

Réalisation d'un Brain-Switch EEG par Géométrie Riemannienne.

Alexandre BARACHANT¹, Stéphane BONNET¹, Marco CONGEDO², Christian JUTTEN²

¹CEA-LETI, MINATEC Campus, 38054 Grenoble, France

²GIPSA-lab, CNRS/UJF-INPG, Domaine Universitaire, 38402 Saint-Martin d'Hères, France

¹prénom.nom@cea.fr, ²prénom.nom@gipsa-lab.grenoble-inp.fr

Résumé – Le domaine des interfaces cerveau-machine représente un réel espoir pour des personnes handicapées moteur afin de commander un effecteur à travers la modulation de leur activité cérébrale. En vue d'une application dans un contexte de vie courante, il est nécessaire de mettre au point un système asynchrone robuste de détection d'une activité mentale spécifique à partir des mesures EEG. La réalisation d'un tel Brain-Switch est abordée dans ce papier à travers l'utilisation des outils de la géométrie Riemannienne pour manipuler et classer les matrices de covariance spatiale, choisies comme descripteur de l'activité mentale sous-jacente. La méthode est testée sur des acquisitions EEG réalisées au sein du laboratoire. Les résultats démontrent une bonne qualité de détection avec un pourcentage de déclenchement volontaire de l'ordre de 92% et de déclenchements réussis de l'ordre de 91%.

Abstract – Brain-Computer Interfaces (BCIs) represent a real hope for people who suffer of motor disabilities. In order to use these interfaces in the real life, it is mandatory to set up a robust asynchronous system for detecting a specific brain pattern within the ongoing EEG activity. In this paper, we address this particular issue through the use of the Riemannian geometry in order to manipulate and classify the spatial EEG covariance matrices, used as brain activity descriptors. The method is evaluated on a dataset realized in our laboratory and results show a good quality of detection with a positive predictive value (PPV) of 92% and a true positive rate (TPR) of 91%.

1 Introduction

Une Interface Cerveau-Machine (ICM) permet à l'utilisateur - handicapé moteur ou non - de s'interfacer avec le monde extérieur par la *seule* utilisation de son cerveau, c'est-à-dire sans recourir à la voie musculaire ou vocale pour exécuter une action. Le travail présenté ici se concentre sur la reconnaissance en ligne d'une activité mentale spécifique par rapport à une activité de fond non spécifique. Les signaux mesurés sont issus d'un système d'électroencéphalographie (EEG) avec plusieurs électrodes de scalp. Un tel paradigme **asynchrone** est connu sous le nom de Brain-Switch (BS). Historiquement introduit par Birch en 2000 [1], le BS est fondamental dans la mise en œuvre pratique d'une ICM. Certains travaux n'hésitent d'ailleurs pas à utiliser une autre technique de mesure de l'activité cérébrale pour rendre cette étape la plus robuste possible, comme en [2] avec l'utilisation de la spectroscopie proche infrarouge. Un Brain-Switch peut permettre soit d'exécuter une commande binaire (de type interrupteur poussoir) soit d'informer le système que l'utilisateur est prêt à vouloir interagir avec une ICM. Dans ce cas, le BS est bien souvent un préalable à un autre type d'ICM qu'il enclenche (ou arrête).

L'imagination d'un mouvement d'un membre induit l'activation de la zone du cortex moteur qui lui est associé. En EEG, cela se traduit par une évolution de la puissance des signaux de certaines électrodes dans une bande de fréquence particulière. De telles réponses à un mouvement imaginé sont bien connues, spécifiques et reproductibles et sont donc couramment utilisées

pour coder les commandes envoyées par l'utilisateur à l'ICM. Ainsi, on peut citer les travaux de G. Pfurtscheller qui utilise l'augmentation de la puissance dans la bande beta [13-35Hz] en fin d'imagerie motrice pour activer un BS [3]. En utilisant de tels principes physiologiques, la tâche de détection induisant le déclenchement du BS se résume à la détection d'un changement caractéristique de la puissance des signaux dans le domaine spatio-fréquentiel au sein de l'activité EEG courante.

La matrice de covariance spatiale contenant à la fois l'énergie des signaux à travers ses éléments diagonaux et l'information spatiale à travers ses éléments non diagonaux est alors un descripteur tout à fait approprié pour représenter et caractériser les signaux EEG dans le cadre de l'imagerie motrice.

L'approche proposée dans ce papier utilise les outils de la **géométrie Riemannienne** pour manipuler les matrices de covariance spatiale, obtenues sur des portions temporelles de signal EEG. Ainsi la variété des matrices symétriques définies positives (SDP) (à laquelle appartiennent les matrices de covariance) est propice à mener des calculs analytiques de moyenne robuste, de distance entre matrices de covariance pour mener à bien la tâche de détection.

2 Méthodes

Désignons par $\mathbf{X} \in \mathbb{R}^{C \times N}$ la portion d'enregistrement EEG (ou *trial*) constituée de N échantillons et C électrodes. En supposant que chaque voie ait été filtrée passe-bande au préalable

comme il est d'usage pour les signaux EEG, ce qui a pour effet de centrer les données, alors la matrice de covariance spatiale locale peut être estimée par la formule usuelle :

$$\mathbf{P} = \frac{1}{N-1} \mathbf{X}\mathbf{X}^T \quad (1)$$

Les outils de la géométrie Riemannienne nous permettent de définir une distance entre deux matrices SDP au sens d'une certaine métrique. On pourra se référer à [4] pour plus de détails. On note par $P(C)$ l'ensemble des matrices symétriques définies positives (SDP) de dimension $C \times C$. La distance Riemannienne entre deux matrices SDP \mathbf{P}_1 et \mathbf{P}_2 est donnée par :

$$\delta_R(\mathbf{P}_1, \mathbf{P}_2) = \left[\sum_{c=1}^C \log^2 \lambda_c \right]^{1/2} \quad (2)$$

avec $\{\lambda_c, c = 1 \dots C\}$ les valeurs propres réelles de $\mathbf{P}_1^{-1}\mathbf{P}_2$. La définition de la distance nous permet également de définir le calcul de la moyenne de M matrices de covariance.

$$\mathfrak{G}(\mathbf{P}_1, \dots, \mathbf{P}_M) = \arg \min_{\mathbf{P} \in P(C)} \sum_{m=1}^M \delta_R^2(\mathbf{P}, \mathbf{P}_m) \quad (3)$$

Ce minimum existe, il est unique et peut être calculé par un algorithme itératif comme dans [5] et reproduit en Algorithme 1. Cette moyenne est communément appelée barycentre de Fréchet-Karcher [4] ou encore moyenne géométrique [6].

Algorithme 1 Moyenne géométrique de M matrices SPD

Entrées : Ω un ensemble de M matrices SPD $\mathbf{P}_m \in P(C)$ et $\epsilon > 0$.

Sortie : \mathbf{P}_Ω l'estimation de la moyenne géométrique.

- 1: Initialise $\mathbf{P}_\Omega^{(1)} = \frac{1}{M} \sum_{m=1}^M \mathbf{P}_m$
 - 2: $t = 1$
 - 3: **repeat**
 - 4: $\mathbf{S} = \frac{1}{M} \sum_{m=1}^M \text{Log}_{\mathbf{P}_\Omega^{(t)}}(\mathbf{P}_m)$
 - 5: $\mathbf{P}_\Omega^{(t+1)} = \text{Exp}_{\mathbf{P}_\Omega^{(t)}}(\mathbf{S})$
 - 6: $t = t + 1$
 - 7: **until** $\|\mathbf{S}\|_F < \epsilon$
-

avec les opérateurs exponentiel et logarithmique définis par :

$$\text{Exp}_{\mathbf{P}}(\mathbf{S}) = \mathbf{P}^{1/2} \text{Exp} \left(\mathbf{P}^{-1/2} \mathbf{S} \mathbf{P}^{-1/2} \right) \mathbf{P}^{1/2}$$

$$\text{Log}_{\mathbf{P}}(\mathbf{P}_m) = \mathbf{P}^{1/2} \text{Log} \left(\mathbf{P}^{-1/2} \mathbf{P}_m \mathbf{P}^{-1/2} \right) \mathbf{P}^{1/2}$$

Méthode de détection de l'activité spécifique

L'activité spécifique que l'on cherche à détecter est par définition reproductible et caractéristique. Les matrices de covariance des essais qui lui sont associées se situent donc dans une région particulière de la variété des matrices SDP. Par opposition, l'activité non spécifique sera constituée de l'ensemble des autres activités possibles, ainsi les matrices de covariance associées peuvent théoriquement compléter le reste de la variété. La méthode de détection comporte une première phase d'apprentissage permettant d'identifier les caractéristiques de l'activité spécifique à partir de données enregistrées et annotées et

une phase de d'utilisation qui utilisera pour la détection les paramètres extrais de la phase d'apprentissage.

Phase d'apprentissage : On dispose de M matrices de covariance, nommées \mathbf{P}_i d'indice $i = 1 \dots M$, ainsi que de leur label associé $y_i \in \{s, \bar{s}\}$ avec s correspondant à l'activité spécifique. En premier lieu, on estime la matrice de covariance moyenne $\mathfrak{G}_s = \mathfrak{G}(\mathbf{P}_i, i | y_i = s)$ des matrices correspondant à l'activité spécifique. On définit ensuite une région d'intérêt $\Omega = \{\mathbf{P} | \delta_R(\mathfrak{G}_s, \mathbf{P}) < \epsilon\}$. La région d'intérêt est définie de manière à contenir la majeure partie des matrices correspondant à l'activité spécifique. Ainsi, on fixe $\epsilon = m + 3s$ avec m et s respectivement la médiane et l'écart type (estimé de manière robuste par la déviation absolue à la médiane) des distances $\{\delta_R(\mathfrak{G}_s, \mathbf{P}_i), i | y_i = s\}$, faisant ainsi l'hypothèse que ces distances suivent une distribution gaussienne. Cette approximation n'est pas rigoureuse mais permet de définir un seuil au delà duquel la probabilité d'appartenance à la classe des activités spécifiques est pratiquement nulle. Pour améliorer la spécificité de la détection, une deuxième étape consiste à estimer la moyenne géométrique $\mathfrak{G}_{\bar{s}} = \mathfrak{G}(\mathbf{P}_i, i | y_i = \bar{s})$ des matrices de covariance des activités non spécifiques contenues exclusivement dans la région d'intérêt. A l'issue de cette procédure d'apprentissage, on a estimé les 3 quantités $\mathfrak{G}_s, \mathfrak{G}_{\bar{s}}$ et ϵ qui seront utilisées lors de la phase d'utilisation.

Phase d'utilisation : On cherche à détecter si une matrice \mathbf{P} de classe inconnue $y \in \{s, \bar{s}\}$ correspond à une activité spécifique. Toute matrice se trouvant à l'extérieur de la région d'intérêt est automatiquement considérée comme correspondant à une activité non spécifique. Si la matrice en question est à l'intérieur de la région d'intérêt, elle appartient à la classe dont la moyenne est la plus proche au sens de la distance Riemannienne. Cette prise de décision se traduit mathématiquement par :

$$y = \begin{cases} s & \text{si } \delta_R(\mathfrak{G}_s, \mathbf{P}) < \epsilon \text{ et } \delta_R(\mathfrak{G}_s, \mathbf{P}) < \delta_R(\mathfrak{G}_s, \mathbf{P}) \\ \bar{s} & \text{sinon} \end{cases} \quad (4)$$

Le Brain-Switch suit un fonctionnement de type bouton poussoir. Le déclenchement est fonction de la sortie du classifieur y . Ainsi l'interrupteur se déclenche lorsque l'on détecte une activité spécifique consécutivement pendant 1 seconde. Une fois l'interrupteur déclenché, il se réamorçe lorsque l'on détecte une activité non spécifique pendant une nouvelle période de 1 seconde.

3 Résultats

On dispose d'enregistrements EEG sur 6 sujets réalisés au laboratoire lors d'une campagne de test précédant le développement de notre algorithme. On utilise 16 électrodes réparties sur l'ensemble du scalp, une fréquence d'échantillonnage de 512 Hz et un filtrage passe bande entre 8 et 30 Hz. Ce filtrage est commun à tous les utilisateurs et permet d'une part de

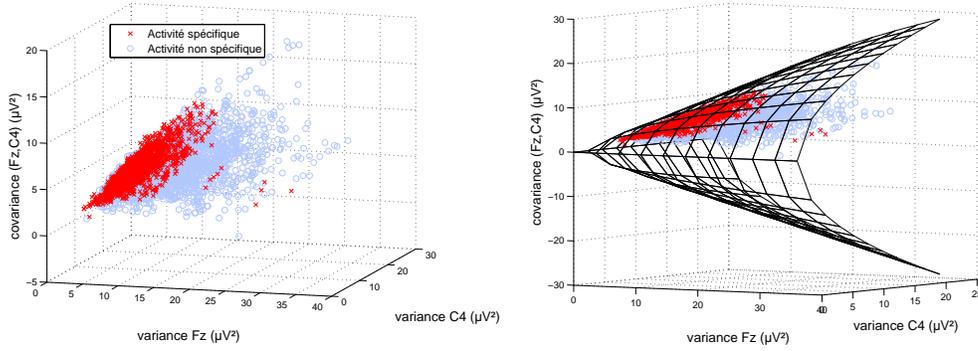


FIGURE 1 – Position des matrices de covariance pour l'utilisateur E, au niveau des électrodes Fz et $C4$. Données d'apprentissage

sélectionner les deux bandes fréquentielles alpha et mu contenant les réponses fréquentielles induites par l'imagerie motrice et d'autre part de supprimer un certain nombre de sources d'artefacts, comme le clignement des paupières, dont les fréquences se situent en dessous de 8 Hz. Une sélection des bandes fréquentielles spécifiques à chaque sujet permettrait d'augmenter les performances du système mais nécessite l'acquisition au-préalable de signaux EEG afin d'en extraire les paramètres fréquentiels optimaux. De la même manière, la position et le nombre d'électrodes utilisées peuvent être optimisés pour chaque sujet à partir du moment où on connaît le type d'imagerie motrice utilisé ou encore si on dispose d'enregistrements EEG permettant d'en extraire les caractéristiques spatiales.

Le paradigme est de type Brain-Switch. Durant la phase d'apprentissage (15 minutes maximum) l'utilisateur effectue, devant un écran, un certain nombre de réalisations d'une activité spécifique (d'une durée de 2 à 10 secondes), basée principalement sur de l'imagerie motrice, et entrecoupés de phase de repos d'une durée similaire. Toutefois ces phases de repos correspondent à un état où le sujet reste concentré et ne représentent donc pas toute la diversité des activités rencontrées lors de l'utilisation d'une ICM en vie courante. Une phase de test de 5 minutes se déroule ensuite où l'utilisateur est soumis au même paradigme.

Les matrices de covariance spatiales sont estimées sur une fenêtre glissante de 1 seconde avec 875 ms de recouvrement (correspondant à un décalage de 64 échantillons). Le choix de ces deux paramètres est un compromis entre performance, latence et réactivité du système dans son utilisation en ligne. En effet, plus la fenêtre est large, meilleure sera l'estimation de la matrice de covariance de (1). En revanche, une fenêtre trop large induit une latence importante ce qui gêne fortement l'utilisateur lors d'une utilisation en ligne de l'algorithme. Le recouvrement influe sur la fréquence de rafraîchissement du retour visuel. Un recouvrement trop faible produit un retour saccadé, et un recouvrement trop important nécessite une puissance de calcul plus élevée pour effectuer l'ensemble des opérations avant l'arrivée de la matrice suivante.

Il est possible d'observer le positionnement des différentes matrices de covariance dans leur espace en se restreignant à deux électrodes (Figure 1). Chaque point décrit la position d'une

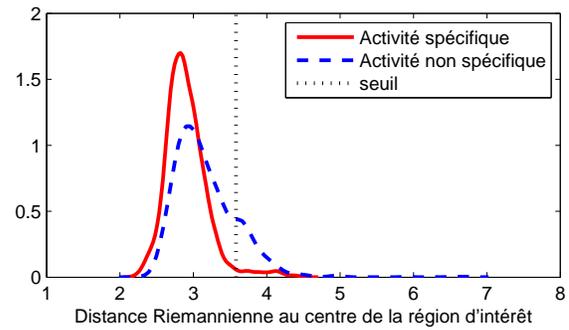


FIGURE 2 – Distribution des distances au centre de la région d'intérêt $\delta_R(\mathcal{G}_s, \mathbf{P}_i)$ et seuil définissant la région d'intérêt.

matrice de covariance. On note que les points correspondant à une activité spécifique sont regroupés, comme attendu, en une zone dense alors que les points correspondant à une activité non spécifique couvrent une zone beaucoup plus large de l'espace. On note enfin que ces points sont proches de la limite de la variété (représentée par un cône), correspondant à une corrélation égale à 1. Le calcul de la matrice de covariance moyenne \mathcal{G}_s de l'activité spécifique permet d'estimer le centre de la région d'intérêt.

La figure 2 illustre la distribution des distances à ce centre d'une part pour les matrices correspondant à une activité spécifique $\{\delta_R(\mathcal{G}_s, \mathbf{P}_i), i | y_i = s\}$ (en trait plein) et d'autre part pour les activités non spécifiques $\{\delta_R(\mathcal{G}_s, \mathbf{P}_i), i | y_i = \bar{s}\}$ (en pointillé). La restriction à une région d'intérêt améliore grandement la robustesse du dispositif, en effet tout point se trouvant en dehors de cette région (en particulier les artefacts) se trouve automatiquement classé en tant qu'activité non spécifique. En revanche, il y a confusion entre les deux types d'activité pour un nombre important de points dans la région d'intérêt, comme en témoigne les distributions montrées à la figure 2. L'utilisation de la matrice moyenne \mathcal{G}_s permet de lever en partie ces ambiguïtés. La figure 3 illustre le processus de classification sur les données de test. Le seuil de décision donné en (4) est représenté par le trait pointillé. La figure 4 montre la suite temporelle des déclenchements pour 10 réalisations d'activité spécifique de durées différentes pour les données de test. Dans

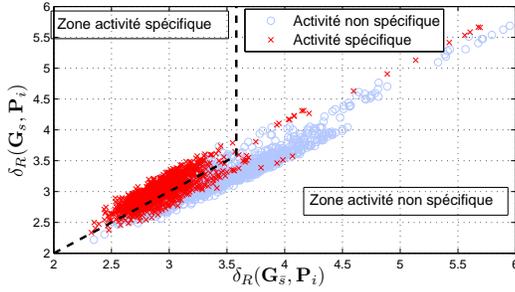


FIGURE 3 – Classification des données de test selon (4) pour l'utilisateur E

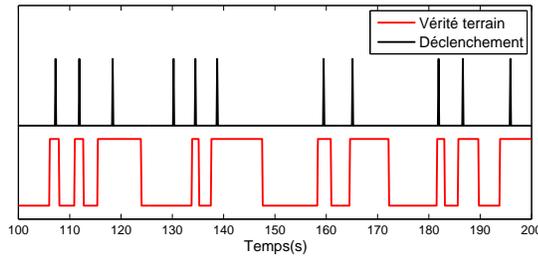


FIGURE 4 – Suite temporelle des déclenchements du Brain-Switch pour les données de test, utilisateur E

cet exemple, toutes les réalisations d'activité spécifique conduisent à un déclenchement et on note un seul déclenchement involontaire. Chez cet utilisateur, le pourcentage de déclenchement volontaire est de 94% tandis que le pourcentage d'essais avec déclenchement réussi est de 100%. Une méthode de référence, consistant en un filtrage spatial par Common Spatial Pattern (CSP) suivi d'une classification par analyse discriminante linéaire de Fisher (LDA), est également appliquée sur les données afin de pouvoir comparer les performances du BS [7]. Cette méthode de référence est mise en oeuvre en utilisant le même filtrage fréquentiel (filtre passe-bande entre 8 et 30 Hz) et les mêmes paramètres temporels (fenêtres de 1 seconde) que la méthode proposée. On utilisera également 3 paires de filtres spatiaux, comme il est généralement d'usage. Les résultats complets sur les 6 utilisateurs sont donnés Table 1 et sont présentés en terme de sensibilité et de précision. La sensibilité, notée TPR , est le rapport entre les déclenchements volontaires (TP) et le nombre total d'essais de déclenchement. La précision, notée PPV , est le rapport entre le nombre de déclenchements volontaires et le nombre total de déclenchement.

La méthode présentée dans ce papier obtient une moyenne de 91% pour le PPV et le TPR ce qui témoigne de la bonne précision et sensibilité de la méthode. La méthode de référence montre un sensibilité (TPR) similaire, en revanche la précision (PPV) est inférieure de 10%. Cette différence s'explique par le fait que la méthode proposée inclut nativement une gestion des artefacts, ce qui n'est pas le cas de la méthode de référence. Les résultats sur les TPR sont cohérents avec nos précédentes observations développées dans [5], à savoir une performance

TABLE 1 – Performances du BS selon les deux méthodes. PPV : précision (% de déclenchements volontaire) , TPR : sensibilité (% d'essais de déclenchement réussi), TP : vrai positif (nombre de déclenchements volontaires)

Utilisateur	CSP + LDA			BS Riemann		
	PPV	TPR	TP	PPV	TPR	TP
A	96	80	24	100	84	26
E	77	96	27	94	100	32
G	92	100	34	94	100	34
I	62	92	23	80	64	16
J	80	95	20	83	100	24
M	74	86	31	100	100	35
moyenne	80.1	91.6	26.5	91.8	91.3	27.8

similaire à la méthode de référence, le gain de cette nouvelle méthode résidant essentiellement dans l'augmentation de la robustesse aux artefacts.

4 Conclusions

Ce papier présente une méthode de classification de signaux EEG basée sur une mesure de distance Riemannienne entre des matrices de covariances spatiales dans le but de réaliser un Brain-Switch. Cette méthode ne nécessite le réglage d'aucun paramètre et gère nativement les artefacts, ce qui facilite sa mise en oeuvre expérimentale. En particulier, la mise en oeuvre adaptative de l'algorithme se résume à la mise à jour des matrices de covariances moyennes et du seuil définissant la région d'intérêt. Les futur travaux se focaliseront sur une prise en compte plus fine de la continuité temporelle des signaux, ce qui permettra d'augmenter la robustesse du dispositif.

Références

- [1] S.G. Mason and G.E. Birch. A brain-controlled switch for asynchronous control applications. *IEEE Transaction on Biomedical Engineering*, 2000.
- [2] S. M. Coyle, T. E. Ward, and C. M. Markham. Brain-computer interface using a simplified functional near-infrared spectroscopy system. *Journal of Neural Engineering*, 2007.
- [3] G. Pfurtscheller and T. Solis-Escalante. Could the beta rebound in the EEG be suitable to realize a "brain switch"? *Clinical Neurophysiology*, January 2009.
- [4] F. Barbaresco. Géométrie différentielle des matrices de covariance & espaces métriques à courbure négative. In *GRETSI*, 2009.
- [5] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten. Riemannian geometry applied to BCI classification. In *LVA/ICA*, 2010.
- [6] M. Moakher. A differential geometric approach to the geometric mean of symmetric Positive-Definite matrices. *J. Matrix Anal. Appl.*, 2005.
- [7] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Muller. Optimizing spatial filters for robust EEG Single-Trial analysis. *Signal Processing Magazine, IEEE*, 25(1) :41–56, 2008.