
Classification en temps-fréquence à partir de données expertisées

Application à la détection des complexes-K

Arnaud Coatanhay* — Cédric Richard** — Luc Staner***

* Laboratoire E³I² – ENSIETA
2, rue François Verny – 29806 Brest cedex 9
coatanar@ensieta.fr

** Laboratoire LM2S – Université Technologique de Troyes
12, rue Marie Curie - BP 2060 - F - 10010 Troyes cedex
cedric.richard@utt.fr

*** Centre de recherche FORENAP
Centre Hospitalier de Rouffach - BP 29 - F - 68250 Rouffach
luc.staner@forenap.asso.fr

RÉSUMÉ. Cet article présente une nouvelle méthodologie pour concevoir un détecteur automatique de complexes-K. Notre approche se fonde sur la détermination d'un classifieur optimal dans le domaine temps-fréquence (transformée de Wigner-Ville) à partir de données d'apprentissage expertisées. Dans un premier temps, nous décrivons comment gérer les avis contradictoires des experts, et comment constituer une base de données fiable. Puis, en justifiant le choix de la représentation temps-fréquence, nous concevons le détecteur optimal par rapport aux données, tout en tenant compte du problème de la complexité. Enfin, par comparaison avec les recherches précédentes, nous quantifions l'amélioration substantielle qu'apporte notre approche.

ABSTRACT. In his paper, we propose a new methodology to design an automated K-complex detector. Our approach is based on optimisation of a linear classifier in the time-frequency domain (Wigner transform) from labelled training data. First, we describe how to manage the contradictions between human experts and how to build a reliable data base. Then, making the choice of Time-Frequency representation, the design of data-driven optimal detector is determined, with take into account the complexity control problem. Finally, compared with previous studies, we estimate the substantial improvement yielded by our original approach.

MOTS-CLÉS : classification, temps-fréquence, réduction de complexité, fusion d'expertises.

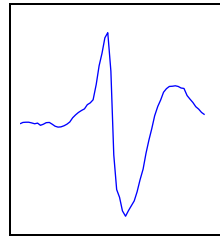
KEYWORDS: classification, time-frequency, complexity regulation, information fusion.

1. Introduction

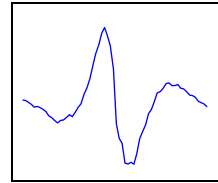
Dans le cadre des activités biomédicales, il arrive très fréquemment que l'observation d'un phénomène physiologique particulier dépende, dans une large mesure, de l'individu considéré. Des éléments comme l'âge, le sexe, une pathologie éventuelle ou la prise d'un traitement pharmacologique peuvent avoir une influence considérable sur la mesure du phénomène à identifier. On constate également souvent que les observations biomédicales peuvent varier aléatoirement chez un même individu d'une manière non négligeable. Dans ces conditions, il devient particulièrement délicat de caractériser un phénomène physiologique par critères paramétriques stricts. D'une part le choix des paramètres objectifs à prendre en compte s'effectue avec un certain arbitraire. D'autre part, il est difficile de déterminer, pour chacun des paramètres, des seuils numériques qui soient à la fois robustes, au regard de la diversité d'une population, et pertinents, pour éviter les faux diagnostics. Pour toutes ces raisons, la lecture et l'interprétation des données biomédicales sont réalisées, le plus souvent, par des experts humains (médecins, électrophysiologistes ou autres) qui peuvent intégrer à la simple mesure leur apprentissage et leur expérience. Bien évidemment, si le phénomène que l'on cherche à observer est complexe, les avis de ces experts deviennent parfois contradictoires. Par ailleurs, et en dépit de toutes ces difficultés, l'automatisation des tâches de décision et d'identification des phénomènes physiologiques s'avère de plus en plus nécessaire, pour épargner aux équipes cliniques certaines tâches répétitives monopolisant des ressources trop importantes, et surtout pour rationaliser au mieux les identifications.

Cette problématique est particulièrement sensible en électro-encéphalographie (EEG) où des électrodes collées à la surface du crâne enregistrent l'activité électrique globale du cerveau. Les signaux ainsi enregistrés correspondent à une mesure, pouvant se dérouler sur une période très longue (parfois plusieurs heures), de processus électriques extrêmement complexes. Au cours de l'enregistrement, certains phénomènes physiologiques s'identifient difficilement, car ils apparaissent très fluctuants dans leur forme et sont soumis à de multiples perturbations (artéfacts électriques, mouvement du sujet etc...). Dans cet article, nous nous proposons de traiter de l'identification automatique d'un phénomène électrique appelé complexe-K qui relève d'une grande importance dans l'étude du sommeil, (cf. [LOO 38] pour la première description du phénomène, ou cf. [BAS 92] pour une analyse plus approfondie). Le complexe-K est une oscillation particulière, voir figure (1(a)), qui s'interprète le plus souvent comme une réponse cognitive à un stimulus externe ou interne. La description de ce phénomène a fait l'objet d'une littérature abondante. Toutefois, il faut savoir que la forme de ce signal subit des variations au cours du temps et d'un individu à l'autre. De plus, ce signal se superpose à l'activité électrique de fond du cerveau ou à des perturbations éventuelles. Enfin, la principale difficulté de la détection des complexes-K réside dans le fait que cette oscillation se confond aisément avec un autre phénomène physiologique appelé onde delta, voir figure (1(b)). Ces ondes delta ont également une forme variable au cours du temps, mais surtout, elles se produisent beaucoup plus fréquem-

ment que les complexes-K (rapport de 10 à 20) lors d'un enregistrement de sommeil normal.



(a) Complexe-K.



(b) Onde delta.

Figure 1. Exemple des deux phénomènes physiologiques.

La classification entre onde delta et complexes-K est réputé être un problème de grande difficulté et entraînant de sérieuses contradictions entre les experts, cf. [BRE 70]. Dans cet article, nous proposons une démarche méthodologique permettant de construire un classifieur automatique de complexes-K à partir de cotations de plusieurs experts humains, et en supposant le minimum *d'a priori* sur la structure des signaux à considérer (approche non-paramétrique). Dans une première étape, nous serons amenés à gérer les contradictions entre les experts et à en faire une synthèse. Nous verrons ensuite comment constituer les bases de données qui serviront à l'apprentissage, et nous justifierons l'espace de représentation que nous avons choisi (le domaine temps-fréquence). Enfin, nous expliciterons la détermination du classifieur optimal, en tenant compte du problème de la complexité.

2. Méthologie

Afin de constituer un ensemble d'apprentissage significatif, nous avons fait coter visuellement, par 6 experts, les complexes-K enregistrés lors des nuits de sommeil de trois jeunes volontaires sains mâles. Pour faciliter la tâche des experts, l'enregistrement EEG a été réalisé par un réseau de 12 électrodes, les différentes voies étant affichées en continu sur un écran. La détection automatique, quant à elle, doit s'effectuer, en accord avec les règles cliniques habituelles, sur une seule électrode posée au sommet du crâne (position centrale). De plus, avant la détection visuelle, des règles communes de cotation ont été préétablies.

2.1. Fusion des expertises

Malgré toutes ces précautions, nous avons pu constater une grande variabilité dans les avis des experts. Si un expert cote en moyenne 804 complexes-K pour les trois

nuits, seulement 292 complexes-K ont été cotés par au moins quatre experts. Pour obtenir une fusion de l'avis de ces experts, des méthodes théoriques générales ont été développées, voir très récemment [VIA 01]. Toutefois, dans le cas présent, les erreurs faites par un classifieur automatique serait surtout préjudiciables s'il ne détectait pas un complexe-K coté par presque tous les experts, ou au contraire s'il retenait des phénomènes cotés par aucun expert. D'une façon pragmatique, nous avons déterminé trois catégories : la première est constituée des phénomènes cotés par au moins trois experts (603 éléments). Cette catégorie regroupent des éléments que l'on peut qualifier presque sûrement de complexes-K. La deuxième catégorie est l'ensemble des éléments cotés par au moins un expert et ne relevant pas de la catégorie précédente. La deuxième catégorie est un ensemble ambigu de phénomènes physiologiques. Ces éléments ne seront pas pris en compte pour évaluer la performance du classifieur, et la catégorie 2 n'interviendra plus dans la suite de cet article. Enfin, la dernière catégorie correspond aux phénomènes cotés par aucun des experts. Ces derniers éléments sont presque sûrement des non complexes-K (principalement des ondes delta).

2.2. Présélection

L'ensemble des complexes-K cotés par tous les experts nous permet de calculer un complexe-K moyen sur une fenêtre temporelle de 2 secondes. Par un filtrage linéaire classique adapté à ce complexe-K moyen, et en choisissant le seuil minimal permettant de retenir l'ensemble des éléments de la première catégorie, on présélectionnera en plus des complexes-K (603 éléments de la première catégorie) 14751 non complexes-K (3ème catégorie) sur les 3 nuits enregistrées. Ayant réalisé une fusion des expertises, notre objectif est donc maintenant de construire un classifieur automatique optimale nous permettant de séparer ces deux classes (complexes-K et non complexes-K).

Les éléments de ces deux classes représentent des parties de l'enregistrement électroencéphalographique sur une période de 2 secondes. Comme l'enregistrement est un signal échantillonné à $32Hz$, les éléments à traiter sont sous la forme de vecteurs de dimension $32 \times 2 = 64$.

2.3. Représentation temps-fréquence

Les études précédentes dans le domaine de la détection automatique de complexes-K, se sont fondées sur l'indentification paramétrique des formes d'ondes ou sur des traitements linéaires des signaux temporels [BAN 92, DA 93, JAN 90, JOB 92, DeB94]. Les performances des classifieurs ainsi obtenus étaient soit très médiocres, soit estimées à partir de données sélectionnées et incompatibles avec la réalité clinique des enregistrements sommeil. Par ailleurs, une préétude [JAC 96] a montré l'intérêt des représentations temps-fréquence pour améliorer la discrimination entre les complexes-K et les ondes delta. Nous avons donc cherché à utiliser la représentation temps-fréquence des éléments des classes plutôt que directement le signal temporel (64 valeurs) pour discriminer les complexes-K des non complexes-K. Concrètement, à un

signal temporel échantillonné $s(k)$, on associe sa transformée de Wigner-Ville discrète $W_s(k, f_j)$, donnée par :

$$W_s(k, f_j) = 2 \sum_{\tau=-N}^N R_s(k, \tau) \cdot e^{-i4\pi f_j \tau}, \quad [1]$$

avec $R_s(k, \tau) = s(k + \tau) s^*(k - \tau)$. La transformée de Wigner-Ville discrète associe donc à un vecteur de 64 valeurs une image de 64×64 soit 4096 points. La classification se fera sur ces éléments de 4096 valeurs. A titre d'exemple, on comprend intuitivement qu'il est plus difficile de distinguer les deux signaux temporels représentés sur la figure (1) que leur transformées de Wigner-Ville données par la figure (2).

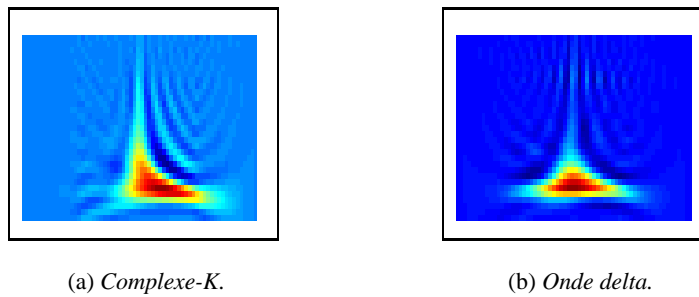


Figure 2. Représentation temps-fréquence des deux signaux de la figure (1).

Il aurait pu être envisagé de choisir une transformée temps-fréquence plus sophistiquée (pseudo Wigner-Ville lissée, transformée de la classe de Cohen, etc...) pour améliorer la lisibilité dans l'espace temps-fréquence. Il faut toutefois rappeler que l'objectif premier de cette représentation temps-fréquence n'est pas ici une extraction paramétrique d'information, mais est l'amélioration de la "séparabilité" statistique de deux classes données. L'avantage de la transformée de Wigner-Ville simple est qu'elle n'élimine pas d'information *a priori* dans le domaine temps-fréquence, contrairement à ce que pourrait entraîner un lissage.

2.4. Classifieur linéaire optimal en temps-fréquence

Pour discriminer les transformées de Wigner-Ville des deux classes, nous avons cherché à construire un filtre linéaire optimal dans le domaine temps-fréquence. On peut faire la remarque qu'un classifieur linéaire dans le domaine temps-fréquence est équivalent à un classifieur quadratique dans le domaine temporel, le temps et la fréquence étant supposés continus, voir [FLA 88]. Toutefois, le fait de considérer un filtre linéaire dans le domaine temps-fréquence nous permet de bénéficier d'une meilleure interprétabilité des résultats.

Pour l'élaboration d'une structure de détection linéaire, une approche classique consiste à maximiser un critère de contraste donné. Se pose dans ce cas, le choix *a priori* d'un critère de contraste. Il est toujours possible de choisir un critère communément admis, par exemple le critère de Fisher, voir [JON 95]. Cependant, l'optimisation par rapport à ce critère ne garantit en rien une séparabilité réelle des populations statistiques, voir [DEV 96]. La pertinence d'un critère de contraste dépend étroitement de la distribution des classes à séparer. Enfin, un problème plus pratique se pose également : l'optimisation des nombreux paramètres de la statistique de détection (4096) demande une puissance de calcul importante. Pour résoudre ces problèmes, Fukunaga [FUK 90] a introduit, dans le contexte de la reconnaissance de forme, une méthode permettant d'optimiser un détecteur linéaire pour une classe étendue de critère de contraste.

On note H_0 et H_1 les hypothèses selon laquelle la transformée de Wigner-Ville correspond respectivement à un complexe-K et à un non complexe-K. La classification linéaire des transformées de Wigner-Ville (images de 64×64 points) se décrit alors synthétiquement :

$$\left\{ \begin{array}{l} \text{Si } \lambda_{TF} = \sum_{k=1}^{64} \sum_{j=1}^{64} g(k, f_j) W_s(k, f_j) \geq \nu \quad \text{alors } H_0, \\ \text{sinon } H_1, \end{array} \right. \quad [2]$$

que l'on peut noter :

$$\lambda_{TF}(\vec{Y}) = \vec{V}^T \vec{Y}, \quad [3]$$

avec :

$$\vec{Y} = [W_s(1, f_1), \dots, W_s(1, f_{64}), \dots, W_s(64, f_1), \dots, W_s(64, f_{64})]^T, \quad [4]$$

$$\vec{V} = [g(1, f_1), \dots, g(1, f_{64}), \dots, g(64, f_1), \dots, g(64, f_{64})]^T, \quad [5]$$

et ν est le seuil de détection.

La construction du filtrage linéaire consiste à trouver les composantes $g(k, f_j)$ du vecteur \vec{V} optimal au regard du critère de contraste retenu. En posant η_0 (resp. η_1) et σ_0 (resp. σ_1) les moments statistiques d'ordre un et deux du classifieur linéaire λ_{TF} conditionnellement à l'hypothèses H_0 (resp. H_1), on a :

$$\eta_i = E \left\{ \vec{V}^T \vec{Y} - \nu | H_i \right\} = \vec{V}^T \vec{M}_i - \nu, \quad [6]$$

$$\sigma_i = \vec{V}^T E \left\{ \left(\vec{Y} - \vec{M}_i \right) \left(\vec{Y} - \vec{M}_i \right)^T | H_i \right\} \vec{V} = \vec{V}^T \Sigma_i \vec{V}, \quad [7]$$

où, conditionnellement à H_0 et H_1 , les vecteurs moyens sont $\vec{M}_i = E \left\{ \vec{Y} | H_i \right\}$, et les matrices de covariances $\Sigma_i = E \left\{ \left(\vec{Y} - \vec{M}_i \right) \left(\vec{Y} - \vec{M}_i \right)^T | H_i \right\}$.

Fukunaga a alors démontré analytiquement que : si le critère de contratsate dépend uniquement des moments d'ordre un et deux du classifieur linéaire (η_0, η_1, σ_0 et σ_1), alors le vecteur optimal \vec{V}_{opt} est de la forme :

$$\vec{V}_{opt} = [\alpha \Sigma_0 + (1 - \alpha) \Sigma_1]^{-1} (M_1 - M_0) , \quad [8]$$

et

$$\alpha = \frac{\partial C / \partial \sigma_0^2}{\partial C / \partial \sigma_0^2 + \partial C / \partial \sigma_1^2} \quad [9]$$

où $C(\nu_0, \nu_1, \sigma_0^2, \sigma_1^2)$ est le critère de contraste, Σ_0 et Σ_1 sont les matrices de covariance, M_0 et M_1 les vecteurs moyennes associés respectivement aux éléments de la classe des complexes-K et de celle des non complexes-K. Sous certaines conditions peu contraignantes pour les citères de contraste considérés ($\partial C / \partial \sigma_0^2$ et $\partial C / \partial \sigma_1^2$ de même signe), le paramètre α est une valeur comprise entre 0 et 1. On doit noter que les critères de contraste pris en compte par ce résultat forment un large ensemble. Et, les critères de contraste les plus usuels, comme celui de Fisher, appartiennent à cet ensemble. A titre d'exemple, l'expression du critère de Fisher est :

$$C_{Fisher} = (\nu_0 - \nu_1)^2 / (P_0 \sigma_0^2 + (1 - P_0) \sigma_1^2) , \quad [10]$$

où P_0 la probabilité *a priori* de l'hypothèse H_0 . Par ailleurs, il est également remarquable que le vecteur optimale \vec{V}_{opt} soit défini à une valeur scalaire α près.

Dans ces conditions, il devient possible de s'affranchir de la notion de critère de contraste. Pour chaque valeur de α , on peut déterminer un classifieur, dont on estime ses performances en terme de taux de bonne détection par rapport au taux de fausse alarme (courbe COR). Le choix final en faveur d'une structure de décision se fait en fonction des objectifs que l'on s'est fixés. Dans notre cas, nous retenons le détecteur minimisant le taux de fausse alarme pour un taux de bonne détection fixé (95%).

2.5. Réduction de complexité

A première vue, la présentation que nous venons de faire, laisserait supposer que le problème de la détermination du classifieur optimal à partir de données expertisées est, à présent, résolu. En réalité, le classifieur que nous avons précédemment construit, est destiné à détecter des complexes-K sur les enregistrements sommeil de n'importe quel sujet. Il faut donc s'assurer de la robustesse de notre classifieur. Ce dernier ne doit pas être trop dépendant des particularités de l'ensemble d'apprentissage.

Il est en particulier connu, voir [VAP 82, VAP 71], que plus un système de décision possède de degrés de liberté, dont le nombre est formellement défini par la dimension de Vapnik-Chervonenkis, plus la taille de l'ensemble d'apprentissage doit être significative. Dans le cas présent les transformées de Wigner-Ville correspondent à des images de 4096 (64×64) points. On peut démontrer que tous ces points ne sont pas linéairement indépendants et que la dimension réelle du système est de 1087,

voir [RIC 98b, RIC 98a]. Cette complexité reste tout de même très élevée, et il est irréaliste de penser obtenir un ensemble d'apprentissage suffisamment conséquent à partir de données cliniques : une technique de régulation de la complexité du problème s'avère nécessaire pour que le classifieur que nous cherchons à construire, soit robuste.

Une méthode, appelée *Optimal Brain Damage (OBD)*, adaptée par Cédric Richard [RIC 99], aux classifieurs linéaires, apporte une solution à cette nécessité de contrôler la complexité. Cette approche nécessite des données indépendantes de celles utilisées pour l'apprentissage proprement dit. Dans notre cas, nous avons constitué cet ensemble à partir des enregistrements sommeils de trois nouveaux jeunes volontaires sains mâles.

D'après la méthodologie précédemment décrite, pour chaque valeur de α nous construisons un classifieur linéaire ayant pour vecteur \vec{V}_α :

$$\vec{V}_\alpha = [\alpha \Sigma_0 + (1 - \alpha) \Sigma_1]^{-1} (M_1 - M_0) . \quad [11]$$

La méthode (OBD) utilise une décomposition en vecteurs et en valeurs propres de la matrice $Mat_\alpha = [\alpha \Sigma_0 + (1 - \alpha) \Sigma_1]^{-1}$. En notant $(v_1, v_2, v_3, \dots, v_{4096})$ les valeurs propres de cette matrice en ordre décroissant, le contrôle de la complexité s'obtient en négligeant toutes les valeurs propres au delà d'un certain rang n_α : on affecte une valeur nulle aux termes de la matrice Mat_α diagonalisée correspondants. Si par diagonalisation, la matrice Mat_α est de la forme :

$$Mat_\alpha = P^1 \cdot \begin{pmatrix} v_1 & & 0 & & \\ & \ddots & & & \\ & & & & \\ 0 & & & & v_{4096} \end{pmatrix} P, \quad [12]$$

où P est une matrice de changement de base. On en déduit une nouvelle matrice $Mat_{(n_\alpha, \alpha)}$ de la forme :

$$Mat_{(n_\alpha, \alpha)} = P^1 \cdot \begin{pmatrix} v_1 & 0 & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & v_{n_\alpha} & \ddots & & \vdots \\ \vdots & & \ddots & 0 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 & 0 \end{pmatrix} P, \quad [13]$$

et un nouveau filtre linéaire $\vec{V}_{(n_\alpha, \alpha)}$:

$$\vec{V}_{(n_\alpha, \alpha)} = Mat_{(n_\alpha, \alpha)} (M_1 - M_0) . \quad [14]$$

La valeur du rang n_α est obtenu, pour chaque valeur de α , en minisant, sur l'ensemble test, le taux de fausse alarme à un taux de bonne détection fixé (95%) du classifieur construit avec $\vec{V}_{(n_\alpha, \alpha)}$.

En fin de compte, l'optimisation du classifieur ne se fait plus sur un paramètre scalaire et un ensemble de données, mais sur deux paramètres scalaires et deux ensembles de données décorrélés.

2.6. Estimation des performances

Pour être tout à fait rigoureux, l'estimation des performances du classifieur a été déterminée à partir d'un troisième ensemble de données décorrélé à l'ensemble d'apprentissage et à l'ensemble test. Ce troisième ensemble a été obtenu, d'une manière identique aux deux premiers ensembles, à partir des enregistrements sommeil de trois nouveaux volontaires sains.

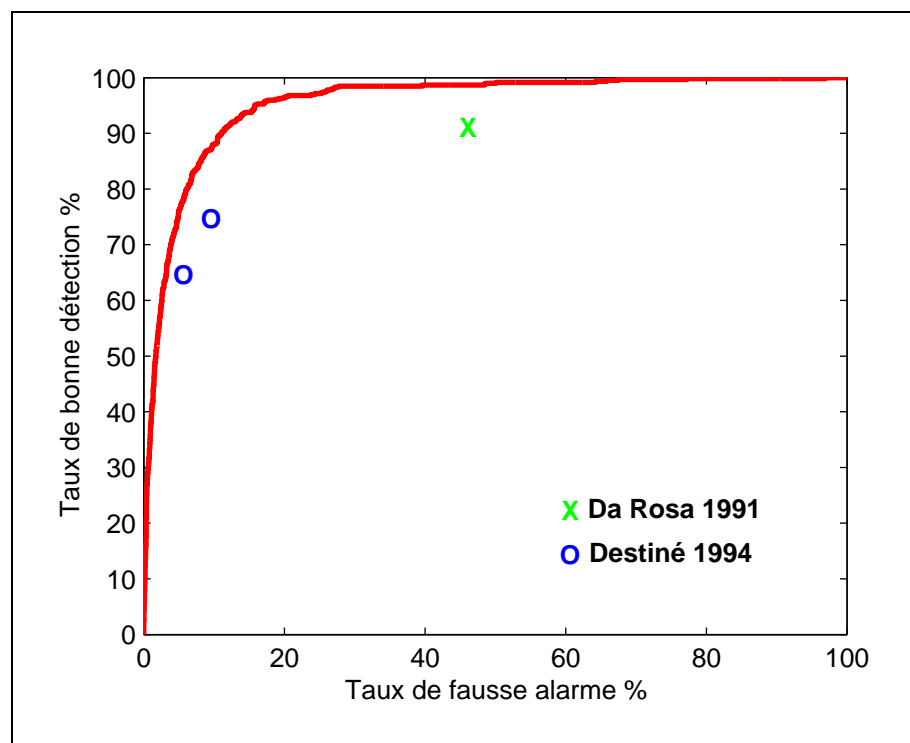


Figure 3. Courbe COR du classifieur.

Même s'il est toujours délicat de comparer des performances sur des données différentes, la figure (3) semble tout de même indiquer une amélioration importante par rapport aux travaux précédents, voir [DA 91, DA 93, DeB94]. Aux points de fonctionnement proposés dans ces articles, on constate que notre classifieur offre des taux de fausse alarme nettement réduits.

En pratique et suivant les applications, l'utilisateur choisira un point de fonctionnement du classifieur en fixant un seuil, c'est à dire en fixant un taux de bonne détection. Le classifieur que nous venons de concevoir a été optimisé pour un taux de bonne détection de (95%). Il est évidemment possible d'optimiser le classifieur à un taux de bonne détection différent.

3. Conclusion

La méthodologie que nous venons de décrire dans cet article, propose un modèle de classification à partir de données expérimentales expertisées. Nous avons montré comment gérer les conflits entre experts et synthétiser un avis global pertinent pour la classification.

L'intérêt du choix de l'espace de représentation des données, ici le domaine temps-fréquence, a aussi été mis en évidence. A cet égard, les liens entre un classifieur linéaire défini sur une transformée de Wigner-Ville et la représentation temporelle du signal ont été rappelés.

De même, la construction d'un classifieur linéaire optimal à partir de deux classes de données, en supposant un minimum d'*a priori* sur la structure statistique des données, a été explicitée. En plus, un soin particulier a été apporté à la gestion de la complexité et à l'obtention d'une classification robuste. On soulignera que les démarches scientifiques proposées, ici, dans le domaine temps-fréquence sont issues des recherches les plus récentes.

Enfin, l'estimation des performances, quantifiée par une courbe COR (Caractéristique Opérationnelle du Récepteur), a permis la comparaison avec les études précédentes. Il est alors clairement apparu que l'intégration des différents traitements que nous appliquons aux données, induisent un gain de sélectivité important.

Même si les méthodes présentées sont bien adaptées à l'application biomédicale particulière qu'est la détection de complexes-K, nous pensons que notre approche peut répondre à une classe très large de problème. Dans beaucoup de domaines où l'interprétation de signaux issus de processus complexes est nécessaire, les avis des experts, avec tous les problèmes que cela engendre, sont les seules références reconnues. Notre méthodologie devient alors une voie possible pour automatiser, classifier et rationaliser les décisions.

4. Bibliographie

- [BAN 92] BANKMAN I. N., SIGILLITO V. G., WISE R. A., SMITH P. L., « Feature-based detection of the K-complex wave in the human electroencephalogram using neural networks », *IEEE Trans. Biomed. Eng.*, vol. 39, n° 12, 1992, p. 1305-1310.
- [BAS 92] BASTIEN C., CAMPBELL K., « The evoked K-Complex : All-or-none phenomenon ? », *Sleep*, vol. 15, n° 3, 1992, p. 236-245.

- [BRE 70] BREMER G., SMITH J. S., KARAKAN I., « Automatic detection of K-complex in sleep electroencephalograms », *IEEE Trans. Biomed. Eng.*, vol. 17, 1970, p. 314-323.
- [DA 91] DA ROSA A. C., KEMP B., PAIVA T., LOPES DA SILVA F. H., KAMPHUISEN H. A. C., « A model-based detector of vertex waves and K complexes in sleep electroencephalogram », *Electroencephalography and clinical Neurophysiology*, vol. 78, 1991, p. 71-79.
- [DA 93] DA ROSA A. C., PAIVA T., « Automatic detection of K-complexes : Validation in normals and dysthymic patients », *Sleep*, vol. 16, n° 3, 1993, p. 239-245.
- [DeB94] *Utilisation des Réseaux de Neurones Artificiels pour la reconnaissance des Grapho-Éléments Phasiques dans le cadre de l'Analyse du Sommeil*, Symposium International 25ans d'Analyse Automatique du Sommeil, 1994.
- [DEV 96] DEVROYE L., GYORFI L., LUGOSI G., *A Probabilistic Theory of Pattern Recognition*, Springer Verlag, New York, 1996.
- [FLA 88] FLANDRIN P., « A Time-Frequency Formulation of Optimum Detection », *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, n° 9, 1988, p. 1377-1384.
- [FUK 90] FUKUNAGA K., *Statistical pattern recognition*, Academic Press, London, 2nd ed. édition, 1990.
- [JAC 96] JACQMIN A., SCHALTENBRAND N., LENGELLÉ R., MUZET A., MACHER J., « Use of time-frequency analysis to discriminate between K-complexes and delta waves », *J. Sleep Res.*, vol. 5, n° 1 (supp.), 1996, page 98.
- [JAN 90] JANSEN B. H., « Artificial neural nets for K-complex detection », *IEEE Eng. Med. Biol.*, vol. 9, 1990, p. 50-52.
- [JOB 92] JOBERT M., POISEAU E., JAHNIG P., SCHULZ H., KUBICKI S., « Pattern recognition by matched filtering : an analysis of sleep and K-complex density under the influence of lormetazepamzopiclone », *Neuropsychobiology*, vol. 26, n° 1-2, 1992, p. 100-107.
- [JON 95] JONES D. L., SAYEED A. M., « Blind Quadratic and Time-Frequency based Detectors from Training Data. », *Actes de IEEE International Conference on Acoustics, Speech and Signal Processing*, 1995, p. 1033-1036.
- [LOO 38] LOOMIS A. L., HARVEY E. N., HOBART G. A., « Distribution of disturbance-patterns in the human electroencephalogram, with special reference to sleep », *J. Neurophysiol.*, vol. 1, 1938, p. 413-430.
- [RIC 98a] RICHARD C., LENGELLÉ R., « On the Dimension of the Discrete Wigner-Ville Transform Range Space. Application to Time-Frequency-Based Detectors Design », *Actes du IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis - IEEE TFST'98*, Pittsburgh, PA, USA, 1998, p. 5-8.
- [RIC 98b] RICHARD C., LENGELLÉ R., « Two Algorithms for Designing Optimal Reduced-Bias Data-Driven Time-Frequency Detectors », *Actes du IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis - IEEE TFST'98*, Pittsburgh, PA, USA, 1998, p. 601-604.
- [RIC 99] RICHARD C., LENGELLÉ R., « Data-driven Design and Complexity Control of Time-Frequency Detectors », *Signal Processing*, vol. 77, n° 1, 1999, p. 37-48.
- [VAP 71] VAPNIK V. N., CHERVONENKIS A. Y., « On the Uniform Convergence of Relative Frequencies of Events to their Probabilities », *Theory of Probabilities and its Applications*, vol. 16, 1971, p. 264-280.

- [VAP 82] VAPNIK V. N., *Estimation of dependences based on empirical data*, Springer Verlag, New York, 1982.
- [VIA 01] VIARDOT G., RICHARD C., LENGELLÉ R., COATANHAY A., « Fusion partiellement supervisée d'avis d'experts. Application à la décision statistique », *Actes du congrès national en traitement du signal GRETSI 2001*, Toulouse, France, 10-13 septembre 2001.