

**DETECTION ET CARACTERISATION EN TEMPS DIFFERE DE SAUTS DE DYNAMIQUE :
APPLICATION A LA SEGMENTATION DE L'EMG****Thibault LAMOTTE⁽¹⁾, Alain PRIEZ⁽²⁾, Evelyne LEPOIVRE⁽³⁾
Jacques DUCHENE⁽⁴⁾, Claude TARRIERE⁽³⁾**

1 : AARIMT, 15 rue de l'école de Médecine, 75006 Paris

2 : IRACBA, Hopital R. Poincaré, 104 bld R. Poincaré, 92380 Garches

3 : Renault, Département Biomédical de l'Automobile, 132 rue des Suisses, 92000 Nanterre

Tel. 33 1 47 77 35 14, Fax. 33 1 47 77 36 36

4 : Université de Technologie de Troyes (UTT), 11 rue M. Curie, 10010 Troyes

RÉSUMÉ

Cet article présente une méthode de détection-caractérisation d'un signal non stationnaire. La détection repose sur une comparaison de variance à l'aide de la loi de Fisher. Dans une première partie, des indices de performance de l'algorithme sont construits, puis la méthode est appliquée à un signal électromyographique (EMG). Dans une seconde partie, une phase de caractérisation en mode supervisé est réalisée. Différentes activités musculaires mais aussi des parasites électriques et mécaniques sont recherchés. Cet ensemble de méthodes autorise la compréhension précise des enregistrements réalisés.

ABSTRACT

This paper presents a detection-characterization method for a non stationary signal. The detection algorithm is based on a comparison of variance. First, a performance analysis of the algorithm is made. Then, this detection method is applied to an electromyographic (EMG) signal. Secondly, a characterization step is realized in a supervised way. Several muscular activities, but also electrical and mechanical noises, are located. This combination of methods allows comprehensive studies of the recorded signals.

1. INTRODUCTION

Afin de créer un indice physiologique du confort apporté par un appui-tête dans une voiture personnelle, des mesures électromyographiques (EMG) des activités des muscles du cou sont effectués en condition réelles de conduite. Lamotte et al. [1] mettaient en évidence la possibilité d'analyse de ces signaux en considérant l'évolution temporelle de percentiles de la fonction de répartition de l'EMG redressé. La variabilité intra-sujet (entre deux jours différents) est assez faible, par contre la variabilité inter-sujet est plus préjudiciable. Ainsi, des résultats difficilement explicables ont été observés. Il est donc nécessaire de mieux comprendre ce qui a été enregistré.

En effet, ces signaux sont non stationnaires. La plupart du temps, seule une ligne de base ne représentant pas une contraction musculaire volontaire, mais un signal composite comprenant le bruit électronique et une activité musculaire tonique, est enregistrée. Cette ligne de base est parfois "contaminée" par l'activité cardiaque. En cas de contraction musculaire, la variance du signal augmente selon l'intensité de l'activité musculaire. Pour la phase de détection, à chaque échantillon, l'alternative suivante est testée :

S0 : il n'y a pas de saut de dynamique.

S1 : il y a un saut de dynamique.

Ensuite, il est nécessaire de caractériser les événements isolés. A priori, trois types d'événements sont observables : soit des parasites 50 Hz (induits par les lignes haute-tension), soit un artefact mécanique ou électrocardiographique (ECG), soit une contraction musculaire volontaire réelle.

En analogique, les signaux sont amplifiés d'un facteur de 80 dB. Compte tenu des caractéristiques spectrales du signal, les

signaux sont échantillonnés à 512 Hz sur 16 bits. La durée des signaux EMG utilisés est de 20 minutes. Pour chaque série de signaux considérés, quatre muscles sont enregistrés.

2. ALGORITHME DE DETECTION

Au vu des caractéristiques du signal brut, il est nécessaire d'implanter un algorithme de détection de saut de dynamique [2]. Excepté l'algorithme SNR (détection de changement de variance, [2]), les méthodes généralement rencontrées nécessitent une modélisation AR du signal. Choisir un ordre de modèle pertinent pour des signaux non stationnaires est difficile. Pour l'algorithme EPL, les résultats sont d'autant meilleurs que le modèle est correctement ajusté [2]. Aussi, pour éviter ces problèmes, un algorithme de comparaison de variance est implanté avec la statistique de Fisher [3]. Les fenêtres utilisées sont : une fenêtre de décision qui estime les caractéristiques locales du signal ($F_{dec}, \hat{\sigma}_{dec}^2$) et une fenêtre d'observation qui donne les caractéristiques globales ($F_{obs}, \hat{\sigma}_{obs}^2$). Les hypothèses testées sont les suivantes :

$$H_0 : \hat{\sigma}_{dec}^2 = \hat{\sigma}_{obs}^2 \text{ contre } H_1 : \hat{\sigma}_{dec}^2 > \hat{\sigma}_{obs}^2.$$

En supposant l'indépendance et le caractère gaussien des échantillons dans chaque fenêtre, le seuil critique est fixé par la loi de Fisher. La loi du paramètre sous H_1 est inconnue.

2.1) Premier essai, sans réinitialisation

En première approche, la fenêtre de décision (F_{dec}) a été choisie d'une taille de 128 points et la fenêtre d'observation comme tout le signal (F_{obs}). Dans ce cas, les fenêtres ne sont



pas indépendantes et F_{obs} contient à la fois des événements sous H_0 et sous H_1 . Le seuil critique ne peut donc pas être fixé par la loi de Fisher. Par observation de la loi mélange des rapports de variance expérimentaux ($X = \frac{\hat{\sigma}_{dec}^2}{\hat{\sigma}_{obs}^2}$, calculé pour

chaque point de la fenêtre d'observation), il est parfois possible de choisir un seuil séparant les événements. En effet, un caractère bimodal de la loi conjointe est observé pour environ la moitié des cas considérés. La figure 1 représente quatre exemples de lois conjointes observées.

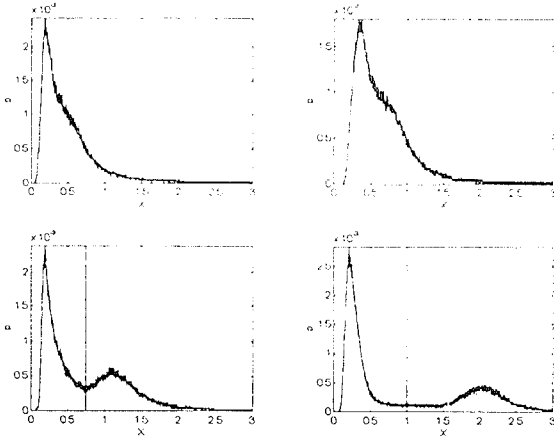


Figure 1 : Graphes avec lois conjointes et seuils choisis (signaux réels)
 X : rapport de variance, p : probabilité

Dans le cas bimodal, le signal peut alors être découpé en deux parties : soit des événements de faible variance, soit des événements de grande variance. Les risques associés sont appréhendés à l'aide de la forme de la loi. Cependant, cette méthode est très limitée :

- Que faire pour les cas unimodaux ?
- Pour les cas bimodaux, le seuil est-il placé en fonction des événements à isoler ?
- Les zones isolées par cette méthode sont parfois non stationnaires donc difficiles à caractériser.

2.2) Méthode retenue, avec réinitialisation

Auparavant, la fenêtre d'observation contenait les différentes composantes du signal. Cependant, la loi théorique est uniquement connue sous H_0 , c'est à dire lorsque la variance est indépendante du temps pour un taux de fausse alarme donné.

F_{dec} est toujours choisie de taille constante T_{dec} , mais F_{obs} est une fenêtre "grandissante" en fonction des échantillons du signal. Au départ, la taille minimale de F_{obs} est égale à T_{dec} . Dans ces conditions, il est alors possible d'utiliser directement l'algorithme proposé par Ragot et al. [3]. De

manière itérative, à chaque échantillon, on teste si $X = \frac{\hat{\sigma}_{dec}^2}{\hat{\sigma}_{obs}^2}$ appartient à l'intervalle de confiance construit sur la loi de Fisher. Si ce n'est pas le cas, alors on réinitialise les fenêtres.

Cette approche s'adapte donc toujours au signal considéré. En sortie de cette méthode, le signal est alors découpé en zones où la variance est stationnaire. Les paramètres à régler pour la méthode sont : T_{dec} , taille de la fenêtre de décision et α taux de fausse alarme théorique.

Pour valider le détecteur, il est proposé de tester dans un premier temps les fausses alarmes (FA) expérimentales sur des séries temporelles sans rupture. Cela conduira à retenir une configuration idéale. Cette configuration sera testée dans un second temps quant aux non-détections (ND) et retards à la détection. Ainsi, des critères de performance de l'algorithme utilisé seront disponibles.

2.3) Sensibilité aux FA, choix de T_{dec}

Le taux de FA expérimentales est calculé comme une fonction de α et de T_{dec} pour des réalisations d'un bruit blanc gaussien de variance unité (30 séries de 1024 points). La moyenne des taux de FA observées est présentée sur le tableau 1. En première approche, les valeurs de FA théoriques sont des valeurs usuelles.

T_{dec}	4	6	8	10	12	14
PFA	0.006	0.004	0.003	0.003	0.011	0.024
5%	(0.28)	(0.47)	(0.63)	(0.63)	(0.15)	(0.05)
PFA	0.002	0.001	$6 \cdot 10^{-4}$	$4 \cdot 10^{-4}$	0.004	0.018
1%	(0.61)	(0.94)	(1.36)	(1.52)	(0.17)	(0.09)

T_{dec}	16	20	30	50	100	128
PFA	0.038	0.079	0.150	0.193	0.397	0.704
5%	(0.07)	(0.11)	(0.22)	(0.39)	(0.82)	(0.59)
PFA	0.030	0.063	0.137	0.185	0.368	0.652
1%	(0.07)	(0.13)	(0.22)	(0.40)	(0.73)	(0.51)

Tableau 1: performance au FA.

PFA : probabilité moyenne de FA.

entre parenthèse coefficient de variation du PFA

Il apparaît que des fenêtres de petite taille (<16 points) procurent un taux de fausse alarme minimum. En effet, plus T_{dec} augmente, plus l'intervalle de confiance de la loi de Fisher est sélectif. Le risque de FA est alors augmenté. Par contre, le choix de α ne semble pas influencer énormément les résultats. Pour des tailles de fenêtre comprises entre 12 et 16 points, les résultats sont très reproductibles mais conduisent à une probabilité moyenne de FA supérieure par rapport à une fenêtre de 8 ou 10 points. Aussi, pour les paragraphes suivants, T_{dec} est choisi égal à 8 points

2.4) Non détection et retards à la détection, choix de α

Des réalisations indépendantes ($N=100$) de signaux synthétiques de 100 points, contenant une rupture à l'échantillon 50, sont créées. Avant la rupture, le signal est un bruit blanc de variance unité. Après la rupture, la variance est multipliée par un coefficient k . Le tableau 2 présente les fausses alarmes observées (PFA), le nombre de non-détections (ND) et les retards à la détection moyens (RD) constatés en fonction du rapport signal à bruit ($RSB=k$).

Le tableau 2 permet de choisir entre ces deux valeurs usuelles de FA théorique : 5% conduit à de meilleurs résultats globaux (FA et ND) que 1%. Etant donné l'existence des FA expérimentales, la méthode procurera parfois un minorant de la durée réelle des zones de stationnarité.

RSB	3	2	1.75	1.5	1.25	0.85
PFA théorique 5 %						
PFA	0.014	0.005	0.004	0.005	0.004	0.004
ND	0	0.02	0.05	0.19	0.40	0.89
RD	-4	0	3	8	13	16
PFA théorique 1 %						
PFA	0.003	0.002	0.001	0.001	$3 \cdot 10^{-4}$	$8 \cdot 10^{-4}$
ND	0.01	0.05	0.25	0.37	0.77	0.99
RD	-3	3	6	9	13	-8

Tableau 2 : performance aux ND et RD,
 PFA : probabilité moyenne de FA,
 ND : taux moyen de non détection (probabilité),
 RD : retard moyen à la détection (en points)

Si la loi du paramètre sous H1 était connue, il serait possible de trouver le meilleur seuil procurant un taux minimal de FA et ND. L'algorithme utilisé ne permet pas ce choix de seuil optimal. Le seuil critique est donc fixé à 5%. Il sera obligatoire de contrôler par la suite (notamment sur les signaux EMG) si les FA et ND dues à la méthode ne sont pas préjudiciables.

2.5) Application à l'EMG

Les paragraphes précédents ont permis de régler l'algorithme. Une fenêtre de décision de 8 points (soit 15,6 ms) et un taux de FA de 5% ont été choisis. Cette méthode autorise alors un découpage du signal en zone de plus grande stationnarité. La durée de ces zones isolées est assez courte : la plupart des zones sont de durée inférieure à la seconde. Sur 16 cas (4 enregistrements de 20 minutes avec 4 muscles), les moyennes et écart-types de statistiques descriptives de ces durées sont : 0.119 (± 0.029) secondes (s) pour le premier quartile, 0.212 (± 0.095) s pour la médiane et 0.436 (± 0.196) s pour le troisième quartile.

3.) METHODE DE CARACTERISATION

Les modes de fonctionnement précédemment énoncés sont recherchés par une analyse de type supervisé. Pour chaque mode, les étapes de caractérisation sont les suivantes :

A) Calcul d'une statistique permettant la classification en fonction de la caractéristique essentielle du mode de fonctionnement considéré.

B) Recherche d'un seuil permettant de classer les événements. De manière itérative, le seuil est déplacé et les événements isolés sont classés par l'expérimentateur de manière supervisée. Dès que le nombre de fausses alarmes devient trop important, la procédure est stoppée. Le nombre de bonnes décisions et mauvaises décisions est alors évalué et conduit à la construction de courbes de type Caractéristique Opérationnelle du Récepteur (COR) pour chaque cas considéré [4].

3.1) Recherche de parasites 50 Hz

Ce point constitue un parasitage ponctuel mais souvent d'amplitude importante.

La statistique (S50) considérée est déduite de l'autospectre de l'événement. C'est le rapport entre l'énergie des

points de l'autospectre situés dans la région autour de 50 Hz et l'énergie totale du spectre. La bande spectrale considérée est limitée par le point immédiatement inférieur et le point immédiatement supérieur à 50 Hz.

Les fenêtres, où S50 est maximale, ont le plus de chances d'être parasitées. La figure 2 représente la courbe COR associée. Un seuil de 0.63 permet d'isoler 12 événements sur 15 avec 6 fausses alarmes. Un taux de détection de 1 nécessite 17 fausses alarmes. Il est généralement constaté que ces fausses alarmes sont constituées par des segments de courte durée, où la résolution spectrale est mauvaise. Par exemple, des activités ECG peuvent alors être considérées comme du 50 Hz... Pour éviter cela, une deuxième statistique égale à S50 multipliée par la durée du segment permet d'améliorer le détecteur : sur les données précédentes, les résultats sont représentés sur la figure 3. Avec cette seconde statistique, la détection des 15 événements n'entraîne plus que 2 FA.

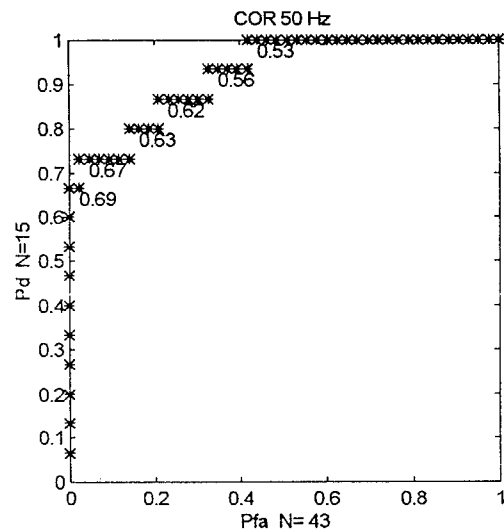


Figure 2 : Courbe COR expérimentale pour les 50 Hz avec S50
 Pd : proportion d'événements détectés
 Pfa : proportion de FA

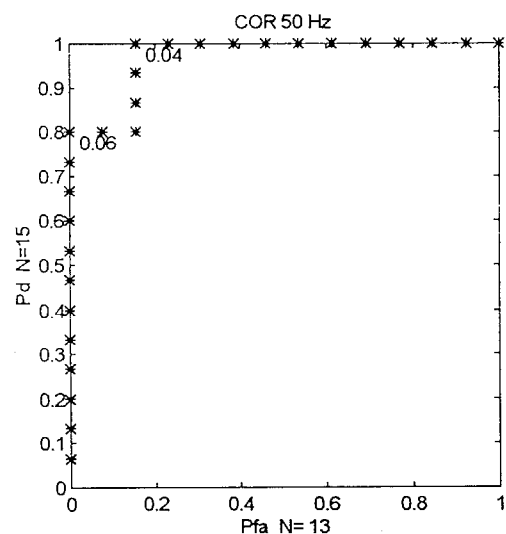


Figure 3 : Courbe COR expérimentale pour les 50 Hz, avec S50 multipliée par la durée du segment
 légende : voir figure 2



3.2) Recherche des artefacts mécaniques et ECG

Un artefact mécanique est défini comme une zone où le signal montre des variations trop lentes pour représenter un phénomène physiologique. Les origines de ces problèmes sont diverses : mouvement relatif peau-électrodes, bruit lié à un mouvement de fil, lien avec un phénomène vibratoire externe,... De même, l'ECG contamine l'EMG par la présence d'activité QRS. Ce parasite se caractérise également par un contenu fréquentiel plus lent que l'EMG. Contrairement au cas précédent, la probabilité d'apparition de ces parasites est plus grande.

En première approche, la statistique retenue est la différence des deux premiers coefficients de la fonction d'autocorrélation normalisée.

La courbe COR est construite à l'aide des segments étiquetés par l'expérimentateur. La position dans le plan de la courbe COR permet d'appréhender la qualité de la statistique et du seuil choisi, à séparer les artefacts mécaniques (figure 3).

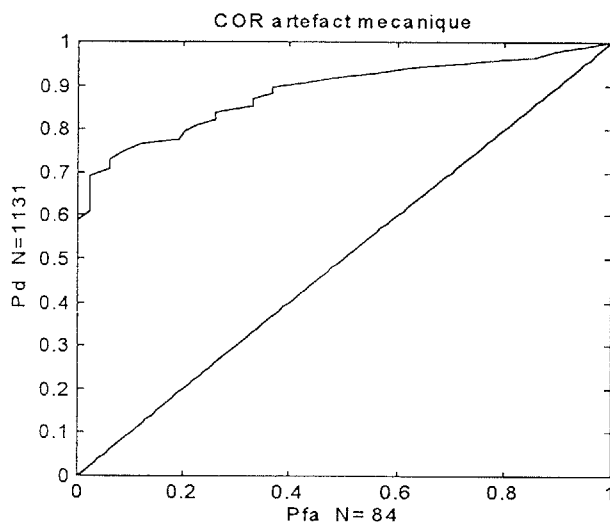


Figure 3 : Courbe COR expérimentale pour les artefacts mécaniques
légende : voir figure 2

Le contrôle des résultats est ici plus difficile :

- Le nombre d'événements à isoler est important (particulièrement pour les séries contaminées par l'ECG). Des oublis sont ainsi possibles.

- Il existe des segments composites : une ligne de base avec un parasitage mécanique (QRS ou offset). Comment classer ces événements ? Ce point constitue d'ailleurs une limite liée au choix du taux de FA théorique choisi : pour ces segments, il serait nécessaire de recalculer le seuil de détection non plus à 5% mais avec une valeur supérieure.

3.3) Recherche des vraies bouffées

Si les étapes précédentes ont correctement fonctionné (avec une probabilité de détection de 1), alors seuls les segments ayant une signification physiologique sont considérés. En visualisant ces segments, il est possible de vérifier les éventuels oublis des étapes précédentes. Cela autorise un contrôle de l'efficacité des méthodes précédentes.

Cependant, il reste dans le signal les fenêtres contenant les lignes de base (bruit électronique et EMG tonique ou postural) et les réelles contractions (EMG phasique). La

différence entre ces deux classes d'activité peut a priori se réaliser par la puissance du signal.

La statistique à considérer est donc la variance de la fenêtre d'observation. Les fenêtres de grande variance contiennent les activités réelles de contraction musculaire. A l'inverse, les lignes de base sont des activités de faible puissance.

Séparer les activités de type contractions volontaires des activités qualifiées de ligne de base semble théoriquement possible. Cependant, les premières observations n'ont pas encore autorisé le choix de ce seuil. Il est évidemment souhaitable de chercher une autre statistique permettant de classer ces deux types d'événements. Quelle que soit cette statistique, la méthode précédemment présentée quant au choix du seuil peut s'appliquer. La statistique la plus pertinente sera alors facilement observée à l'aide des courbes COR expérimentales.

Conclusions :

Ces méthodes de détection-caractérisation permettent une analyse des événements enregistrés sur chaque voie EMG. Après la recherche de zones où le signal EMG est stationnaire à l'ordre 2 par l'algorithme utilisant le test de Fisher, chaque zone a été caractérisée afin d'isoler des modes de fonctionnement connus. Les deux premiers modes (50 Hz et artefacts "lents") sont évidemment des parasites. Les deux derniers modes représentent l'information d'origine musculaire. Après validation des résultats par l'expérimentateur, celui-ci a alors accès aux segments ayant une signification précise pour les études conjointes de physiologie et de confort. Les données déjà présentées [1] peuvent ainsi être analysées une seconde fois de manière plus robuste.

Lorsque la recherche des bouffées sera efficace, il sera possible de connaître "l'arrivée" de ces vraies bouffées et ainsi étudier de possibles co-activations entre les quatre muscles étudiés mais aussi les liens avec des variables cinématiques (mesurées sur le plancher du véhicule et le front du sujet) ou mécaniques (force de pression exercée sur l'appui-tête). Ce lien entre variables physiologiques et mécaniques autorisera un rangement des activités observées selon leur nature : réaction suite à un événement extérieur (par exemple, freinage), mouvement de tête dans une direction connue, appui de la tête sur le coussin de l'appui-tête,... Finalement, pour chaque type d'activités observées, l'analyse de la structure des signaux, et de l'évolution temporelle de cette structure, semblerait autoriser un jugement objectif quant au confort apporté par l'appui-tête.

Références :

- [1] LAMOTTE T., PRIEZ A., LEPOIVRE E., DUCHENE J., TARRIERE C., 1994, Time varying Amplitude Probability Density Function : Application to the study of head rest comfort in a car, IEEE EMBS 16th annual conference, Baltimore
- [2] GDR TDSI 1992, Segmentation des signaux, Traitement du Signal, Supplément vol. 9, numéro 1, pp. 115-141
- [3] RAGOT J., MAQUIN D., DAROUACH M., BLOCH G., 1990, Validation de données et diagnostic, chapitre 3 : Quelques méthodes de détection de changement dans le comportement d'un système, Hermès
- [4] BARKAT M., 1991, Signal detection and estimation, Artech House