

Projet :
Traitement de signal biomédical pour la
caractérisation de l'encodage mémoriel

1 Introduction

1.1 Objectif

Le traitement du signal est une discipline trouvant des applications dans de nombreux domaines :

- traitement du son, de l'image et de la vidéo,
- traitement du signal radar et sonar, de l'analyse de la forme d'onde au pistage d'objets mobiles,
- localisation et navigation GPS,
- communications numériques,
- santé au travers du traitement de signaux biomédicaux (EEG, ECG, EMG, etc.).

Dans ce dernier cas, le traitement du signal peut intervenir à différents niveaux. Il permet d'accompagner le diagnostic d'un médecin ou d'un spécialiste. Il peut être exploité avec des objets connectés grand-public pour obtenir des informations sur des performances sportives. Quant au suivi, également appelé monitoring, il vise à détecter des états physiologiques pour une personne, comme des niveaux extrêmes de charge mentale, une situation de somnolence, de stress ou d'inconscience ou encore un risque d'accident vasculaire cérébral (AVC). La démarche est alors la suivante : comme certaines propriétés des signaux biomédicaux constituent des marqueurs de ces états, également appelés signatures, il s'agit de les extraire.

L'objectif de ce projet est de vous familiariser avec ces signaux biomédicaux et leur étude, ainsi que de mettre en pratique vos connaissances en traitement du signal et processus aléatoires sur une application d'évaluation de la charge mentale d'une personne.

1.2 Évaluation

L'évaluation du travail repose sur un rapport et une note de travail continu. Le travail peut être mené en binôme. Le rapport d'une dizaine de pages maximum doit être dactylographié (sous Word ou latex).

Sous Word, les équations doivent être générées avec l'éditeur d'équations et numérotées. Les commentaires doivent être pertinents et tout résultat justifié. Les programmes développés sous *Matlab* peuvent être mis en annexe. Des références bibliographiques peuvent être introduites et seront regroupées dans une section en fin de rapport.

Le rapport devra également comporter un bilan de l'organisation du projet, qui comprendra en particulier la liste des tâches entreprises, avec une évaluation du temps passé par chacun lors de chaque séance et entre les séances.

Des interfaces *Matlab* peuvent être mises aussi en œuvre et seront comptées comme un bonus dans l'évaluation. Dans ce cas, l'envoi des codes par mail sera nécessaire. La qualité du code produit et une bonne organisation seront des points pris en compte pour la note de travail continu.

Une attention particulière devra être apportée à l'orthographe et à la rédaction de l'introduction et de la conclusion. L'introduction doit amener le sujet que vous abordez et doit poser le problème. Elle doit enfin présenter le plan de votre rapport. La conclusion doit dresser un bilan technique du travail mené. Tout plagiat sera sanctionné, Cf. règlement pédagogique de l'école.

1.3 Contacts

Ce projet est encadré par :
B. Berthelot. bastien.berthelot@fr.thalesgroup.com.

1.4 Analyse fréquentielle d'un signal

1.4.1 Introduction

De manière générale, lors de l'analyse de signaux, la première étape consiste à observer le signal dans le domaine temporel et dans celui fréquentiel. Cela doit être un mécanisme pour le traiteur du signal. Si le signal est non-stationnaire, on préfère une analyse temps-fréquence à une analyse fréquentielle. Cette dernière permet de mettre en valeur la manière dont la richesse fréquentielle du signal évolue au cours du temps. Avant de pouvoir traiter de signaux réels, il est essentiel de pouvoir avoir confiance dans les outils développés. Pour cela, il est classique de les utiliser sur des signaux connus. Par exemple, on connaît l'évolution fréquentielle d'un bruit blanc gaussien centré au cours du temps.

1.4.2 Questions et actions à mener

Fournir les représentations temporelles et temps-fréquence (à l'aide d'un spectrogramme) de deux signaux de taille $M = 2001$. Le premier est un bruit blanc gaussien moyenne nulle et de variance 1. Le second est un bruit blanc de moyenne $\mu = 7$ et de variance $\sigma^2 = 9$. Pour ce faire, on utilisera la fonction **randn** de *Matlab*. Pour le spectrogramme, justifier vos choix sur le taux de recouvrement des segments, la fenêtre choisie et le nombre du points envisagés pour le calcul des transformées de Fourier.

On opéra pour une représentation d'un signal sur une figure comprenant deux zones situées l'une au dessus de l'autre avec la fonction *subplot* de *Matlab*. La représentation temporelle sera située au dessus de la représentation temps-fréquence. Le temps est représenté sur l'axe des abscisses dans les deux cas. Une correspondance des temps entre les deux représentations est demandée.

Commenter les figures obtenues. Que pouvez vous dire de la stationnarité de ces signaux ? Justifiez votre réponse. Que pouvez-vous dire de la puissance portée par ces signaux sur l'ensemble des fréquence ? Quelle action importante déduisez-vous de ces figures lors d'une étude temps-fréquence de signaux ? A l'aide de la fonction *load*, chargez le fichier fourni et représentez un signal EEG dans le domaine temporel et temps-fréquence. Commentez.

2 Régularité d'un signal

2.1 Généralités

Parmi les caractéristiques de base d'un signal, il peut être intéressant d'évaluer la puissance véhiculée par le signal dans certaines bandes de fréquence, nombre de passage par 0, etc. D'autres caractéristiques comme la régularité d'un signal peuvent être d'intérêt. Dans la suite, nous présentons deux méthodes pour la caractériser. La première est une variante de la méthode nommée *detrended fluctuations analysis* (DFA), qui fut développée dans les années 90. Il s'agit de la méthode DFA d'ordre 3. La seconde appelée *detrended moving average* (DMA) fut proposée quelques années plus tard.

Dans la suite, nous présentons les différentes étapes des deux méthodes.

2.2 Detrended fluctuation analysis d'ordre 3 (DFA3)

2.2.1 Mise en œuvre

Soient M échantillons consécutifs d'un signal notés $\{y(m)\}_{m=1,\dots,M}$. Le DFA d'ordre 3 opère en 5 étapes :

1. Le profil, noté $y_{int}(m)$, est tout d'abord défini comme suit :

$$y_{int}(m) = \sum_{i=1}^m (y(i) - \mu_y) \quad (1)$$

où $\mu_y = \frac{1}{M} \sum_{m=1}^M y(m)$ désigne la moyenne temporelle du signal y .

2. Le profil y_{int} est ensuite décomposé en L segments de taille N , notés $\{y_{int,l}(n)\}_{l=1,\dots,L}$ avec $n \in [1; N]$. Le n^{eme} échantillon du l^{eme} segment correspond au $m^{eme} = [(l-1)N + n]^{eme}$ échantillon de y_{int} . La l^{eme} tendance locale, c'est-à-dire la tendance $x_l(n)$ du l^{eme} segment $y_{int,l}(n)$, est modélisée par une régression d'ordre 3 $\forall l \in [1; L], \forall n \in [1; N]$, comme suit :

$$x_l(n) = a_{l,3}[(l-1)N + n]^3 + a_{l,2}[(l-1)N + n]^2 + a_{l,1}[(l-1)N + n] + a_{l,0} \quad (2)$$

Comme le nombre d'échantillons n'est pas nécessairement un multiple de N , on ne considère que LN premiers échantillons du profil.

Puis $\forall l \in [1; L]$, les paramètres $a_{l,3}$, $a_{l,2}$, $a_{l,1}$ et $a_{l,0}$ sont estimés en minimisant le critère suivant :

$$J_{DFA}(a_{l,3}, a_{l,2}, a_{l,1}, a_{l,0}) = \sum_{n=1}^N (y_{int}((l-1)N + n) - x_l(n))^2 \quad (3)$$

Afin d'estimer ces L tendances locales sous *Matlab*, utilisez la fonction *polyfit*.

3. La tendance globale du profil se déduit en concaténant les tendances locales. Cette dernière est soustraite du profil. La racine carrée de la puissance du résidu se déduit alors comme suit :

$$F(N) = \sqrt{\frac{1}{LN} \sum_{l=1}^L \sum_{n=1}^N (y_{int}((l-1)N + n) - x_l(n))^2} \quad (4)$$

4. Les étapes 2 et 3 sont répétées pour différentes valeurs de N . On représente alors les résultats obtenus dans le plan $(\log(F(N)), \log(N))$. Les valeurs de N considérées sont les suivantes :

$$N = [13, 17, 21, 27, 35, 47, 59, 77, 99, 135, 189] \quad (5)$$

D'un point de vue théorique, on peut montrer que :

$$F(n) \propto n^\alpha \quad (6)$$

Pour cette raison, la dernière étape est de rechercher la droite de régression qui est la plus adaptée à la représentation log-log du nuage de points. La pente α de cette droite est la valeur recherchée ; elle caractérise la régularité du signal.

2.2.2 Actions à mener

Fournir des illustrations de chacune des étapes de la méthode, à savoir la représentation du profil et des tendances locales ainsi que l'identification des instants auxquels chaque segment débute ou se termine à l'aide de barres verticales. Pour cela, utilisez un bruit blanc que vous aurez généré.

2.2.3 Test sur signaux synthétiques de régularité connue

Après avoir codé la méthode DFA3 sous *Matlab*, la validation du code se fonde sur l'analyse de signaux synthétiques fournis dont la régularité est connue. Il s'agit d'une démarche classique pour tout algorithme de traitement du signal : tester le code sous *Matlab* dans un cadre où les paramètres sont connus. dans le cas d'un bruit blanc, la régularité théorique est de $\alpha = 0.5$.

Générez cinquante réalisations d'un bruit blanc gaussien centré. Estimez la régularité de ces signaux. Fournissez la moyenne et variance des valeurs obtenues. Répétez cette étape pour cinquante réalisations d'un bruit blanc de moyenne non nulle. Concluez.

2.3 Detrending Moving Average (DMA)

2.3.1 Mise en œuvre

De manière similaire au DFA, partons du profil $y_{int}(m)$. Le DMA opère ainsi :

1. Le profil y_{int} est filtré à l'aide d'un filtre linéaire. La tendance est le signal obtenu en sortie du filtre.
2. La tendance déduite de l'étape précédente est soustraite au profil. La racine carré de la puissance du résidu est alors déduite.
3. Les étapes 2 et 3 sont répétées pour différentes valeurs de N . Comme précédemment, puisque $F(n) \propto n^\alpha$, $\log(F(N))$ est représenté en fonction de $\log(N)$. La dernière étape est le calcul de la pente de la droite de régression du nuage de points.

Dans le cadre de ce projet, on se propose d'appliquer la méthode appelée "backward moving average" pour laquelle la réponse impulsionnelle du filtre numérique, notée $h_{DMA}(n)$ et de longueur N , est définie comme suit :

$$h_{DMA}(n) = \frac{1}{N} \text{ for } n = 0, \dots, N - 1. \quad (7)$$

2.3.2 Questions et actions à mener

Avant de programmer cette approche, nous vous suggérons de répondre aux questions suivantes afin de mieux appréhender la programmation de cette méthode sous *Matlab*.

Veuillez noter que les réponses devront apparaître dans votre rapport sous la forme d'un paragraphe rédigé et non d'une succession de réponses aux questions posées.

1. Caractériser le filtre utilisé : Est-il RIF ou RII ? Causal ? Stable ?
2. Quelle est sa fonction de transfert $H(z)$ du filtre. Donner la valeurs des pôles et des zéros. En déduire la réponse en fréquence du filtre en fonction de N . Commenter.

3. Le filtre est-il à phase linéaire ? Si oui, quel est le retard de groupe ? Cette propriété est-elle importante dans ce cadre d'étude ? Quelle est la conséquence sur la tendance obtenue par rapport au profil du signal ?
4. Quelle est la nature du filtre ? Passe-haut ? Passe-bas ? Réjecteur de Bande ? Commenter en fonction des valeurs de N .

Concernant les actions à mener, il s'agit de :

1. Fournir des illustrations de chacune des étapes de la méthode.
2. Fournir les figures illustrant les réponses fournies obtenues.

2.3.3 Test sur signaux synthétiques de régularité connue

Après avoir codé la méthode DMA sous *Matlab*, la validation du code se fonde sur l'analyse de signaux synthétiques connus. Répétez la méthode décrite en 2.2.3, en estimant la régularité avec la méthode DMA. Concluez.

2.4 Mise en oeuvre sur signaux physiologiques

Une fois les deux approches validées sur les signaux synthétiques, il s'agit de les tester sur des signaux réels fournis.

En particulier, on se propose d'étudier le lien entre la régularité d'un signal EEG et la profondeur d'encodage mémoriel d'un utilisateur. Les signaux fournis sont des signaux EEG issus de deux électrodes. Sept sujets ont effectué les tests dans des conditions¹ donnant *a priori* lieu à deux niveaux de profondeur d'encodage mémoriel : la première phase correspond à une profondeur dite "faible" et la seconde à une profondeur "élevée". Vingt-huit signaux sont donc disponibles et rassemblés dans le fichier *dataEEG2020.mat*. La fonction *load* permet de charger ce fichier. Les signaux sont étiquetés : chacun est associé à une électrode, un sujet et un état. Pour avoir accès au signal issu de l'électrode e , pour la phase p et le sujet s , on utilise la commande : $y = \text{cell2mat}(\text{dataEEG2020e}'e'(p,s))$

1. Estimez la régularité des signaux issus de la première électrode (e1) par la méthode DFA3. Pour le calcul du DFA, prenez les valeurs de N suivantes :

$$N = [13, 17, 21, 27, 35, 47, 59, 77, 99, 135, 189] \quad (8)$$

Analysez les résultats obtenus pour ces valeurs de N . Rassemblez les valeurs de α obtenues dans un tableau. .

2. Effectuer la même démarche pour la seconde électrode. Commentez.
3. Effectuez les deux étapes précédentes avec la méthode DMA. Commentez.
4. Quelle méthode, appliquée à quelle électrode, vous semble la plus pertinente ? Justifiez à l'aide d'une analyse ANOVA.

¹ces tests ont été mis en oeuvre avec le concours de psychologues