Vers un modèle pour unifier musique et son dans une composition multiéchelle

Myriam Desainte-Catherine (myriam@LaBRI.U-Bordeaux.Fr)
Sylvain Marchand (sm@LaBRI.U-Bordeaux.Fr)

SCRIME[1]
LaBRI[2], Université Bordeaux I
351, cours de la Libération
F-33405 Talence cedex – France

Résumé

Nous proposons d'essayer d'unifier l'écriture musicale avec le contrôle sonore pour permettre de composer de façon continue sur plusieurs échelles de temps. Dans ce but, nous présentons un modèle sonore, basé sur des paramètres proches à la fois de la perception et de la terminologie musicale, qui permet une correspondance entre des structures macroscopiques et microscopiques.

1 Introduction

Depuis le début des recherches en informatique musicale, deux courants principaux se développent en parallèle. L'un se base sur les intentions du compositeur pour permettre des constructions d'abstractions musicales de plus en plus sophistiquées. L'autre étudie l'analyse et la synthèse des sons pour manipuler directement les structures intimes des sons.

Le paysage des logiciels d'aide à la composition est à l'image de cette dichotomie. Le compositeur doit maîtriser plusieurs types de logiciels s'il souhaite, en particulier, composer des pièces mixtes, faites de parties instrumentales et électroacoustiques. Dans ces conditions, toute intention d'interaction entre les sons et les macrostructures de sa pièce doit être entièrement gérée par lui-même, au moyen d'allers-retours inconfortables d'une représentation à l'autre, voire même d'un logiciel à l'autre.

Nous envisageons dans cet article une unification des représentations des données des pièces musicales permettant une continuité de représentation et de contrôle de la macrostructure d'une pièce aux microstructures des sons la composant. Elle apporterait au compositeur non seulement un plus grand confort mais surtout un moyen d'expression plus puissant, lui permettant de mélanger des structures jusqu'alors totalement disjointes.

Après un rapide historique de la séparation entre musique et son dans la section 2, nous présentons dans la section 3 le modèle de Synthèse Additive Structurée (SAS) qui permet d'aller vers l'unification du son et de la musique exposée dans la section 4.

2 Historique de la dichotomie musique / son

Actuellement il existe deux branches de recherche parallèles, l'une dans le domaine symbolique de la musique et l'autre dans le domaine du son. Elles considèrent l'ordinateur respectivement comme un outil d'aide à la composition ou comme un instrument de musique numérique. Les recherches sur la modélisation symbolique de la musique sont issues historiquement des premiers travaux de composition automatique par Hiller [1] dans la fin des années 50. Dans ce paradigme, une pièce

[1] Studio de Création et de Recherche en Informatique et Musique Electroacoustique, soutenu par le Conseil Régional d'Aquitaine, l'Université Bordeaux I, le Ministère de la Culture, la Direction Régionale des Actions Culturelles d'Aquitaine, la Mairie de Bordeaux, l'ENSERB et le Conseil Général de la Gironde.

[2] Laboratoire Bordelais de Recherche en Informatique, Unité Mixte de Recherche 5800 du CNRS.

musicale est définie au moyen d'atomes qui sont des notes et les structures musicales sont représentées au moyen de divers paradigmes orientés objet, fonctionnel et résolution de contraintes définis sur les paramètres de la note. Les logiciels de composition dans le domaine symbolique sont nombreux et très performants pour la manipulation de macrostructures musicales. Mais leur puissance s'arrête à la frontière séparant la note du son, et ils sont dans l'impossibilité de pénétrer ces boîtes noires pour contrôler efficacement la microstructure des atomes symboliques.

Cette dichotomie s'explique facilement de par la nature des paradigmes de représentation sonore : les paramètres permettant de les contrôler sont bien éloignés des paramètres musicaux mis en œuvre dans les macrostructures du niveau symbolique.

- La musique concrète [2] se fabrique aujourd'hui à partir de sons échantillonnés que l'on manipule au moyen d'éditeurs de signal audionumérique. Cette représentation n'offre aucun contrôle précis des paramètres musicaux tels que la hauteur, la durée ou le volume, qui n'y sont pas indépendants.
- La synthèse de signal numérique consiste souvent à programmer un générateur dans un langage dérivé de MUSIC V permettant de décrire un algorithme de synthèse. Les paramètres que le compositeur manipule ici sont les interconnexions entre composants tels que des oscillateurs, des additionneurs, des multiplicateurs, des filtres, etc. Les possibilités de synthèse sont innombrables, mais là encore les paramètres sont bien éloignés des grandeurs musicales. Dans les langages comme Csound ou MPEG-4 *Structured Audio* [3] la séparation entre musique et son se veut clairement posée, les fichiers « partition » et « orchestre » étant distincts, même si la frontière entre les deux est loin d'être claire pour le programmeur…
- Les modèles physiques [4] permettent la description de dispositifs de production sonore, tels que la constitution physique d'un instrument et le geste de l'instrumentiste. Si ces modèles sont plus proches des concepts musicaux de par la prise en compte du jeu instrumental, ils ne permettent pas non plus de manipuler des abstractions musicales.

Pour que ces deux branches de recherche parallèles, l'une dans le domaine symbolique de la musique et l'autre dans le domaine du son, puissent se rejoindre et s'enrichir mutuellement, il est nécessaire de trouver une unification permettant la continuité entre les atomes des macrostructures de l'une et la représentation interne de l'autre. Nous présentons dans la section suivante un modèle de son basé sur des critères perceptifs et permettant cette unification. Ses paramètres sont issus à la fois de la musique (hauteur, volume, durée), et de recherches menées précédemment sur le timbre [2, 5, 6, 7, 8].

3 SAS : un modèle sonore basé sur la perception

On peut distinguer trois grandes familles de modèles sonores : les modèles physiques, abstraits et spectraux. Dans les modèles physiques il s'agit de mettre en équations la source sonore et de calculer une solution ; les modèles abstraits proposent une formule mathématique pour le son, contenant si possible des paramètres musicaux ; les modèles spectraux quant à eux paramètrent le son au niveau du récepteur, en tenant compte si possible de la perception.

Afin d'unifier musique et son, nous proposons d'oublier aussi bien l'action physique de l'instrumentiste que la cause physique du son et de les remplacer respectivement par l'intention du premier et la perception du second, et ceci pour parvenir à un point d'accord entre la musique (macroscopique) et le son (microscopique). Le modèle de Synthèse Additive Structurée (SAS) est un modèle sonore, à la fois spectral et abstrait, misant sur la perception pour unifier musique et son. Basé sur la synthèse additive, il impose des contraintes sur les paramètres de cette dernière pour en mettre en évidence d'autres, mais cette fois aussi proches de la perception et de la terminologie musicale que possible.

3.1 Synthèse Additive

La synthèse additive (cf. [9]) est l'une des plus anciennes techniques de synthèse sonore. Elle est très proche de la perception et repose sur le théorème de Fourier qui assure que tout son périodique peut être décomposé en un ensemble de partiels, oscillateurs sinusoïdaux ayant chacun une fréquence et une amplitude propre. Si ces grandeurs varient avec le temps, la synthèse additive retient alors comme paramètres les évolutions temporelles (lentes) des fréquences et amplitudes instantanées de tous les partiels du son, comme préconisé par McAulay et Quatieri [10]. On aboutit alors à une équation pour l'amplitude du signal en fonction du temps de la forme :

$$a(t) = \sum_{p=1}^{P} a_p(t)\cos(\varphi_p(t)) \tag{1}$$

où P est le nombre de partiels et

$$\varphi_p(t) = \varphi_p(0) + 2\pi \int_0^t f_p(u)\,du \quad \text{i.e.} \quad \frac{d\varphi_p}{dt}(t) = 2\pi\, f_p(t) \tag{2}$$

Les fonctions f_p, a_p et φ_p sont respectivement la fréquence, l'amplitude et la phase du p-ième partiel. Les P couples (f_p, a_p) constituent les paramètres du son dans le modèle additif. Ces couples forment des points dans le plan fréquence-amplitude, constituant ainsi la représentation spectrale classique du son illustrée par la figure 1.

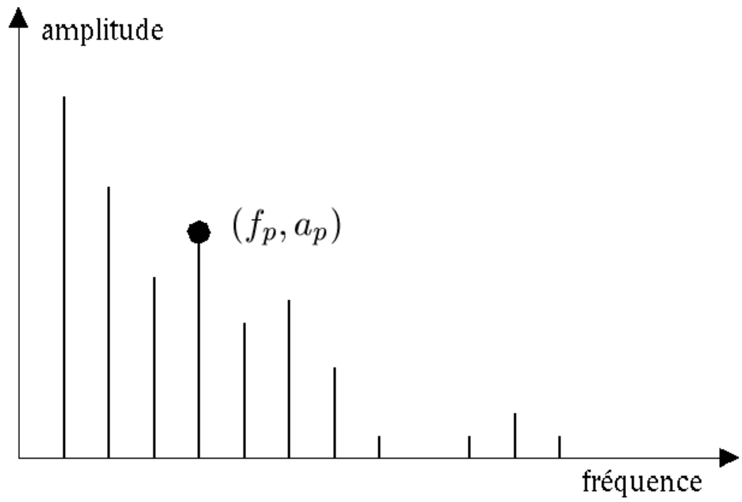

Figure 1 : Décomposition de Fourier d'un son quasi-harmonique.

La synthèse additive est très générale et permet de reproduire avec fidélité une vaste famille de sons. Elle ne permet pas cependant de produire des bruits ou des transitoires rapides. Récemment les modèles SMS [11, 12] (*Spectral Modeling Synthesis*) puis S+T+N [13] (*Sines+Transients+Noise*) ont été proposés pour les prendre en compte. Il faut cependant pouvoir extraire les paramètres du modèle à partir de sons réels lors d'une phase d'analyse. L'analyse est un problème toujours ouvert, mais il existe déjà des méthodes suffisamment précises. Une telle méthode ainsi que des références peuvent être trouvées dans [14, 15]. Il faut également disposer d'une technique de synthèse efficace, car la synthèse nécessite le calcul d'un grand nombre d'oscillateurs. Des algorithmes de plus en plus performants et la montée en puissance des ordinateurs permettent maintenant de produire des centaines voire des milliers d'oscillateurs en temps réel, comme le prouve notre logiciel de resynthèse, *ReSpect*, qui permet de générer 2 oscillateurs par MHz sur un Pentium II Intel, soit 800 oscillateurs à 400 Mhz.

Il reste cependant un problème fondamental : le degré de liberté autorisé par les paramètres, physiquement valides mais musicalement contestables, est si grand qu'il est presque impossible de modifier ou de créer un son « réaliste ». Cet amer constat peut se faire en pratique avec des logiciels de synthèse additive comme *SoftSynth* [16] : il y a trop de paramètres dont la signification est éloignée de la perception.

3.2 Synthèse Additive Structurée (SAS)

Cette section propose le modèle de Synthèse Additive Structurée (SAS) pour remédier au problème décrit précédemment, en structurant les paramètres de la synthèse additive de manière à réintroduire une cohérence perceptive et musicale pour le modèle.

3.3 Paramètres perceptifs

Les paramètres du modèle SAS sont au nombre de quatre :

Amplitude (*amplitude*)

L'amplitude est une fonction A : temps $\rightarrow$ amplitude, calculée à partir des paramètres additifs ainsi : $A(t) = \sum_{p=1}^{P} a_p(t)$. L'amplitude est directement liée à la puissance du son. Elle est perçue logarithmiquement par notre système auditif, et il est aisé de déduire, à partir de cette amplitude, le volume en dB : $20 \log_{10}(A)$.

Fréquence (*frequency*)

La fréquence est une fonction F : temps $\rightarrow$ fréquence, calculée à partir des paramètres additifs dans le cas harmonique ainsi : $F(t) = \mathrm{pgcd}(f_1(t), \cdots, f_p(t))$. Pour les sons harmoniques, F coïncide avec la fondamentale du son (même si cette dernière est manquante ou « virtuelle »). La fréquence est liée à la périodicité du son. Elle est également perçue logarithmiquement et on peut par exemple calculer la hauteur MIDI ainsi : $57 + 12 \log_2(F / 440)$, où 57 correspond au la de l'octave 4, 58 au la dièse, etc.

Couleur (*colour*)

La couleur, est une fonction C : fréquence $\times$ temps $\rightarrow$ amplitude, qui coïncide avec l'enveloppe spectrale du son [8]. La raison d'appeler ce paramètre « couleur » est double, à la fois :

- physique, tant l'analogie entre les spectres audible et visible est grande. En effet, une couleur visible est également définie par son enveloppe spectrale. Les physiciens avaient déjà fait depuis longtemps cette analogie pour désigner les bruits (bruit blanc, bleu, etc.) ;
- musicale, tant le vocabulaire des compositeurs est empreint de cette notion de couleur. On parle en effet de coloration induite par un filtre, un système de diffusion ou une salle de concert. Il s'agit bien en fait de modifications de l'enveloppe spectrale.

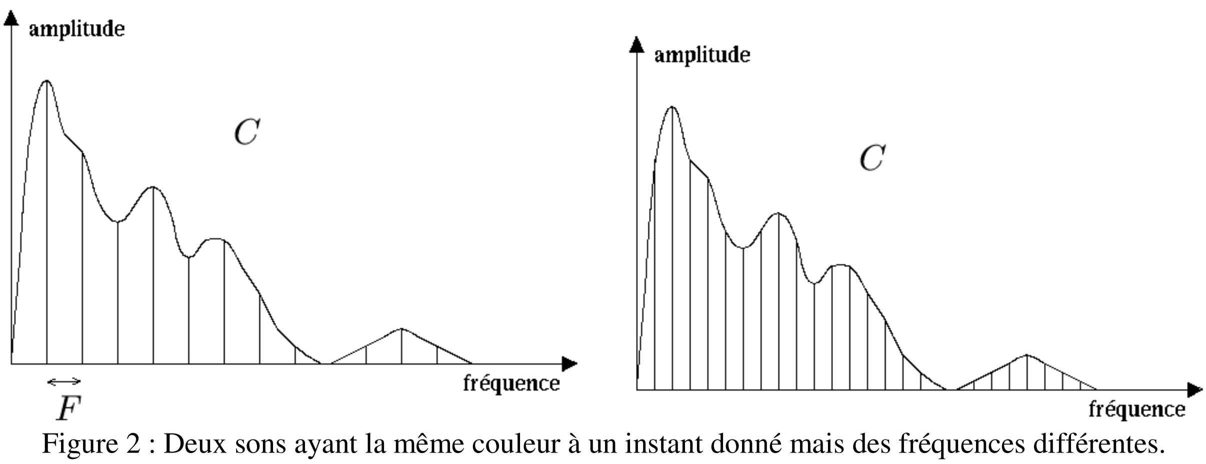

Figure 2 : Deux sons ayant la même couleur à un instant donné mais des fréquences différentes.

Généralement la couleur varie dans le temps, cependant pour certains sons il arrive qu'elle soit quasiment constante. Nous menons, en étroite collaboration avec des psycho-acousticiens, des recherches sur la perception de la couleur. Par exemple il se trouve que la couleur possède une hauteur propre, qui peut dans certains cas prendre le pas sur la fréquence du son et influer sur la hauteur que l'auditeur perçoit effectivement… De plus, il existe de nombreuses caractéristiques pour la couleur, comme la « brillance » [8].

Inharmonicité (*warping*)

L'inharmonicité (distorsion fréquentielle harmonique, en anglais *warping* [17]) est une fonction W : fréquence × temps → fréquence, qui donne la fréquence réelle d'un partiel en fonction de la fréquence théorique qu'il aurait dû avoir si le son avait été harmonique. Elle traduit le fait que les fréquences des harmoniques ne sont pas forcément toutes exactement multiples de la fondamentale F. Pour une fréquence f donnée, on définit la distorsion fréquentielle harmonique pour f ainsi :

$$W_f = \left((rang_f(f_1).f, f_1), \cdots, (rang_f(f_P).f, f_P)\right)$$

avec $rang_f(f_i) = \left[\frac{f_i}{f}\right]$ ([x] désignant l'entier le plus proche de x).

Soit F la fréquence telle que la droite d'équation y = F.x passe au plus près des points $(rang_F(f_i).F, f_i)$. F est alors le paramètre « fréquence » du modèle et l'inharmonicité W est la version interpolée de W_F.

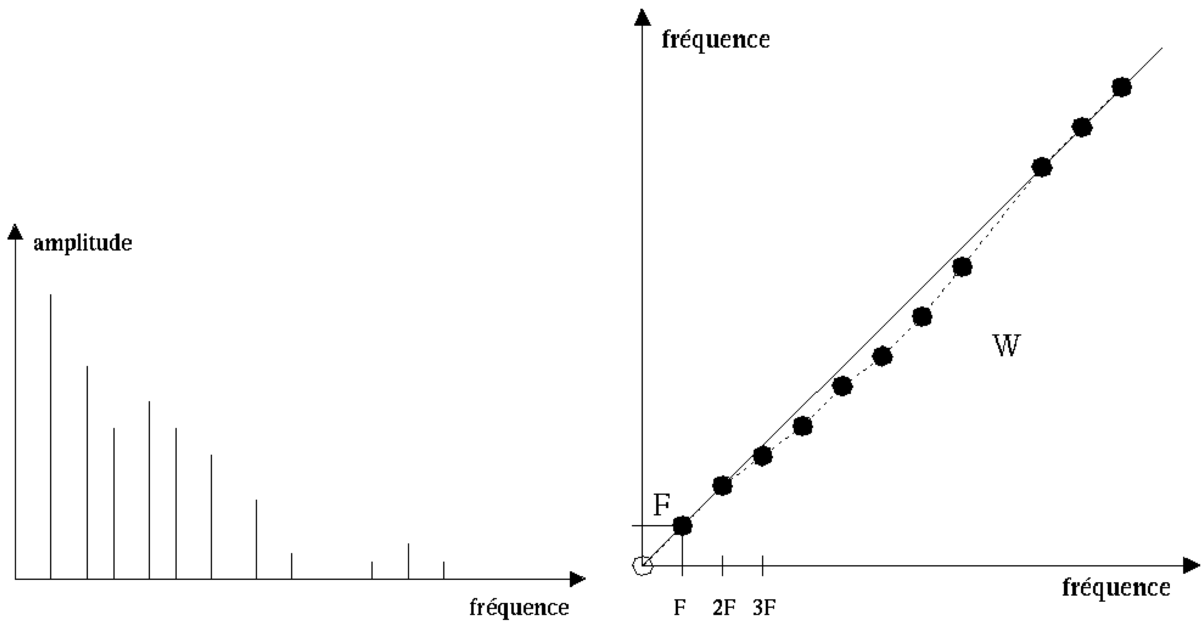

Figure 3 : Spectre d'un son inharmonique et paramètres F et W correspondants.

Pour tous les sons harmoniques W(t) = **Id**, c'est-à-dire que $\forall t, W(f,t) = f$. Il existe cependant des sons présentant une inharmonicité naturelle, comme les pianos, les gongs ou les cloches par exemple.

3.4 Equations du modèle

A partir des quatre paramètres du modèle SAS, on peut calculer l'amplitude du signal en fonction du temps :

$$a(t) = A(t) \frac{\sum_{p=1}^{P} C(W(pF(t),t),t) \cos(\varphi_p(t))}{\sum_{p=1}^{P} C(W(pF(t),t),t)} \tag{3}$$

où $P = \max_t \left\{ \left\lfloor \frac{F_{max}}{F(t)} \right\rfloor \right\}$ (F_{max} est la plus grande fréquence audible) et

$$\varphi_p(t) = \varphi_p(0) + 2\pi \int_0^t W(pF(u), u)\, du \tag{4}$$

Ces équations sont les versions « structurées » des équations 1 et 2, et toutes ces équations nécessitent des temps de calcul sensiblement identiques.

3.5 Limitations du modèle

Le modèle SAS permet de (re)produire une vaste étendue de sons, à condition qu'il s'agisse de sources sonores monophoniques. Une forte réverbération sur une source nuit à sa bonne modélisation. Il est de plus nécessaire d'introduire des structures symboliques comme l'union des sources pour réaliser, par exemple, une polyphonie. Ce modèle se généralise très bien pour les bruits, puisque tout bruit peut être construit comme un bruit blanc filtré (on dit aussi coloré) et ayant une certaine amplitude. L'amplitude et la couleur subsistent donc pour les bruits, et suffisent à les définir. Un bruit blanc a bien une couleur blanche (C = 1), et tous les bruits baptisés par analogie avec un spectre lumineux vérifient cette correspondance terminologique. En revanche les sons très brefs (transitoires) ne peuvent pas être modélisés sous forme spectrale, et ne peuvent donc pas *a fortiori* être produits par ce modèle.

3.6 Extraction des paramètres

Afin de reproduire des sons naturels, il est nécessaire de disposer d'une méthode d'analyse extrayant les paramètres du modèle à partir de ces sons. Pour le modèle SAS, nous utilisons actuellement le logiciel *InSpect* [18] qui extrait les partiels des sons et peut les structurer sous la forme requise. La figure 6 illustre les résultats produits par *InSpect* sur une guitare et un saxophone alto. Pour la guitare on observe bien une décroissance rapide de l'amplitude. La légère augmentation de la fréquence (*portamento*) vers la fin est due au fait que l'instrumentiste avait réalisé un « tiré » sur la corde. Sur le saxophone on voit nettement un *tremolo* (variation sinusoïdale d'amplitude) ainsi qu'un léger *vibrato* (variation sinusoïdale de fréquence). Il se trouve que sur ces exemples la couleur est quasiment constante avec le temps et que ces instruments sont parfaitement harmoniques (d'où W(f) = **Id**). La figure 4 montre des couleurs possibles pour deux voyelles de la langue française, et la figure 5 représente la voyelle « a » chantée sur trois notes.

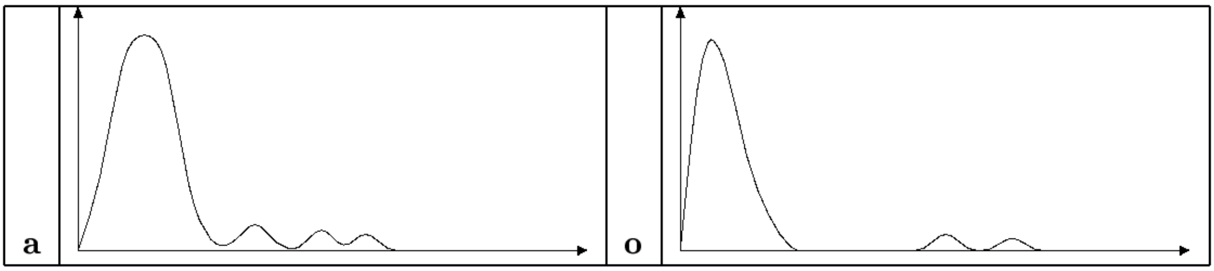

Figure 4 : Exemples de couleurs pour deux voyelles (seule la variation en fréquence C(f) est représentée, C étant quasi-constante dans le temps).

Un projet est en cours au sein du SCRIME pour extraire de même les paramètres d'amplitude (volume) et de fréquence (hauteur) contenus à l'intérieur de fichiers purement musicaux, tels les fichiers MIDI. La sortie du programme obtenu est un ensemble de couples (A_i, F_i) reflétant fidèlement les variations de ces paramètres pour chaque note de la pièce.

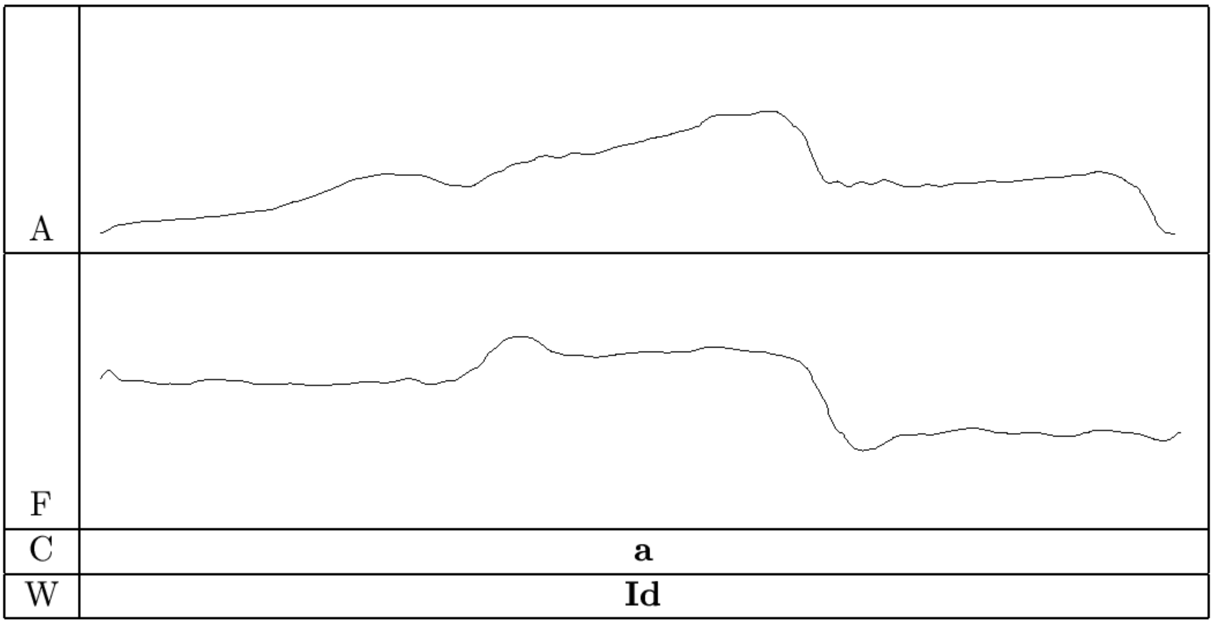

Figure 5 : Exemple de voix chantée dans le modèle SAS.

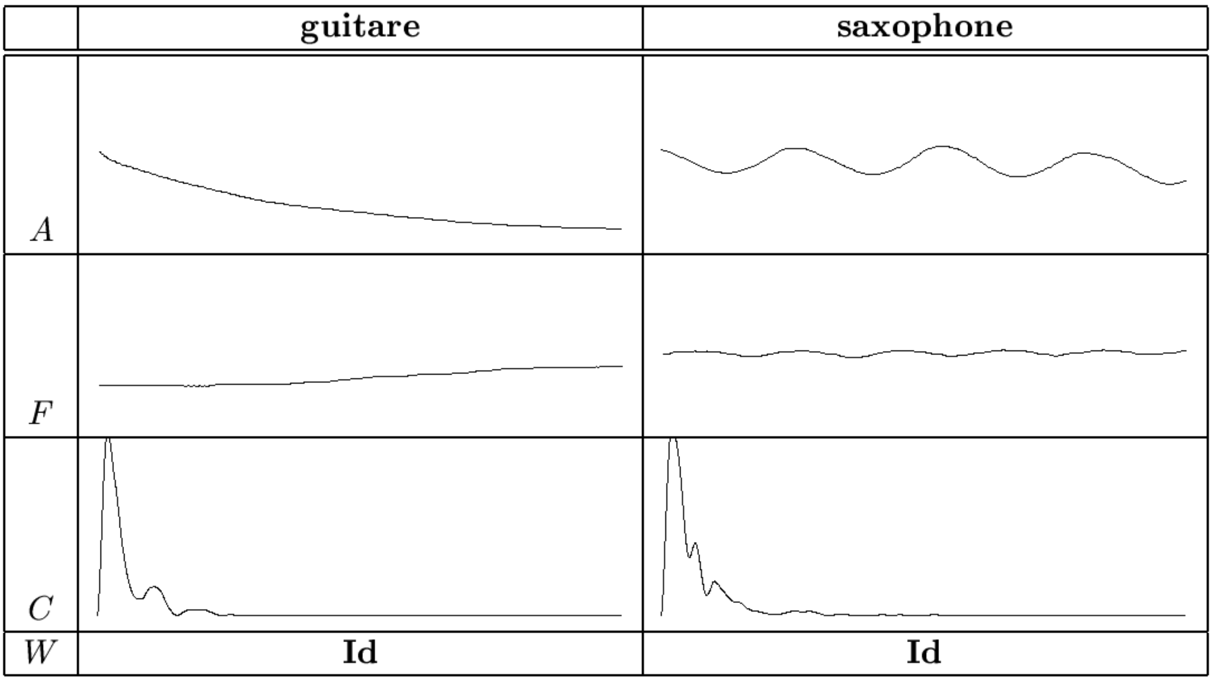

Figure 6 : Exemples d'instruments dans le modèle SAS. De haut en bas sont représentés les quatre paramètres du modèle : l'amplitude et la fréquence en fonction du temps A(t) et F(t), la couleur en fonction de la fréquence C(f), quasi-constante dans le temps pour ces deux exemples, et enfin l'inharmonicité qui vérifie W(f,t)=f car les deux sons sont harmoniques.

On remarque alors que la distinction entre paramètres sonores et musicaux n'est plus très claire. En effet, dans les exemples de la figure 6, le *portamento* effectué par le guitariste est un paramètre à la fois sonore et musical. Il en va de même pour le *tremolo* et le *vibrato* du saxophone. Pour l'exemple de la voix chantée illustré par la figure 5, on peut voir nettement la dynamique et la mélodie du chant dans les paramètres A et F du son.

4 Unifier son et musique

On considère des sources sonores monophoniques, sans acoustique propre et qui restent à spatialiser, la position spatiale étant d'ailleurs un paramètre aussi musical que sonore. Les paramètres sonores du modèle SAS sont très proches des paramètres musicaux et des macrostructures musicales. On peut de plus traduire la plupart des manipulations sonores et musicales sous la forme de variations plus ou moins rapides des paramètres du modèle. Cette section montre la proximité des paramètres sonores, des paramètres musicaux et des macrostructures musicales, ainsi que leur utilisation dans le cadre d'une composition multiéchelle, tout en se limitant cependant au modèle sonore permettant cette composition.

4.1 Temps, tempo et rythme

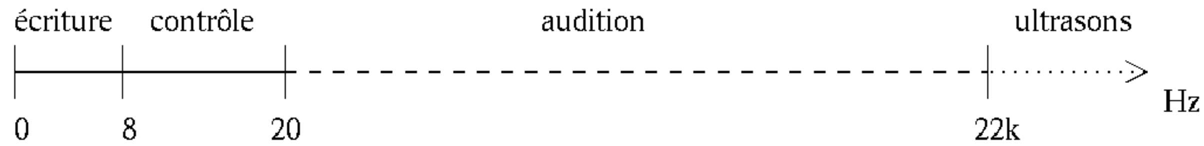

Figure 7 : Echelle des fréquences de variation des paramètres.

On peut schématiquement diviser les fréquences en quatre grands domaines, comme indiqué sur la figure 7. Une telle division a déjà été utilisée par Vaggione [19]. Au-dessus de 22 kHz s'étend le domaine des ultrasons, dont les fréquences sont imperceptibles par le système auditif humain. Entre environ 20 Hz et 22 kHz se situe notre zone d'audition. Ce sont des fréquences trop rapides pour être humainement contrôlées. Les paramètres du modèle SAS sont des grandeurs variant en fonction du temps et peuvent à ce titre être considérés comme des signaux. En aucun cas ils ne peuvent donc varier avec une fréquence très supérieure à 20 Hz, puisque dans ce cas leur variation deviendrait audible et bouleverserait les structures internes du son : c'est le phénomène de modulation, base de méthodes de synthèse sonore célèbres, mais non souhaité ici car il remet en cause la cohérence des paramètres, notamment sur le plan perceptif.

Par conséquent tous les contrôles effectués sur le modèle se situeront en dessous du seuil de l'audition et seront donc « lents ». En dessous de 8 Hz, correspondant à un *tempo* d'environ 500 battements par minute, on peut espérer noter de manière concise les variations des paramètres : c'est la zone de l'écriture, apparaissant naturellement dans le processus de composition. Au-dessus de 8 Hz, on entre dans le domaine du contrôle, de l'interprétation et des « effets ».

Le modèle SAS se place dans les domaines à la fois du contrôle et de l'écriture : du son (échelle microscopique) à la musique (échelle macroscopique) tout n'est qu'une question de vitesse de variation des paramètres perceptifs. De plus le contrôle du temps et de la macrostructure musicale correspondante, le rythme, peut s'y faire très simplement sous la forme d'un étirement ainsi : $t' = k(t)\, t$. On parle de dilatation ou de contraction temporelles suivant que k est inférieur ou supérieur à 1.

4.2 Amplitude, volume et dynamique

Nous avons déjà montré le lien étroit entre amplitude et volume dans la section 3. Les variations les plus rapides de l'amplitude contribuent au « grain » ou à la « rugosité » du son, tandis que les variations les plus lentes constituent la dynamique de la musique. En effet, les termes de *crescendo* et *decrescendo* s'expriment naturellement en terme de croissance ou de décroissance du volume, et donc de l'amplitude. On peut facilement modifier musicalement l'amplitude d'un son avec une formule du type : $A'(t) = k(t)\, A(t)$. Si k est une constante, on réalise ainsi une simple amplification. Si k est une sinusoïde avec une fréquence aux alentours de 10 Hz, l'effet musical obtenu est un *tremolo* à la fréquence correspondante. Si les variations de k sont lentes et monotones (au sens mathématique du terme), on obtient un « fondu » (*fade in* ou *fade out*). Les termes d'amplification, *tremolo*, « fondu », et bien d'autres se ramènent à une variation plus ou moins rapide de l'amplitude. Tout n'est qu'une question d'échelle dans la fréquence de variation du paramètre A.

4.3 Fréquence, hauteur et mélodie

Il existe de même un lien étroit entre la fréquence et la hauteur. Les variations les plus rapides de la fréquence contribuent à la sensation de « scintillement » du son, tandis que les variations les plus lentes constituent la mélodie de la musique. On peut facilement modifier musicalement la fréquence d'un son avec une formule du type : $F'(t) = k(t)\ F(t)$. Si k est une constante, on réalise ainsi une transposition. Si k est une sinusoïde avec une fréquence aux alentours de 10 Hz, l'effet musical obtenu est un *vibrato* à la fréquence correspondante. Si les variations de k sont lentes et monotones (toujours au sens mathématique du terme), on obtient un *glissando* ou un *portamento*. Si ses variations ne sont pas monotones, on peut obtenir une trille. Les termes de transposition, *vibrato*, *glissando*, *portamento*, trille, et bien d'autres se ramènent à une variation plus ou moins rapide de la fréquence. Tout n'est, à nouveau, qu'une question d'échelle dans la fréquence de variation du paramètre *F*.

4.4 Couleur et timbre

Si le volume et la hauteur sont les deux paramètres les plus classiques de la musique occidentale (les seuls notés en plus du temps dans le solfège classique d'ailleurs), il ne faut pas oublier la couleur, fondamentale dans les musiques orientales. Pour les sons harmoniques, la variation de la couleur avec le temps constitue d'ailleurs à elle seule le « timbre », terme qui reste le plus difficile mais aussi certainement le plus intéressant à définir. Sur le plan musical, on peut noter l'importance des couleurs d'orchestres, chaque instrument possédant ses propres couleurs qui se mélangent plus ou moins harmonieusement aux autres. On parle de même de la coloration induite par un système de diffusion ou par une salle de concert. La modification de la couleur (coloration) peut se faire ainsi : $C'(f,t) = k(f,t)\ C(f,t)$. Si k est constant dans le temps, on obtient un filtrage classique où k est la réponse fréquentielle du filtre (sa couleur…). Si k varie, on peut créer des sons hybrides très intéressants.

4.5 Inharmonicité et harmonie

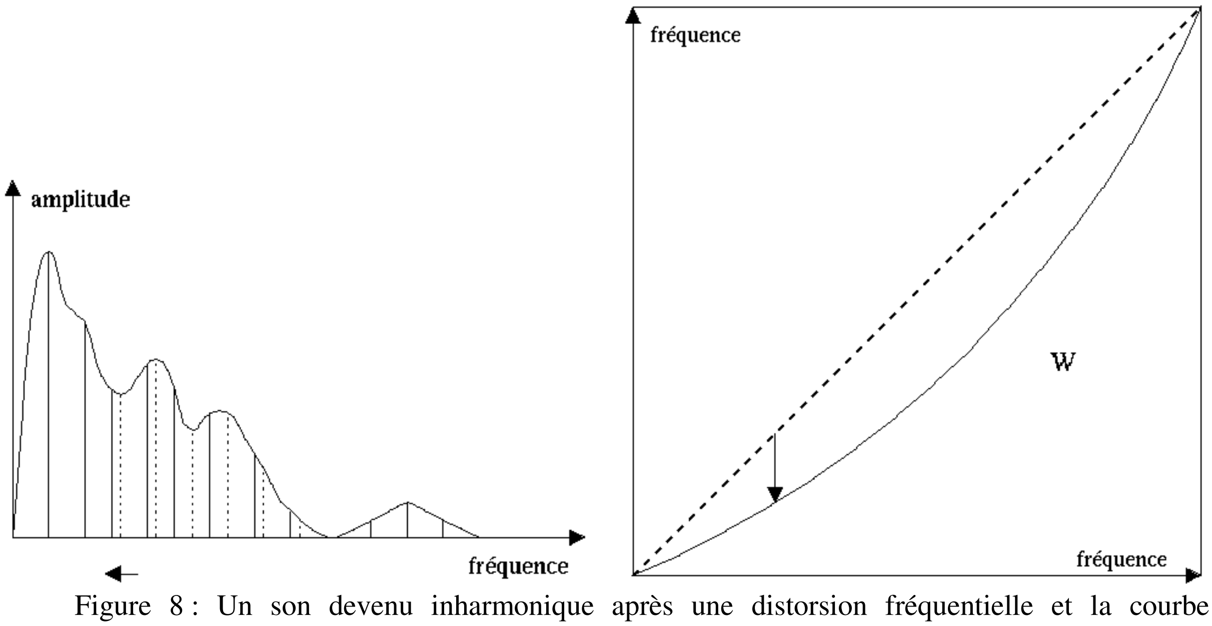

Figure 8 : Un son devenu inharmonique après une distorsion fréquentielle et la courbe d'inharmonicité *W*(f) correspondante.

L'inharmonicité du son joue un rôle important dans l'harmonie musicale [20]. Par exemple il résulte d'un accord quasi-harmonique une impression de fréquence, mais aussi d'inharmonicité. Elles coïncident avec les paramètres *F* et *W* du son résultant. Afin de modéliser un tel accord, on peut par exemple hiérarchiser le modèle en l'étendant avec une structure symbolique de type union. Un accord (une polyphonie) de N sources peut alors s'exprimer ainsi :

$$\bigcup_{i=1}^{N}\{(A_i, F_i, C_i, W_i)\}$$

Pour un accord quasi-harmonique, il est possible de retrouver des paramètres A, F, C et W à partir de ceux des différentes sources, en considérant par exemple les N couples (F_i, A_i) et en utilisant les méthodes décrites dans la section précédente.

L'inharmonicité est probablement le paramètre le moins bien connu des quatre et nécessiterait à lui seul une étude complète.

5 Conclusion

Dans cet article nous avons présenté le modèle de Synthèse Additive Structurée (SAS) favorisant l'unification entre musique et son à un niveau sous-symbolique, en définissant le son comme l'évolution temporelle de paramètres proches à la fois de la perception et de la terminologie musicale.

Il s'adresse à des compositeurs désirant écrire des pièces tout en utilisant la possibilité de modifier intimement la structure des sons, et ce en utilisant le même vocabulaire musical. Il sera utilisé dans des applications ou on ne veut jamais considérer les sons comme des « boîtes noires » sur lesquelles on applique des fonctions abstraites, mais plutôt comme une matière musicale à sculpter à l'aide de paramètres [21].

La rapidité avec laquelle on s'autorise à modifier les paramètres du modèle détermine l'échelle à laquelle on se place musicalement, allant des macrostructures musicales traditionnellement notées au contrôle du timbre lui-même, et ce en continu, c'est-à-dire sans jamais se heurter à une frontière musique-son.

Nous avons montré que cette unification offre la possibilité de faire migrer les contrôles d'un niveau à l'autre, du microscopique au macroscopique, enrichissant la palette d'écriture du compositeur. Il reste à introduire les structures symboliques utiles à l'écriture musicale au niveau du son pour en faciliter la composition.

6 Références

[1] L. A. Hiller et L. M. Isaacson. *Experimental Music*. McGraw-Hill, New York, 1959.

[2] Pierre Schaeffer. *Traité des Objets Musicaux*. Oxford University Press, 1994.

[3] B. L. Vercoe, W. G. Gardner et E. D. Scheirer. Structured Audio: The creation, transmission, and rendering of parametric sound representations. *Proceedings IEEE*, 86(5):922–940, 1998.

[4] Nicolas Castagne et Claude Cadoz. Synthèse de sons par modèles physiques dans l'environnement CORDIS-ANIMA. *Actes des Journées d'Informatique Musicale (JIM98)*, section G3, 1998

[5] David L. Wessel. Low dimensional Control of Musical Timbre. Rapport technique IRCAM numéro 12, 1978.

[6] David L. Wessel. Timbre Space as a Musical Control Structure. *Computer Music Journal*, 3(2):45–52, 1979.

[7] Stephen McAdams. *Spectral Fusion, Spectral Parsing and the Formation of Auditory Images*. PhD thesis, CCRMA, Department of Music, Stanford University, USA, 1984.

[8] Jean-Claude Risset. Timbre et synthèse de sons. *Analyse musicale*, pages 9–19, 1986.

[9] James A. Moorer. Signal Processing Aspects of Computer Music – A Survey. *Computer Music Journal*, 1(1):4–37, 1977.

[10] Robert J. McAulay et Thomas F. Quatieri. Speech Analysis/Synthesis Based on a Sinusoidal Representation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34(4):744–754, 1986.

[11] Xavier Serra et Julius O. Smith. Spectral Modeling Synthesis: A Sound Analysis/Synthesis System Based on a Deterministic plus Stochastic Decomposition. *Computer Music Journal*, 14(4):12–24, 1990.

[12] Xavier Serra. *Musical Signal Processing*, chapitre « Musical Sound Modeling with Sinusoids plus Noise », pages 91–122. Studies on New Music Research. Swets & Zeitlinger, Lisse, the Netherlands, 1997.

[13] Tony S. Verma et Teresa H. Y. Meng. Time Scale Modification Using a Sines+Transients+Noise Signal Model. *Proceedings of the Digital Audio Effects Workshop (DAFX98)*, pages 49–52, 1998.

[14] Myriam Desainte-Catherine et Sylvain Marchand. High Precision Fourier Analysis of Sounds using Signal Derivatives. Rapport technique du LaBRI numéro 120498, 1998. URL http://www.labri.u-bordeaux.fr/Publications.

[15] Sylvain Marchand. Improving Spectral Analysis Precision with an Enhanced Phase Vocoder using Signal Derivatives. *Proceedings of the Digital Audio Effects Workshop (DAFX98)*, pages 114–118, 1998.

[16] Digidesign. *Manuel utilisateur de SoftSynth*.

[17] Gianpaolo Evangelista et Sergio Cavaliere. Dispersive and Pitch-Synchronous Processing of Sounds. *Proceedings of the Digital Audio Effects Workshop (DAFX98)*, pages 232–236, 1998.

[18] Sylvain Marchand. *InSpect*. URL http://www.scrime.u-bordeaux.fr, 1998.

[19] Horacio Vaggione. Transformations morphologiques : quelques exemples. *Actes des Journées d'Informatique Musicale (JIM98)*, section G1, 1998.

[20] Max Mathews et John Pierce. Harmony and Nonharmonic Partials. Rapport technique IRCAM numéro 28, 1980.

[21] Daniel Arfib. *Recherches et applications en informatique musicale*, chapitre 19 : « Des courbes et des sons », pages 277–286. Hermes, Paris, 1998.