HAL open science

Optimizing the use of SSVEP-based brain-computer interfaces for human-computer interaction

Andéol Évain

► To cite this version:

Andéol Évain. Optimizing the use of SSVEP-based brain-computer interfaces for human-computer interaction. Human-Computer Interaction [cs.HC]. Université de Rennes, 2016. English. NNT : 2016REN1S083 . tel-01476185

HAL Id: tel-01476185
https://theses.hal.science/tel-01476185

Submitted on 24 Feb 2017

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.

ANNÉE 2016

THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Bretagne Loire

pour le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1

Mention : Informatique

Ecole doctorale Matisse

présentée par

Andéol Évain

préparée à l'unité de recherche (IRISA - UMR6074)
(Institut de recherche en informatique et systèmes aléatoires - ISTIC)
(UFR Informatique et Électronique)

Optimizing the Use of SSVEP-based Brain-Computer Interfaces for Human-Computer Interaction

Thèse soutenue à Rennes le 6 décembre 2016

devant le jury composé de :

Gernot Müller-Putz
Professeur, Technische Universität Graz / Rapporteur

Michel Beaudouin-Lafon
Professeur, Université Paris-Sud / Rapporteur

Maureen Clerc
Directrice de Recherche, Inria / Examinatrice

Thorsten Zander
Professeur, Technische Universität Berlin / Examinateur

Anatole Lécuyer
Directeur de Recherche, Inria / Directeur de thèse

Nicolas Roussel
Directeur de Recherche, Inria / Examinateur

Géry Casiez
Professeur, Université de Lille / Examinateur

Ferran Argelaguet
Chargé de Recherche, Inria / Examinateur

Principe de Laplace: Le poids des preuves doit être proportionné à l'étrangeté des faits.

Remerciements

Je voudrais remercier mes encadrants Anatole Lécuyer, Ferran Argelaguet, Nicolas Roussel, et Géry Casiez, pour leurs apports et leurs conseils durant cette thèse. Merci d'avoir supporté ma tête de mûle lors de nombreuses discussions. Un merci tout particulier à Ferran, pour son esprit scientifique pointu, et sa diplomatie à toute épreuve.

Merci, également, aux membres de mon Jury: Gernot Müller-Putz, Michel Beaudouin-Lafon, Maureen Clerc, et Thorsten Zander, pour avoir accepté de juger mon travail et mon manuscrit. Merci pour vos commentaires et vos questions toujours très pertinents.

Merci aux stagiaires qui ont accepté de me suivre dans mes recherches: Anthony Strock, Loïs Paulin, et Charles Mille. Je vous souhaite bon courage et bonne chance dans la suite de vos parcours.

Merci aux relecteurs attentifs de ma thèse, dernier rempart entre la grammaire et moi, et que je n'oserai pas lister, parce que je suis trop sûr d'en oublier. Je dois toujours un verre à certains d'entre vous.

Je voudrais également remercier toute l'équipe Hybrid, et sa soeur, l'équipe Mimetic, pour l'ambiance (toujours studieuse) que leurs membres ont su mettre dans le couloir pendant ces trois années, au court de nombreuses pauses café. En particulier, merci à Thomas Boggini, Guillaume-claude Claude, François Lehericey, Tristan Le Bouffant, et Carl-Johan Jorgensen, les meneurs d'ambiance du couloir. Merci également à Jonathan Mercier-Ganady qui a partagé mon bureau pendant deux ans, et un merci tout particulier à Jussi Lindgren, qui a tenu bon l'intégralité de ma thèse, pour son humour finlandais et son expertise scientifique aussi acérés l'un que l'autre.

Merci, enfin, à Nathalie Denis, pour sa gestion sans faille des inscriptions, missions, trajets, conférences, etc. Merci pour son infinie patience vis-à-vis de ceux qui, comme moi, sont plus à l'aise face à des équations ou du C++ que devant un simple formulaire.

Contents

Optimisation de l'utilisation des interfaces cerveau-machines basées sur SSVEP pour l'interaction homme-machine

Les Interfaces Cerveau-Machine (ICO) sont les systèmes mesurant l'activité cérébrale d'un utilisateur, et utilisant ces mesures pour l'interaction homme-machine. L'Interaction Homme-Machine (IHM) est la discipline qui s'intéresse *au design, à l'évaluation, et à l'implémentation de systèmes informatiques interactifs pour l'humain, et à l'étude des principaux phénomènes qui l'entourent* [Hewett et al., 1992]. Cette thèse porte sur l'amélioration des interfaces cerveau-machines basées sur SSVEP, du point de vue de l'interaction homme-machine.

Les interfaces cerveau-machines sont un outil d'interaction parmi d'autres. Pendant longtemps, elles n'ont servi qu'en tant que dispositif d'interaction alternatif pour personnes souffrant de handicaps physiques sévères [Megalingam et al., 2013, Donchin et al., 2000, Blankertz et al., 2006b]. Les ICOs ont peu à peu suscité l'intérêt de domaines d'application plus larges. Avec les avancées de la neurologie et du traitement du signal, ainsi que l'invention de nouveaux paradigmes ICO, elles sont devenues un outil prometteur pour l'interaction homme-machine, dans des domaines variés tels que la réalité virtuelle [Lécuyer et al., 2008], les jeux vidéo [Nijholt et al., 2009, Nijholt and Gürkök, 2013], ou plus généralement comme moyen d'acquérir des information sur l'état mental de l'utilisateur (approche passive) [Lécuyer et al., 2013].

Avec l'extension des domaines d'application des ICO, les méthodes et les concepts de l'IHM deviennent de plus en plus adaptés pour l'amélioration des ICOs. Plutôt que de se concentrer uniquement sur la vitesse et la précision, d'autres critères tels que la simplicité d'installation, le confort de l'utilisateur, ou sa fatigue, doivent être pris en compte. Lors de la conception d'une ICO pour un type d'application défini, deux choix doivent être fait dès le début du développement, et ont un impact important sur les caractéristiques finales de l'interface. Quel type de mesure cérébrale devrait être utilisé, et quel type de motif cérébral devrait être reconnu ?

Electroencéphalographie (EEG)

L'activité cérébrale peut être mesurée par des techniques variant en coût, simplicité d'utilisation, et en résolution spatiale et temporelle. Dans cette thèse, nous nous concentrons sur les ICOs

Figure 1: Utilisateur d'une ICO basée sur SSVEP, portant un casque EEG, et regardant un stimulus clignotant (ici des carrés rouges).

basées sur l'électroencéphalographie (EEG). L'EEG est une technique de mesure externe, c'est-à-dire qu'elle ne demande pas d'opération chirurgicale. Le matériel requis est comparativement peu couteux, et présente une bonne résolution temporelle (typiquement 512 Hz). Ces caractéristiques font de l'EEG un choix naturel pour la plupart des applications d'IHM utilisant des ICOs.

SSVEP

En ce qui concerne le choix du motif cérébral, quand la vitesse d'interaction est un facteur important, les ICOs basées sur SSVEP sont compétitives, présentant une vitesse et une précision parmi les plus élevés des ICOs. Cette thèse se concentre sur les ICOs basées sur SSVEP.

Lorsque que le cerveau reçoit une stimulation visuelle périodique, il répond par une activité à la fréquence de stimulation dans les aires visuelles primaires [Herrmann, 2001]. Cette réponse périodique est appelée *Potentiel évoqué visuellement en régime permanent* (ou plus communément en anglais *Steady-State Visually Evoked Potential* (SSVEP)). Cette réponse peut être détectée dans le signal EEG, sous la forme d'un pic de puissance à la fréquence de stimulation.

Les ICOs basées sur SSVEP exploitent ce phénomène physiologique. Dans une utilisation standard, plusieurs cibles sont présentées à l'utilisateur, et clignotent chacune à une fréquence différente. Quand l'utilisateur concentre son attention visuelle sur l'une ou l'autre de ces cibles (voir Figure 1), le SSVEP en résultant peut être détecté, et la cible est sélectionnée (ou une commande associée est activée) [Lalor et al., 2005, Müller-Putz et al., 2006, Zhu et al., 2010a].

Les ICOs basées sur SSVEP ont fait leurs preuves en termes de vitesse et de précision, comparées aux autres ICOs. A ce jour (en 2016), elles détiennent le record du plus haut débit d'information pour des ICOs basées sur l'EEG [Nakanishi et al., 2014]. La vitesse et la précision des ICOs étant souvent leurs principales limitations, du point de vue de l'IHM, SSVEP apparait comme l'un des paradigmes ICO les plus prometteurs.

Problèmes et objectifs

L'interaction homme-machine (IHM) présente trois composants principaux: Un utilisateur, une interface, et une application sur un appareil électronique. Le rôle de l'interface est de gérer la communication entre l'utilisateur humain et l'application. L'IHM est le domaine de recherche qui vise à comprendre chacun des aspects de l'interaction, et à l'améliorer.

L'essentiel de la recherche sur les ICOs jusqu'ici porte sur le traitement du signal et la classification. Après des années de recherche, la vitesse et la précision des ICOs ont connu d'importantes améliorations. Cependant, afin d'utiliser les ICOs pour des applications réelles, loin des conditions contrôlées de laboratoire, le phénomène d'interaction basé sur des ICO doit être considéré. Notre approche est d'appliquer les méthodes et les concepts de l'IHM aux ICOs. Plutôt que d'améliorer uniquement le coeur des ICOs, nous cherchons à étudier le système interactif entier:

- L'expérience utilisateur ICO: La perception et les réponses qui résultent de l'utilisation d'une ICO (définition inspirée de [Law et al., 2009])
- Les performances des systèmes ICO: L'efficacité totale de l'ICO, incluant sa vitesse, sa précision, et sa disponibilité.
- L'intégration des ICOs dans une application: comment une ICO est utilisée et exploitée par une application.

1. **L'expérience utilisateur ICO**

   Utiliser une ICO est une expérience inhabituelle. Dans les moindres aspects de la vie quotidienne, les interactions entre un humain et son environnement passent par le système nerveux périphérique. Les informations sur le monde sont reçues par des influx sensoriels. Cette information est traitée, et des décisions peuvent être prises. Ces décisions sont transmises sous forme d'activité musculaire, permettant de bouger et d'agir. Cette boucle perception-traitement-action semble si naturelle qu'elle est rarement étudiée, en dehors d'études de psychologie.

   Pourtant, les ICOs bousculent cette boucle. Plutôt que d'agir via les nerfs et les muscles, les utilisateurs d'ICO peuvent agir directement par leur activité cérébrale. Les conséquences de cette caractéristique unique ne sont pas encore bien connues. De nombreuses questions sur l'expérience d'un utilisateur d'ICO restent encore sans réponse. A ce jour, les spécificités de l'expérience utilisateur, l'influence de l'entrainement de l'utilisateur, ou encore la manière dont cet entrainement peut être amélioré reste essentiellement inconnues.

   Dans cette thèse, nous étudions l'influence des erreurs de l'interface sur l'expérience utilisateur, et nous cherchons quel quantité de ressources cognitives sont nécessaires afin d'utiliser une ICO avec succès.

2. **Les performances des systèmes ICO**

   Depuis les premiers travaux des années 1970 [Vidal, 1973], Les ICOs ont connu d'importantes améliorations. Il y a essentiellement deux raisons à ces améliorations: la découverte de nouveaux motifs cérébraux, et leur exploitation pour les ICO d'une part, et les avancées en traitement du signal d'autre part. Le traitement du signal pour les ICOs est responsable de la détection et de la reconnaissance d'un motif cérébral. Les améliorations du traitement du signal visent à réduire le nombre d'erreur de détection. Malgré des recherches approfondies sur l'optimisation du traitement du signal pour les ICOs, des défis majeurs subsistent. La vitesse et la précision peuvent encore être améliorées. Adapter les ICOs à un usage à la volée reste difficile, et la plupart des chaines de traitement du signal nécessite une calibration. Pour l'IHM, cette phase de calibration est un processus contraignant, qui devrait être rendu aussi court et efficace que possible.

Dans cette thèse, nous étudions plus précisément l'impact des conditions de calibration sur les performances finales, afin d'améliorer la précision des ICOs basées sur SSVEP, et de mieux comprendre l'impact des caractéristiques de la calibration.

3. **Intégration des ICOs aux applications**

   Les améliorations des ICOs mettent généralement l'accent sur l'amélioration de la détection d'un motif cérébral. Pourtant, plutôt que d'améliorer uniquement la précision des ICOs, il pourrait parfois être plus efficace d'améliorer son utilisation. Malgré d'importants progrès de la précision de détection, les erreurs des ICOs ne peuvent jamais être complètement éradiquées, et restent plus fréquentes que pour d'autres types d'interface (clavier, souris, mais aussi reconnaissance de gestes ou du langage). Un usage prudent des ICOs devrait permettre de réduire l'impact des erreurs, et ainsi d'améliorer l'interaction dans son ensemble. Les ICOs hybrides constituent une manière prometteuse de gérer un signal d'entrée bruité. En utilisant plusieurs appareils d'interaction, avec des qualités et des défauts complémentaires, il devrait être possible de compenser les erreurs des appareils d'interaction.

   Dans cette thèse, nous proposons une technique d'interaction hybride basée sur le suivi de regard et les ICOs basées sur SSVEP, avec pour objectif une meilleur exploitation de leurs qualités respectives. Nous étudions également comment les ICOs peuvent être intégrée à une application classique: un navigateur web.

Principales contributions de la thèse

Demande Cognitive d'une ICO basée sur SSVEP. La première étape, pour améliorer l'interaction utilisant des ICOs basées sur SSVEP, est de comprendre les besoins et les capacités de l'utilisateur. Lorsqu'un utilisateur interagit avec un ordinateur, il cherche à accomplir une tâche, et l'interaction elle-même n'est qu'un moyen d'y parvenir. Un important critère des techniques d'interaction est donc d'imposer une charge mentale faible à l'utilisateur, afin qu'il puisse se concentrer sur sa tâche principale. La plupart des ICOs nécessitent la pleine concentration de l'utilisateur [Daly et al., 2009]. Cependant, la demande cognitive induite par l'utilisateur d'une ICO basée sur SSVEP est encore inconnue. Quel niveau de concentration est requis pour diriger une ICO basée sur SSVEP ? Un utilisateur peut-il accomplir une tâche secondaire exigeante en même temps ?

Nous avons mené à bien une expérience évaluant la demande cognitive d'une ICO basée sur SSVEP. Lors de cette expérience, les participants devaient réaliser une tâche de mémorisation exigeante, en particulier en mémoire de travail, alors qu'ils utilisaient une ICO basée sur SSVEP en même temps. Nous avons ainsi observé l'impact de la difficulté de la tâche de mémorisation sur la précision de classification SSVEP. L'analyse statistique des résultats n'a révélé aucun impact. Cette observation a été faite à la fois quand les canaux visuels ou auditifs étaient sollicités par la tâche de mémorisation. Ces résultats indiquent que la sélection par SSVEP est une tâche à faible demande cognitive. Elle est sensible à la localisation de l'attention visuelle, mais l'attention peut être divisée entre plusieurs objets co-localisés. Une demande importante d'autres ressources cognitives telles que la mémoire de travail ou l'attention auditive n'affecte pas significativement le niveau d'attention visuelle. Pour l'interaction homme-machine, ces résultats sont encourageants pour l'utilisation de SSVEP, puisqu'ils indiquent que l'interaction basée sur SSVEP ne devrait pas significativement affecter l'attention de l'utilisateur sur sa tâche principale. Ils sont également encourageants pour l'utilisation de SSVEP pour des interfaces hybrides, puisque des ressources cognitives peuvent être allouées aux autres outils ou éléments graphiques de l'interface.

La Frustration de l'Utilisateur d'une ICO basée sur SSVEP Comparée à des interfaces classiques, basées sur une souris ou un clavier, les ICOs présentent un taux d'erreur particulièrement important. Il y a plusieurs raisons à cette imprécision: le bruit dans le signal, la difficulté pour l'utilisateur de produire le motif cérébral correct, les imperfections de l'algorithme de classification, etc. Le taux d'erreur varie grandement selon un certain nombre de facteurs, tel que l'utilisateur, le paradigme ICO, la qualité des électrodes, et bien d'autres. Même en condition parfaite, une précision de 100% ne semble pas réalisable dans un futur proche, ou même peut-être jamais. Pour l'IHM, on sait que les erreurs dans l'interface sont l'une des principales causes de frustration. Cependant, les spécificités des ICOs remettent en question la généralité de ce principe. Y a-t-il une plus grande tolérance aux erreurs pour les ICOs ? L'expérience utilisateur est-elle influencée différemment par les imperfections de l'interface ?

Nous avons mené à bien une expérience évaluant la frustration et la fatigue d'utilisateurs d'une ICO basée sur SSVEP, en fonction du taux d'erreur. Durant cette expérience, les participants devaient réaliser une tâche de sélection à l'aide d'une ICO basée sur SSVEP, alors que la précision de la sélection était artificiellement manipulée, sans que l'utilisateur n'en soit conscient. Des questionnaires ont été utilisés tout au long de l'expérience, afin d'estimer la frustration et la fatigue des participants. Les résultats montrent qu'un taux d'erreur important augmente la frustration de l'utilisateur, et que cette frustration s'accumule au cours du temps. Cette observation est cohérente avec les résultats d'études similaires menées à bien pour d'autres types de systèmes interactifs. Cependant, nous avons également observé que l'augmentation de la frustration ne semble pas critique pour de faibles différences de taux d'erreur.

Optimisation des conditions d'entrainement pour les ICOs basées sur SSVEP. Les ICOs basées sur SSVEP nécessitent l'utilisation de stimulations visuelles présentées à l'utilisateur, afin d'observer et de classifier sa réponse cérébrale. Les propriétés de la stimulation doivent être étudiées avec soin. Par exemple, des caractéristiques telles que la taille des cibles clignotantes, leur couleur, ou encore la distance entre les cibles, ont une influence sur la précision de l'ICO [Cao et al., 2012, NG et al., 2012, Regan, 1966]. Les ICOs basées sur SSVEP nécessitent également une phase de calibration, afin d'adapter le traitement du signal et le classifieur à l'utilisateur. Alors que l'influence des caractéristiques de stimulation durant l'utilisateur de l'ICO a été (partiellement) étudiée, il y a un manque de connaissance sur l'influence de telles caractéristiques pendant la calibration du système. Peut-on optimiser l'entrainement du classifieur avec un choix précis de caractéristiques de stimulation pendant la calibration ?

Nous avons mené à bien une expérience testant diverses conditions d'entrainement, afin de déterminer lesquelles produisent le classifieur le plus stable. Des données SSVEP ont été enregistrées lors d'une simple tâche de sélection à 3 cibles, avec diverses conditions de stimulation, variant suivant deux facteurs: la différence de couleur des cibles, et la distance entre les cibles. L'analyse des résultats indique qu'une grande distance entre les cibles produit un meilleur taux de classification. Ce résultat était attendu, puisque l'on sait qu'une faible distance entre les cibles amène du bruit dans l'activité cérébrale, dû aux plusieurs cibles clignotantes qui amènent des réponses cumulées. De manière plus surprenante, nous avons constaté qu'utiliser des cibles de couleurs différentes pendant l'entrainement du classifieur produit un classifieur plus robuste que d'utiliser uniquement des cibles noires et blanches. Ces résultats montrent qu'avoir des conditions similaires pendant l'entrainement d'un classifieur et son utilisation finale n'est pas toujours préférable. Il est possible d'entrainer un classifieur de manière efficace dans une condition, et de généraliser son usage à des contextes différents.

Vers l'Interaction Hybride et la Fusion d'entrées issues du Regard et du Cerveau
Les ICOs hybrides sont les systèmes interactifs qui utilisent au moins une ICO et un autre outil d'interaction (éventuellement une autre ICO) [Pfurtscheller et al., 2010]. De tels systèmes visent à utiliser plusieurs types d'entrées pour qu'elles compensent mutuellement leurs limitations respectives, tout en exploitant leurs qualités. Combiner les ICOs avec le suivi de regard semble être une association prometteuse, puisque ces deux entrées partagent l'avantage de laisser libre les mains de l'utilisateur. Pour concevoir des systèmes d'interaction mains-libres basés sur le suivi de regard et les ICOs basées sur SSVEP, il n'y a pas de consensus clair sur la manière de combiner les entrées. Quelle entrée devrait être utilisée pour quelle tâche d'interaction ? Est-il plus efficace d'utiliser le suivi de regard pour tout ? Devrait-on utiliser le suivi de regard et l'ICO pour des tâches d'interaction bien distinctes ?

Nous proposons une nouvelle technique pour l'interaction basée sur le suivi de regard et les ICO basées sur SSVEP, recherchant une meilleure exploitation de ces deux entrées complémentaires. Nous proposons d'utiliser une méthode de fusion des canaux d'entré d'ICO et de suivi de regard. Notre méthode peut être utilisée sans minutage imposé, et permet la sélection d'une cible parmi un ensemble de grande taille. Des tests ont été réalisés, dans lesquels les participants devaient sélectionner des cibles à l'aide de notre technique d'interaction, ainsi qu'avec deux autres techniques directement inspirées de l'état de l'art. La distance entre les cibles variait au cours de l'expérience. Les résultats montrent que les performances de notre technique de fusion sont supérieures à celle de la technique hybride séquentielle préexistante, à la fois en vitesse et en sensibilité, et pour toutes les distances testées. Cependant, nous observons également que la technique d'interaction basée sur le suivi de regard seul, utilisant la sélection par *dwell time*, était encore plus rapide et plus fiable que notre technique de fusion. Ce résultat indique que malgré leurs progrès, les ICOs basées sur SSVEP n'ont pas encore un niveau compétitif par rapport aux alternatives d'interfaces mains-libres.

La suite du manuscrit détaille ces contributions. La langue employée est l'anglais, pour pouvoir être lu par un jury de thèse international.

Introduction

The work presented in this manuscript is entitled “Optimizing the Use of SSVEP-based Brain-Computer Interfaces for Human-Computer Interaction”, and aims at improving SSVEP-based Brain-Computer Interfaces from the point of view of Human-Computer Interaction.

The expression *brain-computer interface* (BCI) describes a large family of systems measuring a user cerebral activity, and using this measurement for human-computer interaction. Nonetheless, variations can be observed in what experts call a BCI.

For [Lotte, 2008], “*A BCI is a communication system which enables a person to send commands to an electronic device, only by means of voluntary variations of his brain activity*”. This definition emphasizes the voluntary aspect of the communication. For BCIs as described in this definition, users are aware of the BCI, and they actively try to use it. In the literature, such systems are commonly referred to as *active BCIs*.

For [Nicolas-Alonso and Gomez-Gil, 2012], “*A BCI is a communication system between a user and a computer where the carried message is the measured user's brain activity*”. This definition is more general, as it allows implicit communication. Brain activity can be measured and used for interaction purposes without an explicit decision from the user. BCIs that do not require the user to actively try to use the system are sometimes called *passive BCIs* [Zander and Kothe, 2011].

Finally, we can consider another definition by [Cabestaing and Rakotomamonjy, 2007]: “*Brain-Computer interfaces (BCIs) are direct systems of communication between an individual and a machine, which do not use standard communication channels such as peripheral nerves and muscles*”. As for the previous definitions, the carried message is the cerebral activity. However, here the authors consider that the user does not need any level of motor control in order to use the BCI. The standard vocabulary is to call such systems *independent BCIs*, while BCIs that require any form of motor control (e.g. gaze control) are called *dependent BCIs* [Allison et al., 2007].

Human-Computer Interaction and Brain-Computer Interfaces

Human-Computer Interaction (HCI) is defined by [Hewett et al., 1992] as “*a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them*”. The goal of HCI is to optimize the use of interaction devices, in order to enable interaction between a user and an interactive system. The optimization criteria and the choice of interaction device are driven by the interaction context.

Brain-Computer Interfaces (BCIs) are one of the many existing interaction tools. Long left to the role of alternative interaction means for disabled people [Megalingam et al., 2013, Donchin et al., 2000, Blankertz et al., 2006b], BCIs have slowly earned interest from a wider range of application fields. With the advances of neurology and signal processing, and the invention of new BCI paradigms, they have become a promising tool for Human-Computer Interaction in various fields such as virtual reality [Lécuyer et al., 2008], video games [Nijholt et al., 2009, Nijholt and Gürkök, 2013], or more generally as a way to get information about the user's mental state (passive approach) [Lécuyer et al., 2013].

As the application fields of BCIs expand, the methods and concepts of HCI become more and more relevant to the improvement of BCIs. Instead of focusing solely on speed and accuracy, other criteria such as the simplicity of installation, the user comfort, or the fatigue need to be taken into account.

When designing a BCI for a well defined application (e.g. web browser), two main choices have to be made at an early stage of the software development, and have an important impact on the final user interface characteristics. What type of cerebral measure should be used, and what type of brain pattern should be recognized?

Cerebral activity can be measured by various techniques, varying in cost, ease of use, temporal and spatial resolution. In this work, we focus on BCIs based on electroencephalography (EEG). EEG is a non-invasive measurement technique, meaning that it does not require any surgical operation. EEG requires relatively cheap material, and has a rather good temporal resolution (typically 512 Hz). These characteristics make EEG a natural choice for most HCI applications. As for the choice of brain pattern, this thesis focuses on BCIs based on SSVEP. When the speed of interaction is an important factor, BCIs based on SSVEP (see below) are a competitive possibility, presenting a comparatively high speed and precision.

Steady-State Visually Evoked Potentials (SSVEP)

When the brain is faced with a periodic visual stimulation, it responds with an activity at the frequency of stimulation in the early visual cortical area [Herrmann, 2001]. This periodic response is called the Steady-State Visually Evoked Potential (SSVEP), and can be detected in the EEG signal, under the form of a peak of power at the frequency of interest.

SSVEP-based BCIs exploit this physiological phenomenon. In a typical setup, several targets are presented to the user, and flicker each at a different frequency. When the user focuses its visual attention on one of these targets, the resulting SSVEP can be detected, and the target is selected (or the associated command is activated) [Lalor et al., 2005, Müller-Putz et al., 2006, Zhu et al., 2010b].

SSVEP-based BCIs have been shown to be fast and accurate, for BCI standards, as they currently (in 2016) hold the record of the fastest information transfer rate for EEG-based BCIs [Nakanishi et al., 2014]. As BCIs speed and accuracy are in general their main limitations from an HCI perspective, SSVEP appears as one of the most promising BCI paradigm.

Issues and objectives

Human-Computer Interaction (HCI) features three main components (see Figure 2): a user, an interface, and an application on an electronic device (typically a computer). The interface role is to handle the communication between the user and the application.

Most of the research on BCIs so far has focused on signal processing and classification. After years of research, the speed and accuracy of BCIs have known great improvements. However, in order to bring BCIs to out-of-the-lab applications, the whole interaction process should be considered. The approach followed during this thesis has been to apply the methods and concepts of HCI to BCIs. Instead of improving only the core of BCIs, this thesis aims at studying the whole interactive system (see Figure 3):

- BCI user experience: The user's perceptions and responses resulting from the use of a BCI (definition adapted from [Law et al., 2009]).
- BCI system performance: The global effectiveness of the BCI, including its speed, accuracy, and availability
- Integration of BCIs in applications: The way BCIs are used and exploited in an application.

1. **BCI user experience**

   Using a BCI is an unusual experience. In every aspect of everyday life, all interactions between a human being and his environment pass by his peripheral nervous system. Information about the world is received via sensory input. This information is processed, and decisions can be taken. These decisions are transcribed in muscle activity, allowing moving and performing actions (see Figure 3). However, BCIs are game-changing in this traditional cycle. Instead of performing actions via peripheral nerves and muscles, BCI users are given the possibility of performing actions directly with their brain activity. The consequences of this specificity are not well known yet. Several questions regarding BCI user experience still remain open. As for today, the specificities of user experience, the influence of user training, and the way user training can be improved are still mostly unexplored.

   In this PhD, we study the influence of errors in the interface on user experience, and we explore how much cognitive resources are required in order to successfully drive a BCI.

2. **BCI system performance**

   Since the pioneering work of the 1970s [Vidal, 1973], BCIs have known huge improvements. There are mainly two reasons for these improvements: the discovery of new brain patterns, with their exploitation for BCIs, and the advances in signal processing. The signal processing of BCIs is responsible for the detection and recognition of a specific cerebral pattern (see Figure 3). Improvements in signal processing aim at reducing the number of errors in the detection. Despite the extensive research on optimizing signal processing for BCIs, major challenges remain. Speed and precision can still be improved. Adapting BCIs to a self-paced use remains difficult, and most signal processing chains require a calibration. For HCI, this calibration phase is a constraining process, which should be made as short and efficient as possible.

   In this PhD, we study more specifically the impact of calibration conditions on the end performances, in order to both improve SSVEP-based BCIs accuracy and better understand the impact of calibration conditions.

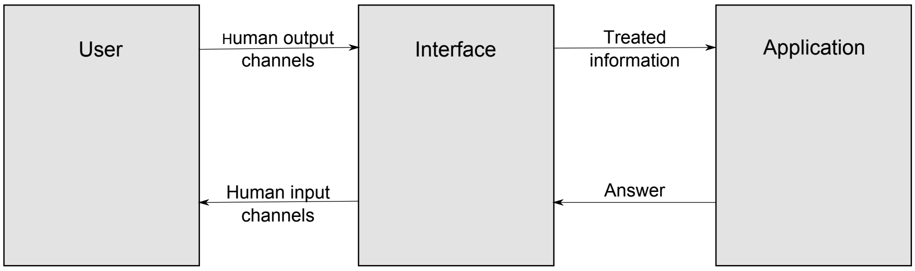


Figure 2: Human-Computer Interaction loop: The interface mediates interaction between the user and the application

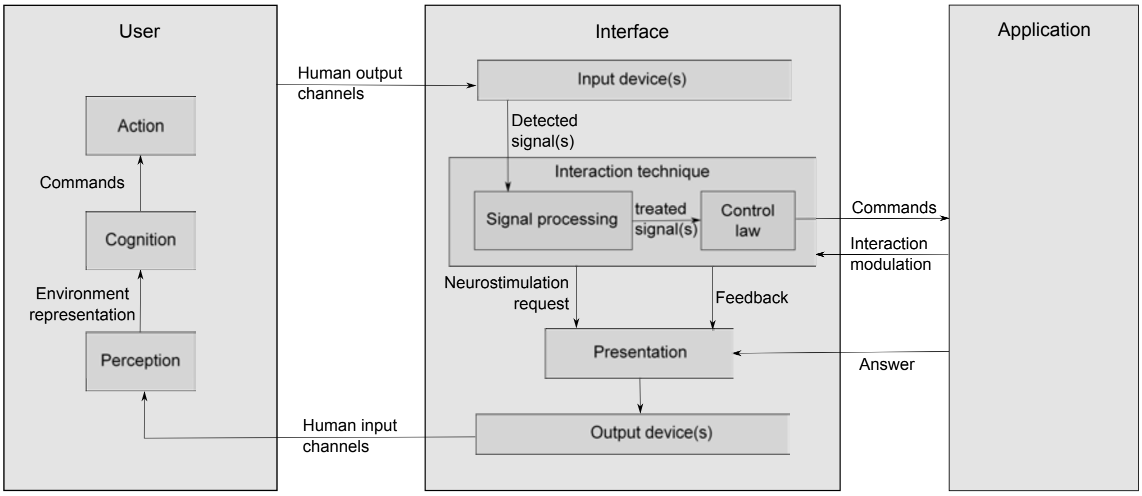


Figure 3: BCI-based interaction loop: A more detailed description of the Human-Computer Interaction loop, emphasizing the components of interest for BCI-based interaction. The user model can be decomposed into perception, cognition, and action processes. Input devices measure human activity, including cerebral activity. This measure is then processed and interpreted into a command. The application recieves the generated command and produces an appropriate response. Output devices provide the required neurostimulation, the feedback from the interface, and the response from the application.

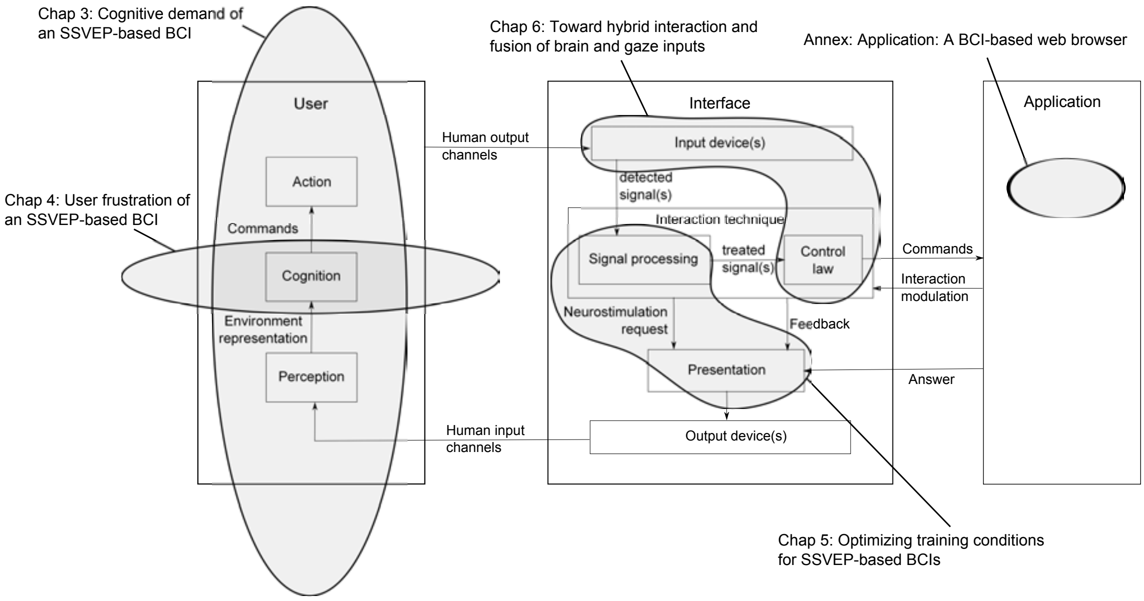


Figure 4: Contributions of this PhD. Chapter 3 and 4 focus on the user experience. Chapter 5 is related to signal processing and stimulation issues. Chapter 6 proposes a control law for hybrid interface combining gaze and EEG inputs. Annexes describe a prototype of application: a web browser based on BCIs.

3. **Integration of BCIs in applications**

   Improvement of BCIs usually emphasizes the improvement of the mental state detection. However, instead of only improving the accuracy of the BCI, it could sometimes be more efficient to improve its use. Despite great progress of the SSVEP accuracy, errors can never be eradicated, and are still more frequent than for most other interfaces systems (keyboard, mouse, but also gesture or speech recognition). Careful use of the BCI output should allow reducing the impact of errors, and improving the interaction as a whole. Hybrid BCIs constitute a promising way of handling noise in an input such as the one given by a BCI. By using several interaction devices, with different assets and flaws, it should be possible to compensate for each other limitations.

   In this PhD, we propose a hybrid interaction technique based on gaze tracking and SSVEP BCIs input, with the goal of better exploiting their respective strengths. We also study how BCIs can be integrated in a classical application: an web browser.

Approach and Contributions

The work presented in this PhD aims at addressing the three main aspects of BCI-based interaction loop: the BCI user experience, the BCI system performances, and the integration of BCIs in applications. Our approach emphasizes user studies and empirical evaluations.

Chapter 1: Fundamentals of Brain-Computer Interfaces. The first chapter presents a summary of previous research related to the topic of this PhD, by focusing on the fundamental principles of BCIs. We first present the main steps of EEG signal acquisition and processing for BCIs, followed by the discussion of the most common brain activity patterns used for brain computer interfaces. A special attention is given to SSVEP-based BCIs.

Chapter 2: A Human-Computer Interaction perspective on Brain-Computer Interfaces. The second chapter introduces several concepts of Human-Computer Interaction and discusses the characteristics of BCIs from the HCI point of view. Moreover, the main interaction paradigms for BCIs are presented and discussed. A short presentation of the main application fields of BCIs concludes this chapter.

Chapter 3: Cognitive demand of an SSVEP-based BCI. The first step in order to improve interaction with SSVEP-based BCIs is to understand the user needs and capabilities. When a user interacts with a computer, he/she aims at accomplishing a task, and the interaction itself is just a requirement for this goal. Thus, an important criterion of interaction techniques is to impose a low mental workload to the user, so that he/she can focus on his/her main task. Most BCIs require the full concentration of the user [Daly et al., 2009]. However, the cognitive demand induced by the use of an SSVEP-based BCI is yet unknown. How much concentration is required to drive an SSVEP-based BCI? Can a user perform a demanding secondary task at the same time?

We present an experiment evaluating the cognitive demand of an SSVEP-based BCI. Participants were asked to perform a demanding task, especially in terms of working memory, while using an SSVEP-based BCI at the same time. We study how the dual task influences the BCI accuracy, and interpret our observations from the point of view of cognitive resources limitation.

Chapter 4: User frustration of an SSVEP-BCI. When compared to classical interfaces such as a computer mouse or a keyboard, BCIs show a particularly high error rate. There are several causes to this imprecision, including noise in the signal, difficulty for the user to produce the correct mental state, and classification algorithm imperfections. The error rate varies greatly depending on a number of factors, such as the user, the BCI paradigm, the electrodes quality, or many others. Even in optimal conditions, a 100% accuracy does not seem achievable any time soon, if ever possible. For Human-Computer Interaction, it is well-known that errors in the interface are one of the primary causes of frustration. However, the specificities of BCIs question the generalization of this principle. Is there a higher tolerance towards errors in BCIs? Is the user experience differently impacted by the interface imperfections?

We present an experiment evaluating the frustration and fatigue of SSVEP-based BCI users depending on the error rate. Participants are asked to perform a selection task using an SSVEP-based BCI, in which the accuracy of the selection is artificially manipulated, and we observe the influence of this accuracy on user experience.

Chapter 5: Optimizing training conditions for SSVEP-based BCIs. BCI systems based on SSVEP require the use of a visual stimulation presented to the user, in order to observe and classify its cerebral response. The properties of the stimulation must be carefully considered. For example, characteristics such as the size of flickering targets, their color, or the distance between targets, have been empirically shown to have an influence on the BCI accuracy

[Cao et al., 2012, NG et al., 2012, Regan, 1966]. SSVEP-based BCIs also require a calibration phase, in order to adapt the signal processing and the classifier to the user. While the influence of stimulus characteristic during BCI use has been partly studied, there is a lack of knowledge on the influence of such characteristics during the system calibration. Can we optimize the classifier robustness with a specific choice of stimulus characteristic during the calibration?

We describe an experiment testing various training conditions, in order to determine the ones leading to the more robust classifier. SSVEP data are recorded in a simple 3-target selection task, with various conditions of target colors and distance between targets. We then analyze how these factors influence the calibration quality.

Chapter 6: Toward hybrid interaction and fusion of brain and gaze inputs. Hybrid BCIs are interaction systems that involve the use of at least one BCI, together with another interaction device (possibly another BCI) [Pfurtscheller et al., 2010]. Such system aims at using several types of inputs that compensate for each other limitations, while fully using their strengths. Combining BCIs with gaze tracking seems a promising association, as both of these inputs share the advantage of being hands-free. When designing a hands-free interaction system using gaze tracking and a SSVEP-based BCI, there is no clear consensus over the optimal combination of inputs. Which input should be used for each interaction task? Is it more efficient to use the gaze tracking for everything? Should one use gaze tracking and BCIs for distinct interaction tasks?

We propose a new interaction technique for target selection, using gaze tracking and a SSVEP-based BCI. This interaction technique aims at combining both input modalities in order to better exploit these two inputs. We introduce a method enabling the fusion of BCIs and gaze tracking inputs, and we analyze the resulting selection speed and sensitivity.

Annexes. Annexes focus on an application example. First, in annex A, a state-of-the-art signal processing method for SSVEP detection and classification is described in details. Potential improvements are discussed and tested. In particular, the signal processing method resulting from these adaptations is both calibration-free and self-paced.

Finally, annex B presents an application of the research carried on in this thesis. It presents the conception and design of a web browser prototype based solely on SSVEP and P300-based BCIs. This hybrid approach aims at optimizing the human-computer interaction for web browsing based on brain-computer interfaces.

Conclusion. The manuscript ends with a general conclusion.

Vocabulary and acronyms

- BCI: Brain-Computer Interface
- BIT: Basic Interaction Task, not to be mismatched with the bit: the unity of information.
- CAR: Common Average Reference
- CCA: Canonical Correlation Analysis
- CSP: Common Spatial Pattern
- EcoG (or ECG): Electrocorticography
- EEG: Electroencephalography
- EOG: Electrooculographie.
- ERD: Event Related Desynchronization
- ERS: Event Related Synchronization
- ErrP: Error Potential
- fNIRS: Functional Near-infrared Spectroscopy
- fMRI: functional Magnetic Resonance Imagery
- GUI: Graphic User Interface
- HCI: Human-Computer Interaction
- LDA: Linear Discriminant analysis
- MEG: Magnetoencephalography
- MI: Motor Imagery
- MRI: Magnetic resonance imagery
- PCA: Principal Component Analysis
- PNS: Peripheral Nervous System
- SCP: Slow cortical potential
- SL: Surface Laplacian
- SSEP: Steady State Evoked Potential
- SSSEP: Steady-state somatosensory evoked potential
- SSVEP: Steady-state visually evoked potential
- SVM: Support Vector Machine

Chapter 1

Fundamentals of Brain-Computer Interfaces

This first chapter presents the main concepts of Brain-Computer Interfaces. First, a categorization of BCIs is presented, followed by the main steps of signal measurement and processing techniques. The cerebral patterns most frequently used for BCIs are then briefly described, before a more detailed presentation of BCIs based on Steady-State Visually Evoked Potentials (SSVEP).

1.1 Categorization of BCIs

BCIs are usually categorized by the type of use that is made of the recognized mental state. They are usually decomposed into three categories: active, reactive, and passive BCIs [Zander et al., 2010b]. The choice of a specific type of BCI strongly depends on the application context.

- For **active BCIs**, the user is explicitly trying to produce the correct mental state. The most common active BCIs are based on motor imagery [Nicolas-Alonso and Gomez-Gil, 2012, Allison et al., 2007]. Active BCIs are based on brain activity which is directly and consciously controlled by the user, i.e. an activity that can be produced by the user without the need for an external stimulation. Producing the desired activity at will often requires some user training.
- For **passive BCIs**, cerebral activity is detected without the need for the user to actively produce it. BCIs of this category are called passive BCIs. Detected mental states include emotional states (dominance, arousal, pleasure) [Heraz and Frasson, 2007], frustration [Reuderink et al., 2009], mental workload [Blankertz et al., 2010], and concentration level [Lee et al., 2009].
- Finally, a **reactive BCI** is one that derives its outputs from brain activity arising in reaction to external stimulation (typically a flash of light). The cerebral response is indirectly modulated by the user for controlling an application [Zander and Kothe, 2011]. The detected brain activity is then called an externally evoked potential. In a typical setup, several stimulations

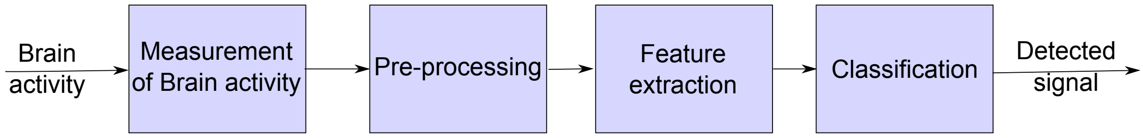


Figure 1.1: General architecture of signal processing for BCIs.

are sent to the user, who focuses on one of them. Based on the difference in brain responses between attended and unattended stimuli, the BCI detects which is the target of interest. The most common reactive BCIs are based on P300 or SSVEP patterns.

1.2 Signal measurement and processing

Historically, the main shortcoming to the development of BCIs is the difficulty to measure brain activity, and to interpret the resulting signal. Over the past century, new sensors have been developed, and great advances have been made in signal processing and machine learning. For BCIs, the signal processing can be decomposed in four steps [Mason and Birch, 2003] (see Figure 1.1):

- Acquisition of a signal depending on mental activity, typically with EEG headset.
- Preprocessing of the signal, aiming at improving the signal quality.
- Feature extraction, to reduce the dimension of the signal.
- Classification of the feature vector / extraction of a mental state.

These steps are common to active, passive, and reactive BCIs. However, the choice of technique for each step strongly depends on the cerebral pattern to be detected, and influences BCI characteristics such as delay, calibration, number of electrodes, etc.

1.2.1 Sensors and activity measurement

Several different sensors have been used to register the brain activity (see below). These sensors measure different kind of physiological signals related to brain activity, and present very different characteristics in terms of cost, spatial resolution, and temporal resolution (see table 1.1):

- functional Magnetic Resonance Imagery (fMRI) measures a blood oxygen level-dependent signal, indirectly measuring the local consumption of energy in the brain [Weiskopf et al., 2004].
- Electrocorticography (EcoG) measures the potential on the surface on the head using electrodes implanted directly on the skull [Leuthardt et al., 2006]. The signal is similar to EEG (see below), but the measure is less sensitive to noise.

Sensors	Cost ($)	temporal resolution	spatial resolution	portability
EEG	$[800 - 30000]$	$1ms$	$[2-3]cm$	high
MEG	$2millions$	$1ms$	$[2-3]cm$	poor
fNIRS	10000	$10s$	$1cm$	high
fMRI	$500000 - 3millions$	$3-6s$	$[2-3]mm$	poor
EcoG	*medical*	$5ms$	$1cm$	poor (invasive)

Table 1.1: Comparison between different types of sensors for brain activity acquisition. For each category, a wide range of devices exists. This table only gives an order of magnitude. Temporal and spatial resolution do not predict sensibility to noise. Different types of devices are sensitive to distinct physiological phenomenon. All data in this summary are order of magnitude and should be taken with caution, especially since further advances could improve them [2].

- functional Near-Infrared Spectroscopy (fNIRS) measures physiological changes in the brain tissues by observing their changes in optical properties. The optical properties are measured by sending near-infrared light toward the brain, typically in the frontal areas, and by observing the reflection [Irani et al., 2007].
- Magnetoencephalography (MEG) measures the magnetic field on the surface of the head. This magnetic field is influenced by simultaneous spike of a high number of neurons with the same orientation [Purves et al., 2012].
- Electroencephalography (EEG) measures the electric potential on the surface of the head. According to [Lotte, 2008], *EEG measures the sum of the post-synaptic potentials generated by thousands of neurons having the same radial orientation with respect to the scalp.*

The diversity in the acquisition methods used to measure brain activity may be explained by their complementary performances (see table 1.1). Some (fMRI, fNIRS) present a better spatial resolution, meaning that they measure changes in small brain areas, while others (EEG, MEG) are more efficient in term of time resolution, but less precise concerning the localization of the activity. Finally, the cost and encumbrance of these systems present strong variations, making the most expensive and precise ones (fMRI, EcoG, MEG) more used in medical context, while cheaper ones (EEG, fNIRS) can be more easily applied to a wider range of applications. The interested reader can refer to [Waldert, 2016] for a discussion on the assets and limitations of invasive versus non-invasive BCIs.

In this review, we focus on systems based on EEG, by far the most used cerebral activity measurement system for BCIs. The dominance of EEG is mainly due to its good portability and comparatively low cost (see figure 1.1). The time resolution of EEG is high (milliseconds time scale), at the cost of a poor spatial resolution (2 to 3 cm range in the best cases).

An EEG headset is relatively cheap, compared to other brain imaging sensors. Medicals headset cost ranges between 5000$ and 30000$ (including the amplifier), whereas Emotiv EPOC Device,

[2] Data were collected from several research publication [Weiskopf et al., 2004, Leuthardt et al., 2006, Irani et al., 2007, Sauvan et al., 2009, George et al., 2014, Müller-Putz et al., 2006, Vaughan et al., 2006, Strangman et al., 2002, Ferrari and Quaresima, 2012]. Finding indication on medical material prices required other sources: `https://www.emotiv.com/`, `http://www.gtec.at/Products`, `https://www.quora.com/How-much-does-an-fMRI-machine-cost`, `https://info.blockimaging.com/bid/92623/mri-machine-cost-and-price-guide`, `http://neurogadget.net/2012/12/15/inexpensive-magnetoencephalography-meg-system-could-be-available-at-every-hospital/6495`

Figure 1.2: Left: g-tec EEG headset, with 32 g.LADYbird active electrodes. Right: Emotiv EPOC low-cost EEG Headset, using dry electrodes [4]

designed for commercial application, only cost 800$ (see Figure 1.2). Medical headsets require applying a conductive gel to ensure the electrical connectivity between the electrodes and the scalp, but some models use a saline solution instead of gel (e.g. Emotiv), or even dry electrodes, at the cost of a generally lower signal quality [Duvinage et al., 2013]. These qualities make EEG better suited to non-medical applications, like video games, or interfaces evaluation (see chapter 2.4).

Traditionally, and in order to facilitate the comparison between studies, electrodes are placed according to the extended international 10-20 system (see figure 1.3), defining reference positions on the scalp [Tanner, 2011]. This system can be further extended to describe the position of more electrodes (see figure 1.4) [Klem et al., 1999]. All the electrode positions names referred to in this PhD use the extended 10-20 system.

The number of electrodes in an EEG headset goes from 1 to 256 [Teplan, 2002], depending on the required precision and the brain activity pattern detected. Each electrode continuously measures a signal of a few micro-volts, and provides a channel of data that can be used for by BCI signal processing and classification. One temporal signal per channel is recorded (see Figure 1.5). This signal is noisy [Teplan, 2002], and of very high dimension (linearly depending of the number of channels and time epoching).

1.2.2 Signal preprocessing

In order to improve the signal-to-noise ratio, some preprocessing steps are used. Preprocessing methods can be decomposed into three categories. Temporal filters, spatial filters, and artifact filters. While temporal and spatial filter aim at reducing general noise, artifact filters focus on some specific type of noise, caused by well identified sources. In order to get the best signal-to-noise ratio, these filters should be used together.

[4] Images from `http://www.gtec.at/Products/Hardware-and-Accessories/g.Nautilus-PRO-Specs-Features` and `https://www.emotiv.com/product/emotiv-epoc-14-channel-mobile-eeg/`

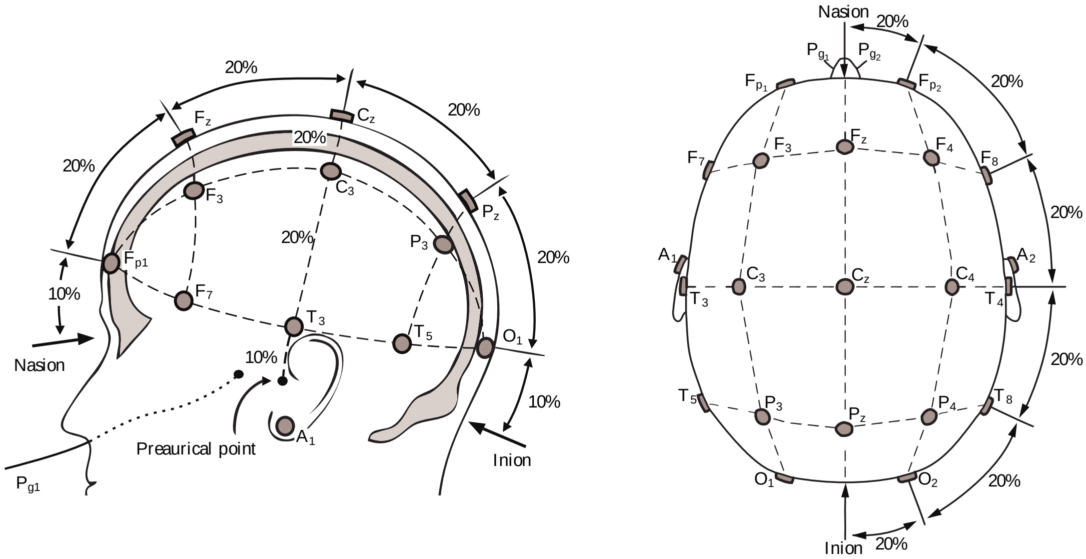


Figure 1.3: Electrodes placement according to the international 10-20 system [Tanner, 2011]. Left: Sagittal view. Right: Axial view

1.2.2.A Temporal filters

The goal of temporal filters is to remove the signal component of irrelevant frequencies. That is: frequencies that are not linked to the mental state to be recognized.

Low-pass filter: A low-pass filter is applied to the time signal, typically with a threshold around 80 or 100 Hz, depending on the application. Few studies have successfully extracted information in higher frequencies using EEG [Ferrez et al., 2006].

High-pass filter: A high-pass filter is applied, (except for SCP-based BCIs, see section 1.3.1), with a threshold around 1 or 2 Hz. Slower variations of potential are linked to cortical excitability [Birbaumer et al., 1990], they can't be controlled without a long and specific training [Birbaumer et al., 2003], and are usually not relevant for BCI.

The exact threshold of the filters can change from one application to another. The filtering itself can be done using direct Fourier Transform, or more advanced technique like finite impulse response or infinite impulse response [Smith, 1997]. These last two techniques aim at learning optimal filters for classification, depending on the signal itself over the last few seconds.

1.2.2.B Spatial filters

In addition to temporal filters, it is common to use some spatial filters, especially when the number of available electrodes is high enough. The use of spatial filter allows reducing the noise by averaging

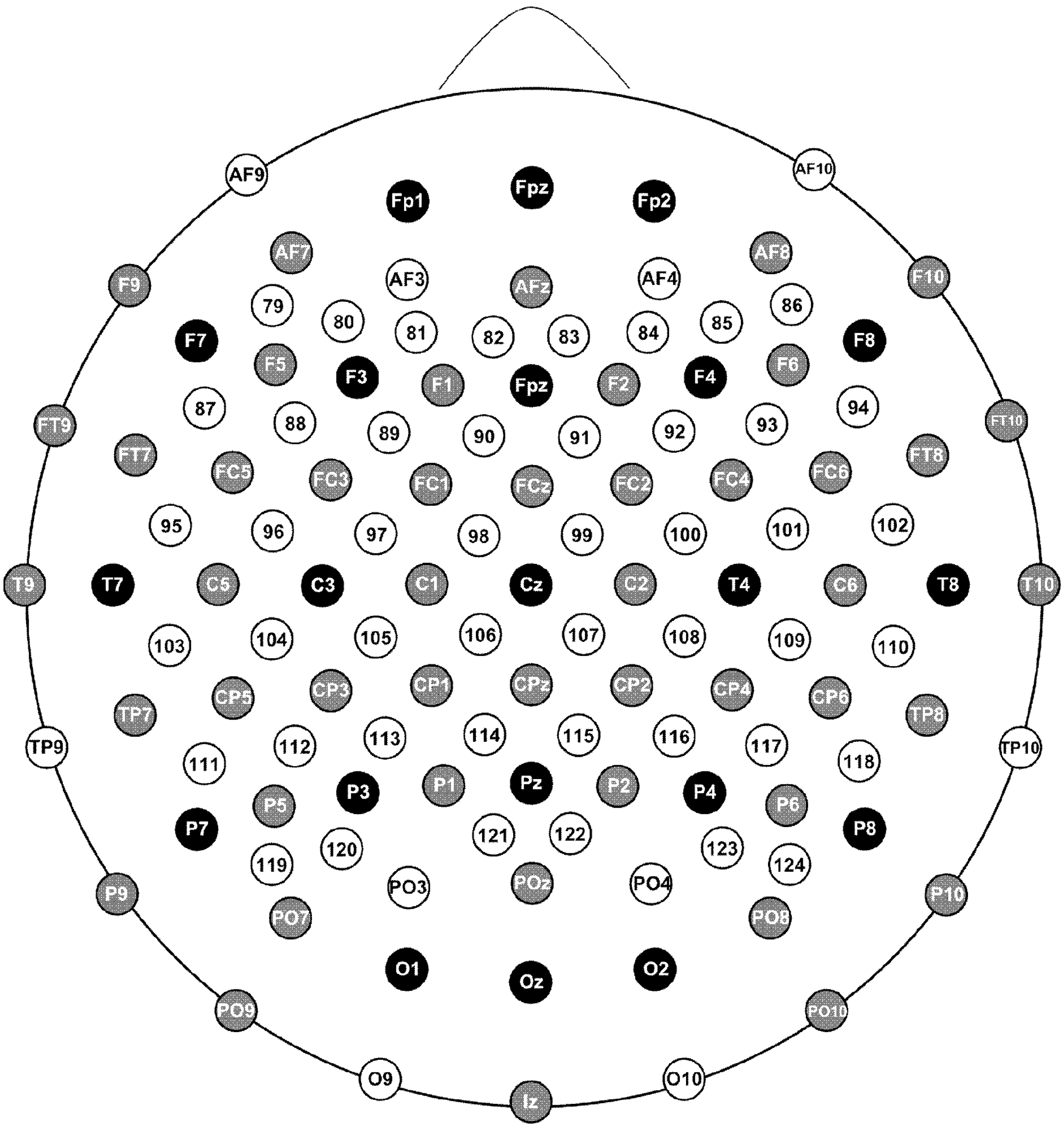


FIG 1

Figure 1.4: Extended 10-20 system for electrodes placement [Klem et al., 1999]

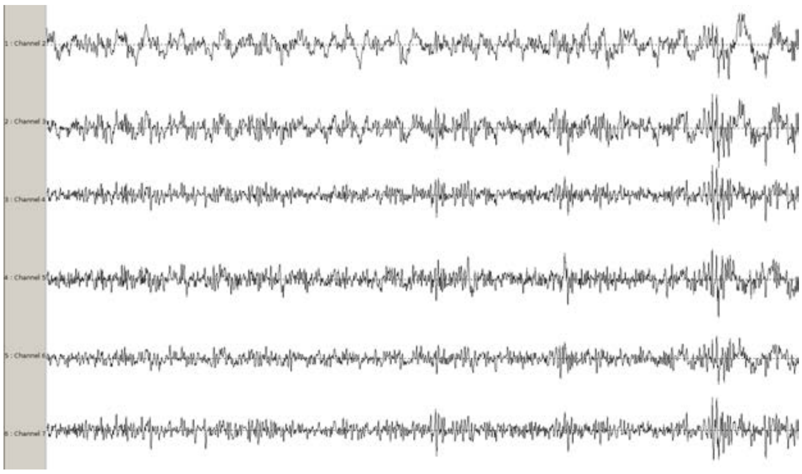


Figure 1.5: Temporal evolution of a typical EEG signal. Each line represents one channel.

the signal over nearby electrodes.

Channel selection: If the number of channels is high, it is often useful to select the most relevant ones, or at least weight them [Singh et al., 1990]. For example, if the BCI looks for a brain pattern taking place in visual areas, located in the occipital lobe, it may be useful to select the electrodes located closer to this area.

Common average reference: It is also possible to increase the signal-to-noise ratio using the common average reference (CAR). It removes the average value of all the electrodes. This average is supposed to be meaningless (and influenced by exterior electrical activity [McFarland et al., 1997].

Surface Laplacian: The Surface Laplacian (LS) filter the signal on each electrode by subtracting the average signal on the closest electrodes, thus reducing locally the background activity. It is generally used together with the CAR. This can be especially useful in the case of a high number of electrodes, since having enough electrodes close to each other is needed to compute the Laplacian precisely [McFarland et al., 1997].

Source reconstruction: The goal of source reconstruction is to reconstruct the information about 3D electromagnetic dipoles given the surface potentials measured by EEG. If source reconstruction is done successfully, it allows working on the 3D-brain activity, instead of using only the surface information. Such an approach presents the advantage of producing interpretable classifier results. [Mattiocco et al., 2006].

Several mathematical models of reconstruction have been proposed, corresponding to different models of the head shape and properties. An advantage of such methods is that they allow getting information about the brain behavior given the information learned by the classifier. This makes such an approach very useful for neuroscience studies. If neuro-anatomical data are available (IRM), they can be taken into account in the reconstruction [Mattiocco et al., 2006].

Common spatial pattern (CSP): Instead of looking for a dimensionality reduction imposed by the programmer, the spatial filter is statistically learned in order to maximize the differentiability between two classes of mental states [Blankertz et al., 2008]. Provided that the amount of data is sufficient, CSP aims at capturing the same information as source reconstruction, without the need for previous knowledge on the use anatomical characteristics.

1.2.2.C Artifact correction

Any signal that is not generated by brain activity is considered an artifact. Some artifacts are physiological, while others might come from the environment.

Eye blinking and muscle contractions (e.g. raw muscles) can cause artifacts whose amplitude is much higher than the signal's. This makes them easily recognizable, and the corresponding part of the signal can be excluded [Fatourechi et al., 2007].

Others are less visible, such as eye movement (gaze direction), heat pulse or muscle contraction (muscles of the face and neck typically generate stronger artifacts than muscles situated further away from the head). Sweat also causes artifacts, by changing the conductivity of the skin. For a more complete survey on artifacts in Brain Computer Interaction, see [Fatourechi et al., 2007].

Finally some artifacts can be generated by the environment. One of the most important ones corresponds to the power supply frequency. Depending on the country, this power can be supplied at 50 or at 60 Hz. Most of the American continent uses 60Hz, with the notable exception of Argentina, while 50Hz is more common elsewhere. Both frequencies are found in Japan [5]. Since this last artifact occurs at a very specific frequency, it is possible to filter out this frequency, by excluding the 49-51 Hz (or 59-61 Hz) band of frequency in the entire signal.

1.2.3 Feature extraction

A feature of a signal is a prominent or distinctive part of this signal. Concretely, features are relevant information extracted from the signal to help classify it. The goal of this step is mainly to reduce the dimensionality of the signal, to allow a training of the classifier with a reasonable amount of data.

For example, for a one-dimensional time signal, sampled at $100Hz$ over 10 seconds, a feature vector could be the values of the spectral density power of this signal at $5Hz$, $10Hz$, and $15Hz$ frequencies. Such a feature vector would have three dimensions, while the original signal would be a 1000 dimension vector (length of the signal times sampling frequency). The main goal of feature extraction is to reduce the signal dimension while keeping as much relevant information as possible, making it more easy to classify. Additionally, non-linear transformation during feature extraction can improve the expressiveness of the following classification.

The features extracted strongly depend on the BCI paradigm. Here are the most common ones:

[5] `https://en.wikipedia.org/wiki/Utility_frequency`

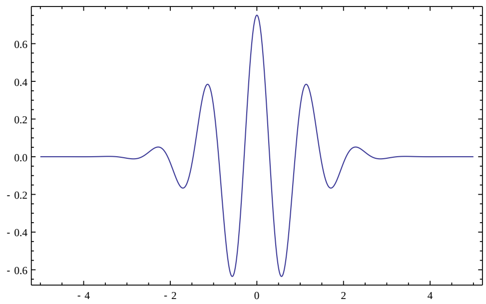


Figure 1.6: Shape of a typical wavelet signal: the Morlet Wavelet [6]

Time Signal: For some applications, it is possible to use directly the potential on every electrode at each time step (see Figure 1.5) [Rebsamen et al., 2007]. The signal then needs to be sub-sampled in order to reduce its dimensionality. The resulting feature vector is still of very high dimensionality (compared to the other classical ones, see below). Typically, an EEG signal can be sub-sampled at $100Hz$, and decomposed in windows of a few seconds. For a dozen of electrodes, this would lead to a feature vector with a dimension of a few thousands.

Band power: Some bands of frequency power in Fourier Space are often used as features [Van Gerven et al., 2009]. Depending on the frequency (and eventually of their localization), they can be related to different mental states. Tables 1.2 to 1.6 present the classical decomposition of cerebral rhythms, with the temporal representation of an EEG channel, filtered at the associated band of frequency:

Wavelet decomposition: While the Fourier transform makes a convolution between the time signal and a perfect sinusoidal, it is possible to compute the convolution product of the signal with "wavelet" signal, which allow extracting frequential information without losing all the time information [Samar et al., 1999]. The signal is then decomposed in a base of functions of the form:

$\Phi_{a,b}(t) = \frac{1}{\sqrt{a}}\Phi(\frac{t-b}{a})$, Φ being a wavelet function define in advance, the aspect of this mother wavelet influence the type of information measured by this technique (see Figure 1.6) [Samar et al., 1999]. The weight of each component of the signal in the wavelet decomposition is used as feature.

Canonical correlation analysis: Canonical correlation analysis (CCA) is a statistical method used on vectors of random variables. It computes the linear combination of variables in each vector to maximize the correlation between these combinations.

This is a frequently used method in data mining to recognize a known pattern. Let $X = (X_1, ..., X_n)$ and $Y = (Y_1, ..., Y_n)$ be two vectors of random variables CCA will find a and b two vectors that maximize $corr(a'X, b'Y)$ by solving the following problem:

$$\max_{a,b} \rho(X,Y) = \frac{a'cov(X,Y)b}{\sqrt{a'cov(X,X)a}\sqrt{b'cov(Y,Y)b}} \tag{1.1}$$

[5] https://en.wikipedia.org/wiki/Morlet_wavelet

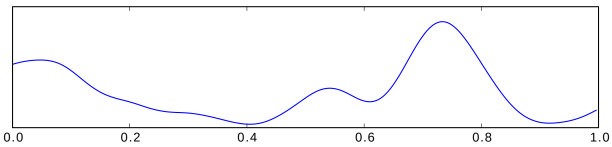


Table 1.2: Delta Rhythms (1 − 4 Hz): Observed almost exclusively in adult in deep sleep [Maquet et al., 1997].

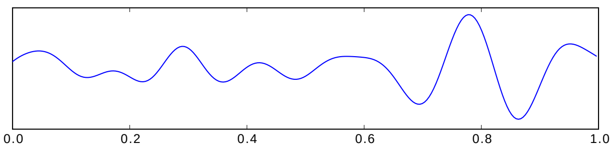


Table 1.3: Theta Rhythm (4−8 Hz): Mainly observed during childhood. It is also seen in drowsiness state, but not in deep sleep. Theta waves take their source in hippocampus, deep under the surface of the brain, which make difficult to detect them using EEG [Cantero et al., 2003].

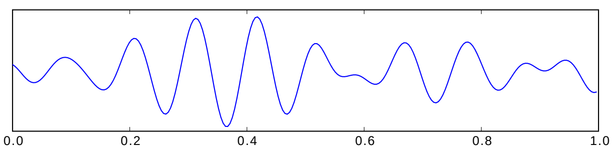


Table 1.4: Alpha Rhythm (8 − 13 Hz): Relates to posterior regions and is linked to relaxation and darkness (closed eyes) [Pivik and Harman, 1995]. This rhythm can often be observed directly by looking at the EEG signal. Mu Rhythm (8 − 13 Hz): This rhythm has the same band of frequency as alpha, but occurs in motor and somatosensory areas. It is blocked when the subject perform movements [Pineda, 2005].

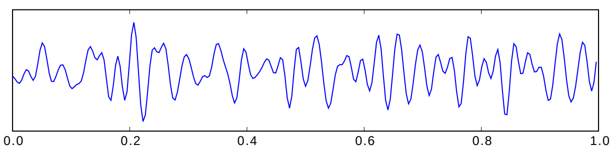


Table 1.5: Beta Rhythm (13 − 30 Hz): Rhythm observed in awaken and conscious person. It is also affected by movement (especially movement changed or inhibition), in the motor areas [Baker, 2007].

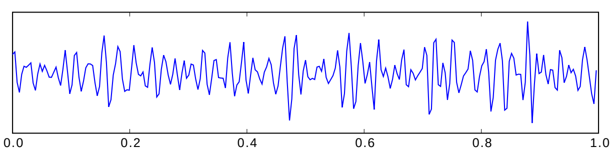


Table 1.6: Gamma (> 30 Hz): Rhythm linked to mental activity, especially to motor functions. It is difficult to detect, and has no clear interpretation [Vanderwolf, 2000].

CCA-based classifiers usually compute CCA between a signal segment of a few second and a set of reference signals. For BCI use, the reference signal would be exact aspect of the expected researched pattern, without noise. One correlation value is computed for each class, giving an indicator of how similar the EEG signal is to the stereotypical reference signal [Lin et al., 2006, Bin et al., 2009].

Higher level features: If the detected brain pattern is well-known, it is possible to use specific features specially adapted to this pattern (e.g. for P300, see below) [Abootalebi et al., 2009].

Hjorth parameters: These values describe the dynamic of a signal, based on its statistical behavior [Obermaier et al., 2001]. They can be used as complementary features.

1.2.4 Classification

The classification step aims at recognizing a class of cerebral activity, using the feature vector as input. Given the variability of the cerebral signal depending on the subject, classifiers generally require a calibration phase [Haw et al., 2006]. The most common approach is to record data offline, in order to calibrate the classifier, and then to use it in an application. In some cases the feature extraction step also requires a calibration (e.g. CSP, see below). However, in some cases, it is possible to train the classifier online, improving its performances little by little as the application is used [Gan, 2006]. In order to achieve the recognition of classes of brain pattern, various classification algorithm are used, each with their advantages and limitations:

Linear Discriminant Analysis (LDA): LDA is computationally cheap. The vector x of features is considered as a point in a multidimensional space. The goal is to define a partition of the feature space into two classes. With LDA, this partitioning is defined by a hyper-plane of separation, with parameters $\mathbf{W}$. Each side of the hyper-plane corresponds to one class (see figure 1.7). The hyper-plane is learned in order to maximize the distances between the two classes' respective means and to minimize the inter-class Variance [Fukunaga, 2013]. If more than two classes have to be discriminated, an LDA can be learned for any pair of classes, and an additional level of classification used to choose given the choice given by these LDA for each pair of classes.

Support Vector Machine (SVM): SVM is similar to LDA, in that it also separates the data space by a hyper-plane. It shows a good resistance to noise, and is relatively stable, but has a high computational cost. The difference with LDA lies in the strategy to choose the hyper-plane. Here it is chosen to maximize the margin, i.e. the distance of the hyperplane to the nearest training examples (see figure 1.8) [Burges, 1998].

Neural networks: Several linear classifiers are used together, regrouped in a network of classifiers separated in several levels. The outputs of the classifiers of level i are used as input for the classifiers of the level $i + 1$. The idea behind this organization is to use classifiers at level i to learn the best features for classifiers of level i+1. Such neural networks can represent high level of information (non-linear separation between classes), but need an important amount of data to be trained (as each sub-classifier adds more parameters to set) [Bishop, 1996]. Additionally, a high

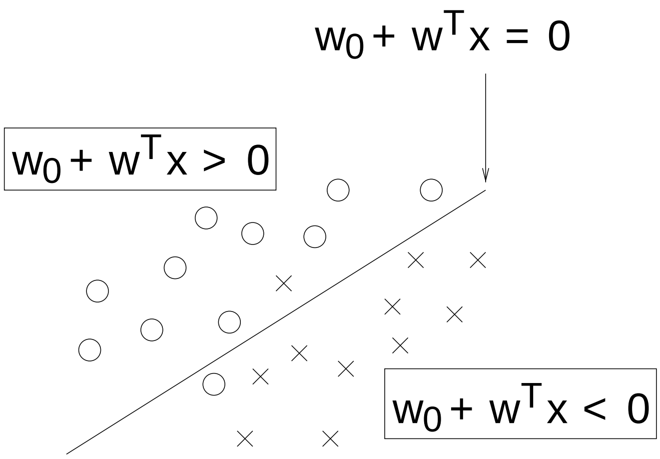


Figure 1.7: Hyper-plane separating two classes. The "crosses" and the "circles" [Lotte, 2008]

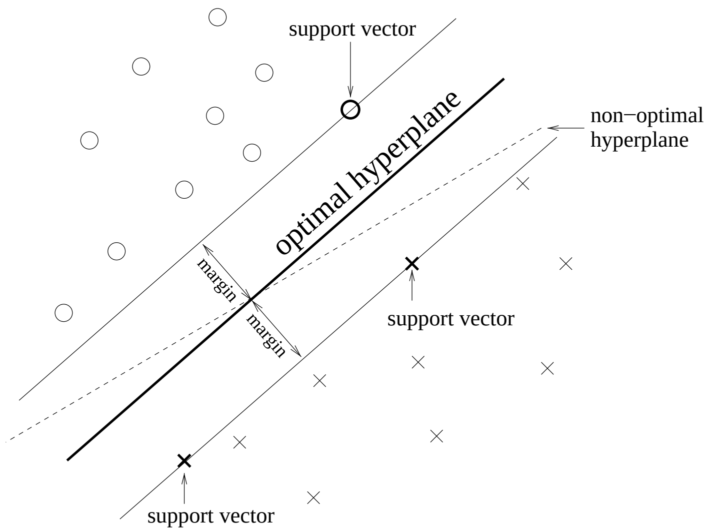


Figure 1.8: Hyper-plane computed from an LDA separating two classes. The "crosses" and the "circles", respecting SVM constraint [Lotte, 2008]

level of expertise is generally necessary to set a relevant network architecture. In recent years, deep neural networks have enable great advances in various machine learning application: image recognition [Krizhevsky et al., 2012], speech recognition [Hinton et al., 2012], or artificial intelligence for the game of go [Silver et al., 2016].

Hidden Markov Models: A specific form of stochastic automaton is used to provide a likelihood of obtaining a feature vector. Each state of the automaton can model the probability of observing a given feature vector. For BCIs, these probabilities usually are Gaussian Mixture Models. Hidden Markov Models are efficient to represent the temporal behavior of a signal [Lotte, 2008].

k-nearest neighbors: For each feature vector, the attributed class is the most represented among the k nearest signal in the training set. Such an approach can be efficient with a lot of data, the main limitation being the need for a relevant distance [Lotte, 2008].

For more information about the classifiers used in BCI, the interested reader is invited to refer to the review on the subject in [Lotte, 2008].

1.3 Brain Patterns used in BCIs

The signal processing tools present in the previous section are very general, and are used in various application fields. However, for BCIs, not any kind of mental state can be classified by these tools. To be detected by an EEG, a cerebral activity pattern must be located close to the surface of the brain, mobilize a high number of neurons, and present a consistent dipole orientation, so that the resulting potential is detectable.

This section presents a summary of brain patterns most commonly used for BCIs.

1.3.1 Slow Cortical Potential

The slow cortical Potential (SCP) is the global potential of the surface of the head. One electrode could be enough to measure it, but it is better to average over more electrodes to reduce the noise. By its very definition, SCP varies very slowly. Historically, it was one of the first BCI paradigms. With hours of training (one hour per day during months), users can learn to control their slow cortical potential. A BCI system can use it to discriminate between 2 classes "high" and "low" cortical potential, and thus sending commands with it [Birbaumer et al., 2003].

By selecting a letter of the alphabet with successive dichotomy, one can write a text using this paradigm, even when all motor control is lost. In a study by [Birbaumer et al., 2003] over 5 patients suffering from ALS, 1 succeeded at writing with an accuracy of 75% in a few weeks. It took months to 2 others, and the last 2 didn't succeed at controlling their slow cortical potential (SCP). The use of preselected words was not appreciated by the patients, even if it improves the writing speed, because it made them feel less free [Birbaumer et al., 2003].

This paradigm is not often used anymore, since others show better performances (see below).

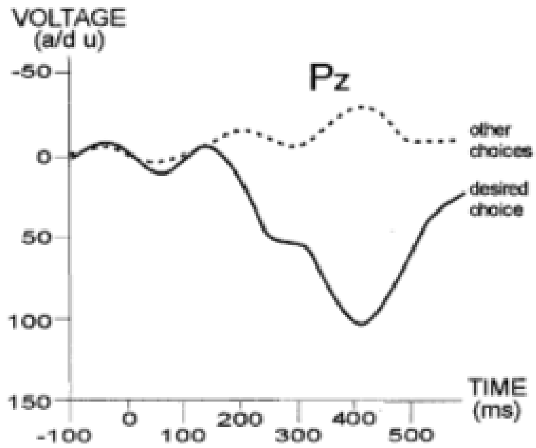


Figure 1.9: P300 potential, averaged over several presentations of the desired stimulus, and of other stimulus [Wolpaw et al., 2002]

1.3.2 P300

When a subject waits for a rare event and the event occurs, it triggers a detectable potential roughly 300ms later, mainly in parietal areas [Wolpaw et al., 2002] (see Figure 1.9).

For a P300 to occur, the user needs to know that the event is going to happen, but to ignore the timing of occurrence. By contrast, unexpected events can trigger different responses (see ErrP below), while an event occurring at a known (preceded by a countdown for example) timing will trigger little reaction. An intuitive and natural example of when a P300 occurs is when a driver waits for a red light to turn green. The driver knows that the event will occur, but does not known the exact time. Additionally, he/she is paying attention to this event. The P300 is larger if the eliciting event is less probable (in the sense of occurrence frequency). It generally needs averaging over several trials to be detectable with a satisfying accuracy.

Considering a matrix of symbols (e.g. letters), the user focuses on the one he/she wants to select. The symbols successively flash, and the user have to count the number of flashes of the desired symbol. (Counting is not necessary to the eliciting of a P300 potential, but it helps keeping the attention of the user.) When a flash occurs, a P300 is triggered in the cerebral activity. This P300 can be detected and associated to the symbol flashing 300 ms ago [Farwell and Donchin, 1988] (see Figure 1.10).

A training phase has to be used to get the precise characteristics of the P300 depending of the user, like its amplitude and delay, but it can be reduced to a short time (around 1 minute) [Rebsamen et al., 2007].

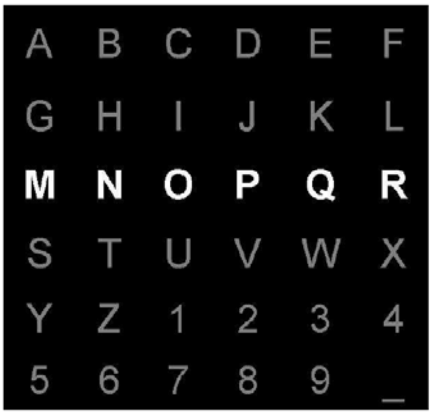


Figure 1.10: Example of a P300-speller interface [Farwell and Donchin, 1988]

1.3.3 Motor Imagery

Some mental tasks block or increase the amplitude of brain rhythms. The decrease or the increase in a rhythm can be detected and used to spot the mental task. Such an increase in brain rhythm is called event-related synchronization (ERS), while a decrease is an event-related desynchronization (ERD) [Pfurtscheller and Da Silva, 1999].

During the preparation of a movement, a strong decrease can be observed in the beta activity over the sensorimotor areas [Wolpaw et al., 2002]. Mu rhythm is also affected by voluntary movement processing. These changes in the brain activity, that occur as a result of motor imagery task, can be detected. It has also been shown that an imaginary movement can produce a weakened version of the same signal characteristic, even with the movement of a lost limb [Blankertz et al., 2006a]. Depending on the member accomplishing the (eventually imagery) movement, the brain area in which the signal is affected differs. For example, a left hand movement creates an activity in the right motor area, while a right hand movement depends on the left area activity [Blankertz et al., 2006a] (see Figure 1.11).

Recognizing a type of motor imagery from the EEG signal requires information on the localization of the detected patterns. EEG signal is limited in spatial resolution, and measures mainly surface activity. Thus, most of the time, only two or three classes of movements are used. These classes correspond to the most easily distinguishable motor area, such as: left or right hand, left or right foot, and tongue moves [Leuthardt et al., 2006]. Another limitation of motor imagery is that it requires an active mental process that not everybody is able to perform [Guger et al., 2003].

An alternative use of the markers used for motor imagery (mu and beta rhythm) is to combine them in order to build a continuous command. After a long training, users can learn to control them [Vaughan et al., 2006].

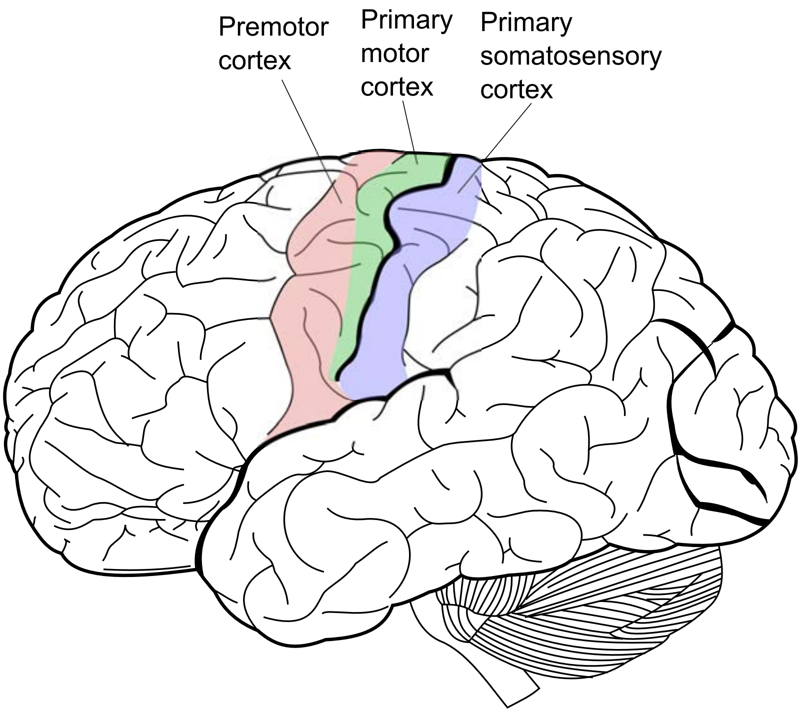


Figure 1.11: Motor and sensorimotor cortical areas

1.3.4 Steady-State Evoked Potentials

When the subject is confronted with a periodic stimulus, the corresponding frequency, along with its harmonics present an increase in power in the cerebral activity of associated sensitive area (sensorimotor, visual or auditory) [Müller-Putz et al., 2006]. This cerebral response is called Steady-State Evoked Potential (SSEP), and can be classified according to the sensorial channel. The cerebral response occurs in a brain area depending on this sensorial channel. It is visible in visual areas for SSVEP, auditory areas for SSAEP, and sensorimotor areas for SSSEP.

While visual SSEP (SSVEP) are the most commonly used for BCIs (please see Section 1.4.4.E for more details on this particular SSEP and its applications to BCIs), Somatosensory (SSSEP) and auditory (SSAEP) evoked potentials have also been shown usable for BCIs [Müller-Putz et al., 2006, Lalor et al., 2005].

For SSSEP use, vibrating stimuli are sent to fingers. The user is asked to focus on the finger of interests. For SSAEP, auditory stimuli are sent to either the left or right ear, and the user focuses on the side of interest. These alternatives enable less speed and precision than SSVEP, and show high variability of performances across users [Müller-Putz et al., 2006] but present the advantage of being usable without any gaze control.

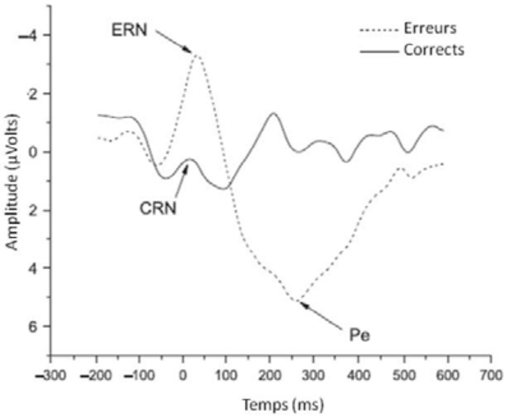


Figure 1.12: Average cerebral answer to trials for correct answers (full line) and incorrect (doted lines) [Perrin, 2012]

1.3.5 Covert Attention

In most ERP-based BCIs (typically P300- or SSVEP- based), the user is asked to look at a target as he is focusing his attention on it. The observed phenomenon is then related to overt attention, meaning that the gaze position matches the visual attention position. By contrast, covert attention describes a shift between the gaze focal point and the attention localization [Treder and Blankertz, 2010, Kelly et al., 2004]. Such a shift can be detected in the EEG signal.

Covert attention is a relatively new mental state for BCIs. Some standard psychology experiments have been reproduced in order to control the direction of visual attention (left or right), and the signal processing tries to recognize this orientation. For now, the classification rate is around 70%, with a high variability between the subjects [Trachel et al., 2013].

Application in adaptive interfaces can be imagined. For example, critical data could be presented where the visual attention is detected. Another approach could be to shift the visual attention of the user toward the right spot before the presentation of the data. Even with a less than perfect recognition, performances on data reading tasks could be improved by such an interactive system.

1.3.6 Error Potentials

When a subject makes a decision (press a button) and immediately realizes he/she has made a mistake, or receives a feedback telling him so, a characteristic potential is triggered in the fronto-central areas (see Figure 1.12), the cingular anterior cortex being suspected to be the origin point of this potential [Perrin, 2012].

Although this error related potential is difficult to detect without averaging, some application could use this phenomenon. It should be noted that there are different kinds of error potential, depending on the condition (through feedback, or by a sudden realization) [Perrin, 2012]. This promising brain pattern has only been discovered relatively recently, and further research might greatly improve its detection.

1.3.7 Other brain patterns used for passive BCI

A few other mental states have been shown to be differentiable using BCIs. It is the case for emotions such as frustration or arousal. Further research could still bring many other useful paradigms for BCI. Such mental states could be used to drive a BCI either in an active way (direct command from the user) or more passively (the user does not try to control those patterns, but the BCI monitor another interaction) [Hjelm, 2003, Lécuyer et al., 2013, Nijholt et al., 2009].

Among those additional brain patterns, the level of alpha band power is particularly informative. When the subject is in a relaxed state, the alpha band of frequency power is stronger, whereas when he/she is well awake, there is almost no alpha rhythm left. At smaller levels, the other band powers also carry a part of information about the concentration level. It is possible to control one's own alpha activity, by trying to relax or to focus, but such a control is relatively imprecise [Hjelm, 2003, Lécuyer et al., 2013]. More classically, some applications could use alpha activity in a passive way, to detect the user's mental state without using it as a direct command.

Finally, the mental workload, a measure of the amount of resources that the brain need to "allocate" to a current task can also be measured by a BCI [George, 2012], allowing to keep track of the difficulty of the task.

1.4 Steady-State Visually Evoked Potential for BCIs

Steady-state visually evoked potentials (SSVEP) are a phenomenon that occurs when the eye is exposed to a periodic stimulation. The resulting cerebral activity can be detected, and exploited for BCIs. The work presented in this PhD focuses on SSVEP-based BCIs. SSVEP is one of the most commonly used brain pattern for BCIs. This section presents the reasons for this success, and the main methods used for SSVEP detection and its use for interaction.

1.4.1 Physiological phenomenon

When the human eye is stimulated by a periodic stimulus, the signal is transmitted by the optic nerve toward the early visual areas (see Figure 1.13). The resulting cerebral response can be detected as a peak of band power in the cerebral activity of early visual areas (in occipital lobe) at the frequency of stimulation. Additionally, a peak of lesser amplitude can also be observed at the harmonics of the stimulation frequency (See figure 1.14).

1.4.2 Electrodes positioning for SSVEP detection

SSVEP is a cerebral pattern that is localized on the visual cortical areas, mainly concentrated on the early visual areas V1 and V2, located in the occipital lobe. Consequently, the detection of

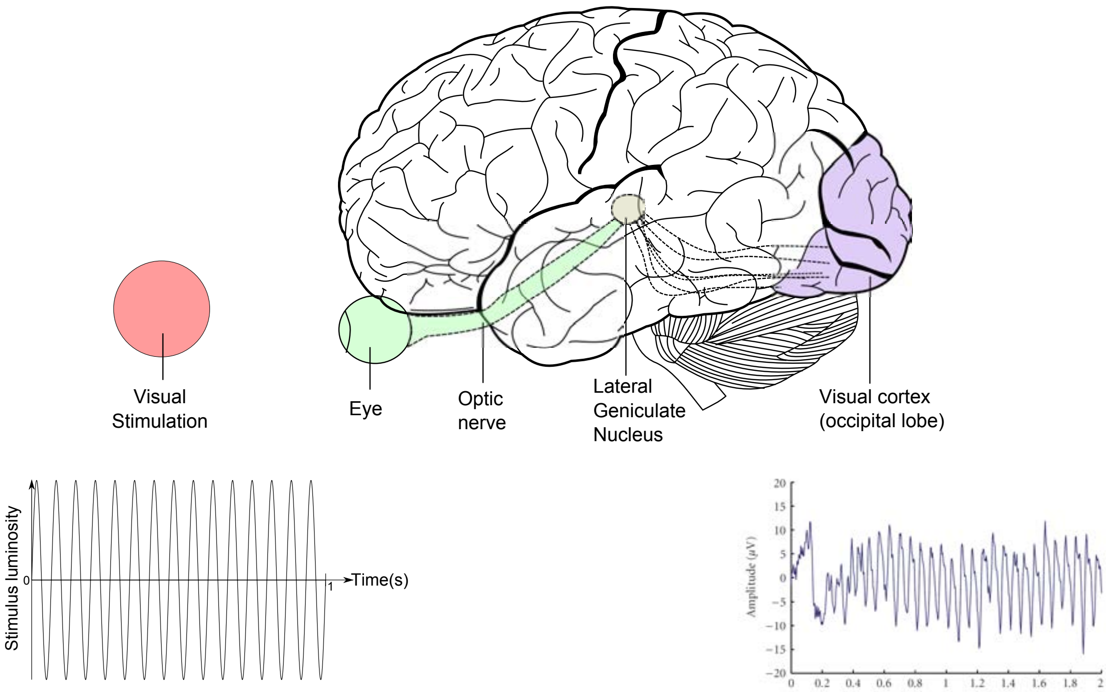


Figure 1.13: A visual stimulation is sent to the eye, with a periodic variation of luminosity, e.g. at 15 Hz (left plot). This signal is perceived and transmitted via the optic nerve, toward the visual cortex (located in the occipital lobe). The response in the resulting cerebral activity presents an enhanced 15 Hz component, called SSVEP (right plot). The cerebral response plot presented here is from [Zhu et al., 2010b]. The rest of this figure is original content.

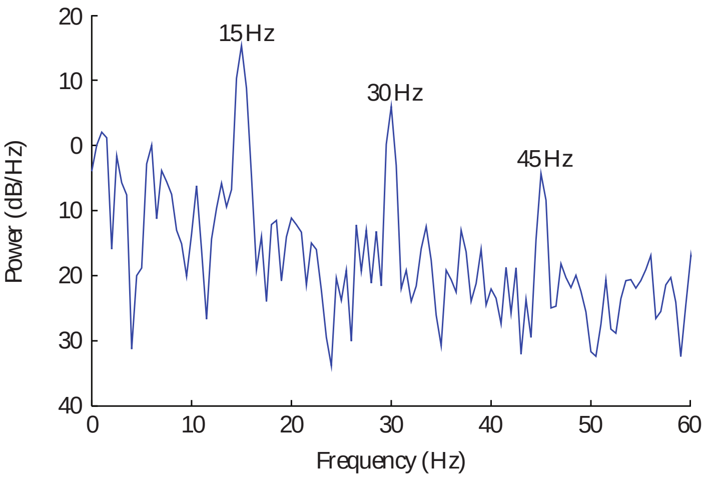


Figure 1.14: Frequency spectrum in visual areas during presentation of a $15Hz$ flickering stimulus [Zhu et al., 2010b]

SSVEP by EEG requires electrodes positioned over the occipital lobe.

According to [Gao et al., 2003], using only sites O1 and O2 is sufficient to provide the necessary information for a functional SSVEP-BCI (see Figure 1.4 for electrodes position reference system). Numerous study use only these sites [Middendorf et al., 2000, Cheng et al., 2002, Müller-Putz et al., 2008, Lee et al., 2010, Brunner et al., 2011, Lalor et al., 2005].

In many cases, at least one additional electrode is positioned at Oz. According to [NG et al., 2012], Oz is the closest to the striate cortex, where neuronal activities that result in SSVEP are anticipated to occur. They also observed that, consistently across subjects, O1, Oz, and O2 are the sites where the SSVEP is strongest. It is also possible to use more electrodes, still focused on visual areas. In [Chen et al., 2014b] and [Chen et al., 2015], for example, electrodes at Pz, PO5, PO3, POz, PO4, and PO6 are used on top of Oz, O1, and O2. In [Cao et al., 2012], electrodes are positioned at POz, P3, P4, Oz, O1 and O2.

To summarize, O1, O2 and Oz are critical locations, while there is no clear consensus over the benefits of placing additional electrodes at other nearby sites.

1.4.3 Signal processing for SSVEP detection

The detection of SSVEP requires several steps of signal processing. In many applications, several flickering stimuli are sent simultaneously, flickering at different frequencies. One of the signal processing goals is then to detect on which target the subject is focusing his/her attention. Historically, the first methods for SSVEP detection were based on the analysis of the cerebral activity spectrum. We first present the main steps of this method. In a second part, we describe an alternative method, based on Canonical Correlation Analysis, which has been shown to be efficient for SSVEP detection.

1.4.3.A Frequency-based signal processing

Since SSVEP is mainly characterized by response in cerebral activity at the frequency of stimulation, a natural approach is to extract the band power in the frequency of interest, and to use them as a classifier feature. The method described below details one possible signal processing chain that falls into this family. This signal processing is used in [Müller-Putz et al., 2005] and in [Legény et al., 2013].

Spatial filter: For each frequency, channels are reconstructed by linear combination of the electrodes output using two fourth-order CSP (common spatial pattern). The role of this step is to choose the channel linear combination that optimizes the detection of specific frequencies, based on training data registered previously.

Frequency power level: The power levels at frequencies of the flickering targets and their first harmonics are calculated by filtering the EEG signal in a narrow band around the target frequency. The signal is divided into epochs of 1 s each, overlapping at 90% (an epoch is created every 100 ms). The signal in each epoch is then squared, averaged over the whole epoch, and a natural logarithm of the resulting value is computed. The final values are used as feature for the classifier [Legény et al., 2013].

Multi-class LDA classifier: LDA is a linear classifier that is calibrated on a training set, and separates two classes. Several LDA classifiers can be combined in order to differentiate more SSVEP targets.

In [Legény et al., 2013], one LDA classifier is trained for each frequency, discriminating between "the frequency is present" and "the frequency is not present". This method enables a self-paced use of the BCI, since it is possible for the resulting multi-class classifier to detect an absence of target.

Alternatively, when building a self-paced interface is not an objective, LDA classifiers can also be combined in a one versus all strategy, where each frequency is tested against the others, and a class is always recognized [Évain et al., 2016].

1.4.3.B Canonical Correlation Analysis:

Using Canonical Correlation Analysis (CCA) as a feature extraction method for SSVEP signal processing has enabled significant improvement of the accuracy, compared to frequency-based approaches [Lin et al., 2006, Bin et al., 2009]. (See 1.2.4 for a description of CCA as a feature extraction method.)

In the case of SSVEP signal processing, a pure sinusoidal signal at the frequency of interest can be used as the reference signal. For each frequency of interest, the correlation between the EEG signal and the reference signal is computed. This single correlation is used as a feature. The recognized target is the one whose frequency of stimulation has the highest correlation with the EEG signal [Lin et al., 2006, Bin et al., 2009]. An advantage of this approach is to get rid of the need for calibration.

Since its first uses in SSVEP-based BCIs [Lin et al., 2006, Bin et al., 2009] CCA became a standard in calibration-free SSVEP-based BCIs. That success can be explained by the very small impact that noise has on its precision, and the comparatively low subject variability it induces. Yet there are some uses of CCA with a calibration method [Wang et al., 2010a, Zhang et al., 2014] in order to improve performance or to adapt to user's reactions in a self-paced context.

1.4.4 Interaction design for SSVEP-based BCIs

The typical SSVEP-based BCI enable the selection of targets or commands. One flickering visual stimulus is associated to each target, and sent to the user. The frequency of stimulation is different for each target. The user focuses on the stimulus associated to the target he wants to select. This attention enhances a SSVEP that can be detected. The frequency of this SSEP can be associated to the stimulus, and the target is selected [Lalor et al., 2005].

There are few BCI-illiterates (people who cannot manage to obtain reasonable precision) for SSVEP [Middendorf et al., 2000]. SSVEP is resilient to artifact as blink movement, partly because eyes are at the opposite side of the head, and because they are confined mostly in lower frequencies

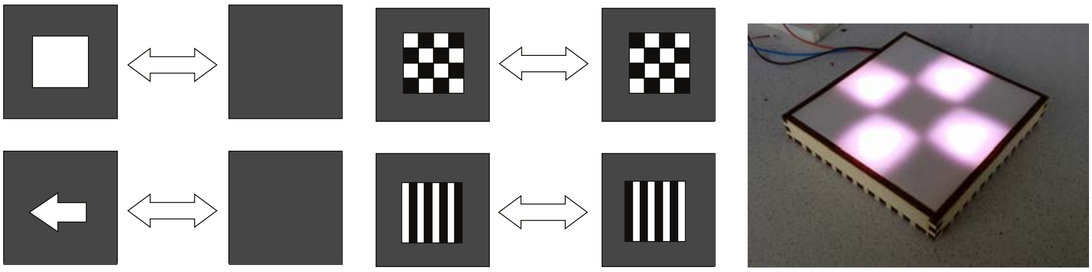


Figure 1.15: Left: graphic alternation [Zhu et al., 2010b]. Middle: Pattern reversal[Zhu et al., 2010b], right: LED tiles [8]

[Lalor et al., 2005]. But the main asset of SSVEP-based BCIs lies in their high speed and precision of selection, compared to other BCIs. In [Chen et al., 2014b], Chen et al. describe an SSVEP-based BCI reaching 172 bits/minute on average across subjects (corresponding to 40 characters/minute spelling). Later, the same research team reported reaching 5.32 bits per second (60 characters per minute) [Chen et al., 2015]. To the best of our knowledge, this is the highest information transfer rate ever achieved by a BCI (ever meaning here at least up until 2016, date of writing of this document).

The performances of an SSVEP-based BCI vary depending on the interaction design. In particular, characteristics of the stimulation can have a major influence on SSVEP response.

1.4.4.A Types of stimulation

Several types of stimulations for SSVEP-based BCIs can be distinguished into three categories, as underlined by the review [Zhu et al., 2010b] (see Figure 1.15):

- Graphic alternation: A simple form (circle, rectangle) is successfully displayed, hidden, or replaced by a second form at the desired frequency.
- Flickering checkerboards: A checkerboard is displayed. Then all black cells become white, and reciprocally, at the desired frequency. This can be considered as a particular case of graphic alternation, with the two alternative forms being two complementary checkerboards.
- Light devices: LEDs can be set to flicker at the desired frequency.

When available, LED-based stimulation presents some advantages: the luminosity and frequency can be controlled more precisely than with a computer screen. However, they require an additional hardware. By contrast, computer screens are limited by their refresh rate. With flickering patterns on a computer screen, the luminosity variations are not controlled as for a LED. Instead, the color of pixels is modified. As a result, LED-based stimulation generally enables a higher bit rate, higher stimulation frequencies, and more variability in the range of frequency [Zhu et al., 2010b].

When computer screens are preferred to LEDs as the stimulation device, on average, graphic alternation on a screen results in a higher bit rate than flickering checkerboards [Zhu et al., 2010b].

[8] Source: https://blog.utwente.nl/designlab/

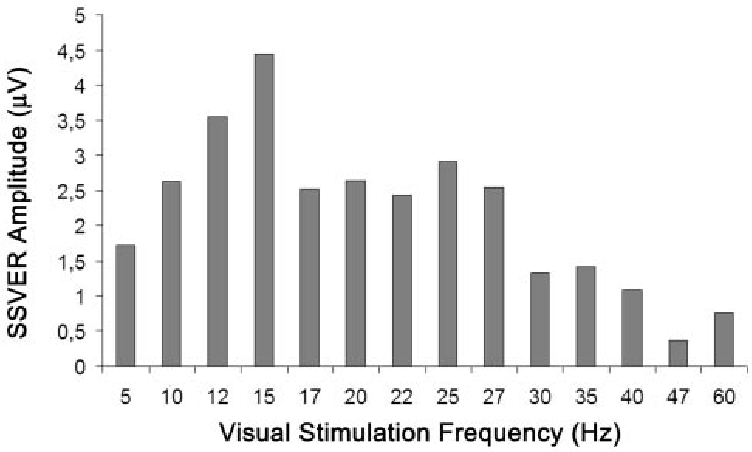


Figure 1.16: Amplitude of SSVEP response, as a function of the stimulation frequency [Pastor et al., 2003]

1.4.4.B Stimulation frequency

SSVEP amplitude depends on the stimulation frequency. It is only visible after a certain frequency, and becomes less strong at higher frequency, probably because of persistence of vision.

[Zhu et al., 2010b] observe that stimulation frequencies in the low (1-12 Hz) and medium (12-30 Hz) frequency bands have been more often applied than those in the higher frequency band (30-60 Hz). Other studies [Herrmann, 2001, Pastor et al., 2003] found that frequencies below 10 Hz and above 25 Hz perform significantly less well than intermediary ones (see Figure 1.16).

1.4.4.C Stimulus color

Little is known about the influence of stimulus color for SSVEP. Several studies have tackled this question, but the comparison of their result is made difficult by differences in protocols. The first major differences come from the stimulation device. Colors of LED stimulation may not be equivalent to computer screen colors, for SSVEP perspectives, as the resulting color spectrum are different. Additionally, luminosity contrast is known to be strongly related to SSVEP amplitude. Not controlling this factor when using a computer screen may induce a bias in favor a colors with a higher contrast (lighter color on a black background). The second difficulty, in studying the influence of colors, is that they are not independent of frequency. The frequency of maximal response depends on the color of stimulation [Tello et al., 2015].

However, by cross-checking the findings of several studies [Cao et al., 2012, Aljshamee et al., 2014, Singla et al., 2013, Tello et al., 2015], recurring observations can be made. The first observation is that white stimulation consistently leads to higher SSVEP amplitudes than colored stimulation. The second most consistent observation across studies is that red stimulation usually leads to higher amplitude than blue or green.

1.4.4.D Stimulus size and inter-stimulus distance

When several stimuli are presented simultaneously, they all stimulate the optic nerve, and they compete for representation in the early visual areas. The level of response depends on the initial excitation (directly linked to size and luminosity of the stimulus), and can be modulated by attention. Under such conditions, it is expected that SSVEP classification would be more difficult when stimuli are small, or when several stimuli are close to each other.

Kian *et al*˙ explored the influence of stimulus size [NG et al., 2012], and concluded that optimal performance is attained when targets are distant of at least 0.087 rad apart, with a size that subtends 0.035 rad of visual angle. Below this size, smaller targets lead to lower SSVEP detection rate.

1.4.4.E Number of SSVEP targets

When using a computer screen as stimulation device, the refresh rate limits the available frequencies, as the only flickering stimulation available should have a frequency of the form $\frac{f}{n}$, with f the refresh rate of the screen, and n an integer. Additionally, in order to remain in the range of frequency that triggers a stronger SSVEP response for humans, using frequencies that are too low or too high is not advised. As a consequence, suitable frequencies on a standard 60 Hz screen are: 6, 6.67, 7.5, 8.57, 10, 12, 15, and 20 Hz. Considering that SSVEP response also occurs at the harmonics of the frequency of stimulation, using both a frequency and one of its harmonics as stimulation frequency could lead to more recognition error. Taking into account all these constraint, most SSVEP-based BCIs are limited to less than 5 targets. However, several techniques enable to bypass these shortcomings, and successful SSVEP-based BCIs use more targets:

- LED: When LED-based stimulation is used, the number of SSVEP targets is not a strong constraint, as each LED can flicker at a different frequency.
- High-frequency screens: While most commercial screens have a refresh rate of 60Hz, using a screen with 120 or 144Hz refresh rate increases the number of available frequencies.
- CCA: It has been shown that using the information contained in harmonics was neither necessary nor clearly useful when using a CCA-based classification. Thus, using targets with flickering stimulation being multiples from each other should not remain a problem when using CCA [Lin et al., 2006, Bin et al., 2009].
- Phase coding: By encoding target information in the phase of the periodic signal, it is possible to use several targets flickering at the same frequency with different phase. This phase is maintained in the SSVEP response, and can thus be detected, provided that timing information and delays are carefully controlled [Manyakov et al., 2013, Chen et al., 2014b]. Based on this principle, SSVEP-based BCIs using up to 40 targets have been proposed [Manyakov et al., 2013].

1.5 Conclusion

This chapter presented the fundamental principles of Brain-Computer Interfaces. It described the common steps of cerebral signal measurement and preprocessing: the types of sensors for activity measurement, the main preprocessing methods and types of features, and the classification methods. Then, a focus was made on SSVEP-based BCI, with a detailed discussion on the main questions related to SSVEP detection. In particular, we saw that SSVEP stimulation characteristics can have a strong influence on the interface accuracy, and although several studies attempted to understand this influence, there is still a lot uncertainty toward the influence of several characteristics. Finally, this chapter covered the most common cerebral patterns used for BCIs, describing what kind of cerebral activity was measured for each of them, and how this activity can be used for BCIs.

Chapter 2

A Human-Computer Interaction perspective on Brain-Computer interfaces

Research on Brain-Computer Interfaces so far usually focused on processing and classifying cerebral signals with the objective of improving the speed and precision of the interface. Progresses in these areas have allowed us to diversify the applications of BCIs. It is now time to work on improving the interactions that occur through these interfaces.

In this chapter, we discuss the relationship between BCIs and Human-Computer Interaction (HCI), and we explore how our knowledge in HCI can be applied to BCIs. This chapter is organized as follows: it starts with a general presentation of the main concepts of Human-Computer Interaction. We then study the most important properties of BCIs in terms of these concepts. A discussion follows, on the problem of choosing the right cerebral pattern for a given interaction and usage context. The most promising recent paradigms of BCI interaction are then described. Finally, the main application fields of BCIs are presented.

2.1 A brief introduction to Human-Computer Interaction

Human-Computer Interaction (HCI) is a discipline concerned with the design, evaluation, and implementation of interactive systems for Human use [Hewett et al., 1992]. In this section, we simply define a few important concepts from this vast research field. Interested readers can refer to [Jacko, 2012] for a much more complete introduction to the subject.

2.1.1 Interactive systems, Interfaces, and Interaction

An *interactive system* is a system whose operations depend on an unpredictable input from an external environment that it does not control [Goldin et al., 2006]. The *interface* is the set of hardware and software that allow a person to operate, control, and supervise an interactive computer system. *Interaction* is the phenomenon that occurs between the user and the system. This phenomenon is the object of study in HCI, which aims to understand it (i.e. observe it, describe it, explain it) and improve it.

2.1.2 Elementary tasks and interaction techniques

The operation and control of an interactive computer system are founded on a set of *elementary tasks* that the user can achieve. Each task can be performed by means of a set of various *interaction techniques*. The elementary tasks are the smallest units of operation possible in a given context. An interaction technique is a certain combination of hardware and software mechanisms that accomplishes a given task. The task represents part of the objective, whereas the technique represents part of the means with which that objective is achieved.

Elementary interaction tasks vary in nature according to the domain of application. For example, Foley et al. list six elementary tasks for graphical interactions: selecting, positioning, orienting, tracing, quantifying, and text input [Foley et al., 1984]. Touching an object on a touch screen or indirectly specifying it by clicking on it with a mouse are two examples of interaction techniques for selecting that object. Operating a physical potentiometer, a virtual potentiometer with the mouse, or entering text are three possible techniques for specifying a numerical value. Voice recognition or keyboard inputs are two possible techniques for entering text.

2.1.3 Theory of action and feedback

The *theory of action* outlined by Norman deconstructs the act of performing a task into seven stages: establishing the goal, forming the intention, specifying the action sequence, executing the action, perceiving the system state, interpreting the state, and evaluating the system state with respect to the goals and intentions [Norman and Draper, 1986] (see Figure 2.1.3). These seven stages are not all necessarily present, and can occur in another order, but this decomposition is nonetheless useful for analyzing and designing interactive systems.

The user's mental picture of certain concepts might be very different from the way that these concepts are implemented by the system. For Norman, there are two gulfs separating the user's conceptions from the system's conceptions: the gulf of execution and the gulf of evaluation. The terms *distance of execution* and *distance of evaluation* describe the effort that the user or the system designer must invest in order to cross these gulfs. The adjectives *semantic* and *articulatory* are used to distinguish efforts related to the meaning of user-system exchanges from efforts related to the form that these exchanges take.

The speed and the form of the *feedback* produced by the system largely determine the capacity of users to perceive, interpret and evaluate their state changes, and thus affect the evaluation distance. For instance, prompt feedback builds a continuous representation of the state of the system, and the effect of actions as they are performed. Prompt feedback also contributes to the sensation that the user is acting directly upon the objects of interest, allowing the user to feel engaged in the task.

Interaction techniques are usually designed around the requirements of the tasks, leveraging as fully as possible the users' cognitive, motor and perceptual skills to reduce the execution and

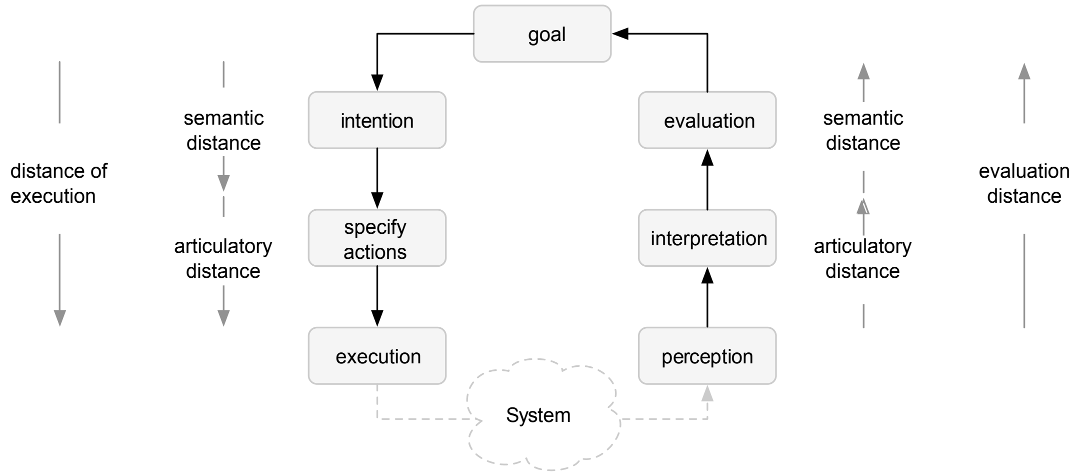


Figure 2.1: The seven stages of user activity when performing a task and the corresponding distances [Norman and Draper, 1986]

evaluation distances. But no matter how carefully the technique is designed, an interaction technique that is perfectly adapted for one task in a certain context may prove unsuitable in others. Text input by voice recognition is undoubtedly preferable to using a keyboard while driving, for example. The question of whether a certain interaction technique is suitable for a certain task in a certain context is the question of *usability*.

2.1.4 Usability

Usability is defined by the ISO 9241 standard as "*the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use*" [ISO, 1998]. *Effectiveness* refers to the capacity of attaining specified goals, *efficiency* describes the resources spent in order to achieve these goals, and *satisfaction* describes how the user perceives the process.

Usability may be evaluated using a variety of different criteria. One might wish that inexperienced users find a system easy to use. Relevant indicators in this case might for example include the percentage of tasks successfully performed on the first attempt, the amount of time required to do so, and the proportion of deliberately performed actions. Alternatively, one might wish a system that is robust to user mistakes. In this case, the percentage of errors corrected by the system, the time taken to do so, and the user's appraisal of the corrections made are of interest. Another possible choice is a system that is easy to learn. We would then look at the number of functions acquired, the time taken to do so, and the user's appraisal of the training process. The choice of usability criteria naturally depends on the domain of application.

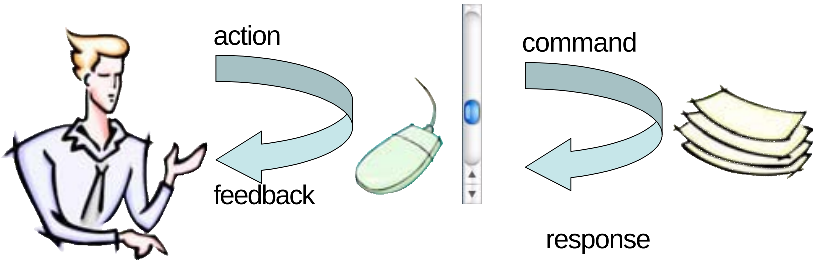


Figure 2.2: Instrumental interaction, as presented in [Beaudouin-Lafon, 2006]: the instrument (center) mediates interaction between the user (left) and the document (right)

2.2 Characteristics of BCIs, from an HCI perspective

Research on Brain-Computer Interfaces usually focuses on lowering the error rate of BCI systems and EEG classification schemes [Plass-Oude Bos et al., 2011]. But several researchers have pointed out that "for general acceptance of this technology, usability and user experience will need to be taken into account when designing (BCI) systems" [Bos et al., 2010] In order to improve the interaction based on BCIs, concepts of HCI should be applied to BCI research. In this section, we present an HCI-oriented approach to BCIs properties. First, a focus is made on the various levels of feedback for BCIs. A presentation of the main characteristics of BCIs follows. Then, we summarize the influence of the type of BCIs on various interaction criteria. Finally, we discuss the resulting choice of BCI, depending on the interaction task to achieve.

2.2.1 Feedback for BCIs

In order to interact efficiently with a machine, the user needs to get some information in return of his/her commands, in order to know how the system reacts. This information can be decomposed in two types (see Figure 2.2):

The **feedback** comes from the interface itself. The delay between the command by the user and the feedback must be extremely short in order to give the user an immediate awareness that his command has been taken into account. For example, when a user moves a mouse, he can almost instantly see the cursor moving on the screen. This feedback gives him confirmation that his commands have been taken into account, as well as the new cursor position.

The **response** of the system may take a longer time to arrive. The actions of the user influence the system, which continuously gives back information about its current state.

In the field of BCIs, the terms feedback and responses are often used indifferently, since most of the tasks achieved so far are very low-level, leading to immediate and intuitive responses. However, as the task at hand become more complex, considering the choice of feedback and responses becomes essential. The user may highly benefit from feedback at different levels in the signal processing chain (see Figure 2.3).

Theoretically, a feedback may come directly **from the acquisition block**. In practice, it is unlikely that this feedback could give useful information to the user, since its visualization is

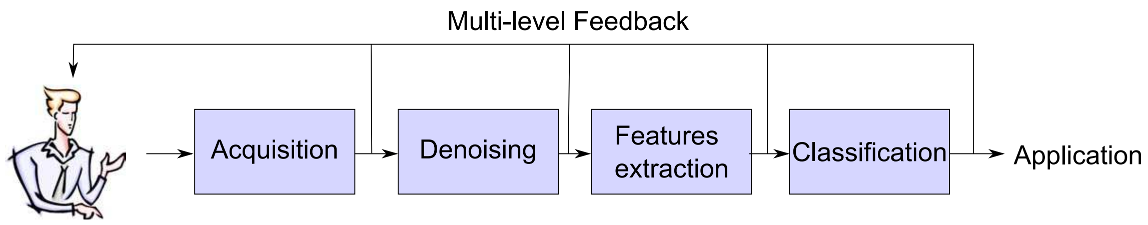


Figure 2.3: Multi-level feedback for a BCI

complex and noisy.

A higher level feedback can be given to the user **after the denoising of the signal**. The visualization of the signal at this stage is quite complex, due to the number of channels and the nature of the signal. However, we can imagine that with adequate training, an experienced user may find this feedback useful, especially to detect electrodes fixation problems at an early stage.

In order to make the feedback more easily interpretable, a **features visualization** can be send to the user. This low-level information is rich, synthetic, and fast to compute. The user can train himself to produce the adequate features, in order to send a clear command. The relevance of such a feedback highly depend on the level of abstraction in the features vector which can present either an high number of low-level features, or a reduced number of high-level ones.

Finally, **the result of the classification** can be send to the user directly, at the same time as it is sent to the application that uses it. This way, the user is aware of the command recognized by the interface as soon as possible, even if the application may take some time to use it. The user can anticipate the next step of interaction.

2.2.2 Properties of BCIs from the perspective of HCI

Brain-computer interfaces measure cerebral activity signals, filter these signals, extract features from them, then classify the vector of features thus obtained. The class that is determined at the end of this chain of processes is then used to activate or configure an interaction technique. The characteristics of the interaction controlled by BCI vary greatly depending on the chosen cerebral pattern. Certain interfaces use potentials triggered by external stimuli, such as P300 or SSVEP. These interfaces often achieve relatively high information transfer rates, but a long exposure to the flickering can be tiring. Conversely, cerebral patterns such as SCP or sensorimotor rhythms can be controlled by experienced users without external stimulation (see Chapter 1 for a more complete review of the most commonly used cerebral patterns with BCIs).

Despite differences in their individual characteristics, there is a set of common properties shared by all BCIs in general, which may be compared with those of more standard interfaces. For example, the information transfer rate of a BCI is always significantly lower than the expected transfer rate achievable with a keyboard or a mouse. We will now present the most important of these aspects, common to all BCIs.

Latency is one of the more obvious properties of BCIs. The latency of most current interfaces is on the order of a few seconds. Indeed, the features of the signals must be measured over a minimum period that is sufficiently large that the cerebral patterns can be classified with reasonable levels of accuracy. For example, at a measured neuronal activation frequency of $5Hz$ (i.e. the period of the signal is $200ms$), we cannot reasonably expect to obtain acceptable levels of precision in less than a second. The acceptability of latency in an interactive system depends on the sensory perception threshold of the user. A latency period of one second very distinctly exceeds the thresholds of auditory and visual perception. Fortunately, in the case of BCIs, the effective threshold is somewhat higher, as our perception of point in time at which a specific mental state was produced is less precise than our perception of the point in time at which a button was pressed, for example. For comparison, interactive graphical systems controlled via mouse and keyboard typically have latencies of less than 100 ms between the physical action performed by the user and the display of updated information on the screen.

The **precision** of a BCI when identifying a command is also relatively low. Errors can arise from the user, who may not be able to adequately produce the required mental state, or, as is more commonly assumed, from the signal processing and classification steps. Most of BCI research until now has focused on improving this precision, and significant progress has been made. Despite this progress, the precision still remains strongly user-dependent, and for certain kinds of cerebral patterns, the classification rate is still fairly low. For a BCI, a classification rate of 90% is considered to be good, whereas in normal conditions virtually every action performed with a mouse or keyboard is accurately registered by the peripheral.

The **number of commands** that may be accessed via a BCI is limited. Even for experienced users, the current levels of classifier precision do not generally allow for large numbers of classes without a drop in the detection rate. However, for some applications, having only 3 or 4 available commands does not necessarily represent a limitation. Note also that the P300 method makes it possible to choose from a large number of commands by successively selecting subsets without requiring the user to explicitly alter their mental state for each command.

The **information transfer rate** between the user and the computer through the BCI, largely limited by the properties of BCIs listed above, only exceeds 100 bpm in very rare cases [Donchin et al., 2000, Middendorf et al., 2000]. For comparison, the information transfer rate that an experienced user can achieve with a keyboard is of the order of 900 bpm (assuming a typing speed of 300 characters per minute and a Shannon entropy of the order of 3 bits/character). With a mouse, the information transfer rate is of the order of a few hundred bits per minute [MacKenzie and Buxton, 1992].

There is significant **variability in the performance** between different BCI users. Some users manage to operate BCIs much more effectively than the average, whereas others, sometimes called "BCI-illiterate" are completely incapable of using them. Even between users that can use BCIs effectively, large degrees of variability are observed from session to session.

BCIs do not require any **motor activity** from the user, as they take input directly from brain activity. They may therefore generally be used by individuals with motor disabilities. Furthermore, other channels of interaction, such as the hands, remain available for controlling other devices (hybrid approach). It can however be difficult to divide attention between multiple different tasks, even if adapted methods of feedback (e.g. tactile) can help [Leeb et al., 2013a]. Also, muscle activity represents a source of noise in the signal with potentially very high amplitude, which can make it more difficult to interpret. Blinking and jaw contractions are particularly problematic. These artifacts can be corrected using an EOG-based method for "cleaning" the signal [Schlögl et al., 2007]. The compatibility of BCIs with other interfaces such as a keyboard or mouse is currently a subject of research [Mercier-Ganady et al., 2013, Leeb et al., 2013b].

The **distance of execution** (see Section 2.1.3) of BCIs is high, as the mental state associated with the desired task must be produced by the user. With practice, the association between accomplishing a goal and producing an intermediate mental task becomes more intuitive. The distance of execution may be reduced by choosing mental tasks that are similar to the physical tasks, e.g. imagining a movement of the left hand to move a cursor to the left.

The **hardware and software** used by BCIs is relatively complex to install and operate. This complexity may slow the propagation of BCIs in domestic environments (e.g. for video games), and is also the subject of research [Duvinage et al., 2012, Frey et al., 2014].

2.2.3 Influence of the cerebral pattern on BCIs properties

For the conception of interaction techniques based on BCIs, the choice of an adapted cerebral pattern is critical. This choice can be guided by several criteria.

The table 2.1 gives a subjective summary of typical qualities and flaws of BCIs associated to various cerebral patterns. All the data in this table should be taken with great care. They are only orders of magnitude and general behaviors. Future advances in research could strongly improve one or another of these characteristics. Despite the number of criteria considered here, this table is not exhaustive. Depending on the application, other criteria could be more relevant than those presented here. We hope that despite its inherent limitations, this summary could help guide a relevant choice of cerebral pattern, for a given application.

First, we look at the most frequently considered characteristics for BCI evaluation. Traditionally, research on BCI focuses on the improvement of speed and recognition rate of the interface commands [Speier et al., 2016, Vaid et al., 2015]. Together with the number of available commands, these criteria can be summarized into one: the information transfer rate. One of the main observations on this criterion is the pre-dominance of BCIs based on evoked potentials (P300 and SSVEP) over other BCI paradigm. P300 allows distinguishing a large number of commands, while still keeping a good precision, compared to other BCIs. Recognizing a high number of targets with SSVEP is possible, and has been successfully achieved, allowing to reach comparable information transfer rates [Chen et al., 2014b, Nakanishi et al., 2014]. These qualities explain why P300 and SSVEP paradigms have been especially successful for BCI-based spellers conception.

Secondly, we observe criteria concerning the time necessary before the effective use of the BCI. Before a user can effectively control a BCI, several steps are required. The material setup itself takes time [Kouroupetroglou, 2014]. This time strongly depends on the number of electrodes in the setup. As a counterpart, using a higher number of electrodes can bring a better precision, provided that the signal processing is adapted, and the used cerebral pattern benefits from it. Once the material is installed, the classifier has to be calibrated. The typical calibration duration also depends on the recognized cerebral pattern [Lotte, 2011]. Finally, the user himself can improve his/her BCI control skills with training [Guger et al., 2003]. Here again, the required time for a user to optimize his/her BCI control depends on the cerebral pattern.

The third group of criteria relates to the BCI performance variability, in term of precision, depending on the user, and across time, for a single user. Between two BCI use sessions, electrodes can be positioned slightly differently, and the mental state of the user can change. In some cases, this leads to a significant variability in performances [McFarland et al., 2011]. There is a lack of comparative studies on the link between detected cerebral pattern and the accuracy variations. Until further research, we are bound to make simple assumptions. Mastering motor imagery or the slow cortical potential (SCP) are skills that require time in order to be acquired, and their level of accuracy vary strongly from one user to another. By contrast, evoked potentials are more robust cerebral patterns, even if the signal quality varies from one user to another. Concerning the performance variability across time for a single user, we distinguish two types of variations: the difference between two sessions, and the variation across time in a single session. P300 and SSVEP, that strongly depend on the user attention, might be very sensitive to the variations across time in a single session, loosing in precision as the user attention deteriorates with fatigue [Lamti et al., 2014b]. By contrast, motor imagery and SCP depends more strongly on a skill learned on a longer term, making variations of performances for a single subject less important in comparison.

The fourth group of criteria considered here describes the comfort of use. BCIs using sensorial stimulation induce fatigue after a prolonged use [Cao et al., 2013, Lamti et al., 2014a, Nijholt et al., 2009]. Furthermore, some cerebral patterns require a certain concentration of the user in order to be produced. The difficulty of use changes from one method to another, and with the user training [Pichiorri et al., 2011]. Finally, the user comfort depends on the coherence between the intermediary mental task to produce and the associated command. This coherence depends on the application, but some cerebral patterns are more easily associated to an interaction task in an intuitive way. For example, the P300 paradigm allows an intuitive selection task [Krusienski et al., 2008], while associating SCP to any mental task is difficult and relatively unintuitive [Birbaumer et al., 2003].

The last group of criteria that we consider here concerns the universality of the cerebral pattern depending on the conditions. Paraplegia does not affect the brain itself. Since the neuronal activity remains mostly unchanged, all BCI paradigms can be used (even motor imagery, the can be relearned if necessary [Neuper et al., 1999]). In most cases of paraplegia, gaze control is maintained, allowing a normal use of SSVEP. This cerebral pattern is more difficult to detect when the gaze cannot be controlled [Allison et al., 2008], even if visual attention alone can be enough to modulate the observed cerebral pattern [Kelly et al., 2005]. People suffering from ALS (Amyotrophic Lateral Sclerosis is a disorder that involves the death of neurons in sensorimotor areas) leads to more severe limitations. While the degenerescence advances into motor areas, motor imagery becomes more difficult. However, it has been shown that motor imagery can still be used to some extent [Pfurtscheller et al., 1998, Kübler et al., 2005]. ALS patients retain gaze control [Liu et al., 2012]. For other types of locked-in syndrome, eyes movements can be limited [Bauer et al., 1979], making SSVEP less precise. In this case, one might prefer SSSEP, if the associated somato-sensorial areas are still functional. Finally, we consider the compatibility of the use of brain patterns for BCI for a simultaneous use with other BCIs on the one hand, and with standard HCI action (requiring a motor activity) on the other hand.

2.2.4 What signal for which command?

We saw in Section 2.1.2 that any interactive task may be decomposed into elementary tasks. Any interaction device is more or less well-suited for certain basic interaction tasks. For example, a keyboard is generally suitable for spelling, while it is not the most convenient way to select a position. Complementarily, a mouse is handy for pointing and selecting, but less efficient for

Paradigm	**SCP**	**P300**	**SSVEP**	**SSSEP**	**MI**	**ErrP**	**Relax/focus**
				Speed et precision			
Speed (commands/minute)	○	●	●	○	●	○	●
Precision	○	●	●	○	●	○	◑
Number of commands	○	●	◑	○	○	○	○
ITR	○	●	●	◑	◑	○	○
				Time required before the BCI effective use			
Installation speed	●	●	●	●	○	○	●
Calibration speed	●	●	●	○	◑	-	●
User Training speed	○	●	●	●	◑	●	◑
Few electrodes are needed	●	◑	●	◑	○	-	◑
				Variability			
Inter-subject robustness	○	●	●	○	○	-	○
Inter-session robustness	◑	○	●	○	◑	-	-
Intra-session robustness	-	○	○	○	●	-	-
				Comfort of use			
Mental fatigue	○	○	●	○	●	●	◑
Sensorial fatigue	●	○	○	○	●	●	●
Mental task/command coherence	○	●	○	○	◑	●	◑
				Universality			
Paraplegia compatibility	●	●	●	●	●	●	●
ALS compatibility	●	●	○	●	○	●	●
Robustness to other cerebral patterns	●	●	○	○	○	●	○
Usable with standard HCI actions	●	◑	●	◑	○	○	○

Table 2.1: Subjective comparison of characteristics of the main cerebral patterns used for BCIs. For each criterion, 2 or 3 levels may be attributed, note from the lowest to the highest value: ○, ◑, and ●. The limits from one level to another are the following (there is only one threshold for criteria on 2 levels, and 2 threshold for criteria on 3 levels): *Speed:* 7 *commands/minute; precision:* 85% *recognition rate; number of commands: almost always less than* 5, *usually less than* 5, *usually more than* 5*; Information transfer rate:*5 *and* 30 *bits per minute ; installation speed:* 5 *minutes (●if the speed is higher, meaning that the installation time is shorter) ; calibration speed:* 5 *minutes; training speed: several months, several hours, or almost instantaneous respectively* ○*,* ◑*, and* ●*. The remaining criteria are subjective evaluations and are not measured by numerical values. For all criteria, the symbol − denotes a lack of documentation on the corresponding cell.*

spelling. In this section, we describe how BCIs deal with the most common elementary tasks. Depending on the nature of the elementary task, certain cerebral patterns are more suitable than others; choosing appropriately allows the distance of execution to be further reduced.

Text entry involves entering sequences of characters into the system. It is typically achieved with a keyboard. This task was one of the first tasks to be realized using BCIs, usually with the P300 [Lotte, 2008], motor imagery [Blankertz et al., 2006a], or SSVEP [Middendorf et al., 2000]. The performance achieved is generally in the magnitude of 7 characters per minute [Donchin et al., 2000]. The top spelling speed ever claimed for a BCI system is 60 characters/minute, with an SSVEP-based BCI [Chen et al., 2015].

Quantification involves specifying a numerical value between some maximal and minimal thresholds. BCIs have not been used often for this task, but producing motor imagery patterns with specific or high amplitude levels has been used as a challenge in video games [Lécuyer et al., 2013, Hjelm, 2003].

Selection involves choosing one or multiple elements from a set of fixed size (e.g. menus, radio buttons, checkboxes) or of variable size (e.g. 2D or 3D targeting, dropdown boxes, or selectboxes). Typically, selection can be done by moving a cursor to the right position using a mouse, or by entering the coordinates using a (real or simulated) keyboard. The coordinate system and the resolution needed may influence the choice of interaction technique for this task. BCIs provide several options for selection. P300 allows users to select one element from a few dozen, striking a good compromise between speed and precision, e.g. one element from 36 in 7.7 seconds, with an 80% success rate [Donchin et al., 2000]. SSVEP also allows users to perform selection tasks. For example, one target can be selected out of 6 in 1 second, with a precision of 86.7% [Middendorf et al., 2000]. For selection by targeting, it is possible to control two combinations of brain rhythms in the sensorimotor regions to control a cursor and select or reject a target highlighted using a third combination [Vaughan et al., 2006]. This technique however requires a large degree of concentration from the user, and a significant amount of practice, e.g. 5 to 15 hours divided into sessions of 24 minutes over several weeks [McFarland et al., 2008].

Manipulation and **transformation** involve modifying the position, orientation, size or shape of an object. In the case where the set of possible manipulations and transformations is small and discrete, these tasks may be performed with a selection task that chooses the desired modification, and so BCIs can be used (see previous paragraph) [Legény et al., 2013].

Navigation involves changing the viewpoint in a virtual setting, for example by repositioning the camera in a 3D environment, or scrolling through the content of a document. An indirect form of navigation with a BCI can be achieved by specifying high-level commands. For example, the destination can be selected using a P300, and then transmitted to an automatic navigation system [Rebsamen et al., 2007]. Motor imagery can provide more direct and refined navigation with low distances of execution. For example, the user can imagine moving the left or right hand to turn in that direction [Leeb et al., 2007].

In addition to the classical tasks listed above, BCIs can also be used to evaluate the user's mental workload and detect certain emotional states such as wakefulness, pleasure, drowsiness, prevalence, or frustration [Heraz and Frasson, 2007, George, 2012]. Few other systems offer this range of possibilities. The main alternatives comprise motion tracking and physiological sensors that can, for example, measure galvanic skin response or cardiac rhythm.

2.3 Paradigms of interaction for BCIs

Interaction with a BCI can take various forms, depending on the type of application. This section first presents the classical BCI interaction loop, with additional precisions added to the traditional model. In a second part, the main paradigms of interaction for BCIs are presented.

2.3.1 The BCI interaction loop

To integrate BCIs into applications, the exchange of information between each component of the interaction must be clearly defined. The simplest way of thinking of the role of the BCI within the structure of a wider interaction is to see it as an external component, as presented in [Hintermüller et al., 2013]. The signal processing chain is the centerpiece of the BCI. The acquisition system registers a signal derived from brain activity. This signal is then filtered and transmitted to the feature extraction block. The extraction block calculates the values of the desired features. The resulting features vector must then be classified. This function is performed by the classification block, which then sends the results to the application. The application is responsible for associating the detected cerebral patterns with the commands to be performed. We will now extend this conceptual model of a BCI to account for feedback, and to make a distinction between the measured brain activity and conscious thought (see Figure 2.4).

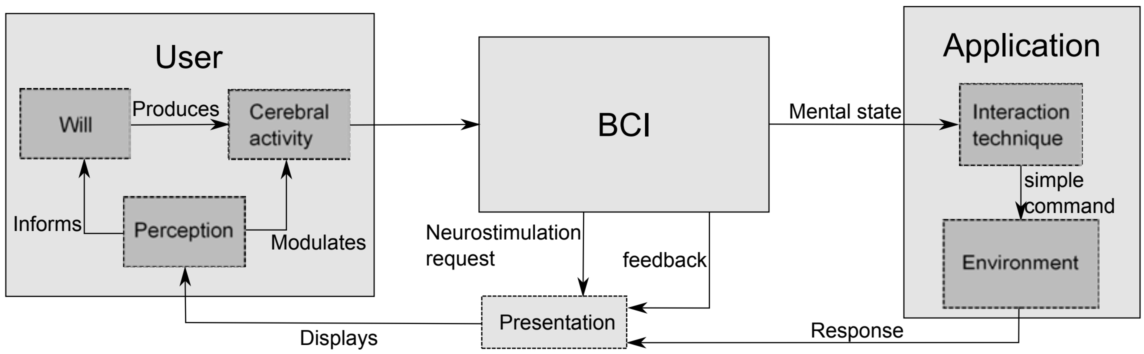


Figure 2.4: **Classical BCI interaction loop**: Components of the BCI interaction loop: The user attempts to produce an intermediate mental state to interact with the machine through the BCI. Given the right stimulation, this mental state will produce a recognizable cerebral pattern for the BCI.

External **stimulation** (generally visual) is necessary for some cerebral patterns. The features that must be extracted to detect the cerebral pattern depend on the timing of this stimulation. Thus, the stimulation block must provide synchronization information to the feature extraction block. This block must also control the visual (or auditory, or tactile) display to the user. Finally, depending on the technique of interaction, it can be useful to adapt the stimulation as a function of the most recently detected cerebral patterns. Hence, the classification block also sends occasional output to the stimulation block.

The **presentation** block is charged of gathering data and sending it to the user. This information includes: neuronal stimulation, feedback, and application-dependent data.

The user's **cognitive and cerebral activity** can be modelized in several ways. The cognitivist approach traditionally proposes a perception-processing-response loop. While this model has a good explanatory power, for neurological processes, it lacks a distinction between conscious and unconscious activity. For BCIs study, a different approach can be proposed: The measurable brain activity can be separated from the user's will. Using a BCI requires the user to produce an intermediate mental state that must be controlled, much like how classical interfaces require an intermediate motor task. Ideally, any stimulation produced by the BCI should only influence the cerebral signal (measured by EEG), without affecting any conscious state, so that the stimulus is not irritating. In practice, if the amplitude of the brain activity is large enough to be picked up by EEG, it also draws the user's attention. Conversely, the application's response must access the user's conscious state, without overly interfering with the EEG signal. To be effective with BCIs, cerebral patterns must be insensitive to this type of noise.

2.3.2 Main paradigms of interaction for BCIs

For purposes of readability, the interaction paradigms outlined in this section are considered at a higher level of abstraction. They are not mutually exclusive, and may be combined with each other.

- **Direct commands** consist in associating each recognized mental state with an explicit command (see Figure 2.4). The role of the BCI is to recognize the mental state from a finite set of classes of mental states (typically two or three classes), and to relay this information to the application. The application systematically relates each possible class to a command that will be executed when the mental state is detected.

  Direct commands are probably the most widely used interaction paradigm for BCIs. Applications allow handicapped persons to operate computers or wheelchairs with direct commands [Rebsamen et al., 2007, Vaughan et al., 2006]. More recently, certain video games used the recognized mental state as the main input [Lalor et al., 2005, Nijholt et al., 2009]. Direct command BCIs are useful in situations where standard interactive devices are ineffective, e.g. because users require the use of their hands for some other task.

- The **hybrid approach** involves using the BCI as a complementary input device in combination with other devices (see Figure 2.5). Two or more interactive devices may be used simultaneously, each sending the information gathered to the application. The BCI is introduced at the same level as the other devices, and provides an additional information channel between the user and the machine. The other devices might for example be a keyboard, a mouse, a joystick [Pfurtscheller et al., 2010] or another BCI [Vaughan et al., 2006]. Hybrid BCIs must be particularly insensitive to noise, because using other devices can generate additional artifacts due to eye movements and muscle contractions. The user must produce a cerebral pattern to control the BCI while simultaneously controlling the other devices as usual.

  Applications with hybrid BCIs use multiple inputs to improve the precision of the interface as a whole, or to specify different parameters of the same command [Fruitet et al., 2011, Li et al., 2010]. More recently, Zander et al. suggest using a gaze-tracking device in combination with a BCI to create a system capable of hands-free targeting and selecting that produces less false positives than traditional gaze-tracking systems, which are based on the eyes' fixation duration [Zander et al., 2010a]. As BCIs are used for increasingly complex tasks, the ability to use multiple inputs more independently could potentially become crucial. For example, in [Leeb et al., 2013b], players can make their character run with a joystick, and jump with a BCI.

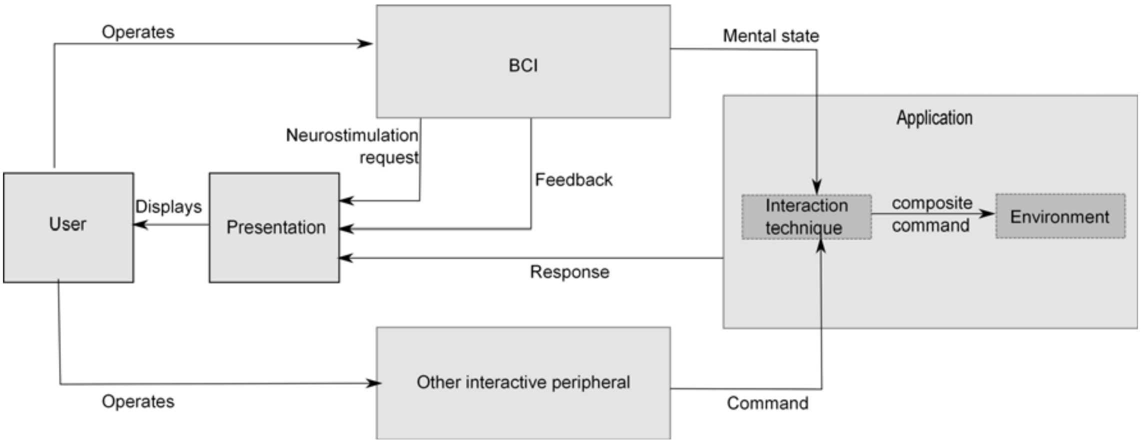


Figure 2.5: **Hybrid interaction** with a BCI: One or more other complementary interaction devices are used simultaneously to accomplish a more complex interactive task

- The **brain switch** involves using a BCI to activate or deactivate another interactive device. BCIs with a single command are sometimes also referred to as *brain switch* BCIs. From the perspective of the interaction, this is a special case of a direct command.

  Two mental states are recognized by the brain switch: a "resting" state and an "action" state. Each time that the "action" state is detected, the other device is toggled. Thus, the brain switch may also be viewed as a special case of hybrid interaction. Indeed, the command to activate or deactivate the device can be transmitted to the application, which controls the device.

  From the user's perspective, it is not necessary to concentrate on the BCI, except when using the specific activation command. When the user does this, he or she performs an intermediate mental task to produce the brain activity associated with the "action" command.

  The need for a way of activating and deactivating BCIs has been raised by Scherer et al. [Scherer et al., 2007]. They suggested using the heartbeat to do this, and showed that with an appropriate amount of training this input can be used as a switch. However, cardiac rhythms can be strongly affected by other phenomena. A controllable cerebral pattern might theoretically prove more reliable for this type of command [George, 2012].

  The question of which patterns are best-suited to brain switch applications remains open. The "action" mental state must be detectable with a good level of precision and a low false positive rate (false positives lead to unwanted activations). On the other hand, since this command is only rarely used (typically once at the beginning of a session to activate the interface, and once at the end to deactivate it), it is acceptable for the activation period to be relatively high, with a large delay (approx. 30 seconds). Conservative approaches with classical cerebral patterns (SSVEP, P300) are good candidates for this interaction paradigm.

- The **passive BCI** approach involves detecting a mental state for purposes other than direct control (see Figure 2.7). The recognized cerebral patterns are sent to the application, which can use the user's mental state as an information parameter for adapting the main interaction

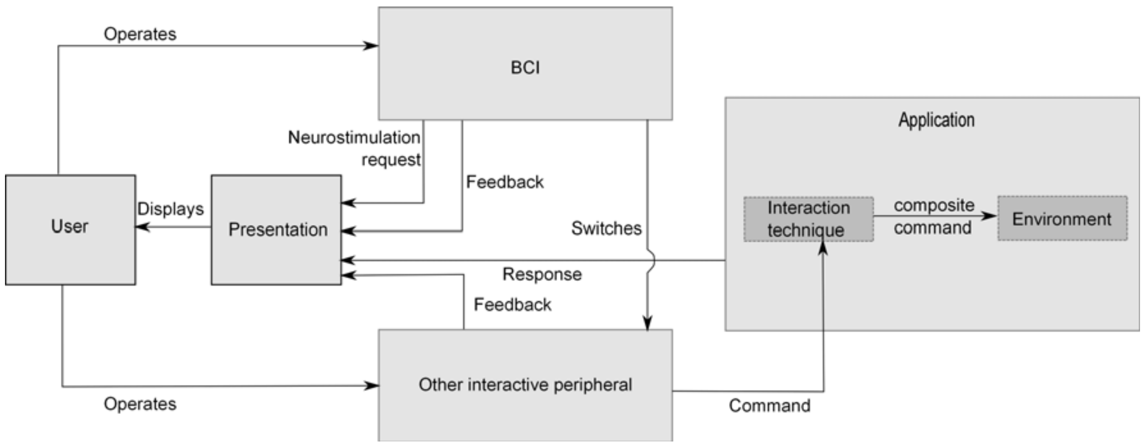


Figure 2.6: **Brain Switch**: The BCI is used to turn ON or OFF another interaction device, possibly another BCI. The same effect can be obtained in a more classical way if the application uses the BCI input to control the other interaction device.

technique [George, 2012]. From the user's perspective, there is no need to consciously control the cerebral patterns produced. Users can concentrate on their primary task rather than focusing attention on the BCI.

The "passive" approach has been used with other physiological markers such as the galvanic skin response [Allanson and Fairclough, 2004] or gaze-tracking [Hyrskykari, 2006]. Cutrell and Tan suggested using BCIs for implicit interactions [Cutrell and Tan, 2008]. Since then, BCIs have been used to detect the level of engagement in a task, the user's mood, certain emotions, error recognition, relaxation and mental workload [Zander and Kothe, 2011, George, 2012]. Detecting these kinds of mental state can be useful for adaptive automated processes (for example dynamic allocation of a task between the user and the machine), implicit markup of multimedia content, video games, and error correction.

Passive BCIs can be useful in any kind of application in which the mental and emotional state of the user is relevant. From the machine's perspective, passive BCIs make it possible to dynamically adapt models of the user-state.

- **Shared control** involves transforming the classes recognized by the BCI into high-level commands that are sent to the application (see Figure 2.8). Shared control is a paradigm of delegation. The machine is responsible for a certain fraction of the system's intelligence, in that high-level concepts and complementary information can be used to determine how a single brain command must be transformed into a more complex, high-level command, or equivalently into several low-level commands.

  From the user's perspective, the number of commands to be sent is low, even for accomplishing complex tasks. The quantity of information transmitted through the BCI is significantly reduced compared to direct command approaches. The user is therefore able to rest while the high-level command is being executed [Lotte et al., 2010].

  It has been shown that shared control can be useful for operating wheelchairs [Philips et al., 2007]. For example, users can select a destination with a BCI based on the P300 paradigm, and a decision-making program equipped with a localization algorithm and object avoidance sensors decides the elementary actions that the wheelchair must perform as a result [Rebsamen et al., 2007].

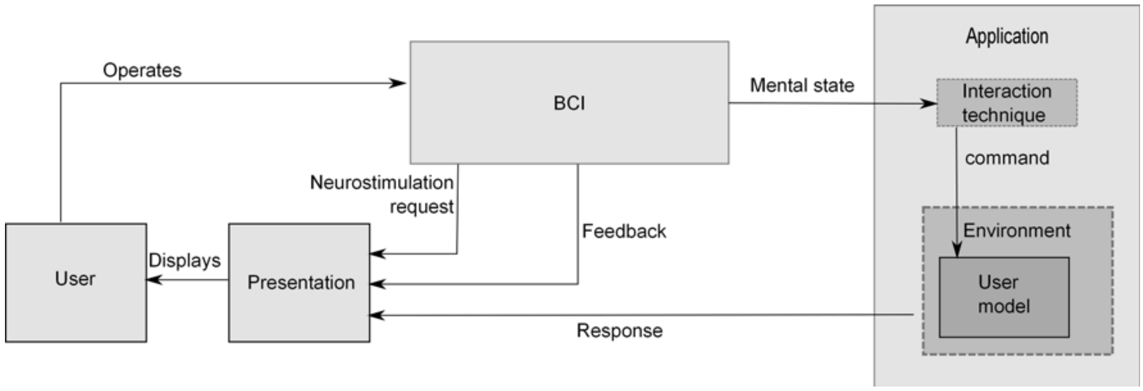


Figure 2.7: **Passive BCI**: The user does not send an explicit command, but the application can construct a model of the user based on his or her mental state

Shared control can significantly speed up the interaction when the machine is capable of anticipating the user's decisions. As artificial intelligence and decision-making algorithms improve, shared control might find new fields of application.

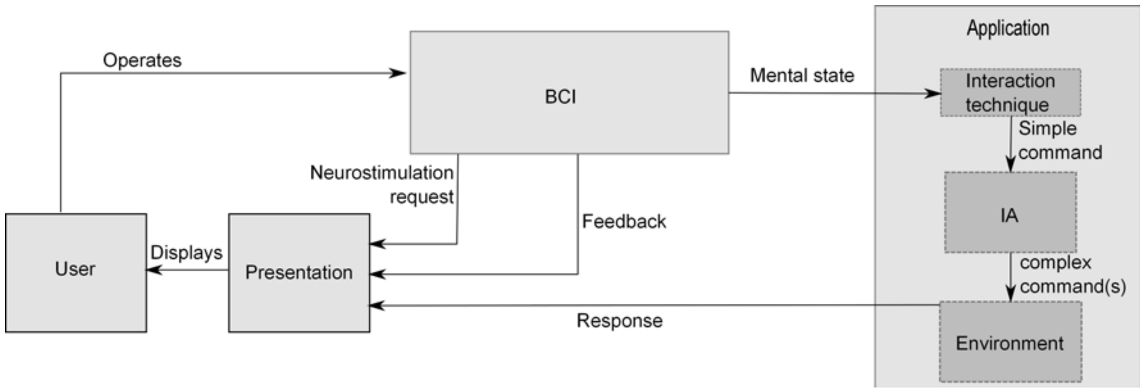


Figure 2.8: **Shared control**: a unique command is analyzed and transformed into a high-level command that may be automatically deconstructed into a series of low-level commands.

- The **multi-user** approach involves using multiple BCI inputs to control a single application. Several mental activities may be combined at different stages of the interaction. The mental state of each user is recognized independently and sent to the application, which combines the recognized mental states in order to produce a command. The purpose of this combination might be to improve the global performance, or to increase the number of available commands. Mental activities can also be combined at the signal processing level, allowing the classifier itself to determine the multi-user command [Bonnet et al., 2013]. Finally, multi-user approaches can be used to search for brain markers ("*hyperscanning*") by observing the similarities between the brain activities of two users in similar situations [Babiloni and Astolfi, 2014].

  An intermediate mental task must be performed by each user to create the correct mental state, and thus transmit the desired command, either in collaboration or in competition with the other user.

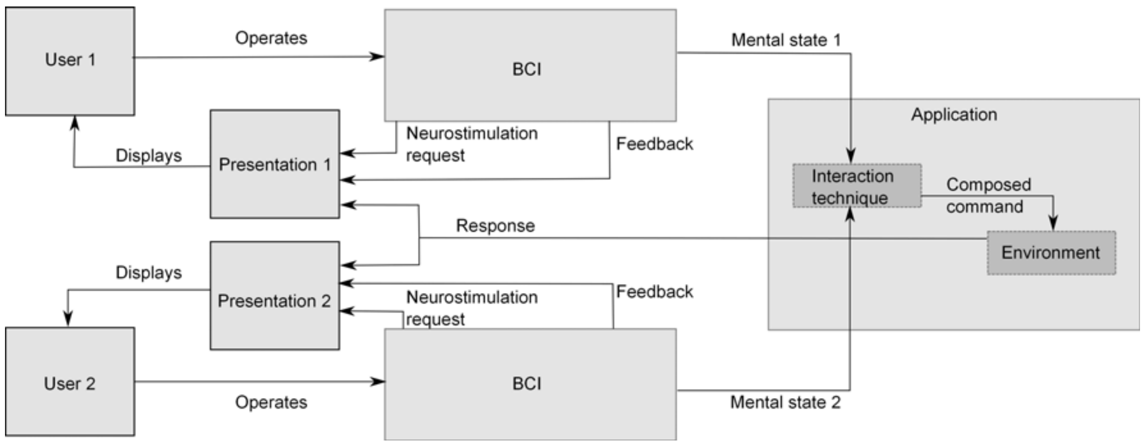


Figure 2.9: **Multi-User brain-computer interaction**: Two or more users are each using a BCI. A comparison block treats the results to allow the cooperation or the competition between the users.

Recently, it was suggested that multi-user BCIs could be developed for applications in video games [Nijholt and Gürkök, 2013]. Two players could attempt to synchronize their cerebral signals directly, or as a mean of achieving higher-level objectives. Competitive "*gameplay*" could also be introduced [Bonnet et al., 2013].

Using the cerebral activity of multiple users could potentially improve the precision of BCIs, as the noise in each individual signal becomes less significant given the information from the other users. Concretely, the classifier could use features extracted from all users in a single classification step. Also, the social presence of another user appears to stimulate the learning process [Richardson et al., 2003].

Future BCIs might belong to one of three possible usage categories. Firstly, BCIs might be used as an alternative to the keyboard or the mouse. It is however not entirely clear whether BCIs will be able to match the performance of these more classical input devices. The nervous signal that directs muscular activity comes directly from the brain, whereas non-invasive technologies only have access to noisy signals. BCIs are disadvantaged at the outset by the clarity of the signal compared to traditional devices (see Figure 2.10). However, even if other tools of interaction are more effective at accomplishing a given task, performing the task with a BCI might provide some inner advantages in term of entertainment [Nijholt et al., 2009]. Secondly, direct and hybrid approaches can be useful in assisting the interactions of disabled users. Finally, for applications aiming at wider audiences, BCIs might be used as complementary input devices that allow the usual channels of interaction to be kept free for other interfaces.

2.4 Applications

2.4.1 BCI for disabled people

Physically disabled people are unable to use the most common interfaces, and their communication skills can be extremely limited. BCIs offer them a new set of interaction techniques

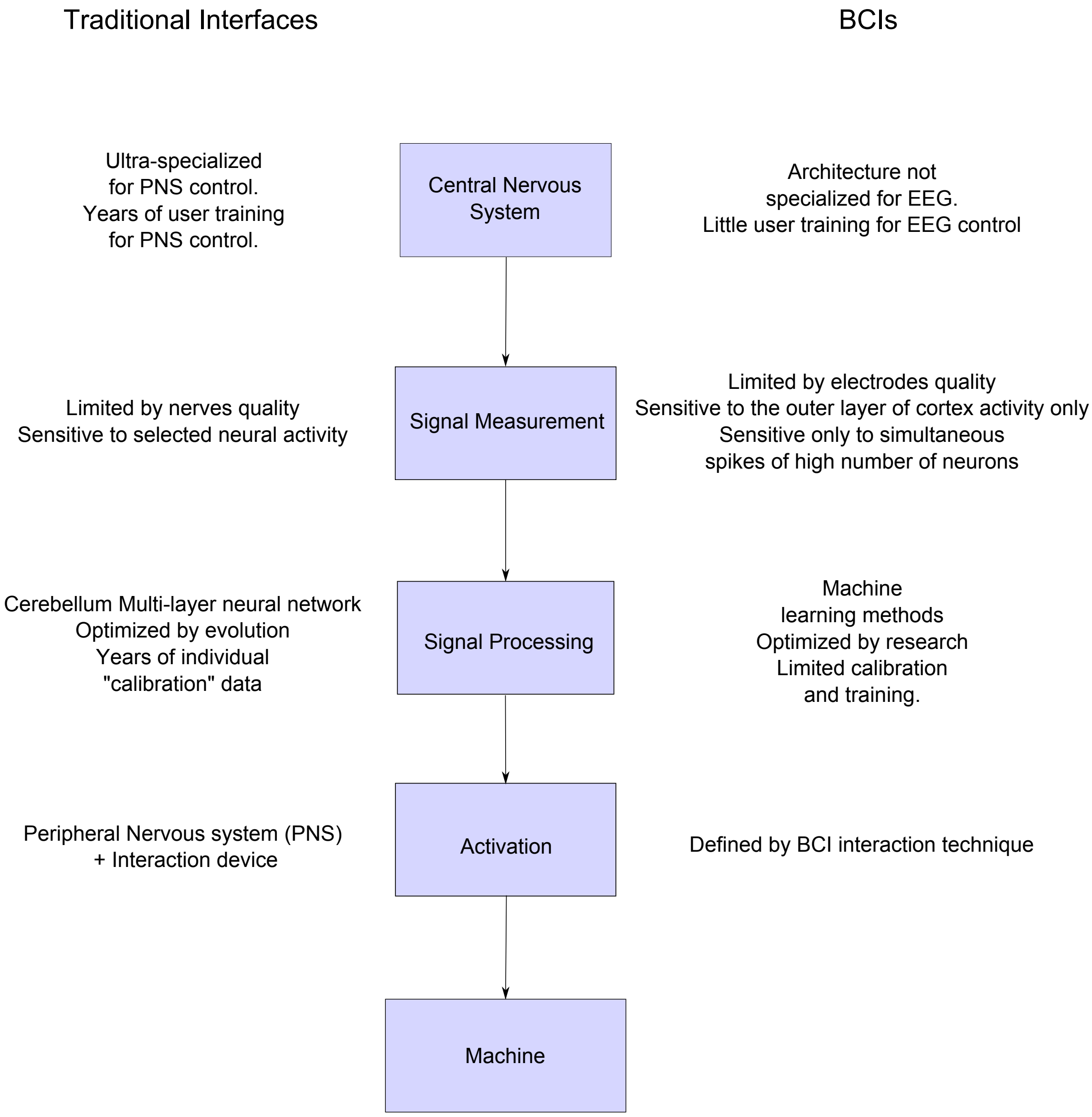


Figure 2.10: **Comparison of brain and motor-based interfaces**

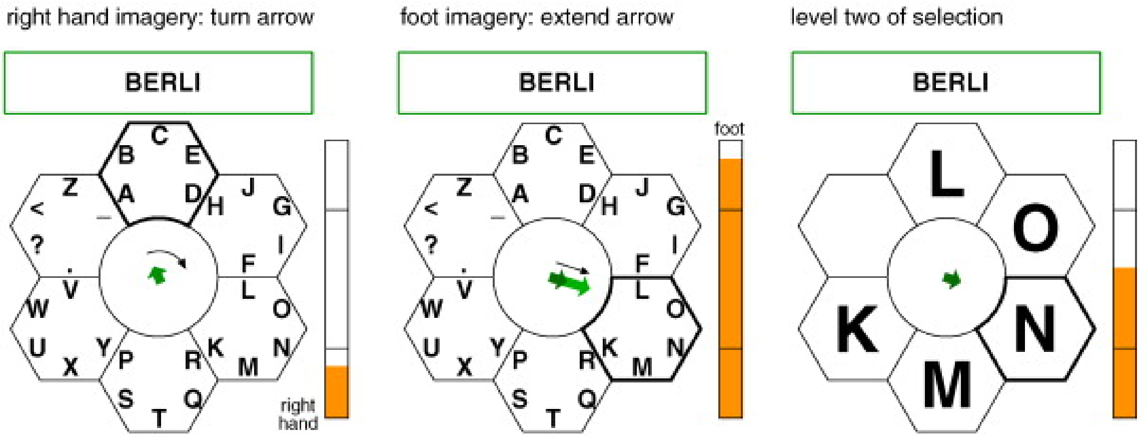


Figure 2.11: The Hex-O-Spell interface [Müller et al., 2008]

[Birbaumer, 2006], providing new communication channels.

2.4.1.A Spellers

One of the most immediate need for strong disabilities is to restore a communication process. BCI spellers provide these new possibilities, without restriction on the expressivity of the text.

The Thought Translation Device: One of the first speller using BCI was based on the control of slow cortical potential [Birbaumer et al., 2003]. By setting their cortical potential to "high" or "low", the subjects were able to select a letter in the alphabet, with successive selection between 2 parts.

Over 5 patients, one managed to reach the threshold of 75% accuracy in a few weeks, 2 others in a few months. With this method, they needed 2 minutes to write a single letter. The use of preselected words could fasten the process, but the patients did not like it, feeling it was making them more constraint by the system. Over the last 2 subjects, one changed his cognitive strategy during the training, which made his performance fall, and he was not able to come back from it. Another died before the end of the study.

Hex-O-Spell: Hex-O-Spell is a speller that uses motor imagery to select letters. The user selects cells containing others cells (see Figure 2.11). When a cell contains only one letter, this letter is written. Using a smart disposition of cells, it is possible to select any letter in only two selection step.

The selection of a cell is done by controlling a rotating arrow. One command, responding to imagery of right foot movement, makes the arrow rotate, while the other, corresponding to imagery of left hand, expands it. When the arrow reaches a cell, this cell is selected.

Using this system, a speed of 5 characters/minute (35bpm) has been achieved in peak for one subject, with no strong decrease of performance over long periods, and with little training [Blankertz et al., 2006b]. Unfortunately, not everyone being able to perform well at motor imagery related tasks, this speller presents high variability of performances among individuals.

The P300-speller: The P300-Speller simply presents a matrix of flickering letters. In order to select a letter, the user focuses on it and counts the number of time the desired letter flickers (see section 1.3.2) [Clerc et al., 2013].

The speed and order of flashes has a strong influence on the end speed and accuracy of the P300 speller. Several studies tackled the problem of flashing sequences, with several strategies being proposed. Letters can flash one by one at a fast pace, or line by line and column by column. Other grouping systems have been proposed [Thomas et al., 2014, Sauvan et al., 2009].

In [Donchin et al., 2000], the authors study the minimal number of flash for a desired precision by averaging. By optimizing this parameters, an offline speed of 7.8 characters/minute was achieved, with an 80% accuracy.

SSVEP spellers: SSVEP-based BCIs enable a high speed and precision of selection. As these criteria are critical for spelling application, several SSVEP-based spellers have been proposed.

The main difficulty in designing an SSVEP-based speller is to overcome the limitation in the number of targets on a computer screen, as the range of frequency is limited. In [Yin et al., 2015], a speller is presented where the row and column are successively selected, and achieved a speed of 41 bits/minute.

In [Chen et al., 2014b], a speller with 40 different stimulation is proposed. This high number of targets is enabled by encoding information in the phase and shape of the signal instead of just in its frequency. The resulting speller is reported to enable a speed of 172 bits/minute on average across subjects (corresponding to 40 characters/minute spelling). Later, the same research team reported reaching 5.32 bits per second (60 characters per minute) [Chen et al., 2015]. To the best of our knowledge, this is the highest information transfer rate ever achieved by a BCI [1].

2.4.1.B Web browsers

As more and more HCI functionalities are embedded in web browsers, prototypes of web browsers based on BCIs have been developped. To the best of our knowledge, the first prototype was proposed by [Karim et al., 2006]. Every command and link can be selected by successive dichotomy in the set of all possible commands, by using an SCP-based BCI.

Later, another web-browser was proposed, using SSVEP-based BCIs [Yin et al., 2009]. The same idea of successive selection in the set of accessible commands was used, with 3 SSVEP classes. Considering that a webpage usually presents more than a hundred hyperlinks, the process of selection can be long.

An alternative approach was proposed by [Sirvent et al., 2010]: instead of mapping the BCIs output to web browsing command, the BCI is used to drive a cursor and emulate a keyboard.

[1] in 2016

Table 2.2: Comparison between existing BCI-based web browsers

Web Browser	navigation commands	hyperlink following
[Karim et al., 2006]	successive dichotomy (SCP)	successive dichotomy (SCP)
[Yin et al., 2009]	SSVEP selection	successive trichotomy (SSVEP)
[Sirvent et al., 2010]	P300 selection	P300-based point-and-click
[Mugler et al., 2010]	P300 selection	P300 selection with link indexing

In this prototype, the only required BCI is P300-based, and all interaction tasks are reduced to selection. While this approach immediately gives access to any web functionality already existing in a standard browser, driving a cursor via selection is a tedious process: the user has to select a direction, and to enter the distance to travel. The cursor is then moved accordingly, and the user iterates this process to refine the position.

Finally, a more user-oriented approach was proposed in [Mugler et al., 2010]. Links are referenced and each link is associated to a code. A P300-based BCI permanently gives access to navigation command, as well as to hyperlink selection, by typing the letters of the code with a P300 speller.

A summary of the assets and limitations of these works is proposed in table 2.2.

2.4.1.C Controlling a wheelchair

Interacting with a computer, and especially being able to spell text is extremely useful for disabled users. Another goal of BCI-based interaction is to allow them to control a wheelchair, thus restoring a certain mobility [Rebsamen et al., 2007, Megalingam et al., 2013].

Several interaction techniques based on BCIs have been proposed. Motor imagery can be used as an intuitive command to direct the wheelchair to the left, right, or up front [Rebsamen et al., 2007] Similar commands of direction are used in [Singla et al., 2014], but SSVEP selection is used instead of motor imagery.

Alternatively, in [Rebsamen et al., 2007], a P300 selection is performed in order to choose a destination, with the wheelchair automatically choosing the path to reach this goal. Using such high level commands and giving a high level of intelligence to the interface allows a safe use, even in case of errors in the BCI. As a counterpart, the set of accessible destination is limited, and the user could have a weakened sensation of control.

2.4.1.D Controlling a mechanical arm

Controlling a mechanical arm typically requires a high number of commands. Using BCI selection methods would be mostly impractical, as selecting each articulation independently would take too much time and effort. However, motor imagery seems suitable for a more intuitive use, since the user only has to focus on the movement he wishes to perform. However, current motor imagery based BCIs based on EEG are limited in precision and in the number of classes (usually 2 or 3 classes) [Naveed et al., 2012].

Thus, controlling a mechanical arm using BCI require to switch from EEG to invasive sensors.

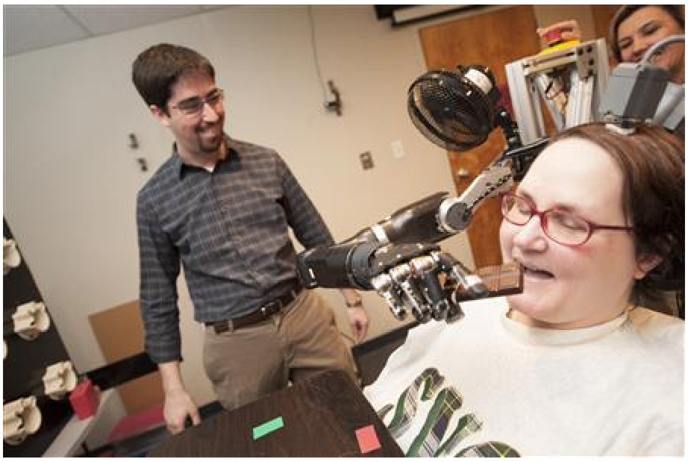


Figure 2.12: A patient controlling a mechanical arm via a BCI to eat chocolate [Collinger et al., 2014]

But by using invasive sensors, such as EcoG, one can achieve the necessary precision. A surgical operation is needed to install electrodes directly into the cortex. The activity detected in the motor cortex is then translated into commands for a robotic arm. After a very long training, one subject managed to eat chocolate or to drink with a straw using a robotic arm driven by an EcoG-based BCI [Collinger et al., 2014].

2.4.2 BCI for games

Most of the time, the limited precision of BCIs and the difficulty to control it are perceived as an interaction limitation. However, learning to use a BCI is often considered as a human skill that can be trained and mastered, especially for motor imagery-based BCIs [Guger et al., 2003, Lotte et al., 2013]. From the video game point of view, such a skill can represent a new challenge to the players. As mastering a mouse can be a key to success in a First Person Shooter game, BCI games can turn the difficulties of using a BCI into game challenges [Nijholt et al., 2009]. MindBalance is a simple game, where the player controls a character using two commands by SSVEP. He/She has to focus on one stimulus or the other in order to keep the character balanced on a rope [Lalor et al., 2005] (see Figure 2.13). The challenge of the game here comes directly from the limits of the classification precision. The next step for BCIs in games might be to propose new multi-player interaction. Players could cooperate in order to drive a BCI, or at the opposite, competitive games could enable players to oppose each other directly on their brain control [Nijholt and Gürkök, 2013].

The second interest of BCIs for videogames lies in the sensation of psychic powers that they could give to the user [Amores et al., 2016, Mercier-Ganady et al., 2015]. Because interacting without moving seems magic to the user, it may be used to improve immersion in fantasy world. Using BCI to enhance immersion was proposed in [Nijholt et al., 2009]. The user controls a character in the famous MMORPG "World of Warcraft", and can use a BCI monitoring the level of stress or relaxation. When the user is relaxed, the character is an elf, and when the user is stressed, the character turns into a ferocious bear. By detecting variation in cerebral rhythms (e.g. the power

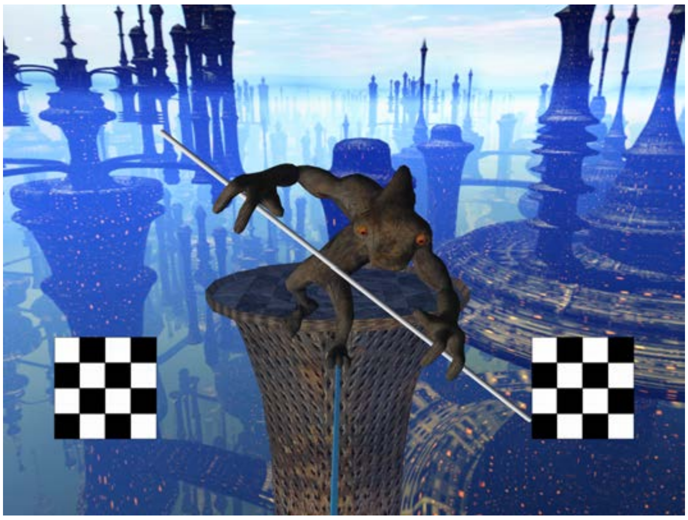


Figure 2.13: The MindBalance application, as display on the user's screen [Lalor et al., 2005]

Figure 2.14: The Dagobah immersive setup, where the user has to focus to lift objects [Lécuyer et al., 2013]

of the alpha band), it is possible to measure the level of concentration or relaxation of the user. In [Hjelm, 2003], a game is presented where the user has to relax in order to win. Similarly, in [Lécuyer et al., 2013], the user has to focus in order to lift a spaceship (see Figure 2.14). This setup was adapted and used in an immersive virtual environment. Overall, monitoring the level of relaxation or concentration of a player provides natural commands for specific gameplay operation, typically emulating psychic powers (telekinesis, metamorphosis).

Finally, another noteworthy potential use of BCIs for games lies in passive interaction. Since passive BCIs can monitor mental states such as frustration, stress, or boredom, it has been proposed to use BCI to adapt the difficulty of the game [George, 2012].

2.4.3 Passive BCI for interface evaluation

Passive BCIs (see section 2.3.2) can measure mental states such as mental workload, or emotional states. Such information can be useful to adapt other interfaces, or to evaluate them.

fNIRS imaging technique has received a particular attention for this type of application. fNIRS allows measuring the oxygenation variations of brain areas. A decrease in oxygenation is a direct trace of mental demand. While its cost is still too high for personal use (compared to EEG), it remains affordable for products or interfaces test phases. Additionally, getting insight from the user brain activity allows avoiding the usual biases of questionnaires. For these reasons, BCIs based on fNIRS have been proposed as a complementary approach for visual interfaces evaluation [Peck et al., 2013].

2.5 Conclusion

This chapter discussed Brain-Computer Interfaces from the point of view of Human-Computer Interaction. First, it introduced some relevant concepts from HCI. Then, a discussion was proposed on the characteristics of BCIs, using these concepts. Emphasis was put on the different types of feedback for BCIs, before commenting on the characteristics of BCIs, both as an interaction modality among others, and depending on the type of BCIs. In particular, the BCI alternatives for specific interaction tasks are presented. The main paradigms of interaction for BCIs were presented, showing how they differ from the standard interaction loop, and the characteristics of each interaction paradigm. A short description of the main fields of application of BCIs concluded this chapter.

As highlighted by this chapter, most research on BCIs so far focused on improving their speed and accuracy, through changes in the signal processing steps. As BCI application field expands from the assistance to disabled people towards videogames and interaction evaluation, it is now time to question the efficient use of existing signal processing methods, as well as the BCI user experience, in order to improve interaction based on BCI, and to exploit this technology in wider interaction contexts.

The new fields of application of BCIs require new interaction paradigms to get the best out of BCIs. The classical BCI context of use, in laboratory conditions, with only one subject without distraction and without any other device to use, if not obsolete, is too restricted for future use. We aim at exploring the BCI user experience, as well as improving BCI control and their efficient use.

Chapter 3

Cognitive demand of an SSVEP-based BCI

When working with BCIs, the standard approach is to focus on the capacity of the signal measurement and processing to detect and classify cerebral activity. Little attention is given to the capacity of the user to produce and control the right cerebral pattern. However, it is known that some users perform generally better than others for this task. Controlling a BCI has previously be described as a *skill that can be learned* [Guger et al., 2003], and [Daly et al., 2009] notes that producing and maintaining the suitable mental stage often requires *intense focus of concentration* from the user. During the interaction process, the attention of the user can be diverted due to other visual or acoustic attentional tasks which could decrease the efficiency of the BCI. Existing studies in motor imagery [Daly et al., 2009, Tamir et al., 2007] and P300-based [Pratt et al., 2011, Brouwer et al., 2012, Kida et al., 2004] BCIs have showed that the attention divergence can have a strong impact on the efficiency of the BCI discouraging the realization of dual tasks. However, while SSVEP-based BCI is one of the fastest and more reliable BCI paradigms [Volosyak, 2011], the cognitive demand it induces and the attentional requirements remain to be explored.

In this chapter, we present the results of a user study which analyzes the impact of a working memory task (N-back task [Owen et al., 2005]) done in parallel with a SSVEP-based BCI target selection task. Participants were instructed to perform a simple selection task (three SSVEP targets) using a classical SSVEP setup (primary task) and at the same time, they had to perform an N-back memory task involving visual or auditory stimuli (secondary task). The load factor of the N-back task was controlled in order to alter its working memory demand. Our results indicate that SSVEP-based BCIs can be used while performing a complex additional mental task. The potential interferences between various types of cognitive resources (working memory, visual attention, auditory attention) do not appear to deteriorate SSVEP detection.

The remainder of this chapter is organized as follows. An overview of the related work on BCIs and mental demand is first presented. The experimental apparatus, methods, and results are then described. A general discussion concludes the chapter.

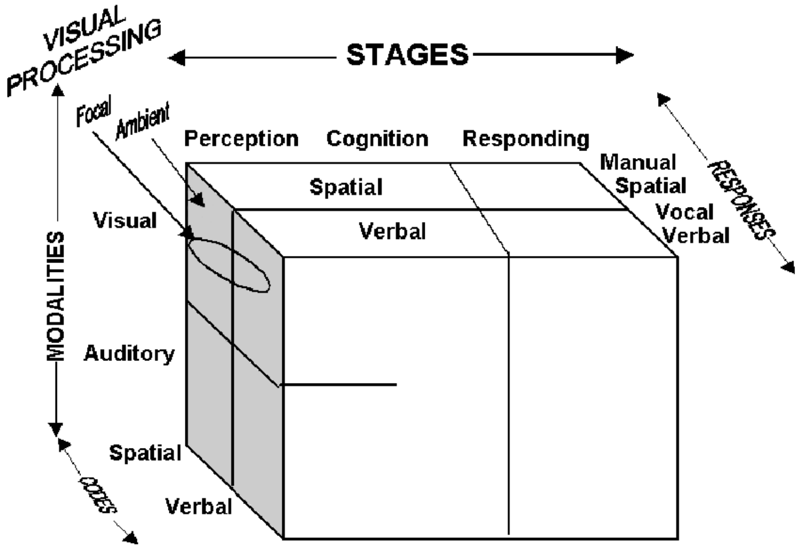


Figure 3.1: The 4-D multiple resources model [Wickens, 2008]

3.1 Related work on the mental demand of BCIs

So far, studies on active BCIs have indicated that they require the full user attention. Performance of BCIs based on the commonly-used P300 cerebral pattern has been shown to decrease when the user has to focus on another concurrent task. For example, Watter et al. studied an N-back task coupled with P300 use [Watter et al., 2001]. They observed that the peak amplitude of P300 signal decreased with increasing memory load, reflecting the reallocation of attention and processing capacity. Hence, performing the memory task negatively affects the P300-based BCI accuracy. Several other studies have since observed similar results with P300, e.g. [Kida et al., 2004, Pratt et al., 2011, Brouwer et al., 2012], and similar observations have also been made with motor imagery [Tamir et al., 2007, Daly et al., 2009].

However, while the influence of several types of cognitive resources on P300 is well documented (even if not fully understood yet), SSVEP-based BCIs have received less attention so far.

Cognitive psychology is still looking for a uniform model of human perception, cognition, and action processes. In this chapter, we propose to consider the 4-D model of multiple resources. The Multiple Resources Theory is an approach to describe the extent to which dual-task performance will lead to decreases in time-sharing ability [Wickens, 2008]. This model considers four dimensions of resources (see Figure 3.1). It classifies the cognitive resources by stages of processing (perception, cognition, and responding), by codes of proceeding (spatial and verbal), and by modality (auditory and visual). The model is generally refined with a fourth dimension, differentiating between focal and peripheral attention. The general idea between these distinctions is that if two tasks use different levels along each of the four dimensions, time-sharing will be better [Wickens, 2008]. Dual tasks should get higher performance if they require resources from different axes. The notion of mental workload can be derived from this model as the demand put on the cognitive resources in general. In particular, the case where the demand is less than the capacity of resources available can be distinguished from the one in which the demand exceeds the capacity, in which case performances are expected to break down.

Within the 4-D model of Multiple Resources, reactive BCIs hold a particular place. They are used as means of action, and should thus put demand on the responding resources. However, they are actually activated by visual attention. In the case of P300, for example, it has been observed that the working memory required by a secondary task competes for attention with visual perception [Agam and Sekuler, 2007]. By contrast, BCIs based on motor imagery follow the traditional perception-cognition-response order of processing stages. Motor imagery should have a high demand in cognition and responding cognitive resources, explaining the need for the user's undivided attention [Daly et al., 2009].

SSVEP is greatly influenced by the gaze focus point, but it has been shown that visual attention alone modulates the evoked signal [Andersen et al., 2008]. However, it is yet unknown how much visual attention is required to operate an SSVEP-based BCI, and how difficult it is to maintain this attention while performing other demanding tasks, visual or not.

In this chapter, we study how SSVEP accuracy evolves, as a function of the verbal working memory required for a secondary task, with either a visual or an auditory input: the two most common HCI modalities.

3.2 Materials and methods

In order to measure the cognitive demand of a task, the classical approach is to use a dual-task paradigm [Baumann et al., 2007, Bruder et al., 2015, Evans and Stanovich, 2013, Owen et al., 2005]. Participants are instructed to perform a primary task (that for which the mental demand is unknown) and a secondary concurrent one of which the cognitive cost can be controlled. We followed this standard procedure in this study. Participants were asked to select one target among three using an SSVEP-based BCI while performing an N-back memory task on letters.

For the N-back memory task, letters were presented in sequence. At each step, the participants had to indicate whether the current letter was the same as N steps before [Owen et al., 2005]. In order to observe the effects of different types and levels of attention, the difficulty of the memory task (load factor) changed across the experiment, as well as the instruction modality (visual or auditory).

3.2.1 Participants

Twenty-six participants were enrolled in this study: 21 men and 5 women, aged between 19 and 41 (mean 26.3, SD 5.8), 20 right-handed and 6 left-handed. Half of the participants performed the memory task with visual stimuli, while the others performed it with auditory ones.

3.2.2 Brain-Computer Interface

EEG data were recorded using 6 passive electrodes out of a non-invasive 16-channel system (g.USBAmp, g.tec company, Austria), with a sampling rate of 512 Hz, combined with the OpenViBE software[1].

[1] http://openvibe.inria.fr

Figure 3.2: Experimental setup. 1: Visual display. 2: g-tec EEG headset. 3: EEG signal amplifier. 4: Laptop for EEG signal processing. 5: Keyboard. 6: Computer handling the stimulations and the instructions (either as displayed or audio text)

Electrodes were positioned according to the extended 10-20 system on `CPz`, `POz`, `Oz`, `Iz`, `O1` and `O2`. A reference electrode was also located on the right hear, and a ground electrode on `AFz`. Signal quality was ensured using an impedance checking of each electrode.

Stimuli were disks flickering between black and white at 10, 12, and 15 Hz, with an opacity of 67% determined through informal tests (Figure 3.5). These made it possible to read the visual stimuli of the N-back task (letters) without difficulty, while still providing good luminosity contrast for the flickering to be perceived. Stimuli were displayed on a DELL^{TM}Ultrasharp^{TM}2007FP 51 cm screen (20.1 inches), with a resolution of 1280×1024 pixels, and a refresh rate of 60 Hz.

3.2.3 Experimental design

Participants were sitting in front of a computer screen, wearing an EEG headset (see Figure 3.2). Prior to the experiment, participants were asked to fill in a written consent form, and a questionnaire collecting statistics about gender, dominant hand, age, and sight. The BCI was then calibrated. During the calibration phase, three targets were displayed, flickering between black and white. The size and the position of targets were the same as in the experiment. For each calibration trial, participants were asked to focus on a given target, indicated by an arrow, while their EEG activity was recorded. For each trial, flickering lasted 7 seconds. Breaks of 5 seconds separated each trial. The arrow indicated the next goal target was displayed during the break, 2 seconds before the flickering starts. No feedback was available during this calibration. The calibration phase contained 18 trials (6 trials for each frequency), for a total duration of 3 minutes

real accuracy	feedback (simulated) accuracy
33%	50%
50%	62.5%
60%	70%
80%	85%
92%	94%
100%	100%

Table 3.1: A comparison of the accuracy given by the positively biased feedback with the real accuracy

and 36 seconds.

The primary and secondary tasks were then explained to participants. They were asked to fill in a form in order to make sure the instructions were correctly understood.

Primary task: On screen, three flickering circular targets were displayed, as shown on Figure 3.5. For each trial, participants had to select the designated target (indicated with an arrow), by focusing their visual attention on it without looking away during 7.2s (this duration was chosen for synchronization constraint with the secondary task, as described later on). At the end of the trial, a visual feedback indicated if the selection was successful, if the target was successfully selected, it turned green, otherwise, the wrongly selected target turned red.

In order to maintain the motivation of participants, we used a positively biased feedback. Positively biased feedback has been shown to keep up the motivation of participants with low accuracy rate in a BCI context [Alimardani et al., 2014]. We used a 25% positive bias, designed so that it has low impact on participants with high accuracy, and high impact on participants with low accuracy. The method to achieve this is described below: When the class was correctly recognized, the feedback was positive. If it was not, there was still a 25% probability for the feedback to be positive. As a result, bias is more important when the accuracy is low. This method ensures that even a participants at random level (33%) gets 50% of positive feedback. At the same time, a participant with close to perfect accuracy is given more faithful feedback (see Table 3.1 for a few examples of the resulting bias).

Secondary task: The secondary task was an N-back letter task (see Figure 3.3). Participants had to memorize sequences of letters presented to them and push a key when (and only when) the current letter was the same as the N-th previous one, before the next letter apparition. In the following example, marked letters are those which should be recognized as repetitions for N=2: L H C **H** S C Q **C** **Q** L C K L H. The value of N determined the number of letters that participants had to remember, which directly determined the difficulty of the task. We used two values for N: 1 and 2, on top of a control condition, in which no letter was presented. These are the most common difficulty levels in experiments relying on the N-back working memory paradigm [Owen et al., 2005].

Each sequence of letters was 4+N letters long. Each letter has a 50% probability of being the same as the N-th previous one. At the end of each sequence, feedback was provided as a score to indicate the number of correct answers over the last 4 letters (the first N being irrelevant, as there

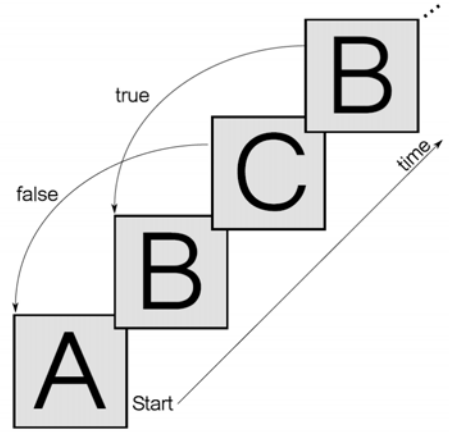


Figure 3.3: 2-back memory task [Schatzschneider et al., 2016]

is no N-th previous one to compare to). Consecutive letters were spaced in time with a randomized interval between 1.6 and 2.0s (as in [Schatzschneider et al., 2016] and [Bruder et al., 2015]), with constraints ensuring that (1) the total duration of the first N letters was N$*$1.8s; and (2) the total duration of the last 4 letters was 7.2s (see Figure 3.4). These constraints ensured the synchronization with the constant pace of the primary task.

As previously said, we used two presentation modalities for the letters: visual and auditory. For the visual condition, letters were displayed for 500ms at the SSVEP target positions, with the stimuli flickering on top of them. The same letter was displayed for every target, so that it could be read wherever the participant looked (see Figure 3.5). For the auditory condition, letters were spoken by a male synthetic voice[2].

Dual-task: Both tasks were performed simultaneously (see Figure 3.4). For each trial, the first N letters (that never require an answer) were presented before the start of the SSVEP stimulation (instruction phase). During the 7.2s of the SSVEP stimulation (SSVEP flickering), 4 letters of the N-back task sequence were presented. At the end of the trial (feedback phase), the feedback for each task was displayed for 2 seconds. In total each trial lasted 12.8s and between each trial one second of break was added.

The study had two main factors, the *difficulty* of the N-back task (within-subjects factor with two levels N= $\{1, 2\}$ and a control level) and the *presentation* of the letters (between-subjects factor with two levels: visual and audio). For each level of difficulty participants performed 44 repetitions, grouped in two bocks. In total, participants did 6 blocks (2 for each level of difficulty). The presentation of the blocks was randomized to avoid ordering effects. A break was given between each block. Participants were given full control over the duration of the break. The full duration of the experiment was about one hour, including installation time and briefing.

The dependent variables were the amount of correct answers for (SSVEP-based) target selection and the N-back task. For the N-back task, the score for each trial ranged from 0 to 4, one point

[2]Synthetic voice from `http://www.fromtexttospeech.com/`

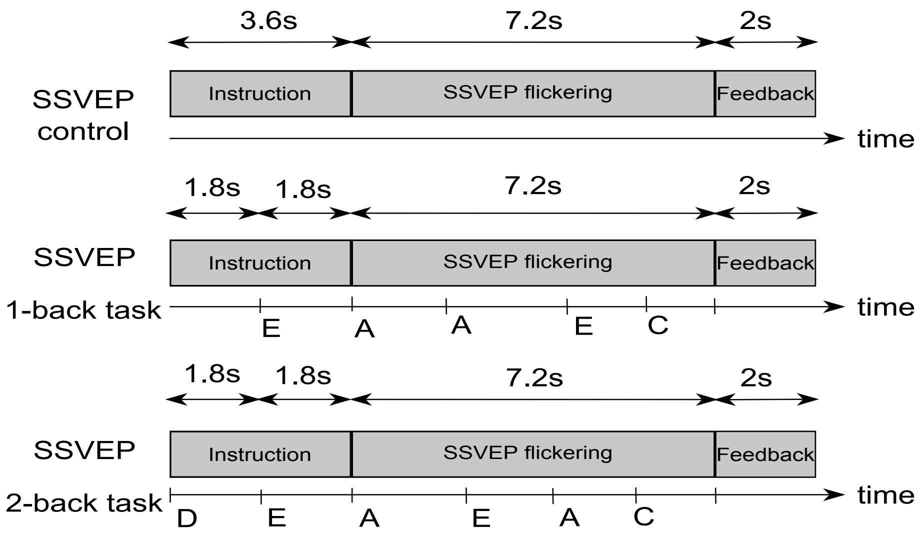


Figure 3.4: Synchronization of SSVEP and N-back memory tasks.

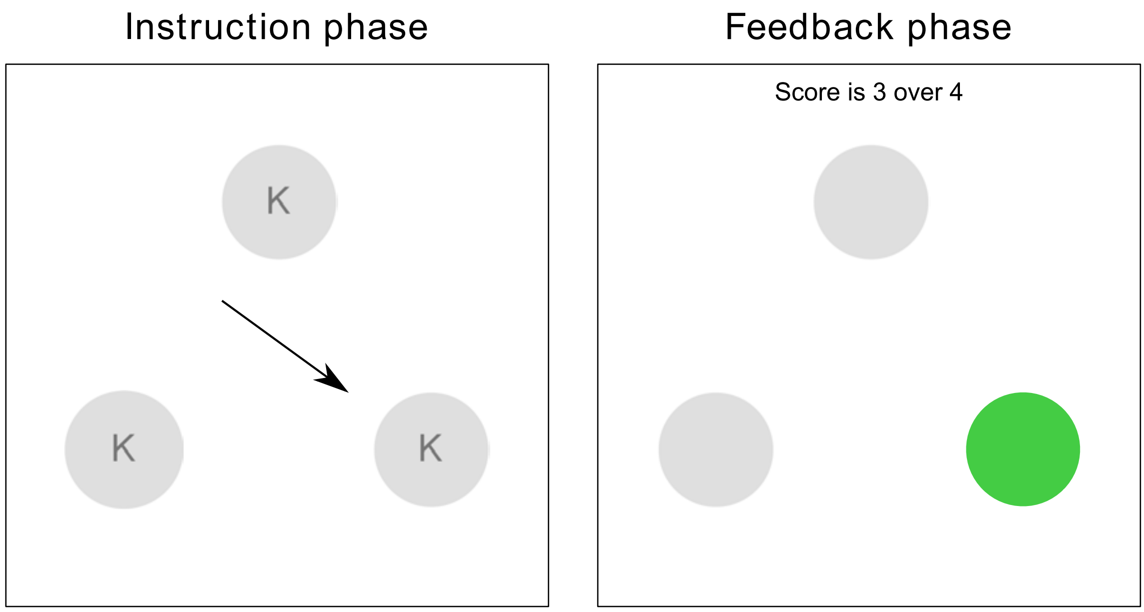


Figure 3.5: Visual display during the instruction (left) and feedback (right) phase. For the auditory conditions, letters are not displayed. The flickering frequencies of the three disks were the same in both conditions (10, 12 and 15 Hz). An arrow indicated the target to select. For both conditions, after each trial, the score of the N-back task is displayed on top of the screen, and the target actually selected with the BCI changed color to indicate selection error (red) or success (green).

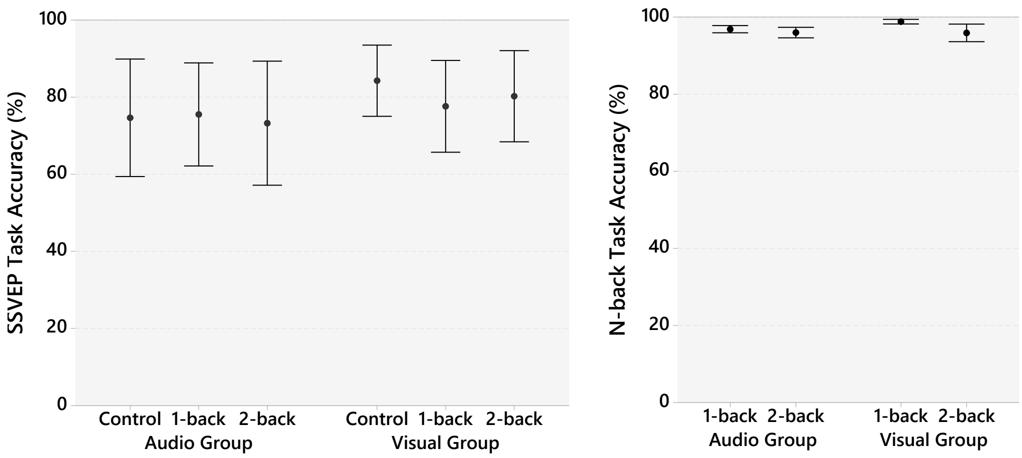


Figure 3.6: Confidence intervals (95%) for the mean accuracy for the SSVEP selection task (Left) and for the N-back task accuracy (Right).

for each good response.

3.2.4 Results

As the target selection accuracy and the N-back task accuracy followed a normal distribution (Anderson-Darling normality test both $p < 0.01$), we performed a mixed two-way ANOVA analysis. Bonferroni post-hoc tests ($\alpha = 95\%$) were used when necessary. Figure 3.6 presents the summary of the results.

SSVEP Target Selection Accuracy The ANOVA difficulty and presentation vs. accuracy did not show any significant effect in difficulty ($F_{2,48} = 1.62$, $p = 0.209$, $\eta_p^2 =$), in presentation ($F_{1,24} = 0.57$, $p = 0.459$, $\eta_p^2 =$) or in their interaction ($F_{2,48} = 2.25$, $p = 0.116$, $\eta_p^2 =$). The standard deviation for both the auditory ($M = 0.74$; $SD = 0.24$) and the visual $M = 0.81$; $SD = 0.18$) groups were high, especially for the auditory group (24%). Eighteen participants had scores higher than 80% (two participants had a perfect score) and six participants had scores lower than 60%. However, additional analysis taking into account this grouping did not show additional significant differences.

N-back Task Accuracy The ANOVA difficulty and presentation vs. accuracy showed a main effect on difficulty ($F_{1,24} = 9.23$, $p > 0.01$, $\eta_p^2 = 0.28$). But there was no main effect in presentation ($F_{1,24} = 1.82$, $p = 0.190$, $\eta_p^2 =$). No two-way interaction effects were found ($F_{1,24} = 2.69$, $p = 0.114$, $\eta_p^2 =$). Post-hoc tests showed the accuracy for the 2-back tasks ($CI = [3.79, 3.89]$) was significantly lower than for the 1-back tasks ($CI = [3.89, 3.93]$). Participants were still proficient on performing the task no matter the difficulty, with accuracy globally above 95%. From the two-way interaction analysis, we observed a tendency for increased performance for the 1-back task and visual presentation, still the results are not conclusive.

3.3 Discussion

We observed that performing a demanding memory task with visual or auditory stimuli did not significantly affect the level of accuracy of an SSVEP-based target selection task. This observation indicates that the visual attention required by SSVEP does not clearly impair the working memory, contrary to what was observed with P300-based BCIs [Pratt et al., 2011, Brouwer et al., 2012, Kida et al., 2004].

In the N-back task with auditory stimuli, the participant's attention had to be divided between auditory attention (to hear the pronounced letters), and visual attention (to acquire a particular flickering target).

As no influence of the level of difficulty was observed, it seems that participants managed to reach a satisfying level of visual attention (for SSVEP selection), auditory attention (to get the sequence of letter) and working memory (to perform the mental task on letters) at the same time.

The visual N-back task is not only demanding in terms of working memory, but also in visual attention. One would expect this demand to interfere with SSVEP, which is known to be sensitive to visual attention [Andersen et al., 2008]. However, according to [Morgan et al., 1996], SSVEP is modulated by the localization of visual attention, but not by its object. Our experiment tends to validate this hypothesis. When participants focused on the target position in order to read the letters of the N-back task, the amplitude of the SSVEP response in the EEG signal was still high, as it does not matter if attention is focused on the flickering or on the letter, as soon as these two stimulations are co-localized.

As predicted by the 4-D Multiple Resources model, multi-tasking is possible with low cost in performance when the demand lays on different cognitive resources. Our study tends to indicate that the overlap between auditive and visual attention does not prevent participant from succeeding in the dual-task.

It can be noted that while P300-based BCIs are also influenced by attention, their compatibility with a secondary task is different. The typical approach to maintain the user's visual attention on the P300 stimulation is to give them a secondary task to perform on this stimulation. The secondary task used for P300 should help maintain visual attention but have a low cognitive cost, so that it does not impair the P300 response (the most common secondary task for P300 is to count the flashes). In particular N-back tasks have been judged non-suited for P300 secondary task [Watter et al., 2001]. Our results show that by contrast, for SSVEP, their processing cost (for N=1 or 2) does not significantly impair the BCI accuracy.

The dual-task performance observed in this study is encouraging for the use of SSVEP in Human-Computer Interaction, as this technique is compatible with various other tasks, demanding different cognitive resources. It particularly encourages the development of hybrid approaches, mixing SSVEP-based BCIs with other interaction devices. The good accuracy obtained in the auditory condition also indicates that the user of an SSVEP-based BCI could listen to music or to someone talking to him/her, while still using the BCI efficiently.

3.4 Conclusion

We explored the mental resources required to operate an SSVEP-based BCI. An experiment was performed where SSVEP-based target selection accuracy was measured while participants had to perform a secondary task at the same time. The secondary task was an N-back memory task with visual or auditory input. Our results indicate that the difficulty of the secondary task did not significantly impact the primary SSVEP-based task. This observation was confirmed with both visual and auditory input for the N-back task.

The high dual-task performance indicates that using an SSVEP-based BCI does not strongly reduce the amount of mental resources available, as it does not significantly impact auditive attention nor working memory. These results are encouraging for HCI, indicating that the user of an SSVEP-based BCI could still think about something else when using the interface. For example, an SSVEP-based BCI user could possibly follow a conversation while using the BCI.

Chapter 4

User frustration of an SSVEP-based BCI

Despite important advances on the reliability of brain activity classification based on EEG, it is still well-known that the error rate of BCI systems remains much higher (typically around 20% [Müller-Putz et al., 2008]), compared to other input devices such are computer mice or keyboards (typically less than 4% [Sears and Shneiderman, 1991, Trewin and Pain, 1999]). On top of this difference in the range of error, BCIs differ from other devices in regard to the interpretation of errors. In the case of interfaces based on a mouse and a keyboard, most errors are attributable to the user. For BCIs, the distinction between errors due to the system and errors due to the user is much less clear. In such conditions, it is yet unknown how a small difference in accuracy impacts user experience.

This is probably one of the reasons why current BCI research is mostly focused on improving the robustness of EEG signal-processing and not on studying the user experience of BCI systems. In particular, we do not know if and how the error rate can influence the user experience and the perceived fatigue. For most interfaces, errors in the interface are a known to be a primary cause of user frustration [Ceaparu et al., 2004]. How does the user frustration correlate with the BCI system accuracy? How does frustration, motivation, or fatigue evolve with the duration of the BCI use? In the BCI community, it is generally assumed that the unreliability of current BCI systems have a negative impact on user experience. In particular, a prolonged use of a BCI could lead to fatigue or boredom [Nijholt et al., 2009]. Some studies have started to explore the effect of error rates for BCIs: in [Fard and Grosse-Wentrup, 2014, ?], artificial errors were added to a keyboard-based interface, in order to emulate the BCI accuracy level. As expected, participants testing this interface reported a higher frustration level when the error rate was high. However, the use of a keyboard to emulate the BCI experience can be put into question. In particular, participants were aware that the errors were artificial. We believe the study of the impact of errors in BCIs requires a more realistic setup.

We aim at exploring the specificities of BCIs' users experience, by reproducing the BCI context more precisely. In this chapter, we report an experiment conducted to assess specifically the influence of error rates and reliability on the perceived frustration, fatigue and motivation of BCI users. Participants were instructed to perform a simple target selection task using a BCI (see Figure 4.1), but the error rate of the interface was artificially controlled. Participants were unaware that

Figure 4.1: Experimental apparatus. A participant wearing an EEG headset tries to select one of the 3 flickering targets by focusing on it. The progression bar, visible on the right of the screen, indicates how many trials have been successful so far.

they had no actual control over the selection task. To measure the evolution of their subjective experience, they had to fill in individual questionnaires after each block of trials.

The remainder of this chapter is organized as follows: We first present our experimental apparatus and methods. Then, the results of our "fake" BCI experiment are reported. The chapter ends with a general discussion and a conclusion.

4.1 Materials and methods

We aimed at assessing the influence of error rate on frustration, fatigue and motivation of BCI users. Participants were asked to perform a selection task using a BCI and wearing an EEG headset. EEG data was recorded but, without the users' knowledge, it was not used during the selection task. Instead, in order to fully control the error rate of the BCI, we used a fake feedback presenting pre-determined results with an error rate varying from 5 to 50%. This range of error rate is not uncommon for BCIs. For example, it was the range of error rate observed in the experiment described in the previous chapter.

4.1.1 Participants

Twelve participants were enrolled in this study: 7 men and 5 women, aged between 18 and 60 (mean 29, SD 12), 11 right handed and 1 left handed. One of them presented a small degree of color blindness; the others had normal or corrected vision. None of the participants had ever presented signs of epileptic seizure. All of them were ignorant of the purpose of the experiment and of the fact that they had no control over the selection task.

4.1.2 Brain-Computer Interface

The used BCI relied on SSVEP. This family of BCIs is promising for HCI, because it enables high information transfer rates [Middendorf et al., 2000]. However, the flickering stimulation can be tiring and uncomfortable for the user [Lamti et al., 2014a].

During the experiment, EEG data was not used on-line. However, in order to give to the participants the impression of being in the same condition as in a real BCI session, the EEG headset was installed normally, and the EEG data was recorded (for off-line analysis). Electrodes were positioned according to the extended 10-20 system on CPZ, POZ, OZ, IZ, O1 and O2. Additionally, a reference electrode was located on the right hear, and a ground electrode on AFZ. Signal quality was ensured using an impedance checking of each electrode.

4.1.3 Experimental design

Participants were sitting in front of a computer screen, wearing an EEG headset (see Figure 4.1). Three flickering square targets were displayed on screen. For each trial, the task consisted in selecting one target (indicated with an arrow), by simply focusing on it without looking away during 4 seconds. A fake feedback based on visual and auditory cues was provided at the end of each trial, indicating either a success or a failure. The visual feedback was a single word displayed at the center of the screen (*success*; *failure*). The auditory feedback was either a buzzer sound (failure) or a game-like reward sound (success). Additionally, a vertical progress bar was displayed on the right of the screen. This progress bar grew for each success, and stayed at the same level in case of failure. In a controlled laboratory experiment, the number of variable factors needs to be kept to a minimum, in order to avoid potential uncontrolled bias. However, this constraint makes the task seem artificial, and the participants have little incentive to success, compared to a normal use of an interface. Pre-experiments revealed that without an appropriate feedback, participants quickly lose interest in the task. Our choice of a rewarding or punishing feedback aims at maintaining the user engagement, and emulating the cost of errors encountered in real use.

Participants had to achieve blocks of 20 consecutive trials. This size of block enable to give meaningful accuracy levels, while still keeping each block to a reasonable duration. Error rates changed depending on the block. The 4 levels of error rate were: 50% (10 errors over 20 trials), 35% (7/20), 20% (4/20), and 5% (1/20). There were 3 repetitions for each of the four conditions, for a total of 12 blocks. This number of blocks was chosen as a trade-off between the need for a long experiment, allowing the study of the effect of a prolonged use, and the need for a duration bearable to the users. At the end of each block, participants had to fill in a short questionnaire to gather their state. In order to avoid ordering effects, the order of the blocks was randomized. We ensured that each error rate was presented exactly once for each third of the experiment.

Each trial lasted 8 seconds (2 seconds for instructions, 4 seconds of flickering, 2 seconds of pause), and each block lasted 2 minutes and 40 seconds. The duration of the experiment was around one hour, including the time of installation and briefing.

4.1.4 Collected data

The questionnaires filled at the end of each block collected subjective data on the user perspective, for each participant. They were given as little information as possible, in order to avoid experimentalist bias. Using Likert scales, participants had to grade after each block: their frustration during the last block (*instant frustration*), their frustration since the beginning of the experiment (*global frustration*), their fatigue since the beginning of the experiment, their motivation at this stage of the experiment (*motivation*), and whether they found the interface effective during the last block (*effectiveness*). The definitions of each term (frustration, fatigue, motivation, and effectiveness) were provided on the questionnaire to ensure a good understanding of the questions. Likert scales for instant frustration, global frustration, fatigue and motivation were scaled on 7 levels, as recommended in [Urdapilleta et al., 2001] : absent (1), barely perceptible (2), faintly present (3), light (4), marked (5), pronounced (6), strongly pronounced (7). Effectiveness was rated on 5 levels, as proposed in [Vagias, 2006] : strongly disagree (1), disagree (2), neither agree nor disagree (3), agree (4), strongly agree (5). The full questionnaires are provided in Annex C.

Our hypothesis was that frustration increased with the error rate of the previous block, and over time. Motivation was checked mainly to assert that the feedback is playing its incentive role, and that users are still focusing on the task. We are also interested in a potential discouraging effect of an high error rate. Finally, fatigue is measured in order to evaluate how it evolve with the time of SSVEP stimulation exposition. Is there a sharp time limit, or a more continuous increase?

4.2 Results

The results of the ratings provided by participants at the end of the 12 blocks of trials are summarized in Table 4.1. In order to test for two-way interactions, the data were transformed using Aligned Rank Transform (ART) [Wobbrock et al., 2011]. After ART, a standard two-way ANOVA analysis with Tukey post-hoc tests ($\alpha < 0.05$) was performed.

Instant frustration. The two-way ANOVA error rate and repetition against instant frustration ratings showed a main effect on error rate ($F_{3,33}$=23.0;$p<0.001$;η_p^2=0.67) and repetition ($F_{2,22}$=7.2;$p<0.001$;η_p^2=0.4), no interaction effects were observed. For all levels of error rate (except between 35% and 20%) all pairwise comparisons showed significant results (all $p < 0.05$). Regarding the repetition factor, only repetitions 1 and 3 were found to be significantly different. As expected lower levels of error rate result in lower ratings of instant frustration, but instant frustration also increases with repetitions.

Global frustration. The two-way ANOVA error rate and repetition against global frustration ratings showed a main effect on error rate ($F_{3,33}$=7.0;$p<0.001$;η_p^2=0.39) and repetition ($F_{2,22}$=6.7;$p<0.001$;η_p^2=0.38), no interaction effects were observed. Post-hoc tests showed that global frustration ratings at error levels of 50% and 20% are significantly higher than at 5%. Regarding the repetition factor, the global frustration for the second and third repetition was significantly higher than that of the first repetition. We could observe here that the level of global frustration rises quite fast and then seems to stabilize.

Effectiveness. The two-way ANOVA error rate and repetition against effectiveness ratings showed only a main effect on error rate ($F_{3,33}$=17.3;$p<0.001$;η_p^2=0.61). For all levels of error rate (except between 35% and 20%) all pairwise comparisons showed significant results (all $p < 0.05$).

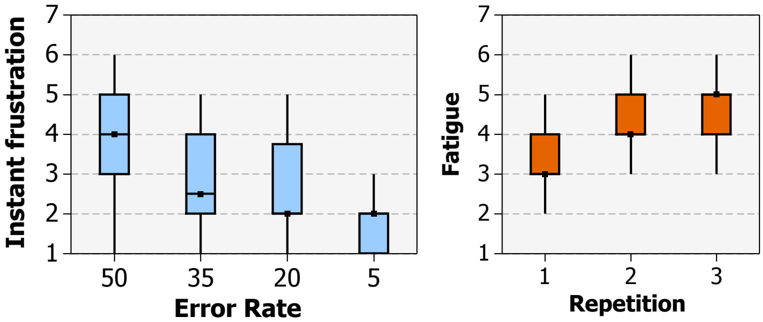


Figure 4.2: Experimental results. (Left) Frustration ratings as function of BCI error rate. (Right) Fatigue ratings as function of repetition.

	Instant frustration			Global frustration			Fatigue			Motivation			Effectiveness		
Error rate	Q1	M	Q3	Q1	M	Q3	Q1	M	Q3	Q1	M	Q3	Q1	M	Q3
50%	3	4	5	2	2	4	3	4	5	5	5	6	2	3	4
35%	2	2.5	4	2	2	4	3	4	5	5	6	6	3	4	4
20%	2	2	3.25	2	2	3.25	3	4	5	5	5	6	3	4	4
5%	1	2	2	2	2	2	3	4	5	5	6	6	3.75	4	4
Repetition	Q1	M	Q3	Q1	M	Q3	Q1	M	Q3	Q1	M	Q3	Q1	M	Q3
1	1	2	4	2	2	3	3	3	4	5	6	7	3	4	4
2	2	2	4	2	2	4	4	4	5	5	5	6	3	3	4
3	2	3	4	2	2	4	4	5	5	4	5	6	2	3	4

Table 4.1: Participants answers to questionnaires depending on error rate and repetition. First quartile (Q1), median (M) and third quartile (Q3) are provided. Instant frustration, global frustration, fatigue and motivation are rated on a 7-point Likert-scale. Effectiveness is rated on a 5-point Likert-scale.

Thus, the participants are consistent through all the experiment and the perceived effectiveness does not seem to depend on the repetition. Unsurprisingly, there is a strong negative correlation between error rate and perceived effectiveness of the BCI system.

Fatigue. The two-way ANOVA error rate and repetition against fatigue ratings only showed a main effect on repetition ($F_{2,22}$=41.3;$p<0.001$;η_p^2=0.79). Post-hoc tests showed that fatigue ratings significantly increase at each repetition.

Motivation. The two-way ANOVA error rate and repetition against motivation ratings showed a main effect on both error rate ($F_{3,33}$=3.4;$p<0.05$;η_p^2=0.23) and on repetition ($F_{2,22}$=6.7;$p<0.01$;η_p^2=0.36). No interaction effects were observed. Post-hoc tests showed a significant difference between error rate levels 5% and 50%, and between the first and third repetitions. Not surprisingly, motivation strongly drops after performing a block with 50% error rate, but also lowers at the end of the experiment.

The EEG signals of the participants were analyzed after the experiment in order to compute the real performance of the participants during the experiment. The signal processing technique is similar to [Legény et al., 2013] and to the one described in Chapter 3. We found an average classification rate of 50.4% ($SD = 16.8\%$), which is above chance level (33%) but still a bit lower than what has been observed in other SSVEP experiments with similar settings [Lalor et al., 2005, Legény et al., 2013, Middendorf et al., 2000]. Several technical factors could explain this drop in performance. The targets flickering frequencies were not controlled as carefully as they would have been for a "real" SSVEP experiment. Additionally, some participants might have tried to disrespect the instructions, if they were suspicious about the sham feedback. Interestingly, two subjects performed especially well (respectively 94.2% and 69.5%). One of them did not have any suspicion on the feedback, while the other noticed in the open questions: “To really know if the system is efficient, I would have to look into the wrong "square" to test if it always gives a failure, which obviously I could not do during the experiment”. It is very possible that the good results of these two participants were possible thanks to a good engagement in the task and respect of the instructions, while other participants might have been more distracted. It is also possible that the lack of an honest feedback prevented users from learning by trial-and-error how to get the right mental state.

4.3 Discussion

Our results show that errors of the BCI system are a primary source of user frustration, as revealed by the evolution of both global and instant frustration indicators. Similar results had been observed in other HCI contexts [Ceaparu et al., 2004], and our study confirms that BCIs are no exception. Even though a rather good user acceptance has been observed with BCI technology [Plass-Oude Bos et al., 2011], the tolerance towards errors is still limited, and user frustration increases over time. It is worth noting that our maximum error rate of 50% ends up with a median rating of *4*, i.e., a *light frustration* sensation. Thus the participants of our experiment did not express very strong feelings of frustration overall. This could be due to the fact that they were all highly motivated volunteers with a positive bias towards the experiment.

It is possible that the user expectation regarding accuracy influences the resulting frustration. Another possibility is that if the cost of errors had been higher, the frustration feeling could have been stronger. This aspect could be specifically targeted in future work. It is also worth noting that the difference of user frustration between error rates of 20% and 35% is non-significant. Thus, even if accuracy seems to be an important factor of user frustration, a small improvement of this accuracy (e.g. inferior to 5%) does not seem strongly influential. For the conception of BCI-based interaction systems, we would therefore advise to favor the comfort over the BCI accuracy, if the potential gain in accuracy is small, possibly even up to 15%.

Fatigue increases significantly over time, and seems to be less influenced by the error rate. The SSVEP context and the flickering targets could annoy some of the participants and generate visual fatigue after some time [Middendorf et al., 2000]. Besides, participants' motivation drops rapidly, and is also influenced by the error rate. This observation could be interpreted as a tendency for the participants to get discouraged quickly when confronted with a lot of errors, even if they do not feel more tired. Finally, *effectiveness* ratings are well correlated to the error rate, suggesting that our participants were able to keep track of their performance.

At the end of the experiment, participants could answer open questions about the BCI effectiveness and their potential frustration or fatigue. Many comments confirmed that the experiment is perceived as tiring, e.g. , “*fatigue comes late, but all of a sudden*". Some participants complained about the duration but also the repetitiveness of the experiment, e.g. , “*wish to sleep!*". Several noticed the link between failures and frustration: "*there is a bit of frustration sometimes when the system does not indicate the correct fixation, and it causes fatigue and loss of concentration*".

Beyond those results, we were able to observe the importance of engagement for user experience. In pre-experiments, we observed that participants did not feel frustrated by the results, because they did not care about it. They just felt bored. Frustration likely depends on the engagement in the task. In "real life", human-computer interaction is always directed towards a goal. However, in a strictly controlled experimental environment, the goal is artificial, and the users may not be motivated about it. In order to test more accurately user experience in real life, motivation has to be induced in the controlled experiment. One solution is to measure user experience over longer periods of time, while they accomplish a real-life task. But the complexity of a BCI setup, along with the difficulty of controlling external factors, makes this method difficult to apply. In our case, we were able to motivate participants to get good results by using a "carrot and stick" feedback, even though the success or failure was uncorrelated to the users' actions. One participant commented: "The progress bar gives the incentive to success and the (fail) noises are very frustrating".

The last two questions of our post-hoc questionnaire controlled if the participants had noticed the fake feedback. From their answers, it seems that five participants did not suspect anything at all, and had the impression that the responses were always consistent with their performance. In particular, the participant who performed the best in the post-experiment EEG analysis did not have any suspicion at all. The other participants had more nuanced answers and could have been more distracted. But nobody claimed openly that he/she was sure of facing a fake feedback. Besides, it seems that they all decided to play the game and did not try to check this during the experiment. The success of the sham feedback is encouraging, but has some drawbacks, as some participants had suspicions. Thus, we would only recommend using a sham feedback only if a full control on error rate is necessary in the experiment. Otherwise, biased feedback, as in [Alimardani et al., 2014] or in Chapter 3, can be an interesting alternative.

One peculiarity of errors in BCI is that the source is difficult to identify. Is the user mistaken, or did the interface failed to recognize a well-done task? While using a keyboard, you can hit the wrong key, resulting in the wrong letter being typed. In that case, the error comes from the user, even if you can blame the interface for not being easier to use. If the written letter does not correspond to the pressed key, the error comes from the interface. But when using a BCI, errors from the interface are much more common, and impossible to distinguish from user errors. We observe that participants were often assimilating a *failed* trial with an error of the system. For example, one participant answered "*there were times at which I fixed the correct square, but the system did not find it*". Thus, there seems to be a natural tendency to attribute success to oneself, and failure to the system.

4.4 Conclusion

This chapter presented a study on the influence of error rate on the perceived experience of BCI users. An empirical analysis was performed, evaluating the user experience of participants faced with various levels of error rate. To do so, we designed a *fake* SSVEP-based BCI experiment, in which the results of a target selection task were artificially simulated with different error rates. The error rate changed across the experiment, and questionnaires were used during the experiment to regularly asses the user experience.

Our results show that high error rates increases user frustration, and that this frustration accumulates over time. This observation is consistent with previous studies on other devices, and indicates that BCI user experience is similar to standard interfaces, on this aspect. However, we also observe that the increase in frustration does not seem critical for small variations of error rates. Thus, for the future of BCIs, we advise to favor user comfort over small improvements in accuracy. Besides, the use of a fake feedback and simulated experimental conditions was barely noticed by the participants and could inspire future studies on BCI user experience in tightly-controlled conditions.

Our findings confirmed that errors of command recognition can increase the user frustration. Additionally, we observed that a prolonged use of a BCI can be tiring, and decrease the motivation. For future work, we would like to study the influence of other potential factors such as the feedback delay, or the error cost. Then, we would like to design specific interaction techniques for BCI systems that would take into account the error rate and the cost of errors, in order to improve the BCI user experience. In particular, we believe error recovery systems could greatly reduce the impact of errors in the interface on users frustration.

Chapter 5

Optimizing training conditions for SSVEP-based BCIs

In a great number of BCI systems, a stimulation is presented to the user in order to observe and classify his/her cerebral response. This is notably the case for SSVEP-based BCI, which rely on external flashing or flickering stimuli. In such setups, the properties of the stimulation must be carefully considered. In the case of SSVEP for example, one has to choose: the size of the flickering targets, the distance between them, their flickering frequencies, and even their color or difference in color (using different colors for different targets). Target size and frequencies, as well as inter-target distance, were found to strongly influence the resulting classification accuracy of the BCI [Cao et al., 2012, NG et al., 2012, Regan, 1966].

In some cases, the characteristics of the stimulation vary between the training and the testing phases. SSVEP-based BCIs are sometimes trained in stimulation conditions which differ from the ones of the online use. For instance, in [Lalor et al., 2005], only one flickering target at a time is presented to the user during the training phase, while two different targets are presented during the test phase of the experiment. On top of that, the targets used during training are much larger than the online ones. In other cases, the properties of the stimulation are kept rigorously identical during both phases. For example, in [Legény et al., 2013], the flickering targets displayed during the training phase are identical to the targets used during the online BCI-based video-game. But very little is actually known regarding the potential influence of a difference in stimulation conditions between training and end use of exogenous BCI. In other words, is it preferable to train a BCI classifier in a stimulation context which is close to the end use context, i.e. the one in which the classifier will be used later?

It has already been observed that some SSVEP stimulation conditions, (thereafter referred to as "*easy condition*") could lead to a better classification accuracy [Cao et al., 2012, NG et al., 2012] (such as a larger distance between targets), whereas other stimulation characteristics (thereafter referred to as "*hard condition*") could lead to worse performance (resp. close distance). Therefore, is it preferable to train the BCI classifier in the *easy condition* or in the *hard condition*? In other words, is it better to "learn the easy way" or to "learn the hard way"? Limiting noise in the data with the *easy condition* could result in a better classifier. But on the other hand, training the system in the *hard condition* could result in a more robust classifier, as "he who can do more, can do less". Besides, in the Machine Learning community, it is generally considered that the training has to be done in conditions as close as possible to the end use context (from now on called "*end use condition*"). One last possibility is that mixing different conditions during the training, with both *hard condition* and *easy condition* trials, would result in a more robust classifier.

Thus, this chapter presents a study exploring which stimulation condition are optimal in order to train the a robust classifier in an SSVEP context. To sum up, four strategies for the optimal training conditions are possible: the *easy condition*, the *hard condition*, the *end use condition*, and the *mix condition*. We focused on two parameters of SSVEP visual stimulation: distance between targets, and difference in color between targets. Distance between targets is known to influence the accuracy of classification [NG et al., 2012]. With a short distance between targets, SSVEP responses for several frequencies are observable in cerebral activity, thus making classification harder [NG et al., 2012]. Visual selective attention is known to be affected by color, and this effect can be observed in the measurable cerebral activity [Hillyard and Anllo-Vento, 1998]. Thus, it is plausible that using targets of different colors might have an effect on SSVEP response.

The remaining of this chapter is organized as follows: first, the materials and method of the experiment are presented. Then, the signal processing for SSVEP classification is detailed. The results are then presented and analyzed. A discussion of the interpretation of these results follows, before the conclusion of this chapter.

5.1 Materials and method

We aim at exploring the influence of training conditions on the end classification accuracy. EEG signals were registered in several conditions of SSVEP stimulation. These datasets were used offline to train classifiers in various conditions, and to test their accuracy.

More precisely, we explored the effect of distance and difference in color between targets. SSVEP data were registered in different conditions, considering two factors. The first factor is the distance between targets, which can take three different values. The second factor is the difference in color between targets. In the "monochrome" condition, targets are all black with a white background, while in the "color" condition, they each have a different color, with the same white background. For each of the conditions of distance and color, one data set was recorded. This data set was partitioned into 6 folds. For each fold, a classifier was trained on $5/6^{th}$ of the original data set. The resulting classifier was then tested on the last $6th$, and on the data from all the other conditions. For each condition, the rate of correct classification is measured by averaging over the folds. We could then study the variations of accuracy of a classifier, depending on both the training conditions and the testing conditions.

Participants: 12 healthy participants volunteered to participate in the study: 10 men and 2 women, aged from 23 to 37 years old (average 28 years old, standard deviation 3.9 years). All of them happened to be right-handed, and had a normal or corrected vision. Before the experiment, each participant signed a written consent form, and filled a questionnaire collecting statistics about gender, dominant hand, age and sight. One participant reported not having followed the instructions. Thus, another participant passed the experiment to compensate for the discarded data.

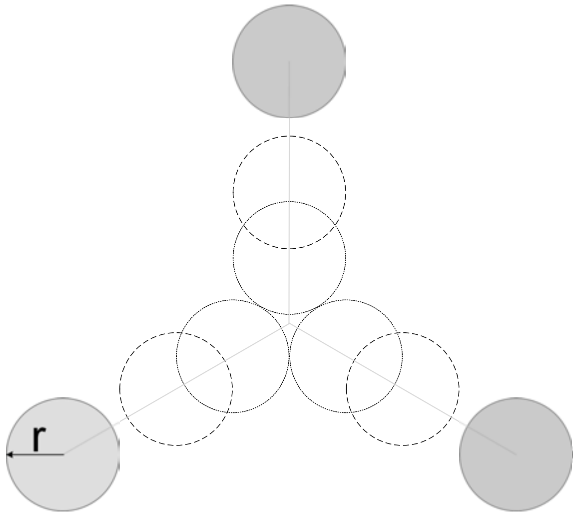


Figure 5.1: Different distances between targets (three possible distances: $2r$, $4r$, and $8r$).

Data acquisition: EEG data were recorded using 6 passive electrodes out of a 16-channel system (g.USBAmp, g.tec company, Austria), with a sampling rate of 512 Hz. Electrodes were placed according to the extended 10-20 system [Jasper, 1958], concentrated on the occipital lobe, in order to focus on the visual cortical areas. Electrodes were positioned on CPz, POz, Oz, Iz, O1 and O2. A reference electrode was located on the right ear, and an additional ground electrode was located on AFz. Channels were amplified and band-pass filtered between 2 and $60Hz$. A notch filter was applied to exclude frequencies between 48 and $52Hz$, corresponding to the power supply frequency band. Electrode impedance was checked to be below 1 kilo-ohm to ensure signal quality.

Design: Inter-stimuli distances and color conditions varied across the experiment, with 3 levels of distance (See Figure 5.2), and 2 levels of color. All the other characteristics of the stimuli were kept constant.

On a 60 Hz screen, it is possible to display frequencies at 8.57 Hz or below, and at 10, 12, 15, 20, 30 or 60 Hz. Frequencies lower than 10 Hz and above 25 Hz were excluded from this study because of their weaker SSVEP response [Herrmann, 2001, Pastor et al., 2003, Zhu et al., 2010b]. Finally, SSVEP brain response is observable at the harmonics of the stimulation frequencies. Thus, a 20 Hz frequency could interfere with the detection of a 10 Hz one. We chose to use 3 targets flickering respectively at 10, 12 and 15 Hz. The screen vertical synchronization was used to control each frequency.

In order to reach an optimal detection, stimuli should be at least 0.035 rad wide [NG et al., 2012]. On top of that, a large inter-target distance is preferable, up to 0.087 rad [NG et al., 2012]. The targets had to be large enough to allow a good target detection, but small enough to be displayed close to each other without overlapping. We chose to use targets with a diameter of 0.05 rad of visual angle. Considering that the participants would sit between 50 and 70 cm of the screen, 0.05 rad of visual angle corresponds to 3.1 cm on the screen. We used a DELL™Ultrasharp™2007FP 51 cm screen (20.1 inches), with a resolution of 1280 × 1024 pixels, and a refresh rate of 60 Hz, resulting in a diameter of 58 pixels (3.1 cm) for each stimulus.

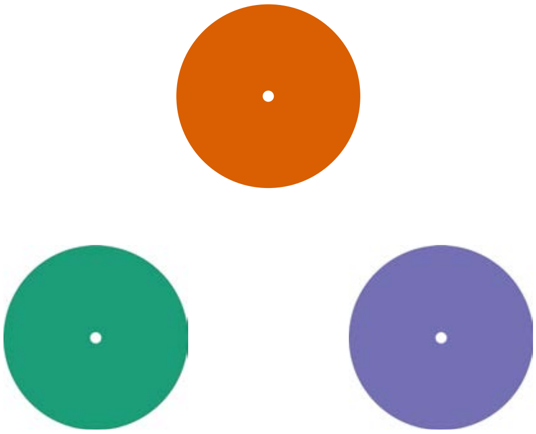


Figure 5.2: Targets shape, and their respective colors in the colored condition: left target is green, right target is blue, and upper target is orange.

The distance between targets varied during the experiment, with 3 possible distances: In the "close" condition of distance, the targets were touching each other, without overlapping (3.1 cm between the centers of 2 stimuli, exactly 58 pixels). In the "medium" condition of distance, the targets were twice as distant, at 6.3 cm of each other. Finally, in the "far" condition, the distance is again doubled, reaching 12.6 cm between the centers of each target (See Figure 5.2).

In the "monochrome" condition, the stimuli were flickering in pure black and white (RGB: 0, 0, 0 and RGB: 255, 255, 255), one at 10 Hz, one at 12 Hz, and the last one at 15 Hz. In the "colored" condition, the frequencies of the stimuli were the same, but they were each associated with a specific color, and flickered between white and this color. In order to exhibit a possible effect of a difference in color between targets, colors were chosen to be as different as possible from each other, in the sense of attention selection. The chosen colors should also present a good contrast with the white background (as luminosity contrast is known to be essential to SSVEP responses [Galloway, NR, 1990]), that is, a similar level of gray. The GeoVista center of Penn State University proposed several graphic palettes specially designed for attention differentiation [1]. Among those palettes, the final choice of colors was the one presenting the highest level of gray, and guaranteed to be colorblind-friendly [Gardner, 2005]. It was composed of a green (RGB: 27, 158, 119), an orange (RGB: 217, 95, 2), and a purple (RGB: 117, 112, 179) (see Figure 5.1). Some interactions between color and frequency of stimulation have been discovered [Cao et al., 2012, Regan, 1966], but they are not well understood yet. In this study, we chose not to focus on those aspects, in order to maintain a low number of factors, and thus keep a good statistical power. Therefore, we associated each frequency to a color, and kept it constant. This association was chosen in order to get the strongest SSVEP response, according to the present knowledge. The SSVEP response to a red stimulation is stronger at 12 Hz [Cao et al., 2012, Regan, 1966]. Thus, we associated the 12 Hz target with the color with the strongest red component (orange). Other colors have not been shown to have such a strong dependency toward the frequency of stimulation for SSVEP response [Regan, 1966]. The green was associated with the 10 Hz target, and the purple with the 15 Hz one.

Experimental procedure: Participants were seated in a comfortable chair, in front of a computer screen, on which flickering stimuli were displayed, while their cerebral activity was recorded.

[1] available on `http://colorbrewer2.org/`

The experiment was divided into eight blocks, separated by one minute breaks.

Before each trial, instructions were displayed on the screen for 3 seconds, indicating which of the 3 targets participants had to focus on. The targets then flickered during 7 seconds, and a simulated feedback on the selected target was then displayed.

In order to keep the attention of the participants, it is important to give them a feedback indicating what target has been recognized [Legény et al., 2011]. However, a training phase for the feedback could induce a bias toward the condition of feedback training. In order to avoid an augmentation of the number of conditions, a fake feedback was used. It has been shown that even a fake feedback can help to keep the attention of a participant during BCI tasks [Gonzalez-Franco et al., 2011]. Thus, after each trial, a fake feedback was displayed during 2 seconds, giving a positive response with a realistic accuracy (80%). This level of accuracy has been reached in pre-experiments, as well as in previous studies on SSVEP [Bi et al., 2013, Legény et al., 2011]. The users were not aware that the feedback was simulated.

Each block of acquisition contained 18 trials, and lasted 3 minutes and 36 seconds. The whole experiment was composed of 8 blocks. Between each block, participants were given at least a one minute break that could be extended upon participants will. All participants did all of the blocks (balanced design), with randomized order of conditions. The distance between the targets and the colors of the targets changed across blocks. Finally, the participants were given a post-experiment questionnaire. In particular, this questionnaire allowed us to check if the participants trusted the feedback. The total duration of the experiment was less than 45 minutes.

Each block was characterized by the distance between the targets D, and the color of the targets C. There were 3 distance levels: close (D_c), medium (D_m) and far (D_f), and 2 color levels: monochrome (C_m) and colored targets (C_c). Each of the 6 resulting joint conditions of D/C corresponded to one block. On top of these 6 blocks, 2 additional blocks mixed all conditions with equal representation. The order of the final 8 blocks was randomized for each participant. In every block, each frequency was targeted 6 times. Each block was divided into 6 sub-blocks in which every frequency was targeted once, in order to avoid any bias in the frequency distribution across the block. In the blocks that mixed the conditions, each combination of target, color and distance was presented exactly once.

In summary, the experimental design was 12 participants $\times$ 8 blocks (3 distances $\times$ 2 colors + 2 mixed) $\times$ 18 trials = 1,728 total trials.

5.2 Signal processing

The signal processing chain used for this study is the same as the one described in chapters 3 and 4, and followed classical methods for each step [Quan et al., 2013].

Feature extraction: A measure of the spectral density in each frequency of interest is computed as in [Legény et al., 2013]. Each channel was processed through a 0.5 Hz-wide band-pass filter at the 3 fundamental frequencies of stimulation (10, 12 and 15 Hz), and their first harmonic (respectively 20, 24 and 30 Hz). For each frequency of stimulation, the six channels are processed through two fourth order common spatial pattern (CSP) filters, in order to optimize the detection of this specific frequency [Legény et al., 2013]. The resulting filtered signals are then decomposed in 0.5 seconds moving windows, with 0.1 seconds moving steps. Let $S(f)$ be the signal filtered around frequency f. The energy spectral density is computed as the average of $S^2(f)$ over the time window, and a natural logarithm of this estimated density is computed and used as feature for the following classification algorithm.

Classification: A three-class LDA classifier was trained, combining 3 two-class LDA classifiers, each of them discriminating one class versus all the others. For each class i of stimulation (defined by its frequency), a two-class LDA classifier is learned, discriminating signals of class i against all the others. This classifier provides d_i an oriented distance to the hyperplane of separation. When used with two classes, such a classifier decides for class i if $0 \leq d_i$. In order to combine several two-class LDA classifiers to classify between more classes, the chosen class is the one maximizing $d_i - \sum_{j \neq i} d_j$. The resulting 3-class classifier takes a decision every 0.1 seconds, and thus has a relatively low recognition rate (71.8% on average during this experiment), since its decision is based on only 0.5 seconds of signal. In order to get one single classification for each 7-seconds trial, a voting step could be added, by deciding for the class the most consistently detected over the 7 seconds. Such a voting system allow to reach recognition rates close to 100% on 7-seconds trials. The resulting accuracies are strongly discrete, since the dataset were only 18 trials long. Thus, to avoid discretization of measurements, the following performance analysis uses the accuracy measured on 0.5-seconds windows.

5.3 Data analysis and results

The following analysis focuses on the effect of the difference in color and distance of the SSVEP stimuli on the classification performance. For each pair of blocks (characterized by its level of Distance and Color), we will refer to it as training block when its data is used to train the classifier and as test block when it is used to test the classifier. The matrix of performance P provides for each cell (x,y) (row, column) the average classification rate for the classifier trained using the x training block, and tested on the y test block (See Table 5.1). The accuracy for each cell has been computed by 6-fold trial-based cross-validation. For each block, the dataset is divided into 6 folds. The folding does not separate data from the same trial, in order to avoid overlapping in samples. It also ensures that classes are equally represented between folds. For each fold, $5/6^{th}$ of the dataset is used to train a classifier. This classifier is tested on the last 6^{th} (to get accuracy when $y = x$), and on the dataset from the other conditions (to get accuracy when $y \neq x$). Accuracies are then averaged over the folds. This method allows to compute the diagonal of P, without introducing a bias due to the length of the training set, and without recording each condition twice, which would have made the experiment twice as long, and would not have been acceptable for the participants, especially since SSVEP stimulation causes a strong visual fatigue.

Before the analysis, let us first introduce the notation used. Each block was characterized by the distance D between targets (close (D_c), medium (D_m) and far (D_f)) and their color C (monochrome (C_m) and colored targets (C_c)). We write each condition as "Distance condition/Color condition". For example, D/C_c represents the set of all blocks with colored targets, while D_m/C_c only represents the block with medium distance and colored targets. Mixed blocks were noted as Mix_1 and Mix_2 (Mix represents both blocks).

Regarding the overall performance rates, classification accuracy was 71.8%. As said in the previous section, this accuracy is measured on 0.5 seconds sliding windows for each trial (7 seconds).

condition	D_c/C_m	D_m/C_m	D_f/C_m	D_c/C_c	D_m/C_c	D_f/C_c	$Mix1$	$Mix2$
D_c/C_m	76.0	73.7	73.3	61.7	66.1	63.3	71.1	68.3
D_m/C_m	72.1	76.7	75.1	64.3	68.4	66.7	74.1	68.3
D_f/C_m	71.9	76.8	75.8	66.1	69.0	68.4	73.5	70.6
D_c/C_c	70.7	72.2	73.8	70.6	70.4	70.7	74.4	70.9
D_m/C_c	70.6	75.4	74.9	69.6	74.6	71.7	76.4	72.7
D_f/C_c	73.0	74.2	75.0	71.2	73.1	74.0	75.2	72.1
$Mix1$	72.9	76.3	75.2	68.5	71.3	69.5	76.3	71.7
$Mix2$	72.8	75.5	75.7	68.0	70.9	69.4	74.7	71.4

Table 5.1: Performance matrix averaged over the 12 participants. D_c, D_m and D_f are respectively the close, medium, and far distance levels. C_m and C_c are the monochrome and colored levels of color. Each cell contains the accuracy obtained with a classifier trained in the condition specified by the line, tested in the condition specified by the column.

5.3.1 Single conditions analysis

We performed a four-way ANOVA with a full factorial design considering the training ($x = \{D/C\}$) and testing ($y = \{D/C\}$) conditions. When needed, Tukey post-hoc pairwise-tests were performed (α=0.95). As the Mix condition cannot be attributed a level of distance or color, it was considered in a second analysis.

The four-way ANOVA only showed a significant main effect on test distance ($F_{2,22}$=5.3;$p < 0.05$; η_p^2=0.325). Post-hoc tests showed that $y = \{D_c/C\}$ resulted in a significantly lower accuracy rate ($\overline{x}$=69.8; σ=14.7) than $y = \{D_m/C\}$ ($\overline{x}$=72.5; σ=14.3).

Regarding higher interaction effects, there was a significant two-way interaction effect between training and testing distance conditions ($F_{4,44}$=5.9;$p < 0.01$;η_p^2=0.349), as well as between training and testing color conditions ($F_{1,11}$=26.63;$p < 0.001$;η_p^2=0.71) (see Figure 5.3). The interaction between training and testing distance conditions can be interpreted as that training with ($x = \{D_c/C\}$) provides similar performances for all ($y = \{D/C\}$) while training with ($x = \{D_m/C\}$) and ($x = \{D_f/C\}$) results in similar performances for ($y = \{D_c/C\}$) but increased performances with ($y = \{D_m/C\}$) and ($y = \{D_f/C\}$). Regarding training and testing color, post-hoc tests showed that training the classifier with monochrome targets ($x = \{D/C_m\}$) resulted in a significantly lower classification accuracy for colored targets ($y = \{D/C_c\}$). On average, classifiers trained on monochrome targets presented an accuracy 8.6% lower when tested on colored targets.

We also analyzed the effect of training and testing the classifier with the same conditions ($y = x$) and the opposite case ($y \neq x$). We performed a two-way ANOVA dataset (($y = x$), ($y \neq x$)) and training condition vs accuracy. The results show that the accuracy of a classifier trained in the same condition as the end use ($y = x$) have a significantly better average accuracy than the ones which are not ($y \neq x$) ($F_{1,11}$=43.50;$p < 0.001$;η_p^2=0.79). However, this effect is observed only when comparing the diagonal with all the training options at once, and does not remain significant when focusing on more specific conditions.

In particular, when comparing with the classifier trained on colored targets ($x = D/C_c$), we found no significant differences in accuracy. These results suggest that training a classifier in the end use condition gives good results, but that similar results could be obtained with another training condition (here with a classifier trained on colored targets at medium or far distance) and, as it will be presented later in section 6, such other classifier can be used for a wide range of end

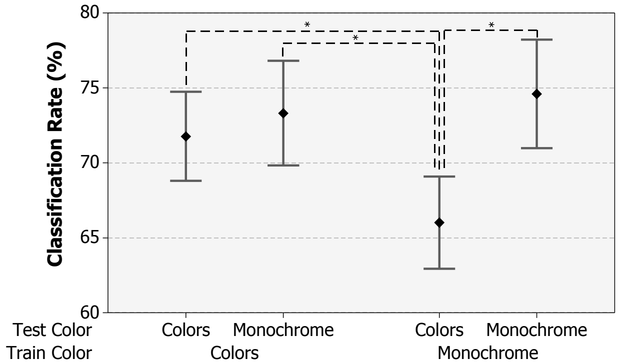


Figure 5.3: Classifier accuracy (CI 95%) depending on the training and testing color. Left: classifiers trained with colored targets tested on colored and monochrome targets. Right: classifiers trained with monochrome targets tested on colored and monochrome targets.

use conditions, meaning without the need for re-computing a classifier after every change in the end use condition.

5.3.2 Mix condition analysis

We tested the quality of the classifier when trained with the *Mix condition* x=$\{Mix\}$ and tested against the different combinations of color and distance y=$\{D/C\}$.

The two-way ANOVA test distance and test color vs accuracy showed a main effect on distance ($F_{2,22}$=6.19;$p<0.01$;η_p^2=0.36). Post-hoc tests showed that the mean accuracy of D_c ($\overline{x}$=70.55; σ=16.4) was significantly smaller than for D_m ($\overline{x}$=73.5; σ=15.22). We also observed a tendency on test color ($F_{1,11}=5.9$, $p>0.08$, $\eta_p^2=0.249$) which suggests that y=$\{D/C_m\}$ would result in higher classification rates.

In addition, we tested the accuracy of the classifier when training it with the different "single" conditions x=$\{D/C\}$ and tested against the *Mix conditions* y=$\{Mix\}$. The two-way ANOVA training distance and training color vs accuracy showed a strong tendency on training color ($F_{1,11}=4.62$, $p>0.055$, $\eta_p^2=0.296$) and a tendency on training distance ($F_{2,22}=2.88$, $p=0.078$, $\eta_p^2=0.207$). This results are in line with the results of the single condition analysis as the tendencies suggest that ($x=\{D/C_c\}$) would result in higher classification rates than ($x=\{D/C_m\}$), and ($x=\{D_c/C\}$) would result in lower classification rates than ($x=\{D_m/C\}$) and ($x=\{D_f/C\}$).

In short, in order to get a general-purpose SSVEP classifier able to adapt to various conditions, training a classifier on distant targets is more efficient than training them on short distance. This effect could be explained by the noise generated by the interference between targets. We also saw that training a classifier on colored targets seems to be more efficient, when the end use presents colored targets, and as efficient when the end use presents monochrome targets.

5.3.3 Subjective evaluation

Participants filled a questionnaire aiming at evaluating perceptual differences between the conditions of stimulation.

Concerning the detection of the fake feedback, a 7-point Likert-scale was used to rate the precision of the target detection and the control of the target detection. For all participants, with all conditions, the feedback had 80% accuracy. The precision of the target detection was rated 5.8 on average ($SD = 0.4$), and control of the detection was rated 5.3 ($SD = 1.2$). This high level of control of the detection indicates that most subjects did not suspect the fake feedback. One participant rated 5 for the detection and 2 for the control. In addition, this participant stated that *"the errors are surprising: sometimes no mistake while I had the impression of having a bad focus, and when there is a mistake, the detected target is not the one that was distracting me"*. This participant could have discovered the fake feedback. However, his data resulted in a good accuracy (71% of correct classification on average), without any drop of performance on specific blocks, as it should be the case for someone trying to not follow the instruction. The other participants all rated control of the detection at 5 or higher.

Concerning subjective preference between color and monochrome targets, opinions differ across the subject. They were invited to report any impression with their own words. Some participants preferred the colored targets, while others preferred the monochrome ones. For example, one participant stated: *"About black and white: I don't know if the precision is actually better, but it is undeniably more comfortable to use"*, while another one said the opposite: *"Overall, the flickering in black and white targets seemed more annoying to bear with"*. Most participants did not comment about the distance between targets. This suggests that the distance between targets did not have a meaningful influence on user comfort.

To sum up with, difference in color between targets seems to have contrasted effects depending on the participants, while the distance between targets has little influence on comfort. This observation should encourage future applications to use distant targets for both training phase and online mode whenever it is possible, since the resulting classification rate is higher. As for color, considering that using colored target during end use results in a similar classification rate, and that the comfort related to color seems to be participant-dependent, the choice of using colored or monochrome targets during the end use could be left to the user.

5.3.4 Results summary

We observed an effect of the distance between targets for the testing condition. When SSVEP targets are closer to each other, they are harder to classify. This result is consistent with previous research [NG et al., 2012]. Additionally, an interaction effect was observed between training distance condition and test distance condition. It can be interpreted as follows: A classifier trained at close distance is equally accurate for targets at close or far distance, while a classifier trained with targets at far distance is more accurate for an end use at far distance, and equally accurate at close distance. In short, the distance condition seems to follow a rule of "learning the easy way".

The color condition behaves quite differently. No significant effect of color on accuracy was found neither during the training phase, nor during the end use. However, an interaction effect was observed between the training color condition and the testing color condition. It can be interpreted as follows: a classifier trained with monochrome targets is efficient for an end use with monochrome targets, but is less accurate in the presence of colored targets. On the opposite, a classifier trained with targets of different colors is accurate during the end use for both monochrome and colored targets. In short, the color condition seems to follow a rule of "learning the hard way".

The main lesson, here, is that a classifier trained in the right conditions (medium or far distance, and different colors), can be used with a broader range of targets without any loss of accuracy compared to a specialized classifier trained on the exact same conditions as the end use.

5.4 Discussion

In machine learning, it is generally considered that it is preferable to train the classifier in conditions as close as possible to the end use ones. Since all parameters cannot be controlled, this general principle is most of the time restricted to the parameters considered to be the most important ones. For example, in an SSVEP context, the frequency of flickering is always conserved between training and end use. The core principle of SSVEP requires keeping such parameter constant. In addition, the targets color is generally conserved, as well as the distance between targets and the shape of those targets.

The same reasoning leads to mixing conditions during the training when the end use requires various conditions. However, our experiment suggests that this approach is not optimal, and that some conditions are more favorable for a robust training, even when extended to less favorable conditions.

We observed an interaction effect on the accuracy obtained depending on the color condition. In order to get a good accuracy when the end use context contains monochrome targets, both conditions of color during the training lead to similar performance. However, if the end context contains colored targets, training the classifier on monochrome targets becomes less efficient. Thus, if the end context is uncertain, we would recommend using targets of different colors for the training phase.

This effect can be explained, considering that adding color to the stimulation is equivalent to adding a noise to the data that does not change the class of each trial. Since the color does not change the frequency of stimulation on which the detection is based, the class actually stays unchanged. Thus, color might be a relevant class invariant. Previous work in machine learning showed that adding a relevant noise to training data can improve the generalization power of a classifier [Bishop, 1995, Zhao, Wenyi and Chellappa, Rama and Phillips, P Jonathon, 1999]. Understanding precisely how the neural structure of visual areas is related to this change when using targets of different colors still requires investigation. In particular, it would be interesting to explore whether the effects discovered in this study can be generalized to different sets and configurations of colors and frequencies of stimulation.

We observed an effect of the distance between targets during the end use. Targets separated by larger distances are easier to distinguish than targets close to each other. This is consistent with previous studies [NG et al., 2012] that showed that when several targets flicker at different frequencies in the visual field of the user, all frequencies are represented in the cerebral activity, and compete for representation. The attention of the user then selects one frequency to improve its representation. Despite this attention selection, discrimination between close targets is still harder to achieve. Distant targets influence less the representation of the desired target, since the visual receptive cells are mainly condensed in the center of the visual field. Thus, for the distance parameter, the *easy condition* is when the targets are far from each other, while the *hard condition* corresponds to the closest distance.

We observed an interaction effect between training and testing distance conditions. Overall, an increase of the distance between targets during the test condition leads to an improved classification. However, this improved performance is less visible when the classifier is trained with closer targets. In contrast, a classifier trained with further away targets presents a stronger classification improvement between closer and further away targets. In short, it seems that training on the *easy condition* can be better than training a classifier on the *hard condition*. Concerning inter-stimulus distance, learning the easy way is more efficient than learning the hard way.

Let us now consider the *end use condition*: from our observations, in some cases, learning on conditions as close as possible to the *end use condition* is efficient. However, it does not seem to be always the case. For example, when the end use targets are close to each other, classifiers trained with targets at far distances are as efficient as the ones trained at short distances, with the additional advantage of being better for far distance targets. Similarly, classifiers trained with colored targets are as good as the ones trained on monochrome targets when it comes to classify monochrome targets, while classifiers trained on monochrome targets are less accurate to classify colored targets.

The comparison to the *Mix condition* reveals that the *mix condition* is not more efficient than the *far condition*, which corresponds to the *easy condition*. To sum up with, the distance parameter follows the rule: learning the easy way.

Taken together, our results show that training a classifier in the end use condition is a relatively efficient approach. However, it is possible to choose the training conditions in order to ensure a robust accuracy for all end use conditions. Namely, training a classifier with targets at medium or far distance, and of different colors, leads to an accuracy similar to that of a training in the end use condition. In other words, by training a classifier in optimal conditions, the need for re-calibration when the end use conditions change is avoided. The classifier can be trained once and for all.

5.5 Conclusion

In this chapter, we studied the influence of several factors of stimulation during classifier calibration. These two factors were the distance between targets, which was known to have an effect on the end use accuracy, and the difference in color between targets. While stimulus colors had already been studied for the general case, the combination of several stimuli with different colors used at the same time, like they are presented in this chapter, had never been tested before.

Our results show that having similar conditions during the training of a classifier and its end use is not always better. It is actually possible to efficiently train a classifier on a specific condition and to generalize its use to different contexts. In some cases, it can even result in a better classification rate than when the classifier is trained in the end use context. In our SSVEP experiment, we found that distant targets are easier to distinguish, and are preferable as training conditions as well. Concerning the difference in color between targets, we found that targets of different colors are not significantly more difficult to distinguish. Last, we observed that using different colors for targets during the training phase results is a classifier efficient in a wider range of end use contexts.

While a lot remains unknown on the influence of training condition (other parameters, such as single color, frequency, and shape of targets, still require investigation), some guidelines can already be formulated, based on these observations. When the end SSVEP application allows it, we advise to use distant targets with different colors during the training phase, and distant targets for the end use, while the end use color has a low influence on accuracy, and could be let as a choice for the user.

Chapter 6

Toward hybrid interaction and fusion of brain and gaze inputs

Selection is one of the basic interaction tasks of the Foley decomposition [Foley et al., 1984]. Selection generally takes various forms: pushing a button, selecting an element in a menu, selecting an icon, etc. The most common interaction techniques for selection are based on the point-and-click paradigm. In HCI, the most common selection paradigm is the point-and-click. However, various interaction contexts require hands-free interaction. Among them are interaction for virtual reality, for mobile devices, or for disabled people. Traditional devices enabling point-and-click techniques (mouse, touchpad, etc) are not applicable to hands-free interaction. Other input devices then need to be used. Alternatives comprise gaze tracking [Zhu and Yang, 2002, Velichkovsky et al., 1997], speech recognition [Miniotas et al., 2006], and BCIs [Gürkök et al., 2011]. Hands-free interaction techniques can rely on a point-and-click paradigm, but in this specific context, the "clicking" is often as problematic as the "pointing" [Velichkovsky et al., 1997, Zander et al., 2010a]. Gaze tracking, speech recognition, and BCIs all share the particularity of presenting a relatively high error rate, compared to a keyboard or a mouse, for example.

Gaze-driven interfaces have been widely considered for hands-free interaction. But while controlling a cursor through gaze seems intuitive, it suffers from several limitations. The first limitation is physiological. The real gaze position oscillates quickly around the center of the fixation point, making it by essence limited to a precision around 1° of visual angle [Zhu and Yang, 2002, Kammerer et al., 2008]. The second limitation is technological. The tracking quality can vary greatly depending on many factors, such as the gaze tracking system, the user, the luminosity, etc. The third limitation lays in the interaction techniques. While gaze is relatively intuitive for pointing, it lacks a natural activation command (similar to the click of a computer mouse). The most used technique for selection using gaze tracking is the dwell time. A target is selected when the gaze position stays on the target for more than a certain time (the dwell time). However, dwell time is prone to false positives (i.e. unwanted selections) when the user looks at something that he did not want to select. This issue is usually referred to as the "Midas Touch problem" [Velichkovsky et al., 1997].

BCIs and gaze tracking show complementary advantages in the context of hands-free interaction: gaze tracking allows quickly defining a region where potential targets of interest can be selected, while SSVEP is suitable for selecting one target in a small set. Using two measures for related types of input should enable a better reliability of the resulting measure, by checking the consistency of both channels. Thus, BCIs could help improve the precision of gaze-based interaction, on top of providing a click command. However, there is surprisingly little previous work that tried to combine these two input modalities [Zander et al., 2010a].

In this chapter, we propose a new approach based on input fusion, designed for improving selection time and accuracy. This approach is illustrated by a fully functional interaction technique that was compared to the state of the art. We show that this fusion technique outperforms the previously existing sequential technique for BCI and gaze tracking hybrid selection.

The remainder of this chapter is organized as follows: first, the specific related work on target selection techniques based on gaze tracking, and on hybrid gaze tracking BCIs is provided. Second, our new approach for hybrid interaction based on input fusion of gaze tracking and BCIs is detailed. Then, a controlled experiment is described, in which the proposed technique is evaluated and compared with two previously existing selection techniques combining gaze and BCI inputs. Section 6.4 presents the results of the experiment followed by its discussion in Section 6.5. Finally, Section 6.6 provides the concluding remarks.

6.1 Related work on Hybrid Interaction Systems using Brain and Gaze inputs

This section presents related work on existing gaze- and SSVEP-based methods for target selection.

6.1.1 Gaze-based interaction

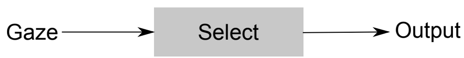
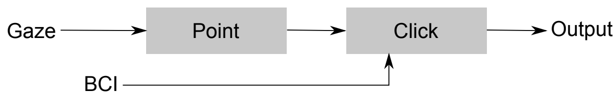
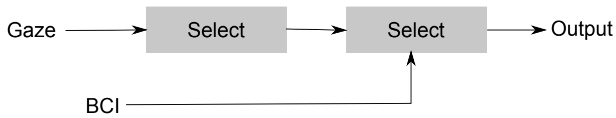
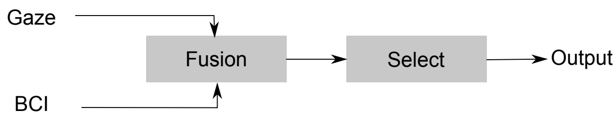
In order to improve dwell-based techniques, several methods have been proposed such as the *Fish-eye methods* [Ashmore et al., 2005]. Fish-eye methods magnify (zoom in) the area around the gaze position, thus decreasing the required selection precision, but without addressing the Midas touch problem. However, the omnipresence of the visual deformation can degrade the exploration of the graphical interface. A potential solution is to zoom in only when potential targets are available [Ashmore et al., 2005, Istance et al., 2008]. Another solution relies on designing user interfaces specifically suited for gaze-based selection such as hierarchical menus [Kammerer et al., 2008].

Most of the time, SSVEP-based interfaces are limited to a small number of targets (commonly three targets), although some attempts were successful at using more targets, in a synchronous context [Wang et al., 2010b, Chen et al., 2014a, Manyakov et al., 2013].

6.1.2 Gaze and EEG based hybrid interaction

The concept of *Hybrid BCI* was originally introduced in [Pfurtscheller et al., 2010] and it was defined as a system "composed of two BCIs, or at least one BCI and another system" that fulfills four criteria: "(i) the device must rely on signals recorded directly from the brain; (ii) there must be at least one recordable brain signal that the user can intentionally modulate to effect goal-directed behavior; (iii) real time processing; and (iv) the user must obtain feedback".

Table 6.1: Comparison of interaction techniques for gaze and/or BCI-based selection. A, B, C: State-of the art approaches. D: Our novel approach based on input fusion.

	Gaze-only approach	
A		Selection based on gaze only, usually with dwell time. [Zhu and Yang, 2002, Velichkovsky et al., 1997, Majaranta et al., 2009, Majaranta and Räihä, 2007]
	Hybrid Sequential approach (task attributes separation)	
B		Hybrid interaction with sequential processing: the gaze moves a cursor and the BCI selects the target [Zander et al., 2010a]
	Hybrid Sequential approach (two-step selection)	
C		Hybrid interaction with sequential selection. A first selection is done with the gaze only, to determine the set from which a second selection is performed with the BCI only, as in [Kos' Myna and Tarpin-Bernard, 2013] and [Choi et al., 2013].
	Hybrid Simultaneous approach (our approach based on input fusion)	
D		Hybrid interaction based on input fusion. Both inputs are combined to perform a single selection task.

In the past few years, it has been proposed to combine BCIs with a keyboard [Nijholt and Tan, 2008], a computer mouse [Mercier-Ganady et al., 2013], or a joystick [Leeb et al., 2013b]. Several types of BCIs can also be used at the same time [Fruitet et al., 2011, Li et al., 2010]. In [Gürkök et al., 2011], participants can switch at will between a SSVEP-based BCI and a speech recognition system. For a more complete review on hybrid BCIs, the interested reader can refer to [Pfurtscheller et al., 2010].

Hybrid interaction techniques can be broadly classified in two categories: sequential and simultaneous processing [Pfurtscheller et al., 2010] (see Table 6.1). Hybrid BCIs based on sequential processing use two or more inputs to accomplish two or more interaction tasks. Each input is then responsible for one task. Hybrid BCIs based on simultaneous processing can fuse several inputs in order to achieve a single interaction task [Müller-Putz et al., 2011].

Although the idea of combining BCI and gaze-tracking has been already proposed, it has been marginally explored. Existing works have mainly focused on P300 [Choi et al., 2013] and motor imagery [Zander et al., 2010a] BCIs. Regarding P300 paradigms, [Choi et al., 2013] combined gaze tracking with a P300-based BCI for a spelling application. Compared to a P300 speller, the number of accessible characters and the detection accuracy are improved . In contrast, Zander *et al.* proposed to control a 2D cursor with the gaze, and to emulate a mouse "click" with a motor-imagery based brain switch [Zander et al., 2010a]. They found that interaction using only gaze tracking was a bit faster, but that BCI-based click is a reasonable alternative to dwell time.

Later, [Kos' Myna and Tarpin-Bernard, 2013] proposed to use both gaze tracking and SSVEP-based BCI for a selection task in the context of a videogame. The gaze tracking allowed for a first selection task (selecting an object), followed by BCI-based selection for a second task (selecting a transformation to apply to the previously selected object). The findings of this study indicated that selection based only on gaze was faster and more intuitive.

So far, attempts at creating hybrid interfaces using EEG and gaze tracking inputs for target selection have focused on sequential techniques, and proposed ways to separate the selection into secondary tasks. While [Zander et al., 2010a] separates the task (selection attribute) into pointing and clicking, both [Choi et al., 2013] and [Kos' Myna and Tarpin-Bernard, 2013] use a two-step selection. We hypothesize that a simultaneous approach, based on the fusion of information from gaze-tracking and SSVEP-based BCI at a low level of abstraction, could better handle the uncertainty of inputs, and achieve better performance.

6.2 Combining gaze and BCI inputs for target selection

This section details the proposed gaze and SSVEP-based hybrid interaction technique that allows simultaneous processing, as defined in [Pfurtscheller et al., 2010]. Both inputs are combined at a low level of abstraction for a single selection task. It also presents two already-existing hands-free selection techniques that will be used as references for comparison. Our hypothesis is that the combination of both inputs will lead to better accuracy and higher speed than previous hybrid methods. We also hypothesized that this hybrid approach can outperform dwell time approaches for high density targets, as gaze-based interaction using dwell time is especially sensible to inter-target distance [Miniotas et al., 2006].

6.2.1 Novel approach for combining brain and gaze inputs: the fusion

6.2.1.A Concept

The general idea of our new approach is to combine gaze and EEG inputs at a lower level, in order to build a single, more precise, selection command (see Table 6.1). When the user wants to select a target, both the gaze position and the cerebral activity will be combined in order to estimate the desired target.

Any interaction system using target selection can be considered as an estimator of the fact F: "the user is trying to select *this* target". The main idea of the input fusion is to build a probabilistic model of F, while taking into account the specificity of the inputs uncertainty. For each input, a model of the distribution of errors is proposed. The resulting models are combined in order to build a higher level estimation of F's likelihood. This information is accumulated over time, until a target is selected when a certainty threshold has been reached. The goal is to allow the user to simply look at any target in order to select it while keeping false positives as low as possible.

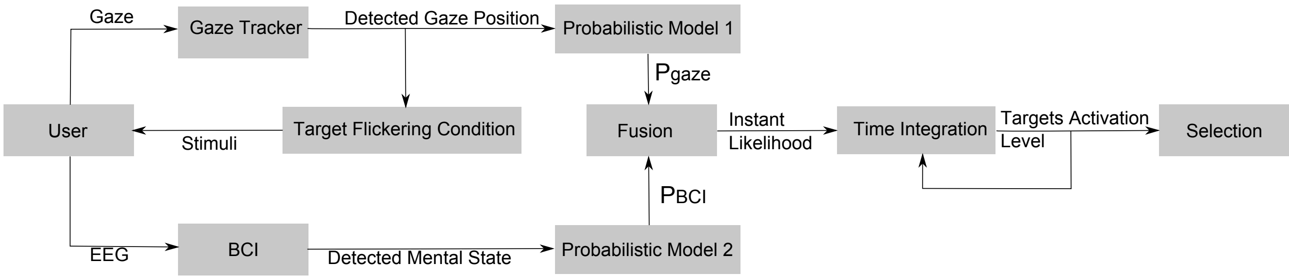


Figure 6.1: Schematic diagram of simultaneous processing for the proposed hybrid interaction technique based on the fusion of gaze and EEG inputs

6.2.1.B Main components

The components of the proposed hybrid interaction technique are depicted in Figure 6.1. The use of SSVEP-based BCI requires a visual stimulus to be associated with each target. In order to overcome the issue of the number of available frequencies of stimulation, a *target flickering condition* is used: only the targets close enough to the detected gaze position flicker.

For both the gaze and the BCI information, a probabilistic model estimates the probability for each target that the user is trying to select (Probabilistic models 1 and 2 on Figure 6.1). These two simple models of error can be combined, in order to fuse the inputs and build an estimate of the likelihood of F at any time. The resulting likelihood is integrated over time in order to reach a satisfying certainty level and to select the target. An activation level is associated to each potential target. For each time step, this activation level is increased by the current likelihood $P(F)$. Additionally, a decrease C of the activation level over time is needed, so that the activation level remains confined. A target is selected when the activation level of the target reaches a predefined threshold.

6.2.1.C Target flickering condition

In order to define what it means to be "close enough to the detected gaze position", the threshold of distance is fixed as a function of the standard deviation of the gaze detection accuracy σ. Let us build rules for deciding which target should flicker, by iteratively adding constraints to avoid unwanted behaviors.

- R_0, the rule of supplies limitation: *a maximum of 3 SSVEP targets can flicker at the same time.*
- R_1, the rule of comparative distance: *Targets closer to the detected gaze position have higher priority.*
- R_2, the "out-of-reach" rule: *Targets do not flicker if their distance to the detected gaze position is higher than* $3 * \sigma$.

The rule of supplies limitation (R_0) constraints the number of targets flickering at the same time, for a better comfort of use. R_1, the rule of comparative distance is a simple way of choosing which target should be flickering. However, it seems unnecessary to have a target flicker if it is very far from the gaze, in the case where the target density is low. Hence, the rule R_2 can be added.

Stopping at these rules would be enough to have a well-defined flickering condition, but an issue remains: if more than 3 targets are closer than the threshold, there would be a phenomenon where the targets that are flickering constantly change, making it extremely difficult for the user to focus on any of them. When two targets are almost equidistant to the gaze position, one might flicker, then the gaze move slightly, and the other takes priority. Then the gaze moves again just by a little, and the first one takes priority again, and so on. In order to avoid this phenomenon, R_1 is replaced by R_3 and R_4:

- R_3, the rule of immediate vicinity: *Targets which are closer than σ to the gaze position flickers, except when this rule conflicts with R_0, in which case priority rule R_1 applies.*
- R_4, the rule of preservation: *Targets that were already flickering before will keep flickering, as long as this rule does not conflict with rules R_0, R_2, and R_3.*

This set of rules ensures that there will never be more than 3 targets flickering at the same time, targets close to the gaze position will flicker, and a certain stability of the flickering targets is maintained. Flickering conditions are checked at the gaze tracking refresh rate of $60Hz$.

6.2.1.D Gaze precision model

Gaze precision is modeled with the idea that big errors in position detection are unlikely, while minor imprecision is perfectly plausible. Additionally, gaze tracking usually shows a better precision on the horizontal axis than on the vertical one. We thus choose to model gaze precision by P_{gaze} (see Equations 6.1 and 6.2), a two-dimensional Gaussian distribution, centered on the detected gaze position (x_g, y_g), where the horizontal (σ_1) and vertical σ_2 variances can be determined experimentally.

$$N(\sigma, \mu, x) = \frac{1}{\sigma\sqrt{2 * \pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{6.1}$$

$$P_{gaze}(x, y) = N(\sigma_1, x_g, x) * N(\sigma_2, y_g, y) \tag{6.2}$$

6.2.1.E BCI precision model

The BCI precision is modeled by P_{BCI}, a uniform distribution over all targets $(T_i)_{i\in\{1,...n\}}$, except for the target associated to the stimulus corresponding to the detected frequency f, which is given a higher probability p (probability that the classifier gives the correct class at any time). P_{BCI} is p for the target chosen by the SSVEP classifier, and the remaining $1 - p$ uniformly distributed across the other targets (see Equation 6.3).

$$P_{BCI}(T_i) = \begin{cases} p & \text{if } T_i \text{ flickers at frequency } f \\ \frac{1-p}{n-1} & \text{otherwise} \end{cases} \tag{6.3}$$

6.2.1.F Combining the probabilistic models for target selection

The final likelihood estimation is built by combining P_{gaze} and P_{BCI}:

$$P(F) = P_{gaze}(F) * P_{BCI}(F) \tag{6.4}$$

Finally, the target activation levels evolve as a function of $P(F)$. At each time step (every 100ms, following the SSVEP classification output), activation levels are increased by $P(F)$, and decremented by a constant C. This decrease over time ensures that targets can eventually return at rest. It is chosen to be linear in order to keep the number of parameters of the technique small. Formally, at each time step and for each target, the activation level is increased by $P_{BCI} * P_{gaze} - C$.

6.2.2 State-of-the-art techniques

For comparison purposes, we describe (and implement) two already existing techniques in the formalism of the probabilistic models described above.

6.2.2.A Gaze-only with dwell time

The most commonly used approach for hands-free selection relying on gaze tracking remains the dwell time [Majaranta et al., 2009, Majaranta and Räihä, 2007]. A target is selected when the gaze is detected to stay on the target for more than the *dwell time*. With the formalism of the fusion technique description (see Figure 6.1), it can be seen as the trivial case with no stimulation or BCI, and the fusion of inputs is the raw detection of gaze position (see Equations 6.5 and 6.6). For any target T_i, centered in (x, y):

$$P_{gaze}(x, y) = \begin{cases} 1 & \text{if } (x_g - x_i)^2 + (y_g - y_i)^2 < \sigma^2 \\ 0 & \text{otherwise} \end{cases} \tag{6.5}$$

$$P(F) = P_{gaze}(F) \tag{6.6}$$

The activation level of a target rises at a constant rate when the gaze is detected to be fixed on it, and go back to 0 when the gaze is away. A target is selected when a threshold of activation level (the *dwell time*) is reached.

For this experiment implementation, the detected gaze position is considered to be "on" the target when the distance between the detected gaze position and the center of the target is smaller than σ, the standard deviation of the gaze detection. In other words, the gaze position is treated as a disc of radius σ rather than a point. σ was measured in pre-experiments (see Section 6.3.2). It is also the radius of the feedback circle. The target activation level decreases linearly in time when the gaze is away, in order to keep the number of parameters small. The decrease rate was chosen by optimizing the system sensitivity in pre-experiments (see section 6.3.5). Refresh rate of the activation level was set to 20Hz.

6.2.2.B Hybrid approach based on sequential processing

State-of-the art techniques for hybrid gaze and BCI-based target selection (see section 6.1.2) use gaze to determine which targets are eligible, and the BCI to trigger the selection (see Table 6.1). For the sake of comparison, we propose here a reproduction of this approach.

In the formalism used to describe the fusion technique (see Figure 6.1), it can be seen as the case when the gaze is used to determine the target flickering condition, as in the fusion technique, but is ignored for the instant likelihood estimation (see Equation 6.7). The target flickering condition is the same as the one used for the fusion technique (see section 6.2.1.C).

$$P(F) = P_{BCI}(F) \tag{6.7}$$

If the user is focusing on one flickering stimulus, the SSVEP-based BCI will detect it. As for the fusion technique and the dwell time technique, an activation level is associated to each target, and is increased by $(P(F) - C)$ at each time step. The decrease rate C was chosen by optimizing the system sensitivity in pre-experiments (see section 6.3.5). Activation level is updated at each SSVEP classifier output, leading to a refresh rate of $10Hz$.

6.2.3 Signal processing

In this subsection we provide the underlining algorithms and methods used in signal processing and classification for both gaze and EEG signals in all three presented interaction techniques. The gaze signal processing is common to all three techniques, while EEG signal processing is applied for both hybrid techniques.

6.2.3.A Gaze signal processing

The raw coordinates output given by the gaze tracker is noisy. The resulting trajectory is discontinuous and irregular. In order to get a smooth trajectory, this trajectory is passed through a low-pass exponential filter which rate evolves depending on the variance of the position, as in [Casiez et al., 2012]. This method allows for rapid shift, while still stabilizing the noisy detection when the gaze is fixed.

6.2.3.B EEG signal processing

Feature extraction. Features of interest are extracted from the EEG data: the frequencies of interest were the three possible frequencies of stimulation: 10, 12 and 15Hz. A measure of the spectral density of the frequency of interest, as well as of its first harmonic, was computed as in [Legény et al., 2013]. The filtered signals for the 6 channels are processed through two fourth-order common spatial pattern (CSP) filters, in order to optimize the detection of these specific frequencies. The resulting filtered signals are then decomposed in 0.5 second moving windows, with 0.1 second moving steps. If $S(f)$ is the signal filtered around frequency f, the energy spectral density is computed as the average of $S^2(f)$ over the time window, and a natural logarithm of this estimated density is computed and used as feature for the following classification algorithm.

Classification: Classification is done as in [Legény et al., 2013]. A three-class LDA classifier is trained, combining 3 two-class LDA classifiers, each of them discriminating one class versus all the others. For each class i of stimulation (defined by its frequency), a two-class LDA classifier is learned, discriminating signals of class i against all the others. This classifier gives an oriented distance d_i to the hyperplane of separation. When used with two classes, such a classifier decides for class i if $0 \leq d_i$. In order to combine several two-class LDA classifiers to classify between more classes, the chosen class is the one maximizing d_i. Previous studies [Évain et al., 2016] have found that such a classifier shows a precision of the order of 65% for each 0.5 second time window. A better accuracy can be reached by adding a voting step. In this study, the voting step is included into the interaction technique (see section experiment procedure).

6.3 Experimental Evaluation

We performed an experiment evaluating the performances of the three techniques presented in the previous section: two previously existing techniques, and our novel fusion-based technique. During this experiment, participants were asked to perform a selection task using the three different techniques. We observed the sensitivity of each technique, together with their respective speed of selection.

6.3.1 Participants

Twelve participants took part in our experiment (3 women), all right-handed, aged between 22 and 43 (mean 28, SD 6.2). None of them had vision problems, none of them wore glasses. Two participants (not counted in the twelve) were excluded from the experiment because the eye-tracking system was not working properly, these participants were replaced.

Moreover, 5 additional participants took part in a first pre-experiment in order to assess the gaze tracking accuracy. Finally, 15 additional participants performed a second pre-experiment which aimed at choosing the optimal parameters for each of the three interaction techniques considered in the experiment (5 participants each).

6.3.2 Material and apparatus

Considering that this study presented no risk for either physical or mental health of participants, and that all the participants were healthy adults, this study was exempted from an ethic committee approval. Nevertheless, this study was carried out in accordance with the recommendations of the Declaration of Helsinki. Prior to the experiment, all participants were asked to read and fill a consent form stating their rights and the objective of the experiment. In particular, this consent form recalled that participation was not remunerated, that they had the right to withdraw without prejudice at any moment and without having to give a reason, that all collected data would be anonymized and used exclusively for research purposes.

Gaze data was measured using a Facelab 5.0 gaze tracker and a chin rest was used to maintain the proper positioning of participant's head (see Figure 6.2). The variance of the detection on both directions was measured in a pre-experiment on 5 participants. During pre-experiment, gaze

Figure 6.2: Experimental setup. 1: Visual display. 2: Gaze tracker. 3: Chin rest. 4: EEG headset. 5: EEG signal amplifier. 6: Laptop with OpenViBE software for EEG signal processing.

tracking was first calibrated. Participants were then asked to look at a fixed point, while the detected position was recorded. Distance between the eyes and the screen was about $70cm$. We measured a horizontal standard deviation of $\sigma_1 = 0.78cm$ ($0.64°$ of visual angle) and a vertical standard deviation of $\sigma_2 = 1.49cm$ ($1.22°$ of visual angle). Without separating the dimensions, the resulting standard deviation is $\sigma = 1.68cm$. These findings are well in the order of magnitude of the gaze tracker performances [1].

EEG data was acquired using 6 electrodes out of a 16-channel system (g.USBAmp, g.tec company, Austria), with a sampling rate of 512 Hz. Electrodes of interest were concentrated above the visual cortex, at position CPz, POz, Oz, Iz, O1, and O2 according to the extended 10-20 system. A reference electrode was located on the right ear, and an additional ground electrode was located on AFz. Channels were amplified and band-pass filtered between 2 and $60Hz$. A notch filter was applied to exclude frequencies between 48 and $52Hz$, corresponding to the power supply frequency band. Electrode impedance was checked to be below 1 kilo-ohm to ensure signal quality. Signal processing was done on OpenViBE running on a dedicated machine [Renard et al., 2010]. Visual display was provided by a DELL^{TM}Ultrasharp^{TM}2007FP 51 cm screen (20.1 inches), with a resolution of 1280×1024 pixels, and a refresh rate of 60 Hz.

For each participant, the system had to be calibrated. Since BCI calibration takes more time, and that gaze tracking shows various performance depending on the participant, gaze tracking was calibrated first. For gaze tracking calibration, participants were asked to keep their eyes on a point moving on the screen. Gaze tracking calibration took about 30 seconds. BCI was then calibrated. Three SSVEP targets were displayed on the screen. Targets had 3.2cm of diameter,

[1] `http://www.ekstremmakina.com/EKSTREM/product/facelab/index.html`

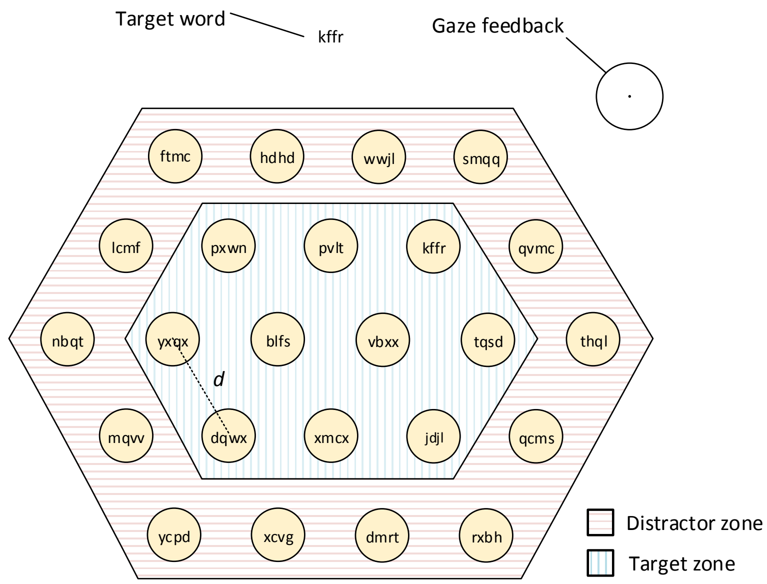


Figure 6.3: Design of the experimental task: The user has to look for the goal word displayed at the top of the screen. Then, the user has to select the target with the exact same word. The detected gaze position is displayed under the form of a circle and a central point (visual feedback). For all trials the size of the targets remained constant, and only the length of the target word and the separation (d) between targets varied. The targets at the outer circle were distractors in which the target word was never placed.

and were disposed as the corner of an equilateral triangle, 14cm from each other. Participants were instructed to look at one of them while their brain activity was registered. Targets flickered during 7 seconds, followed by a 4-second break during which the next target was indicated to the participant. The full BCI calibration took around 3 minutes. After the BCI calibration, gaze tracking calibration was again checked, in order to ensure that it was still valid. If not, gaze tracking could be re-calibrated.

6.3.3 Task

Once the calibration was done, the core of the experiment could start. For each trial, a *goal* word was displayed at the top of the screen. Several targets were displayed below on a hexagonal layout (see Figure 6.3). Random words were displayed on each target. These words were generated as in [Zander et al., 2010a], i.e. they were sequences of random letters uniformly distributed over the consonants. One of the target words matched the goal. Participants' task was to find the target with the *goal*, and to select it. The selection technique varied depending on the block, and was always explained to the participant beforehand.

When a target was close enough to the detected gaze position, a flickering stimulation was superposed on the target (except for the gaze only dwell time interaction technique). Stimuli were circles flickering between black and white at frequency 10, 12, and 15Hz. A fully opaque stimulation would hide the text behind it, but a good contrast of luminosity was needed for SSVEP detection. Thus, stimuli were given an opacity of 2/3. A maximum of three targets could flicker at the same time. Stimulation size was $3.6cm$ of diameter. For a participant seated at $70cm$ of the screen, this size corresponds to $3°$ of visual angle. It was found to be a good trade-off between the size of the stimulation and the SSVEP accuracy in [NG et al., 2012].

At any time, a feedback indicates the detected gaze position to the user, under the form of a circle. The radius of this circle corresponds to the measured standard deviation of gaze tracking accuracy ($1.6cm$). Additionally, the goal target was never on the outer layer, in order to avoid changes in the number of neighboring distractors. Participants were aware of this particularity.

Evaluation criteria

The performance of an interaction technique for selection is usually measured by Fitts' law. This law is a descriptive model of human movement. It predicts that the time required to rapidly move to a target area is a function of the ratio between the distance to the target and the width of the target. This model is well suited to measure pointing speed, and has thus been widely used for point-and-click selection techniques where the 'pointing' is critical, while the 'clicking' is not. However, when the error rate is more critical than the pointing time, alternatives to Fitts' law can be used as metrics.
For each selection task, three outcomes are possible:

- True Positive (TP): The participant succeeds in selecting the right target.
- False Positive (FP): The participant accidentally selects a distractor.
- Miss: The participant could not select anything after a time limit of 10 seconds. This time limit allows avoiding cases where the participant does not manage to select a target, and the trial lasts too long.

The hit-false rate (also called sensitivity, or d') is widely used as a metric of precision [Stillman, 1993, Verde et al., 2006]. This measure takes into account TP, FP, and Miss at the same time. d' is a measure of sensitivity defined as $d' = (\#TP - \#FP)/(\#TP + \#FP + \#Miss)$. d' is equivalent to counting 1 point for each success, -1 for each error, and 0 when no selection is done (the result being normalized by the number of trials). Most interaction contexts complete the property $A0$: *A false positive leads the user to select an undo command, and thus having one more command to issue before being able to try again.* For any interaction context under $A0$, d' is the expectancy of the number of effective commands issued by trial, as a function of the true positive rate, the false positive rate, and the miss rate. A sensitivity d' lower than 0 indicates that interaction is not possible in practice within assumption $A0$.

6.3.4 Experimental design

Three independent variables were considered in the experiment. First, the interaction technique which had three levels: gaze-only (see 6.2.2.A), sequential (see 6.2.2.B), and fusion (see 6.2.1). Second, the length of the target word. Words could be either 4 or 7 letters long (as in [Zander et al., 2010a]). Finally, the distance between targets, which had three levels: small, medium and long. Since the targets have a diameter of $2.68cm$, the three distances were:

- Short distance: $d_1 = 2.68cm$ apart. Targets are touching each other. Density is $d_1 = pi/2\sqrt{3} = 90.7\%$.
- Medium distance: $d_2 = \sqrt{2} * d_1 = 3.79cm$ apart. Density is $d_2 = d_1/2 = 45.3\%$.

- Long distance: $d_3 = \sqrt{2} * d_2 = 2 * d_1 = 5.35cm$ apart. Density is $d_3 = d_1/4 = 22.7\%$.

The experiment was divided into three blocks. For each block, a different interaction technique was used. The order of the interaction techniques was counterbalanced between the participants. For each block, 9 trials were performed for each factor combination, leading to a total of 81 trials per block, performed in a fully randomized order. The full duration of the experiment was about 45 minutes, including the breaks and calibration time.

Expectations were that longer goal words and short distances would have a negative impact on performance, resulting in more erroneous selections, and longer task completion times. We expected the gaze-only interaction technique to perform better than the hybrid approach based on sequential processing. The fusion technique was expected to perform better than the sequential technique, and possibly even better than the gaze-only.

6.3.5 Parameters optimization

For each of the interaction techniques considered (see Section 6.2), a different set of parameters have to be defined. As we could not find optimal configurations in the literature, in order to ensure an optimal configuration we performed a pre-experiment in which we tested several parameter configurations.

The optimized parameters were:

- The **selection threshold** T describes the activation level at which a target is effectively selected (see Section 6.2). In the case of the dwell time technique, this threshold is simply the dwell time.
- The **decrease rate** C describes the rate of linear decrease of the activation level, when the inputs do not point toward this target (parameter C for each technique in Section 6.2).

The task used for the optimization was the same as the one described in Section 6.3.3. An exploratory search of the parameters optimzing the sensitivity measure d' was performed. The "threshold" and "decrease" parameters varied across the experiment. For this pre-experiment, the word length was fixed to 7 letters, and the distance between targets was fixed to $3.79cm$.

A dynamic search was used to quickly find the optimal parameters. The goal is to find the pair (T_i, C_i) that maximizes d'. For each participant, tests were performed to compute d' for 9 pairs of parameters, combining 3 values of T, and 3 values of C.

- The first participant was tested on values in $\{T_0/2, T_0, 2 * T_0\} * \{C_0/2, C_0, 2 * C_0\}$. The resulting d' were registered.
- For each participant after that, tested values were $\{T_opt/2, T_opt, 2 * T_opt\} * \{C_opt/2, C_opt, 2 * C_opt\}$, with (T_opt, C_opt) being the pair of parameters that resulted in the highest d' on average on all the previous participants ((T_opt, C_opt) was updated after each participant).

This searching method allows converging to the best parameters order of magnitude, and refine the evaluation precision for the most likely optimum. Search is stopped when the current value of the pair (T_opt, C_opt) is based on at least 5 participants.

Method	Smallest possible activation time (sec)	Smallest possible deactivation time (sec)
Gaze-only	1	1
Sequential Hybrid	3.33	5
Fusion	0.4	1.3

Table 6.2: Results of pre-experiments and optimization: Thresholds of activation, and decrease activation level rate for all three interaction techniques. Sequential BCIs need to integrate information for a long time before taking an informed decision. At the opposite, the fusion hybrid integration can take a decision quickly if the inputs are consistent, but activation decreases faster.

6.3.5.A Pre-experiment results

The optimal parameters found during this pre-experiment can be interpreted by the resulting smallest possible time of target activation, and the smallest possible time of deactivation. Selection threshold T and decrease rate C are parameters unsuited for comparison, as they depend on the technique refresh rate r, and on the maximal activation level increment $maxI$ achievable at each time step, which depends on the probabilistic model. For the sake of interpretability, we provide the smallest possible activation time, computed as $T/(r*maxI)$, and the smallest possible deactivation time, computed as $T/(r*C)$. A high minimal time of activation, or a small time of deactivation, denotes that the optimization of parameters resulted in a rather conservative approach. The resulting optimal parameters found in this pre-experiment (see Table 6.2) were used for the main experiment.

Overall, the sequential technique is prone to generate false positives, because of the limited precision of the BCI, the raw accuracy of the BCI classifier p is estimated at 65%, based on previous studies with similar design and signal processing [Évain et al., 2016]. Thus, in order to optimize the sensitivity, the chosen parameters were quite conservative, with a high activation threshold. Information has to be integrated over a long time to allow selection. Dwell time was fixed at 1 second, which is consistent with previous research [Ashmore et al., 2005]. For our fusion technique, both inputs need to be consistent in order to get a significant rise in activation level. The parameters chosen by sensitivity optimization allow a very quick selection when these inputs are consistent. As a compensation to avoid too many false positives, the activation level decreases faster (short memory) than for the sequential technique when the inputs are not consistent.

6.4 Data Analysis and results

6.4.1 Task Performance

Results of the experiment are displayed in Table 6.3. We performed a three-way ANOVA considering as factors the interaction technique, the inter-targets distance and the task difficulty vs the sensitivity (d'). All factors being within-subjects (see Figure 6.4). All three factors were found to have a significant influence on d'. The ANOVA showed a main effect on all three factors: interaction technique $F_{2,22} = 34.45$, $p > 0.001$, $\eta_p^2 = 0.76$, inter-targets distance $F_{2,22} = 61.13$, $p > 0.001$, $\eta_p^2 = 0.85$, and task difficulty $F_{2,11} = 8.19$, $p > 0.05$, $\eta_p^2 = 0.43$. In addition, we also observed an interaction effect between technique and distance $F_{2,44} = 4.00$, $p > 0.01$, $\eta_p^2 = 0.27$.

Table 6.3: Success rates for the three interaction techniques, and the resulting sensitivity (see Annex for individual results).

Technique	**Correct**	**Miss**	**Error**	**Sensitivity** (d')
Gaze	82.5%	6.8%	10.7%	0.73
Sequential	20.8%	73.7%	5.6%	0.18
Fusion	55.5%	29.9%	14.6%	0.44

Post-hoc tests (Bonferroni, $\alpha < 0.05$) showed that the sensitivity for the fusion technique ($M = 0.44, SD = 0.37$) was significantly higher than for the sequential technique ($M = 0.18, SD = 0.24$), while remaining lower than the gaze-only technique ($M = 0.73, SD = 0.33$).

Regarding inter-targets distance, the close condition ($M = 0.20, SD = 0.36$) leaded to the lowest d' (compared to medium: ($M = 0.56, SD = 0.36$) and far: ($M = 0.59, SD = 0.32$)). Medium and far conditions were not significantly different. Finally, d' was significantly lower for 7-letters words ($M = 0.41, SD = 0.38$) than for 4-letters words ($M = 0.49, SD = 0.39$). Post-hoc tests did not show any conclusive interaction effect between technique and distance.

Additionally, Table 6.3 provides the breakdown for the trial outcome for each interaction technique. We observe that the decreased performance for the sequential technique is mainly due to the increased number of misses. Individual results are provided in annex D).

6.4.2 Selection Time

We conducted a three-way ANOVA analysis of the interaction technique, the inter-targets distance, and the task difficulty vs the selection time for successful trials. We found a main effect of the interaction technique $F_{2,22} = 110.10$, $p > 0.001$, $\eta_p^2 = 0.91$ and the task difficulty $F_{2,11} = 7.95$, $p > 0.05$, $\eta_p^2 = 0.42$. There was no main effect of distance and no significant interaction effects were found.

Post-hoc pairwise comparisons revealed that the selection time with the gaze-only technique

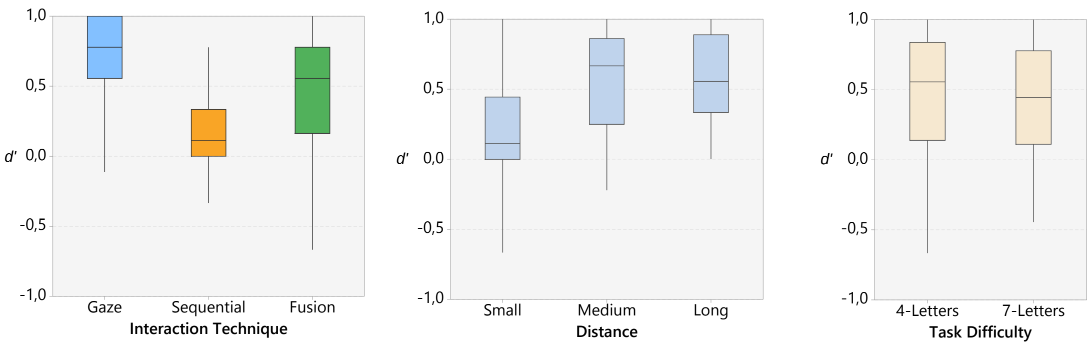


Figure 6.4: Boxplots for the sensitivity results, (left) technique, (center) distance and (right) task difficulty.

Technique	Neighbor Errors	Non-Neighbor Errors
Gaze-only	$M = 5.58s, SD = 1.59s,\ N = 59$	$M = 3.33s, SD = 1.53s,\ N = 10$
Sequential	$M = 7.45s, SD = 1.38s,\ N = 36$	None
Fusion	$M = 6.12s, SD = 2.06s,\ N = 44$	$M = 1.97s, SD = 0.98s,\ N = 40$

Table 6.4: Errors analysis: means and standard deviations of error timing according to the relationship between the correct target and the selected target.

($M = 5.45s, SD = 0.85s$) was significantly smaller than with the fusion technique ($M = 6.23s, SD = 1.10s$), itself smaller than with the sequential technique ($M = 8.19s, SD = 0.95s$). In addition 4-letter tasks leaded to significantly faster selections ($M = 6.35s, SD = 1.58s$) than 7-letter tasks ($M = 6.72s, SD = 1.33s$).

6.4.2.A Error Selection Time

The task for each trial can be subdivided in two sub-tasks. First, the user has to search the goal target (search sub-task) and second, the user has to issue the selection trigger (selection sub-task). Selection errors can occur during both of these phases. During the search sub-task, users can involuntarily select one target just by staring too much time on it (e.g. while reading it).These are typical Midas touch errors. In contrast, errors during the second sub-task come from precision limitations, and do not fall within the Midas touch. The user can wrongly select one target while he is trying to select another one.

False positive trials, when a target that is not the correct goal is selected, can be very instructive. In particular, we noticed that most of these errors occurred on targets that are neighbors of the goal. This particularity indicates that most errors occur while the participant has already found the right target, and is trying to select it. We decomposed false positive into two categories:

- Search phase errors: A target has been selected by mistake when the participant was searching for the correct target. This type of error should occur in the first few seconds of the trial, and can be on any target.
- Selection phase errors: A target has been selected by mistake when the participant had already found the correct target, and was trying to select it. Almost all of these errors should be selection of a neighbor target, and they can occur a long time after the trial start.

Figure 6.5 presents the mean selection time for the erroneous selections. We observe that, for the gaze-only technique, most errors (89%) occur during the selection sub-task. For the sequential technique, all errors resulted in the selection of a target in the neighborhood of the main target and happened after 5 seconds (see Table 6.4). This technique is very conservative, and prevents early errors. As a counterpart, a lot of trials end up with the time limit. Finally, for the fusion technique, we observe that the number of errors is evenly split between the search and the fusion sub-tasks.

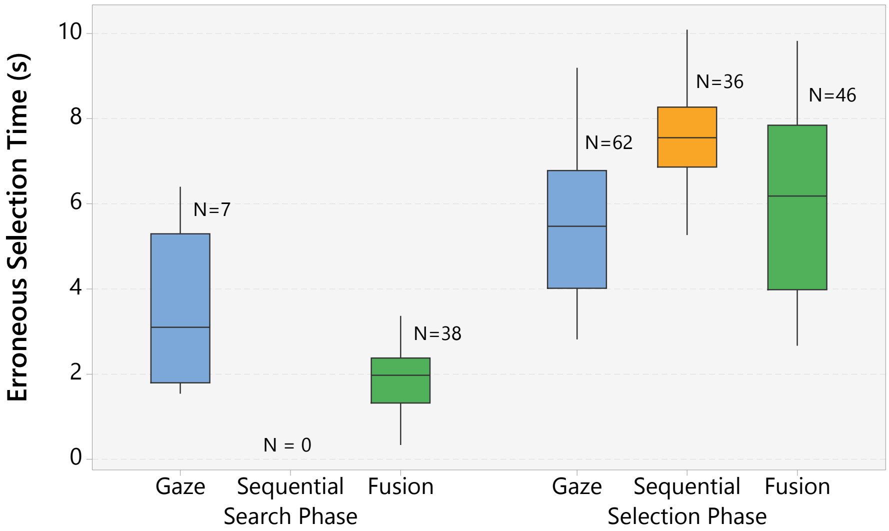


Figure 6.5: Boxplot of the mean error selection time grouped by selection technique and whether the selected target was a neighbor or not of the real target. The number of errors, and the individual values for each error are also provided in the plot.

6.4.3 Questionnaires results

Participants were asked to rate the fatigue, mental effort, effectiveness, feeling of control, and visual comfort associated to each technique using a 7-point Likert scale. We performed a Friedman test on the rating of fatigue, mental effort, effectiveness, control, and visual comfort, at $p < 0.05$ level. All criteria were significantly influenced by the interaction technique.

Post-hoc pairwise comparisons (Wilcoxon) revealed that that fatigue was higher with the sequential technique than with the gaze-only or fusion techniques (both $p < 0.05$). No significant differences appeared between gaze-only and fusion techniques. Mental effort was judged to be lower for the gaze-only technique than for the sequential and fusion techniques (all $p < 0.01$). No significant differences appeared between sequential and fusion techniques. Effectiveness was rated higher for the gaze-only technique, followed by fusion, and then sequential. All effects were significant ($p < 0.05$). Control was rated higher for the gaze-only technique, followed by fusion, and then sequential. All effects were significant ($p < 0.05$). Finally, visual comfort was rated higher for the gaze-only technique, followed by fusion, and then sequential. All effects were significant ($p < 0.05$).

Furthermore, participants were also asked to rank the interaction techniques by preference. All participants preferred our fusion technique over the sequential technique. Eleven of them ordered the techniques as $gaze - only > fusion > sequential$, while one chose $fusion > gaze - only > sequential$. Overall, participants showed a strong preference related to the performance of each technique. Namely, they preferred the gaze-only, followed by the fusion and then the sequential technique. Open comments indicated that this choice is influenced both by the accuracy and speed, and because the flickering stimulation is judged to be tiring by the participants.

6.5 Discussion

The novel approach that we proposed, based on input fusion, was found to be faster and more reliable than the approach based on sequential input processing. This observation is present for all tested levels of target density and task complexity. This demonstrates the feasibility of hybrid interfaces using gaze detection and SSVEP-based BCI simultaneously to achieve a single task.

Several factors can influence the accuracy and the selection speed. We observed that closer distance between targets leads to more selection errors. This effect was expected, as both gaze selection and SSVEP-based BCIs are known to be sensitive to the distance between targets. The difficulty of the search task, manipulated by the length of the *goal* word, was found to influence the selection time. This effect had already been observed in [Zander et al., 2010a]. However, *goal* complexity seems to have little influence on the system sensitivity, meaning that a high complexity task requires more time, but has a similar end result.

For hands-free selection techniques, usually, a trade-off has to be made between the difficulty of the selection and its accuracy. While a conservative interaction technique avoids unwanted selections, it increases the difficulty to select the desired target. At the opposite, other choices of interaction parameters may lead to easy selections, but can also be responsible for a lot of unwanted selections. Interaction parameters thus need to be set carefully. In this study, these parameters were chosen by optimizing a sensitivity measure d', focusing on task completion. This optimization ensures a balance between speed and robustness, and allows a fair comparison between interaction techniques. We observed that the sequential hybrid technique is best suited to gather information over a long period, and takes action only when enough information is gathered. By contrast, the fusion-based hybrid selection takes very quick decision when the inputs are consistent, and forgets quickly about old data if inputs stop pointing toward selection. Finally, the dwell time technique was found to be more efficient for all criteria. Interestingly, a study comparing interaction based on gaze only with a hybrid interaction based on both gaze and speech recognition found similar results: "Contrary to our expectations, an input device solely based on eye gazes turned out to be superior to the combined gaze- and speech-based device" [Kammerer et al., 2008]. In order to efficiently use a BCI input, we believe that the classification accuracy needs to be improved. With the current level of BCI reliability, a hybrid interaction technique would need to give a very low weight to the BCI input, and to favor the gaze tracking. However, even if a gain in accuracy can be obtained, the discomfort caused by the flickering might overcome the benefit.

In order to understand the current limits of the interaction techniques, we observed the characteristics of errors. The user task considered in this study can be decomposed into two phases. First, participants look for the *goal* target. During this phase, they may accidentally select unwanted targets (Midas Touch). When this first phase is over, users try to select a desired and well identified target. During this second phase, they may accidentally select another target. Typically a neighbor of the desired one, especially if the distance between targets is small. This problem is closely related to the gaze tracking accuracy limitation. We could observe that in practice, most errors occur during the second phase. This finding modulates the importance of the factors of interest. In particular, *goal* complexity influences the completion time of the first phase. We believe that this is why *goal* complexity has a strong influence on the overall task complexion time, but not on the system sensitivity. At the opposite, targets density does not influence the search time, but it is a critical factor for selection errors during the second phase.

6.6 Conclusion

This chapter proposed a new approach for hybrid brain-and-gaze interfaces, based on the fusion of inputs. This approach considers each input as a probability distribution, thus being able to handle different levels of precision, and combines them in order to build a single, more precise, estimator. We proposed a self-paced interaction technique based on this approach, enabling a hands-free selection among a high number of targets.

A user evaluation was conducted for assessing the performance benefit granted by our technique. Volunteers performed a hands-free selection task with two techniques directly inspired from previous work: gaze-only selection, and sequential hybrid selection, as well as with our novel technique. The performance of each technique was analyzed taking into account their accuracy, false positive rate, and speed. We found that our fusion technique was faster and more accurate than the previously existing hybrid techniques based on sequential processing. However, this improved speed and accuracy remains lower than those of interaction based on gaze only.

In order to outperform the techniques based on gaze only, future hybrid interfaces for target selection could be based on similar fusion approach, rather than on sequential selection techniques. In particular, progresses in signal processing can be directly included within our model. A possible trail for future improvements of hands-free selection techniques could be the use of SSVEP-based BCIs allowing a selection among a high number of targets, similar to [Wang et al., 2010b, Chen et al., 2014a, Manyakov et al., 2013]. However, user comfort when using such systems needs to be addressed. Additionally, these systems have yet to be tested in a self-paced context. One idea could be to use gaze tracking as a "double-check" modality, in order to correct SSVEP false positives. Considering our results in this study, we would advise the use of a fusion-based approach instead. Another way to limit false positives for a self-paced SSVEP detection is proposed in Annex A. This new signal processing chain is strongly inspired from [Chen et al., 2014a], but is adapted to a self-paced use.

Another trail of improvement is to explore other types of BCIs. Reactive BCIs (as SSVEP) can be uncomfortable, and, as this study shows, require improvements before becoming a useful addition to gaze-based interaction. Active BCIs require training from the user, and lack accuracy. Finally, passive BCIs could be used together with gaze tracking, in a hybrid interaction setting. Their potential contribution has yet to be explored.

Chapter 7

Conclusion

The work described in this manuscript, entitled "Optimizing the Use of SSVEP-based Brain-Computer Interfaces for Human-Computer Interaction", aims at improving SSVEP-based BCIs, from the point of view of HCI. Research on BCIs usually focuses on their speed and accuracy. However, in order to develop out-of-the-lab applications based on BCIs, the whole interaction process needs to be considered. We aimed at improving the interaction based on BCIs with three main objectives. First, we tried to better understand the BCI user experience. In particular, we studied the influence of error rate in the interface on user experience, and we explored how much cognitive resources are required in order to successfully drive a BCI. Second, we proposed improvements on BCI system performance, by studying the impact of calibration conditions on end performance. Finally, we studied interaction techniques for BCIs. In particular, we studied hybrid interaction, and proposed a method for fusion of gaze tracking and SSVEP-based BCI inputs. The main contributions of this PhD are described hereafter:

Chapter 3 studied the **cognitive demand** of an SSVEP-based BCI. Participants were asked to perform a dual task, consisting in selecting a given target with an SSVEP-based BCI, while performing a demanding memory task at the same time. We studied the impact of the memory task difficulty on SSVEP classification accuracy. We observed no significant impact. This observation was made both when visual or auditory channels were solicited by the memory task. We discussed these observations in term of **cognitive models**, in order to explain how SSVEP selection interacts with other cognitive processes. The findings of this study indicate that SSVEP selection is a low-demanding cognitive task. As previous work found, SSVEP selection is sensitive to the localization of visual attention, but this attention can be split between co-localized objects. A high demand on other cognitive resources such as the working memory or the auditory attention does not significantly affect the level of visual attention. For human-computer interaction, these findings are encouraging for the use of SSVEP, as they indicate that interaction based on SSVEP should not impair significantly the user attention on his/her main task. They are also encouraging for the use of SSVEP in hybrid interfaces, as cognitive resources can be allocated to other devices or GUI elements.

Chapter 4 studied the **frustration of SSVEP-based BCI users**, with respect to the error rate of the interface. Empirical analyses were performed, evaluating the user experience of participants faced with various levels of BCI error rate. To do so, we designed a *fake* SSVEP-based BCI experiment, in which the results of a target selection task were artificially simulated. The simulated error rate changed across the experiment, while the participants were made to believe they were the one controlling the interface. Questionnaires were used along the experiment, to assess the frustration and fatigue of the participants. Our results show that high error rates increase user frustration, and that this frustration accumulates over time. This observation is consistent with similar studies carried on for other interaction devices. However, we also observed that the increase in frustration does not seem critical for small differences of error rate.

Chapter 5 tackled the question of **stimulus characteristics for SSVEP classifier training**. For most BCIs systems, and in particular for most SSVEP-based BCIs, a classifier is used to identify a specific cerebral pattern. This classifier needs to be calibrated before use, in order to take into account the high inter-subject variability for BCIs. We performed an empirical experiment to test which type of SSVEP stimuli during calibration would lead to the most robust classification. Two stimulus characteristic were tested: the **distance between targets**, and the **difference of colors between targets**. We found that a high distance between targets leads to a better average classification accuracy. This was expected, as it is known that a short distance between targets leads to noise in the cerebral activity, as several flickering targets trigger cumulated responses. More surprisingly, we observed that using targets of different colors during the classifier training leads to a more consistently accurate classifier than using only black and white targets. These results indicate that having similar conditions during the training of a classifier and its end use is not always better. It is possible to efficiently train a classifier on a specific condition, and to generalize it to different contexts.

Chapter 6 focused on **Hybrid interaction and fusion of brain and gaze inputs**. We developed an interface based on both **gaze tracking and SSVEP-based BCI**, and proposed a new method based on **input fusion** to efficiently use these two inputs. Our method is **self-paced**, and allows the **selection of a single target among a large set.** Extended performance tests were performed, in which participants were asked to select targets using our interaction technique, as well as two other techniques directly inspired from the state of the art. Distance between targets varied across the experiment. The results showed that our fusion method out-performed the pre-existing sequential hybrid approach for all distances, both in term of speed and sensitivity. However, we also observed that interaction based on gaze-only, using the standard *dwell time* selection, was still even faster and more reliable than our fusion method. This result tends to indicate that despite the recent progresses in signal processing, SSVEP-based BCIs have yet to reach a level of competitiveness with alternative hands-free interfaces.

Additionally, in Annex, we presented a practical application: a BCI web browser. A state of the art signal processing method for SSVEP detection and classification was implemented and improved. We designed a prototype including a set of interaction techniques for the various tasks required by web browsing. This prototype is based on a hybrid BCI, combining P300 and SSVEP paradigms. Taken together, our results show the feasibility of a web browser fully based on BCIs, and somehow illustrate how HCI concepts apply to BCI applications.

Future work and perspectives

Several open questions are raised by the work presented in this manuscript. We discuss thereafter the main limitations of the techniques and studies that were presented, and propose possible short-term and long-term trails for improvement.

Cognitive demand of an SSVEP-based BCI (Chapter 3)

Studying different memory task. Our experiment proposed a memory task with various difficulties: from the control task when the only goal is to perform a selection with an SSVEP-based BCI, up to a dual task with the BCI-based selection tasks and a 2-back memory task. Our findings could be extended by a replication of our experiment enriched with a 3-back memory task. Additionally, compatibility of SSVEP use with different types of other cognitive tasks could be explored. In particular, influence of spatial memory use (as opposed to verbal memory) still remains unknown. Lastly, a split of visual attention localization could be tested. One hypothesis is that a co-localization of visual attention for SSVEP and the memory task is required to keep a satisfying level of accuracy.

Exploring the cognitive process behind SSVEP. Previous work concluded that SSVEP is sensitive to visual attention. According to the multiple resource theory, one would expect a decrease of accuracy of the SSVEP selection when another mental resource is used. However, multiple resource theory does not provide estimation of the amplitude on this effect. Further research is required to accurately measure the interaction between the various cognitive resources, and refine the 4-D model by giving it numerical prediction capabilities. The elaboration of such a model necessitates a high amount of empirical data. It could be achieved by collaborative research between experts of various field of expertise (neurology, psychology, signal processing, etc), combining the results of several studies.

User frustration of an SSVEP-based BCI (Chapter 4)

Comparing SSVEP-based BCIs to other types of BCIs and other devices. We observed that user frustration was rated generally low by BCI users. Is that a rating bias, or are BCIs errors generally less frustrating than those of other devices? A complementary experiment could be carried on, comparing SSVEP-based BCIs to other types of BCIs (P300, motor imagery) and to other interfaces (keyboard, joystick). In the case were such an experiment would effectively show that BCIs are less frustrating than other devices, a long-term experiment should explore if this is just a novelty effect, or a phenomenon that passes the test of time.

Exploring the influence of error cost. In our experiment, only the error rate was considered as a factor. However, we could observe that the feedback seemed important, because it gives an incentive to succeed the task. This observation suggests that the cost of the error might be of great influence on the user frustration. Cost of errors can vary depending on several factors. First the user engagement in the task gives a subjective importance to the success or failure of the task. This incentive to success could be artificially controlled by using various types of feedback, an absence of consequence of the task leading to a low incentive to success, while gamification methods could be used to emulate the user motivation [Amir and Ralph, 2014]. Additionally, the cost of errors depends on the time needed to correct them. In particular, the availability of an undo command could reduce the cost of errors.

Optimizing training conditions for SSVEP-based BCI (Chapter 5)

Extending results to different contexts. The influence of two parameters during SSVEP classifier training were explored: the distance between targets, and the difference in colors between targets. While our results are a first step toward a better understanding of training conditions influence, several other parameters could be explored. In particular, parameters that could have an effect comprise the number of targets, their size, shape, color, and flickering frequency. Systematic exploration of the influence of these parameters could lead to the definition of standard guidelines for SSVEP calibration design.

Generalizing training set characteristics effect. We studied the influence of stimulus characteristic for SSVEP classifier training. Similar studies could tackle the influence of various characteristics of the classifier training set. General rules for machine learning could be derived from these observations.

Toward Hybrid interaction and fusion of brain and gaze inputs (Chapter 6)

Improving the fusion. Our method showed improvement compared to the "usual" sequential hybrid approach, but still failed to out-perform the interface based on gaze only. Several trails for improvement could help achieve this goal. First, the weight of the brain and gaze inputs can be changed, possibly dynamically, in order to give more importance to the more reliable input. Another possible improvement lies on the probability models used for both the gaze tracking and the SSVEP detection. Other models of errors could be tested. In particular, the use of a priori knowledge on the user intention (what target is likely to be selected next, considering the current context), could help reduce false prediction, based on a Bayesian method. Finally, the probability models could evolve during use, based on unsupervised learning methods, in order to correct classification bias, and adapt to the uncertainty of each input.

Adapting to other BCIs. One explanation for the fact that our hybrid approach does not manage to outperform gaze alone could be that the SSVEP stimulation itself makes the gaze tracking less precise. Should that be the case, applying the same fusion method to motor imagery-based BCIs could be tested. As a counterpart, the cognitive demand for the user could be higher. Alternatively, a passive BCI could potentially be used, provided that it can detect a mental state characteristic of the user intention to select. Even with a low precision, such a detection could induce a positive bias to gaze selection.

The variety of the perspectives opened by the contributions of this PhD illustrate the magnitude of the challenge of bringing BCIs out of laboratories, and into the real world of usable applications. There are little doubts that the signal processing and classification methods for EEG can still improve greatly into the near future, along with progress in neuroscience and machine learning. Before BCIs can play a major role in the general public daily life, a lot of innovations focusing on the interaction as a whole are still required.

Publications

The work performed during this PhD lead to several publications (one book chapter, one conference paper, and two journal papers):

- ÉVAIN, Andéol, ROUSSEL, Nicolas, CASIEZ, Géry, ARGELAGUET, Ferran, and LÉCUYER, Anatole. Brain-computer interfaces for human-computer interaction. In: *Brain-Computer Interfaces 1: Methods and Perspectives*, Edited by CLERC, Maureen, BOUGRAIN, Laurent, and LOTTE, Fabien.
- EVAIN, Andéol, ARGELAGUET, Ferran, CASIEZ, Géry, ROUSSEL, Nicolas, and LÉCUYER (2016). Design and evaluation of fusion approach for combining brain and gaze inputs for target selection. *Frontiers in Neurosciences (section neuroprosthetics)*, vol. 10, p. 454-468.
- ÉVAIN, Andéol, ARGELAGUET, Ferran, CASIEZ, Géry, ROUSSEL, Nicolas, and LÉCUYER, Anatole. Do the stimuli of an SSVEP-based BCI really have to be the same as the stimuli used for training it?. *Brain-Computer Interfaces*, 2016, vol. 3, no 2, p. 103-111.
- ÉVAIN, Andéol, ARGELAGUET, Ferran, STROCK, Anthony, ROUSSEL, Nicolas, CASIEZ, Géry, and LÉCUYER, Anatole. Influence of Error Rate on Frustration of BCI Users. In: *Proceedings of the International Working Conference on Advanced Visual Interfaces.* ACM, 2016. p. 248-251.
- ÉVAIN, Andéol, ARGELAGUET, Ferran, CASIEZ, Géry, ROUSSEL, Nicolas, and LÉCUYER, Anatole. Can I Think of Something Else when Using a BCI? Cognitive Demands of an SSVEP-based BCI. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, 2017.

Annexes

Appendix A

An EEG signal processing method for self-paced and calibration-free SSVEP detection

In this first annex, we discus the signal processing methods for SSVEP detection and classification. Starting with a state-of-the-art method for classification, an iterative approach is given to provide a signal processing method that is calibration-free, self paced, and present a better sensitivity than the state of the art.

This annex is organized as follows: first, the most common signal processing method for SSVEP classification is presented. Leads for improvements are discussed and tested. Then, the benefits and counter-parts of a calibration phase are discussed. Off-line tests are performed in order to assess these benefits. Methods to reduce the amount of false positives during SSVEP detection are then discussed. An already-existing method based on alpha wave detection is presented, and improvements of this method are proposed and tested. Once combined, these signal processing and classification techniques create a fully self-paced and calibration-free method for SSVEP detection, with a better sensitivity than state-of-the-art methods.

A.1 CCA-based SSVEP detection for self-paced use

The best classification accuracy for SSVEP signal processing in the literature has been achieved by Canonical Correlation Analysis (CCA, see Section 1.4.3.B). CCA [Hotelling, 1936] is a multivariable statistical method. It is an extension of classical correlation enabling its use on vectors of random variables instead of random variables (see Figure A.1). CCA computes the best linear combinations of variables in each vector to maximize the correlations between those combinations. This method is frequently used method in data mining to recognize a known pattern. Let $X = (X_1, ..., X_n)$ and $Y = (Y_1, ..., Y_n)$ be two vectors of random variables, CCA computes the two vectors a and b that maximize $corr(a'X, b'Y)$, by solving the Equation A.1.

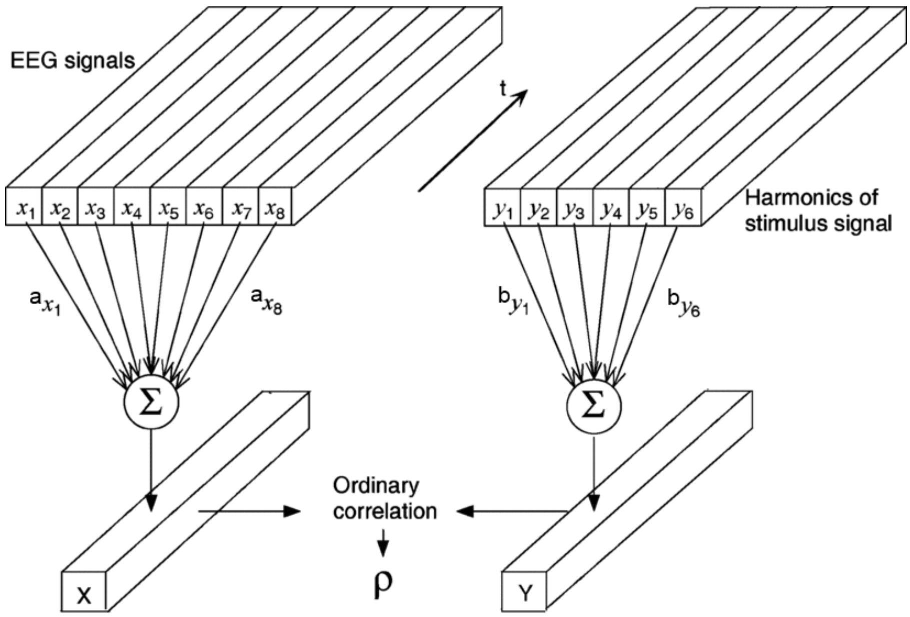


Figure A.1: CCA use for EEG signal analysis [Friman et al., 2001]. The signals x_1, x_2, etc. is are compared to the reference signals y_1, y_2, etc. (typically the harmonics of the stimulus signal for SSVEP classification), producing a single correlation measure p.

$$\max_{a,b} \rho(a'X, b'Y) = \frac{a'cov(X,Y)b}{\sqrt{a'cov(X,X)a}\sqrt{b'cov(Y,Y)b}} \tag{A.1}$$

For SSVEP detection, the reference signals are frequently sine and cosine of the tested frequency. This algorithm can be considered as a black box providing, for each frequency of stimulation f_i, a correlation $p(f_i)$ that can be consider as an index of likelihood for SSVEP response (a high value of $p(f_i)$) indicates that the user is probably trying to select the target associated to frequency f_i).

For synchronized SSVEP detection, a simple decision method is to select the target i with the highest $p(f_i)$. It has been used with a good precision, compared to alternative detection methods [Bin et al., 2009], and has since then grown increasingly popular. This method is suited for segmentation-paced context. This pipeline is described in Figure A.2, with X the input signal, Y the reference signal, and

$$Y_f = \begin{pmatrix} sin(2\pi ft) \\ cos(2\pi ft) \\ \bullet \\ \bullet \\ sin(2N_H\pi ft) \\ cos(2N_H\pi ft) \end{pmatrix}, \text{ and } \rho_k = CCA(X, Y_{f_k}).$$

For a self-paced use, it must be possible to "not select". The decision of SSVEP selection can be taken when $p(f_i)$ goes past a pre-defined threshold T_i. While segmentation-paced detection can

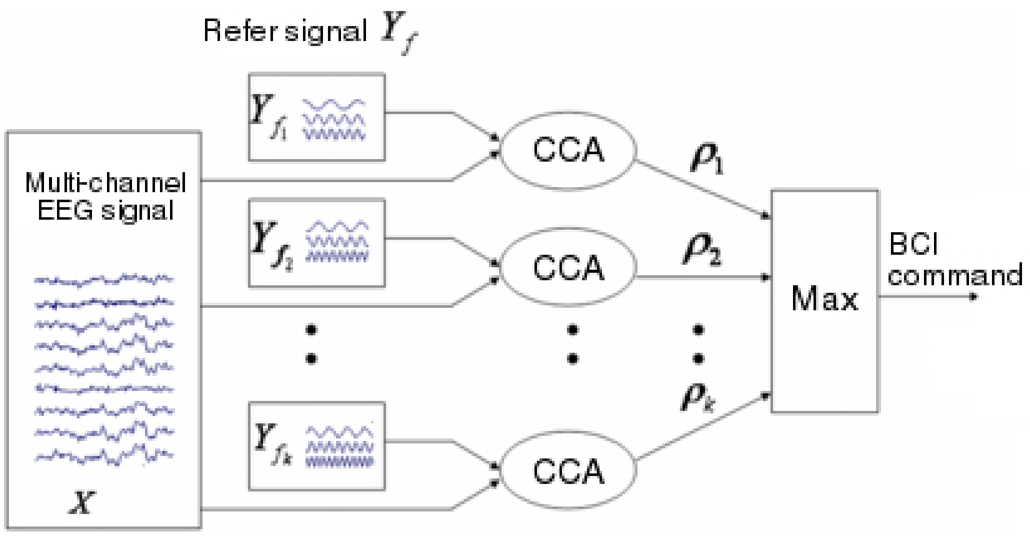


Figure A.2: Description of CCA-based pipeline for segmentation-paced classification, with with k targets, as described by [Bin et al., 2009]

be performed without calibration, using CCA in a self-paced context requires setting the values of the thresholds T_i. Adding this calibration step for CCA-based signal processing has been proposed for example in [Zhang et al., 2014] or [Wang et al., 2010a]. Several alternative methods could be explored. Instead of using a threshold on the raw value of the correlation for each frequency, other measures could be used, such as $T_i - average(T_j)$ (distance to mean), or $T_i/average(T_j)$ (ration to mean).

A.2 Off-line tests for discriminative features

We performed off-line tests to assess the interest of these features, regarding the ability to discriminate stimulation from idle state.

Our quality measure is the percentage of non-overlapping parameter's values between the idle and activity class for the whole record, once extreme values are taken of. It is computed as follows. Let A be the stimulated values set and I be the non-stimulated values set. Let D_Q be Q's first decile and D_I be I's last decile. Then the quality measure M is defined by:

$$M = \frac{\#\{x \in A | x > D_I\} + \#\{x \in I | x < D_Q\}}{\#A + \#I}.$$

The resulting differentiability measures are presented in figure A.3, for various discrimination features. Upper and lower borders of these boxes represent the first and last quartiles of statistical series. Red lines represent medians and red crosses represent outliers.

These tests encourage the use of the *Max Value* as feature for SSVEP detection, as in the literature. Thus, based on CCA values p_i, the decision method based on the *Max Value* feature is

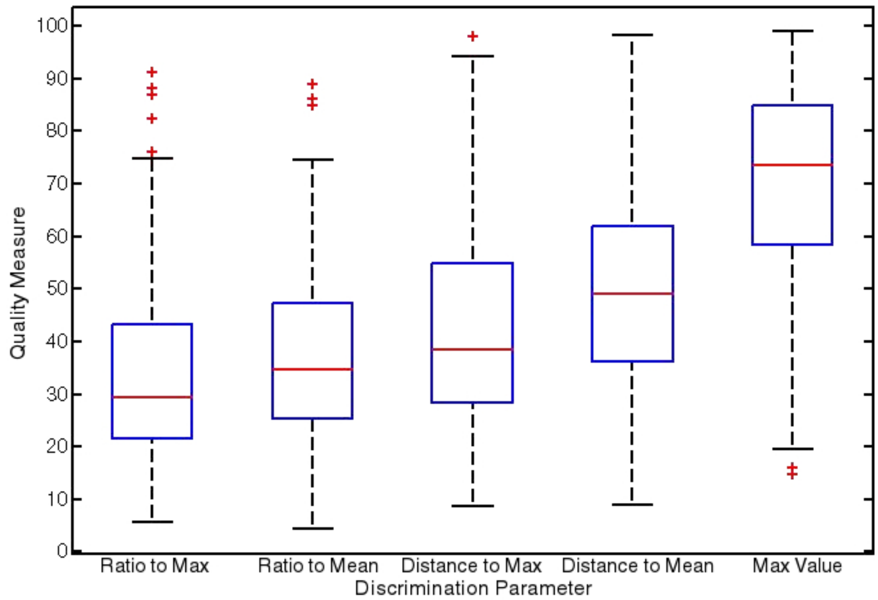


Figure A.3: Box plot of the quality measure for several discrimination features

defined solely by the values of T_i. Tests on the same data revealed the optimal T_i values do not change much depending on i. Thus, a robust heuristic (in order to avoid over-fitting) would be to keep a single T for all SSVEP targets i.

To sum up these tests results, we empirically tested the performances of various methods for detecting SSVEP, in a context allowing idle state, and found the optimal method to be the one based on a single threshold on correlation value, with the same threshold used regardless of frequency. These tests validate the method most widely used in the literature.

A.3 Calibration-based vs calibration-free approaches

Initially, CCA was proposed as a method that does not require a calibration [Bin et al., 2009]. Later on, it was proposed to enhance its accuracy by adding a calibration step [Zhang et al., 2014, Wang et al., 2010a]. It should be considered that calibration approach usually has some drawbacks. The calibration phase itself is time-consuming. Calibration may be unsuccessful, started over from the beginning. Finally, calibrated parameters can shift over time, requiring frequent recalibration. As a counterpart, calibration is expected to improve the detection accuracy. This gain should be clear enough to overcome the drawbacks. We decided to perform tests in order to evaluate the gain provided by the calibration step.

With a calibration, the optimistic hypothesis would be that the optimal value of T can be found for each subject (personalized optimization). By contrast, without calibration, the same value of T has to be used for all subjects (value by default, optimized on a set of subjects).

We performed empirical tests on the data gathered during the experiment described in chapter

3. An optimal threshold is computed on the data from half of the participants. Then, sensitivity obtained with this common threshold are compared to the one obtained with a personalized threshold. The results are showed on figure A.4.

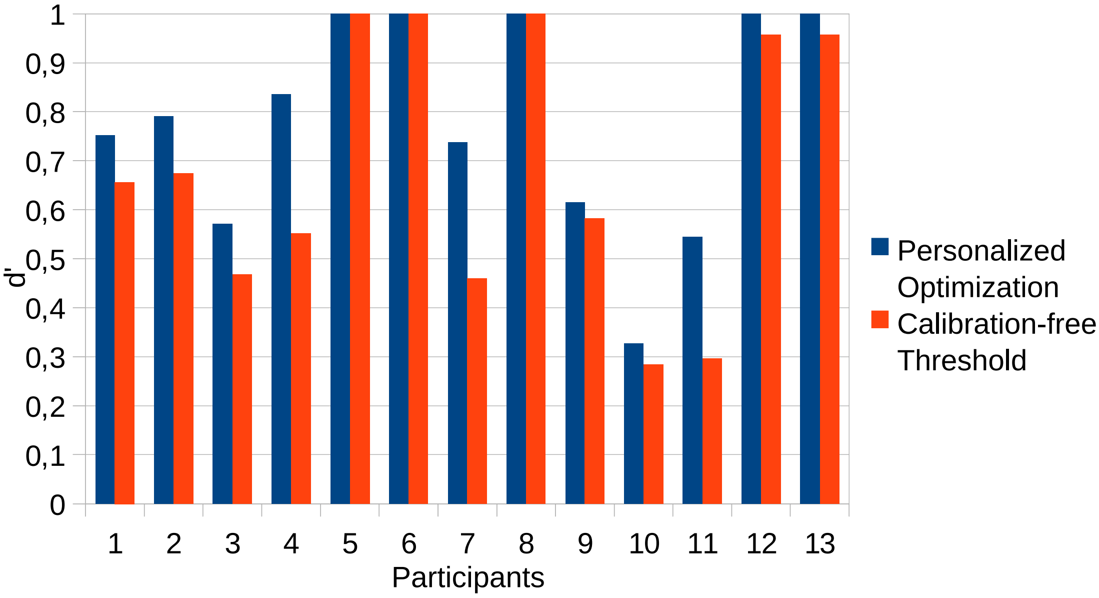


Figure A.4: Normalized sensitivity measure d' for 13 participants, computed with a personalized optimized threshold (blue), and with a common pre-determined threshold (red).

By nature, the personalized optimization can only lead to equal or higher results. However, tests revealed that the difference between the two methods is small A.4. Thus, the advantage provided by the calibration is limited, and the calibration is not necessary.

A.4 Using Alpha wave detection for CCA false positive correction

A high proportion of the false positives encountered when using CCA signal processing are due to alpha rhythm. Alpha rhythm is a cerebral rhythm that typically has a frequency of 10Hz, and diffuse from the occipital lobe. This makes it easily missmatched with an SSVEP response to a 10Hz stimuli.

However, SSVEP responses occur not only at the fundamental frequency of stimulation, but also at its harmonics. By contrast, alpha rhythm presents no harmonics activity [Wang et al., 2010a]. This phenomenon can be exploited to reduce the false positive, as proposed by [Wang et al., 2010a].

When the CCA signal processing chain detects the 10 Hz SSVEP response, two cases are considered:

- The 10Hz detection was just alpha wave. The detection is considered to be a false positive due to alpha rhythm and discarded.
- The 10Hz detection was not an alpha wave. The detection is validated.

This decision process is summarized in figure A.5.

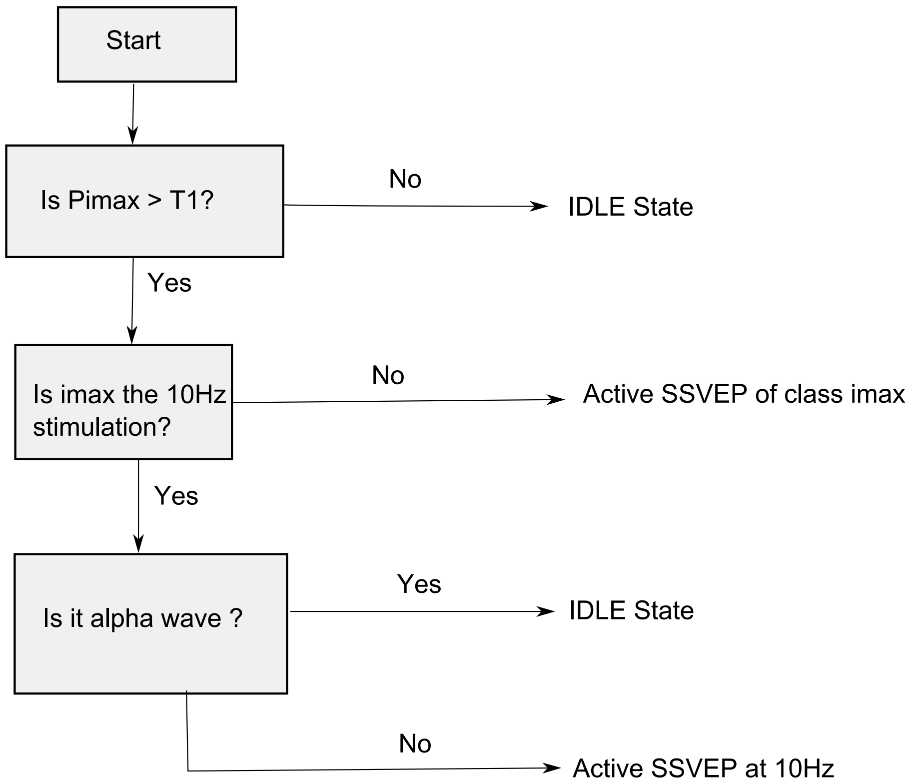


Figure A.5: Idle state detection Pipeline: For each frequency of stimulation i, CCA values P_i are computed. Let $imax$ be the frequency maximimizing P_i. When P_imax passes a pre-defined threshold T_1, the corresponding SSVEP activation is detected. $10Hz$ detection is an exception, and an additional alpha wave detection is performed for reducing false positives.

In [Wang et al., 2010a], the alpha wave detection is performed based on the SNR of the 10Hz frequency first harmonic (20Hz). If the SNR at 20Hz is higher than a pre-defined threshold θ (and only then), the SSVEP activation is confirmed.

We propose to exploit the P_{imax} value as an additional feature. SSVEP activation is detected if (and only if) equation A.2 is fulfilled, with F a decision function to define.

$$P_{imax} > F(SNR(20Hz)) \tag{A.2}$$

The method by [Wang et al., 2010a] can be expressed in this formalism by defining the decision function F as in Equation A.3:

$$F(x) = \begin{cases} +\infty & \text{if } x < \theta \\ 0 & \text{otherwise} \end{cases} \tag{A.3}$$

We propose to use an alternative decision functions F (Equation A.4), using T_h and T_l a high and a low threshold, and x the spike detection value at 20 Hz:

$$F(x) = \begin{cases} T_h & \text{if } x < \theta \\ T_l & \text{otherwise} \end{cases} \tag{A.4}$$

The function defined by Equation A.4 is represented in red on Figure A.6. This decision function (from now on called the *step function*) is discontinuous, reflecting the binary recognition of alpha wave. However, the detection of alpha wave is not a solid certainty, but rather an uncertain diagnostic. Thus, we hypothesized that better results could be achieved by reflecting this uncertainty in the decision process. In order to test this hypothesis, we used a continuous function as F. Tests were performed with a smooth threshold function: a sigmoid (see blue curve on Figure A.6).

This function is defined as:

$$\sigma(x) = T_l + (T_h - T_l)\frac{\theta^n}{\theta^n + x^n} \tag{A.5}$$

With T_l and T_h the low and high thresholds, θ the inflexion point abscissa, x the spike detection value, and n to set the slope of the function.

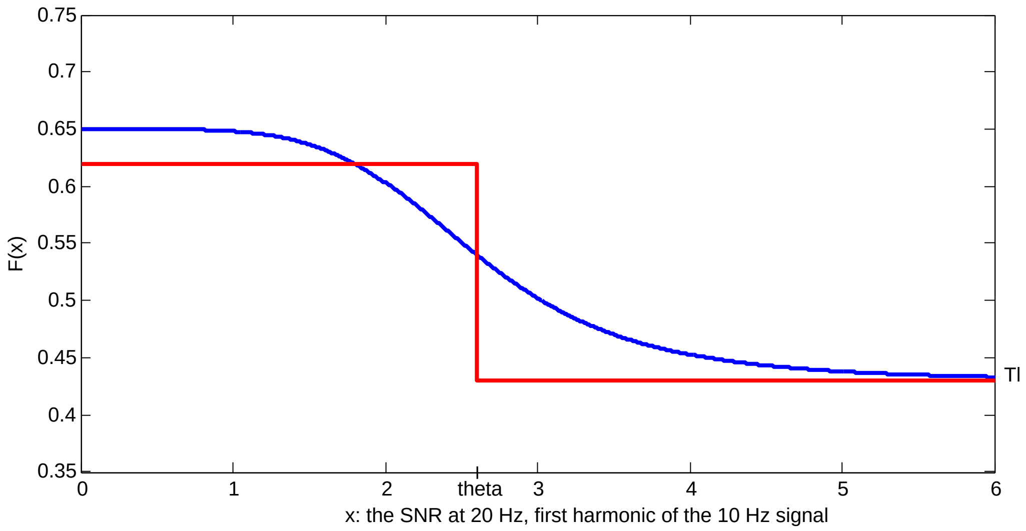


Figure A.6: Decision frontiers: the *step function* (red) and the *sigmoid* (blue)

The *step function* decision method requires the optimization of 3 parameters: the *high threshold*, the *low threshold*, and the *step position*. By contrast, our *continuous threshold* decision method is defined by 4 parameters: a *high threshold*, a *low threshold*, and two parameters θ and n.

These two methods were tested and compared, after optimization of all the parameters. All parameters (T_1, T_h, T_l, θ, n) were calibrated on the data from the *visual group* of the experiment described in chapter 3. Tests of the resulting sensitivity were then performed on a second set of participants (*Auditory group* from chapter 3). Any over-fitting effect is thus avoided.

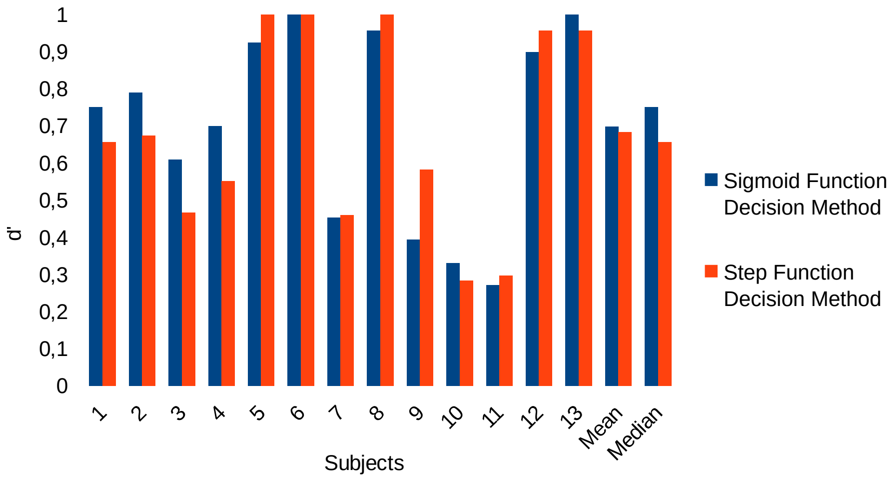


Figure A.7: Comparison of normalized sensitivity d' obtained with the sigmoid (continuous) decision function (blue), and the step decision function (red), for the 13 participants, and the resulting mean and median sensitivity (right).

The results of these tests are shown on figure A.7. They reveal that our *continuous* decision methods leads to significantly higher sensitivity for 4 participants, inferior sensitivity for 1 participant, and similar results for the remaining 8. The global mean and median sensitivity are higher with the *continuous* decision method.

This variability of performances depending on subjects is likely due, at least in part, to the difference in alpha rythm position and amplitude. Alpha wave frequency repartition among the population has been reported to follow a normal distribution centered on 9.9 Hz, with a standard deviation of 1.0 Hz [Wieneke et al., 1980]. Likewise, alpha wave detection step is only useful for subjects whose alpha rhythm is close to 10 Hz. For the same reason, it is likely that alpha wave detection would not bring any benefit when applied to the 12Hz SSVEP detection. Additional tests were performed and confirmed this hypothesis.

A.5 Conclusion

In most BCIs signal processing chains, a calibration phase is needed in order to take into account subject variability. Additionally, most signal processing chains are meant to segmentation-paced used. We combined and improved existing methods for SSVEP detection, and proposed a CCA-based signal processing method for SSVEP detection that is both self-paced and calibration-free. Additionally, improvements were made to alpha wave detection for false positive reduction leading to a better overall sensitivity than previously existing methods.

Appendix B

Application: A BCI-based web browser

Most BCI applications remain confined to laboratory experiments, in very controlled and simplified environments. Even in the few cases when they are applied outside the lab, the interaction context is usually simplified and the possibilities are restricted. For example, there are several BCI spellers available, but they usually enable text spelling only, without editing and formatting tools.

Web browsing is typically an application that presents complex and various interaction tasks, and becomes increasingly important for disabled people interaction, as more and more computer functionalities are embedded in web browsing. Several interaction techniques for web browsing with BCIs have been proposed before (see Section 2.4.1.B). Each of these browser are based on a single type of BCIs (either SCP, SSVEP, or P300). In this chapter, our objective is to develop a functional prototype of BCI-based web browser. We developed an interaction technique based on hybrid BCIs, aiming at optimizing the use of BCIs for Human-Computer Interaction.

In the remainder of this chapter, the approach for the BCI web browser conception is first described, detailing the targeted interaction tasks, the choice of cerebral patterns, and the resulting interaction technique. The prototype design and use is then presented, with the current state of advancement. The chapter ends with a discussion on future improvements and a conclusion.

B.1 Required tools and guidelines for Web browsing

In order to design an interaction technique for Web Browsing, the first step is to identify the main relevant tasks. Fortunately, the taxonomy of web browsing action has been thoroughly studied before (for example in [Catledge and Pitkow, 1995], [Weinreich et al., 2008], and [Jankowski, 2011]). These studies revealed that Web Browsing was originally mainly performed with two sub-tasks: *link following* and *back*. Since the early years of the internet, these actions remained central, and were joined by form completing (see Figure B.1).

A more complete taxonomy has been proposed by [Jankowski, 2011] (see Figure B.2). For the specific case of web accessibility for law bandwidth input, a set of guidelines has been proposed by

	Catledge et al.	Tauscher et al.	Weinreich et al.
Time of study	1994	1995 - 1996	2004 - 2005
Link following	45.7%	43.40%	43.50%
Direct access	12.60%	13.20%	9.40%
New window/tab	0.20%	0.80%	10.50%
Submit	-	4.40%	15.30%
Back	35.7%	31.70%	14.30%
Reload	4.30%	3.30%	1.70%
Forward	1.50%	0.80%	0.60%
Other	-	2.30%	4.80%

Figure B.1: Comparing Web usage studies [Weinreich et al., 2008]

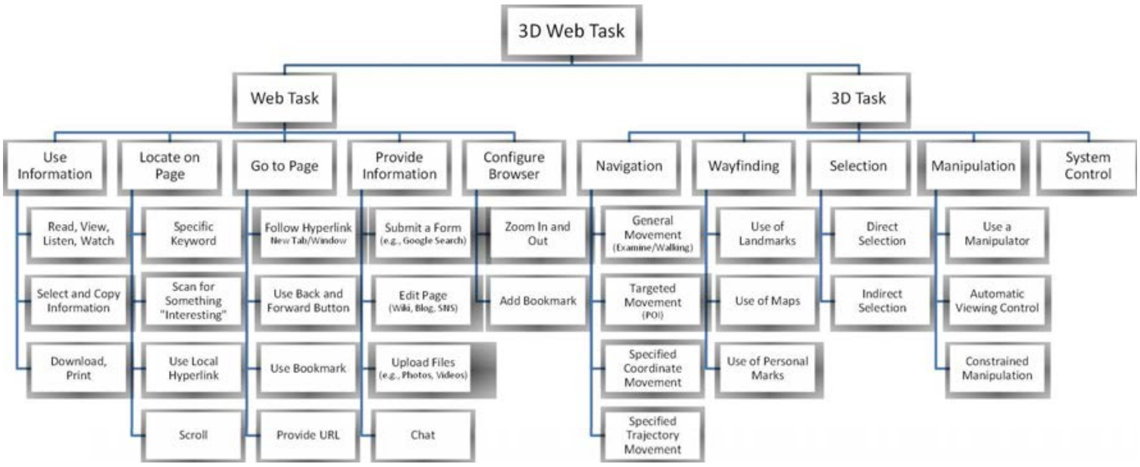


Figure B.2: Taxonomy of 3D web user tasks [Jankowski, 2011]

[Mankoff et al., 2002]. He proposes 3 primary objectives:

1. The currently selected link is visible.
2. The user can read and navigate text even if it contains no link.
3. The user can traverse the history list forward and backward.

In addition to these primary concerns, he describes secondary issues:

4. The user can access her bookmarks, and add to them.
5. The user can go quickly to a point of interest with a minimal number of signals.
6. The user is given alternatives for entering text and dealing with form elements.
7. The user is given enough information about link targets to make inform decisions about whether to follow them.

Respecting all these guidelines in a BCI context is a challenge in itself. As a first step, we aim at implementing a prototype respecting the 3 primary objectives of [Mankoff et al., 2002].

B.2 Objectives and interaction choices

As a first step, we aim at implementing a BCI-based web browser that enable the most critical and common interaction tasks, while following the Mankoff criteria as well as possible. We extracted four critical tasks for web browsing: *reading elements*, *following hyperlink*, *going back*, and *scrolling*. It can be noted that reading elements does not require any specific adaptation for BCIs, as displaying text is all that is required. However, the importance of this task implies that the interface should be self-paced, and present a low false positive rate.

Other important tasks are to *Submit a form*, and *provide URL*. The other tasks described by [Jankowski, 2011] can be considered to be of lesser importance for now. Additionally, considering the high error rate of BCIs, we predict that the possibility to cancel the last elementary interaction step will be of greater importance than what has been observed for "usual" web browsing. Hence, we decided to provide a web browser enabling:

- Hyperlink following.
- Scroll (up and down)
- A *back* command.
- An *undo* command.

These tasks are divided in two categories. Some tasks that can be achieved by selecting a command in a restricted set (thereafter called navigation commands): scroll up, scroll down, back and undo. One the other side, the Hyperlink following requires selecting an element (typically a link) in a set of variable size. This division aim at guiding the choice of BCI used for each task. Fast access to frequently used commands can be implemented, while the selection of a Hyperlink requires a specific interaction technique.

B.3 Choice of cerebral patterns

As discussed in 2.2.3, the cerebral pattern detected by a BCI has a critical influence on the BCI characteristics. In order to build a BCI adapted to various tasks, we propose to use a hybrid paradigm: different tasks can be handled by different BCIs.

The number of navigation commands is limited, and they require a fast access. Eligible cerebral patterns for these requirements are motor imagery and SSVEP (see Section 2.2.3). Additionally, we know that SSVEP use is compatible with dual cognitive task (see Chapter 3). By contrast, motor imagery requires a strong focus of the user, and is prone to false positive in a self-paced interaction context. Hence, we set our choice on SSVEP.

Hyperlink following presents different constraints. Using SSVEP selection for selecting a hyperlink would require having all the links permanently flickering at the same time. Such a display would be visually exhausting. Additionally, the number of link in a single webpage is usually very high (typically more than a hundred), preventing the attribution of a single frequency for each target. Finally, since hyperlinks in a webpage are often close to each other, the distance between stimuli would be too short for a precise detection (see Section 1.4.4.D). For hyperlink following, being able to handle a high number of targets is a criterion more important than speed of selection. Hence, we propose to use a link selection based on P300.

B.4 Interaction technique

We propose a web browser presenting two functioning modes. In the first functioning mode (thereafter called *navigation mode*), the webpage is displayed, and SSVEP selection is available for the navigation commands.

The navigation commands are:

- Scroll up: moves the view of the webpage toward the top.
- Scroll down: moves the view of the webpage toward the bottom.
- Back: Goes to the previously visited webpage. If selected n times in a row, goes to the n^th previously visited webpage.
- Undo: Cancel the last navigation command, or last hyperlink following. If selected twice in a row, cancel its own effect (thus leaving the browser state unchanged) [1].
- Select link: switch the web browser to Hyperlink selection mode (see below).

The second mode of functioning is the hyperlink selection mode (see Figure B.3). In this mode, a code is associated to each hyperlink visible on the webpage. Each code is displayed next to the association link on the webpage. A P300 speller is enabled. In order to select a link, the user

[1] It is worth noting that the *Undo* command is not equivalent to the *previous page* command. Let us consider the case of a user in page P_1. From there, he goes to page P_2, and then $P3$. Then our user selects two consecutive *previous page* commands. He would end up in P_1. By contrast, if after reaching P_3, he selects one *previous page* command, followed by one *Cancel* command, he would end up in P_3.

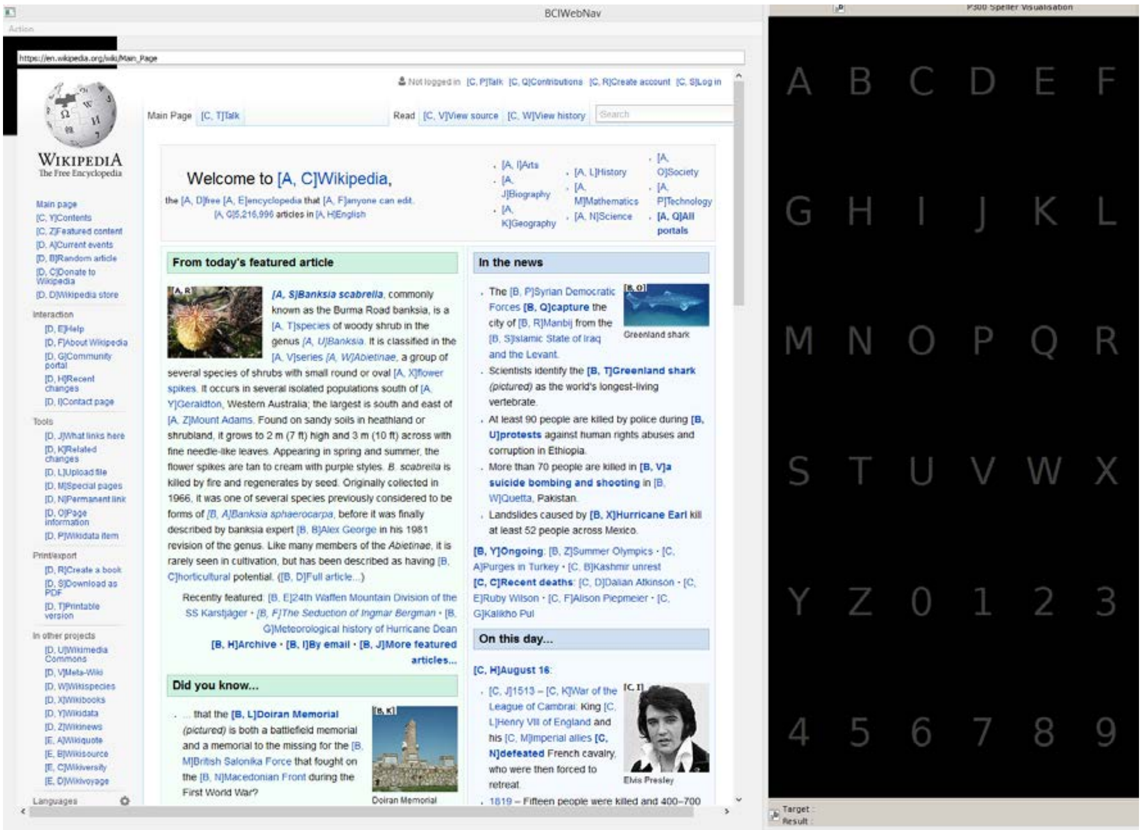


Figure B.3: The Web Browser design, in the hyperlink selection mode.

types the code of the desired link in this speller. This method is similar to the one described in [Mugler et al., 2010].

B.5 Prototype design

During the navigation mode, the webpage is displayed in the middle of the screen leaving the left and right sides to SSVEP selection. Navigation commands are accessible via *SSVEP buttons* placed around the screen (see Figure B.4). Each button is visible as a flickering SSVEP stimulation, with a symbol displayed in the center, representing the associated command. For each button, a feedback on SSVEP detection is given as a green filling-up bar that gives on indication on the level of activation. The self-paced SSVEP detection is done using a novel signal processing method (described in annex).

In order to limit the visual fatigue and the false positive rate, SSVEP buttons are only displayed when the current context requires it. In many cases, it does not make sense for the user to try to select a button. For example, the *scroll up* button is not displayed when the current view is already at the top of the webpage. Similarly, the *back* button is not displayed when the application just started.

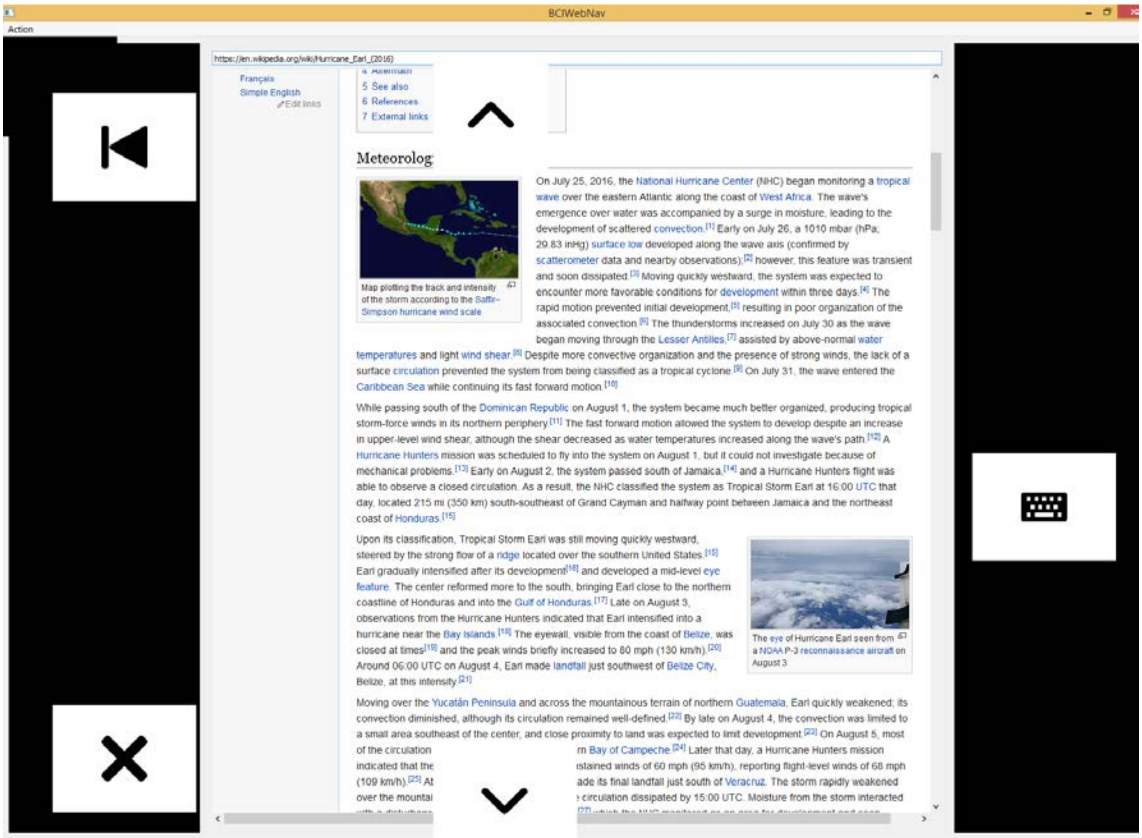


Figure B.4: The Web Browser design, in the navigation mode.

The SSVEP buttons are:

- Scroll up: this button is displayed on the top of the screen. It presents an arrow pointing upward. It is displayed only when the current view is not at the top of the webpage. The scroll is discrete, with a step of 60% of the screen height.
- Scroll down: this button is displayed at the bottom of the screen. It presents an arrow point downward. It is displayed only when the current view is not at the bottom of the webpage. The scroll is discrete, with a step of 60% of the screen height.
- Back: This button is displayed in the top left corner. It presents a *back* symbol inspired from music or video players. It is only displayed if the pile of visited webpages is not empty.
- Undo: This button is displayed in the bottom left corner. It presents a cross. It is not displayed when the application just started.
- Select link: This button is displayed on the right of the screen. It presents a keyboard symbol.

In the hyperlink selection mode (see Figure B.3), the webpage is moved to the left of the screen, and a P300 speller appears on the right. The code associated to each hyperlink is written directly as text in the webpage (see Figure B.3), preceding its links. In order to select a link, the user types the code of the desired link in the P300 speller. For most webpages, a code is composed of 2 letters (as the number of accessible links is usually between 27 and $26^2 = 676$).

The full setup is presented in Table B.1. The user is facing the screen on the right, and wears the EEG headset. Tables B.1 to B.9 describe in details an example of the web browser use. We follow a fictional user through the steps of hyperlink following with our BCI web browser prototype.

B.6 Future improvements

Our web browser can be evaluated in regards to Mankoff's guidelines for web browsing (see Section B.1). The primary objectives are respected. However, the secondary concerns are not yet implemented:

Mankoff's guideline number 4 states that *the user can access her bookmarks, and add to them.* A bookmark system could be added to our prototype. The most straightforward approach would be to add navigation commands (available with SSVEP). One navigation command would mark the current page, and another would access a bookmark selection menu. However, adding two SSVEP commands would come with an increased risk of false positive. Since adding a bookmark is not a common interaction task, it would be acceptable to provide this functionality in the P300 selection, under the form of a dedicated P300 keyboard key. Similarly, accessing a bookmark can be assimilated to accessing a webpage, provided that each bookmark is treated as a permanently accessible link.

Mankoff's guideline number 5 states that *the user can go quickly to a point of interest with a minimal number of signals.* While this consideration has been taken into account during the design of our prototype, improvements could be made to the scroll up and scroll down system, in the case of long pages. In the current version, browsing a long webpage can take a long time, as the scroll down command has to be selected numerous times. A system of links at a different position in the

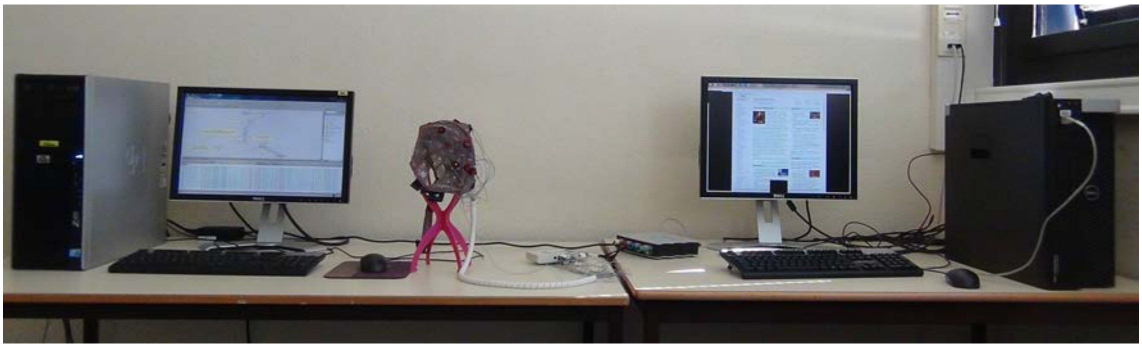


Table B.1: Setup: A computer is dedicated to SSVEP signal processing and monitoring (left). An EEG headset (middle left) is used to measure the cerebral activity, and is connected to an amplifier (middle right) that relays the signal to the server machine (right). The main machine (right) handle is used as acquisition serveur, for the P300 signal processing, and displays the web browser interface.

Table B.2: The user is wearing an EEG headset. He is currently on the front page of wikipedia. He wants to select a link to go to the geography portal. He focuses on the *Select link* SSVEP button, on the right of the screen.

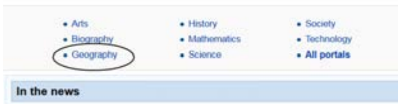


Table B.3: The link that the user wants to select.

Table B.4: The web browser is switched to hyperlink selection mode. A P300 keyboard appears, and codes appear next to each hyperlink.

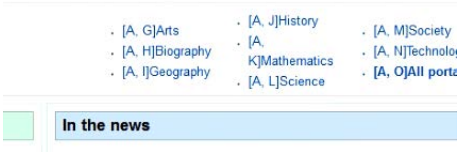


Table B.5: Codes appear next to each possible link. The user looks next to the geography portal link, and read his code: *A*, *I*

same page could be added to the browser to improve this aspect.
A second limitation to Mankoff's guideline number 5 is the imperfection of the undo command. The goal of the undo function is to reduce the cost of errors in the interface. Thus, it should be accessible at any time, and lead to a correction of the last atomic error. But for now, the undo command cancels the last full command. In the context of BCI-based interaction, a single command sometimes requires several sub-steps to be selected. Typically, during link selection, three basic interaction tasks are performed: selecting the *select link* SSVEP button, typing the first letter of the code, and typing the second letter of the code. Each of these sub-steps requires a significant time, and presents a risk of error. In the current version of the prototype, an error at the last step of this selection requires the user to *undo* the full hyperlink following, and to start again. While the original tasks required 3 basic interaction tasks to be achieved, a single error in one of the 3 steps leads to the need of 7 basic interaction tasks (3 steps comprising one mistake, one *Undo* to select, the 3 again to succeed in the original task). By comparison, if the undo command canceled only the last basic interaction task, instead of the last full command, an error in the 3 steps of the selection would only bring the total amount of steps to 5 (3 steps comprising 1 mistake, one *Undo* to select, and one step to succeed in the previously missed step). In summary, the *Undo* command can be improved by properly canceling only the last basic interaction task, instead of the full command.

Mankoff's guideline number 6 states that *the user is given alternatives for entering text and dealing with form elements*. Form filling is a major interaction task for web browsing, and has not been yet been implemented in our prototype. The interaction technique enabling it would straightforward: selecting a form could be done just as selecting a link, with selection codes associated to each form, on top of each link. This selection could lead to a spelling application (P300 speller) to fill the form. The tools for this interaction technique are in place.

Finally, Mankoff's guideline number 7 states that *the user is given enough information about link targets to make inform decisions about whether to follow them*. Our current prototype keeps displaying the name of the link during selection. The code associated to each hyperlink during selection does not replace the text describing it, but is rather added to it (see Figure B.3). However, by comparison to standard web browsers, the potential mouseover text is lost. The issue of displaying mouseover texts during hyperlink selection mode without them overlapping is left to

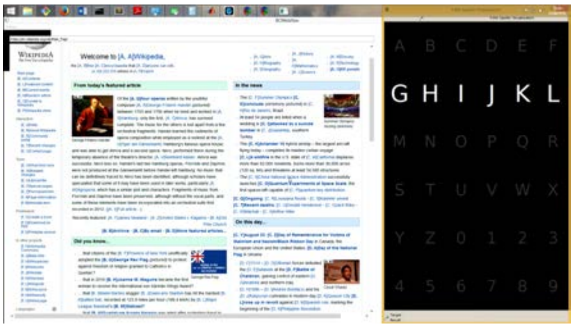


Table B.6: The P300 stimulation begins. Rows and columns of letters are flashing. The user focuses on the *A*, and counts the number of times it flashes.

Table B.7: The letter *A* is selected. A green feedback confirms the selection.

Table B.8: The P300 keyboard then start flashing again. The user repeats the operation to provide the second letter of his code: *I*

Table B.9: The selection of the second code letter succeeded. The web browser switch back to navigation mode, and opens to geography portal page.

future research.

B.7 Conclusion

This chapter presented a prototype of web browser fully based on BCIs. The essential interaction tasks for web browsing were first delimited. An interaction technique was proposed to tackle these interaction tasks, while taking into account the particularities of BCI-based interaction. A prototype was finally developed, proposing a design for a fully BCI-based web browser. By contrast with previously existing BCI-based web browsers, a hybrid approach was privileged. For selection among a limited set of options, a SSVEP-based BCI was used, because of its higher speed. At the opposite, for hyperlink selection, a P300-based was used, in order to reduce the number of steps. , enabling to reduce the number of steps necessary for each interaction task.

Several improvements have yet to be implemented. In particular, the *Undo* command can be improved to better limit the cost of errors. The web browser prototype could also be enriched with new functionalities, such as form filling.

Appendix C

Chapter 4 Annex: Questionnaires

Starting next page, the following of this annexe are the questionnaires filled in by all participants of the experiment described in chapter 4. It contains one pre-experiment questionnaire, one questionnaire filled in after each block i, and one post-experiment questionnaire. As our participants were native french speakers, the questionnaires are written in this language.

Objectifs

Durant cette expérience nous allons évaluer différents systèmes d'interface cerveau-ordinateur de type "SSVEP" (basées sur des cibles qui clignotent à l'écran). Nous allons évaluer les performances des différents algorithmes et également recueillir vos appréciations.

La tâche à effectuer

Vous devez sélectionner une cible parmi trois cibles affichées à l'écran, en utilisant une interface cerveau-ordinateur. Pour sélectionner une cible vous devrez simplement fixer votre attention sur celle-ci. Au début de chaque essai, une flèche bleue vous indiquera la cible que vous devez sélectionner. Votre objectif est donc à chaque essai de fixer la cible pointée par la flèche. Après chaque essai, le résultat de la sélection par le système est affiché pendant 2 secondes : "réussite" ou "échec".
L'expérience est divisée en plusieurs blocs contenant chacune plusieurs essais. Entre chacun des blocs vous aurez une pause pour pouvoir vous reposer et remplir un petit questionnaire.

Quelques définitions

Pour que l'on soit d'accord sur les termes employés lors des différents questionnaires que vous allez devoir remplir, voici trois définitions.

Fatigue : Diminution des forces de l'organisme, généralement provoquée par un travail excessif ou trop prolongé, ou liée à un état fonctionnel défectueux

Frustration : État de quelqu'un qui est frustré, empêché d'atteindre un but ou de réaliser un désir

Efficacité : Caractère d'une chose qui produit l'effet attendu

Motivation : Raisons, intérêts, éléments qui poussent quelqu'un dans son action

Recommandation

L'objectif est de réussir à sélectionner le plus grand nombre possible de cibles sans échec. Il faut rester concentrer sur cet objectif pendant toute la durée de l'expérience. Les données du casque à électrodes sont enregistrées en permanence et tout au long des essais. Il faut donc maintenir votre niveau de performance le plus élevé possible tout au long de l'étude.

Session :
ID :

Questionnaire pré-expérience

En relisant attentivement les définitions données ci-dessus veuillez répondre au questionnaire suivant.

Genre:
Age:
Main directrice (Gaucher, Droitier, Ambidextre) :
Vision (Normale, Corrigée) :
Daltonien (Oui, Non) :
Antécédents épileptiques (Si oui préciser) :

Expérience BCI (Si oui préciser) :

Evaluez votre sentiment de **frustration** avant le début de l'expérience:

Absent	A peine perceptible	Faiblement présent	Léger	Marqué	Prononcé	Très prononcé

Evaluez votre sentiment de **fatigue** avant le début de l'expérience :

Absent	A peine perceptible	Faiblement présent	Léger	Marqué	Prononcé	Très prononcé

Evaluez votre **motivation**:

Absent	A peine perceptible	Faiblement présent	Léger	Marqué	Prononcé	Très prononcé

Session :
ID :

Bloc *i*

Evaluez votre sentiment de **frustration** éprouvé durant ce bloc:

Absent	A peine perceptible	Faiblement présent	Léger	Marqué	Prononcé	Très prononcé

Evaluez votre sentiment de **frustration** depuis le début de l'expérience:

Absent	A peine perceptible	Faiblement présent	Léger	Marqué	Prononcé	Très prononcé

Evaluez votre sentiment de **fatigue** depuis le début de l'expérience:

Absent	A peine perceptible	Faiblement présent	Léger	Marqué	Prononcé	Très prononcé

Le système présenté vous semble-t-il **efficace** ?

Pas du tout d'accord	Pas d'accord	Ni en désaccord ni d'accord	D'accord	Tout à fait d'accord

Evaluez votre **motivation** durant ce bloc:

Absent	A peine perceptible	Faiblement présent	Léger	Marqué	Prononcé	Très prononcé

Session :
ID :

QUESTIONS OUVERTES

Remarques éventuelles sur le système et son utilisation ?

Remarques éventuelles sur l'efficacité du système ?

Remarque sur la frustration ou la fatigue éventuellement ressenties ?

Avez-vous trouvé que le système répondait de manière cohérente à vos intentions ?

Session :
ID :

Avez-vous perçu que les résultats (échec/réussite) n'étaient en fait pas du tout liés à vos performances réelles ?

Certains objectifs de l'expérience ne vous ont pas été donnés avant, afin de ne pas influencer vos résultats pendant l'expérience.

A l'heure actuelle les performances des interfaces cerveau-machines ne sont pas toujours très bonnes. Les erreurs commises par ces interfaces sont parfois très nombreuses avec des taux d'erreurs importants. Ceci peut grandement nuire à l'acceptabilité et à la diffusion de ces technologies. Nous souhaitons donc mieux comprendre les conditions dans lesquelles ces technologies peuvent être acceptées par des utilisateurs, et les conditions qui génèrent un sentiment de fatigue ou de frustration de l'utilisateur. A termes ceci pourrait servir de référence aux futurs concepteurs de ces technologies.

Dans cette expérience nous étudions le lien éventuel entre les taux erreurs de telles interfaces, la frustration et la fatigue engendrées sur l'utilisateur. Dans ce cadre, afin d'avoir un contrôle sur le taux d'erreur, nous avons entièrement simulés vos résultats affichés (échec/réussite) par des tirages aléatoire. Cependant votre activité cérébrale était elle réellement mesurée et enregistrée afin de pouvoir analyser en détail les données et l'influence du taux d'erreur plus finement en post-traitement.

Nous vous remercions grandement pour votre participation à cette expérience et nous vous invitons à ne pas parler des objectifs réels de celle-ci. Si aider la science intéresse vos proches, n'hésitez pas à leur proposer de participer !

Avez vous des questions ?

Appendix D

Chapter 6 Annex: individual results

The table D.1 presents the individual results for the experiment described in chapter 6.

Table D.1: Individual results for hit, miss, and false positive rates, along with the resulting sensitivity d'

Subject	Method	Hit	Miss	FP	d'
Subject 1	Gaze-only	0.8	0.09	0.11	0.69
	Sequential	0.33	0.65	0.02	0.31
	Fusion	0.41	0.43	0.17	0.24
Subject 2	Gaze-only	0.81	0.04	0.15	0.67
	Sequential	0.13	0.83	0.04	0.09
	Fusion	0.69	0.28	0.04	0.65
Subject 3	Gaze-only	0.69	0.19	0.13	0.56
	Sequential	0.19	0.78	0.04	0.15
	Fusion	0.59	0.31	0.09	0.50
Subject 4	Gaze-only	0.91	0.02	0.07	0.83
	Sequential	0.28	0.70	0.02	0.26
	Fusion	0.80	0.11	0.09	0.70
Subject 5	Gaze-only	0.59	0.24	0.17	0.43
	Sequential	0.14	0.78	0.08	0.06
	Fusion	0.17	0.46	0.37	-0.20
Subject 6	Gaze-only	0.89	0.00	0.11	0.78
	Sequential	0.06	0.89	0.06	0.00
	Fusion	0.63	0.30	0.07	0.56
Subject 7	Gaze-only	0.91	0.02	0.07	0.83
	Sequential	0.24	0.67	0.09	0.15
	Fusion	0.52	0.33	0.15	0.37
Subject 8	Gaze-only	0.89	0.02	0.09	0.80
	Sequential	0.17	0.78	0.06	0.11
	Fusion	0.73	0.18	0.10	0.63
Subject 9	Gaze-only	0.94	0.00	0.06	0.89
	Sequential	0.35	0.59	0.06	0.30
	Fusion	0.52	0.26	0.22	0.30
Subject 10	Gaze-only	0.87	0.02	0.11	0.76
	Sequential	0.31	0.65	0.04	0.28
	Fusion	0.46	0.41	0.13	0.33
Subject 11	Gaze-only	0.81	0.06	0.13	0.69
	Sequential	0.48	0.52	0.00	0.48
	Fusion	0.65	0.26	0.09	0.56
Subject 12	Gaze-only	0.91	0.02	0.07	0.83
	Sequential	0.20	0.67	0.13	0.07
	Fusion	0.74	0.20	0.06	0.69

List of Figures

Bibliography

[Abootalebi et al., 2009] Abootalebi, V., Moradi, M. H., and Khalilzadeh, M. A. (2009). A new approach for EEG feature extraction in p300-based lie detection. *Computer Methods and Programs in Biomedicine*, 94(1):48 – 57.

[Agam and Sekuler, 2007] Agam, Y. and Sekuler, R. (2007). Interactions between working memory and visual perception: an ERP/EEG study. *Neuroimage*, 36(3):933–942.

[Alimardani et al., 2014] Alimardani, M., Nishio, S., and Ishiguro, H. (2014). Effect of biased feedback on motor imagery learning in bci-teleoperation system. *Frontiers in systems neuroscience*, 8:52.

[Aljshamee et al., 2014] Aljshamee, M., Mohammed, M. Q., Riaz-Ul-Ahsan Choudhury, A. M., and Luksch, P. (2014). Beyond pure frequency and phases exploiting: Color influence in ssvep based on bci. *Computer Technology and Application*, 5(2).

[Allanson and Fairclough, 2004] Allanson, J. and Fairclough, S. H. (2004). A research agenda for physiological computing. *Interacting with Computers*, 16(5):857 – 878.

[Allison et al., 2008] Allison, B. Z., McFarland, D. J., Schalk, G., Zheng, S. D., Jackson, M. M., and Wolpaw, J. R. (2008). Towards an independent brain–computer interface using steady state visual evoked potentials. *Clinical neurophysiology*, 119(2):399–408.

[Allison et al., 2007] Allison, B. Z., Wolpaw, E. W., and Wolpaw, J. R. (2007). Brain–computer interface systems: progress and prospects. *Expert review of medical devices*, 4(4):463–474.

[Amir and Ralph, 2014] Amir, B. and Ralph, P. (2014). Proposing a theory of gamification effectiveness. In *Companion Proceedings of the 36th International Conference on Software Engineering*, pages 626–627. ACM.

[Amores et al., 2016] Amores, J., Benavides, X., and Maes, P. (2016). Psychicvr: Increasing mindfulness by using virtual reality and brain computer interfaces. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 2–2. ACM.

[Andersen et al., 2008] Andersen, S. K., Hillyard, S. A., and Müller, M. M. (2008). Attention facilitates multiple stimulus features in parallel in human visual cortex. *Current Biology*, 18(13):1006 – 1009.

[Ashmore et al., 2005] Ashmore, M., Duchowski, A. T., and Shoemaker, G. (2005). Efficient eye pointing with a fisheye lens. In *Proceedings of Graphics interface 2005*, pages 203–210. Canadian Human-Computer Communications Society.

[Babiloni and Astolfi, 2014] Babiloni, F. and Astolfi, L. (2014). Social neuroscience and hyperscanning techniques: Past, present and future. *Neuroscience & Biobehavioral Reviews*, 44:76 – 93.

[Baker, 2007] Baker, S. N. (2007). Oscillatory interactions between sensorimotor cortex and the periphery. *Current Opinion in Neurobiology*, 17(6):649 – 655. Motor systems / Neurobiology of behaviour.

[Bauer et al., 1979] Bauer, G., Gerstenbrand, F., and Rumpl, E. (1979). Varieties of the locked-in syndrome. *Journal of neurology*, 221(2):77–91.

[Baumann et al., 2007] Baumann, M. R., Rösler, D., and Krems, J. F. (2007). Situation awareness and secondary task performance while driving. In *International Conference on Engineering Psychology and Cognitive Ergonomics*, pages 256–263.

[Beaudouin-Lafon, 2006] Beaudouin-Lafon, M. (2006). Interaction homme-machine. In Akoka, J. and Comyn-Wattiau, I., editors, *Encyclopédie de l'Informatique et des Systèmes d'Information*. Vuibert.

[Bi et al., 2013] Bi, L., Jie, K., an Fan, X., and Li, Y. (2013). A SSVEP brain-computer interface with the hybrid stimuli of SSVEP and P300. In *ICME International Conference on Complex Medical Engineering (CME)*, pages 211–214.

[Bin et al., 2009] Bin, G., Gao, X., Yan, Z., Hong, B., and Gao, S. (2009). An online multi-channel ssvep-based brain–computer interface using a canonical correlation analysis method. *Journal of neural engineering*, 6(4):046002.

[Birbaumer, 2006] Birbaumer, N. (2006). Breaking the silence: Brain-computer interfaces (bci) for communication and motor control. *Psychophysiology*, 43(6):517–532.

[Birbaumer et al., 1990] Birbaumer, N., Elbert, T., Canavan, A. G., and Rockstroh, B. (1990). Slow potentials of the cerebral cortex and behavior. *Physiological Reviews*, 70(1):1–41.

[Birbaumer et al., 2003] Birbaumer, N., Hinterberger, T., Kubler, A., and Neumann, N. (2003). The thought-translation device (ttd): neurobehavioral mechanisms and clinical outcome. *IEEE transactions on Neural Systems and rehabilitation engineering*, 11(2):120–123.

[Bishop, 1996] Bishop, C. (1996). Neural network for pattern recognition. *Oxford University Press.*

[Bishop, 1995] Bishop, C. M. (1995). Training with noise is equivalent to Tikhonov regularization. *Neural Computation*, 7(1):108–116.

[Blankertz et al., 2006a] Blankertz, B., Dornhege, G., Krauledat, M., Muller, K., Kunzmann, V., Losch, F., and Curio, G. (2006a). The berlin brain-computer interface: Eeg-based communication without subject training. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2):147–152.

[Blankertz et al., 2006b] Blankertz, B., Dornhege, G., Krauledat, M., Shröder, M., Williamson, J., Murray-Smith, R., and Müller, K.-R. (2006b). The berlin brain-computer interface presents the novel mental typewriter hex-o-spell. In *Proceedings of the 3rd International Brain-Computer Interface Workshop and Training Course, Graz*, pages 108–109.

[Blankertz et al., 2010] Blankertz, B., Tangermann, M., Vidaurre, C., Fazli, S., Sannelli, C., Haufe, S., Maeder, C., Ramsey, L., Sturm, I., Curio, G., et al. (2010). The berlin brain–computer interface: non-medical uses of bci technology. *Frontiers in neuroscience*, 4.

[Blankertz et al., 2008] Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., and Muller, K.-R. (2008). Optimizing spatial filters for robust eeg single-trial analysis. *Signal Processing Magazine, IEEE*, 25(1):41–56.

[Bonnet et al., 2013] Bonnet, L., Lotte, F., and Lécuyer, A. (2013). Two Brains, One Game: Design and Evaluation of a Multi-User BCI Video Game Based on Motor Imagery. *IEEE Transactions on Computational Intelligence and AI in games*, 5(2):185–198.

[Bos et al., 2010] Bos, D.-O., Reuderink, B., van de Laar, B., Gurkok, H., Muhl, C., Poel, M., Heylen, D., and Nijholt, A. (2010). Human-computer interaction for bci games: Usability and user experience. In *Proc. of Cyberworlds*, pages 277–281.

[Brouwer et al., 2012] Brouwer, A.-M., Hogervorst, M. A., Van Erp, J. B., Heffelaar, T., Zimmerman, P. H., and Oostenveld, R. (2012). Estimating workload using EEG spectral power and ERPs in the n-back task. *Journal of neural engineering*, 9(4):045008.

[Bruder et al., 2015] Bruder, G., Lubas, P., and Steinicke, F. (2015). Cognitive resource demands of redirected walking. *IEEE transactions on visualization and computer graphics*, 21(4):539–544.

[Brunner et al., 2011] Brunner, C., Allison, B., Altstätter, C., and Neuper, C. (2011). A comparison of three brain–computer interfaces based on event-related desynchronization, steady state visual evoked potentials, or a hybrid approach using both signals. *Journal of neural engineering*, 8(2):025010.

[Burges, 1998] Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, 2(2):121–167.

[Cabestaing and Rakotomamonjy, 2007] Cabestaing, F. c. and Rakotomamonjy, A. (2007). Introduction aux interfaces cerveaux machines. In *Colloque GRETSI*.

[Cantero et al., 2003] Cantero, J., Atienza, M., Stickgold, R., Kahana, M., Madsen, J., and Kocsis, B. (2003). Sleep-dependant theta oscillations in the human hippocampus and neocortex. *The journal of neuroscience*, 23(34):10897–10903.

[Cao et al., 2012] Cao, T., Wan, F., Mak, P. U., Mak, P.-I., Vai, M. I., and Hu, Y. (2012). Flashing color on the performance of SSVEP-based brain-computer interfaces. In *IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1819–1822.

[Cao et al., 2013] Cao, T., Wong, C. M., Wan, F., and Hu, Y. (2013). Objective indication of fatigue in SSVEP-based BCI through EEG spectral analysis. *5th international brain-computer interface (BCI) meeting.*

[Casiez et al., 2012] Casiez, G., Roussel, N., and Vogel, D. (2012). 1€ filter: a simple speed-based low-pass filter for noisy input in interactive systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2527–2530. ACM.

[Catledge and Pitkow, 1995] Catledge, L. D. and Pitkow, J. E. (1995). Characterizing browsing strategies in the world-wide web. *Computer Networks and ISDN systems*, 27(6):1065–1073.

[Ceaparu et al., 2004] Ceaparu, I., Lazar, J., Bessiere, K., Robinson, J., and Shneiderman, B. (2004). Determining causes and severity of end-user frustration. *International Journal of Human-Computer Interaction*, 17:333–356.

[Chen et al., 2014a] Chen, X., Chen, Z., Gao, S., and Gao, X. (2014a). A high-ITR SSVEP-based BCI speller. *Brain-Computer Interfaces*, 1(3-4):181–191.

[Chen et al., 2015] Chen, X., Wang, Y., Nakanishi, M., Gao, X., Jung, T.-P., and Gao, S. (2015). High-speed spelling with a noninvasive brain–computer interface. *Proceedings of the National Academy of Sciences*, 112(44):E6058–E6067.

[Chen et al., 2014b] Chen, X., Wang, Y., Nakanishi, M., Jung, T.-P., and Gao, X. (2014b). Hybrid frequency and phase coding for a high-speed SSVEP-based BCI speller. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3993–3996. IEEE.

[Cheng et al., 2002] Cheng, M., Gao, X., Gao, S., and Xu, D. (2002). Design and implementation of a brain-computer interface with high transfer rates. *IEEE transactions on biomedical engineering*, 49(10):1181–1186.

[Choi et al., 2013] Choi, J.-S., Bang, J. W., Park, K. R., and Whang, M. (2013). Enhanced perception of user intention by combining EEG and gaze-tracking for brain-computer interfaces (BCIs). *Sensors*, 13(3):3454–3472.

[Clerc et al., 2013] Clerc, M., Mattout, J., Maby, E., Devlaminck, D., Papadopoulo, T., Guy, V., and Desnuelle, C. (2013). Verbal communication through brain computer interfaces. In *Interspeech-14th Annual Conference of the International Speech Communication Association-2013*.

[Collinger et al., 2014] Collinger, J. L., Kryger, M. A., Barbara, R., Betler, T., Bowsher, K., Brown, E. H. P., Clanton, S. T., Degenhart, A. D., Foldes, S. T., Gaunt, R. A., Gyulai, F. E., Harchick, E. A., Harrington, D., Helder, J. B., Hemmes, T., Johannes, M. S., Katyal, K. D., Ling, G. S. F., McMorland, A. J. C., Palko, K., Para, M. P., Scheuermann, J., Schwartz, A. B., Skidmore, E. R., Solzbacher, F., Srikameswaran, A. V., Swanson, D. P., Swetz, S., Tyler-Kabara, E. C., Velliste, M., Wang, W., Weber, D. J., Wodlinger, B., and Boninger, M. L. (2014). Collaborative approach in the development of high-performance brain-computer interfaces for a neuroprosthetic arm: Translation from animal models to human control. *Clinical and Translational Science*, 7(1):52–59.

[Galloway, NR, 1990] Galloway, NR (1990). Human brain electrophysiology: Evoked potentials and evoked magnetic fields in science and medicine. *The British journal of ophthalmology*, 74(4):255.

[Zhao, Wenyi and Chellappa, Rama and Phillips, P Jonathon, 1999] Zhao, Wenyi and Chellappa, Rama and Phillips, P Jonathon (1999). *Subspace linear discriminant analysis for face recognition.* Citeseer.

[Cutrell and Tan, 2008] Cutrell, E. and Tan, D. (2008). BCI for passive input in HCI. In *Proceedings of CHI*, volume 8, pages 1–3.

[Daly et al., 2009] Daly, J. J., Cheng, R., Rogers, J., Litinas, K., Hrovat, K., and Dohring, M. (2009). Feasibility of a new application of noninvasive brain computer interface (BCI): a case study of training for recovery of volitional motor control after stroke. *Journal of Neurologic Physical Therapy*, 33(4):203–211.

[Donchin et al., 2000] Donchin, E., Spencer, K., and Wijesinghe, R. (2000). The mental prosthesis: assessing the speed of a P300-based brain-computer interface. *IEEE Transactions on Rehabilitation Engineering*, 8(2):174–179.

[Duvinage et al., 2012] Duvinage, M., Castermans, T., Dutoit, T., Petieau, M., Hoellinger, T., Saedeleer, C. D., Seetharaman, K., and Cheron, G. (2012). A P300-based quantitative comparison between the Emotiv Epoc headset and a medical EEG device. *Biomedical Engineering*, 765:2012–764.

[Duvinage et al., 2013] Duvinage, M., Castermans, T., Petieau, M., Hoellinger, T., Cheron, G., and Dutoit, T. (2013). Performance of the Emotiv Epoc headset for P300-based applications. *Biomedical engineering online*, 12(1):1.

[Évain et al., 2016] Évain, A., Argelaguet, F., Casiez, G., Roussel, N., and Lécuyer, A. (2016). Do the stimuli of an SSVEP-based BCI really have to be the same as the stimuli used for training it? *Brain-Computer Interfaces*, 3(2):103–111.

[Evans and Stanovich, 2013] Evans, J. S. B. and Stanovich, K. E. (2013). Dual-process theories of higher cognition advancing the debate. *Perspectives on psychological science*, 8(3):223–241.

[Fard and Grosse-Wentrup, 2014] Fard, P. R. and Grosse-Wentrup, M. (2014). The Influence of Decoding Accuracy on Perceived Control: A Simulated BCI Study. *ArXiv e-prints*.

[Farwell and Donchin, 1988] Farwell, L. A. and Donchin, E. (1988). Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and clinical Neurophysiology*, 70(6):510–523.

[Fatourechi et al., 2007] Fatourechi, M., Bashashati, A., Ward, R. K., and Birch, G. E. (2007). EMG and EOG artifacts in brain computer interface systems: A survey. *Clinical Neurophysiology*, 118(3):480 – 494.

[Ferrari and Quaresima, 2012] Ferrari, M. and Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fnirs) development and fields of application. *NeuroImage*, 63(2):921 – 935.

[Ferrez et al., 2006] Ferrez, P., GalánMoles, F., Buttfield, A., Gonzalez Andino, S., Grave de Peralta Menendez, R., and Millán (2006). High frequency bands and estimated local field potentials to improve single-trial classification of electroencephalographic signals. *3rd Int. BCI Workshop and Training Course*.

[Foley et al., 1984] Foley, J., Wallace, V., and Chan, P. (1984). The human factors of computer graphics interaction techniques. *IEEE Computer Graphics and Applications*, 4(11):13–48.

[Frey et al., 2014] Frey, J., Gervais, R., Fleck, S., Lotte, F., and Hachet, M. (2014). Teegi: Tangible EEG interface. In *UIST-ACM User Interface Software and Technology Symposium*, pages 301–308.

[Friman et al., 2001] Friman, O., Cedefamn, J., Lundberg, P., Borga, M., and Knutsson, H. (2001). Detection of neural activity in functional MRI using canonical correlation analysis. *Magnetic Resonance in Medicine*, 45(2):323–330.

[Fruitet et al., 2011] Fruitet, J., Clerc, M., and Papadopoulo, T. (2011). Preliminary study for an offline hybrid BCI using sensorimotor rhythms and beta rebound. *International Journal of bioelectromagnetism*, 13(2):70–71.

[Fukunaga, 2013] Fukunaga, K. (2013). *Introduction to statistical pattern recognition*. Academic press.

[Gan, 2006] Gan, J. Q. (2006). Self-adapting BCI based on unsupervised learning. In *3rd International workshop on brain-computer interfaces*, pages 50–51.

[Gao et al., 2003] Gao, X., Xu, D., Cheng, M., and Gao, S. (2003). A BCI-based environmental controller for the motion-disabled. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):137–140.

[Gardner, 2005] Gardner, S. D. (2005). *Evaluation of the colorbrewer color schemes for accomodation of map readers with impaired color vision*. PhD thesis, Pensylvania State University, College of Earth and Mineral Sciences.

[George, 2012] George, L. (2012). *Contribution to the Study of Active and Passive Brain-Computer Interfaces for Interacting With Virtual Environments Based on EEG and Mental Workload*. PhD thesis, INSA Rennes.

[George et al., 2014] George, L., Marchal, M., Glondu, L., and Lécuyer, A. (2014). Combining brain-computer interfaces and haptics: Detecting mental workload to adapt haptic assistance. In *Brain-Computer Interface Research*, pages 19–28. Springer.

[Goldin et al., 2006] Goldin, D., Smolka, S. A., and Wegner, P. (2006). *Interactive Computation: The New Paradigm*. Springer-Verlag New York.

[Gonzalez-Franco et al., 2011] Gonzalez-Franco, M., Yuan, P., Zhang, D., Hong, B., and Gao, S. (2011). Motor imagery based brain-computer interface: A study of the effect of positive and negative feedback. In *IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 6323–6326.

[Guger et al., 2003] Guger, C., Edlinger, G., Harkam, W., Niedermayer, I., and Pfurtscheller, G. (2003). How many people are able to operate an EEG-based brain-computer interface (BCI)? *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 11(2):145–147.

[Gürkök et al., 2011] Gürkök, H., Hakvoort, G., and Poel, M. (2011). Modality switching and performance in a thought and speech controlled computer game. In *Proceedings of the 13th international conference on multimodal interfaces*, pages 41–48. ACM.

[Haw et al., 2006] Haw, C. J., Lowne, D., and Roberts, S. (2006). User specific template matching for event detection using single channel EEG. In *Graz conference on brain-computer interfaces*.

[Heraz and Frasson, 2007] Heraz, A. and Frasson, C. (2007). Predicting the three major dimensions of the learner's emotions from brainwaves. *International journal of computer science*, 25:323–329.

[Herrmann, 2001] Herrmann, C. S. (2001). Human EEG responses to 1-100 hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental Brain Research*, 137(3-4):346–353.

[Hewett et al., 1992] Hewett, T. T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M., Perlman, G., Strong, G., and Verplank, W. (1992). *ACM SIGCHI curricula for human-computer interaction*. ACM.

[Hillyard and Anllo-Vento, 1998] Hillyard, S. A. and Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences*, 95(3):781–787.

[Hintermüller et al., 2013] Hintermüller, C., Kapeller, C., Edlinger, G., and Guger, C. (2013). *Brain-Computer Interfaces systems : recent progress and future prospects*, chapter BCI integration: Application Interfaces. inTech.

[Hinton et al., 2012] Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A.-r., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T. N., et al. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6):82–97.

[Hjelm, 2003] Hjelm, S. I. (2003). The making of brainball. *Interactions*, 10:26–34.

[Hotelling, 1936] Hotelling, H. (1936). Relations between two sets of variates. *Biometrika*, 28:321–377.

[Hyrskykari, 2006] Hyrskykari, A. (2006). *Eyes in attentive interfaces: Experiences from creating iDict, a gaze-aware reading aid.* Ph.D., University of Tampere.

[Irani et al., 2007] Irani, F., Platek, S. M., Bunce, S., Ruocco, A. C., and Chute, D. (2007). Functional near infrared spectroscopy (fNIRS): An emerging neuroimaging technology with important applications for the study of brain disorders. *The Clinical Neuropsychologist*, 21(1):9–37. PMID: 17366276.

[ISO, 1998] ISO (1998). Ergonomic requirements for office work with visual display terminals – part 11: Guidance on usability. International standard 9241-11, ISO/TC 159.

[Istance et al., 2008] Istance, H., Bates, R., Hyrskykari, A., and Vickers, S. (2008). Snap clutch, a moded approach to solving the midas touch problem. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, pages 221–228. ACM.

[Jacko, 2012] Jacko, J. A., editor (2012). *The Human-Computer Interaction handbook.* CRC Press.

[Jankowski, 2011] Jankowski, J. (2011). A taskonomy of 3d web use. In *Proceedings of the 16th International Conference on 3D Web Technology*, pages 93–100. ACM.

[Jasper, 1958] Jasper, H. H. (1958). The ten twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology*, 10:371–375.

[Kammerer et al., 2008] Kammerer, Y., Scheiter, K., and Beinhauer, W. (2008). Looking my way through the menu: The impact of menu design and multimodal input on gaze-based menu selection. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, pages 213–220. ACM.

[Karim et al., 2006] Karim, A. A., Hinterberger, T., Richter, J., Mellinger, J., Neumann, N., Flor, H., Kübler, A., and Birbaumer, N. (2006). Neural internet: Web surfing with brain potentials for the completely paralyzed. *Neurorehabilitation and Neural Repair*, 20(4):508–515.

[Kelly et al., 2004] Kelly, S., Lalor, E., Finucane, C., and Reilly, R. (2004). A comparison of covert and overt attention as a control option in a steady-state visual evoked potential-based brain computer interface. In *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, volume 2, pages 4725–4728. IEEE.

[Kelly et al., 2005] Kelly, S. P., Lalor, E. C., Reilly, R. B., and Foxe, J. J. (2005). Visual spatial attention tracking using high-density SSVEP data for independent brain-computer communication. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 13(2):172–178.

[Kida et al., 2004] Kida, T., Nishihira, Y., Hatta, A., Wasaka, T., Tazoe, T., Sakajiri, Y., Nakata, H., Kaneda, T., Kuroiwa, K., Akiyama, S., et al. (2004). Resource allocation and somatosensory P300 amplitude during dual task: effects of tracking speed and predictability of tracking direction. *Clinical Neurophysiology*, 115(11):2616–2628.

[Klem et al., 1999] Klem, G. H., Lüders, H. O., Jasper, H., Elger, C., et al. (1999). The ten-twenty electrode system of the international federation. *Electroencephalogr Clin Neurophysiol*, 52(3).

[Kos' Myna and Tarpin-Bernard, 2013] Kos' Myna, N. and Tarpin-Bernard, F. (2013). Evaluation and comparison of a multimodal combination of BCI paradigms and eye-tracking in a gaming context. *IEEE Transactions on Computational Intelligence and AI in Games (T-CIAIG)*, pages 150–154.

[Kouroupetroglou, 2014] Kouroupetroglou, C. (2014). *Enhancing the Human Experience through Assistive Technologies and E-Accessibility.* IGI Global.

[Krizhevsky et al., 2012] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105.

[Krusienski et al., 2008] Krusienski, D. J., Sellers, E. W., McFarland, D. J., Vaughan, T. M., and Wolpaw, J. R. (2008). Toward enhanced P300 speller performance. *Journal of neuroscience methods*, 167(1):15–21.

[Kübler et al., 2005] Kübler, A., Nijboer, F., Mellinger, J., Vaughan, T. M., Pawelzik, H., Schalk, G., McFarland, D. J., Birbaumer, N., and Wolpaw, J. R. (2005). Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface. *Neurology*, 64(10):1775–1777.

[Lalor et al., 2005] Lalor, E. C., Kelly, S. P., Finucane, C., Burke, R., Smith, R., Reilly, R. B., and McDarby, G. (2005). Steady-state VEP-based brain-computer interface control in an immersive 3D gaming environment. *EURASIP Journal on Applied Signal Processing.*

[Lamti et al., 2014a] Lamti, H. A., Khelifa, M. M. B., Alimi, A. M., and Gorce, P. (2014a). Effect of fatigue on ssvep during virtual wheelchair navigation. *Journal of theorical and applied information technology*, 65:1–10.

[Lamti et al., 2014b] Lamti, H. A., Khelifa, M. M. B., Alimi, A. M., and Gorce, P. (2014b). Influence of mental fatigue on P300 and SSVEP during virtual wheelchair navigation. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1255–1258.

[Law et al., 2009] Law, E. L.-C., Roto, V., Hassenzahl, M., Vermeeren, A. P., and Kort, J. (2009). Understanding, scoping and defining user experience: a survey approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 719–728. ACM.

[Lécuyer et al., 2013] Lécuyer, A., George, L., and Marchal, M. (2013). Toward adaptive VR simulators combining visual, haptic, and brain-computer interfaces. *IEEE Computer Graphics and Applications*, 33(5):18–23.

[Lécuyer et al., 2008] Lécuyer, A., Lotte, F., Reilly, R. B., Leeb, R., Hirose, M., Slater, M., et al. (2008). Brain-computer interfaces, virtual reality, and videogames. *IEEE Computer*, 41(10):66–72.

[Lee et al., 2009] Lee, C.-h., Kwon, J.-w., Kim, G.-d., Hong, J.-e., Shin, D.-S., and Lee, D.-h. (2009). A study on EEG based concentration transmission and brain computer interface application. *Journal of the Institute of Electronics Engineers of Korea SC*, 46(2):41–46.

[Lee et al., 2010] Lee, P.-L., Sie, J.-J., Liu, Y.-J., Wu, C.-H., Lee, M.-H., Shu, C.-H., Li, P.-H., Sun, C.-W., and Shyu, K.-K. (2010). An SSVEP-actuated brain computer interface using phase-tagged flickering sequences: a cursor system. *Annals of biomedical engineering*, 38(7):2383–2397.

[Leeb et al., 2013a] Leeb, R., Gwak, K., Kim, D.-S., and del R Millan, J. (2013a). Freeing the visual channel by exploiting vibrotactile BCI feedback. In *Engineering in Medicine and Biology Society (EMBC)*, pages 3093–3096.

[Leeb et al., 2013b] Leeb, R., Lancelle, M., Kaiser, V., Fellner, D. W., and Pfurtscheller, G. (2013b). Thinking penguin: Multimodal brain-computer interface control of a VR game. *IEEE Transactions on computational intelligence and AI in games*, 5(2):117–128.

[Leeb et al., 2007] Leeb, R., Lee, F., Keinrath, C., Scherer, R., Bischof, H., and Pfurtscheller, G. (2007). Brain-computer communication: Motivation, aim, and impact of exploring a virtual apartment. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 15:473–482.

[Legény et al., 2011] Legény, J., Viciana-Abad, R., and Lécuyer, A. (2011). Navigating in virtual worlds using a self-paced SSVEP-based brain-computer interface with integrated stimulation and real-time feedback. *Presence*, 20(6):529–544.

[Legény et al., 2013] Legény, J., Viciana-Abad, R., and Lécuyer, A. (2013). Toward contextual SSVEP-based BCI controller: Smart activation of stimuli and control weighting. *IEEE Transactions on Computational Intelligence and AI in Games*, 5(2):111–116.

[Leuthardt et al., 2006] Leuthardt, E., Miller, K., Schalk, G., Rao, R. P. N., and Ojemann, J. (2006). Electrocorticography-based brain computer interface-the seattle experience. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 14(2):194–198.

[Li et al., 2010] Li, Y., Long, J., Yu, T., Yu, Z., Wang, C., Zhang, H., and Guan, C. (2010). A hybrid BCI system for 2-d asynchronous cursor control. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 4205–4208.

[Lin et al., 2006] Lin, Z., Zhang, C., Wu, W., and Gao, X. (2006). Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs. *IEEE Transactions on Biomedical Engineering*, 53(12):2610–2614.

[Liu et al., 2012] Liu, S. S., Rawicz, A., Rezaei, S., Ma, T., Zhang, C., Lin, K., and Wu, E. (2012). An eye-gaze tracking and human computer interface system for people with als and other locked-in diseases. *Journal of Medical and Biological Engineering*, 32(2):111–116.

[Lotte, 2008] Lotte, F. (2008). *Study of Electroencephalographic Signal Processing and Classification Techniques towards the use of Brain-Computer Interfaces in Virtual Reality Applications.* PhD thesis, INSA Rennes.

[Lotte, 2011] Lotte, F. (2011). Generating artificial EEG signals to reduce BCI calibration time. In *5th International Brain-Computer Interface Workshop*, pages 176–179.

[Lotte et al., 2013] Lotte, F., Larrue, F., and Mühl, C. (2013). Flaws in current human training protocols for spontaneous brain-computer interfaces: lessons learned from instructional design. *Frontiers in Human Neuroscience.*

[Lotte et al., 2010] Lotte, F., Van Langhenhove, A., Lamarche, F., Ernest, T., Renard, Y., Arnaldi, B., and Lécuyer, A. (2010). Exploring Large Virtual Environments by Thoughts using a Brain-Computer Interface based on Motor Imagery and High-Level Commands. *Presence: Teleoperators and Virtual Environments*, 19(1):54–70.

[MacKenzie and Buxton, 1992] MacKenzie, I. S. and Buxton, W. (1992). Extending fitts' law to two-dimensional tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 219–226.

[Majaranta et al., 2009] Majaranta, P., Ahola, U.-K., and Špakov, O. (2009). Fast gaze typing with an adjustable dwell time. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 357–360.

[Majaranta and Räihä, 2007] Majaranta, P. and Räihä, K.-J. (2007). Text entry by gaze: Utilizing eye-tracking. *Text entry systems: Mobility, accessibility, universality*, pages 175–187.

[Mankoff et al., 2002] Mankoff, J., Dey, A., Batra, U., and Moore, M. (2002). Web accessibility for low bandwidth input. In *Proceedings of the fifth international ACM conference on Assistive technologies*, pages 17–24. ACM.

[Manyakov et al., 2013] Manyakov, N. V., Chumerin, N., Robben, A., Combaz, A., van Vliet, M., and Van Hulle, M. M. (2013). Sampled sinusoidal stimulation profile and multichannel fuzzy logic classification for monitor-based phase-coded SSVEP brain–computer interfacing. *Journal of neural engineering*, 10(3):036011.

[Maquet et al., 1997] Maquet, P., Degueldre, C., Delfiore, G., Aerts, J., Péters, J.-M., Luxen, A., and Franck, G. (1997). Functional neuroanatomy of human slow wave sleep. *The Journal of Neuroscience*, 17(8):2807–2812.

[Mason and Birch, 2003] Mason, S. G. and Birch, G. E. (2003). A general framework for brain-computer interface design. *IEEE transactions on neural systems and rehabilitation engineering*, 11(1):70–85.

[Mattiocco et al., 2006] Mattiocco, M., Babiloni, F., Mattia, D., Bufalari, S., Sergio, S., Salinari, S., Marciani, M., and Cincotti, F. (2006). Neuroelectrical source imaging of mu rhythm control for BCI applications. In *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE*, pages 980–983.

[McFarland et al., 2008] McFarland, D., Krusienski, D., Sarnacki, W., and Wolpaw, J. (2008). Emulation of computer mouse control with a noninvasive brain-computer interface. *Journal of Neural Engineering*, 5:101–110.

[McFarland et al., 1997] McFarland, D. J., McCane, L. M., David, S. V., and Wolpaw, J. R. (1997). Spatial filter selection for EEG-based communication. *Electroencephalography and Clinical Neurophysiology*, 103(3):386 – 394.

[McFarland et al., 2011] McFarland, D. J., Sarnacki, W. A., Townsend, G., Vaughan, T., and Wolpaw, J. R. (2011). The P300-based brain–computer interface (bci): effects of stimulus rate. *Clinical neurophysiology*, 122(4):731–737.

[Megalingam et al., 2013] Megalingam, R. K., Thulasi, A. A., Krishna, R. R., Venkata, M. K., BV, A. G., and Dutt, T. U. (2013). Thought controlled wheelchair using EEG acquisition device. In *3rd International Conference on Advancements in Electronics and Power Engineering*.

[Mercier-Ganady et al., 2013] Mercier-Ganady, J., Loup-Escande, E., George, L., Busson, C., Marchal, M., and Lécuyer, A. (2013). Can we use a brain-computer interface and manipulate a mouse at the same time? In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, pages 69–72.

[Mercier-Ganady et al., 2015] Mercier-Ganady, J., Marchal, M., and Lécuyer, A. (2015). BC-invisibility power: introducing optical camouflage based on mental activity in augmented reality. In *Proceedings of the 6th Augmented Human International Conference*, pages 97–100. ACM.

[Middendorf et al., 2000] Middendorf, M., McMillan, G., Calhoun, G., and Jones, K. (2000). Brain-computer interfaces based on steady-state visual evoked response. *IEEE Trans. Rehab. Eng.*

[Miniotas et al., 2006] Miniotas, D., Špakov, O., Tugoy, I., and MacKenzie, I. S. (2006). Speech-augmented eye gaze interaction with small closely spaced targets. In *Proceedings of the 2006 symposium on Eye tracking research & applications*, pages 67–72. ACM.

[Morgan et al., 1996] Morgan, S., Hansen, J., and Hillyard, S. (1996). Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proceedings of the National Academy of Sciences*, 93(10):4770–4774.

[Mugler et al., 2010] Mugler, E. M., Ruf, C. A., Halder, S., Bensch, M., and Kubler, A. (2010). Design and implementation of a P300-based brain-computer interface for controlling an internet browser. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(6):599–609.

[Müller et al., 2008] Müller, K.-R., Tangermann, M., Dornhege, G., Krauledat, M., Curio, G., and Blankertz, B. (2008). Machine learning for real-time single-trial EEG-analysis: From brain-computer interfacing to mental state monitoring. *Journal of Neuroscience Methods*, 167(1):82 – 90.

[Müller-Putz et al., 2008] Müller-Putz, G., Scherer, R., Brunner, C., Leeb, R., and Pfurtscheller, G. (2008). Better than random: A closer look on BCI results. *International Journal of Bioelectromagnetism*, 10(EPFL-ARTICLE-164768):52–55.

[Müller-Putz et al., 2006] Müller-Putz, G., Scherer, R., Neuper, C., and Pfurtscheller, G. (2006). Steady-state somatosensory evoked potentials: suitable brain signals for brain-computer interfaces? *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14:30–37.

[Müller-Putz et al., 2011] Müller-Putz, G. R., Breitwieser, C., Cincotti, F., Leeb, R., Schreuder, M., Leotta, F., Tavella, M., Bianchi, L., Kreilinger, A., Ramsay, A., et al. (2011). Tools for brain-computer interaction: a general concept for a hybrid BCI (hBCI). *Frontiers in neuroinformatics*, 5(30):30.

[Müller-Putz et al., 2005] Müller-Putz, G. R., Scherer, R., Brauneis, C., and Pfurtscheller, G. (2005). Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components. *Journal of neural engineering*, 2(4):123–130.

[Nakanishi et al., 2014] Nakanishi, M., Wang, Y., Wang, Y.-T., Mitsukura, Y., and Jung, T.-P. (2014). A high-speed brain speller using steady-state visual evoked potentials. *International journal of neural systems*, 24(06):1450019.

[Naveed et al., 2012] Naveed, K., Iqbal, J., and ur Rehman, H. (2012). Brain controlled human robot interface. In *Robotics and Artificial Intelligence (ICRAI), 2012 International Conference on*, pages 55–60. IEEE.

[Neuper et al., 1999] Neuper, C., Schlögl, A., and Pfurtscheller, G. (1999). Enhancement of left-right sensorimotor EEG differences during feedback-regulated motor imagery. *Journal of Clinical Neurophysiology*, 16(4):373–382.

[NG et al., 2012] NG, K. B., Bradley, A. P., and Cunnington, R. (2012). Stimulus specificity of a steady-state visual-evoked potential-based brain-computer interface. *Journal of Neural Engineering*, 9(3).

[Nicolas-Alonso and Gomez-Gil, 2012] Nicolas-Alonso, L. F. and Gomez-Gil, J. (2012). Brain computer interfaces, a review. *Sensors*, 12(2):1211–1279.

[Nijholt et al., 2009] Nijholt, A., Bos, D. P.-O., and Reuderink, B. (2009). Turning shortcomings into challenges: Brain–computer interfaces for games. *Entertainment Computing*, 1(2):85–94.

[Nijholt and Gürkök, 2013] Nijholt, A. and Gürkök, H. (2013). Multi-brain games: Cooperation and competition. In *Universal Access in Human-Computer Interaction. Design Methods, Tools, and Interaction Techniques for eInclusion*, pages 652–661. Springer.

[Nijholt and Tan, 2008] Nijholt, A. and Tan, D. F. (2008). Brain-computer interfacing for intelligent systems. *Intelligent Systems, IEEE*, 23(3):72–79.

[Norman and Draper, 1986] Norman, D. A. and Draper, S. W., editors (1986). *User centered system design: new perspectives on Human-Computer Interaction*. Lawrence Erlbaum.

[Obermaier et al., 2001] Obermaier, B., Guger, C., Neuper, C., and Pfurtscheller, G. (2001). Hidden markov models for online classification of single trial EEG data. *Pattern recognition letters*, 22(12):1299–1309.

[Owen et al., 2005] Owen, A. M., McMillan, K. M., Laird, A. R., and Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human brain mapping*, 25(1):46–59.

[Pastor et al., 2003] Pastor, M. A., Artieda, J., Arbizu, J., Valencia, M., and Masdeu, J. C. (2003). Human cerebral acivation during stady-state visual-evoked responses. *The Journal of Neuroscience*, 23(37):11621–11627.

[Peck et al., 2013] Peck, E. M. M., Beste, F. Y., Ottley, A., Jacob, R. J. K., and Chang, R. (2013). Using fNIRS brain sensing to evaluate information visualization interfaces. *CHI*.

[Perrin, 2012] Perrin, M. (2012). *Coadaptation cerveau machine pour une interaction optimale : application au P300-speller*. PhD thesis, Universite Claude Bernard - Lyon 1.

[Pfurtscheller et al., 2010] Pfurtscheller, G., Allison, B. Z., Bauernfeind, G., Brunner, C., Solis Escalante, T., Scherer, R., Zander, T. O., Mueller-Putz, G., Neuper, C., and Birbaumer, N. (2010). The hybrid BCI. *Frontiers in Neuroscience*, 4(3).

[Pfurtscheller and Da Silva, 1999] Pfurtscheller, G. and Da Silva, F. L. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical neurophysiology*, 110(11):1842–1857.

[Pfurtscheller et al., 1998] Pfurtscheller, G., Neuper, C., Schlogl, A., and Lugger, K. (1998). Separability of eeg signals recorded during right and left motor imagery using adaptive autoregressive parameters. *IEEE transactions on Rehabilitation Engineering*, 6(3):316–325.

[Philips et al., 2007] Philips, J., del R.Millan, J., Vanacker, G., Lew, E., Galan, F., Ferrez, P., Van Brussel, H., and Nuttin, M. (2007). Adaptive shared control of a brain-actuated simulated wheelchair. In *IEEE International Conference on Rehabilitation Robotics*, pages 408–414.

[Pichiorri et al., 2011] Pichiorri, F., Fallani, F. D. V., Cincotti, F., Babiloni, F., Molinari, M., Kleih, S., Neuper, C., Kübler, A., and Mattia, D. (2011). Sensorimotor rhythm-based brain–computer interface training: the impact on motor cortical responsiveness. *Journal of neural engineering*, 8(2):025020.

[Pineda, 2005] Pineda, J. (2005). The functional significance of mu rhythms: translating "seeing" and "hearing" into "doing". *Brain research reviews*.

[Pivik and Harman, 1995] Pivik, R. and Harman, K. (1995). A reconceptualization of EEG alpha activity as an index of arousal during sleep: all alpha activity is not equal. *Journal of sleep research*, 4(3):131–137.

[Plass-Oude Bos et al., 2011] Plass-Oude Bos, D., Gürkök, H., van der Laar, B., Nijboer, F., and Nijholt, A. (2011). User experience evaluation in BCI: Mind the gap! *International Journal of Bioelectromagnetism*, 13(1):48–49.

[Pratt et al., 2011] Pratt, N., Willoughby, A., and Swick, D. (2011). Effects of working memory load on visual selective attention: behavioral and electrophysiological evidence. *Frontiers in human neuroscience*, 5:57.

[Purves et al., 2012] Purves, Augustine, Fitzpatrick, Hall, Lamantia, McNamara, and Williams (2012). *Neurosciences*. De Boeck.

[Quan et al., 2013] Quan, L., Kun, C., Qingsong, A., and Xie, S. Q. (2013). Review : Recent development of signal processing algorithms for SSVEP-based brain computer interfaces. *Journal of biomedical and Biological Engineering*, 34:299–309.

[Rebsamen et al., 2007] Rebsamen, B., Chee, L. T., Zeng, Q., and Ang Jr, M. H. (2007). Controlling a wheelchair indoors using thought. *IEEE Computer Society*.

[Regan, 1966] Regan, D. (1966). An effect of stimulus colour on average steady-state potentials evoked in man. *Nature*, 210:1056–1057.

[Renard et al., 2010] Renard, Y., Lotte, F., Gibert, G., Congedo, M., Maby, E., Delannoy, V., Bertrand, O., and Lécuyer, A. (2010). Openvibe: an open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments. *Presence: teleoperators and virtual environments*, 19(1):35–53.

[Reuderink et al., 2009] Reuderink, B., Nijholt, A., and Poel, M. (2009). Affective pacman: A frustrating game for brain-computer interface experiments. In *International Conference on Intelligent Technologies for Interactive Entertainment*, pages 221–227. Springer.

[Richardson et al., 2003] Richardson, J. C., D, P., and D, K. S. P. (2003). Examining social presence in online courses in relation to students' perceived learning and satisfaction. *Journal of Asynchronous Learning Networks*, 7:68–88.

[Samar et al., 1999] Samar, V. J., Bopardikar, A., Rao, R., and Swartz, K. (1999). Wavelet analysis of neuroelectric waveforms: A conceptual tutorial. *Brain and Language*, 66(1):7 – 60.

[Sauvan et al., 2009] Sauvan, J.-B., Lécuyer, A., Lotte, F., and Casiez, G. (2009). A performance model of selection techniques for P300-based brain-computer interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 2205–2208, New York, NY, USA. ACM.

[Schatzschneider et al., 2016] Schatzschneider, C., Bruder, G., and Steinicke, F. (2016). Who turned the clock? effects of manipulated zeitgebers, cognitive load and immersion on time estimation. *IEEE transactions on visualization and computer graphics*, 22(4):1387–1395.

[Scherer et al., 2007] Scherer, R., Müller-Putz, G. R., and Pfurtscheller, G. (2007). Self-initiation of EEG-based brain-computer communication using the heart rate response. *Journal of Neural Engineering*, 4:23–29.

[Schlögl et al., 2007] Schlögl, A., Keinrath, C., Zimmermann, D., Scherer, R., Leeb, R., and Pfurtscheller, G. (2007). A fully automated correction method of EOG artifacts in EEG recordings. *Clinical Neurophysiology*, 118(1):98 – 104.

[Sears and Shneiderman, 1991] Sears, A. and Shneiderman, B. (1991). High precision touchscreens: design strategies and comparisons with a mouse. *International Journal of Man-Machine Studies*, 34(4):593–613.

[Silver et al., 2016] Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al. (2016). Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587):484–489.

[Singh et al., 1990] Singh, H., Hines, E., Stocks, N., and Syan, C. (1990). Engineering the brain signal : Preprocessing. *Psychological reviews*.

[Singla et al., 2013] Singla, R., Khosla, A., and Jha, R. (2013). Influence of stimuli color on steady-state visual evoked potentials based BCI wheelchair control. *Journal of Biomedical Science and Engineering*, 6(11):1050.

[Singla et al., 2014] Singla, R., Khosla, A., and Jha, R. (2014). Influence of stimuli colour in SSVEP-based BCI wheelchair control using support vector machines. *Journal of medical engineering & technology*, 38(3):125–134.

[Sirvent et al., 2010] Sirvent, J. L., Azorín, J. M., Iáñez, E., Úbeda, A., and Fernández, E. (2010). P300-based brain-computer interface for internet browsing. In *Trends in practical applications of agents and multiagent systems*, pages 615–622. Springer.

[Smith, 1997] Smith, S. W. (1997). *The Scientist and Engineer's Guide to Digital Signal Processing.* ISBN.

[Speier et al., 2016] Speier, W., Arnold, C., and Pouratian, N. (2016). Integrating language models into classifiers for BCI communication: a review. *Journal of neural engineering*, 13(3):031002.

[Stillman, 1993] Stillman, J. A. (1993). Response selection, sensitivity, and taste-test performance. *Perception & Psychophysics*, 54(2):190–194.

[Strangman et al., 2002] Strangman, G., Boas, D. A., and Sutton, J. P. (2002). Non-invasive neuroimaging using near-infrared light. *Biological psychiatry*, 52(7):679–693.

[Tamir et al., 2007] Tamir, R., Dickstein, R., and Huberman, M. (2007). Integration of motor imagery and physical practice in group treatment applied to subject with parkinson's disease. *Neurorehabilitation and Neural Repair*, 21(1):68–75.

[Tanner, 2011] Tanner, A. (2011). *Automatic seizure detection using a two-dimensional EEG feature space.* PhD thesis, degree of Master of Science in Technology.

[Tello et al., 2015] Tello, R. J. M. G., Müller, S. M. T., Ferreira, A., and Bastos, T. F. (2015). Comparison of the influence of stimuli color on steady-state visual evoked potentials. *Research on Biomedical Engineering*, 31(3):218–231.

[Teplan, 2002] Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement science review*, 2(2):1–11.

[Thomas et al., 2014] Thomas, E., Daucé, E., Devlaminck, D., Mahé, L., Carpentier, A., Munos, R., Perrin, M., Maby, E., Mattout, J., Papadopoulo, T., et al. (2014). Coadapt P300 speller: optimized flashing sequences and online learning. In *6th International Brain Computer Interface Conference*.

[Trachel et al., 2013] Trachel, R., Brochier, T., and Clerc, M. (2013). Enhancing Visuospatial Attention Performance with Brain-Computer Interfaces. In *ACM SIGCHI Conference on Human Factors in Computing Systems*.

[Treder and Blankertz, 2010] Treder, M. S. and Blankertz, B. (2010). (c) overt attention and visual speller design in an erp-based brain-computer interface. *Behavioral and brain functions*, 6(1):1.

[Trewin and Pain, 1999] Trewin, S. and Pain, H. (1999). Keyboard and mouse errors due to motor disabilities. *International Journal of Human-Computer Studies*, 50(2):109–144.

[Urdapilleta et al., 2001] Urdapilleta, I., Saint-Denis, C., Nu, C. T., and de Kermadec, F. H. (2001). *Traité d'évaluation sensorielle*. Dunod.

[Vagias, 2006] Vagias, W. M. (2006). Likert-type scale response anchors.

[Vaid et al., 2015] Vaid, S., Singh, P., and Kaur, C. (2015). EEG signal analysis for BCI interface: A review. In *2015 Fifth International Conference on Advanced Computing & Communication Technologies*, pages 143–147. IEEE.

[Van Gerven et al., 2009] Van Gerven, M., Bahramisharif, A., Heskes, T., and Jensen, O. (2009). Selecting features for BCI control based on a covert spatial attention paradigm. *Neural Networks*, 22(9):1271 – 1277.

[Vanderwolf, 2000] Vanderwolf, C. (2000). Are neocortical gamma waves related to consciousness? *Brain Research*, 855(2):217 – 224.

[Vaughan et al., 2006] Vaughan, T., McFarland, D., Schalk, G., Sarnacki, W., Krusienski, D., Sellers, E., and Wolpaw, J. (2006). The wadsworth BCI research and development program: at home with BCI. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14:229–233.

[Velichkovsky et al., 1997] Velichkovsky, B., Sprenger, A., and Unema, P. (1997). Towards gaze-mediated interaction: Collecting solutions of the "midas touch problem". In *Human-Computer Interaction INTERACT'97*, pages 509–516. Springer.

[Verde et al., 2006] Verde, M. F., Macmillan, N. A., and Rotello, C. M. (2006). Measures of sensitivity based on a single hit rate and false alarm rate : The accuracy, precision, and robustness of d', az, and a'. *Perception & psychophysics*, 68(4):643–654.

[Vidal, 1973] Vidal, J. J. (1973). Toward direct brain-computer communication. *Annual review of Biophysics and Bioengineering*, 2(1):157–180.

[Volosyak, 2011] Volosyak, I. (2011). SSVEP-based Bremen-BCI interface - boosting information transfer rates. *Journal of neural engineering*, 8(3):036020.

[Waldert, 2016] Waldert, S. (2016). Invasive vs. non-invasive neuronal signals for brain-machine interfaces: Will one prevail? *Frontiers in Neuroscience*, 10:295.

[Wang et al., 2010a] Wang, N., Qian, T., Zhuo, Q., and Gao, X. (2010a). Discrimination between idle and work states in BCI based on SSVEP. In *Advanced computer control (ICACC), 2010 2nd international conference on*, volume 4, pages 355–358. IEEE.

[Wang et al., 2010b] Wang, Y., Wang, Y., and Jung, T.-P. (2010b). Visual stimulus design for high-rate SSVEP BCI. *Electronics letters*, 46(15):1.

[Watter et al., 2001] Watter, S., Geffen, G. M., and Geffen, L. B. (2001). The n-back as a dual-task: P300 morphology under divided attention. *Psychophysiology*, 38(06):998–1003.

[Weinreich et al., 2008] Weinreich, H., Obendorf, H., Herder, E., and Mayer, M. (2008). Not quite the average: An empirical study of web use. *ACM Transactions on the Web (TWEB)*, 2(1):5.

[Weiskopf et al., 2004] Weiskopf, N., Mathiak, K., Bock, S., Scharnowski, F., Veit, R., Grodd, W., Goebel, R., and Birbaumer, N. (2004). Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI). *Biomedical Engineering, IEEE Transactions on*, 51(6):966–970.

[Wickens, 2008] Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(3):449–455.

[Wieneke et al., 1980] Wieneke, G., Deinema, C., Spoelstra, P., Van Leeuwen, W. S., and Versteeg, H. (1980). Normative spectral data on alpha rhythm in male adults. *Electroencephalography and clinical neurophysiology*, 49(5-6):636–645.

[Wobbrock et al., 2011] Wobbrock, J. O., Findlater, L., Gergle, D., and Higgins, J. J. (2011). The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proc. of CHI*, pages 143–146.

[Wolpaw et al., 2002] Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., and Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6):767 – 791.

[Yin et al., 2015] Yin, E., Zhou, Z., Jiang, J., Yu, Y., and Hu, D. (2015). A dynamically optimized SSVEP brain–computer interface (BCI) speller. *IEEE Transactions on Biomedical Engineering*, 62(6):1447–1456.

[Yin et al., 2009] Yin, J., Jiang, D., and Hu, J. (2009). Design and application of brain-computer interface web browser based on VEP. In *2009 International Conference on Future BioMedical Information Engineering (FBIE)*, pages 77–80.

[Zander et al., 2010a] Zander, T. O., Gaertner, M., Kothe, C., and Vilimek, R. (2010a). Combining eye gaze input with a brain-computer interface for touchless human-computer interaction. *International Journal of Human-Computer Interaction*, 27(1):38–51.

[Zander and Kothe, 2011] Zander, T. O. and Kothe, C. (2011). Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general. *Journal of Neural Engineering*, 8(2):025005.

[Zander et al., 2010b] Zander, T. O., Kothe, C., Jatzev, S., and Gaertner, M. (2010b). Enhancing human-computer interaction with input from active and passive brain-computer interfaces. In *Brain-computer interfaces*, pages 181–199. Springer.

[Zhang et al., 2014] Zhang, Y., Zhou, G., Jin, J., Wang, X., and Cichocki, A. (2014). Frequency recognition in SSVEP-based BCI using multiset canonical correlation analysis. *International journal of neural systems*, 24(04):1450013.

[Zhu et al., 2010a] Zhu, D., Bieger, J., Molina, G. G., and Aarts, R. M. (2010a). A survey of stimulation methods used in SSVEP-based BCIs. *Computational Intelligence and Neuroscience*, 2010.

[Zhu et al., 2010b] Zhu, D., Bieger, J., Molina, G. G., and Aarts, R. M. (2010b). A survey of stimulation methods used in SSVEP-based BCIs. *Intell. Neuroscience*, 2010:1:1–1:12.

[Zhu and Yang, 2002] Zhu, J. and Yang, J. (2002). Subpixel eye gaze tracking. In *Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on*, pages 124–129.