

CLASSIFICATION ET AIDE AU DIAGNOSTIC EN E.M.G.

S. SITBON - F. GASMI - F. CASTANIE (\*)  
 JP. CARILLO - J. PERISSE - JJ. MOOG (\*\*)

(\*) GAPSE - ENSEIHT - Rue camichel 31071 TOULOUSE  
 (\*\*) Faculté de médecine de Toulouse

RESUME

L'objectif de cette communication est de présenter des méthodes de classification automatique du signal électromyographique. Le praticien dénombre douze états pathologiques ( myogène ou neurogène ), diagnostic établi sur le signal enregistré au moyen d'électrode aiguille. Les performances du classifieur ont été également testées sur les signaux de surface. La procédure de classification utilise tout d'abord une modélisation paramétrique autorégressive des signaux à traiter.

L'algorithme de classification procède ensuite à la reconnaissance de la classe par la recherche de la distance minimale entre le vecteur "inconnu" et les représentants moyens de chaque classe.

Au stade actuel de l'étude, la validité des méthodes a été vérifiée sur des signaux appartenant à la base de donnée.

De bons résultats ont été obtenus, notamment avec la métrique de 'Mahalanobis' appliquées aux coefficients cepstraux. On notera la validité de la classification sur les signaux de surface.

I INTRODUCTION

L'électromyographie est aujourd'hui un examen de routine. L'étude des manifestations électriques d'un muscle en contraction permet le diagnostic des pathologies du muscle et du système nerveux périphérique. Notre objectif est de fournir un outil d'aide à la décision dans la détermination de ces pathologies.

II RAPPEL SUR LE SIGNAL ELECTROMYOGRAPHIQUE

II-1 Structure d'un muscle strié

Le muscle squelettique est composé d'un ensemble de sous-unités indépendantes - les Unités Motrices (U.M.). Chaque U.M. est elle-même formée de plusieurs fibres musculaires

II-2 Principe de l'excitation

Chaque U.M. est activée par un signal électrique impulsionnel véhiculé par une terminaison nerveuse (axone). A la suite de chaque impulsion nerveuse, une U.M. émet un signal électrique, le M.U.A.P. ( Potentiel d'Action d'Unité Motrice ). Le signal temporel produit par chaque U.M. est donc un train de potentiel d'action, le M.U.A.P.T. (Train de Potentiel d'Action d'Unité Motrice )

La force musculaire ainsi produite est une fonction du nombre d'U.M. actives et de leur fréquence d'activation.

II-3 Allure du signal recueilli

Selon le mode de recueil, on est en présence de deux types de signaux. Une *électrode aiguille* enregistre l'activité des U.M. proches de sa surface active; à bas niveau de force, l'allure des M.U.A.P.T. se distingue aisément; à haut niveau de contraction le signal à l'aspect d'un bruit. Les *électrodes de surface* de par leur nature, captent l'activité d'un nombre important d'U.M. Le signal observé à l'aspect d'un bruit dès qu'une contraction existe.

ABSTRACT

This paper deals with the efficiency of some automatic classification algorithms working on the electromyographic signal. There are twelve distinguishable pathologic states observed on the signal recorded by a concentric needle electrode ( myopathic or neurogenic ).

The first step of the classification procedure is an autoregressive modelling. Thus, the representation space is order modelling dependent and based upon the autoregressive or derived coefficients - cepstral, reflection.

The classification is realized by searching the smaller distance between the signal to be recognized and the average vector of each class.

Several metrics have been employed; they include the classical euclidian, chi-deux, cepstral and weighted by a matrix of variance or covariance.

The best result were obtained with the Mahalanobis algorithm, acting on cepstral parameters.

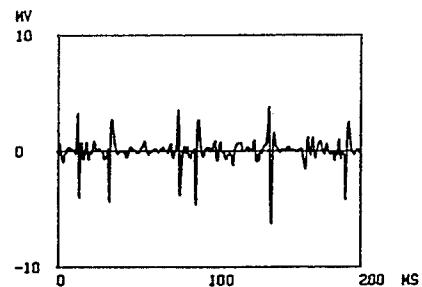


fig. 3a : piquée, contraction faible.

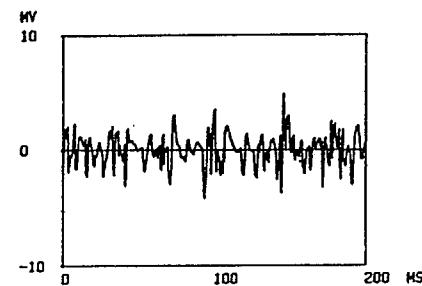


fig. 3b : piquée, contraction forte

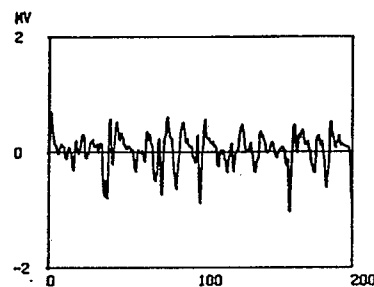


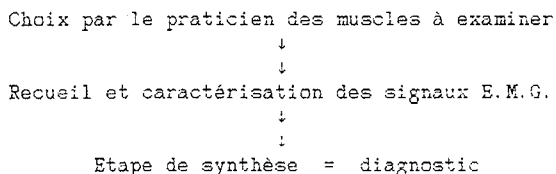
fig. 3c : surface, contraction faible



III OBJECTIFS DE LA CLASSIFICATION AUTOMATIQUE

III-a Situation au cours d'un examen

Le déroulement d'une consultation suit le schéma suivant :



Nous nous proposons d'automatiser l'étape numéro deux. Actuellement, le médecin travaille uniquement sur des signaux recueillis par électrode piquée; les méthodes de classifications seront éprouvées sur les signaux de surface également, ce mode de recueil étant bien mieux perçu par les patients.

III-b Pathologies à distinguer

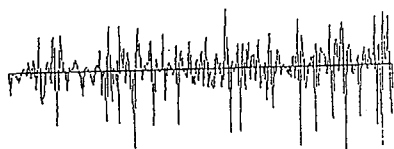
L'électromyographie dénombre principalement treize états caractéristiques de l' E.M.G. :

- stade normal
- myogène
- neurogène stade 2/5, 3A/5, 3B/5, 3C/5, 4

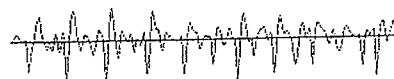
L'aggravation du stade neurogène se traduit par un appauvrissement du tracé ( diminution du nombre d'U.M. actives ).

Les stades 3A, 3B et 3C se divisent chacun en stade aigu ( potentiel de dénervation rajoutés et activité de repos), subaigu (idem sans activité de repos) et chronique ( M.U.A.P. de forme simplifiée, durée normale ou augmentée);

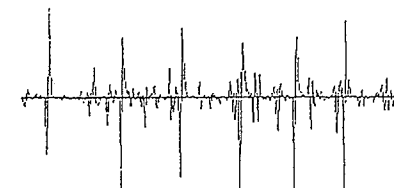
tracé normal



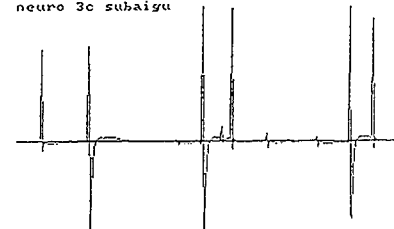
neuro 3A chronique



neuro 3b chronique



neuro 3c subaigu



IV CARACTERISATION DES SIGNAUX

IV-a Méthode choisie

Tout algorithme de classification demande la construction d'un vecteur de paramètres caractérisant le signal. Pour ce faire, nous avons opté pour des méthodes de modélisation paramétrique autorégressives. Elles permettent de prendre en compte l'allure générale du signal par le biais d'un vecteur de paramètres de faible dimension. Celui-ci est composé des coefficients autorégressifs du modèle ou de ses dérivés : coefficients de réflexion (Lévinson-Durbin), cepstraux.

Deux algorithmes de modélisation autorégressive ont été utilisés, méthode de Lévinson-Durbin (Corrélation), algorithme de Morf (Covariance).

IV-b Choix de la tranche

Les extrémités de la tranche du signal sont positionnées de manière à ne pas couper une forme d'onde ( risque d'artéfact dans la modélisation ).

Le praticien opère sur une fenêtre de 400 millisecondes, valeur conservée pour la modélisation.

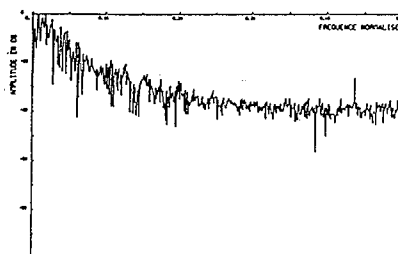
La bande passante à -25 dB étant inférieure à 2.5 kHz pour tous les signaux en notre possession, les caractéristiques suivantes ont été retenues :

OBSERVATION = 400 ms , fe = 5 kHz (2000 points)

IV-c Validation du modèle

- par concordance entre spectre obtenu par périodogramme et spectre autorégressif. Dès l'ordre 5, l'allure de la densité spectrale de puissance est correctement approximée (spectre lissé); à partir de l'ordre 20, une bonne correspondance est obtenue.

- l'erreur de prédiction en fonction de l'ordre fournit quelques indications; l'amélioration ne dépasse pas cependant cinq pour cent entre les ordres cinq et trente.



DSP OBTENUE PAR FFT

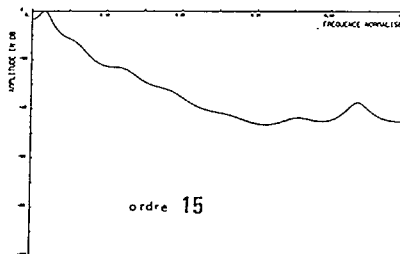


fig.4:périodogramme sur 2048 points et spectre AR (Morf)

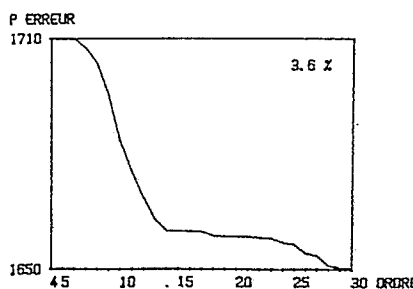


fig.5:erreur de prédiction en fonction de l'ordre

**V CLASSIFICATION**

Dans une première étape, l'efficacité de plusieurs algorithmes de classification a été éprouvée, sur une base de données constituée de signaux expertisés. Puis l'algorithme le plus performant est utilisé pour la reconnaissance de signaux connus non inclus dans la base de données.

**V-a Processus de décision**

Il fait appel au calcul d'une distance (métrique). A chaque classe J de la base de donnée est associé un vecteur moyen Y<sub>J</sub> (centre d'inertie de la classe).

Un vecteur X<sub>i</sub> est déclaré appartenant à une classe J lorsque :

$$d(X_i, Y_J) = \min d(X_i, Y_L) \quad (5.1)$$

L = 1, K n° de la classe

avec X<sub>i</sub> = (c<sup>1</sup>, c<sup>2</sup>, ..., c<sup>p</sup>)<sup>T</sup> p : ordre du modèle  
Toutes ces métriques suivent la forme suivante :

$$d(X, Y) = (X - Y)^T \cdot P \cdot (X - Y)$$

avec P : matrice de pondération symétrique définie positive.

**V-b Distances utilisées**

Diverses métriques ont été mises en oeuvre :

- euclidienne : P = I<sub>p</sub> , p = ordre du modèle

- pondération par les variances :

\* totales : P = diag { ..., 1/σ<sub>i</sub><sup>2</sup>, ... } avec

$$\sigma_i^2 = 1/N \cdot \sum_{k=1}^N (x_{ki} - \bar{x}_i)^2 \quad \text{et} \quad \bar{x}_i = 1/N \cdot \sum_{j=1}^N x_{ji} \quad (5.2)$$

N : nombre total de vecteurs  
K : nombre de classes  
x<sub>ji</sub> : i<sup>ème</sup> composante du vecteur J.

\* interclasses : P = diag { ..., 1/λ<sub>i</sub><sup>2</sup>, ... }

$$\lambda_i^2 = 1/K \cdot \sum_{k=1}^K (y_{ki} - \bar{y}_i)^2 \quad \text{et} \quad \bar{y}_i = 1/K \cdot \sum_{j=1}^K x_{ji} \quad (5.3)$$

\* intraclases : P = diag { ..., 1/μ<sub>i</sub><sup>2</sup>, ... }  
avec μ<sub>i</sub><sup>2</sup> = σ<sub>i</sub><sup>2</sup> - λ<sub>i</sub><sup>2</sup> ( th. de Huyghens )

- métrique de 'Mahalanobis' : P = (W)<sup>-1</sup>  
W est la matrice de covariance intraclasse obtenue ici de façon indirecte : W = T - B

T : matrice de covariance totale (t<sub>ij</sub>)

$$t_{ij} = 1/N \cdot \sum_{k=1}^N (x_{ki} - \bar{x}_i) \cdot (x_{kj} - \bar{x}_j) \quad (5.4)$$

et B : matrice de covariance interclasse (b<sub>ij</sub>)

$$b_{ij} = 1/K \cdot \sum_{k=1}^K (y_{ki} - \bar{y}_i) \cdot (y_{kj} - \bar{y}_j) \quad (5.5)$$

- métrique du chi-deux :

$$d(X, Y) = \sum_{j=1}^p 1/m_j \cdot (x^j/x - y^j/y)^2 \quad (5.6)$$

avec m<sub>j</sub> : J<sup>ème</sup> composante du vecteur moyen total

$$\text{et } x = 1/p \cdot \sum_{j=1}^p x^j \quad \text{et} \quad y = 1/p \cdot \sum_{j=1}^p y^j \quad (5.7)$$

distances cepstrales :

\* simple : d(X, Y) = (c<sub>x</sub><sup>0</sup> - c<sub>y</sub><sup>0</sup>)<sup>2</sup> + 2 \cdot \sum\_{i=1}^p (c<sub>x</sub><sup>i</sup> - c<sub>y</sub><sup>i</sup>)<sup>2</sup> (5.8)

\* pondéré : d(X, Y) = \sum\_{i=1}^p I<sub>i</sub> \cdot (c<sub>x</sub><sup>i</sup> - c<sub>y</sub><sup>i</sup>)<sup>2</sup> (5.9)

**VI RESULTATS**

**VI-a Reclassement de la base de donnée :**

Lors de l'impression de cet article, la base de donnée est encore incomplète. Les cinq classes

suivantes sont disponibles : signaux sains, neurogène 3A chronique, 3A subaigu, 3B chronique, 3C subaigu, soit au total 24 signaux. Ils sont tous recueillis sur le muscle Pédieux.

Les tests ont été effectués sur les coefficients autorégressifs du modèle (ak), les coefficients de réflexion (kp) fournis par l'algorithme de Lévinson-Durbin et les coefficients cepstraux déduits du cepstre complexe.

La classification est équivalente avec les deux types d'algorithme de modélisation - Lévinson Durbin et Morf.

L'influence d'une fenêtre de Hamming appliquée aux signaux est négligeable sur la classification.

Les performances obtenues avec chaque métrique sont détaillées dans le tableau ci-après.

p \	EU	C2	VT	VI	MA	CE	DP
5	54	54	58	58	72	-	-
	-	63	63	63	79	63	63
	58	58	58	63	75	-	-
10	54	56	63	67	92	-	-
	-	63	67	71	92	67	71
	58	63	63	63	88	-	-
15	46	67	67	67	93	-	-
	-	67	67	71	96	75	75
	58	63	67	67	93	-	-
20	54	63	67	67	79	-	-
	-	63	71	79	96	75	79
	50	63	67	67	83	-	-
25	58	63	63	67	46	-	-
	-	63	67	79	75	71	79
	58	58	63	63	46	-	-

Pour chaque ordre, la première ligne correspond aux paramètres ak, la seconde aux ck, la troisième aux kp.

- EU : distance euclidienne
- C2 : distance du chi-deux
- VT : variance totale
- VI : variance intraclasse
- MA : mahalanobis
- CE : distance cepstrale
- CP : distance cepstrale pondérée

Tableau n°1 : pourcentage de bien classés sur des signaux recueillis par électrode piqués.

p \	EU	C2	VT	VI	MA	DC	DP
10	54	54	54	58	75	-	-
	-	54	58	58	79	63	63
	58	54	58	58	79	-	-
15	54	58	58	63	75	-	-
	-	58	58	63	79	67	63
	58	58	63	63	75	-	-
20	58	58	58	63	71	-	-
	-	58	58	63	75	67	71
	54	54	58	58	71	-	-

Tableau n°2 : classification des signaux de surface

La distance de Mahalanobis appliquée aux coefficients cepstraux permet d'obtenir les meilleurs résultats. D'autres métriques telles que la distance cepstrale pondérée et la pondération par les variances intraclases ont des performances moindres mais significatives.

Il semble exister un ordre optimal de modélisation - situé autour de la valeur quinze.

On notera les résultats obtenus sur la classification des signaux de surface



## VII CONCLUSION

Ces premiers résultats de classification automatique sur des signaux électromyographiques sont encourageants. A cette étape de notre travail, les méthodes de reconnaissance n'ont pas encore été appliquées à des signaux non inclus dans la base de données, faute d'un nombre suffisant d'enregistrements.

Il nous semble intéressant d'adjoindre au vecteur caractéristique certains paramètres temporels tirés du level-crossing sur le signal ou l'erreur de prédiction, dont le facteur discriminant est dans certains cas évident.

Le processus de reconnaissance a été validé sur des signaux de surface, les recherches seront donc poursuivies dans ce sens.

## BIBLIOGRAPHIE

[1] Agnel J.F. 'Elimination de fouillis radar et détection séquentielle des cibles', thèse de doctorat-ingénieur, ENST, février 85

[2] Calamel J. Michel C. Guiheneuc P. 'Reconnaissance et analyse automatique des PUMs', V<sup>ème</sup> journée d'électromyographie, Nantes, juin 86

[3] Carrault G. Toulouse P. Le Bars R. Pinel J.F. 'Intérêt de l'analyse du signal interférentiel', V<sup>ème</sup> journée d'électromyographie, Nantes, juin 86

[4] Coatrieux J.L. 'Interference electromyogramme processing - Experimental and simulated EMG AR modeling', Electromyogr. clin. Neurophysiol., 1983, 23, p.481-490.

[5] De Luca C.J. 'Physiology and mathematics of myoelectric signals', IEEE trans. BME vol.26, n°6, juin 79

[6] Diday, Lemaire, Pouget, Testu, 'Eléments d'analyse des données', Edition Dunod, p.309-334

[7] Gray A.H. Markel J. 'Distance measure for speech processing', IEEE ASSP-24, n°5, octobre 76

[8] Grenier Y. 'Modélisation et reconnaissance de la parole', dans 'Outils et modèles mathématiques pour l'automatique, l'analyse et le traitement du signal', édition du CNRS tome 2, 1982, p.617-637.

[9] Kay S.M. Marple S.L. 'Spectrum analyse - a modern perspective', Proceeding of the IEEE, vol 69, n°11, nov.81, p.1380-1417.

[10] Lucas M.F. 'Identification d'un modèle statistique de signal d'électromyographie de surface' V<sup>ème</sup> journée d'électromyographie, Nantes, juin 86

[11] Richfield E. Cohen B. 'Review of quantitative and automated needle electromyographic analyses', IEEE Trans.BME, vol.28, n°7, juin 81.

[12] Villemur C. Castanié F. Georgel B. 'Modélisation paramétrique et classification automatique de signaux de forme transitoire', GRETSI juin 87.