

OBJECTIFS DE LA SECONDE SEANCE DE TP

Le but de cette séance est d'appliquer des notions vues en cours sur les filtres numériques mais également d'aller un peu plus loin par l'application et l'observation.

Des programmes Matlab (fichiers avec l'extension .m) vous sont fournis.
Pour interpréter les figures, n'hésitez pas à « zoomer » et/ou à agrandir la fenêtre de la figure.

Dans la partie A, vous devez comprendre les commandes fournies et compléter le programme.
Le but de cette partie est d'appliquer des filtres numériques sur un signal audio afin, d'une part, supprimer un instrument (la contrebasse) puis, d'autre part, rehausser la voix.

Dans la partie B, toutes les commandes sont fournies, vous devrez les comprendre. Lors de cette partie, nous nous familiariserons avec les notions de base de la synthèse vocale en modélisant la prononciation de sons « voisés » : les voyelles. Vous devrez proposer des solutions pour améliorer les résultats.

RENDU DU TRAVAIL

Vous travaillerez en binôme. Chaque binôme enverra son rapport par mail à l'enseignant(e) qui l'a encadré lors de la séance et mettra obligatoirement son binôme en copie du mail.

Contenu du travail à remettre:

Le rapport devra présenter le contexte, ce qui a été fait, les résultats/observations obtenu(e)s et surtout expliquer et justifier ces dernier(e)s. Bien entendu, vous y répondrez également aux questions particulières posées dans le sujet.

Format du travail à remettre:

Vous préciserez dans le nom du document joint par mail, les 2 noms du binôme et le groupe.

Exemple : Dupont_Durand_G2e.pdf.

Seul le format « pdf » sera accepté.

Délai pour la remise du rapport :

Vous avez 1 semaine maximum pour envoyer votre rapport.

Tout rapport remis après cette échéance pourra ne pas être corrigé et le binôme se verra attribuer la note « 0 ».

Les enseignants intervenant dans l'unité :

S. Dupont-Legendre

sylvain.dupont-legendre@esiee.fr

N. Madaoui

nadia.madaoui@esiee.fr

A. Ugon

adrien.ugon@esiee.fr

PARTIE A : FILTRAGE D'UN SIGNAL AUDIO

Pour cette partie, nous disposons d'un programme (*Filt_PartieA.m*) que nous devons comprendre et compléter. Le programme est bien commenté afin de nous aider à comprendre chaque étape. Vous disposez également en annexe des explications sur les fonctions « Matlab » utilisées.

INTRODUCTION :

On dispose d'un extrait de chanson (fichier : « extrait.wav »).

En plus de la voix de la chanteuse (Natalie Cole), parmi les instruments présents, il y a une contrebasse que nous voulons supprimer de l'extrait.

Ce que personnellement je trouve très dommage... mais faisons ce sacrifice pour la science...

La contrebasse est un son grave → situé aux basses fréquences.

A0

La première chose à faire est de récupérer et d'observer le signal à traiter.

Le signal est en stéréo : la variable « son » est donc une matrice de N lignes (nombre d'échantillons) et de 2 colonnes (voie droite et voie gauche).

Nous ne traiterons qu'une seule voie. Rien ne nous empêche par la suite d'écouter ce que cela donne avec une voie traitée et une voie non traitée... voire de traiter les 2 voies... Bref, vous pouvez vous « amuser » et faire tous les tests que votre curiosité peut suggérer.

A1 : SUPPRESSION DE LA CONTREBASSE

Sur le spectre du signal, la contrebasse sera représentée par le pic principal situé entre 80 et 83 Hz. En « zoomant » sur le spectre de l'extrait, vous déterminerez cette fréquence que l'on notera **Fcb**.

Vous devez :

- calculer les pôles et/ou zéros du filtre qui va supprimer cette fréquence **Fcb**
- donner les coefficients du filtre
- représenter le comportement du filtre dans les domaines fréquentiel et complexe
- filtrer le signal x
- et, finalement, observer (courbes) et écouter le résultat y_1 du filtrage. Que constatez-vous ?

A2 : REHAUSSEMENT DE LA VOIX

Après le filtrage précédent, la voix de Natalie Cole est nettement moins « agréable » !

Sa voix est située à une fréquence **FNat** comprise entre 200 et 400 Hz.

Nous allons essayer de la « récupérer » et, au passage, les basses fréquences des instruments qui ont été atténuées par le premier filtre.

Vous devez :

- choisir la fréquence **FNat**
- de la même manière que pour le filtre précédent, vous devez faire les différentes étapes qui permettront de créer et d'observer le nouveau filtre qui va rehausser fortement les fréquences autour de **FNat** et filtrer le signal y_1
- et, finalement, observer (courbes) et écouter le résultat y_2 du filtrage. Que constatez-vous ?

En comparant les courbes (temporelles, fréquentielles et spectrogrammes) des 3 signaux (x , y_1 et y_2) pouvez-vous montrer la cohérence avec ce que vous entendez sur les 3 signaux ? Commentez, justifiez.

PARTIE B : SYNTHESE VOCALE

Pour cette partie, nous disposons d'un programme (*Filt_PartieB.m*) que nous devons comprendre. Le programme est bien commenté afin de nous aider à comprendre chaque étape. Pour vous familiariser avec la synthèse vocale, vous disposez de quelques explications succinctes en annexe ainsi que le document *Biblio1_Voix.pdf* parmi les fichiers fournis.

INTRODUCTION :

Dans cette partie nous allons simuler la prononciation des voyelles (a, e i, o, u) par un homme (fréquence fondamentale F_0 autour de 200 Hz) ou par une femme (fréquence fondamentale F_0 autour de 400 Hz).

Ici, chaque voyelle est caractérisée par 2 fréquences. Le son créé correspondra à une suite d'impulsions de fréquence F_0 dont on accentue fortement les fréquences f_1 et f_2 de la voyelle que nous voulons prononcer.

B1 : CHOIX DU FILTRE

Quel type de filtre devons-nous utiliser et de quel ordre doit-il être ? Expliquer pourquoi.

B2 : COMPREHENSION DES ETAPES DE LA SYNTHESE

Commentez les étapes nécessaires à la création des sons et commentez les résultats/courbes.

B3 : AMELIORATIONS POSSIBLES

Par rapport à la conclusion faite à la fin du programme *Filt_PartieB.m*, proposez des pistes et expliquez en quoi elles peuvent améliorer les résultats.

B4 : MISE EN ŒUVRE DES AMELIORATIONS

Mettez en œuvre les propositions faites à la question **B3** et montrez les améliorations obtenues.

Remarque : vous pouvez vous enregistrer en train de prononcer les voyelles et comparer votre « prononciation » avec les signaux « synthétiques »...

ANNEXE

PARTIE A

Nous utilisons les fonctions Matlab suivantes que nous devons comprendre avec l'aide en ligne de Matlab :

- **poly** (racine1, racine2, ...) : donne les coefficients du polynôme dont les racines sont les paramètres d'entrée de la fonction.
- **freqz** (B,A) : trace le gain ($|H(f)|$) et le déphasage ($\varphi(f)$) du filtre dont les coefficients de $H(z)$ sont les paramètres d'entrée. B (respectivement A) est le vecteur contenant les coefficients du numérateur (respectivement du dénominateur).
- **zplane** (B,A) : trace dans le plan complexe les pôles et les zéros de $H(z)$. Les paramètres d'entrée sont les même que pour *freqz*.
- **filter** (B,A,x) : calcule la sortie du filtre dont les coefficients sont donnés par les vecteurs B et A pour un signal d'entrée x.

Remarque : lors du tracé des spectrogrammes, nous ferons attention à ne tracer qu'une partie du signal (le choix a été fait pour observer les instruments et la voix). Si nous traçons le spectrogramme pour tous les échantillons, Matlab va avoir un peu de mal...

PARTIE B

La voix :

Dans le domaine du traitement de la parole, les voyelles sont classées comme des sons « voisés » alors que les consonnes sont considérées comme des sons « non voisés ».

La production de chaque type de son (voisé ou non voisé) est modélisé de manière différente.

On considère que toutes les voyelles (sons voisés) sont générées par une excitation périodique. La voyelle prononcée sera différenciée d'une autre par ses « formants », généralement au nombre de 3. Les formants sont les fréquences auxquelles nous observons une surtension (l'énergie est concentrée en ces 3 fréquences). Vous trouverez ci-dessous un tableau récapitulant les 3 formants pour chaque voyelle de la langue française.

Pour produire le son d'une voyelle, il nous faut donc :

- 1- un signal d'excitation périodique (un peigne de Dirac dont la fréquence F_0 correspondra au timbre de la voix que l'on veut simuler). Ce sera le signal d'entrée du filtre et simulera la vibration des cordes vocales par l'air.
- 2- un filtre qui accentuera l'énergie aux fréquences correspondantes aux formants de la voyelle. Il simulera le conduit vocal lors de la prononciation de la voyelle.

Le signal filtré correspondra alors au son (voyelle) dont on a simulé la « prononciation ».

Attention, ce modèle ne s'applique qu'aux sons voisés. Pour les consonnes, le modèle est différent autant en terme du choix du signal d'excitation que dans la modélisation du conduit vocal.

		F1	F2	F3
voy. fermées	i	308	2064	2976
	y	300	1750	2120
	u	315	764	2027
voy. mi- fermées	e	365	1961	2644
	ø	381	1417	2235
	o	383	793	2283
voy. mi- ouvertes	ɛ	530	1718	2558
	œ	517	1391	2379
	ɔ	531	998	2399
voy.ouv.	a	684	1256	2503

Tableau 1: valeurs formantiques moyennes des voyelles orales du français (d'après Tubach, 1989)

A voir également le document : Biblio1_Voix.pdf.