



HAL
open science

BCI exploitant les SSVEP : étude de la stabilité du classifieur dans le temps et des effets de l'apprentissage humain sur les performances de classification

Alban Duprès, François Cabestaing, José Rouillard

► To cite this version:

Alban Duprès, François Cabestaing, José Rouillard. BCI exploitant les SSVEP : étude de la stabilité du classifieur dans le temps et des effets de l'apprentissage humain sur les performances de classification. Handicap 2014, Jun 2014, Paris, France. hal-01361912

HAL Id: hal-01361912

<https://hal.science/hal-01361912>

Submitted on 7 Sep 2016

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.

BCI exploitant les SSVEP : étude de la stabilité du classifieur dans le temps et des effets de l'apprentissage humain sur les performances de classification

Alban Duprès
Université de Lille1
alban.dupres@univ-lille1.fr

François Cabestaing
Université de Lille1
francois.cabestaing@univ-lille1.fr

José Rouillard
Université de Lille1
jose.rouillard@univ-lille1.fr

Résumé—Les interfaces cerveau-ordinateur (ou BCI pour Brain-Computer Interfaces) exploitant les potentiels évoqués visuels stables (ou SSVEP pour Steady-State Visual Evoked Potentials) permettent le contrôle d'une application par l'intermédiaire de stimuli visuels clignotant à différentes fréquences. Cette modalité d'interaction pourrait permettre à des personnes souffrant d'un handicap moteur sévère d'améliorer leur qualité de vie en regagnant une autonomie partielle. D'après les travaux relatés dans la littérature, chaque session d'utilisation d'une BCI-SSVEP intègre une phase de calibrage visant notamment à régler les paramètres d'un classifieur. Notre objectif est d'évaluer si les mêmes paramètres peuvent être utilisés pendant plusieurs sessions, ce qui éviterait d'introduire systématiquement une phase de calibrage, très contraignante pour l'utilisateur. Pour ce faire, nous analysons la stabilité dans le temps des résultats de classification. D'autre part, les données acquises sont mises à profit pour étudier dans une seconde étude les effets de l'apprentissage humain sur les performances de l'interface et de confirmer ou non l'état de l'art à ce sujet. D'après la littérature, les BCI-SSVEP fonctionnent correctement dès la première utilisation et leurs performances ne s'améliorent pas avec l'expérience de l'utilisateur.

Mots clés—Interface cerveau-ordinateur (SSVEP); apprentissage; évaluation; classifieur.

I. INTRODUCTION

Une BCI est un système permettant une communication directe entre une machine (bras robotisé, fauteuil roulant) ou un ordinateur et le cerveau d'un utilisateur, sans passer par l'intermédiaire des nerfs ou muscles périphériques. Ce type d'interface peut s'avérer utile pour des patients handicapés par une maladie dégénérative (myopathie de Duchenne, sclérose latérale amyotrophique) ou une lésion haute de la moelle épinière, en leur permettant de retrouver une autonomie partielle et d'améliorer leur qualité de vie.

D'autres technologies éprouvées ont précédé les BCI dans ce domaine : suivi du regard [1] ou de la tête [2], interface exploitant l'électromyographie [3] ou l'électro-oculographie [4], boule de commande (ou trackball) [5], joystick, contacteurs, système de reconnaissance de voix [6]. Les travaux relatés dans la littérature semblent indiquer que ces techniques permettent pour l'instant un contrôle plus fiable et plus rapide

que l'approche BCI [7], [8].

Il serait logique de privilégier ces techniques dans le cadre de la palliation du handicap au détriment des BCI. Néanmoins, les techniques précédemment mentionnées présentent chez des patients très lourdement handicapés un certain nombre de limites auxquelles les BCI ne sont pas confrontées. La principale réside dans la nécessité de disposer d'une activité musculaire résiduelle, même minimale. Or certaines personnes souffrant d'un syndrome d'enfermement (ou LIS pour Locked-In-Syndrom) sont incapables d'interagir avec un ordinateur ou une machine par le biais de ces techniques. De plus, même pour certains patients capables d'utiliser des dispositifs usuels d'assistance, une BCI peut présenter des avantages dans certaines applications. Par exemple, le pilotage d'un fauteuil roulant est plus intuitif et mieux adapté avec une BCI exploitant l'imagination motrice [9], qu'avec un système de suivi du regard.

Pour des personnes totalement paralysées, mais dont les capacités cognitives sont intactes, les BCI sont le seul moyen de communication avec leur environnement. C'est pourquoi elles ont leur place dans le cadre de l'assistance du handicap moteur sévère et ne doivent pas être négligées lors de la prise en charge de patients lourdement handicapés.

Une interface cerveau-ordinateur intègre six étapes fonctionnelles (voir figure 1) :

- 1) *mesure de l'activité cérébrale*. Des systèmes d'acquisition, plus ou moins invasifs, sont disponibles : électroencéphalographie (EEG), électrocorticographie [10], interface neuronale directe [11], magnéto-encéphalographie. Pour notre étude, nous utilisons l'électroencéphalographie de surface, où les signaux cérébraux sont enregistrés par des électrodes placées sur le scalp.
- 2) *pré-traitement des signaux*. Cette étape vise à éliminer le bruit contenu dans les signaux, lequel peut provenir du réseau électrique (50 Hz), de l'activité musculaire ou oculaire et de l'activité cérébrale de fond (rythme α) dont la fréquence couvre celle des indices électrophysiologiques. Cette étape est indispensable pour mettre en évidence les motifs électrophysiologiques

liés à une tâche mentale particulière.

- 3) *extraction de caractéristiques*, qui consiste à transformer les signaux filtrés en vecteurs de caractéristiques. Ces derniers contiennent des valeurs pertinentes pour la classification des signaux. Par exemple, on peut définir la puissance du signal pour une fréquence particulière comme valeur caractéristique.
- 4) *classification*, qui détermine la classe à laquelle appartient l'état mental mesuré au travers du signal en fonction des valeurs du vecteur de caractéristiques. L'étape de classification peut s'effectuer en temps-réel pour une application mais également hors-ligne pour tester les performances de différents classificateurs.
- 5) *commande de l'ordinateur*, qui s'effectue lorsque le système a identifié grâce à l'étape de classification la tâche mentale réalisée par le sujet, comme par exemple imaginer un mouvement des pieds. Chaque tâche mentale est associée à une commande spécifique.
- 6) *retour perceptif (ou feedback)*, qui permet enfin à l'utilisateur de savoir comment le système a interprété son état mental.

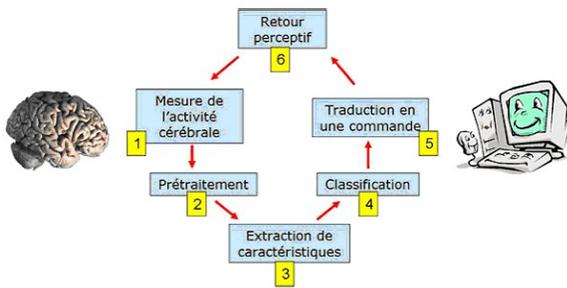


FIGURE 1: Les six étapes impliquées dans le fonctionnement d'une BCI (source [12]).

On parle d'interface asynchrone lorsque les motifs électrophysiologiques sont générés par l'utilisateur lui-même ou synchrone lorsqu'ils sont provoqués par une stimulation externe. Les BCI asynchrones fonctionnent selon deux modes : l'imagination motrice (ou IM) et les potentiels corticaux lents (ou SCP pour Slow Cortical Potentials). Les BCI synchrones exploitent deux types de motifs cérébraux : les potentiels évoqués stables (SSVEP) et transitoires (P300). [13] présente une description détaillée des BCI et de leurs différents paradigmes. Dans la suite de notre étude nous nous focaliserons sur les BCI basées sur les SSVEP.

Le SSVEP est un signal qui apparaît au niveau du cortex occipital en réponse à une stimulation visuelle répétitive. Le lobe occipital est le centre d'intégration de l'information visuelle. Ce type de signal a comme propriété d'osciller à la même fréquence que celle de la stimulation, de telle sorte que l'on peut identifier le stimulus que l'utilisateur a observé et ainsi déclencher la commande associée.

L'utilisation d'une BCI-SSVEP nécessite une première phase dite de calibrage pour ajuster le classifieur aux

spécificités de chaque utilisateur. La phase de calibrage requiert une concentration importante de la part de l'utilisateur afin d'obtenir un classifieur robuste et de maximiser les performances du système. De plus, cette phase de calibrage est renouvelée à chaque utilisation. Cela a pour conséquence une fatigue supplémentaire, un ennui et une perte de temps pour l'utilisateur. A notre connaissance aucun article ne semble avoir abordé ce problème de redondance de la phase de calibrage. C'est pourquoi nous allons étudier la stabilité du classifieur dans le temps. L'objectif de la première partie de l'article est d'évaluer si les coefficients d'un classifieur spécifiquement appris pour une personne peuvent être réutilisés par cette même personne plusieurs jours après leur calcul, en vue de réduire le nombre de phases de calibrage et ainsi de limiter les inconvénients liés à cette étape.

D'autre part, d'après la littérature, les BCI de type SSVEP fonctionnent correctement dès la première utilisation [14], [15]. Aucune amélioration des performances ne semble être observée avec l'expérience de l'utilisateur, contrairement à d'autres paradigmes tels l'imagination motrice [16] (IM) ou les potentiels corticaux lents [17] (ou SCP pour Slow Cortical Potentials) pour lesquels plusieurs semaines voire plusieurs mois sont nécessaires pour contrôler de manière fiable une application. C'est pourquoi, profitant des données recueillies pour répondre à la première problématique, nous étudierons les effets de l'apprentissage humain sur les performances d'une BCI-SSVEP afin de confirmer ou non l'état de l'art.

Enfin, nous cherchons à évaluer, sur des données recueillies sur des personnes valides, si les BCI-SSVEP peuvent être pertinentes pour la palliation du handicap. L'objectif est de tester si le SSVEP pourrait concurrencer les autres paradigmes de BCI, sachant que l'utilité, dans le cadre de l'assistance aux personnes handicapées, de l'IM comme du P300 a déjà été démontrée dans [18] et [19].

II. ACQUISITION DES DONNÉES ET PROTOCOLE EXPERIMENTAL

L'expérience a été réalisée par seize sujets sains (12 hommes et 4 femmes) âgés de 20 à 52 ans. Le consentement éclairé des sujets a été recueilli oralement. De plus, ils ont été informés au préalable du contenu de l'expérience et de la possibilité d'arrêter la session à tout moment. Deux participants (sujets 9 et 12) avaient déjà utilisé une interface cerveau-ordinateur de type SSVEP avant cette expérience. Sept sujets ont une vision normale et les neuf autres possèdent un dispositif de correction de la vue. Les participants sont installés confortablement dans un fauteuil pour qu'ils soient le plus relâché possible et pour limiter les artefacts musculaires. Il leur a également été demandé d'éviter de cligner des yeux pendant l'expérience pour ne pas enregistrer d'artefacts oculaires. On fixe la distance personne/écran à 70 cm et la luminosité de la pièce est diminuée pour amplifier le signal SSVEP. En effet, nos premières tentatives en plein jour n'ont pas permis de générer un SSVEP suffisamment ample pour contrôler le système de manière fiable.

On installe sur la tête de chaque sujet un bonnet à électrodes (GAMMAcap, g.tec), de telle sorte que les électrodes Ag/AgCl soient positionnées selon le système international 10/20 [20]. Quatre canaux monopolaires sont enregistrés au niveau du lobe occipital : O_z , O_1 , O_2 et PO_z , avec comme référence une électrode placée sur le lobe de l'oreille droite et comme masse une électrode frontale. Un gel est appliqué entre la peau et les électrodes pour augmenter la conduction du signal électrique. Les signaux sont amplifiés, filtrés entre 0.1 et 100 Hz — le 50 Hz étant éliminé pour supprimer les artefacts liés au réseau électrique — puis échantillonnés à la fréquence F_e , avec $F_e = 512$ Hz.

Le système d'enregistrement est composé d'un amplificateur de signaux physiologiques 16 canaux (g.USBamp, g.tec) et d'un ordinateur portable Dell utilisant le système d'exploitation Windows XP. Un écran différent de celui de l'ordinateur d'acquisition est utilisé pour afficher les stimuli. L'ensemble des interfaces graphiques présentées dans cet article sont gérées par le logiciel OpenViBE¹, qui permet également de contrôler en temps réel l'acquisition du signal, le traitement du signal et l'envoi de commandes pour le contrôle de l'application. La figure 2 présente la salle d'expérimentation et l'installation du matériel nécessaire à la mise en place d'une BCI, utilisés lors de nos expériences.

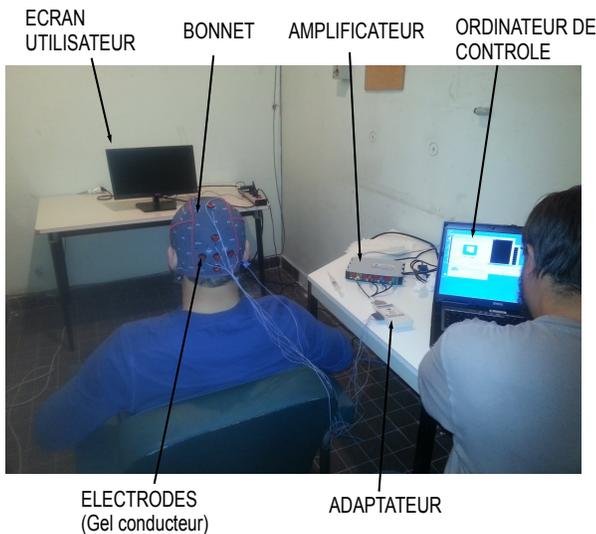


FIGURE 2: Mise en place d'une BCI.

Chaque sujet a réalisé trois sessions sur dix jours, de telle sorte que la seconde et la troisième session soient menées respectivement deux et dix jours après la première, comme le montre le tableau I. La première session est constituée d'une phase de calibration suivie d'une phase, dite en ligne, durant laquelle le sujet contrôle une application ludique. Les sessions 2 et 3 sont composées de trois phases : une première phase en ligne utilisant le classifieur calculé lors de la première session puis une phase de calibration suivi d'une seconde phase en ligne avec le classifieur nouvellement appris. La durée de

chaque session varie de 30 à 45 minutes en fonction du temps nécessaire à l'utilisateur pour finir l'application.

Session 1 (<i>Jour 0</i>)	Session 2 (<i>Jour 2</i>)	Session 3 (<i>Jour 10</i>)
$C_1 A_1$	$A'_1 C_2 A_2$	$A''_1 C_3 A_3$

Tableau I: Contenu de chaque session. Les Lettres C et A signifient respectivement calibration et Application. A'_1 et A''_1 correspondent respectivement à l'application en ligne réalisée lors des sessions 2 et 3 avec le classifieur calculé lors de la première session.

La phase de calibration permet de calculer les coefficients du classifieur utilisés lors de l'interaction, en temps-réel, avec l'application (phase en ligne). Les données enregistrées pendant cette phase sont également utilisées pour évaluer les performances hors-ligne du classifieur. On parle de performances hors-ligne lorsque le classifieur est testé sur des données pré-enregistrées et non pas sur des données acquises en temps-réel. Pour l'étape de calibration quatre cibles sont présentées à l'écran. Trois d'entre elles clignotent à une fréquence constante tandis que la dernière cible reste figée, comme le montre la figure 3. Les fréquences de stimulation des cibles sont 6, 10 et 15 Hz. Ces fréquences ont été sélectionnées de telle sorte que chacune d'entre elles soit un sous-multiple de 60 Hz, ce qui correspond à la fréquence de rafraîchissement de l'écran. Les cibles sont de couleur blanche sur fond noir afin d'accroître le contraste. Ceci permet de stimuler davantage la rétine de l'utilisateur et donc d'augmenter l'amplitude du SSVEP. On demande aux participants de focaliser leur attention visuelle sur les cibles indiquées par une flèche jaune. Chaque essai dure 7 secondes et un temps de repos de 4 secondes est observé entre chaque essai. Un total de 32 essais (huit par cible) est réalisé par chaque sujet.

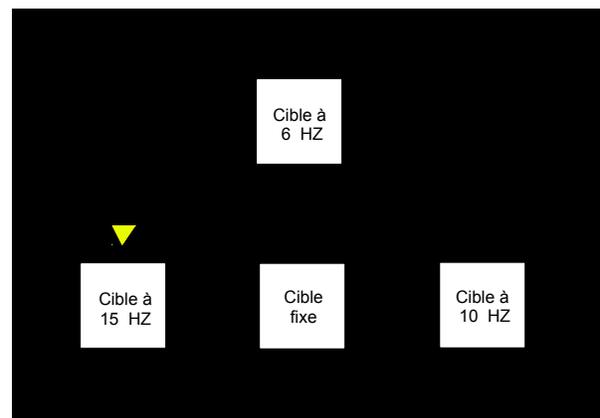


FIGURE 3: Phase de calibration.

La phase en ligne permet de contrôler une application ludique, proposée dans le logiciel OpenViBE. L'objectif de l'application est d'atteindre des cibles grâce à un "vaisseau spatial". Les cibles sont représentées par des cercles pleins,

1. openvibe.inria.fr

tandis que le “vaisseau spatial” est schématisé par un cercle vide sur lequel sont disposés de manière équidistante trois stimuli, 2 carrés et un triangle comme le représente la figure 4. En focalisant son regard sur le carré de gauche ou de droite l'utilisateur est capable de faire pivoter le “vaisseau spatial” respectivement dans le sens horaire ou anti-horaire. En regardant le triangle, l'utilisateur peut tirer afin d'atteindre la cible qui disparaît, puis réapparaît à une position aléatoire sur l'écran. L'application prend fin une fois les huit cibles atteintes.

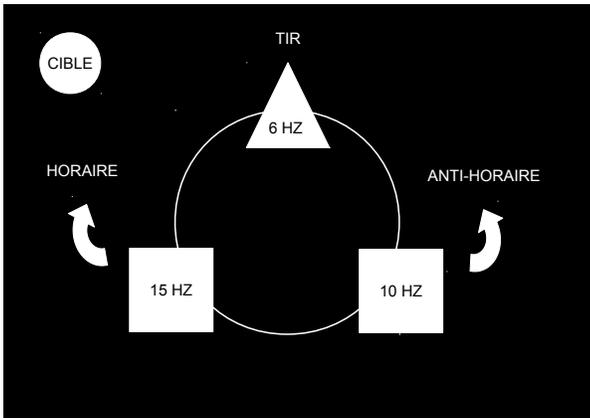


FIGURE 4: Phase en ligne.

III. TRAITEMENT DU SIGNAL ET CLASSIFICATION

Cette étape porte sur l'ensemble de la chaîne de traitement du signal et de classification. Le traitement du signal permet d'éliminer au maximum les artefacts associés au signal et de conserver uniquement l'information pertinente. L'étape de classification utilise cette information pour identifier les motifs cérébraux spécifiques du mode SSVEP et contrôler une application en ligne.

La littérature a montré l'utilité des filtres CSP [21] et des classifieurs LDA [14] dans un problème de classification binaire dans le cadre des SSVEP. C'est pourquoi nous les utiliserons dans notre étude. Leur calibrage est spécifique à chaque utilisateur.

La première étape consiste à filtrer le signal brut autour de chaque fréquence de stimulation (6, 10 et 15 Hz) par l'intermédiaire d'un filtre passe-bande de type Butterworth (ordre 4). La bande passante autour de chacune des fréquences est $\delta = \pm 0.25$ Hz. Ces données filtrées permettent d'optimiser les filtres CSP (Common Spatial Pattern), en extrayant une fenêtre de temps de 7 secondes, 1 seconde après le début de chaque essai (ou stimulation). On obtient 8 fenêtres de temps (ou paquets de données) pour chaque fréquence de stimulation, y compris pour la fréquence nulle. Les données extraites sont ensuite séparées en deux classes. Une classe correspond à la fréquence d'intérêt (par exemple 10 Hz) et l'autre classe aux autres fréquences. Les filtres spatiaux sont calibrés spécifiquement pour chaque fréquence de stimulation. De plus, on conserve les deux meilleurs filtres spatiaux,

correspondant aux filtres ayant les valeurs propres les plus élevées.

Une fois les filtres spatiaux ajustés, ils sont utilisés pour calculer les valeurs caractéristiques pour le classifieur LDA (Linear Discriminant Analysis). Un classifieur LDA est calibré pour chaque fréquence de stimulation. Le calibrage du classifieur nécessite d'extraire, à partir des signaux filtrés autour de chaque fréquence de stimulation puis spatialement par CSP, des fenêtres de temps de 7 secondes, 1 seconde après le début de chaque essai (ou stimulation). On obtient 8 paquets de données pour chaque fréquence de stimulation comme dans le cas des filtres CSP. Puisqu'on utilise deux filtres CSP, un paquet de données est une matrice à 2 colonnes et N lignes, avec N le nombre d'échantillon contenu dans une fenêtre de 7 secondes tel que $N = F_e \times 7$. Ces données sont à nouveau séparées en deux classes en fonction de la fréquence à laquelle ils appartiennent et pour laquelle on calibre le classifieur. Par exemple, si on calibre le classifieur pour la fréquence à 6 Hz, les paquets de données enregistrés lors de la stimulation à 6 Hz appartiennent à la classe 1 et ceux enregistrés pendant les autres stimulations appartiennent à la classe 2. Puis pour chaque paquet de chaque classe, on extrait des fenêtres d'échantillon d'une demie-seconde à chaque 1/10ème de seconde. On obtient 66 fenêtres d'échantillon pour chaque paquet. En effet, par paquet, on extrait dix fenêtres pour chaque seconde — les fenêtres sont extraites tout les 1/10ème de seconde — sauf pour la dernière seconde (seconde 7) où on peut extraire que 6 fenêtres d'une demie-seconde. Une fenêtre est une matrice à 2 colonnes, les signaux ont été filtrés par deux filtres CSP, et l lignes, avec $l = F_e \times 0.5$ c'est-à-dire le nombre d'échantillon contenu dans une fenêtre de temps d'une demie-seconde. Les valeurs contenues dans chaque fenêtre sont élevées au carré puis moyennées sur chaque colonne, de telle sorte que chaque paquet P ($P \in \mathbb{R}^{66 \times 2}$) est constitué de 66 couples de valeurs caractéristiques (X, Y). Une dernière opération consiste à transformer les valeurs X et Y en $\log(X+1)$ et $\log(Y+1)$. Ces valeurs caractéristiques sont ensuite utilisées pour calculer les coefficients de chaque classifieur.

Le logiciel OpenVibe utilise une méthode de validation croisée pour calculer les classifieurs. Les valeurs caractéristiques sont séparées en dix groupes indifféremment de leur classe. Le classifieur est calibré sur 9 groupes puis testé sur un. Cette méthode est réalisée 10 fois de telle sorte que le classifieur soit testé sur chaque groupe. Le classifieur pour lequel on obtient le meilleur résultat est conservé pour la phase en ligne. Pendant la phase en ligne, les filtres CSP et les classifieurs LDA analysent en continu les données enregistrées au cours de la dernière seconde, avec un décalage de 0.1 seconde. La sortie du classifieur LDA est basée sur la distance des valeurs caractéristiques par rapport à un point, une droite, un plan ou un hyperplan selon la dimension de la matrice de valeurs caractéristiques. Etant donné qu'on obtient un classifieur pour chaque fréquence de stimulation, la prise de décision en ligne se fait par vote automatique, en tenant compte des sorties de chacun d'entre eux.

Pour calibrer hors-ligne le classifieur, on procède à une méthode différente que dans OpenVibe. On élimine pour chaque fréquence les paquets P aberrants, c'est-à-dire les paquets pour lesquels le point moyen est du mauvais côté de la droite définie par les coefficients du classifieur. Ces coefficients sont calculés à partir des 31 paquets restants, sept appartenant à la fréquence pour laquelle on calcule le classifieur et 24 appartenant aux autres fréquences. Les paquets qui ne sont pas éliminés sont conservés pour calibrer le classifieur. Afin de tester ses performances en intra-session, on divise à nouveau les données qui ont permis son calibrage en deux groupes. Un groupe permet de calculer un classifieur et l'autre de tester ses performances. Pour tester ses performances en inter-session, on applique directement le classifieur à l'ensemble des données de l'autre session.

IV. RÉSULTATS

Cette partie présente les performances moyennes en ligne et hors-ligne obtenues sur l'ensemble des participants pendant les trois sessions. Le temps (seconde) et l'aire sous la courbe ROC (Receiver Operating Characteristic) ou AUC (Area Under the Curve) correspondent respectivement aux indicateurs de performance en ligne et hors-ligne. L'indicateur de performance hors-ligne est défini par une valeur v , avec $0 \leq v \leq 1$, de telle sorte que la performance du classifieur soit proportionnelle à la valeur de v . Un classifieur avec une valeur $v = 1$ est parfaitement exact, il ne se trompe jamais. Si $v = 0.5$ le classifieur donne des résultats aléatoires, le taux de bonne classification est égal à celui d'un tirage au sort. Ces résultats permettront de comparer les performances obtenues sous différentes conditions et de répondre aux problématiques posées dans cet article. Etant donné que la taille de l'échantillon est inférieure à 30, les performances sont comparées à partir d'un test non-paramétrique de Wilcoxon, avec un risque $\alpha = 5\%$.

A. Performances de classification hors-ligne

La figure 5a compare sous forme d'histogramme les performances du classifieur 1 (classifieur calibré avec les données de la session 1) sur les données de la session 1 (C_1) avec ses performances sur les données de la session 2 (C'_1) et 3 (C''_1).

La figure 5b compare sous forme d'histogramme les performances du classifieur 1 sur les données de la session 2 (C'_1) et 3 (C''_1) avec les performances respectivement du classifieur 2 et 3 (classifieur calibré avec les données de la session 2 et 3) sur les données de la session 2 (C_2) et 3 (C_3).

La figure 5c compare sous forme d'histogramme les performances du classifieur 1 sur les données de la session 1 (C_1) avec les performances respectivement du classifieur 2 et 3 sur les données de la session 2 (C_2) et 3 (C_3).

L'ensemble des performances présentées correspondent à la moyenne des performances de chaque individu.

Pour étudier la stabilité du classifieur dans le temps on réalise deux tests statistiques. Le premier, comme le montre la figure 5a, compare les performances entre la session 1 et la session 2 puis entre la session 1 et la session 3. Cette

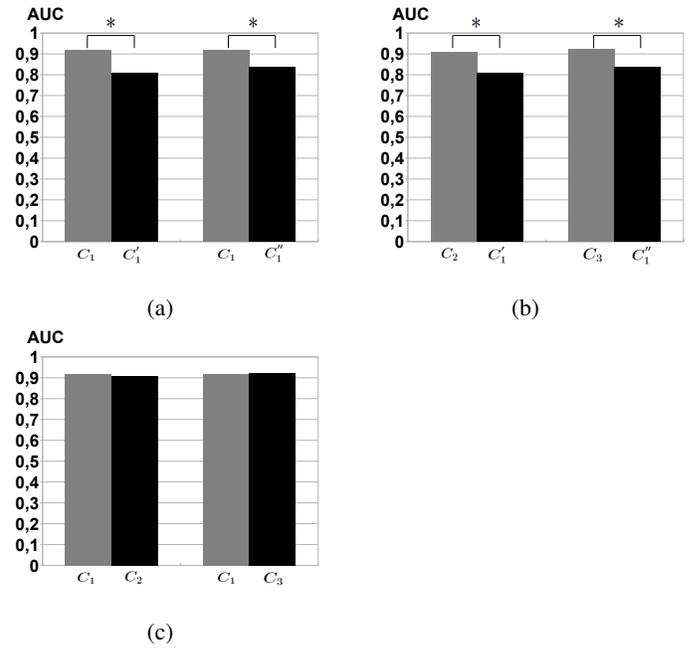


FIGURE 5: Comparaison sous forme d'histogramme des performances hors-ligne de différents classifieurs sur différents jeux de données. Les valeurs (AUC) correspondent à la moyenne des performances sur l'ensemble des sujets. C_1 correspond aux performances du classifieur 1 (classifieur calibré sur les données de la session 1) sur les données de la session 1. C'_1 et C''_1 correspondent aux performances du classifieur 1 sur les données respectivement de la session 2 et 3. C_2 et C_3 correspondent aux performances respectivement du classifieur 2 et 3 (classifieurs calibrés sur les données respectivement de la session 2 et 3) sur les données de la session 2 et 3. Le symbol “*” signifie que la différence entre les deux valeurs comparées est significative avec un risque $\alpha = 5\%$.

comparaison prend en compte les performances d'un même classifieur (classifieur 1) testé sur des données de calibrage enregistrées à plusieurs jours d'intervalle. On évalue ainsi la fiabilité du classifieur à différents instants et sous différentes conditions expérimentales. En effet, différents paramètres varient d'une session à une autre : position des électrodes, motivation, fatigue du sujet, etc. Un test de Wilcoxon indique une différence significative entre les performances du classifieur sur les données de la session 1 et de la session 2 ($p\text{-value} = 3,051.10^{-5}$). Une différence significative est également observée entre les performances du classifieur sur les données de la session 1 et de la session 3 ($p\text{-value} = 9,15.10^{-5}$). Ce résultat est appuyé par un second test statistique, voir figure 5b, qui compare les performances du classifieur 1 avec les performances respectivement du classifieur 2 et 3 sur les données de la session 2 et 3. L'idée est de comparer les performances d'un ancien classifieur avec celles d'un classifieur nouvellement calculé. Dans ce cas les classifieurs sont testés sur les mêmes données de calibrage et donc sous les mêmes conditions expérimentales. Un test de Wilcoxon indique une

différence significative entre les performances du classifieur 1 et 2 sur les données de la session 2 ($p\text{-value} = 3,051.10^{-5}$) et entre les performances du classifieur 1 et 3 sur les données de la session 3 ($p\text{-value} = 3,051.10^{-5}$).

Enfin, pour étudier les effets de l'apprentissage humain sur les performances de l'interface, on compare les performances entre la session 1 et la session 2 puis entre la session 1 et la session 3, voir figure 5c. Il s'agit des performances de différents classifieurs testés sur différentes données de calibrage. En réalisant un test de Wilcoxon on n'observe pas de différence significative entre les performances du classifieur 1 et 2 sur les données respectivement de la session 1 et 2 ($p\text{-value} = 0,175$) puis entre les performances du classifieur 1 et 3 sur les données respectivement de la session 1 et 3 ($p\text{-value} = 0,632$).

B. Performances en ligne

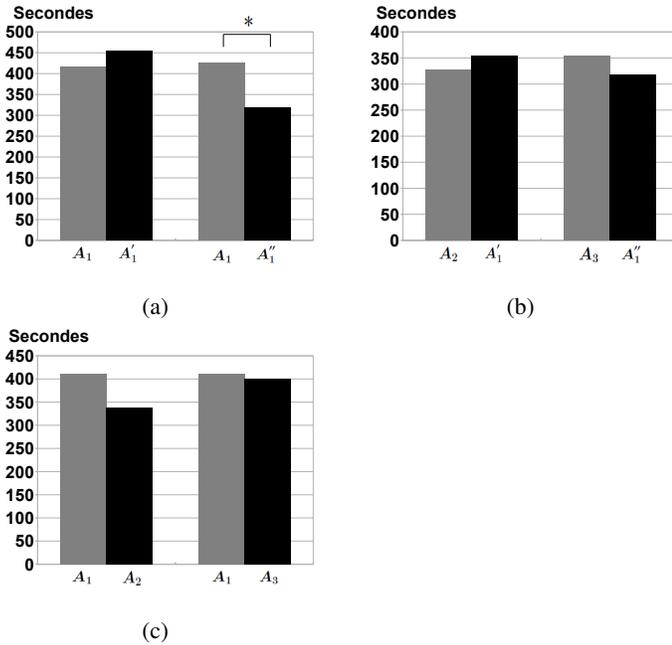


FIGURE 6: Comparaison sous forme d'histogramme des performances en ligne de différents classifieurs à différentes sessions. Les valeurs, en secondes, correspondent à la moyenne des performances de l'ensemble des sujets. A_1 , A'_1 et A''_1 correspondent aux temps pour finir l'application respectivement lors de la session 1, 2 et 3 en utilisant le classifieur 1. A_2 et A_3 correspondent aux temps pour finir l'application respectivement lors de la session 2 et 3 en utilisant respectivement le classifieur 2 et 3. Le symbole "*" signifie que la différence entre les deux valeurs comparées est significative avec un risque $\alpha = 5\%$.

La figure 6a compare sous forme d'histogramme les performances en ligne du classifieur 1 lors de la session 1 (A_1) avec ses performances lors de la session 2 (A'_1) et 3 (A''_1).

La figure 6b compare sous forme d'histogramme les performances en ligne du classifieur 1 lors de la session 2 (A'_1) et

3 (A''_1) avec les performances respectivement du classifieur 2 et 3 lors de la session 2 (A_2) et 3 (A_3).

La figure 6c compare sous forme d'histogramme les performances en ligne du classifieur 1 lors de la session 1 (A_1) avec les performances du classifieur 2 (A_2) et 3 (A_3) respectivement lors de la session 2 et 3.

L'ensemble des performances présentées correspondent à la moyenne des performances de chaque individu.

Les résultats présentés figure 6 ne prennent pas en compte les sujets n'ayant pas de contrôle sur l'application. En effet, deux sujets (sujets 3 et 5) n'avaient aucun contrôle sur l'application quel que soit le classifieur utilisé et la session. Deux autres sujets (sujets 12 et 13) n'avaient aucun contrôle sur l'application lors des sessions 2 et 3 en utilisant le classifieur 1. Enfin, un dernier sujet (sujet 14) n'avait pas de contrôle sur l'application lors de la session 3 en utilisant le classifieur 1.

Pour étudier la stabilité du classifieur dans le temps on réalise les mêmes tests statistiques que ceux réalisés dans la section IV-A. Dans un premier temps, on compare les performances entre la session 1 et la session 2 puis entre la session 1 et la session 3, voir figure 6a. On évalue la fiabilité du classifieur (classifieur 1) à différents instants et sous différentes conditions expérimentales. Un test de Wilcoxon n'indique pas de différence significative entre les performances obtenues à la session 1 et celles obtenues à la session 2 ($p\text{-value} = 0,339$). Néanmoins on observe une différence significative entre les performances obtenues à la session 1 et celles obtenues à la session 3 ($p\text{-value} = 0,0136$). Ce résultat est appuyé par un second test statistique qui compare les résultats obtenus lors des deux phases en ligne des sessions 2 et 3, voir figure 6b. On compare ainsi les performances en ligne du premier classifieur avec celles des classifieurs nouvellement calculés. Dans ce cas les classifieurs sont testés sous les mêmes conditions expérimentales. Un test de Wilcoxon n'indique pas de différence significative entre les performances du classifieur 1 et 2 lors de la session 2 ($p\text{-value} = 0,51$) et entre les performances du classifieur 1 et 3 lors de la session 3 ($p\text{-value} = 0,365$).

Enfin, pour étudier les effets de l'apprentissage humain sur les performances de l'interface, on compare les performances entre la session 1 et la session 2 puis entre la session 1 et la session 3, voir figure 6c. Il s'agit des performances de différents classifieurs testés dans différentes conditions expérimentales. En réalisant un test de Wilcoxon on n'observe pas de différence significative entre les performances du classifieur 1 et 2 respectivement lors de la session 1 et 2 ($p\text{-value} = 0,104$), ni même entre les performances du classifieur 1 et 3 respectivement lors de la session 1 et 3 ($p\text{-value} = 1$).

V. DISCUSSION

Dans un premier temps on s'intéresse aux effets de l'apprentissage humain sur les performances de l'interface. Les tests statistiques ne montrent pas de différence significative entre les performances obtenues hors-ligne par différents classifieurs calculés à différentes sessions (voir figure 5c). On constate

des résultats similaires en ce qui concerne les performances réalisées en ligne (voir figure 6c). Ces résultats semblent donc indiquer que l'apprentissage humain n'a aucun effet sur les performances du classifieur. En effet, les résultats obtenus hors-ligne ou en ligne ne s'améliorent pas de manière significative d'une session à une autre, tandis que le participant acquiert davantage d'expérience avec le temps. De plus, on n'observe pas une amélioration des résultats obtenus hors-ligne ou en ligne par les sujets ayant déjà eu une expérience avec les SSVEP (sujets 9 et 12) par rapport aux résultats des autres sujets.

On constate que l'ensemble des sujets, à l'exception des sujets 3 et 5, ont terminé l'application dans un intervalle de temps satisfaisant (≤ 13 minutes) lorsqu'ils ont utilisé un classifieur nouvellement calibré. Cela tend à prouver que les BCI de type SSVEP fonctionnent correctement dès la première utilisation. Néanmoins, comme pour les autres paradigmes (IM, SCP), on observe des cas pour lesquels l'interface ne fonctionne pas même après plusieurs sessions, comme pour les sujets 3 et 5. On trouve fréquemment dans la littérature le terme "illiteracy" pour caractériser ce type de phénomène [22]. Néanmoins une personne dite "illettrée" pour un paradigme ne l'est pas nécessairement pour un autre. C'est pourquoi il est intéressant de proposer aux participants une interface dite "hybride", associant plusieurs types de BCI [23] ou une BCI avec un autre canal de contrôle [24].

Intéressons nous dans un second temps à l'analyse de la stabilité du classifieur. Les tests statistiques réalisés sur les résultats obtenus hors-ligne (voir section IV-A) semblent indiquer une certaine instabilité du classifieur dans le temps. En effet, on observe une baisse significative de performance pour un même classifieur évalué sur les données de différentes sessions (voir figure 5a). Ce constat est identique lorsqu'on compare sur les données d'une même session les performances d'un classifieur plus ancien avec celles d'un classifieur nouvellement appris (voir figure 5b). Cependant, ces résultats obtenus hors-ligne ne sont pas confirmés par les résultats réalisés en ligne (voir figures 6a et 6b). En effet, les tests statistiques ne semblent pas montrer de différences significatives. Une amélioration des résultats est même observée (voir figure 6a), et les performances en ligne du classifieur 1 semblent s'être améliorées de la session 1 à la session 3. Les résultats obtenus en ligne et hors-ligne sont donc contradictoires. Les résultats hors-ligne semblent se dégrader lorsque le classifieur est réutilisé sur d'autres données de calibrage. Or cette dégradation des résultats hors-ligne ne semble pas affecter les résultats en ligne puisque ces derniers restent constant et s'améliorent même dans un cas. Cela peut s'expliquer par le fait que les performances en ligne ne dépendent pas uniquement de la fiabilité du classifieur mais aussi en grande partie de la stratégie de jeu adoptée par l'utilisateur. En effet, on observe un phénomène d'inertie lors de la phase de jeu : le "vaisseau spatial" semble continuer sa rotation un certain temps après l'arrêt de la stimulation visuelle. Cela peut s'expliquer par le fait que les neurones du lobe occipital restent activés un certain temps après l'arrêt

de la stimulation. Ce phénomène d'inertie est contraignant pour contrôler correctement l'application. Les performances en ligne dépendent donc pour beaucoup de la capacité à anticiper ce phénomène.

Ces données semblent donc montrer qu'un classifieur peut être réutilisé plusieurs fois après la phase de calibrage, malgré une dégradation de ces performances hors-ligne. Néanmoins, il est important de rappeler que deux sujets (sujets 12, 13) n'ont pas réussi à réutiliser en ligne le classifieur et un dernier (sujet 14) n'a pas réussi à le réutiliser lors de la troisième session. De plus, l'expérience a été menée sur dix jours, on ne peut donc pas conclure sur la fiabilité du classifieur au-delà de cette période. Enfin, l'étude ne tient pas compte de la parité homme/femme.

VI. CONCLUSION

Les BCI de type SSVEP présentent comme avantage de fonctionner théoriquement de manière optimale dès la première utilisation [14], [15]. Aucune phase d'apprentissage n'est nécessaire de la part de l'utilisateur contrairement aux potentiels corticaux lents [17] ou à l'imagination motrice [16]. En effet ces deux types de BCI nécessitent un processus cognitif important pour apprendre à contrôler les motifs cérébraux propres à chaque mode, tandis que le mode SSVEP fait intervenir uniquement un processus perceptif. La personne doit focaliser son attention visuelle sur un stimulus particulier qui excite la rétine et provoque un SSVEP au niveau du lobe occipital. On comprend donc par la nature même de son mécanisme que le paradigme SSVEP fonctionne sans apprentissage. Cependant on peut se demander si, avec l'habitude de réaliser la même expérience, l'utilisateur arrive de manière consciente ou non à davantage se focaliser sur les stimuli et donc à obtenir de meilleurs résultats. Les données dont nous disposons ne semblent pas indiquer d'amélioration significative des performances avec l'augmentation de l'expérience du sujet. Ces résultats tendent donc à confirmer ceux de la littérature.

La concentration visuelle importante exigée par le mode SSVEP, notamment durant la phase de calibrage de manière à maximiser les performances du classifieur, induit chez le sujet une fatigue précoce et peut entraîner une certaine lassitude. C'est pourquoi l'article a abordé la stabilité du classifieur dans le temps. Nos résultats semblent indiquer une dégradation des performances hors-ligne pour un même classifieur d'une session à une autre. Néanmoins, cela n'affecte nullement les performances en ligne tant que cette dégradation des performances n'excède pas un certain seuil (80%), qui permet de conserver un contrôle correct sur l'application. Cela tend à démontrer que le classifieur peut être réutilisé sans de nouveau réaliser une phase de calibrage. On limite ainsi les contraintes liées à cette étape.

Les résultats semblent montrer que le paradigme SSVEP permet d'obtenir, pour certains utilisateurs, un contrôle rapide et satisfaisant de l'interface chez des sujets novices. Néanmoins même si on peut diminuer les contraintes en limitant les phases de calibrage (voir section IV), ce paradigme

reste à long terme fatigant pour l'utilisateur. En effet, nous avons observé chez certaines personnes des phénomènes de clignement et de picotement des yeux dus à la répétition des stimuli. Si ce paradigme devient contraignant chez des personnes saines après quelques minutes (≈ 30 min) d'utilisation, on peut en déduire qu'il n'est pas approprié à un contexte d'utilisation prolongée tel que seront amenés à l'utiliser les personnes handicapées. Le mode SSVEP peut être une stratégie intéressante, dans un premier temps, pour permettre au patient de regagner rapidement en autonomie. Mais, si l'application le permet, un passage progressif vers un paradigme moins contraignant et fatigant, comme par exemple l'imagination motrice, serait plus judicieux. Le mode SSVEP peut également être intégré à un système hybride dans lequel il ne représente pas la composante principale de l'interface. Cela limiterait les contraintes liées à ce paradigme tout en conservant ses avantages.

RÉFÉRENCES

- [1] Nataliya Kosmyna and Franck Tarpin-Bernard. Une combinaison de paradigmes d'interaction cerveau-ordinateur et suivi du regard pour des interactions multimodales. In *Actes de la conférence Ergonomie, Interaction Homme-Machine (Ergo'IHM 2012)*, page online, 2012.
- [2] Matthew R Williams and Robert F Kirsch. Evaluation of head orientation and neck muscle EMG signals as command inputs to a human-computer interface for individuals with high tetraplegia. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 16(5) :485–496, 2008.
- [3] Hyuk Jeong and JinSung Choi. An EMG-controlled graphic interface considering wearability. In *Proceedings of Human Computer Interaction, INTERACT'03*, pages 958–961, 2003.
- [4] Chun Sing Louis Tsui, Pei Jia, John Q Gan, Huosheng Hu, and Kui Yuan. EMG-based hands-free wheelchair control with EOG attention shift detection. In *Proceedings of the IEEE International Conference on Robotics and Biomimetics, ROBOT'07*, pages 1266–1271, 2007.
- [5] Hsieh-Ching Chen, Chia-Ling Chen, Chang-Ching Lu, and Ching-Yi Wu. Pointing device usage guidelines for people with quadriplegia : a simulation and validation study utilizing an integrated pointing device apparatus. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 17(3) :279–286, 2009.
- [6] Mark S Hawley, Pam Enderby, Phil Green, Stuart Cunningham, Simon Brownell, James Carmichael, Mark Parker, Athanassios Hatzis, Peter O'Neill, and Rebecca Palmer. A speech-controlled environmental control system for people with severe dysarthria. *Medical Engineering & Physics*, 29(5) :586–593, 2007.
- [7] Febo Cincotti, D Mattia, F Aloise, S Bufalari, MG Marciari, G Schalk, G Oriolo, A Cherubini, F Davide, and F Babiloni. Non-invasive brain-computer interface system to operate assistive devices. In *Proceedings of the 29th IEEE Annual International Conference on Engineering in Medicine and Biology Society, EMBS'07*, pages 2532–2535, 2007.
- [8] Guenter Edlinger, Christoph Kapeller, Arnau Espinosa, Sergi Torrellas, Felip Miralles, and Christoph Guger. Multi-modal computer interaction for communication and control using EEG, EMG, EOG and motion sensors. In *Universal Access in Human-Computer Interaction. Design Methods, Tools, and Interaction Techniques for eInclusion*, pages 633–641. 2013.
- [9] Johan Philips, J del R Millan, Gerolf Vanacker, Eileen Lew, Ferran Galán, Pierre W Ferrez, Hendrik Van Brussel, and Marnix Nuttin. Adaptive shared control of a brain-actuated simulated wheelchair. In *Proceedings of the 10th IEEE International Conference on Rehabilitation Robotics, ICORR'07*, pages 408–414, 2007.
- [10] K Jerbi, O Bertrand, B Schoendorff, D Hoffmann, L Minotti, P Kahane, A Berthoz, and J-P Lachaux. Online detection of gamma oscillations in ongoing intracerebral recordings : From functional mapping to brain computer interfaces. In *Proceedings of the Joint Meeting of the 6th International Symposium on Noninvasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging, NFSI-ICFBI'07*, pages 330–333, 2007.
- [11] Philip R Kennedy, M Todd Kirby, Melody M Moore, Brandon King, and Adon Mallory. Computer control using human intracortical local field potentials. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 12(3) :339–344, 2004.
- [12] Fabien Lotte. Les interfaces cerveau-ordinateur : Conception et utilisation en réalité virtuelle. *Revue Technique et Science Informatiques*, 31(3) :289–310, 2012.
- [13] Jonathan R Wolpaw, Niels Birbaumer, Dennis J McFarland, Gert Pfurtscheller, and Theresa M Vaughan. Brain-computer interfaces for communication and control. *Clinical neurophysiology*, 113(6) :767–791, 2002.
- [14] Edmund C Lalor, Simon P Kelly, Ciarán Finucane, Robert Burke, Ray Smith, Richard B Reilly, and Gary Mcdarby. Steady-state VEP-based brain-computer interface control in an immersive 3D gaming environment. *EURASIP journal on applied signal processing*, 2005 :3156–3164, 2005.
- [15] Luca Piccini, Sergio Parini, Luca Maggi, and Giuseppe Andreoni. A wearable home BCI system : preliminary results with SSVEP protocol. In *Proceedings of the 27th IEEE Annual International Conference on Engineering in Medicine and Biology Society, EMBS'05*, pages 5384–5387, 2006.
- [16] Gert Pfurtscheller, C Guger, G Müller, G Krausz, and C Neuper. Brain oscillations control hand orthosis in a tetraplegic. *Neuroscience letters*, 292(3) :211–214, 2000.
- [17] Niels Birbaumer, Andrea Kubler, Nimr Ghanayim, Thilo Hinterberger, Jouri Perelmouter, Jochen Kaiser, Iver Iversen, Boris Kotchoubey, Nicola Neumann, and Herta Flor. The thought translation device (TTD) for completely paralyzed patients. *Rehabilitation Engineering*, pages 190–193, 2000.
- [18] Robert Leeb, Doron Friedman, Gernot R Müller-Putz, Reinhold Scherer, Mel Slater, and Gert Pfurtscheller. Self-paced (asynchronous) BCI control of a wheelchair in virtual environments : a case study with a tetraplegic. *Computational intelligence and neuroscience*, 2007, 2007.
- [19] Eric W Sellers and Emanuel Donchin. A P300-based brain-computer interface : initial tests by ALS patients. *Clinical neurophysiology*, 117(3) :538–548, 2006.
- [20] American Clinical Neurophysiology Society et al. Guideline 5 : Guidelines for standard electrode position nomenclature. *American journal of electroneurodiagnostic technology*, 46(3) :222, 2006.
- [21] Sergio Parini, Luca Maggi, Anna C Turconi, and Giuseppe Andreoni. A robust and self-paced BCI system based on a four class SSVEP paradigm : algorithms and protocols for a high-transfer-rate direct brain communication. *Computational Intelligence and Neuroscience*, 2009, 2009.
- [22] Clemens Brunner, Brendan Z Allison, Dean J Krusienski, Vera Kaiser, Gernot R Müller-Putz, Gert Pfurtscheller, and Christa Neuper. Improved signal processing approaches in an offline simulation of a hybrid brain-computer interface. *Journal of neuroscience methods*, 188(1) :165–173, 2010.
- [23] Brendan Z Allison, Clemens Brunner, Vera Kaiser, Gernot R Müller-Putz, Christa Neuper, and Gert Pfurtscheller. Toward a hybrid brain-computer interface based on imagined movement and visual attention. *Journal of neural engineering*, 7(2) :026007, 2010.
- [24] Robert Leeb, Hesam Sagha, Ricardo Chavarriaga, and J del R Millan. Multimodal fusion of muscle and brain signals for a hybrid-BCI. In *Proceedings of the IEEE Annual International Conference on Engineering in Medicine and Biology Society, EMBC'10*, pages 4343–4346, 2010.