

INSTRUMENT DE SYNTHÈSE GRANULAIRE DANS MAX/FTS

Par Todor Todoroff

todor@musique.fpms.ac.be

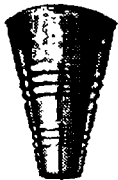
Avant de décrire le fonctionnement de l'instrument de synthèse granulaire, nous définirons les objectifs généraux que nous croyons nécessaires à la réalisation d'outils d'aide à la composition adaptés aux besoins des compositeurs de musique électroacoustique et acousmatique. Nous insisterons notamment sur l'importance à accorder au fonctionnement en temps réel, car les possibilités de commande gestuelle qui en découlent permettent d'envisager différemment la génération de morphologies sonores. Nous proposerons également un certain nombre de méthodes qui facilitent la génération de familles morphologiques cohérentes, en donnant une plus grande cohérence à l'évolution des différents paramètres de l'instrument de synthèse.

Alors que de nombreuses techniques de synthèse et de transformation sonore ont atteint un stade de maturité, nous voulons focaliser notre attention sur le développement de méthodes de commande plus intuitives et mieux adaptées aux besoins des compositeurs de musique acousmatique. Pour cela, il faut prendre en compte trois facteurs clefs: la conception de l'interface graphique, le fonctionnement en temps réel et l'accès gestuel aux valeurs des paramètres. Car la combinaison de ces points permet une approche de type instrumentale des outils de synthèse, de transformation et de projection du son.

1.1 L'interface graphique

La présentation d'un programme est primordiale et dépasse largement le simple objectif de la convivialité: elle conditionne profondément l'usage que le compositeur en fera. L'expérience montre en effet que l'interface graphique et le mode d'accès d'un outil orientent la manière dont il sera exploité et auront donc une influence sur le type de musique qui sera composée. L'interface graphique devrait permettre au compositeur de visualiser l'ensemble des paramètres importants du processus en cours d'un seul coup d'oeil, en évitant autant que possible la multiplication des menus et des fenêtres qui accaparaient une partie de son attention et réduiraient l'intensité de son écoute musicale. Les possibilités graphiques de MAX/FTS (la version tournant sur l'ISPW, Ircam Signal Processing Workstation) sont à ce jour trop limitées pour atteindre ce but et il nous faudra créer des objets graphiques plus performants lorsque la nouvelle version de MAX/FTS le permettra. Ces objets de visualisation et de commande graphique seront multidimensionnels. Ils s'inscriront dans la lignée de la fenêtre d'interpolation du

isea95@er.uqam.ca T : (514) 990-0229



ISEA95
montréal

SYTER (SYstème Temps Réel du Groupe de Recherches Musicales, Paris) et de la formalisation rencontrée dans le logiciel Animal [Lindemann, E & de Cecco, M.].

1.2 Le temps réel

Tout comme l'aspect visuel, le temps que doit attendre le compositeur entre l'instant où il modifie ses paramètres et celui où il peut écouter le résultat sonore influence son attitude. Un temps d'attente trop long ralentit le processus créatif et freine la recherche des paramètres optimaux. Le compositeur a alors tendance à accepter un résultat approximatif ou à abandonner sa recherche, alors qu'un temps de réaction court ou immédiat aurait au contraire stimulé sa curiosité. Dans la situation idéale du temps réel, le musicien n'hésite plus à essayer de nouvelles combinaisons de paramètres car il entend immédiatement le résultat musical qui en résulte.

1.3 L'interface gestuelle et l'approche instrumentale

Historiquement, les compositeurs de musique acousmatique ont été habitués à manipuler physiquement leurs instruments: enregistreurs à bande magnétique, matériel de studio doté de boutons, de potentiomètres, etc. Ils ont acquis une pratique compositionnelle basée sur l'accès simultané aux divers paramètres des appareils d'enregistrement et de traitement de son. Il ne suffit donc pas de leur offrir un programme puissant; il faut également, sous peine de limiter leur créativité, penser l'interface de manière à leur permettre un accès simultané aux différents paramètres. Or, si la plupart des logiciels disponibles font un effort de présentation, ils souffrent souvent d'un manque d'accès aisé aux valeurs de leurs paramètres. En effet, la solution souvent adoptée, qui consiste à afficher à l'écran des potentiomètres virtuels manipulés à l'aide de la souris, ne permet pas de modifier plusieurs paramètres simultanément. Ce point est cependant d'une importance capitale car la musique concrète

fait largement appel à la notion de morphologie, c'est à dire de variation continue des caractéristiques des sons.

Cette approche conditionne la musique que ces outils contribuent à composer car ils conduisent à assembler des séquences de sons stables plutôt que d'explorer les innombrables possibilités de transformations évolutives des sons. Nous voulons au contraire inciter les compositeurs à travailler les évolutions morphologiques des sons. Car c'est une des spécificités importantes de la musique acousmatique.

Si on propose au compositeur un programme fonctionnant en temps réel et une bonne interface gestuelle, tel un boîtier de potentiomètres MIDI qu'il peut manipuler simultanément ou encore une interface de type "Data Glove" (nous utilisons le gant *Power Glove* de Mattel qui, associé à l'interface *SensorLab* de STEIM, envoie à l'ordinateur des informations sur la position de la main et la flexion des doigts), on lui ouvre le monde des instruments virtuels. Un monde où il pourra construire son instrument imaginaire et choisir la manière dont il désire en jouer. Un monde qui lui révélera de nouvelles possibilités expressives.

GÉNÉRATION DE MORPHOLOGIES SONORES COHÉRENTES

L'expérience acquise en composant avec des instruments analogiques et numériques nous a convaincu de l'intérêt musical qu'il y a à corrélérer l'évolution temporelle des différents paramètres des outils de synthèse et de transformation sonore. Il semble en effet que l'auditeur perçoive l'existence des lois qui guident l'évolution des matériaux sonores. Cela se comprend lorsque l'on sait que les structures cérébrales qui régissent notre perception auditive se sont presque essentiellement formées à l'écoute de sons d'origine naturelle ou mécanique. Or ceux-ci présentent le plus souvent un haut degré de corrélation entre

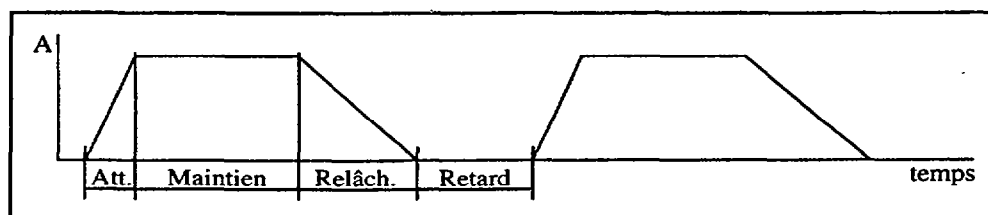


Fig. 1A: Mode normal

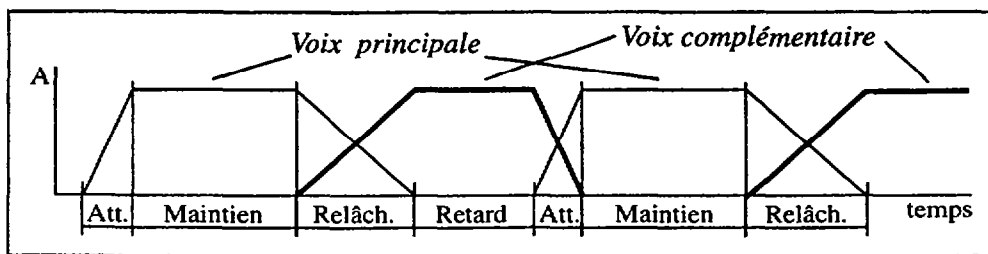


Fig. 1B: Mode fondu-enchaîné

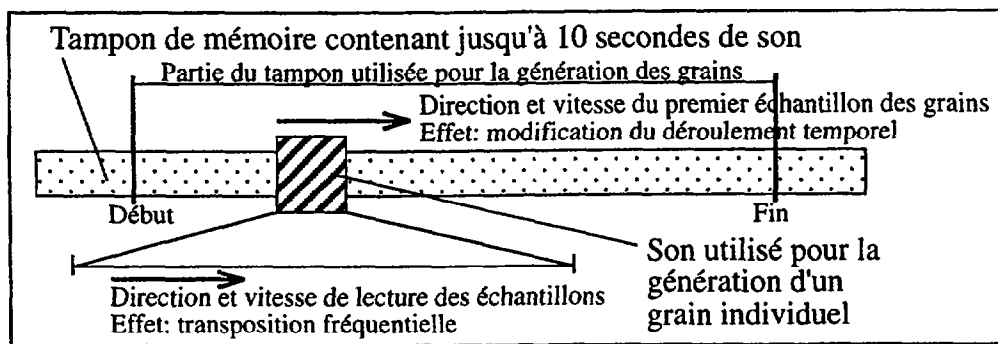


Fig. 2: Transposition et modification du déroulement temporel à partir du tampon

amplitude, hauteur et timbre. Corrélation que l'on peut attribuer au fait que ces paramètres ne sont que les diverses facettes d'une dissipation d'énergie mécanique sous forme acoustique. C'est ainsi, par exemple, que James Beauchamp a pu synthétiser avec beaucoup de réalisme des sons de trompette en exploitant les lois qui lient amplitude et richesse harmonique dans les sons naturels.

Bien que les sons d'origine numérique ne soient plus tenus de respecter ces règles de dissipation d'énergie, il nous semble utile de réduire dans certains cas le domaine de variation des paramètres d'un processus sonore en créant artificiellement des lois de variations conjointes qui réduisent le nombre de degrés de liberté du système et créent différents sous-domaines qui correspondent chacun à des familles distinctes de morphologies. Ces lois implicites sont perçues par l'auditeur et contribuent à donner plus de crédibilité aux morphologies ainsi générées.

Nous pensons qu'une architecture logique à deux niveaux peut aider les compositeurs à générer des morphologies cohérentes sans limiter le potentiel sonore des instruments:

(1) au niveau inférieur, on implante des algorithmes de synthèse ou de transformation sonore qui permettent de modifier les paramètres de manière totalement indépendante, en s'efforçant d'étendre leur plage de variation au-delà des limites habituelles.

(2) au niveau supérieur, on crée plusieurs instruments qui commandent de manière différente l'algorithme sous-jacent, en limitant les domaines d'excursion de certaines variables et/ou en créant des liens entre certaines d'entre elles. Ceci équivaut à ouvrir plusieurs fenêtres sur le monde de toutes les possibilités sonores d'un algorithme: chacune permettant de l'examiner suivant un autre point de vue. Chaque instrument aide l'utilisateur à explorer avec plus d'efficacité un sous-domaine particulier régi par des lois spécifiques. En bout de chaîne, l'auditeur percevra intuitivement l'existence de ces lois et sera tenté de croire que le son qu'il entend aurait pu avoir une origine réelle.

Il convient bien entendu de donner aux compositeurs les moyens de modifier aisément les instruments du niveau supérieur en fonction de leurs goûts et de leurs besoins. Ils pourront ainsi choisir librement les limites auxquelles il désirent se mesurer. Un peu comme l'utilisateur d'une ligne de retard numérique,

qui peut soit décider d'explorer séparément les effets phasing, chorus et écho, soit passer continûment d'un effet à l'autre. Dans le cas de l'instrument de synthèse granulaire, on pourra choisir n'importe quel domaine qui ira de la synthèse granulaire à la génération de structures rythmiques en passant par le brassage et la fragmentation.

Instrument de synthèse granulaire

3.1 Généralités

Ce principe de synthèse est bien connu: il consiste à accumuler un grand nombre de courts événements sonores, les grains, afin de générer des structures sonores complexes. L'instrument se distingue cependant par la grande plage de variation des paramètres (la durée des grains peut varier de 10 millisecondes à plusieurs secondes) et par l'approche adoptée pour sa commande. L'accès est immédiat: on sélectionne le son sur lequel on désire travailler et l'on peut sans plus attendre le transformer en manipulant un boîtier MIDI de seize potentiomètres qui permettent de modifier en temps réel tous les paramètres de la synthèse: durée, enveloppe, amplitude, fréquence, durée du silence entre grains consécutifs et valeurs des facteurs aléatoires. Les échantillons nécessaires à la génération de chaque grain sont prélevés dans un son de 10 secondes stocké en mémoire qui peut être parcouru à vitesse variable ou de manière aléatoire. On peut générer de cette manière jusqu'à 32 voix de synthèse granulaire par carte ISPW. Elles sont ensuite réparties dans un espace stéréophonique, quadraphonique, hexaphonique ou octophonique au choix. Cet algorithme, qui autorise un contrôle précis des microstructures et une densité de grains importante, permet de générer une grande variété de trames sonores et de structures rythmiques et spatiales suivant les valeurs des paramètres et la nature sonore de l'échantillon de départ.

De plus, l'accès direct à tous les paramètres par l'intermédiaire d'un clavier et d'un boîtier de potentiomètres permet de modéliser en temps réel les évolutions morphologiques: transposition, raréfaction-densification, regroupement-dispersion, vitesse de lecture du fichier son, etc.

Il est également possible de commander d'un seul geste des évolutions complexes et coordonnées des paramètres de

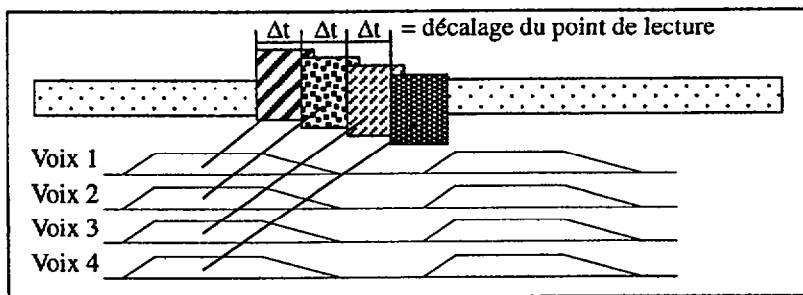


Fig. 3: Décalage du point de lecture des voix successives

manière à générer des morphologies sonores suivant les principes que nous avons exposé plus haut. Nous utilisons l'interface "Circ" développée à l'IRCAM par Gerhard Eckel qui permet, à l'aide de la souris, d'interpoler entre plusieurs configurations de paramètres localisées sur la circonférence d'un cercle. Nous étendons également ce concept en nous servant du Data Glove pour parcourir un espace tridimensionnel, sous-domaine de l'espace à 16 dimensions de toutes les combinaisons possibles des paramètres.

Enfin, l'accès MIDI autorise l'écriture et l'édition de l'évolution temporelle de tous les paramètres dans n'importe quel séquenceur du commerce, tandis que l'environnement MAX permet la programmation de structures de commande algorithmiques sophistiquées.

3.2 Fonctionnement

Chacune des 32 voix de synthèse granulaire permet de générer une succession de grains déterminés par trois catégories de paramètres: les paramètres d'enveloppe, ceux de transposition et ceux qui permettent de choisir l'échantillon sonore à utiliser pour la génération des grains. Des paramètres globaux gèrent l'activation, la désactivation et la synchronisation des 32 voix. D'autres servent à sélectionner les sons sur lesquels on désire travailler, à sauver ou à rappeler des configurations de paramètres, ainsi qu'à interpoler entre un certain nombre de ces configurations.

3.2.1 Les paramètres d'enveloppe et les deux modes de fonctionnement

Ils sont classiques: temps d'attaque, de maintien et de relâchement ainsi que retard entre le déclenchement des grains successifs. Deux modes de fonctionnement sont disponibles: le mode normal (voir figure 1A), qui permet de gérer 32 voix indépendantes, et un mode fondu-enchaîné (voir figure 1B)

qui permet de créer un continuum sonore en associant les voix deux à deux, voix principales et voix complémentaires, en faisant en sorte que chaque grain disparaisse dans un fondu-enchaîné qui coïncide avec la naissance d'un autre. Dans ce cas, les temps d'attaque, de maintien, de relâchement et de retard de la voix complémentaire correspondent respectivement aux temps de relâchement, de retard, d'attaque et de maintien de la voix principale.

Les temps de maintien et de retard, ainsi que l'amplitude de chacun des grains peuvent être soumis à des variations aléatoires réglables en intensité.

Remarquons que nous avons opté pour des variations non linéaires des paramètres de temps, de manière à permettre de couvrir toute la plage de variation utile à l'aide des 128 valeurs qu'autorisent les contrôleurs continus MIDI. Ceci correspond bien aux attentes perceptives: le passage d'une durée de 5 ms à 6 ms est significatif, celui de 2000 ms à 2001 ms ne l'est pas.

3.2.2 Les paramètres de transposition

Le son prélevé (voir ci-dessous) dans un son préalablement stocké en mémoire vive peut être transposé dans une gamme de plus ou moins deux octaves, par pas de demi-tons et de cents, à l'aide de potentiomètres ou à l'aide d'un clavier et de la molette de Pitch-Bend. La moitié supérieure de la course du potentiomètre correspond à une lecture de l'échantillon dans le sens normal, alors que la moitié inférieure correspond à la lecture à l'envers.

On peut aussi superposer un facteur de variation aléatoire de la transposition.

3.2.3 Les paramètres de gestion des échantillons

Le son utilisé pour générer chacun des grains n'est pas synthétique; il est prélevé dans un tampon mémoire de 10

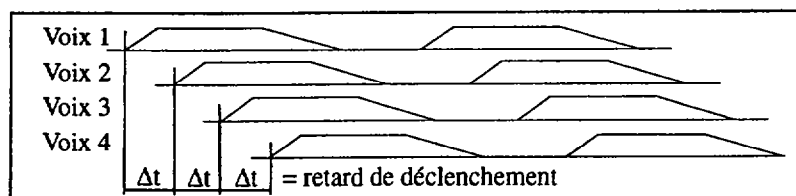


Fig. 4: Retard de déclenchement des voix successives

secondes où l'utilisateur a préalablement stocké le son sur lequel il désire travailler. Nous décrivons ici les paramètres qui permettent de choisir les fragments de son utilisés pour construire les grains. Tout d'abord on définit des points de début et de fin afin de délimiter la zone utile du son pour, par exemple, isoler un mot ou une syllabe dans une phrase. Un paramètre détermine ensuite le sens et la vitesse de déplacement du point à partir duquel les échantillons sont lus entre ces limites. La valeur 1 correspond à un déroulement temporel identique à celui du son original. Les valeurs supérieures à 1 provoquent la lecture accélérée du son, tandis que les valeurs positives inférieures à 1 réalisent des extensions temporelles. Il est possible de stopper complètement le déplacement (gel) ou de choisir des vitesses extrêmement faibles (il n'y a pas de limite inférieure), de manière à réaliser des extensions temporelles très impressionnantes. Les valeurs négatives correspondent à une lecture du son "à l'envers". Remarquons que ce sens et cette vitesse de déplacement sont distincts des paramètres similaires qui régissent la transposition (voir figure 2). Il est ainsi possible, dans le domaine temporel, de modifier indépendamment la fréquence d'un son et son déroulement temporel. La qualité sonore la plus naturelle est obtenue dans le mode fondu-enchaîné. Mais les paramètres d'enveloppe, indépendants des facteurs de transposition fréquentielle et de distorsion temporelle, peuvent modifier profondément le timbre: des grains très courts, sans variation aléatoire de durée, permettent par exemple de créer des effets similaires à la modulation en anneau.

Un autre paramètre permet d'associer à chaque voix de synthèse un décalage croissant du point de lecture dans le tampon mémoire (voir figure 3). On obtient ainsi des effets de phasing (filtrage en peigne), de chorus ou d'écho en fonction de l'importance du décalage. L'effet croît en intensité au fur et à mesure que l'on introduit des voix de synthèse supplémentaires, un peu comme lorsque l'on augmente le feedback d'une ligne de retard mais sans risque d'accrochage.

Enfin, un paramètre dose l'intensité de la variation aléatoire du point de lecture. Les petites valeurs permettent de générer des effets de chorus riches et naturels sans devoir recourir à la transposition. En augmentant la valeur de ce paramètre on obtient un effet de brassage aléatoire dans une zone du tampon de mémoire qui peut être étendue à l'ensemble des 10 secondes disponibles. Une amélioration future consistera à écrire un module externe qui simulera un tampon mémoire circulaire et permettra de traiter des sons directs.

3.2.4 Synchronisation des voix de synthèse et création de structures rythmiques

Comme nous l'avons indiqué plus haut, les voix de synthèse sont a priori totalement indépendantes. Lorsqu'une voix doit générer un nouveau grain, elle consulte les variables centrales où elle trouvera les dernières valeurs des paramètres, modifiées en permanence par les mouvements des potentiomètres. Il n'y a donc pas de processus central qui gère le déclenchement des

grains de chacune des voix. Les grains resteront cependant parfaitement synchronisés pour autant que les facteurs de variation aléatoire des temps de maintien et de retard soient nuls.

Il suffit donc de déclencher les voix à des moments précis pour créer des structures rythmiques stables. Et il suffit de jouer sur les facteurs aléatoires précités pour désynchroniser les voix plus ou moins progressivement.

Un ensemble de boîtes de message MAX, numérotées de 0 à 32, permettent d'activer et de désactiver les différentes voix individuellement ou globalement à l'aide de la souris. On peut également utiliser les touches d'un clavier de commande MIDI ou les boutons d'un boîtier MIDI programmable. Et il est à tout moment possible de resynchroniser les grains sans interrompre la synthèse. Une variable règle le retard entre l'activation des voix successives de manière à créer des structures rythmiques (voir figure 4). Selon l'assignation des voix de synthèse aux sorties audio, diverses structures spatiales se superposent aux structures rythmiques.

4 CONCLUSIONS

Cet instrument dépasse les limites habituelles de la synthèse granulaire. Il couvre un vaste champ opératoire qui inclut le brassage, la fragmentation et la génération de structures rythmiques et spatiales complexes. Son accès gestuel et sa facilité d'utilisation le mettent à la portée d'un plus grand nombre de compositeurs.

C'est une première étape dans la réalisation d'une gamme d'instruments qui offriront la possibilité de contrôler différentes techniques de transformation sonore de manière plus intuitive et plus naturelle. Nous pensons que cette approche favorisera la génération de morphologies sonores plus variées et plus expressives qui enrichiront la palette sonore des compositeurs.

© Todoroff 1995

Faculté Polytechnique de Mons
& Conservatoire Royal de Musique de Mons
Recherche financée en Belgique par la Région Wallonne

Références

- Jones, D. and Parks, T., "Generation and Combination of Grains for Music Synthesis", *CMJ Vol. 12 (2), Summer 1988*, pp. 27-34.
Lindemann, E. & de Cecco, M., "Animal: Graphical Data Definition and Manipulation in Real Time", *CMJ, Vol. 15, No. 3, Fall 1991*, pp. 78-100.
Roads, C., "Granular Synthesis of Sound" In C. Roads and J. Strawn, Editors, 1985, *Foundations of Computer Music, Cambridge, Massachusetts: MIT Press*, pp. 145-159.
Teruggi, D., "The Morpho Concepts: trends in software for Acousmatic Music Composition", *Proc. ICMC 1994, Aarhus*, pp. 213-215.
Todoroff, T., "Outils de synthèse granulaire et de spatialisation du son à commande gestuelle dans l'environnement MAX/FTS", *JIM '95, Paris, Institut Blaise Pascal, LAFORIA 95/13*, pp. 101-108.
Todoroff, T., "Real Time Granular Morphing and Spatialisation of Sounds with Gestural Control within MAX/FTS", *Proc. ICMC 1995, Banff*, pp. 315-318.
Truax, B., "Time-Shifting of Sampled Sound with a Real-Time Granulation Technique", *Proc. ICMC 1993, Tokyo*, pp. 82-85.