

FuRIA : un nouvel algorithme d'extraction de caractéristiques pour les interfaces cerveau-ordinateur utilisant modèles inverses et modèles flous

FuRIA: a new feature extraction algorithm for brain-computer interfaces using fuzzy and inverse models

Fabien Lotte^{1,2}, Anatole Lécuyer², Bruno Arnaldi^{1,2}

¹IRISA-INSA Rennes, 20 avenue des buttes de coesmes, 35043 Rennes Cedex, France
fabien.lotte.irisa.fr,

²IRISA-INRIA Rennes, avenue du général Leclerc, Campus Universitaire de Beaulieu, 35042 Rennes Cedex, France,
anatole.lecuyer.irisa.fr,
bruno.arnaldi.irisa.fr

Manuscrit reçu le 15 avril 2008

Résumé et mots clés

Cet article propose un nouvel algorithme d'extraction de caractéristiques pour les Interfaces Cerveau-Ordinateur (ICO) basées sur l'électroencéphalographie. Cet algorithme utilise les modèles inverses ainsi que le nouveau concept de Région d'Intérêt (RI) floue. Il peut automatiquement identifier les RI pertinentes pour la discrimination ainsi que les bandes de fréquences dans lesquelles ces RI sont les plus discriminantes. Les activités calculées dans ces RI peuvent ensuite être utilisées comme caractéristiques d'entrée pour n'importe quel classifieur. Une première évaluation de l'algorithme, utilisant une Machine à Vecteurs Supports (SVM) comme classifieur, est présentée sur le jeu de données IV de la « BCI competition 2003 ». Les résultats s'avèrent prometteurs avec une précision sur l'ensemble de test allant de 85 % à 86 % contre 84 % pour le gagnant de la compétition sur ces données. Enfin, nous montrons que combiner ce nouvel algorithme avec des systèmes d'inférence flous permet de concevoir des ICO potentiellement interprétables.

interface cerveau-ordinateur (ICO), extraction de caractéristiques, modèle inverse, localisation de sources, classification, interprétabilité, ensemble flou, électroencéphalographie (EEG).

Abstract and key words

When using Brain-Computer Interfaces (BCI) based on ElectroEncephaloGraphy (EEG), the identification of mental tasks relies on two main points: feature extraction and classification [MAM⁺06, BFWB07, LCL⁺07]. Feature extraction aims at describing EEG signals by a few relevant values called "features", whereas classification aims at automatically assigning a class to these features. In this paper we focus on feature extraction, as the BCI community has stressed the need to explore new feature extraction algorithms [MAM⁺06].

Recently, inverse models have been revealed as promising feature extraction algorithms for BCI [QDH04, GGP⁺05, WGW05, CLL06]. Such models aims at computing the activity in the whole brain volume, by using only scalp EEG signals and a head model representing the brain as a set of voxels (volume elements). The activity computed in a few brain regions has been used as features for BCI systems.

Despite good results, some limitations remain. Indeed, it seems that current methods cannot conciliate genericity, i.e., the capability to deal with any kind of mental task, and the fact of generating few features. On one hand, methods that are generic and automatic tend to generate a large number of features, as they extract several features for each voxel [GGP⁺05]. The activity in neighboring voxels can be correlated and, as such, it would be more appropriate to gather these voxels in brain regions. On the other hand, methods that generate few features have been proposed, but they are not generic anymore as they need a priori knowledge on the mental tasks used, and are currently limited to motor imagery-based BCI [QDH04, WGWB05]. Recently, we have proposed a method which is generic and which generates few features, as voxels whose activity is correlated are gathered into Regions Of Interest (ROI) [CLL06]. However, this method is not completely automatic and is limited to the use of two ROI whose spatial extension is hard to define [CLL06]. In this paper, we propose a generic feature extraction algorithm which can automatically identify any number of relevant ROI and can properly define their spatial extension thanks to the new concept of fuzzy ROI. This algorithm is known as FuRIA, which stands for “**F**uzzy **R**egion of **I**nterest **A**ctivity”.

The FuRIA algorithm

FuRIA is a trainable feature extraction algorithm. Indeed, it can automatically identify what are the relevant ROI to discriminate mental tasks, and what are the frequency bands in which these ROI react. The FuRIA training is performed in 5 steps:

Statistical analysis: This step consists in comparing, using a statistical test, the activity in each voxel, possibly averaged over a given time window, between the different classes. This analysis is performed for each of the N_f frequencies considered. Once the analysis is achieved, we obtain N_f statistical maps, each one revealing which voxels are significant for discrimination.

Clustering: For each statistical map computed at the previous step, a clustering algorithm is used in order to gather significant voxels into ROI. To perform this clustering, a 4-dimensional vector $[x, y, z, s]$ is associated to each significant voxel. The x, y, z coordinates are the 3D coordinates of the voxel in the head model used and s is the voxel statistic obtained at the previous step. Once the clustering is achieved, all voxels whose associated vector belongs to the same cluster are gathered into the same ROI.

Fusion: Once the relevant ROI are identified for each frequency, we have to identify which ROI are relevant in which frequency band. To do so, ROI obtained at similar positions (i.e., with an overlap of more than 50%), for consecutive frequencies, are merged, simply using the union set operator, and identified as reactive in the frequency band resulting from the concatenation of these frequencies.

Adaptation: Because of the use of the union as the fusion operator, a ROI Ω identified as reactive in the frequency band $[f_a; f_b]$ can contain voxels which were significant for a single frequency belonging to $[f_a; f_b]$ but which are not significant anymore for the whole $[f_a; f_b]$ frequency band. Consequently, we perform another statistical analysis for each ROI, in order to remove voxels that are not significant anymore.

Fuzzification: In the ROI defined so far, it can be noticed that all voxels do not have, statistically, the same discriminative power. However, all these voxels contain a given amount of information. Moreover, the brain regions of a given subject are not well defined but are rather naturally fuzzy. Hence, identifying which voxels should be kept in a ROI, and which ones should be removed is a difficult problem. We believe that all voxels that contain information should be in the ROI, but that those which contain less information should be less in the ROI. Thus, we propose the concept of fuzzy ROI, which relies on the fuzzy set theory of Zadeh [Zad96]. A standard (crisp) ROI Ω_c is defined by the set of voxels it contains. The activity γ_{Ω_c} in this ROI Ω_c is classically computed from the activity γ_v of the voxels v of this ROI (see Equation 1). A fuzzy ROI Ω_f is not defined by a set of voxels anymore, but by a fuzzy membership function μ_{Ω_f} . This function provides the membership degree (in $[0; 1]$) of any voxel into the ROI Ω_f . The activity in a fuzzy ROI is computed as described in Equation 2, where N_v is the number of voxels in the head model used. This formalism enables use to weigh each voxel according to the importance of its contribution within the ROI, and, as such, to use efficiently the available information. To complete the training of FuRIA, all ROI Ω are turned into fuzzy ROI by associating to each of them a fuzzy membership function μ_{Ω} which uses the statistics obtained at the previous step.

Once the training of FuRIA is completed, a set of discriminant fuzzy ROI should have been identified together with the corresponding reactive frequency bands. The activity in these ROI and in these frequency bands, possibly averaged over time, can be used as features for any classifier.

Evaluation

In order to evaluate FuRIA, we used sLORETA (standardized low resolution electromagnetic tomography) as the inverse model [PM02], a multiple comparison randomization t-test as the statistical analysis [HBWF96] and the “Mean Shift” algorithm as the clustering algorithm [CM02]. Concerning the fuzzy membership functions, it appears that, due to the low resolution of sLORETA, significant voxels with a relatively high probability of type I error are much more numerous than significant voxels with a relatively low probability of type I error. Consequently, for each ROI Ω , we used a membership function μ_{Ω} which assigns a membership degree which decreases exponentially with the voxel statistics, as described in Equation 10. In this equation, N_{Ω_s} is the number of voxels in the ROI Ω ; p_{min} and p_{max} are respectively the minimum and maximum probability, among the ROI voxels, of correctly rejecting the null hypothesis H_0 ; and p_v is the probability of correctly rejecting H_0 for voxel v .

In order to evaluate FuRIA, we have used the data set IV of the “BCI competition 2003” [WZL+04]. These data contain EEG signals recorded 500 ms before a subject hits a keyboard key with a finger from his left or right hand. The goal of the competition participants was to identify, for each trial, the hand that was used.

We trained the FuRIA algorithm on the training set provided, by considering the frequencies between 3 Hz and 45 Hz, with a step of 1 Hz, and by considering the time window comprising the last 250 ms of each trial. FuRIA has found 5 discriminant fuzzy ROI (see Figure 1), hence generating 5-dimensional feature vectors.

These features have been classified using a linear support vector machine. The recognition rates obtained after 100 training/testing repetitions are displayed in Table 2 and compared with two other methods: the method of the competition winner [WZL+04] and a method using inverse models that we proposed previously [CLL06]. The results obtained without the fuzzification step (step 5 of the training, see previous section) are also displayed as a comparison. We can notice that the method that uses FuRIA with fuzzification reached the best results with a recognition score ranging from 85% to 86%. These results suggest that the features extracted with FuRIA are very discriminant, as they are not numerous (there are only 5 features) and still enable us to reach good classification results. It is also interesting to notice that when using FuRIA without fuzzification, the results are significantly lower. This stresses the advantage of using efficiently the information contained in any significant voxel, thanks to the concept of fuzzy ROI.

In addition to this evaluation, we also proposed a first step towards the design of an interpretable BCI system. To this end, we proposed to classify FuRIA features, which are interpretable features, with a Fuzzy Inference System (FIS), which is an interpretable classifier. Indeed, FIS are a set of fuzzy “if-then” rules that can be learnt from data. By combining FuRIA features with FIS, we can obtain a set of rules describing which activity in which brain regions and frequency bands is related to which mental task. An evaluation of this method on the EEG data mentioned above suggested that the obtained rules were consistent with the literature (see Figure 3).

brain-computer interface (BCI), feature extraction, inverse model, source localization, classification, interpretability, fuzzy set, electroencephalography (EEG).

Remerciements

Les auteurs souhaitent remercier Oscar Yanez-Suarez pour son aide et ses conseils à propos de l'algorithme Mean Shift ainsi que Morgane Rosendale, Stéphanie Gerbaud, Yann Renard et Marco Congedo pour leur aide à la préparation de ce manuscrit. Les auteurs souhaitent également remercier les rapporteurs de cet article pour leurs remarques et leurs suggestions pour améliorer la méthode proposée. Ce travail a été soutenu par l'Agence Nationale de la Recherche à travers le projet Open-ViBE (ref. ANR05RNTL01601).

1. Introduction

Les Interfaces Cerveau-Ordinateur (ICO) sont des systèmes de communication permettant à un individu d'envoyer des commandes à un ordinateur uniquement par le biais de son activité cérébrale [WBM⁺02, PGN06, Bir06, CR07], celle-ci étant traitée et mesurée par le système [WLA⁺06, MBF⁺07]. Les ICO ont permis à des personnes atteintes de handicaps sévères d'utiliser, uniquement grâce à des variations volontaires de leur activité cérébrale, divers appareils de suppléance fonctionnelle tels que des prothèses [GHHP99, MPP08], des éditeurs de texte [SD06, BKG⁺00] ou bien des fauteuils roulants [RBG⁺07, VML⁺07]. Les ICO ont également un nombre grandissant d'applications dans le domaine de la robotique [MRMG04, LLA07b] ou bien du multimédia et de la réalité virtuelle [LSF⁺07, LLR⁺06].

Mesurer l'activité cérébrale dans une ICO peut se faire de différentes manières [WLA⁺06], mais la méthode la plus utilisée reste l'ÉlectroEncéphaloGraphie (EEG). En effet, l'EEG est une méthode d'imagerie cérébrale peu coûteuse, non-invasive, portable et fournissant une bonne résolution temporelle [NdS05]. Ainsi dans cet article, nous nous intéresserons uniquement aux ICO utilisant l'EEG. L'objectif du traitement des signaux EEG est de détecter ou de distinguer différents états mentaux (aussi appelés « tâches mentales »), et d'associer une commande à chacun de ces états, c'est-à-dire à chaque « classe » de signaux. L'identification de ces tâches mentales se fait le plus souvent par des méthodes de reconnaissance de formes et repose donc sur deux étapes clés qui sont l'extraction de caractéristiques [BFWB07] et la classification [LCL⁺07]. Le but de la première étape est de décrire les signaux EEG par un faible nombre de valeurs pertinentes appelées « caractéristiques » tandis que la deuxième a pour objectif d'attribuer automatiquement une classe à ces caractéristiques (la classe de l'état mental identifié). Dans cet article nous nous intéressons à l'extraction de caractéristiques, domaine pour lequel l'exploration de nouveaux algorithmes s'avère nécessaire [MAM⁺06]. En effet, les performances des algorithmes actuels restent encore relativement modestes et il serait intéressant de pouvoir les interpréter c'est-à-dire d'extraire de la connaissance des caractéristiques utilisées dans les ICO [MAM⁺06].

Dans cet article, nous proposons un nouvel algorithme d'extraction de caractéristiques s'appuyant sur l'utilisation de modèles inverses. De tels modèles peuvent reconstruire l'activité dans le cerveau, en utilisant uniquement les signaux EEG enregistrés en surface du crâne et un modèle de la conductivité de la tête. L'algorithme proposé se nomme FuRIA, acronyme de *Fuzzy Region of Interest Activity*.

Cet article est organisé de la façon suivante : la partie 2 présente un état-de-l'art de l'utilisation des modèles inverses pour la conception d'interfaces cerveau-ordinateur. L'algorithme FuRIA est présenté dans la partie 3. Les parties 4 et 5 présentent respectivement les choix d'implémentation de l'algorithme ainsi qu'une évaluation de celui-ci sur des données de la « BCI

compétition 2003 ». À cette occasion nous montrerons que l'utilisation de FuRIA en collaboration avec des systèmes d'inférence flous permet la conception d'ICO interprétables. En effet, il est ainsi possible de comprendre ce que le système a appris automatiquement, et donc d'en extraire de nouvelles connaissances. Enfin, la partie 6 propose une conclusion sur ces travaux ainsi que quelques perspectives.

2. Interfaces cerveau-ordinateur et modèles inverses : un état-de-l'art

Bien que de nombreuses méthodes d'extraction de caractéristiques aient été explorées pour les ICO [BFWB07], ce n'est que très récemment que les modèles inverses ont été considérés comme une méthode potentiellement intéressante [QDH04, GGP⁺05, CLL06]. Cette partie propose un état-de-l'art de l'utilisation de ces modèles inverses pour l'extraction de caractéristiques dans les ICO.

Les modèles inverses permettent d'avoir accès à une nouvelle information, à savoir l'activité dans le volume cérébral. Cette information ayant de solides bases physiologiques, elle apparaît comme un nouveau type de caractéristiques particulièrement intéressant. Les études menées jusqu'à présent ont donc cherché à évaluer l'utilité de ces modèles inverses comme extracteur de caractéristiques et ont obtenu des premiers résultats très prometteurs [LLA07a, BGM07, BCM⁺07, NKM08, KLH05]. Différents types de modèles inverses ont été explorés pour les ICO, tels que les modèles distribués, estimant l'activité d'un grand nombre de dipôles, modélisés par des voxels (éléments de volume), distribués dans tout le cortex ou le cerveau. Dans cette catégorie on trouve par exemple les modèles inverses LORETA/sLORETA [CLL06, LNM06], ELECTRA [GGP⁺05] ou encore la méthode de la norme minimale pondérée en profondeur [BCM⁺07]. Des auteurs ont également utilisé des modèles à dipôle équivalent, estimant la position, l'intensité et l'orientation d'une ou de deux sources principales, modélisées chacune par un dipôle [QDH04, KLH05].

Dans le but de concevoir des ICO, les modèles inverses sont généralement utilisés de deux façons :

- Comme méthode directe d'extraction de caractéristiques. Dans ce cas, la valeur de l'activité cérébrale, reconstruite dans certaines Régions d'Intérêt (RI) [CLL06, LNM06] ou bien la position des sources d'activité [QDH04, KLH05, WGWB05] sont directement utilisées comme caractéristiques servant à déterminer la tâche mentale effectuée.
- Comme méthode de prétraitement précédant l'extraction de caractéristiques. Dans ce cas, le modèle inverse est utilisé pour reconstruire les signaux correspondant au décours temporel de l'activité de chaque voxel. L'étape d'extraction de caractéris-

tiques est ensuite effectuée sur ces nouveaux signaux [GGP+05, BGM07, BCL+07, NKM08].

Ces méthodes ont permis d'obtenir des résultats très satisfaisant, généralement aussi bons voire meilleurs que ceux existant dans la littérature. En effet, de plus en plus d'indices tendent à montrer que l'utilisation de modèles inverses permettrait d'augmenter les performances du système, en terme de taux de reconnaissance correcte. Ainsi, il a été observé qu'extraire des caractéristiques dans le domaine des sources (obtenues grâce au modèle inverse) serait plus efficace que de les extraire directement dans le domaine des électrodes (signaux bruts) [GGP+05, BGM07, NKM08]. Ces bons résultats peuvent s'expliquer si l'on considère les modèles inverses comme des filtres spatiaux basés sur des a priori physiologiques. Ils vont donc permettre de se concentrer sur l'information utile en éliminant le bruit provenant des autres régions cérébrales, non liées aux tâches mentales effectuées par le sujet.

En dépit de ces résultats prometteurs, quelques limitations subsistent dans l'utilisation des modèles inverses pour les ICO. Il s'avère en effet difficile de concilier à la fois la généralité, c'est-à-dire la capacité de traiter n'importe quel type de tâche mentale, et le fait de générer peu de caractéristiques. En effet, plusieurs méthodes basées sur les modèles inverses nécessitent de fortes connaissances *a priori* concernant les mécanismes neuro-physiologiques associés aux tâches mentales utilisées, et ne sont donc pas du tout généralistes [QDH04, WGWB05, KLH05, LNM06, BCM+07]. Avec ces méthodes, il faut définir au préalable et à la main quelles sont les RI à utiliser. Ces méthodes sont, pour l'instant, limitées à l'utilisation de tâches mentales faisant intervenir les aires motrices du cerveau. Le fait d'utiliser des régions prédéfinies pose de plus le problème de la spécialisation. En effet, il est bien connu que chaque sujet possède ses propres spécificités, en termes de caractéristiques spatiales (régions cérébrales mises en jeu), et fréquentielles (bandes de fréquences dans lesquelles ont lieu les variations d'intérêt) [dRMFn+02]. Ainsi, une méthode non généraliste, car basée exclusivement sur des connaissances *a priori*, aura des difficultés à s'adapter aux spécificités de chaque sujet, et n'aura donc vraisemblablement pas des performances optimales.

Plusieurs méthodes généralistes et automatiques, toutes basées sur des modèles inverses distribués, ont été proposées [GGP+05, BGM07, NKM08]. Leur principal inconvénient vient du fait que pour parvenir à cette généralité, ces méthodes doivent extraire une ou plusieurs caractéristiques par voxel, ce qui génère globalement un très grand nombre de caractéristiques. En effet les modèles de tête généralement utilisés comportent des centaines ou des milliers de voxels. Ainsi, l'utilisation de techniques de sélection de caractéristiques [GGP+05, NKM08] ou de classificateurs parcimonieux [BGM07] s'avère indispensable. Bien que cette solution soit satisfaisante et donne de bons résultats, le nombre de caractéristiques utilisées reste généralement relativement élevé, notamment en comparaison avec le nombre de caractéristiques utilisées par les méthodes non généralistes [LNM06, QDH04]. De plus, dans ces méthodes, tous les voxels

sont gérés de manière indépendante, alors qu'ils sont généralement dépendants les uns des autres et devraient donc être regroupés au sein de régions cérébrales.

En 2006, Congedo *et coll.* ont proposé une méthode qui est à la fois généraliste et génératrice de peu de caractéristiques étant donné que les voxels sont regroupés en Régions d'Intérêt [CLL06]. Cette méthode, combinant un filtre spatial dépendant des données et le modèle inverse sLORETA, a permis d'obtenir des résultats comparables à ceux du gagnant de la «BCI competition 2003». Elle n'est cependant pas complètement automatique et reste limitée à l'utilisation de deux RI dont l'extension spatiale est difficile à déterminer [CLL06].

3. L'algorithme FuRIA

Dans ce papier, nous proposons l'algorithme FuRIA, qui est un algorithme d'extraction de caractéristiques basé sur les modèles inverses distribués. Il est généraliste et automatique, peut utiliser n'importe quel nombre de RI pertinentes et définir leur extension spatiale grâce à l'utilisation du nouveau concept de RI flou. FuRIA est un algorithme d'extraction de caractéristiques avec apprentissage. En effet, il peut apprendre, à l'aide d'exemples, quelles sont les RI pertinentes pour discriminer des tâches mentales données ainsi qu'apprendre les bandes de fréquences dans lesquelles ces RI réagissent. Cette partie décrit le fonctionnement de cet apprentissage ainsi que l'utilisation de FuRIA pour l'extraction de caractéristiques.

3.1 Architecture générale de l'algorithme d'apprentissage

Pour utiliser FuRIA, une phase d'apprentissage est nécessaire. Cet apprentissage a pour objectif d'identifier les RI et les bandes de fréquences dont l'activité permet de discriminer le mieux possible les différentes tâches mentales. L'apprentissage est effectué hors-ligne et est composé de 5 étapes successives :

1. **Analyse statistique** : Une analyse statistique comparant l'activité entre les différentes classes (tâches mentales) est effectuée pour chaque fréquence et chaque voxel. L'objectif est d'identifier les voxels dont l'activité pour une fréquence donnée peut discriminer les classes.
2. **«Clustering»** : Un algorithme de «clustering» est effectué afin de regrouper les voxels dont l'activité est statistiquement discriminante sous forme de RI.
3. **Fusion** : Les RI obtenues à des positions similaires, pour des fréquences consécutives, sont fusionnées et identifiées comme réactives dans la bande de fréquence résultant de la concaténation de ces fréquences.
4. **Adaptation** : Pour chaque RI obtenue pendant la fusion, une autre analyse statistique est effectuée, afin d'enlever les voxels qui ne sont plus statistiquement discriminant dans la bande de fréquence identifiée à l'étape précédente.

5. «**Fuzzification**» : Chaque RI est «fuzzifiée», c'est-à-dire transformée en RI floue (cf. Partie 3.1.5) afin de définir correctement son étendue spatiale.

Les parties suivantes détaillent ces différentes étapes.

3.1.1 Analyse statistique

Cette étape consiste à comparer, en utilisant un test statistique, l'activité de chaque voxel, moyennée sur une fenêtre temporelle donnée, entre les différentes classes. Cette analyse est effectuée pour chacune des N_f fréquences considérées. Pour cela les signaux EEG sont filtrés dans une bande de fréquence de 2 Hz autour de chaque fréquence considérée, et un modèle inverse est utilisé pour calculer l'activité des différents voxels. Une fois l'analyse effectuée, nous obtenons N_f cartes de statistiques, chacune révélant quels sont les voxels dont l'activité est statistiquement différente entre les classes pour la fréquence correspondante. Nous appelons ces voxels «voxels significatifs». Lors de la conception d'ICO binaires (utilisant seulement deux tâches mentales), la famille des tests t peut être par exemple utilisée, tandis que pour concevoir des ICO multi-classes, la famille des ANOVA peut être utilisée.

3.1.2 Clustering

Pour chaque carte de statistiques calculée à l'étape précédente, un algorithme de clustering est utilisé dans le but d'agréger les voxels significatifs sous forme de RI. Pour utiliser cet algorithme, un vecteur à 4 dimensions $[x, y, z, s]$ est associé à chaque voxel significatif. Les coordonnées x, y, z sont les coordonnées 3D du voxel dans le modèle de tête choisi, et s la statistique du voxel obtenue à l'étape précédente. Le but est de trouver les voxels significatifs qui sont spatialement regroupés (c'est-à-dire les voxels proches ou voisins), et qui ont des statistiques communes. Ce dernier point permet principalement de ne pas agréger ensemble un voxel avec une statistique positive et un voxel avec une statistique négative. L'utilisation des coordonnées spatiales des voxels significatifs en entrée de l'algorithme de clustering a pour objectif d'identifier quels sont les groupes de voxels significatifs qui sont voisins ou proches les uns des autres. En effet, les voxels significatifs voisins ou proches appartiennent très vraisemblablement à la même source d'activité. Il est cependant à noter qu'aucune contrainte topologique de type voisinage n'est utilisée dans cet algorithme.

Une fois le clustering effectué, tous les voxels dont les vecteurs associés appartiennent au même cluster, sont agrégés dans la même RI. Les RI très petites, contenant seulement un ou deux voxels par exemple, peuvent être supprimées. En effet, physiologiquement, il est assez peu probable de trouver une source d'activité ne contenant qu'un ou deux voxels. Ceci est d'autant plus vrai lorsque le modèle inverse utilisé est un modèle basse résolution tel que sLORETA [PM02]. En effet, de tels modèles ont tendance à trouver des sources relativement «étalées» spatialement.

À la fin de cette étape, un ensemble de RI pertinentes a été identifié pour chaque fréquence. Idéalement, l'algorithme de clustering utilisé doit pouvoir identifier automatiquement n'importe quel nombre de clusters et ne doit pas faire d'hypothèses trop fortes quant à la forme de ces clusters.

3.1.3 Fusion

Après avoir identifié quelles étaient les RI pertinentes pour chaque fréquence individuelle, il faut trouver quelles RI sont pertinentes dans quelles bandes de fréquences. Pour ce faire, chaque RI Ω est associée à la fréquence f à laquelle elle a été trouvée, puis la procédure suivante est effectuée :

1. Parmi tous les couples $(\Omega; f)$, choisir deux couples $(\Omega_1; f_1)$ et $(\Omega_2; f_2)$ tels que le chevauchement entre Ω_1 et Ω_2 soit élevé et les bandes de fréquences f_1 et f_2 soient consécutives ou se chevauchent. Nous pouvons considérer qu'un chevauchement est élevé si $card(\Omega_1 \cap \Omega_2) > 0.5 * \min(card(\Omega_1), card(\Omega_2))$ avec $card(\Omega)$ étant le nombre de voxels dans Ω ;
2. Remplacer les couples $(\Omega_1; f_1)$ et $(\Omega_2; f_2)$ par l'unique couple $(\Omega_1 \cup \Omega_2; f_1 \cup f_2)$. Cela signifie que nous avons choisi l'Union comme opérateur de fusion de RI ;
3. Retourner au point 1, jusqu'à ce que plus aucune RI ne puisse être fusionnée.

3.1.4 Adaptation

En raison de l'utilisation de l'Union comme opérateur de fusion, une RI Ω identifiée comme réactive dans la bande de fréquence $[f_a; f_b]$ peut contenir des voxels qui étaient significatifs à une fréquence comprise entre f_a et f_b mais qui ne le sont plus dans $[f_a; f_b]$. Par conséquent, pour chaque RI trouvée à l'étape précédente, nous effectuons une autre analyse statistique pour les voxels appartenant à la RI, dans la bande de fréquence qui lui est associée, afin de retirer ceux qui ne sont plus significatifs. De plus, les statistiques obtenues à cette étape vont être utilisées lors de la prochaine et dernière étape.

3.1.5 Fuzzification et concept de RI floue

Dans les RI définies jusqu'à présent, nous pouvons remarquer que tous les voxels n'ont pas, statistiquement, le même pouvoir discriminant. Cependant, ces voxels portent tous plus ou moins d'information. De plus, pour un sujet donné, les régions cérébrales correspondant à des fonctions cérébrales particulières ne sont pas bien définies : les limites de ces régions sont naturellement floues. De ce fait, identifier quels sont les voxels qui doivent être gardés dans une RI est un problème difficile.

Nous pensons que tous les voxels contenant de l'information devraient être dans la RI, mais que ceux qui en contiennent moins, devraient être «moins» dans la RI. Nous proposons ainsi le concept de RI floue, en s'appuyant sur la théorie des ensembles flous développée par Zadeh [Zad96].

Une RI classique Ω_c est définie par l'ensemble des voxels qu'elle contient. L'activité γ_{Ω_c} de cette RI Ω_c est classiquement

calculée comme suit :

$$\gamma_{\Omega_c} = \sum_{v \in \Omega_c} \gamma_v \quad (1)$$

où γ_v est l'activité du voxel v .

Une RI floue Ω_f n'est plus définie par un ensemble de voxels, mais par une fonction d'appartenance floue μ . Cette fonction donne le degré d'appartenance, dans $[0; 1]$, de n'importe quel voxel existant dans la RI Ω_f . De cette façon, les RI floues ont des limites qui ne sont plus strictement définies et sont donc naturellement «floues». Du point de vue de la théorie des ensembles flous de Zadeh, une RI floue est donc un ensemble flou de voxels. L'activité dans une RI floue est calculée de cette façon :

$$\gamma_{\Omega_f} = \sum_{v=1}^{N_v} \mu(v) \gamma_v \quad (2)$$

où N_v est le nombre total de voxels dans le modèle de tête utilisé. Ce formalisme permet de pondérer, au sein de la RI, chaque voxel en fonction de son pouvoir discriminant, et ainsi, d'utiliser efficacement toute l'information disponible. En guise de récapitulatif, le tableau 1 propose une comparaison entre RI classiques et RI floues, à travers leur définition et le calcul de l'activité dans ces RI.

Pour terminer l'apprentissage de FuRIA, toutes les RI Ω sont

Tableau 1. Comparaison entre régions d'intérêt classiques et régions d'intérêt floues.

	RI classique Ω_c	RI floue Ω_f
Définition	$\Omega_c = \{v v \in \Omega_c\}$	$\Omega_f = \{v \mu_{\Omega_f}(v) > 0\}$
Activité γ	$\gamma_{\Omega_c} = \sum_{v \in \Omega_c} \gamma_v$	$\gamma_{\Omega_f} = \sum_{v=1}^{N_v} \mu_{\Omega_f}(v) \gamma_v$

transformées en RI floues en associant à chacune d'entre elles une fonction d'appartenance floue μ_{Ω} . Cette fonction utilise les statistiques calculées lors de l'étape précédente afin d'identifier le pouvoir discriminant de chaque voxel. L'expression de cette fonction μ_{Ω} dépend du modèle inverse et du test statistique utilisé. Dans la partie 4 nous proposons une fonction d'appartenance floue adaptée aux algorithmes que nous avons utilisés. Naturellement, une grande variété de fonctions d'appartenance floues peuvent être utilisées en fonction de l'objectif visé.

3.2 Extraction de caractéristiques avec l'algorithme FuRIA

Une fois l'apprentissage automatique de FuRIA terminé, un ensemble de RI floues Ω_i , pertinentes pour la discrimination, a été identifié, ainsi que les bandes de fréquence f_i correspondant à ces RI. L'extraction de caractéristiques consiste alors, pour chacun des N_{Ω} couples (Ω_i, f_i) , à filtrer les signaux EEG dans la bande de fréquence f_i puis à utiliser un modèle inverse pour

calculer l'activité γ_i dans la RI Ω_i . Cette activité est ensuite moyennée sur une fenêtre temporelle donnée. Enfin, les valeurs de ces activités sont ensuite concaténées en un vecteur de caractéristiques $[\gamma_1, \gamma_2, \dots, \gamma_{N_{\Omega}}]$ qui peut alors être utilisé avec n'importe quel classifieur, dans le but d'identifier la tâche mentale effectuée par l'utilisateur.

4. Implémentation de l'algorithme

Grâce à sa décomposition en plusieurs étapes, FuRIA est un algorithme modulaire qui peut être utilisé avec différentes implémentations pour chacune de ces étapes. Ainsi, différents modèles inverses, différents algorithmes de clustering, différents tests statistiques ou différentes fonctions d'appartenance floues peuvent être utilisés. Cette partie présente les méthodes que nous avons utilisées dans notre implémentation afin d'évaluer l'algorithme FuRIA.

4.1 Le modèle inverse sLORETA

Nous avons choisi d'utiliser le modèle inverse sLORETA (standardized Low Resolution Electromagnetic Tomography) proposé par Pascual-Marqui [PM02]. sLORETA est une solution linéaire, distribuée, discrète et instantanée du problème inverse. Le fait que cette méthode soit linéaire est une propriété essentielle étant données les fortes contraintes temps réel imposées lors de l'utilisation en ligne d'une ICO.

Ce modèle inverse a l'avantage d'avoir de très bonnes capacités de localisation (ce qui est essentiel si l'on cherche à interpréter les RI obtenues) et de ne pas avoir de biais de localisation en présence de bruit physiologique ou de bruit dû aux mesures [PM07]. Enfin, il a été démontré expérimentalement que l'algorithme sLORETA était adapté pour la conception d'ICO basée sur l'EEG [CLL06].

Pour comprendre comment sLORETA propose une solution au problème inverse, considérons d'abord le problème dit «direct» [MML+04, PM99]. Celui-ci permet de calculer quels seront les signaux EEG mesurés en fonction de l'amplitude et l'orientation des dipôles dans le volume cérébral, de la position des électrodes de mesure et des propriétés de conductivité de la tête :

$$m = Kc \quad (3)$$

où m est le vecteur de dimension N_e des mesures instantanées des signaux EEG (une mesure pour chacune des N_e électrodes), c le vecteur de dimension $3N_v$ représentant l'activité et l'orientation de chacun des N_v voxels (dipôles) à un instant donné. Dans c , les 3 composantes associées à chaque voxel correspondent à l'intensité du courant dipolaire selon chacune des 3 coor-

données spatiales. Enfin, K est une matrice représentant le modèle des milieux conducteurs (c'est-à-dire le modèle de tête) ainsi que la position des électrodes de mesures. Résoudre le problème inverse pour obtenir l'estimation \hat{c} de l'activité des sources consiste donc à résoudre le système d'équations suivant :

$$\hat{c} = Tm \quad (4)$$

où T est l'inverse généralisée de K , que l'on appelle matrice de transfert. Ce système d'équations admet une infinité de solutions, étant donné que le nombre de voxels utilisés est très supérieur au nombre d'électrodes. De plus, différentes configurations de sources peuvent donner lieu aux mêmes mesures en surface. Une solution régularisée par moindres carrés est la suivante :

$$T = K^T(KK^T + \alpha_R X)^+ \quad (5)$$

où α_R est un paramètre de régularisation positif et X la matrice de centrage, pour exprimer les données en référence moyenne commune. L'exposant T exprime la transposition, et l'exposant $+$ la pseudo-inverse matricielle. Dans le but d'améliorer la capacité de localisation, sLORETA s'appuie sur une standardisation de cette solution. Ainsi l'activité γ_v d'un voxel v est obtenue de la manière suivante avec sLORETA :

$$\gamma_v = \hat{c}_v^T S_v^{-1} \hat{c}_v \quad (6)$$

où $S = TK$, avec S_v étant le v -ième bloc 3×3 de S . De même \hat{c}_v est le v -ième triplet de \hat{c} .

Congedo a démontré que l'équation 6 pouvait se réécrire sous forme quadratique [Con06] :

$$\gamma_v = m^T Q_v m \quad (7)$$

où $Q_v = T_v^T S_v^{-1} T_v$. Avec cette formulation, l'activité γ_Ω dans une RI Ω est donc calculée comme suit :

$$\gamma_\Omega = \sum_{v \in \Omega} \gamma_v = m^T Q_\Omega m \quad \text{avec} \quad Q_\Omega = \sum_{v \in \Omega} Q_v \quad (8)$$

Ainsi, calculer l'activité γ_{Ω_f} dans une RI floue Ω_f peut être fait grâce à un simple produit matriciel :

$$\gamma_{\Omega_f} = m^T Q_{\Omega_f} m \quad \text{avec} \quad Q_{\Omega_f} = \sum_{v=1}^{N_v} \mu(v) Q_v \quad (9)$$

Cette formulation permet donc de calculer l'activité dans une RI floue de manière très rapide étant donné que Q_{Ω_f} peut être calculée hors-ligne et que sa taille ne dépend que du nombre d'électrodes et pas du nombre de voxels considérés.

Dans cette étude, nous avons utilisé un modèle de tête sphérique standard contenant 2394 voxels, disponible dans le logiciel LORETA-key [PM]. Ce modèle de tête est composé de trois sphères concentriques qui représentent le scalp (la peau), les os du crâne et le cerveau. Concernant la régularisation dans

sLORETA, nous avons utilisé $\alpha_R = 1000$, cette valeur ayant été identifiée comme étant la plus adaptée lors d'une étude précédente pourtant sur les mêmes données [CLL06].

4.2 Test statistique

Lors de cette évaluation nous avons utilisé comme test statistique un test t avec permutations multiples, comparaisons multiples, tel que proposé par Holmes et coll. [HBWF96]. Cette méthode a l'avantage d'être non-paramétrique, de permettre le calcul exact de l'erreur de type I et d'avoir été conçue pour identifier les zones actives dans des images obtenues par Imagerie par Résonance Magnétique (IRM). Elle apparaît donc comme parfaitement adaptée à notre problématique. Lors de nos évaluations, nous avons considéré qu'un voxel était significatif si la probabilité d'erreur de type I qui lui était associée était inférieure à $\alpha = 0.01$.

4.3 Algorithme de clustering

Comme algorithme de clustering, nous avons choisi d'utiliser l'algorithme «Mean Shift», proposé par Comaniciu et Meer [CM02]. En effet cet algorithme est également non-paramétrique et peut identifier n'importe quel nombre de clusters de formes quelconques. De plus, il a été récemment utilisé pour segmenter des images IRM, ce qui est un problème relativement proche du notre [JAPAMBYS06]. Cet algorithme nécessite le choix d'une valeur pour son hyperparamètre H . Ce paramètre, appelé paramètre de lissage, permet de déterminer la résolution du clustering. Ainsi, une valeur basse de H donnera lieu à un grand nombre de petits clusters, tandis qu'une valeur élevée donnera lieu à un petit nombre de grands clusters. Lors de nos évaluations, nous avons utilisé $H = 2$.

4.4 Choix des fonctions d'appartenance floues

Durant l'étape de fuzzification, nous avons eu à choisir les fonctions d'appartenance floues les plus adaptées. Comme sLORETA possède une faible résolution spatiale, les voxels significatifs avec une relativement forte probabilité d'erreur de type I sont beaucoup plus nombreux que les voxels significatifs avec une relativement faible probabilité d'erreur de type I. Par conséquent, pour chaque RI Ω , nous avons choisi une fonction μ_Ω donnant un degré d'appartenance qui décroît exponentiellement avec la statistique du voxel :

$$\mu_\Omega(v) = \begin{cases} e^{-\frac{1}{2} \left(\frac{p_{max} - p_v}{\sigma_{max}} \right)^2} & v \in \Omega \\ 0 & v \notin \Omega \end{cases} \quad (10)$$

avec $\sigma_{max} = 1 / (N_{\Omega_s} \sum_{p_v > 0} (p_{min} - p_v)^2)$

N_{Ω_s} est le nombre de voxels dans la RI Ω ; p_{min} et p_{max} sont respectivement la probabilité minimale et maximale, parmi les voxels de la RI, de rejeter correctement l'hypothèse nulle H_0 ; et p_v est la probabilité de rejeter correctement H_0 pour le voxel v . Plus la RI est large, plus le nombre de voxels avec une faible valeur de p_v est grand lorsque l'on utilise sLORETA. Cela a été pris en compte en rajoutant le terme N_{Ω_s} au dénominateur de σ_{max} .

Une contrainte de FuRIA est la nécessité d'utiliser un nombre d'électrodes suffisamment grand pour pouvoir obtenir une solution correcte au problème inverse. La phase d'apprentissage de FuRIA peut également s'avérer longue si l'analyse statistique utilisée est coûteuse en temps de calcul, comme c'est le cas pour les tests avec permutations que nous avons utilisés. L'apprentissage s'effectuant hors-ligne, cette limitation n'est cependant pas critique.

5. Évaluation

Cette partie présente une première évaluation de l'algorithme FuRIA, sur des données de la «BCI competition 2003». Tout d'abord les données EEG utilisées sont présentées puis les résultats obtenus sont décrits. Enfin, nous verrons comment exploiter ces résultats afin de concevoir une ICO interprétable.

5.1 Jeu de données EEG

Nous avons utilisé le jeu de données IV de la «BCI competition 2003» [BMC⁺04]. Ces données contiennent des signaux EEG enregistrés alors qu'un sujet tapait sur un clavier avec un doigt de sa main gauche ou droite. Les signaux EEG ont été échantillonnés à 100 Hz, enregistrés avec 28 électrodes et contenaient les 500 ms précédant la frappe effective sur le clavier. Les données contenaient un ensemble d'apprentissage composé de 314 essais, étiquetés avec leur classe d'appartenance, ainsi qu'un ensemble de test comportant 100 essais, non étiquetés. Ainsi, pour la compétition, les participants devaient identifier, en s'aidant de l'ensemble d'apprentissage, la main qui allait être utilisée pour chaque essai de l'ensemble de test. Il s'agissait donc d'un problème de classification binaire où les classes étaient

«gauche» et «droite». Le gagnant était celui dont l'algorithme permettait d'obtenir le plus grand taux de reconnaissance correcte.

5.2 Résultats de l'extraction de caractéristiques avec FuRIA

Nous avons entraîné l'algorithme FuRIA sur l'ensemble d'apprentissage fourni. Nous avons considéré les fréquences comprises entre 3 Hz et 45 Hz, avec un pas de 1 Hz. Nous nous sommes focalisés sur la fenêtre temporelle comprenant les dernières 250 ms pour chaque essai, c'est-à-dire que l'activité calculée dans les différentes RI était moyennée sur cette fenêtre temporelle. Les filtrages passe-bande ont été effectués à l'aide de filtres à Réponse Impulsionnelle Finie (RIF). Suite à son apprentissage, FuRIA a identifié cinq RI flous discriminantes, générant ainsi des vecteurs de caractéristiques de dimension 5. Ces RI et leurs bandes de fréquences correspondantes sont présentées sur la Figure 1.

Nous pouvons remarquer que FuRIA a trouvé deux RI dont les fréquences correspondantes sont dans le rythme μ (13 Hz) et trois RI avec des fréquences dans le rythme β (15-16 Hz, 15-36 Hz, 18-28 Hz). De plus, les RIs trouvées se situent au niveau des aires motrices des cortex gauche et droit. Ceci s'avère être très cohérent avec la littérature concernant les effets de l'intention de mouvements sur l'activité cérébrale. En effet, on sait que cela provoque des variations (en anglais appelées «Event Related Desynchronisation» lors de diminutions de l'amplitude du signal et «Event Related Synchronisation» lors d'augmentations de cette amplitude) de l'amplitude des rythmes μ et β au niveau des cortex moteurs et sensorimoteurs [Pfu99]. Ainsi, ces observations montrent un avantage de l'extraction de caractéristiques basée sur les modèles inverses, et donc un avantage de FuRIA : les caractéristiques extraites (et apprises) semblent interprétables, puisqu'elles nous renseignent sur les régions cérébrales et les bandes de fréquences impliquées par les tâches mentales utilisées. Ainsi, un tel algorithme pourrait être utilisé pour tester et évaluer différentes tâches mentales tout en extrayant, par la même occasion, de la connaissance sur le fonctionnement cérébral relatif à ces tâches. Il convient de noter que nous avons utilisé au sein de sLORETA un modèle de tête standard et non réaliste (car modélisant simplement la tête comme trois sphères concentriques). Afin d'avoir une localisa-

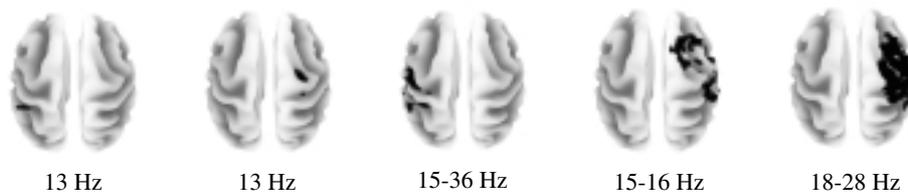


Figure 1. Les RI flous (en noir) et leurs fréquences correspondantes qui ont été automatiquement obtenues par FuRIA. Ces images ont été obtenues à l'aide du logiciel LORETA-key [PM]. Le cerveau est représenté en vue de dessus, la partie frontale orientée vers le haut.

tion de sources plus précise et plus exacte, et donc une meilleure interprétabilité, il serait intéressant de travailler avec un modèle de tête réaliste correspondant à l'anatomie du sujet, c'est-à-dire avec un modèle de tête généré à partir d'images IRM (Imagerie par Résonance Magnétique) de la tête du sujet.

5.3 Classification

Pour classifier les caractéristiques extraites en utilisant FuRIA, nous avons utilisé une machine à vecteurs supports (SVM) linéaire, parfois aussi appelée Séparateur à Vastes Marges [Bur98, Cor02]. Ce classifieur est réputé pour son efficacité et ses très bonnes capacités de généralisation. Dans le domaine des ICO, il a fait preuve de son efficacité dans de nombreuses études et est désormais l'un des classifieurs les plus populaires [LCL⁺07]. Ce classifieur a été entraîné sur l'ensemble d'apprentissage et ensuite utilisé pour classifier les caractéristiques de l'ensemble de test. La valeur optimale du paramètre de régularisation du SVM a été sélectionnée par validation croisée. Les taux de reconnaissance obtenus après 100 répétitions d'apprentissage/test sont rapportés dans le tableau 2 et comparés avec deux autres méthodes : celle du gagnant de la compétition [WZL⁺04] et une méthode utilisant les modèles inverses que nous avons proposée précédemment [CLL06]. Les résultats obtenus avec FuRIA sans l'étape de fuzzification (cf partie 3.1.5) sont aussi rapportés dans le but d'évaluer l'apport de la fuzzification.

Tableau 2. Comparaison de différentes méthodes en termes de taux de reconnaissance sur l'ensemble de test du jeu de données IV de la «BCI competition 2003».

Méthode	Taux de reconnaissance
Gagnants «BCI competition 2003» [WZL+04]	84 %
Travaux précédents [CLL06]	83 %
FuRIA sans fuzzification	80-82 %
FuRIA avec fuzzification	85-86 %

Nous pouvons remarquer que la méthode utilisant FuRIA avec fuzzification atteint le plus haut score avec un taux de reconnaissance allant de 85 % à 86 %. Ces résultats suggèrent que les caractéristiques extraites avec FuRIA sont très discriminantes, étant donné qu'elles sont peu nombreuses mais permettent de bons résultats de classification. Nous pouvons aussi remarquer que lorsque l'on utilise FuRIA sans fuzzification, les résultats sont moins bons, ce qui montre l'intérêt d'utiliser le concept de RI flou. En effet ce concept semble permettre d'utiliser plus efficacement l'information utile, et donc d'augmenter les capacités de reconnaissance du système. Enfin, il est intéressant de remarquer que dans les méthodes proposées dans [CLL06] et

[WZL+04], des filtres spatiaux dépendants des données ont été utilisés pour augmenter les performances. Ces filtres sont respectivement les motifs spatiaux communs («Common Spatial Pattern» en anglais) et la décomposition en sous-espaces communs («Common Sub-Space Decomposition» en anglais). Cela suggère que combiner FuRIA avec des filtres tels que ceux décrits dans [CLL06] pourrait encore améliorer les résultats.

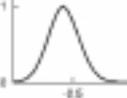
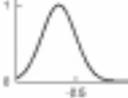
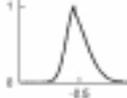
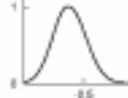
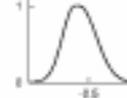
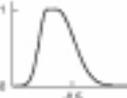
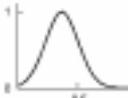
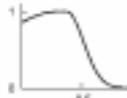
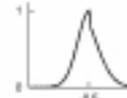
5.4 Vers des interfaces cerveau-ordinateur interprétables

Dans la partie 5.2 nous avons mis en avant le fait que les caractéristiques apprises et extraites à l'aide de FuRIA semblaient interprétables. En ceci FuRIA répond à un des problèmes posés par la communauté ICO, à savoir le manque de caractéristiques dont on peut extraire de la connaissance [MAM+06]. Les caractéristiques extraites avec FuRIA nous permettent de savoir quelles sont les RI et les bandes de fréquences mises en jeu par un certain nombre de tâches mentales. Elles ne nous permettent pas de savoir quelles valeurs de l'activité dans ces régions et fréquences sont associées à quelles tâches mentales. Dans le but de faire un pas de plus vers la conception d'interfaces cerveau-ordinateur interprétables, nous proposons de classifier les caractéristiques obtenues avec FuRIA à l'aide de Systèmes d'Inférence Flous (SIF), plutôt qu'avec des SVM. Les SIF sont un ensemble de règles «si – alors» floues, pouvant être apprises automatiquement à partir de données. Ces règles sont de la forme suivante :

Si X_1 est A_{j1} et X_2 est A_{j2} et ... alors la classe est Cl_i

où X_k est le $k^{\text{ième}}$ élément du vecteur de caractéristiques et A_{jk} est une fonction d'appartenance floue, décrivant comment la caractéristique X_k est distribuée pour la classe Cl_i . Plus précisément, A_{jk} est une fonction qui associe à une valeur de X_k (ici normalisée entre -1 et 1) un degré d'appartenance floue dans $[0; 1]$. Cette formalisation sous forme de règles fait des SIF des classifieurs également interprétables, si les caractéristiques qu'ils utilisent en entrée le sont également. De plus, l'intérêt des SIF pour les ICO a été démontré récemment [Lot06, LLLA07]. Les SIF s'avèrent donc être un candidat idéal pour classifier des caractéristiques extraites avec FuRIA. Nous avons donc utilisé le SIF étudié dans [LLLA07], et l'avons entraîné sur l'ensemble d'apprentissage utilisé plus haut. Le tableau 3 représente le fonctionnement de toute notre ICO, c'est-à-dire les caractéristiques et les règles apprises automatiquement, qui sont au nombre de deux. Dans ce tableau, les caractéristiques X_1, \dots, X_5 , sont représentées par la RI et la bande de fréquence dans laquelle la caractéristique est extraite. La première ligne du tableau correspond donc à la forme générale d'une règle floue, tandis que les deux lignes suivantes représentent les deux règles floues apprises. Elles décrivent donc les distributions des caractéristiques (ou fonctions d'appartenance floues A_{jk}) pour chacune des deux classes gauche et droite. L'abscisse des fonctions

Tableau 3. Règles floues et caractéristiques obtenues automatiquement par apprentissage.

	Si	et	et	et	et	alors
						la classe est
	13 Hz est	13 Hz est	15-36 Hz est	15-16 Hz est	18-28 Hz est	
Règle 1						Gauche
Règle 2						Droite

A_{jk} correspond à la valeur de la caractéristique X_k normalisée entre -1 et 1 , et l'ordonnée correspond au degré d'appartenance flou associé (dans $[0; 1]$, 1 représentant le degré d'appartenance maximum).

D'après la littérature on sait que l'intention de mouvement des mains est censée provoquer une baisse de la puissance dans les rythmes μ et β , contralatéralement à la main utilisée [Pfu99]. Cette dynamique semble effectivement se retrouver dans ce tableau. Regardons par exemple la 1ère et la 3ème caractéristique, correspondant toutes deux à l'activité dans le cortex moteur gauche, respectivement dans les rythmes μ (13 Hz) et β (15-36 Hz). On remarque que les valeurs de ces caractéristiques sont plus élevées (c'est-à-dire que la fonction d'appartenance floue associée est centrée sur une valeur plus élevée) dans la règle 1 qui décrit la classe «gauche», que dans la règle 2 décrivant la classe «droite». Combiner FuRIA avec des systèmes d'inférence flous semble donc permettre de concevoir une ICO interprétable. Ceci peut s'avérer utile afin de valider le système à l'aide des connaissances disponibles dans la littérature ou bien au contraire pour extraire de nouvelles connaissances de ce qu'a appris automatiquement une ICO basée sur de nouvelles tâches mentales.

En terme de performances, ce système obtient un score de 85 % sur l'ensemble de test, c'est-à-dire un score similaire à celui obtenu avec la SVM et légèrement meilleur que le score obtenu par le gagnant de la compétition. Ce système est donc également efficace en terme de taux de reconnaissance.

6. Conclusion et travaux futurs

Dans cet article, nous avons proposé un algorithme d'extraction de caractéristiques pour les interfaces cerveau-ordinateur appe-

lé FuRIA (*Fuzzy Region of Interest Activity*). Cet algorithme utilise des modèles inverses ainsi que le nouveau concept de Régions d'Intérêt (RI) floues. FuRIA peut automatiquement identifier quelles sont les RIs pertinentes pour discriminer des classes, ainsi que les bandes de fréquences dans lesquelles ces régions réagissent. L'activité dans ces RIs peut ensuite être utilisée en tant que caractéristique en entrée de n'importe quel classifieur. FuRIA a été évalué sur le jeu de données IV de la «BCI competition 2003» en utilisant une Machine à Vecteurs Supports (SVM) comme classifieur. Les résultats sont prometteurs étant donné que cette méthode a obtenu un meilleur score que les gagnants de cette compétition. Enfin, en utilisant un système d'inférence flou à la place de la SVM, nous sommes parvenus à concevoir une ICO interprétable, dont on pourrait extraire de la connaissance, répondant ainsi à l'un des problèmes soulevés par la communauté ICO [MAM+06].

Travaux futurs : Il serait intéressant d'évaluer FuRIA sur de nouveaux jeux de données, concernant d'autres sujets et de nouvelles et plus nombreuses tâches mentales. Il serait également intéressant d'évaluer la pertinence des RI apprises, en terme de performances et de signification neurophysiologique. Dans ce but, nous pourrions comparer les RI automatiquement obtenues par FuRIA et les performances résultantes avec celles de RI définies par des experts en neurophysiologie. Il pourrait également s'avérer utile d'étudier l'influence de la résolution spatiale (nombre d'électrodes ou de voxels utilisés) sur les performances du système. Une autre piste d'amélioration possible concerne l'utilisation de contraintes topologiques dans l'algorithme d'apprentissage de FuRIA. En effet, pour l'instant, les degrés d'appartenance flous attribués au voxels ne dépendent que du voxel concerné et pas de ses voisins ni de sa position dans la RI floue. Ainsi, utiliser des contraintes topologiques pourrait permettre de choisir des degrés d'appartenance plus finement, par exemple en s'appuyant sur la position du voxel dans l'atlas cérébral du sujet (obtenu suite à une imagerie par

résonance magnétique par exemple) ou bien en lissant la fonction d'appartenance floue pour que des voxels voisins aient des degrés d'appartenance flous proches. Ceci permettrait par exemple de limiter l'influence du bruit lors de l'apprentissage et peut-être d'augmenter les performances. Enfin, nous proposons d'enrichir l'algorithme en rajoutant la prise en compte de la dimension temporelle, en plus des dimensions spatiales et fréquentielles. Ceci permettrait d'extraire plus d'information et donc d'augmenter éventuellement les performances.

Références

- [BCM*07] F. BABILONI, F. CINCOTTI, M. MARCIANI, S. SALINARI, L. ASTOLFI, A. TOCCI, F. ALOISE, F. DE VICO FALLANI, S. BUFALARI and D. MATTIA, «The Estimation of Cortical Activity for Brain-Computer Interface: Applications in a Domestic Context», *Computational Intelligence and Neuroscience*, **2007** (2007).
- [BFWB07] A. BASHASHATI, M. FATOURECH, R. K. WARD and G. E. BIRCH, «A Survey of Signal Processing Algorithms in Brain-Computer Interfaces Based on Electrical Brain Signals», *Journal of Neural Engineering*, **4**, (2007), n° 2, R35-57.
- [BGM07] M. BESSERVE, L. GARNERO and J. MARTINERIE, «De l'estimation à la classification des activités corticales pour les interfaces cerveau-machine», in *Proc. GRETSI*, 2007.
- [Bir06] N. BIRBAUMER, «Breaking the silence: Brain-computer interfaces (BCI) for communication and motor control», *Psychophysiology*, **43** (2006), n° 6, p.517-532.
- [BKG*00] N. BIRBAUMER, A. KÜBLER, N. GHANAYIM, T. HINTERBERGER, J. PERELMOUTER, J. KAISER, I. IVERSEN, B. KOTCHOUBEY, N. NEUMANN and H. FLOR, «The Thought Translation Device (TTD) for completely paralyzed patients», *IEEE Transactions on Rehabilitation Engineering*, **8** (2000), p. 190-193.
- [BMC*04] B. BLANKERTZ, K. R. MÜLLER, G. CURIO, T. M. VAUGHAN, G. SCHALK, J. R. WOLPAW, A. SCHLÖGL, C. NEUPER, G. PFURTSCHELLER, T. HINTERBERGER and M. SCHRÖDER and N. BIRBAUMER, «The BCI competition 2003: Progress and perspectives in detection and discrimination of EEG single trials», *IEEE Transactions on Biomedical Engineering*, **51** (2004), n° 6, p.1044-1051.
- [Bur98] C. J. C. BURGESS, «A Tutorial on Support Vector Machines for Pattern Recognition», *Knowledge Discovery and Data Mining*, **2** (1998), p.121-167.
- [CLL06] M. CONGEDO, F. LOTTE, A. LÉCUYER, «Classification of Movement Intention by Spatially Filtered Electromagnetic Inverse Solutions», *Physics in Medicine and Biology*, **51** (2006), p. 1971-1989.
- [CM02] D. COMANICIU, P. MEER, «Mean Shift — A Robust Approach toward Feature Space Analysis», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24** (2002), n° 5, p. 603-619.
- [Con6] M. CONGEDO, «Subspace Projection Filters for Real-Time Brain Electromagnetic Imaging», *IEEE Transactions on Biomedical Engineering*, **53** (2006), n° 8, p. 1624-1634.
- [Cor02] A. CORNUÉJOLS, «Une introduction aux SVM», *Bulletin de l'AFIA (Association Française d'Intelligence Artificielle)*, **51** (2002).
- [CR07] F. CABESTAING and A. RAKOTOMAMONJY, «Introduction aux interfaces cerveau-machine (BCI)», in *21ème Colloque sur le Traitement du Signal et des Images, GRETSI'07*, 2007, p.617-620.
- [dRMFn*02] J. DEL R. MILLÁN, M. FRANZÉ, J. MOURIÑO, F. CINCOTTI and F. BABILONI, «Relevant EEG features for the classification of spontaneous motor-related tasks», *Biological Cybernetics*, **86** (2002), n° 2, p. 89-95.
- [GGP*05] R. GRAVE DE PERALTA MENENDEZ, S. GONZALEZ ANDINO, L. PEREZ, P.W. FERREZ and J. DEL R. MILLÁN, «Non-Invasive Estimation of Local Field Potentials for Neuroprosthesis Control», *Cognitive Processing, Special Issue on Motor Planning in Humans and Neuroprosthesis Control*, **6** (2005), p. 59-64.
- [GHP99] C. GUGE, W. HARKAM, C. HERTNAES and G. PFURTSCHELLER, «Prosthetic control by an EEG-based brain-computer interface (BCI)», *Proc. AAATE 5th European Conference for the Advancement of Assistive Technology*, 1999.
- [HBWF96] A. P. HOLMES, R. C. BLAIR, J. D. WATSON and I. FORD, «Nonparametric analysis of statistic images from functional mapping experiments», *Journal of Cerebral Blood Flow and Metabolism*, **16** (1996), p. 7-22.
- [JAPAMBYS06] J.R. JIMENEZ-ALANIZ, M. POHL-ALFARO, V. MEDINA-BANUELOS and O. YANEZ-SUAREZ, «Segmenting Brain MRI using Adaptive Mean Shift», *28th International IEEE EMBS Annual Conference*, 2006, p. 3114-3117.
- [KLH05] B. KAMOUI, Z. LIU and B. HE, «Classification of Motor Imagery Tasks for Brain-Computer Interface Applications by means of Two Equivalent Dipoles Analysis», *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **13** (2005), p. 166-171.
- [LCL*07] F. LOTTE, M. CONGEDO, A. LÉCUYER, F. LAMARCHE and B. ARNALDI, «A Review of classification algorithms for EEG-based Brain-Computer Interfaces», *Journal of Neural Engineering*, **4** (2007), p.R1-R13.
- [LLA07a] F. LOTTE, A. LÉCUYER and B. ARNALDI, «FuRIA : un nouvel algorithme d'extraction de caractéristiques pour les interfaces cerveau-ordinateur utilisant modèles inverses et modèles flous», in *Colloque GRETSI*, 2007, p. 665-668.
- [LLA07b] F. LOTTE, A. LÉCUYER and B. ARNALDI, «Les Interfaces Cerveau-Ordinateur : Utilisation en Robotique et Avancées Récentes», *Journées Nationales de la Recherche en Robotique*, 2007.
- [LLA07] F. LOTTE, A. LÉCUYER, F. LAMARCHE and B. ARNALDI, «Studying the Use of Fuzzy Inference Systems for Motor Imagery Classification», *IEEE Transactions on Neural System and Rehabilitation Engineering*, **15** (2007), n° 2, p. 322-324.
- [LLR*06] F. LOTTE, A. LÉCUYER, Y. RENARD, F. LAMARCHE and B. ARNALDI, «Classification de Données Cérébrales par Système d'Inférence Flou pour l'Utilisation d'Interfaces Cerveau-Ordinateur en Réalité Virtuelle», in *Actes des Premières Journées de l'Association Française de Réalité Virtuelle, AFRV06*, 2006, p. 55-62.
- [LNM06] R. LEHEMBRE, Q. NOIRHOMME and B. MACQ, «Inverse Problem Applied to BCI's : Keeping Track of the EEG's Brain Dynamics Using Kalman Filtering», in *Proc. of the 3rd International Brain Computer Interface Workshop*, 2006, p. 32-33.
- [Lot06] F. LOTTE, «The use of Fuzzy Inference Systems for classification in EEG-based Brain-Computer Interfaces», in *Proc. of the 3rd international Brain-Computer Interface workshop*, 2006, p. 12-13.
- [LSF*07] R. LEEB, R. SCHERER, D. FRIEDMAN, F. LEE, C. KEINRATH, H. BISCHOF, M. SLATER and G. PFURTSCHELLER, «Towards brain-computer interfacing», ch. Combining BCI and Virtual Reality: Scouting Virtual Worlds, MIT Press, G. Dornhege, R. Millan Jdel, T. Hinterberger, D. J. McFarland & K. R. Müller éd., 2007.
- [MAM*06] D. J. MCFARLAND, C. W. ANDERSON, K.-R. MULLER, A. SCHLOGL and D. J. KRUSIENSKI, «BCI meeting 2005-workshop on BCI signal processing: feature extraction and translation», *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **14** (2006), n° 2, p. 135-138.
- [MBF*07] S.G. MASON, A. BASHASHATI, M. FATOURECHI, K.F. NAVARRO and G.E. BIRCH, «A Comprehensive Survey of Brain Interface Technology Designs», *Annals of Biomedical Engineering*, **35** (2007), n° 2, p. 137-169.
- [MML*04] C.M. MICHEL, M.M. MURRAY, G. LANTZ, S. GONZALEZ, L. SPINELLI and R. GRAVE DE PERALTA, «EEG source imaging», *Clin Neurophysiol*, **115** (2004), n° 10, p. 2195-2222.
- [MPP08] G.R. MULLER-PUTZ and G. PFURTSCHELLER, «Control of an Electrical Prosthesis With an SSVEP-Based BCI», *IEEE Transactions on Biomedical Engineering*, **55** (2008), n° 1, p.361-364.
- [MRMG04] J. R. MILLÁN, F. RENKENS, J. MOURINO and W. GERSTNER, «Noninvasive brain-actuated control of a mobile robot by human EEG», *IEEE Transactions on Biomedical Engineering*, **51** (2004), n° 6, p. 1026-1033.

- [NdS05] E. NIEDERMEYER and F. LOPES DA SILVA, *Electroencephalography: basic principles, clinical applications, and related fields*, 5th éd., Lippincott Williams & Wilkins, ISBN 0781751268, 2005.
- [NKM08] Q. NOIRHOMME, R.I. KITNEY and B. MACQ, «Single Trial EEG Source Reconstruction for Brain-Computer Interface», *IEEE Transactions on Biomedical Engineering*, **55** (2008), n° 5, p. 1592-1601.
- [Pfu99] G. PFURTSCHELLER, «EEG event-related desynchronization (ERD) and event-related synchronization (ERS)», *Electroencephalography: Basic Principles, Clinical Applications and Related Fields*, 4th ed. (1999), p. 958,967.
- [PGN06] G. PFURTSCHELLER, B. GRAIMANN and C. NEUPER, «Wiley Encyclopedia of Biomedical Engineering», ch. EEG-Based Brain-Computer Interface System, John Wiley & Sons, Inc., 2006.
- [PM] R.D. PASCUAL-MARQUI, «LORETA/sLORETA website».
- [PM99] R.D. PASCUAL-MARQUI, «Review of methods for solving the EEG inverse problem», *International Journal of Bioelectromagnetism*, **1** (1999), p. 75-86.
- [PM02] R.D. PASCUAL-MARQUI, «Standardized low resolution brain electromagnetic tomography (sLORETA) : technical details», *Methods and Findings in Experimental and Clinical Pharmacology*, **24D** (2002), p. 5-12.
- [PM07] R.D. PASCUAL-MARQUI, «Discrete, 3D distributed, linear imaging methods of electric neuronal activity. Part 1: exact, zero error localization», Tech. report, arXiv:0710.3341v2, 2007.
- [QDH04] L. QIN, L. DING and B. HE, «Motor Imagery Classification by Means of Source Analysis for Brain Computer Interface Applications», *Journal of Neural Engineering*, **1** (2004), n° 3, p. 135-141.
- [RBG+07] B. REBSAMEN, E. BURDET, C. GUAN, H. ZHANG, C. LEONG TEO, Q. ZENG, C. LAUGIER and M. H. ANG Jr., «Controlling a Wheelchair Indoors Using Thought», *IEEE Intelligent Systems*, **22** (2007), n° 2, p. 18-24.
- [SD06] E.W. SELLERS and E. DONCHIN, «A P300-based brain-computer interface: initial tests by ALS patients», *Clin Neurophysiol.*, **117** (2006), n° 3, p. 538-548.
- [VML+07] G. VANACKER, J.R. MILLÁN, E. LEW, P.W. FERREZ, F. GALÁN MOLES, J. PHILIPS, H. VAN BRUSSEL and M. NUTTIN, «Context-Based Filtering for Assisted Brain-Actuated Wheelchair Driving», *Computational Intelligence and Neuroscience*, 2007 (2007), p. Article ID 25130, 12 pages.
- [WBM+02] J. WOLPAW, N. BIRMAUMER, D. MCFARLAND, G. PFURTSCHELLER and T. VAUGHAN, «Brain-computer interfaces for communication and control», *Clinical Neurophysiology* **113** (2002), n°6, p. 767-791.
- [WGW05] M. G. WENTRUP, K. GRAMANN, E. WASCHER and M. BUSS, «EEG source localization for brain-computer-interfaces», in *2nd International IEEE EMBS Conference on Neural Engineering*, 2005, p. 128-131.
- [WLA+06] J. WOLPAW, G. LOEB, B. ALLISON, E. DONCHIN, O. DO NASCIMENTO, W. HEETDERKS, F. NIJBOER, W. SHAIN and J. N. TURNER, «BCI meeting 2005-workshop on signals and recording methods», *IEEE Transaction on Neural Systems and Rehabilitation Engineering* **14** (2006), n°2, p. 138-141.
- [WZL+04] Y. WANG, Z. ZHANG, Y. LI, X. GAO, S. GAO and F. YANG, «BCI competition 2003-data set IV : an algorithm based on CSSD and FDA for classifying single-trial EEG», *IEEE Transactions on Biomedical Engineering* **51** (2004), n°6, p. 1081-1086.
- [Zad96] L. A. ZADEH, «Fuzzy sets», *Fuzzy sets, fuzzy logic, and fuzzy systems : selected papers by Lotfi A. Zadeh* (1996), p. 19-34.



Fabien Lotte

Fabien Lotte est chercheur doctorant à l'Institut de Recherche en Informatique et Systèmes Aléatoires (IRISA) et à l'Institut National des Sciences Appliquées (INSA) à Rennes, France. Ses recherches s'intéressent au traitement et à la classification de signaux électroencéphalographiques pour les interfaces cerveau-ordinateur afin d'utiliser de telles interfaces dans des applications de réalité virtuelle.



Anatole Lécuyer

Anatole Lécuyer est chercheur à l'Institut National de Recherche en Informatique et en Automatique (INRIA), au sein de l'équipe BUNRAKU, à Rennes, France. Ses axes de recherche principaux sont la réalité virtuelle, l'interaction 3D, la perception visuo-haptique, les interfaces haptiques, et les interfaces cerveau-ordinateur. Il est le coordinateur du projet ANR multi-partenaires OpenViBE sur les interfaces cerveau-ordinateur et la réalité virtuelle.



Bruno Arnaldi

Bruno Arnaldi est professeur à l'Institut National des Sciences Appliquées (INSA) de Rennes, France, et effectue ses recherches à l'Institut de Recherche en Informatique et Systèmes Aléatoires (IRISA) où il a dirigé le projet SIAMES (projet commun INRIA/CNRS) de 1992 à 2006. Ses activités de recherche concernent l'informatique graphique, plus précisément l'animation/simulation de systèmes physiques et la réalité virtuelle. Il est membre du Network Management Committee du réseau d'excellence (NoE) INTUITION, coordinateur de la plateforme ANR Part@ge et est actuellement président de l'association française de réalité virtuelle (AFRV).

