Extraction de potentiels évoqués P300 pour les interfaces cerveau-machine

Bertrand RIVET, Antoine SOULOUMIAC CEA, LIST, Laboratoire d'Electronique et de Traitement du Signal, Gif sur Yvette, F-91191, France bertrand.rivet@cea.fr,antoine.souloumiac@cea.fr

Résumé – Les interfaces cerveau-machine permettent une communication sans contrôle musculaire directement entre le cerveau de l'utilisateur et un ordinateur. Cet article traite de l'estimation des potentiels évoqués P300 pour les interfaces cerveau-machine de type *P300 Speller* qui permettent d'épeler des mots. Nous proposons ici de rehausser ces potentiels P300 à partir des signaux électroencéphalogrammes bruts grâce à une méthode non-supervisée d'estimation du sous-espace signal. La classification des signaux estimés est alors faite par un classifieur de type séparateur à vaste marge linéaire (SVM). Nous montrons finalement que la méthode proposée permet d'augmenter le débit de mots épelés tout en simplifiant l'apprentissage du classifieur.

Abstract – Brain-computer interfaces (BCI) enable non-muscular powerful channel for communicating thanks to direct communication between the user's brain and a computer. In this article, the presented BCI concerns the P300 speller which enables people to write a text on a computer by estimating P300 evoked potentials. In this work, an unsupervised algorithm is introduced for signal subspace estimation: the raw EEG are thus enhanced by projection on the estimated subspace. The detection of P300 potentials is performed by a linear support vector machine (SVM). The proposed method is finally shown to be efficient since it increases the number of spelt characters per minute.

1 Introduction

Les interfaces cerveau-machine ("Brain Computer Interface" ou BCI en anglais) permettent une communication directe entre le cerveau de l'utilisateur et un ordinateur grâce à l'analyse des signaux électroencéphalographiques qui reflètent l'activité du cerveau [1, 2]. Ce type d'interfaces homme-machine, qui a été rendu possible par les avancées faites dans la compréhension du fonctionnement du cerveau, fournit un nouveau moyen de communication non musculaire. Ces interfaces permettent ainsi notamment aux personnes atteintes de problèmes neuromusculaires ou du syndrome d'enfermement ("locked-in syndrom" en anglais) de pouvoir interagir et communiquer avec leur environnement. Pour cela, les utilisateurs peuvent par exemple apprendre à contrôler certaines ondes de leur cerveau ou la BCI peut exploiter les réponses naturelles du cerveau à des stimuli externes (voir [1] pour une revue de divers types de marqueurs électrophysiologiques).

La BCI présentée ici permet l'épellation d'un mot. Elle est fondée sur la détection de l'apparition d'un potentiel évoqué dans les électroencéphalogrammes (EEG) en réponse à un certain stimulus visuel. Il s'agit donc de discriminer les épochs contenant ou non le potentiel évoqué en réponse aux stimuli cibles. Malheureusement, d'une part le rapport signal sur bruit des signaux EEG est très faible et d'autres part les signaux EEG peuvent être contaminés par des artéfacts de type musculaires ou oculaires, ainsi divers auteurs ont proposé de rehausser le rapport signal sur bruit en effectuant une analyse en composantes indépendantes des signaux EEG [3, 4]. L'inconvénient de ces méthodes est qu'elles sont supervisées au sens où il est nécessaire, une fois effectuée la séparation, de sélectionner les composantes contenant les potentiels évoqués cibles.

Dans cet article, nous proposons d'estimer le sous-espace contenant les potentiels évoqués cibles par une méthode non supervisée reposant uniquement sur l'information temporelle *a priori* dont nous disposons pendant la phase d'apprentissage (*i.e.* les instants des stimuli). Après avoir décrit la BCI que nous étudions au paragraphe 2.1, le principe d'estimation du sousespace signal est présenté (paragraphe 2.2) ainsi que la méthode de prédiction du caractère épelé (paragraphe 2.3). Enfin, nous terminons par des résultats expérimentaux (paragraphe 3) et une conclusion (paragraphe 4).

2 Méthodologie

L'objectif de cette étude est de proposer une solution simple et non-supervisée pour l'estimation du sous-espace P300 en vue d'augmenter les performances de classification.

2.1 L'interface cerveau-machine "P300 speller"

L'interface cerveau-machine, que nous utilisons ici, est le "*P300 speller*" introduit par Farwell et Donchin [5, 6]. Celle-ci permet d'épeler un mot : pour cela l'utilisateur fixe une matrice 6×6 composée de caractères (Fig. 1). Pour chacun des caractères épelés, les stimuli visuels sont générés par l'illumination de chaque ligne et de chaque colonne dans un ordre aléatoire. L'utilisateur doit alors compter le nombre de fois où la ligne et la colonne contenant le caractère qu'il souhaite épeler s'illumination des lignes et colonnes contenant le caractère à épeler génère l'apparition d'un potentiel P300 (*i.e.* un potentiel positif environ 300ms après le stimulus, *cf.* Fig. 1(b)).

Il s'agit alors de détecter les stimuli qui ont entraîné l'apparition d'un potentiel P300 pour connaître le caractère épelé. Afin de rendre plus robuste l'estimation des potentiels P300, chaque caractère est typiquement épelé 15 fois.



FIG. 1 – Interface *P300 Speller*. Fig. 1(a) : matrice d'épellation des mots avec une ligne illuminée. Fig. 1(b) : potentiels évoqués P300 enregistrés sur l'aire pariétale (électrode C_z) après moyennage sur tous les épochs correspondant aux stimuli cibles (trait continu) et aux stimuli non cibles (trait discontinu).

2.2 Estimation du sous-espace P300

Étant donné que les signaux EEG mesurés sur le scalp de l'utilisateur sont une image de l'activité électrique du cerveau, ils regroupent non seulement les potentiels P300, que l'on cherche à détecter, mais aussi d'autres signaux émis par le cerveau ou encore des artéfacts musculaires ou oculaires. Il en résulte que le rapport signal sur bruit du potentiel P300 est très faible.

De façon à augmenter le rapport signal sur bruit des potentiels P300, nous allons estimer à partir d'une base d'apprentissage et de façon itérative les principales composantes du sousespace P300 (*i.e.* sous-espace contenant les potentiels P300) en prenant en compte la connaissance *a priori* dont nous disposons : c'est-à-dire les instants des stimuli visuels.

Soit $\phi_1(t)$ le signal temporel typique d'un potentiel P300 appelé ici noyau. Pour la base d'apprentissage (pour laquelle on connaît les stimuli qui ont engendré un potentiel P300), il est alors possible de construire un signal de référence des potentiels P300 $r_1(t)$ tel que

$$r_1(t) = \sum_{j=1}^{J} \phi_1(t) * \delta(t - \tau_j),$$
(1)

où $\delta(t)$ est l'impulsion de Dirac, τ_j les instants (connus) des stimuli qui ont engendré un potentiel évoqué P300 et J le nombre total de ces stimuli.

Soit $\mathbf{x}(t) \in \mathbb{R}^{N_c}$ le vecteur colonne des observations, où N_c est le nombre de capteurs. Le but est d'estimer le vecteur colonne \mathbf{b}_1 tel que

$$\mathbf{b}_{1} = \arg\min_{\mathbf{g}} \sum_{t=0}^{T-1} \|\mathbf{g}^{T} \mathbf{x}(t) - r_{1}(t)\|^{2},$$
(2)

où .^T est l'opérateur de transposition. La solution de ce problème est classiquement donnée par

$$\mathbf{b}_1 = (X^T X)^{-1} X^T \mathbf{r}_1, \tag{3}$$

où $X = [\mathbf{x}(0), \dots, \mathbf{x}(T-1)]^T$ et $\mathbf{r}_1 = [r_1(0), \dots, r(T-1)]^T$. Or le noyau $\phi_1(t)$, qui a servi à construire le signal de référence $r_1(t)$, est *a priori* inconnu. Pour surmonter cette difficulté, nous proposons un schéma récursif d'estimation du noyau $\phi_1(t)$.

Soient $\phi_1^{(k-1)}(t)$ l'estimation au pas k-1 du noyau $\phi_1(t)$ et $r_1^{(k-1)}(t)$ l'estimation du signal de référence des potentiels P300 (1) obtenu à partir du noyau $\phi_1^{(k-1)}(t)$. Au pas k, la résolution du problème (2) fournit d'après (3) un vecteur $\mathbf{g}^{(k)}$

$$(X^T X)^{-1} X^T \mathbf{r}_1^{(k-1)}$$
. Ainsi,
 $s_1^{(k)}(t) = \mathbf{g}^{(k)^T} \mathbf{x}(t)$ (4)

est l'estimation du signal $r_1^{(k-1)}(t)$ au sens de (2). Pour mettre à jour l'estimation du noyau $\phi_1(t)$, nous effectuons un épochage du signal $s_1^{(k)}(t)$ sur les potentiels P300 :

$$\forall j \in \{1, \cdots, J\},\ f_j^{(k)}(t) = \left(s_1^{(k)}(t) \times \Pi_{T_{300}}(t - \tau_j)\right) * \delta(t + \tau_j),\ (5)$$

où $\Pi_{T_{300}}(t)$ est la fonction porte valant 1 sur son support $[0, T_{300}]$ et nulle ailleurs (typiquement $T_{300} = 600$ ms). Finalement, nous cherchons maintenant le noyau $\phi_1^{(k)}(t)$ qui soit le plus corrélé possible en moyenne avec tous les épochs $f_j^{(k)}(t)$ contenant les potentiels P300. Ce problème d'optimisation peut s'écrire comme

$$\phi_1^{(k)} = \arg \max_{\|\phi_1\|_2 = 1} \sum_{j=1}^{J} \left| < \mathbf{f}_j^{(k)}, \phi_1 > \right|^2, \tag{6}$$

avec $\mathbf{f}_{j}^{(k)} = [f_{j}^{(k)}(0), \cdots, f_{j}^{(k)}(T_{300} - 1)]^{T}$ et $\phi_{1}^{(k)} = [\phi_{1}^{(k)}(0), \cdots, \phi_{1}^{(k)}(T_{300} - 1)]^{T} < \cdot, \cdot >$ est le produit scalaire entre deux vecteurs. Après des manipulations algébriques simples, il vient que la solution optimale $\phi_{1}^{(k)}$ du problème (6) est le vecteur propre associé à la plus grande valeur propre de la matrice $\sum_{j=1}^{J} \mathbf{f}_{j}^{(k)} \mathbf{f}_{j}^{(k)^{T}}$. Après convergence, ce schéma récursif fournit donc le vecteur $\mathbf{b}_{1} = \mathbf{g}^{(k)}$ support de la composante principale du P300.

Pour estimer un sous-espace P300 de dimension supérieure à un, nous proposons une méthode itérative. Soit \mathbf{b}_i la $i^{\text{ème}}$ composante du sous-espace P300 et soit V_i l'espace engendré par { $\mathbf{b}_1, \dots, \mathbf{b}_i$ }. Pour estimer \mathbf{b}_{i+1} , l'algorithme récursif décrit aux paragraphes précédents est appliqué sur $\mathbf{x}_{i+1}(t)$, où $\mathbf{x}_{i+1}(t)$ est le vecteur des observations $\mathbf{x}(t)$ projeté sur V_i^{\perp} , avec V_i^{\perp} l'espace orthogonal à V_i . L'algorithme itératif complet pour extraire I composantes des potentiels P300 est résumé dans l'algorithme 1.

Nous avons observé sur de nombreuses expériences que le schéma récursif (étapes 6 à 13 de l'algorithme 1) converge en peu d'itérations vers une solution stable indépendamment de l'initialisation du noyau $\phi_i^{(0)}(t)$.

2.3 Prédiction du caractère épelé

Pour faire la discrimination des signaux EEG entre les épochs contenant des potentiels P300 (donc cibles) et les épochs qui n'en contiennent pas, nous proposons d'utiliser un classifieur de type séparateur à vaste marge linéaire (SVM) [7].

Soit $\mathbf{u}^{(l/c)}$ le vecteur caractérisant le stimulus correspondant à l'illumination d'une ligne l ou d'une colonne c. Celui-ci est obtenu à partir de la concaténation des épochs sources $\mathbf{f}_i^{(l/c)} = [f_i^{(l/c)}(0), \dots, f_i^{(l/c)}(T_{300})]^T$, pour $1 \le i \le I$, avec

$$f_i^{(l/c)}(t) = \left(\left(\mathbf{b}_i \, \mathbf{x}(t) \right) \times \Pi_{T_{300}}(t - \tau_{c/l}) \right) * \delta(t + \tau_{c/l}),$$

où \mathbf{b}_i sont les vecteurs définissant le sous-espace P300 obtenus par l'algorithme 1 et $\tau_{c/l}$ est l'instant du stimulus correspondant à la ligne l ou la colonne c. On a alors

$$\mathbf{u}^{(l/c)} = \left[\mathbf{f}_1^{(l/c)^T}, \dots, \mathbf{f}_I^{(l/c)^T}\right]^T.$$
(7)



FIG. 2 – Rehaussement des P300 par projection sur le sous-espace P300. Fig. 2(a) : le tracé du haut correspond aux EEG bruts enregistrés sur l'électrode C_z , le tracé du bas montre le signal $s_1(t) = \mathbf{b}_1 \mathbf{x}(t)$. Les pointillés verticaux correspondent au début des stimuli visuels cibles et les flèches correspondent à une durée de 300ms. Fig. 2(b) : projection des filtres spatiaux \mathbf{b}_i sur les électrodes.

Algorithm 1 Estimation itérative du sous-espace P300. 1: $\mathbf{x}_1(t) = \mathbf{x}(t), X_1 = X.$ 2: for i = 1 à *I* do k = 03: $\kappa = 0$ initialiser $\phi_i^{(0)}(t)$ par un bruit blanc calculer $r_i^{(0)}(t) = \sum_{j=1}^J \phi_i^{(0)}(t) * \delta(t - \tau_j)$ while convergence pas atteinte **do** 4: 5: 6: $k \leftarrow k+1$ 7: $\mathbf{g}^{(k)}_{(l)} = (X_i^T X_i)^{-1} X_i^T \mathbf{r}_i^{(k-1)}$ 8: $\mathbf{g}^{(k)} = \mathbf{g}^{(k)} \mathbf{x}_i(t)$ $\mathbf{g}^{(k)}(t) = \mathbf{g}^{(k)^T} \mathbf{x}_i(t)$ epochage de $s_i^{(k)}(t) \Rightarrow \mathbf{f}_j^{(k)}$ pour $j \in \{1, \cdots, J\}$ (5) 9: 10: trouver $\phi_i^{(k)} = \arg \max_{\|\phi_i\|_2=1} \sum_{j=1}^J \left| < \mathbf{f}_j^{(k)}, \phi_i > \right|^2$, qui est le vecteur propre associé à la plus grande valeur propre de $\sum_{j=1}^J \mathbf{f}_j^{(k)} \mathbf{f}_j^{(k)^T}$ $r_i^{(k)}(t) = \sum_{j=1}^J \phi_i^{(k)}(t) * \delta(t - \tau_j)$, end while 11: 12: 13: $\mathbf{b}_i = \mathbf{g}^{(k)}$ 14projection de $\mathbf{x}_i(t)$ sur l'espace $V_i^{\perp} \Rightarrow x_{i+1}(t), X_{i+1}$ 15: end for 16:

Les paramètres du SVM linéaire sont alors appris à partir des couples $\{\mathbf{u}^{(l/c)}, y^{(l/c)}\}$, où $y^{(l/c)}$ est la classe associée au stimulus : $y^{(l/c)} = 1$ si à la ligne l ou la colonne c est cible et $y^{(l/c)} = 0$ sinon.

Soit $h_{l/c}$ le score de la ligne l ou de la colonne c obtenu par le SVM après une séquence complète d'illumination de toutes les lignes et colonnes (soit 12 stimili correspondant aux 6 lignes et 6 colonnes). Le score total $H_{l/c}(r)$ de la ligne/colonne l/c après r répétitions est finalement donné par

$$H_{l/c}(r) = H_{l/c}(r-1) + h_{l/c},$$
(8)

où $H_{l/c}(0) = 0$. Le caractère prédit à la $r^{\text{ème}}$ répétition est alors celui qui correspond à l'intersection de la ligne l et de la colonne c ayant obtenues les scores $H_{l/c}(r)$ maximaux.

3 Résultats expérimentaux

La méthodologie proposée, pour estimer le sous-espace P300 (paragraphe 2.2) et pour reconnaître le caractère épelé (paragraphe 2.3), a été appliquée sur les données de la compétition BCI 2003 [8] qui sont fournies par Wadsworth Center [9]. Les données correspondent à 64 capteurs EEG échantillonnés à 240Hz. Les EEG bruts ont été préalablement filtrés entre 1 et 20Hz par un filtre d'ordre 4. Plusieurs expériences ont été menées pour analyser le bénéfice apporté par notre approche.

3.1 Estimation du sous-espace P300

Dans une première expérience (Fig. 2), l'estimation du sousespace P300 (paragraphe 2.2) a été appliqué sur la base de données d'apprentissage contenant 39 caractères, fournissant ainsi I filtres spatiaux \mathbf{b}_i dont les trois premiers sont tracés Fig. 2(b). Il est intéressant de remarquer que, pour les premières dimensions du sous-espace P300, les électrodes ayant les plus fortes pondérations sont principalement situées au dessus des zones occipitales et pariétales.

De plus, comme le montre la figure 2(a), la méthode proposée permet de rehausser les potentiels évoqués P300. Le tracé du haut montre les signaux EEG bruts enregistrés sur l'électrode C_z (zone pariétale) : il est difficile de détecter la présence de P300 (déviation positive 300ms après les stimuli cibles représentés ici par les pointillés verticaux). En revanche, la projection des observations $\mathbf{x}(t)$ sur la première composante du sous-espace P300, définie par \mathbf{b}_1 , permet de rehausser les P300 comme le montre le tracé du bas.

3.2 Prédiction du caractère épelé

Dans une seconde série d'expériences, nous comparons les résultats de prédiction des caractères épelés obtenus par la méthode proposée au paragraphe 2.3 avec ceux obtenus par diverses méthodes. Les résultats de classification sont présentés au tableau 1 où est reporté le nombre d'erreurs de prédiction en fonction du nombre de répétitions et de la méthode utilisée.

Dans la première expérience, tous les capteurs ou deux jeux de 10 capteurs pré-sélectionnés manuellement ont été utilisés pour la prédiction des caractères épelés. Dans chacun des cas, le vecteur de caractéristiques à l'entrée du SVM est obtenu par la concaténation des épochs correspondants aux capteurs retenus : il est ainsi de taille (nombre de capteurs)×(taille des épochs). Notons tout d'abord qu'une sélection d'un certain nombre de capteurs permet d'améliorer les performances, comme on peut le voir sur les deux premières lignes du tableau 1. Cependant, le choix des capteurs à utiliser n'est pas facile à réali-

	Nombre de séquences utilisées						
	1	2	3	4	5	10	15
64 canaux (tous les canaux)	20	15	9	7	6	1	0
10 capteurs pré-sélectionnés (arrière du crâne)	16	10	7	2	1	1	0
10 capteurs pré-sélectionnés (avant du crâne)	27	24	22	19	19	12	10
Première composante du sous-espace P300	9	6	3	2	2	0	0
3 premières composantes du sous-espace P300	6	4	2	0	0	0	0

(a) Performances de prédiction

(b) Position des capteurs



TAB. 1 – Nombre de caractères erronés dans les mots de la base de test en fonction du nombre de séquences utilisées et des algorithmes (pas d'optimisation du paramètre de contrainte du SVM).

ser : ainsi, ne choisir qu'un nombre restreint de capteurs situés plutôt à l'arrière du crâne permet d'améliorer les performances, alors qu'en sélectionnant des capteurs situés plutôt à l'avant du crâne les performances sont nettement moins bonnes.

Dans une deuxième expérience, seules les premières composantes $s_i(t)$ du sous-espace P300 sont utilisées de façon à prédire le caractère épelé. Le vecteur de caractéristiques à l'entrée du SVM est alors obtenu par la concaténation des épochs correspondants aux sources estimées $s_i(t)$: il est ainsi de taille (dimension du sous-espace P300)×(taille des épochs). On constate que notre traitement (*cf.* paragraphe 2.2) préalable à la classification permet d'augmenter de façon significative les performances de prédiction (*cf.* les deux dernières lignes du tableau 1) même en n'utilisant qu'un nombre limité de dimensions du sous-espace P300. Ainsi, par exemple, après seulement quatre répétitions la méthode proposée permet d'obtenir 100% de bonne prédiction avec les trois premières composantes du sous-espace P300.

4 Conclusion

Dans cet article, nous avons proposé une nouvelle méthode non-supervisée de traitement des signaux EEG pour les interfaces cerveau-machine de type P300 speller. La méthode proposée permet d'estimer le sous-espace signal P300 par une approche fondée sur un modèle qui n'utilise que la seule information a priori dont on dispose : les instants des stimuli. Les potentiels évoqués P300 sont ainsi rehaussés par projection des signaux EEG bruts sur le sous-espace P300 estimé. La méthode proposée permet donc d'obtenir un vecteur de caractéristiques à l'entrée du classifieur de taille plus petite que les méthodes standard de la littérature. De plus, les résultats ont montré tout l'intérêt de la méthode proposée. En effet, pour un nombre fixé de répétitions, une meilleure prédiction des caractères épelés est obtenues ce qui permet en autre d'augmenter significativement le débit de caractères épelés : de 2 caractères à la minute, à 6 caractères à la minute.

Finalement, nos futurs travaux prendront en compte les variabilités des potentiels évoqués P300 par exemple leur latence et leur amplitude ou bien grâce à l'emploi de classifieurs plus complexes tels que ceux présentés dans [10].

Références

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control." *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, June 2002.

- [2] L. R. Hochberg, M. D. Serruya, G. M. Friehs, J. A. Mukand, M. Saleh, A. H. Caplan, A. Branner, D. Chen, R. D. Penn, and J. P. Donoghue, "Neuronal ensemble control of prosthetic devices by a human with tetraplegia," *Nature*, vol. 442, no. 7099, pp. 164–171, Jul. 2006.
- [3] N. Xu, X. Gao, B. Hong, X. Miao, S. Gao, and F. Yang, "BCI Competition 2003–Data Set IIb : Enhancing P300 Wave Detection Using ICA-Based Subspace Projections for BCI Applications," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1067–1072, June 2004.
- [4] H. Serby, E. Yom-Tov, and G. Inbar, "An improved P300-based brain-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 13, no. 1, pp. 89–98, March 2005.
- [5] L. A. Farwell and E. Donchin, "Talking off the top of your head : toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, December 1988.
- [6] E. Donchin, K. Spencer, and R. Wijesinghe, "The mental prosthesis : assessing the speed of a P300-based braincomputer interface," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 174–179, June 2000.
- [7] K.-R. Muller, S. Mika, G. Ratsch, K. Tsuda, and B. Scholkopf, "An introduction to kernel-based learning algorithms," *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 181–201, March 2001.
- [8] B. Blankertz, K.-R. Müller, G. Curio, T. M. Vaughan, G. Schalk, J. R. Wolpaw, A. Schlögl, C. Neuper, G. Pfurtscheller, T. Hinterberger, M. Schröder, and N. Birbaumer, "The BCI Competition 2003 : Progress and Perspectives in Detection and Discrimination of EEG Single Trials," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1044–1051, June 2004.
- [9] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000 : A General-Purpose Brain-Computer Interface (BCI) System," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1034–1043, June 2004.
- [10] A. Rakotomamonjy, V. Guigue, G. Mallet, and V. Alvarado, "Classification d'EEG pour les Interfaces Cerveau-Machine," in *Proc. GRETSI*, Louvain la neuve, September 2005, pp. 719–722.