

**Projet signal**  
**Filière Télécommunications, 2ème année, 2019-2020**  
**Rehaussement d'un signal de parole monovoie**  
**par décomposition en valeurs singulières**

## **1 Introduction**

### **1.1 Objectif**

Le traitement du signal permet non seulement de manipuler du son, de la parole et de l'audio, mais aussi :

- ✓ d'analyser des signaux biomédicaux,
- ✓ d'aborder des problématiques dans le secteur de l'aéronautique liées aux signaux sonar et radar, à la localisation et à la navigation GPS, à la poursuite de missiles,
- ✓ d'étudier des systèmes de communications mobiles,
- ✓ de mettre en œuvre des systèmes combinant différents capteurs, etc.

Il est à noter que l'extraction de signatures joue aussi un rôle important dans les approches par machine learning.

Le traitement de la parole est l'une des applications les plus classiques du traitement du signal ; il ne requiert pas trop d'expertise *a priori*, mais reste complexe à maîtriser. L'objectif de ce projet en traitement du signal est de vous familiariser avec ces signaux de parole et de mettre en pratique les enseignements fondamentaux de traitement du signal vus en première année et ce premier semestre, notamment en cours de « processus aléatoires », de « signal continu », de « traitement numérique du signal » et de « filtres numériques et estimation ».

Ce projet vous permet de renforcer vos compétences en traitement du signal en vue d'une découverte de cette thématique ou d'une orientation envisagée pour l'option de 3<sup>ème</sup> année I2SC.

Le projet sera effectué en binôme pendant quatre à cinq séances de durée différente selon les groupes.

Le projet sera mené sous *Matlab*.

### **1.2 Evaluation**

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 *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** et du déroulement 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 poser le problème. Elle doit enfin présenter le plan de votre rapport. La conclusion doit dresser un bilan technique du travail mené.

Un entretien oral lors de la dernière séance sera pris en compte dans la note. Tout plagiat sera sanctionné, Cf. règlement pédagogique de l'école.

### **1.3 Contacts**

Ce projet est encadré par E. Grivel. [eric.grivel@enseirb-matmeca.fr](mailto:eric.grivel@enseirb-matmeca.fr)

## **2 Projet à mener**

### **2.1 Introduction**

Ce projet est consacré au **rehaussement de la parole à partir d'un unique microphone fondé sur une décomposition en valeurs propres des matrices de données bruitées.**

Pour évaluer sa pertinence, on dispose d'un signal non bruité que l'on perturbe par un bruit synthétique ; puis on le rehausse.

Remarque : du point de vue méthodologique, il est préférable d'avancer étape par étape dans un projet de traitement du signal. Pour commencer, on se place dans un cas d'école où le signal de parole non bruité est disponible ; il sert ainsi de référence. Cela permet de valider une méthode (ici celle de rehaussement) en analysant ses limites, ses avantages et ses inconvénients. Une fois ce cas d'école mené, les traiteurs du signal se placent dans des hypothèses de plus en plus réalistes jusqu'à traiter le cas où les données ne sont plus synthétiques, mais réelles.

Le bruit est supposé aléatoire de nature Gaussienne, centré, blanc et stationnaire. C'est le cas le plus « simple » du fait des propriétés de corrélation et spectrales. On pourrait ensuite traiter le cas d'un bruit synthétique coloré stationnaire, celui d'un bruit coloré synthétique non-stationnaire et enfin le cas d'un bruit réel.

Afin d'analyser subjectivement la qualité du signal rehaussé, on peut écouter les signaux et donner un avis sur la qualité du rehaussement. Cela donne lieu à des campagnes d'écoute et d'évaluation connues sous le nom de « Mean Opinion Scores ». Pour se faire aussi un avis, on peut représenter les signaux non bruité, bruité et rehaussé dans le domaine temporel et dans le domaine fréquentiel. Comme le signal de parole est non-stationnaire, une analyse temps-fréquence du signal est nécessaire. La méthode la plus simple est alors le spectrogramme.

Afin d'analyser objectivement la qualité du rehaussement, on utilise le gain en rapport signal à bruit, c'est-à-dire la différence en dB entre les rapports signal à bruit avant et après traitement. Cette notion a déjà été introduite en cours puis revue en TD.

Remarque 2 : ce projet permet de faire la synthèse de plusieurs notions :

- ✓ génération et caractérisation de signaux synthétiques ;
- ✓ intérêt du fenêtrage et gestion de ce traitement ;
- ✓ processus aléatoires.

## 2.2 ***Préliminaire 1 : bruiteur un signal de parole selon un rapport signal à bruit donné***

- ✓ Générer un bruit blanc de moyenne nulle et de variance  $\sigma^2$  ; représenter sa fonction d'autocorrélation théorique et celle estimée à partir de différents estimateurs dont on rappellera les définitions et les conditions d'utilisation. Représenter le spectre de puissance d'une réalisation de ce bruit blanc ; le comparer à la densité spectrale de puissance du bruit. Commenter.
- ✓ Qu'appelle-t-on *Périodogramme de Daniell*, *Périodogramme de Bartlett* et *Périodogramme de Welch* ? Quels sont leurs intérêts ? Faire des recherches bibliographiques et programmer ces périodogrammes sous *Matlab* ; les comparer au spectre de puissance et à la densité spectrale de puissance dans le cas d'un bruit blanc.
- ✓ Qu'appelle-t-on *Corrélogramme* ? Faire des recherches bibliographiques et le programmer sous *matlab*. Le comparer au spectre de puissance et à la densité spectrale de puissance dans le cas d'un bruit blanc.

## 2.3 ***Préliminaire 2 : bruiteur un signal de parole selon un rapport signal à bruit donné***

- ✓ Mettre en œuvre un traitement visant à bruiteur le signal de parole avec un bruit additif pour un rapport signal à bruit (RSB) donné.  
Développer une fonction *Matlab*. On utilisera la fonction *randn*. On définit le RSB en fonction des échantillons du signal et du bruit additif. Les cas suivants sont à tester : 5 dB, 10 dB et 15 dB.
- ✓ Fournir, sur une même figure, la représentation temporelle et le spectrogramme du signal de parole original à l'aide des fonctions *spectrogram* et *subplot*, puis sur une seconde figure la représentation temporelle et le spectrogramme du signal de parole bruité. Faire en sorte que les axes temporels coïncident pour la représentation temporelle et le spectrogramme. Le temps définit l'axe des abscisses. Il est donc exprimé en secondes.
- ✓ Dans le rapport, commenter l'allure des spectrogrammes.

## 2.4 ***Préambule à la méthode de traitement d'un signal de parole bruité par un bruit blanc gaussien : procédure dite d'addition-recouvrement***

Le signal de parole étant non-stationnaire, on peut être amené quand le bruit est aléatoire et/ou non-stationnaire à effectuer un traitement « trame par trame », à l'aide d'un fenêtrage, pour se placer dans des hypothèses de quasi-stationnarité. Mettre en place une procédure permettant :

- ✓ de décomposer le signal en trames avec un recouvrement de 50% ou sans recouvrement,
- ✓ de reconstruire **exactement** le signal global à partir des trames (à l'exception du début et de la fin du signal). Pour vous assurer que le signal global est exactement reconstruit, partez d'un signal synthétique, appliquez la méthode d'addition-recouvrement et comparez le signal reconstruit au signal d'origine.

Remarque : une réflexion sur le fenêtrage (et notamment le choix de la fenêtre) est ici à mener. Elle devra aussi apparaître dans le rapport.

En outre, proposer une méthode permettant d'estimer la variance du bruit additif.

## 2.5 ***Présentation de la méthode de traitement d'une trame d'un signal de parole bruité par un bruit blanc gaussien :***

Partant de l'hypothèse que le signal et le bruit sont décorrélés, l'idée de cette méthode de rehaussement développée par Jensen est de dissocier l'espace engendré par les observations

bruitées en deux sous-espaces, l'un relatif au signal, l'autre se rapportant au bruit blanc additif. L'algorithme opère en trois étapes :

1. on construit une matrice Hankel  $H_y$  de taille  $L \times M$  à partir des données bruitées :

$$H_y = \begin{bmatrix} y(1) & y(2) & \cdots & y(M) \\ y(2) & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ y(L-1) & y(L) & \cdots & y(N-1) \\ y(L) & y(L+1) & \cdots & y(N) \end{bmatrix} \quad (1)$$

Une matrice de Hankel a tous ses coefficients, membres d'une anti-diagonale, égaux.  $L$  et  $M$  vérifient  $L + M = N + 1$  ; on prendra  $L > M$  .

2. on estime au sens des moindres carrés la matrice de Hankel du signal, en extrayant les  $K$  valeurs singulières dominantes de la matrice Hankel  $H_y$  des observations bruitées :

$$H_y = U \Sigma V^T = \begin{bmatrix} U_{1,K} & U_{K+1,M} \end{bmatrix} \begin{bmatrix} \Sigma_{1,K}^s & 0 \\ 0 & \Sigma_{K+1,M}^b \end{bmatrix} \begin{bmatrix} V_{1,K}^T \\ V_{K+1,M}^T \end{bmatrix} \quad (2)$$

où  $U \in \mathbb{R}^{L \times M}$  ,  $\Sigma \in \mathbb{R}^{M \times M}$  et  $V \in \mathbb{R}^{M \times M}$  , et où  $U_{1,K} \in \mathbb{R}^{L \times K}$  et  $V_{K+1,M} \in \mathbb{R}^{M \times M-K}$  .

avec

$$\Sigma = \begin{bmatrix} \Sigma_{1,K}^s & 0 \\ 0 & \Sigma_{K+1,M}^b \end{bmatrix} = \begin{bmatrix} \sigma_1 & 0 & \cdots & \vdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots & \cdots & \cdots & \vdots \\ \vdots & \ddots & \sigma_K & \vdots & \cdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \sigma_{K+1} & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & \sigma_M \end{bmatrix} . \quad (3)$$

3. Etant donné la relation (2), on déduit une estimation de la matrice de données du signal notée  $H_s$  , au sens des moindres carrés, notée  $H_s^{LS}$  définie comme suit :

$$H_s^{LS} = U_{1,K} \Sigma_{1,K}^s V_{1,K}^T . \quad (4)$$

On peut aussi déduire une estimation de  $H_s$  au sens du minimum de variance, notée  $H_s^{MV}$  définie par :

$$H_s^{MV} = U_{1,K} F_{MV} \Sigma_{1,K}^s V_{1,K}^T , \quad (5)$$

où

$$F_{MV} = \begin{bmatrix} 1 - \frac{\sigma_{noise}^2}{\sigma_1^2} & 0 & \dots & 0 \\ 0 & 1 - \frac{\sigma_{noise}^2}{\sigma_2^2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 - \frac{\sigma_{noise}^2}{\sigma_K^2} \end{bmatrix} \quad (6)$$

Puis, pour restaurer la structure de la matrice Hankel de données du signal, et obtenir une estimation du signal, on effectue la moyenne des anti-diagonales des matrices  $H_s^{LS}$  ou  $H_s^{MV}$ . Il suffit alors de considérer les échantillons de la première colonne et de la dernière ligne pour obtenir la trame rehaussée.

- [Jen94] Jensen S. H., Hansen P. C., Hansen S. D. and Sorensen J., "A Signal Subspace Approach for Noise reduction of Speech Signals," EUSIPCO '94, pp. 1174-1177, 1994.
- [Jen95] Jensen S. H., Hansen P. C., Hansen S. D. and Sorensen J., "Reduction of Broad band Noise in speech by Truncated QSVD," IEEE Trans. On Speech and Audio Processing, vol. n°3, n°6, pp. 439-448, November 1995.
- [Jen99] Jensen J., Jensen S. and Hansen H., "Exponential Sinusoidal Modeling of Transitional Speech Segment," ICASSP '99, March 15-19, 1999, Arizona, USA, vol. n°1, pp73-476.

## 2.6 Cahier des charges

A partir d'un signal de parole échantillonné à 8 KHz (fourni), vous devrez réaliser sous *Matlab* les actions suivantes :

1. Créer votre base de données des signaux bruités à traiter. Cf. section 2.2.
2. Effectuer le traitement de rehaussement demandé, trame par trame, en appliquant sur chaque trame de  $N$  échantillons la méthode décrite dans la section 2.5.
  - observer, trame après trame, sur une figure, les représentations temporelles de la trame de parole bruitée, de la trame de parole originale et de la trame de parole rehaussée. On prendra trois couleurs différentes pour les représentations.
  - appliquer la méthode d'addition recouvrement présentée à la section 2.4 pour reconstruire le signal global et fournir sur une même figure la représentation temporelle et le spectrogramme du signal de parole rehaussé dans sa totalité.
  - recommencer les étapes avec une autre séquence de bruit blanc additif.
3. Effectuer une analyse critique des résultats obtenus :
  - calculer le RSB sur les signaux rehaussés (cela est possible dans le cas présent car vous disposez du signal de parole non bruité) ; fournir un tableau récapitulatif des gains en RSB obtenus quand le RSB avant traitement vaut 5dB, 10dB et 15dB. Afin d'avoir une analyse représentative, donner des valeurs moyennées sur un nombre de réalisations jugé suffisant. A l'écoute du signal rehaussé, que pouvez-vous entendre ? Quel phénomène observez-vous ? Expliquez ce phénomène.