Étude Théorique : TP

Elwalid Aboulaakoul

07/12/2024

Partie 1 : Analyse théorique

Calcul de N

Nous savons que :

$$N = \text{durée} \times \text{Fréquence d'échantillonnage}$$

Si d n'est pas un entier, on utilise la fonction `ceil` pour arrondir au plus proche entier supérieur. Ainsi, pour $Fs = 44100$ Hz et $d = 2.5$ s :

$$N = \lceil 44100 \times 2.5 \rceil = 110250$$

Observations sur f_0

Lors de la lecture du son :

- Pour $f_0 = 300$ Hz : le son est relativement grave.
- Pour $f_0 = 1200$ Hz : le son est beaucoup plus aigu.

Période fondamentale T_0

La période fondamentale T_0 est donnée par :

$$T_0 = \frac{1}{f_0}$$

Pour $f_0 = 440$ Hz, on trouve :

$$T_0 = \frac{1}{440} \approx 0.00227 \text{ s.}$$

Le signal $x(t)$ a été affiché sur l'intervalle $[0, 4T_0]$.

Calcul de l'énergie totale $E_x(t)$

L'énergie totale d'un signal $x(t)$ est définie par l'intégrale suivante :

$$E_x = \int_{-\infty}^{+\infty} |x(t)|^2 \, dt$$

Dans notre cas, le signal $x(t)$ est limité dans le temps, car sa durée est $d = 2$ secondes. L'énergie totale peut donc être calculée comme suit :

$$E_x = \int_0^d |x(t)|^2 \, dt$$

En remplaçant $x(t)$ par son expression sinusoïdale :

$$x(t) = \sin(2\pi f_0 t)$$

On obtient :

$$E_x = \int_0^d \sin^2(2\pi f_0 t) \, dt$$

En utilisant l'identité trigonométrique $\sin^2(\theta) = \frac{1}{2}(1 - \cos(2\theta))$, cela devient :

$$E_x = \frac{1}{2} \int_0^d \big(1 - \cos(4\pi f_0 t)\big) \, dt$$

En séparant les termes de l'intégrale :

$$E_x = \frac{1}{2} \int_0^d 1 \, dt - \frac{1}{2} \int_0^d \cos(4\pi f_0 t) \, dt$$

La première intégrale donne :

$$\frac{1}{2} \int_0^d 1 \, dt = \frac{1}{2} \cdot d$$

La seconde intégrale est nulle car $\cos(4\pi f_0 t)$ est périodique et sa moyenne sur une période entière est nulle. Par conséquent :

$$E_x = \frac{1}{2} \cdot d = \frac{1}{2} \cdot 2 = 1$$

L'énergie totale du signal est donc :

$$E_x = 1$$

Le signal $x(t)$ est-il à énergie finie ?

Oui, le signal $x(t)$ est à énergie finie car son énergie totale E_x est finie et vaut 1.

Calcul de la puissance moyenne $P_x(t)$

La puissance moyenne totale d'un signal $x(t)$ est définie comme :

$$P_x = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |x(t)|^2 \, dt$$

Cependant, comme notre signal est limité dans le temps ($d = 2$ secondes), la puissance moyenne se calcule sur cette durée :

$$P_x = \frac{1}{d} \int_0^d |x(t)|^2 \, dt$$

En reprenant les calculs précédents :

$$P_x = \frac{1}{d} \cdot E_x = \frac{1}{2} \cdot 1 = \frac{1}{2}$$

La puissance moyenne totale du signal est donc :

$$P_x = \frac{1}{2}$$

Partie 2 : Observations temporelles

Période fondamentale M_0

La période fondamentale en termes d'échantillons est donnée par :

$$M_0 = \frac{Fs}{f_0}$$

Pour un signal de fréquence $f_0 = 800$ Hz et une fréquence d'échantillonnage $Fs = 8000$ Hz :

$$M_0 = \frac{8000}{800} = 10 \text{ échantillons.}$$

Distance euclidienne entre deux périodes

En supposant que $x[n]$ est périodique, la distance euclidienne entre deux périodes successives est donnée par :

$$D = \sqrt{\sum_{n=1}^{M_0} (x[n] - x[n + M_0])^2}$$

Énergie et puissance pour un signal de durée $d = 1000$ s

L'énergie totale est donnée par :

$$E_x = \frac{1}{2} \cdot d$$

En remplaçant $d = 1000$ s :

$$E_x = \frac{1}{2} \cdot 1000 = 500$$

La puissance moyenne reste identique pour un signal périodique :

$$P_x = \frac{1}{2}$$

Echantillonnage et critière de Nyquist

Dans cette étude, nous analysons le critère de Nyquist et l'effet de l'aliasing en utilisant deux fréquences $f_0 = 800\,\text{Hz}$ et $f_0 = 7200\,\text{Hz}$. Nous travaillons avec une fréquence d'échantillonnage $F_s = 8000\,\text{Hz}$ et explorons différentes valeurs de F_s pour illustrer expérimentalement le critère de Nyquist.

1. Valeur minimale selon Nyquist

Le critère de Nyquist stipule que la fréquence d'échantillonnage doit être au moins deux fois plus grande que la fréquence maximale contenue dans le signal pour éviter l'aliasing. Ainsi, pour un signal de fréquence $f_0 = 800\,\text{Hz}$, la fréquence minimale d'échantillonnage est donnée par :

$$F_s^{\text{min}} = 2 \times f_0 = 1600\,\text{Hz}.$$

2. Illustrations expérimentales

Pour vérifier expérimentalement ce critère, on choisit différentes valeurs de fréquence d'échantillonnage :

$$F_s = [1000, 2000, 4000, 8000, 16000]\,\text{Hz}.$$

Pour chaque F_s, on génère un signal sinusoidal :

$$x_n(t) = \sin(2\pi f_0 t),$$

où $t = \frac{n}{F_s}$ et n est l'indice d'échantillon.

Les signaux sont ensuite écoutés. Lorsque $F_s < F_s^{\text{min}}$, on observe une dégradation due à l'aliasing.

3. Fréquences apparentes

Pour un signal avec une fréquence f_0, la fréquence apparente perçue après échantillonnage est donnée par la relation modulaire :

$$f_{\text{apparent}} = |f_0 - k \cdot F_s|\,,$$

où $k = \text{round}\left(\frac{f_0}{F_s}\right)$.

Pour $f_0 = 800\,\text{Hz}$ et $f_0 = 7200\,\text{Hz}$, et $F_s = 8000\,\text{Hz}$, les calculs montrent :

$$f_{\text{apparent, 800}} = \text{mod}(800, 8000) = 800\,\text{Hz},$$
$$f_{\text{apparent, 7200}} = \text{mod}(7200, 8000) = 800\,\text{Hz}.$$

Cela signifie que les deux fréquences produisent le même signal apparent en raison de l'aliasing.

4. Conclusions

- Le critère de Nyquist est respecté lorsque $F_s \geq 2 \cdot f_0$. En dessous de cette valeur, des artefacts d'aliasing apparaissent. - Les signaux avec $f_0 = 800\,\text{Hz}$ et $f_0 = 7200\,\text{Hz}$ ont la même fréquence apparente lorsque $F_s = 8000\,\text{Hz}$, ce qui illustre le phénomène d'aliasing.

Synthèse d'une note

La fréquence f_{note_0} d'une note donnée peut être calculée avec la formule suivante :

$$f_{\text{note}_0} = 440 \times 2^{\frac{\text{note}-69}{12}}$$

Où :

- 440 Hz est la fréquence de la note A4 (la quatrième note "A" dans l'octave),
- note est le numéro de la note que l'on veut jouer,
- 69 correspond à la note A4 dans la table de fréquence.

Exemple de calcul

Calculons la fréquence d'une note note = 72, qui correspond à la note C5 :

$$f_{\text{note}_0} = 440 \times 2^{\frac{72-69}{12}} = 440 \times 2^{\frac{3}{12}} \approx 523.25\,\text{Hz}$$

Synthèse d'une mélodie

Une mélodie est une suite de notes jouées successivement. Pour la créer, on génère un signal pour chaque note en utilisant une sinusoïde, puis on les assemble en fonction de leur durée.

Fonction de génération de mélodie

La fonction `create_melody.m` prend en entrée :

- Un vecteur de durées $\mathbf{d} = [d_1, d_2, \ldots, d_n]$,
- Un vecteur de numéros de notes $\mathbf{note} = [n_1, n_2, \ldots, n_n]$,
- La fréquence d'échantillonnage F_s.

Elle renvoie un signal x et un vecteur temps t, où x est la concaténation des sinusoïdes correspondant à chaque note, et t est le vecteur temps associé.

Exemple de modification de la mélodie

- Pour jouer une octave en dessous, on divise la fréquence de chaque note par 2 :
$$f_{\text{new}} = \frac{f_{\text{note}_0}}{2}$$
- Pour jouer la mélodie 50% plus vite, on réduit la durée de chaque note de moitié :
$$d_{\text{new}} = 0.5 \times d$$

Étude dans le domaine fréquentiel

4.1 Cas d'une note

1. Résolution en fréquence de la TFD

La résolution en fréquence de la transformée de Fourier discrète (TFD) est déterminée par la durée du signal d. Si d est la durée du signal, la résolution en fréquence Δf est donnée par :

$$\Delta f = \frac{1}{d}$$

Dans notre cas, avec une durée de 2 secondes, la résolution en fréquence est :

$$\Delta f = \frac{1}{2} = 0.5\,\text{Hz}$$

2. Notes observables sur la TFD

Les notes dont les fréquences sont suffisamment séparées par rapport à la résolution en fréquence peuvent être observées correctement dans la TFD. Si les fréquences des notes sont trop proches, elles peuvent se superposer, rendant leur identification difficile.

4.2 Cas d'une mélodie

1. Identification des notes dans une mélodie

Lorsque plusieurs notes sont jouées successivement, leur spectre peut être analysé en identifiant les pics correspondant aux fréquences des notes. Les pics seront visibles dans le spectre de la TFD et permettront de retrouver les notes jouées.

Filtrage et séparation de sources

Dans cette partie, nous allons traiter le problème de la séparation de sources par filtrage dans le domaine fréquentiel. L'objectif est de séparer deux notes successives à l'aide de filtres passe-bas et passe-haut.

Transformation de Fourier et filtrage

La méthode que nous allons utiliser consiste à appliquer un filtrage linéaire dans le domaine fréquentiel, ce qui correspond à une multiplication de la transformée de Fourier du signal avec la fonction de transfert du filtre. Les fonctions de transfert des filtres sont définies comme suit :

- **Filtre passe-bas** :

$$H_{\mathrm{LP}}(f) = \begin{cases} 1 & \text{si } |f| < f_c \\ 0 & \text{sinon} \end{cases}$$

- **Filtre passe-haut** :

$$H_{\mathrm{HP}}(f) = \begin{cases} 0 & \text{si } |f| < f_c \\ 1 & \text{sinon} \end{cases}$$

- **Filtre passe-bande** :

$$H_{\mathrm{BP}}(f) = \begin{cases} 1 & \text{si } f_{c1} < |f| < f_{c2} \\ 0 & \text{sinon} \end{cases}$$

Dans ces équations, f_c représente la fréquence de coupure du filtre.

Cas de deux notes successives

Nous allons créer un signal sonore composé de deux notes successives : une note aiguë et une note grave. Ces notes seront synthétisées à des fréquences f_1 et f_2 (choisies respectivement pour la note aiguë et la note grave). Le signal sera échantillonné à une fréquence de $f_s = 8000\,\text{Hz}$.

La transformée de Fourier du signal sera ensuite calculée, et le module au carré de cette transformée sera tracé pour observer les deux notes. Nous proposerons une fréquence de coupure f_c permettant de séparer les deux notes.

Création des filtres

Nous allons créer deux filtres : un filtre passe-bas et un filtre passe-haut. La fréquence de coupure f_c sera choisie de manière à séparer les deux notes présentes dans le signal. Ces filtres seront ensuite appliqués à la transformée de Fourier du signal.

Filtrage et reconstruction du signal

Après avoir appliqué les filtres, nous effectuerons une transformée de Fourier inverse pour récupérer les signaux filtrés. Le signal filtré avec le filtre passe-bas isolera la note grave, tandis que le filtre passe-haut isolera la note aiguë.

Les résultats seront présentés sous forme de graphiques pour comparer les signaux originaux et filtrés.

5.2 Séparation de sources supervisée

1. Création du signal x_1

Nous avons créé un signal x_1 d'une durée de 10 secondes, échantillonné à $F_s = 8000\,\text{Hz}$, contenant des notes comprises entre 50 et 71 Hz. Ce signal est constitué de plusieurs sinusoïdes correspondant à des notes musicales graves. Après l'écoute du signal x_1, nous avons constaté que la tonalité générale était dans la gamme des notes basses.

2. Création du signal x_2

Un deuxième signal x_2 a été créé de la même manière, mais avec des notes comprises entre 72 et 90 Hz, représentant des notes plus aiguës. Ce signal a également été échantillonné à $F_s = 8000\,\text{Hz}$, avec une durée de 10 secondes. Après avoir écouté le signal x_2, nous avons remarqué que les notes étaient dans la gamme des sons aigus.

3. Fréquence de coupure f_c

Le signal x est la somme des deux signaux x_1 et x_2, soit $x = x_1 + x_2$. Afin de séparer ces deux signaux, nous devons choisir une fréquence de coupure f_c qui sépare les fréquences basses (présentes dans x_1) des fréquences aiguës (présentes dans x_2).

En observant les gammes de fréquences des deux signaux, nous pouvons définir une fréquence de coupure f_c autour de 71 Hz, ce qui permettra de filtrer efficacement les deux signaux. Les fréquences inférieures à 71 Hz seront attribuées au signal x_1, et les fréquences supérieures à 71 Hz au signal x_2.

4. Reconstruction des signaux y_1 et y_2

En utilisant des filtres passe-bas et passe-haut avec une fréquence de coupure $f_c = 72\,\text{Hz}$, nous avons reconstruit les signaux y_1 et y_2 à partir du signal x. Le signal y_1 a été estimé en appliquant un filtre passe-bas, et le signal y_2 a été estimé en appliquant un filtre passe-haut.

5. Comparaison des signaux et calcul de la distance euclidienne

Après écoute des signaux reconstruits y_1 et y_2, nous avons comparé ces signaux aux sources réelles x_1 et x_2. La distance euclidienne entre x_1 et y_1, ainsi qu'entre x_2 et y_2, a été calculée. La distance euclidienne est une mesure de la différence entre deux signaux, définie comme suit :

$$\text{Distance euclidienne} = \sqrt{\sum_{n=1}^{N}(x(n) - y(n))^2}$$

Où $x(n)$ et $y(n)$ sont les valeurs des signaux respectifs à l'instant n.

Les distances calculées entre les signaux originaux et les signaux reconstruits étaient relativement faibles, ce qui indique que la reconstruction a été efficace.

6. Observation des spectres des signaux

Les spectres des signaux x_1 et y_1, ainsi que ceux des signaux x_2 et y_2, ont été observés. Les spectres montrent la répartition des fréquences dans les signaux, et ils confirment que les filtres ont correctement séparé les composantes basses et aiguës des signaux.

Le spectre de x_1 montre principalement des fréquences inférieures à 72 Hz, tandis que le spectre de x_2 montre des fréquences supérieures à 72 Hz. Les spectres de y_1 et y_2 devraient être similaires aux spectres respectifs de x_1 et x_2, respectivement.

7. Tests avec d'autres fréquences de coupure

Nous avons testé différentes valeurs de fréquence de coupure f_c. Par exemple, en choisissant $f_c = 80\,\text{Hz}$, nous avons observé que la séparation des signaux s'effectuait de manière similaire, bien que les résultats aient légèrement changé en raison de l'impact de la fréquence de coupure sur la gamme des fréquences filtrées.

Il est important de noter que le choix de la fréquence de coupure f_c a un impact direct sur la qualité de la séparation. Une fréquence de coupure trop élevée ou trop basse pourrait entraîner une mauvaise séparation des sources.

Conclusion

Le projet a permis de mettre en œuvre une méthode supervisée de séparation de sources audio, où l'information a priori sur les signaux a été utilisée pour séparer deux mélodies distinctes. Nous avons créé deux signaux musicaux x_1 et x_2 représentant respectivement des notes graves et aiguës. Après les avoir additionnés pour former un signal composite x, nous avons appliqué des filtres passe-bas et passe-haut pour estimer les signaux originaux x_1 et x_2.

La séparation des sources a été réalisée efficacement grâce au choix approprié de la fréquence de coupure, f_c, permettant de distinguer les fréquences graves des fréquences aiguës. L'utilisation de filtres spécifiques à chaque gamme de fréquence a conduit à des signaux reconstruits y_1 et y_2 proches des sources originales, comme l'indiquent les faibles distances euclidiennes entre les signaux originaux et les signaux estimés.

L'observation des spectres des signaux avant et après séparation a montré que les filtres appliqués ont permis de conserver les caractéristiques principales des sources, tout en supprimant les interférences entre les deux gammes de fréquences.

Enfin, les tests avec différentes fréquences de coupure ont permis de confirmer que le choix de f_c est crucial pour la qualité de la séparation. Un ajustement fin de cette fréquence est nécessaire pour obtenir des résultats optimaux, en fonction des caractéristiques spécifiques des signaux à traiter. Ce projet a ainsi démontré l'efficacité des techniques de filtrage pour la séparation de sources en audio, et souligne l'importance de l'information a priori dans les méthodes supervisées.