Fabrication et utilisation d'hybrides sonores

Daniel ARFIB
CNRS-LMA
31, Chemin Joseph Aiguier
13402 Marseille Cedex 20
arfib@lma.cnrs-mrs.fr

Résumé

Un hybride sonore est un composite de deux sons (naturels ou synthétiques) qui possède des propriétés de ses géniteurs. Cet article répertorie les moyens classiques de créer ces hybrides: techniques électroacoustiques (mixage, modulation en anneau et en bande latérale unique), séparation source-résonance (vocodeur à canaux, prédiction linéaire, cepstre), représentations temps/fréquences (spectrogrammes et phasogrammes) et modèles paramétriques. L'utilisation des hybrides est évoquée, avec le cas particulier du métamorphose sonore. (morphing)

1.les techniques électroacoustiques

1.1. le mixage

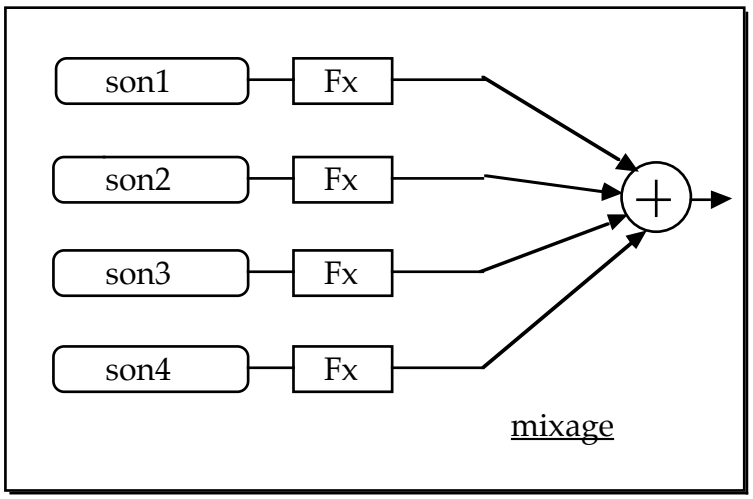

fig. 1: mixage de sons

Le mixage - en particulier numérique - est le mode le plus primaire pour faire d'un ensemble de sons un complexe sonore dont les sources sont difficiles à décrypter.
L'opération mathématique est:

$$res(t)=\sum_i son_i(t)$$

Deux exemples simples démontreront son usage: un choeur qui fait interagir des voix dont la fréquence est très proche, et un cluster d'orchestre. Dans ces deux cas, on obtient un résultat sonore que l'oreille ne différencie pas. C'est l'inverse de la ségrégation qui permet de distinguer des instruments différents dans un duo par exemple.

1.2. le modulateur en anneau

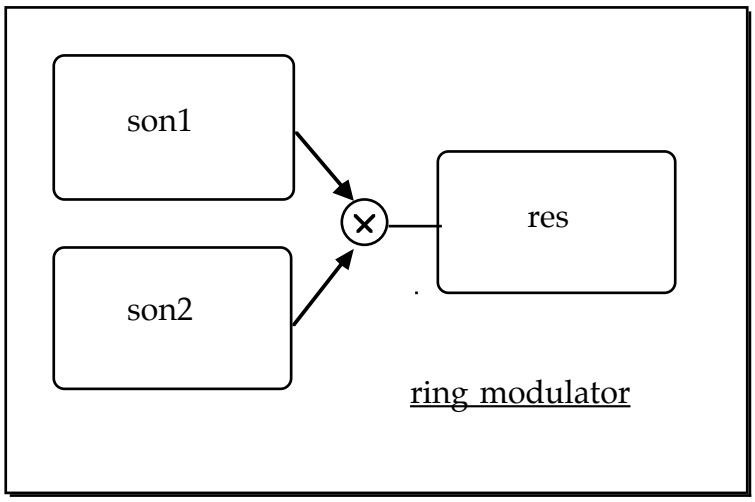

fig. 2: modulation en anneau

Le modulateur en anneau fait usuellement interagir un son naturel avec un son synthétique de forme simple (sinusoïde, triangle, carré) mais peut aussi se former entre deux sons naturels.

L'opération mathématique est:

res(t)=son1(t)*son2(t)

dans laquelle * représente la multiplication, échantillon par échantillon, des deux sources sonores. La multiplication temporelle étant réciproque de la convolution fréquentielle, le résultat par exemple d'une modulation en anneau d'une sinusoïde de fréquence fp avec un son harmonique de fréquence $n*f_0$ donne des raies aux fréquences:

$fp \pm n* f_0$

Autrement dit le spectre est déplacé en symétrique à droite et à gauche de la porteuse, et la partie gauche négative se replie dans le domaine positif. Ceci peut conduit à de l'inharmonicité et parfois du frottement entre des raies voisines, un phénomène souvent exploité en électroacoustique.

1.3. modulation en bande latérale unique.

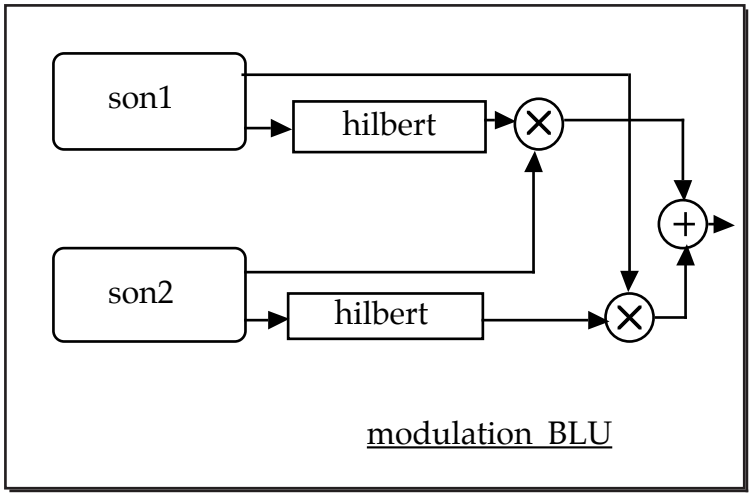

fig. 3: modulation en bande latérale unique

Si l'on désire translater un signal (c'est à dire ajouter un décalage à toutes ses composantes), ce qui est différent de transposer, qui est une multiplication de toutes les raies par un même facteur, on peut faire appel à la modulation en bande unique (BLU).

Celle-ci se réalise en fabriquant un signal jumeau du signal original, dont les phases sont toutes décalées de 90 degrés. Cela s'appelle encore la transformée de Hilbert ou la fabrication du signal analytique. Le moyen le plus simple de la réaliser est de fabriquer le signal déphasé en filtrant le signal original par un filtre de Hilbert qui peut être non récursif ou récursif. Dans les deux cas, ce n'est qu'une approximation, mais suffisante pour l'usage musical.

L'opération mathématique est ($son1_h$ est le signal son1 déphasé de 90 degrés)

$$res(t)= \cos(2\pi ft)*son1(t)+\sin(2\pi ft)*son1_h(t)$$

On peut bien sûr réaliser une modulation de deux sons en prenant leurs transformées de Hilbert et en calculant

$$res(t)= son1_h(t)*son2(t)+son1(t)*son2_h(t)$$

On obtient alorsdes raies dont les fréquences sont les sommes des fréquences initiales.

2.les croisements de sons.

Une grande classe d'interactions est basée sur le fait qu'un son possède deux propriétés que l'on peut distinguer: un aspect tonal (les valeurs en fréquence des partiels) et un aspect spectral (l'évolution en amplitude de ces partiels) Cela est très typé pour les sons vocaux: on peut distinguer nettement l'action des cordes vocales qui ajustent la fréquence de la voix, que l'on peut dénommer "source", et celle de la cavité buccale qui détermine l'articulation que l'on peut dénommer "résonance". On peut prendre pour autre exemple le violon électronique de Max Mathews où le son d'une corde frottée (sans table d'harmonie) est modelé par une boîte électronique constituée de filtres résonants.

Les croisements utilisent souvent des méthodes de séparation source-résonance. Un croisement entre deux sons consiste alors à utiliser la "source" d'un son avec la "résonance" d'un autre. On peut ou non retirer au premier son sa propre résonance (on le "blanchit" en ce cas), et on modèle ensuite ce son par l'enveloppe spectrale évolutive du second.

2.1.vocodeur à canaux.

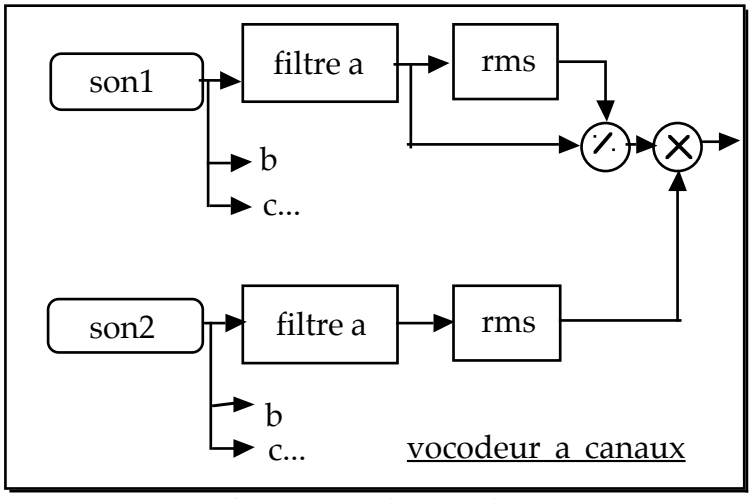

Fig. 4: schéma d'un vocodeur

Cette technique utilise deux bancs de filtres, afin de pouvoir détecter les enveloppes spectrales par l'analyse des amplitudes de chaque bande. On substitue alors dans chaque bande l'amplitude du signal filtré du son 1 par celle du son 2. La partie la plus délicate est évidemment la détection de cette amplitude, qui est usuellement réalisée par RMS (root mean square), c'est à dire le filtrage passe-bas du signal au carré suivi d'une racine carrée.

L'opération mathématique est:

$$res(t)=\sum s_i(t)*rms2_i(t)/rms1_i(t)$$

où $rms1_i$ et $rms2_i$ représentent les amplitudes des signaux dans la bande i du banc de filtre.

On notera qu'il faut en pratique tenir compte d'un seuil pour le dénominateur de la division, c'est à dire que lors du "blanchiment" il ne faut pas sur-amplifier une voie qui n'aurait pas de signal.

2.2. prédiction linéaire:

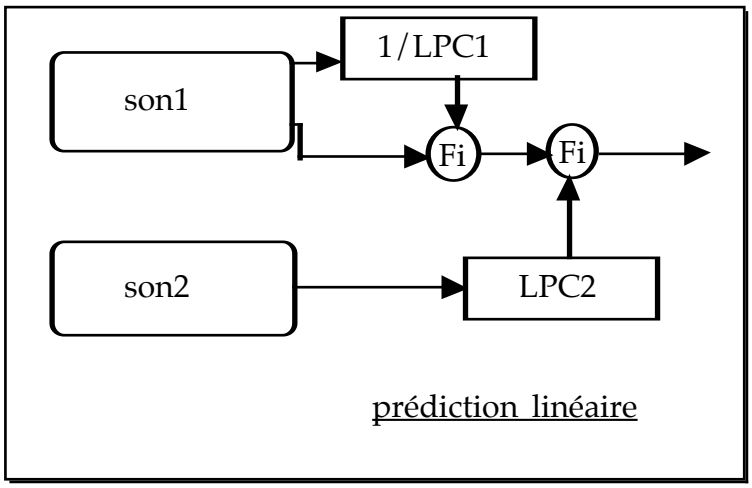

Fig. 5: principe du croisement en prédiction linéaire

Une autre façon de procéder est d'utiliser un algorithme de prédiction linéaire. Très utilisée en traitement de la parole, cette technique s'adapte aux sons musicaux s'ils sont monodiques (source unique). Elle consiste à trouver à tout moment le meilleur filtre qui permette la séparation source-résonance (l'algorithme de Levinson est classique dans ce domaine; il permet le calcul d'un filtre non récursif de longueur donnée à partir de la donnée d'un extrait du signal sonore). La source est obtenue par antifiltrage (en remplaçant les pôles par des zéros) du son avec un filtre évolutif, et le résultat final est obtenu en filtrant cette source par le filtre issu du son numéro 2

2.3. traitement homomorphique (cepstre)

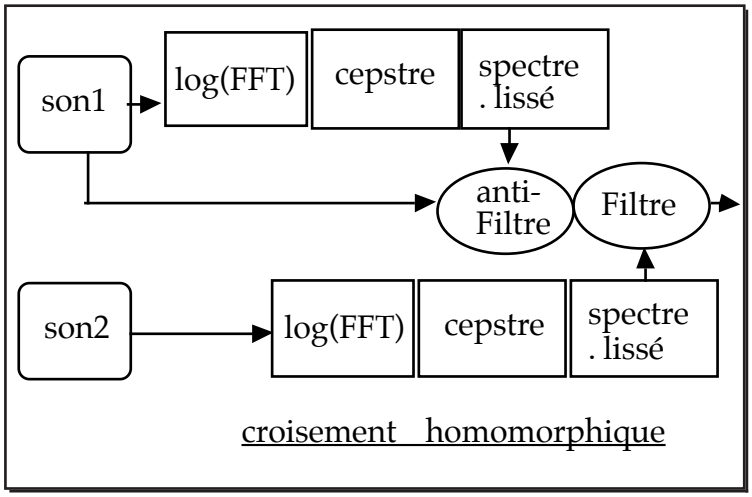

Fig. 6: utilisation du cepstre dans un croisement de sons

Dans le même ordre d'idées de détection source-résonance, on trouve la méthode du cepstre, qui est à mon sens une des plus stables et performantes. La technique de base consiste, après avoir obtenu la FFT du signal auquel est assignée une fenêtre temporelle, à décomposer ce spectre sous la forme d'un produit de deux spectres.

On peut le réaliser en prenant le logarithme de ce spectre, et en le décomposant en somme de deux spectres. En remarquant que la résonance est plutôt "molle" alors que la source est typiquement une suite

d'impulsions correspondant aux harmoniques, il suffit de filtrer ce logarithme du spectre. C'est là qu'intervient le "cepstre", qui et la FFT du logarithme de la FFT.

Pour croiser deux sons, on blanchit le premier en lui retirant plus ou moins sa coloration spectrale, et on lui surimpose la deuxième. Cela peut être obtenu, avec les précautions nécessaires, par la méthode de "l'overlap and add" (décomposition du signal en grains se recouvrant, modification des grains, sommation de grains modifiés successifs)

3. les mutations sonores

Pour obtenir des hybrides sonores, il est possible d'utiliser les représentations temps/fréquence de ces sons, et de faire la resynthèse d'un son à partir d'une représentation composite. Deux représentations temps/fréquence sont très importantes: en ondelettes et en Gaborettes. Dans ce dernier cas, on utilise souvent la technique de transformée de Fourier à fenêtre glissante, mais il est intéressant de voir ces représentations sous une forme générale:

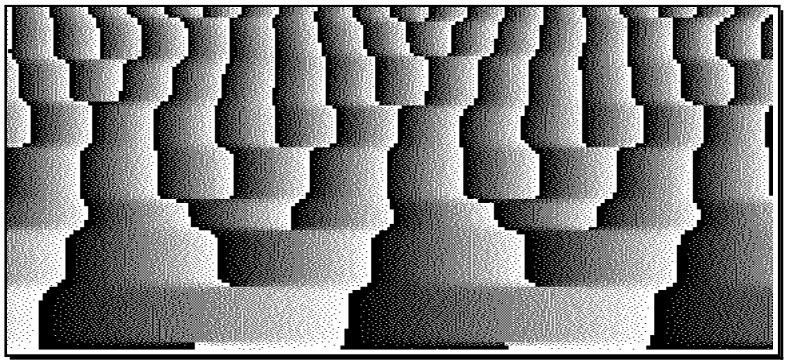

Fig. 7: Phasogramme en grille fine d'un son de flûte

Une transformée temps/fréquence est obtenue en prenant le produit scalaire d'un signal avec une base d'ondelettes ou de Gaborettes. Ces représentations ont plusieurs propriétés:
- elles sont à valeur complexe, donc décomposables en deux diagrammes: amplitude (sonogramme) et phase (phasogramme)
- la connaissance des valeurs sur une grille (dyadique pour la Transformée en ondelettes, rectangulaire pour une transformée de Gabor) permet la resynthèse exacte du signal original
- Une image arbitraire a peu de chances d'être la transformée d'un signal. On dit alors qu'elle n'est pas une transformée valide. Elle devrait pour cela satisfaire à des relations entre les valeurs prises sur les points de la grille (condition au noyau reproduisant).
- forcer une resynthèse sonore à partir d'une image arbitraire est équivalent à appliquer le noyau reproduisant de la transformée considérée. Ce noyau est, en un point, la transformée de la gaborette (ondelette) de resynthèse vue à travers les gaborettes (ondelettes) d'analyse.

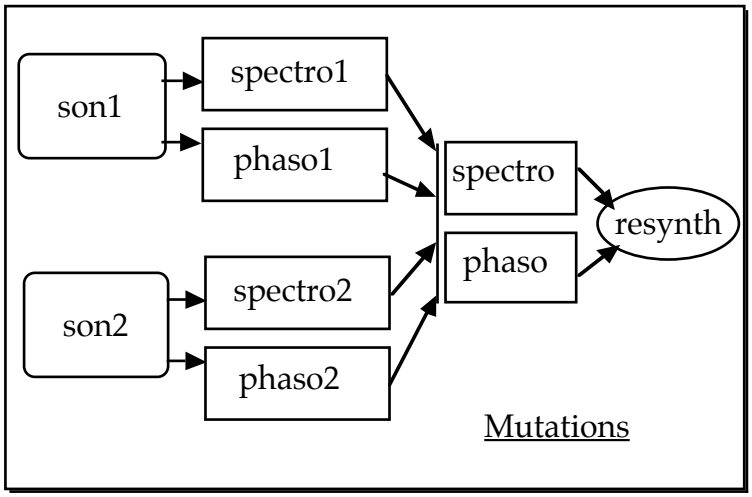

Fig. 8: principe des mutations spectrales

Ceci étant rappelé, on peut construire des hybrides sonores en créant une représentation sonore à partir de deux analyses, en sachant toutefois que cette représentation a peu de chances d'être valide (au sens mathématique), ce qui veut dire en termes sonores que des artefacts seront peut être ajoutés par le processus de resynthèse.

3.1. opérations sur les modules.

Plusieurs opérations sont réalisables:
- multiplication des modules (c'est à dire addition de leurs valeurs en décibels). Cela correspond a un 'ET' logique: seules seront conservées les zones où il y de l'énergie dans les deux sons.
- addition des modules, cela correspond à un 'OU logique mais qui n'est pas équivalent à un mixage car ce ne sont pas les valeurs complexes que l'on ajoute mais les modules.
- masquage d'un son par un autre (l'équivalent d'un "OU exclusif")
- utilisation du cepstre réel du module pour modifier l'autre module (analogue au traitement homomorphique) ou toute autre fonction perturbant le module

3.2.opérations sur les phases

Le diagramme des phases est d'importance considérable pour la resynthèse d'hybrides. Il conditionne étroitement la perception de hauteur. De plus des erreurs par rapport à la validité de la transformée peuvent être drastiques dans leur résultat. Voici quelques exemples possibles:
- préservation de la phase d'un seul signal
dans ce cas la resynthèse aura tendance à préserver la fréquence de ce son puisque la reconstruction par le noyau reproduisant consolidera les raies correspondant aux harmoniques.
- imposition d'un diagramme des phases arbitraires
On tend à imposer ici une structure déterminée qui correspond selon le cas à un son harmonique, ou inharmonique, ou bruité.
- mixage des phases
On tient ici à garder un peu des deux sons. Mais c'est une opération non linéaire qui n'est pas un mixage. De plus il faut alors souvent dérouler la phase car sa définition modulo 2π ne peut plus être conservée

3.3. résultat des hybridations par représentations T/F

Parmi les plus classiques, on peut noter
- robotisation d'un son par imposition d'une phase d'un son synthétique
- bruitage aléatoire de la phase ou du module
- synthèse croisée par la méthode du cepstre du module

4. Métamorphose sonore (morphing)

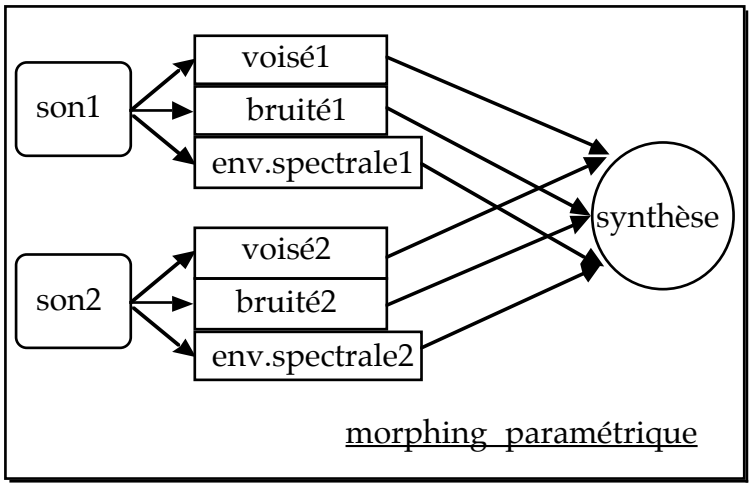

Fig. 9: resynthèse paramétrique d'un hybride

L'utilisation musicale d'hybrides sonores ne dépend pas seulement des techniques utilisées, mais et surtout de leur contrôle. Le son a ceci de particulier qu'il s'appuie sur le temps et que cette dimension est un acteur fondamental de la perception sonore.

Alors que l'image fixe est hors du temps et sur deux dimensions similaires de l'espace, le son possède une évolution temporelle et de matière sonore multidimensionnelle qui rend l'idée de métamorphose un peu différente de l'image. En fait on peut vouloir plusieurs types de métamorphose (morphing):
- fabriquer une boucle répétitive où il y a évolution successive d'un son vers un autre
- ou bien sur un seul son passer entre le début et la fin d'un premier son à un autre
- ou bien sur deux séquences sonores faire évoluer dynamiquement ou statiquement des aspects du son. Par exemple on peut désirer garder la fréquence fondamentale du premier et le "timbre" du second, ou ne transformer que l'aspect temporel (attaques).

Dans toutes ces utilisations musicales possibles interviennent deux techniques:
- distinguer sur un son ou une séquence sonore une série de zones de forces (composantes principales) qui permettent de reconstruire un son
- créer des points d'ancrage temporels qui permettent les interpolations
On peut alors indiquer la proportion, évoluant dans le temps, de chaque indice de resynthèse.

5 Conclusion.

L'utilisation d'hybrides sonores en musique peut rapidement donner lieu à des clichés si ceux-ci ne sont pas utilisés à partir de concepts de composition. En fait c'est le matériau idéal pour composer le son, et non seulement avec le son.

références

[d'Alessandro]C. d'Alessandro, C. Demars: *Représentations temps-fréquence du signal de parole*, Traitement du Signal, vol 9, n°2, pp153-173

[Arfib 1990] D. Arfib, *Transformations exotiques de sons par ordinateur,* 1er Congrès Français d'Acoustique, Lyon, C2, 1990, pp 853-856.

[Arfib 1991] D. Arfib . *Analysis, transformation, and resynthesis of musical sounds with the help of a time-frequency representation.* In G. de Poli, A. Picciali & C. Roads, editors, The representation of musical sounds, M.I.T Press, 1991, pp 87-118.

[Arfib 1993] D.Arfib, N. Delprat. *Musical Transformations using the modification of time-frequency images*, computer Music Journal, vol 17, n°2, summer 93, pp 66-72

[Bargar 1995]Bargar, Holloway, Rodet, Hartman: *Defing spectral surfaces,* proceedings of the 1995 ICMC Banff, pp 373-376

[Galas 1990] Galas T., Rodet X.: *An improved cepstral method for deconvolution of source-filter systems with discrete spectra*, Proceedings ICMC, Glasgow, pp 82-84, 1990

[Moorer 1978] J.A Moorer *The use of the phase vocoder in computer music applications*, J. Audio Eng. Soc., vol 26 (1), 1978

[Pielemeier 1996] Pielemeier WJ, Wakefield, Simoni: *Time-Frequency Analysis of musical signals*, proc of the IEEE, vol. 84, n°9, sept96, pp 1216 1229.

[Robinson 1997] T. Robinson: *Speech Analysis* (une introduction aux techniques LPC, cepstre, etc...) http://svr-www.eng.cam.ac.uk/~ajr/SpeechAnalysis/index.html

[Slaney 1996] Slaney M.,Covell M. and Lassiter B.:*Automatic Audio Morphing*, Procceedings of the 1996 International Conference on Acoustics, Speech, and Signal Processing, IEEE, Atlanta, GA, May 7-10, 1996.

*