Classification de potentiels évoqués P300 par géométrie riemannienne pour les interfaces cerveau-machine EEG.

Alexandre BARACHANT[1], Marco CONGEDO[1], Gijs VAN VEEN[1], Christian JUTTEN[1]

[1]Equipe ViBS (Vision and Brain Signal Processing), GIPSA-lab, CNRS, Université de Grenoble.
Domaine Universitaire, F-38402 Saint Martin d'Hères, France.
alexandre.barachant@gmail.com

Résumé – Cet article présente une nouvelle méthode de classification pour les potentiels évoqués dans le cadre des interfaces cerveau machine (ICMs) EEG. À travers une estimation spécifique des matrices de covariance, ce travail étend l'utilisation de la géométrie Riemannienne, jusqu'alors limitée aux ICMs fondées sur l'imagination motrice, à la classification des potentiel évoqués. En comparaison aux méthodes de l'état de l'art, la méthode présentée offre une augmentation des performances tout en diminuant le nombre de données de nécessaire à la calibration du classifieur et en offrant une meilleure généralisation entre les sessions d'enregistrement.

Abstract – This paper presents a new classification methods for Event Related Potentials (ERP) based on an Riemannian geometry framework. Through a new estimation of covariance matrices, this work extend the use of Riemannian geometry, which was previously limited to SMR-based BCI, to the problem of classification of ERPs. As compared to the state-of-the-art, this new method increases performance, reduces the number of data needed for the calibration and features good generalisation across sessions and subjects.

1 Introduction

Une interface cerveau machine (ICM) est un dispositif permettant à son utilisateur d'interagir directement avec un ordinateur par l'intermédiaire de son activité cérébrale. Il existe deux principaux types d'ICM. Les ICMs fondées sur l'imagerie motrice exploitent les modulations volontaires de l'activité cérébrale produites par l'imagination d'un mouvement d'un membre. Elles permettent à l'utilisateur d'interagir de manière asynchrone. Les ICM fondées sur les potentiels évoqués exploitent les réponses induites par des stimulations, principalement visuelles, pour déterminer les intentions de l'utilisateur. Cet article porte sur ce deuxième type d'ICM, et plus particulièrement sur les ICMs exploitant le potentiel évoqué P300. L'ICM P300 permet à son utilisateur de sélectionner un élément parmi un ensemble d'éléments proposés à l'écran. L'application principale étant le *P300 speller* [11] qui permet d'épeler des mots. Cette ICM est fondée sur la détection de l'apparition d'un potentiel évoqué en réponse à un certain stimulus visuel et plus particulièrement 300 ms après un stimulus rare au sein d'une série de stimuli fréquents. Il s'agit donc de discriminer les essais contenant ou non le potentiel évoqué en réponse aux stimuli cibles. Les éléments sont présentés à l'utilisateur sous forme d'une matrice, généralement de taille 6×6, dont les lignes et les colonnes vont être mises en surbrillance (stimulation) alternativement. Un potentiel évoqué apparaît alors lorsque la ligne ou la colonne contenant l'élément que l'utilisateur souhaite sélectionner est stimulée. La détection de ce potentiel permet donc de remonter au choix de l'utilisateur.

De nouvelles méthodes de classification ont récemment été introduites pour traiter le cas des ICMs fondées sur l'imagerie motrices [1, 2, 3]. Ces méthodes utilisent les matrices de covariance estimées sur des portions de signaux EEG comme caractéristiques du signal. Étant donné un modèle de mélange instantanée des sources cérébrales eléctroencéphalographiques [4], la matrice de covariance contient toute l'information spatiale jusqu'à l'ordre deux. De plus, suite au filtrage fréquentiel passe-bande appliqué préalablement aux signaux, sa diagonale (variance) contient une partie de l'information fréquentielle. Elle est un descripteur tout à fait approprié pour représenter les changements spatio-fréquentiels des signaux EEG induit par l'imagination d'un mouvement d'un membre. Ces matrices sont ensuite classées en utilisant des outils empruntés à la géométrie riemannienne, respectant ainsi la structure particulière des ces matrices symétriques et définies positives. Outre l'augmentation des performances, ces méthodes permettent de se passer de l'étape de filtrage spatial en exploitant directement l'information spatiale contenue dans les matrices de covariance. Les conséquences sont une simplification de la chaîne de traitement permettant, entre autre, d'envisager des implémentations adaptatives et en-ligne des algorithmes de classification. Il est donc tentant d'utiliser ces méthodes pour les ICMs utilisant les potentiels évoqués P300. Malheureusement, l'estimation de la matrice de covariance spatiale ne tenant pas compte de la structure temporelle des signaux, leur classification ne permet pas d'atteindre des performances suffisantes. Ce travail introduit une estimation supervisée d'une "super" matrice de covariance permettant de prendre en considération la forme d'onde du poten-

tiel évoqué et ainsi de bénéficier des avantages apportés par les méthodes de classification précédemment introduites.

2 Méthode

Soit $\mathbf{X}_i \in \mathbb{R}^{C\times N}$ une portion de signal EEG d'indice i, correspondant à N échantillons de signal consécutif au $i$ème stimulus, observé avec C électrodes. Soit $\mathcal{I}^{(T)}$ l'ensemble des indices correspondant à la classe cible, on définit la réponse moyenne du potentiel évoqué P300 par :

$$\mathbf{P}_1 = \frac{1}{|\mathcal{I}^{(T)}|}\sum_{i\in\mathcal{I}^{(T)}} \mathbf{X}_i \tag{1}$$

Pour chaque essai $\mathbf{X}_i$, on construit ensuite un essai $\tilde{\mathbf{X}}_i \in \mathbb{R}^{2C\times N}$ par la concaténation de $\mathbf{P}_1$ et $\mathbf{X}_i$:

$$\tilde{\mathbf{X}}_i = \begin{bmatrix} \mathbf{P}_1 \\ \mathbf{X}_i \end{bmatrix} \tag{2}$$

Ces nouveaux essais sont utilisés pour estimer les matrices de covariance $\tilde{\mathbf{\Sigma}}_i \in \mathbb{R}^{2C\times 2C}$ sur lesquelles sera appliquée la classification.

$$\tilde{\mathbf{\Sigma}}_i = \frac{1}{N-1}\tilde{\mathbf{X}}_i\tilde{\mathbf{X}}_i^T = \begin{bmatrix} \mathbf{\Sigma}_{P_1} & \mathbf{C}_{P_1,X_i}^T \\ \mathbf{C}_{P_1,X_i} & \mathbf{\Sigma}_i \end{bmatrix} \tag{3}$$

Ces matrices de covariance peuvent être décomposées en 3 différents blocs. $\mathbf{\Sigma}_{P_1}$ la matrice de covariance de $\mathbf{P}_1$, commune à tout les essais et donc inutile pour la discrimination des classes. $\mathbf{\Sigma}_i$ la matrice de covariance de $\mathbf{X}_i$, contenant une partie de l'information pertinente (l'information spatiale des sources) mais ne suffisant pas à elle seule pour obtenir des performances suffisantes. Enfin $\mathbf{C}_{P_1,X_i}$, la matrice de covariance croisées (non symétrique) entre $\mathbf{P}_1$ et $\mathbf{X}_i$, contient la majeure partie de l'information utile à la classification (la corrélation temporelle du potentiel évoqué avec son prototype). En effet, si l'essai $\mathbf{X}_i$ contient un potentiel évoqué synchrone avec $\mathbf{P}_1$ cela se traduira par une augmentation significative de la covariance croisée au niveau des électrodes concernées. À l'inverse, en l'absence de potentiel évoqué ou d'une réponse non synchrone, la covariance croisée sera proche de zéro.

Suite à l'estimation de ces matrices de covariance, une classification par distance minimale aux centres des classes (MDM) est appliquée [1]. Les matrices de covariance étant par nature symétriques et définies positives, elle appartienne a une variété riemannienne. Cette variété est munie d'une métrique permettant la définition d'une distance entre les matrices. Ainsi, pour deux matrices de covariance $\mathbf{\Sigma}_1$ et $\mathbf{\Sigma}_2$, la distance est définie par [5] :

$$\delta_R(\mathbf{\Sigma}_1,\mathbf{\Sigma}_2) = \|\log\left(\mathbf{\Sigma}_1^{-1/2}\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{-1/2}\right)\|_F = \left[\sum_{c=1}^{C}\log^2\lambda_c\right]^{1/2} \tag{4}$$

avec $\lambda_c, c = 1\dots C$ les valeurs propres réelles de $\mathbf{\Sigma}_1^{-1/2}\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{-1/2}$

Pour chaque classe, cible et non-cible, le barycentre des matrices de covariance (moyenne géométrique riemannienne) est estimée et la classification se fait sur la base de la différence des distances riemanniennes entre une matrice à classer et les deux barycentres. Le barycentre de I matrices de covariance est définit comme la matrice minimisant la somme des distances riemannienne au carré [6] :

$$\mathfrak{G}\left(\mathbf{\Sigma}_1,\dots,\mathbf{\Sigma}_I\right) = \underset{\mathbf{\Sigma}\in P(n)}{\arg\min}\sum_{i=1}^{I}\delta_R^2\left(\mathbf{\Sigma},\mathbf{\Sigma}_i\right). \tag{5}$$

Cette moyenne (géométrique) est également appelée moyenne de Fréchet ou barycentre de Karcher [7]. La solution de cette fonction est obtenue par l'intermédiaire d'une procédure itérative [8], dont une implémentation en Matlab est disponible sous licence libre [1].

La classification est ensuite simplement effectué a partir de la différences des distances riemannienne au centre des classes :

$$s_i = \delta_R(\tilde{\mathbf{\Sigma}}_i,\bar{\mathbf{\Sigma}}_{\bar{T}}) - \delta_R(\tilde{\mathbf{\Sigma}}_i,\bar{\mathbf{\Sigma}}_T), \tag{6}$$

avec $\bar{\mathbf{\Sigma}}_T$ et $\bar{\mathbf{\Sigma}}_{\bar{T}}$ les matrices de covariance moyennes de la classe cible et non-cible, respectivement. La méthode complète est résumée par l'Algorithme 1

Algorithm 1 Classification MDM

Entrée : Les essais $\mathbf{X}_i$ et les labels y_i.
Entrée : Un essai $\mathbf{X}$ de classe inconnue y.
Sortie : s le score de classification de $\mathbf{X}$.

1: Estimation de $\mathbf{P}_1$, Eq.(1)
2: Construction des essais $\tilde{\mathbf{X}}_i$ et $\tilde{\mathbf{X}}$, Eq.(2)
3: Estimation des matrices de covariance $\tilde{\mathbf{\Sigma}}_i$ et $\tilde{\mathbf{\Sigma}}$, Eq.(3)
4: $\bar{\mathbf{\Sigma}}_T = \mathfrak{G}\left(\tilde{\mathbf{\Sigma}}_i | y_i = T\right)$, barycentre de matrices cibles, Eq.(5)
5: $\bar{\mathbf{\Sigma}}_{\bar{T}} = \mathfrak{G}\left(\tilde{\mathbf{\Sigma}}_i | y_i = \bar{T}\right)$, barycentre des matrices non-cibles.
6: $s = \delta_R(\tilde{\mathbf{\Sigma}},\bar{\mathbf{\Sigma}}_{\bar{T}}) - \delta_R(\tilde{\mathbf{\Sigma}},\bar{\mathbf{\Sigma}}_T)$, score de classification Eq.(6)
7: **return** s

3 Résultats

Les performances de cette méthode sont évaluées sur 2 jeux de données et sont comparées à deux méthodes de l'état de l'art. La première méthode, nommée XDAWN, consiste en un filtrage spatial par XDAWN [9], une décimation puis une classification par analyse discriminante linéaire régularisée [10] sur les caractéristiques construit par agrégation des signaux après filtrage spatial par les deux filtres spatiaux les plus discriminant. La deuxième méthode, nommée SWLDA, consiste en une décimation, une agrégation des signaux de chaque électrode, une sélection de modèle par stepwise-régression puis une classification par LDA sur les dimensions précédemment sélectionnées [11].

Les deux jeux de données ont été enregistrés lors de l'utilisation du jeu *brain invaders* [12], inspiré du célèbre jeu japonais *Space Invaders*. Le but de ce jeu est de détruire un alien cible parmi un ensemble d'autre d'aliens en utilisant un paradigme de type *P300 speller*. Les aliens sont arrangés sous forme de matrice 6×6. Les aliens sont flashé par groupe de 6 avec un interval

1. http ://github.com/alexandrebarachant/covariancetoolbox

aléatoire tiré selon une distribution exponentielle de moyenne correspondant à un intervalle de 200ms. Chaque alien est flashé deux fois au cours d'une répétition, cette dernière est donc constitué de 12 flashs. À la fin de chaque répétition, l'alien dont le score de classification est le plus élevé est détruit. Si cet alien correspond à l'alien cible, le jeu avance au niveau suivant. Dans le cas contraire une nouvelle répétition commence (dans la limite de 8 maximum) et ses score de classification seront ajouté à ceux de la répétition précédente. Le jeu comporte 12 niveaux, dont la difficulté varie grâce à l'ajout de distracteurs ou de cible en mouvement [13].

3.1 Jeu de donnée I

Le premier jeu de données comporte 24 sujets ayant effectués une session décomposée en une phase de calibration de 10 minutes environ suivie d'une phase de jeu de durée équivalente. Les signaux EEG, mesurés sur 17 électrodes, sont échantillonnés à 128Hz par un amplificateur Nexus. Un filtrage fréquentiel passe-bande entre 1 et 20 Hz est appliqué préalablement. Les résultats sont évalués hors ligne sur un critère d'aire sous la courbe roc (AUC), témoignant de la qualité de la discrimination des deux classes. Pour chaque sujet, on fait varier la quantité de données utilisée pour la calibration des algorithmes, les performances étant évaluées sur les données de test correspondant à la phase de jeu en ligne. Les résultats sont présentés Figure 1. Les 3 algorithmes convergent asymptotiquement vers des performances équivalentes. Néanmoins, la différence entre MDM et XDAWN est statistiquement significative (t-test for paired sample, $t(23) = 4.11\,;\, p < 0.0005$). La méthode SWLDA est la méthode qui converge le moins rapidement. En effet, compte tenu de la grande dimension de l'espace des caractéristiques, la sélection de modèle nécessite plus de données pour être efficace. Grâce au filtrage spatial qui réduit la dimension de l'espace, la méthode XDAWN offre de bonne performance pour un nombre réduit de données de calibration. Enfin, la méthode MDM nécessite moins de données d'entraînement à performances équivalentes grâce à la prise en compte de l'information spatio-temporelle directement au niveau du classifieur.

3.2 Jeu de donnée II

Le deuxième jeu de données comporte 4 sujets ayant effectué 6 sessions enregistrées au cours de 2 mois. Chaque session consiste en une partie en-ligne de *brain-invaders*, sans phase de calibration, le classifieur étant calibré sur les données des sessions précédentes, à l'exception de la première session où un classifieur générique (établit à partir des données d'autres utilisateurs) a été utilisé. Les signaux EEG ont été enregistrés grâce à l'amplificateur g.USBamp de g.tec et un système de 16 électrodes sèches g.Sahara. Les électrodes sèches offrent une qualité de signal moins bonne et une plus grande variabilité entre les sessions en comparaison aux électrodes humides. Un filtrage fréquentiel passe-bande entre 1 et 20 Hz est également appliqué. Ce jeu de données a pour but d'évaluer la généralisation des classifieurs entre sujet ou d'une session à l'autre pour un même sujet. Les performances sont évaluées hors ligne sur un critère d'AUC. Les résultats sont donnés en Figure 2. Les performances entre sujets sont données pour le nombre de sessions $n = 0$ sur l'axe des abscisses. Pour ce faire, les différentes chaînes de classification sont entraînées à partir des données de 3 sujets, et appliquées sur le sujet restant. Il apparait ainsi que la méthode MDM offre la meilleure généralisation entre les sujets. À l'inverse, la méthode XDAWN se généralise très mal, en raison de la difficulté d'estimer des filtres spatiaux sur des données hétérogènes. Les performances entre sessions sont données pour $n > 0$. Les classifieurs sont entraînés sur n sessions, et ap-

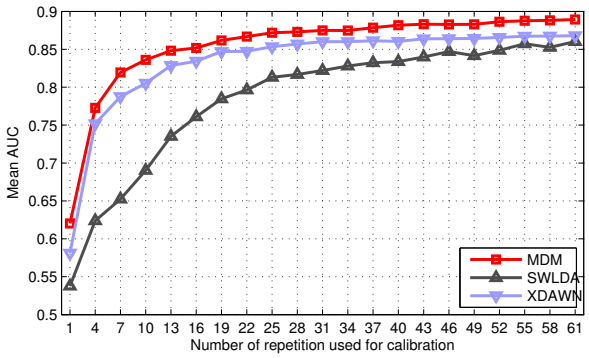


FIGURE 1 – Performances des différents algorithmes de classification en fonction de la quantité de données utilisée pour la calibration pour les 24 sujets du jeu de données I. 1 répétition = 12 essais, 2 cibles et 10 non-cibles. Performances évaluées suivant le critère d'aire sous la courbe roc (AUC) sur les données de test. À performances égales, la méthode proposée (MDM) a besoin de moins de données de calibration.

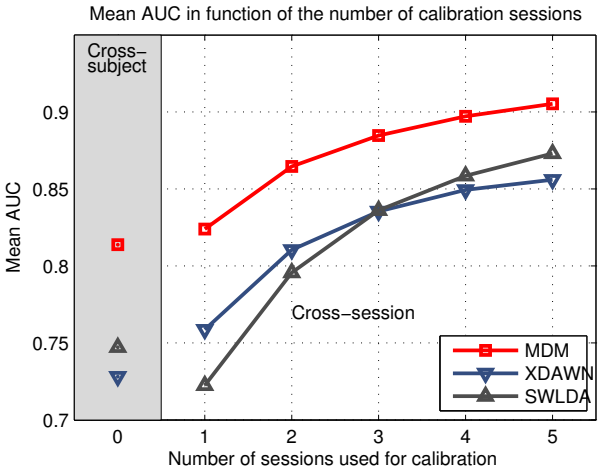


FIGURE 2 – Performances en fonction d'un nombre croissant de sessions utilisées pour la calibration, évaluées suivant le critère d'aire sous la courbe roc. La méthode MDM se généralise mieux entre les sessions.

pliqués sur les $6 - n$ sessions restantes, et ce pour l'ensemble des combinaisons possibles de n sessions parmi 6. Ces résultats ne prennent donc pas en compte le possible entraînement de l'utilisateur au cours du temps, mais seulement l'effet d'ajout de données dans la base d'apprentisage. Il apparaît ainsi que la méthode XDAWN se généralise relativement mal à d'autres sessions, et tend à plafonner autour de 0.85 d'AUC. A partir de 3 sessions, la méthode SWLDA offre des résultats supérieurs à XDAWN, et semble rejoindre asymptotiquement les performances de la MDM. Elle reste néanmoins en deçà de la méthode MDM qui offre une très bonne généralisation même avec très peu de sessions d'entraînement. Ainsi, à partir de deux session d'entraînement, la MDM offre de meilleurs performances que les autres méthodes pour le nombre maximal de session disponible. On notera également que la méthode MDM entraînée à partir des données des autres sujets offre des performances similaire aux méthodes de l'état de l'art entraînées sur deux sessions propre au sujet.

4 Conclusion

Cet article introduit une nouvelle méthode de classification pour les potentiels évoqués en EEG. Cette méthode se fonde sur une estimation supervisée des matrices de covariance de chaque essai, produites par concaténation de l'essai et de la réponse P300 prototype. Cette opération permet la prise en compte de la forme d'onde du potentiel évoqué et il devient alors possible d'utiliser ces matrices de covariance comme caractéristiques pour la classification. La géométrie riemannienne est ensuite utilisée pour classer ces matrices.

En comparaison aux méthodes de l'état de l'art, la méthode proposée offre de meilleures performances, permet de réduire le nombre de données nécessaires à la calibration et se généralise bien entre différentes sessions d'enregistrement ou d'un sujet à l'autre. Grâce à ces propriétés et de part la simplicité de la méthode, il est possible de réaliser une implémentation adaptative et en ligne de cet algorithme, où le classifieur serais initialisé à partir de paramètres génériques extrais depuis une base de données d'utilisateur puis adapté au spécificité individuelle du sujet. Une telle implémentation, permettant de supprimer la phase de calibration préalable à l'utilisation d'une ICM, est actuellement en cours de validation dans notre laboratoire.

Cependant, cette approche montre ses limites lorsque le nombre d'électrodes est important ($C > 64$). D'une part car la moindre manipulation des matrices de covariance suivant la métrique Riemannienne requière au moins une décomposition en valeurs propres, ce qui augmente le coup calculatoire de la méthode. D'autre part car la variété Riemannienne étant de grande dimension ($C(C+1)/2$), il est fréquent d'être confronté à la "malédiction de la dimension", se traduisant par des problèmes de sur-apprentissage ou de convergence des algorithmes.

Enfin, l'introduction de cette méthode d'estimation des matrices de covariance permet l'unification de la chaîne de classification pour les différents types d'ICMs (Imagerie motrice, P300, SSVEP, Potentiels d'erreur), réduisant ainsi la fracturation du domaine. Cette unification trouve un intérêt tout particulier dans l'analyse des potentiels d'erreur, où la présence d'une réponse asynchrone en plus du potentiel évoqué a été récemment mise en évidence. Il est désormais possible de traiter conjointement ces deux réponses, grâce à une estimation spécifique de la matrice de covariance, là ou deux chaîne de classification distinctes aurait été employé.

Références

[1] A. Barachant, S. Bonnet, M. Congedo et C. Jutten. *Multiclass Brain-Computer Interface Classification by Riemannian Geometry*. IEEE Transactions on Biomedical Engineering, Vol. 59, No. 4, p. 920-928, 2012.

[2] A. Barachant, S. Bonnet, M. Congedo et C. Jutten. *Classification of covariance matrices using a Riemannian-based kernel for BCI applications*. NeuroComputing, Vol. 112, pp. 172-178, 2013

[3] A. Barachant, S. Bonnet, M. Congedo et C. Jutten. *Réalisation d'un Brain-Switch EEG par Géométrie Riemannienne*. 23ème colloque GRETSI, 2011.

[4] M. Congedo, C. Gouy-Pailler, C. Jutten *On the blind source separation of human electroencephalogram by approximate joint diagonalization of second order statistics* Clinical Neurophysiology, vol. 119, pp. 2677-2686, 2008.

[5] W. Förstner and B. Moonen *A Metric for Covariance Matrices* Technical report, Dept. of Geodesy and Geoinformatics, Stuttgart Univ., 1999

[6] M. Moakher *A Differential Geometric Approach to the Geometric Mean of Symmetric Positive-Definite Matrices* SIAM J. Matrix Anal. Appl., Vol. 26, No. 2, pp. 735–747, 2005

[7] F. Barbaresco *Géométrie différentielle des matrices de covariance & espaces métriques à courbure négative* 22ème colloque GRETSI, 2009.

[8] X. Pennec, P. Fillard, N. Ayache *A Riemannian framework for tensor computing* International Journal of Computer Vision, Vol 66, No. 1, pp 41-66, 2006

[9] B. Rivet, A. Souloumiac, V. Attina et G. Gibert. *xDAWN Algorithm to Enhance Evoked Potentials : Application to Brain Computer Interface*. IEEE Transactions on Biomedical Engineering, Vol. 56, No. 8, pp. 2035-2043, 2009.

[10] B. Blankertz, S. Lemm, M. Treder, S. Haufe et KR. Müller. *Single-trial analysis and classification of ERP components - a tutorial*. Neuroimage, Vol. 56, No. 2, pp. 814-825, 2010.

[11] D. J. Krusienski, E. W. Sellers, F. Cabestaing, S. Bayoudh, D. J. McFarland, T. M. Vaughan, et J. R. Wolpaw. *A comparison of classification techniques for the P300 Speller* Journal of Neural Engineering, vol. 3, no. 4, pp. 299-305, 2006.

[12] Congedo M, Goyat M, Tarrin N, Varnet L, Rivet B, Ionescu G, Jrad N, Phlypo R, Acquadro M, Jutten C. *"Brain Invaders" : a prototype of an open-source P300-based video game working with the OpenViBE platform*. Proceedings of the 5th International BCI Conference, Graz, Austria, pp. 280-283, 2011.

[13] G.F.P. van Veen. *Brain Invaders, fiding the paradox of control in a P300-Game through the use of distractions* Master thesis, University of Twente, 2013