

Illustration pédagogique du concept de corrélation : mesure de la vitesse du son et synthèse de voyelles

C. Durieu *
Mis en ligne le 22 février 2002.

Résumé

La fonction de corrélation est un outil important en traitement du signal. Sont présentées ici deux applications très simples de la fonction de corrélation permettant aux étudiants d'exploiter des signaux réels qu'ils enregistrent eux-mêmes. La première application concerne la mesure du retard entre deux signaux et la deuxième la synthèse de voyelles. Cette dernière application est également l'occasion d'étudier des signaux dans le domaine temporel et le domaine fréquentiel.

Mots-clés : corrélation, équation de Yule-Walker, mesure de retard, modèle autorégressif, synthèse de voyelles.

© EDP Sciences, 2002.

Niveau de connaissances requis. Notions de base sur les signaux et systèmes. Quelques éléments sur les signaux aléatoires.

Niveau des étudiants. Deuxième cycle EEA, physique. Après adaptation, ces manipulations peuvent être effectuées en partie en classes préparatoires ou en IUT.

* **Cécile Durieu** est maître de conférences au département EEA de l'ENS de Cachan depuis 1989. Elle y enseigne l'automatique et le traitement du signal en licence, maîtrise et préparation aux agrégations. Sa principale activité de recherche, menée au sein du LESiR, concerne actuellement l'estimation de paramètres et d'état à l'aide de techniques ellipsoïdales.

Adresse postale : Département EEA, ENS Cachan, 61 avenue du président Wilson, 94235 Cachan, France.
email : cecile.durieu@eea.ens-cachan.fr

Article available at <http://www.j3ea.org> or <http://dx.doi.org/10.1051/bib-j3ea:2002004>

1. Motivations

De nombreuses notions enseignées en traitement du signal sont très souvent abstraites pour les étudiants et difficiles à appréhender ; ce cours est souvent perçu comme étant très mathématique et peu appliqué. Le but de cet article est de présenter deux exemples d'application très simples de la fonction de corrélation. La première application concerne la détermination de la vitesse du son et la deuxième la synthèse de voyelles [1,2,3]. Chaque illustration se déroule sur une séance de travaux pratiques de 4 heures.

Un des intérêts de ces manipulations est que les étudiants participent directement en enregistrant eux-mêmes des sons qu'ils émettent. De plus, ils perçoivent la synthèse de voyelles comme un jeu car, à la fin de la séance de travaux pratiques, une écoute successive des signaux enregistrés et des signaux synthétisés est effectuée. Cette expérience a été initiée au département EEA de l'ENS de Cachan et se déroule actuellement dans le cadre de la maîtrise EEA de l'Université de Paris Sud-Orsay.

2. Vitesse du son

La mesure de la vitesse du son s'effectue en mesurant le retard τ_r entre les signaux $s_1(t)$ et $s_2(t)$ reçus par deux microphones situés dans la direction d'émission de la source et séparés d'une distance d_0 connue. Un schéma du dispositif est représenté sur la figure 1. Le retard pourrait être mesuré très facilement avec un signal émis très bref, par exemple un claquement. Cependant, afin de justifier l'utilisation de la fonction de corrélation, le signal émis est un « che » qui est maintenu assez longtemps pour faciliter les acquisitions. Un exemple de signal reçu par les microphones est représenté sur la figure 2 ($d_0 = 0,5 \text{ m}$ et $\tau_r = 1,46 \text{ ms}$ à 20°C). À partir d'une visualisation des signaux il est impossible d'estimer le retard. Le spectre (périodogramme normalisé) des signaux est relativement riche comme le montre la figure 3.

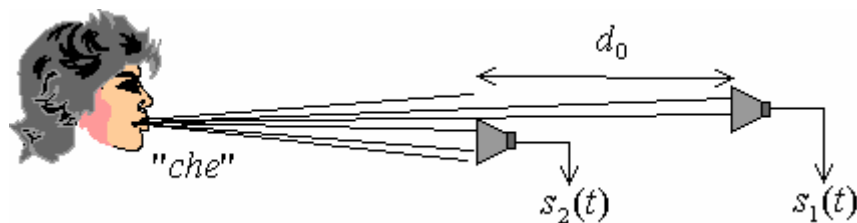


Figure 1. Dispositif.

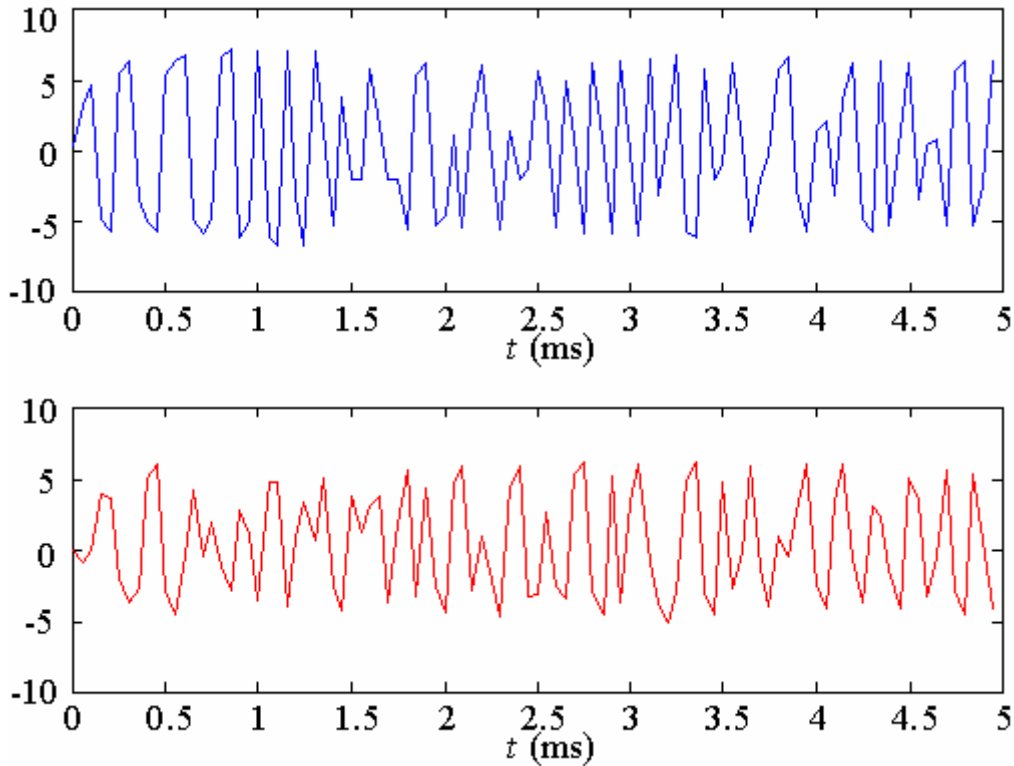


Figure 2. Zoom des signaux enregistrés ($s_1(t)$ en haut et $s_2(t)$ en bas).

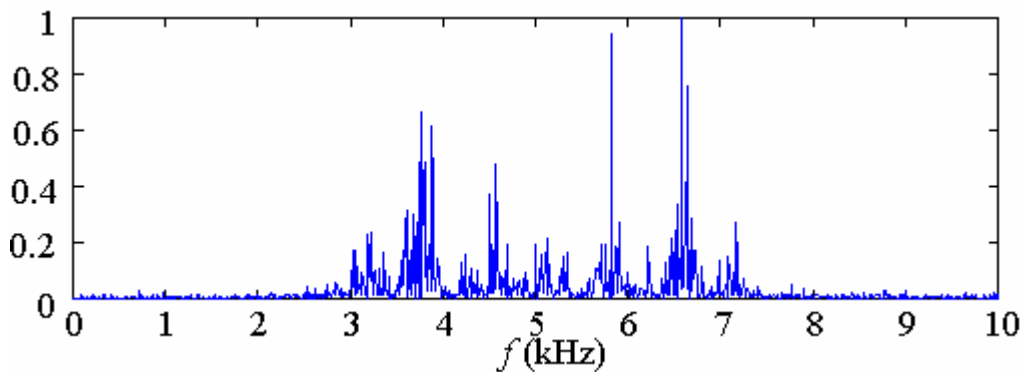


Figure 3. Spectre des signaux.

Pour l'étude théorique, le signal à analyser est supposé être une réalisation d'un processus aléatoire stationnaire.

La fonction d'intercorrélation théorique $c_{12}(\tau)$ entre les signaux $s_1(t)$ et $s_2(t)$ présente un maximum en $\tau = \tau_r$. En pratique, la fonction de corrélation n'est pas connue. Une approche répandue (mais empirique) consiste à estimer la fonction de corrélation à partir des échantillons disponibles du signal puis à rechercher la position du maximum. L'estimateur retenu est non biaisé. Les signaux étant enregistrés sur N points ($n=1, \dots, N$) à la fréquence $1/T_e$, $\hat{c}_{12}(\tau)$ est alors estimée en $\tau = kT_e$ par

$$\hat{c}_{12}[k] = \frac{1}{N-k} \sum_{n=k+1}^N s_1[n]s_2[n-k] \text{ pour } k \in [0, N-1]$$

et

$$\hat{c}_{12}[k] = \frac{1}{N-|k|} \sum_{n=1}^{N-|k|} s_1[n]s_2[n-k] \text{ pour } k \in [-N+1, 0]$$

avec $s_i[n]=s_i(nT_e)$ ($i=1,2$).

Un exemple de fonction d'intercorrélacion estimée est représenté sur la figure 4. La détermination de la position du maximum, qui est assez étroit (ce qui justifie le son choisi), permet d'en déduire la vitesse du son ou encore la

$$v = v_0 \sqrt{1 + T/273}$$

température de la pièce en utilisant la loi empirique , T étant la température de la pièce

$$v_0 = 331,3 \text{ m s}^{-1}$$

(exprimée en degré Celsius) et v_0 la vitesse à 0°C (dans les conditions ordinaires). Signalons que la vitesse trouvée expérimentalement est correcte dans la majorité des cas.

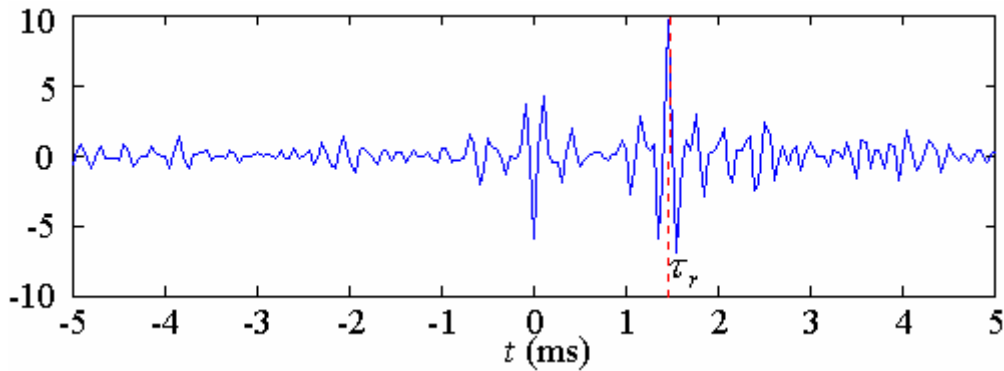


Figure 4. Intercorrélacion entre $s_1(t)$ et $s_2(t)$.

Cette manipulation est l'occasion d'illustrer que la fonction d'intercorrélacion mesure la ressemblance entre deux signaux. Cette fonction est programmée dans le domaine temporel par les étudiants afin qu'ils comprennent mieux sa signification. Cette partie pourrait être étoffée en envisageant également une programmation dans le domaine fréquentiel à base de Transformée de Fourier Rapide. Une fonction ainsi écrite est donnée aux étudiants afin qu'ils vérifient leur programme et surtout qu'ils comparent les temps d'exécution. Afin d'affiner la recherche de la position du maximum, un sur-échantillonnage de la fonction de corrélacion estimée pourrait être envisagé.

En début de séance de travaux pratiques l'effet de bord, dû à la fenêtre d'observation, est mis en évidence en étudiant des signaux synthétiques dont la fonction de corrélacion théorique est connue. Les étudiants estiment également puis comparent la fonction de corrélacion de différents bruits blancs filtrés.

3. Synthèse de voyelles

En tout début de séance de travaux pratiques, les étudiants enregistrent une voyelle qui sert à l'étude. Un exemple de signal (« a ») servant à illustrer la synthèse est représenté sur la figure 5. En première approximation on peut considérer qu'une voyelle est un motif élémentaire qui est répété K fois. La période de répétition T dépend du timbre du locuteur et vaut environ 10 ms pour un homme et 5 ms pour une femme. La fréquence $1/T$ correspond au fondamental de la voix. Les paramètres T et K peuvent être identifiés directement et simplement sur le signal

enregistré ($T \cong 4,7 \text{ ms}$ et $K \cong 28$ pour le signal enregistré sur la figure 5).

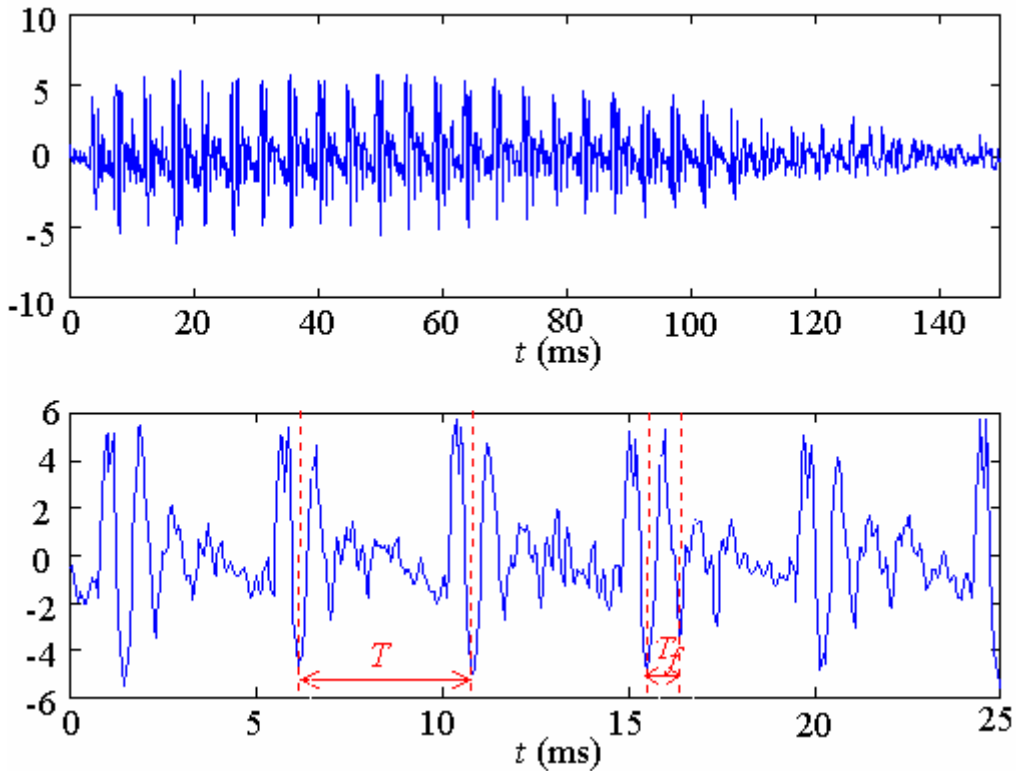


Figure 5. Voyelle « a » enregistrée et zoom.

A l'intérieur du signal on peut également mettre en évidence d'autres fréquences (3 à 5) que l'on appelle

$$1/T_f \cong 1,1 \text{ kHz}$$

formants ; par exemple la fréquence est présente dans le signal enregistré sur la figure 5. Une étude spectrale (Fig. 6) de la voyelle confirme ceci. Le spectre (périodogramme) est constitué de raies très fines centrées sur des fréquences qui sont des multiples de $1/T$. Pour l'exemple traité le spectre présente un maximum pour la fréquence $1/T_f$. Cette analyse permet de faire le lien entre la représentation temporelle et la représentation fréquentielle des signaux.

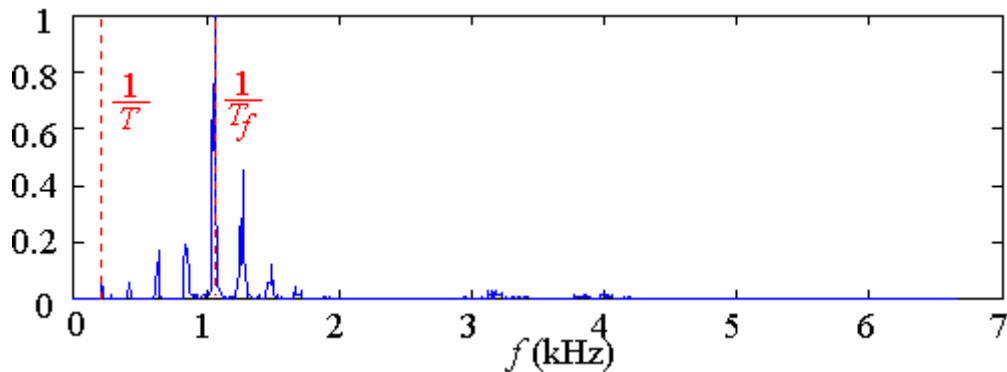


Figure 6. Spectre de la voyelle.

Pour la synthèse le signal à analyser est supposé être une réalisation d'un processus aléatoire stationnaire et autorégressif, c'est-à-dire engendré par le passage d'un processus aléatoire blanc et stationnaire dans un filtre linéaire invariant de fonction de transfert $H(z)$ tous pôles d'ordre d . Le problème se réduit alors à un problème d'estimation de paramètres : les d coefficients a_i du filtre. Ces paramètres et la fonction d'autocorrélation $c[k]$ du signal échantillonné $s[k]$ sont reliés par les équations dites de *Yule-Walker*..

Dans le cas purement théorique où l'on connaîtrait la fonction de corrélation, il suffirait de résoudre ce système d'équations linéaires pour en déduire les coefficients du filtre. Mais, tout comme précédemment, la fonction de corrélation n'est pas connue et l'approche consiste toujours à estimer la fonction de corrélation à partir des échantillons disponibles du signal, à reporter ces valeurs estimées dans les équations de Yule-Walker, puis à résoudre ces dernières. Il faut évidemment choisir un estimateur de la fonction de corrélation. Si un estimateur non biaisé est retenu, la matrice de corrélation ainsi obtenue n'est plus garantie définie positive et, si le signal traité est *mal choisi*, le filtre identifié peut être instable. On retient alors un estimateur biaisé de la fonction de corrélation garantissant la stabilité du filtre. $c[k]$ est alors estimé par

$$\hat{c}[k] = \frac{1}{N} \sum_{n=k+1}^N s[n]s[n-k], \quad k = 1, \dots, d$$

Reste à déterminer l'ordre du filtre. Pratiquement un ordre 16 ou 20 est retenu. Le filtre autorégressif identifié et permettant la synthèse peut ensuite être étudié : lien entre la réponse impulsionnelle (Fig. 7), la réponse en fréquence (Fig. 8) et la position des pôles (Fig. 9).

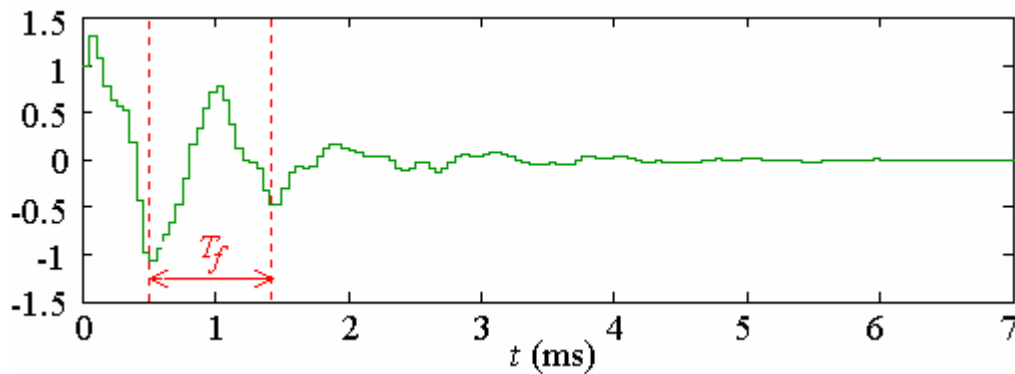


Figure 7. Réponse impulsionnelle du filtre.

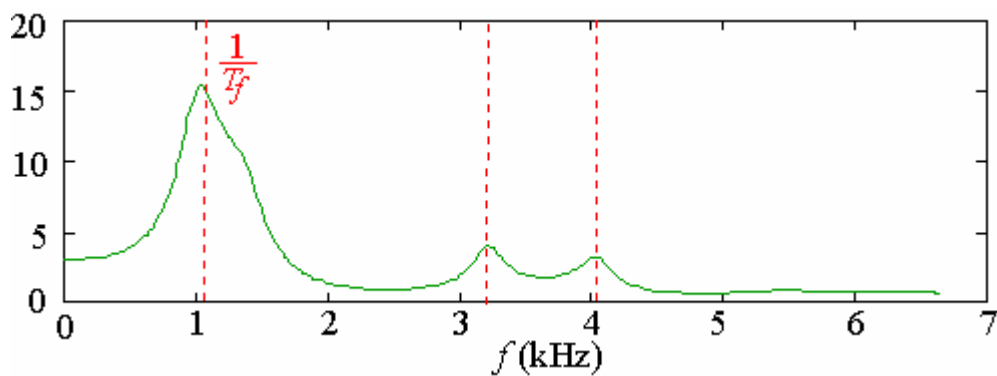


Figure 8. Réponse en fréquence du filtre.

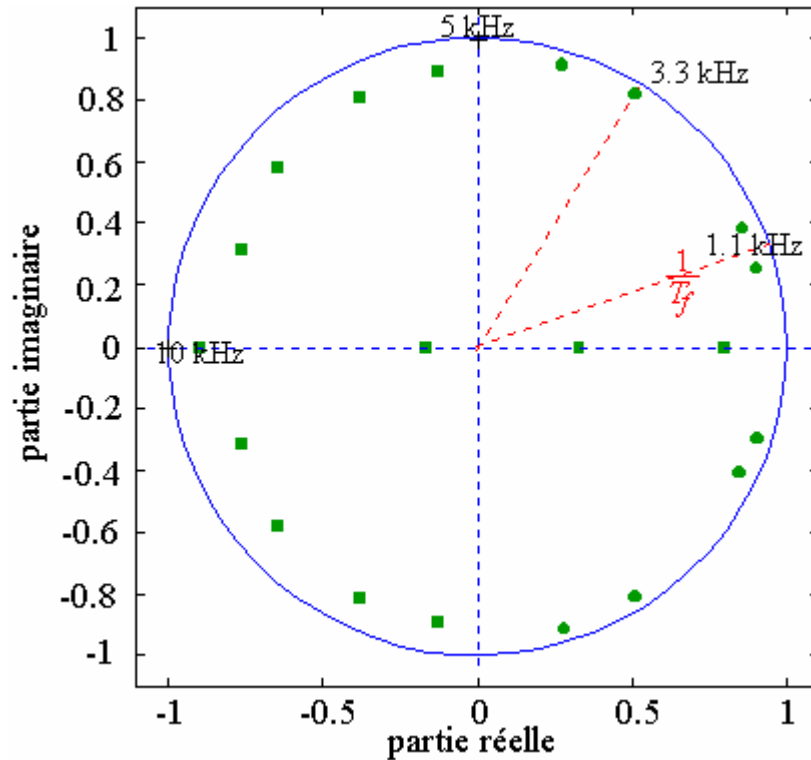



Figure 9. Pôles du filtre.

Pour la synthèse proprement dite le signal d'excitation du filtre est une série de K impulsions espacées de T ou encore distantes de T/T_e échantillons. La démarche expérimentale est représentée sur la figure 10. Le signal synthétisé est représenté sur la figure 11 et son spectre (périodogramme) sur la figure 12. La voyelle enregistrée et la voyelle synthétisée sont assez proches (en temps et en fréquence). Une écoute successivement des voyelles enregistrées puis synthétisées illustre cette manipulation et permet de reconnaître facilement les voyelles synthétisées ( entendre les voyelles).

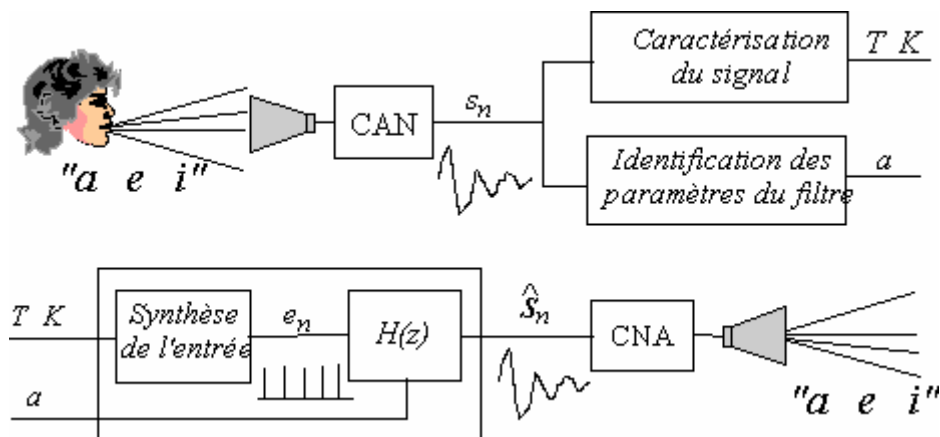


Figure 10. Synoptique de la synthèse de voyelles.

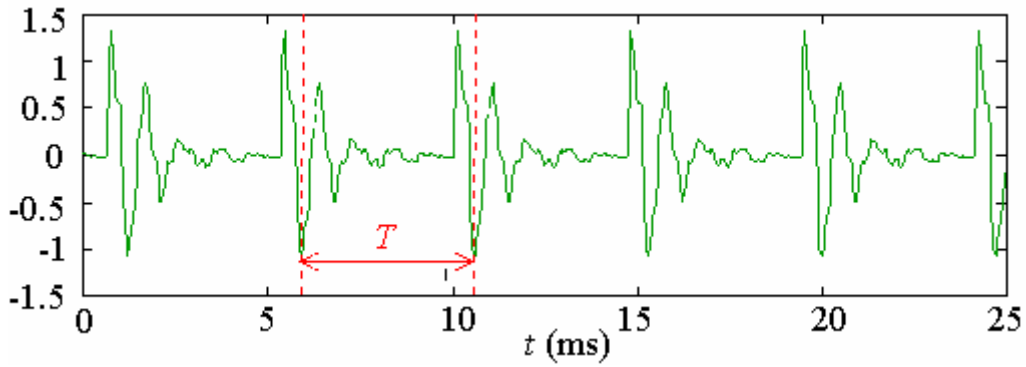


Figure 11. Voyelle synthétisée.

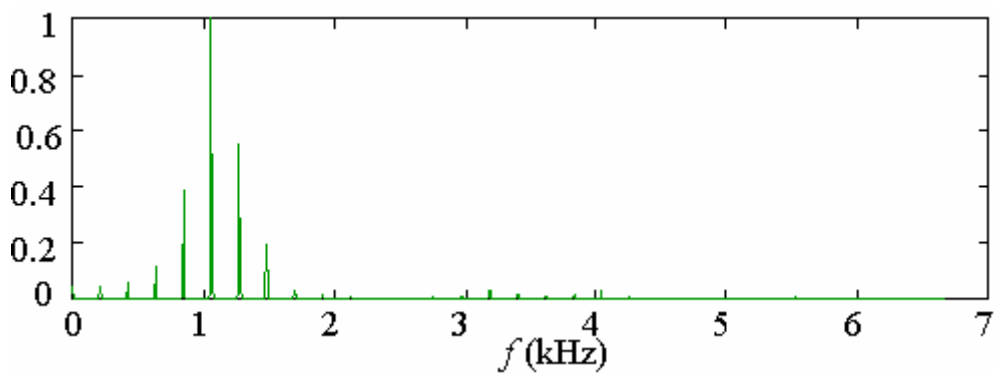


Figure 12. Spectre de la voyelle synthétisée.

Signalons que le signal est enregistré sur environ 3000 points et que la synthèse ne nécessite la connaissance que d'une vingtaine de paramètres (les d coefficients a_i du filtre, T et K), soit une réduction de plus de 100 de l'encombrement mémoire nécessaire. Une extension naturelle de cette étude est la reconnaissance de voyelles basée sur leurs caractéristiques (spectres, formants, pôles dominants, coefficients du filtre).

4. Matériel

Le matériel nécessaire est très réduit :

- Pour la mesure de la vitesse du son : une maquette comprenant deux microphones avec leur électronique associée (qui est très réduite) ainsi qu'un ordinateur équipé d'une carte d'acquisition (fréquence d'échantillonnage supérieure ou égale à environ 16 kHz).
- Pour la synthèse de voyelles : un ordinateur avec un microphone et un haut parleur intégrés ou, à défaut, un microphone avec son électronique associée, un haut parleur et une carte d'acquisition (fréquence d'échantillonnage supérieure à 16 kHz).

Le traitements de données est effectué, par exemple, à l'aide du logiciel Matlab™.

5. Conclusion

Ces manipulations (qui se déroulent en tout début d'année, le cours de traitement de signal venant de commencer) intéressent beaucoup les étudiants. Le concept de corrélation y est appréhendé sous un angle très pratique. Les applications sont simples mais riches et les outils présentés en cours et travaux dirigés sont illustrés de façon attrayante.

Références bibliographiques

- [1] J.P. Tubach, (CALLIOPE), *La parole et son traitement automatique* (Masson, 1989).
- [2] R. Boite et M. Kunt, *Traitement de la parole* (Presses Polytechniques Romandes, 1987).
- [3] M. Ferretti et F. Cinare, *Synthèse et reconnaissance de la parole* (Éditests : diffusion PSI, 1983).