

# Classification multivoie de signaux électromyographiques par ondelettes adaptées

Aude MAITROT, Marie-Françoise LUCAS, Christian DONCARLI

Institut de Recherche en Communication et Cybernétique de Nantes (IRCCyN)  
1, rue de la Noë, BP 92101, 44321 Nantes Cedex 03, France  
{Maitrot, Lucas, Doncarli}@irccyn.ec-nantes.fr

**Résumé** – Cet article présente une méthode de classification supervisée de signaux multivoies, basée sur l'utilisation d'ondelettes adaptées aux différentes voies. Pour une voie donnée, l'espace de représentation est construit à partir d'une décomposition dépendant d'une ondelette mère paramétrée, et il est optimisé pour minimiser la probabilité d'erreur de classification estimée sur une population d'apprentissage. La procédure de classification multivoie comporte deux étapes : classification voie par voie, puis fusion de ces décisions en tenant compte des probabilités d'erreur associées. L'application visée concerne la reconnaissance de mouvements à partir de signaux d'électromyographie enregistrés sur 8 voies. Les premiers résultats obtenus montrent le gain apporté par l'optimisation de chaque ondelette mère, puis par la procédure de fusion proposée.

**Abstract** – This paper addresses the supervised classification of multi-channel signals, and presents a method based on channel adapted wavelets. For one channel, the representation space is built from a wavelet decomposition depending on a parameterized mother wavelet, and is optimized by minimizing the misclassification error estimated on the learning set. The multi-channel classification procedure involves two steps : mono-channel classification of each channel, and combination of these decisions thanks to the associated error probabilities. Application concerns movement recognition from electromyographic signals recorded on 8 channels. First results show the benefit brought by each mother wavelet optimization, and by the proposed decision combination procedure.

## 1 Introduction

Les signaux d'électromyographie de surface (EMGs) sont générés par l'activité électrique d'un muscle lors de sa contraction et sont détectés sur la peau au moyen d'électrodes de surface. Parmi les nombreuses applications liées à l'étude de ces signaux, la plupart requièrent une phase de classification : aide au diagnostic médical, reconnaissance de mouvements pour le contrôle de prothèse, classification des troubles neuromusculaires... Nous nous intéressons ici à la classification de mouvements et nous proposons une méthode de classification supervisée à partir de signaux multivoies.

L'utilisation de plusieurs électrodes, réparties sur une zone d'intérêt, génère des signaux multivoies. Les différentes classes de signaux correspondent à différents mouvements, mettant en jeu différents groupes musculaires. Chaque voie correspond à l'enregistrement de l'activité d'un certain ensemble de fibres musculaires différent d'une voie à l'autre. La procédure de classification multivoie doit donc s'effectuer en deux temps : classer voie par voie puis fusionner ces décisions.

Concernant la nature des signaux, chaque voie est constituée par la superposition de trains de formes d'onde, de spectres souvent très similaires. Ces formes d'onde inconnues, mais caractérisant les classes, apparaissent à des instants inconnus distribués aléatoirement au sein d'une même classe. Les techniques classiques de représentation - modèles type AR ou représentations temps-fréquence (à

base de sinus) - sont peu adaptées pour ce type de signaux [1]. L'espace de représentation adéquat doit être sensible aux formes d'onde. Ainsi, les ondelettes, si on les adapte aux signaux, semblent a priori être un bon outil pour classer les EMGs. Pour rendre l'espace de représentation insensible aux instants d'apparition des formes d'onde, on utilise comme descripteurs les marginales de chaque échelle de la transformée discrète en ondelettes dyadiques (DWT : Discrete dyadique Wavelet Transform) du signal.

L'optimisation de l'espace de représentation en fonction des signaux s'effectue voie par voie. Pour chaque voie, nous utilisons une ondelette mère paramétrée que nous optimisons grâce à un critère de sélection défini pour un contexte de classification. À partir de l'ensemble d'apprentissage, nous estimons la probabilité d'erreur de classification, puis minimisons cette probabilité vis-à-vis des paramètres de l'ondelette mère.

Nous disposons donc des procédures de décision (optimisées) voie par voie accompagnées de leur probabilité d'erreur. La fusion de ces procédures fournit la règle de décision multivoie.

Ce papier est organisé de la manière suivante. Nous présentons en Section 2 l'espace de représentation paramétré ainsi que le processus de décision dans cet espace, puis son optimisation, pour une voie. Nous montrons ensuite comment pondérer les décisions issues de chaque voie, pour fournir la décision finale. Nous présentons des résultats sur signaux réels dans une quatrième partie.

## 2 Classification supervisée monovoie 2.2 Règle de décision

Dans cette section, nous traitons la voie  $x_k$  du signal comportant  $M$  voies,  $x = \{x_1[n], \dots, x_k[n], \dots, x_M[n]\}$ ,  $n = 1, \dots, N$ . Pour ne pas alourdir les notations, nous omettons l'indice  $k$  se rapportant à la voie concernée.

### 2.1 Choix d'un espace de représentation paramétré

Nous présentons la méthode utilisée pour paramétrer l'ondelette mère, dans le cadre de l'analyse multirésolution (AMR). On note  $\theta$  le vecteur des paramètres.

Dans le cas d'ondelettes orthogonales, l'ondelette mère définissant la DWT est caractérisée par le filtre échelle  $h$ . Le filtre ondelette  $g$  se déduit directement de  $h$  :  $g[n] = (-1)^n h[1-n]$ . Pour générer une AMR,  $h$  doit remplir certaines conditions. Dans le cas d'un filtre à réponse impulsionnelle finie de longueur  $L_h$ , il doit satisfaire  $L_h/2 + 1$  conditions suffisantes pour assurer l'existence et l'orthogonalité de la fonction échelle et de l'ondelette mère [2, 3]. Il reste  $L_h/2 - 1$  degrés de liberté disponibles pour construire le filtre  $h$ .

La paramétrisation de Pollen décrite dans [4] permet de construire  $h$  en fonction de  $L_h/2 - 1$  nouveaux paramètres libres. Par exemple, si  $L_h = 6$ ,  $\theta$  a 2 composantes,  $\theta = [\alpha, \beta]$ , et  $h$  s'écrit :

$$\begin{aligned} i = 0, 1 : h[i] &= 4\sqrt{2}[(1 + (-1)^i \cos \alpha + \sin \alpha)(1 - (-1)^i \cos \beta - \sin \beta) + (-1)^i 2 \sin \beta \cos \alpha] \\ i = 2, 3 : h[i] &= 2\sqrt{2}[1 + \cos(\alpha - \beta) + (-1)^i \sin(\alpha - \beta)] \\ i = 4, 5 : h[i] &= 1/\sqrt{2} - h(i-4) - h(i-2) \end{aligned}$$

Pour des longueurs différentes de  $L_h$ , les expressions de  $h$  sont données dans [5, 6].

Les  $N$  coefficients  $\{d_x(s, u)\}_{s=1, \dots, L; u=0, \dots, N/2^s - 1}$ , de la décomposition par DWT d'un signal discret  $x$  de longueur  $N$  sur  $L$  niveaux sont calculés grâce à l'algorithme de Mallat [7].

Afin de rendre l'espace de représentation insensible aux instants d'apparition des formes d'onde composant les signaux, nous utilisons comme descripteurs les marginales normalisées de chacun des niveaux de la décomposition définies par :

$$m_x(s) = \sum_{u=0}^{N/2^s - 1} c_x(s, u), \quad s = 1, \dots, L \quad (1)$$

$$c_x(s, u) = \frac{|d_x(s, u)|}{\sum_{s=1}^L \sum_{u=0}^{N/2^s - 1} |d_x(s, u)|} \quad (2)$$

où  $L$  est le niveau maximum de décomposition ( $L = \lceil \log_2(N) \rceil$ ). Les descripteurs représentant le signal  $x$  sont les composantes du vecteur  $M_x = [m_x(1), \dots, m_x(L)]$ . Ainsi, le vecteur  $M_x$  représente le signal par la contribution de chaque échelle calculée avec une onde analysante qui peut être adaptée aux signaux. Au lieu de permettre ce choix, une transformée de Fourier impose les sinus.

Dans cet espace de représentation, comme la normalisation (2) assure :

$$\begin{cases} m_x(s) \geq 0, & s = 1, \dots, L \\ \sum_{s=1}^L m_x(s) = 1 \end{cases}$$

nous pouvons nous servir de la distance de Kullback classiquement utilisée pour la comparaison entre distributions :

$$D(M_x, M_y) = \sum_{s=1}^L \left[ m_x(s) \log \frac{m_x(s)}{m_y(s)} + m_y(s) \log \frac{m_y(s)}{m_x(s)} \right] \quad (3)$$

Dans l'espace (1) muni de la distance (3), la règle de décision utilisée est la règle du plus proche représentant. Soit  $\{\omega_i\}_{i=1, \dots, N_c}$ , l'ensemble des  $N_c$  classes de signaux. Le représentant, noté  $R_i$ , de la classe  $\omega_i$  est le vecteur moyen des  $M_x$ , calculé sur les signaux d'apprentissage de la classe  $\omega_i$ . La décision " $x$  appartient à  $\omega_i$ " est obtenue pour  $i$  tel que :

$$i = \underset{j}{\text{Argmin}} D(M_x, R_j) \quad (4)$$

### 2.3 Critère de qualité

L'optimisation de l'espace de représentation nécessite la définition d'un critère adapté à l'objectif de classification. Nous utilisons comme critère la probabilité d'erreur de classification, à minimiser. Cette probabilité,  $P_e^\theta$ , est calculée de manière empirique sur l'ensemble d'apprentissage, avec régularisation par la méthode du leave-one-out. Soit  $\Omega_i$  l'ensemble des signaux d'apprentissage de la classe  $\omega_i$ . Pour chaque  $x$  de  $\Omega_i$ , nous appliquons la règle du plus proche représentant (4), où les représentants des classes  $\omega_j, j \neq i$ , sont calculés comme en 2.1, mais où le représentant de la classe  $\omega_i$  est calculé sans prendre en compte la décomposition de  $x$ . Une erreur de classification est comptabilisée lorsque le signal  $x$  de  $\Omega_i$  est désigné comme appartenant à  $\omega_j, j \neq i$ . La probabilité d'erreur pour la classe  $\omega_i$ ,  $P_e^\theta(\omega_i)$ , est alors le nombre d'erreurs divisé par le nombre de signaux d'apprentissage de  $\Omega_i$ .

La probabilité d'erreur globale,  $P_e^\theta$ , est la moyenne des probabilités d'erreur pour chaque classe :

$$P_e^\theta = \frac{1}{N_c} \sum_{i=1}^{N_c} P_e^\theta(\omega_i) \quad (5)$$

A l'issue de cette étape, on dispose donc pour chaque voie  $x_k$  d'une ondelette optimale définie par  $\hat{\theta}_k$ , et d'une procédure de décision accompagnée de sa probabilité d'erreur calculée selon (5) et notée  $P_e^{\hat{\theta}_k}$ .

## 3 Classification supervisée multivoie

Cette section présente la méthode permettant de prendre en compte les décisions issues des  $M$  classifications monovoies. Nous utilisons le formalisme suivant. La voie  $x_k$  a fourni la décision notée  $\alpha_k$ . Les  $\alpha_k$  prennent leur valeur dans l'ensemble  $\{\Delta_j\}_{j=1, \dots, N_c}$  où  $\Delta_j$  est un vecteur de taille  $N_c$  et  $\Delta_j = [0, \dots, 0, 1, 0, \dots, 0]$ , avec le 1 en jème

position. Ainsi la décision " $\alpha_k$  appartient à  $\omega_i$ " est notée  $\alpha_k = \Delta_j$ .

Disposant de la probabilité d'erreur de classification pour chaque classe et chaque voie  $P_e^{\hat{\theta}_k}(\omega_i)$ , on peut alors exprimer la distribution des  $\alpha_k$  conditionnellement à  $\omega_j$  :

$$p(\alpha_k/\omega_j) = 1 - P_e^{\hat{\theta}_k}(\omega_j) \quad \text{si } \alpha_k = \Delta_j \quad (6)$$

$$= P_e^{\hat{\theta}_k}(\omega_j) \quad \text{si } \alpha_k \neq \Delta_j \quad (7)$$

qui peut s'écrire :

$$p(\alpha_k/\omega_j) = (1 - P_e^{\hat{\theta}_k}(\omega_j))^{1 - \frac{1}{\sqrt{2}}\|\alpha_k - \Delta_j\|} (P_e^{\hat{\theta}_k}(\omega_j))^{\frac{1}{\sqrt{2}}\|\alpha_k - \Delta_j\|} \quad (8)$$

Le vecteur contenant l'ensemble des  $M$  décisions pour le signal  $x$  est noté  $\alpha = [\alpha_1, \dots, \alpha_M]$ . En supposant les voies indépendantes, on peut écrire :

$$p(\alpha/\omega_j) = \prod_{k=1}^M p(\alpha_k/\omega_j) \quad (9)$$

La règle de décision multivoie correspond alors au choix de la classe dont la probabilité conditionnelle est maximale, soit :

$$"x \text{ appartient à } \omega_i" \text{ pour } i = \underset{j}{\text{Argmax}} p(\alpha/\omega_j) \quad (10)$$

Notons que cette procédure, entièrement déterminée sur l'ensemble d'apprentissage, n'est possible que parce que l'on dispose pour chaque voie de la probabilité d'erreur.

## 4 Résultats sur signaux réels

### 4.1 Protocole expérimental et résultats

A titre illustratif, nous avons appliqué cette méthode sur des signaux réels d'électromyographie. Ces signaux ont été enregistrés au moyen de  $M = 8$  électrodes, placées autour de l'avant-bras, pour  $N_c = 6$  classes de mouvement : flexion et extension, pronation et supination du poignet, ouverture et fermeture de la main. Le filtre  $h$  optimisé est de longueur  $L_h = 6$ . Il y a donc 2 paramètres à optimiser. Techniquement, la recherche de l'optimum est faite sur une grille d'échantillonnage des paramètres.

Les signaux considérés sont de 512 points échantillonnés à 2048 Hz. La population d'apprentissage est constituée de 60 signaux par classe et on dispose d'un ensemble test distinct de 120 signaux par classe. Nous obtenons les résultats résumés dans le tableau 1. Nous avons également testé la classification en utilisant une ondelette standard (Daubechies 3, dont le filtre d'échelle est de longueur 6). Enfin, la classification spectrale correspond à la classification des spectres de puissance des signaux en utilisant la distance de Kullback. Les résultats présentés correspondent à un seul individu.

### 4.2 Discussion

Ces résultats montrent plusieurs choses. Tout d'abord, l'optimisation d'une ondelette mère par voie permet d'obtenir de meilleures performances que l'utilisation d'une ondelette standard commune à toutes les voies (d'autres

TAB. 1 – Résultats de la classification multivoie sur 6 classes, avec 60 signaux (512 points) par classe d'apprentissage et 120 par classe de test. Le critère correspond à la probabilité d'erreur de classification monovoie estimée sur l'ensemble d'apprentissage. "Tx MC" signifie le taux de mal classés calculé sur l'ensemble test.

Filtre $h$ optimisé par voie								
Voies	1	2	3	4	5	6	7	8
Critère	0.44	0.47	0.40	0.46	0.49	0.54	0.36	0.36
Tx MC (%)	60	70	58	70	71	86	51	66
Tx MC (%)	21							
Ondelette Daubechies 3								
Critère	0.48	0.52	0.42	0.50	0.56	0.60	0.40	0.47
Tx MC (%)	70	76	56	69	89	88	54	86
Tx MC (%)	45							
Classification spectrale								
Tx MC (%)	48							

ondelettes standard donnent des résultats similaires). La classification basée sur une représentation spectrale des signaux est également très nettement moins performante.

Les résultats monovoies montrent qu'une seule voie ne permet pas de classer correctement un signal parmi 6 classes, ce qui semble normal. En effet, une électrode peut être placée sur une zone ne permettant pas de distinguer plusieurs des 6 mouvements effectués. En revanche, la fusion des décisions issues de chaque voie, utilisant l'information de la probabilité d'erreur, améliore grandement la capacité de classification.

Les 8 ondelettes optimales correspondant aux 8 voies sont représentées figure 1. Bien que nous n'ayons pas imposé de contrainte de régularité, 7 ondelettes sur 8 sont assez régulières. L'ondelette de la voie 7 n'est pas régulière, mais le critère estimé sur l'apprentissage (0.36), et le taux de mal classés obtenu sur la population test (51 %) pour cette voie montrent que c'est pourtant l'ondelette assurant le meilleur résultat. Ceci illustre l'intérêt de l'adaptation de l'ondelette aux signaux et au contexte : le but étant de classer, il faut choisir l'ondelette permettant de classer le mieux possible.

Enfin, la régularisation du critère par leave-one-out favorise la cohérence entre les performances prévues (critère calculé sur l'apprentissage) et celles obtenues sur la population test. Plus précisément, les taux de mal classés, en monovoie, obtenus sur l'ensemble test sont supérieurs aux prévisions représentées par les critères, quelle que soit l'ondelette utilisée (optimisée ou standard). Des écarts apparaissent déjà dans le cas d'une classification bichasse monovoie sur signaux simulés [8], il est donc d'autant plus normal de retrouver cette tendance dans le cas 6 classes sur signaux réels : la population d'apprentissage ne peut représenter de manière exhaustive toutes les caractéristiques des signaux susceptibles d'influencer la classification. Cependant, on remarque que ces écarts sont cohé-

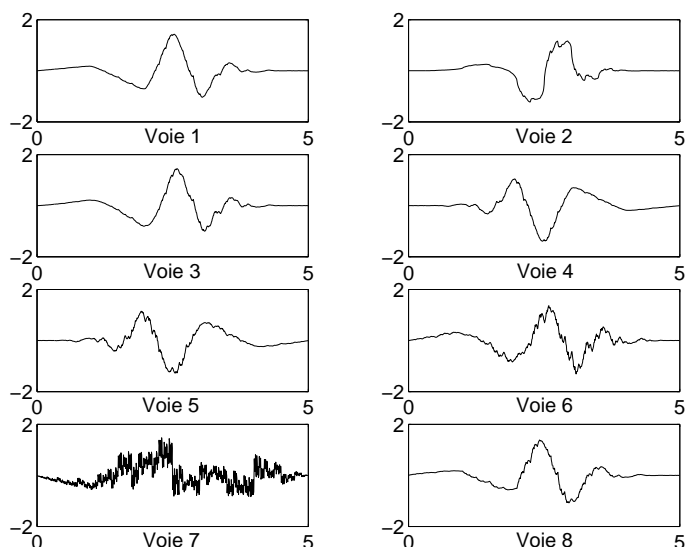


FIG. 1 – Ondelettes optimales de chaque voie, optimisées sur l’ensemble d’apprentissage par minimisation du critère de probabilité d’erreur.

rents : pour 6 voies sur 8 (1,2,5,6,7,8), les résultats obtenus sur la population test avec les ondelettes optimisées sont meilleurs que ceux obtenus avec l’ondelette standard, ce qui explique la très nette amélioration des performances du classifieur multivoie.

Remarque : dans ce papier, on n’utilise à dessein pas d’autre information que celle apportée par les marginales de la DWT, afin de montrer l’apport de l’optimisation de l’espace de représentation (*via* l’ondelette).

## 5 Conclusion

Nous avons proposé une méthode de classification supervisée de signaux multivoies, basée sur l’utilisation d’ondelettes adaptées aux différentes voies. Elle s’appuie sur l’optimisation d’une ondelette mère par voie, en minimisant un critère de probabilité d’erreur de classification, afin de construire un espace de représentation directement adapté aux signaux à traiter. La règle du plus proche représentant définit une règle de décision pour chaque voie, et le critère choisi permet de combiner ces procédures monovoies pour fournir une décision finale. Des tests sur signaux réels d’électromyographie montrent que cette optimisation, ainsi que la procédure de fusion des décisions monovoies, améliorent les performances de classification par rapport à l’utilisation d’ondelettes classiques. Cette méthode pourrait s’appliquer sur d’autres types de signaux où la représentation par décomposition en ondelettes est pertinente.

D’autres types de paramétrisation de l’ondelette sont possibles. Les combinaisons linéaires d’ondelettes ont été étudiées dans [8] et on peut également envisager une construction par lifting. Concernant la procédure de classification, l’utilisation des SVM (Support Vector Machines) permettrait d’optimiser également le classifieur lui-même.

## Références

- [1] K. Englehart, B. Hudgins, P. Parker, M. Stevenson. *Classification of the myoelectric signal using time-frequency based representations*. Med. Eng. Phys., 21, pp.431-438, 1999.
- [2] W. Lawton. *Tight frames of compactly supported affine wavelets*. Journal of Mathematical Physics, vol. 31(8), pp. 1898-1901, 1990.
- [3] W. Lawton. *Necessary and sufficient conditions for constructing orthonormal wavelet bases*. Journal of Mathematical Physics, vol. 32(6), pp. 1440-1443, 1991.
- [4] D. Pollen.  *$SU_I(2, F[z, 1/z])$  for  $F$  a subfield of  $C$* . Comm. Pure Appl. Math., vol. 45, p.485, 1990.
- [5] C. S. Burrus, R. A. Gopinath, and H. Guo. *Introduction to Wavelets and Wavelet Transforms - A Primer*. Prentice Hall, 1997, pp. 53-66.
- [6] I. W. Selesnick. *Maple and the parameterization of orthogonal wavelet bases*. Online, Oct. 1997. Available : <http://taco.poly.edu/selesi/theta2h/>
- [7] S. Mallat. *A theory for multiresolution signal decomposition : the wavelet representation*. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. II, no. 7, pp. 674-693, 1989.
- [8] A. Maitrot, M.F. Lucas, C. Doncarli. *Design of wavelets adapted to signals and application*. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Philadelphie, USA, mars 2005.