Conservatoire National Supérieur de Musique et de Danse de Paris

Formation Supérieure aux Métiers du Son

Mémoire de Fin d'Etudes :

REALISATION D'UN MIXAGE DE *POP-MUSIC* INTERACTIF

Présenté par :
Nicolas Déflache

Sous la Direction de :
François Pachet – Sony CSL Paris

Session de Décembre 2001

Remerciements

J'aimerais tout d'abord remercier les acteurs de la Formation Supérieure aux Métiers du Son, professeurs et équipe pédagogique, qui m'ont accompagné pendant mon travail autour de ce mémoire, mais aussi tout au long de mes études au Conservatoire. Merci à Laurent Millot pour sa relecture. Merci à Daniel Zalay pour son soutien et ses conseils en toute occasion.

Merci à Catherine de Boishéraud et à l'équipe du Service Audiovisuel qui ont su répondre à mes besoins techniques inhabituels et accueillir mes longues expérimentations.

Un grand merci à Julien Perraudeau pour les nombreuses nuits passées à écrire, enregistrer et monter la musique servant de support à ce travail. Merci aussi d'avoir bien voulu partager tes idées qui m'ont permis d'avancer dans ma réflexion sur l'interactivité. Grand merci aussi à Mathieu Coupat pour la même raison et pour son soutien.

Un grand merci à Olivier Delerue pour son idée, sa sympathie et son aide précieuse lors de la mise au point du système et au moindre problème informatique.

Enfin, merci à Sony CSL qui m'a permis d'expérimenter sur MusicSpace, et à François Pachet, mon directeur de mémoire, pour sa patience et pour la rigueur de sa démarche expérimentale.

Résumé

Nous avons pour objectif de permettre l'écoute active d'un enregistrement sonore. Nous étudions la place de l'interactivité dans l'évolution culturelle. Certaines musiques actuelles présentent des caractéristiques propices à son introduction, en particulier dans la *pop-music*. Nous répertorions les choix que nous pourrions donner à l'utilisateur, et réfléchissons aux conditions de réussite d'un *mixage interactif* : comment préserver la cohérence d'une mixage fragile aux points de vue technique et stylistique ?

Nous décrivons ensuite la réalisation d'un mixage interactif autour de *MusicSpace*, un programme développé par Sony CSL. Un titre de *pop-music* a été enregistré pour l'occasion. Lors de son écoute, l'auditeur peut agir sur les caractéristiques musicales ou sonores de l'enregistrement en déplaçant à la souris l'un des 4 curseurs appelés « Handles », qui interagissent entre eux. Une analyse de ce mixage souligne points forts et points faibles du procédé et donne quelques directions pour l'améliorer.

Mots Clés : Interactivité, Musique Interactive, Pop-Music, MusicSpace, Sony.

Abstract

Our purpose is to enable the active listening of a sound recording. We tried to figure out how interactivity may integrate in the cultural evolution of music. We make the list of choices that could be let to the user, and find out the conditions we should fulfill to succeed in making an *interactive mix*: how to preserve the integrity of the mix on a technical and stylistic point of view ?

Then we describe the production of an interactive mix with *MusicSpace,* a software developped by Sony CLS. A *pop-music* track has been recorded for that purpose. While listening, the user may change the musical caracteristics of the mix, or the technical caracteristics of the sound picture, by dragging one of the 4 available Handles. Some Handles interact with some others. An analysis of the mix highlights the good and bad points of the process, and gives directions to make it better.

Keywords : Interactivity, Interactive Music, Pop-Music, MusicSpace, Sony

Table des Matières

INTRODUCTION

CHAPITRE 1 L'INTERACTIVITE : JUSTIFICATION CULTURELLE 7

1.1 La musique se fait interactive 8

1.2 Des musiques ambivalentes 8

1.3 Les paramètres musicaux mobiles en *pop-music* 9

1.4 Les paramètres sonores mobiles en *pop-music* 10

1.5 Regroupement de paramètres 12

1.6 Des liens entre les points d'écoute 12

CHAPITRE 2 LA COHERENCE D'UN MIXAGE DE *POP-MUSIC* 14

2.1 Cohérence stylistique : les exigences culturelles 15

2.2 Cohérence de l'image sonore : les exigences techniques 15

CHAPITRE 3 REALISATION D'UN MIXAGE INTERACTIF AVEC « MUSICSPACE » 17

3.1 L'installation Interactive utilisée 18

3.2 La musique 19

3.3 Prise de Son, Montage et Mixage 21

3.4 Elaboration des « Handles » 21

3.5 Objets contrôlés par plusieurs Handles 28

CHAPITRE 4 ANALYSE DU MIXAGE OBTENU 35

4.1 La précision des réglages 36

4.2 Les Déplacements dans l'espace sonore 36

4.3 Fonctions de transfert du déplacement 38

4.4 Borner les déplacements des Handles 39

4.5 Des Handles à l'action clairement audible 40

4.6 Respect du style et reconnaissance de la signature 41

4.7 Problèmes survenant lors des Changements d'Orchestration ou d'Arrangement 41

CONCLUSION **43**

BIBLIOGRAPHIE & DISCOGRAPHIE **44**

ANNEXES

Annexe I **Installation autour de MusicSpace** **45**

Annexe II **Notions de prise de son** **50**

Annexe III **L'enregistrement Sonore : Interprétation, Adaptation et Fixation** **56**

Annexe IV **Contenu de la bande pré-mixée** **57**

Introduction

Interactivité :
Activité engageant des actions réciproques entre un utilisateur et un système informatique.

A chaque étape de la production traditionnelle d'un morceau de musique, compositeur, musicien puis ingénieur du son font continuellement des choix, pour aboutir au produit sonore qu'ils ont imaginé. L'auditeur découvre l'enregistrement dans un état figé, identique à celui présenté à un autre auditeur.

Nous nous intéresserons à la possibilité de réaliser des œuvres sonores à l'écoute desquelles l'auditeur ne serait pas passif : il pourrait influer sur le contenu proposé. Différentes écoutes du même morceau de musique seraient possibles. Différentes images de la même scène sonore seraient disponibles. L'auditeur pourrait choisir selon ses goûts, ou selon les propriétés acoustiques d'un lieu. Plus encore, on pourrait lui laisser la liberté d'entendre une image ou une musique non entièrement décrite par ses compositeurs : définie par certains pôles et par les moyens de passer d'un pôle à l'autre, elle comporterait une part d'imprévu, tout en restant fidèle au désir de l'artiste.

Pour réaliser ce projet, nous devons:
- être capable de construire des points d'écoute variés
- établir des règles de passage d'une scène sonore à une autre
- avoir les moyens de mettre en œuvre cette transition.

Compte tenu de l'originalité de l'idée, il nous a paru intéressant de réaliser un premier produit sonore de ce type. Nous avons contacté un laboratoire de recherche en informatique qui nous a fourni un programme capable de répondre à une demande abstraite de l'auditeur par un ensemble d'actions concrètes et coordonnées, et ainsi de mettre en œuvre cette interactivité.

Nous montrons en première partie que l'interactivité trouve sa place dans une évolution culturelle. Nous tentons de déterminer ce qui, dans le *pop-music,* pourrait être rendu interactif : nous répertorions les caractéristiques de la musique qui pourraient changer d'un mixage à l'autre.

Nous réfléchissons ensuite aux conditions que doit remplir le « mixage interactif » pour être correct du point de vue de l'ingénieur du son, et de celui du musicien. A l'issue de l'analyse de mixages existants se dégagent des critères techniques vérifiés dans la plupart des mixages de *pop-music*. Ils sont expliqués par des nécessités musicales et par le comportement perceptif de l'oreille.

Suit la description de notre travail avec « MusicSpace », lors de la production du « mixage interactif » d'un morceau de *pop-music*. Après un bref descriptif de l'outil et de l'installation mise en place, nous relatons l'enregistrement de la musique choisie. Puis nous décrivons la construction de chacun des Handles, ces contrôles laissés à l'auditeur. En dernier lieu sont décrites les actions croisées de Contraintes sur les Objets, qui font la force de MusicSpace.

Enfin, nous analysons le résultat et vérifions son adéquation avec les objectif proposés aux Chapitres 1 et 2. Nous en tirons des conclusions sur l'outil employé, sur la validité du fait interactif. L'analyse permet de dégager des directions pour une amélioration future du concept et de l'outil.

Chapitre 1

L'Interactivité : justification culturelle

Nous avons recherché si l'interactivité était présente dans la musique existante. Les ordinateurs et les sons électroniques se généralisent, et des systèmes de commande de ces sons sont partout développés. Il semble donc déjà possible d'agir sur le déroulement de sons ou de séquences en temps réel. Mais le procédé musical existe t'il ? Existe t'il des compositeurs qui, plutôt que d'écrire toute leur musique note à note, utilisent des règles, la conduisant à passer d'un état à un autre ?

Nous souhaitons permettre à l'auditeur de faire passer une musique d'un pôle à l'autre, lors de son écoute. Mais est-on capable de composer une musique qui se transforme de la sorte ? Dans la musique existante, et en particulier dans la *pop-music,* nous avons recherché des musiciens qui savent faire cohabiter et même interagir des éléments musicaux à première vue éloignés, et dont le style musical est basé sur cette interaction. Nous nommons leur musique « ambivalente », car elle se présente sous deux aspects différents, parfois plus.

Par l'analyse d'enregistrements, nous passons en revue les paramètres musicaux que le compositeur pourrait faire évoluer dans une œuvre interactive, ainsi que les choix sonores pris par l'ingénieur du son, qui pourraient y être intégrés.

Nous soulignons enfin la nécessité de grouper les paramètres mis en valeur, pour créer des contrôles signifiants et plus riches.

1.1 La musique se fait interactive

Dans la musique contemporaine, il existe des oeuvres musicales dans lesquelles le compositeur laisse à l'interprète une responsabilité sur le contenu musical. Dans « Jeux » de Gilbert Amy, écrite en 1973, deux hautbois puisent dans des réservoirs de notes à l'enchaînement structuré, qu'ils jouent selon des rythmes fixés en respectant certaines règles : ils sont placés à la limite de l'improvisation. Utilisant des représentations de plus en plus abstraites, les compositeurs contemporains définissent ainsi de moins en moins *directement* la musique qu'ils écrivent.

Parallèlement, la musique contemporaine se fait électroacoustique, incluant des sons *concrets* ou des dispositifs de traitement. Ces sons enregistrés, ces effets programmés, conduisent au contraire la musique vers la rigidité – qui est d'ailleurs l'un des problèmes majeurs lors de l'utilisation de « machines musicales ». On cherche donc à rendre vie à ces sons. Avec des logiciels tels que *Max*, largement utilisé par les compositeurs du monde entier, on permet à l'interprète d'agir sur le dispositif électroacoustique, par l'intermédiaire d'outils interactifs.

Il existe en France des ateliers mettant à la disposition des compositeurs contemporains des systèmes automatiques comme interface pour la diffusion de leurs œuvres : déclenchement des sons d'une installation sonore par le public, par exemple. A l'ACROE de Grenoble, spécialisé dans les rapports entre le geste et l'audition, on a mis au point un système informatique capable de modéliser le son d'un instrument. Un instrumentiste joue sur une interface maison. Le système génère le son en conséquence. Si l'on change les paramètres du générateur sonore, par exemple l'acoustique de la salle virtuelle, l'instrumentiste modifie son jeu en conséquence.

L'interactivité peut également s'opérer avec le public. Eric Redlinger a mis au point à Banff, au Canada, un dispositif permettant de piloter *Max* avec des *SMS* émis depuis un réseau de téléphones portables. La diffusion dépend ainsi des choix des auditeur. La musique entendue comporte une part d'imprévu. Introduisant des règles, le compositeur en définit les limites, adaptant son oeuvre à la situation interactive, pour la faire entrer dans sa pensée musicale: il est compositeur de musique interactive.

Lors de la prise de son ou du mixage de musiques de ce type, l'ingénieur du son est confronté à des problèmes inhabituels. Il a la charge d'adapter la volonté du compositeur aux possibilités permises par les techniques – interactives – actuelles. Interprétant l'œuvre de manière auditive, il pourrait également contribuer à l'œuvre, en cédant à l'auditeur des décisions prises habituellement lors de la prise de son.

1.2 Des musiques ambivalentes

Des œuvres contemporaines présentent plusieurs facettes, entendues l'une après l'autre ou simultanément. Agissant sur l'orchestration, l'harmonie ou encore le type d'écriture, le compositeur nous donne à entendre un même matériau en le présentant sous différents angles. Dans « Jazz Music sur Stockhausen », Jonathan Pontier nous fait entendre la pensée musicale de Stockhausen, oscillant entre l'austérité du sérialisme et l'atmosphère débonnaire des clubs de jazz new-yorkais, puis en réalisant la fusion.

Nous sommes très intéressés par certains courants actuels de la *pop-music* qui, pensons-nous, s'abreuvent de musique savante, introduisant des sons et des structures musicales de plus en plus complexes dans une musique à l'origine tournée vers les sens. Eux aussi ne se contentent plus d'une musique à un seul visage. Dans « I've seen it all », qu'elle interprète avec Thom Yorke (C), Björk superpose un orchestre symphonique à une boucle réalisée à partir du rythme régulier des roues d'un train sur la voie. Le titre exploite cette ambivalence comme un jeu, alternant des passages en quatuor à cordes simple avec des passages orchestraux accompagnés de la boucle transformée par un filtrage mobile, donnant à la musique un caractère résolument moderne, malgré sa grande ressemblance avec un duo de comédie musicale.

1.3 Les paramètres musicaux mobiles en *pop-music*

Dans cette étude, nous nous intéresserons plus particulièrement aux *musiques populaires existant principalement au travers de leur fixation enregistrée*. Il peut s'agir de Rock, de Pop, de Chanson mais aussi dans certains cas de Jazz, de Musique de Films ou de Musique contemporaine. Ces Musiques sont ici regroupées sous le terme générique et malheureusement inexact de *« pop-music »*.

1.3.1 Instrumentation

Dans la Musique de Film, on observe fréquemment qu'un même thème est exposé dans tout le film sous des présentations différentes qui à chaque fois s'adaptent au caractère de la scène. La chose la plus facile à faire pour atteindre ce résultat est de le faire jouer par des instruments différents.

Dans la partition d' « Amélie Poulain » de Jeunet (E), Yann Tiersen utilise 3 fois le thème de la « Valse d'Amélie ». Dans la première, l'accordéon éclairé au vibraphone prédomine, donnant un caractère gai. Un bouzouki finit par introduire une connotation exotique. La seconde, est introduite au vibraphone. Celui-ci est donc passé de son rôle de coloration à un rôle de soliste : il a changé de fonction. La musique glisse vers un accompagnement à l'orchestre de chambre. Le rythme est ponctué à la contrebasse et aux timbales, le tempo est plus lent : le caractère en est plus grave. La troisième version est simplement jouée au piano.

Pour densifier la musique, on peut ajouter de nouveaux instruments. Il existe plusieurs manières de les introduire : tous à la fois, un par un, par une augmentation de niveau étudiée, ou bien en utilisant un arrangement de transition - le musicien chargeant progressivement son jeu, par exemple...

Les possibilités de ré-instrumentation d'un titre de *pop-music* s'avèrent innombrables. Ainsi, il est sans doute possible de faire jouer deux versions d'un même titre avec deux sets instrumentaux différents. Pourquoi ne pas laisser à l'auditeur le choix du set écouté ?

1.3.2 Arrangement

Les musiciens de Jazz ont assez exhaustivement exploré la réécriture de l'arrangement d'un titre pour le mettre au goût du jour. Les standards de Jazz sont pour la plupart extraits de comédies musicales écrites pour orchestre symphonique et chantées dans les théâtres de Broadway. Ils ont pourtant été joués dans des arrangements pour des formations aussi éloignées que le Big-Band et le quartet, ce qui laisse supposer un type d'écriture totalement différent.

La manière d'arranger est déterminante vis à vis du résultat musical. Reprenons l'exemple d' « Amélie Poulain » : la dernière version, plus lente, est à 6/8, alors que les deux premières étaient à 3/4. Ce passage du binaire au ternaire contribue à donner à la musique un caractère posé.

Il paraît clair que les arrangeurs sont capables d'écrire des versions différentes d'un titre. Il peuvent agir sur la mesure, le tempo, le mode de jeu, la densité instrumentale (complémentarité rythmique, doublures...), et bien d'autres paramètres. Le résultat aura un caractère introverti ou extraverti, gai ou triste, rugueux ou tendre... Il pourrait être intéressant d'opposer ces arrangements.

1.3.3 Interprétation

Différences résultant d'un changement de technique (batterie aux balais ou aux baguettes, ajout de la charley...), d'intention (jeu vertical ou horizontal) ou de mode de jeu (guitare arpégée ou en accords). Sous ce contrôle peuvent également figurer des variantes sur les solos improvisés.

1.4 Les paramètres sonores mobiles en *pop-music*

Il s'agit de transformations modifiant les caractéristiques auditives de l'image sonore, sans toucher au contenu musical. Ces éléments sonores peuvent être moteurs d'une construction artistique.

1.4.1 Eclatement des sources

Les vieux disques de Jazz étaient enregistrés avec un seul microphone placé au centre du studio, à égale distance des musiciens. Lors d'un solo, l'instrumentiste s'approchait, réalisant une balance naturelle. Ce type de son est passé dans le contexte culturel et est encore apprécié par certains jazzmen, qui réalisent leurs enregistrements presque en monophonie.

Dans le mixage 5.1 d'un quintet de Chick Corea (H), les solistes sont disposés de manière écartée à l'avant de l'auditeur : trompette à gauche de l'enceinte gauche, basse au centre et sax sur l'enceinte droite. Les éléments de batterie sont disposés tout autour, mais restent assez localisables. Le piano est partout.

A une extrême, on a une image stéréo étroite : voix simplement au centre, instruments solistes et d'accompagnement peu larges et peu écartés. A l'autre extrême, l'image Surround est plus ou moins écartée. On dose la quantité d'arrière et la largeur des sources. Les deux images appartiennent pourtant au même style musical : le paramètre « degré d'éclatement des sources » existe donc, et sa valeur est fixée lors du mixage. Il semble que ce choix puisse sans problème stylistique être donné à l'auditeur.

1.4.2 Acoustique Virtuelle

Dans « There's more to life than this » (B), Björk chante sur une ambiance "Boite de nuit" marquée. Après quelques phrases, on l'entend se déplacer et entrer dans les toilettes, ou elle continue à chanter dans une ambiance différente – que l'on reconnaît à la réverbération portée sur la voix. Au loin, on entend la musique venant de la piste de danse, qui l'accompagne. L'arrangement de cet accompagnement est presque techno, donc dans un style différent du style habituel de la chanteuse. Ainsi, Björk utilise l'acoustique virtuelle pour conférer un caractère dramatique à sa musique, et pour éviter de brouiller sa signature.

Sur le refrain de « Beautiful Day » (D), U2 crée deux plans sonores très marqués entre le lead proche et les chœurs soumis à la réverbération très large qui fait partie de la signature du groupe. La profondeur du second plan est encore accentuée sur le pont par la présence de voix en échos, soumises au même traitement. Juste après le pont, on entend les chœurs du refrain en solo : on a l'impression de basculer un instant dans cet espace lointain, qui s'avère artificiel lorsqu'on perçoit des effets de phasing. Au refrain suivant, on perçoit ces chœurs un peu différemment et l'impression qui se dégage du refrain est changée.

On voit que la création d'un espace sonore fait partie du travail de composition d'un titre de rock. Le passage dans une autre acoustique peut changer la perspective qu'on a de la musique. Par la quantité et le type de réverbération, on peut modifier la taille apparente de la salle réelle ou virtuelle, ainsi que sa couleur. Les premières réflexions peuvent grossir une source, lui donner plus de chaleur ou de largeur. L'égalisation peut traduire l'éloignement. D'autres effets rendent les sons plus diffus ou plus doux.

On voit que les compositeurs créent ces espaces virtuels, et en programment la modification au cours de la chanson. L'effet « modification de l'acoustique virtuelle » existe, et il semble que son déroulement puisse être commandé par l'auditeur.

1.4.3 Implication de l'auditeur

On distingue deux écoles de mixage surround: le mixage de type Direct/Ambiance , dans lequel l'auditeur est spectateur, et le Direct-Sound-All-Around, dans lequel il est impliqué dans la scène sonore [1].

En Direct-sound-all-around, des sources sont localisables devant, derrière, et à l'intérieur de la tête. Certaines sources, larges et diffuses, sont de localisation floue. L'auditeur est enveloppé. On autorise des mouvements de source croisés. Les sources avant et arrières sont disposées de manière fusionnelle, et complétées d'effets de phase ou de delays. Dans la version 5.1 de l'album studio des Boys 2 Men (F), mixé selon l'école Direct-sound-all-around, l'auditeur est entouré : une voix Ad Lib est placée à l'arrière, le chœur est tout autour, tandis que le lead est à l'avant. Certaines voix se positionnent devant les enceintes. Le piano est enregistré avec 5 micros différents, panoramiqués sur les 5 enceintes.

En Direct/Ambiance, les sources sont positionnées principalement à l'avant. On recrée une ambiance de salle tout autour, en utilisant des réverbérations sur 5 canaux. L'auditeur est entouré par les sons. Les sources arrières n'attirent pas son attention. Les timbres des sources avant et arrière sont différents, pour éviter les effets de fusion. Les Eagles présentent leur hit « Hotel California » en situation de concert (G). Les sources principales telles que lead, guitares et batterie sont situées à l'avant. Il n'y a pas de son placé en extra largeur. A l'arrière on trouve des nappes de synthé, les cris du public, de la réverbération ou bien des congas. Pourtant on trouve des sources de fonction similaires à différents endroits de l'espace : le piano placé à l'avant est joué à la manière des cordes placé à l'arrière. Les cordes élargissent le son du piano par effet de fusion partielle avant/arrière, ce qui place ce mixage à l'intermédiaire entre Direct/Ambiance et Direct-Sound-All-Around.

Il semble que l'on soit capable de construire un mixage surround selon deux principes distincts, sans que l'un ou l'autre le soit attaché à un type de musique. Pourquoi ne pas laisser le choix du type de mixage à l'auditeur ?

1.4.4 Distance Perçue

Les instruments peuvent être enregistrés à différentes distances, ce qui modifie leur timbre (voir Annexe II.1 et II.3), et la perception que l'on a de leur environnement. Il est vrai que certaines pièces musicales demandent une acoustique particulière : la musique classique est en général enregistrée de plus loin que la *pop-music.* Pourtant, on peut trouver la même œuvre enregistrée, selon le goût du directeur artistique, à des distances variées.

Lorsque l'auditeur se trouve entouré de sources sonores, on peut lui donner le choix d'élargir le cercle des sources qui l'entourent. L'auditeur étant placé hors du groupe d'instrumentistes, il peut choisir la distance à laquelle écouter le concert. Un éloignement se traduit par une diminution du rapport Champ direct/Champ réverbéré, une modification du timbre, et les sources paraissent plus étroites.

1.4.5 Mouvements de sources autour de l'auditeur

Le DVD des Boys 2 Men cité plus haut présente beaucoup de mouvements. Des sources sont introduites à l'arrière ou sur les côtés, puis se déplacent en traversant la zone d'écoute. Dans une image large, on peut introduire des mouvements, comme des balancements avant/arrière ou gauche/droite, par orchestration (alternance tambourin/Charley) ou par changements de panoramiques. Cet effet spectaculaire est également très utilisé au cinéma.

Les compositeurs de l'image sonore sont ainsi capables de faire bouger les sources, et on pourrait imaginer laisser à l'auditeur le choix de rendre ces mouvements plus ou moins intenses.

1.5 Regroupement de paramètres

Lors du mixage, l'ingénieur du son a accès à un ensemble de paramètres de bas niveau. On ne peut pas laisser à l'auditeur final la gestion de ces paramètres, car les notions mises en jeu sont spécialisées. La fonction de l'outil de diffusion interactive est de grouper des paramètres de bas niveau, qui répondront par la suite à un contrôle d'un niveau plus élevé.

La recherche puis le regroupement de paramètres perceptifs ont déjà été réalisés dans le domaine musical. A Stockholm, Friberg & Al. [2] ont répertorié des paramètres de bas niveau concernant l'interprétation d'une pièce de musique classique. Ils ont décrit un ensemble de règles du type « le musicien insère une micro-pause entre les phrases », « lorsque la note est plus aiguë, le musicien a tendance à jouer plus fort », ou bien, « lorsque la note est plus aiguë, le soliste joue un peu haut ». Ces paramètres sont ensuite regroupés sous des commandes d'un niveau plus élevé du type « expressivité du jeu ». Dans un jeu expressif, le musicien est sensé exagérer les micro-pauses, se retenir de jouer fort dans les passages aigüs mais par contre les jouer vraiment haut, par exemple. Les auteurs ont mis au point un programme informatique appelé *Director Musices* qui fait jouer un instrument MIDI selon une partition préalablement analysée et des critères d'interprétation déterminés en temps réel par l'auditeur. C'est la modélisation d'une interprétation. Si *Director Musices* n'a pas la prétention d'égaler une vraie interprétation, le résultat est cependant remarquable.

Comme décrit précédemment, le mobile envisageable en *pop-music* interactive se situe au niveau :
- musical (modification ou addition de parties instrumentales)
- sonore (modifications de paramètres techniques du mixage)

Chacun des paramètres envisagés ne se retrouve pas forcément sous un seul curseur portant son nom : il est souvent intéressant de grouper des paramètres de nature différente, mais dont la modification d'ensemble va dans le même sens musical.

On pourrait imaginer par exemple un unique curseur, permettant une alternance entre deux mixages :
- intime : sources peu écartées, direct-sound-all-around, acoustique matte, arrangement doux, sources très présentes, solo tendre…
- extraverti : image large, acoustique réverbérée, direct-ambience, sources lointaines, arrangement dense…

1.6 Des liens entre les points d'écoute

Mettons à la disposition de l'auditeur des curseurs ayant un rôle déterminé. Il est libre de les manier à sa guise. On souhaite que quelque soit la position des curseurs, la musique soit toujours également intéressante, et le résultat sonore de qualité égale. Si un curseur commande le passage d'un état à un autre de l'un des paramètres décrit au 1.3 ou au 1.4, comment allons-nous réaliser ce passage ? Dans la musique existante, nous avons entendu des transitions par *superposition* ou par *alternance*.

On fait une transition par Superposition lorsqu'on ajoute de nouveaux éléments, sans modifier la musique existante. On peut prendre l'exemple du vibraphone venant colorer l'accordéon au deuxième couplet de la « Valse d'Amélie ».

Il y a transition par Alternance lorsqu'un groupe de sources ou un réglage de paramètres remplace un autre groupe ou réglage. C'est le cas lors du passage d'une acoustique à une autre, chez Björk. Lorsque les deux versions peuvent être entendues simultanément sans interférer, on peut procéder à un *tuilage*. Il suppose une complémentarité. On peut encore doser la rapidité d'évolution des paramètres, quel est le point de croisement et à quel moment il est atteint.

Lorsque le tuilage n'est pas possible ou pas souhaité, on fait une transition en *cut*. C'est le cas lorsque les chœurs lointains sont mis en valeur sur un refrain de « Beautiful Day ».

Conclusion de la 1ière partie

La volonté d'inclure des sons, des styles, des techniques d'écriture d'origines diverses dans la *pop-music* a toujours existé. Par l'analyse d'enregistrements existants, nous avons tenté de montrer qu'il existe souvent plusieurs manières d'arranger, d'orchestrer, d'interpréter, d'enregistrer ou de mixer la musique en restant dans un style donné. Il semble qu'il existe actuellement des compositeurs dont la signature [3] consiste à mettre en valeur cette ambivalence, et à mettre en place le plus possible d'interactions entre deux pôles musicaux. Cette pratique est en voie de passer rapidement dans le contexte : des artistes comme Radiohead ou Björk sont très médiatisés.

Pour l'instant, lors du mixage de ces musiques, les paramètres de bas niveau, tels que des égalisations, des envois d'effets ou des modifications de balance, sont automatisés un à un sur séquenceur ou « automation » de console. Le procédé est très long : il est difficile de rendre vivantes les transitions entre deux pôles éloignés du mixage car beaucoup de paramètres doivent être changés en même temps. La mise au point d'un outil de mixage pourrait aider ces musiciens à fabriquer ces transitions, et à augmenter encore les interactions. Les paramètres de bas niveau pourraient être regroupés sous des curseurs au comportement moins technique et plus musical.

Par ailleurs il existe une tendance des compositeurs à laisser une marge de manœuvre par rapport au résultat auditif qu'ils attendent de leur musique. Ils en définissent le style mais pas toujours précisément la réalisation, qui dépend de règles. Ces compositeurs souhaiteraient peut-être franchir le pas et formaliser le mode de reproduction de leurs œuvres, profitant d'un nouveau mode d'écoute actif : l'écoute d'un mixage interactif.

Chapitre 2

La cohérence d'un mixage de *pop-music*[1]

Nous voudrions permettre à l'auditeur de faire son choix entre différentes orchestrations et différents arrangements, en changeant la nature des sources ou en accordant à certaines une présence optionnelle dans le mixage. Nous voudrions également donner à l'auditeur la liberté d'agir sur l'espace sonore, changeant la position des sources, la manière de les enregistrer ou les traitements qui leurs sont appliqués. Quels sont les critères qui nous permettront de dire que le « mixage interactif » réalisé est réussi ?

Au niveau stylistique, l'importance culturelle du son dans les enregistrements de *pop-music* impose une certaine prudence lors de l'élaboration d'un mixage interactif. Peut-on faire n'importe quel son sur n'importe quelle musique ? Existe t'il, comme l'affirment certains, plusieurs manières de mixer un titre ? Quelles doivent-être les règles présentes dans le mixage interactif ?

D'un point de vue technique, notre objectif est que l'image sonore de tout mixage interactif vérifie les mêmes critères de cohérence que celle produite lors d'un mixage traditionnel. Nous répertorions les critères de cohérence de l'image sonore en *pop-music*. Nous décrivons les modifications auxquelles nous projetons de soumettre l'image interactive qui pourraient remettre en cause cette cohérence.

[1] Dans le sens indiqué au 1.3 : par *pop-music,* nous entendons « *musiques populaires existant principalement au travers de leur fixation enregistrée* ».

2.1 Cohérence stylistique : les exigences culturelles

Un style en *pop-music* consiste en un ensemble de stéréotypes, dont la définition se fait à partir du contexte musical présent à une époque donnée. D'après Emiliano Flores [3], l'évolution de la pop repose sur le passage de la « signature » des nouveaux artistes (qui apportent de nouvelles idées) dans le contexte musical, par le biais de la médiatisation. Les règles stéréotypées du contexte sont transgressées, générant de nouvelles signatures.

La *pop-music* est composée, arrangée et orchestrée en studio ou « home studio », au cours du processus d'enregistrement. Les compositeurs utilisent plus que dans toute autre musique les techniques sonores comme outil d'écriture.

En conséquence, pour qu'une musique soit de la *pop-music*, il est nécessaire :
- Si elle se réclame d'un style, qu'elle respecte les codes sonores attachés à son style
- Si elle est indépendante, qu'elle définisse des éléments de signature sonore
- Que tout ce qui la constitue fasse partie soit de sa signature particulière, soit du contexte

L'existence de ces exigences justifie le métier de directeur artistique.

2.1.1 Le respect des codes sonores

L'histoire du Rock est marquée par les évolutions de la facture instrumentale et de la technologie. De l'électrification du Blues à l'usage des plug-ins de ProTools dans le R&B actuel, les disques sont témoins de leur époque. Dans les années 30, on entend de la distorsion sur les disques de Blues Acoustique de Robert Johnson : la prise de son s'adapte à un média à la dynamique faible et compresse la musique avec les moyens du bord, c'est à dire à même le support d'enregistrement. Dans l'album « Music » de Madonna (A) on entend un *autotune* sur presque tous les titres. Il est utilisé ici non comme un moyen de corriger la justesse vocale, mais comme renfort de l'expression musicale. Utilisant les techniques de studio qui leur étaient contemporaines, les musiciens ont créé un son en même temps que leur musique, laissant l'auditeur faire progressivement un amalgame culturel entre type de son et style musical. L'enregistrement a ainsi apporté une richesse supplémentaire à la musique rock, et on peut dire que le son fait partie de cette musique [4].

Pour enregistrer la pop-music dans un style donné, on doit donc respecter les codes musicaux qui lui correspondent, mais aussi reproduire les codes sonores correspondant à ce style. Par exemple, on peut donner à la musique une connotation psychédélique en utilisant une « electric sitar » branchée dans un chorus et réverbérée. Le musicien pourra jouer en glissés sur une gamme indienne.

2.1.2 Définir une signature

La responsabilité de l'artiste est de faire des choix artistiques, qui constituent sa signature. Lors de la fabrication d'un mixage interactif, l'artiste doit s'assurer qu'il donne suffisamment de règles pour définir complètement sa signature.

2.2 Cohérence de l'image sonore : les exigences techniques

En *pop-music*, on n'est pas attaché, comme en musique Classique, à l'obtention d'un timbre naturel. Ainsi, les sources sont souvent enregistrées en champ proche. Par ailleurs, il est fréquent d'enregistrer les instruments séparément, pour travailler l'interprétation et le mixage de manière plus fine. L'objectif de l'ingénieur du son est en effet d'obtenir une « image sonore » qui soit à la fois cohérente et qui permette aussi l'identification d'éléments constitutifs de l'ensemble. Par l'habile combinaison des sources sonores, il atteint une image à la richesse sonore maximale.
A l'autre bout de la chaîne, les auditeurs ont des habitudes d'écoute précises quant à la qualité sonore. Bien qu'il soit très difficile de décrire de manière exhaustive et précise ce qui caractérise

un mixage de qualité, nous proposons ici quelques critères[2] qualifiés de « techniques » car concernant strictement l'ingénieur du son, par opposition aux critères musicaux évoqués plus haut. Ces critères sont discutables, et ne revêtent pas une forme définitive ou officielle. Néanmoins, on peut affirmer que c'est l'existence même de tels critères, indépendante des musiques mises en jeu, qui justifie le métier d'ingénieur du son.

a) L'Equilibre Spectral

L'équilibre spectral concerne la répartition des fréquences dans le signal mixé. Cette répartition est complexe pour plusieurs raisons :

- Chaque source en isolation peut avoir un spectre complexe ;
- Le mixage des sources fait intervenir un effet psycho-acoustique qui a une influence sur la perception de ce spectre : le masquage de fréquence ;
- Les spectres des sources individuelles évoluent dans le temps.

Un mixage équilibré spectralement devra en outre utiliser tout le spectre audible. Cet équilibre spectral est fragile : il est le résultat de la combinaison fine de toutes les sources, et nécessite une bonne connaissance des propriétés spectrales des instruments et de leurs règles de cohabitation.

b) L' Homogénéité (globale)

L'image sonore doit décrire un espace cohérent. Les sources doivent avoir des couleurs et des propriétés acoustiques *compatibles* (par exemple réverbération, égalisation, localisation, dynamique, etc.). Cette cohérence donnera une unité à l'image et permettra de réaliser des fusions. On peut par exemple vouloir que certaines sources se complètent au point de ne faire plus qu'une. La basse, qui est le lien entre rythmique et harmonie, doit ainsi parfois être fondue avec la grosse caisse. De même, il peut être souhaitable de fondre un saxophone, au début de son solo, dans la dernière note du chant. Une boucle rythmique peut se mélanger à la batterie, pour créer un timbre et un *feeling* original.
L'homogénéité dépend du choix des instruments, de la manière dont ils sont joués et enregistrés et des traitements qui leur sont appliqués.

c) La Balance (globale)

La balance d'un morceau dénote l'équilibre relatif des niveaux des différentes sources sonores. Une différence dans la balance peut modifier complètement notre perception du titre. Une « bonne » balance permet d'atteindre une richesse maximale lorsque les différents groupes fonctionnels (lead, harmonie, rythmique) se complètent. La réalisation d'une telle balance est aussi sensible aux effets de masquage mentionnés plus haut.

d) L'Intelligibilité (globale)

Certaines sources ont des besoins d'intelligibilité particuliers. Par exemple, un chanteur doit rester – en principe – intelligible. Pour réaliser cette contrainte il faut tenir compte des propriétés acoustiques particulières des voix chantées. Il faut ainsi éviter de masquer les fréquences les mieux perçues par l'oreille, et éventuellement limiter les effets (réverbération). Cette contrainte est générale, et peut s'appliquer à bien d'autres sources sonores.

e) La Définition (individuelle)

Il s'agit de la précision avec laquelle les timbres seront perçus. On peut souhaiter doter certaines sources de certaines propriétés sonores particulières, indépendamment des autres sources. Par exemple, on peut vouloir une voix douce, pure ou bien au contraire, saturée ou donnant l'impression de venir d'un tube en métal.
Les moyens techniques pour obtenir cette définition doivent bien sur prendre en compte les contraintes globales énumérées ci-dessus. Il peut arriver que la résolution de l'équilibre spectral nuise à la définition d'une source isolée : on doit alors faire des compromis. Il est parfois possible d'utiliser les effets de masque pour cacher la mauvaise définition d'une source : un son jugé mal défini entendu isolément peut paraître très beau dans l'ensemble du mixage.

---

[2] Ces critères sont la synthèse d'un savoir-faire enseigné au Conservatoire et glané au gré des expériences en studio. Nous avons tenté de les expliquer en se servant des notions exposées en Annexe II.

Chapitre 3

Réalisation d'un mixage interactif avec « MusicSpace »

Le *Sony Computer Science Laboratory - Paris* a entamé son travail sur le « mixage interactif » dans le cadre d'une réflexion sur l'écoute active. Imaginé entre 1998 et 2001 par et l'équipe de Sony, et en particulier par Olivier Delerue dans le cadre de sa thèse, le projet de « MusicSpace » est de permettre à l'auditeur de mixer sans rien connaître à la technique, ni à la musique. Il réagit à son écoute par le maniement de potentiomètres, pour obtenir un résultat plus proche de ce qu'il a envie d'entendre. La machine réalise au mieux son désir tout en respectant certaines règles, établies entre des objets sonores.

Le labo a mis « MusicSpace » à notre disposition. Le programme est encore en cours de développement. Nous décrivons ici le travail pratique que nous avons réalisé d'Octobre 2000 à Juin 2001 autour de cet outil.

En premier lieu, nous faisons une description concise du système de « mixage interactif » mis en place. Nous exposons les principes de fonctionnement de MusicSpace, et expliquons la manière dont nous avons décidé de l'utiliser.

Nous décrivons ensuite la musique que nous avons choisi de rendre interactive, ainsi que le processus d'enregistrement et de post-production traditionnels que nous avons suivi.

Nous abordons enfin le travail avec MusicSpace lui-même : définition des objets nécessaire au mixage, conception et programmation dans le logiciel des contrôleurs qui seront mis à la disposition de l'auditeur ; construction des règles qui régiront les rapports entre objets et contrôleurs.

3.1 L'installation Interactive utilisée

On trouvera une description plus complète de MusicSpace et de notre installation en Annexe I.

3.1.1 Principe de MusicSpace

MusicSpace [5] [6] est une télécommande pour un système de diffusion sur 360 degrés, permettant la spatialisation d'un ensemble de sources issues d'un multipiste. Chaque piste est représentée par un Objet. Les Objets sont disposés dans un espace graphique à 2 dimensions. Pour chacun, MusicSpace envoie des données de position angulaire et de distance à un spatialiseur externe.

- Représentation graphique: schématisation 2D de l'espace sonore disponible entre 4 enceintes disposées en carré.

- Avatar: représentation de l'auditeur, par rapport à laquelle les déplacements d'objets seront évalués.

Figure 3-1 Avatar

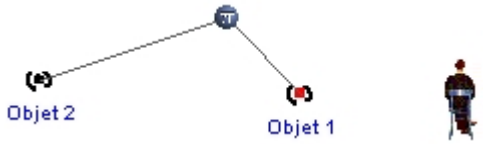
- Objet: représente une source sonore ou un quelconque paramètre de la console ou d'un module d'effets.

Figure 3-2 Objet

Un ensemble de règles appelées Contraintes permet de définir les interactions des Objets entre eux et par rapport à l'Avatar. Lorsque l'un est déplacé, les positions des autres sont modifiées en respectant ces règles.

- Contrainte: impose une relation entre les mouvements des objets qui lui sont soumis. Il ne s'agit pas toujours d'un lien linéaire ou non, mais d'un ensemble de conditions qui doivent être vérifiées. C'est par le biais des contraintes que seront exprimées les règles de composition de l'image.

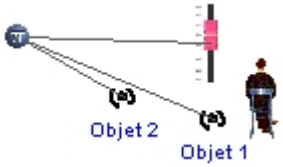
Les déplacements des Objets 1 & 2 sont liés par une contrainte « rapport des distances constant »



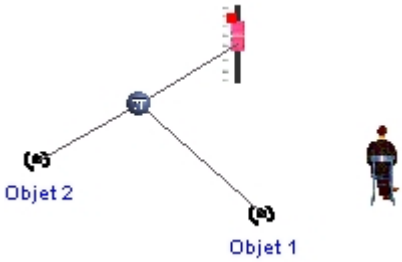
Lorsqu'on éloigne l'Objet 1 de l'Avatar, l'Objet 2 en est aussi éloigné dans le même rapport. Sa position angulaire p/r l'Avatar demeure inchangée

Figure 3-3 La Contrainte « rapport des distances constant » appliquée à 2 Objets

- Handle: Objet accessible à l'utilisateur, utilisé pour transmette des mouvements à d'autres Objets, cachés, par le biais d'un ensemble de contraintes.



Handle et objets sont contraints en « rapport des distances constantes »



Les 2 objets s'éloignent avec le Handle

Figure 3-4 Handle agissant sur 2 objets

3.1.2 Mode de fonctionnement

Nous avons choisi d'utiliser le mode MIDI de MusicSpace. Notre outil de diffusion multicanale est une console 03D. Nous utilisons des Objets Volume 03D spécifiques dont la distance à l'Avatar contrôle le gain sur chaque tranche. D'autres Objets Panoramique 03D pilotent les panoramiques surround. Nous utilisons des Objets MIDI génériques pour doser les départs auxiliaires. Nous disposons par ailleurs d'une Lexicon 300 et d'un Eventide H3000. Certains de leurs paramètres seront pilotés par l'intermédiaire d'Objets MIDI génériques.

3.2 La musique

Nous avons travaillé sur « Gena », une chanson rock-pop écrite par Fabrice Ducognon et Julien Perraudeau et enregistrée spécialement pour l'étude. A travers sa production interactive, nous avons tenté de mettre en valeur les idées apportées par les concepteurs de MusicSpace, de soulever les problèmes et de réaliser nos envies d'ingénieur du son face à un tel outil.

La musique se décompose en 3 groupes fonctionnels: lead, harmonie et rythmique. Afin de s'affranchir des problèmes stylistiques, on a choisi un groupe que les musiciens eux-mêmes classaient difficilement.

Nous avons enregistré le titre sous deux arrangements distincts, nommés « Acoustique » et « Electrique ». Nous avons cherché à ce que chacun apporte un éclairage différent, et à ce que les deux puissent être joués simultanément. Nous avons distingué les arrangements par :
- l'orchestration
- l'écriture
- la décomposition rythmique de base.

Solistes
Ces sources sont entendues dans les 2 arrangements.
Voix lead
Guitare lead (introduction)
Accordéon
Trombone Solo

Arrangement « Acoustique »
Contrebasse
Batterie (aux balais)
2 Guitares Acoustiques

Arrangement « Electrique »
Basse Synthé
Boîte à rythmes
Fender Rhodes

Figure 3-5 Répartition des sources par arrangement

Les deux arrangements reposent sur une rythmique ternaire. Mais pour l'arrangement « Electrique », la décomposition est deux fois plus rapide. Pour passer de l'arrangement « Acoustique » à l'arrangement « Electrique », on utilise une équivalence.

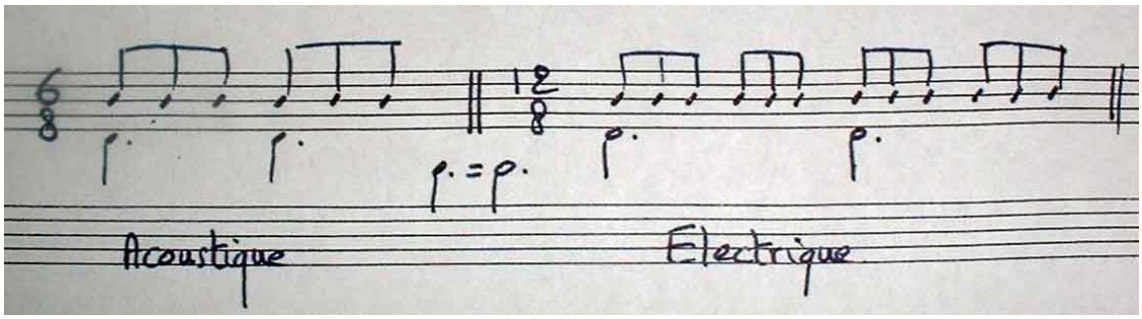



Figure 3-6 Extrait de partition rythmique

Par ailleurs, nous avons enregistré des sources pouvant venir renforcer l'écriture sur les refrains et les parties instrumentales (c'est à dire partout sauf sur les couplets). Leur présence n'est pas obligatoire. Chacun des arrangements possède ses sources optionnelles propres.

Sources Optionnelles
Basses d'Accordéon
Percussions (Bidons métalliques + Tom étouffé très grave)
Trombones en accords

Arrangement « Acoustique »
Guitare rythmique

Arrangement « Electrique »
Fender Rhodes son Saturé
Bruit de Surface Vinyle
H3000

Figure 3-7 Les Sources Optionnelles

3.3 Prise de Son, Montage et Mixage

La prise de son et le mixage ont été adaptés à une écoute en salle Schaeffer, dont l'acoustique est très colorée, comportant notamment des modes propres marqués ainsi qu'une réverbération longue dans les graves.

Réalisée conjointement à l'écriture progressive de l'arrangement, la prise de son s'est étalée sur plusieurs mois. Travail d'étudiant, elle comporte des défauts, mais nous a permis de faire de nombreux essais de prises, de montage et de traitements : remise en place d'une batterie ; échantillonnage; écriture d'une boucle sur séquenceur et sa mise en place sur un clic irrégulier ; prises dans différentes acoustiques : rôle des réflexions et de la réverbération en prise de proximité; égalisation et pré-mixage dès la prise pour gagner du temps et des pistes ; dé-corrélation des micros en prise « stéréo » de proximité (trouver un équilibre entre une richesse sonore et une image homogène) ; nécessité de construire le mix dans l'espace, les fréquences et le temps ; problèmes de transposition d'un Fender Rhodes enregistré trop bas (quand le time stretch fonctionne mal !) ; traitements d'une voix lead (compression, EQ, autotune...)...

L'espace multicanal est très large et on perçoit souvent des sources trop ponctuelles. Dès que nous en avons eu conscience, nous avons enregistré les instruments au moins en Stéréo : Piano "Rhodes" sur 4 canaux, Guitares Acoustiques sur 2 canaux chacune. Là où ça n'a pas été fait à la prise, on a préparé des pré-mixages comportant des delays courts, donnant l'illusion d'une source large. Mais ce type de traitement pose toujours des problèmes de timbre.

3.4 Elaboration des « Handles »

Le programme final devrait dans l'idéal permettre à l'auditeur d'agir sur ce qu'il écoute sans rien connaître à la technique. Nous avons choisi quelques contrôles dont la réalisation nous paraissait possible avec le matériel disponible. Ces contrôles ont été matérialisés dans MusicSpace sous la forme de "Handles". Leur déplacement provoque le déplacement d'un ou plusieurs objets, qui modifient la valeur de certains paramètres et donc changent la réponse de la console ou des effets pilotés. L'utilisateur a accès seulement à ces Handles, et ne voit pas les Objets pilotés.

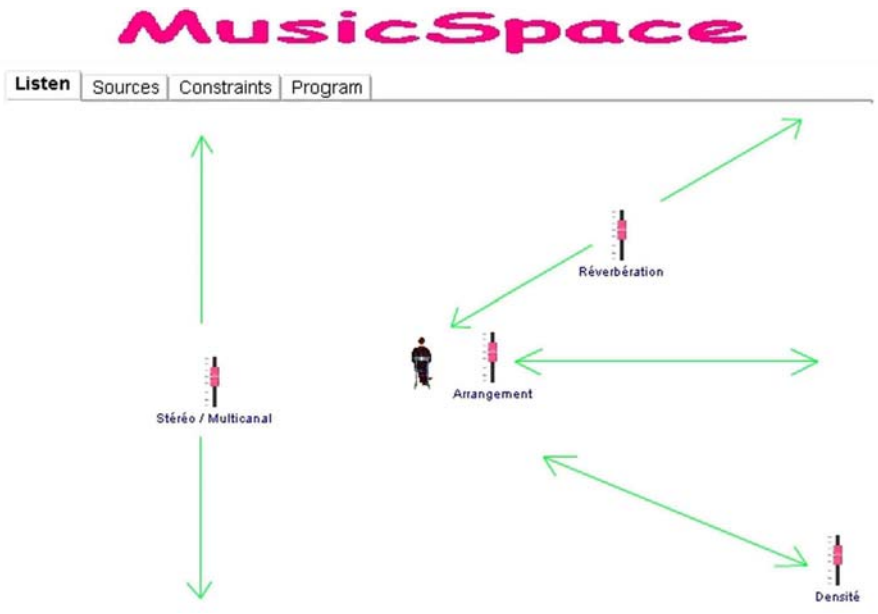


Figure 3-8 Interface Utilisateur et sens de déplacement des Handles

Des Handles de 4 types différents ont été mis en place:
- pilotage des paramètres d'un effet
- remplacement d'une source par une autre
- ajout de nouvelles sources
- modification de panoramiques

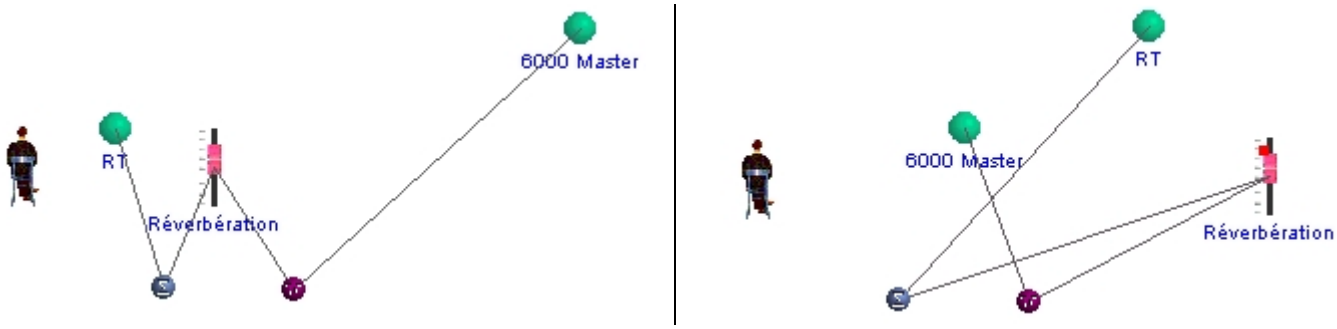
3.4.1 Handle « Réverbération »

Nous proposons ici un choix entre une acoustique sèche et une acoustique à réverbération longue bien que discrète.

En acoustique « sèche », Handle en position rapprochée, seule la Lexicon 300 est présente, élargissant et colorant légèrement la voix lead et les solistes, par un programme d'ambiance introduisant des réflexions précoces.

En acoustique réverbérée, la Lexicon fournit toujours les réflexions précoces, mais aussi une réverbération dont la longueur augmente avec l'éloignement du Handle. Cette réverbération, dont la durée (*RT* pour *Reverberation Time*) maximale est de 3,8 secondes, est appliquée à la voix lead. Elle est plutôt sourde, avec un shape moyen, jouant le rôle de pre-delay. A cela s'ajoute une réverbération longue de la M6000, en 4 canaux. La M6000 est appliquée à la voix lead, à l'harmonie et à la rythmique. Son niveau augmente lorsque le Handle s'éloigne de l'Avatar. Il semble que la réverbération multicanale ne brouille presque pas l'image, ce qui permet d'en mettre beaucoup sans être gêné.

Les paramètres des 2 modules d'effets sont contrôlés en MIDI :
- l'Objet « RT » pilote le temps de réverbération de la 300. Il est lié au Handle par une contrainte « rapport des distances constant », et se déplace dans le même sens que lui.
- l'Objet « 6000 Master » pilote le niveau général du départ Aux vers la M6000. Il est lié au Handle par une contrainte « produit des distances constant », et se déplace en opposition avec celui-ci.



Ambiance sèche — Salle réverbérée

Figure 3-9 Fonctionnement du Handle "Réverbération"

Pour résumer, l'éloignement du Handle par rapport à l'Avatar provoque:
- l'allongement du temps de réverbération de la 300
- l'augmentation du niveau réverbération fournie la M6000

Evaluation quantitative

Durée du mouvement : 8 secondes.
Nombre de paramètres modifiés simultanément : 2.
Types de paramètres modifiés : paramètre d'un effet et niveau de départ auxiliaire.
Nombre d'évènements MIDI envoyés par MusicSpace : 88.

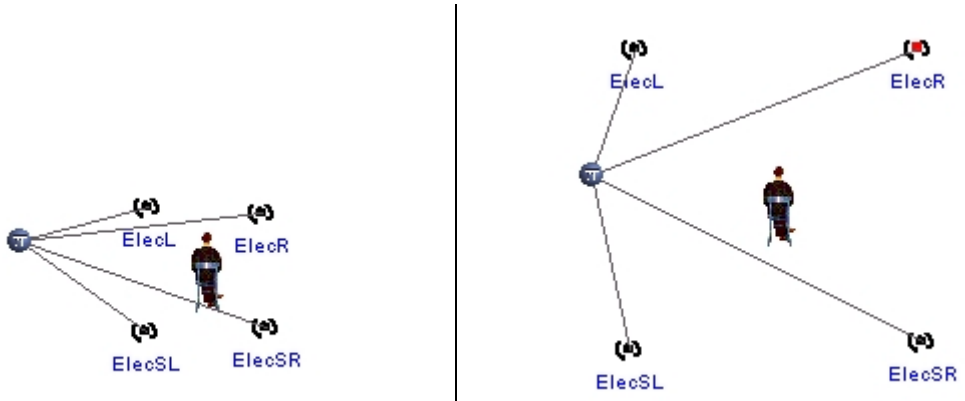
Démonstration
On peut entendre l'effet du Handle « Réverbération » sur le disque d'accompagnement.
1. Sec vers Réverbéré vers Sec sur Arrangement Acoustique
2. Sec vers Réverbéré vers Sec sur Arrangement Electrique

3.4.2 Handle « Arrangement »

Le Handle "Arrangement" remplace les sources de l'arrangement « Acoustique » par celles de l'arrangement « Electrique » et vice-versa. Les équilibres de niveau de chacun sont respectés.

On prépare d'abord, dans MusicSpace, une représentation mettant en jeu les objets correspondant aux sources. Celles-ci sont disposées autour de l'Avatar en fonction de la direction et du niveau approximatif auquel on veut les entendre. Le niveau de la source décroît lorsque celle-ci s'éloigne de l'Avatar. Plaçons les sources de l'arrangement « Electrique ». Une contrainte « rapport des distances constant » relie entre eux les objets appartenant à cet arrangement.

Pour les raisons que nous exposons au 4.1, nous avons regroupé certaines des sources de chaque arrangement sous un même objet. Ainsi, toutes les sources de l'arrangement « Electrique » sont pré-mixées sous les objets « Elec L », « Elec R », « Elec SL » et « Elec SR ».



Les objets sont proches: le niveau des sources est élevé

Lorsqu'on éloigne l'un d'eux, tous les objets s'éloignent dans la même proportion

Figure 3-10 Contrainte sur les objets de l'arrangement « Electrique »

On effectue la même opération avec les objets de l'arrangement « Acoustique ». Ils sont pré-mixées sous les objets « Acou L », Acou R », « GtrL », « GtrR », « GtrSL » et GtrSR ».

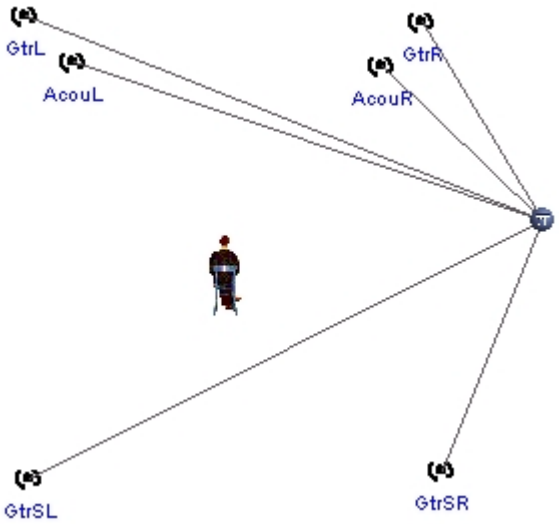
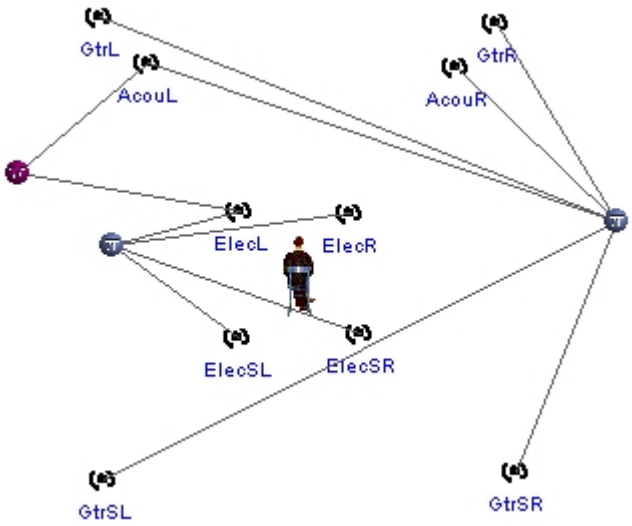


Figure 3-11 Contrainte sur les objets de l'arrangement « Acoustique »

On applique ensuite une contrainte « produit des distances constant » aux objets appartenant à des arrangements différents. Ils se déplacent ainsi en opposition.



Lorsque les objets d'un arrangement s'approchent, ceux de l'autre arrangement s'éloignent dans les même proportions.

Figure 3-12 Contrainte sur des objets d'arrangements différents

Visuellement, le résultat correspond à ce que l'on veut entendre. Les niveaux de chaque source sont ajustés pour aboutir à une balance dans chaque arrangement. Pour cela, on utilise les *trims d'entrée* de la console – on ne change pas la distance des Objets à l'Avatar, pour conserver la clarté de représentation.

On attache le Handle à l'une des sources. On utilise pour cela une contrainte « produit des distances constant ». La position du Handle la plus éloignée doit correspondre à la disparition de l'un des arrangements. Dans cette position, l'un des groupes d'objets renvoie à la 03D un paramètre plaçant son potard de gain à –infini. La position la plus rapprochée doit correspondre à celle ou l'autre groupe d'objets devient inaudible.

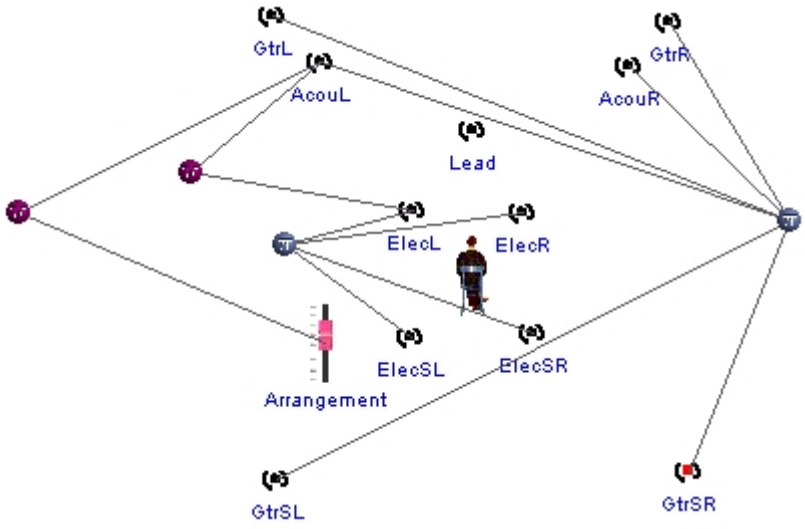


Figure 3-13 Le Handle "Arrangement" et les sources qu'il pilote

Le Handle « Arrangement » pilote aussi des sources optionnelles, qui lui sont liées de la même manière que les sources ci-dessus. Ces objets sont par ailleurs pilotées par le Handle « Densité », comme nous le verrons au 3.4.3. La gestion des mouvements d'un même objet par plusieurs Handles sera traitée au 3.5.

Evaluation quantitative

Durée du mouvement : 8 secondes.
Nombre de paramètres modifiés simultanément : 15 (incluant les sources optionnelles).
Type de paramètre modifié : gain de tranche.
Nombre d'évènements MIDI envoyés par MusicSpace : 545.

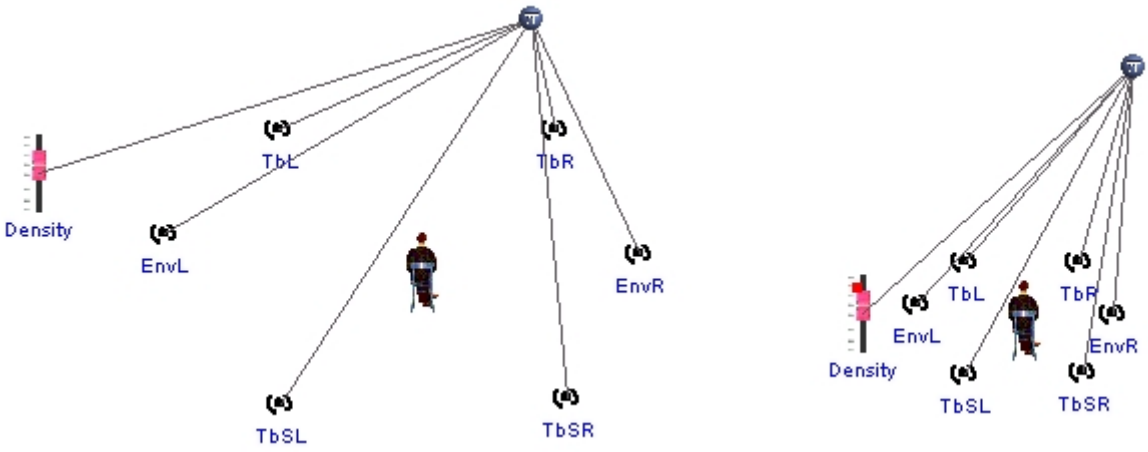
Démonstration
On peut entendre l'effet du Handle « Arrangement » sur le disque d'accompagnement.

1. Electrique vers Acoustique
2. Acoustique vers Electrique
3. Electrique vers Acoustique avec Densité élevée
4. Acoustique vers Electrique avec Densité élevée

3.4.3 Handle « Densité »

Ce Handle introduit des sources dont la fonction est de densifier l'arrangement. Nous décrivons d'abord l'introduction des « sources optionnelles », listées en Figure 3-7, qui sont présentes quelque soit l'arrangement. Il s'agit des trombones (nommées « TbL, R, SL et SR »), ainsi que des basses d'accordéon et des percussions (regroupées sous les objets « EnvL » et « EnvR »).

Comme lors de la fabrication du Handle "arrangement", on prépare d'abord une représentation graphique signifiante. Les sources sont disposées autour de l'Avatar en fonction de la direction et du niveau approximatif auquel on veut les entendre. Les objets sont reliés par une contrainte « rapport des distances constant ». Passée une certaine distance à l'Avatar, le potard de la 03D est à -infini et la source n'est plus entendue. Le Handle doit alors être dans sa position la plus éloignée.



On n'entend pas les sources trop éloignées

Le rapprochement du Handle provoque celui des sources

Figure 3-14 Principe du Handle « Densité »

Utilisant les *trims* de la 03D, on fait une balance entre les sources optionnelles, dans leur position finale (près de l'Avatar).

Le Handle « Densité » rapproche également des sources propres aux arrangements « Acoustique » et « Electrique » (voir Figure 3-20). Elles sont soumises également à l'action du Handle « Arrangement ». Leur intégration est décrite au 3.5.2.

Evaluation quantitative

Durée du mouvement : 10 secondes.
Nombre de paramètres modifiés simultanément : 12.
Types de paramètres modifiés : gain de tranche et niveau de départ auxiliaire.
Nombre d'évènements MIDI envoyés par MusicSpace : 545.

Démonstration
On peut entendre l'effet du Handle « Densité » sur le disque d'accompagnement.

5. Augmentation de Densité sur Arrangement Electrique
6. Augmentation de Densité sur Arrangement Acoustique
7. Diminution de Densité sur Arrangement Electrique
8. Diminution de Densité sur Arrangement Acoustique

3.4.4 Handle « Stéréo / Quad »

Ce curseur effectue un simple repliement des sources arrières sur les enceintes avant, faisant sentir clairement la différence entre les deux types de diffusion stéréo ou quadriphonique.

Les panoramiques de chaque tranche de la console sont pilotés en MIDI par l'intermédiaire d'Objets « position ». Le Handle contraint tous les Objets « position » situés à l'arrière de l'Avatar à se déplacer vers l'avant. Il utilise pour cela une contrainte « translation ».

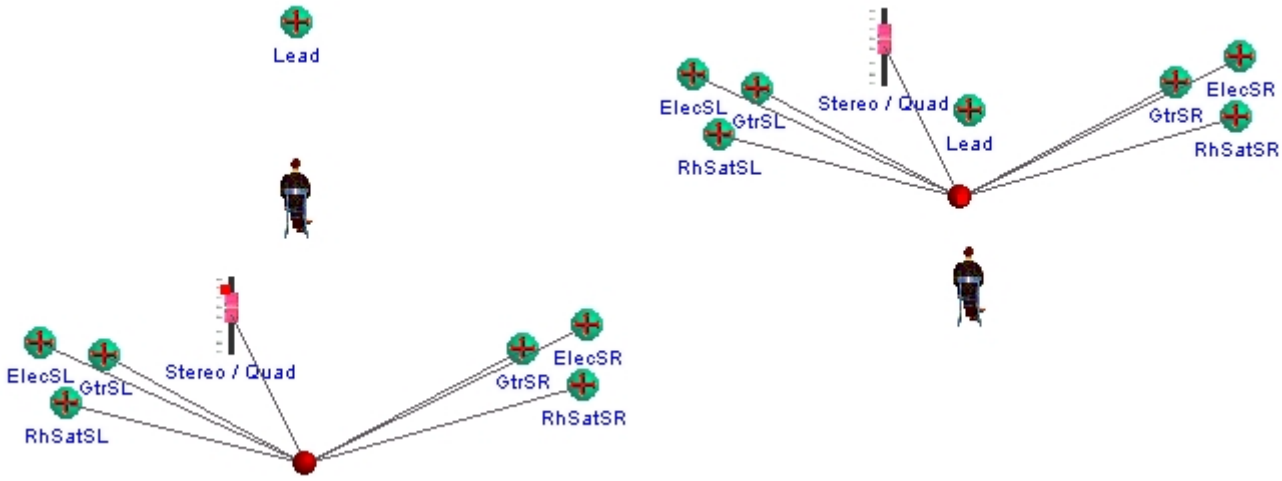


Figure 3-15 Le Handle "Stéréo - Quad"

Ce cas particulier est un peu simpliste, mais nous avons voulu le réaliser car les effets à base de mouvements de sources ont été à la base de l'idée de MusicSpace.

Evaluation quantitative

Durée du mouvement : 18 secondes.
Nombre de paramètres modifiés simultanément : 12.
Type de paramètre modifié : Panoramiques.
Nombre d'évènements MIDI envoyés par MusicSpace : 2914.

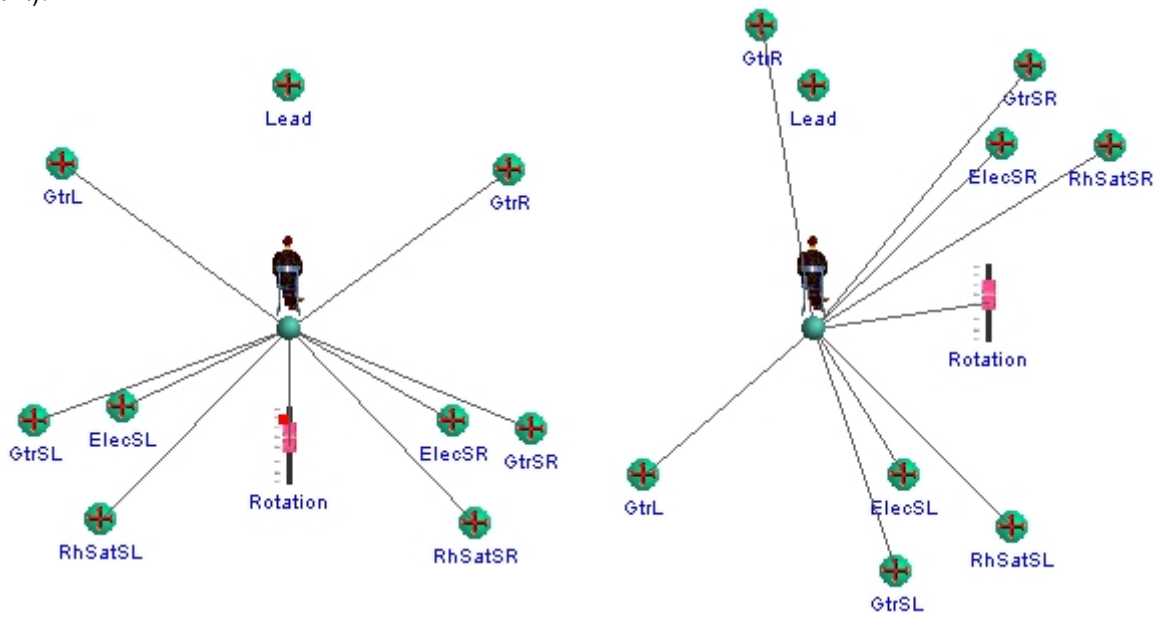
3.4.5 Autres tentatives de Contrôles

Bien qu'ils ne soient pas utilisés dans le « mixage interactif » présenté, nous avons expérimenté autour de ces quelques contrôles.

3.4.5.1 Handle « Rotation autour de l'auditeur »

L'objectif est de faire tourner l'ensemble de l'orchestre autour de l'auditeur. On peut exclure les solistes de ce mouvement, afin de conserver un repère. On peut imaginer disposer une version du mixage à l'avant, l'autre à l'arrière, et effectuer la transition par rotation. On pourrait également se servir de l'effet pour attirer l'attention de l'auditeur sur une source, placée en zone frontale.

Après avoir disposé les Objets « position » de chaque source autour de l'Avatar, nous introduisons une contrainte « angulaire », qui impose la conservation des rapports angulaires (Objet 1; Avatar; Objet 2).



Le Handle et les Objets « position » sont disposés autour de l'Avatar. On leur applique une contrainte angulaire.

On fait tourner le Handle dans le sens trigonométrique : les rapports angulaires entre objets sont conservés.

Figure 3-16 Objets soumis à une contrainte angulaire

3.4.5.2 Handle « Présence »

En tant qu'ingénieur du son, nous avons d'abord cherché à ce que le mobile interactif porte sur le son. L'auditeur choisirait entre divers placements de micros, apportant plus ou moins de présence ou plus ou moins d'espace ; il agirait sur des quantités de réverbération, sur la localisation, sur les paramètres des effets sonores ou, pour adapter la diffusion à sa situation d'écoute, sur le taux de compression général.

Ces premiers essais ont été réalisés sur des formations orchestrales classiques, en effectif moyen. Je disposai d'enregistrements multi-microphoniques, comportant un couple pour une image d'ensemble lointaine, des microphones placés dans le champ réverbéré, et des appoints de proximité sur chaque instrument. Nous avons travaillé notamment sur une oeuvre de Bernat Vivancos enregistrée en public lors des « Journées de la Composition 1999 », au Conservatoire.

Un Handle basé sur un relais de microphone éloigné à microphone rapproché a été mis au point. Nous l'avons baptisé « Présence ».

Lorsque l'auditeur se trouve entouré de sources sonores, il lui est possible d'élargir le cercle des sources qui l'entourent. Chaque source reste dans le même axe, mais sa largeur et son timbre

sont modifiés, ainsi que le rapport Champ Direct/Champ Réverbéré: l'instrument moins présent est perçu comme étant plus distant.

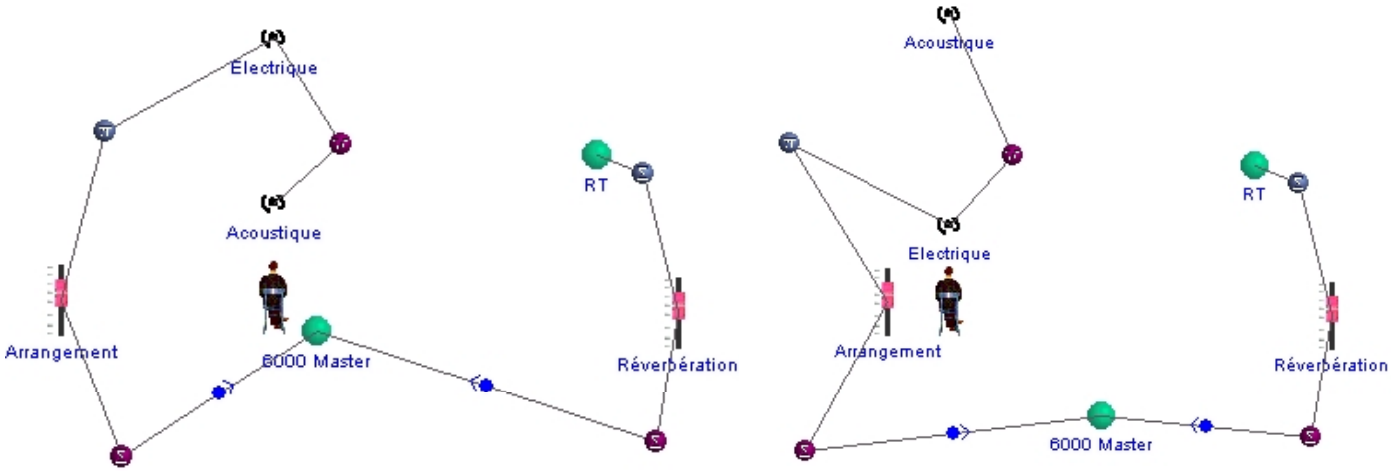
Sur une prise de son au couple stéréophonique (ou multicanale), on superpose une prise de proximité en multi-microphonie. Le Handle modifie le rapport entre les 2. L'absence de réverbération dans la prise de proximité est compensée par l'ajout de micros dans le champ réverbéré.

3.5 Objets contrôlés par plusieurs Handles

3.5.1 Action croisée des Handles « Arrangement » et « Reverb »

Comme lors d'un mixage traditionnel, il est souvent nécessaire de compenser la modification d'un paramètre par des ajustements sur les autres.

Par exemple, on constate qu'un ajustement de la réverbération est nécessaire lors du passage d'un arrangement à l'autre. Dans la version acoustique, les guitares, très denses, masquent les réverbérations. Celles-ci sont beaucoup plus apparentes dans la version électronique, qui est moins chargée. Il est nécessaire que le Handle « arrangement » fasse un ajustement sur le niveau de réverbération: nous introduisons donc une contrainte entre ce Handle et l'Objet contrôlant le niveau de M6000.

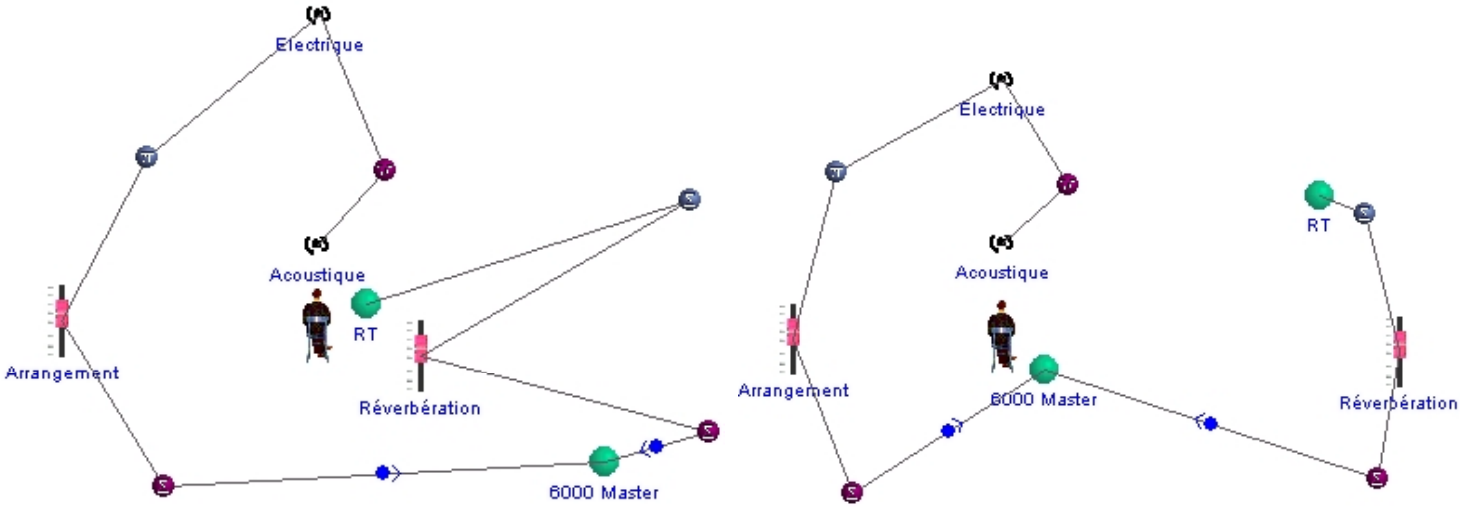


Placé en position éloignée, le Handle « Réverbération » donne un RT long et un niveau de M6000 élevé

Le passage à l'arrangement « électrique » provoque l'éloignement de l'Objet « 6000 Master ». Nous n'avons pas modifié la position du Handle « Réverbération »

Figure 3-17 Contrainte entre le Handle « arrangement » et les paramètres de réverbération

Cette contrainte possède un sens de propagation. En effet, l'Objet « 6000 Master » est déjà contraint par ailleurs : il obéit aux déplacements du Handle « réverbération ». On souhaite que le passage de l'arrangement "acoustique" à l'arrangement "électrique" provoque une diminution du niveau de réverbération. Mais on ne souhaite pas qu'une modification des paramètres de la réverbération entraîne un changement dans l'arrangement. Par ailleurs, pour plus de clarté de représentation, on souhaite que la quantité de réverbération soit modifiée au changement d'arrangement, sans que le Handle « réverbération » soit pour autant déplacé. Le sens de propagation d'une contrainte est matérialisé dans MusicSpace par des petites flèches bleues.



L'éloignement du Handle « Réverbération » provoque le rapprochement de l'Objet « 6000 Master », mais ne se propage pas vers le Handle « Arrangement »

Figure 3-18 Propagation des contraintes entre le Handle « Réverbération », l'Objet « 6000 Master » et le Handle « Arrangement »

3.5.2 Action croisée des Handles « Densification » et « Arrangement »

Certaines sources optionnelles sont également sujettes aux actions croisées de deux Handles distincts. Comme nous le montrons Figure 3-24, lorsque l'on souhaite densifier l'arrangement « Acoustique », ce sont les sources optionnelles acoustiques qui doivent s'approcher. Mais si l'on souhaite densifier l'arrangement « Electrique », la modification doit porter sur les sources optionnelles électriques. Par exemple, la guitare rythmique de l'arrangement « Acoustique » est remplacée par un Fender Rhodes. De même, dans l'arrangement « Electrique » uniquement, le Handle « Densité » pilote aussi le niveau d'envoi auxiliaire de la voix lead vers un *phaser* du H3000 qui, diffusé à l'arrière, donne plus de présence à la voix et donne l'illusion qu'elle se rapproche de l'auditeur.

Comme lors de la fabrication du Handle « Arrangement », on regroupe les sources optionnelles correspondant à chaque arrangement. Pour l'arrangement « Acoustique », ce sont « GtcSL » et « GtcSR ». Pour l'arrangement « Electrique », ce sont « RhSatSL », « RhSatSR », « Vinyle » et « Phaser Lead ». Les sources de chaque groupe sont liées par des contraintes « rapport des distances constant ».

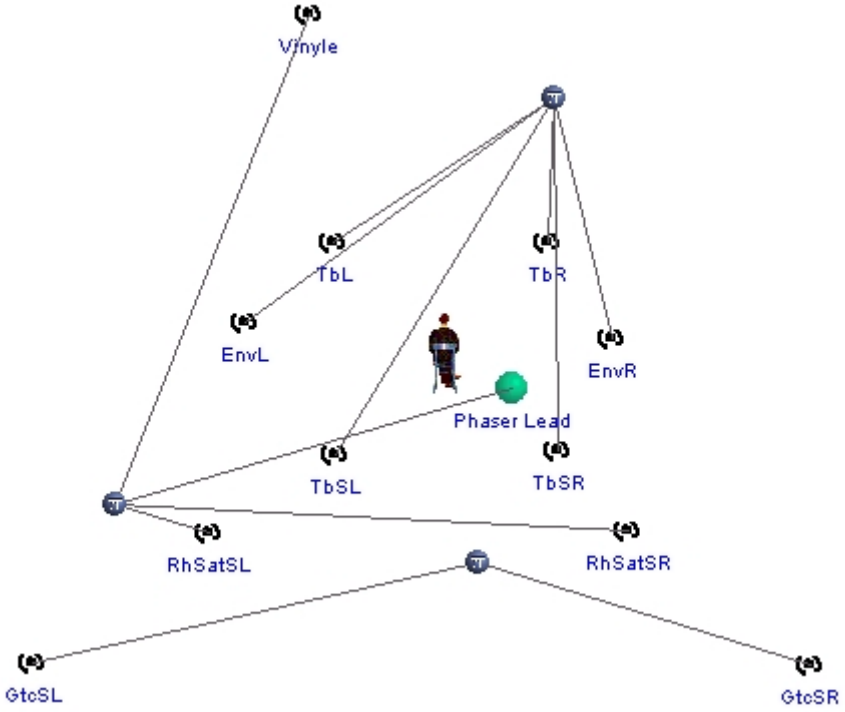


Figure 3-19 Sources affectées par le Handle "Densité", groupées par arrangement

On inclue le Handle dans la contrainte « rapport des distances constant » de chaque groupe d'objets.

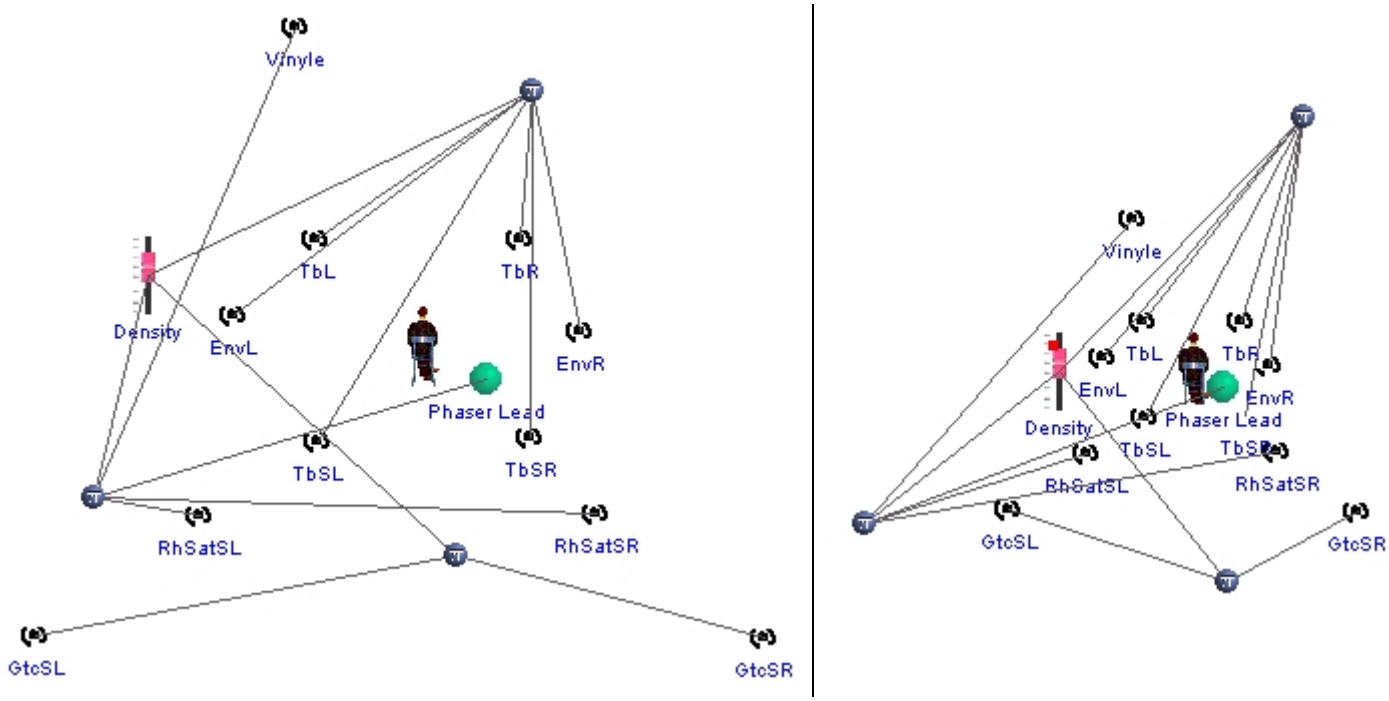


Figure 3-20 Action du Handle "Densité"

Figure 3-21, on fait en sorte que le déplacement des objets d'un groupe ne se propage pas aux autres groupes : à nouveau, l'objectif est atteint grâce à l'action magique des petites flèches bleues. Ainsi, les mouvements du Handle se propagent aux « Gtc », mais celui des « Gtc », qui aura lieu au changement d'arrangement, ne se propage pas au Handle.

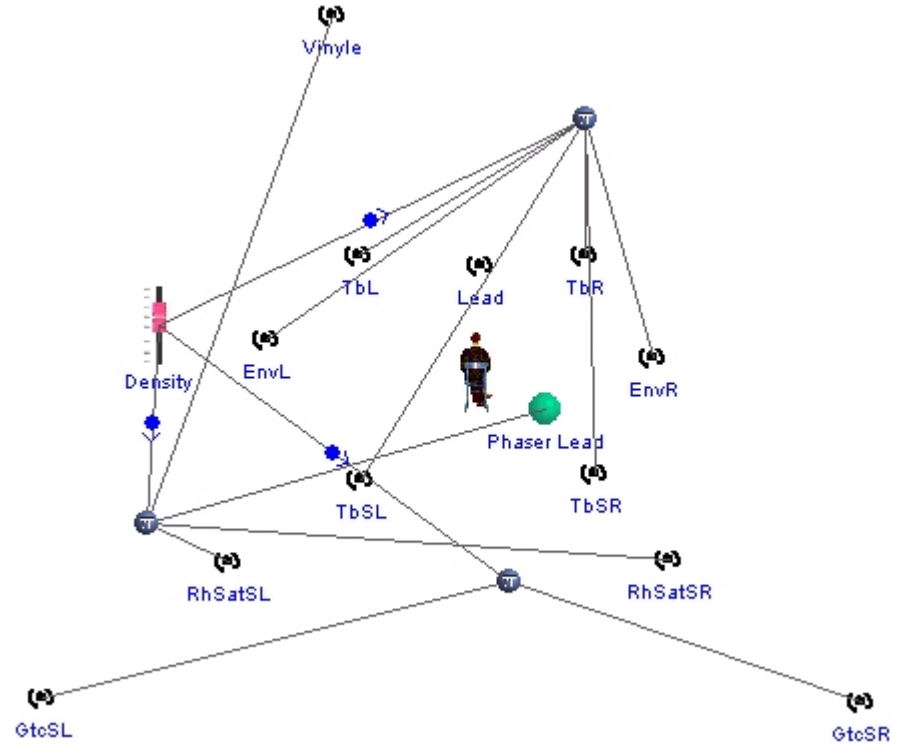
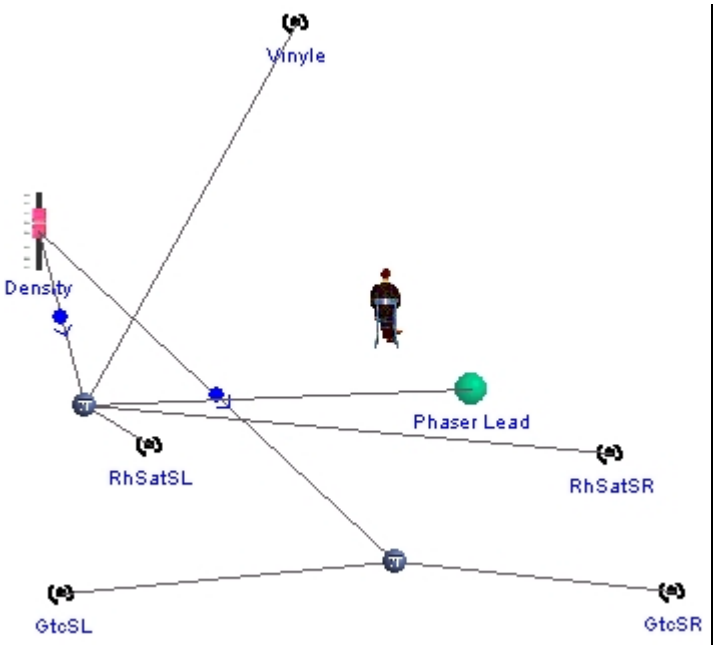
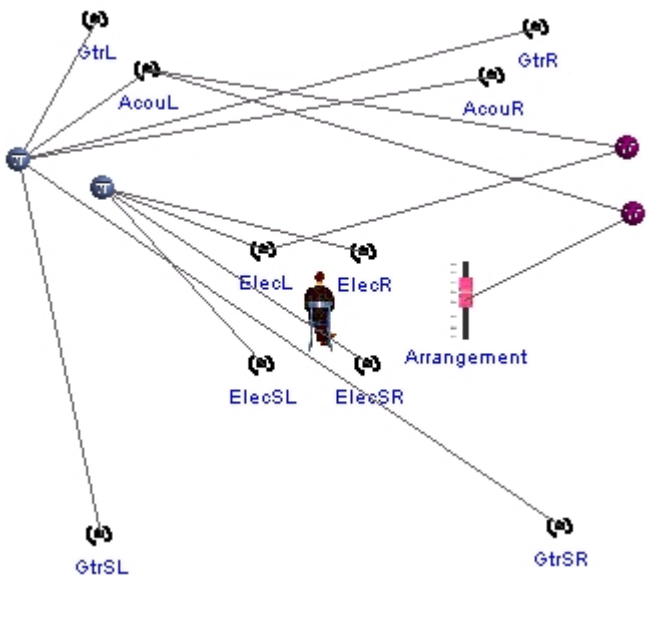


Figure 3-21 Définition du sens de propagation du mouvement du Handle « Densité »

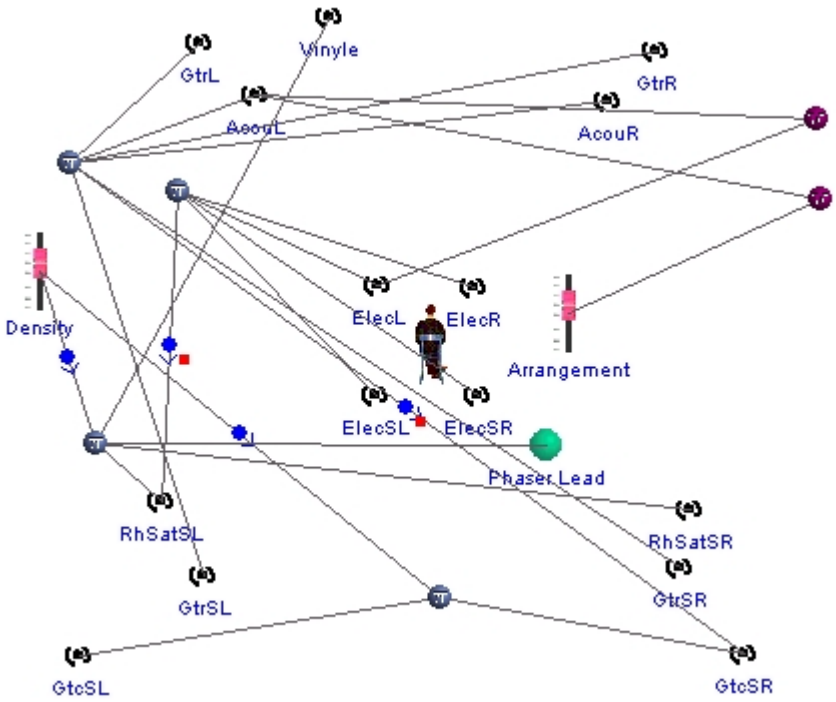


L'action du Handle « Densité »
(pour plus de clarté, les objets présents dans les 2 arrangements ont été retirés)



L'action du Handle « Arrangement »
(voir 3.4.2)

Figure 3-22 Les Handles à combiner



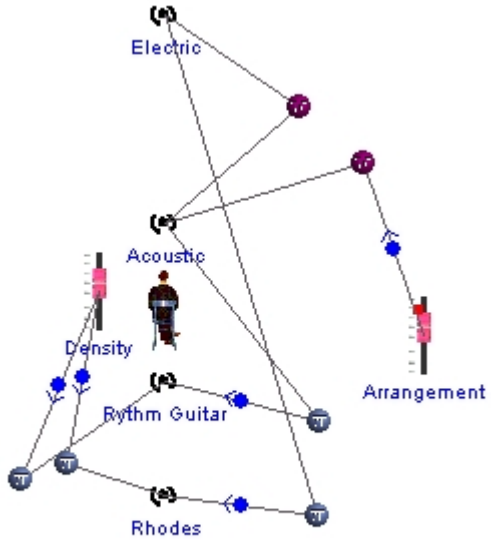
Les flèches bleues marquées d'un point rouge empêchant le mouvement des objets principaux de l'arrangement lorsque les sources optionnelles sont déplacées.

Figure 3-23 Combinaison des Handles "Arrangement" et "Densité"

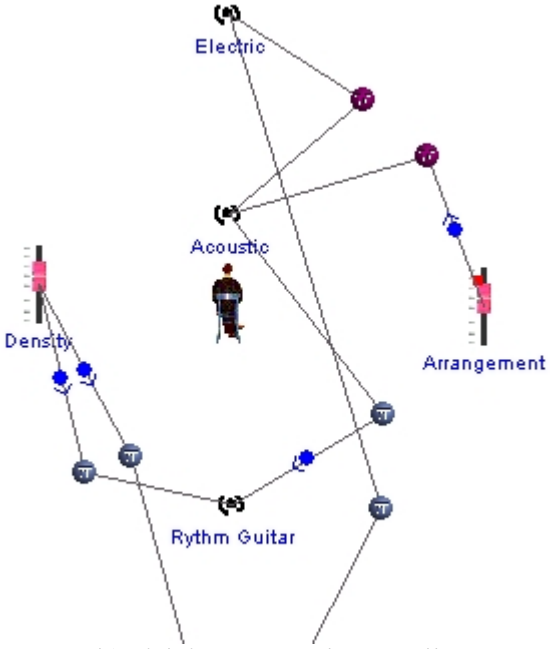
Nous souhaitons superposer le mécanisme de la Figure 3-21 avec celui de exposé en Figure 3-13. Sur la Figure 3-23, on observe que les contraintes groupant les objets d'un même arrangement sont aussi appliquées aux objets « optionnels ». Deux nouvelles flèches de propagation apparaissent, empêchant le mouvement des objets principaux de l'arrangement lorsque les sources optionnelles sont déplacées.

La Figure 3-24 présente le fonctionnement conjugué de ces deux Handles. Pour plus de clarté, nous n'avons gardé qu'un objet représentatif de chaque groupe : les objets principaux de chaque arrangement sont appelés « Acoustic » et « Electric », et on a conservé seulement « Rythm Guitar » et « Rhodes » comme source densificatrice de chacun de ces arrangement, respectivement.

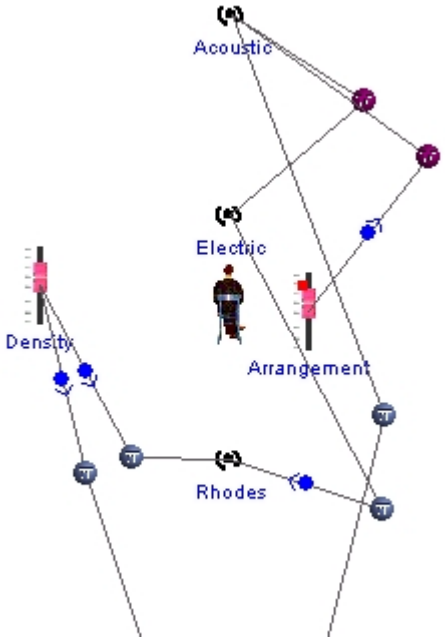
On choisit un arrangement en plaçant le Handle « arrangement » dans l'une des positions extrêmes. En position « non densifiée », on veille à ce que le rapprochement maximal du Handle « arrangement » ne rende pas audible une des sources de densification. Par ailleurs, on veille à ce que le rapprochement maximal du Handle « densification » conserve les sources optionnelles de l'autre arrangement hors de la zone audible.



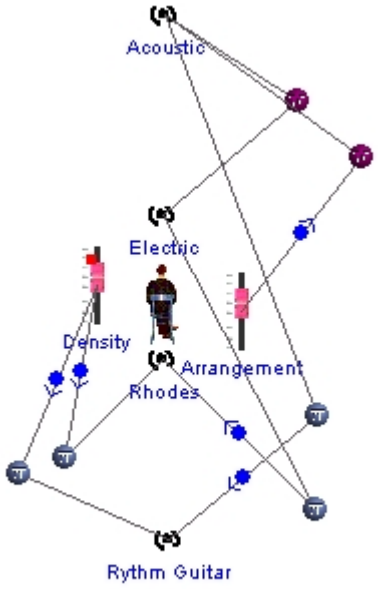
a) L'arrangement « Acoustique » est densifié par la Guitare Rythmique



b) L'éloignement du Handle « Densification » ne se propage pas au Handle « Arrangement »



c) Les mouvement du Handle « Arrangement » sont sans effet sur le Handle « Densification ». Ils rapprochent l'Objet « Rhodes », mais pas suffisamment pour le rendre audible



d) L'arrangement « Electrique » est densifié par le Rhodes. La Guitare Rythmique est placée suffisamment loin pour être inaudible

Figure 3-24 Contraintes croisées entre les Handles « Arrangement » et « Densité »

Conclusion de la 3ième partie

Nous avons fait écouter le résultat à un panel d'étudiants de la FSMS et d'ingénieurs du son. Ce type de public ressent la spécificité de l'outil, permettant avec un seul curseur et instantanément de faire bouger un nombre de boutons trop important pour que l'opération ne soit effectuée par un seul opérateur.

Nous avons présenté la chanson interactive produite à des musiciens et des compositeurs, pour recueillir leurs impressions. Les musiciens sont-ils intéressés ? Auraient-ils peur, laissant un choix, de perdre leur signature ? L'écoute de ce « mixage interactif » permet d'engager la conversation. Souvent, la proposition semble faire jaillir des idées de compositions originales.

Nous voilà maintenant à l'issue d'une expérience qui demeure à notre connaissance unique en son genre. L'enregistrement réalisé n'est pas disponible sous une forme facilement diffusable : le système de reproduction nécessite une console 03D, des effets de studio, et le mixage a été adapté aux spécificités acoustiques de la salle Schaeffer... Pourtant, les moyens techniques de fabriquer une version diffusable existent. Il serait ainsi possible de permettre à l'auditeur d'écouter le titre chez lui, avec tout le temps qu'il désire, afin de lui permettre de dépasser le stade du jeu, suscité par une technologie nouvelle.

Chapitre 4

Analyse du mixage obtenu

Nous avons réalisé 4 Handles parmi les propositions faites au Chapitre 1. Ici apparaissent les difficultés que nous avons rencontrées et parfois résolues durant ce travail. Le mixage obtenu répond-il aux exigences que nous avons formulées au Chapitre 2 ?

Un mixage de *pop-music* demande souvent une longue mise au point aboutissant à un résultat fragile. MusicSpace, encore en cours de développement, résout les problèmes dans leur principe. Nous examinons à quel niveau son manque de précision fait le plus défaut.

La localisation est au cœur de nombreux débats sur le multicanal. Nous nous intéressons aux modifications de l'image sonore par changement de panoramique. Les nombreuses contraintes proposées dans MusicSpace pour changer la perception spatiale sont-elles utilisables ?

Nous étudions ensuite deux notions utiles lorsqu'on veut relier des paramètres de nature différente et en effectuer le pilotage : la fonction de transfert et les bornes.

Enfin, nous décrivons les problèmes que nous avons eu lorsque nous nous posions la question de savoir sur quoi pouvait porter l'interactivité : comment faire en sorte que l'auditeur novice entende les modifications, sans pour autant les exagérer au point d'atteindre le style.

4.1 La précision des réglages

L'un des critères de cohérence d'une image sonore en *pop-music* est le maintien d'une balance, souvent avec une précision de moins d'un déciBel. Beaucoup des contrôles de « MusicSpace » sont assez imprécis, ce qui rend le mixage difficile à mettre au point : l'espace sonore se présente graphiquement comme un écran vide. L'ingénieur du son n'a pas de repère et ne peut connaître exactement la position ou le niveau de la source diffusée : il est difficile, par exemple, de placer un son exactement sur une enceinte.

Au cours de notre pratique, nous avons résolu ce problème en procédant à des pré-mixages chaque fois que des équilibres précis étaient nécessaires. Lors de ces pré-mixages, les sources ont été groupées par recoupement entre zone de diffusion et type de paramètre piloté. Par exemple, on trouve un pré-mixage « Batterie + Contrebasse », assurant un bon équilibre entre ces éléments rythmiques. On trouve aussi un pré-mixage des sons diffusés à l'arrière dans l'arrangement électronique.

Une liste des pistes réelles, sur chaque arrangement, avec leur affectation spatiale, se trouve en Annexe IV.

Ces pré-mixages nous ont du même coup permis d'alléger la représentation graphique dans MusicSpace. Ils permettent de faire figurer un moins grand nombre d'objets et de contraintes sur l'écran : les Objets « Acoustique L » et « Acoustique R » apparaissent à la place d'une liste de tous les instruments acoustiques qui seront traités de la même façon.

Enfin, les pré-mixages ont diminué le nombre de pistes nécessaires – la 03D possédant 26 entrées convenables. Au delà de notre expérimentation, ce procédé est à considérer si l'on souhaite rationaliser les moyens techniques nécessaires à la diffusion du produit.

4.2 Les Déplacements dans l'espace sonore

L'idée originale des concepteurs de MusicSpace était de fournir un programme capable de situer n'importe quelle source mobile par rapport à l'Avatar, et de piloter en conséquence un système de diffusion sur 360°. Il est possible grâce au système de contraintes de faire des groupes de sources, et de les soumettre à des déplacements graphiques complexes. Tout au long de notre étude, nous avons tenté de mettre en valeur cette idée, en nous interrogeant sur :

- le type de déplacements admissibles
- la qualité de leur perception

Les Contraintes de MusicSpace permettent d'automatiser des rotations ou des translations entre n'importe quels points de l'écran, ou même de commander des mouvements aléatoires. L'auditeur peut lui aussi être mobile par rapport à l'espace graphique et donc par rapport aux sources, rendant théoriquement possible l'effet « déplacement de l'auditeur au sein de l'orchestre ».

La perception de l'espace en quadriphonie est pourtant assez fragile. Beaucoup de publications actuelles ont pour objectif d'en améliorer la précision. Même si le fait que les sources soient mobiles améliore quelque peu la localisation, nous avons constaté qu'un déplacement aléatoire se soldait souvent par une perception peu subtile du type « avant – arrière – gauche – droite ». Afin d'assurer la perception de mouvements plus intéressants, nous avons essayé d'organiser les mouvements proposés : nous avons choisi de fixer la position de l'avatar à sa position centrale originale. La contrainte « rotation autour d'un point » a été utilisée en « rotation autour de l'*auditeur* ». La contrainte « translation » a été utilisée pour des déplacements sur une *ligne avant-arrière*. Plus faciles à concevoir, ces déplacements sont perçus à l'identique par un plus grand nombre d'auditeurs. Nous nous sommes également limités au déplacement d'un petit nombre de sources.

Notre réalisation la plus simple est celle du Handle « Stéréo - Quad », qui replie vers l'avant les panoramiques des sources situées à l'arrière. Il fonctionne relativement bien. Nous aurions souhaité aller plus loin : partir d'une image monophonique, passer par la stéréo puis l'extra largeur, puis par une image multicanale « direct-ambience », pour aboutir à une image

multicanale « direct-sound-all-around ». Nous avons tenté de réaliser un effet de rotation autour de l'auditeur (3.4.5.1).

Nous nous sommes heurtés aux problèmes perceptifs suivants :
- un changement de position change beaucoup la perception timbrale des sources, et donc la définition et la balance, car l'influence de l'effet de masque est changée
- l'homogénéité est mal conservée : un mauvais positionnement peut nuire à la fusion des timbres. De plus, comme en stéréo, lorsqu'une source s'approche d'un haut-parleur, elle paraît plus forte.
- l'oreille sépare plus difficilement les sons lorsque les sources sont superposées : ceci nuit à l'intelligibilité.

Ces mouvements mettent ainsi en danger l'équilibre spectral, l'homogénéité, la balance et l'intelligibilité (voir 2.2). La réalisation de l'effet de rotation fût un échec, le mixage ne conservant pas une balance correcte.

En revanche, l'effet de rotation d'un espace entier autour de l'auditeur est saisissant. De plus, l'angle d'écoute multicanal étant supérieur à l'angle en stéréo, on a plus de place pour positionner les sources, qui se masquent moins entre elles : dans certains cas, l'intelligibilité est moins difficile à atteindre car il y a moins de conflits. C'est pour cela qu'on dit parfois que mixer en 5 canaux est plus facile qu'en stéréo. En contrepartie, il est plus difficile de remplir l'espace entier : certaines sources paraissent trop ponctuelles. La prise de sources en stéréo est souvent nécessaire.

Pour résoudre ces problèmes, il faudrait utiliser le système de contraintes pour :
- contrôler les déplacements
- changer l'égalisation des sources au cours de leur déplacement, pour compenser le détimbrage
- réajuster la balance pendant la rotation.

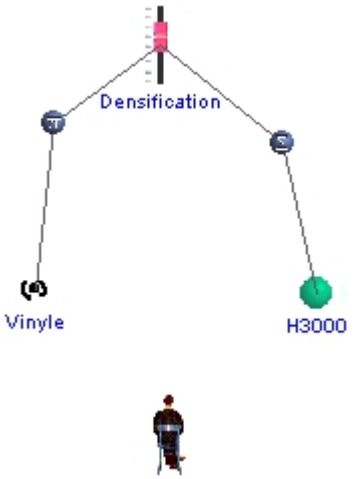
Nous nous sommes ici heurtés aux limites techniques de MusicSpace. Le système de contraintes s'est avéré ne pas être assez performant, et pas assez précis :
- Il s'agit de modifications au dB près. Au cours de la production, les équilibres demandant autant de précision ont été effectués par pré-mixages, hors de MusicSpace.
- Les déplacements sous contrainte « angulaire » se sont révélés ne pas être parfaitement reproductibles, lorsqu'on leur combinait d'autres contraintes. Un déplacement de l'avatar a beaucoup d'influence sur l'ensemble des paramètres, et rend également la reproductibilité aléatoire.
- Pour corriger des niveaux et des timbres en même temps que l'on modifie les panoramiques, il faudrait disposer d'une contrainte permettant de lier un déplacement angulaire à un éloignement ou un rapprochement. Cette contrainte n'existe pas. De plus, l'interface ne serait pas adaptée : on aurait besoin d'un très grand nombre d'objets, et leurs déplacements n'auraient pas de sens par rapport à la représentation graphique.

4.3 Fonctions de transfert du déplacement

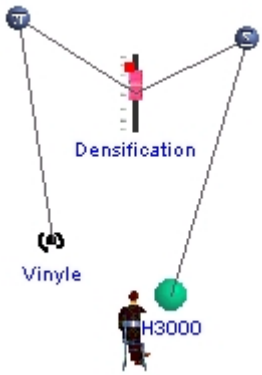
Nous avons constaté la nécessité de régler la rapidité avec laquelle les sources contraintes se déplacent les unes par rapport aux autres (voir 12). Nous appelons la fonction définissant la distance parcourue par une source en fonction de la distance parcourue par une autre source une « fonction de transfert ».

Ainsi lorsque nous avons construit le Handle « Arrangement », nous avons constaté qu'un silence apparaissait à la transition entre les deux arrangements : les sources électriques s'éloignaient alors que les sources acoustiques n'étaient pas encore suffisamment proches pour être entendues. Or, nous voulions faire une transition par tuilage.

Sur « MusicSpace », trois types de contraintes permettent d'agir sur la rapidité des déplacements relatifs entre les sources : les contraintes en produit de distances constantes, en ratio de distances constantes, et les contraintes en somme et différence de distances constantes. Nous n'avons pu définir précisément leur action : dans notre pratique, le choix s'est fait intuitivement, par essai et erreurs.



2 objets sont soumis à 2 contraintes différentes et alignés à la même distance de l'Avatar



L'objet soumis à la contrainte « différence des distances constante » arrive le premier

Figure 4-1 Variation de la vitesse de déplacement en fonction de la contrainte utilisée

Pour régler la transition entre les deux arrangements, nous avons ainsi choisi des associations de contraintes permettant un tuilage sans rupture. Dans le mixage final, lorsque le Handle est placé dans une position intermédiaire, les deux arrangements peuvent être entendus en même temps, bien que la présence simultanée de toutes ces sources donne une musique un peu chargée. Ils ne sont pas à proprement parler complémentaires, mais chacun est intéressant en lui même, et leur cohabitation fonctionne. Le passage de l'un à l'autre provoque un effet saisissant.

Le même type de problème s'est posé avec le Handle « réverbération ». Nous souhaitions modifier simultanément les paramètres de 2 modules d'effet – le RT de la Lexicon 300 et le niveau général de la TC M6000. Pour que les 2 réverbérations se modifient dans le même rapport, nous avons utilisé une fonction de transfert intégrée au module de commande MIDI de la 300: ainsi, nous avons défini dans quelle mesure le RT est affecté par les déplacements de l'Objet « RT ».

Enfin, on a parfois souhaité qu'une source intervienne un peu plus tôt que les autres lors du rapprochement du Handle : on l'a placée un peu plus près de l'Avatar. Ici, l'effet du H3000 intervient avant que le Rhodes et le bruit de Vinyle ne se fassent entendre (Figure 4-2).

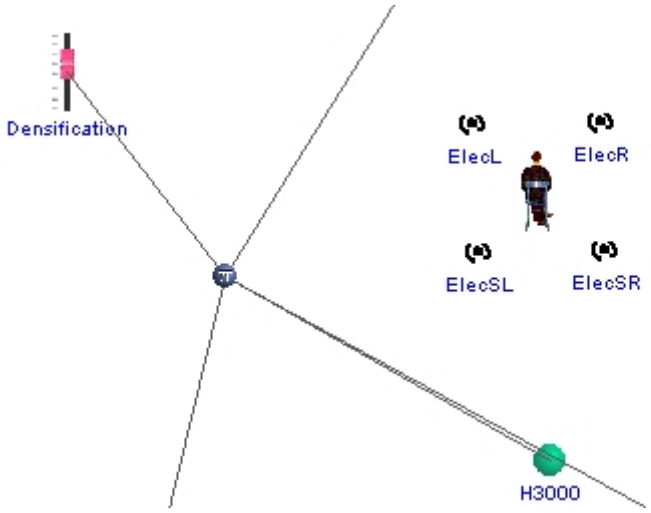


Figure 4-2 Densification progressive de l'arrangement "Electrique"

En résumé, les « fonctions de transfert » permettent de :
- Faire évoluer conjointement des paramètres de types différents.
- Doser la rapidité d'apparition de certaines sources en fonction de la disparition des autres (Handle "arrangement"), ou de l'apparition des autres (Handle "densification").

Il serait possible d'introduire dans MusicSpace une contrainte possédant un paramètre, permettant de doser la rapidité de déplacement d'une source contrainte par rapport aux autres.

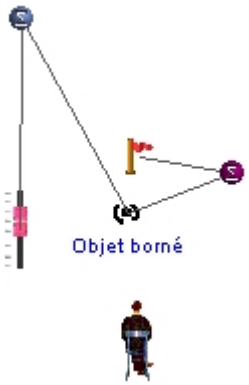
4.4 Borner les déplacements des Handles

Dans MusicSpace, les contraintes jouent le rôle des règles, tandis que l'amplitude des déplacements représente en quelque sorte leur souplesse. Or on a besoin de pouvoir dire : « ceci est la maximum/le minimum de réverbération souhaité » ; ou « pour respecter la balance, le niveau de cette source doit être compris entre ces limites », par exemple. Chaque déplacement doit se faire entre deux bornes : il est nécessaire d'empêcher l'auditeur de dépasser certaines limites. Pour le moment, la borne est matérialisée d'un côté par les limites de l'écran, de l'autre par la position centrale de l'Avatar.

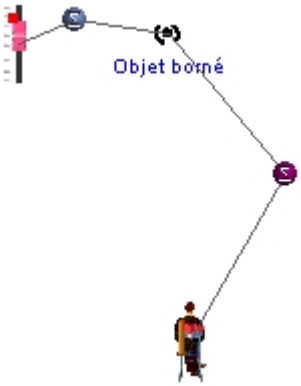
Or l'écran n'est pas sphérique mais rectangulaire, ce qui donne une borne différente selon que l'on se dirige vers un coin, vers les bords haut et bas, ou vers les bords gauche et droite. C'est déroutant pour l'auditeur.

Il serait intéressant de pouvoir définir les bornes entre lesquelles les Objets peuvent se déplacer, et donc les paramètres être modifiés.

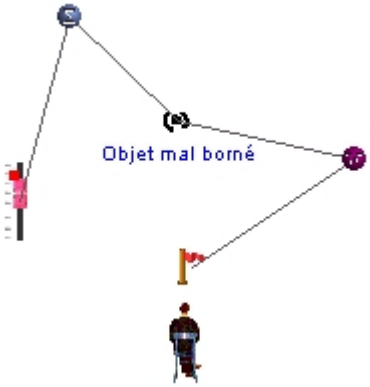
Lors de notre pratique, nous avons imaginé une solution pour créer ces bornes avec les outils existants. Nous lions un objet "Flag" (uniquement graphique) à l'un des objet du groupe devant être "borné", avec une contrainte "somme" ou "différence des distances constante". Lorsqu'il heurte l'Avatar, le Flag s'immobilise, et la contrainte interdit le mouvement des objets qui lui sont liés. Un Objet source peut parfois jouer le rôle du Flag. Ce truc ne fonctionne pas avec une contrainte en « produit » ou « ratio de distances », pour lesquelles l'avatar ne représente plus une borne minimale - car le rapport n'ayant pas de limite.



L'Objet se déplace en opposition avec le Flag. On utilise une contrainte « somme des distances constante »



L'Objet est borné d'un côté par sa propre collision avec l'Avatar, de l'autre par la collision du Flag avec l'Avatar



La contrainte Flag/Objet est en « rapport des distances constantes ». L'Objet n'est pas borné à l'extérieur car le rapport n'a pas de limite : le Flag se déplace sans fin vers l'Avatar

Figure 4-3 Création d'une borne avec un Flag

4.5 Des Handles à l'action clairement audible

Nous avons réalisé, sur une œuvre contemporaine jouée par un orchestre de chambre, un Handle agissant sur la « Présence » perçue des instruments. Ce contrôle fonctionne bien. Toutefois, le résultat est assez subtil et seule une oreille exercée apprécie vraiment le prodige…

Quelques démonstrations du mixage interactif de *pop-music* en cours de réalisation ont révélé la nécessité que chaque Handle ait un effet clairement audible à tout moment. Il faut également que les effets produits par le déplacement de chaque Handle varient aussi peu que possible tout au long de la chanson. Si ce n'est pas le cas, l'auditeur même averti est perdu avec les Handles et est incapable de déduire l'action de chacun d'entre eux.

Pour cette raison, nous nous sommes orientés vers des transformations plus évidentes, affectant non seulement le son mais aussi la musique.

Nous avons rencontré les difficultés suivantes :

- certains des paramètres n'ont de sens que sur une partie de la chanson. Par exemple, il est risqué de densifier beaucoup l'orchestration du couplet, qui risque d'être entendu plus chargé que le refrain (si, par exemple, l'auditeur revient sur un réglage « Densité » faible entre-temps).
- Pour certains paramètres, on ne peut se permettre que des variations subtiles : une exagération détruirait l'image ou l'équilibre. Si en plus l'auditeur sait que « parfois, certains curseurs n'ont pas d'effet », il n'entend vraiment rien.

Pour résoudre ces problèmes, on pourrait :

- faire disparaître les curseurs lorsqu'ils sont inactifs (les contraintes pour le faire existent plus ou moins, mais leur emploi est extrêmement complexe)
- rendre les contrôles plus efficaces, en regroupant des Handles, jusqu'à aboutir, par exemple, à un choix sur seulement une dimension (voir 1.5). Cela diminuerait le nombre de curseurs disponibles mais augmenterait leur intérêt.

4.6 Respect du style et reconnaissance de la signature

Une fois une première version de la bande multipiste réalisée, on a tenté de la mixer selon différentes optiques. La difficulté en fût grande pour plusieurs raisons :

- La densité de l'arrangement laissait peu de place à des traitements sonores apparents. Nous avons réalisé que les différences entre les mixages proposés se devaient d'être énormes, pour être convaincantes et entendues par l'auditeur non-spécialiste.
- Malgré la relative originalité de la musique proposée, des connotations stylistiques demandaient un traitement sonore encore trop défini. Ceci confirme ce que nous affirmions au 2.1.1 : au cours de notre pratique, nous n'avons pas réussi à proposer deux versions d'un mixage différant uniquement sur le plan de la balance ou des traitements sonores.

Les décisions prises par l'auditeur peuvent-elles être du même ordre que celles prises par l'artiste ? Cette question reste ouverte. Toutefois, nous pensons que l'auditeur ne doit pas être l'interprète. On ne lui donne pas à choisir entre différentes options, dont l'une serait meilleure que les autres. Comme pour toute composition, il faut savoir dire non à certaines solutions, moins intéressantes, car ce n'est pas le rôle de l'auditeur de faire les choix esthétiques. Il s'agit plutôt de poser plusieurs problèmes, dont chacun aura sa solution. L'auditeur choisit la question qu'il veut voir traitée par le compositeur. Le compositeur y répond par avance au moyen de règles pré-établies. Composer une oeuvre interactive, ce n'est pas reporter ses choix à plus tard, ni fournir un travail de composition incomplet. Une activité artistique se définit par ses choix. Ne pas faire de choix laisse toutes les possibilités ouvertes mais limite l'imagination à l'écriture de l'œuvre, ce qui nuit à sa richesse. Un aspect de la maîtrise artistique se caractérise par la facilité et la précision de représentation mentale de l'œuvre finie. Faire les choix le plus tôt possible permet de l'améliorer.

Doit-on proposer des points de vue se distinguant au point de vue stylistique ? Donner à entendre différentes « versions » de sa musique présente le danger, pour l'artiste, de rendre sa signature méconnaissable. Etudiant l'importance des caractéristiques de la musique pop sur la reconnaissance d'une signature, Emiliano Flores [3] montre que dans la *pop-music* enregistrée, la voix est l'élément prédominant de la signature. L'auditeur reconnaît le groupe principalement grâce à la voix du chanteur. Ainsi, il est possible que nous ayons une certaine liberté stylistique vis-à-vis de l'accompagnement.

Pourtant, lors des écoutes de la chanson après sa production, aucun auditeur ne s'est plaint de ne pas pouvoir reconnaître un style musical, et ce malgré les différences importantes introduites dans les arrangements. L'interactivité ne semble pas avoir mis en danger la signature du groupe.

4.7 Problèmes survenant lors des Changements d'Orchestration ou d'Arrangement

Remplaçons un instrument par un autre. L'arrivée d'un nouvel environnement autour d'un timbre instrumental provoque un effet extraordinaire. L'effet de masque donne l'illusion que ce timbre est modifié. Mais l'effet de masque a aussi un effet négatif, puisqu'il modifie l'équilibre spectral perçu de la source, et en conséquence celui du mixage ainsi que sa balance. Il peut être nécessaire d'introduire une égalisation sur une source au passage de l'une à l'autre présentation.

Un changement d'arrangement, au niveau rythmique, ou un changement d'interprétation, peut modifier la densité sonore: il peut être nécessaire de ré-ajuster les réverbérations, ou même des éléments de balance. Par exemple, lorsque le flot sonore se fait moins dense, on peut privilégier des éléments harmoniques au release plus long, découvrant leur déroulement sonore.

Introduisons des sources optionnelles : l'ajout de ces sons modifiera l'équilibre spectral, qu'il faudra peut-être corriger. Lorsque l'ensemble des instruments accompagnateurs se fait plus dense, il est nécessaire de ré-égaliser la voix lead pour la faire ressortir.

Conclusion de la 4ième partie

Compte tenu de l'originalité du produit, on pourrait effectuer une grande variété d'expériences pour l'évaluer. Elles peuvent être d'ordre ergonomique, culturel, musical ou technique. Observant l'auditeur qui se trouve confronté au produit, nous pouvons évaluer la qualité de l'interface, d'un point de vue ergonomique. L'action de chaque curseur est elle clairement définie ? La présentation graphique est-elle un bon choix ? Afin d'évaluer la composante culturelle dans les choix du mixeur, on pourrait comparer les choix des utilisateurs novices avec ceux effectués par des ingénieurs du son. Puis nous pourrions observer la réaction de l'auditeur à la situation interactive. A t'il envie de jouer avec la machine ? Les images sonores produites sont-elles réellement originales et différentes d'un utilisateur à l'autre ? D'un point de vue strictement technique, nous pourrions juger enfin la cohérence des images intermédiaires, obtenues par la moyenne des auditeurs.

Conclusion

Nous avons essayé de définir au mieux ce que nous entendions par « interactivité dans un enregistrement sonore ». Nous nous sommes posés la question du rôle de l'ingénieur du son, déterminant quel type de décisions il doit prendre et pourquoi. Nous nous sommes confrontés à plusieurs affirmations plus ou moins répandues, dont nous ne savions que penser lorsque nous avons entamé notre travail :

- « il existe plusieurs manières de mixer une même bande »
- « un mixage de *pop-music* est fragile »
- « on ne peut pas théoriser le mixage »

Il s'est avéré qu'il n'est pas si évident de faire deux mixages distincts réussis de la même bande. S'il est vrai que deux mixeurs appliquent chacun leurs techniques, leurs deux versions seront pourtant probablement équivalentes au point de vue stylistique. Des modifications plus profondes auront leur origine au niveau musical : lorsque le mixeur dispose de pistes redondantes, il a la possibilité de faire des choix musicaux.

Il est vrai qu'un mixage de *pop-music* est long à mettre au point dans le détail. Nous avons également passé beaucoup de temps à rééquilibrer la balance ou a retoucher des égalisations, trouvant des compromis pour que notre mixage interactif fonctionne dans tous les cas de figure. Pourtant il existe une marge de manœuvre quand à la perfection sonore nécessaire, d'autant plus importante que nous introduisons des éléments nouveaux : l'auditeur captivé par les modifications de la musique ne porte aucune attention à la cohérence de l'image sonore.

Notre 1ière partie est bien un début de théorisation du mixage. Nous faisons l'inventaire des paramètres de bas niveau mobiles. Dans un travail plus approfondi, on pourrait quantifier leur variation et leur influence, automatiser leur réglage, trouver des liens entre eux et induire des paramètres d'un niveau plus élevé.

Il reste encore beaucoup à faire pour mettre au point l'outil dont nous aurions besoin. Le système de la résolution de contrainte semble intéressant. Il reste à lui trouver une présentation qui pourra s'adapter aux nombreux cas de figure du mixage tout en restant légère.

Par l'interactivité, nous remettons partiellement en cause une caractéristique forte de la *pop-music*: sa fixation enregistrée sous une forme définitive. Une oeuvre d'art traditionnelle est l'image de son temps. Les critères esthétiques qui ont réglé son élaboration évoluent. L'interactivité peut permettre à l'œuvre de traverser le temps, évoluant avec lui : le compositeur peut permettre son *adaptation* à un désir d'écoute différent, selon le degré de liberté accordé à chaque paramètre. L'œuvre peut ainsi s'adapter aux mouvements culturels.

Bibliographie

[1] **Bergame Périaux – « Cohérence de l'Image Sonore en Mixage Variétés 5.1 », Mémoire CNSMDP, 2000**

[2] **Anders Friberg and Al. – "Generating Musical Performances with Director Musices", Computer Music Journal 24:3, pp. 23-29, Fall 2000**

[3] **Emiliano Flores – « Signature et contexte dans les musiques pop enregistrées », Mémoire ENSLL, 2000**

[4] **Mathieu Coupat, « Angel de Massive Attack – mémoire d'analyse musicale », CNSMDP, 1999**

[5] **Olivier Delerue - « MusicSpace : Documentation », Sony CSLP**

[6] **F. Pachet, O. Delerue – « On-The-Fly Multi Track Mixing », 109th AES Convention, Los Angeles 2000**

[7] **Moore - "an introduction to the Psychology of Hearing", Academic Press, 1994**

[8] **Mike Williams et le Guillaume Le Dû – « Multi-Microphone Analysis for Multichannel Sound Recording », AES Preprint 4997 (A-5)**

[9] **Christian Braut - « Le livre d'or de la norme MIDI », 1994**

[10] **Jean-Marc Lyzwa - « Fragments/Pour mémoire », notes de travail, CNSMDP 1999**

[11] **Itai M. Neoran – « Surround Sound Mixing Using Rotation, Stereo Width and Distance Pan Pots », 109th AES Convention, Los Angeles 2000, preprint #5242**

Discographie

(A) Madonna - "Music", 2001

(B) Björk - "Debut", 1993

(C) Björk - "Selma Songs", 2001

(D) U2 - "All that you can't leave behind", 2000

(E) Yann Tiersen - "Le fabuleux Destin d'Amélie Poulain", 2001

(F) Boyz II Men - "II", CS, 1994

(G) Eagles - "Hell Freezes Over", CS, 1994

(H) Chick Corea - "*", CS, 199***

Annexe I Installation autour de MusicSpace

I.1 Description de MusicSpace

I.1.1 Interface graphique

L'interface proposée actuellement permet de placer les sources et l'auditeur sur un repère plan à deux dimensions. Dans un espace graphique, des « Objets » sonores sont placés à différentes distances d'une « Avatar » représentant l'auditeur. La position de ces « Objets » est repérée en coordonnées polaires autour de l'« Avatar » : à la distance graphique correspond la distance perçue, tandis que l'angle représente la direction dans laquelle le son est entendu entre les enceintes.

L'interface est ainsi adaptée à représenter et à faire évoluer le positionnement des sources dans l'image sonore.

I.1.2 Pilotage de systèmes spatialiseurs

MusicSpace fournit la position de l'objet sur le repère 2D. Ces systèmes traduisent la donnée en informations auditives : déplacement entre deux enceintes, par exemple. Ils sont capables a priori de combiner les effets d'un ensemble de traitements.

I.1.2.1 La spatialisation avec DirectX

MusicSpace fournit des informations de position sous la forme de coordonnées cartésiennes, ainsi qu'une information de niveau sonore corrélée à la distance objet-avatar : lorsqu'on est loin de l'avatar, le niveau est plus faible.

DirectX reproduit la situation visuelle sur un système quadriphonique en monophonie dirigée. Ce spatialiseur est très limité : on a seulement accès aux courbes de volume. Peut-être serait-il possible d'intégrer le pilotage de la réverb et d'un égaliseur sommaire. Mais le résultat sonore a peu de chances d'être probant : peu de paramètres accessibles ; effets plus adaptés à une spatialisation rapide de sources simples (par exemple en VRML), qu'à un vrai mixage musical. Par ailleurs, DirectX ne permet pas de positionner entre 2 enceintes diamétralement opposées, effet très utilisé lors du mixage multicanal.

En conséquence, il est nécessaire que tous les traitements appliqués aux sources soient enregistrés sur le disque dur, disponibles sur des pistes séparées. Le nombre de pistes nécessaires est rapidement important.

I.1.2.2 La spatialisation avec le Spat de l'IRCAM

On transmet en MIDI des informations de position (angle, distance, orientation de la source et de l'auditeur). Le Spat met en oeuvre des traitements complexes (réverbération, filtrage, ajout de premières réflexions, traitements transauraux, HRTF) reproduisant la réponse naturelle d'une salle.

I.1.2.3 Pilotage MIDI

Il est possible de commander n'importe quel appareil recevant des messages MIDI : console ou module d'effets. L'esclave pourrait ne pas concerner la spatialisation, mais par exemple le filtrage, ou la compression....

I.1.3 Les contraintes

MusicSpace repose sur la résolution de systèmes de contraintes. Si, de par la nouvelle position d'un objet, les règles par lesquelles les autres objets qui lui sont liés ne sont plus satisfaites, ces objets doivent être déplacés.

a) Contraintes sur la distance

Ce premier ensemble de contraintes régit les rapports de distance des objets, l'angle restant libre. On trouve les Contraintes en :

- Différence des distances constante
- Somme des distances constante
- Produit des distances constant
- Rapport des distances constant

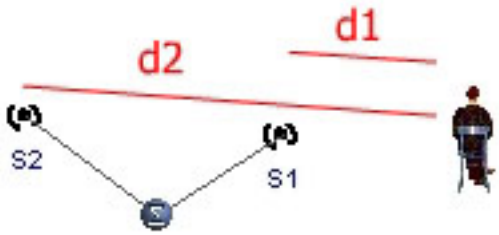


Figure 1 Distances prises en compte par la contrainte

b) Contraintes sur l'angle

Ces contraintes régissent les conservations angulaires lors de mouvements rotatifs autour de l'Avatar.

I.2 La source sonore multipiste

On a le choix entre les deux modes de fonctionnement de MusicSpace :

- en mode « autonome », les sources sont sur le disque dur. Tous les traitements doivent être effectués par le PC qui délivre en sortie un signal quadriphonique par l'intermédiaire de DirectX.
- le fonctionnement en MIDI, mis au point lorsque le programme s'appelait encore « MidiSpace ». Il permet de piloter n'importe quel instrument MIDI externe.

Les deux modes de fonctionnement n'utilisent pas les mêmes Objets dans la représentation graphique. Il y a des différences dans la manière d'utiliser le logiciel et dans les fonctions disponibles.

Il y a plusieurs inconvénients à travailler avec la version autonome actuelle. Aucun pilotage de réverbération ou d'égalisation ne sont implémentés dans MusicSpace – bien que la présence d'algorithmes basiques dans DirectX rendent cette implémentation possible. DirectX est pourtant limité en matière de traitements : on peut modifier peu de paramètres. Il est plutôt destiné à des applications dans lesquelles la qualité sonore importe peu. Il est possible de contourner cette limitation par une post-production poussée : sur le disque dur seraient placés des pré-mixages et les enregistrements des traitements nécessaires. Cette méthode, qui nécessite de lire un grand nombre de pistes simultanément, est probablement viable, pour peu que le disque dur soit assez rapide, et que la lecture de toutes les pistes soit synchronisée à quelques échantillons près. Son manque de souplesse pose toutefois des problèmes majeurs : toute l'image sonore doit être pensée au préalable (pré-mixages) ; un grand nombre de sources plus ou moins équivalentes présentes sur l'écran (sources traitées par tel ou tel effet), sans un sens véritable dans la représentation ; limitations des traitements possibles (seuls les traitements assimilables à un changement de gain sont envisageables : impossible par exemple de faire entendre une réelle augmentation du temps de réverbération ou de pré-delay). Nous pensons ainsi que le mode « autonome » existant est mal adapté à nos expérimentations.

Utilisant le pilotage MIDI, il nous sera possible d'accéder à une grande partie des traitements disponibles lors d'une production musicale professionnelle.

Cette démonstration expérimentale donnera une idée de ce qui pourra être fait avec un système comme MusicSpace dans un futur proche, lorsque des traitements plus élaborés seront disponibles à l'intérieur même d'un PC. Dans un deuxième temps, nous pourrons réaliser une version réduite de la production, à base de pré-mixages, qui sera autonome et pourra servir de démonstration pour MusicSpace lui-même.

I.3 L'installation proposée

I.3.1 Moyens Hardware

Nous disposons d'une bande 24 pistes, qui alimente les 24 entrées d'une console numérique Yamaha 03D. Celle ci nous permet d'accéder en MIDI pour chaque tranche aux niveaux, aux panoramiques en Surround, à 4 envois d'effets et aux égaliseurs. Par ailleurs, nous utilisons 3 modules d'effets externes : Lexicon 300, Eventide H3000 et TC Electronics M6000. Les retours d'effets et les 4 sorties audio de la O3D entrent dans une SSL Axiom MT, qui sert à la diffusion quadriphonique. Les paramètres de la 300 et du H3000 sont pilotables en MIDI.

Le pilotage des égaliseurs et des effets constitue un détournement de MusicSpace. En effet, la représentation graphique n'a pas de signification pour ces objets, ce qui a parfois pour conséquence de rendre les contraintes difficilement utilisables.

I.3.2 La commande MIDI

Console de dimensions modestes, la O3D possède pourtant des caractéristiques intéressantes : grand nombre de bus en départ, et panpot en surround intégré. Son implémentation MIDI est très complète : il est possible de piloter absolument tous les paramètres internes, sur chaque tranche, auxquels on accède par control change.

Seul est à redouter un engorgement du bus MIDI, qui permet un débit de seulement 31250 bauds. L'interface étant sérielle asynchrone, la transmission d'un octet a une forme spéciale, et il faut transmettre 10 bits pour 8 utiles. Ce débit correspond ainsi à environ 3000 octets MIDI par seconde. [9]

Les paramètres des machines MIDI sont en général accessibles par SysEx, qui présentent l'avantage d'être particuliers à chaque appareil. C'est le cas de la 03D. L'inconvénient des SysEx est leur longueur.

F0 – ID – data – F7
ID=manufacturer ID
Data=at least instrument ID, controller # and value

Figure 2 Exemple de Message SysEx

Par chance, tous les paramètres de la 03D sont aussi pilotables par Control Change. Ces messages sont parmi les plus légers:

Bn – cc – vv
N=channel #
cc=controller #
vv=value

Figure 3 Exemple de message Control Change

Il y a plus de paramètres accessibles sur la O3D que de Control Change existants : c'est pourquoi un utilitaire de mapping est présent sur la console. Il a été nécessaire de sélectionner les paramètres qui seront utiles pour la production proposée. Nous avons donc élaboré un tableau de correspondances.

Control Change	Paramètre	Tranche
1 à 24	Trim de Gain	1 à 24
25	Aux3 Send	24
26	Aux3 Send	7
27	Aux3 Send	8
28	Aux2 Send	24
29	Aux1 Send	1
30	Aux1 Send	9
31	Aux1 Send	11
32	Aux1 Send	13

Control Change	Paramètre	Tranche
33	Aux1 Send	19
34	Fader	Master Aux1
35	Aux1 Send	24
36	Aux1 Send	3
37	Aux1 Send	4
38	Aux1 Send	5
39	Aux1 Send	6
40 à 61	Pan	3 à 24
64 à 87	Fader	1 à 24

Figure 4 Correspondance entre numéro de Control Change et paramètre sur la 03D

I.3.3 Des objets pour piloter la Yamaha03D

Il existe dans MusicSpace un objet capable d'envoyer des messages de Control Change. Ses caractéristiques sont les suivantes :

- entrée : coordonnées (X,Y) ou (angle, distance)
- paramètres : bornes de variation (entre l'avatar et la bordure de l'écran, ou le contraire).
- sortie : 1 ou 2 Control Change indépendants de 0 à 127

Pour faciliter le travail dans MusicSpace, nous avons demandé à Olivier Delerue de créer des objets spécifiques pour la 03D. Nous disposons maintenant de deux Objets séparés pour le Volume et le Panoramique.

Des objets « Volume O3D » ont été ajoutés. La position de la source est représentée en coordonnées polaires autour de l'avatar, tandis que le volume est représenté par la distance. Ils sont capables d'envoyer les Control Change correspondant aux mouvements des faders sur 24 canaux.

Conjointement, des objets « Position O3D » ont été créés. Ils envoient les Control Change correspondant au panoramique en coordonnées cartésiennes.

On a imaginé une contrainte permettant de lier ces deux objets. La contrainte « en étoile » fait la correspondance entre la position cartésienne de l'objet « Position 03D » et la position polaire de l'objet « Volume 03D », dans les 2 sens.

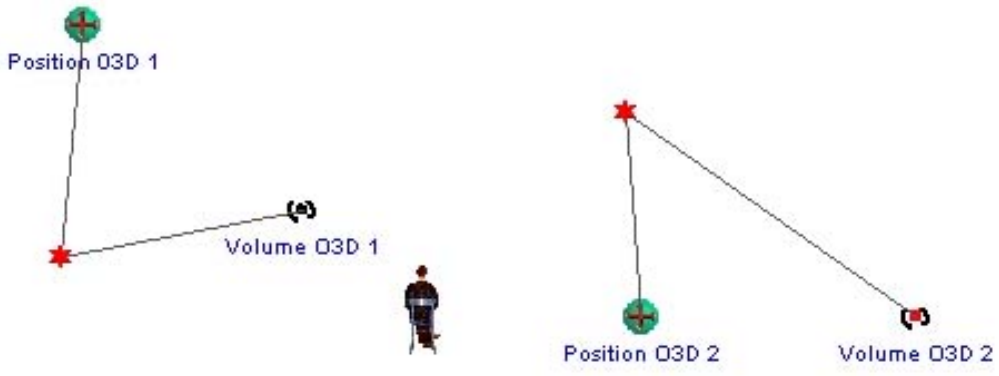


Figure 5 Les nouveaux objets 03D et la contrainte qui les relie

Afin d'améliorer la clarté de la représentation graphique, la taille des objets et celle de la police utilisée pour les nommer ont été fortement diminuées.

Annexe II Notions de Prise de Son

Nous abordons ici les notions qui nous ont été utiles pour comprendre les raisons pour lesquelles l'image sonore pouvait perdre sa cohérence. Dans un contexte ou chaque source sonore peut-être réutilisée de manière plus ou moins prévisible, il importe de prendre quelques précautions à la prise de son.

II.1 La perception de la distance

Il est actuellement très difficile de donner à l'auditeur l'impression qu'il écoute une source lointaine, si on dispose seulement d'un enregistrement réalisé en proximité.

L'audition distante [7] se traduit acoustiquement par des différences dans :
- le niveau sonore
- le spectre (filtrage des aigüs)
- l'existence, la densité, le timbre et l'enveloppement des 1ières réflexions
- le rapport Champ Direct / Champ Réverbéré
- la décorrélation interaurale du champ réverbéré
- le timbre instrumental

On sait actuellement à peu près émuler le comportement d'une salle, moyennant des processeurs puissants pour recréer les premières réflexions et la réverbération. Mais il n'est pas possible de capter le vrai timbre d'un instrument en proximité : celui-ci a besoin d'un peu de distance pour se développer, notamment dans les graves. Le rayonnement instrumental n'est pas homogène et une prise rapprochée n'en capte qu'une partie. De plus, ce rayonnement possède une certaine largeur qui n'est pas restituée en monophonie.

C'est pour ces raisons que l'on utilise des microphones distants, en particulier des couples, pour des enregistrements reproduisant une situation naturelle. Aussi, il semble utile d'apprendre à intégrer un couple à un mixage 5.1.

II.2 Les images fantôme

En écoute binaurale, une image fantôme est perçue entre 2 haut-parleurs diffusant des signaux comparables, que l'oreille a fusionné pour créer une image virtuelle. Des différences d'amplitude ou temporelles permettent de tirer l'image vers l'un ou l'autre des haut-parleurs.

La monophonie dirigée permet ainsi une localisation solide, en stéréo. On fait varier l'amplitude diffusée par chaque enceinte suivant une « loi de panoramique stéréo» simple, qui assure la constance de l'énergie globale, et l'homogénéité du niveau sonore dans toutes les directions.

$L=\cos\ \alpha \qquad R=\sin\ \alpha$

$L^2+R^2=1$

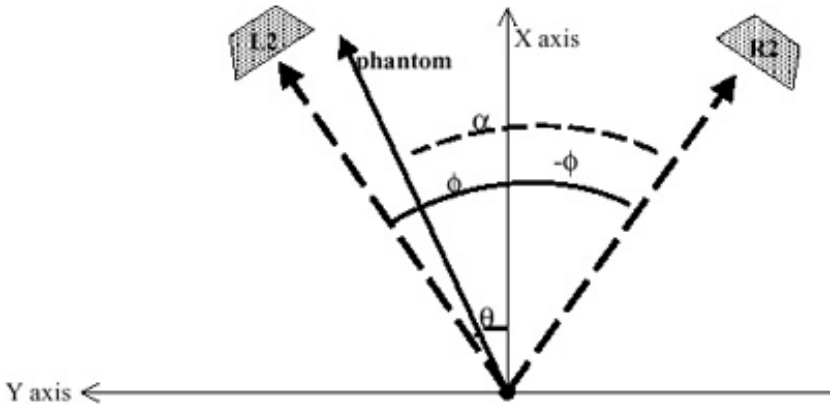


Figure II-1 Panoramique stéréo [11]

Lorsqu'il écoute un système de reproduction en surround, l'auditeur peut avoir 2 attitudes différentes : il peut conserver sa tête immobile, constamment dirigée vers les enceintes frontales. C'est le cas au cinéma, et dans les enregistrements musicaux où le sujet est placé devant l'auditeur. Il peut également choisir de tourner la tête, pour faire face aux sources qui l'entourent : ceci peut arriver si l'on fait figurer des sources importantes sur les côtés ou à l'arrière, qui orienteront l'attention de l'auditeur.

De ces deux attitudes d'écoute découlent deux manières fondamentalement différentes d'enregistrer et de construire l'image sonore. En effet, elles conduisent à deux modes de perception distincts. La localisation d'une source latérale est essentiellement monaurale, alors que celle d'une source frontale est aussi binaurale [7].

II.3 Panoramique en surround de couples stéréophoniques

L'un des objectifs premiers de l'interface de MusicSpace était de rendre les sources sonores mobiles sur 360°. L'une des contraintes permet de lier deux sources à la manière d'un couple stéréophonique. Ce type de déplacement a t'il un sens ?

Il serait intéressant d'adapter le principe d'un panoramique en monophonie dirigée à un panoramique en « stéréophonie dirigée », permettant de placer un couple n'importe où sur 360°, diffusé sur 2 enceintes ou plus. Par l'étude critique de quatre méthodes, nous recherchons ici une manière d'approcher ce résultat.

II.3.1 Positionnement de l'image fantôme par accentuation de l'un des canaux du couple

Afin de donner un déséquilibre minimal à ces 2 canaux, on peut identifier la position de la source fantôme principale présente dans le couple stéréo, et s'en servir pour

sélectionner les 3 enceintes qui serviront à la diffusion. C'est une version améliorée de la monophonie dirigée, permettant de reproduire au moins partiellement le rayonnement instrumental dans sa largeur, bien qu'elle introduise des distorsions trop importantes pour être applicable à des couples présentant des sources stéréo multiples.

On pourrait parfaitement utiliser une des lois de panoramique surround, pour tenir compte de la position frontale, latérale, ou arrière des enceintes choisies.

II.3.2 Positionnement par panoramique surround de chacun des 2 canaux du couple

Cette approche, à première vue assez naturelle, est celle utilisée par la contrainte « couple stéréo » de MusicSpace [5]. Elle conduit à se servir du panoramique de la console pour placer sur 360° chacun des canaux d'un couple, qui se retrouvent dirigés entre deux enceintes. On imagine créer ainsi des haut-parleurs virtuels, sur lesquels seront dirigées les sources réelles, et entre lesquels devraient apparaître les sources fantômes. On pourrait par exemple essayer de les placer en extra largeur.

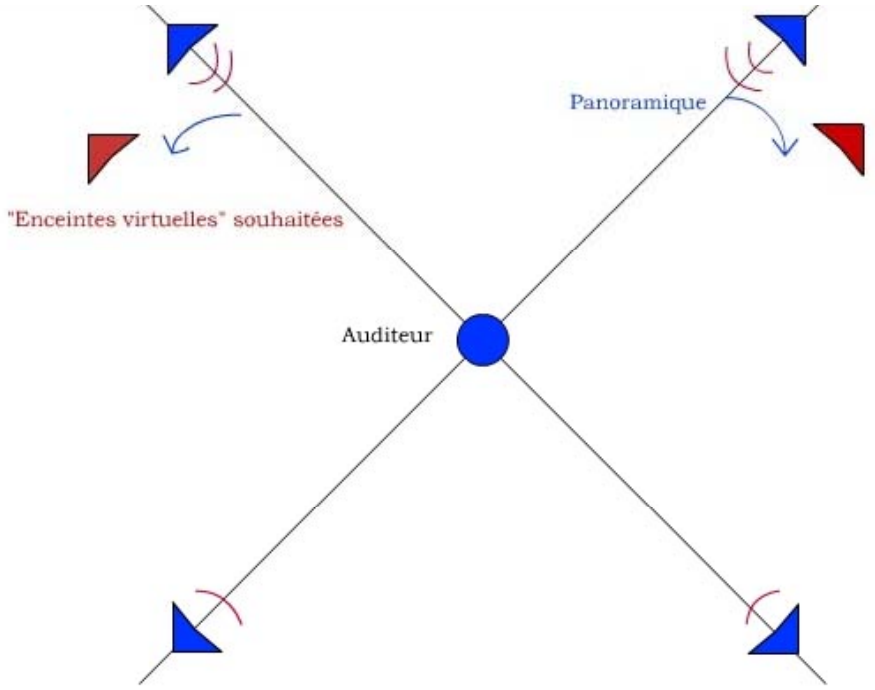


Figure II-2 Tentative de placement d'un couple en extra-largeur

La pratique démonte rapidement cette idée. Une source fantôme qui devrait être centrale n'est plus localisable. L'explication est simple : une source qui devrait être entendue en centre gauche, donc légèrement à droite de l'enceinte gauche, se retrouve diffusée dans des proportions variables à la fois par les enceintes droite, gauche, surround gauche et surround droite, c'est à dire sur les 4 enceintes, ce qui est déroutant... Il paraît clair que cette solution est incohérente. La perception binaurale, sur laquelle repose le principe du couple, est extrêmement sensible et ne peut être généralisée de cette manière [10].

II.3.3 Séparation de la source fantôme principale et de son environnement par codage M-S

Pour un plug-ins Waves de Pro-Tools, Neoran [11] propose de coder le signal du couple en M-S, puis de panoramiquer M en monophonie dirigée. S, qui ne contient

alors plus d'information de localisation, peut être diffusé par n'importe lesquelles des autres enceintes selon les besoins… Si on réutilise ainsi le signal du couple, il ne s'agit plus à proprement parler de stéréophonie. Les qualités du timbre capté en proximité risquent d'être compromises par le codage M-S, qui inverse la phase du signal. Mais cette solution a le mérite d'introduire l'idée de la séparation instrument/environnement.

II.3.4 Adaptation de la prise de son à la situation d'écoute

L'écoute binaurale obéit à des règles différentes selon la position de la source autour de la tête. La localisation dans l'image stéréo étant essentiellement binaurale, on doit adapter la prise de son aux conditions d'écoute pour obtenir une image fantôme réellement équilibrée.

Selon Williams et Le Dû [8], la captation doit faire correspondre parfaitement les différences de temps et d'intensité. Quelle que soit la position de la source sur 360°, ils se placent en situation d'écoute frontale et proposent une méthode de calcul de l'angle, de la distance et des offsets temporels à donner aux microphones pour obtenir cette correspondance en tenant compte de l'espacement irrégulier des enceintes en 5.1. Ainsi, à une captation correspond un couple d'enceintes donné, devant lequel l'auditeur se place de manière frontale.

En conclusion, il n'y a pas de solution simple quant au panoramique de couples stéréophoniques. De façon sûre la solution proposée dans MusicSpace ne donne pas de bons résultats. La seule méthode réellement rigoureuse est, pensons-nous, d'enregistrer la source en l'écoutant dans la position dans laquelle l'auditeur l'entendra, panoramiquant les microphones sur des enceintes. Une modification interactive de la position surround de la source pourra alors se faire par tuilage de ces enregistrements.

II.4 L'effet « cocktail party »

Dans une situation d'écoute difficile, dans laquelle le signal utile est partiellement couvert par d'autres signaux parasites – bruits ou instruments accompagnateurs -, les oreilles travaillent en coopération. Par comparaison de leurs deux signaux, le cerveau est capable d'extraire le signal utile. L'efficacité du procédé dépend de la corrélation des signaux arrivant sur les deux oreilles [7].

Des modifications de positionnement des sources peuvent donc avoir une influence sur le pouvoir séparateur de l'oreille.

II.5 L'effet de Masque

La membrane basilaire détecte les vibrations transmises jusqu'à l'oreille interne. Sa constitution rend les hautes fréquences plus difficilement détectables en présence de basses fréquences. Ainsi, la cohabitation de sons de différentes natures se fait au prix d'un masquage. De fait, les sons graves sont généralement « masquants » des sons aigüs, qui sont « masqués ». Ce phénomène change l'équilibre spectral d'un son complexe, surtout au voisinage des harmoniques du masquant. Par ailleurs, il fait apparaître des sons différentiels, ainsi que des battements. Les transitoires sont peu affectés.

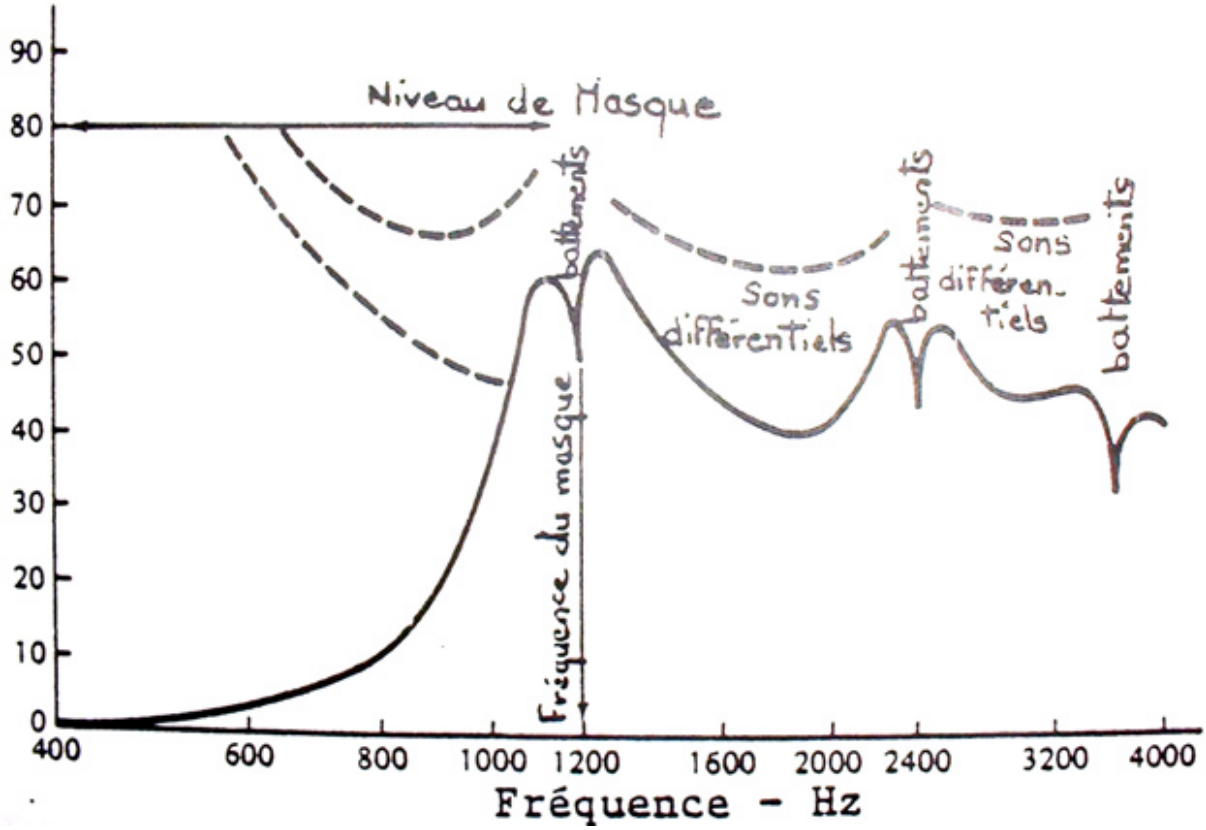


Figure II-3 Elévation du seuil de masquage en présence d'un masquant à 1200Hz (d'après Moore [7])

En conséquence, un timbre est modifié lorsqu'il est mixé avec des timbres d'autres natures. Ceci peut conduire à une mauvaise définition, à une mauvaise intelligibilité, mais surtout à une modification de la balance.

L'orchestrateur et l'arrangeur sont conscients de ces problèmes : l'un les résout au niveau fréquentiel, en utilisant les bonnes associations sonores, et en évitant des empilements d'instruments masquant la mélodie. L'autre les résout au niveau temporel, en construisant la musique en sections entre lesquelles les instruments se répondent. L'ingénieur du son en tient compte lorsqu'il traite ses sons, imbriquant les sources entre elles. Une méthode consiste à éviter de trop grandes variations dans l'influence que les sources ont les unes sur les autres : il s'aide d'égaliseurs, pour agir sur le timbre et le caractériser, et de compresseurs, pour limiter les variations de niveau. Chaque source dispose ainsi de son propre espace temporel et spectral.

II.6 Utilisation de microphones distants

Nous nommons ici « microphones distants » des capteurs placés dans une zone où effet de salle et effet de proximité sont tous deux limités.

Laissons la localisation de côté pour se préoccuper du timbre. Enregistrons un instrument à l'aide de deux microphones distants. Comme nous l'avons mentionné précédemment, ceux-ci reproduiront un timbre plus naturel que des capteurs de proximité, et permettront de reproduire le rayonnement dans toute sa largeur. Mais la largeur de l'instrument - donc celle du rayonnement, est limitée. Nous voulons garder une certaine cohérence entre les signaux délivrés par ces deux capteurs, afin que l'oreille puisse les confondre. Nos microphones seront donc nécessairement assez proches l'un de l'autre, et nous les retrouvons placés... en couple stéréo.

Que nos microphones soient placés *rigoureusement* en couple stéréo A-B n'a pas d'importance. Nous recherchons ici le meilleur timbre possible sans se préoccuper outre mesure de la localisation. Un timbre plus riche sera obtenu en écartant les microphones d'une distance en rapport avec la largeur du rayonnement instrumental.

Selon la source enregistrée, un seul point est alors à surveiller : respecter au maximum la phase du signal, sur tout le spectre. Pour éviter le détimbrage, il ne faut pas mélanger électriquement des sons partiellement hors-phase, et simplement éviter de placer sur la même enceinte des signaux présentant ces risques [13].

- Les paires non coïncidentes sont hors-phase pour toutes les longueurs d'onde du double de leur écartement.
- Des omnidirectionnels sont plus sensibles au problème que des cardios dirigés en opposition.
- Utiliser les delays électroniques pour annihiler les delays acoustiques.

On peut ainsi parfaitement utiliser des microphones distants si l'on évite dans la mesure du possible d'en panoramiquer les composants sur la même enceinte. On utilisera avantageusement la forte densité d'enceintes frontales en 5.1.

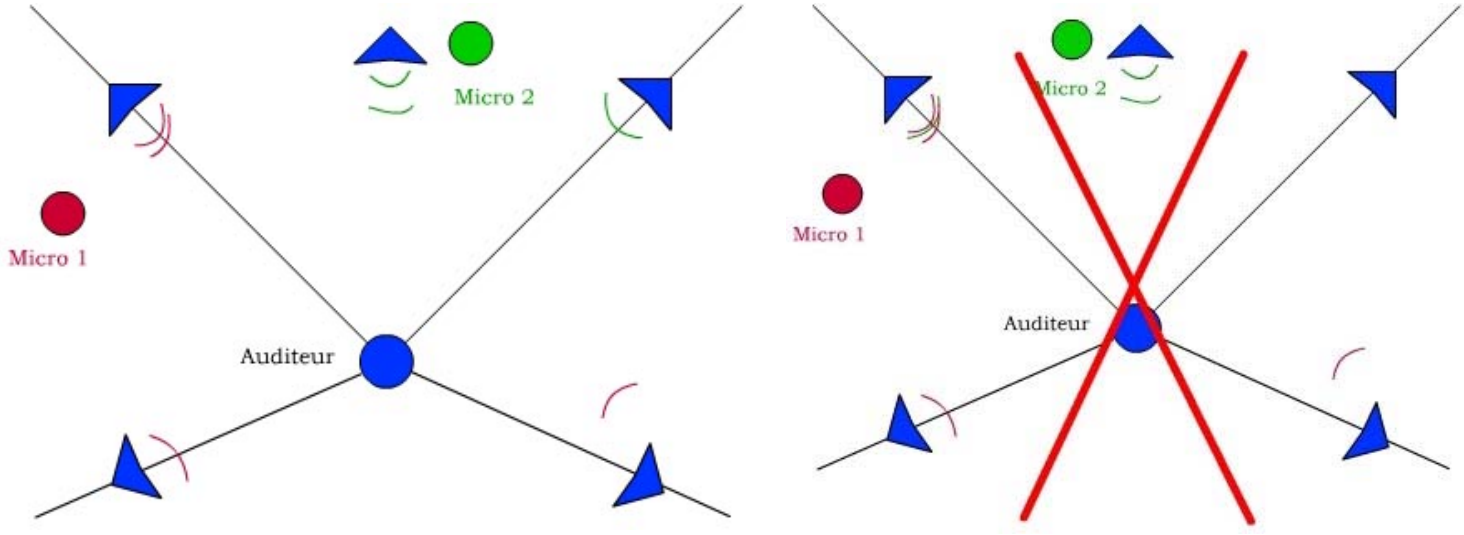


Figure II-4 (a) Positionnement de Sources Corrélées
(b) Positionnement Incorrect de Sources Corrélées

Annexe III L'enregistrement Sonore : Interprétation, Adaptation et Fixation

Lorsque l'œuvre est transmise de manière indirecte, par la description – partition musicale ou vidéo chorégraphique, sa création laisse la place à *l'interprétation*.

Le fait d'enregistrer une pièce musicale existante en tant que telle est une *adaptation*. Il n'y a qu'une manière d'adapter l'évènement sonore au support : celle qui respecte les codes attachés au style et au mode d'interprétation, lorsqu'on les connaît. Lorsque l'auditeur s'attend à un type de son particulier, présent dans la culture de chacun, l'ingénieur du son doit le restituer: recherche de la quantité de réverbération optimale pour un orchestre par 2 de Mozart ou de la présence adaptée à un lied de Schubert ; adaptation de la largeur d'une batterie au jeu du batteur et à l'effectif de la formation enregistrée.

Dans le cas des musiques à base de sons enregistrés, le stade de l'enregistrement est celui de la *fixation*. Fixée sur un support, l'oeuvre d'art atteint un état définitif. Cet état, qui n'évolue pas dans le temps, est la conséquence d'une esthétique.

Annexe IV : Contenu de la bande pré-mixée

Source	Enveloppement	Arrangement	Position	Piste Premix	Nom piste
Gt c 4V	Enveloppée	Acoustique	SL - R	19	
Gt c 87	Enveloppée	Acoustique	SR - L	20	
Rhodes Sat SL	Enveloppée	Electronique	SL	21	
Rhodes Sat SR	Enveloppée	Electronique	SR	22	
Accordéon Accords	Enveloppée	Toutes	L-R	23	
Accordéon Basses	Enveloppée	Toutes	L-R	9-10	Env
Percus L	Enveloppée	Toutes	L	9	Env
Percus R	Enveloppée	Toutes	R	10	Env
Poums	Enveloppée	Toutes	C	9-10	Env
Tb a L	Enveloppée	Toutes	L	15	
Tb a R	Enveloppée	Toutes	SL - R	16	
Tb b L	Enveloppée	Toutes	R	17	
Tb b R	Enveloppée	Toutes	SR - L	18	
Batterie L	Toutes	Acoustique	L	1	Acou
Batterie R	Toutes	Acoustique	R	2	Acou
Contrebasse	Toutes	Acoustique	C	1-2	Acou
Gt a 4V	Toutes	Acoustique	L	11	
Gt a 87	Toutes	Acoustique	R	12	
Gt b 4V	Toutes	Acoustique	SR-L	13	
Gt b 87	Toutes	Acoustique	SL-R	14	
Gt intro 4V	Toutes	Acoustique	L	1	Acou
Gt intro 87	Toutes	Acoustique	R	2	Acou
Basse Synth	Toutes	Electronique	C	3-4	ElecAv
Bcle 1	Toutes	Electronique	L	3	ElecAv
Bcle 2	Toutes	Electronique	R	4	ElecAv
Bcle 3	Toutes	Electronique	SL	5	ElecAr
Bcle 4	Toutes	Electronique	SR	6	ElecAr
Rhodes L	Toutes	Electronique	L	3	ElecAv
Rhodes R	Toutes	Electronique	R	4	ElecAv
Rhodes SL	Toutes	Electronique	SL	5	ElecAr
Rhodes SR	Toutes	Electronique	SR	6	ElecAr
Accordéon Solo	Toutes	Toutes	L-R	7-8	Solo
Lead	Toutes	Toutes	C	24	
Solo Tb L	Toutes	Toutes	L	7	Solo
Solo Tb R	Toutes	Toutes	R	8	Solo