

Université de Montréal

**The use of real-time interactive music systems
in music composition and performance**

**par
Zack Settel**

Faculté de musique

**Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de
Docteur en musique (D.Mus.)
option composition**

Juin 2001

© Zack Settel, 2001

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée :

**The use of real-time interactive music systems
in music composition and performance**

présentée par :

Zack Settel

A été évaluée par un jury composé des personnes suivantes :

Thèse acceptée le : _____

SUMMARY

This text presents the author's approach to music composition for live performance, and his application of real-time "interactive" computer-based techniques in the realization and execution of his compositions. The body of this text examines the author's work, discussing goals, strategies and methods, with a primary focus on technical, practical and aesthetic issues. The opera *L'enfant des glaces* along with several of the author's other compositions are drawn upon in the various discussions, including implementation algorithms, diagrams, and flow-charts, as well as excerpts from recordings and scores. Though the underlying techniques are presented in depth, the overall discussion remains focused on creating and performing music.

Nowadays one finds the term "interactivity" appearing quite routinely in program notes, interviews and other descriptions of music, in which technology has played a role in its production and/or execution. Live music is by definition, interactive; the interactivity exists on many levels between both the performer and the audience, as well as the performers themselves. The author addresses the subject of "interactivity" in his music from the point of view of a composer, and a performer of his own and other composer's works. The focus on interactivity in the author's work is applied to collaborative-creation, composing, rendering pieces in concert, and expressive (*virtuoso*) performance. Interactivity is a vital element in the creative process. It is a key source of inspiration in the author's work, sometimes dangerous, sometimes disappointing, but always alive. It can play a significant role in determining the character of the resulting music, and how it touches the performers and the audience. Incorporating methodologies and techniques to foster interactivity in the production and execution of his music is one of the author's greatest endeavors. The author's approach to composition often incorporates the use of technology (e.g. sound/music generation and processing) in several stages in the composing and execution of a work. Sometimes the technology is applied simply as a predetermined means to reach a particular end, while other times it is drawn upon as a catalyst in the creative process. The author's principal interest in the use of technology is mainly as a source of inspiration and discovery in the creative process. A major part the text is devoted to the definition and description of the various stages in the author's approach to composition and execution of the author's works where technology is employed in the creative process.

It is in the initial stages of creating music that questions and ideas stemming from the technology can serve as a source of "compositional curiosity", motivation or inspiration. Many of the author's

works, in their incipience, have been nourished by the curiosity accompanying novel techniques and their inherent musical potential. Often in the author's experience in composing music, the choice of musical material and its use ("compositional language" if you like) in a particular piece follow from some form of discovery (or rediscovery). The discovery can be on the level of sound, following from say, a synthesis technique; it can be on the level of behavior, following from some new control algorithm or event detection scheme; or it can be on the level of timbre (sound and behavior) via signal processing or via a particular coupling of a given synthesis/signal processing system to a given control system. This discovery element functions as a source of inspiration in the making of the piece, leading to the harmonic, contrapuntal, structural, or even conceptual organization of the piece. A long-term approach of the author has been to work with (often new) techniques that promote discovery. Instrument design is an essential part of the author's compositional approach. The author examines the relationship between instruments and the music written for them. Nine elaborated examples of instruments designed for, and used in particular compositions of the author are presented along with accompanying musical material (excerpts from the scores). They are all based on the author's model for instrument design featuring event detection, signal processing and sound generation.

The author discusses his approach to composition, and the underlying techniques, methods and aesthetics. As with other disciplines such as desktop publishing, important thresholds have been crossed that change the nature of the way that we create. The strong emphasis on the "instrument" throughout this text reflects the author's longtime preoccupation with composing with instruments for live performance. In the author's approach, the instrument is present in the composition process at its very inception—as a source of musical expression, wonderment, curiosity, discovery, and inspiration. The approach to electroacoustic instrument design presented in this text provides for the invention and reinvention of new instruments, which themselves can lead to new musical expressions, from which compositions eventually emerge—which in turn, inspire the invention of still newer instruments. This iterative process embodies composition and performance in a continuous cycle of creation and expression, where the underlying techniques serve simply as a vehicle for the imagination and as a source for the musical spirit.

RÉSUMÉ EN FRANÇAIS

Ce texte traite de l'approche compositionnelle de l'auteur, orientée sur la performance en direct, et des formes spécifiques d'application de techniques informatiques « interactives » en temps réel qu'il a développées pour la réalisation et l'interprétation de ses compositions. Le texte portera essentiellement sur le travail de l'auteur; les objectifs, stratégies et méthodes mises en œuvre y seront abordées et les questions d'ordre technique, pratique et esthétique feront l'objet d'une attention particulière. Cette discussion s'appuiera sur des exemples tirés notamment de l'opéra *L'enfant des glaces* mais également d'autres compositions de l'auteur – ces exemples comprennent des algorithmes d'implémentation, des diagrammes, des organigrammes, des extraits sonores et des partitions. Bien que les techniques utilisées sont discutées en profondeur, l'objet principal du texte demeure la création et la réalisation musicale.

De nos jours, le terme « interactivité » est fréquemment employé dans les notes de programme, entretiens et autres discussions des pratiques musicales où la technologie joue un rôle déterminant dans la production et/ou la réalisation. Plus souvent qu'autrement, il sert à décrire un processus impliquant des relations bidirectionnelles – habituellement plus complexes, par exemple, que l'exécution juste, du point de vue de l'intonation, d'une note de violon, qui est un exemple de relation bidirectionnelle où l'« interaction » consiste à la correction de la hauteur à l'aide du doigt et de l'oreille. Le processus implique souvent des musiciens qui jouent en direct avec un dispositif électronique géré par des logiciels qui « écoutent » et répondent au jeu des interprètes. Pour sa part, l'auteur a toujours privilégié la performance en direct, laquelle est interactive par définition, l'interactivité s'appliquant ici aussi bien aux rapports des interprètes entre eux qu'à ceux qui s'établissent entre les interprètes et le public. Dans sa musique, l'auteur aborde la question de l'« interactivité » tant du point de vue du compositeur que de celui de l'interprète – de sa musique et de celle des autres. La notion d'interactivité recouvre alors aussi bien la création collective que la composition, l'exécution en concert et la performance expressive (virtuose). Dans *Interactive Music Systems*, Robert Rowe décrit plusieurs formes d'interactions caractéristiques de la pratique compositionnelle de l'auteur. D'après la description de Rowe, l'interaction dans la musique de l'auteur serait « induite par la performance » (*performance-driven*), basée sur le « paradigme instrumental » et ferait une large place à des « méthodes de réponses » transformatrices, génératives et par mise en séquence. À partir de cette terminologie, l'auteur décrit la façon dont les concepts de Rowe s'appliquent à sa musique. Pour l'auteur, l'interactivité est au cœur même du processus créatif : source d'inspiration, elle comporte une part de risque et peut entraîner des déceptions, mais

elle confère à la musique son caractère vivant. Elle joue ainsi un rôle déterminant tant dans la musique qui en résulte que dans la réception par les interprètes et le public. Le développement de méthodes et de techniques favorisant l'interactivité est, pour l'auteur, un des objectifs les plus importants de sa démarche.

Un chapitre complet est consacré aux nombreuses techniques que l'auteur a explorées, développées et intégrées dans son travail au cours des quatre dernières années. Elles sont présentées ici de façon détaillée et seront reprises dans les chapitres ultérieurs. Ces techniques et leurs implémentations logicielles ont été développées sur des plates-formes courantes, tels que des ordinateurs personnels très performants permettant l'utilisation de logiciels « populaires » comme Max/MSP, ou encore des processeurs ultra-rapides pour les systèmes non interactifs (en temps différé). La propriété essentielle de ces deux types de systèmes est la brièveté de la période d'attente ou de feedback : le temps nécessaire pour que la production d'un geste ou la formalisation d'une idée soit entendue et évaluée.

La discussion proprement technique débute avec la détection d'événement (*event detection*), une technique d'analyse qui permet d'identifier certaines caractéristiques d'un signal audio (provenant notamment d'une source musicale). La détection d'événement produit des signaux de contrôle dérivés du jeu des interprètes (ou analogues à leur modes de jeu). Ces signaux de contrôle sont utilisés pour contrôler la production sonore et le traitement du signal – l'auteur applique rarement la technique de suivi de partition (*score following*) qui consiste à la synchronisation des événements produits par l'ordinateur avec le jeu instrumental en direct. Dans le travail de l'auteur, la détection d'événement constitue généralement la source principale de contrôle pour les procédés de traitements. Une section de ce chapitre est consacrée à la présentation de plusieurs techniques de détection (détection d'enveloppe, de passage à zéro (*zero crossing*), de spectre et de hauteur), suivie d'une section consacrée aux diverses techniques de génération de son développées par l'auteur. Les générateurs sonores sont souvent contrôlés dynamiquement par des signaux de contrôle, lesquels sont produits instantanément par l'analyse en temps réel du jeu de l'interprète (au moyen notamment de détecteurs d'enveloppe ou de hauteur). Les structures de contrôle sont ainsi conçues pour offrir un contrôle potentiel maximal du déclenchement et de l'évolution dynamique des sons générés. La section suivante traite des techniques de mise en boucle d'échantillon (*sample looping*), d'échantillonnage contrôlé par enveloppe (*envelope-gated sampling*) et de modélisation physique, et est suivie d'une présentation de techniques de traitement de signal développées par l'auteur et intégrées dans ses compositions. Dans bien des cas, notamment avec les techniques génératives

décrites ci-dessus, le contrôle dynamique des processus de traitement (dérivé du jeu de l'interprète) est l'aspect le plus important de l'implémentation; ce contrôle est déterminant pour la qualité et le caractère du résultat sonore. Les modules de traitement eux-mêmes sont plutôt standard; aussi, la discussion ne portera pas tant sur leur conception que sur leurs propriétés spécifiques de traitement et leurs structures de contrôle. Enfin, des techniques telles que l'harmonisation, la distorsion par réinjection, la transposition linéaire de fréquence (*frequency shifting*), la modulation en anneau et la réverbération seront présentées.

Les techniques de traitement de signal dans le domaine spectral constituent depuis longtemps un des principaux champs d'intérêt de l'auteur; aussi, une section complète leur est consacrée. Ces techniques regroupent à la fois la production sonore et le traitement du signal, et sont basées sur des approches fort différentes des techniques traditionnelles de traitement dans le domaine temporel. Cette section couvre les méthodes de l'auteur pour le traitement spectral du signal, une description de son environnement de développement, ainsi que diverses approches de traitement et de génération de son. Un exposé des principes de base de la convolution sera suivi d'une étude des diverses applications développées et implémentées par l'auteur, la plupart ayant été intégrées aux dispositifs de traitement de signal en temps réel employés dans ses compositions ainsi que pour la composition en différé de bandes sonores. Cette discussion traitera également des stratégies et techniques visant à améliorer le rendement sonore des techniques de synthèse croisée en s'attaquant notamment au problème du filtrage par convolution – bien que ce problème ne soit pas exclusif au filtrage par convolution mais concerne de façon générale la synthèse croisée – dont les bandes de filtres sont très pointues et très étroites. La notion de « compatibilité spectrale » est également abordée dans cette discussion de la synthèse croisée. Le résultat d'une opération de synthèse croisée (opération multiplicative où deux signaux sont multipliés dans le domaine spectral pour en produire un troisième) dépend fortement des similarités des deux sources, puisque que le spectre sortant contient uniquement des composantes communes aux deux signaux-sources. Cette section se termine par la présentation d'une version du vocodeur de phase de Miller Puckette, initialement développé pour la plate-forme ISPW en 1994, et transféré vers Max/MSP en 1999. Depuis, l'auteur a enrichi l'implémentation de Puckette en y intégrant une fonctionnalité d'étirement du bruit, une technique permettant de conserver les caractéristiques du contenu bruiteux du spectre, indépendamment du facteur d'étirement.

Le travail de l'auteur sur la spatialisation sonore est ensuite discuté à travers un exposé sur les techniques de spatialisation multipiste qu'il a développées. Cette approche permet de positionner

plusieurs sources sonores distinctes dans un champ de diffusion multi-voix (système de haut-parleurs) et de produire des mouvements dynamiques indépendants pour chaque source sonore. Cet exposé traitera d'abord de la spatialisation de base pour s'arrêter ensuite sur une application spécifique où le rythme et la spatialisation sont mis en relation directe. C'est en travaillant sur un système de spatialisation multi-voix que l'auteur a découvert, « par accident », une technique de « spatialisation rythmique » qu'il a depuis appliquée à plusieurs œuvres. Cette technique est destinée au travail sur des musiques à métriques très définies (*groove*). Le principe inhérent à cette technique consiste à prendre plusieurs sources sonores (voix) dont les rapports métriques sont explicites et à leur imprimer des mouvements indépendants dans l'espace et ce, en maintenant des rapports métriques explicites. Le déplacement peut être aléatoire pour chaque voix, ou asservi au déplacement d'une voix « maîtresse ».

Dans l'approche compositionnelle de l'auteur, la technologie – par le biais de la génération et du traitement de son/musique – joue un rôle à divers stades de la composition et de l'exécution de l'œuvre. Si parfois elle représente simplement un moyen prédéterminé pour atteindre une fin particulière, elle peut également être une sorte de catalyseur dans le processus créateur. Parfois, elle permet aussi à l'auteur de diffuser ses œuvres dans des contextes où les moyens technologiques sont fortement valorisés – indépendamment de critères tels que le contenu musical – par exemple, lors de l'International Computer Music Conference (ICMC) ou de l'International Symposium on Electronic Arts (ISEA). Quoi qu'il en soit, l'intérêt premier de l'auteur pour la technologie demeure le fait qu'elle constitue une source d'inspiration et de découverte au sein même du processus créateur. Dans cette perspective, la technologie peut être employée pour ses capacités analytiques d'exploration du son : elle permet alors de percevoir davantage (de découvrir), de la même façon qu'un microscope nous permet de discerner des détails habituellement imperceptibles, ou encore de la même façon qu'une caméra peut nous permettre de décomposer le mouvement en une série d'images, nous permettant ainsi de comprendre ce qui, autrement, serait trop fluide pour être saisi. La technologie peut être appliquée également à la génération de son et de musique. Le principe à la base de la représentation sonore d'interfaces (*auditory display*) consiste à rendre audibles des structures « cachées » par mappage des données structurelles aux paramètres de contrôle de la génération du son. Quelle qu'en soit son application, la technologie, lorsqu'elle est approchée sous l'angle de ses possibilités créatrices, devient un élément important de l'environnement de création. Une partie substantielle de ce texte est consacrée à la définition et la description des diverses étapes de l'approche de l'auteur à l'égard de la composition et de la performance de ses œuvres où la technologie contribue au processus créateur.

C'est dans cette mesure que des questions et des problèmes soulevés par l'utilisation de la technologie peuvent s'allier à l'impulsion créatrice et stimuler la « curiosité compositionnelle ». Ainsi par exemple, l'apparition des unités de délai de signal a permis l'exploration de nouvelles possibilités polyphoniques pour les instruments monophoniques tels que la clarinette ou le saxophone, ouvrant ainsi la voie à de nouveaux modes d'expression pour les interprètes et les compositeurs. L'« idée » d'écrire une pièce polyphonique pour flûte est directement liée à l'émergence de cette nouvelle possibilité technologique. *Marimba Phase* de Steve Reich fut sans aucun doute inspiré de ses premières expériences avec les techniques de mise en boucle de bande magnétique. Cette forme de curiosité est à l'origine de plusieurs œuvres de l'auteur, l'incitant à explorer de nouvelles techniques et les possibilités musicales qui leur sont inhérentes. Une fois qu'un objectif compositionnel est bien défini, qu'une idée ou un concept est établi, la tâche compositionnelle demeure, par laquelle la structure et le contenu sont réalisés et développés. Dans le travail de l'auteur, le choix des matériaux et leur utilisation, le « langage musical » pourrait-on dire, découlent souvent de découvertes – dans certains cas, il peut également s'agir de choses qui sont découvertes à nouveau. Ces découvertes peuvent concerner le son lui-même, par exemple l'emploi d'une technique de synthèse; elles peuvent s'appliquer à des comportements liés à la mise en œuvre d'un nouvel algorithme de contrôle ou de détection d'événement; elles peuvent aussi s'appliquer directement au timbre à travers l'élaboration d'un dispositif spécifique de traitement de signal ou d'une configuration particulière d'un système de contrôle du traitement de signal. Cet élément de découverte est fondamental pour la réalisation de l'œuvre, sur le plan de son organisation harmonique aussi bien que contrapuntique, structurelle et conceptuelle. Bien plus qu'une simple étape secondaire liée à la réalisation technique d'une œuvre, le travail d'interaction avec les techniques électroacoustiques informatiques décrites précédemment s'est avéré essentiel dans le processus compositionnel de l'auteur. Dans le présent exposé, plusieurs exemples tirés de compositions récentes de l'auteur viendront illustrer les rapports entre les méthodes utilisées et la réalisation musicale d'idées qui trouvent leur origine dans son approche compositionnelle.

La conception d'instrument est un autre aspect important de l'approche compositionnelle de l'auteur. Mais plutôt que s'en tenir à une présentation générale de ce domaine, il s'agira ici de décrire l'approche spécifique que l'auteur a développée quant à la conception et l'utilisation de nouveaux instruments destinés à la composition et la performance, c'est-à-dire aux rapports entre les instruments et les œuvres composées spécifiquement pour ces instruments. À cet égard, le piano offre un exemple intéressant: son apparition est liée aux travaux de perfectionnement du clavecin

qui ont permis l'introduction du concept de gamme dynamique indépendante pour chaque note d'un clavier. Le développement du répertoire a connu son essor véritable une fois que l'instrument a fait l'objet d'une appropriation par les interprètes et les compositeurs. Bref, par sa sonorité et sa nature, le piano a incité les interprètes et les compositeurs à découvrir de nouvelles possibilités musicales, lesquelles se sont progressivement constituées en un genre: la « musique de piano » (que l'on associe par exemple à Mozart, Chopin ou Debussy). À la rigueur, on pourrait dire que l'instrument recelait un potentiel musical particulier, notamment en raison de sa conception technologique, potentiel qui attendait en quelque sorte d'être « découvert ». L'introduction du saxophone et de la guitare électrique ont donné lieu à des phénomènes semblables d'apparition de musiques liées directement à ces instruments ; par exemple, le jeu et la musique de Charlie Parker et Jimi Hendrix ne peuvent être dissociés des caractéristiques propres à leur instrument. Aussi, afin d'illustrer ce rapport entre les instruments et la musique qui leur est liée, l'auteur présente neuf exemples d'instruments conçus spécifiquement pour des compositions particulières, accompagnés d'exemples musicaux tirés des partitions. Ces instruments ont tous été élaborés à partir d'un modèle instrumental, développé par l'auteur, intégrant la détection d'événement, le traitement de signal et la génération du son en direct. Bien que ces instruments présentent des ressemblances entre eux, du point de vue des objectifs de jeu, telles que les possibilités de l'accompagnement par soi-même, ou encore l'extension des possibilités timbrales, chaque exemple illustre une forme spécifique d'implémentation. Enfin, l'auteur propose sa propre évaluation de son travail en s'efforçant de dégager autant les réussites que les échecs qui ont jalonné sa pratique de la composition et de la performance. Cette évaluation s'appuie principalement sur des exemples tirés de son opéra *L'enfant des glaces*, qui rassemble les préoccupations essentielles de l'auteur et constitue ainsi un exemple privilégié pour la réflexion sur son approche compositionnelle et son utilisation de la lutherie électronique en direct.

Dans ce texte, l'auteur a présenté son approche compositionnelle et ses implications techniques, méthodologiques et esthétiques. À l'instar d'autres disciplines telles que la publication assistée par ordinateur, des étapes importantes ont été franchies, qui transforment la nature de notre travail de création. Tel fut le cas pour l'auteur, dans le cadre de sa propre pratique. Avec le dépassement progressif des limites imposées par la technique, les problèmes liés au traitement de signal et aux techniques de contrôle s'amenuisent et l'attention se tourne davantage vers le potentiel de découverte, du point de vue de la création musicale proprement dite. L'attention accordée à « l'instrument », tout au long de ce texte, témoigne bien des préoccupations déterminantes dans la démarche compositionnelle axée sur la performance en direct chère à l'auteur, pour qui l'instrument

est présent tout au long du processus créateur, à la fois comme source d'expression, d'étonnement, de curiosité, de découverte et d'inspiration. Cette approche de l'instrument électroacoustique ouvre la voie à l'invention – et la « réinvention » – de nouveaux instruments, lesquels sont susceptibles d'inspirer de nouvelles formes expressives qui, à nouveau, peuvent conduire à l'invention de nouveaux instruments. Ce processus itératif intègre la composition et la performance en un cycle continu de création et d'expression où les moyens techniques ne sont que le support de l'imagination et fournissent en quelque sorte une impulsion à l'esprit musical.

TABLE OF CONTENTS

Summary	ii
Résumé en Français.....	iv
Table of contents.....	xi
Table of figures	xiii
Sound examples.....	xv
Abbreviations	xvi
Introduction	1
L'enfant des glaces.....	2
The Drama	3
<i>The sources.....</i>	<i>3</i>
<i>The synopsis</i>	<i>3</i>
The Form.....	4
The Libretto.....	5
Compositional Approach.....	8
<i>Background and context.....</i>	<i>8</i>
<i>The use of electroacoustic systems: compositional strategies.....</i>	<i>8</i>
The hybrid instrument	9
Writing for the hybrid instrument.....	9
Auto-accompaniment using vocal self-harmonization.....	10
Extending the voice's timbral range	12
L'enfant des glaces: wide-range voice extension instrument	13
The Accompaniment	17
The machine: rhythm and harmony	17
Integration with melodic material	18
Auto-evaluation	19
<i>Pre-composition</i>	<i>19</i>
Using electroacoustic instruments.....	19
<i>Score design</i>	<i>20</i>
<i>Performance logistics.....</i>	<i>20</i>
Rehearsal considerations	20
Synchronization: performers and electronics	21
Techniques.....	22
Event detection	23
<i>Envelope tracking</i>	<i>24</i>
<i>Zero-cross tracking.....</i>	<i>26</i>
<i>Spectrum tracking.....</i>	<i>27</i>
<i>Pitch tracking.....</i>	<i>27</i>
<i>The detection interface</i>	<i>28</i>
Sound generation techniques	28
<i>Groove looping.....</i>	<i>28</i>
Simple looping.....	29
Looping with greater complexity	31
<i>Gated sampling.....</i>	<i>34</i>
<i>Physical modeling: "Strang".....</i>	<i>37</i>

Signal processing techniques	39
<i>Harmonizer/delay</i>	40
Multi-channel harmonizer	40
Constant (absolute pitch) harmonizer.....	40
<i>Feedback distortion</i>	40
<i>Frequency shifting and ring modulation</i>	42
<i>Reverb</i>	42
<i>Signal routing</i>	43
Frequency domain techniques for sound generation and processing	45
<i>Some background</i>	45
<i>The basic plumbing</i>	46
<i>Windowing</i>	47
Bin-specific table lookup.....	48
<i>Algorithms and basic operations</i>	49
<i>Flavors of convolution</i>	49
<i>Frequency-domain DSP applications</i>	52
Filtering	52
Frequency-dependent spatialization	53
Band-limited amplitude-dependent dynamics processor.....	53
Band-limited frequency-dependent dynamics processor.....	55
Cross synthesis	56
Improvements to cross synthesis: spectral compatibility	58
Spectral smoothing for input with dissimilar spectra.....	59
Spectral energy scaling for input with dissimilar spectra	60
Controlling the degree of spectral intersection	62
Spectral energy scaling input signals with too-similar spectra	63
Phase Vocoder: improvements for noise stretching	64
<i>Audio-rate control of FFT-based processing</i>	66
Control limitations.....	66
Generating spectral processing functions (SPF).....	67
Generating complex spectral envelopes via spectrum generation.....	67
Generating complex forms using waveform generators.....	68
Applying the techniques	70
Dynamic control of processing parameters via input signal analysis.....	71
Multi-channel sound spatialization	74
<i>Sound source positioning and simulation</i>	74
<i>Rhythm-synchronized spatialization</i>	77
Conclusion	80
Bibliography	82
Discography	86
Appendixes	87
Libretto source texts and translations of Nerval and Quevedo	I
<i>French</i>	I
<i>Pular</i>	I
<i>Armenian</i>	I
<i>Finnish</i>	II
<i>Japanese</i>	II
<i>Persian</i>	II
<i>Russian</i>	III

TABLE OF FIGURES

Figure 1: Translation to suit the drama, L'enfant des glaces, section C2.3.....	6
Figure 2: Libretto to suit the electronic interaction, L'enfant des glaces, section B1.3	7
Figure 3: Translation to suit the song, L'enfant des glaces, section C5.3,C5.4	7
Figure 4: Electroacoustic instrument.....	9
Figure 5: Three-note fixed-transposition harmonizer instrument.....	10
Figure 6: L'enfant des glaces, section B1.....	11
Figure 7: L'enfant des glaces: resulting harmony in section B1	11
Figure 8: L'enfant des glaces: vocal timbre extension instrument.....	12
Figure 9: L'enfant des glaces, section B1.3.....	13
Figure 10: L'enfant des glaces, section B3.1.....	13
Figure 11: L'enfant des glaces: complex extended voice instrument	14
Figure 12: L'enfant des glaces: tenor solo for extended voice instrument, section A "L'homme 6" .	16
Figure 13: L'enfant des glaces: soprano solo for extended voice instrument, section C2.2	17
Figure 14: Harmonic content for accompaniment in L'enfant des glaces, section C2.2.....	18
Figure 15: Harmonic content for accompaniment in L'enfant des glaces, section B2.3.....	18
Figure 16: "the melody": the soprano's part from L'enfant des glaces, section B2.3.....	18
Figure 17: View of the basic amplitude follower	25
Figure 18: 10 millisecond noise waveform	26
Figure 19: Event tracking interface	28
Figure 20: Simple two-voice looping engine	30
Figure 21: Complex voice element (from two-voice looping engine) with quad output	33
Figure 22: Live instrument timbral extension via envelope follower and gated sampler	35
Figure 23: MSP implementation of a five-voice gated sampler.....	36
Figure 24: String model with parameters	38
Figure 25: Live performance instrument based on "strang" object.....	39
Figure 26: Feedback distortion unit.....	41
Figure 27: Standard reverberator interface from Jimmies DSP library.....	43
Figure 28: Signal routing and mixing implementation.....	44
Figure 29: Sample-by-sample output of the FFT object.....	47
Figure 30: Windowing function generator	47
Figure 31: Typical two-overlap windowing of the input and output signals.....	48
Figure 32: Using a lookup table to obtain bin-specific attenuation values in a graphic equalizer implementation.....	49
Figure 33: Simple convolution retaining the phases and amplitudes of each input source.....	50
Figure 34: Convolution of phase/amplitude (left) and amplitude-only (right) spectra, used in filtering implementations.....	50
Figure 35: Convolution of phase-only (left) and amplitude-only (right) spectra	51
Figure 36: Phase rotation, used in spatialization implementations	51
Figure 37: Filtering with a user-specified spectral envelope.....	52
Figure 38: Panning by frequency band.....	53
Figure 39: Gating low amplitude noise	54
Figure 40: Gating stable frequency components	56
Figure 41: Vocoding implementation.....	57

Figure 42: Vocoding with performer-analogous inflection and spectral emphasis.....	57
Figure 43: “Incompatible spectra” rendered compatible.....	59
Figure 44: Energy scaling to increase the degree of spectral intersection.....	60
Figure 45: Dynamics processing to boost weaker spectral components:	61
Figure 46: Dynamics processor implementation in MSP	61
Figure 47: The amplitude spectrum "intersection" parameter.....	63
Figure 48: Exponential combined energy from cross synthesis.....	64
Figure 49: A signal is filtered (via convolution) using a Spectral Processing Function (SPF).....	67
Figure 50: A complex spectral form depending on only three parameters	68
Figure 51: 1-8 Lookup table operations to generate a SPF	69
Figure 52: Generating a complex SPF.....	70
Figure 53: Applying SPFs for filtering and spectral panning.....	71
Figure 54: Self-modifying signal via the waveform generation technique	72
Figure 55: Self-modifying signal via the spectral generation technique.....	73
Figure 56: Interface for spatialization	75
Figure 57: Schematic overview of the spatializer	77
Figure 58: Beat-triggered panning via algorithm	78
Figure 59: Beat-triggered panning via Master/Slave model.....	78

SOUND EXAMPLES

Sound examples #1 through #23 are provided on CD #1 in Appendixes, p. 87. The sound example's number corresponds to the CD's track number.

sound example 1: L'enfant des glaces, section C2.3	6
sound example 2: L'enfant des glaces section C5.3-4	7
Sound example 3: L'enfant des glaces section B1	11
Sound example 4: L'enfant des glaces section B1.3	13
Sound example 5: L'enfant des glaces section A3 "L'homme 6"	16
Sound example 6: L'enfant des glaces section C2.2	17
sound example 7: L'enfant des glaces, section B2.3 (whispers)	18
Sound example 8: GO LoopEngine1	30
Sound example 9: L'enfant des glaces: groove examples	33
Sound example 10: Bash	34
Sound example 11: Machine1	34
Sound example 12: VoicePerc1	37
Sound example 13: L'enfant des glaces section A: la femme	54
Sound example 14: Pitch stripping	56
Sound example 15: Punjar section 1	58
Sound example 16: dynamics_example6	59
Sound example 17: dynamics_example7	62
Sound example 18: Stretched_whisper	66
Sound example 19: dynamics_example8	71
Sound example 20: dynamics_example9	71
Sound example 21: dynamics_example12	73
Sound example 22: L'enfant des glaces section B3.1	77
Sound example 23: L'enfant des glaces outtake1	79

ABBREVIATIONS

db:	Decibel
DSP:	digital signal processing
FFT:	fast Fourier transform
FM:	frequency modulation
Fq:	frequency
Hz:	Hertz
IFFT:	inverse fast Fourier transform
ISPW:	IRCAM signal processing workstation
MAX:	programming environment named after Max Mathews
MIDI:	musical instrument digital interface
ms:	milliseconds
MSP:	Max signal processing
rms:	root mean squared
SMPTE:	society of motion and television engineers
SPF:	spectral processing function
sqrt:	square root

INTRODUCTION

The use of real-time technology in music composition and live performance has long captured the imagination of composers. The application of analog electronics, and now, real-time interactive¹ digital computing technologies to music composition, performance, and music distribution has occasionally provided significant new opportunities for musical expression, around which an accompanying aesthetic has emerged. Over the years the author has drawn on interactive real-time technologies as source for new approaches to creating and performing music. Often, this has resulted in: 1) alternative ways to deal with conventional problems in creating music (e.g. form or variation in composition, or in performance), 2) the discovery of new modes of compositional expression, via unprecedented kinds of compositional material (e.g. cannons via digital delay) in quality or degree, or 3) new ways of hearing music. Specifically, these technologies have been applied by the author to the areas of: electroacoustic instrument design², precomposition³ (material generation), and live accompaniment rendering and performance mediation⁴.

This text presents the author's approach to music composition for live performance, and his application of real-time "interactive" computer-based techniques in the realization and execution of his compositions. The body of this text examines the author's work in the areas mentioned above, discussing goals, strategies and methods, with a primary focus on technical, practical and aesthetic issues. The author's recent opera *L'enfant des glaces*⁵ along with several of the author's other compositions are drawn upon in the various discussions, including implementation algorithms, diagrams and flow-charts, as well as excerpts from recordings and scores. Though the underlying techniques are presented in depth, the overall discussion remains focused on creating and performing music.

¹ a system, in which, the exchange of information determines the quality of subsequent information to be exchanged. In electroacoustic music, this implies real-time systems that respond to player's musical input.

² C. Lippe, "Music for piano and computer: A Description", 1997.

³ I. Xenakis, *Formalized Music*, 1971.

⁴ R. Rowe, *Interactive music systems: machine listening and composing*, 1993.

⁵ Electroacoustic opera composed by Zack Settel, performed by Pauline Vaillancourt (soprano) and Jean Maheux (tenor) and recorded in Montreal, September 2000. Recording provided with dissertation in Appendixes, p. 87.

L'ENFANT DES GLACES

L'enfant des glaces, premiered in September 2000, is a one-hour opera for two singers (tenor and soprano) and electronics (termed *electropera*⁶). It is the principal object of this dissertation. The drama for the work was inspired by a recent archeological discovery⁷ of Inca sacrificial remains, as well as by texts of Quevedo and Narval, which were used as source material for the construction of the libretto. The principal theme of the work is man's mysterious existence in time, space, and spirit. The general aesthetic setting underlying the drama, music and scenography—especially the visual aspects (e.g. image projection, décor, and lighting), tends towards the “extra worldly”: the infinite, ethereal, timeless, and mysterious. Electronic audiovisual techniques, with their “zero-mass” and highly malleable content, are used extensively throughout the work to support this aesthetic. There is also a secondary aesthetic setting which is in direct opposition to the one discussed above: it is one of “this world”: finiteness, mass, time, gravity and noise; it is reflected by the use of mechanical objects and décor (e.g. noisy machines, metal grids, etc.) in the set design and sound generation. Both settings create a dichotomy in which contrasting modes of existence are put into dramatic opposition. The libretto and the music were created by the composer, and were based on the dramatic synopsis for the work, created by Pauline Vaillancourt⁸.

⁶ the term denotes an operatic work combining the use of the voice and electroacoustic systems

⁷ described in the “Andes Expedition”, by explorer Johan Reinhard, 1997 (National Geographic, Nov. 1999)

⁸ soprano in the opera and artistic director of the opera production company Chants Libres, (Montréal)

THE DRAMA

The drama, or story behind the opera, is based on three principal sources. Ultimately, it took the form of a two-page synopsis, around which the structure of the opera was conceived.

The sources

In 1995, an archeological team led by anthropologist Johan Reinhard unearthed a remarkably well preserved (frozen) teen-aged victim of Inca Sacrifice, known as the Ice Maiden, in the Andes on top of Peru's Mount Ampato. In 1997 another expedition led by Reinhard returned there to discover additional burial sites and victims. Like the Ice Maiden, the victims were young, in their early teens, and even better preserved. These children were regarded as ambassadors of their people, beautifully dressed, decorated, and sacrificed five centuries ago at "the top of the world" to voyage beyond the earthly realm in time and space. The omniscient eternal visitor, or phantom, is an important theme in the work, from which the title of the opera, *L'enfant des glaces* (or "ice child") is based. He/she effortlessly passes through the centuries and the cultures, witnessing the cycles of life and death. He/she regards our present-day world with a considerable objective detachment (*recul*), yet is completely present in every moment.

Two texts, one by French novelist and poet Gérard De Nerval (1808–1855), and one by Spanish writer and poet Francisco de Quevedo (1580-1645), were drawn upon for their thematic content and general suitability (lyrical quality) for the libretto. Both texts embody a sense of timelessness, other-worldliness, and the cyclical nature of existence.

The synopsis

A two-page synopsis describing the dramatic action and general tone of the work provides the dramatic foundation for the opera. It tells the story of two individuals and their union. There is the woman (*la femme*), a spirit-like being who is (re)born into the world of today, while belonging to no world herself; and the man (*l'homme*), who is very much a product of, and totally immersed in the world of today. His world is dominated by the machine. During the course of the opera, both discover each other, form a couple, and finally break apart. In the first of the opera's three (approximately 20 minute) sections, *la femme*, like the Ice Child, emerges from a frozen crypt, arriving from some other time and culture, while *l'homme* is wholly immersed in the many activities of his rather material world, with which he is constantly (pre)occupied. Both characters remain separate for the entire section. During this section, *la femme* enters into the world of today and begins to acquire an earthly presence; and she becomes verbal. Meanwhile, *l'homme* goes about his everyday activities, frantic, obsessive, excessive and raucous, oblivious to her presence.

In the middle section, *la femme* continues her assimilation into the world of today, finally gaining a singing voice and the use of language. It is during this section that both *l'homme* and *la femme* become aware of each other. Both of their worlds are greatly rocked by the presence of each other, and the force of love which intensifies. By the end of the section, *la femme* has crossed over into the man's world (his side of the stage) and the two have formed a couple. They are both fully engaged in, and taken by the rhythm and currents of the world of today. He invites her to "ride the machine" together. The section ends with a love scene atop his machine.

The final section begins with the couple in blissful play. However, the machine gradually races out of control towards self-destruction, shattering the couple, and eventually destroying *l'homme*. *La femme* finds herself alone, serene and deeply moved. Meanwhile, the destroyed man crosses over into the world, from which, the woman emerged: the mysterious world of the Ice Child. As the opera ends, he enters *her* crypt, and passes into another world as the phantom, who makes an apparition at the close of the opera. Thus, *l'homme* and *la femme* are reunited in the world of the Ice Child.

THE FORM

The below structure for the opera was built around the synopsis. Each major section, A,B, and C, is approximately 20 minutes in duration. The work's total duration is approximately one hour and varies with each performance, depending on the singer's pacing.

Section A "The Emerging"

- A1 (4') Music of the Enfant.
- A2 (6') Woman's emergence
- A3 (10') Man's world

Section B "The Encounter"

- B1 (5'30") Man / Woman discovers the mirror
- B2 (4': 2'+1') Discovery of the other: Man and Woman
- B3 (10') Man and Woman meet and couple

Section C "Celebration and Transformation"

- C1 (5') Celebration of the couple
- C2 (2') Buildup
- C3 (3') Breakddown
- C4 (2') The Man crosses over, recontacts *The Source*
- C5 (3") The Woman integrated in the Man's world
- C6 (4') Final duet, the couple is reunited

Note: for simplicity and clarity, additional levels of subsections are not shown.

The principal objective in structuring the opera was to support the dramatic rendering of the story. Since there were no a priori structural constraints (e.g. fixed libretto, or strict dramatic scenario), the sections, subsections, and durations were chosen to provide for contrasting scenes and respites for the two singers. In section A, extensive dramatic juxtaposition of the singers' spaces and musical material is used to emphasize the contrasting worlds of the two characters, l'homme and la femme. As the work progresses, their worlds meet and merge into a common space and musical form. However, by the end of the opera, both characters are again spatially separated on the stage, but continue to share a common musical space.

THE LIBRETTO

The libretto was constructed by the composer. As mentioned above, source texts were taken from texts by Nerval, Quevedo, and Reinhard⁹. However, the text of the latter was employed only briefly in two subsections of section A; the other two authors' texts provided the basis for the rest of the libretto. Two paragraphs were chosen from each author and translated into 6 languages: Armenian, Finnish, Japanese, Persian, Pular and Russian. Underlying the choice of these different languages was a concern for the sonic (*musical*) qualities of the text. The languages were selected based on their relatively exotic nature, and their inherent and specific phonetic character. Thus, the libretto could be assembled from material which was largely unfamiliar to (most of) the public, and highly differentiated in terms of phonetic content or *color*. There libretto has no semantic function, only a musical one—to be experienced in the way one experiences a libretto sung in an unfamiliar language.

Thus, the “raw material¹⁰” for the libretto consisted of 12 paragraphs of highly differentiated texts. Since there was no concern for the semantic integrity of the texts, the composer took liberties in reordering the words of the source texts while constructing the phrases. However, relatively little mixing among the source translations occurs within phrases, or even groups of phrases. Rather, a particular source translation would be drawn upon to suit the dramatic setting in a given section. For example, in section C2.3 the couple is spinning out of control, the machine is whirling around faster and faster, clothes are tearing, metal is shearing—the world is going to hell. The use of the Russian

⁹ article written for *National Geographic Magazine* (Nov. 1999)

¹⁰ the complete source texts (with translations) of Nerval and Quevedo are provided in Appendix 4.

translation in the man’s libretto below was chosen to evoke a particular kind of desperation, drunken insanity, and abandon towards complete folly—Russian style.

The image shows a musical score for section C2.3 (42:32) Duo. It consists of three systems of music, each with a vocal line (PB or B) and a bass line (B). The lyrics are in Russian, with English translations provided below them. The score includes performance instructions such as 'crescendo poco a poco', 'ritardando', and 'f'.

System 1 (Measures 149-156):
 PB: [triplets] ta wi ko wi tye de ge de ge si ka ti ka mmm on = ki de ge de ge de ge si ka ti ka ti ka si ka di
 B: hurried, agitated, alarmed, exultant/boisterous becoming shudder
 f Ya yest péchínka!

System 2 (Measures 157-164):
 PB: (crescendo poco a poco) ge de ge de ge de ge dou dou dou dou dou dou tic tic
 B: crescendo poco a poco
 Ya yest péchínka! ve mirazhnie

System 3 (Measures 156-164):
 PB: dou dou dou dou dou dou dou dou dou dou dou dou dou dou ya yest pé chín ka de ge de ge de ge de ge di
 B: **8** (41:07)
 l'imp. mnisestvoch téhiskovsk

Figure 1: Translation to suit the drama, *L'enfant des glaces*, section C2.3

sound example 1: *L'enfant des glaces*, section C2.3

Earlier in the opera, Japanese translations are drawn upon for the man’s solo in section B1.3. The text is ideal for the electronics, which easily recognize and take advantage of the short percussive spoken text material and sustained voiced material (vowels). Below, the libretto is shown in service of the underlying music.

42 *plutôt calme*
dsp: gated physmod resonance w/metro

h. *mp* ko mi won no wo nai ko ko mi won no wo nai

47
h. ko mi won no wo ko mi won no wo nau

53
h. ko ko mi won no wo nai ko ko o ron ki

Figure 2: Libretto to suit the electronic interaction, *L'enfant des glaces*, section B1.3

Sound example 4: *L'enfant des glaces* section B1.3

In the final section of the opera, section C5.3, the action is winding down; *l'homme* has died, or more precisely, passed on to another world. *La femme* is left to sing a final aria (accompanied by *l'homme*, from the other world). The sheer beauty of the Persian text—especially its suitability for the softly repeated quazi-monotone phrases at the end of the section, made it a clear choice over the other translations. Here's an excerpt:

363 14 C5.3 (52:29) Chant continues for 1:50
etc.....

PB *molto rubato, sans vibrato*
en gâr ké za mine aze ma dâ rache_ va sar gar dâne dar â sé mân en gâr ké za mine aze ma dâ rache_

367 15 C5.4 (52:39) Final Duo
etc.....

PB (sans vibrato aux notes tenues)
va sar gar dâne dar â sé mân_ ra - - che va sar gardâne dar â sé mân_ en

B. moitié sur le souffle, sans vibrato
mp a - o - a - -

Figure 3: Translation to suit the song, *L'enfant des glaces*, section C5.3,C5.4

sound example 2: *L'enfant des glaces* section C5.3-4

COMPOSITIONAL APPROACH

Background and context

The music for *L'enfant des glaces* is a culmination of my developed approach to composition. The music borrows from many sources. Certainly I have borrowed from my former teacher Morton Subotnick, whose approach to live electronics (mentioned above) embraced the “instrument”, putting performers and technicians into sometimes-precarious performance situations. This was often to the delight of the aware public, where gesture and musical interpretation are brought into the musical foreground, and combined with the danger, excitement and occasional beauty and/or mystery of novel electroacoustic techniques. There is something quite compelling in the experience of live electronic music that denies the usual response to an encounter with something unexpected or unfathomable. When we see movies where gravity is defied, or somebody comes back from the dead, we typically imagine that “the trick” was done off-line, in some studio, with effects. But when the experience occurs live, there is wonderment. Nobody goes to the movies to see a magic act. Though Subotnick’s music goes well beyond the level of “tricks” or illusions, it is the “magic” in performance that I have been most impressed by. In *L'enfant des glaces*, the magic of the soloists playing live on their electroacoustic instruments was an important source of compositional inspiration, and central to the aesthetic setting.

The music of Jimi Hendrix is an excellent example of technique in the service of musical expression. The material played by Hendrix on his guitar (often with harmonic distortion, delay and wah-wah) cannot be effectively rendered otherwise. In this sense, the music played by Hendrix, and his instrument, are both one in the same: the instrument (as a source) embodies the music that flows from it. In my music, I have worked extensively with this kind of model, where marriages of musical material and “source” instruments provide the basic inspiration and material for the composition. The music in *L'enfant des glaces* resulted directly from this approach.

The use of electroacoustic systems: compositional strategies

Electroacoustic sound generation played a large role in the music for the opera, given the limited number of voices and absence of acoustic instruments in the score. The electronics were utilized in two capacities: 1) as support for “live” instrumental sections (i.e. in solos, or cadenzas), based on the “hybrid instrument” that combines electronic and acoustic sound sources; and 2) as accompaniment, either prerecorded (i.e. “tape style”), or generated live, to offer durational flexibility.

The hybrid instrument

The hybrid instruments I designed for the opera tended to be complex, typically involving acoustic instruments with electronic extensions for signal analysis, signal processing, and sound generation. The acoustic instrument is used as both a sound source and as a controller, providing control signals derived via event detection (see p. 23). Here's how it looks on paper:

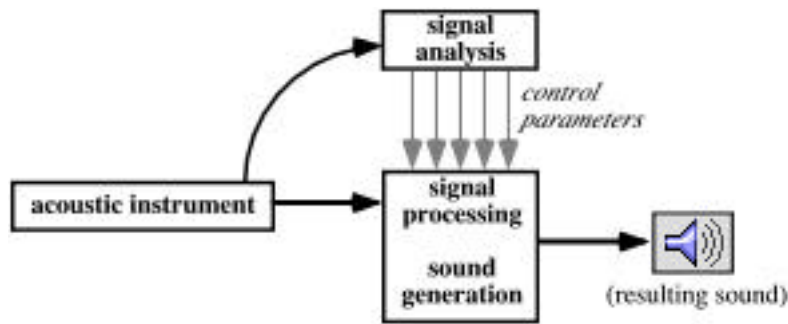


Figure 4: Electroacoustic instrument

Note: The use of acoustic instruments in the design provides a “built-in” degree of instrumental complexity and richness, and a high-resolution “musical interface” for controlling the instrument via event detection (see p. 23). It’s a good way to get around the limitations of (usually low and) finite-resolution controllers (i.e. all electronic controlling devices, and particularly those with 7-bit MIDI implementations). In this way, the control resolution can approach that of the audio (e.g. the envelope follower, that produces a 32-bit intensity parameter). Additionally, the acoustic instrument’s dynamic spectrum (or fundamental aspects of it) is often present in the instrument’s output, usually transformed. Thus, the player’s technique and initial expression, as manifest in the dynamic spectrum, are retained. An excellent simple example of this is singing into a ring modulator; while the output sound is completely transformed, all changes (no matter how fine) to the singer’s spectrum at the input of the ring modulator will produce corresponding degrees of change in the modulator’s output spectrum. From this point of view, the processing technique is highly transparent to the singer. Finally, the acoustic instrument’s spectrum can be combined with additional sound sources, whose behavior is analogous in one or more ways to the acoustic instrument’s behavior, and a hybrid sound results. For example, an FM generator’s frequencies can be mapped from the acoustic instrument’s detected pitch, while the FM generator’s dynamic spectrum is spectrally gated (as in vocoding) by the acoustic instrument’s spectrum. The resulting sound’s behavior is quite similar to that of the acoustic instrument. However the spectral content can be radically different (see pp. 56).

Writing for the hybrid instrument

The hybrid instrument described above offers the possibility of an expanded compositional range (i.e. potential for massive sound, self accompaniment, complex contrapuntal passages, etc.). This made it possible to compose for widely ranging musical formats such as, cannons, cadenzas, solos, “tutti” sections, etc. Some of the hybrid instruments used in the opera offer a rather high degree of polyphony, and contrapuntal complexity. Certain sections of the opera include music which is highly polyphonic, and quite dynamic—not at all characteristic of an *a cappella* two-voice duo. The examples below present the instruments in a compositional context.

Auto-accompaniment using vocal self-harmonization

This opera incorporates several methods for self-accompaniment, including one that allows for self-harmonization. The technique consists of a three-voice harmonizer that provides three simultaneous transpositions of the live voice. However, the transposition values are fixed, not relative (i.e. by interval, as is usually the case); this is accomplished by using the singer's instantaneous pitch (continuous output of a pitch tracker) to continuously (re)calculate the correct transposition for the harmonizer such that the result will always sound at a particular pitch. Another way of viewing this is:

$$\text{HarmTransp} = \text{currentPitch} - \text{targetPitch}$$

No matter what the singer's pitch is, or how it dynamically changes (e.g. *glissando*), the harmonizer will continuously force-transpose the singer's voice to the specified constant pitch. Using the three-voice harmonizer instrument shown below, three-note harmonic fields can be specified to accompany the voice, providing a stable harmonic structure. The instrument:

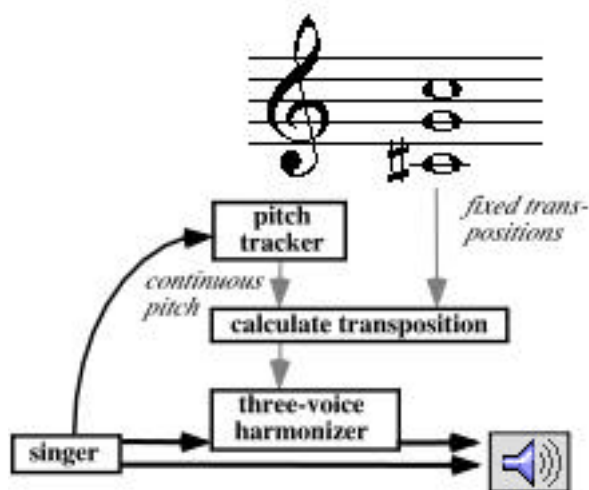


Figure 5: Three-note fixed-transposition harmonizer instrument

A key part of the instrument's functionality is its ability to capture the fixed transposition (pitch) values it uses "on the fly", according to the singer's input. Rather than storing these pitches in some sort of electronic score or cue list, they are defined as notes in the performer's written score, sung by the performer, and captured by the instrument for immediate use *in context*. The rule is: notes that precede silences are captured one at a time. Up to three notes can be utilized at a given moment, potentially forming a triad of transposed material. Thus, as music is played into the instrument, the

instrument will capture and accumulate up to three fixed-transposition values, which are used immediately by the instrument. In this sense, the writing (strategies) in sections featuring the use of this instrument can be likened to canonic writing, where previous material is combined with current material. However, in this case, there is no predefined second or third voice. Rather, the instrument, based on its design, processes the singer's input, and produces the additional voices. This instrument is used in section B1. The singer's part is below:

② (21:14) femme song B
 23 [dsp: hp harm series/no tape]
 f. mf na mi no dhè di'bou wo wa wi'ko mi na no wi ko
 30
 f. si wa bou dhè di ki'ta si wa wi ka'di ko na

Figure 6: *L'enfant des glaces*, section B1

Thus, when the music above is being performed (i.e. sung into the instrument), the result is that the first note, **c**, is taken and sustained (harmonized) with the material in the score through mm. 25. Diads sound. The rest in mm. 25 signals the instrument to capture the next pitch, **b**. The notes after that rest, in mm. 25 and 26, now sound as triads (two transpositions: **c** and **b**, plus the notes in the score). Then in mm. 26-27 an additional rest is encountered. The pitch that follows in mm. 27, **a**, is captured. The notes after that rest, in mm. 27 and 28, now sound as four-voice chord (three transpositions: **c** and **b**, and **a**, plus the notes sung in the score). When the rest in mm. 29 is encountered, the oldest transposition, **c**, will be discarded and the following pitch, **c**, will be captured and used in subsequent transpositions. The resulting harmony is shown below:

23
 electronics
 soprano
 27
 bass

Figure 7: *L'enfant des glaces*: resulting harmony in section B1
 Sound example 3: *L'enfant des glaces* section B1

Musically, this mode of operation resembles a canon: previous pitch material is combined with current material. When examining the passage above, one can see that any pitch that follows a rest will force the output of one of the three harmonizer voices to that pitch. The pitch will remain active in the musical context until it is replaced by another, three “rest-following” pitches later.

Extending the voice’s timbral range

An instrument using performer-based amplitude gating (see pp.34) was developed for use in the opera. The gated material (looped samples) is fed into a physical model unit generator (see Physical Modeling, p. 37). The reason for using an envelope-gated looping sampler is that the physical model unit generator is essentially a collection of potentially resonant band pass filters. If the singer’s signal were to be processed live, the open-air feedback risk would be major. To avoid that problem (making life much more comfortable in the concert hall tech pit and on stage) the player’s amplitude envelope is used to gate looped sample material, which in turn, is used as an excitation for the filter bank. The slightest utterances by the singer are translated into corresponding resonances. The instrument is structured as follows:

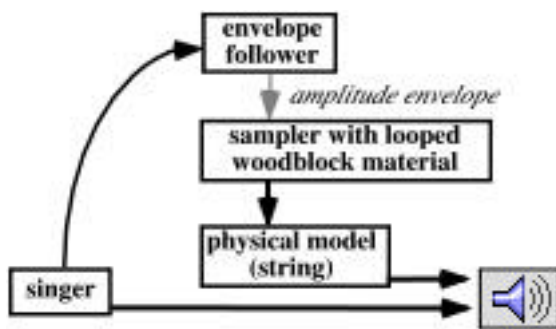


Figure 8: *L'enfant des glaces*: vocal timbre extension instrument

The material below was composed for the instrument. Note the similarity of the writing to what one might expect to see written for a resonant percussion instrument such as the celeste or vibraphone. Silences (rests) allow the resonant material, triggered by the singer’s sharp attacks, to decay before subsequent material is introduced (triggered). The occasional whole notes allow the triggered material to be sustained. They too, can be likened to tremelo material in writing for percussion instruments. Japanese was chosen for the libretto in this section (section B1.3) for its inherently staccato and percussive quality, which works very well with the instrument’s resonance properties.

42 *plutôt calme*
[dsp: gated physmod resonance w/metro]

h. *mp* ko mi won no wo nai ko ko mi won no wo nai

47

h. ko mi won no wo ko mi won no wo nau

53

h. ko ko mi won no wo nai ko ko o ron ki

Figure 9: *L'enfant des glaces*, section B1.3

Sound example 4: *L'enfant des glaces* section B1.3

Note the use of percussive spoken and sustained sung text; both are colored by the resonances of the physical model. The resonances were “tuned” to a specific harmonic context, shown below:

electronics

These resonances anticipate the following section, B3.1, where they are translated into the sung pitch material:

46 *mf* ki nou wa sa ri a shi ta wa ma da ko nai kyou to i e ba se wa shi na ku su gi te shi mau

8 (30.05)

Figure 10: *L'enfant des glaces*, section B3.1

L'enfant des glaces: wide-range voice extension instrument

The voice processing instruments used in the opera varied in terms of complexity and versatility. In most sections, the instruments performed specific musical functions, such as the ones described above for self-harmonizing or “tuned” resonances. However, in certain sections of the opera, the musical range was wider, and at the same time, less specific (not precisely notated). Though not improvised *per se*, these sections allowed for a fair amount of expressive interpretation or virtuosity. In the music of these sections, a certain disregard for the score is encouraged, based on

the idea of the virtuoso performer: i.e. that he or she, given the right musical situation (including the instrument), will do a *much* better job of rendering the music than the composer would on paper. This sort of freer reign makes certain demands on the instrument design: the instrument must be able to *simultaneously* support several different kinds of *modes de jeux* (i.e. quiet/loud, percussive/sustained material, pitched/unpitched, rapidly/slowly changing, etc.), responding immediately to the player's musical demands.

The wide-range voice instruments developed for the singers in the opera reflect years of development, revision and refinement. They combine several methods for self-accompaniment including those discussed above. The following instrument provides a good example of what's involved:

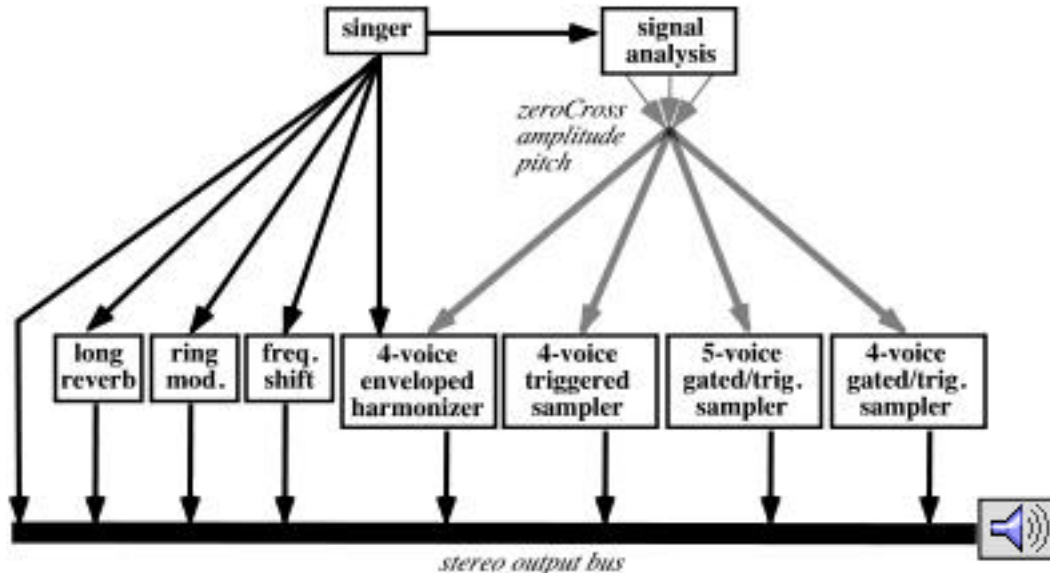


Figure 11: *L'enfant des glaces*: complex extended voice instrument

The basic design strategy is to dispose each of the component unit generators to respond independently of one another, and in a complementary fashion—otherwise said: to set them all up so they keep out of each other's hair¹¹. Thus, some delay the input, some sound only when the singer is singing (envelope-gated), others play independently of the singer, but can be muted by the vocalist singing: “ssshhhhhh”, etc. As shown above, the reverb, ring modulator and frequency shifter are not controlled dynamically.

¹¹ D. Rokeby, “The Construction of Experience: Interface as Content”, 1998.

However, the other unit generators are controlled dynamically and are described below.

- 4-voice enveloped harmonizer: each of its voices has an amplitude envelope generator (sharp attack, slow decay), which is triggered by the singer's sharp attacks. The envelope's duration is approximately 9 seconds. The transpositions are chosen each time at random, ranging from -18 to +6 semitones. The delays are fixed and range from 0.4 to 4 seconds.
- 4-voice triggered sampler, with widely ranging sample material, including percussion, machine sounds, sonar beeps, watch ticks and more. Notes are triggered by the singer's sharp attacks with an amplitude intensity corresponding to the singer's. The duration, ranging from 0.7 to 7 seconds, is controlled by an envelope generator (sharp attack, slow decay), whose values, along with the note's pitch and playback offset in the sample, are chosen for each note at random. The signal from the zerocrossing (noise) detector (see p. 26) is used to silence the sampler's output, and to alternate the sample material associated with each voice.
- 5-voice gated and triggered sampler, with various percussion sample material (wood, metal and skins). Notes are triggered by the singer's sharp attacks, and gated by his amplitude envelope.
- 4-voice gated and triggered sampler, with periodic machine sample material. Notes are triggered by the singer's sharp attacks with an amplitude intensity corresponding to the singer's. The pitch, ranging from -18 to +6 semitones, and playback offset in the sample, are chosen each time at random.

The following section of the opera for solo tenor voice (a cadenza of sorts) is unaccompanied. Using the instrument above, the singer follows the score below, but is encouraged to “do better”: i.e. to employ some virtuosity and “wing it” a bit in structured improvisation. The score:

129
kal fè kal fè bha do to ka ka dian dhou dhé

137
ka kambou wa si ta wi ko kal fé bhado to ka ko dé

143
(dé) ko dé ti na ti na shhhh!

Figure 12: *L'enfant des glaces*: tenor solo for extended voice instrument, section A “L’homme 6”

Sound example 5: *L'enfant des glaces* section A3 “L’homme 6”

The score consists of material (sharp attacks, swept intervals via grace notes, rests, *staccotti*, *crescendi*, sibilant text (i.e. “shhhh”), long notes, etc.) that will provoke a number of various and highly differentiated responses from the instrument. The instrument is also highly sensitive to temporal aspects of the singer’s *mode de jeux*, such as the length of sustained notes, and rests. Given the use of constrained random parameter generation, and the instrument’s extreme sensitivity to the singer’s input, the music is never the same from one performance to another. While it is not improvisation *per se*, the music in this section is still highly “interpretable”, adding to the piece an element of danger, excitement, and liveliness — a compositional goal, of sorts.

Later in the opera, there is a solo section (though accompanied by a playback soundtrack) for the soprano. Again, a wide-range voice extension instrument is employed in a solo/virtuoso context. The instrument is similar to the one shown above (on p.13), but more complex. The additional complexity involves the use of “gap-fill” reverb, which ebbs forward (increases in intensity) when the singer’s amplitude is low (when she is not singing); thus, its gain is inversely proportional to that of the singer. One of the gated sampler’s output gain is also inverted, so that its amplitude intensity is inversely proportional to that of the singer. These inversions were necessary to make “room” in the overall texture for different sonic materials that would otherwise clutter up the sound

tuned resonances (see Physical modeling: “Strang”, pp.37) in order to articulate a particular tonal/harmonic framework. For example, in section C2.2 the harmony, shown below, is “embedded” in the machine sounds, which provide the underlying harmonic accompaniment (both vertical/chordal and horizontal/melodic) for the voices.



Figure 14: Harmonic content for accompaniment in *L'enfant des glaces*, section C2.2

Sound example 6: *L'enfant des glaces* section C2.2

Integration with melodic material

In section B2.3, the soprano’s part is harmonically fused with the accompaniment (both shown below). The latter consists of processed whispering. As in the example above, the processing makes use of a physical model algorithm in order to bring out tuned resonances of a particular harmonic character. The resulting harmonic content of the accompaniment, and the singer’s part are shown below respectively. Both function harmonically, in the traditional sense of melody (the singer) and accompaniment (the tape).

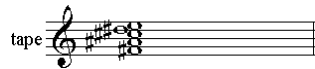


Figure 15: Harmonic content for accompaniment in *L'enfant des glaces*, section B2.3

5 *whispered, but very intense*

h. *pp* wa ta shi wa ka ko ni no ko sa re mi rai o machi gen zai ni ta ta zu un za ri su ru ko to da

♩ = 80 [manual cue]

8 **B2.3 (28:52) Femme Chant A**

sans vibrato

f. *mp* wa o ni wa o nai o ni no wa

17

f. o ni wa o nai ni no wa wa o

Figure 16: "the melody": the soprano’s part from *L'enfant des glaces*, section B2.3
 sound example 7: *L'enfant des glaces*, section B2.3 (whispers)

AUTO-EVALUATION

The following discussion, or auto-critique, examines successes and failures I have encountered in the composition and execution of *L'enfant des glaces*. It is a piece which embodies much of my work, reflecting my developed approach to composition and the use of live electronics.

Pre-composition

In general, the software tools I have developed for generating material for composition have proved to be quite useful. Unlike certain other approaches¹², I make no attempt to generalize my composition software (i.e. “tool box” approach) for reuse in subsequent pieces. Rather, the tools are programmed or recycled each time one is required for a particular piece. The MAX / Max Signal Processing¹³ programming environment provides enough programming abstraction¹⁴ to allow for the development of low-level object libraries that can be drawn upon from application to application; only the high-level functionality is reprogrammed each time. While this approach may not be totally efficient from a computer science point of view, it does keep things fresh and “non-systematic”; the composer is obliged to invent (or reinvent) the means of production for each piece, rather than to make choices based on preset functionalities. Thus, I am satisfied with the add-hoc approach to computer-assisted composition that I have adopted; it allows for modes of discovery and inspiration, while reducing the tendency to rely on pre-existing systems for creation.

Using electroacoustic instruments

The problem of instrumental depth¹⁵ in instrument design remains as a primary concern of mine. The electroacoustic instruments designed for the voices in the opera represent years of evolution and refinement, used in both composed and improvised music. As suggested earlier in this text, designing an instrument for improvisation is a great challenge; I am still far from having produced one that has the instrumental depth to “go the distance” in an extended (longwinded) free-improvisation situation. The opera *L'enfant des glaces* includes several instances of these instruments, which are designed for structured—in some cases, very loosely structured improvisation (see *L'enfant des glaces*: wide-range voice extension instrument, pp.13). Their timbral depth is sufficient for their intended use in the opera, their musical range is constrained variously by the score, and their potential for expression is considerable. Overall, I am pleased with my work in their design. However, these electroacoustic instruments are still too limited to be used

¹² e.g. IRCAM’s “Patchwork” composing environment, and K. Essel, “Real Time Composition Library”, author of composing software ©Ircam, 1992.

¹³ D. Zicarelli, creator of the Max Signal Processing (MSP) environment, *Cycling74*, 1997.

¹⁴ Template definition for functionality and data structures and libraries to containing them.

¹⁵ responsiveness, the degree to which the instrument can consistently render a sonic response specific to a particular instrumental manipulation.

“as-is” for extended free improvisation. In terms of instrumental depth, these instruments are much closer to the all-electronic “gadget” type, than to even the electro-mechanical sort. There is much room for improvement.

Note: The past 40 years has shown that our modern “all-electronic” instruments (i.e. instruments with no mechanical sound generation components) have not generated and sustained performance and usage activity nearly as well as their acoustic or electromechanical (e.g. electric guitar, Hammond tone-wheel organ, or Fender Rhodes electric piano) counterparts. Instead, these instruments tend to enjoy brief periods of widespread use and popularity (usually following from a hit tune in which one was featured), and then they disappear from current use. Some well-known examples of this are: the Moog synthesizer, which was highly popularized by Walter (now Wendy) Carlos’ use of it in the *Switched on Bach* album; or, the electric clavinet, popularized by Billy Preston and Stevie Wonder; or the Yamaha DX7¹⁶ and the Korg M1¹⁷, both of which were widely used by composers of film and TV music during the 80’s and early 90’s. These instruments have *already* become museum pieces! It is curious how acoustic instruments, and to a lesser extent, electromechanical instruments, tend to “escape” this particular type of fate, or at least take a much slower route to extinction.

Score design

The score for the opera (see Appendixes, pp.87) is for the most part, fully notated. However, in certain sections, this level of precision has proven to be impractical or unnecessary, and graphical notation, textual instructions, or notated “guide” material is used. In practice, this has worked well, although players can be reluctant to abandon the suggested “guide” material and forge out and improvise on their own. To discourage this tendency in future works I will likely provide an appendix section to the score that contains examples (“guide” material) for training the performers, while using textual and graphic instructions in the score for the structured improvisation parts.

The representation of the electronic part in the score was sufficient both in performance and rehearsal. For obvious reasons (indeterminate behavior, etc.), the score does not attempt to represent the output of the electroacoustic instruments beyond the occasional textual comment describing an aspect of their expected behavior.

Performance logistics

Rehearsal considerations

The specially designed electroacoustic instruments used in the piece generally require learning (playing experience); the score demands a degree of instrumental command—especially in the sections calling for *molto espressivo* interpretation, and even more so in the ones calling for structured improvisation. One cannot expect to rehearse the work in the standard way, where rehearsal time is allocated for learning the music; time must also be reserved for learning the instrument. The performers also must be willing to make the effort. Based on my experience with

¹⁶ FM-based synthesizer, introduced in the early 80’s

¹⁷ playback sampler/synthesizer, introduced in the early 90’s.

the opera (and other works), it is clear that the performers have to be trained on the electroacoustic instruments used in the work, otherwise they tend to avoid taking interpretative liberties, let alone veering off the notated (example) path as indicated when structured improvisation is required. The parts of the score for structured improvisation include notated “guide” material that catalogue, to some extent, the kinds of material that the instrument is most sensitive to (designed for). The solution I adopted was to schedule special sessions with each performer and a “coach” (someone who understands how the instrument behaves and who can help the player master it). Using the structured improvisation guides, the player is able to work with the instrument and the various material that it was designed for. After two or three one-hour sessions, the performer usually is familiar enough with the instrument’s behavior to be able to play it—not just operate it.

Synchronization: performers and electronics

Synchronization between the electronics and the performers was managed via manual cuing by a musical assistant¹⁸ during the performance—a “human score follower” of sorts. In parts of the opera, the players perform with/to a fixed-duration playback part, while in other parts, the electronic accompaniment is strictly live; in the latter case, the electronics are synchronized to the performers, and not the playback clock. The integration of both “synch-to-playback” and “synch-to-live” provided structure, flexibility and freedom, as required in certain sections for dramatic and musical purposes. Thus, the role of the musical assistant was an important one, since he alone was ultimately responsible for the coordination of the electronics (live and playback) and the performers, occasionally cuing the electronics during performance, or re-cuing them during rehearsal. In addition, the musical assistant was able to function as a coach during sectional rehearsals, helping the performers to understand and work with the electronics.

¹⁸ a person with musical and technical training, and (hopefully) great patience; sometimes employed “seasonally” by IRCAM’s production department, where I first encountered the term.

TECHNIQUES

This chapter is devoted to the various techniques I have researched, developed, and integrated in my work over the past four years. These techniques are presented here in detail, and are referred to in previous chapters. Unless otherwise noted, both the discussed techniques and underlying programming were developed by myself. While a thorough reading of this chapter is not essential to the development of the previous chapters, it is nonetheless suggested that the reader be acquainted with material presented¹⁹.

There are two important underlying concerns in the techniques supporting my work: cost/availability and “realtime” performance. Since a good deal of my music is written for performance with live electronics, I have favored the use of techniques and systems that are widely-available and relatively inexpensive, i.e. systems that I can provide, or ideally, that concert producers can easily provide. Thus the techniques discussed below have been developed on common platforms, such as high-performance personal computers running “popular” software, such as Max/MSP. In general, the systems I have utilized or developed, have been real-time performance systems, or in the case of non-causal systems (e.g. offline renderers), fast-working systems. The essential quality of both of these types of systems is the turn-around or feedback time: the time it takes to make a gesture, or implement an idea, and appreciate (hear) the result.

One of the most feedback-sensitive tasks that comes to mind is voicing or timbre processing. While John Chowning developed (and patented) his famous Frequency Modulation (FM) Synthesis approach, the “real” breakthrough in FM was made with the help of David Bristow, who collaborated with Chowning²⁰ in Yamaha’s development of the DX-7 synthesizer-Yamaha’s famous pioneering implementation of FM synthesis on a massively large-scale level. Bristow developed many of the initial FM voices in Japan, using a room-sized real-time prototype system, developed by the DX-7’s “founding” engineer Mr. H. Suzuki. The making of a voice (or a patch) for that system involves the configuration of over a thousand of parameters (exactly 1024). While FM theory explains how to produce a particular spectrum, the approach for producing good-sounding voices (i.e. voices whose time-varying spectra approach the perceivable complexity and

¹⁹ notes: 1) In the following discussion, the distinction between control events/signals and audio signals is made based on function: thus, a signal that is used to control a sound generator or processor is referred to as a control signal. 2) Frequency-domain-based techniques are discussed in a separate section, and not in the sections for sound generation or sound processing.

²⁰ D. Bristow & J. Chowning, coauthors of *FM Theory and Applications*, 1986.

richness of acoustic instruments) is more like alchemy: a mix of theory, practice, and good luck; it's because of this that the ability to modify parameters continuously in realtime is so indispensable. According to both Chowning and Bristow, and certainly based on my own experience, it's safe to say that in FM synthesis-based voice editing, the convergence on a finished voice (either preconceived or via serendipity) occurs "on-the-fly". Successive modifications to the voice's behavior are evaluated by the ear, while the voice creator compares the resulting sound with the sound he or she anticipated (predicted). As soon as the time interval between successive modifications to a voice becomes too great, the approach breaks down, and progress occurs at a snail's pace. This bears on the quality of the result, since voice creation often involves following a feeling or sense, implying a sort of fleeting or ephemeral "window of discovery". After minutes or hours of waiting, the "creator" can get lost. Several of Bristow's better voices were the result of serendipity via real-time voice editing.

While the above example showing the importance of real-time functionality in a music system is fairly simple, and perhaps self-evident, there are several other music systems, which I have used that are similarly reliant on real-time functionality. Since FM voicing and digital signal processing (DSP) are fairly similar in terms of parameter specification (i.e. voicing), The ability to edit DSP patches in realtime is crucial to me. The importance of real-time functionality in other music systems I have developed is less obvious, but nonetheless indispensable. Take composition (music generating) systems as an example. While I never used one of them in the generation of a large-scale form, they have been very useful in the rendering of compositional material (e.g. harmonic or rhythmic, on smaller time scales, say on the order of minutes. Real-time control of these systems can provide the composer with direct dynamic control over the process that is generating the musical material; the composer can pilot, influence or guide the production of a part, or aspects of the composition-in-progress in realtime. And of course, it facilitates discovery via serendipity, which continues to be an important source of compositional material and inspiration in my work.

EVENT DETECTION

Event detection is a technique, which is used to identify features in an audio signal (presumably from a musical source). The features can be physical, such as signal intensity, psychoacoustic, such as instrument pitch, or musical, as in the case of a "sharp attack". The features can be continuous sample streams (e.g. time-varying amplitude) or discrete events, as with the start of a note. When these features are recognized, the event detector outputs corresponding events. Event detection

provides a way to obtain control signals based on (or analogous to) the playing of live performers. I often use these signals to control sound generation and sound processing (though rarely for synchronization of the electronics to the live playing, e.g. score following). In my works, event detection, or tracking, often provides the main control sources for the electronics. The principal techniques for event detection used in my work are described below.

Envelope tracking

Even small latencies (10 ms or 20 ms) in an event detection system can be noticeable when, for example, the derived control signals are mapped on to sound generators (e.g. sluggish triggering of percussion sounds based on the player's sharp attacks). Worse, the latencies are also "felt" by the performer who triggers the sluggish action, thereby encumbering the player's performance. Thus, for control mappings to audio where there is a tight temporal coupling of performer and resulting audio, it is necessary to use low-latency event detectors. This rules out pitch followers, which by nature require at least a period and a half of captured waveform before they are comfortable about pronouncing their guess as to what the pitch is. This also rules out many frequency-domain analysis-based techniques²¹ since they need to collect signal (512 samples of it equals 11 ms at 44.1 kHz sampling rate) before analyzing and processing the data. But time-domain techniques such as amplitude envelope following require little, or no "state" (buffered samples). The envelope follower's control signal output responds rapidly and reliably to even the smallest changes in the input signal. The resulting control signal is potentially rich in nuance (dynamic amplitude resolution) and sensitivity, while being virtually instantaneous (extremely low-latency, i.e. 2 ms or less) in terms of responsiveness.

Below is a view of the envelope follower. For flexibility and economy, it has two outputs: one control rate, and one audio rate. The control output is used for subsequent analysis (feature recognition), while the audio output is most often used for direct gating purposes.

²¹ these techniques use the Fast Fourier Transform (FFT) to convert time-domain signals into frequency-domain (complex) signals.

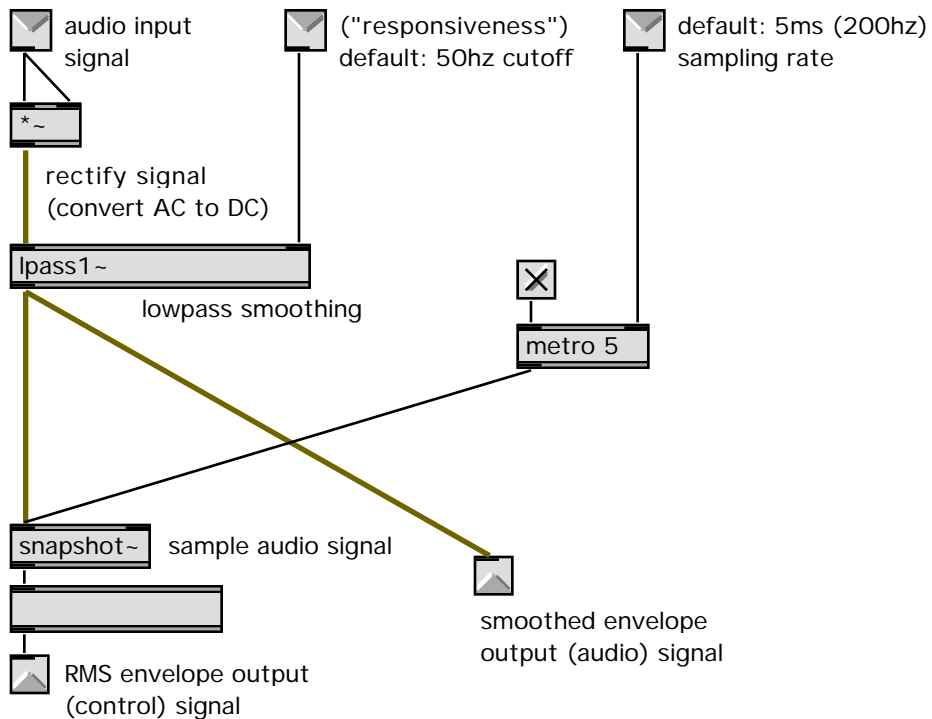


Figure 17: View of the basic amplitude follower

With the RMS (root mean squared) envelope control information, it is possible to recognize additional features in the audio input signal (performer's playing). The following features are recognized and produce corresponding control information:

- “silences” (rest event)
- large positive amplitude changes (spike event)
- RMS amplitude (amplitude value)
- smoothed audio envelope (envelope signal)

A rest (relative silence) can generate a very useful control event, often used for resetting a sound generator, or checking registers, etc. The ‘rest’ event is generated as soon as the value of the RMS envelope drops below a certain threshold.

The ‘spike’ event is triggered when the positive change of the RMS envelope, between successive samples (5 ms interval) is greater than a certain amount. Described another way:

$$a' = (a[t] - a[t-n]). \text{ If } a' > \text{thresh, then bang}$$

This provides for detection of attacks in the audio input signal (the performer's playing), or *subito* dynamic level increases (say, from *mp* to *ff*).

The 'amplitude' event provides an instantaneous amplitude value, which is often consulted when an attack is reported and the current amplitude value is needed to generate new parameters for a given sound generator or processor.

The 'envelope signal' is used for gating since it's more economical to do gating directly from the envelope follower's audio output than to do so using the RMS amplitude values. Plus the envelope signal is squared (see above, no square root), and sounds better to me when used as a gating function.

Zero-cross tracking

This technique is a very simple: it's based on a counter that is incremented each time the amplitude value of its input signal goes positive or negative (i.e. crosses the value 0). Below is a 10 ms plot of a noise sample showing how often a noisy signal crosses the value of 0.

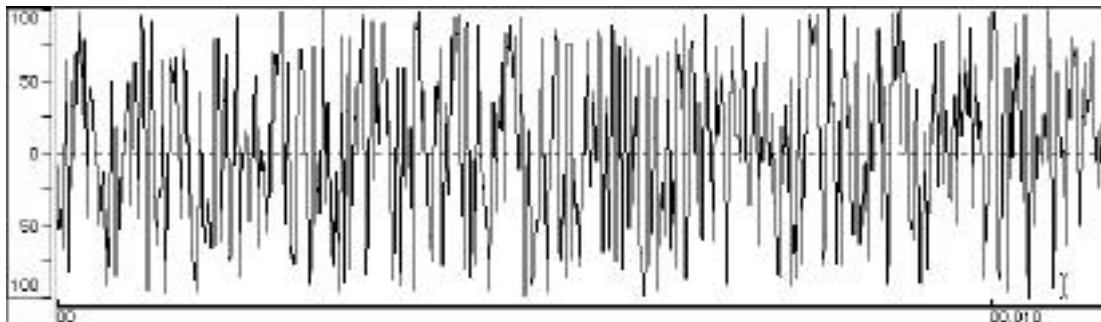


Figure 18: 10 millisecond noise waveform

The zero-crossing counter is examined and then reset to zero periodically (say every 10 ms). If at any time the examined value of the counter exceeds a certain threshold, a 'zero' event is triggered:

- zerocrossing density (zero event)

In practice, using the ‘zerocross’ object²², the ‘zero’ event signals a high zero crossing density in the input waveform. Since noisy signals tend to cross zero often, the ‘zero’ event can be a reliable detector of the presence of noise in the input signal. In speech, it makes for an excellent detector of the “v”, “f” or “z” sibilant (or hissing) quality of sound.

Spectrum tracking

Though an interesting and potentially useful technique for detecting changes of spectral brightness (the spectral energy’s center of mass), my work with the spectral centroid²³ has not yet been fully integrated in my compositions. The robustness of this technique is limited by the fact that the quality of brightness of an acoustic instrument is often a function of both register and loudness. Using a pitch follower in conjunction with the “centroid” object²⁴, it is possible to decouple the two, but the responsiveness of the brightness tracking suffers from the pitch-tracker’s latency, and its sometimes unstable (presence of inharmonic energy) output. The only control output by the centroid object is:

- spectral center of mass (frequency value)

Pitch tracking

I have developed a pitch following interface around the ‘pt’ object²⁵. The ‘pt’ object provides two kinds of pitch information: event-driven MIDI pitch events, and polled (every 7 ms.) instantaneous frequency. The control output is:

- note event (MIDI pitch value)
- instantaneous frequency (frequency value)

For the most part, the MIDI events are not used, since the envelope follower does a much faster and more stable job of recognizing note onsets (except for legato passages, in which case the MIDI events are relied on). Unlike the MIDI events, with 7-bit integer pitch resolution from 0 to 127, the

²² ISPW/MSP object written by Francois Dechelle, IRCAM, 1991.

²³ spectral center of mass as a measure of brightness; D. Wessel, D. Bristow & Z. Settel, “Control of Phrasing and Articulation in Synthesis”, 1987.

²⁴ MSP object written by Ted Apel of UCSD, 1998.

²⁵ pitch tracker written by Miller Puckette, IRCAM, 1990

resolution of polled frequency values is 32-bits. When a note onset is detected by the envelope follower and its pitch needs to be known, the instantaneous frequency value at that moment is used.

The detection interface

My event detection interface allows for the adjustment of the responsiveness (sensitivity) of the various event detectors to live performance situations for each performer.

Tracking settings

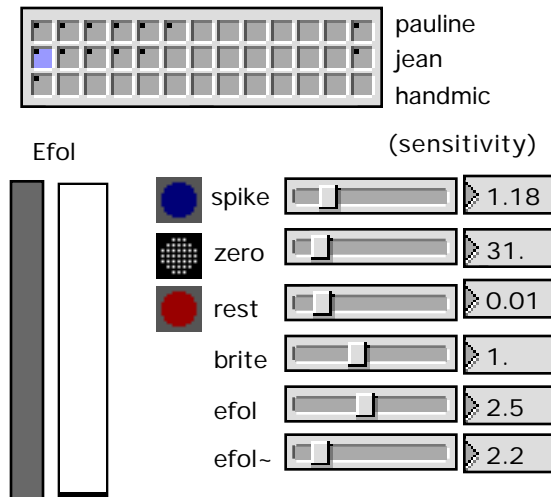


Figure 19: Event tracking interface

SOUND GENERATION TECHNIQUES

This section presents approaches to the generation of sound that I have developed for my compositions. Often, the sound generators are controlled dynamically by control signals instantaneously produced by live analysis of performer's playing (e.g. envelope or pitch followers). Thus, the sound generators' control structures are designed to provide a maximum amount of potential control (or "shaping") over the triggering and dynamic evolution of the generated sounds.

Groove looping

Looping of sampled audio has been used in a number of different ways. Xenakis on his UPIC system provided for looping of various "grains" (very short samples, on the order of a waveform) as

a way to generate complex waveforms of reasonable complexity²⁶. Based on a technique much documented by C. Roads, Cort Lippe²⁷ developed a granular synthesizer on the ISPW system (and now in Max/MSP) that allowed for continuous looping on a given sample stream; the loop size, and offset into the sample stream, and the rate of change of the offset play a large role in determining the quality of the output. In “techno” music, the loops tend to be larger, often involving periodic sound material, looped at rhythmic boundaries, on the order of beats or measures. Among these three examples, the main difference is the typical duration of the loop (i.e. the loop size). I have used looping techniques in a number of pieces. Two implementations are discussed below; the first, developed for the piece *GO* is relatively simple; while the second, developed for later work and particularly my opera, *L'enfant des glaces*, is a good bit more complex.

Simple looping

In the dance piece *GO*, it was necessary to develop a music that was periodic, though without a pronounced rhythmic groove. I decided to use a technique popularized by Steve Reich, consisting of pairing like or similar loops, and exploiting playback phasing between them. Since the source material in the loops is so important in determining the resulting character of the music, I chose to use (piano) samples from John Cage’s prepared piano sonatas.

In the figure below a simple two-voice looping engine is shown. Both voices are sample-synchronous. Each voice includes an independent ring modulator. The looping engine is controlled by the preset object (the large grid), which recalls previously edited parameter sets. The principal control parameters for each of the two voices are:

- loop speed (transposition)
- loop amplitude
- modulation frequency (randomly chosen each loop period)
- modulator output gain
- panning position (randomly chosen each loop period)
- source sample name

Note that in this implementation, there are no parameters for loop offset, or loop length; loops play from the beginning, and are as long as the soundfile they are sourced from.

²⁶ Sequential Circuit’s Prophet VS synthesizer [1985] was based on this idea.

²⁷ C. Lippe, "Real-time Granular Sampling Using the IRCAM Signal Processing Workstation", 1994.

At the heart of the looping engine voices is the “groove” object, which provides for sample looping. No amplitude windowing²⁸ of the loops is done. Instead, the samples used for looping are prepared to be “edgewise continuous” by closely matching slope and amplitude values at their beginning and end. The output of each loop voice is split: one copy is routed to a ring modulator, whose output is then mixed with the other (dry) copy and routed to the panning element. The amount of dry and modulated signal from the loop voice is controlled with attenuators. The loop voice also signals each time it loops, so that actions can occur on each loop period. A random selection of a new ring modulation frequency and panning position occurs each loop period.

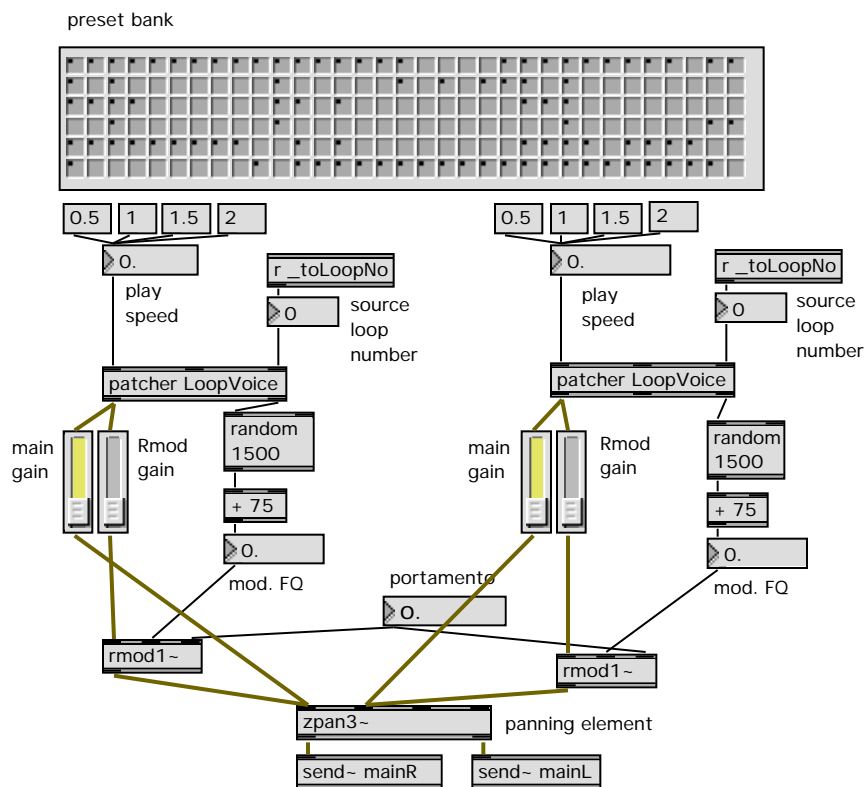


Figure 20: Simple two-voice looping engine

The following contains excerpts from the piece *GO*. Several different loop sources are featured.

Sound example 8: *GO LoopEngine1*

²⁸ attenuation of a signal segment’s beginning and end samples.

Looping with greater complexity

In my more recent work, I have made use of looping to produce music with pronounced rhythmic grooves, say in a quazi-techno style, as well as “industrial” music—music that draws on machine-like sound sources. In the preparation of the soundtrack for the opera, *L’enfant des glaces*, it was necessary for me to develop a technique that would allow for the convenient generation of both of these kinds of music. A new approach to looping material, as well as additions to the looping software were developed. The first steps were towards adapting the looper to deal more finely with rhythmic loop source material. Specifically, this meant, adding control over the loop’s playback offset index (read position in the source sample) as well as its duration, using a modulus read pointer; thus, if the duration is less than the length of the source sample, the read pointer will loop back to the specified offset index. With these additional parameters, offset index and duration, and the modulus read pointer, it is possible to define sub-regions of the source sample that can be continuously looped. In the following discussion, I will refer to these as “subloops”.

One important discovery I made using rhythmic source sample material, such as complete one or two-bar percussion grooves, was that using samples with “edgewise” continuity was not only unnecessary, but often produced results that did not sound as good as when using the raw samples, whose endpoints were discontinuous. This is because the clicks resulting from the discontinuities actually fuse with, and add to, the transient content at the start of the sample; they are not heard as clicks. As long as the subloops are chosen such that their offset index is located on a transient, one may loop with impunity, without concern for clicks. And of course, given the rhythmic nature of the source sample material, it is easy to predict where the transients will be located: on downbeats and upbeats as a rule²⁹. Thus, the offset index, and the duration parameters are specified in terms of units whose length is an integer subdivision of the source sample length: i.e. the units relate to beats within the source sample. For the offset index, the best choice for the unit ended up being 1/24, while for the duration parameter, a power-of-2 subdivision worked the best. Thus, the offset index is quantized to $(n * \text{sampleLength}) / 24$, while the duration param is quantized to the $\text{sampleLength}/n$, e.g. 1/1, 1/2, 1/4, 1/8, 1/16, 1/32, etc. Again, thanks to the modulo operator, choosing an offset index $23/24^{\text{th}}$ of the sample length, and a duration equal to the sample length (i.e. 1/1) would simply mean that the resulting loop would play until the end of the sample, and then

²⁹ the transients themselves are rarely located exactly on the expected rhythmic subdivisions, but are very often close enough to yield the expected transient spectral energy when indexed.

jump back to the beginning of the sample and continue. One requirement for working with rhythmic material in this way is that the source samples used in each of the voice elements be of the same length and related (compatible; i.e. whole number ratios) in meter and tempo.

Below is a view of a voice element of the two-voice looping engine and its additional parameters, used in the production of the soundtrack for *L'enfant des glaces*. Note that unlike the looping engine used in the work *GO*, the panning is now for a quad (4-channel) field, and can be triggered after N loops, abruptly relocating the looper's output to a new output channel.

Parameters:

- Offset index;
- Duration.

Panning at loop boundaries:

- Pan_trigger_after_N_loops;
- Pan interval (time it takes to complete the pan).

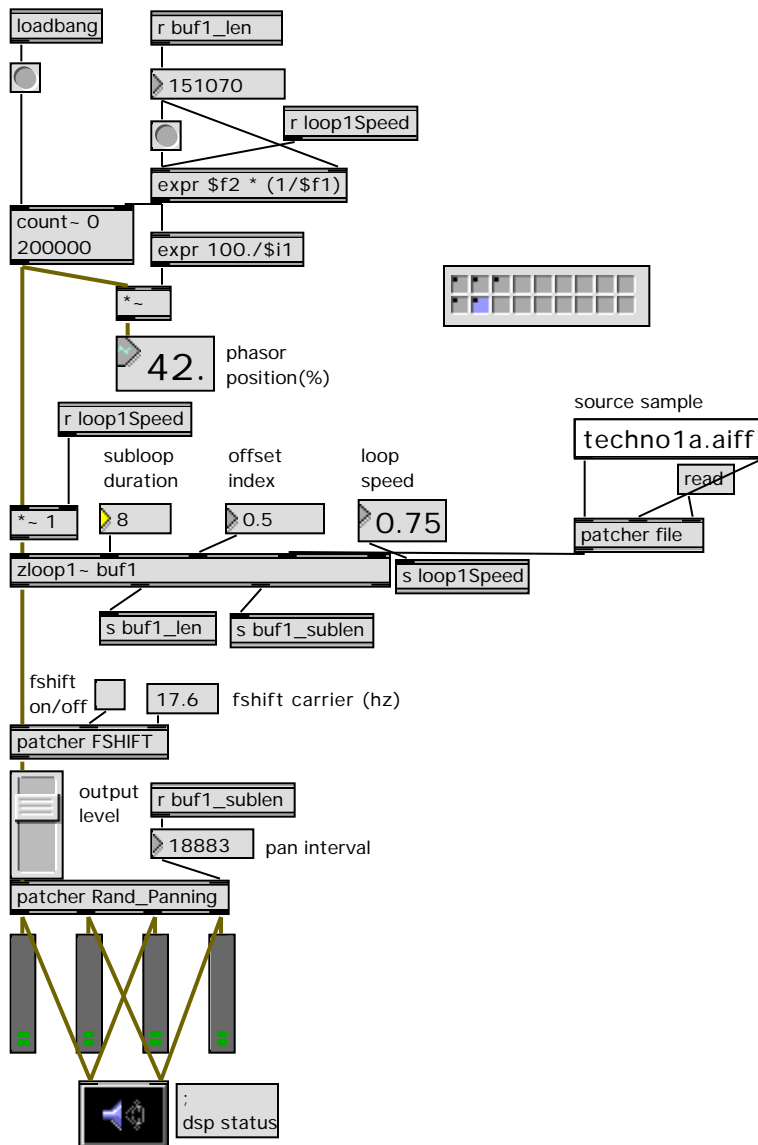


Figure 21: Complex voice element (from two-voice looping engine) with quad output

The looping element above, when used in the two-voice looping engine, was very useful in the generation of rhythmic (groove) material based on the rhythmic source material. The ability to pan rapidly on rhythmic boundaries (beats and bars) added an additional weight to the sense of the groove. I will discuss this further in the section on spatialization and beat-cued panning. Extensive use of the looping engine was made in the realization of the sound track for *L'enfant des glaces*. An excerpt is provided in the sound example below:

Sound example 9: *L'enfant des glaces*: groove examples

Another important discovery I made while using this looping technique was that industrial-style machine sounds could be easily generated when using noisy source sample material of a non-periodic nature, such as random percussion sounds of various timbres; below is a sound example of one of the sources.

Sound example 10: Bash

Machine sounds are strongly characterized by their periodicities or cycles; even noisy aperiodic material, when repeated, becomes regular, resembling the sound and character of industrial machinery. Using this kind of source sample material with the looping engine, I was able to develop a kind of “arbitrary machine sound generator”, which I used extensively in the production of the soundtrack for *L'enfant des glaces*. The best results were often obtained when using two different source samples, one in each voice element of the two-voice looping engine. The following “industrial machine” sound example was generated using the above source sample as well as another one, which contained random percussion sounds.

Sound example 11: Machine1

Gated sampling

The gated sampler and envelope follower are at the heart of an extremely effective real-time instrument processing technique, which I have used extensively in several pieces for acoustic instruments and live processing.

The basic design is illustrated below:

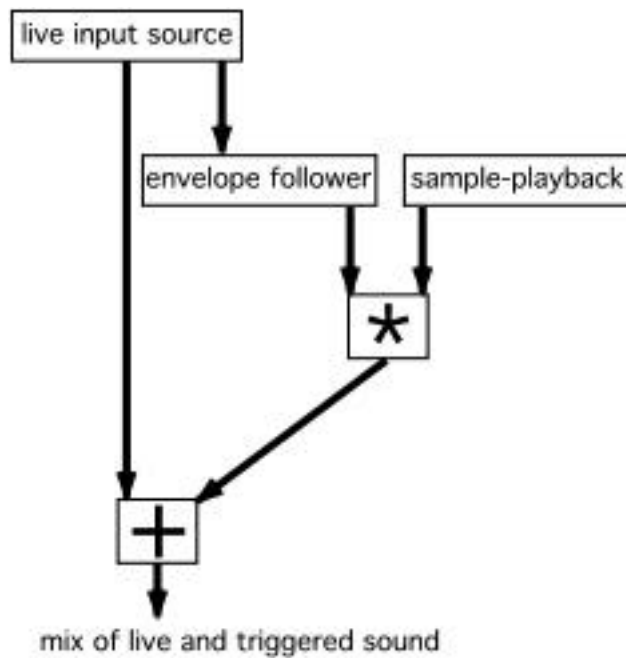


Figure 22: Live instrument timbral extension via envelope follower and gated sampler

Unlike other tracking techniques, envelope following is instantaneous; it requires no buffering. Thus, the control signal outputted by the envelope follower has no latency. When fed input from a live instrument, the envelope follower responds immediately to changes in the player's dynamic level, offering a very high degree of responsiveness and control over any sound generator linked to (i.e. gated by) the envelope follower. The gated sampler, when coupled to a live instrument via an envelope follower, becomes a potential source of timbral extension (additional timbre), which tends to fuse with the live instrument's timbre since both sound sources share the same (the instrument's) analogous amplitude envelope.

The following figure describes a gated sampler as implemented in Max/MSP:

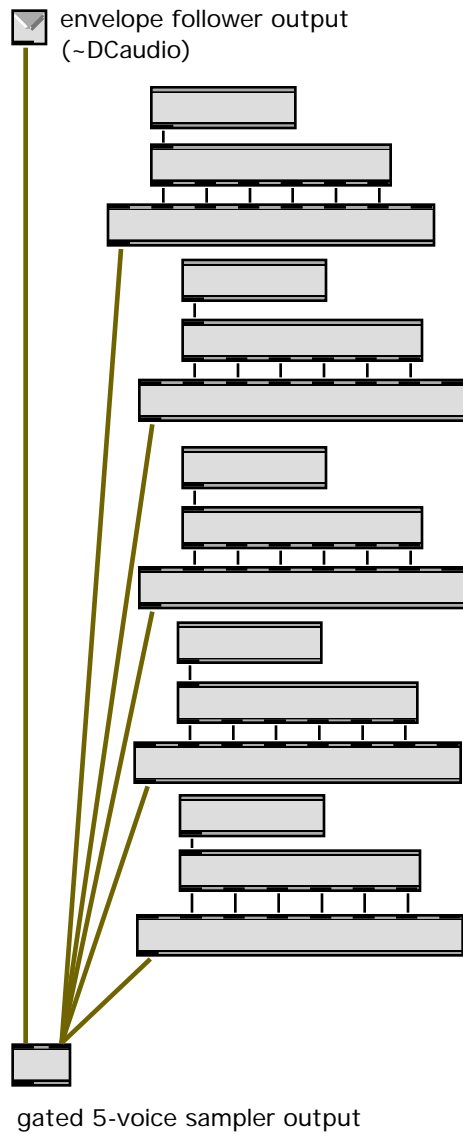


Figure 23: MSP implementation of a five-voice gated sampler

This technique of providing timbral extension for live instruments is surprisingly effective, and particularly useful in electroacoustic instrument design. The method consists of using source sample material that contains no silences, and allowing the material to loop continuously in the sampler. The looping source sample material can be thought of as an ever-present potential sound/timbre source that is elicited by the envelope follower (live instrument). As described above, the output level of the sampler is controlled by, *and analogous to*, the live instrument's dynamic level. Thus, the louder the playing, the more of the sampler's output is heard; no playing means no sampler

output. All changes to the player's dynamic level, no matter how abrupt, extreme or subtle, are applied instantly and continuously to the output level of the sampler, making sudden attacks and silences potentially quite dramatic.

The sound example below illustrates this tight coupling between acoustic (live) and electronic sound sources.

Sound example 12: VoicePercl

Physical modeling: “Strang”

A rich and surprisingly supple technique for real-time processing is based on physical modeling. A particular excitation (input) signal is fed into the “model” (think of it as a collection of resonators), which resonates according to its own characteristics and the quality (amplitude and spectral content) of the input signal. Physical models come in a variety of implementations, based on different approaches, ranging from the “wave-guide” models³⁰, developed by Julius Smith (Stanford), to the resonance-models³¹ developed by Adrian Freed (CNMAT) to the physics-based representations³² developed by Jean Marie Adrien and others (IRCAM). In all the approaches, an attempt to simulate the behavior of acoustic sound generators (such as musical instruments) is sought. What follows is a brief description of the model (Modalys) that I have integrated into my real-time processing environment:

“Modalys objects describe vibrating structures defined by their geometrical characteristics (e.g. strings, plates, or membranes) and the physical parameters of the (homogeneous) material they are made from. Based on this description, Modalys calculates the (intermediate) modal representation used in the synthesis process (frequency, damping factor, and mode shape for each mode of vibration).”³³

The “Strang” object, developed by Richard Dudas (IRCAM 1998) for Max/MSP, is based on IRCAM's Modalys software, which followed from the work of several researchers at IRCAM working over the years in the area of physical model-based synthesis. The object is based on a physical model for a string; the string can be put into motion (vibration) via excitation signals, which are fed to it via its audio input. The object (model) features several parameters, which can be

³⁰ J.O. Smith, “Physical modeling using digital waveguides”, 1992.

³¹ T. Jehan, A. Freed & R. Dudas, “Musical Applications of New Filter Extensions to Max/MSP”, 1999.

³² J.M. Adrien, “The Missing Link: Modal Synthesis”, 1991.

³³ F. Iovino, R. Caussé & R. Dudas, “Recent work around Modalys and Modal Synthesis”, 1987.

varied dynamically (though with some care when in extreme ranges). The reference point and the excitation point (points along the string's length where the string's sound is tapped, or where the excitation is applied) can also be specified. Given the wide range of physical description (potential timbre), the possibility to feed the object arbitrary excitation signals (ranging from clicks to continuous speech), and a robust dynamic parameter set, the Strang object makes an excellent signal generator/processor. And it offers the possibility to work within a harmonic context, due to the tunable resonances that it offers. Below is a view of the object and its parameters:

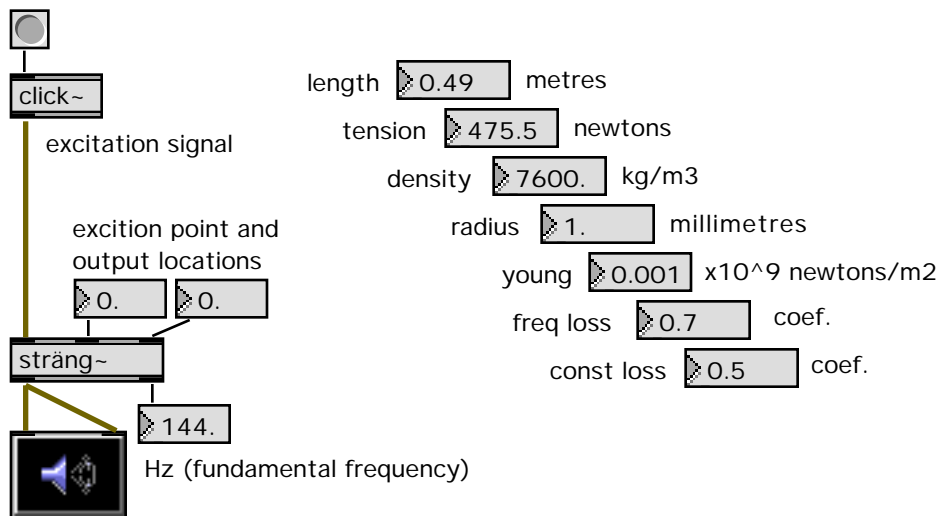


Figure 24: String model with parameters

The real-time DSP processing environment used in *L'enfant des glaces* incorporates the “strang” (string) object as a sound generator and as a sound-processing element. In both cases, care is taken to adjust the parameters such that the resulting resonances are structured to, and in tune with, the harmonic context of the composition. The timbres produced range in quality from wet to dry and from “metal / semi-inharmonic” to “gut / harmonic”. The timbres can vary continuously within those ranges.

When implemented as a sound generator, a click is fed into the model's audio input (as shown above), causing the model to resonate. When implemented as a processing element, the “plumbing” is more complicated: a signal based on the real-time audio input from the players is fed into the model's audio input. However, if one uses the player's direct audio input as an excitation signal for the model, an audio feedback loop can occur since the model's potentially highly resonant output (from the speakers and monitors in the performance space) can be picked up by the players' microphones and reenter the system. To overcome this problem, the model's audio (excitation)

input is provided by a looping sampler (using continuous sample source material), which is gated by the live player's dynamic amplitude envelope (see gated sampler description above). The resulting sound, due to the direct real-time amplitude control by the performer, follows directly from his playing: the louder the playing, the greater the resonance. Below is the processing element implementation:

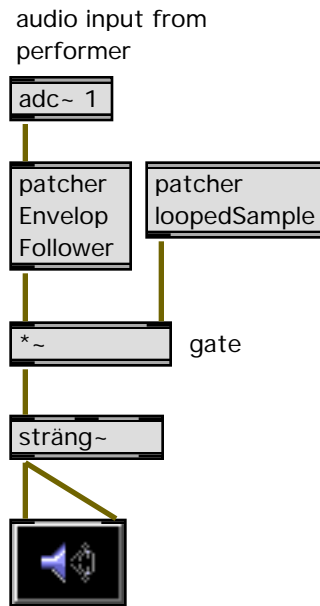


Figure 25: Live performance instrument based on "sträng" object

The following sound example, taken from the *L'enfant des glaces*, illustrates the high degree of performer control and responsiveness this processing technique offers when used with the voice.

Sound example 4: *L'enfant des glaces* section B1.3

SIGNAL PROCESSING TECHNIQUES

The following signal processing techniques have been developed and incorporated in my compositions. Quite often, as with the sound generation techniques discussed above, the real-time dynamic control of these processors, derived from the performer's playing, is the most important part of the implementation, and ultimately responsible for the quality or character of the output sound. Since the processing modules themselves are quite standard, the discussion will tend to focus on their essential processing qualities and control structures.

Harmonizer/delay

Multi-channel harmonizer

This processing technique consists of a four-voice harmonizer, whose voice element outputs are “panable” among the four output channels in the audio configuration. The voice element of the harmonizer is based on the fairly standard model of alternating-read variable delays. Its parameter set includes:

- transposition (in cents)
- delay (ms)
- feedback (scaler)
- output assignment (multichannel pan)

The multi-voice harmonizer allows for a given input signal to be transposed, delayed, and panned independently by each one of the four-harmonizer voices. It uses an envelope follower to detect note onsets.

Constant (absolute pitch) harmonizer

This harmonizer is built on the one above but takes fixed transposition values, rather than relative ones (intervals). It is intended for processing input signals with a definite pitch. With the addition of a pitch tracker, the harmonizer is constantly “aware” of an input signal’s pitch, and is thus able to calculate the necessary transposition in order to force the pitch of its output to the specified fixed transposition value. A full description of this technique is provided elsewhere in the text, among other “offbeat” custom DSP implementations (see p. 8).

Feedback distortion

Major feedback distortion can be produced, and finely controlled using the following technique, which consists of a boost unit, a clip unit, a pair of flangers with a feedback section, and a pair of resonant filters.

Here's how it looks:

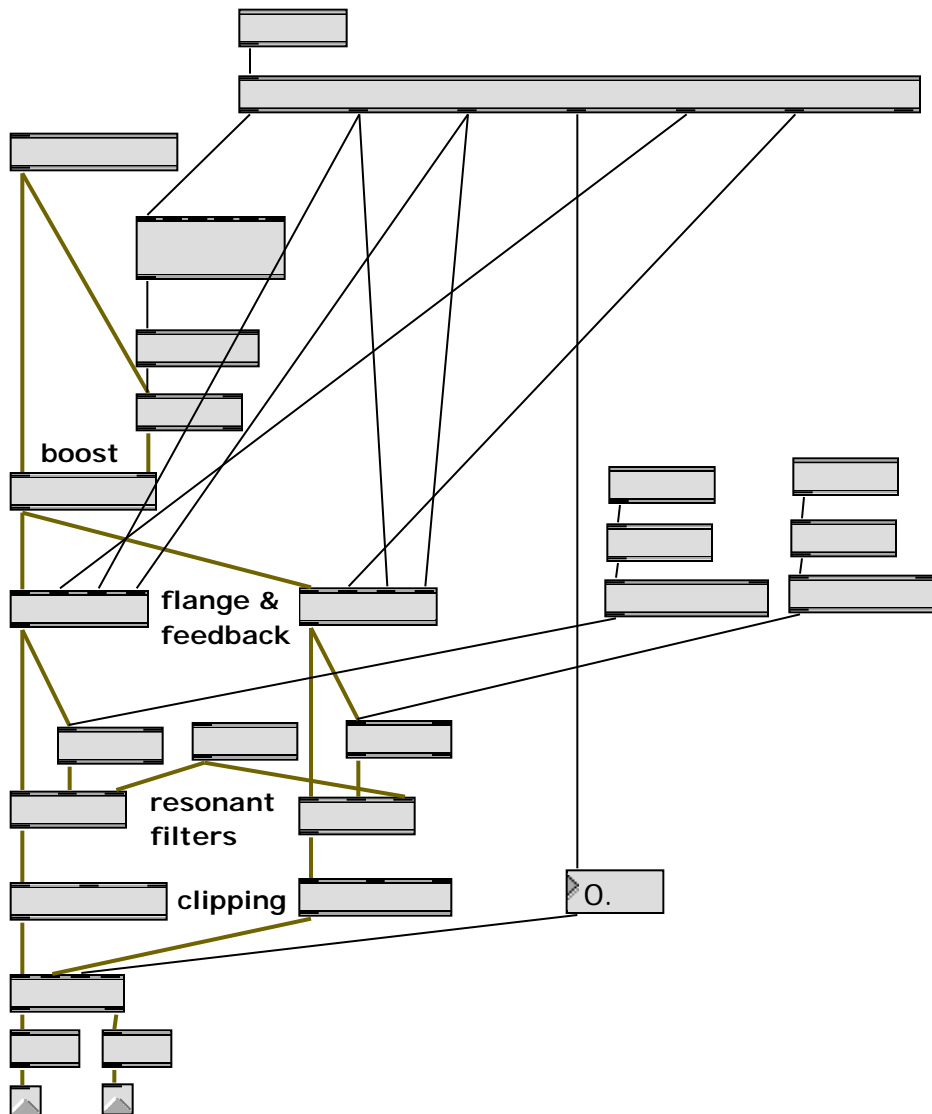


Figure 26: Feedback distortion unit

As you can probably see, the signal is first boosted, then flanged. Two independent flangers are used in order to provide emphasis and de-emphasis of harmonically unrelated partials of the input signal. The flangers include a feedback section, adding the possibility for closed-circuit feedback, outside of the open-air variety that involves microphones and speakers. Then, the boosted, flanged, and “feedbacked” signal is further boosted/cut using a pair of independent resonant filters. Finally, the whole resonant mess is clipped, using a very elementary clipping unit, thus providing distortion. By carefully choosing the right boost and feedback levels, and suitable resonance frequencies for

the flangers and resonant filters, major “harmonic” (tuned) feedback distortion can be produced—even at extremely low sound pressure levels (“conversation” level or less!).

Frequency shifting and ring modulation

I am assuming that the reader knows how these relatively basic DSP techniques work. Thus, I won’t describe their low-level implementation which is described in the Jimmies³⁴ DSP library. Rather, I will simply make the observation that both techniques are extremely interesting for the generation of secondary spectra, and better yet, they incur very little latency. This makes them ideal processing units for remapping spectral energy, especially when the input signal is essentially inharmonic (like drums or bells). In such a case, the use of these units makes a fine alternative for “transposition” to the harmonizer (whose latency is not only substantial but varying). One additional point: linear transposition of spectral energy is a very cool thing. It seriously alters harmonic spectra, while leaving inharmonic spectra relatively intact. Thus, when you send a tune such as *Giant Step*³⁵ through a ring modulator, you end up with an extremely interesting transformation of Coltrane’s saxophone, while Art Taylor’s drum kit passes through almost untouched. Wow.

Reverb

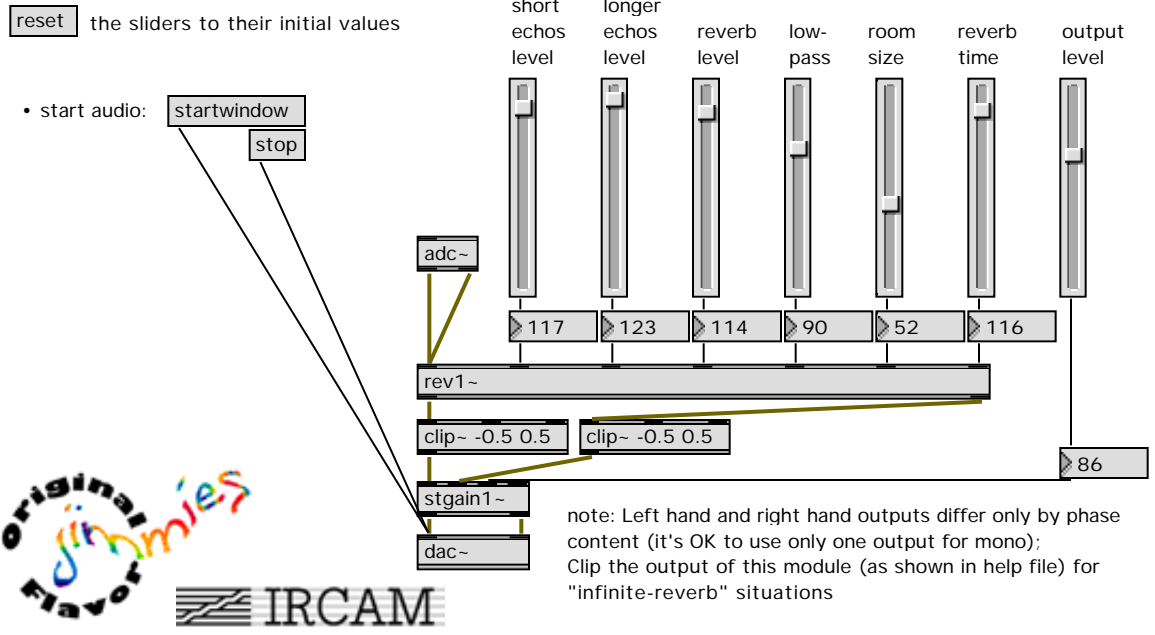
While reverberation is almost always included in the electronics for my works, only a short description of the technique is provided, since it is a very popular and well-known type of signal processing. Below is a view of the reverberator I use, based on an algorithm by Miller Puckette, and implemented for the Jimmies Library by the author. Note the six parameters (not including the output level) that can be dynamically controlled. For an example of dynamic parametric control of this reverberator, refer to the text following Figure 12 (p. 16) regarding “gap-fill” reverb.

³⁴ Jimmies signal processing library, IRCAM, Zack Settel, 1994.

³⁵ J. Coltrane, *Giant Steps* from album *Giant Steps*, 1960

rev1~ A simple reverberator.

note: all parameters normalized to < 0 - 127 >

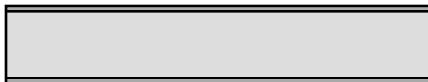


The "jimmies" by Zack Settel, © 1994-98 IRCAM.
 The jimmies were made possible in part by a brain cell donation from Miller Puckette and Jean-Marc Jot. Graphic design by Xoaz.

Figure 27: Standard reverberator interface from Jimmies DSP library

Signal routing

Signal routing and mixing are handled using abstractions in MSP to manage signal assignment, addition and multiplication (ramped); they receive messages from the MAX control environment. Below is an example abstraction of this functionality, coded in MSP:



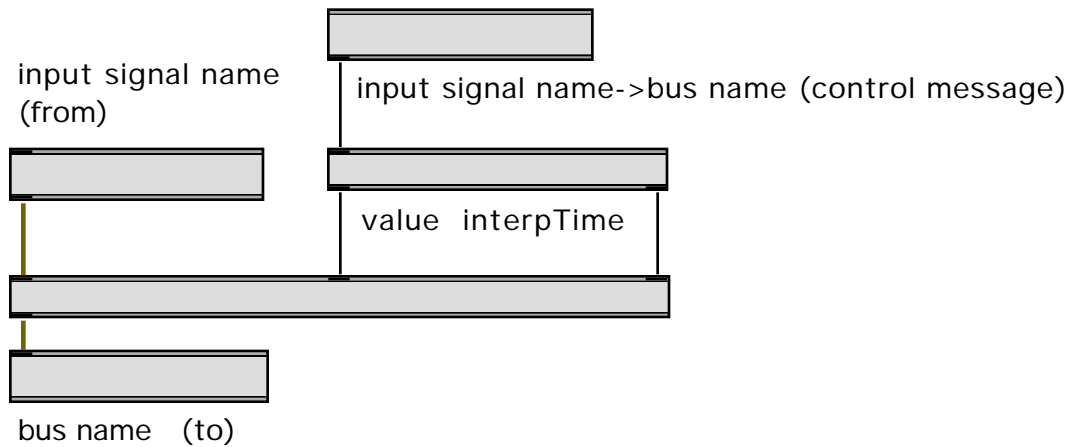
Note the three arguments: the first, "dspA~", is the source signal to be routed; the second, "reverb1~", is the destination bus; and the third dspA->reverb1, is the receiver (in Max) where messages are sent to control the gain and ramp time of the source signal assigned to the bus.

The messages look like this:

```
dspB->clip1 0 1000;
// fade up to 0 db (unity gain) in 1000 ms
;
dspB->clip1 -99 300;
// fade down to -99 db in 300 ms
```

A glimpse of the abstraction's inner works is shown below:

`assign1 ~.abs --assign to bus module`



Thus, the global definition (with six source signals) of the bus sending to `reverb1~` looks something like this:



Figure 28: Signal routing and mixing implementation

FREQUENCY DOMAIN TECHNIQUES FOR SOUND GENERATION AND PROCESSING

Frequency-domain DSP techniques have been a primary focus of mine during the past several years. An entire section has been devoted to this subject, since it encompasses both sound generation and sound processing (and all that lies in between), and uses very different approaches than the ones used in conventional time-domain DSP. In this section I present my methods for frequency-domain DSP, a description of the programming environment, as well as the various approaches to sound processing/generation and implementations that I have developed to date.

Note: Equations are only provided occasionally in order to supplement the discussions, which are often accompanied by graphic pseudo-code figures, or actual code examples in Max/MSP (similar to graphic pseudo-code).

Some background

The Fast Fourier Transform (FFT) is a powerful general-purpose algorithm widely used in signal analysis. FFTs are useful when the spectral information of a signal is needed, such as in pitch-tracking or vocoding³⁶ algorithms. The FFT can be combined with the Inverse Fast Fourier Transform (IFFT) in order to resynthesize signals based on the FFT's analysis. This application of the FFT/IFFT is of great interest in electroacoustic music because it allows for a high degree of control of a given signal's spectral information (an important aspect of timbre), allowing for flexible and efficient implementation of signal processing algorithms. Real-time implementations of the FFT and IFFT are of particular interest since they may be used to provide musicians with a highly responsive and straightforward means for generating and controlling sound in live-performance situations. This section on frequency-domain DSP techniques presents musical applications developed in Max/MSP that make use of FFT/IFFT-based resynthesis for timbral transformation in realtime. Techniques for filtering, cross-synthesis, noise reduction, dynamic spectral shaping, dynamics processing, and phase vocoding are presented along with control structures that allow for fine timbral modification and dynamic control of complex sound transformations using few parameters. Emphasis is also placed on developing control structures fed by control signals based on live input from performers.

The development environment supporting my implementations, Max/MSP, has evolved from the Max software originally developed by Miller Puckette for the Ircam Signal Processing Workstation

³⁶ the imposition of a given signal's spectral amplitude envelope onto that of another signal

(ISPW)³⁷. The standard environment provides all of the necessary unit generators³⁸ necessary for real-time frequency domain DSP. The FFT object provided in MSP is based on Miller Puckette's ISPW implementation³⁹ and stores time-domain signals as buffers of samples upon which the FFT analysis is done. For the purpose of discussion, the examples given here make use of buffers of 1024 samples, though buffers of up to 4096 are used in practice. Unlike time-domain signals, a frequency-domain signal is represented by a succession of spectral "frames". Like frames in a movie, the frames of FFT data represent a "snapshot" of a brief segment of an audio signal. A frame consists of a certain number of equally spaced frequency bands called "bins" (spectral components). The number of bins is equal to the size of the FFT buffer, thus the frames of FFT data have 1024 bins. Each bin describes the energy in a specific part of the audio signal's frequency range.

The basic plumbing

The FFT object outputs each frame, bin-by-bin, using three parallel sample streams (signals) running at the sampling rate. The data is in *rectangular* form, as opposed to the more common *polar* form, where a given spectral component (bin) is represented by a phase value and a magnitude value. In the FFT object, each bin is represented by three samples consisting of *real* (X) and *imaginary* (Y) values, and the bin number (index). At any given instant, each of the FFT's three signal outlets, shown below, produce a sample corresponding to the nth bin of the current FFT frame. The IFFT is the complement of the FFT and expects, as input, real and imaginary values in the same format as FFT output.

Note: Since we are more used to thinking about spectral components in polar form, the following description of *rectangular to polar* conversion is provided:

For a FFT's given bin[n]:

$$\begin{aligned} \text{phase}[n] &= \text{arctangent}(\text{imaginary}[n], \text{real}[n]); \\ \text{magnitude}[n] &= \text{squareRoot}(\text{imaginary}[n]^2 + \text{real}[n]^2); \end{aligned}$$

³⁷ E. Lindemann *et al.*, "The Architecture of the IRCAM Music Worskstation", 1991

³⁸ a module that generates or modifies audio signals.

³⁹ M. Puckette, "FTS: A Real-time Monitor for Multiprocessor Music Synthesis", 1991.

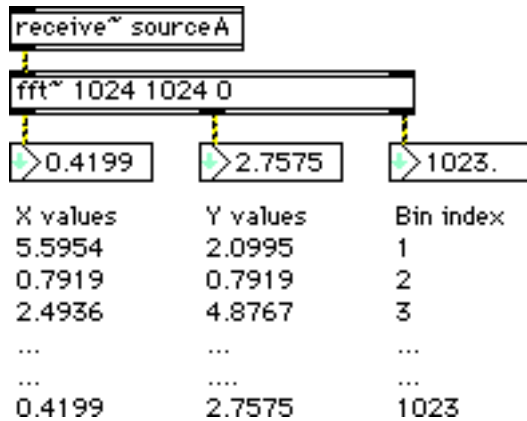


Figure 29: Sample-by-sample output of the FFT object

As seen in the figure above, the index values provide a synchronization signal, making it possible to identify bins within a frame, and recognize frame boundaries. The index values can be used to access bin-specific data for various operations, such as attenuation or spatialization, and to read lookup tables for windowing.

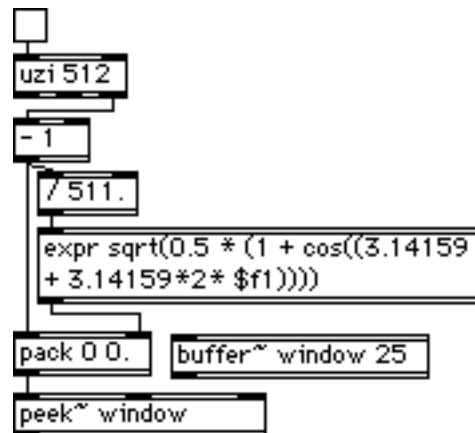


Figure 30: Windowing function generator

Windowing

It is necessary when modifying spectral data to apply an envelope (window) to the time-domain input/output of an FFT/IFFT pair, and to overlap multiple frames⁴⁰. For simplicity's sake, the windowing operation shown below corresponds to a two-overlap implementation (two overlapping FFT/IFFT pairs); it is easily expanded for use in a four or eight-overlap implementation. Because of the flexibility of MAX and MSP, arbitrary windowing functions can be conveniently generated (see

⁴⁰ Rabiner & Gold, 1975

figure above). In the figure below, note the use of the FFT frame index to read the lookup table-based windowing function in synchronization with the frame. The frame index is scaled between 0 and 1 in order to read the windowing function stored in an oscillator.

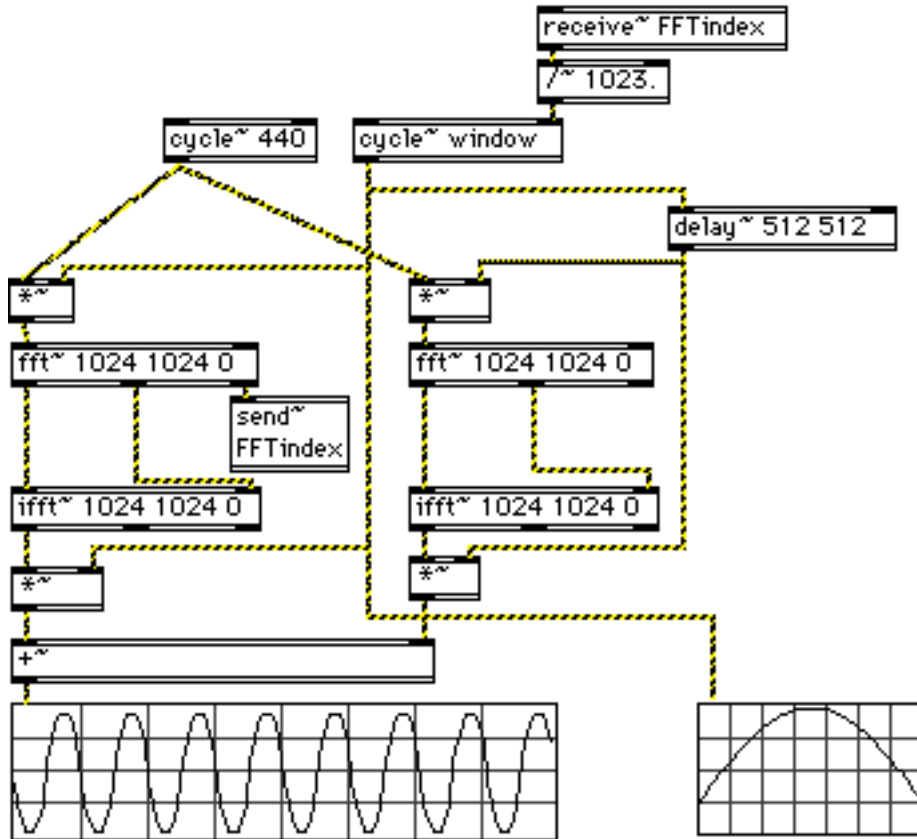


Figure 31: Typical two-overlap windowing of the input and output signals

Bin-specific table lookup

The ability to access specific spectral components (bins) is central to performing frequency domain operations. The following figure illustrates how this can be accomplished using lookup tables. A 1024-point FFT, using a lookup table, offers control of 512 equalization bands across the frequency spectrum. Note that the FFT's frame index signal is used to read the lookup table without rescaling the index values.

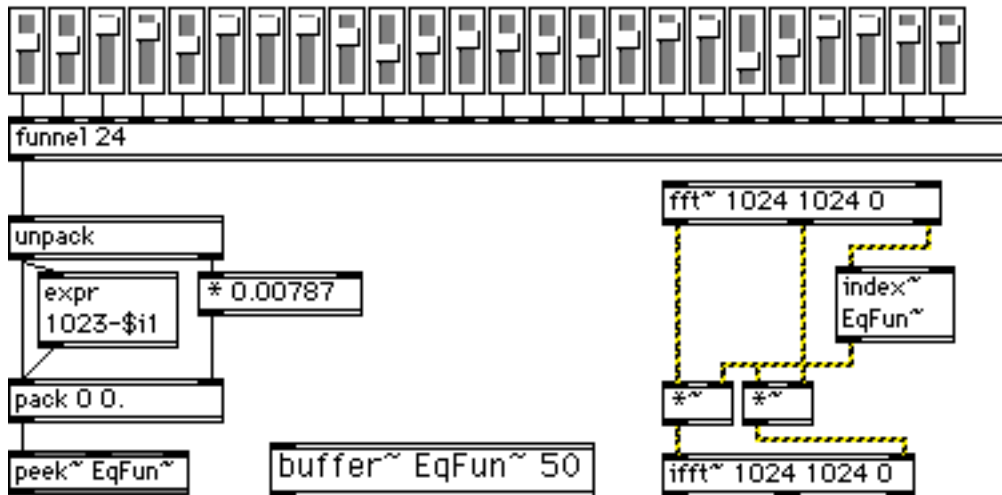


Figure 32: Using a lookup table to obtain bin-specific attenuation values in a graphic equalizer implementation

Algorithms and basic operations

All of the frequency-domain processing applications discussed in this section modify incoming signals and are based on the same general DSP configuration. Using an overlap-add technique, the DSP configuration includes the following steps: (1), windowing of the input signals, (2) transformation of the input signals into the spectral domain using the FFT, (3) operations on the signals' spectra, (4) resynthesis of the modified spectra using the IFFT, (5) and windowing of the output signal. Operations in the spectral domain include applying functions (often stored in tables), convolution (complex multiplication), addition, and square root (used in obtaining an amplitude spectrum); the data are always in the form of rectangular coordinates (X: real and Y: imaginary). Due to the inherent delay introduced by the FFT/IFFT process, smaller (say, 512 point) FFTs are used for live signal processing when responsiveness is important.

Note: The operations (even on phase angles) are always performed on rectangular (X,Y) coordinates, and not polar (magnitude,angle); this keeps calculation overhead lower, and in some instances, provides advantages (e.g. addition of phase without phase unwrapping).

Flavors of convolution

Multiplying one frequency-domain signal by another (convolution) involves the operation of complex multiplication. The math looks like:

$$\begin{aligned} X' &= \cos(a) * \cos(b) - \sin(a) * \sin(b) \\ Y' &= \cos(b) * \sin(a) + \sin(b) * \cos(a) \end{aligned}$$

This operation is at the heart of many frequency-domain processing techniques, and provides the basis for cross-synthesis, filtering, spatialization, phase vocoding, denoising and numerous other applications. A technique often used with convolution is the reduction of a signal's spectrum to its amplitude and/or phase information. Four examples of convolution are shown below; each one corresponds to particular kinds of signal processing applications.

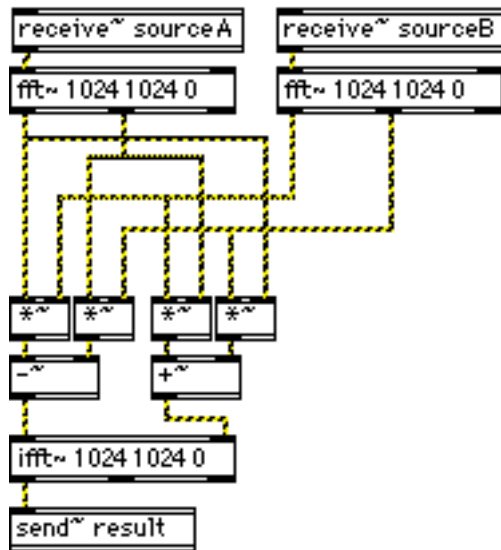


Figure 33: Simple convolution retaining the phases and amplitudes of each input source

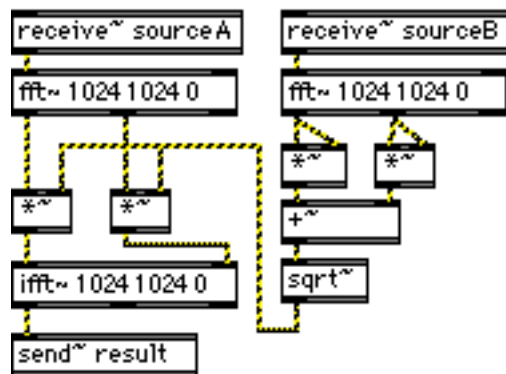


Figure 34: Convolution of phase/amplitude (left) and amplitude-only (right) spectra, used in filtering implementations

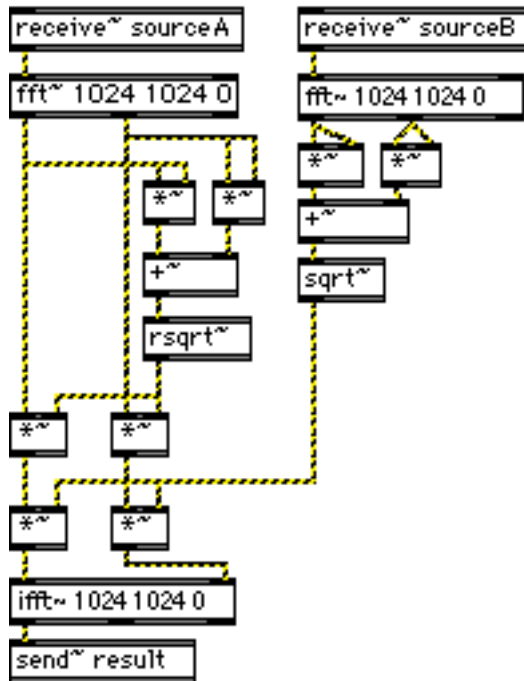


Figure 35: Convolution of phase-only (left) and amplitude-only (right) spectra
 Note: “rsqrt” refers to the reciprocal square root operation (i.e. $1/\sqrt{x}$)

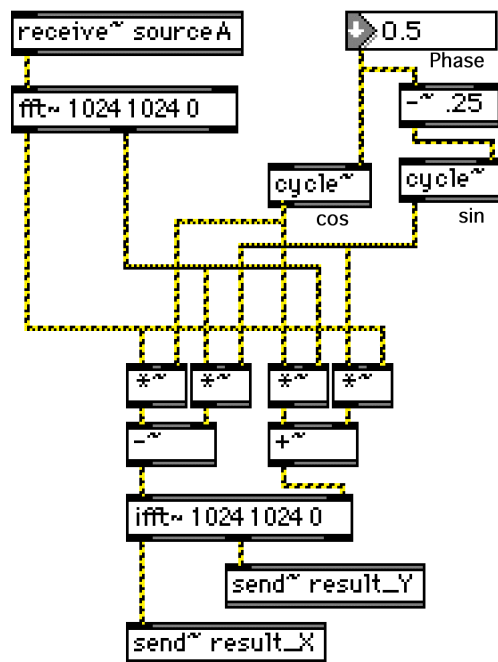


Figure 36: Phase rotation, used in spatialization implementations

Differences in the choice of spectral domain operations, kinds of input signals used, and signal routing determine the nature of a given application; small changes to the topology of the basic DSP

configuration can result in significant changes to its functionality. Much of the code remains the same for diverse applications. Though functionally dissimilar, the convolution-based implementations above differ only slightly: the routine for spatialization closely resembles the one for filtering. The only significant difference is the operation performed on the data. The operation for filtering involves multiplying the input signal's magnitude by a value stored in a lookup table, while the operation for spatialization involves multiplying the input signal by a complex sinusoid, whose phase value is stored in a lookup table.

Frequency-domain DSP applications

The following applications have been developed and implemented; many of them have been incorporated in the live signal processing for my compositions, and/or the composing of their (non-realtime) soundtracks.

Filtering

Highly detailed time-varying spectral envelopes can be produced and controlled by relatively simple means. A look-up table can be used to describe a spectral envelope in the implementation of a 512-band graphic EQ (the bandwidth being $SR/512$, e.g. 86 Hz at SR equal 44.1 kHz). The spectrum of the input signal is convolved, point-by-point, with the data in the look-up table, producing a filtered signal. Because we are able to alter the spectral envelope in real time at the control rate (up to 1 kHz), it's possible to modify the spectral envelope graphically or algorithmically, and hear the results immediately.

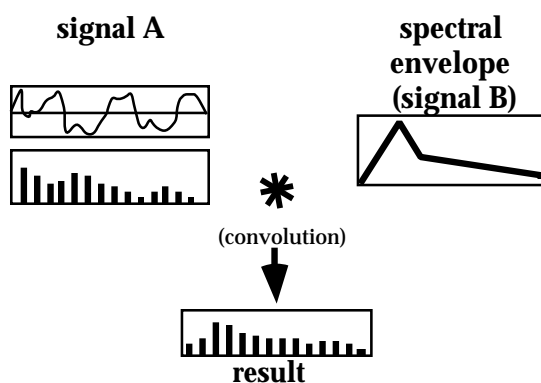


Figure 37: Filtering with a user-specified spectral envelope

Frequency-dependent spatialization

In the frequency domain, the phases of a given signal's frequency component can be independently rotated in order to change the component's energy distribution in the real and imaginary part of the complex output signal. Since the real and imaginary parts of the IFFT's output can be assigned to separate output channels, which are in turn connected to different loud-speakers, it is possible to control a given frequency's energy level in each loud-speaker using phase rotation. Thus it is possible to specify, graphically (or algorithmically) the panning (phase offset) for up to 512 independent frequency bands equally spaced from 0 to $SR/2$ Hz. Unlike in filtering application discussed above, the processed (rotated) signal often has real AND imaginary components, as shown below:

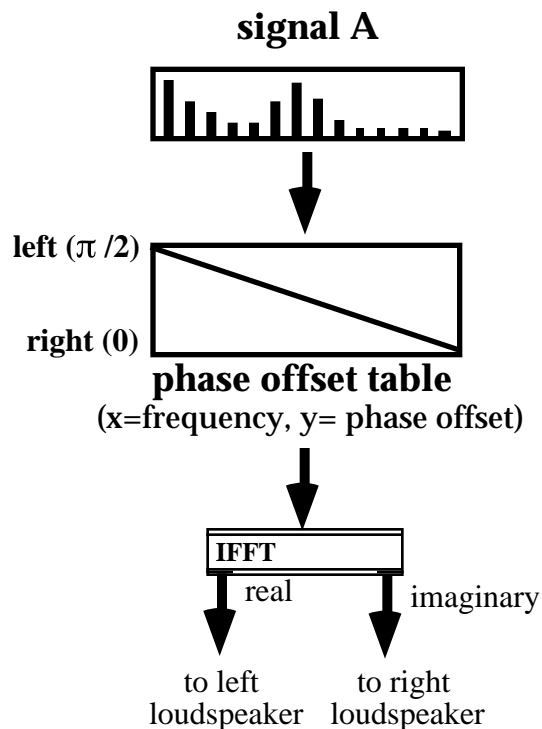


Figure 38: Panning by frequency band

Band-limited amplitude-dependent dynamics processor

In the spectral domain, the energy of a given signal's frequency components can be independently modified. The noise reduction algorithm I have used is based on a technique⁴¹ that allows independent amplitude gating threshold levels to be specified for up to 512 frequency band-limited

⁴¹ A. Moorer & Berger, "Linear-Phase Bandsplitting: Theory and Applications", 1984

regions of a given signal. With a user-defined transfer function, the energy in a given frequency range can be altered based on its intensity and a user-specified threshold level. This technique, besides being potentially useful for noise reduction, can be exaggerated in order to create unusual spectral transformations of input signals, resembling extreme chorusing effects. Using non-linear transfer functions, it is possible to modify the relative intensities of the input's frequency components, allowing for example, masked⁴² or less important components to be emphasized and brought to the aural foreground (see sound example below). In the figure below, a periodic signal at unity amplitude is mixed with low-amplitude noise; the signal is then processed:

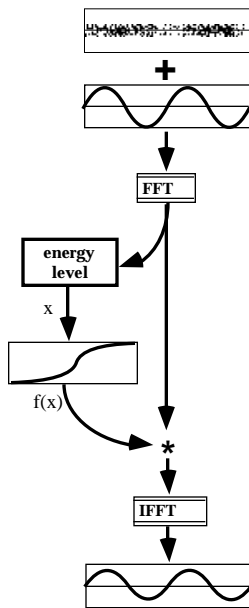


Figure 39: Gating low amplitude noise

The following sound example, taken from a real-time work, illustrates the use of a non-linear transfer function. The resulting spectral distribution of energy is highly transformed from that of the original input.

Sound example 13: *L'enfant des glaces section A: la femme*

⁴² when a spectral component is “shadowed” or “hidden” (not heard) by neighboring components.

Band-limited frequency-dependent dynamics processor

Unlike the dynamics processor described above, this module functions independently of input energy; its output amplitude depends only on the stability of each of the frequency components of the input signal. Using a technique borrowed from phase-vocoding⁴³, time-varying frequency differences of components in a given band-limited region of the input signal's spectrum are examined in terms of stability. A spectral component is considered to be stable when its frequency difference (or stability index) from moment to moment (FFT frame) is small. Another way to see it is, for a given bin_t:

$$\text{Frequency}_t = \phi_t - \phi_{t-1}$$

$$\text{FrequencyDiff}_t = \text{Frequency}_t - \text{Frequency}_{t-1}$$

or just,

$$\text{stability index} = \text{frequencyDiff}_t = \phi''_t \text{ (second derivative, phase)}$$

Note: recall that frequency = phase_t - phase_{t-1}

As a rule, components contributing to “pitch” (including inharmonic partials) in the input signal tend to be stable. Based on a given component's stability index, its output energy can be modified, allowing for attenuation or boosting of pitch-contributing (stable) components or components containing noise (unstable).

⁴³ J. Gordon & J. Strawn, “An introduction to the phase vocoder”, 1987.

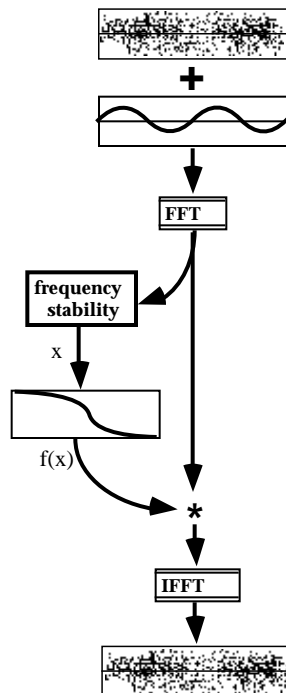


Figure 40: Gating stable frequency components

Sound example 14: Pitch stripping

Note: in the above sound example, a solo flute passage is presented, followed by a “pitch stripped” version of it.

Cross synthesis

In this cross synthesis implementation, two input signals, "A" and "B", are specified; signal A's spectrum is convolved with the amplitude spectrum of signal B. Thus, the pitch/phase information of signal A and the time varying spectral envelope of signal B are combined to form the output signal. Favorable results are produced when Signal A has a relatively constant energy level and broadband spectrum, and when signal B has a well defined time varying spectral envelope. For example, when wishing to transform spoken or sung text, one can assign the text material to signal B while specifying a pulse train, noise source or some other constant-energy broadband signal as signal A. Since the frequency information (pitch, harmonicity, noise content, etc.) of signal A is retained in the output, unusual effects can be produced when frequency related changes occur in signal A. In the following figure describing a vocoder, text can be decoupled from the speaker's or singer's "voice quality", allowing one to modify attributes of the voice such as noise content, inharmonicity, and inflection, independently of the text material.

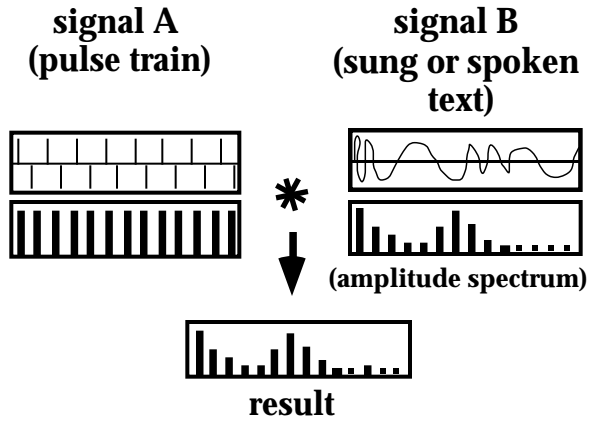


Figure 41: Vocoding implementation

A simple FM pair may be used to provide an easily controlled, constant-energy broadband spectrum for use in cross synthesis as signal A. In some cases, the "musical" relationship between signal A and signal B can become much more unified if certain parameters of signal B are used to control signal A. Specifically, real-time continuous control parameters can be derived from signal B and used to control signal A. For example, the pitch of signal B can be tracked and applied to signal A (FM) to control the two oscillators' frequencies. Envelope following of signal B can yield expressive information, which can be used to control the intensity of frequency modulation (FM index) of signal A. In experiments incorporating the above, a mezzo soprano's voice was assigned to signal A, while her pitch and intensity were mapped onto signal B (FM), producing striking results akin to harmonization and frequency shifting.

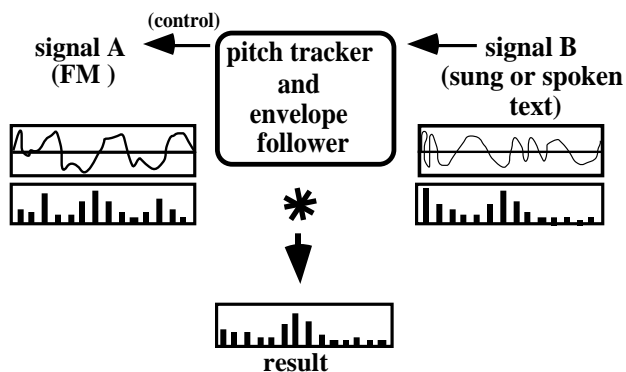


Figure 42: Vocoding with performer-analogous inflection and spectral emphasis

The following sound example illustrates the use of a saxophone as signal B:

Sound example 15: Punjab section 1

It should be noted that interesting transformations can be produced by simply convolving signal A's spectrum with signal B's spectrum. In this case, the phase (frequency) and spectral envelope information from each signal figures in the output signal. Transformations of broadband sounds, akin to, but more pronounced than flanging, can be produced when convolved with the signal of a high-index, inharmonically-tuned FM pair, whose frequency parameters are controlled by the pitch of the first signal.

Finally, an interesting way to view the Cross Synthesis process described above is to consider signal A as being filtered by the spectral envelope of signal B. Given the rather high resolution, stability, response time and economy of such a filter, one is tempted to look for "B" signals, whose relatively complex spectral envelopes can be controlled with few parameters. FM synthesis provides an excellent source for broadband signals of considerable spectral complexity, with low-dimensional control. Interesting and dynamically (non-linearly) changing "filter" functions can be obtained by acting on the FM synthesis parameters of Index and Ratio; where roughly speaking, the index determines the filter's bandwidth, and the ratio determines placement of the notches and peaks (this idea is further developed in the section on "Generating complex spectral envelopes via spectrum generation", p. 67).

Improvements to cross synthesis: spectral compatibility

One serious problem has emerged with the above approach to "filtering" (though the problem is general to cross synthesis): the filter function's notches tend to be quite deep and wide, while its peaks tend to be tall and narrow (as in high "Q"). In convolution, a multiplicative operation, the similarity of both signal's spectral energy distribution (where the notches and peaks lie) is key to the outcome of the operation, since the resulting spectrum will contain only components which are common (have energy) to both input spectra (signal A and the filtering function given by Signal B). Otherwise said, if you superimpose both spectra, the only parts that will sound are those that overlap. Thus the idea of "spectral compatibility" emerges when considering the suitability of two given signals for cross synthesis.

Spectral smoothing for input with dissimilar spectra

I developed a technique to de-emphasize spectral notches and peaks by smearing the energy in peaks into the adjacent notches. An amplitude spectrum with tall peaks and deep notches (the case for a pitched sound) makes a good candidate for spectral smoothing. With the FFT implementation in MSP, it is quite easy to apply a second-order low-pass filter to an amplitude spectrum. The filter's cutoff frequency and damping parameters control the degree of spectral smoothing, averaging the energy across empty bins, thereby reducing the sharp notches or peaks which can be harmonically distributed across the spectrum of pitched signals. The smoothed spectrum shown below will combine via convolution (multiplication) much more effectively with the spectrum of another sound—particularly when the other sound also has pronounced peaks and notches distributed differently across its spectrum. For example, this technique can prove useful in crossing the amplitude spectrum of a singer with the spectrum of a pitched instrument, such as a saxophone—especially when they are sounding pitches with few harmonics in common.



Figure 43: “Incompatible spectra” rendered compatible.

Top: notched amplitude spectrum of sound A

Middle: sound A’s smoothed spectrum (to reduce notches/zeros) is more compatible with sound B’s spectrum

Bottom: notched amplitude spectrum of sound B.

The amplitude spectrum of signal A is low-pass filtered, thus reducing the sharp notches or peaks typically distributed harmonically across the spectrum of pitched signals. Thus, the degree of spectral energy intersection between signal A and signal B is augmented. The lowpass filter's cutoff frequency value provides the cross synthesis implementation described above with an additional interesting and useful parameter for spectral "smoothing". A sound example demonstrating this technique is provided.

Sound example 16: dynamics_example6

Spectral energy scaling for input with dissimilar spectra

The “incompatible spectra” cross-synthesis problem described above can also be addressed using another approach I have developed. Like the technique discussed above, the basic goal is to eliminate zeros or notches in the input signal(s). But rather than redistributing energy *across* a given spectrum, the approach entails boosting spectral energy where it is low, via bin-independent input amplitude scaling (frequency-domain dynamics processing). Input scaling provides yet another way to increase the degree of spectral energy intersection in cross synthesis, and yields an additional parameter to control it. The approach is shown graphically below:

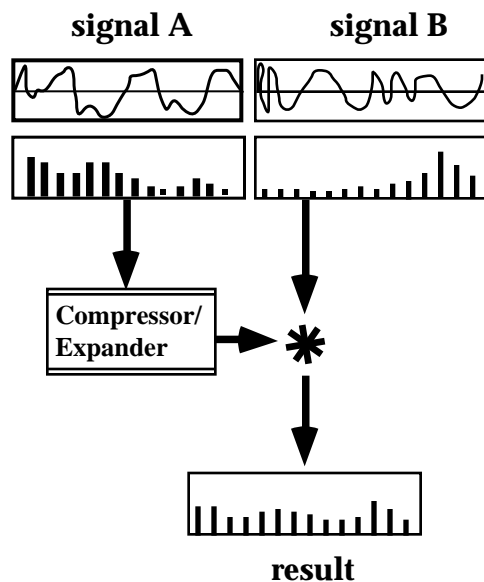


Figure 44: Energy scaling to increase the degree of spectral intersection

By choosing a compression/expansion function that boosts or cuts the energy of weaker components in one of the input signals, the degree of spectral intersection of the two inputs can be specified. The compression/expansion ratio parameter (or intensity of compression/expansion) provides dynamic control over the "degree of spectral intersection" of the two input signals. This parameter is particularly useful when cross-synthesizing dissimilar sounds such as tuba and percussion.

As mentioned above, boosting weaker components of an amplitude spectrum is a technique which increases the potential degree of spectral intersection in cross-synthesis. Below is a description of a

compressor/expander that operates independently on each frequency bin of a spectrum⁴⁴. As shown below, each bin's amplitude is used as an index into a lookup table containing a compression/expansion function; the value read from the table is then used to alter or scale the particular bin's amplitude. The degree of dynamic processing is controlled by biasing or scaling the lookup table index.

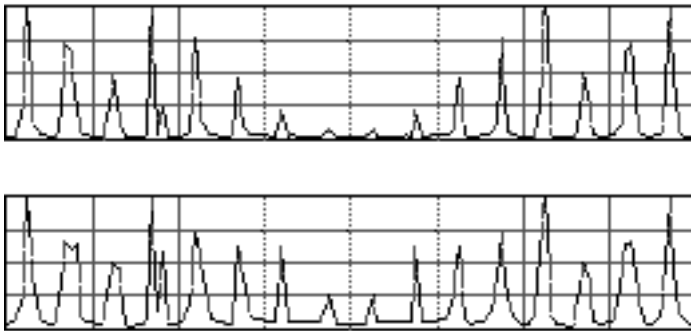


Figure 45: Dynamics processing to boost weaker spectral components:
 Top: input signal's spectrum
 Bottom: input signal's expanded spectrum

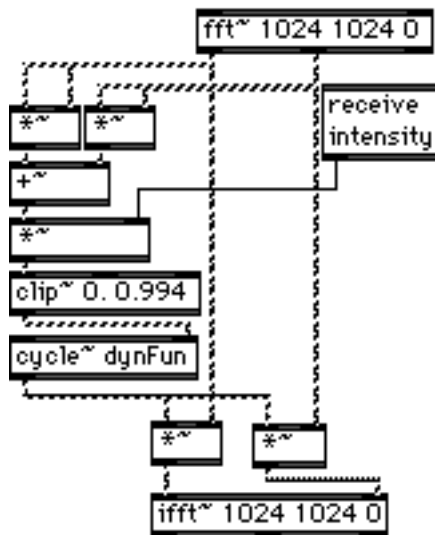


Figure 46: Dynamics processor implementation in MSP

Note: compression/expansion functions may also be used to attenuate stronger amplitude components. This can be an effective check against extreme amplitude levels, which result when crossing two sounds with similar spectral energy distributions. This problem is discussed later in this section.

⁴⁴ C. Lippe, "FFT-based Resynthesis for Timbral Transformation in a Compositional Context", 1994.

Finally, it is possible to maximize a spectrum's potential for intersection with another spectrum by forcing the amplitude information in each bin of the spectrum to a constant value of one (unity). Phase information remains unmodified (though not always defined). The resulting constant-amplitude phase-only spectrum will combine with the spectrum of another sound wherever there is energy in the other sound's spectrum (see "a phase-only constant-amplitude spectrum is crossed with an amplitude spectrum" in *Flavors of convolution*, p 49).

In any spectrum, the phase of an "empty" (near 0 amplitude) bin is undefined. Consider two spectra with dissimilar spectral energy distributions. When forcing the amplitude of one sound's empty bins to unity and performing cross-synthesis with the amplitude spectrum of another sound, the resulting spectrum will tend to contain components whose phase is undefined (random). Thus, the resulting sound will contain little, if any, pitch or stable frequency information. A sound example illustrating this technique is included.

Sound example 17: dynamics_example7

Controlling the degree of spectral intersection

By combining the dynamics processing technique with the constant-amplitude forcing techniques described above, the degree of amplitude forcing towards unity can be continuously controlled. Thus, it is possible to specify how much of a given spectrum's original amplitude information, if any, will be used in a cross-synthesis operation with another spectrum.

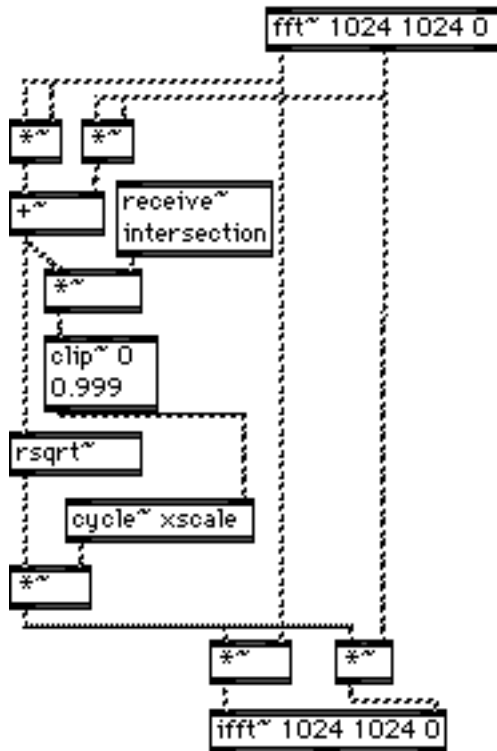


Figure 47: The amplitude spectrum "intersection" parameter

Spectral energy scaling input signals with too-similar spectra

We have examined two approaches that address the problem of “spectral incompatibility” in cross synthesis. There remains one more problem, which occurs when the spectra of two input signals for cross synthesis are *too* similar. When two input signals of similar intensities and spectral distributions (e.g. two singers' voices) are convolved, the resulting spectral distribution can be quite different. In such a case, strong lower mid range components can become much louder, while weaker higher frequencies virtually disappear. A good worst-case example of this occurs when convolving a signal by itself.

The resulting energy of a given component is its square, as shown below:

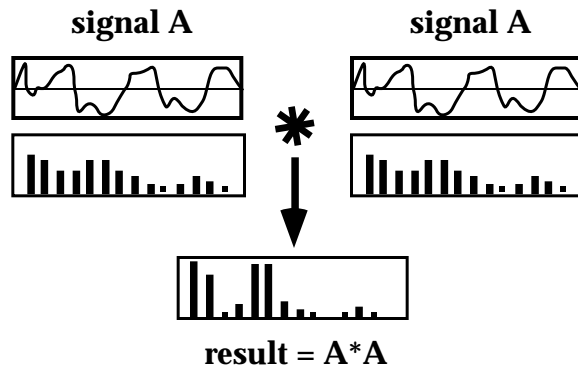


Figure 48: Exponential combined energy from cross synthesis

By scaling the input signal(s), the "original" spectral distribution can be preserved. If in the above example, we take the square root of the output signal, a given component's energy is now the arithmetic mean of the energy of its input signals, and not the square. In this case we recover the original input signal.

$$\text{result} = \text{squareRoot}(A*A) = A$$

The algorithms for cross-synthesis shown earlier in this section can be expanded to include additional processing units for amplitude scaling (dynamics processing) and additional control parameters. The compressor/expander based on the energy-dependent dynamics processor mentioned above is applied to one of the two input signals before convolution. The choice of compression/expansion functions determines the way the energy in the output signal will be scaled; certain choices (such as a square root function) help preserve the spectral distribution of the input signals. Thus, two input signals of similar intensities and spectral distributions will combine to form a spectrum within which the relative energies of the components remain almost unchanged.

Phase Vocoder: improvements for noise stretching

A version of Miller Puckette's phase vocoder, originally written in Max/ISPW in 1994, was ported to Max/MSP in the summer of 1999. Since then, I have added a feature to Puckette's implementation that provides for noise stretching; it allows "noisy" components of the spectrum to retain their character, independent of the stretch factor. The assumption is that unlike the phase of stable components (pitched or inharmonic), the phase of noise components changes irregularly from

moment to moment. In the implementation of the phase vocoder, you can say that the phases of noise components change randomly within a certain range from one window of input to the next. Thus, noise components in the phase vocoder can be characterized by their rate and quality of change in time, i.e. noise is change. Now, when the phase vocoder is reconstructing a signal with a stretch factor of 1 or less, the duration is shortened and the rate of change for spectral components is as great or greater than in the original signal. However, when the stretch factor is greater than one, the inverse occurs: the phase propagation of each component slows down according to the stretch factor—as does the rate of spectral change. Noise stops sounding like noise and becomes some sort of weird semi-stable spectral mass, composed of many various non-harmonically related tones. This is why slowed-down phase-vocoded whispering sounds terrible—or, at least, quite unlike slowed-down whispering.

In order to make a “better” sounding noise stretcher, I introduce, in the resynthesis stage, a degree of random phase offset for each bin proportional to the stretch factor. Thus, I am able to maintain a suitable spectral rate and quality of change for the noisy signal being stretched and resynthesized. Here’s the method in pseudo-code for a given bin_k:

```
phase[k] = lastPhase[k] + stretchFactor * random(0 to 3.14159)
```

Note: Puckette’s implementation of the phase vocoder operates on spectral data in rectangular form, thus the algorithm for phase offset involves the multiplication of the (complex) signal by a complex random number with unity magnitude, thereby rotating the phase by a given amount. The pseudo-code for the operation is as follows:

```
randPhase = stretchFactor * random(0 to 3.14159);
realRand = cos(randPhase);
imaginaryRand = sin(randPhase);

// complex multiply to produce random phase instability
newX = (realCurrent * realRand) - (imaginaryCurrent * imaginaryRand);
newY = (imaginaryCurrent * realRand) + (realCurrent * imaginaryRand);
```

Ideally the degree of random phase offset should be a function of both the stretch factor and the instantaneous degree of noisiness in the signal being stretched. A crude attempt at this was done using a zerocrossing tracker as a noise detector; the greater the density of zero crossings (noise content) in the signal being analyzed, the greater the degree of noise regeneration (random phase offsetting) in the resynthesis stage of the phase vocoder. However, a better solution would be to perform bin-independent noise detection and noise regeneration. Using the technique of phase instability detection, discussed above in Band-limited frequency-dependent dynamics processor (p. 55), it is possible for the degree of noise regeneration in each bin to be proportional to its

measured noise content (phase instability). The included sound example demonstrates the stretching of a noisy signal. The sound example features a whispered phrase followed by four stretched versions: stretch factor equals .2 *normal*, stretch factor equals .2 *noiseEnhanced*; stretch factor equals .01 *normal*, stretch factor equals .01 *noiseEnhanced*.

Sound example 18: Stretched_whisper

Audio-rate control of FFT-based processing

Control limitations

As is often the case with audio programming environments, the Max/MSP environment has two run-time schedulers: the Max "control" scheduler, which is timed on the basis of milliseconds, and the MSP "signal" scheduler, which is timed at the audio sampling rate⁴⁵. In FFT-based processing applications, where changes to the resulting spectrum are infrequent, MSP's control objects may be used to provide control parameters for the processing. This is both precise and economical, but limited in terms of the spectral rate of change. For example, our implementation of FFT-based narrow-band "graphic" equalization uses a lookup table describing a filtering function, which is updated at the control rate. A given frequency-domain input signal is convolved with this function and spectral processing is accomplished. However, updating lookup tables at the control rate has bandwidth limitations. The rapidity with which a lookup table can be altered is limited by the bandwidth of the control system. Using 512 sliders to control individual FFT bins, drawing a filter shape for a lookup table with the mouse, or changing the lookup table data algorithmically provides only limited time-varying control of the filter shape. In addition, the amount of control data represented in a lookup table is large and cumbersome. Significant and continuous modification of a spectrum, as in the case of a sweeping band-pass filter, is not possible using MSP's control objects, since they can not keep up with the task of providing 1024 parameter changes at the FFT frame rate of 43 times a second (at the audio sampling rate of 44,100 samples per second).

Keeping in mind that the input signal to be modified is FFT data, a more dynamic approach to filtering is to update lookup tables containing a filter function at the signal rate (audio sampling rate). The term "Spectral Processing Function" (SPF) will be used frequently in the following text. It refers to a lookup table-based function (actually a signal), whose length is that of the FFT. For each window of input signal (FFT data) we receive in real-time, we generate a corresponding SPF

⁴⁵ M. Puckette, "FTS: A Real-time Monitor for Multiprocessor Music Synthesis", 1991.

with which the input may be convolved. Dynamically, the SPF can describe a particular sequence of forms (or spectral envelopes), which determine the time-varying intensity of spectral processing by frequency component via convolution; a "form" describes the action of the spectral processor. Thus, my approach to dynamic spectral processing focuses on efficiently generating forms with a potentially high degree of detail or complexity, whose descriptions are simple (generated using few parameters, for low dimensional control) and intuitive⁴⁶.

Generating spectral processing functions (SPF)

I will present two methods for generating SPFs: the first involves the use of a complex waveform generator's spectral envelope via the FFT, while the second generates the spectral processing envelope directly using function generators (waveforms and operators). In each case, low-dimensional parametric control of complex forms is achieved. In both cases, a time-varying SPF is generated for, and convolved with each window of input signal to perform filtering or band-limited panning as shown below:

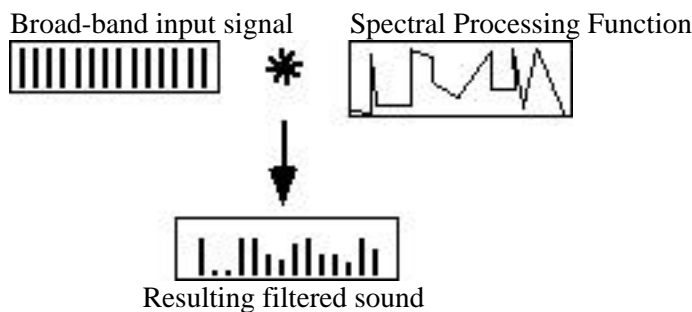


Figure 49: A signal is filtered (via convolution) using a Spectral Processing Function (SPF)

Generating complex spectral envelopes via spectrum generation

When hunting around for a suitable method to generate complex "forms" (or spectral envelopes) using simple techniques, the classic FM pair algorithm (generating complex spectra using frequency modulation), described earlier in this section, leaps to mind. The SPF is simply the spectral envelope of the signal generated via FM. It provides a rich source of possible forms, whose shapes and complexity are determined using the FM algorithm's few parameters (carrier FQ, carrier: modulator ratio, modulation index). The intuitive mapping of the FM parameter values to the resulting spectral form makes this method extremely easy to use. In this implementation, a

⁴⁶ Z. Settel & C. Lippe, "Real-time Frequency-Domain Digital Signal Processing on the Desktop", 1998

particular set of static or time-varying parameters is specified for the FM algorithm; the parameters determine the shape of the resulting SPF. With each window of input signal to be processed, the algorithm's output is transformed via an FFT into a SPF (spectral envelope), which is then applied to the corresponding window of input signal via convolution. Thus, the maximum rate of change for the SPF (signal processing parameters) is given by the frame rate (function of window length) of the FFTs; a 1024-point FFT, with a sampling rate of 44.1 kHz, would translate to a spectral update rate of approximately 43 Hz.

Needless to say, other sources for generating complex spectra, such as amplitude modulation, additive synthesis, or waveshaping may be used in a similar implementation. However, the clear advantage of using FM lies in the simple control of the highly complex spectra it offers. For example:

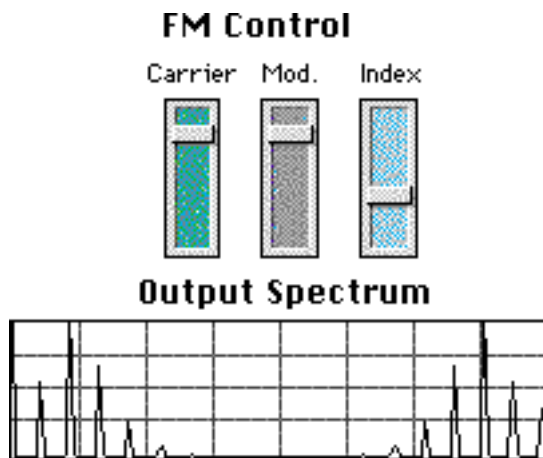


Figure 50: A complex spectral form depending on only three parameters

Generating complex forms using waveform generators

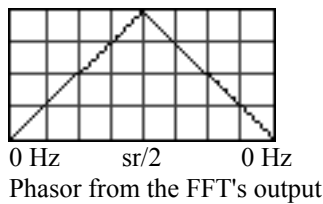
The use of wave tables and basic table lookup operations provides the foundation in the (classic) approach to waveform generation and synthesis techniques I use to generate Spectral Processing Functions (SPF). Techniques such as FM, AM, waveshaping, phase modulation and pulse-width modulation, all have the potential to provide complex, evolving waveforms, which can then be used as SPFs to provide a high level of flexibility and detail for spectral processing techniques such as filtering or spectral panning. The use of table lookup operations such as inversion, scaling, offsetting, wrapping, and nonlinear distortion (waveshaping), provides powerful means for dynamically modifying these SPFs. For example, nonlinear indexing of lookup tables can be

employed to provide dynamic control for a constant-Q bandpass filter, where control on a nonlinear frequency scale is required. Most important, the parameters of these waveform-based techniques are few, familiar and easy to understand.

Note: Given the dynamic nature of the SPF produced by waveform generation, the reader may wonder what the SPF's limitations are regarding the rate of (spectral) change, since the SPF can change more rapidly than the FFT's frame rate. However in practice, the complex waveform generator is used in the manner of a Low frequency oscillator (LFO), never changing at rates greater than 30 or 40 Hz.

Let's now focus on the lookup table operations underlying these techniques for the generation of a Spectral Processing Function. As discussed above, the FFT used in these implementations provides a phasor that is used as an index for table lookup.

Note that the index is mirrored around the sampling rate divided by 2, following from the symmetrical (real) spectrum output by the FFT.



The following sequence of operations is performed on a phasor to generate a SPF; note how its shape (form) is progressively changed by each operation.

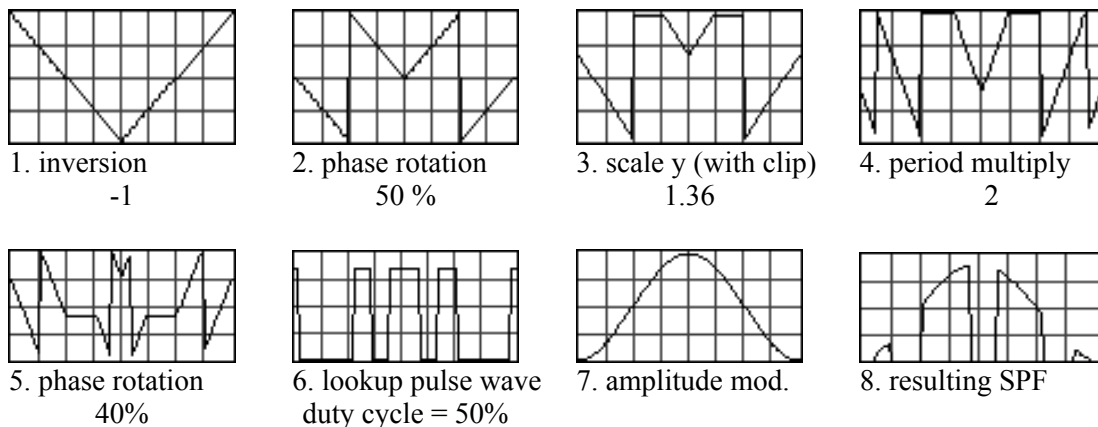
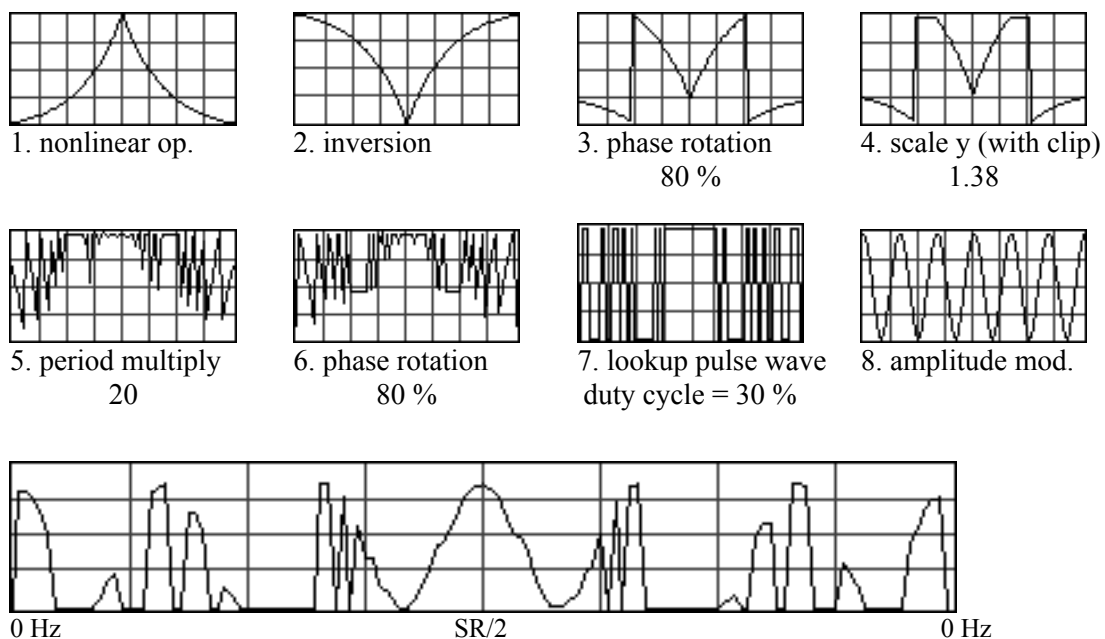


Figure 51: 1-8 Lookup table operations to generate a SPF

Note that step 8 is the result of step 6 multiplied by the function in step 7. Also note that steps 3, 6, and 7 entail a loss of information; their placement in the sequence of operations cannot be arbitrary.

While the above operations and parameter settings will produce a static SPF, it is important to note that the "phase rotation" operation also accepts a frequency parameter (LFO); a non-zero value will cause the phase rotation value to change smoothly and constantly. Thus, the resulting SPF will vary periodically due to the constant change of its phase rotation parameter. This "LFO-based" technique provides control for periodic spectral modulation, and is useful in implementations such as swept band-pass filters (or spectral panners), comb filters, or "phase shifters".

To generate a SPF of greater complexity, I start with the same phasor above and then apply the following operations:



Resulting Spectral Processing Function (SPF)

Note that the resulting SPF is the result of step 7 multiplied by the function in step 8.

Figure 52: Generating a complex SPF

Applying the techniques

The above techniques for dynamic spectral processing lend particularly well to applications such as high-resolution dynamic filtering, and spectral panning. I have implemented such applications, and have experimented with spectral resolution, making use of 1024 to 4096 point FFTs, twice or four times overlapped. Given the implicit tradeoff between time and frequency resolution, and considerations of computational cost, the use of two overlapped 1024 point FFTs is preferable for real-time performance situations when audio signal latency (delay from input to output) is an issue. However, a longer FFT of 4096 points is generally preferable, despite the relatively low spectral

update rate of 10 Hz (10 FFT frames per second), since a long window allows for a (correspondingly long) highly detailed spectral processing function (SPF). For example, the "period multiply" operation shown above requires a long SPF when a greater number of periods is specified.

In the case of high resolution filtering applications, the SPF specifies the degree of attenuation to be applied to each component (band-limited region or bin) of the input spectrum. The input signal's spectrum is convolved with the SPF and filtering is accomplished. The nature of the filter is determined by the form of the SPF. However, in the case of an application for spectral panning (band-limited panning), the SPF is used to specify the degree of phase rotation, via convolution with a complex sinusoid, to be applied to the components of the input spectrum. This operation results in a change to a given component's energy distribution in its real and imaginary parts. The real and imaginary outputs of the IFFT are mapped to a corresponding stereo output. Dynamic spectral panning occurs when the phase of the input signal's spectral components is rotated by a changing amount.

The SPF is applied in each application as shown below:

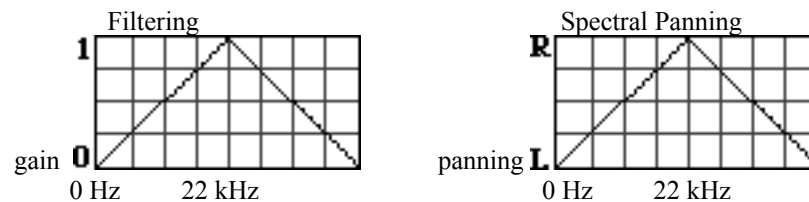


Figure 53: Applying SPFs for filtering and spectral panning

The filter above defines a highpass. The spectral panning distributes (pans) the energy of the input signal's components continuously from left to right, based on the component's frequency. For example, components near 0 Hz are panned Left, components near 12 kHz are panned L and R, and components near 22 kHz are panned right. Sound examples of filtering and spectral panning are provided.

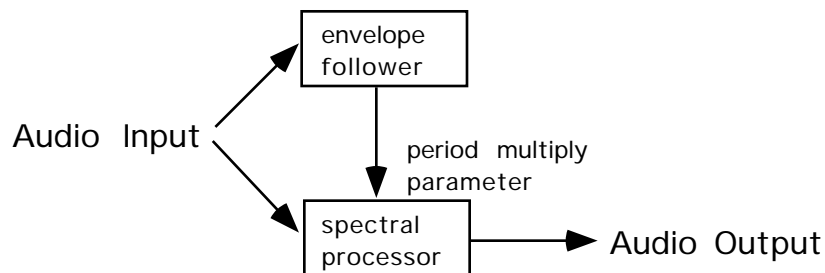
Sound example 19: *dynamics_example8*

Sound example 20: *dynamics_example9*

Dynamic control of processing parameters via input signal analysis

Compressor/limiters are examples of signal processors that analyze their input signal to determine how it is in turn to be modified. A signal passing through such a processor can be thought of as a self-modifying signal. When an analysis stage (such as an envelope or pitch follower) is added to the spectral processing implementation, the resulting input analysis information may be used to dynamically control one or more parameters of the spectral processing. Since the interface for spectral processing requires only a small number of parameters, and since each parameter can significantly alter the shape of the SPF (offering a wide potential range of transformational possibilities), the mapping of few (or even just one) input-derived control streams to processing parameters can provide a very high degree of signal self-modification.

Musically, the ability to control the degree and quality of a signal's spectral transformation via aspects of the same signal's (performer's) dynamics or inflection can be compelling. In live performance applications, this mapping of musical gesture to resulting timbre can be quite tangible and inspiring to the performer. A particularly effective mapping of this sort can be produced using the Waveform Generation technique discussed earlier, where the control derived from the input intensity is mapped to the 'Period Multiply' parameter of the processing shown below:



Which, as a function of input signal intensity, can produce a range of SPFs such as:

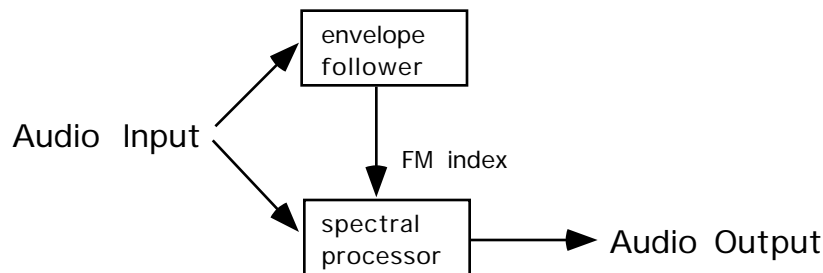


Figure 54: Self-modifying signal via the waveform generation technique

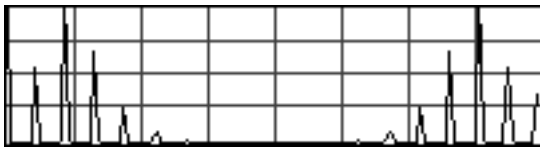
Thus, the more intense the input signal, the greater number of peaks and notches in the SPF. In a filtering application, the effect is that increasing input dynamic level translates to increasing

separation (as opposed to fusion) of spectral components into discretely perceived sinusoids. In a spectral panning application, increasing dynamic level translates to increasing stereo separation of neighboring components, enhancing the perception of sound source location. In practice, the value of the period multiply parameter may reach 40 or 50, allowing for extremely fine spectral processing.

Alternatively, a similar result can be achieved using the Spectral Generation technique discussed above. In this case, a frequency modulation pair is used to generate the SPF, where the control derived from the input intensity is mapped to the "modulation index" parameter of the processing shown below.



The more intense the input signal, the greater the index of modulation, and thus, the greater the number of peaks and notches in the SPF.



SPF generated from the spectrum of an FM pair with moderate index value

Figure 55: Self-modifying signal via the spectral generation technique

The following sound example demonstrates this technique.

Sound example 21: dynamics_example12

Needless to say, there are many other possible mappings of input features to spectral processing parameters. In each case, a given aspect of the input signal will potentially cause the spectrum of that same signal to be modified in a significant way. The challenge is, of course, to recognize the practical and musical sense of certain choices on mappings, which in turn, can serve a musical

purpose in performance or composition. The approach described above of playing an instrument while processing its sound has proven to be a very effective way to evaluate the mappings and discover their particular musical tendencies. As with dynamics processors, the choices of input signal, analysis mapping and processing technique produce results that cover a wide range of timbral possibilities. My ongoing work in this area is concerned with discovering and establishing musically effective combinations of these choices.

MULTI-CHANNEL SOUND SPATIALIZATION

During the past four years I have been developing techniques for multi-channel sound source spatialization. The approach entails placing one or more sound sources in a multi-point (speaker) diffusion field, and providing for dynamic and independent movement of the sound sources. The basic spatializer is presented below, followed by a particular application where rhythm and panning are coupled.

Sound source positioning and simulation

The terms “sound source positioning”, and “simulation” respectively refer to: 1) the distribution of a sound’s energy among several loudspeakers in a particular diffusion space via panning, and 2) the application of additional signal processing steps to provide Doppler shift, delay, filtering, reverbation and attenuation for modeling the behavior of sounds located in acoustic spaces (Chowning⁴⁷, Moore⁴⁸, and Kendall⁴⁹). The model is a geometric one, where sound energy is represented by vectors in a two dimensional quadrilateral space (three was too expensive to warrant the fuss), behaving according to laws on the physics of sound.

⁴⁷ Sound source Simulation Chowning.

⁴⁸ Sound source Simulation Moore.

⁴⁹ Sound source Simulation Kendall.

Below is a view of the interface and its parameters:

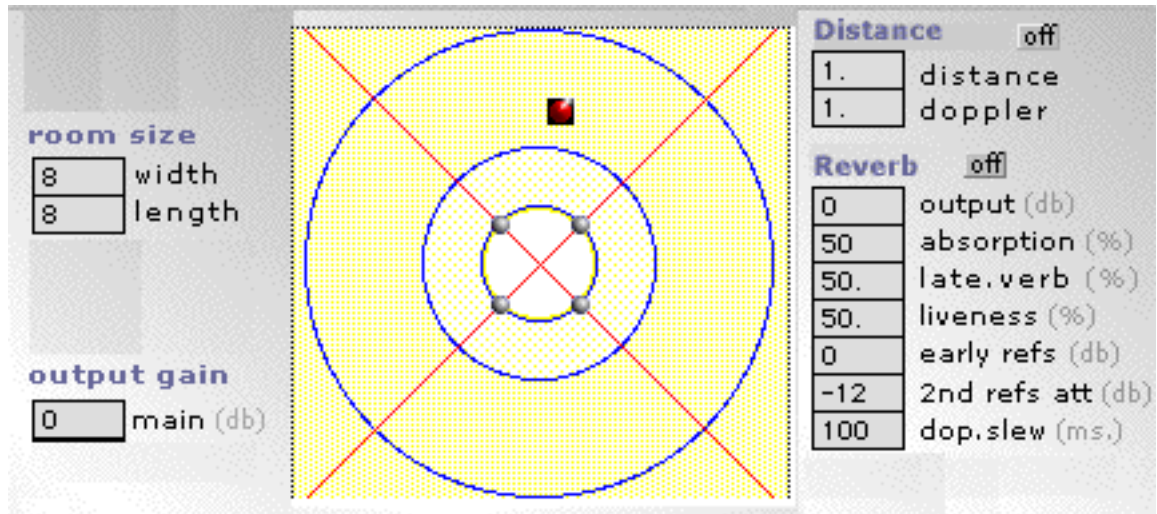


Figure 56: Interface for spatialization

The speakers are represented by small gray spheres, and the sound's position by a small square cursor. The listener is assumed to be located in the center of the speaker field. The rings on the interface are placed every -6 db from the listener, representing a doubling of distance from the center (e.g. at 1, 2, and 4 meters). For flexibility and computational economy, a modular approach was adopted; thus, panning, distance and Doppler simulation, and multi-channel reverberation can be combined, as needed, for particular needs of sound spatialization. As shown in the interface above, the following three “levels” of simulation are available:

1) *Panning* (always enabled) distributes the energy of the direct signal in the speaker field as follows:

If the sound source position is within the circle, on which the speakers are located, the resulting energy is a weighted sum, distributed among all the speakers.

If not, the sound source position is outside of the circle, and the energy is distributed among only the nearest two adjacent speakers, based on the source position's angle of incidence with respect to the listener (center).

2) *Distance* provides distance-scaled attenuation (i.e. $1/\text{distance}^2$) of the direct signal and Doppler shift, via variable delay. When movement about the listener is not circular, the distance between the sound source and the listener changes, as well as the corresponding delay time for the variable

delay, producing transposition (Doppler). The two parameters shown in this section of the interface control the intensity of the effect of distance and Doppler, respectively.

3) *Reverb*, like *Distance*, provides additional cues via signal processing for space simulation according to a geometric model for sound behavior. It includes filtering, early reflections, and a recirculating reverb section. The reflections are copies of the direct sound, which are filtered, based on the quality of the surface in the room model. They are delayed and attenuated according to each one's length of travel, and panned according to each one's angle of incidence with respect to the listener. The implementation provides four first-order⁵⁰ (one for each wall) and eight second-order reflections. To avoid clicking when dynamically changing the position of the sound source, interpolated delays are used. The recirculator section's energy is distributed evenly among the speakers; the weighting of the recirculator's (reverberation) energy with respect to the direct energy increases with distance. The interface above (Figure 56: Interface for spatialization) provides parameters to configure the room size, and quality; some are explained below:

- **'output'**: the reverb section's overall output attenuation.
- **'absorption'**: translates to the cutoff frequency of the lowpass filter, applied to the signal in the early reflection and recirculator sections of the reverb. This parameter corresponds to the "surface quality" of the walls in the modeled space. For example, carpeted walls would have a much greater absorption (lower cutoff frequency) than glass walls.
- **'late.verb'**: attenuation for the recirculator's output. The recirculator provides a resonant "shadow" of the input sound, providing additional cues about the resonance of the space.
- **'liveness'**: like reverb time: amount of recirculation (feedback).
- **'early refs'**: attenuation for early reflections (both 1st and 2nd order). In the model, the reflections provide cues about the dimensions and quality of the space.
- **'2nd refs att'**: cut/boost for 2nd order early reflections.
- **'dop.slew'**: interpolation time parameter for the early reflection delay units: larger values reduce the side effect of Doppler, when the sound source position movement moves and reflection path lengths change.

The parameters shown in the above interface can be stored and recalled as presets, including the on/off state of the two simulation levels, *distance* and *reverb*, which are not always required. For

⁵⁰ another term for "first-order reflections" is "first reflections", similarly "second-order reflections" are referred to as "second reflections", etc.

example, in the application for Rhythm-synchronized spatialization, discussed below, only panning is used.

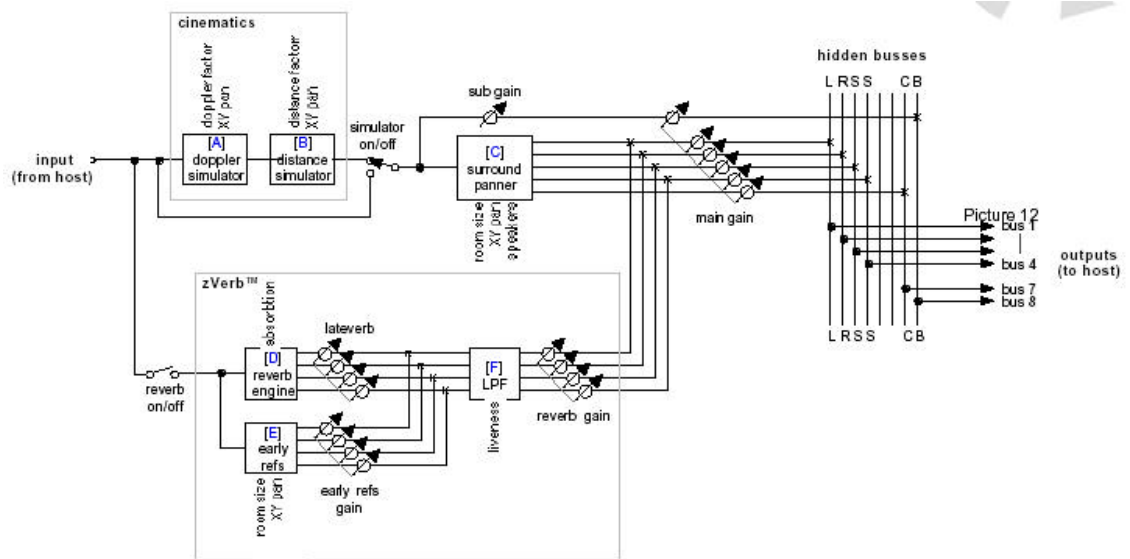


Figure 57: Schematic overview of the spatializer

The following sound example, an excerpt from the opera *L'enfant des glaces*, was realized using the spatializer in full simulation mode (i.e. panning, distance and reverb).

Sound example 22: *L'enfant des glaces* section B3.1

Rhythm-synchronized spatialization

While working on polyphonic spatializers (i.e. multiple sources with independent spatial trajectories), I stumbled accidentally on a technique for working with rhythm that I have since employed in several works. The technique requires a multi-channel diffusion system, and is intended to operate on music with a definite beat or groove. The idea is to take two or more rhythmically related sound sources (“voices”), and move them independently about the listening space, “in time” with the music. The independent movement can be chosen at random for each voice, or it can be based on (slaved to) the movement of one of the voices (master).

Both approaches are shown below:

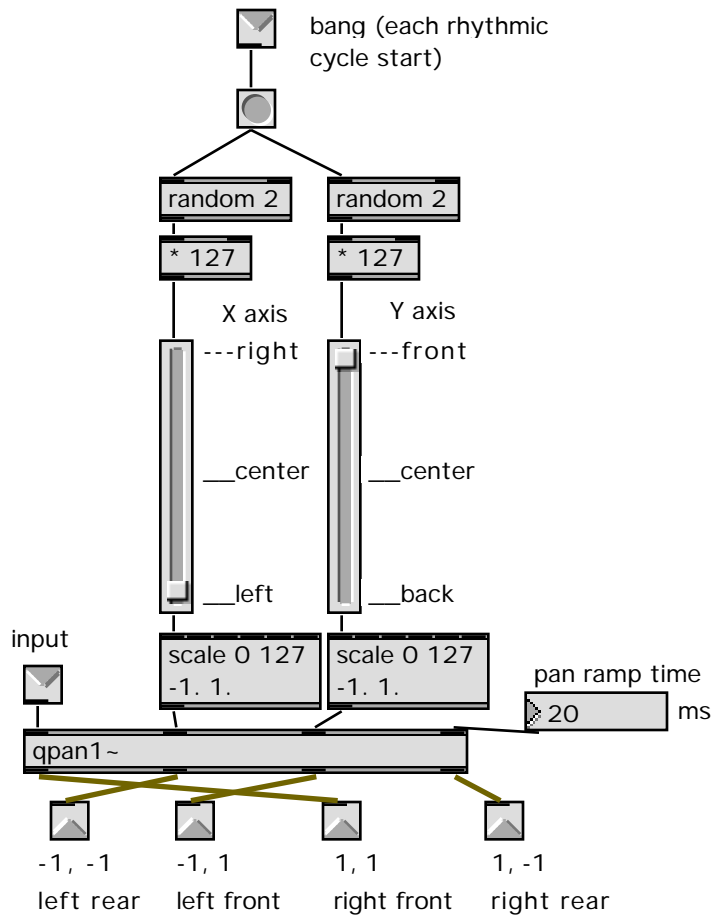


Figure 58: Beat-triggered panning via algorithm

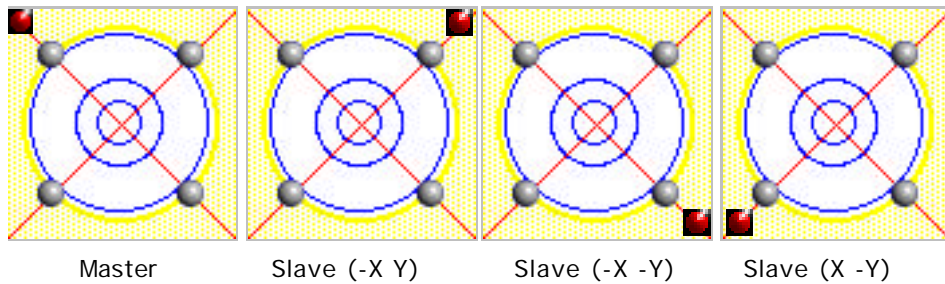


Figure 59: Beat-triggered panning via Master/Slave model

Note: The square cursor represents the sound's position in the quad field (gray spheres).

The movement occurs quickly (at once) at key points (e.g. downbeats) in the rhythmic structure of the music. The result is a multi-voice rhythmic groove that is articulated spatially. It's probably appropriate to mention Henry Brandt⁵¹ at this point, since I was inspired by his thinking about spatial rendering of musical elements—specifically harmony. He suggested that spatially distributing the independent voices of chord changed (as in, negated) the chord's harmonic effectiveness. In the case of rhythm, I have found that the “rhythm-synched” independent spatialization of the voices in a groove tends to reinforce the groove, adding to it an additional dimension of rhythmic interest. The following sound example has been mixed down to stereo (since multi-channel audio file formats are still largely unsupported); nonetheless, even in a stereo field, the effect described above is still appreciable.

Sound example 23: L'enfant des glaces outtake1

⁵¹ informal conversations with Brandt, during his lecture visit to CalArts in 1984.

CONCLUSION

Human endeavors are dynamic processes, which undergo stages of development and evolution. The early days of flying have long since passed; aviation is now a mature, sophisticated and well-documented/understood field. I regard the practice of electroacoustic music as being at a pivotal point in its evolution having recently crossed an important threshold, thanks in large part to the computer music revolution (especially real-time digital audio and music technology). This threshold, both quantitative and qualitative, is marked by the size of the community of creators using electroacoustic techniques in the production of their music, and the resulting new aesthetics and approaches to music production that have emerged. In live electroacoustic performance the “pivotal point” is even more apparent. Since the arrival of real-time digital signal processing on the desktop, composers can make use of advanced techniques (such as real-time frequency domain signal processing,) while providing a reasonable degree of concert-time stability/security and portability for their pieces. We, composers and performers, still “crash and burn” in concert, but the reliability level of live electroacoustic systems and their greatly simplified logistics have tremendously advanced their use in live performance. At the same time, the sophisticated techniques ushered in by this revolution (discussed in the chapter Techniques, p. 2), have enkindled inspiration and excitement in composers of live electroacoustic music.

In my experience making music over the years, there have been serious shortcomings in live electroacoustic music practice, which I have confronted and struggled with. To an important extent, these shortcomings have been due to, or at least followed from, limitations in the underlying real-time technologies of recording, synthesis/signal processing, event detection, control and, to a certain degree, graphical music representation (i.e. the score). For example, my piece *Hommage to Hydjik Nelmen* called for an 8-second variable delay. In 1984, this was not an obvious thing to manage, let alone provide. The technology was not available; I had to build an “it-will-never-work”⁵² contraption involving tape recorders, pulleys and fishing weights. And the safety margin for use in performance was extremely narrow. While it worked very well (and was quite amusing), the piece was played only once. It was such an ordeal to realize that I decided to suspend working in that direction. Without a doubt, advances in real-time digital systems—particularly regarding cost and computing power, have been key in reducing or eliminating these kind of practical limitations, freeing up my energies and imagination for matters of greater compositional importance.

⁵² Barry Schrader’s comment to me as his student.

In this text, I have discussed my approach to composition, and the underlying techniques, methods and aesthetics. In many ways I too am at a pivotal point with my music. I have learned (and relearned) to maintain a balance between innovation and practicability, ideally producing works that are interesting *and* performable (even in the composer's absence). As with other disciplines such as desktop publishing, advances in both technology and practice have changed the nature of the way we create. During the course of my doctoral work, I have developed new approaches to creating music. Concerns with real-time signal processing and control techniques for live performance are no longer priorities. My focus is ever greater on engendering discovery in the process of creating music. The strong emphasis on the "instrument" throughout this text reflects my longtime preoccupation with composing with instruments for live performance. In my approach, the instrument is present in the composition process at its very inception—as a source of musical expression, wonderment, curiosity, discovery, and inspiration. The approach to electroacoustic instrument design presented in this text provides for the invention and reinvention of new instruments, which themselves can lead to new musical expressions, from which compositions eventually emerge—which, in turn, inspire the invention of still newer instruments. This iterative process embodies composition and performance in a continuous cycle of creation and expression, where the underlying techniques serve simply as a vehicle for the imagination and as a source for the musical spirit.

BIBLIOGRAPHY

J. M. Adrien, *Étude de Structures Complexes Vibrantes, Applications à la Synthèse par Modèles Physiques* (Ph.D. Thesis, Université de Paris VI, Paris, 1988)

J. M. Adrien, "The Missing Link: Modal Synthesis" in *Representations of Musical Signals* edited by G. De Poli, A. Picalli & C. Roads (Cambridge, Massachusetts: MIT Press, 1991)

J. Chowning, "The Synthesis of Complex Audio Spectra by means of Frequency Modulation" (1973) *Journal of the Acoustical Society of America*, 21(7), pp. 526-534

J. Chowning, "The Simulation of Moving Sound Sources" preprint of the *Audio Engineering Society 38th Convention, 1970, May 4-7*, New York

J. Chowning & D. Bristow, *FM Theory and Applications, By Musicians for Musicians*, (Japan: Yamaha Music Foundation, 1986)

P. Depalle & X. Rodet, "De la voix aux instruments" (1993) *Cahiers de l'IRCAM*, No. 2, pp. 121-139

M. Dolson, "The phase vocoder: a tutorial" (1986) *Computer Music Journal*, 10(4), pp. 14-27

G. Eckel, F. Iovino & R. Caussé, "Sound synthesis by physical modelling with Modalys", ISMA: International Symposium of Music Acoustics, July 1995

F. Iovino, R. Caussé & R. Dudas, "Recent work around Modalys and Modal Synthesis" (1997) *International Computer Music Conference*, Greece, September 1997

J. Gordon & J. Strawn, "An introduction to the phase vocoder" (1987) *Proceedings, CCRMA*, Department of Music, Stanford University, February 1987

R. Haddad & T. Parsons, *Digital Signal Processing, Theory, Applications and Hardware* (New York: Computer Science Press, 1991)

- T. Jehan, A. Freed & R. Dudas, “Musical Applications of New Filter Extensions to Max/MSP” (1999) *Proceedings of the International Computer Music Conference*, Beijing, China
- G. Kendall & L. Martens, “Simulating the Cues of Spatial Hearing in Natural Environments” (1984) *Proceedings of the 1984 International Computer Music Conference*, (William Buxton: 1985)
- E. Lindemann, M. Starkier & F. Dechelle, “The Architecture of the IRCAM Music Workstation” (1991) *Computer Music Journal*, 15(3), pp. 41-49
- C. Lippe, “Real-time Granular Sampling Using the IRCAM Signal Processing Workstation” (1994) *Contemporary Music Review*, vol. 10, (United Kingdom, Harwood Academic Publishers) pp. 149-155
- C. Lippe, “FFT-based Resynthesis for Timbral Transformation in a Compositional Context” (1994) *Proceedings of International Colloquium in New Music Research*, Ghent, pp. 239-243
- C. Lippe, “Getting the Acoustic Parameters from a Live Performance” (1994) *Proceedings of the 3rd International Conference for Music Perception and Cognition*, Liege, pp. 328-333
- C. Lippe, “Music for piano and computer: A Description” (1997) *Information Processing Society of Japan SIG Notes*, vol. 97, no. 122, pp. 33-38
- C. Lippe & M. Puckette, “Musical Performance Using the IRCAM Workstation” (1991) *Proceedings of the 1991 International Computer Music Conference*, San Francisco: International Computer Music Association
- G. W. McNally, “Dynamic Range Control of Digital Audio Signals” (1984) *Journal of the Audio Engineering Society*, vol. 32
- F. R. Moore, “A General Model for Spatial Processing of Sounds” (1983) *Computer Music Journal*, vol. 7, no. 3, pp. 6–15

- A. Moorer & Berger, "Linear-Phase Bandsplitting: Theory and Applications" (1984) *76th AES Convention 1984*, Audio Engineering Society (preprint #2132), New York
- J. Morrison & J. M. Adrien, "Control Mechanisms in the MOSAIC Synthesis Program" (1991) *Proceedings of the 1991 International Computer Music Conference*, pp. 19-22
- J. Morrison & J. M. Adrien, "Mosaic: A Framework for Modal Synthesis" (1993) *Computer Music Journal*, vol. 17, no. 1, pp. 45-56
- R. Nieberle & M. Warstat, "Implementation of an analysis/synthesis system on a DSP56001 for general purpose sound processing" (1992) *Proceedings of the 1992 International Computer Music Conference*, San Francisco: International Computer Music Association
- M. Orth, "Interface to Architecture: Integrating Technology into the Environment in the Brain Opera" (1997) *ACM Conference*, DIS '97
- Y. Potard, P. F. Baisnée & J.-B. Barrière, "Méthodologie de synthèse du timbre: l'exemple des modèles de résonance" in *Le timbre, métaphore pour la composition* edited by J.B. Barrière (IRCAM, Paris: Christian Bourgois, 1991)
- M. Puckette, "The Patcher" (1988) *Proceedings of the 1988 International Computer Music Conference*, San Francisco: International Computer Music Association
- M. Puckette, "FTS: A Real-time Monitor for Multiprocessor Music Synthesis" (1991) *Music Conference*, San Francisco: Computer Music Association, pp. 420-429
- J. Reinhard, "Frozen in Time" (1999) *National Geographic Magazine*, vol. 196 no. 5, pp. 36-55
- D. Rokeby, "The Construction of Experience: Interface as Content" in *Digital Illusion: Entertaining the Future with High Technology* edited by Clark Dodsworth Jr. (Addison-Wesley, 1998)
- R. Rowe, *Interactive music systems: machine listening and composing* (Cambridge, Mass.: MIT Press, 1993)

- Z. Settel & C. Lippe, "Real-time Timbral Transformation: FFT-based Resynthesis" (1994), vol. 10, *Contemporary Music Review* (United Kingdom, Harwood Academic Publishers), pp. 171-179
- Z. Settel & C. Lippe, "Real-time Musical Applications using Frequency-domain Signal Processing" (1995) *IEEE ASSP Workshop Proceedings*, Mohonk, New York
- Z. Settel & C. Lippe, "Real-time Frequency-Domain Digital Signal Processing on the Desktop" (1998) *International Computer Music Conference Proceedings*, Ann Arbor, Minnesota
- J. O. Smith, "Physical modeling using digital waveguides, Part I" (1992) *Computer Music Journal*, 16(4):74-91
- I. Xenakis, *Formalized Music* (Indiana: Indiana University Press, 1971; revised ed. Stuyvesant: Pendragon Press, 1992)
- D. Wessel, "Timbre Space as a Musical Control Structure" (1979) *Computer Music Journal*, vol. 3, no. 2
- D. Wessel, D. Bristow & Z. Settel, "Control of Phrasing and Articulation in Synthesis" (1987) *International Computer Music Conference Proceedings*
- D. Zicarelli, "Music Technology As A Form of Parasite" (1992) *International Computer Music Conference Proceedings*, pp. 69-72

DISCOGRAPHY

- J. Cage, 1946-8, *Sonatas and Interludes for Prepared Piano*, WERGO WER60150-50
- J. Coltrane, 1960, "Giant Steps", *Giant Steps*, ATLANTIC B000003489
- J. Hunt, 1979, "Lattice, Transform Stream", *Lattice*, CRI CD713
- J. Hunt, 1979, "Chimanzzi, Bitom", *Ground*, OD OD disk #9
- C. Lippe, 1993, "Music for Sextet and ISPW", *Computer Music Series vol. 21*, CENTAUR CRC 2255
- C. Lippe, 1992, "Music for Clarinet and ISPW", *Computer Music Series vol. 24*, CENTAUR CRC 2310
- S. Martirano, 1992, "SATBehind Demo", *Computer Music Series vol. 22*, CENTAUR CRC 2266
- H. Partch, 1963-4, "And on the seventh day, petals fell in petalumacri", *The Music of Harry Partch*, CD 7000
- T. Rai, 1992, "Three Inventions", *Computer Music Series vol. 21*, CENTAUR CRC 2255
- Z. Settel, 1990, "Skeewit Chupp", *Electro Clips*, DIFFUSION IMEDIA IMED9004
- Z. Settel, 1993, "Hok Pwah", *Computer Music Series vol. 23*, CENTAUR CRC 2302
- M. Subotnick, 1968, "The Wild Bull", *Silver Apples on the Moon/The Wild Bull*, WERGO WER20352 1994
- M. Subotnick, 1981, "Ascent Into Air", *CalArts Twentieth-Century Players* conducted by Stephen Mosko, NONESUCH 78020-1
- A. Vinao, 1985, "Son Entero", WERGO WER 2019-50

APPENDIXES

Appendix 1: Score for the electroacoustic opera *L'enfant des glaces*

Appendix 2: CD #1 of sound examples #1 through #31

Appendix 3: CD #2 recording of the electroacoustic opera *L'enfant des glaces*

Appendix 4: Libretto source texts and translations, of Nerval and Quevedo

LIBRETTO SOURCE TEXTS AND TRANSLATIONS OF Nerval AND QUEVEDO

French

“ À travers les nuages chassés par le vent, je vis plusieurs lunes qui passaient avec une grande rapidité. Je pensai que la terre était sortie de son orbite et qu'elle errait dans le firmament comme un vaisseau démâté se rapprochant ou s'éloignant des étoiles qui grandissaient ou diminuaient ”

- Nerval-

“ Hier s'en est allé, demain n'est pas encore;

Aujourd'hui passe sans le moindre répit:

Je suis un Il Fut, Il Sera, Il est, bien las. ”

Quevedo –

*Pular*⁵³

“ Hak koudé doulè dhè hendou doun woni è radha dè, mi yi i leb bi wondhi è fedjudè tin na tin na. Mi dho sikkou dè hari leydi din iwi ka horè lawol mou oun è hari hi n'di djindoudè ka kam bou wa si tawi kö lana kalfè bhadotoka ma wothotoka kodè wonnou dhè è diandhoudhè ma fandhoudhè. ”

-Nerval-

“ Hanki fedji, djangö fèdjali tawö;

Handè no fedjoudè wa si tawi kö fousse:

Kö mi, won nö, wo naï, kö ö, ronki mödja. ”

-Quevedo-

*Armenian*⁵⁴

“ Ampéri midjiov vor kaminé vogsoum, yes téssa zanazan loussinège voge antsnoumèin chat arague. Metatsomeign yes vorg yerkirné itchel irg ouretsrits yèw mi navé dartsel arants kaymi vorg taparoumé astrérov moténalov nerantz yev erranalov voroumk mezanomein yev pokranoum. ”

-Nerval-

“ Yéréka tcheka vare derr tchi batsvats;

Antsnoumé aillsora arants hanguesti:

Hokissé vourr ka, vorg yérélé

Yiev vorr kalini havitian houknatz. ”

-Quevedo-

⁵³ translation by Mamadou Cellou Barry

⁵⁴ translation by Anaït Khemroien

Finnish⁵⁵

“ Tuulin puhaltamien pilvien läpi, näin monet kuut, jotka kulivat erittäin nopeasti. Luulin, että maa poistui kiertoradalta ja vaelti taivaassa kuin mastoton laiva, tavan takaa keskellä tähtiä, jotka kasvoivat tai vähetivät. ”

-Nerval-

“ Eilinen lähti, huomina ei ole vielä;

Tämä päivä vierähtää pysähtymättä:

Olen menneisyys, tulevaisuus, nykyhetki, melko raukea. ”

-Quevedo-

Japanese⁵⁶

“ Kase ni chitta kumoma ni taishita ikioi de tonde yuku ikutsumo no tsuki o mita. Chikyuu wa kidou o hazure, hobashira no oreta fune yoroshiku, ookiku nattari chiisaku nattari suru hoshiboshi ni chikazuki, mata toozakari tsutsu oozora o samayotte iru no dewa nai ka to omo eta. ”

-Nerval-

“ Kinou wa sari, ashita wa mada ko nai.

Kyou to ieba sewashi naku sugite shimau :

Watashi wa kako ni nokosare, mirai o machi, genzai ni tatazumu.

Unzari suru koto da. ”

-Quevedo-

Persian⁵⁷

“ Bar farâzé abre ha bar bâd rafté, aghmâri ké bé sor at migozarand didamme. Engâr ké zamine aze madârache gardide va sargardâne dar âsémân einé yek kachtie bi dakal ké bé sêtârégâne dour va nazdik michode. ”

-Nerval-

“ Dirouz rafte, fardâ hanouz nayâmadé;

Emrouz bé zoudi migozaré:

Manne ghable, baad, alân, hastam, khasté. ”

-Quevedo-

⁵⁵ translation by Tadashi Tokieda

⁵⁶ translation by Tadashi Tokieda

⁵⁷ translation by Maneli Pirzadeh

Russian⁵⁸

“ Tchérez ablaka ganimouillè viétram, ya vidziel loun beskagnéchnest pralitayouchir mima vagnavieniè oka. Kazales chtow zimlia sachla se orbité i bloujdayet po nièbosvodov karablom bizz matchti vestichiaias i prachiaias so zvieudami ovelitchivaya ir i oumingchaya. ”

-Nerval-

“ Vtchiéra ouchlow a zaftra niè nastâlce;

I gnignèschni gniè oumalitz razbieg:

Ya yést péchinka ve mirazdagniè

I imya mniè oustavchi tchièlœviek. ”

-Quevedo-

⁵⁸ translation by Anaït Khemroien