

Résumé - L'exploration physiologique du sommeil implique l'étude de plusieurs signaux acquis durant celui-ci. Les enregistrements polysomnographiques permettent d'analyser à la fois l'organisation du sommeil en stades et cycles, mais aussi, et de manière plus fine, la microstructure des signaux recueillis. Dans le cadre de l'EEG en particulier, il s'avère indispensable d'effectuer la reconnaissance des phénomènes transitoires (pointes vertex, fuseaux, complexes K, etc.). L'objet de cet article est de proposer une solution pour la détection de l'un d'entre eux : le complexe K. Le choix de ce problème se justifie par une assez grande difficulté à le résoudre en raison des similitudes marquées de cet événement avec d'autres phénomènes observés en sommeil profond : les bouffées d'ondes Delta. Dans un premier temps, nous exposons une méthodologie pour la synthèse de détecteurs linéaires à partir d'une base d'apprentissage, et nous l'appliquons au cas d'un récepteur opérant conjointement dans les domaines temporel et temps-fréquence. Cependant, il apparaît que les détecteurs résultant de processus d'apprentissage présentent un biais important lorsque la quantité de données disponibles est relativement faible. Pour remédier à cela, nous exposons deux méthodes qui permettent de contrôler la complexité des détecteurs. Notre approche est alors validée sur le problème de la détection du complexe K : il en résulte que la solution obtenue est l'une des plus performantes parmi celles que l'on peut trouver dans la littérature.

Détection / phénomènes transitoires / EEG de sommeil / analyse temps-fréquence / apprentissage

Abstract - Automated recognition of waveforms such as Delta and K-complexes is an important component of sleep EEG analysis. The K-complex is a key feature that contributes to sleep stages assessment. However, its automated detection is still difficult due to the similarities which exist between this event and some other waveforms. This paper is organized as follows. First, we introduce a method of designing linear receivers from training data. This approach is then used to train a detection structure which can be interpreted as joint linear filtering operations in the time and time-frequency domains. However, it appears that detectors designed from training data can have a large bias when the number of training data is relatively small. To remove this difficulty, we introduce two methods of matching the receivers complexity, which can yield a substantial improvement in their performances. Finally, the efficiency of our approach is explored for the detection of K-complexes in sleep EEG.

Detection / transient signal / sleep EEG / time-frequency analysis / learning strategy

I. INTRODUCTION

L'étude des phénomènes transitoires de l'EEG a débuté dans les années 30 [1] avec la première description des fuseaux de sommeil. L'objectif de ces travaux était alors de contribuer à la définition des stades de sommeil. A l'heure actuelle, l'analyse des phénomènes transitoires dans l'EEG a plusieurs objectifs dont l'un des plus importants est probablement de comprendre les mécanismes générateurs et le rôle physiologique de ces signaux. Dans ce cadre, il s'avère indispensable de proposer des outils de traitement automatique de l'information. L'objet de cet article est de proposer une solution pour la détection d'un signal particulier : le complexe K. Le choix de ce problème se justifie par une assez grande difficulté à le résoudre en raison d'une ressemblance importante de cet événement avec des phénomènes non stationnaires observés en sommeil profond : les bouffées d'ondes Delta, notamment dans le cas où le nombre de dérivations enregistrées est faible. Un certain nombre de travaux sur le complexe K et la non stationnarité de l'EEG de fond justifient l'intérêt de l'analyse temps-fréquence, à la fois pour la caractérisation des signaux [2-4], et pour la détection de ces phénomènes [5, 6].

La synthèse d'une structure de détection nécessite la connaissance des propriétés statistiques de l'observation, conditionnellement aux hypothèses. Cependant, de nombreuses applications mettent en jeu des phénomènes complexes pour lesquels l'information *a priori* est faible. Pour pallier cela, nous tirons profit des informations disponibles en élaborant directement le détecteur sur une base d'apprentissage constituée de signaux expertisés, par maximisation d'un critère de contraste. Nous montrons alors qu'en considérant la famille des critères de contraste ne dépendant que des moments conditionnels d'ordre 1 et 2 de la statistique de détection, il est possible de sélectionner automatiquement celui conduisant au détecteur à probabilité d'erreur minimale. Dans le cadre de cet article, nous avons recours à cette méthode pour l'apprentissage d'une structure de détection opérant conjointement dans les temporel et temps-fréquence. En effet, nous montrons que la combinaison de telles informations peut être profitable lorsque l'on ne dispose d'aucune connaissance *a priori* sur la statistique des hypothèses en compétition.

Cependant, il apparaît que la complexité d'un détecteur conditionne fortement ses performances lorsqu'il résulte d'un processus d'apprentissage : les récepteurs possédant un trop grand nombre de degrés de liberté présenteront un faible pouvoir de généralisation, tandis que ceux dotés d'une complexité insuffisante seront incapables d'intégrer toute l'information discriminante véhiculée par les données d'apprentissage. Ce phénomène a été en particulier étudié par Vapnik et Chervonenkis [7], qui ont théoriquement établi des liens entre la probabilité d'erreur de généralisation d'un détecteur, sa complexité et la taille de la base d'apprentissage. En nous appuyant sur ces travaux, nous proposons deux méthodes permettant de contrôler efficacement la complexité des structures de

détection, ce qui a pour effet d'améliorer significativement leurs performances.

Finalement, l'une de ces méthodes est mise en œuvre pour tenter d'apporter une solution au problème de la détection du complexe K dans l'EEG de sommeil.

II. UNE CLASSE GENERALE DE DETECTEURS

1. Eléments d'analyse temps-fréquence

i) *Définition :*

La distribution de Wigner-Ville (WV), qui a toujours joué un rôle majeur dans l'analyse TF des signaux non stationnaires, est définie ainsi [8] :

$$W_x(t, \omega) = \int_{-\infty}^{\infty} R_x(t, \tau) e^{-j\omega\tau} d\tau \quad (1)$$

où $R_x(t, \tau) = x(t + \tau/2) x^*(t - \tau/2)$ est la fonction d'auto-corrélation locale de x .

Cette distribution se distingue par le nombre important de propriétés qu'elle vérifie, parmi lesquelles on peut citer la conservation de l'énergie du signal, le caractère réel et la vérification des contraintes de distributions marginales. Cependant, l'un des reproches que l'on peut formuler à son égard, en plus de sa non positivité, est relatif à la présence de termes interférentiels sur les représentations TF qui nuisent particulièrement à l'analyse des signaux multi-composantes [8]. Pour pallier cette difficulté, on peut alors envisager de filtrer la fonction R_x , ce qui nous amène tout naturellement à définir les distributions de la classe de Cohen¹ :

$$C_x(t, \omega) = \iint_{-\infty}^{\infty} F(\tau, \nu) R_x(t, \tau) e^{-j(\omega - \nu)\tau} d\tau d\nu \quad (2)$$

où F est le noyau caractérisant la représentation.

Cette classe, qui englobe par définition l'ensemble des distributions bilinéaires covariantes par translation dans le plan TF, regroupe des représentations aussi variées que le spectrogramme, les distributions de Rihaczek, Choï-Williams, etc.

ii) *Discrétisation :*

Une discrétisation possible des distributions C_x de la classe de Cohen est fournie par :

$$C_x(n, m) = \sum_{p=1-M}^L \sum_{q=-P}^L F(p, q) R_x(n, m) \quad (3)$$

où F est le noyau à support fini $\mathcal{D}_F = \{(p, q) \in \mathcal{L} \cdot \mid \dots\}$ qui caractérise la distribution. $R_x(n, m) = x(n + m) x^*(n - m)$ représente la fonction d'auto-corrélation locale du signal discret \underline{x} .

¹ Ce rôle central joué par la fonction d'auto-corrélation locale dans la définition de la classe de Cohen est totalement arbitraire. Cette dernière aurait également pu être définie à partir de la distribution de WV [8].

Enfin, pour s'affranchir du problème de repliement spectral dû à la nature bilinéaire de R_x , ce qui théoriquement nécessiterait de sur-échantillonner le signal d'un facteur 2 par rapport à la fréquence de Shannon, on aura systématiquement recours au signal analytique associé à x [8].

2. Détecter dans les domaines temporel et temps-fréquence

i) Détecteur linéaire dans le domaine temporel

Le problème de détection considéré est le suivant. Etant donné un signal discret x connu sur l'intervalle (D) , où $x = [x(1) \dots x(d)]^T$, il s'agit de choisir l'une des hypothèses suivantes :

$$\begin{cases} H_0: & x(k) = b(k) \\ H_1: & x(k) = s(k) + b(k) \end{cases}, k \in \{1, \dots, d\}, \quad (4)$$

où s est le signal à détecter et b un bruit additif. Dans le cadre de l'application faisant l'objet de cet article, s pourra représenter le signal transitoire à détecter et b l'activité EEG de fond. Le choix entre H_0 et H_1 se ramène généralement à la comparaison d'une grandeur réelle $l(x)$, appelée statistique de détection, à un seuil n . A titre d'exemple, lorsque s est déterministe et connu et que b est un bruit blanc gaussien de variance connue s^2 , on peut montrer que la statistique de détection suivante est équivalente au rapport de vraisemblance :

$$\lambda_T(x; s) = \sum_{k=1}^d s(k) x(k). \quad (5)$$

Cette structure de détection, appelée filtre adapté, utilise une forme *linéaire* des échantillons $x(k)$. On notera enfin que (5) peut être obtenu en maximisant le rapport signal sur bruit (RSB) en sortie du filtre dans le cas d'un détecteur linéaire, lorsque le bruit n'est pas gaussien.

ii) Détecteur linéaire dans le plan temps-fréquence

Dans la perspective d'une structure de détection opérant dans le plan TF, le problème de détection (4) peut être reformulé ainsi :

$$\begin{cases} H_0: & C_x(k, f_i; F) = C_n(k, f_i; F) \\ H_1: & C_x(k, f_i; F) = C_{s+n}(k, f_i; F) \end{cases}, k \in \{1, \dots, d\}, i \in \{1, \dots, d\}. \quad (6)$$

Par analogie avec la statistique linéaire (5) définie dans le domaine temporel, on peut considérer la classe suivante des détecteurs linéaires opérant dans le domaine TF :

$$l_{TF}(x, \xi) = \sum_{k=1}^d \sum_{i=1}^d \xi \quad (7)$$

où g est une "référence" TF à déterminer en fonction des hypothèses. Dans le cas où l'on ne dispose d'aucune connaissance *a priori* sur H_0 et H_1 , on pourra par exemple rechercher g de façon à maximiser le critère de contraste de Fisher entre les hypothèses H_0 et H_1 définies en (6) [9]. Dans le cadre de la Section II.3, une approche plus générale est présentée.

Enfin, on notera que la statistique de détection (7) conduit nécessairement à une solution sous

optimale du problème (4) lorsque s est déterministe et connu, et que b est un bruit blanc gaussien : d'après (3), l_{TF} est en effet une fonction purement *quadratique* des échantillons de l'observation.

iii) Intérêt de la détection conjointe temporelle/temps-fréquence

La solution de certains problèmes de détection fait parfois intervenir conjointement les échantillons de l'observation x sous une forme linéaire et quadratique. Considérons à titre d'exemple le test d'hypothèses suivant :

$$\begin{cases} H_0: & x(k) = \rho \cdot b(k) \\ H_1: & x(k) = s(k) + b(k) \end{cases}, k \in \{1, \dots, d\}, \rho \text{ réel} \quad (8)$$

où b est un bruit blanc gaussien, centré, de variance σ^2 connue, et ρ un paramètre réel. Le signal s est supposé déterministe et connu. On peut aisément montrer que la statistique de détection optimale au sens du rapport de vraisemblance s'exprime ainsi [10] :

$$\Lambda(x) = \frac{1}{2\sigma^2} \sum_{k=1}^d x_k + \sum_{k=1}^d x_k^2 \quad (9)$$

Par conséquent, $l(x)$ apparaît comme la somme des fonctions $l_T(x)$ et $l_{TF}(x)$ faisant respectivement intervenir les échantillons $x(k)$ sous une forme linéaire et quadratique. En utilisant la relation de Moyal [8], on montre également que (9) peut s'interpréter comme un détecteur opérant conjointement dans les domaines temporel et temps-fréquence (T-TF) puisque l'on a :

$$\Lambda(x) = \frac{1}{2d\sigma^2} \sum_{k=1}^d \sum_{i=1}^d v_{k,i} x \quad (10)$$

On remarquera que la "référence" TF utilisée est constante sur l'ensemble du plan TF, ce qui fait de $l_{TF}(x)$ un détecteur d'énergie, et que la "référence" temporelle n'est autre que le signal à détecter.

Cet exemple souligne ainsi l'intérêt que présente la classe des détecteurs

$$k=1 \quad i=1 \quad (11)$$

par rapport à des solutions purement temporelles ou TF, lorsque l'on ne dispose d'aucune information *a priori* sur H_0 et H_1 , et qu'il est alors nécessaire d'imposer une structure au récepteur.

3. Conception de détecteurs à partir d'une base d'apprentissage

Nous exposons maintenant une méthode permettant de synthétiser un détecteur linéaire à partir d'une base d'apprentissage, indépendamment des lois statistiques suivies par les signaux. Cette méthode a d'abord été présentée par Fukunaga [11] dans le contexte de la Reconnaissance des Formes, puis réintroduite par Richard et Lengellé [12] pour la génération automatique de détecteurs TF à partir d'ensembles de réalisations des hypothèses H_0 et H_1 .

En utilisant la statistique (11), le problème de détection standard (4) peut être reformulé ainsi :

$$\left\{ \text{sinon } H_0, \right. \quad (12)$$

avec les notations suivantes : n est un seuil à déterminer,

$$\underline{\mathbf{I}} = \left[\cup_{\mathbf{x}}(\mathbf{I}, \mathbf{I}_1, \Gamma) \dots \cup_{\mathbf{x}}(\mathbf{I}, \mathbf{I}_d, \Gamma), \right. \\ \left. \underline{\mathbf{V}} = \left[\mathcal{G}(\mathbf{I}, \mathbf{I}_1) \dots \mathcal{G}(\mathbf{I}, \mathbf{I}_d) \right]. \right.$$

L'élaboration de la structure de détection (12) consiste à déterminer $\underline{\mathbf{V}}$ et n optimaux au sens d'un critère à définir. Soit $f(\eta_0, \eta_1, \sigma_0^2, \sigma_1^2)$ une mesure de la séparabilité des classes H_0 et H_1 , où η_i et σ_i^2 sont les moments conditionnels d'ordre 1 et 2 de la statistique de détection \mathbf{I} :

$$\text{avec } \underline{\mathbf{M}}_{\mathbf{I}_i} = \mathbf{E} \left[\underline{\mathbf{I}} | \Pi_i \right], \quad \underline{\mathbf{L}}_i = \mathbf{E} \left[\dots \right]$$

Ce critère de contraste devant être optimal pour le récepteur recherché, ses dérivés partielles par rapport à $\underline{\mathbf{V}}$ et n doivent être nulles :

$$\left\{ \frac{\partial f}{\partial \underline{\mathbf{V}}} = \frac{\partial f}{\partial \sigma_1^2} \frac{\partial \sigma_1^2}{\partial \underline{\mathbf{V}}} + \frac{\partial f}{\partial \sigma_2^2} \frac{\partial \sigma_2^2}{\partial \underline{\mathbf{V}}} \right. \quad (13)$$

La résolution de ce système conduit à une solution analytique du problème (12) particulièrement intéressante, puisque $\underline{\mathbf{V}}_{\text{opt}}$ vérifie l'équation suivante :

$$\left[\dots \right] \underline{\mathbf{V}} = \left(\frac{\underline{\mathbf{V}} \underline{\mathbf{L}}_1}{\dots} - \frac{\underline{\mathbf{V}} \underline{\mathbf{L}}_1}{\dots} \right) \quad (14)$$

Ainsi, la statistique linéaire $\underline{\mathbf{V}}_{\text{opt}}$ ne dépend de f que par l'intermédiaire d'un paramètre a , $a \in [0, 1]$. Nous retenons dans notre cas la valeur de a et le seuil n qui minimisent la probabilité d'erreur de généralisation E_{gen} définie ainsi :

$$\lambda(\underline{\mathbf{X}}; \underline{\mathbf{V}}, \nu) > 0 \quad (15)$$

où P_i désigne la probabilité *a priori* de H_i et $p_i(\underline{\mathbf{X}})$ la densité de probabilité conditionnelle de $\underline{\mathbf{X}}$. En pratique, lorsque (15) ne peut être aisément calculé, les approximations de E_{gen} fournies par les méthodes de resubstitution, de bootstrap ou encore de validation croisée peuvent être utilisées [11]. Nous adopterons cette dernière approche en estimant la probabilité d'erreur de généralisation par comptage des échantillons mal classés sur un ensemble de données indépendantes de l'ensemble d'apprentissage, dit ensemble test.

En conclusion, la méthodologie qui vient d'être présentée conduit au détecteur linéaire optimal au sens du meilleur critère de contraste $f(\eta_0, \eta_1, \sigma_0^2, \sigma_1^2)$, sans avoir à l'exhiber. En particulier, le

récepteur obtenu offre des performances supérieures (ou égales) à celles de détecteurs établis par maximisation du RSB ou du critère de Fisher, ces derniers correspondant à $a = 1$ et $a = P_0$ [12].

III. CONTROLE DE LA COMPLEXITE

1. "Structural Risk Minimization" [13]

Obtenir une probabilité d'erreur de généralisation E_{gen} satisfaisante avec un détecteur déterminé au moyen d'une base d'apprentissage nécessite d'adapter la complexité de celui-ci au nombre d'exemples dont on dispose. Ainsi, les récepteurs dotés d'un nombre trop important de degrés de liberté auront un faible pouvoir de généralisation. Dans le cas contraire, ces derniers seront incapables d'intégrer toute l'information discriminante présente dans la base d'apprentissage. Entre ces extrêmes, il existe une complexité optimale, dépendant notamment de la quantité de données disponibles, pour laquelle E_{gen} est minimale. La recherche de cet extremum, qui a été en particulier étudiée par Vapnik et Chervonenkis [7], est maintenant brièvement exposée.

Soit c un ensemble de détecteurs, et soit VC_c sa VC-dimension. Ce paramètre caractérise la complexité des récepteurs de c puisqu'il est défini par le plus grand nombre d'échantillons dont ils puissent réaliser toutes les dichotomies. Dans le cas des détecteurs linéaires, il est aisé de voir que ce nombre maximal d'échantillons est de 2 dans R , de 3 dans R^2 , etc. Ce résultat peut se généraliser à R^N pour lequel on a $VC_c = N + 1$. Par conséquent, $VC_c = d^2 + 1$ lorsque c représente la classe des détecteurs TF définie en (7). Cependant, la détermination analytique de VC_c reste difficile dans le cas général. L'intérêt de connaître VC_c réside dans le fait que ce paramètre permet de déterminer un intervalle de confiance de E_{gen} pour tout détecteur de c , connaissant sa probabilité d'erreur E_{app} sur la base d'apprentissage, et la taille N de celle-ci. Avec une probabilité de $1 - h$, et pour toute structure de détection de c , on démontre en effet que l'inégalité suivante est satisfaite [7] :

$$(16)$$

Le principe de *Structural Risk Minimization* proposé par Vapnik consiste à rechercher le meilleur compromis possible entre les termes antagonistes E_{app} et $D(.)$: la réduction de VC_c va en effet entraîner la décroissance de $D(.)$ et la croissance de E_{app} , et inversement. Pour cela, on construit au sein de la classe c une séquence de sous-ensembles imbriqués χ_r de VC-dimension VC_{χ_r} :

$$(17)$$

Afin de déterminer le détecteur de VC-dimension optimale sur c , Vapnik propose l'approximation suivante [13] :

$$(\underline{v}, \underline{v})_{opt} = \text{argmin}_{\{(\underline{v}, \underline{v}) \in \chi\}} E_{gen} \quad (18)$$

où les détecteurs $(\underline{V}, n)_{r, \text{opt}}$ de VC-dimension croissante avec r sont déterminés sur chaque ensemble χ_r au moyen d'un algorithme d'apprentissage. Enfin, lorsque trop peu de données sont disponibles pour estimer E_{gen} par validation croisée, Vapnik suggère d'utiliser la borne définie en (16).

2. Contrôle de la complexité des détecteurs TF

Nous présentons maintenant 2 stratégies permettant de faire varier la VC-dimension des structures opérant dans le plan TF, ce qui nous conduit naturellement à construire la séquence (17). Ces approches sont évidemment applicables aux détecteurs opérant conjointement dans les domaines temporel et TF.

i) Partitionnement du plan TF [14]

La première approche que nous proposons consiste à imposer à $g(k, f_i)$ définie en (7) d'être constante par régions, ces régions $\{A_1, \dots, A_r\}$ constituant une partition du plan TF. Par cette heuristique, nous espérons exploiter la structure particulière des représentations TF (corrélation entre positions TF adjacentes, inexistence d'impulsion dans le plan TF, etc.). Les ensembles χ_r définis en (19), de VC-dimension égale à $r + 1$, définissent la structure de type (17) préconisée par Vapnik. Dans ce contexte, on peut donc envisager d'appliquer le principe de *Structural Risk Minimization*.

$$\chi_r = \{ (\wedge_{\text{TF}}(\cdot, \underline{g}), \underline{v})_r \mid \underline{g}(k, f_i) = \mu \} \quad (19)$$

La recherche du détecteur optimal $(l_{\text{TF}}(\cdot; \underline{g}), n)_{r, \text{opt}}$ de c_r intervenant en (18) consiste à optimiser la partition $\{A_1, \dots, A_r\}$. Dans ce cas, le critère à minimiser est une estimation de la probabilité d'erreur E_{gen} du détecteur $(l_{\text{TF}}(\cdot; \underline{g}), n)_r$ associé à $\{A_1, \dots, A_r\}$ et obtenu par l'algorithme d'apprentissage (14). Ces opérations sont résumées en Figure 1. L'algorithme utilisé pour l'optimisation de la partition est décrit en détails dans [15].

Finalement, le détecteur $(l_{\text{TF}}(\cdot; \underline{g}), n)_{r, \text{opt}}$, $0 \leq r \leq d^2$, retenu est celui qui présente la probabilité d'erreur E_{gen} estimée la plus faible.

ii) Principe d'Optimal Brain Damage [16]

Une autre façon de contrôler la complexité du détecteur $(l_{\text{TF}}(\cdot; \underline{g}), n)$ défini en (7) est de décimer un certain nombre de paramètres $g(k, f_i)$. On rappelle que ces derniers constituent les composantes du vecteur \underline{V} solution du système linéaire (14). Définissons les meilleurs candidats pour cette opération comme étant ceux qui minimisent l'accroissement de l'erreur quadratique SE_a suivante :

$$\Delta SE_{\alpha} = \|\underline{Z}_{\alpha} \cdot \underline{V} - \underline{IV}\|, \quad \text{OU} \quad \underline{Z}_{\alpha} = \underline{U}\underline{Z} \quad (20)$$

On peut montrer que dans la base des vecteurs propres normés de S_a , l'accroissement ΔSE_{α}^i de SE_a dû à la décimation de la $i^{\text{ème}}$ composante de \underline{V} vérifie :

$$\Delta SE_{\alpha}^i = [\mu_{\alpha}^i]^2 [(\Phi_{\alpha}^i)^T \underline{V}]^2, \quad (21)$$

où μ_{α}^i (resp. Φ_{α}^i) désigne la $i^{\text{ème}}$ valeur propre (resp. le $i^{\text{ème}}$ vecteur propre normé) de S_a .

En conclusion, les composantes de \underline{V} qui correspondent aux accroissements DSE_a les plus importants sont considérées comme prioritaires durant le processus de sélection des variables. Cette méthode est basée sur le principe d'*Optimal Brain Damage*, une procédure communément utilisée dans le cadre de la détermination de la structure des réseaux de neurones artificiels. Dans [16], une autre méthode est également proposée, basée cette fois sur le principe de *Weight Decay* [13].

IV. DETECTION DU COMPLEXE K DANS L'EEG DE SOMMEIL

1. Présentation du problème

Le complexe K constitue avec le spindle l'un des principaux marqueurs du début de sommeil puisqu'il apparaît dès le stade 2. Il est généralement d'aspect biphasique, comprenant une onde négative rapide suivie d'une onde positive de grande amplitude (Figure 2). Il dure environ 1 seconde, présente un support fréquentiel compris entre 1 et 4 Hz, et se distingue par son amplitude de l'activité EEG de fond du stade 2. Cependant, en raison d'une similitude marquée avec d'autres phénomènes non stationnaires observés en sommeil profond tels que les bouffées d'ondes Delta (Figure 3), il reste très difficile à isoler en stades 3 et 4. Ceci justifie l'intérêt de méthodes statistiques pour tenter d'apporter une réponse satisfaisante à ce problème de détection.

2. Description des données

L'acquisition des signaux EEG utilisés dans le cadre de cette étude a été effectuée sur la dérivation Cz à une fréquence d'échantillonnage de 128 Hz. Une base de complexes K a d'abord été constituée par un groupe de 5 experts. Ces électrophysiologistes confirmés, qui ont effectué visuellement la cotation des signaux, ont travaillé individuellement sur les enregistrements de 3 nuits de sommeil. Du fait d'une certaine variabilité des résultats obtenus par ces cotateurs, nous n'avons retenu pour cette étude que les événements sélectionnés par la majorité d'entre eux afin de constituer un ensemble de référence pour le complexe K. Ensuite, une base de phénomènes présentant des similitudes marquées avec ce dernier (bouffées d'onde Delta, etc.) a été constituée. Tous ces événements transitoires ont alors été synchronisés avant d'être segmentés sur des durées de 2 secondes. Finalement, les données ont été décimées d'un facteur 4, une fréquence d'échantillonnage de 32 Hz étant compatible avec les supports fréquentiels du complexe K, des ondes rapides qui peuvent se superposer à lui, et des bouffées d'ondes Delta.

En conclusion, une base de données composée de 609 complexes K et de 1196 ondes présentant des similitudes avec ces derniers, que l'on désignera désormais de façon impropre par ondes Delta, a été constituée pour la présente étude.

3. Résultats et discussion

Afin d'élaborer le détecteur de complexes K , nous avons effectué l'apprentissage de la structure de détection (I_{T-TF}, n) opérant conjointement dans les domaines temporel et TF au moyen de la méthode présentée au paragraphe (II.3). Afin d'optimiser la VC-dimension de ce détecteur, nous avons arbitrairement adopté l'heuristique conduisant au partitionnement du plan TF. Cependant, les approches basées sur les principe d'*Optimal Brain Damage* et de *Weight Decay* nous ayant donné de bons résultats sur des données simulées [16], nous envisageons de les tester ultérieurement sur ce problème.

Les partitions du plan TF $\{A_1, \dots, A_r\}$ définies en (19) ont été optimisées pour r variant de 10 à 90 sur un ensemble de 300 complexes K et de 600 ondes Delta, en suivant le mode opératoire décrit par la Figure 1. La Figure 4 montre que la partition obtenue pour $r=30$ conduit à un minimum pour l'estimation de E_{gen} . Sur la base de cette partition optimale, nous avons alors utilisé les 309 complexes K et les 596 ondes delta restants pour enchaîner 250 phases d'apprentissage et de test : 220 complexes K et 500 ondes delta ont été tirés aléatoirement à chaque fois pour la synthèse du détecteur, les autres échantillons servant à l'estimation de sa courbe COR.

La Figure 5 illustre la contribution des parties temporelle (I_T) et temps-fréquence (I_{TF}) à la valeur finale de la statistique I_{T-TF} , par l'intermédiaire de la projection des données sur le plan défini par I_T et I_{TF} . Cette représentation démontre explicitement l'intérêt de combiner les informations provenant des domaines temporel et TF pour résoudre notre problème. Dans un premier temps, l'efficacité de notre approche a été évaluée en comparant les résultats obtenus à ceux d'un détecteur opérant dans le domaine temporel puis à ceux d'un détecteur TF à complexité optimisée. Pour ce dernier, le paramètre r a été fixé à 30 selon les résultats présentés en Figure 4. Les courbes COR moyennes résultant des 250 apprentissages décrits ci-dessus, représentées en Figure 6, indiquent les pertes de performances dues à l'usage de statistiques de détection purement linéaires ou quadratiques pour résoudre ce problème. Dans un second temps, les performances de notre détecteur ont été comparées à celles de solutions déjà existantes. Comme cela est montré dans le Tableau 1, les meilleurs résultats ont été rapportés dans [17] : un réseau de neurones a permis d'obtenir 90% de bonnes détections pour 8.1% de fausses alarmes, à comparer aux 4% de fausses alarmes que nous avons obtenu pour le même niveau de sensibilité. Tout en restant prudent sur la comparaison de ces résultats qui dépendent fortement de la constitution de la base de données, nous pouvons affirmer que la solution à laquelle nous avons abouti est l'une des plus performantes.

V. CONCLUSION

Dans cet article, nous avons exposé une méthode permettant de synthétiser des récepteurs linéaires à partir d'une base d'apprentissage, et nous l'avons mise en oeuvre sur une statistique de

détection utilisant conjointement les informations provenant des domaines temporel et TF. Cependant, il est bien connu en Reconnaissance des Formes que les classifieurs obtenus par apprentissage peuvent présenter un biais important, particulièrement lorsque la quantité de données disponibles est relativement faible. Pour remédier à ce problème, nous avons exposé 2 méthodes qui reposent sur le principe de *Structural Risk Minimization*, et qui permettent de contrôler la complexité des détecteurs de façon à améliorer leurs performances.

Notre approche a alors été validée sur le problème de la détection automatique du complexe K dans l'EEG de sommeil. Il résulte de la présente étude que le détecteur obtenu est l'un des plus performants par rapport aux solutions que l'on peut trouver dans la littérature. De plus, il semble que les performances de celui-ci puissent encore être améliorées en combinant les informations provenant de plusieurs dérivations. Enfin, on notera que la méthodologie présentée dans le cadre de cet article est immédiatement transposable à tout autre problème de détection de phénomènes transitoires (pointes vertex dans l'EEG, ...)

VI. BIBLIOGRAPHIE

- [1] A. L. LOOMIS, E. N. HARVEY, and G. A. HOBART. Potentials Rhythms of the Cerebral Cortex during Sleep. *Science* 1935, 81, 597-598.
- [2] J. C. BRAUN, D. F. HANLEY, and N. V. THAKOR. Detection of Neurological Injury Using Time-Frequency Analysis of the Somatosensory Evoked Potentials. *Electroencephalography and Clinical Neurophysiology* 1996, 100, 310-318.
- [3] H. I. CHOI, W. J. WILLIAMS, and H. ZAVERY. Analysis of Event Related Potentials: Time-Frequency Energy Distributions. *Biomedical Sciences Instrumentation* 1987, 23, 251-258.
- [4] N. SCHALTENBRAND, R. LENGELLE, and J.-P. MACHER. Neural Network Model: Application to Automatic Analysis of Human Sleep. *Computers and Biomedical Research* 1993, 26, 157-171.
- [5] Z. TANG and N. ISHII. Detection of the K-Complex using a New Method of Recognizing Waveform Based on the Discret Wavelet Transform. *IEICE Trans. Inf. & Syst.* 1995, 78, 77-85.
- [6] A. CIMETIERE. "Méthodes Temporelles et Temps-Fréquence pour la Reconnaissance Automatique des Complexes K de l'EEG de Sommeil," . Compiègne : UTT, 1997.
- [7] V. N. VAPNIK and A. Y. CHERVONENKIS. On the Uniform Convergence of Relative Frequencies of Events to their Probabilities. *Theory of Probabilities and its Applications* 1971, 16, 264-280.
- [8] P. FLANDRIN. Temps-Fréquence. Paris, F : HERMES, 1993.

- [9] D. L. JONES and A. M. SAYEED. Blind Quadratic and Time-Frequency based Detectors from Training Data. IEEE International Conference on Acoustics, Speech and Signal Processing 1995, 1033-1036.
- [10] H. V. POOR. An Introduction to Signal Detection in Noise. New York, NY: Springer-Verlag, 1994.
- [11] K. FUKUNAGA. Introduction to Statistical Pattern Recognition. London: Academic Press, 1990.
- [12] C. RICHARD and R. LENGELLE. Une Nouvelle Approche pour la Détection Linéaire Optimale dans le Plan Temps-Fréquence. Seizième Colloque GRETSI 1997, 659-662.
- [13] V. N. VAPNIK. Estimation of Dependences Based on Empirical Data. New York, NY: Springer Verlag, 1982.
- [14] C. RICHARD and R. LENGELLE. Structural Risk Minimization for Reduced-Bias Time-Frequency-Based Detectors Design. IEEE International Conference on Acoustics, Speech, and Signal Processing 1998, 2397-2400.
- [15] M. SCHOENAUER, L. KALLEL, and F. JOUVE. Mechanical Inclusions Identification by Evolutionary Computation. Revue Européenne des Eléments Finis 1996, 5, 619-648.
- [16] C. RICHARD and R. LENGELLE. Two Algorithms for Designing Optimal Reduced-bias Data-Driven Time-Frequency Detectors. Accepté à IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis 1998.
- [17] I. N. BANKMAN, V. G. SIGILLITO, R. A. WISE, and P. L. SMITH. Feature-Based Detection of K-Complex Wave in the Human Electroencephalogram using Neural Networks. IEEE Transactions on Biomedical Engineering 1992, 39, 1605-1610.
- [18] A. C. DA ROSA, B. KEMP, and T. PAIVA. A Model-based Detector of Vertex Sharp Waves and K-Complexes in Sleep Electroencephalogram. Electroencephalography and Clinical Neurophysiology 1991, 78, 71-79.
- [19] J. DESTINE, B. BECKERS, M. FOMBELLIDA, R. POIRRIER, D. DIVE, and G. FRANCK. Utilisation des Réseaux de Neurones Artificiels pour la Reconnaissance de Grapho-Elements Phasiques dans le Cadre de l'Analyse du Sommeil. Symposium International 25 ans d'Analyse Automatique du Sommeil 1994.

LEGENDES DES TABLEAUX ET FIGURES

Tableau 1 Performances de quelques détecteurs de complexes K existants.

Figure 1 Principe du partitionnement du plan TF.

Figure 2 Exemples de complexes K.

Figure 3 Exemples de bouffées d'ondes Delta.

Figure 4 Erreur de généralisation estimée en fonction de la taille r de la partition du plan TF. Détecteur TF (continu) ; Détecteur T-TF (discontinu).

Figure 5 Projection des données sur le plan défini par l_T et l_{TF} . Complexes K (-) ; Bouffées d'onde Delta (i).

Figure 6 Détecteurs de complexes K. Courbes COR obtenues avec le détecteur T-TF (continu), le détecteur TF (discontinu), et le détecteur T (pointillé). Performances des détecteurs de Da Rosa (u), Destiné (²) et Bankman (l).

TABLEAUX ET FIGURES

Tableau 1

	approche	$P_{\text{bonnes détections}} (\%)$	$P_{\text{fausses alarmes}} (\%)$
Da Rosa <i>et al.</i> [18]	modélisation	89	49
Destiné <i>et al.</i> [19]	réseau de neurones	64 (resp. 76)	4 (resp. 12)
Bankman <i>et al.</i> [17]	réseau de neurones	90 (resp. 95)	8.1 (resp. 14.1)
Cette étude	détection T-TF	90 (resp. 95)	4 (resp. 7.5)

Figure 1

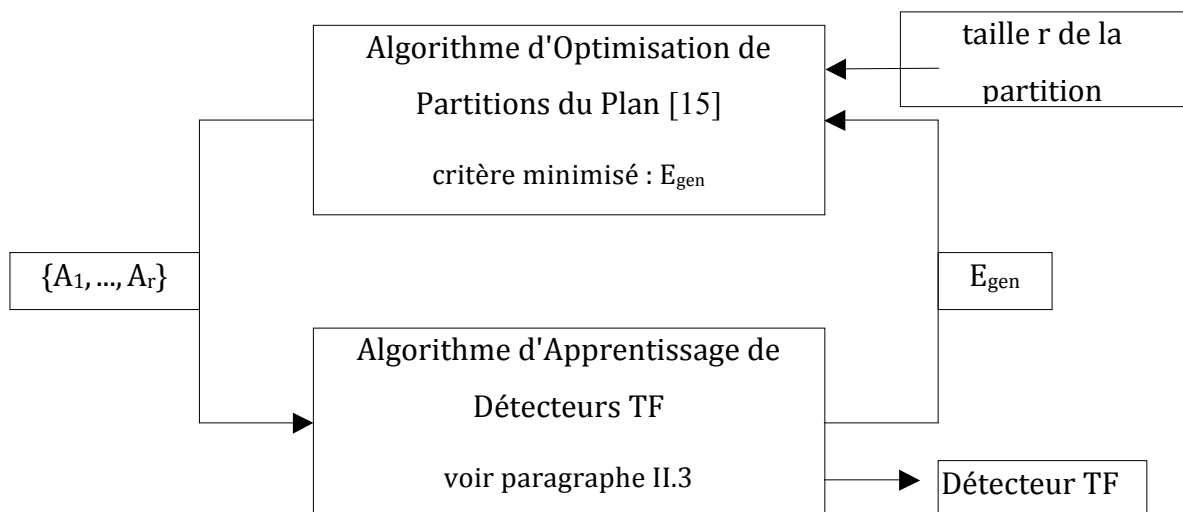


Figure 2

Title: figure1.eps
 Creator: MATLAB, The Mathworks, Inc.
 CreationDate: 03/29/98 16:40:17

Figure 3

Title: figure2.eps
Creator: MATLAB, The Mathworks, Inc.
CreationDate: 03/29/98 16:40:33

Figure 4

Title: figure3.eps
Creator: MATLAB, The Mathworks, Inc.
CreationDate: 03/29/98 16:40:27

Figure 5

Title: figure4.eps
Creator: MATLAB, The Mathworks, Inc.
CreationDate: 03/29/98 16:40:23

Figure 6

Title: fig5.eps
Creator: MATLAB, The Mathworks, Inc.
CreationDate: 04/01/98 09:55:59