriminant function associated to KNN is of the form:

$$f_{KNN}^*(x) = \sum_{n \in N_K(x)} y_n$$

where $N_K(x)$ is the index of the K nearest neighbors of x in the training set. From an algorithmic point of view, KNN does not require the iterative learning phase since there is no need to fit a model to the data; however the whole training data set has to be kept in memory

to classify the new data. Compared to linear methods the decision boundary $f_{KNN}^*(x) = 0$ is strongly nonlinear as shown on an example Fig. 2 c).

Of course, there are a wide variety of other existing techniques that have been used, like neural networks (Hiraiwa et al., 1990, Anderson et al., 1998, Haselsteiner et al., 2000) or Hidden Markov Models (Obermeier et al., 2001). A complete review of the existing techniques is out of the focus of this paper and the reader can refer to (Duda and Hart, 2001) or (Hastie et al., 2001) for a general view of the field of Pattern Recognition and (Lotte et al., 2007) for an exhaustive review of the algorithms already used for EEG based BCI.

2.4. Validating classifier performance

Measuring classification accuracy

Training a classifier aims at minimizing classification error generally quantified as the ratio of the number of well-classified samples to the total number of samples. In order to interpret it, one should take the chance level as a reference. When trying to classify one sample by randomly distinguishing between two balanced classes, one may expect a mean accuracy of 50%. By analogy, dealing with a three class problem, chance should not go past about 33% asymptotically. However, given that such a level of accuracy might well be considered as insufficient, other criteria have been introduced. One criterion, known as kappa, represents the rate of well-classified examples after subtraction of the asymptotic chance accuracy (Townsend et al., 2006). A further measure of classifier performance is Area Under the ROC curve (AUC). The receiver operating characteristic (ROC) curve represents the evolution of the false positive rate versus the true positive rate which results from thresholding a discriminant function as a function of threshold values (Green and Swets., 1974, Duda and Hart, 2001). A value of the discriminant function above the threshold predicts class 1, whereas a discriminant function below the threshold yields a class 2 prediction. In practice, this measure allows for a control of the acceptable false negative rate of a diagnosis. In summary, the area under the ROC curve is a global measure of the discriminant power of the discriminant function, regardless of the threshold to be chosen and if the area under this curve reaches 1 (i.e. AUC=1) the discrimination is considered to be perfect.

Finally, for BCI applications, it is obvious that the performance of a classifier also depends on the time-length of the data to be classified and the number of class to be recognized. For example, taking two classifiers with the same accuracies, if the first performs classification on only 1 s time-windows whereas the other uses 2 s windows, then the first can transmit 2 times more information during the same period of time. Similarly, for equal accuracies, a four class classifier sends two times more binary information than a two class classifier with the same number of trials. To allow comparison of classification results associated to these experimental parameters, the notion of information transfer rate (ITR) has been borrowed to communication theory (Shannon 1964) and can be computed using classifier accuracy p , the number of classes N and the number of trial T according to :

$$ITR = \frac{1}{T} (\log_2 N + p \log_2 p + (1 - p) \log_2 \frac{(1 - p)}{N - 1})$$

This equation enables the comparison of all kinds of BCI paradigms: according to the literature the current reachable limit of ITR is approximately 25 bits/minute for non invasive BCI's (Krepki et al.).

Generalization

Generalization is the crucial notion to truly quantify the performance of a classifier: it corresponds to the ability of a trained classifier to classify accurately new data (out of the training set). In order to illustrate that point, consider a set of learning algorithms with a parameter enabling to increase their fit to the training data (for example by using more and more non-linear discriminant functions). By increasing this fit, classification error on the training data will steadily decrease since the algorithm's aim is to minimize it. But the classifier error on another data set (validation data) will stop decreasing and even start increasing when the increased fit to the training set no more represent robust properties of the data to classify: this is called overfitting or overtraining. This scenario is illustrated by Fig. 3 where the validation error curve appears useful to select the best step at which the learning process should better be stopped. Due to this issue, classifier performance has always to be computed on a validation or test dataset to assess generalization ability.

Although they appear under various names in the literature, such as cross validation, leave-one-out, leave-k-out, Jack knife (Stone, 1974, Hastie et al., 2001), existing generalization ability assessment frameworks share a common principle. They all consists in depriving the training set of some data samples in order to provide two kinds of sets, training sets and validation sets, with several sampling of the whole data set. For each sampling, trained classifier accuracies are estimated with the items of data that were not used by the learning algorithm.

Complexity

A last issue the authors would like to address here is the problem of complexity, a notion related to multiple mathematical quantities. In statistical learning, complexity quantifies the ability of a set of discriminant functions to separate with a perfect accuracy a training set when the clouds of points of the respective classes are very intricate. According to this definition, the discriminant functions of a KNN classifier are clearly more complex than LDA's affine functions. The more complex the discriminant functions of a classifier are, the better they can fit training data and thus the increasing fit to the data can be interpreted as an increase in complexity. As previously mentioned, an overly good fit to the training data may severely impair generalization accuracy of the method. In particular, for affine classifiers such as LDA or linear SVM, complexity increases with the dimension of the feature space. Thus, in MEEG classification, a high dimension due to the huge number of features used to describe each segment of signal may cause problems. This problem is sometimes referred as the curse of dimensionality, arising from the fact that the more dimensions one introduces into the classification problem, the more data is needed to tackle the subsequent increase of possible variability (degrees of freedom). The point is that including an additional dimension would require also adding many more than just one sample to each data set. One should therefore pay attention to the concept of feature selection, which allows us to control the dimension of the feature space. A standard way to select variables is univariate selection: a reduced number of features are drawn from the whole set according to their individual discriminative power. This discriminative power is usually quantified by a Fisher T test. More advanced multivariate selection methods take into account the whole set of features simultaneously in order to choose the optimized subset of variables for the classifier (Garett et al., 2003). Another method that can be used to limit classifier complexity is Principal Component Analysis (PCA), which reduces the high dimensional data to a few components used for classification.

3. Application to visuomotor MEG data

We will now illustrate and investigate the use of such procedures by applying linear SVM classification to continuous epochs of MEG data recorded during ongoing visuomotor coordination and during resting states. The aim of this analysis was to assess the ability of the classifier to differentiate between the two conditions and thereby demonstrate its application to real data and indirectly investigate possible physiological processes reflected by the parameters shown to provide best discrimination between the two conditions.

3.1 Data

Behavioral task

The visuomotor experiment required subjects to continuously manipulate a track-ball to compensate the random rotations of a cube projected on a display screen. The visuomotor (VM) task was alternated with a resting condition (R) during which the subjects relaxed while looking at a motionless cube. The subjects were cued to switch between the two conditions every 8 to 12 s, yielding continuous epochs of steady-state MEG data. All subjects gave informed consent and the study was approved by the local medical ethics committee (Jerbi et al. 2007).

Recordings and Preprocessing

The cerebral activity was recorded with a whole-head MEG system (151 sensors, VSM MedTech Ltd.) and digitized at 1.25 KHz. The acquired data was first low-pass filtered (100 Hz cut-off), down-sampled to 312.5 Hz and subject to visual inspection. All data segments contaminated by eye-blinks or unwanted swallowing, coughing or movement artifacts were rejected and heart-beat artifacts were corrected using template-matching. All 8-12s continuous MEG records of each condition (VM or R) were split into non-overlapping 1s-epochs. This yielded about 150 to 300 artifact-free trials for each of the two conditions per subject.

3.2 Methods

Feature quantification

After multiplication by a Hamming window (to avoid edge effects), the 1-s segments of raw MEG signal $z(t)$ corresponding to the k -th time window W_k was band-pass filtered in 6 frequency bands using zero-phase FIR filters (computed by frequency sampling (Oppenheim 89)). The Spectral power was determined for each MEG sensor using Equation 1 in each one of the following standard physiological frequency bands: delta (2-4Hz), theta (5-8Hz), alpha (8-12Hz), beta (15-30Hz), gamma1 (30-60Hz) and gamma2 (60-90Hz). In addition to spectral power, we also computed phase synchrony (i.e. phase locking value, (Lachaux 1999)) between each pair of sensors in each frequency band using Equation 2.

Classification process

The many types of features that have been computed to quantify the brain state are separated into 12 subsets: 6 subsets containing spectral power features in each of the 6 frequency bands

and 6 subsets containing phase synchrony features, also in each frequency band. An exhaustive comparison of these 12 feature subsets has been carried out by using one feature type at a time to predict the brain state (i.e. the behavioral condition) with one of the 3 classifiers : LDA, SVM and KNN. The regularization constant of the SVM was set to $C=.1$ (based on a preliminary investigation) and the number of neighbors for KNN has been set to 5. The predictive power of the power and synchrony features was assessed using two methods:

- 1) Ten fold cross-validation (intra-session): the training set of each session of each subject were split into 10 parts, learning was performed on 9 parts and the classifier was tested on the remaining part.
- 2) Inter-session cross-validation: for each subject, learning was performed on one full session and tested on the remaining session(s). (At least two sessions were available for each subject).

Moreover, the quantity of information about the mental state given by the classifier was evaluated by two quantities:

- 1) Classification accuracy: this is the average between the percentages of good classified points in each class. This measure gives an average rate of 50% for a random guess of the classifier even if the proportion of each class in the dataset is unbalanced.
- 2) Area under the ROC curve (AUR): this quantity gives values of the class information of the discriminant function resulting from the classifier regardless of the threshold to be applied to actually predict the class. The advantage of this method is that it allows for an evaluation of the classifier independently of any possible offset between the features of the test set and the training set. For a random guess the value of AUR is .5 and a perfect (i.e. systematically correct) guess gives $AUR=1$. In our results, AUR is computed from the test set to evaluate the generalization ability of the classifier.

3.3 Results

SVM versus LDA and feature selection

To gain preliminary insight into the effect of the number of features (i.e. possible discrimination parameters) on classification accuracy, we first performed variable selection on the features linked to the beta power. Using an intra-session 10 fold cross validation, only the most discriminative features according to a t-test on the training set were fed into a linear SVM. For comparison, the same procedure was carried out with an LDA classifier and a 5 nearest neighbor classifier (KNN, $K=5$). The classifier accuracy results of all three algorithms are shown in Fig. 4. Our findings suggest that, when using a small training set (150 time windows), the accuracy of linear SVM and KNN remains stable as the number of selected variables grows. In contrast, LDA accuracy drastically decreases when more than 100 features are selected. Linear SVM and KNN may therefore be considered to be more robust than the LDA approach with this high dimensional data. Our results also indicate that the drop in LDA accuracy is less important when a larger training set (300 points) is used. This is in line with the idea that high dimensional data is better classified using more data points. Moreover, the results in Fig. 4 also suggest that even when using linear SVM, it might not be necessary to select a high number of features in order to obtain an optimal accuracy with the data at hand. According to these first results from Fig. 4, in the rest of the study investigating the MEG

visuomotor data, we chose to use 36 features (with a Fisher t-test) in each subset before the training of a linear SVM classifier.

Classifier validation

To assess the stability of classifier performance across time, we estimated the classifier accuracy of SVM derived from each feature subset both within the same session and in between sessions using 10 fold cross-validation and inter-session cross-validation respectively. The average classification results across the 3 subjects are given Fig. 5. Our results show that for the MEG visuomotor data used here, the best accuracy was obtained using beta activities as discriminant features. This result did not depend on the type of feature (power or synchrony) nor the type of validation used (inter or intra session). In particular, the highest score (86%) was obtained for beta power in intra session cross-validation (Fig. 5). Note that nearly all features are less predictive in inter session than in intra session validation. In particular, the classification accuracy based on phase synchrony in the beta and gamma bands was reduced (up to 25%) in inter-session validation. Furthermore, while the predictive power was equivalent for spectral power features and phase synchrony in intra-session validation, spectral power was clearly more predictive in intersession.

The information content of each feature set was also investigated with AUR which we computed for the values of the discriminant functions on the test set resulting from each previously computed SVM classifier. The mean AUR across three subjects for each feature subset are given in Fig. 6. The same general tendency as in Fig. 5 is observed except that the ROC area does not drop in intersession compared to intra session in particular for the beta and gamma1 power features. This confirms that, with the data at hand, beta power clearly carries discriminative information which can be used to infer the mental state from data in another recording session.

The same two cross-validations were also carried out to compute the classification accuracy of feature subsets that combine both spectral power and phase synchrony for a given frequency range. The comparison of the results obtained with classifications based only on power, only on synchrony and on the combination of both are given in Fig. 7. Our finding suggest that whether this combination is beneficial or not to the classifier strongly depends on the frequency band: in the three lower frequency bands, including phase synchrony seems to decrease classifier accuracy achieved with spectral power alone, but conversely the combination of power and synchrony in higher frequency bands outperforms the results obtained with each taken separately. Moreover, combining the selected features from all frequency bands also improves classifier accuracy, reaching 97% as shown on the right of Fig. 7. This accuracy leads to an excellent ITR of 48 bits/minutes.

Distribution of classifier coefficients

The final point we investigated in this study using the MEG data set is the spatial distribution of classifier coefficients across the sensor-space. When using a binary linear classifier such as LDA or SVM, the discriminant function is an affine function of the features. Therefore when dealing with features of comparable amplitude, one would expect that the features associated to the coefficients with the highest absolute value are the most discriminant features for the classification. If this coefficient is positive the increase of corresponding feature produces the decision 'class 1' (i.e. VM) while its decrease classifies the data as 'class 2' (Rest). Such

interpretations are not completely rigorous as the function is highly multivariate but it may provide useful information on the task-specific features.

To illustrate this with an example, we computed the scalp distribution of the SVM classifier coefficients obtained for delta, beta and gamma 2 power as well as with alpha synchrony for one representative subject on Fig. 8. The upper row shows the topographic distribution of the classifier coefficients related to spectral power features. The beta band power classifier topography shows that this frequency range is associated with negative classifier coefficients in frontal central areas (blue blob in the central topography). In other words, a decrease in beta power over these sensors draws the prediction towards classifying the state as an ongoing visuomotor state (VM) whereas an increase is associated with the resting state (Rest). Moreover, positive coefficients, which correspond to labeling a data sample VM (i.e. a brain state corresponding to visuomotor control) show prominent peaks over frontal areas contralateral to the moving limb in both the delta and gamma2 frequency bands. In contrast to the rather central beta suppression, the location of these delta and high gamma peaks most likely corresponds to sensors most sensitive to neural activation originating in the contralateral sensorimotor cortex. Furthermore, although one might be tempted to interpret the positive coefficients over posterior regions as being task-related and reflecting induced power increase in all three frequency bands during visuomotor coordination, this should be done with great caution. Indeed, gamma power modulations over posterior occipital areas are thought to be involved in visual processing, but one cannot exclude that the posterior modulations at the bottom of the sensor array are not at least in part caused by muscle artifacts. Muscle artifacts are known to be broad band (extending also over gamma, beta and even lower frequencies) and may well be selectively present during the visuomotor task but not during the resting condition and thus picked up by our SVM classifier as a discriminant feature. This said, the classifier coefficient peaks over central and contralateral sensorimotor areas are too far from muscle activities to be contaminated by any such artifacts and they clearly represent task-related neural power modulations. Interestingly, these findings are inline with our previous results on task-related power modulation in visuomotor control that were not obtained by any classification procedure but rather by contrasting power topographies across conditions both at sensor level and source level (Jerbi et al. 2004, 2005). This converging evidence from analysis of task-related cortical power modulations on the one hand and linear classification based on multi-frequency power on the other, underline the fact that changes in oscillatory power represent a strong candidate for classification methods because they represent fundamental modulations in the neurophysiological processes associated with various states. As far as delta and alpha synchrony is concerned (middle and lower panel of Fig. 8), the spatial distribution of positive and negative coefficients seems to be more complex than the power topographies. This may be in part due to the high number of features available. The delta synchrony distribution shows a high number of positive coefficients for synchronization couples between the left motor cortex and other areas, also in line with previous results (Jerbi et al., 2007). Conversely, it is noteworthy that right motor cortex yields negatives synchrony coefficients with other distant areas. Whether this reflects the disengagement of the ipsilateral motor cortex (versus the engagement of the contralateral motor cortex) remains to be checked on a complete group of subjects. Also interesting are the positive alpha synchronizations coefficients between contralateral parietal cortex and contralateral frontal areas that could reflect the enhanced interactions between visual and frontal areas during the visuomotor task.

3.4 Discussion

The results reported here illustrate several important issues related to the classification of MEG states. First of all, we have shown that the number of chosen features can influence classification performance drastically. This is especially true for LDA and any other model-based classifier that requires the estimation of many parameters of the probability distribution of the data. In practice, the number of parameters to estimate has to be small compared to the number of data points in the training set. In LDA for example, all the elements of the covariance matrix of the data have to be estimated, which is thus proportional to the square of the number of features. Using LDA in high dimensions thus require a large amount of data points in the training set. Moreover, even though SVM seems to be less sensitive to the dimension of the data set used here, this is not necessarily true in general. In fact, if we had chosen a larger regularization constant C , classification would have dropped due to the high complexity of the classifier. Finally, a non linear classifier like KNN can achieve competitive classification accuracy levels and may very well be chosen for MEEG classification, but does not clearly outperform a linear SVM. It thus seems that, in the high dimensional feature space considered here, linear classifiers are robust and flexible enough to perform the classification task with acceptable accuracy.

After having chosen a reasonable number of parameters and a reliable classifier, the quality of the results is still subject to the validation method. We have shown for example that predicting the mental state of data from one session after training the classifier on another session is more difficult than predicting across the same session using cross-validation. This can be due to many factors related to the reproducibility of the experiment and session specific artifacts or parameter variations (e.g. variable head position within the MEG helmet). This may lead to significant changes in the MEG recordings that could then impair classification results. In some cases, a simple change of the threshold of the discriminant function may be sufficient to improve classification. This is precisely the case we discussed when we observed that in inter-session validation the ROC area did not drop off whereas classifier accuracy did (i.e. ROC accounts for offset variations related to variable baseline levels in between sessions). Such simple corrections might not be sufficient when using phase synchrony measurements as discriminant features, especially in the higher frequency ranges which can be particularly extremely sensitive to artifacts. A further possible cause for the drop in classification accuracy between recording sessions is a change in the state of the subjects themselves, due to learning, a change of strategy or simply alertness and motivation levels. This is particularly relevant for continuous mental states which are internally paced and thus less stable than exogenous activities triggered by a stimulus. In general, the difficulty to predict one session using another is a crucial and challenging issue which has to be examined carefully for applications such as Brain Computer Interfaces.

In this paper, we discussed the use of power and phase synchrony to perform data classification. A crucial limitation of phase synchrony analysis is that it is extremely difficult to differentiate between true physiological coupling between distinct neural assemblies and false synchronization that appears due to the spreading on the scalp of the electromagnetic field of a unique cortical source. According to Fig. 7 combining power and synchrony features in the upper frequency bands improves classification accuracy which is an argument to say that the recorded synchrony does not carry exactly the same information as power, and thus may not uniquely be explained by diffusion. Whether, and if so to which extent, synchrony between neighboring sensors reflects local amplitude modulations of the same neuronal assembly is not the aim of this specific study and it remains of course an open issue. What is more important, however, is that long-range synchronization between widely

separated sensors is less likely to represent a single source especially if one can show that the coupling between the two sensors is not zero-phase (Nolte et al. 2004). Finally, it might be noteworthy to recall that although insight into the dynamics of local and large-scale synchrony and differentiating the role of the two would be well appreciated from a basic physiological stand point, the ultimate goal of a classification method is above all to separate two or more conditions in the most robust and efficient way possible.

The scalp distributions of classifier coefficients obtained here with the visuomotor MEG data set are in agreement with previous works in the field of ERD and ERS analysis in particular in the sensorimotor cortex using non-invasive (Pfurtscheller and Lopes Da Silva, 1999) and invasive (Crone et al. 1998a, 1998b) recordings. Although the spatial distribution of classifier coefficients was reported here with real visuomotor MEG data in a subset of frequency bands for illustrative purposes, the obtained topographies are in remarkable agreement with our knowledge of the modulation and distribution of cortical oscillatory activity during sensorimotor behavior. First, the distribution of the central negative coefficients for beta power is in line with the known suppression of beta power during movement over bilateral sensorimotor cortices. The power suppression, which seems to be enhanced over the contralateral cortex, probably involves primary sensory, motor and premotor areas, including Brodmann Area 4 (BA4) and most likely the supplementary motor area (SMA) (Jerbi et al. 2004, 2005). Furthermore, in the light of the positive coefficients revealed here in the high gamma range over the contralateral sensorimotor area, and bearing in mind previous findings on the presence of focal gamma activity in the activated motor cortex (Crone et al. 1998b), it is also tempting to suggest that the bilateral beta band activations might be less task-specific or functionally relevant than the gamma power modulations. More studies will be needed in order to systematically address the functional relationship between beta suppression and gamma power increase during motor tasks. Finally, a further frequency range that showed strong positive classification coefficients between visuomotor and resting conditions is the delta (2-4 Hz) band. Although surprisingly low, this frequency band has recently been shown to host slow sensorimotor oscillations related to the ongoing control of hand speed (Jerbi et al. 2007). Whether an increase in delta-range cortical power changes in the motor cortex represents modulation of an intrinsic sensorimotor rhythm or whether it is rather a direct reflection of task parameters is still an open question. Similarly, it would be difficult to draw qualitative conclusions by visual inspection of the complex synchronization topographies (middle and lower panel of Fig. 8), and our observations about enhanced contra-lateral synchronization in the delta and alpha band have to be confirmed by future work on a larger population of subjects. What matters here, is that the classification method we used identified the sensorimotor delta power as one of the major features that best separates the two conditions, a finding fully in line with a separate analysis of the data. Again, these converging findings from basic research analysis and data classification procedures confirm the fact that successful classification inherently relies on distinguishing task-related physiological phenomena. Moreover, although classification techniques can be improved by taking neuroscience *a priori* knowledge into account (such as frequency bands or regions of interest), they also hold the potential of indirectly enhancing our understanding of brain dynamics. Indeed, the features revealed by such methods to be most relevant for the optimization of classification may lead to new hypotheses about the dynamics of human cognition in specific tasks.

Conclusion

Beyond reviewing the basic principles of BCI-oriented data classification, we used MEG data to illustrate the limitations and differences between several approaches and to shed some light on the effect of various user-defined parameters. Furthermore, the successful MEG data classification reported here in offline analysis, demonstrates that the real-time detection of visuomotor activity using power and synchrony measurements is feasible and suggests that classification of motor or visuomotor imagery (without actual movement) can be obtained using the same methodology developed here. This is particularly the case, since motor imagery is known to activate (although to a lesser extent) predominantly the same areas involved in real movement execution (Beisteiner et al., 1995; Leocani et al., 1999) and the subject's ability to enhance the corresponding oscillatory modulations can also be trained. On the whole, the accuracy of the classification achieved here during a continuous task (i.e. not event-related) suggests the possibility of achieving a high transmission rate asynchronous BCI. In addition, the agreement between the extracted classification features on one hand and the knowledge of the physiological processes underlying the visuomotor control data on the other, suggests that basic neuroscience research can also benefit from M/EEG data classification methods as they may yield novel insights to the study of the mechanisms underlying ongoing brain dynamics during mental tasks.

Acknowledgements

The authors would like to thank Dr Mario Chavez and Frédérique Amor for their useful comments on the manuscript.

Bibliography

- ANDERSON C, STOLZ E, SHAMSUNDER S (1998). Multivariate Autoregressive Models for Classification of Spontaneous Electroencephalogram During Mental Tasks. *IEEE Trans. Biomed. Eng.* 45: 277-286.
- BARRET G (1996). Event-related potentials (ERPs) as a measure of complex cognitive function. *Electroencephalogr Clin Neurophysiol* 46: 53-63.
- BASHASHATI A, FATOURECHI M, WARD RK, BIRCH GE (2007) A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals. *J. Neural. Eng.* 4 :R32-R57
- BEISTEINER R, HOLLINGER P, LINDINGER G, LANG W, BERTHOZ A. (1995) Mental representations of movements. Brain potentials associated with imagination of hand movements. *Electroenceph Clin Neurophysiol*, 96:183-193
- BORISOFF JF, MASON SG, BASHASHATI A, BIRCH GE (2004). Brain-Computer Interface Design for Asynchronous Control Applications: Improvements to the LF-ASD Asynchronous Brain Switch. *IEEE Trans. Biomed. Eng.* 51: 985-992.
- BOSTANOV V (2004). {BCI} competition 2003–data sets ib and iib: feature extraction from event-related brain potentials with the continuous wavelet transform and the t-value scalogram. *IEEE Trans. Biomed. Eng.* 51: 1057-1061.
- BROVELLI A, DING M, LEDBERG A, CHEN Y, NAKAMURA R, BRESSLER S (2004). Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by granger causality. *Proc Natl Acad Sci USA* 101: 9849-9854.
- CARLSON TA, SCHRATER P, HE S (2003). Patterns of activity in the categorical representation of objects. *J. Cogn. Neurosci.* 15: 704–717.
- CRONE N, MIGLIORETTI D, GORDON B, LESSER R (1998a). Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. ii. event-related synchronization in the gamma band. *Brain* 121: 2301-2315.
- CRONE N, MIGLIORETTI D, GORDON B, SIERACKI J, WILSON M, UEMATSU S, LESSER R (1998b). Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. i. alpha and beta event-related desynchronization. *Brain* 121: 2271-2299.
- DUDA RO, HART PE, STORK DG (2001). Pattern Classification.
- FARWELL LA, DONCHIN E (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroenceph. Clin. Neurophysiol.* 70: 510-523.
- FRACKOWIAK RSJ, FRISTON KJ, FRITH CD, DOLAN RJ, MAZZIOTTA JC (1997). Human Brain Function.
- FUCHS T, BIRBAUMER N, LUTZENBERGER W, GRUZELIER JH, KAISER J (2003). Neurofeedback Treatment for Attention-Deficit/Hyperactivity Disorder in Children: A Comparison with Methylphenidate. *Applied Psychophysiology and Biofeedback* 28: 1-12.

- GARRETT D, PETERSON DA, ANDERSON CW, THAUT MH (2003). Comparison of linear, nonlinear, and feature selection methods for EEG signal classification. *IEEE Trans. Neural Syst. Rehabil. Eng.* 11: 141-144.
- GEORGOPOULOS AP, LANGHEIM FJ, LEUTHOLD AC, MERKLE ANMSPMTISEBR (2005). Magnetoencephalographic signals predict movement trajectory in space. *Exp. Brain Res.* 167: 132-135.
- GRAIMANN B, PFURTSCHELLER G (2006) Quantification and visualization of event-related changes in oscillatory brain activity in the time-frequency domain In: NEUPER C, KLIMESCH W (eds) *Progress in Brain Research* 159: 79-97.
- GREEN DM, SWETS JA (1974). *Signal Detection Theory and Psychophysics*.
- GYSELS E, CELKA P (2004). Phase Synchronisation for the recognition of Mental Tasks in a Brain Computer Interface. *IEEE Trans. Neural Syst. Rehabil. Eng.* 12: 206-415.
- GYSELS E, RENEVEY P, CELKA P (2005). SVM-based recursive feature elimination to compare phase synchronization computed from broadband and narrowband EEG signals in Brain-Computer Interfaces. *Signal Processing* 85: 2178-2189.
- HASELSTEINER E, PFURTSCHELLER G (2000). Using time-dependent neural networks for EEG classification. *IEEE Trans. Rehab. Eng.* 8: 457-463.
- HASTIE T, TIBSHIRANI R, FRIEDMAN J (2001). *The Elements of Statistical Learning*.
- HAYNES JD, REES G (2006). Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.* 7: 523-534.
- HIRAIWA A, SHIMOHARA K, TOKUNAGA Y (1990). EEG topography recognition by neural networks. *IEEE Eng. Med. Biol. Mag.* 9: 39-42.
- JERBI K, BAILLET S, LACHAUX JP, PANTAZIS D, LEAHY R, GARNERO L (2005). Modulations of Power and Synchronization of Neural Activity during Sustained Visuomotor Coordination, a MEG Study. *Proc. 11th Int. Conf. on Human Brain Map.*
- JERBI K, LACHAUX JP, N'DIAYE K, PANTAZIS D, LEAHY RM, GARNERO L, BAILLET S (2007). Coherent neural representation of hand speed in humans revealed by MEG imaging. *Proc. Nat. Ac. Sci.* 104: 7676-7681.
- JERBI K, LACHAUX J-P, BAILLET S, GARNERO L (2004). Imaging Cortical Oscillations during Sustained Visuomotor Coordination In MEG. *Proc. IEEE International Symposium on Biomedical Imaging: From Nano to Macro.*
- KEIRN ZA, AUNON JI (1990). A new mode of communication between man and his surroundings. *IEEE Trans. Biomed. Eng.* 37: 1209-1214.
- KREPKI R (2006). Berlin Brain-Computer Interface. The HCI communication channel discovery. *Int. J. Human-computer Studies* 65: 460-477.
- KUBLER A, KOTCHOUBEY B, KAISER J, WOLPAW JR, BIRBAUMER N (2001). Brain-computer communication: unlocking the locked in. *Psychol. Bull.* 127: 358-375.
- LACHAUX JP, RODRIGUEZ E, MARTINERIE J, VARELA FJ (1999). Measuring phase synchrony in brain signals. *Hum. Brain Mapp.* 8: 194-208.
- LACONTE S, STROTHER S, CHERKASSKY V, ANDERSON J, HU X (2005). Support vector machines for temporal classification of block design fMRI data. *Neuroimage* 26: 317-329.
- LAL TN, SCHRODER M, HILL J, PREISSEL H, HINTERBERGER T, MELLINGER J, BOGDAN M, ROSENSTIEL W, HOFMANN T, BIRBAUMER N, SCHOLKOPF B (2005) A Brain Computer Interface with Online Feedback based on Magnetoencephalography. In : DE RAEDT L, WROBEL S (eds) *Proceedings of the 22nd International Conference on Machine Learning*. pp : 465 – 472.
- LEBEDEV M, NICOLELIS MA (2006). Brain-machine interfaces: past, present and future. *Trends Neurosci.* 29: 536-546.

- LEOCANI L, MAGNANI G, COMI G (1999). Event Related Desynchronisation during execution, imagination and withholding of movement. Handbook of Electroencephalography and Clinical Neurophysiology. PFURTSCHELLER GALDS, F. H. 6: 245-268.
- LOPES DA SILVA FH (2006). Event related oscillations: what about phase? In: NEUPER C, KLIMESCH W (eds) Progress in Brain Research 159: 1-17.
- LOTTE F, CONGEDO M, LECUYER A, ARNALDI B (2007). A review of classification algorithms for EEG-based Brain-Computer Interfaces. J. Neural. Eng. 4: R1-R13.
- MIDDENDORF M, MCMILLAN G, CALHOUN G, JONES KS (2000) Brain-computer interfaces based on the steady-state visual-evoked response IEEE Trans. Rehabil. Eng. 8 : 211-214
- MILLAN JR, RENKENS F, MOURINO J, GERSTNER W (2004) Noninvasive brain-actuated control of a mobile robot by human EEG. IEEE Trans. Biomed. Eng. 51 : 1026-33.
- MULLER KR, MIKA S, RATSCH G, TSUDA K, SCHOLKOPF B (2001) An introduction to kernel-based learning algorithms. IEEE Tran. Neur. Netw 12 : 181-202
- MULLER-GERKING J, PFURTSCHELLER G, FLYVBJERG H (1999). Designing optimal filters for single-trial EEG classification in a movement task. Clin. Neurophysiol. 110: 787-798.
- NOLTE G, BAI O, WHEATON L, MARI Z, VORBACH S, HALLETT M (2004). Identifying true brain interaction from EEG data using the imaginary part of coherency. Clin Neurophysiol 115: 2292-2307.
- OBERMEIER B, GUGER C, NEUPER C, PFURTSCHELLER G (2001). Hidden Markov models for online classification of single trial EEG. Pattern Recognit. Lett. 22: 1299-1309.
- OPPENHEIM AV, SCHAEFER RW (1989). Digital signal processing.
- PFURTSCHELLER G (2006). The cortical activation model (CAM) In: NEUPER C, KLIMESCH W (eds) Progress in Brain Research 159: 19-27.
- 19-27.
- PFURTSCHELLER G, NEUPER C, FLOTZINGER D, PREZENGER M (1997). EEG-based discrimination between imagination of right and left hand movement. Electroencephalogr. Clin. Neurophysiol. 103: 642-651.
- PFURTSCHELLER G, LOPES DA SILVA FH (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin. Neurophysiol. 110: 1842-1857.
- PIKOVSKI A, ROSENBLUM M, KURTHS J (2001). Synchronization: A Universal Concept in Nonlinear Science.
- WEI Q, WANG Y, GAO X, GAO S (2007) Amplitude and phase coupling measures for feature extraction in an EEG-based brain-computer interface. J. Neural Eng. 4: 120-129
- QUYEN MLV, SOSS J, NAVARRO V, ROBERTSON R, CHAVEZ M, BAULAC M, MARTINERIE J (March 2005). Preictal state identification by synchronization changes in long-term intracranial EEG recordings. Clin. Neurophysiol. 116: 559-568.
- R. JMD, RENKENS F, MOURINO J, GERSTNER W (2004). Noninvasive Brain-Actuated Control of a Mobile Robot by Human EEG. IEEE Trans. Biomed. Eng. 51: 1026-1033.
- RODRIGUEZ E, GEORGE N, LACHAUX JP, MARTINERIE J, RENAULT B (1999). Perception shadow : long distance gamma band synchronisation and desynchronisation on the human scalp. Nature 397: 430-433.

- SALMELIN R, HARI R (1994). Spatiotemporal characteristics of sensorimotor neuromagnetic rhythms related to thumb movement. *Neuroscience* 60: 537-550.
- SCHERER R, MÜLLER GR, NEUPER C, GRAIMANN B, PFURTSCHHELLER G (2004). An Asynchronously Controlled EEG-Based Virtual Keyboard : Improvement of the Spelling Rate. *IEEE Trans. Biomed. Eng.* 51: 979-984.
- SHANNON CE, WEAVER W (1964). *Mathematical Theory of Communication*.
- STONE M (1974). Cross-validatory choice and assesment of statistical predictions. *J. Roy. Statist. Soc.* 36: 111-147.
- STREHL U, LEINS U, GOTH G, KLINGER C, HINTERBERGER T, BIRBAUMER N (2006). Self-regulation of Slow Cortical Potentials: A New Treatment for Children With Attention-Deficit/Hyperactivity Disorder. *Pediatrics* 118: 1530-1540.
- TALLON-BAUDRY C, BERTRAND O (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn. Sci.* 3: 151-162.
- TALLON-BAUDRY C, BERTRAND O, DELPUECH C, PERMIER J (1997). Oscillatory gamma-band (30-70 Hz) activity induced by a visual search task in humans. *J. Neurosci.* 17: 722-734.
- TAYLOR DM, TILLERY SI, SCHWARTZ AB (2002) Direct cortical control of 3D neuroprosthetic devices. *Science* 296:1829–1832
- TOWNSEND G, GRAIMANN B, PFURTSCHHELLER G (2006) A comparison of common spatial patterns with complex band power features in a four-class BCI experiment *IEEE Trans Biomed Eng* 53 : 642- 651
- VAPNIK V (1998). *Statistical Learning Theory*.
- VARELA F, LACHAUX JP, RODRIGUEZ E, MARTINERIE J (2001). The Brainweb : Phase synchronisation and Large-scale integration. *Nature Rev. Neurosci.* 2: 229-239.
- VARELA FJ (1995). Resonant cell assemblies : a new approach to cognitive function and neuronal synchrony. *Biol. Res.* 28: 81-95.
- WALTRE DO (1968). Coherence as a mesure of relationship between EEG records. *Electroenceph. Clin. Neurophysiol.* 24: 282.
- WESSBERG J, NICOLELIS MA (2004) Optimizing a linear algorithm for real-time robotic control using chronic cortical ensemble recordings in monkeys. *J Cogn Neurosci* 16:1022–1035
- WOLPAW JR, BIRBAUMER N, MCFARLAND DJ, PFURTSCHHELLER G, VAUGHAN TM (2002). Brain-computer interfaces for communication and control. *Clin. Neurophysiol.* 113: 767-791.

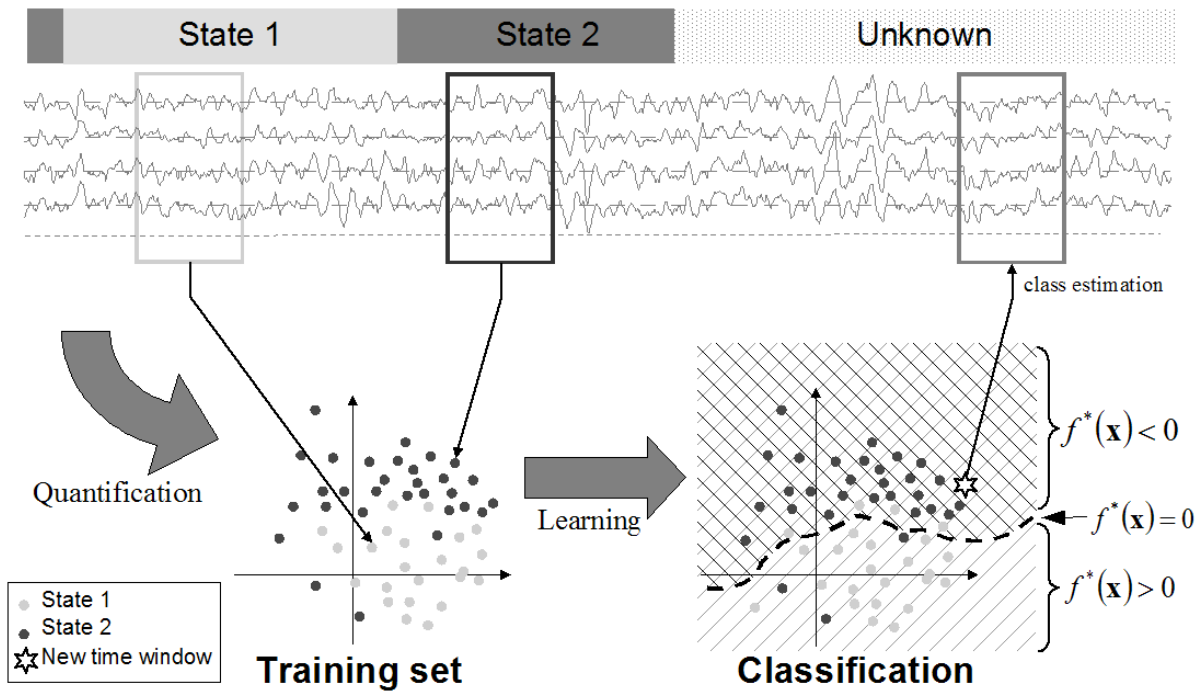


Figure 1: Principle of a binary classification of MEEG recordings for a continuous experiment. First, signals are segmented in successive time windows, then features are computed representing each time window as a point in a multidimensional space. Then, the classification algorithm separates the features' space in two domains according to the sign of a discriminant function $f^*(x)$, enabling to predict the class of a time window according to its feature vector.

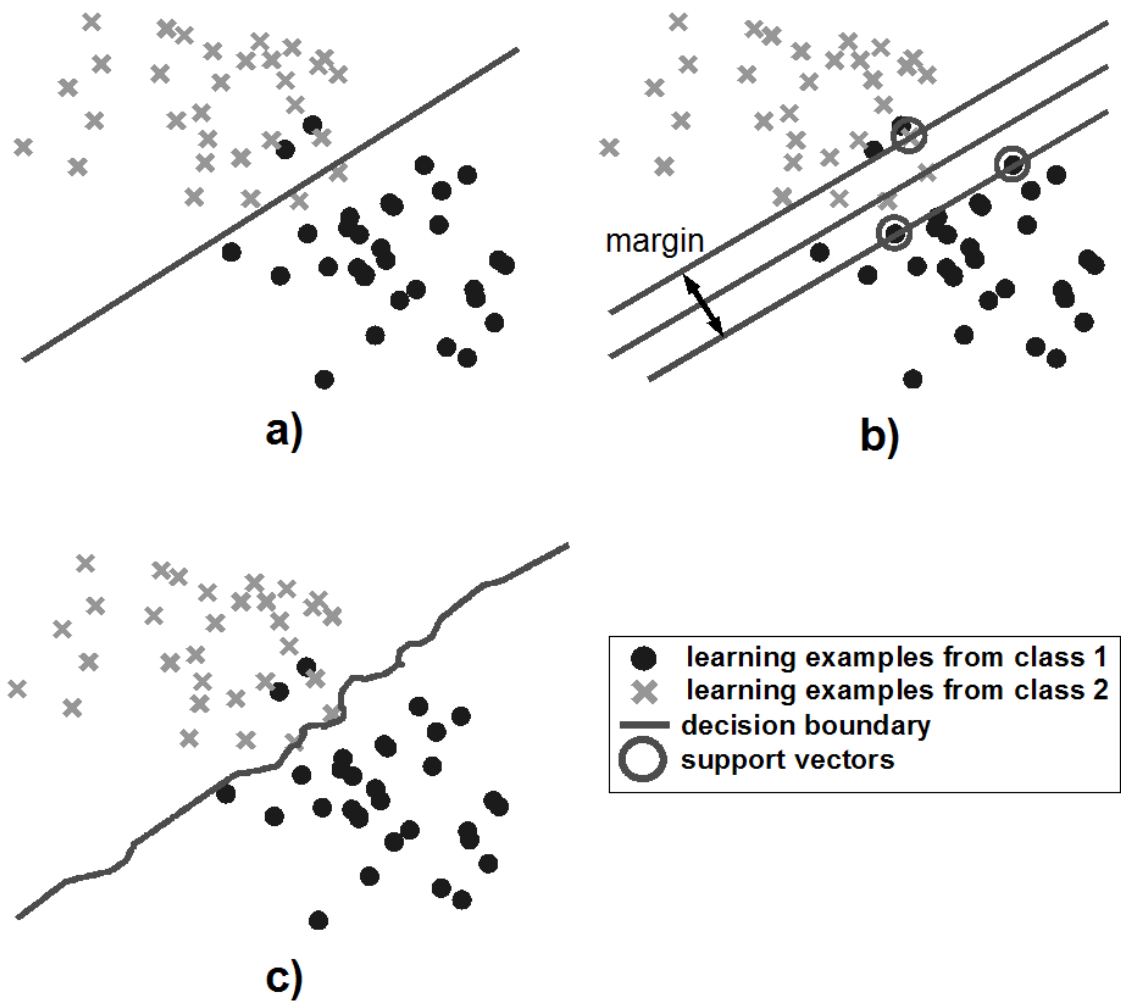


Figure 2: Examples of decision boundaries obtained with the same training set and different classification algorithms: a) LDA classifier, b) SVM classifier, c) classifier with five nearest neighbors using the Euclidian distance.

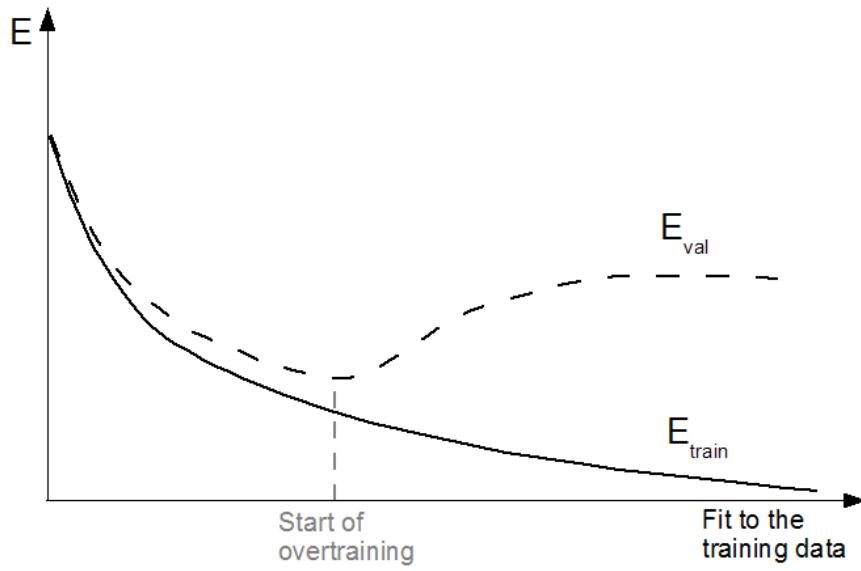


Figure 3: Schematic evolution of learning error (E_{learn}) and validation error (E_{val}) as a function of the fit to learning data. After a regular decrease, E_{val} reaches a minimum and a better fit to learning data leads to a worse performance on validation data.

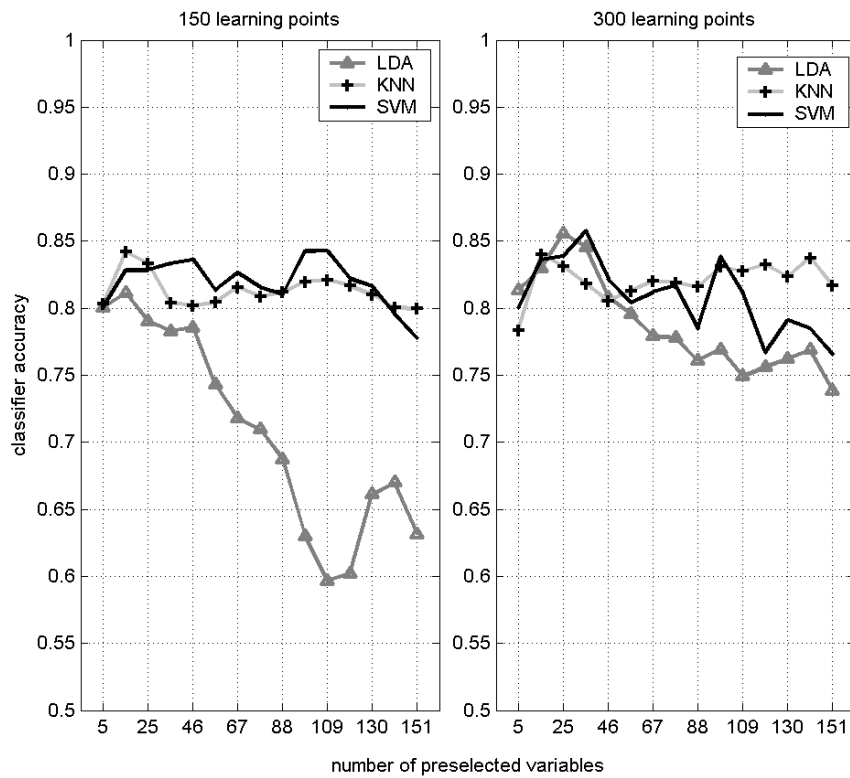


Figure 4: Classifier accuracy as a function of the number of features and training samples. Average classifier accuracy in inter-session cross-validation for the 3 subjects as a function of the number of selected power features (from 5 to 151 variables) for LDA, KNN and SVM for two different sizes of the learning set: 150 points (left panel) and 300 points (right panel). Note that these results were obtained by only investigating beta power features as a preliminary analysis aimed at comparing the three techniques.

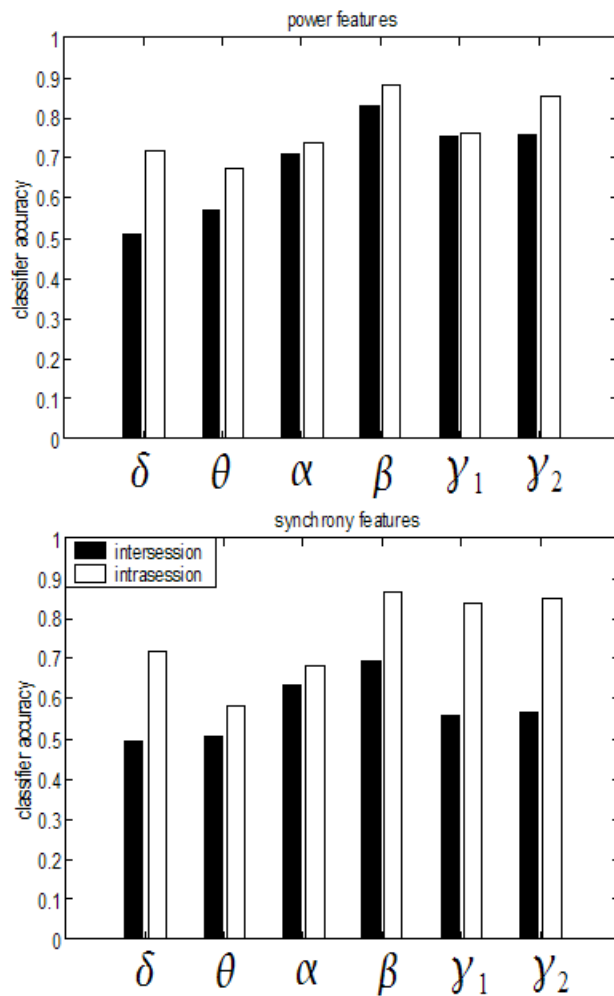


Figure 5: Inter-session versus intra-session validation. Classification accuracy of the classifiers derived from power (upper panel) and synchrony (lower panel) features in different frequency bands and for two validation methods: intra-session (using ten fold cross validation) and inter-session (learning on one session and testing on the other).

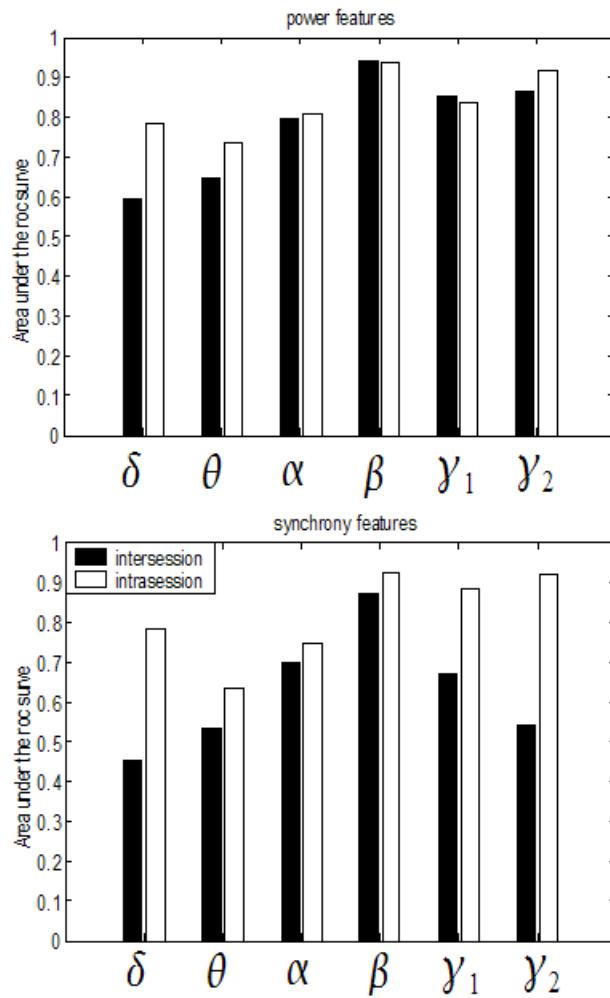


Figure 6: Area under the ROC curve of the decision function of classifiers derived from each feature subset in intra and inter-session cross-validation. Upper panel and lower panels depict the results obtained using power and synchrony features respectively in all six selected frequency bands. The mean classifier accuracy obtained using the selected features from all the frequency bands are also represented on the right.

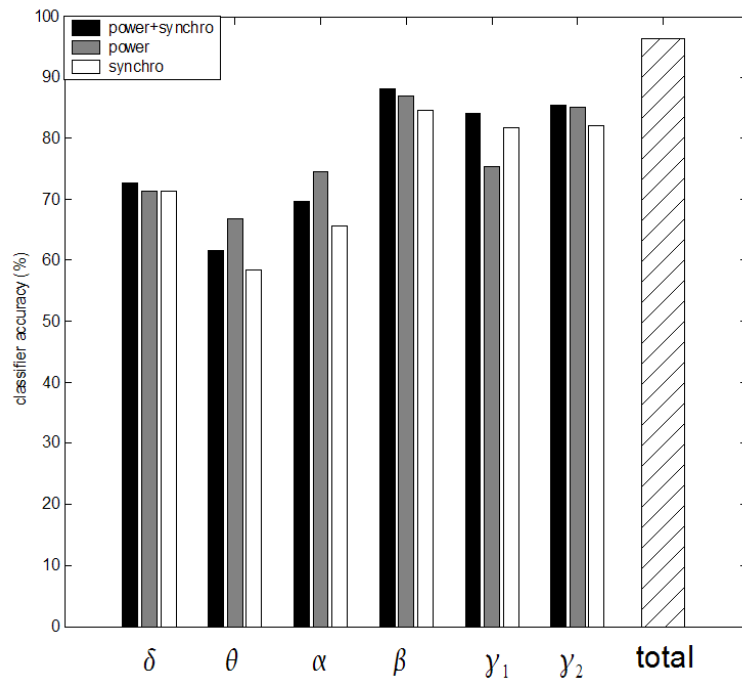


Figure 7: Mean classifier accuracy obtained by combining power and synchrony features, compared to power-only or synchrony-only classification. The histograms depict the relationship between classifier accuracy (in all three modes) and the selected frequency band. The bar on the right side shows classification accuracy obtained using the selected features from all the frequency bands combining power and synchrony.

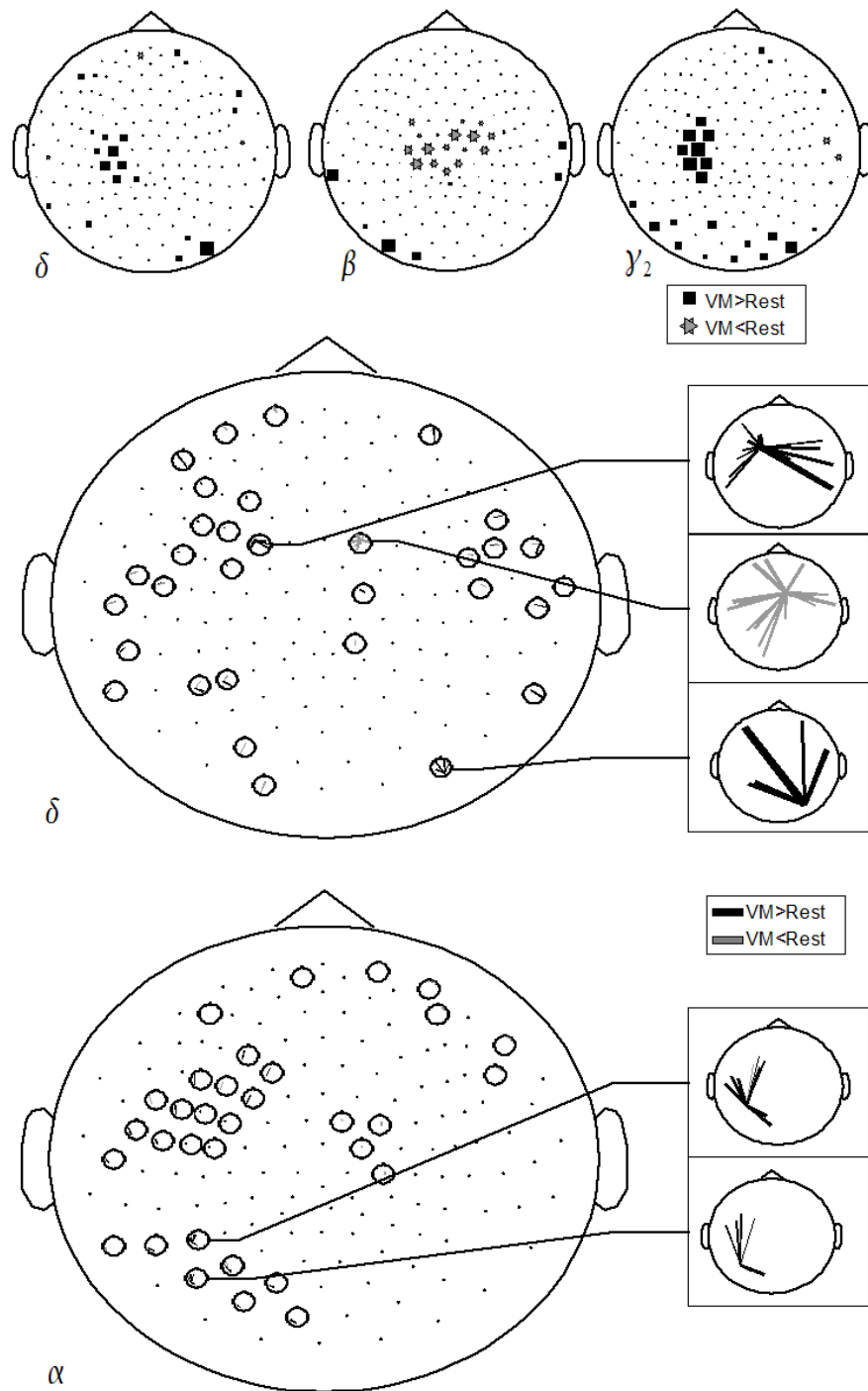


Figure 8: Spatial distribution of SVM classifier coefficients for power and synchrony features in various frequency bands for one subject. Upper row: The distribution of delta (δ), beta (β) and gamma 2 (γ_2) power classifiers coefficients: positive coefficients are represented by black squares and negative ones by gray stars (marker size is proportional to amplitude). Middle and lower panel: distribution of delta and alpha synchrony classifiers coefficients. Note that each small topography on these panels represents the non-zero coefficients of phase synchrony involving the sensor at the current scalp position. Positive coefficients are represented by black lines linking the current sensor and the other sensors, negative ones are represented by grey lines. The close-ups indicate interesting sensors showing many long distance synchronies.

Annexe C : OObci, une librairie orientée objet pour les ICM

Pour tester l'implémentation de différentes méthodes, des techniques de validation croisée sont implémentées sur les données. La chaîne de traitement (prétraitement, quantification, classification) étant redondante dans l'ensemble de nos traitements, il a été choisi de structurer l'organisation des programmes de validation des différents traitements. Ainsi, une boîte à outils orientée objet a été implémentée au format Matlab pour permettre le traitement des données. Il existe une classe d'objet correspondant à chaque étape de traitement :

- les données d'enregistrement sont représentées par une classe *rawdata* qui contient l'ensemble des informations sur un jeu de données : le chemin d'accès aux fichiers d'enregistrement, la position temporelle de l'ensemble des fenêtres d'analyse de chaque enregistrement, la fréquence d'échantillonnage, etc... Les méthodes qui s'appliquent à cette classe permettent d'extraire les signaux associés à différentes fenêtres d'analyse au format Matlab. Différents formats de données sont associés à des classes qui héritent de *rawdata* : les données issues du système BrainAmp (.eeg) correspondent à la classe *bamdata*, les données au format matlab (.mat) à la classe *matdata* et les données simulées à partir d'un modèle prédéfini à la classe *simdata*.
- A un jeu de données, on associe une méthode de quantification des activités à l'aide d'une classe de données quantifiées *dstructq*. Une instance de cette classe contient à la fois le jeu de données de classe *rawdata* et les informations concernant la méthode de quantification stockées dans une structure *quant*. La classe *bssalg* correspond à une quantification qui fait appel à un ajustement automatique de paramètres : c'est le cas lorsqu'on utilise une technique de séparation de sources (BSS : Blind Source Separation ou séparation de sources aveugle). Il est possible d'avoir accès à la quantification d'une fenêtre d'analyse soit par calcul direct à partir des données *rawdata*, soit par lecture dans un fichier généré automatiquement.
- la classe *statlearn* correspond à des algorithmes d'apprentissage statistiques et la classe *classifier*, qui en hérite, aux algorithmes de classification. Ceux-ci regroupent les informations sur la méthode de classification ainsi que les données (quantifiées) d'apprentissage et de test sous forme d'objets *dstructq*. La puissance de cette classe est que la validation du classifieur est automatisée grâce à des méthodes de cross-validation à travers les sessions ou les jours d'enregistrements.

- enfin des méthodes d’affichage de la distribution spatiale des résultats sont réalisées à partir de la classe *headtopo* et ces héritières qui incorporent les données spécifiques aux sujets (position des électrodes, surface du cortex, etc...)

Ce cadre de travail permet d’agir indépendamment sur les jeux de données, leur quantification et leurs classifieurs, garantissant ainsi la portabilité des méthodes implémentées sur de nouveaux jeux de données. Un exemple rapide montre la puissance de la boîte à outils pour programmer de manière concise la validation d’un classifieur :

```
% chargement des données quantifiées
bciquant=dstructq(learn_async_1_quant);
% modification des bandes de fréquences d’analyse (Hz)
bciquant.quant.v_freq=[4 8 12; 8 12 18];
% utilisation d’une matrice de filtrage spatial de type ’référence moyenne
% commune’ pour le prétraitement des données
bciquant.quant.W=M_comRef;
% création du classifieur svm linéaire sur les données quantifiées
classifobj=classifier(’svm_l’,bciquant);
% estimation du taux de bien classés pour la discrimination classe 3 versus
% classe 4 à l’aide d’un cross validation à travers les sessions
Tpct=crossval(classifobj,3,4,’session’);
```

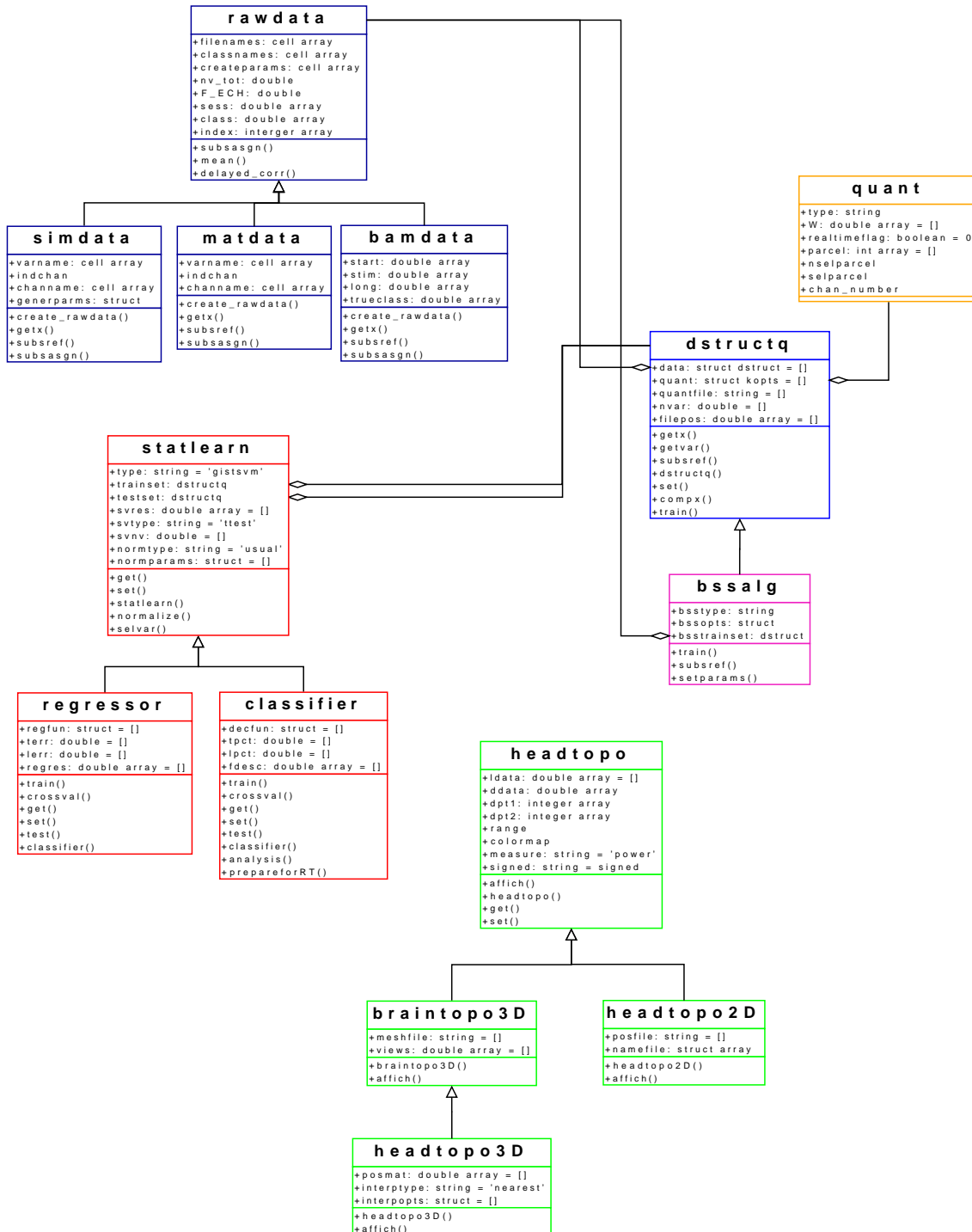



FIGURE 8.1 – Structure de la librairie orientée objet pour le traitement des données ICM

Index

- aire sous la courbe de ROC, 27
- ALD, 25, 81
- ALS, 3
- Analyse
 - en composantes indépendantes (ACI), 7, 33
 - en composantes principales (ACP), 7, 33
- Analyse linéaire discriminante, 8, 25
- BCI, 1
- BEM, 38
- BOLD, 54
- Boundary Element Method, 38

- calcul mental, 21
- champ sensibilité, 37
- cohérence, 16
- Common Spatial Patterns, 33, 81
- courbe de ROC, 27
- CSP, 33, 81

- dipôle de courant, 36
- décussation, 2

- EEG, 5
- ensemble
 - d'apprentissage, 23
- event-related
 - desynchronization, 14
 - synchronization, 14

- FEM, 38
- Finite Element Method, 38
- fonction discriminante, 23
- frontière de décision, 23

- généralisation, 29

- homoncule, 19

- ICM, 1
 - asynchrone, 1
 - synchrone, 1
- imagerie
 - auditive, 20
 - mentale, 18
 - motrice, 18
 - visuo-spatiale, 20
- Imagerie par Résonance Magnétique, 54
- intégrales de frontières, 38
- IRM, 54
- IRMf, 5

- K nearest neighbors, 26
- K plus proches voisins, 26
- KNN, 26
- KPPV, 26

- L2-AROM, 91
- LDA, 25
- lead field, 37
- LIS, 2
- locked-in syndrome, 2

- MEG, 5
- minimum norme, 41
- Modèle de têtes, 38

- neurones pyramidaux, 36

- patterns spatiaux communs, 33, 81
- PE, 13
- Phase Locking Value, 16
- PLV, 16
- potentiel évoqué, 13
- problème direct, 35

problème inverse, 39

ROC, 27

Résonance Magnétique Nucléaire, 54

sclérose latérale amyotrophique, 3

score de Fisher, 31

Support Vector Machine, 25

sur apprentissage, 29

SVM, 25

- de norme 0, 91

synchronie de phase, 16

syndrome d'enfermement, 2

Séparateur à vaste marge, 25

T1 (pondération en), 54

taux de bien classés, 27

tri de variables, 31

validation croisée, 30, 62

- “k fois”, 30
- inter-jours, 62
- inter-sessions, 62
- leave one out, 30

variable ranking, 31

éléments finis, 38

Bibliographie

- Altman A and Gondzio J. Regularized symmetric indefinite systems in interior point methods for linear and quadratic optimization, 1998. URL citeseer.ifi.unizh.ch/altman98regularized.html. 26
- Anderson C, Stolz E, and Shamsunder S. Multivariate autoregressive models for classification of spontaneous electroencephalogram during mental tasks. *IEEE Trans. Biomed. Eng.*, 45(3) :277–286, 1998. 15, 18, 27
- Arridge SR. Optical tomography in medical imaging. *Inverse Problems*, 15 :R41–R93, 1999. 7
- Baillet S, Mosher J, and Leahy R. Electromagnetic brain mapping. *IEEE Signal. Proc Mag*, 18 :14–30, 2001. 37, 42
- Barret G. Event-related potentials (ERPs) as a measure of complex cognitive function. *Electroencephalogr Clin Neurophysiol*, 46 :53–63, 1996. 13
- Bartlett P and Mendelson S. Rademacher and gaussian complexities : Risk bounds and structural results. *Journal of Machine Learning Research*, 3 :463–482, 2002. 31
- Bashashati A, Fatourechi M, Ward R, and Birch GE. A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals. *J. Neural. Eng.*, 4 :R32–R57, 2007. 13, 14
- Beisteiner R, Hollinger P, Lindinger G, Lang W, and Berthoz A. Mental representations of movements. brain potentials associated with imagination of hand movements. *Electroenceph. Clin. Neurophysiol.*, 96 :183–193, 1995. 19
- Bell A and Sejnowsky TJ. The independent components of natural scenes are edge filter. *Vision Research*, 36 :287–314, 1997. 33
- Bellman R. *Dynamic Programming*. Princeton University Press, 1957. 31
- Belouchrani A, Abed-Meraim K, Cardoso JF, and Moulines E. A blind source separation technique using second-order statistics. *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, 45(2) :434–444, Feb. 1997. 33, 75, 76

- Besserve M, Garnero L, and Martinerie J. De l'estimation à la classification des activités corticales pour les Interfaces Cerveau Machines. In *21ème colloque GRETSI sur le traitement du signal et des images*, 2007a. [93](#), [94](#)
- Besserve M, Jerbi K, Garnero L, and Martinerie J. Prediction of cognitive states using MEG and blind source separation. *International Congress Series*, 1300 :205–208, 2007b. [115](#)
- Birbaumer N, Ghanayim N, Hinterberger T, Iversen I, Kotchoubey A, and Kubler A. A spelling device for the paralyzed. *Nature*, 398 :297–298, 1999. [15](#), [16](#)
- Blumer A, Ehrenfeucht A, Haussler D, and Warmuth MK. Learnability and the vapnik-chervonenkis dimension. *Journal of the ACM*, 36 :929–865, 1989. [31](#)
- Borisoff JF, Mason SG, Bashashati A, and Birch GE. Brain-computer interface design for asynchronous control applications : Improvements to the LF-ASD asynchronous brain switch. *IEEE Trans. Biomed. Eng.*, 51(6) :985–992, 2004. [2](#)
- Bostanov V. BCI competition 2003-data sets ib and iib : feature extraction from event-related brain potentials with the continuous wavelet transform and the t-value scalogram. *IEEE Trans. Biomed. Eng.*, 51 :1057–1061, 2004. [25](#)
- Brillinger DR. Time series : Data analysis and theory. In *Classics in Applied Mathematics*. SIAM, Philadelphia., 2001. [62](#)
- Brovelli A, Ding M, Ledberg A, Chen Y, Nakamura R, , and Bressler S. Beta oscillations in a large-scale sensorimotor cortical network : directional influences revealed by granger causality. *Proc Natl Acad Sci USA*, 101 :9849–9854, 2004. [13](#)
- Butz M, Timmermann L, Gross J, Pollok B, Dirks M, Hefter H, and Schnitzler A. Oscillatory coupling in writing and writer's cramp. *J Physiol Paris*, 99 :14–20, 2006. [36](#)
- Cantero JL, Atienza M, and Salas RM. Human alpha oscillations in wakefulness, drowsiness period, and REM sleep : different electroencephalographic phenomena within the alpha band. *Neurophysiol. Clin.*, 32(1) :54–71, 2002. 0987-7053 Journal Article Review. [14](#)
- Cardoso J and A. Souloumiac A. Blind beamforming for non gaussian signals. *IEE - Proceedings -F*, 140 :362–370, 1993. [33](#)
- Cardoso J and Pham DT. Séparation de sources par l'indépendance et la parcimonie. *Colloque GRETSI*, 2003. [83](#)
- Cardoso JF and Souloumiac A. Jacobi angles for simultaneous diagonalization. *SIAM Journal on Matrix Analysis and Applications*, 17(1) :161–164, 1996. [75](#), [76](#)

- Cassim F, Monaca C, Szurhaj W, Bourriez J, Defebvre L, Derambure P, and Guieu J. Does post-movement beta synchronization reflect an idling motor cortex? *Neuroreport*, 12 :3859–3863, 2001. [14](#)
- Chavez M, Besserve M, Adam C, and Martinerie J. Towards a proper estimation of phase synchronization from time series. *Journal of Neuroscience Methods*, 154 :149–160, 2006. [62](#)
- Cincotti F, Mattia D, Aloise F, Bufalari S, Astolfi L, De Vico Fallani F, Tocci A, Bianchi L, Marcian MG, Gao S, Millan J, and Babiloni F. High-resolution EEG techniques for brain-computer interface applications. *Journal of Neuroscience Methods*, 167 : 31–42, 2008. [44](#)
- Comon P. Independant component analysis - a new concept? *Signal processing*, 36 : 287–314, 1994. [33](#), [76](#)
- Congedo M, Lotte F, and Lécuyer A. Classification of movement intention by spatially filtered electromagnetic inverse solutions. *Phys. Med. Biol.*, 51 :1971–1989, 2006. [44](#)
- Connolly J, Andersen R, and Goodale M. FMRI evidence for a parietal reach region in the human brain. *Exp Brain Res*, 153 :140–145, 2003. [116](#)
- Cottureau B, Jerbi K, and Baillet S. Multiresolution imaging of MEG cortical sources using an explicit piecewise model. *NeuroImage*, 38 :439–451, 2007. [42](#)
- Curran E, Sykacek P, Stokes M, Roberts W S J ans Penny, Johnsrude I, and Owen AM. Cognitive tasks for driving a brain-computer interfacing system : A pilot study. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 12 :48–54, 2003. [15](#), [16](#), [19](#), [116](#)
- Dale A, Liu A, Fischl B, Buckner R, Belliveau J, Lewine J, and Halgren E. Dynamic statistical parametric mapping : combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron*, 26 :55–67, 2000. [43](#)
- Dale A and Sereno M. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction - a linear approach. *J. Cogn. Neurosci*, 5 :162–176, 1993. [89](#), [91](#)
- David O and Garnero L. Time-coherent expansion of MEG/EEG cortical sources. *NeuroImage*, 17 :1277–1289, 2002. [44](#)
- Decety J and Michel F. Comparative analysis of actual and mental movement times in two graphic tasks. *Brain Cogn*, 11 :87–97, 1989. [19](#)
- Deecke L and Kornhuber H. *Human freedom, reasoned will, and the brain. The Bereitschaftspotential-Story*, volume 33. Kluwer Academic / Plenum Publishers, 2003. [14](#)

- Dehaene S. Sources of mathematical thinking : behavioral and brain imaging evidence. *Science*, 284 :970–974, 1999. 21
- Donoghue JP. Connecting cortex to machines : recent advances in brain interfaces. *Nat. Neurosci.*, 5 :1085–1088, 2002. 1, 18
- Donoho DL. High-dimensional data analysis : the curses and blessings of dimensionality. *Aide-Memoire of the lecture in AMS conference : Math challenges of 21st Century*, 2000. 116
- Dossevi A, Cosmelli D, Garnero L, and Ammari H. Multivariate reconstruction of functional networks from cortical sources dynamics in MEG/EEG. *to appear in IEEE Trans. Biomed. Eng.*, 2008. 44
- Duda R, Hart P, and Stork D. *Pattern Classification*. Wiley-Interscience, 2001. 23, 25, 27, 29
- Ermer JJ, Mosher J, Baillet S, and Leahy R. Rapidly recomputable EEG forward models for realistic head shapes. *Phys Med Biol*, 46(4) :1265–1281, Apr 2001. 39
- Farwell L and Donchin E. Talking off the top of your head : toward a mental prosthesis utilizing event-related brain potentials. *Electroenceph. Clin. Neurophysiol.*, 70 :510–523, 1988. 13
- Fisher RA. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7 :179–188, 1936. 25
- Friston KJ. Functional and effective connectivity in neuroimaging : a synthesis. *Hum. Brain Mapp.*, 2 :56–78, 1994. 16
- Fuchs M, Wagner M, Kohler T, and Wischmann H. Linear and nonlinear current density reconstructions. *J. Clinical Neurophysiol.*, 16 :267–295, 1999. 42
- Fuchs T, Birbaumer N, Lutzenberger W, Gruzelier J, and Kaiser J. Neurofeedback treatment for attention-deficit/hyperactivity disorder in children : A comparison with methylphenidate. *Applied Psychophysiology and Biofeedback*, 28 :1–12, 2003. 4
- Garrett D, Peterson DA, Anderson CW, and Thaut MH. Comparison of linear, nonlinear, and feature selection methods for EEG signal classification. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 11 :141–144, 2003. 25
- Gavit L, Baillet S, Mangin JF, Pescatore J, and Garnero L. A multiresolution framework to MEG/EEG source imaging. *IEEE Trans. Biomed. Eng.*, 48 :1080–1087, 2001. 42
- Gloor P. Neuronal generators and the problem of localization in electroencephalography : Application of volume conductor theory to electroencephalography. *J Clin Neurophysiol*, 2 :327–354, 1985. 36

- Grasman RPPP, Huizenga HM, Waldorp LJ, Böcker KBE, and Molenaar PCM. Frequency domain simultaneous source and source coherence estimation with an application to MEG. *IEEE Trans. Biomed. Eng.*, 51 :45–55, 2004. 44
- Grave de Peralta Menendez R, Gonzalez Andino S, Lucas Perez L, Ferrez P, and del R. Millan J. Non-invasive estimation of local field potentials for neuroprosthesis control. *Cogn Process*, 6 :59–64, 2005. 44
- Gray C and Singer W. Stimulus specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. USA*, 86(5) :1698–1702, 1989. 13
- Green D and Swets J. *Signal Detection Theory and Psychophysics*. Wiley, New York, 1974. 29
- Gribonval R and Nielsen M. Sparse decompositions in unions of bases. *IEEE Trans. Inf. Th.*, 49 :3320–325, 2003. 83
- Gross J, Kujala J, Hamalainen M, Timmermann L, Schnitzler A, and Salmelin R. Dynamic imaging of coherent sources : studying neural interactions in the human brain. *Proc. Natl. Acad. Sci. U. S. A.*, 98 :694–699, 2001. 43, 44
- Gross J, Timmermann L, Kujala J, Dirks M, Schmitz F, Salmelin R, and Schnitzler A. The neural basis of intermittent motor control in humans. *Proc Natl Acad Sci USA*, 99 :2299–2302, 2002. 36
- Guyon I and Elisseeff A. An introduction to variable and feature selection. *J. Mach. Learn. Res.*, 3 :1157–1182, 2003. 32, 33, 93
- Gysels E and Celka P. Phase synchronisation for the recognition of mental tasks in a Brain Computer Interface. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 12(4) :206–415, 2004. 17
- Gysels E, Renevey P, and Celka P. SVM-based recursive feature elimination to compare phase synchronization computed from broadband and narrowband EEG signals in Brain-Computer Interfaces. *Signal Processing*, 85(11) :2178–2189, 2005. 17
- Haselsteiner E and Pfurtscheller G. Using time-dependent neural networks for EEG classification. *IEEE Trans. Rehab. Eng.*, 8 :457–463, 2000. 27
- Hastie T, Tibshirani R, and Friedman J. *The Elements of Statistical Learning*. Springer, 2001. 26, 27, 30
- Haueisen J, Ramon C, Czapski P, and Eiselt M. On the influence of volume currents and extended sources on neuromagnetic fields : a simulation study. *Annals of Biomed. Eng.*, 23 :728–739, 1995. 39
- Hiraiwa A, Shimohara K, and Tokunaga Y. EEG topography recognition by neural networks. *IEEE Eng. Med. Biol. Mag.*, 9 :39–42, 1990. 27

- Hämäläinen M, Hari R, Ilmoniemi R, Knuutila J, and Lounasmaa O. Magnetoencephalography. theory, instrumentation and applications to the noninvasive study of human brain function. *Rev. Modern Phys.*, 65 :1993, 1993. [12](#)
- Hori J, Sakano K, and Saitoh Y. Development of a communication support device controlled by eye movements and voluntary eye blink. *IEICE Trans. Inf. Syst.*, E89-D :1790–1797, 2006. [4](#)
- Hérault J, Jutten C, and Ans B. Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage mono supervisé. *Colloque GRETSI*, pages 1017–1020, 1985. [33](#)
- Hyvärinen A and Oja E. A fast fixed-point algorithm for independent component analysis. *Neural Computation*, 9 :1483–1492, 1997. [33](#), [76](#)
- Jeannerod M and Decety J. Mental motor imagery : a window into the representational stages of action. *Curr. Opin. Neurobiol.*, 5 :727–732, 1995. [19](#)
- Jenrich RI. Stepwise discriminant analysis. In Enslein K, Ralston A, and Wilf HS, editors, *Statistical methods for digital computers*. John Wiley and Sons, 1960. [32](#)
- Jerbi K, Lachaux J, N'Diaye K, Pantazis D, Leahy R, Garnero L, and Baillet S. Coherent neural representation of hand speed in humans revealed by meg imaging. *Proc. Nat. Ac. Sci.*, 104 :7676–7681, 2007. [13](#), [44](#), [69](#), [77](#), [133](#)
- Kandel E, Schwartz JH, and Jessell T. *Principles of Neural Science*. McGraw Hill, 2000. [11](#)
- Kawamichi H, Kikuchi Y, and Uenoc S. Spatio-temporal brain activity related to rotation method during a mental rotation task of three-dimensional objects : An MEG study. *Neuroimage*, 37 :956–965, 2007. [20](#)
- Keirn ZA and Aunon JJ. A new mode of communication between man and his surroundings. *IEEE Trans. Biomed. Eng.*, 37 :1209–1214, 1990. [15](#), [18](#), [116](#)
- Klimesch W. EEG alpha and theta oscillations reflect cognitive and memory performance : a review and analysis. *Brain Res. Brain Res. Rev.*, 29(2-3) :169–95, 1999. 0165-0173 Journal Article Review. [14](#), [69](#)
- Kosslyn SM, Ganis G, and Thompson WL. Neural foundations of imagery. *Nat Rev Neurosci*, 2 :635–642, 2001. [18](#), [19](#), [20](#)
- Krepki R. Berlin brain-computer interface. the HCI communication channel discovery. *Int. J. Human-computer Studies*, 65 :460–477, 2006. [14](#), [15](#), [29](#)
- Lachaux J, Rodriguez E, Martinerie J, and Varela F. Measuring phase synchrony in brain signals. *Hum. Brain Mapp.*, 8 :194–208, 1999. [13](#), [16](#)

- Lachaux JP, Rodriguez E, Le Van Quyen M, Lutz A, Martinerie J, and Varela Fa. Studying single-trials of phase synchronous activity in the brain. *International Journal of Bifurcation and Chaos*, 10 :2429–2439, 2000. [70](#)
- Laurent F, Besserve M, Florence G, and Martinerie J. Apport de la reconstruction de sources en EEG pour la détection d'états de fatigue mentale. *Colloque GRETSI*, pages 629–632, 2007. [120](#)
- Le Van Quyen M, Soss J, Navarro V, Robertson R, Chavez M, Baulac M, and Martinerie J. Preictal state identification by synchronization changes in long-term intracranial EEG recordings. *Clin. Neurophysiol.*, 116(3) :559–568, 2005. [17](#)
- Lebedev M and Nicolelis M. Brain-machine interfaces : past, present and future. *Trends Neurosci.*, 29 :536–546, 2006. [1](#), [6](#)
- Leeb R, Friedman D, Muller-Putz G, Scherer R, Slater M, and Pfurtscheller G. Self-paced (asynchronous) BCI control of a wheelchair in virtual environments : A case study with a tetraplegic. *Computational Intelligence and Neuroscience*, 2007. [18](#)
- Lopes Da Silva F. Event related oscillations : what about phase? In *Event-Related Dynamics of Brain Oscillations*, chapter 1, pages 19–27. Elsevier, 2006. [14](#)
- Lotte F, Congedo M, Lécuyer A, and Arnaldi B. A review of classification algorithms for EEG-based Brain-Computer Interfaces. *J. Neural. Eng.*, 4 :R1–R13, 2007a. [27](#)
- Lotte F, Lécuyer A, and Arnaldi B. FuRIA : A novel feature extraction algorithm for brain-computer interfaces using inverse models and fuzzy region of interest. *Proc. of the 3rd IEEE-EMBS International Conference on Neural Engineering*, pages 175–178, 2007b. [44](#)
- McFarland D, McCane L, Davi S, and Wolpaw J. Spatial filter selection for EEG-based communication. *Electroencephalogr. Clin. Neurophysiol.*, 103 :386–394, 1997. [7](#), [60](#)
- Meijs J, Weier O, Peters M, and van Oosterom A. On the numerical accuracy of the boundary element method. *IEEE Trans. on Biomed. Eng.*, 36 :1038–1049, 1989. [39](#)
- Middendorf M, Mcmillan G, Calhoun G, and Jones KS. Brain-computer interfaces based on the steady-state visual-evoked response. *IEEE Trans. Rehabil. Eng.*, 8 : 211–214, 2000. [15](#)
- Millan del R. J, Renkens F, Mouriño J, and Gerstner W. Noninvasive brain-actuated control of a mobile robot by human eeg. *IEEE Trans. Biomed. Eng.*, 51(6) :1026–1033, 2004. [15](#), [18](#)
- Mosher J, Lewis P, and Leahy R. Multiple dipole modelling and localization from spatiotemporal meg data. *IEEE Trans. Biomed. Eng.*, 39 :541–557, 1992. [41](#)

- Muller K, Mika S, Ratsch G, Tsuda K, and Scholkopf. An introduction to kernel-based learning algorithms. *IEEE Tran. Neur. Netw.*, 12 :181–202, 2001. [26](#)
- Muller-Gerking J, Pfurtscheller G, and Flyvbjerg H. Designing optimal filters for single-trial EEG classification in a movement task. *Clin. Neurophysiol.*, 110 :787–798, 1999. [7](#), [33](#)
- Naeem M, Brunner C, Leeb R, Graimann B, and Pfurtscheller G. Separability of four-class motor imagery data using independent components analysis. *J. Neural Eng.*, 3 :208–216, 2006. [34](#)
- Naranjo J, Brovelli A, Longo R, Budai R, Kristeva R, and Battaglini P. EEG dynamics of the frontoparietal network during reaching preparation in humans. *Neuroimage*, 34 :1673–1682, 2007. [116](#)
- Nelson LA. The role of biofeedback in stroke rehabilitation : Past and future directions. *Topics in Stroke Rehabilitation*, 14 :59 – 66, 2007. [4](#)
- Neuper C, Muller GR, Kubler A, Birbaumer N, and Pfurtscheller G. Clinical application of an EEG-based brain-computer interface : a case study in a patient with severe motor impairment. *Clin Neurophysiol*, 114 :399–409, 2003. [15](#)
- Neuper C and Pfurtscheller G. Motor imagery and ERD. In Pfurtscheller G and Lopes da Silva FH, editors, *Handbook of Electroencephalography and Clinical Neurophysiology*, volume 6, chapter 20, pages 303–325. Elsevier, 1999. [19](#)
- Nunez P. Electric fields of the brain : the neurophysics of EEG. New York, 1981. [12](#), [36](#)
- Nunez PL and Silberstein RB. On the relationship of synaptic activity to macroscopic measurements : does co-registration of EEG with fMRI make sense? *Brain Topogr*, 13 :79–96, 2000. [36](#)
- Obermeier B, Guger C, Neuper C, and Pfurtscheller G. Hidden markov models for online classification of single trial EEG. *Pattern Recognit. Lett.*, 22 :1299–1309, 2001. [27](#)
- Pantazis D, Nichols TE, Baillet S, and Leahy RM. Spatiotemporal localization of significant activation in MEG using permutation tests. *Inf. Process. Med. Imaging*, 18 :512–23, 2003. 1011-2499 Journal Article. [32](#), [42](#), [96](#)
- Pascual-Marqui R. Standardized low resolution brain electromagnetic tomography (sLORETA) : technical details. *Methods Find. Exp. Clin. Pharmacol.*, 24 :5–12, 2002. [43](#), [44](#)
- Pfurtscheller G. The cortical activation model (cam). In *Event-Related Dynamics of Brain Oscillations*, chapter 2, pages 19–27. Elsevier, 2006. [15](#)

- Pfurtscheller G and Aranibar A. Evaluation of event-related desynchronization (ERD) preceding and following voluntary self-paced movement. *Electroenceph. Clin. Neurophysiol.*, 46 :138–146, 1979. [14](#)
- Pfurtscheller G and Lopez Da Silva F. Event-related EEG/MEG synchronization and desynchronization : basic principles. *Clin. Neurophysiol.*, 110 :1842–1857, 1999. [14](#)
- Pfurtscheller G, Neuper C, Flotzinger D, and Prezenger M. EEG-based discrimination between imagination of right and left hand movement. *Electroencephalogr. Clin. Neurophysiol.*, 103 :642–651, 1997. [15](#)
- Pikovski A, Rosenblum M, and Kurths J. *Synchronization : A Universal Concept in Nonlinear Science*. Cambridge Univ. Press, 2001. [17](#)
- Qin L, Ding L, and He B. Motor imagery classification by means of source analysis for brain-computer interface applications. *J. Neural Eng.*, 1 :135–141, 2004. [44](#)
- Quiroga R, Snyder L, Batista A, Cui H, and Andersen R. Movement intention is better predicted than attention in the posterior parietal cortex. *Journal of neuroscience*, 26 :3615–3620, 2006. [6](#)
- Rakotomamonjy A. Variable selection using SVM-based criteria. *Journal of Machine Learning Research*, 3 :1357–1370, 2003. [32](#)
- Rodriguez E, George N, Lachaux J, Martinerie J, and Renault B. Perception shadow : long distance gamma band synchronisation and desynchronisation on the human scalp. *Nature*, 397 :430–433, 1999. [13](#), [17](#)
- Rosenblatt M. *Stationnary sequences and random fields*. Birkhauser, 1985. [70](#)
- Salenius M S, Kajola, Thimpson WL, Kosslyn S, and Hari R. Reactivity of magnetic parieto-occipital alpha rhythm during visual imagery. *Electroenceph. Clin. Neurophysiol.*, 95 :453–462, 1995. [20](#)
- Salmelin R and Hari R. Spatiotemporal characteristics of sensorimotor neuromagnetic rhythms related to thumb movement. *Neuroscience*, 60 :537–550, 1994. [14](#), [43](#)
- Sasaki K, Tsujimoto T, Nishikawa S, and Nishitani NTI. Frontal mental theta wave recorded simultaneously with magnetoencephalography and electroencephalography. *Neurosci. Res.*, 26 :79–81, 1996. [21](#)
- Scherer R, Müller GR, Neuper C, Graimann B, and Pfurtscheller G. An asynchronously controlled EEG-based virtual keyboard : Improvement of the spelling rate. *IEEE Trans. Biomed. Eng.*, 51(6) :979–984, 2004. [1](#), [18](#), [25](#)
- Scherg M and Cramon D. Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model. *Electroenceph. Clin. Neurophysiol.*, 62 :32–44, 1985. [40](#)

- Stinear CM, Byblow WD, Steyvers M, Levin O, and Swinnen SP. Kinesthetic, but not visual, motor imagery modulates corticomotor excitability. *Exp. Brain. Res.*, 168 : 157–164, 2006. [19](#), [117](#)
- Tallon-Baudry C and Bertrand O. Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn. Sci.*, 3(4) :151–162, 1999. [14](#)
- Tesche C and Karhu J. Somatosensory evoked magnetic fields arising from sources in the human cerebellum. *Brain Res.*, 744 :23–31, 1997. [36](#)
- Tesche C and Karhu J. Theta oscillations index human hippocampal activation during a working memory task. *Proc. Natl. Acad. Sci. USA*, 97 :919–924, 2000. [36](#)
- Thevenet M, Bertrand O, Perrin F, Dumont T, and Pernier J. The finite element method for a realistic head model of electrical brain activities : preliminary results. *Clin. Phys. Physiol. Meas*, 12 :89–94, 1991. [39](#)
- Thirion B, Duchesnay D, Hubbard E, Dubois J, Poline JB, Lebihan D, and Dehaene S. Inverse retinotopy : Inferring the visual content of images from brain activation patterns. *Neuroimage*, 33 :1104–1116, 2006. [6](#)
- Timmermann L, Gross J, Kircheis G, Haussinger D, and Schnitzler A. Cortical origin of mini-asterixis in hepatic encephalopathy. *Neurology*, 58 :295–298, 2002. [36](#)
- Tononi G. Consciousness, information integration, and the brain. *Prog. Brain Res.*, 150 :109–126, 2005. [17](#)
- Townsend G, Graimann B, and Pfurtscheller G. A comparison of common spatial patterns with complex band power features in a four-class BCI experiment. *IEEE Trans Biomed Eng*, 53 :642– 651, 2006. [27](#)
- Van Veen B, Van Drongelen W, Yuchtman M, and Suzuki A. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans Biomed Eng*, 44 :867–880, 1997. [41](#)
- Vapnik V. *Statistical Learning Theory*. Wiley-Interscience, New York, 1998. [25](#), [26](#), [63](#), [93](#), [94](#)
- Vapnik V and Chervonenkis A. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and its Applications*, 16 :264–280, 1971. [31](#)
- Varela F. Resonant cell assemblies : a new approach to cognitive function and neuronal synchrony. *Biol. Res.*, 28 :81–95, 1995. [17](#)
- Varela F, Lachaux J, Rodriguez E, and Martinerie J. The brainweb : Phase synchronisation and large-scale integration. *Nature Rev. Neurosci.*, 2 :229–239, 2001. [13](#), [16](#)

- Wackermann J, Lehmann D, Michel C, and Strik W. Adaptive segmentation of spontaneous EEG map series into spatially defined microstates. *International Journal of Psychophysiology*, 14 :269–283, 1993. [117](#)
- Wallstrom GL, Kass RE, Miller A, Cohn JF, and Fox NA. Automatic correction of ocular artifacts in the EEG : a comparison of regression-based and component-based methods. *Int. J. Psychophysiol.*, 53 :105–19, 2004. [60](#)
- Wei Q, Wang Y, Gao X, and Gao S. Amplitude and phase coupling measures for feature extraction in an EEG-based brain-computer interface. *J. Neural Eng.*, 4 :120–129, 2007. [17](#)
- Wessberg J, Stambaugh C, Kralik J, Beck P, Laubach M, Chapin J, Kim J, Biggs S, Srinivasan M, and Nicolelis M. Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature*, 16 :361–365, 2000. [6](#)
- Weston J, Elisseeff A, Schölkopf B, and Tipping M. Use of the zero norm with linear models and kernel methods. *J. Mach. Learn. Res.*, 3 :1439–1461, 2003. ISSN 1533-7928. [32](#), [93](#)
- Wilhelm B, Jordan M, and Birbaumer N. Communication in locked-in syndrome : Effects of imagery on salivary pH. *Neurology*, 67 :534–540, 2006. [3](#)
- Wolpaw JR, McFarland DJ, Neat GW, and Forneris CA. An EEG-based brain-computer interface for cursor control. *Electroenceph. Clin. Neurophysiol.*, 78 :252–259, 1991. [15](#), [16](#)
- Wolpert DH. Stacked generalization. *Neural Netw.*, 5(2) :241–259, 1992. [119](#)
- Yvert B, Bertrand O, Echallier J, and Pernier J. Improved forward EEG calculations using local mesh refinement of realistic head geometries. *Electroenceph. Clin. Neurophysiol.*, 95 :381–392, 1995. [39](#)
- Zago L, Pesenti M, Mellet E, Crivello F, Mazoyer B, and Tzourio-Mazoyer N. Neural correlates of simple and complex mental calculation. *Neuroimage*, 13 :314–327, 2001. [21](#)
- Zatorre RJ, Halpern AR, Perry DW, Meyer E, and Evans AC. Hearing in the mind’s ear : a pet investigation of musical imagery and perception. *J. Cogn. Neurosci.*, 8 : 29–46, 1996. [21](#)