



GABRIEL RIMOLDI DE LIMA

MODELAGEM INTERATIVA APLICADA À SÍNTESE E
ESPACIALIZAÇÃO NO DOMÍNIO MICROTEMPORAL

Campinas
2013



UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE ARTES

GABRIEL RIMOLDI DE LIMA

MODELAGEM INTERATIVA APLICADA À SÍNTESE E ESPACIALIZAÇÃO NO DOMÍNIO MICROTEMPORAL

Dissertação apresentada ao Instituto de Artes da
Universidade Estadual de Campinas, como exigência
parcial para obtenção do título de Mestre em Música.

Área de concentração: Fundamentos Teóricos.

Orientador: Prof. Dr. Adolfo Maia Jr.

Este exemplar corresponde à versão final de
Dissertação defendida pelo aluno Gabriel Rimoldi de
Lima, e orientada pelo Prof. Dr. Adolfo Maia Jr.

Campinas
2013

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Artes
Eliane do Nascimento Chagas Mateus - CRB 8/1350

R468m Rimoldi, Gabriel, 1987-
Modelagem Interativa aplicada à Síntese e Espacialização Sonora no Domínio Microtemporal / Gabriel Rimoldi de Lima. – Campinas, SP : [s.n.], 2013.

Orientador: Adolfo Maia Júnior.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Artes.

1. Espacialização Sonora. 2. Síntese Granular. 3. Música por Computador. 4. Meios Eletrônicos (Música). I. Maia Junior, Adolfo, 1957-. II. Universidade Estadual de Campinas. Instituto de Artes. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Interactive Modeling for Sound Synthesis and Spatialization in Microtemporal Domain

Palavras-chave em inglês:

Sound Spatialization

Granular Synthesis

Computer Composition

Electronic Devices (Music)

Área de concentração: Fundamentos Teóricos

Titulação: Mestre em Música

Banca examinadora:

Adolfo Maia Júnior [Orientador]

Jônatas Manzolli

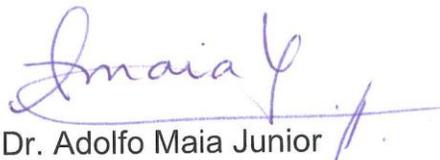
Maurício Alves Loureiro

Data de defesa: 28-08-2013

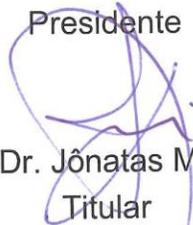
Programa de Pós-Graduação: Música

Instituto de Artes
Comissão de Pós-Graduação

Defesa de Dissertação de Mestrado em Música, apresentada pelo Mestrando Gabriel Rimoldi De Lima - RA 114515 como parte dos requisitos para a obtenção do título de Mestre, perante a Banca Examinadora:



Prof. Dr. Adolfo Maia Junior
Presidente



Prof. Dr. Jônatas Manzolli
Titular



Prof. Dr. Mauricio Alves Loureiro
Titular

RESUMO

RIMOLDI, G. (2013). *Modelagem Interativa aplicada à Síntese e Espacialização no Domínio Microtemporal*. Dissertação de Mestrado, Instituto de Artes, UNICAMP, Campinas.

Este trabalho apresenta um estudo sobre modelos interativos aplicados à síntese e espacialização sonora em tempo real. Mais especificamente, aborda sobre o desenvolvimento de métodos matemáticos e algorítmicos para geração e controle espacial de fluxos granulares através de parâmetros sonoros extraídos do sinal sonoro dos grãos. Abordamos inicialmente aspectos teóricos que fundamentam o campo da pesquisa proposta e, em seguida, apresentamos um modelo computacional desenvolvido para controle espacial de fluxos granulares. Discutimos primariamente acerca de elementos conceituais sobre a espacialização sonora, relacionados à escuta espacial, contextos históricos da espacialização na música, técnicas de auralização e sistemas de espacialização em tempo real e, em seguida, apresentamos um estudo sobre o domínio microtemporal como modelo de representação musical, abordando aspectos referentes à análise, síntese sonora e composição musical. Com base nestes estudos, desenvolvemos um software em Pure Data (PD) para espacialização de sons granulares em tempo real que denominamos GTSpa – *Granular Timbre Spatialization*. No modelo elaborado, um vetor de características extraído através de descritores acústicos é atribuído a cada grão, que por sua vez é virtualmente localizado na espacialização sonora em função do mesmo. Através de três distintas propostas interativas, criamos módulos interativos em que o usuário pode manipular diferentes propriedades da espacialização e da síntese sonora. Propomos ainda dois modelos de distribuição espacial automática dos grãos sonoros, baseando-se em estratégias adaptativas do sistema em função da interação proposta pelo usuário. Por último, apresentamos algumas aplicações artísticas que utilizam o sistema GTSpa como suporte para criação e que, deste modo, servem como mecanismo de avaliação das ferramentas desenvolvidas.

Palavras-chave: Espacialização Sonora; Síntese Granular; Sistemas Interativos; Música Eletroacústica;

ABSTRACT

RIMOLDI, G. (2013). *Interactive Modeling for Synthesis and Spatialization in Microtemporal Domain*. Dissertação de Mestrado, Instituto de Artes, UNICAMP, Campinas.

This dissertation presents a study on interactive models applied to the real time synthesis and sound spatialization. More specifically addresses on the development of mathematical and algorithmic methods for generation and spatial control of granular flows through similarities measures between the grains. Initially approached theoretical aspects underlying the field of proposed research and then present a computational model developed for spatial control of

granular flows. We discuss primarily about conceptual elements on the sound spatialization, related to spatial hearing, historical contexts of spatial music, techniques and systems for real time auralization, and then present a study of microtemporal representation as support for analysis, sound synthesis and music composition. Based on these studies, we developed a software in Pure Data (PD) for real time granular spatialization called GTSpat - Granular Timbre spatialization. In this model, one feature vector extracted by audio features are assigned to each grain, which is virtually located in the spatialization as a function thereof. Via three distinct interactive proposals, we create modules that the user can manipulate the different properties of sound synthesis and spatialization. We also propose two models of the automatic spatial distribution of the grains, based on adaptive strategies of the system depending on the proposed user interaction. Finally, we present some artistic applications using the system for creating GTSpat as support and, thus, serve as a mechanism for assessing the developed tools.

Keywords: Sound Spatialization; Granular Synthesis; Interactive Systems; Electroacoustic Music;

SUMÁRIO

1. INTRODUÇÃO	1
1.1. Objetivos	2
1.2. Premissas	3
1.3. Estrutura da Dissertação	3
2. ESPACIALIZAÇÃO SONORA: CONCEITOS E CONTEXTOS	5
2.1. Escuta Espacial	5
2.2. Atributos Espaciais	7
2.3. Espacialização na Música	10
2.4. Técnicas de Auralização	21
2.5. Sistemas de Espacialização em Tempo Real.....	28
3. MICROSSOM: ANÁLISE, SÍNTESE E COMPOSIÇÃO	33
3.1. Precedentes Históricos	34
3.2. A Teoria de Gabor	35
3.3. Síntese Granular Analógica	38
3.4. Síntese Granular Digital	41
3.5. Síntese Concatenativa	43
3.6. Microcomposição	44
3.7. Espacialização de Sons Granulares	50
3.8. Considerações	53
4. O SISTEMA GTSPAT	55
4.1. Módulo de Análise dos Segmentos Sonoros	56
4.2. Módulo de Síntese e Espacialização	61
4.3. Módulos de Interação	63
4.3.1. GTSpa <i>t Scenes</i>	64
4.3.2. GTSpa <i>t Navigation</i>	65
4.3.3. GTSpa <i>t Trajectories</i>	67
4.4. Controle Adaptativo da Distribuição dos Grãos	70
4.4.1. Controle Adaptativo do Vetor de Dados	71
4.4.2. Controle Adaptativo Através de Ressíntese	77
5. APLICAÇÕES DO SISTEMA GTSPAT	81
5.1. <i>interferência</i> (2012)	81
6. CONSIDERAÇÕES FINAIS.....	91
7. BIBLIOGRAFIA	95
8. APÊNDICE	99
8.1. Conceitos e Fórmulas dos Descritores utilizados no Sistema GTSpa <i>t</i>	99
8.2. Mensagens Processadas pelo Sistema GTSpa <i>t</i> (via comandos MIDI ou OSC)	102

*“Marco Polo descreve uma ponte, pedra por pedra.
– Mas qual é a pedra que sustenta a ponte?– perguntou Kublai Khan.
– A ponte não é sustentada por esta ou aquela pedra – responde Marco –, mas pela curva que estas formam.
Kublai Khan permanece em silêncio, refletindo. Depois acrescenta:
– Por que falar das pedras? Só o arco me interessa.
Polo responde:
– Sem pedras o arco não existe”*

(CALVINO, I. As Cidades Invisíveis, 1990, p.72)

AGRADECIMENTOS

Ao professor Adolfo Maia, pela paciência, confiança e enorme competência dedicada a este trabalho.

Ao professor Jônatas Manzolli pelo seu contagiante entusiasmo pela música que em tantos momentos contribuiu para a consolidação deste trabalho.

Ao professor José Augusto Mannis pelas ricas contribuições trazidas durante o exame de qualificação.

Ao Instituto de Artes (IA) e ao Núcleo Interdisciplinar de Comunicação Sonora (NICS) pelo apoio institucional e acadêmico.

À CAPES e FAPESP pelo financiamento desta pesquisa.

Aos colegas de Pós-Graduação, em especial Adriano Monteiro, Fernando Falci, Tiago de Mello, Felipe Castellani, Igor Maia, Ivan Simura e Clayton Mamedes pelas discussões, contribuições e incentivos prestados direta e indiretamente a este trabalho.

Ao Stephan Schaub, pela sempre prontidão em dividir seus conhecimentos.

Ao Thiago de Mello e todo pessoal do NME pela disponibilidade e abertura.

Ao Bernardo, Pedro e Lucas pela fraterna convivência e por terem sido o lar que precisei durante esses anos. Às queridas amigas Andreia, Ana Paula, Sheyla e Poliana pela bela amizade cultivada e pelo “cuidado de mãe” com que cada uma, à sua maneira, me acolheu.

Aos também queridos amigos de Uberlândia, que não deixaram a distância diminuir nossa amizade, em especial Fernando, Daniele, Gaspar, Jean, Paulo Agenor e Fabio.

À Luciana, por todo incentivo e carinho.

Aos meus familiares, especialmente meus irmãos Rodrigo e Ana Carolina, pela inspiração pessoal, afetiva e profissional que muito me amparam. Ao Arthur e Rejaine por cuidarem daqueles que tanto amo.

Ao meu companheiro Hugo Cezar, o psicólogo mais *expert* sobre música eletroacústica, que esteve tão presente nesta empreitada, em tantas outras que já tentei e quiçá em todas as outras que ainda tentarei.

Aos meus pais, engrenagem mestra de todo meu afeto.

LISTA DE FIGURAS

Figura 1: Incidência de uma fonte sonora posicionada a uma distância r , ângulo azimutal φ e ângulo de elevação θ em relação ao ouvinte.	6
Figura 2: Basílica de São Marcos, Veneza, Itália.	11
Figura 3: Stockhausen manipulando o sistema de mesa giratória utilizado na composição de <i>Kontakte</i> (1960)	18
Figura 4: pavilhão alemão da EXPO 70, em Osaka, Japão, coordenado por K. Stockhausen	19
Figura 5: Curvas de Lissajous no plano de espacialização de <i>Turenas</i> (1972), de J. Chowning.	20
Figura 6: Sistema Vetorial de Panorama por Amplitude (SVPA).	23
Figura 7: Sistema Soundfield SPS422B para captura de campo sonoro em <i>B-format</i> (detalhe para a capsula interna do microfone)	25
Figura 8: Interface gráfica do Spatialisateur, do IRCAM para Max/MSP.	28
Figura 9: Software <i>Zirkonium</i> , da ZKM.	29
Figura 10: Interface gráfica do ViMiC (<i>Virtual Microphone Control</i>) para Max/MSP	30
Figura 11: Espaço <i>Tempo-Frequência</i> de Gabor	37
Figura 12: <i>Pithoprakta</i> (1955-56), comp. 52-59, rascunho do próprio compositor.	39
Figura 13: Sequência de telas (<i>screens</i>) utilizadas para controle microestrutural de <i>Analogique B</i> (1959), de I. Xenakis, com valores discretos de frequência (Δf), intensidade (Δg) e duração (Δt).	40
Figura 14: Fluxograma de Síntese Concatenativa	44
Figura 15: Sonograma de <i>Pictor Alpha</i> (2003), de Curtis Roads.	45
Figura 16: Sonograma de fragmento (1'7" a 1'48") de <i>Volt Air, Part II</i> (2003), de Curtis Roads.	46
Figura 17: Sonograma de <i>Wings of Nike</i> (1988), <i>II - Ein kleines Scherzo</i> , de Barry Truax (frequências em distribuição logarítmica).	47
Figura 18: Detalhe da visualização do sistema de espacialização granular de Wilson (2008). As esferas representam os <i>boids</i> e os cubos os alto-falantes	51
Figura 19: Sistema <i>EVOGrain</i> , desenvolvido por Souza e Maia (2011).	52
Figura 20: Sistema <i>CataRT</i> , do IRCAM, desenvolvido em MAX/MSP por Schwarz e colaboradores (2005)	53
Figura 21: Arquitetura geral do sistema GTSPat	56
Figura 22: Módulo de análise do sistema GTSPat	58
Figura 23: Envoltórias de amplitude do sistema GTSPat (em vermelho), com valores de k , α e β pertencentes ao intervalo $[0,1]$: senoidal (a), gaussiana (b), rexpodec (c), expodec (d), retângulas (e) e ADSR (f).	60
Figura 24: Interface do Módulo de análise do sistema GTSPat	61
Figura 25: Submódulo de análise em tempo real do sistema GTSPat.	61
Figura 26: Módulo de Síntese e Espacialização do Sistema GTSPat	63

Figura 27: Interface de visualização dos grãos sonoros e objetos da cena espacial no GTSpac <i>Scenes</i> ..	64
Figura 28: Interface de controle do GTSpac <i>Scenes</i>	65
Figura 29: Espaço de Navegação e Navegador (quadrado azul) do sistema GTSpac <i>Navigation</i>	67
Figura 30: Interface de controle do sistema GTSpac <i>Navigation</i>	67
Figura 31: Representação vetorial das possíveis trajetórias a partir do ponto (x_0, y_0) comparadas à trajetória "alvo" (em vermelho).	69
Figura 32: Interface de controle do módulo GTS Trajectories.	70
Figura 33: Diagrama de interação no submódulo adaptativo pelo controle automático do vetor de características do sistema GSTpat <i>Navigation</i>	72
Figura 34: Distribuição normalizada dos dados extraídos dos descritores centroide espectral (1), variância espectral (2), obliquidade espectral (3), curtose espectral (4), taxa de cruzamento por zero (5) e energia média (6).....	75
Figura 35: Exemplo de distribuição de posições no eixo X (1) e eixo Y (2) do navegador bidimensional do sistema GTSpac <i>Navigation</i> durante 594 disparos.....	76
Figura 36: Distribuição inicial dos grãos (primeiro quadro), trajeto percorrido pelo navegador bidimensional (segundo quadro) e redistribuição dos grãos através de adaptação do sistema GTSpac (terceiro quadro).	77
Figura 37: Diagrama de interação do submódulo de adaptação por ressíntese do GSTpat <i>Navigation</i> ..	78
Figura 38: Distância entre os grãos selecionados para ressíntese e grãos com alto percentual de disparo no sistema GTSpac <i>Navigation</i>	79
Figura 39: Processo de síntese cruzada do módulo adaptativo do sistema GTSpac.	80
Figura 40: Fluxograma de informações (áudio e dados) da obra <i>interferência</i> (2012)	81
Figura 41: Mapas de distribuição espacial dos grãos sonoros das amostras (1-4) ao longo da peça <i>interferência</i> (2012), subdividida em seis momentos (a-f).	87
Figura 42: Espectrograma da parte eletrônica (4 canais) de performance de <i>interferência</i> (2012).....	89

LISTA DE TABELAS

Tabela 1: Atributos espaciais de acordo com Rumsey (2002) e Kendall (2010).....	8
Tabela 2: Atributos de percepção espacial de acordo com Berg e Rumsey (2003).....	9
Tabela 3: Classificação geral de configurações espaciais em Música (Trochimezyk, 2001).	10
Tabela 4: Presença da espacialização em algumas obras anteriores ao século XX	12
Tabela 5: Presença da espacialização em algumas obras instrumentais ao século XX	15
Tabela 6: Número de canais necessários por ordem para implementação do sistema Ambisonics	26
Tabela 7: Equações de codificação Ambisonics até 3ª ordem (FARIA, 2005).	26
Tabela 8: Taxonomia de modelos de síntese sonora digital, baseada em (TOLONEN <i>et al</i> , 1998)	33
Tabela 9: Média, grau de dispersão e grau de assimetria na distribuição dos dados extraídos de cada descritor acústico.....	75
Tabela 10: Média, grau de dispersão e grau de assimetria na distribuição de posições nos eixos X e Y do navegador bidimensional do sistema GTSpac	76
Tabela 11: Distâncias entre os dados de cada descritor e a dados da trajetória espacial do navegador para cada eixo da espacialização	76
Tabela 12: Definição de parâmetros de análise das amostras sonoras utilizadas em <i>interferência</i> (2012)	82
Tabela 13: Comandos enviados ao sistema GTSpac via interface MIDI	83
Tabela 14: Comandos enviados ao sistema GTSpac através dos pedais de expressão da interface MIDI	83
Tabela 15: Partitura Algorítmica utilizada no módulo GTS <i>Scenes</i> em <i>interferência</i> (2012).	84
Tabela 16: Lista de instruções da Partitura Algorítmica dadas ao improvisador em <i>interferência</i> (2012)	88

Capítulo 1

INTRODUÇÃO

Este trabalho apresenta um estudo sobre modelos interativos de síntese e espacialização sonora. Abordamos métodos matemáticos e algorítmicos para controle, em tempo real, de síntese sonora e espacialização, baseando-se em representações do material sonoro cujos elementos básicos têm a duração da ordem de 10 a 100 milissegundos. Dizemos então que temos uma representação em nível microtemporal ou simplesmente *microssons*.

Nossa proposta de pesquisa constitui-se de duas partes que se complementam. Na primeira, teórica e descritiva, pretendemos trazer para compositores e músicos interessados na criação com suporte eletroacústico uma contribuição sobre modelos formais e abordagens artísticas emergentes que incorporem a espacialização como elemento interativo, levantando aspectos técnicos e estéticos relacionados à mesma. Ainda nesta primeira parte, apresentamos uma revisão sobre métodos de síntese, análise e composição baseados em representações do microssom. Mais especificamente, abordaremos diferentes métodos utilizados na síntese granular, apresentando ainda o estado da arte sobre os modelos formais de espacialização sonora dos grãos.

A segunda parte de nossa proposta trata da criação e desenvolvimento de modelos computacionais que, baseando-se no escopo teórico levantado na primeira parte deste trabalho, explorem algumas propriedades da espacialização como elemento interativo. Com isso, apresentamos o desenvolvimento de um sistema interativo aplicável à espacialização sonora de fluxos granulares, o qual denominamos GTSpat – *Granular Timbre Spatialization*. Neste sistema, a distribuição dos grãos na espacialização sonora resulta de uma função adotada, que compara os grãos através de características objetivas relacionadas a aspectos múltiplos da percepção do timbre. Estas características são representadas por parâmetros reais, isto é, medidas obtidas da extração de informações do sinal sonoro digitalizado, conhecidas na literatura como *Descritores Acústicos*.

1.1. Motivação

Mesmo que implícita a qualquer de forma produção sonora, a espacialização nem sempre é concebida e/ou percebida como um primeiro plano nos contextos musicais de criação e performance. A formalização do espaço como parâmetro estrutural nestes contextos é um fenômeno relativamente recente. Com algumas exceções, sua aplicação mais sistemática remonta-se a partir de meados do século XX, quando diversos compositores abordaram alguns atributos do espaço como primeiro plano no desdobramento do discurso musical (TROCHIMCZYK, 2001). Com a música eletroacústica, as possibilidades de exploração artística do espaço tornaram-se ainda mais amplas, sobretudo pela extensão dos domínios até então estipulados pela gestualidade da música instrumental.

Em nosso contexto, observamos a prática comum dentre os compositores em considerar a espacialização uma prática *a posteriori* das decisões composicionais relativas a outros aspectos da obra musical. Apesar de sua interdependência em relação aos outros processos composicionais, a espacialização é em geral uma prática abordada de modo menos sistemático. Nas últimas duas décadas, o grande desenvolvimento e a maior acessibilidade das tecnologias voltadas para a difusão multicanal, tem renovado o interesse de compositores e pesquisadores sobre a investigação de métodos formais para o controle espacial (OTONDO, 2008).

Nossa principal motivação foi desenvolver ferramentas composicionais que tratassem a espacialização sonora como um processo conjugado, ou par a par, com outros aspectos da criação musical. Deste modo, investigamos possíveis soluções para o controle espacial dos sons através de medidas associadas ao timbre, relacionando assim as qualidades timbrística de cada segmento sonoro e sua localização espacial na difusão sonora em suporte eletrônico digital.

1.2. Objetivos

Este trabalho tem como objetivos fundamentais:

- a. Avaliar o estado da arte em espacialização sonora aplicada a sistemas musicais interativos e, de maneira particular, as possíveis abordagens de controle da síntese e espacialização em nível microtemporal.
- b. Construir um modelo interativo de espacialização aplicado à síntese granular, levando em conta o conteúdo informacional dos grãos mensurado através de parâmetros extraídos de suas formas de onda digitalizadas e definindo uma conveniente função distância.

1.3. Premissas

Neste trabalho adotamos as seguintes premissas:

a. *Utilização de softwares de código aberto:*

Para o desenvolvimento computacional deste trabalho escolhemos a utilização de plataformas livres e/ou abertas. Elegemos assim a plataforma Pure Data (PD), por tratar-se de um ambiente aberto de programação, pela sua ampla difusão nos meios acadêmicos e artísticos, bem como pela plasticidade da plataforma, que permite a construção de ferramentas de processamento digital de áudio em tempo real com baixa latência de resposta. A portabilidade para diversos sistemas operacionais e a fácil comunicação com diversos outros softwares, através dos protocolos OSC (*Open Sound Control*) e MIDI (*Musical Instrument Digital Interface*), foram também fatores preponderantes na escolha do software utilizado.

b. *Interface intuitiva:*

De maneira geral, o propósito deste trabalho é o desenvolvimento de ferramentas digitais que sejam úteis tanto para compositores e instrumentistas, interessados na criação musical através de suporte tecnológico, como também para o público em geral, tornando possível uma exploração mais intuitiva das propostas interativas apresentadas pelo sistema. Estabelecemos assim um conjunto de parâmetros que poderiam ser manipulados pelo usuário com um curto tempo de aprendizado sem que, no entanto, isto implicasse uma limitação do sistema, ou seja, a prática continuada deveria assim resultar em possíveis melhorias no uso das ferramentas propostas.

1.4. Estrutura da Dissertação

Esta dissertação está estruturada em seis partes, conforme descrito abaixo.

CAPÍTULO 1 – *Introdução*: apresentação das propostas, premissas e objetivos de pesquisa e introdução dos conceitos que possibilitaram a construção do modelo computacional desenvolvido.

CAPÍTULO 2 – *Espacialização Sonora: Conceitos e Contextos*: revisão de literatura sobre aspectos conceituais acerca da espacialização e sua recorrência em contextos musicais. Neste capítulo abordaremos os principais conceitos relacionados à “escuta espacial” e uma contextualização histórica da espacialização no contexto da música ocidental, sobretudo na música eletroacústica. Relataremos ainda sobre os principais métodos computacionais de auralização sonora e, mais detalhadamente, a

técnica Ambisonics, utilizada na implementação do presente modelo. Por último, apresentaremos um *estado da arte* sobre a espacialização sonora em tempo real.

CAPÍTULO 3 – *Microsossom: análise, síntese e composição*: uma revisão da literatura da área sobre métodos e técnicas baseadas em representações microtemporais da música. Iniciaremos este capítulo com a apresentação dos fundamentos teóricos da *Teoria Granular de Gabor* e, em seguida, relataremos sua repercussão como proposta de criação sonora nos contextos analógicos e, sobretudo, digitais. Prosseguiremos discutindo diferentes desdobramentos da poética de criação em nível microtemporal e, por último, apresentaremos alguns trabalhos recentes dedicados à espacialização de sons granulares.

CAPÍTULO 4 – *O sistema GTSpat*: apresentação do modelo de síntese e espacialização granular desenvolvido por nós. Relatamos inicialmente a arquitetura geral do modelo elaborado e, em seguida, os procedimentos utilizados em cada um de seus módulos. Por último, discutimos sobre a utilização de ferramentas adaptativas para o controle de distribuição dos grãos.

CAPÍTULO 5 – *Aplicações: sistema GTSpat*: relatamos algumas aplicações artísticas do sistema nos contextos da criação mista que avaliam as potencialidades das ferramentas computacionais desenvolvidas.

CAPÍTULO 6 – *Considerações Finais*: comentamos sobre o desenvolvimento da pesquisa, análise dos resultados e conclusões obtidas.

Esta dissertação é acompanhada ainda por um CD-ROM contendo duas pastas. Na primeira, constam os patches e sub-patches do sistema GTSpat acompanhados de um arquivo de texto com breves instruções sobre o software. Na segunda pasta, incluímos uma versão da obra *interferência* (2012), composta através de suporte do sistema GTSpat e que constitui, portanto, uma experiência de teste e validação do modelo computacional como parte dos resultados da pesquisa desenvolvida.

Capítulo 2

ESPACIALIZAÇÃO SONORA: CONCEITOS E CONTEXTOS

2.1. Escuta Espacial

A informação sonora é espacial por natureza. Constituído de ondas de pressão longitudinais (na direção do movimento) que se deslocam através do ar, o som expande-se a partir de sua fonte, decaindo sua intensidade à medida que se espalha através do espaço em redor. A partir do momento que é emitida pela sua fonte, a onda sonora em geral é modificada pela sua interação mecânica com corpos e outros meios físicos contidos neste espaço, por processos de reflexão, difração ou absorção, chegando por fim ao aparelho auditivo humano. Este, por sua vez, envia informações sinápticas ao cérebro, que é especializado em reconhecer não apenas informações qualitativas referentes à fonte emissora, tais como o tipo de fonte, modo de emissão e sua localização espacial, como também informações referentes ao espaço ou meio em que a onda se propagou (MALHAM, 1998).

A percepção espacial do som baseia-se na interpretação de um conjunto de referenciais extraídos do campo sonoro ao redor do ouvinte. Esses referenciais advêm, majoritariamente, da diferença entre as informações captadas pelos ouvidos direito e esquerdo e também pelas reflexões e absorções ao redor da cabeça, torso e pino.

Em 1907, o físico John William Strutt (Lord Rayleigh) propôs uma teoria de localização espacial do som baseado em dois índices que exprimem diferenças do modo de recepção sonora dos ouvidos esquerdo e direito. De acordo com o modelo proposto por Rayleigh, conhecido como *Teoria Duplex*, sons puros (senoidais) podem ser espacialmente localizados através de dois indicadores, a Diferença Interaural de Tempo (*Inter-aural Level Difference* – ILD) e a Diferença Interaural de Intensidade (*Inter-aural Time Difference* – ITD). Estes indicadores são chamados de índices biaurais (BLAUERT, 1996; MOORE, 2003).

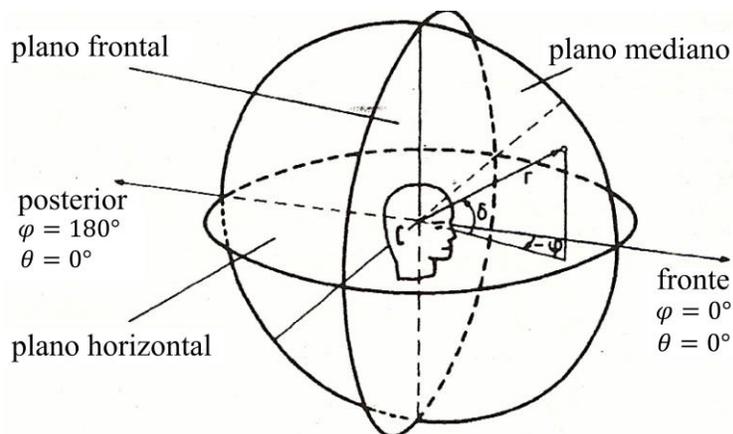


Figura 1: Incidência de uma fonte sonora posicionada a uma distância r , ângulo azimutal φ e ângulo de elevação θ em relação ao ouvinte.¹

Tal qual sugere o nome, a ILD indica a atenuação do sinal entre os ouvidos, ocasionada tanto pela diferença de intensidade de pressão do som resultante da diferença de distância entre cada um dos ouvidos e a fonte sonora, bem como pela atenuação do som no ouvido contralateral à fonte pela cabeça humana (*Head Shadow*). O efeito de atenuação da cabeça é relevante apenas para frequências acima de 300Hz, dado que a difração ocorre para comprimentos de onda que se aproximam do tamanho da cabeça humana. A ITD indica a diferença de fase (tempo de chegada) da onda entre os ouvidos. A distância entre os dois ouvidos faz com que o som alcance o ouvido ipsilateral antes que o ouvido contralateral.

Os índices biaurais são, no entanto, insuficientes na precisão de localização espacial do som, sobretudo na percepção de elevação e na diferenciação de sons oriundos da região frontal e posterior. Neste caso, índices monoaurais que modelam a refração e reflexão do som ocasionada pela cabeça, torso e, sobretudo, pela pina são elementos essenciais na localização espacial do som.

A pina, porção externa e visível do aparelho auditivo, funciona como espécie de funil, pelo qual as ondas sonoras se refletem até alcançar o ouvido médio. A maneira com que a pina reflete as ondas sonoras afeta a variedade de estímulos sonoros que chegará até o ouvido médio, como uma espécie de filtro em que certas frequências serão atenuadas ou amplificadas dependente da direção em que incidem sobre o ouvido. Além da atenuação ou amplificação de frequências, o som refletido está ainda

¹ Fonte: BLAUERT, 1996: 14.

sujeito a um minúsculo atraso de tempo entre o som direto e o sinal refletido. As mudanças resultantes no perfil do espectro sonoro variam de acordo com o azimute e elevação da fonte sonora em relação à pina (BLAUERT, 1996; WARREN, 1999).

Como referido anteriormente, os índices de ILD e ITD não são suficientes para uma precisa localização em todo o plano azimutal. A razão disto deve-se à ineficiência da distinção entre as localizações nos hemisférios frontal e posterior que apresentam a mesma distância entre os receptores direito e esquerdo (por exemplo, a distinção entre uma fonte localizada a 45° ou a 135° em relação ao receptor). Para solucionar esta ambiguidade, conhecida como “cone de confusão”, uma pequena inclinação da cabeça pode modificar tanto a amplitude quanto a fase das ondas que chegam a cada ouvido, traduzindo assim uma informação de orientação vertical como mecanismo de localização horizontal (MILLS, 1972).

O conjunto de mudanças espectrais sofridas entre a emissão sonora e sua recepção pelo canal auditivo, incluindo as distorções causadas pela presença da cabeça, do torso e da pina, é comumente denominado de funções de transferência relacionadas à cabeça (*Head-Related Transfer Function – HRTF*). A modelagem destas funções pode ser uma tarefa extremamente complexa, visto que cada pessoa tem proporções físicas próprias que afetarão o modo de recepção sonora. Um banco de dados de HRTFs pode ser obtido, no entanto, através de medições realizadas com o auxílio de um manequim², dentro de uma câmara anecóica, com a fonte sonora em alguns pontos numa esfera ao redor do mesmo.

Uma análise mais aprofundada sobre a teoria da escuta espacial pode ser encontrada em (BLAUERT, 1996; MOORE, 2003). No tópico a seguir, discutiremos acerca de alguns atributos espaciais do som relevantes para o estudo da espacialização.

2.2. Atributos Espaciais

Na concepção de um sistema de espacialização sonora, devem ser considerados diversos aspectos qualitativos referentes à recepção do som produzido por alto-falantes. Diversas pesquisas em engenharia de áudio dedicam-se ao estudo da percepção de atributos espaciais do som, o que tem conduzindo a um conjunto de terminologias que auxiliam o desenvolvimento da arquitetura de sistemas

² Na literatura este manequim é comumente conhecido como *Knowles Electronics Manikin for Acoustic Research* (KEMAR).

de espacialização (ZACHAROV e KOIVUNIEMI, 2001). Dentre os principais atributos estabelecidos encontra-se a *espacialidade*.

O conceito de espacialidade foi introduzido em meados da década de 70, associado à ideia de percepção de imersão do ouvinte em relação aos sons do ambiente. Referido também na literatura como *impressão espacial*, a espacialidade é atualmente definida como um atributo subjetivo que agrega um conjunto de características, dentre elas a Largura Aparente da Fonte e a Envolvência do Ouvinte.

A Largura Aparente da Fonte (*Apparent Source Width – ASW*), está relacionada às reflexões laterais iniciais (primeiras reflexões) e refere-se à maneira que a largura da fonte sonora são percebidos pelo ouvinte. Esta medida é também associada à percepção de localização espacial do som. Um alto grau de ASW indica que a fonte sonora é ampla e alongada sobre o plano horizontal ou azimutal, o que, por consequência, implica uma maior dificuldade na precisão de localização da fonte sonora. Inversamente, um baixo grau de ASW indica uma fonte sonora mais estreita e, portanto, implica uma melhor localização da mesma.

O grau de Envolvência do Ouvinte (*Listener Envelopment – LEV*), relaciona-se às reflexões tardias que constituem a reverberação. Quando o som reverberado aparenta chegar de todas as direções de maneira homogênea, o grau de envolvimento é elevado, o que significa que a energia sonora reverberada é bastante difusa.

Rumsey (2002) propõe uma taxonomia bastante concisa de atributos espaciais que se subdividem em duas categorias, a primeira de atributos *dimensionais* e a segunda de atributos *imersivos*. Da primeira categoria têm-se os atributos de *distância*, *largura* e *profundidade*, enquanto que da segunda tem-se *presença* e *envolvimento*. Abordando especificamente a taxonomia de Rumsey ao contexto da música eletroacústica, Kendall (2010) acrescenta ainda à primeira categoria os atributos de *direção* e *altura*, como podemos observar na Tabela 1.

Tabela 1: Atributos espaciais de acordo com Rumsey (2002) e Kendall (2010)

Dimensional	Imersivo
<i>Direção</i>	<i>Presença</i>
<i>Distância</i>	<i>Envolvimento</i>
<i>Extensão (largura, profundidade e altura)</i>	

Em seu artigo, Rumsey prossegue ainda afirmando que tais atributos podem ser instanciados em distintos níveis, que vão do *micro* ao *macro*. Para isso, propõe um paradigma da espacialização estruturada em quatro níveis, sendo eles *fonte sonora*, *conjunto*, *sala* e *cena*. Para o autor, estes níveis se organizam de maneira hierárquica, ou seja, as *fontes sonoras* são agrupadas em *conjuntos*, dispostos por sua vez numa determinada *sala*, a qual por sua vez é instanciada sob uma determinada *cena*.

Apesar de tratar de maneira mais específica da espacialidade voltada a gravações convencionais, os conceitos tratados por Rumsey podem facilmente se adequar ao contexto da música eletroacústica consonante aos conceitos abordados pela literatura especializada. Berg e Rumsey (2003) apresentam ainda um conjunto de 13 atributos espaciais que foram por eles utilizados em testes, conforme apresentado na Tabela 2.

Tabela 2: Atributos de percepção espacial de acordo com Berg e Rumsey (2003)

Atributo	Descrição
Naturalidade	Similaridade do som produzido pelos alto-falantes em relação à sua emissão natural.
Presença	A experiência de estar no mesmo ambiente acústico que a fonte sonora.
Preferência	Se o som como um todo lhe agrada.
Conteúdo de baixas frequências	Nível sonoro de frequências graves.
Largura do conjunto	Largura percebida do conjunto de todas as fontes sonoras. O ângulo ocupado pelo conjunto desconsiderando-se a reverberação da sala.
Largura da fonte individual	Largura percebida de uma fonte sonora individual.
Localização	A facilidade de se perceber a posição de uma fonte sonora.
Distância da fonte	Distância da fonte sonora.
Envoltório da fonte	A extensão com que uma fonte sonora envolve o ouvinte, desconsiderando a reverberação.
Largura da sala	A largura ou ângulo ocupado pelos sons oriundos das reflexões das fontes sonoras, desconsiderando o som direto.
Tamanho da sala	Percepção das dimensões da sala.
Nível sonoro da sala	Nível de sons gerados na sala como resultante da ação da fonte sonora, sem se considerar o som direto.
Envoltória da sala	A extensão com que o som vindo das reflexões da fonte sonora na sala envolve o ouvinte.

2.3. Espacialização na música

A espacialização é um aspecto inerente de qualquer prática musical, dado que para qualquer experiência sonora são necessários minimamente uma fonte sonora, um ouvinte e o ambiente onde ambos estão situados. A ênfase dada à perspectiva aural pela tradição musical, da qual resulta a histórica legitimação das estruturas formadas por alturas e durações como identidade única e inequívoca da obra musical (GOODMAN, 1976), coloca a concepção do espaço físico como um elemento secundário na composição. Podemos, no entanto, observar diversas recorrências ao longo da história que propõem uma mudança de perspectiva, apontando a projeção espacial, a direção e localização sonora como importantes elementos da estrutura musical. A esta prática, denomina-se música espacializada (*spatial music*). Trochimczyk (2001) apresenta uma classificação de tipologias de espacialização aplicadas à música, conforme Tabela 3.

Tabela 3: Classificação geral de configurações espaciais em Música (Trochimczyk, 2001).

Ambiente acústico	Espaço fechado da sala de concerto Espaço fechado de qualquer tipo Ar livre (diferentes planos acústicos) Espaço variável (performers e audiência móveis) Espaço privado, virtual (fones de ouvido)
Tipos de sons espaciais	Reais (fontes vocais e/ou instrumentais) Virtuais (fontes eletroacústicas) Mistos (ambas as fontes sonoras)
Categoria de mobilidade	Performers e audiência estáticos Performers móveis e audiência estática Performers estáticos e audiência móvel Performers e audiência móveis

2.3.1. De Gabrieli a Mahler: precedentes históricos

A espacialização sonora é uma prática extremamente antiga. As primeiras referências na história ocidental acerca da prática musical espacializada situam-se ainda no período da Idade Média³, sobretudo com prática da Salmodia Antifonal. Nesta, os salmos eram cantados em alternância de duas

³ Embora a prática de se cantar salmos em atos públicos seja extremamente antiga, possivelmente desde o século IV, os registros mais antigos desta encontram-se nos séculos IX e X. Para o leitor interessado, consultar: PSALMODY, LATIN. In: The New Harvard Dictionary of Music, ed. Don Michael Randel. Cambridge MA and London: The Belknap Press of Harvard University Press, 1986.

metades de um mesmo coro, sendo comum que cada metade se posicionasse uma de frente para a outra, localizadas em lados opostos do altar.

Em Veneza, em meados do século XVI, a construção de dois órgãos posicionados nas galerias laterais da Basílica de São Marcos (Figura 2) foi um importante propulsor para o desenvolvimento da escrita policoral renascentista. Seu mestre de capela, o compositor flamengo Adrian Willaert (c.1490-1562) instituiu a técnica dos *cori spezzati* (coros espaçados), que tratava de posicionar dois ou mais corais em regiões diferentes da Basílica, propiciando um diálogo vívido entre a polifonia da escrita musical e o espaço arquitetônico da igreja. Esta técnica foi bastante difundida entre os compositores da Escola Veneziana, principalmente por Giovanni Gabrieli (c.1557-1612). Em suas obras, Gabrieli inovou pela utilização de diferentes combinações de vozes e grupos de instrumentos para cada coral, em formações quase sempre heterogêneas, reforçando assim o contraste entre as partes da polifonia espacialmente distribuídas.

Pela emergência do estilo pré-clássico na primeira metade do século XVIII, a escrita policoral passou a ser associada ao *stilo antico* e a prática antifonal recorrente apenas como uma espécie de anacronismo, sobretudo na composição formas litúrgicas corais. Com a expansão dos corpos orquestrais e a tendência homofônica da música no período Clássico e Romântico, a espacialização manteve-se num plano secundário, reclusa na maioria das vezes à adequação do modelo arquitetônico italiano dos teatros da época, no qual o público experimentava um único bloco sonoro proveniente de uma mesma direção (ZVONAR, 2010).



Figura 2: Basílica de São Marcos, Veneza, Itália⁴.

⁴ Extraída de: <http://www.wga.hu>. Acesso em 20 fev. 2013.

Discretas utilizações do espaço podem ser posteriormente encontradas em obras de Berlioz, Verdi e Mahler, com inserções de efeitos teatrais específicos, mantendo ainda a informação frontal da projeção sonora como atrativo principal. Na Tabela 4 encontram-se algumas obras anteriores ao século XX que apresentam algum tipo de espacialização definida.

Tabela 4: Presença da espacialização em algumas obras anteriores ao século XX

Compositor	Obra	Ano	Descrição
C. W. Gluck (1714-1787)	<i>Orfée et Euridice</i>	1762	Uma segunda orquestra “derrière le théâtre” composto por cordas e harpa surge em algumas cenas do 2º ato.
W. A. Mozart (1756-1791)	<i>Serenade n° 8 (Notturmo), K.286</i>	1777	Efeito de eco entre as quatro orquestras.
	<i>Don Giovanni, K527</i>	1787	Três grupos tocam separadamente (fosso da orquestra, no palco e atrás do palco)
L.van Beethoven (1770-1827)	<i>Fidelio: Abertura Leonora, Op. 5</i>	1805	Trompete tocado fora do palco
H. Berlioz (1803-1869)	<i>Symphonie Fantastique: 3º movm, “scène aus champs”</i>	1830	Oboé e Corne Inglês atrás do palco tocam um tema pastoral num diálogo de pergunta e resposta.
	<i>Grande messe des morts, Op. 5, “Requiem”: Tuba Mirum</i>	1837	Quatro grupo de metais separados localizados nos quatro pontos cardeais anunciam a entrada do coro.
G. Verdi (1813-1901)	<i>Requiem</i>	1874	Metais fora do palco
G. Mahler (1860-1911)	<i>Symphony n° 2: V</i>	1893/94	Em duas pequenas seções, uma pequena banda militar (4 trompetes, 4 trompas, címbalo, triângulo e tímpano) tocam em lugar distante do palco.

2.3.2. Espacialização na música do século XX

Frente às mudanças estéticas da música do final do século XIX, surgem novas propostas de exploração do espaço físico como elemento composicional. A espacialização ganha progressiva notoriedade na música do século XX, equiparando-se em diversos momentos a outros parâmetros da composição musical, tais como o ritmo e a harmonia.

O compositor experimentalista americano Charles Ives (1874-1954) foi um importante precursor neste contexto. Ives concebe a superposição de diferentes camadas sonoras distribuídas no espaço usando diferentes grupos de instrumentistas tocando simultânea e independentemente. É recorrente em sua música a presença de grupos funcionalmente distintos e fisicamente separados, cada qual executando uma música “autossuficiente”, seja sob os aspectos harmônicos, rítmicos ou melódicos (ZVONAR, 2010). Em sua conhecida obra *The Unanswered Question* (1930-35), Ives recorre à separação espacial como estratégia de diferenciação entre as três camadas sonoras que acontecem simultaneamente. O trompete solista posiciona-se à frente do palco, enquanto que o quarteto de sopros, que tenta exaustivamente responder à “pergunta musical” executada pelo solista, posiciona-se ao fundo do palco e o quarteto de cordas, que executa durante toda a peça uma harmonia coral contínua e bastante esparsa, deve se posicionar preferencialmente fora do palco. O contraste musical destas distintas camadas, peculiares em seus timbres, texturas e harmonias, é reforçado por meio da distribuição espacial dos instrumentistas. É interessante observamos que, neste caso, o compositor aborda o espaço não apenas como um articulador da complexidade musical, mas também como um elemento expressivo de uma descrição metamusical. As experimentações de Ives serão de grande influência sobre os que podemos denominar como “compositores espaciais”, em especial Henry Brant e Pierre Boulez (HARLEY, 1994).

Henry Brant (1913-2008) refere-se ao uso da espacialização sonora como estratégia para “desembaraçar” texturas densas. A proposta criativa de Brant fundamenta-se em sobreposições de múltiplas texturas no tempo, mas não no espaço. Além da clarificação de texturas densas, a dispersão espacial na musica de Brant tem por consequência o aumento da amplitude total e das ressonâncias sonoras do espaço físico.

Tal abordagem encontra paralelos no pensamento musical de John Cage (1912-1992), que concebe a separação espacial como ênfase do que ele denomina por *coexistência de dissimilaridades*, eventos não relacionados justapostos temporalmente. Para Cage (1961), a fusão do som foi um aspecto essencial ao longo da história da música ocidental. Tradicionalmente, os músicos de um conjunto devem estar posicionados o mais próximo possível uns dos outros, de modo que suas ações produzam “um único objeto no tempo” (p.39). Em direção oposta, a proposta estética de Cage emerge da presença de simultaneidades que não se reduzem a uma unidade sonora. Cada performance, única e irrepitível, resulta do encontro entre diversos elementos autônomos que se desdobram no tempo e no espaço. A

separação espacial dos músicos é um método conveniente a fim de garantir a independência entre os materiais musicais. Cage diz que:

Esta separação permite que os sons se propaguem de seus próprios centros e interpenetrem de forma tal que não seja obstruído pelas convenções da teoria e harmonia europeia sobre as relações e interferências dos sons. (...) a ação dos músicos é resultante de um processo, nenhuma harmoniosa fusão dos sons é essencial. (CAGE, 1961: 39-40, *tradução nossa*)⁵

Tanto para Brant como para Cage a segregação espacial de fluxos sonoros distintos atua como articuladora da complexidade musical de simultaneidades, ou nas palavras de Brant, uma maneira de “tornar a complexidade inteligível” (BRANT apud HARLEY, 1997). Harley (1997) observa que a complexidade na música de Brant emerge não de operações composicionais complexas, mas da justaposição de materiais distintos e aparentemente simples espacialmente separados. Enquanto a “consonância” entre os materiais sonoros separados espacialmente garantiu certa coesão à experiência renascentista, em Brant e Cage observamos uma relação mais imbricada entre materiais dissonantes entre si e a sua espacialização, estendendo assim as possibilidades do uso do espaço como um elemento efetivo de percepção, conduzindo assim a uma escuta musical mais geral e flexível.

A compreensão da modelagem sonora na música eletroacústica, principalmente a partir do surgimento dos primeiros sistemas estereofônicos, possibilitou a emergência de uma nova concepção espacial com a qual os compositores podem planejar trajetórias espaciais sonoras, através da manipulação dos envelopes dinâmicos do som. Esta técnica encontrou rapidamente paralelos na escritura instrumental, como podemos observar em *Gruppen*, para três orquestras e *Carré*, para quatro orquestras e coros, ambas de Stockhausen, e *Terretektorh*, para oitenta e oito músicos espalhados pela audiência, de Iannis Xenakis.

Oposto ao experimentalismo de Brant, o compositor alemão Karlheinz Stockhausen (1928-2007) argumenta em seu texto *Musik im Raum* (1959) que a única característica espacial que merece destaque especial na prática composicional é a direção, uma vez que pode ser estruturada serialmente. Em *Gruppen*, escrita pouca antes do referido texto, Stockhausen estabelece uma escala de localidades obtidas pela divisão do círculo em partes simétricas. Do conjunto de segmentos espaciais estabelecidos,

⁵ “This separation allows the sounds to issue from their own centers and to interpenetrate in a way which is not obstructed by the conventions of European harmony and theory about relationships and interferences of sounds. (...) the action of the players is productive of a process, no harmonious fusion of sound is essential” (CAGE, 1961, p.39-40).

o compositor sistematiza a distribuição espacial do som através de operações seriais, que indicam a localização e o movimento do som no espaço.

Na Tabela 5, apresentamos um panorama de obras do século XX que exploram a espacialização como um elemento estrutural da obra.

Tabela 5: Presença da espacialização em algumas obras instrumentais ao século XX

Compositor	Obra	Ano	Descrição
K. Stockhausen (1928-2007)	<i>Gruppen</i>	1955-57	Para três orquestras dispostas nas posições frontal e laterais da audiência
	<i>Carrè</i>	1959-60	Para quatro orquestras e coros dispostos em forma de quadrado ao redor da audiência
I. Xenakis (1921/22-2001)	<i>Terretektorh</i>	1965/66	Para 88 músicos dispersos entre a audiência
	<i>Peserphassa</i>	1969	Para seis percussionistas posicionados em círculo ao redor da audiência
L. Berio (1925-2003)	<i>Cyrclès</i>	1960	Primeira obra em que o solista se move de um lugar ao outro no decorrer da obra.
H. Brant (1913-2008)	<i>Millennium II</i>	1954	10 trombones e 10 trompetes localizados nas paredes laterais da audiência, 8 trompas, 2 tubas e 4 percussionistas no palco e uma voz localizada num ponto distante dos dois outros grupos.
	<i>Windjammer</i>	1969	Para quinteto de sopro. Inclui planos de trajetos que os músicos devem percorrer no decorrer da obra.
T. Musgrave (n.1928)	<i>Concerto para clarinete e orquestra</i>	1969	O solista é requerido atravessar por toda a orquestra, intermitentemente ativando e/ou participando de pequenos grupos de concertino.
P. Boulez (n. 1925)	<i>Ritual</i>	1974	Orquestra dividida em oito grupos espacialmente localizados.
P. Murray Schafer (n. 1933)	<i>Credo</i> , da obra <i>Apocalypsis</i>	1976/77	Para 12 coros localizados em círculo.
Claire Polin (1926-1995)	<i>A Klockwrk Diurnal</i>	1977	Para saxofone, trompa e fagote. Os músicos sentam-se sobre cadeiras giratórias que devem se mover silenciosamente durante a obra.

2.3.3. Espacialização na Música Eletroacústica

A emergência da música realizada a partir de suporte eletrônico propôs novos paradigmas no que concerne às relações espaço-temporais de produção, recepção e compreensão da arte. No final do século XIX e início do subsequente, as invenções do rádio, da telefonia e dos processos eletromecânicos de gravação ampliaram os territórios da escuta, que até então era orientada exclusivamente pela recepção das fontes sonoras fisicamente localizadas no espaço.

Com o desenvolvimento das técnicas de edição em estúdio e a criação dos primeiros sintetizadores em meados do século XX, os alto-falantes tornam-se mediadores imprescindíveis entre os processos elétricos e os sons por estes gerados. O desvio de foco do gesto instrumental para uma atitude acusmática⁶, pela qual a atenção multimodal da percepção (sobretudo condicionada pela visão) é direcionada exclusivamente para o elemento acústico, exigiu por parte do ouvinte o estabelecimento de novas estratégias de escuta.

Para o compositor e teórico neozelandês Denis Smalley, a espacialização tem papel fundamental na estruturação da escuta da música eletroacústica, na medida em que é capaz de “salientar as propriedades do som inerentes em sua espectromorfologia e relações estruturais”⁷ (SMALLEY, 1991: 123, *tradução nossa*). Smalley aborda ainda uma diversidade de espaços que devem ser considerados em relação à música eletroacústica, que vão desde o *espaço composto*, ou seja, o espaço idealizado pelo compositor para a obra, até o *espaço de escuta*, no qual esta é experienciada. Para o autor, a experiência espacial da música advém da sobreposição destes espaços, ora *consonantes*, ora *dissonantes* entre si.

Chion (1988) também concebe a articulação do espaço na música eletroacústica (ou como referido pelo autor, na “arte dos sons fixados”) através de dois níveis: interno e externo. O primeiro refere-se à espacialidade intrínseca da espectromorfologia, às qualidades espaciais presentes na gravação sonora, enquanto que o segundo refere-se à maneira com que o som é difundido no espaço durante a performance.

⁶ O termo acusmática, derivado da palavra grega *akousmatikós* (disposto a ouvir), faz referência à Pitágoras, quem discursava por detrás de uma cortina para impedir que seus alunos o vissem e assim, se distraíssem com informações visuais. A primeira referência do termo ao contexto musical foi feito por Schaeffer em seu *Tratado dos Objetos Musicais*. Bayle (1993) propôs a utilização do termo a referir-se à música eletrônica produzida em suporte fixo diferenciando-se da música eletrônica produzida através de sintetizadores, mas que era interpretada em forma presencial.

⁷ “(...) a means of enhancing the sounding properties inherent in spectro-morphologies and structural relations” (SMALLEY, 1991:123)

Embora a gravação magnética tenha sido inventada entre as décadas de 20 e 30, o acesso a esta tecnologia manteve-se bastante restrito até o final da Segunda Guerra Mundial. Somente em 1948, Pierre Schaeffer (1910-1995), engenheiro da *Radiodiffusion-Télévision Française* (RTF), apresentou as primeiras obras criadas a partir de gravações em fita magnética. Em parceria com o jovem compositor Pierre Henry, Schaeffer criou uma série de obras para fita magnética, dentre elas a conhecida *Symphonie pour un Homme Seul* (1950).

Gravações multicanais tornaram-se disponíveis e diversos compositores da RTF passaram a utilizar até cinco fitas monofônicas roteadas para um sistema de quatro canais de alto-falantes. Um mecanismo construído por Jacques Poullin (por sugestão de Schaeffer) denominado *potentiomètre d'espace*⁸ (1951) era utilizado para controlar a difusão sonora através de bobinas de indução. Poullin descreve duas maneiras distintas de projeção sonora a partir deste mecanismo, *statisches Relief* (revelos estáticos) e *kinematisches Relief* (relevos dinâmicos). A primeira refere-se à projeção simultânea de distintas partes da composição em diferentes pontos no espaço, enquanto que a segunda envolve a manipulação de trajetórias do som entre os alto-falantes (HARLEY, 1994). A experiência da escuta de imagens sonoras movimentando-se pelo espaço durante os concertos de Música Concreta no início da década de 50 repercutirá de maneira significativa nos modelos de espacialização utilizados por Boulez e Stockhausen⁹.

A proposta de tratamento da direção e dos movimentos espaciais do som como um elemento estruturante da obra musical influenciou também na composição eletroacústica de Stockhausen. Em sua obra *Gesang der Jungelinge* (1956), o compositor reitera o tratamento espacial abordado anteriormente na obra orquestral *Gruppen*, evidenciando o espaço como “uma nova dimensão da experiência musical” (STOCKHAUSEN, 1961, p.68, *tradução nossa*). A obra, em geral considerada a primeira composta em multicanal, utilizava uma máquina de quatro canais mais uma segunda máquina monofônica para a geração do quinto canal. Na concepção original do compositor, este quinto canal deveria ser projetado através de um alto-falante suspenso sobre a audiência, o que, por motivos de logística, não ocorreu. Posteriormente à estreia da obra, Stockhausen remixou a obra para quatro canais.

⁸ Em algumas referências este é também denominado como *pupitre d'espace* (CHADABE, 1997, p.31-32).

⁹ Vale ressaltar que ambos os compositores tiveram contato com o estúdio da RTF entre 1952-53.

Provavelmente a primeira composição de fato quadrifônica foi *Kontakte* (1960), obra também composta por Stockhausen. Os canais foram distribuídos em quatro alto-falantes dispostos em Frontal, Direito, Esquerdo e Posterior. Para criar a movimentação sonora, o compositor utilizou um sistema de mesa giratória, que consistia num alto-falante colocado acima de uma tábua giratória e rodeado por quatro microfones que captavam a rotação dos sons por meio da manipulação do próprio compositor (Figura 3).



Figura 3: Stockhausen manipulando o sistema de mesa giratória utilizado na composição de *Kontakte* (1960) ¹⁰

Entre as décadas de 60 e 70, o surgimento da arte conceitual, dos *happenings*, bem como o desenvolvimento de diversos projetos intermidiáticos que aliavam inovações tecnológicas e vanguardismo artístico, redimensionam o campo de possibilidades de utilização do espaço na composição musical. *Vortex* (1957-59) foi um dos primeiros projetos multimídia realizados pela dupla Jordan Belson e Henry Jacobs. Luzes e sons eram projetados no espaço do planetário Morrison, em São Francisco. O sistema dispunha de aproximadamente 36 alto-falantes e os movimentos de rotação eram coordenados por um console especialmente desenvolvido

¹⁰ Extraída de <http://www.stockhausen.org/>. Acesso em 20 janeiro 2013.

Em 1958, o destaque da Feira Mundial de Bruxelas foi o *Poème Electronique*, um pavilhão multimídia idealizada por Le Corbusier, encomendado da empresa holandesa Philips, com projeto arquitetônico desenvolvido por Iannis Xenakis e músicas compostas por Edgard Varèse e também por Xenakis. Ao longo de toda a superfície do parabolóide hiperbólico da construção arquitetônica, cerca de 400 alto-falantes combinados em 20 diferentes amplificadores projetavam os sons controlados por meio de uma fita de 15 canais.

A EXPO 70, feira ocorrida em Osaka, no Japão, concentrou diversas obras multimídias de compositores e artistas de vanguarda. No pavilhão japonês, Xenakis apresentava sua obra para 12 canais *Hibiki Hana Ma*, projetada em 800 falantes localizados ao redor, acima e embaixo das cadeiras da audiência. No pavilhão alemão, Stockhausen e mais 20 solistas apresentavam performances duas vezes ao dia durante 183 dias num auditório esférico de 28 metros de diâmetro. Stockhausen controlava a projeção sonora, distribuindo os sons em padrões circulares e esféricos através de um conjunto de 55 alto-falantes disposto em sete anéis concêntricos por toda a esfera (Figura 4).

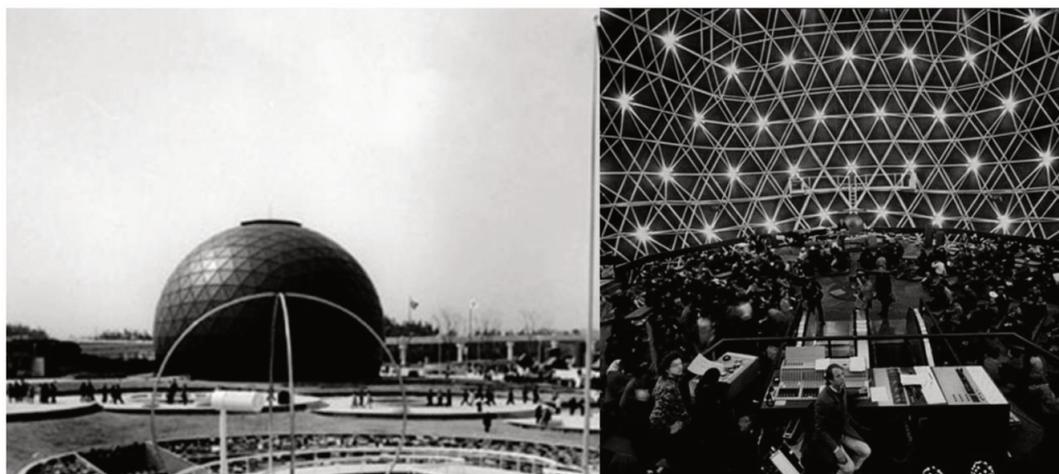


Figura 4: pavilhão alemão da EXPO 70, em Osaka, Japão, coordenado por K. Stockhausen ¹¹

Enquanto a grande maioria das obras eletrônicas era ainda produzida através de suportes analógicos, um emergente desenvolvimento da música realizada com auxílio dos computadores digitais tornava-se notável. Durante a década de 70, o compositor John Chowning (n. 1934) trabalhou no desenvolvimento de um algoritmo computacional capaz de controlar a localização horizontal, distância

¹¹ Extraída de <http://www.stockhausen.org/>. Acesso em 20 janeiro 2013.

e velocidade do som num sistema quadrifônico. O modelo apresentado por Chowning (1977) é pioneiro na aplicação de equações que simulam o efeito Doppler¹² para criar movimentos em diferentes velocidades e acelerações das trajetórias sonoras. O programa desenvolvido foi utilizado na espacialização da conhecida obra *Turenas* (1972), a primeira a utilizar a técnica de síntese por Modulação de Frequência (FM). As trajetórias espaciais realizadas na obra aproximam-se das curvas de Lissajous¹³, conforme Figura 5.

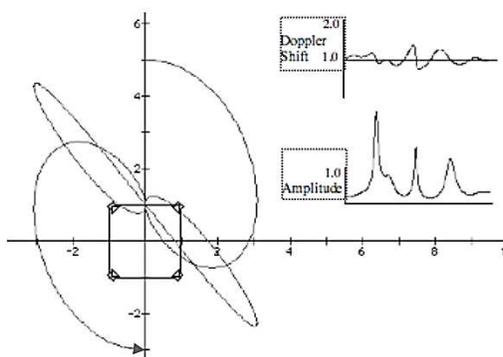


Figura 5: Curvas de Lissajous no plano de espacialização de *Turenas* (1972), de J. Chowning¹⁴.

A relação entre o comportamento do espectro sonoro e sua localização espacial é elemento de interesse na obra *Archipelago* (1980-82), do compositor Roger Reynolds (n.1934). Nesta obra, para 32 instrumentistas e *tape*, Reynolds, a partir de seus experimentos realizados no IRCAM juntamente com Thierry Lancino, dividiu o espectro do som de um oboé entre dois canais, endereçando as parciais pares para um canal e as ímpares para o outro canal da estereofonia. Discretas modulações de frequência (FM) eram adicionadas a cada um dos canais. Se a frequência moduladora fosse a mesma para ambos os canais, obtinha-se a fusão dos sinais. Se, porém, diferentes frequências modulassem cada um dos sinais, o som era dividido em dois objetos sonoros independentes no espaço (HARLEY, 1994; TOPPER et al., 2002).

¹² O efeito Doppler consiste na característica observada em sinais periódicos pela qual a frequência, percebida por um observador de uma fonte em movimento relativo, aumenta ou diminui conforme a fonte se aproxima ou se afasta do observador.

¹³ Também conhecidas por figuras de Lissajous ou curva de Bowditch, trata-se de curvas complexas harmônicas no plano cartesiano, gerados pela modulação entre duas funções senoidais em defasagem.

¹⁴ Extraída de (CHOWNING, 2011: 9);

O uso de espacialização como ferramenta de separação do espectro é também presente na obra octofônica do compositor inglês Jonathan Harvey (1939-2012), *Mortuos Plango, Vivos Voco* (1980). A composição, que tem como material subjacente a gravação dos sinos da Catedral de Winchester e a voz do filho do próprio compositor, é seccionada em oito seções. Cada seção é baseada e enunciada por uma das oito frequências mais graves das parciais harmônicas do sino gravado (HARVEY, 1981). A organização espacial dos sons desdobra-se de diversas maneiras ao longo da obra, sobretudo pela partição dos componentes espectrais do som entre os oito alto-falantes. Como apontado pelo próprio compositor, “as parciais do sino são distribuídas estatisticamente no espaço, produzindo a impressão sonora de estar dentro do sino, enquanto que a voz do menino tem a tendência de atuar como um espírito livre se movendo através da audiência” (HARVEY, 1990, *tradução nossa*)¹⁵.

Outra importante abordagem do espaço na música eletroacústica refere-se à difusão em tempo real. Diversos compositores e engenheiros têm se dedicado ao desenvolvimento de extensivos sistemas de difusão envolvendo um grande número de alto-falantes. Em contraste à espacialização fixada na composição em estúdio, a proposta destes sistemas baseia-se na difusão sonora proveniente de um número de alto-falantes, em geral heterogêneos, cada qual com suas características e resposta em frequência. Em 1974, François Bayle criou uma orquestra de alto-falantes, denominada *Acousmonium*. Sua construção consiste em 80 alto falantes de diversos tamanhos localizados em diferentes alturas e distâncias do proscênio. O *Birmingham Electro-Acoustic Sound Theatre* (BEAST), dirigido por Jonty Harrison, utiliza mais de 32 canais de alto-falantes, arranjados em pares, cada qual com características e posicionamentos especificados.

2.4. Técnicas de auralização

2.4.1. Biaural

Os métodos biaurais baseiam-se na geração de dois canais de som, um para cada ouvido, contendo informações temporais e espaciais necessárias para que o sistema auditivo perceba ou possa reconstruir um espaço acústico bi ou tridimensional. Em geral, gravações biaurais são feita a partir de dois microfones acoplados a uma cabeça de manequim (KEMAR), recriando a exata recepção de cada

¹⁵ “The bell’s partials are statically distributed in space producing the sonic impression of being inside the bell, whereas the boy’s voice has the tendency to act as a free moving spirit throughout the concert hall” (HARVEY 1990)

um dos ouvidos ¹⁶. Métodos de síntese podem também modelar a gravação binaural através do uso de *HRTFs*.

Para a reprodução sonora em sistemas binaurais faz-se necessária a utilização de fones de ouvido ou de alto-falantes muito próximos aos ouvidos (dipolo estéreo), o que torna esta técnica bastante invasiva e restrita quanto ao número de ouvintes simultâneos (apenas um). Além disto, o uso de *HRTFs* genéricas (gravação com o uso do KEMAR) empobrece a informação espacial, podendo em alguns casos confundir o sistema auditivo. Em contrapartida, o uso de *HRTFs* personalizadas oferecem resultados bastante satisfatórios e muito próximos à realidade.

Diversas pesquisas recentes têm abordado técnicas e sistemas para reprodução binaural, como também o modelamento, cálculo e processamento de *HRTF*'s. Um recente levantamento de banco de *HRTF* para melhor modelamento tridimensional das fontes sonoras foi realizado por Kim (2005), que viabilizou o desenvolvimento de um sistema adaptativo de espacialização baseada na posição arbitrária do ouvinte, apresentado por Kim, Kong e Jang, (2008). Outros experimentos sobre modelagem de *HRTF* têm sido conduzidos por Plaskota e Dobrucki (2008) e Haraszy (2012).

2.4.2. Sistema Vetorial de Panorama por Amplitude (SVPA)

O Sistema Vetorial de Panorama por Amplitude (SVPA), originalmente em inglês *Vector Based Amplitude Panning* (VBAP), consiste num método em que o mesmo sinal sonoro é distribuído com diferentes amplitudes utilizando um arranjo bi ou tridimensional de autofalantes. Trata-se de uma generalização do modelo estereofônico do panorama por amplitude (*Amplitude Panning*) que permite combinações de três ou mais alto-falantes em localizações arbitrárias (PULKKI, 2001).

¹⁶ A gravação binaural pode ser simplificada pela utilização de uma esfera com dois microfones laterais ou mesmo um disco circular posicionando um microfone de cada lado (MALHAM, 1998). Microfones binaurais em miniatura, encontrados comercialmente, possibilitam ainda que o usuário utilize sua própria cabeça como modelo para gravação ao posicionar cada microfone em um dos ouvidos.

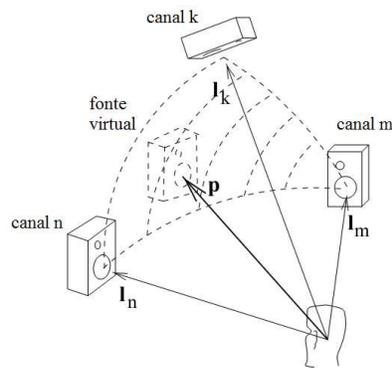


Figura 6: Sistema Vetorial de Panorama por Amplitude (SVPA)¹⁷.

Apesar da grande vantagem do livre posicionamento dos alto-falantes, o sistema possui um *sweet spot*¹⁸ bastante reduzido, visto que o ganho para cada canal é calculado a partir da posição do ouvinte. O SVPA também não leva em consideração alguns parâmetros psicoacústicos importantes, como a dependência da amplitude em função da frequência.

2.4.3. Wave Field Synthesis

A técnica Wave Field Synthesis (WFS), formalizada por Berkhout baseia-se no princípio de Huygens de reconstrução da frente de onda a partir da superposição de diversas ondas. A proposta desta técnica é sintetizar a frente de onda que é produzida por fontes primárias num ambiente acústico específico, através de fontes secundárias ou, mais especificamente, um conjunto de alto-falantes densamente distribuídos no contorno da área de audição. O campo sonoro reconstruído é válido para toda a região de reprodução, o que garante a esta técnica um *sweet spot* ampliado, abrangendo assim um maior número de ouvintes (BERKHOUT, DE VRIES e VOGEL, 1993).

A gravação a partir do WFS é realizada por meio de uma matriz de pares de microfones que circundam a fonte sonora a ser gravada. A reprodução é realizada por uma matriz idêntica do processo de gravação, substituindo o par de microfones por um par de alto-falantes. O processo de gravação pode também ser simulado computacionalmente, a partir do cálculo da velocidade e pressão da propagação das ondas de uma cena virtual até os microfones.

¹⁷ Figura extraída de (PULKKI, 2001).

¹⁸ Trata-se do ponto ótimo de projeção sonora, geralmente localizado no ponto ou região central de posicionamento dos alto-falantes.

A WFS possibilita um maior e mais preciso controle sobre o modo de propagação das ondas no meio e sobre os fenômenos de interferência entre si e entre os obstáculos, atestando assim bastante realismo na construção de cenas sonoras virtuais. Pelo grande número de sinais gerados e pelo denso conjunto de alto-falantes, esta técnica requer um alto desempenho computacional além do alto custo de equipamento.

2.4.4. Ambiphonics

Ambiphonics é uma técnica mista de gravação e reprodução que combina mecanismos de simulação tridimensional utilizados pela estereofonia tradicional, Ambisonics e técnicas binaurais. O fundamento da técnica é separar o som direto, que é reproduzido por um par de alto-falantes (denominados Ambiopolos) com cancelamento de interferência e a ambiência, que por sua vez é reproduzida por outro conjunto de alto-falantes (GLASGAL, 2004).

O sistema comporta de um a dois usuários (*sweet spot* reduzido), sendo assim voltado para o uso doméstico em *home theaters*. Seu formato de renderização pode ser incluído em mídias como SACD e DVD-Audio, o que, aliado ao baixo custo de gravação e reprodução, sugere uma possível sucessão dos sistemas estéreo e 5.1 de uso doméstico.

2.4.5. Ambisonics

O sistema Ambisonics foi desenvolvido a partir dos princípios apontados por Alan Blumlein, que no final da década de 20 desenvolveu uma técnica de gravação que utilizava dois microfones direcionais para captação em estéreo (MALHAM, 1998). Esta técnica, conhecida como “MS”, utiliza um par de microfones (*mid* e *side*), o primeiro onidirecional, captando a região esférica ao seu redor e o segundo, de padrão figura de oito, apontado lateralmente em eixo horizontal. Pela somatória e diferença dos sinais captados, ter-se-ia assim a codificação do sinal em estéreo.

Deste princípio decorreram desenvolvimentos ulteriores, sobretudo por Peter Fellgett, Peter Craver e Michael Gerzon, que foram documentados por Gerzon (1973), tendo este desenvolvido um protótipo composto por um microfone onidirecional e três microfones “figura de oito”, cada qual direcionado em um eixo. Este microfone possibilitava a gravação de campos sonoros, levando em conta a reconstituição do sinal sonoro e os parâmetros psicoacústicos de sua reprodução. Posteriormente, um microfone específico foi desenvolvido, sem que houvesse assim a necessidade de quatro microfones

para a captação, conforme podemos observar na Figura 7. Atualmente, este microfone é comercializado pela empresa Soundfield.

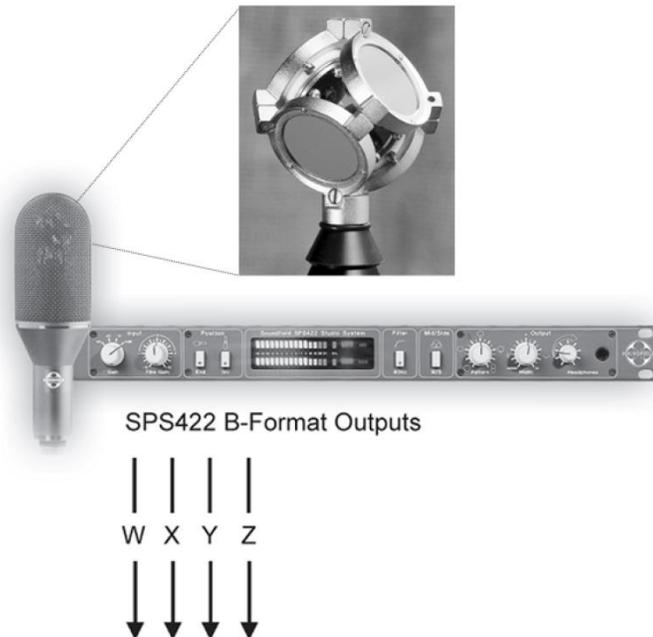


Figura 7: Sistema Soundfield SPS422B para captura de campo sonoro em *B-format* (detalhe para a capsula interna do microfone) ¹⁹.

O sinal captado é deste modo codificado para os quatro canais do sistema Ambisonics (comumente denominados por W, X, Y e Z) e enviado a um decodificador, que determina através de uma matriz de ganhos, qual a amplitude de cada sinal codificado deverá ser endereçada a cada alto-falante do sistema de reprodução. No processo de decodificação, cada alto-falante recebe uma soma ponderada de todos os canais em função de sua posição.

Ao sistema de codificação em quatro canais indica-se como de 1ª ordem, comumente conhecido como *B-format*. Sistemas de maior ordem são possíveis, conforme podemos observar na Tabela 6. Pela ausência de modelos comerciais capazes de gravar em ordens superiores, sobretudo devido à dificuldade de se montar uma capsula que comporte nove, dezesseis ou mais transdutores, a gravação através de sistemas Ambisonics de alta ordem tem se restringido ainda aos meios acadêmicos, que na

¹⁹ Extraído de www.soundfield.com/. Acesso em 17 janeiro 2013.

última década foram responsáveis pelo desenvolvimento de diversos protótipos de segunda, terceira e quarta ordem.

Tabela 6: Número de canais necessários por ordem para implementação do sistema Ambisonics

n° de canais	Canais	Ordem Ambisonics
3	WXY	1ª ordem (sem elevação)
4	WXYZ	1ª ordem (com elevação)
5	WXYUV	2ª ordem (sem elevação)
9	WXYZRSTUV	2ª ordem (com elevação)
7	WXYUVPQ	3ª ordem (sem elevação)
16	WXYZRSTUVKLMNOPQ	3ª ordem (com elevação)

A codificação do Ambisonics pode, no entanto, ser realizada por meio de síntese, pela qual um sinal monofônico é virtualmente posicionado no espaço através da manipulação dos ganhos para cada canal. Através do formalismo matemático da técnica, é possível determinarmos o ganho para cada canal através de equações relativamente simples. Deste modo, torna-se viável a implementação de sistemas Ambisonics de ordem superior através de algoritmos que emulam a captação de fontes sonoras espacialmente localizadas. Na Tabela 7, estão listadas as equações em coordenadas polares de codificação Ambisonics de até terceira ordem, onde s_i é o sinal mono não codificado da i -ésima fonte sonora, com $i = \{1, 2, 3, \dots, k\}$, φ é o ângulo azimutal e θ o ângulo de elevação que determinam a direção da mesma.

Tabela 7: Equações de codificação Ambisonics até 3ª ordem (FARIA, 2005).

Ordem	Canal	Equação
0	W	$\sum_{i=1}^k s_i \frac{1}{\sqrt{2}}$
1	X	$\sum_{i=1}^k s_i \cos \varphi_i \cos \theta_i$
	Y	$\sum_{i=1}^k s_i \sin \varphi_i \cos \theta_i$
	Z	$\sum_{i=1}^k s_i \sin \theta_i$
2	R	$\sum_{i=1}^k s_i [(3 \sin^2 \theta - 1)/2]$
	S	$\sum_{i=1}^k s_i \cos \varphi_i \cos 2\theta_i$
	T	$\sum_{i=1}^k s_i \sin \varphi_i \sin 2\theta_i$
	U	$\sum_{i=1}^k s_i \cos 2\varphi_i \cos^2 \theta_i$
	V	$\sum_{i=1}^k s_i \sin 2\varphi_i \cos^2 \theta_i$

3	K	$\sum_{i=1}^k s_i [\frac{1}{2} \text{sen}^2 \theta_i (5 \text{sen}^2 \theta_i - 3)]$
	L	$\sum_{i=1}^k s_i [\frac{11}{8} \cos \varphi_i \cos \theta_i (5 \text{sen}^2 \theta_i - 1)]$
	M	$\sum_{i=1}^k s_i [\frac{11}{8} \text{sen} \varphi_i \cos \theta_i (5 \text{sen}^2 \theta_i - 1)]$
	N	$\sum_{i=1}^k s_i \cos 2\varphi_i \text{sen} \theta_i \cos^2 \theta_i$
	O	$\sum_{i=1}^k s_i \text{sen} 2\varphi_i \text{sen} \theta_i \cos^2 \theta_i$
	P	$\sum_{i=1}^k s_i \cos 3\varphi_i \cos^2 \theta_i$
	Q	$\sum_{i=1}^k s_i \text{sen} 3\varphi_i \cos^2 \theta_i$

As equações de decodificação dependerão do número e da disposição dos alto-falantes bem como da quantidade de canais codificados. Em geral, a decodificação é feita a partir de matrizes de ganho previamente calculadas²⁰. Pela complexidade dos cálculos dessas matrizes, variando de acordo com a configuração de alto-falantes utilizada, é comum o emprego de matrizes previamente calculadas de configurações padrões de alto-falantes.

O Ambisonics apresenta grandes vantagens em relação a outras técnicas de espacialização. Uma das grandes vantagens do sistema Ambisonics é a separação dos processos de codificação e decodificação, sobretudo pela flexível portabilidade dos sistemas de reprodução, tendo em vista não ser necessário que a formação de alto-falantes seja definida previamente ao processo de codificação. Além de sua fácil implementação e custo computacional relativamente baixo, quando em altas ordens o sistema comporta um amplo *sweet spot* e é capaz de (re)criar campos sonoros com bastante realismo.

Pelas equações de codificação apresentadas na Tabela 7, podemos notar que as coordenadas polares (θ, φ) são as únicas variáveis da matriz de ganho para cada canal. Em função disto, Ambisonics representa uma abordagem bastante conveniente na determinação da informação direcional de fontes sonoras. Informações sobre a distância são, no entanto, desconsideradas em abordagens mais tradicionais desta técnica. Matematicamente, assume-se que todas as fontes sonoras encontram-se à superfície de uma esfera de raio unitário. Na literatura podemos, entretanto, encontrar diversas abordagens que associam simulação de distâncias à técnica Ambisonics (DANIEL, 2003).

²⁰ A formulação matemática das matrizes vai além do escopo deste trabalho e, portanto, não será detalhada aqui. Para o leitor interessado, consultar (FURSE, 1999), além de informações que podem ser encontradas no website “Ambisonics.Net” em <http://www.ambisonic.net>. Acesso em: 23 jul. 2013.

2.5. Sistemas de espacialização em tempo real

Neste tópico abordaremos acerca de recentes trabalhos sobre espacialização sonora em tempo real.

MidiSpace é um projeto desenvolvido pela Sony francesa durante a década de 90 destinado ao controle espacial bidimensional de arquivos gerados em protocolo MIDI. O sistema, descrito por Pachet e Delerue (1998), possui uma interface gráfica na qual o usuário pode controlar a posição de cada canal e também a posição de um *avatar* que simula a posição do ouvinte.

O projeto *Spatialisateur*, iniciado em 1991 em parceria da *Espaces Nouveaux* e o IRCAM²¹ pretendia o desenvolvimento de um processador acústico virtual útil tanto para criação musical, pós-produção ou performances ao vivo. O objeto desenvolvido para Max/MSP, *Spat~*, associa a reconstrução de localização das fontes sonoras e efeitos relacionados à sala. Diversos destes objetos podem ser associados em paralelo, possibilitando o processamento simultâneo de diversos sinais, cada qual facilmente configurado para um tipo de técnica de codificação e configuração de alto-falantes (Figura 8).

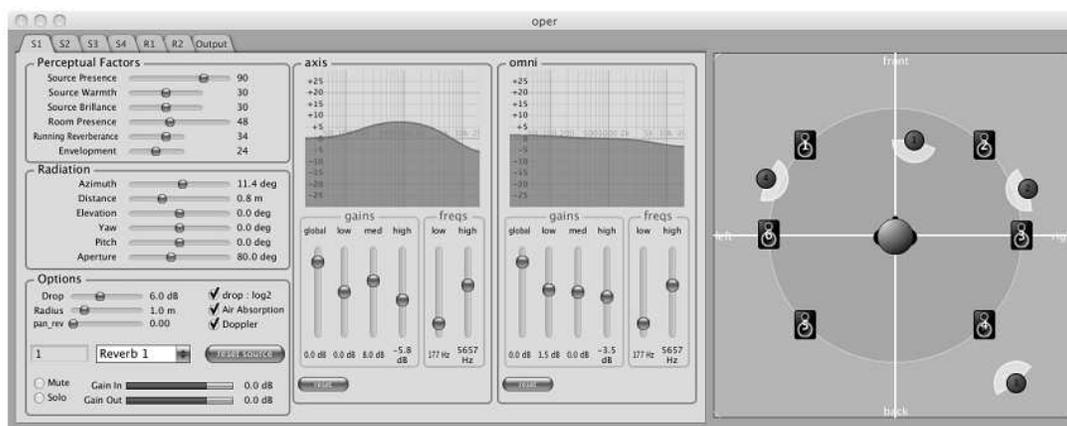


Figura 8: Interface gráfica do Spatialisateur, do IRCAM para Max/MSP²².

*Zirkonium*²³ é um software aberto de espacialização sonora desenvolvido para Mac OS, pelo Instituto de Música e Acústica da ZKM, Alemanha (Figura 9). Sua codificação espacial é realizada

²¹ <http://www.ircam.fr/>

²² Extraída de: <http://forumnet.ircam.fr>. Acesso em 10 fev. 2013.

²³ <http://www.zkm.de/zirkonium>

através de um Sistema Vetorial de Panorama por Amplitude (SVPA), disponibilizando ao usuário uma flexível configuração da quantidade e disposição dos alto-falantes. O software possibilita também a composição, armazenamento e execução de “coreografias” espaciais. Zirkonium pode também comunicar-se com outros softwares via OSC (como por exemplo Logic Pro, Max/MSP, Pure Data, Super Collider, Nuendo, etc).

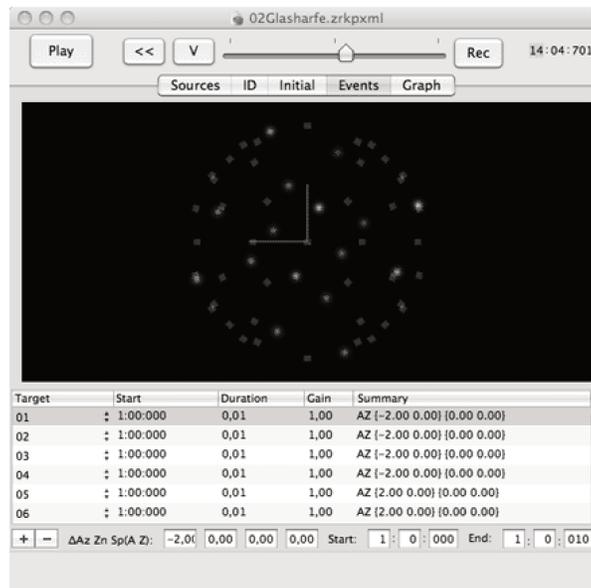


Figura 9: Software *Zirkonium*, da ZKM.

Como resultado do projeto *Gesture Control of Spatialization*, iniciado em 2005 pelo IDMIL²⁴ da Universidade McGill, o sistema *ViMiC (Virtual Microphone Control)* é uma ferramenta desenvolvida para MAX/MSP para controle da espacialização sonora para situações de concerto, instalações imersivas e, em especial, para audiências maiores e pouco centralizadas (Figura 10). Seu princípio baseia-se numa série de microfones com diversos padrões de diretividade que são posicionados dentro de uma sala virtual tridimensional (PETERS et al, 2011). O usuário pode também controlar as propriedades da sala virtual (tamanho e número de reflexões). O sistema suporta até 24 canais, sendo que não há a necessidade de uniformidade e equidistância na distribuição de alto-falantes na audiência. Pela facilidade de troca de informações via protocolo OSC, o sistema pode ser controlado por interfaces gestuais. O mapeamento gestual é feito através do protocolo *Gesture Description*

²⁴ <http://www.idmil.org>

Interchange Format (GDIF), desenvolvido por Marshall *et al.* (2006), que padroniza o modo de transmissão e armazenamento da informação gestual, captada por uma placa com oito sensores de movimento.

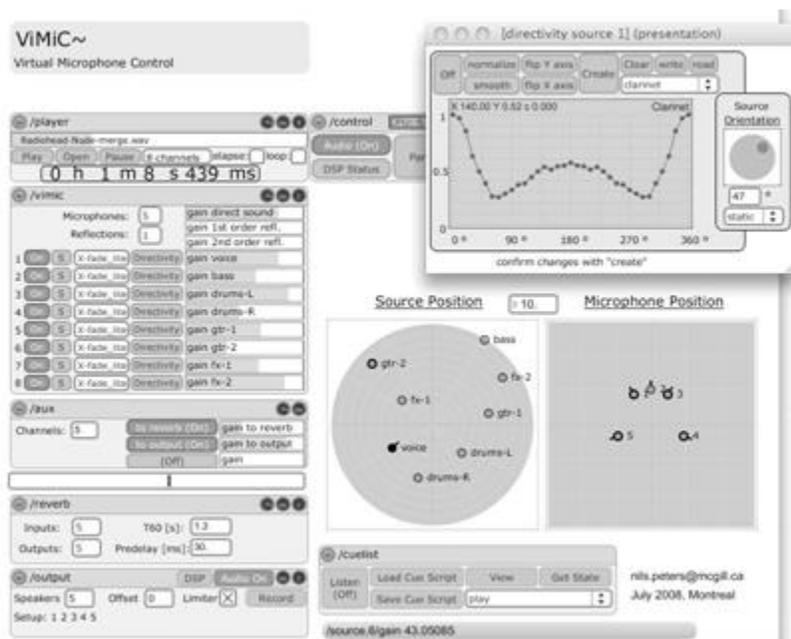


Figura 10: Interface gráfica do ViMiC (*Virtual Microphone Control*) para Max/MSP

Challenge Bodies é um sistema de código aberto para PD, desenvolvido pelos departamentos de música e informática da University of Regina, no Canadá. O sistema suporta a espacialização de até 10 fontes sonoras simultâneas distribuídas numa matriz de sete alto-falantes frontais. A disposição dos alto-falantes do sistema segue uma configuração não convencional, com cinco alto-falantes no plano mediano frontal e dois com elevação. Sua interface de controle consiste em duas janelas: a primeira, implementada em GEM²⁵, permite visualizar a disposição dos alto-falantes e modificar a posição das fontes sonoras no espaço; a segunda janela é um patch [rtss.pd], o qual controla a amplitude de cada fonte e a amplitude geral.

Audience é um sistema de espacialização desenvolvido por Faria (2005) para auralização imersiva implementado em PureData. Baseia-se numa arquitetura de quatro camadas responsáveis pela composição da cena acústica, simulação de propagação acústica do som, codificação e decodificação

²⁵ <http://gem.iemat>

espacial, respectivamente. O sistema comporta a codificação/decodificação de Ambisonics de até terceira ordem. Dispõe de uma série de ferramentas que permitem ao usuário controlar o posicionamento de diversas fontes sonoras além de poder alterar dinamicamente a posição do ouvinte (*sweet spot*).

SMuSIM é um sistema interativo de controle da espacialização também desenvolvido para PD. A primeira versão do projeto destina-se à espacialização quadrifônica e utiliza a técnica de panorama por amplitude para espacialização de até quatro fontes sonoras simultâneas. O foco da pesquisa, apresentada por Bozzolan e Cospito é o estudo de diferentes interfaces de controle, de fácil acessibilidade e baixo custo. Para isto, utilizam o controle através de mouse/teclado, controladores de jogos e câmeras de vídeos.

Capítulo 3

MICROSSOM: ANÁLISE, SÍNTESE E COMPOSIÇÃO

O entendimento acerca da natureza do fenômeno acústico tem historicamente repercutido em nossos modos de concepção, produção e performance musical. No início do século XX, os avanços da teoria acústica, aliados obviamente ao surgimento de novas tecnologias baseadas na eletricidade e à utilização de sinais eletromagnéticos, propiciaram meios para a geração de sons sem a utilização de instrumentos acústicos ou processos mecânicos. Desde a invenção dos primeiros sintetizadores analógicos e, sobretudo, o advento da tecnologia digital, diversos métodos de síntese sonora foram desenvolvidos, cada qual sob uma particular representação do fenômeno acústico. Na Tabela 8, listamos, resumidamente, os principais modelos de síntese sonora digital, baseado na taxonomia apresentada por Tolonen e colaboradores (1998).

Tabela 8: Taxonomia de modelos de síntese sonora digital, baseada em (TOLONEN *et al*, 1998)

Algoritmos abstratos	Modulação de Frequência (<i>FM</i>); Modelador de ondas (<i>Waveshaping</i>); Osciladores controlados por voltagem (<i>VCO</i>) amplificadores (<i>VCA</i>) ou filtros (<i>VCF</i>); <i>Karplus-Strong</i> .
Síntese por sons gravados (processados ou não)	Amostragem; <i>Wavetable</i> (temporal); <i>Wavelet</i> (temporal); Síntese Granular.
Modelos Espectrais	Aditiva; <i>Phase-Vocoder</i> ; Síntese por Formantes; Subtrativa (Filtragem); Ressíntese (Transformada Inversa de Fourier).
Modelos Físicos	<i>Karplus-Strong</i> (estendido); Síntese Modal;

Dentre tais representações, podemos encontrar aquelas que se baseiam em níveis temporais extremamente breves, situadas nos limiares perceptivos humanos dos parâmetros sonoros (altura, duração e intensidade). A este tipo de representação, que se verte tanto como técnica de análise como

também de criação sonora, denomina-se **microssom**. Esta abordagem tem origens na teoria do *quantum* acústico do físico húngaro, prêmio Nobel de física, Denis Gabor (1890-1979), o qual mostrou que um som de qualquer complexidade e duração pode ser representado, em analogia à Representação de Fourier, como uma soma ponderada de um conjunto de “células acústicas”, também denominadas grãos sonoros. A síntese granular, um “clássico” exemplo de criação em nível microtemporal, baseia-se assim no processo inverso, isto é, na construção, ou síntese, de morfologias sonoras (ou estruturas complexas) a partir de um conjunto de grãos sonoros ordenados discretamente no tempo.

Nas últimas décadas, podemos encontrar diversos métodos de síntese sonora que tem como fundamento o modelo granular, dos quais podemos citar:

- Síntese algorítmica (BERG, 1985; MANZOLLI, 1993);
- Métodos baseados em representações granulares do som, sobretudo síntese granular assíncrona (*Asynchronous Granular Synthesis – AGS*) e quase-assíncrona (*Quase-Asynchronous Granular Synthesis – QAGS*), controlados através de modelos estatísticos e probabilísticos (ROADS, 2001; TRUAX, 1988);
- Sistemas dinâmicos não lineares (DI SCIPIO, 1994a; TRUAX, 1990);
- Síntese granular síncrona (*Pitch-Synchronous Granular Synthesis – PSGS*) (POLI & PICCIALLI, 1991);
- Síntese concatenativa (SWARZ, 2005).

Nos tópicos a seguir apresentamos um breve histórico do debate teórico acerca da natureza do som, em seguida, abordamos o conceito fundamental da teoria do *quantum* acústico de Gabor.

3.1. Precedentes históricos

O princípio de que sons contínuos podem ser decompostos em porções temporais menores tem suas origens no pensamento atomístico da filosofia antiga. Leucippus e Demócrito (aprox. séc. V a.C.) defendiam que toda matéria era constituída de átomos separados por um espaço vazio (*kenon* ou *vácuo*). Eles especularam que qualquer substância poderia ser subdividida em partes cada vez menores até que se chegue a uma partícula indivisível denominada átomo.

As teorias atomísticas do som foram reavivadas por diversos pensadores e cientistas do século XVI e XVII, dentre eles Pierre Gassendi, Robert Hooke e Isaac Beeckman. O estudioso holandês Isaac

Beeckman (1588-1637) propôs uma teoria “corpúscular” do som, em que qualquer objeto em vibração, como uma corda, perfura o ar em redor em corpúsculos esféricos de ar ou partículas que são projetadas em todas as direções pela vibração. Beeckman atribuiu as variações de timbre ao tamanho, forma, velocidade e densidade destas partículas sonoras (ROADS, 2001).

Opondo-se à representação do som enquanto partícula, outra corrente teórica, também originada na filosofia antiga grega, propunha a interpretação do fenômeno acústico análoga ao comportamento ondulatório do movimento da água. Experimentos foram realizados por diversos cientistas em meados do século XVI no intuito de provar a hipótese da natureza ondulatória do som, dentre eles Robert Boyle e Marin Mersenne. Esses experimentos foram precursores no entendimento de comportamentos bastante peculiares do som, tais como a reflexão e a difração.

No século XIX, enquanto diversos experimentos tornavam a teoria atomística cientificamente aceitável, a teoria acústica ainda se manteve focada na natureza ondulatória do som. Em grande parte, esta tendência deve-se ao trabalho de Jean-Baptiste Fourier (1768-1830), que propôs a ideia de que qualquer sinal periódico (inclusive o sonoro) complexo pode ser decomposto pela soma ponderada de simples ondas senoidais de diferentes amplitudes, frequências e fases.

3.2. A Teoria de Gabor

Em meados da década de 40, Gabor propôs formalmente o conceito do *Quantum Acústico*, definido como a “mínima partícula” de som detectável pelo aparelho auditivo humano (GABOR, 1947)²⁶. Assim, em sua teoria, Gabor propõe que qualquer som complexo e de “longa duração” poderia ser representado pela sucessão de discretas unidades de energia acústica, os *quanta* sonoros.

Diferentemente da abordagem assumida por Fourier, a solução proposta por Gabor relaciona a função densidade de energia $s(t)$ associada ao sinal em relação ao tempo à função densidade de energia $S(f)$ do espectro (densidade espectral ou distribuição de energia nas componentes espectrais) representado no domínio das frequências. Gabor pondera que uma representação exclusiva do sinal sonoro como função $s(t)$ é uma interpretação errônea do fenômeno sonoro, ao menos sob a perspectiva da psicoacústica, dado que nenhum sinal pode ser detectado num intervalo de tempo infinitesimal ou instantâneo. A recepção e codificação do sinal sonoro pelo aparato auditivo humano requer a exposição

²⁶ O conceito de partícula de som proposto por Gabor tem origens na Teoria quântica da luz, a qual afirma que a luz, mesmo sendo uma onda eletromagnética, é constituída de corpúsculos ou unidades de energia denominadas *fótons* ou *quanta* de luz.

do mesmo pelo menos durante certo intervalo de tempo. Do mesmo modo, uma representação exclusiva do sinal através da soma de ondas senoidais de duração infinita, tal qual propõe o modelo tradicional da Teoria de Fourier, embora matematicamente correta, não corresponde a real experiência auditiva humana.

Deste modo, Gabor mapeia a função $s(t)$ restrita a uma determinada duração Δt e a função $S(f)$ a uma banda espectral Δf . A análise do espectro através de intervalos temporais finitos, procedimento este conhecido atualmente na literatura como *análise por janelamento*, implica, entretanto, uma *relação de incerteza* entre a resolução de tempo e frequência. Em outras palavras, quanto maior a precisão na medição de certa frequência f_0 , sob um desvio de erro Δf , maior será o intervalo temporal Δt necessário, o que implica dizer que Δf e Δt são quantidades inversamente proporcionais²⁷. Assim, quanto maior a precisão no domínio das frequências, maior o tamanho da janela temporal necessário para análise e, portanto, menor a acuidade de informação no domínio do tempo (GABOR, 1947). Gabor expressa essa relação através da desigualdade:

$$\Delta t \cdot \Delta f \geq 1 \quad (3.1)$$

O princípio da teoria de Gabor pode ser exemplificado através de um sinal elementar s constituído de um oscilador em qualquer frequência audível, modulada por envelope de duração finita:

$$s(t) = e^{-\alpha^2(t-t_0)^2} e^{i\alpha 2\pi f_0 t}, \quad (3.2)$$

onde f_0 é a frequência da função senoidal, t e t_0 os tempo inicial e final da função temporal gaussiana e α uma constante (GABOR, 1947: 591).

O comportamento gaussiano do sinal representado do domínio do tempo, determinado pela expressão $e^{-\alpha^2(t-t_0)^2}$ da equação 3.2, implica, pelo Teorema da Transformada de Fourier, um comportamento análogo à função $S(f)$, resultante da Transformada do sinal para o domínio das frequências. Como podemos observar na Figura 11, a modulação do sinal sonoro por uma função gaussiana no domínio do tempo, de largura ou abertura Δt , implica uma modulação do sinal também no

²⁷ É interessante observarmos que este raciocínio fundamenta-se sob um argumento similar ao princípio da Incerteza de Heisenberg da Mecânica Quântica, pelo qual a exata precisão da posição de uma determinada partícula é diretamente proporcional à incerteza de sua exata velocidade.

domínio das frequências por uma função similar, de largura Δf . As larguras Δt e Δf , relativas respectivamente à duração temporal e banda frequencial efetivas de representação do sinal, são dadas pelas equações 3.3.

$$\Delta t = \frac{\sqrt{\pi}}{\alpha} \quad \Delta f = \frac{\alpha}{\sqrt{\pi}} \quad (3.3)$$

Pelas equações apresentadas acima, podemos observar claramente a proporção inversa entre as grandezas, que pelo seu produto, $\Delta t \cdot \Delta f = 1$, satisfazem minimalmente a princípio da desigualdade de Heisenberg. A relação dessas entre estas duas grandezas é representada por Gabor através de eixos ortogonais no plano cartesiano bidimensional, sendo o *Tempo* representado no eixo vertical e a *Frequência* no eixo horizontal, conforme Figura 11. Esta representação é conhecida como Espaço *Tempo-Frequência* de Gabor. Neste diagrama, o quantum acústico é representado pelo retângulo cinza, cuja área é dada por $A = \Delta t \cdot \Delta f$. Prova-se que em unidades adequadas, a área A é, essencialmente, a energia do sinal sonoro.

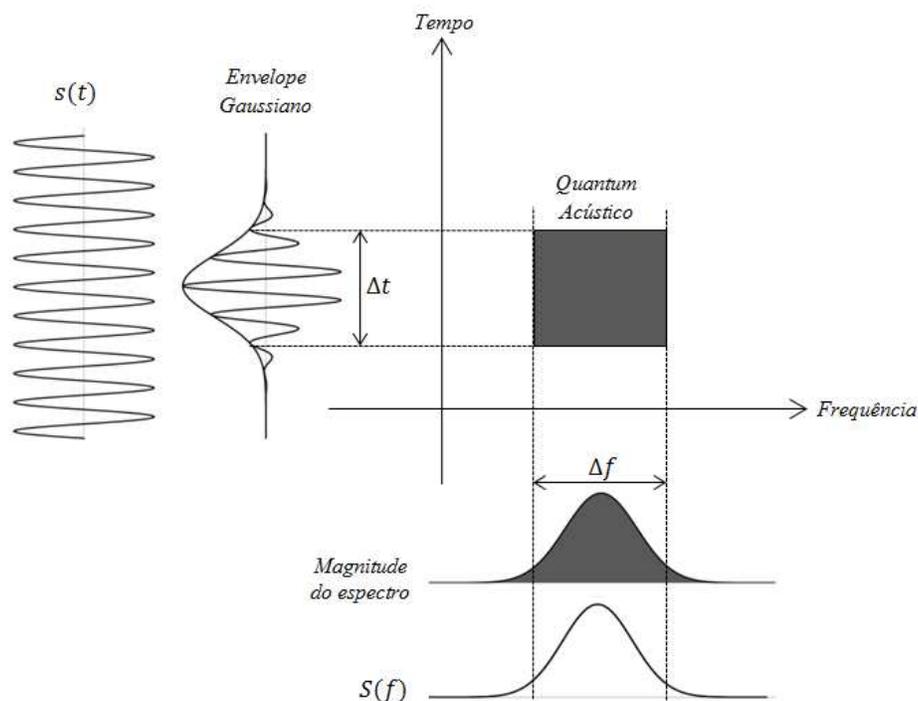


Figura 11: Espaço *Tempo-Frequência* de Gabor

3.3. Síntese Granular Analógica

Impulsionados pela possibilidade de representação do som através dos princípios quânticos estabelecidos a partir da teoria de Gabor, diversos compositores apropriaram-se destas ideias como proposta de criação sonora²⁸.

Iannis Xenakis foi um dos primeiros a explorar musicalmente a ideia da Síntese Granular. Em seu livro *Musique Formelles*, publicado em 1963²⁹, Xenakis transpõe seu pensamento acerca da criação sonora a partir de grãos.

Todo o som é uma integração de grãos, de partículas sonoras elementares som, de *quanta* sonoros. Cada uma dessas partículas elementares tem uma natureza tripla: a duração, a frequência e intensidade. Todo som, toda variação sonora, mesmo que contínua, é concebida como um conjunto de grãos elementares suficientemente numerosos e adequadamente organizados no tempo. Deste modo, qualquer complexo sonoro pode ser analisado em séries de sons puramente senoidais—ainda que as variações destes sejam infinitamente próximas, curtas e complexas. No ataque, corpo e decaimento de um som complexo, milhares de sons puros aparecem num intervalo de tempo Δt suficientemente curto. Uma avalanche de sons puros é necessária para se criar um som complexo³⁰. (XENAKIS, 1963 : 61, *tradução nossa*)

O uso de complexos sonoros constituídos por um grande conjunto de grãos sonoros é elemento recorrente na escritura de Xenakis. Desde suas primeiras obras orquestrais – *Metastasis*, *Piktoprakta* e *Achorripsis* – podemos observar a preocupação do compositor em estabelecer formas globais (macroestruturas) através de *micro-transformações* dos materiais (sons orquestrais) disponíveis.

Na Figura 12 apresentamos um esboço gráfico utilizado como suporte pelo Xenakis para a composição dos *pizzicati glissandi* nos instrumentos de corda de *Pithoprakta* (1955, comp. 52-59).

²⁸ Diversos autores atribuem esta tendência de articulação do discurso sonoro em níveis temporais cada vez menores como decorrência da estética atomística da música pós-weberniana, difundida, sobretudo, pelas escolas de Darmstadt e da música eletrônica de Köln (THOMSON, 2004). Esta tendência atomística tem repercussões em outras abordagens da música pós-serialista, em especial na *micropolifonia* do compositor György Ligeti e também na música espectral de Gerard Grisey.

²⁹ Vale ressaltar que o capítulo de *Musique Formelles* ao qual Xenakis se refere à ideia da síntese granular já havia sido publicado anteriormente numa tradução para o alemão em: XENAKIS, I. Auf der Suche einer Stochastischen Musik. *Gravesaner Blätter*, n°11-12, 1958, p. 98-111/112-122.

³⁰ "Tout son est une intégration de grains, de particules élémentaires sonores, de quanta sonores. Chacun de ces grains élémentaires a une triple nature: la durée, la fréquence et l'intensité. Tout son, toute variation sonore même continue est conçue comme un assemblage de grains élémentaires suffisamment nombreux et disposés dans le temps d'une façon adéquate. Donc : tout complexe sonore est analysable en séries de sons purs sinusoïdaux même si les variations de ces sons sinusoïdaux sont infiniment rapprochées, brèves et complexes. Dans l'attaque d'un son complexe, dans son corps, dans sa chute, des milliers de sons purs apparaissent dans un intervalle de temps Δt assez court. Des hécatombes de sons purs sont nécessaires à la création d'un son complexe" (XENAKIS, 1963 : 61).

Cada linha, representada por um instrumento de corda, movimenta-se função das alturas (eixo vertical) e do tempo (eixo horizontal). Xenakis determina os parâmetros locais de cada glissando (altura do ataque, duração e velocidade) através de distribuições probabilísticas globais representadas através do gráfico.

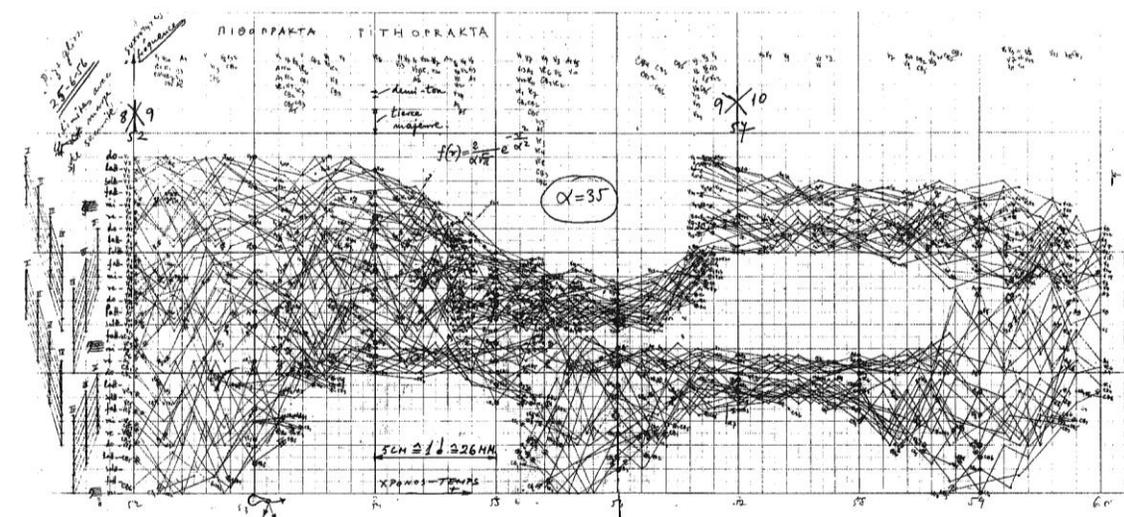


Figura 12: *Pithoprakta* (1955-56), comp. 52-59, rascunho do próprio compositor³¹.

Frente às possibilidades de manipulação sonora através da gravação e edição de fitas eletromagnéticas, Xenakis explorou a articulação do material em níveis temporais ainda menores. *Concrete PH* (1958, para fita eletromagnética) configura-se como um dos primeiros exemplos de síntese granular analógica. A obra foi composta no estúdio do GRM e estreada como um interlúdio do espetáculo *Poème Électronique* concebido por Le Corbusier para o Pavilhão Philips da Expo'58 de Bruxelas (*Exposition universelle et internationale de Bruxelles*). De acordo com Di Scipio (1998), a construção da obra envolveu três etapas distintas. A primeira compreendeu a gravação em fita magnética dos material sonoro que seria utilizado, a dizer, estalidos provenientes da queima de carvão. Na segunda etapa, pequenos fragmentos foram extraídos da gravação original e isolados do seu contexto. Cada fragmento correspondia a um estalido extraído da gravação e de duração variável inferior a poucos centésimos de segundo³². Por último, na terceira etapa, a justaposição de inúmeros

³¹ Extraído de: XENAKIS, I. *Musique*. Architecture, Tournai, Casterman, 1971, p.167.

³² É interessante notarmos que, dada a natureza do material gravado, caracterizado pela ampla distribuição de energia por todo o espectro, a percepção de frequência torna-se dependente da duração de cada fragmento, de forma que quanto menor o impulso de som, maior a banda de frequência perceptível.

fragmentos extraídos da fita resultou a criação de uma textura maior. Diferentemente de outras obras do compositor, este processo deriva de uma manipulação intuitiva e direta com o material concreto e, portanto, não sinaliza nenhuma formalização do controle macroestrutural da obra.

Em *Analogique B* (1959)³³, Xenakis utilizou processos estocásticos markovianos para controle macroestrutural da síntese granular. O modelo formal consistia primeiramente em dividir o tempo em unidades iguais de curtíssima duração, as quais ele denominou telas (*screens*). Cada tela era representada por um conjunto de parâmetros atribuídos a um grão sonoro. O sequenciamento dos grãos era determinado por um conjunto de matrizes que indicavam a probabilidade de transição entre as telas (Figura 13). Os sons utilizados para a criação eletrônica basearam-se em sons senoidais gravados em fita analógica que eram editados pelo compositor (XENAKIS, 1992; ROADS, 2001).

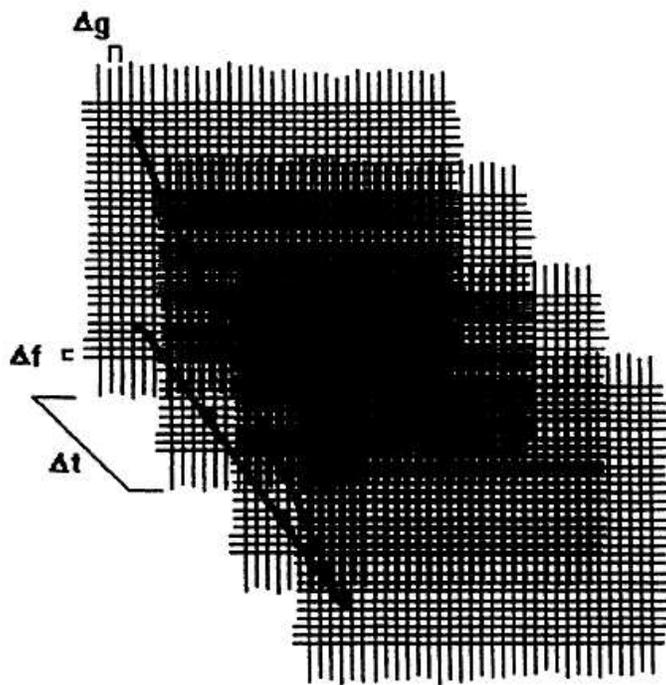


Figura 13: Sequência de telas (*screens*) utilizadas para controle microestrutural de *Analogique B* (1959), de I. Xenakis, com valores discretos de frequência (Δf), intensidade (Δg) e duração (Δt).³⁴

³³ *Analogique B* foi composta como complemento à obra instrumental *Analogique A* (1958), para nove instrumentos de cordas (3 violinos, 3 violoncelos e 3 contrabaixos), que deriva de processos estocásticos similares à primeira. No segundo capítulo de seu livro *Musique Formelles*, o compositor descreve exaustivamente a formalização da qual resultou a composição de ambas as obras (XENAKIS, 1963).

³⁴ Extraída de Di Scipio (1998: 216).

Em seu artigo publicado em 1957, intitulado “... *how times passes...*”, Stockhausen propõe uma visão composicional unificada das diferentes camadas temporais da estrutura musical (STOCKHAUSEN, 1957). Na concepção do compositor, ritmo e altura seriam representações de um mesmo fenômeno – a periodicidade – diferidas apenas por suas respectivas escalas temporais. Um pulso consistiria deste modo numa altura de frequência infrassônica, bem como a altura num pulso de período muito curto, ao ponto de ser percebido como um evento contínuo.

Em sua obra *Kontakte* (1958-60)³⁵, Stockhausen aborda uma unidade no tratamento dado aos diversos parâmetros da composição (STOCKHAUSEN, 1962). Diversos geradores de impulsos analógicos, cada qual conectado a outros componentes (filtros passa-banda, gravadores de fita em loop e circuitos de realimentação com diferentes arquiteturas), geram sequências de impulsos em diferentes velocidades (intervalos regulares ou irregulares) e densidades, estabelecendo transformações simétricas³⁶ no domínio das alturas e das durações.

Procedimentos similares foram utilizados também por outros compositores, como podemos observar em *Scambi*, de Henri Pousseur e *Syntaxis*, de Bruno Maderna, ambas compostas em 1957.

3.4. Síntese Granular Digital

O surgimento da tecnologia digital e, particularmente, o desenvolvimento de algoritmos computacionais capazes de gerar síntese sonora foram fatores preponderantes para a emergência do microssom. Na década de 70, sistemas como o SAWDUST, desenvolvido por Herbert Brün (1918-2000), e o SSP (*Sound Synthesis Program*), desenvolvido por Gottfried M. Koenig (n.1926), foram pioneiros pela utilização de pequenas amostras sonoras como elemento para síntese digital (THOMSOM, 2004).

Motivado pelas ideias de Xenakis, o compositor americano Curtis Roads (n.1951) foi um dos precursores de aplicação da síntese granular no contexto digital. Em seu primeiro estudo experimental, denominado *Klang I*, realizado em 1974 através de um processador duplo Burroughs B6700 da Universidade da Califórnia (USDC), Roads gerou 766 grãos de som para compor pouco menos que um

³⁵ A obra possui duas versões distintas, a primeira acusmática para sistema quadrifônico e a segunda mista, para piano percussão e sons eletrônicos. Uma terceira versão foi realizada em 1961, como teatro musical, posteriormente incorporada à segunda versão da obra.

³⁶ Vale ressaltar que neste período Stockhausen interessava-se pelos desdobramentos da técnica serial para outros parâmetros da composição e o estabelecimento de simetrias entre as dimensões perceptivas demonstrava-se, ao menos conceitualmente, uma promissora ferramenta para serialização das durações e timbres.

minuto de música, especificando os parâmetros (frequência, amplitude e duração) para cada um dos grãos.

Já em seu segundo experimento, Roads escreveu um programa denominado PLFKLANG para gerar uma maior quantidade de grãos através de especificações de mais alto nível. *Prototype* (1975) resulta da manipulação de sete parâmetros de nuvens de grãos que estabelecem controles estocásticos globais da composição, tal qual concebido por Xenakis (ROADS, 2001).

Desde então, Roads dedicou-se a uma sistemática exploração das possibilidades de síntese obtida através de grãos. Em seu livro *Microsound* (ROADS, 2001), o compositor apresenta um grande compêndio das contribuições trazidas por ele mesmo e por outros pesquisadores nas últimas décadas à área da síntese granular.

A grande maioria dos trabalhos desenvolvidos por Roads nas décadas de 80 e 90 referem-se a métodos de síntese sonora em tempo diferido. O compositor canadense Barry Truax (n.1947) foi o pioneiro em realizar síntese granular em tempo real com o auxílio de um processador DMX-1000 controlado por um microcomputador DEC LSI-11, no Instituto de Sonologia de Haia, na Holanda (TRUAX, 1988). Sua primeira composição realizada com suporte deste software foi a célebre *Riverrun* (1986, quadrifônico; 2004, octofônico). A obra, homônima à primeira palavra de *Finnegans Wake* do escritor irlandês James Joyce, sugere uma relação metafórica com um rio, em seu contínuo fluxo constituído de inúmeras gotas de água. Na composição, pequenas unidades sonoras se acumulam gradualmente, criando longas texturas e massas sonoras. Pela natureza em tempo real do software utilizado, o compositor dispõe de um controle mais incisivo sobre a gradação de diferentes parâmetros ao longo do tempo.

Truax foi também o primeiro a utilizar amostras sonoras pré-gravadas para criar texturas granulares, processo este conhecido como *granulação*. Sua obra *Wings of Nike* (1987), utiliza dois fonemas gravados por voz masculina e feminina, “tuh” e “ah” respectivamente, como material base principal.

Mais recentemente, podemos encontrar distintas abordagens de síntese granular recorrentes em obras de diversos compositores, tais como Agostino Di Scipio, Ludger Brümmer, Manuel Rocha, Ramon Gonzalez-Arroyo, Eduardo R. Miranda, Damián Keller, Mara Helmuth, Kim Cascone, dentre outros.

Observamos também um extensivo emprego de modelos abstratos, formais ou matemáticos, para controle de parâmetros relacionados à síntese granular. Maia e Miranda (2005) utilizaram Conjuntos Fuzzy e Cadeias de Markov, a Teoria dos Grafos foi proposta como modelo de controle por Valle e Lombardo (2003), o algoritmo *boids*, que emula comportamentos espaciais de deslocamento de animais em “bandos”, foi utilizado por Blackwell e Young (2004), Kim-Boyle (2006) e Barreiro (2011) e algoritmos genéticos foram utilizados por Fornari e colaboradores (2006) e também por Souza e Maia (2009).

No tópico a seguir, detalharemos uma abordagem peculiar de síntese que deriva do método granular, a síntese concatenativa.

3.5. Síntese Concatenativa

A *Síntese Concatenativa* é um método de produção sonora baseado na concatenação de um conjunto de pequenos segmentos sonoros para constituir um determinado perfil sonoro. O método empregado utiliza um enorme banco de sons gravados, que são segmentados em *unidades*, e um algoritmo de seleção que encontra uma sequência destas *unidades* que melhor se assemelhe ao som ou frase a ser sintetizada, que por sua vez é denominada por *alvo* (SCHARZ, 2006).

Essas *unidades*, apesar de não se restringirem aos limites microtemporais (< 100ms), sendo comum o emprego de *unidades* de durações variáveis de até 1000 milissegundos, têm função correlata aos grãos nos modelos granulares “clássicos” apresentados anteriormente.

A comparação entre o som *alvo* e cada uma das *unidades* é realizada através da descrição das fontes sonoras, seja por meio da recuperação de informações extraídas do sinal sonoro *alvo* (características descritivas relativas ao espectro sonoro ou da forma de onda, por exemplo) ou descrições de alto nível atribuídas a elas (descritores simbólicos ou descrições associadas a modelos cognitivos ou psicoacústicos, por exemplo).

Como podemos observar na Figura 14, um conjunto de características são extraídas tanto do banco de sons armazenados quanto do som *alvo*. A partir da segmentação temporal do conjunto de dados extraídos do som *alvo* em diversas durações, um algoritmo calcula uma *distância* entre o vetor de características das *unidades* armazenadas no banco de sons e os vetores de características de cada segmento do som *alvo*. À medida que o algoritmo encontra similaridades entre os vetores analisados, sob um dado limiar, as *unidades* correlatas são então endereçadas à saída de áudio do sistema.

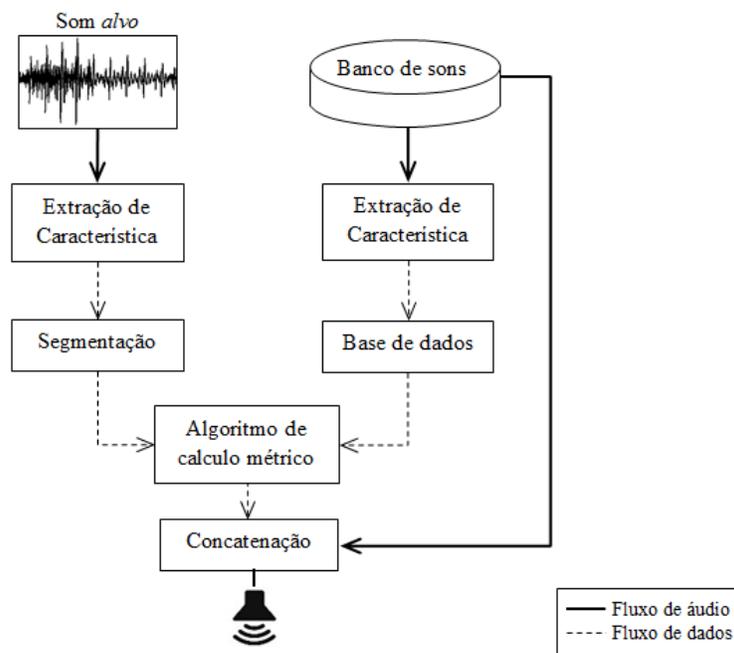


Figura 14: Fluxograma de Síntese Concatenativa

Na década de 90, este método foi empregado como síntese de voz falada, através da concatenação de um grande número de fonemas previamente gravados (PRUDON, 2003) e também utilizado para síntese de voz cantada (RODET, 2002). Nas últimas décadas, diversas aplicações musicais tem utilizado este método, tais como os sistemas *Caterpillar* (SCHWARZ, 2003), *Soundmosaic* (HAZEL, 2001), *MATConcat* (STURM, 2004), *SoundSpotter* (SPEVAK & FAVREAU, 2002) e *CataRT* (SCHARTZ, 2005).

3.6. Microcomposição

O surgimento das ferramentas computacionais e dos dispositivos digitais representa mais que uma mera expansão dos recursos técnicos aplicados à criação musical, mas a emergência conceitual de um profundo remodelamento da ação do compositor frente às relações estabelecidas entre forma e material (DI SCIPIO, 1994).

Frente a essas mudanças do paradigma composicional, apresentaremos neste tópico um breve panorama sobre as principais abordagens da composição no contexto do microsom. De maneira

particular, discutiremos alguns aspectos referenciais da composição em domínio microtemporal presentes nas obras de quatro importantes compositores: Curtis Roads, Barry Truax, Horacio Vaggione e Agostino di Scipio.

A extensiva e sistemática exploração de diferentes métodos de síntese granular propiciou para Roads o desenvolvimento de uma ampla e rica taxonomia de estruturas granulares. O principal aspecto que ressaltaremos sobre a abordagem composicional de Roads refere-se à ideia de contrastes. Em suas obras, podemos notar em diversos momentos a presença de perfis bem delineados, morfologias sonoras que se articulam através de diversos níveis temporais, que vão desde o micro até o mesoestrutural.

Na Figura 15, apresentamos um sonograma³⁷ dos quatro segundos iniciais de *Pictor Alpha* (2003). Pelo pequeno trecho apresentado na ilustração, podemos observar uma clara segmentação do material em subsequências temporais. Cada fragmento constitui-se como um objeto sonoro que se diferencia por seus comportamentos espectrais e morfológicos.

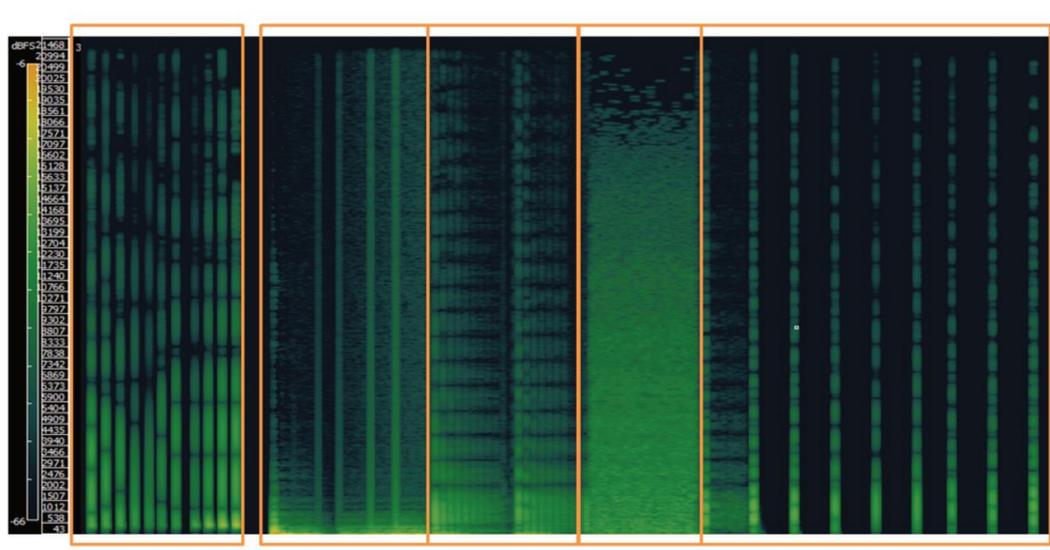


Figura 15: Sonograma de *Pictor Alpha* (2003), de Curtis Roads.

No caso de *Pictor Alpha*, os objetos sonoros são basicamente constituídos pela repetição de pequenas células. Em algumas outras peças, no entanto, cada fragmento é delineado por processos distintos de transformação dos grãos ao longo do tempo. Na Figura 16, um sonograma de um pequeno

³⁷ Todos os sonogramas utilizados neste capítulo foram realizados com o auxílio do software *Sonic Visualizer*, disponível gratuitamente em <http://www.sonicvisualiser.org/>.

fragmento da segunda parte de *Volt Air* (2003) exemplifica este procedimento. O aspecto global da peça constitui-se na combinação de morfologias mesotemporais – nuvens granulares resultantes de micro transformações (no caso, diferentes procedimentos de filtrações) do material sonoro.

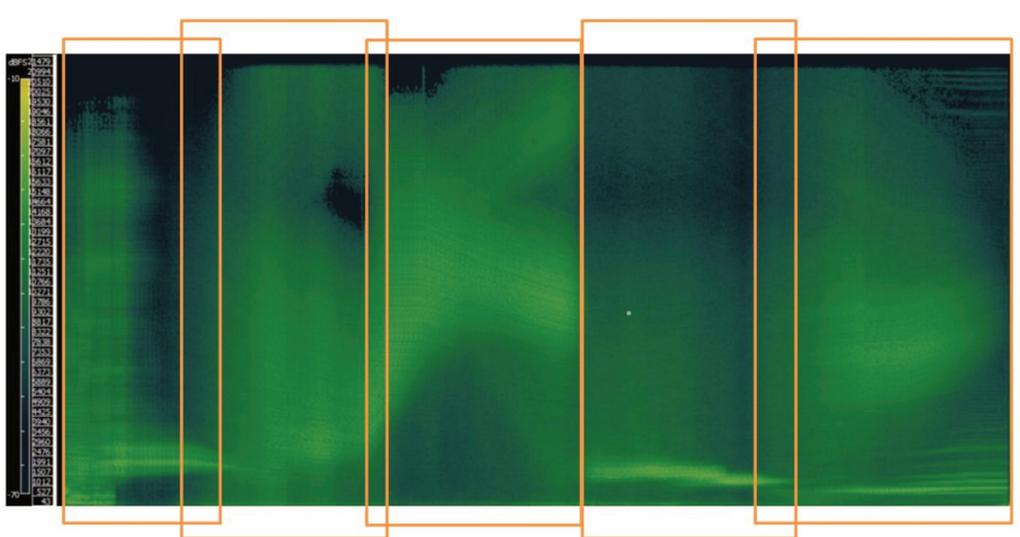


Figura 16: Sonograma de fragmento (1'7'' a 1'48'') de *Volt Air, Part II* (2003), de Curtis Roads.

Na obra de Barry Truax, a microcomposição resulta de desdobramentos texturais, de processos de transformação dos materiais sonoros ao longo tempo. Ao contrário dos exemplos que apresentamos anteriormente na peça de Roads, é raro encontrarmos rupturas bem delineadas em alguma obra de Truax. Como decorrência clara da abordagem em tempo real desenvolvida por ele, os aspectos globais de sua obra derivam da modificação de certo conjunto de parâmetros que graduam a textura sonora de um determinado estado até outro.

Na Figura 17, apresentamos um sonograma do segundo movimento da obra *Wings of Nike, Ein kleines Scherzo*. Pela visualização, podemos observar a ausência de grandes contrastes na obra como um todo, prevalecendo um grande perfil crescente e decrescente ao longo da peça. O Gráfico 1 representa a análise do fluxo espectral³⁸ da amostra analisada. Nele, observamos uma curva de tendência também em perfil de arco crescente e decrescente, o que indica menor variabilidade do

³⁸ O Fluxo Espectral é um descritor acústico que calcula a diferença de magnitude entre quadros da FFT do sinal amostrado. Valores baixos indicam um sinal mais estático, ou seja, com poucas variações de energia, enquanto que valores altos indicam maior variação do envelope espectral (PEETERS, 2003).

espectro nas regiões iniciais e finais, e maior variabilidade espectral na região central. Este processo resulta da técnica empregada pelo compositor, que inicia a obra apresentando grãos de maior duração e esparsos no tempo, direcionando assim a escuta para cada grão individualmente. A duração e esparsidade dos grãos no início da obra estabelece ao ouvinte referências em relação à origem dos materiais sonoros utilizados (conforme descrito no item 3.4). No decorrer do movimento, o vínculo com a fonte torna-se cada vez menos evidente, processo este ocasionado pelo efeito de mascaramento³⁹, ocasionado pelo aumento da densidade dos grãos e diminuição da duração dos mesmos. Por fim, a reaparição de grãos de maior duração distribuídos de maneira mais esparsa no tempo sugere uma espécie de retorno ao estado inicial da obra.

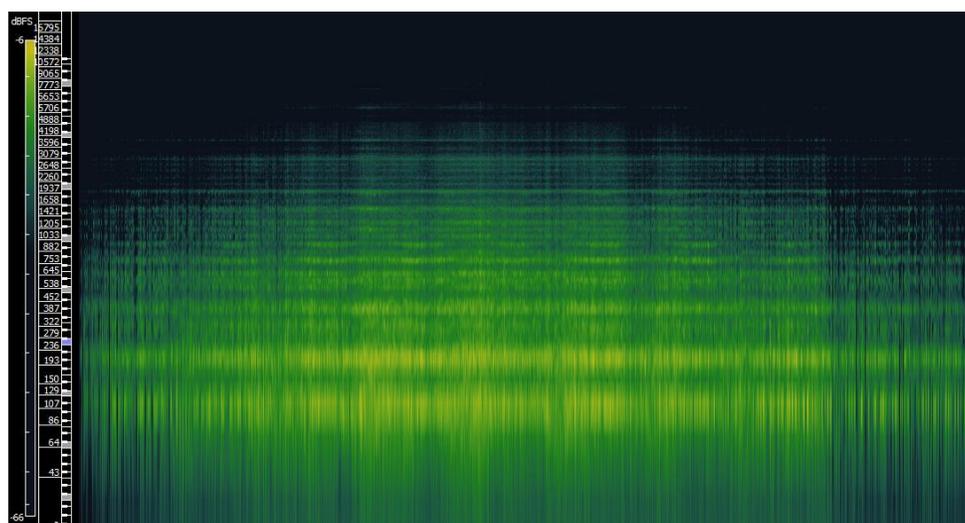


Figura 17: Sonograma de *Wings of Nike* (1988), II - *Ein kleines Scherzo*, de Barry Truax (frequências em distribuição logarítmica).

³⁹ Trata-se de um fenômeno psicoacústico que ocorre quando o aparelho auditivo humano não é capaz de distinguir dois eventos sonoros isoladamente devido a um curto distanciamento temporal entre os mesmos (intervalos próximos ou inferiores a 50ms). Em síntese granular, quando a densidade de grãos supera a quantidade de 20 grãos por segundo, ocorre uma fusão dos mesmos, que passam a ser percebidos como uma textura contínua.

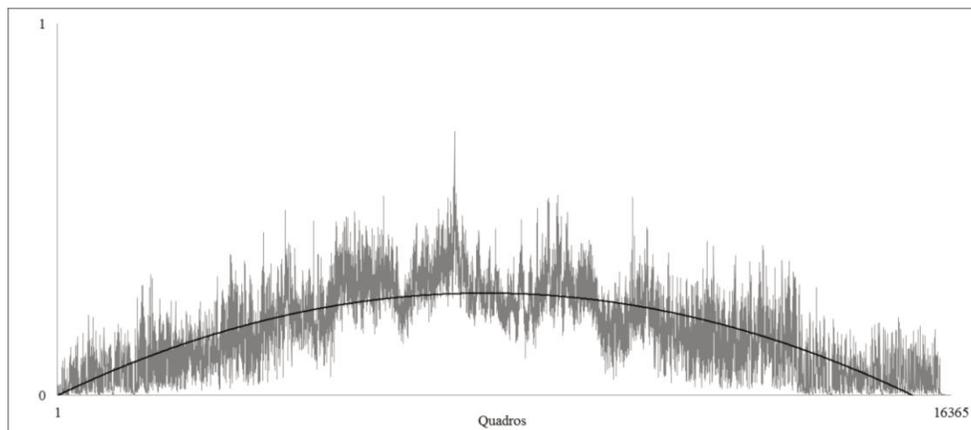


Gráfico 1: Fluxo Espectral em *Wings of Nike, II - Ein kleines Scherzo* de Barry Truax (eixo vertical: magnitude normalizada do espectro).

Outra importante abordagem da composição no domínio microtemporal é a do compositor argentino Horácio Vaggione (n.1943). Para ele, o microtemporal instancia a emergência da complexidade dinâmica dos sons, o domínio no qual “estruturas dissipativas de energia sonora”⁴⁰ manifestam-se em diferentes níveis de descontinuidade (VAGGIONE, 2003).

A principal técnica explorada pelo compositor é conhecida como *micromontagem*. O termo, derivado do conceito de *montagem* aplicado ao cinema, refere-se à construção de padrões sonoros complexos a partir da edição de pequenos fragmentos sonoros. A técnica assemelha-se à técnica de granulação, porém, ao contrário dessa que geralmente é aplicada através de processos automáticos, aquela resulta do exaustivo trabalho de manipulação e edição dos grãos individualmente⁴¹ (ROADS, 2002).

Thema (1985, para saxofone baixo e eletrônica), *Tar* (1987, para clarinete baixo e eletrônica) e *SCIR* (1988, para flauta baixo e eletrônica) são exemplos de aplicação dessa técnica. É interessante notarmos que os materiais sonoros utilizados na parte eletrônica dessas peças, e de uma maneira geral em toda a produção de obras mistas do compositor, são amostras previamente gravadas dos próprios instrumentos. As aglomerações dos materiais amostrados em padrões de diferentes tamanhos

⁴⁰ A terminologia utilizada por Vaggione (2003) se relaciona ao conceito de *estruturas dissipativas* postuladas pelo químico Ilya Prigogine.

⁴¹ Notamos que, apesar da diferenciação descrita por Roads (2002), os termos apresentam bastante ambiguidade e denotam práticas muito semelhantes.

estabelecem distintos pontos de intersecção para com a escritura, atuando assim a eletrônica como uma espécie de extensão do próprio instrumento.

Um importante aspecto da composição de Vaggione é a sua abordagem multiescalar. O compositor explora as diversas possibilidades de articulação do material em simultâneas escalas temporais. Sua obra *Schall* (1995, acusmática) é um exemplo magistral dessa abordagem. Todo o material utilizado na obra consiste de amostras sonoras gravadas de piano que foram granuladas e processadas. O conjunto limitado de elementos sonoros é interposto em distintas escalas que transitam entre durações micro a mesotemporais, em que pequenas notas e gestos sonoros podem ser percebidos.

Na visão do compositor italiano Agostino Di Scipio (n.1962), a microcomposição é entendida como uma maneira de “permitir propriedades morfológicas globais da estrutura musical emergem de condições locais do material sonoro”⁴² (DI SCIPIO, 1994a, p. 205, *tradução nossa*). Inspirado nos modelos probabilísticos e estocásticos da música de Xenakis, Di Scipio utiliza funções iterativas (DI SCIPIO, 1994a) e sistemas dinâmicos não lineares (DI SCIPIO, 2002) para o controle da síntese granular em suas obras. Seu modelo de composição pauta-se na ideia da macroforma como uma estrutura emergente do som em si e de sua organização interna, de forma que estruturas complexas emergem de simples interações locais entre os componentes do sistema (SOLOMOS, 2006).

A manipulação composicional instancia-se neste caso em operações no domínio *sub-simbólico*, em que a ação do compositor incide apenas indiretamente sobre os aspectos globais (macroforma) da composição. Esta abordagem criativa é pautada de princípios da auto-organização, sobretudo inspirada pelos modelos das ciências naturais, em que a estrutura (padrões de formas definidas) e a ordem (organização) de um sistema surgem como um fenômeno interno, sem a intervenção de um agente externo, resultante da interação entre seus componentes, em estado menos complexo e menos ordenado.

Na seção a seguir, relataremos o estado da arte acerca da espacialização sonora no domínio microtemporal.

⁴² “The task of microcompositional strategies can be described as one of letting global morphological properties of musical structure emerge from the local conditions in the sonic matter” (DI SCIPIO, 1994a., p.205)

3.7. Espacialização de sons granulares

Como tratado no capítulo anterior, a espacialização sonora pode configurar-se sob múltiplas escalas temporais. De maneira peculiar, a espacialização no domínio micro, em que uma localização aural específica é atribuída a cada partícula ou grão sonoro, mostra-se como procedimento mais idiomático ao contexto da síntese granular.

Novamente, o emprego de mecanismos macro para controle dos parâmetros de síntese torna-se indispensável. Dado que inúmeros grãos são necessários para a formação de um determinado segmento sonoro, é comum o emprego de controle macroestrutural, que determine, dentre outros parâmetros, a localização específica para cada grão no espaço de projeção sonora.

Podemos encontrar na literatura o uso frequente de métodos matemáticos elementares na distribuição espacial de síntese granular como, por exemplo, através de distribuições aleatórias ou estocásticas dos grãos (ROADS, 2001). Estes, no entanto, apresentam na maioria das vezes pouca ou nenhuma relação com os outros aspectos sonoros e/ou musicais em contexto. Para além dos procedimentos aleatórios e/ou estocásticos, recentes trabalhos têm adotado outras estratégias de auralização dos grãos sonoros.

Kim-Boyle (2006), Wilson (2008) e Barreiro (2011) têm utilizado o algoritmo *boids* para o controle de espacialização granular em sistemas multicanal. O algoritmo, desenvolvido por Craig Reynolds⁴³, simula comportamentos emergentes complexos de deslocamento espacial de “bandos” através de três regras elementares: separação, alinhamento e coesão. O modelo de Reynolds tem se estendido a outras regras arbitrárias, tais como a presença de atratores e a definição de limites espaciais.

Wilson (2008) desenvolveu um sistema de granulação espacial de enxames (*Spatial Swarm Granulation*) que utiliza o algoritmo *boids* para controlar dinamicamente a localização espacial de fluxos granulares. O sistema é baseado em três classes implementadas em SuperCollider⁴⁴ que controlam a posição e o movimento de diferentes fluxos granulares num espaço definido por um grande número de alto falantes localizados virtualmente (Figura 18).

⁴³ Para maiores detalhes sobre o modelo algorítmico, consultar informações contidas na página do Craig Reynolds na internet em <http://www.red3d.com/cwr/boids/>. Acesso em 10fev2013.

⁴⁴ Consultar <http://supercollider.sourceforge.net>.

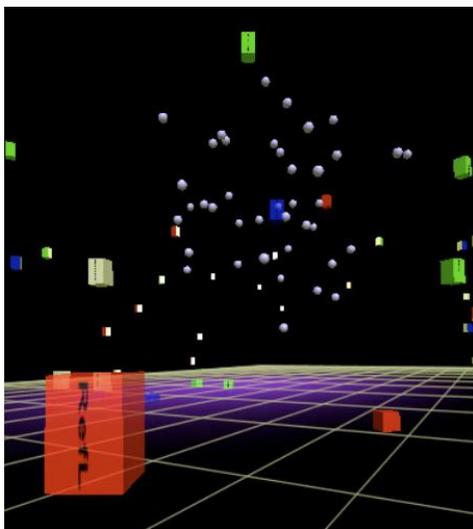


Figura 18: Detalhe da visualização do sistema de espacialização granular de Wilson (2008). As esferas representam os *boids* e os cubos os alto-falantes .

O controle espacial de segmentos sonoros via modelos de Algoritmos Genéticos foi utilizado Souza e Maia (2009) e também por Fornari e colaboradores (FORNARI *et al*, 2007). No modelo de Souza e Maia (2009), a posição panorâmica de cada grão no sistema estereofônico é resultante dos processos de *reprodução* e *mutação* do sistema evolutivo. Através de uma interface bastante intuitiva, denominada *EvoGrain*, o usuário pode determinar regiões “alvo” (representada pelos quadrados vermelhos e azuis na Figura 19) para um dado conjunto de parâmetros de síntese. Uma função de adequação (*fitness*) classifica os grãos sonoros gerados em relação à respectiva região “alvo” e seleciona os mais “adaptados” para realizar troca genética com os demais e assim gerar novos grãos sonoros.

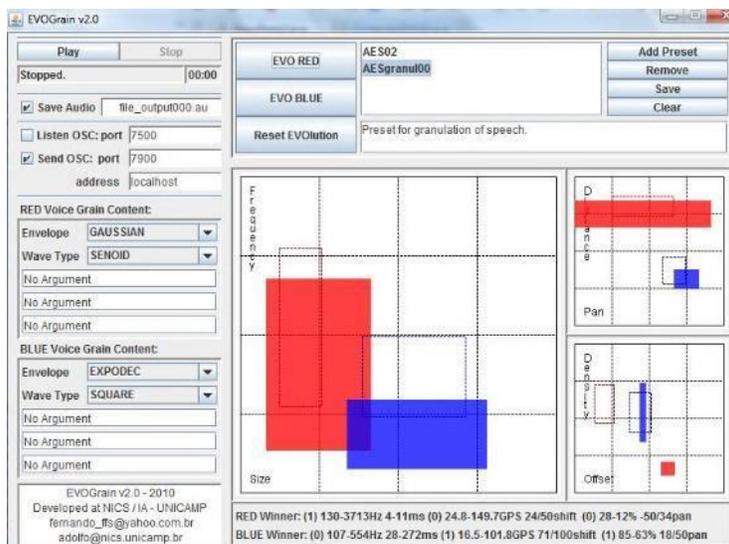


Figura 19: Sistema *EVOGrain*, desenvolvido por Souza e Maia (2011).

Diversos trabalhos em síntese concatenativa têm explorado a posição espacial como um parâmetro de controle. Nesse caso, atribui-se uma localização espacial virtual para cada segmento sonoro que se correlaciona à classificação do mesmo em relação aos dados extraídos dos descritores acústicos.

No sistema CataRT, do IRCAM, implementado em MAX/MSP, os *corpus sonoros* são distribuídos em um espaço bi ou tridimensional⁴⁵ em função da classificação dos descritores escolhidos pelo usuário. Através de uma interface (Figura 20), o usuário pode “navegar” neste espaço dos descritores, indicando o ponto ou região alvo de seleção dos grãos para síntese. Para cada grão selecionado atribui-se então uma posição aural no sistema multicanal em função correlata à sua posição no espaço dos descritores. Einbond & Schwarz (2010) descrevem algumas possibilidades de espacialização dos grãos a partir do sistema CataRT. Os grãos são agrupados em zonas, *subcorpus* sonoros que são espacializados de maneira independente. A síntese resultante pode ser espacializada sob a metáfora de uma única fonte ou múltiplas fontes sonoras. No primeiro caso, cria-se a ilusão de uma fonte sonora única deslocando-se no espaço enquanto que, no segundo, múltiplas fontes estáticas são metaforicamente excitadas por um corpo que se move no espaço.

⁴⁵ Neste sistema, a terceira dimensão é representada graficamente através de uma escala de cores.

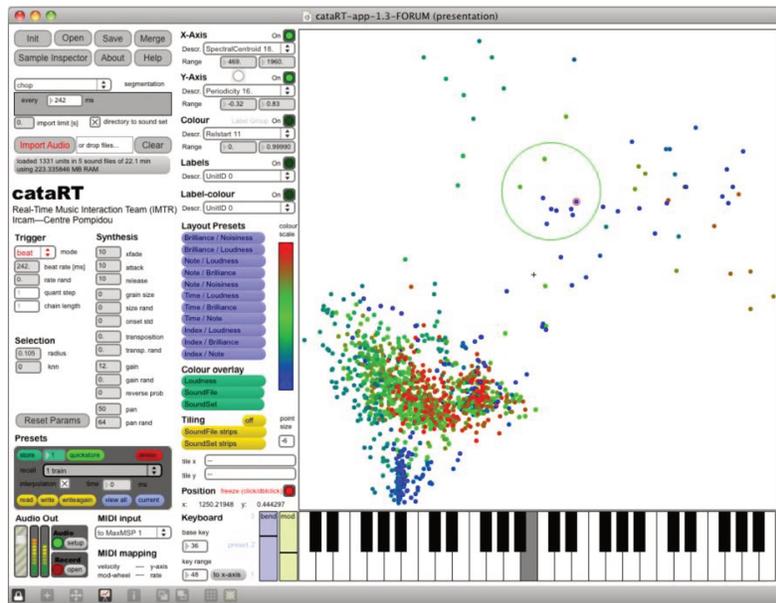


Figura 20: Sistema *CataRT*, do IRCAM, desenvolvido em MAX/MSP por Schwarz e colaboradores (2005)

3.8. Considerações

Neste capítulo apresentamos um panorama do microssom como abordagem aplicada à análise, síntese e composição musical. Relatamos inicialmente acerca do escopo teórico ao qual este está fundamentado, a teoria dos *quanta* acústico de Gabor e, em seguida, da decorrência do mesmo como técnica de síntese sonora nos contextos analógico e digital. Discutimos ainda algumas possibilidades de tratamento do microssom como paradigma composicional, ressaltando a abordagem utilizada por quatro compositores: Roads, Truax, Vaggione e Di Scipio. Ressaltamos que há muitas outras abordagens de síntese e controle de sons granulares, algumas delas simples modificações das que apresentamos acima. Por fim, apresentamos um estado da arte acerca da espacialização sonora no domínio microtemporal.

Capítulo 4

O SISTEMA GTSPAT

Neste capítulo apresentaremos nosso modelo formal para a espacialização de sons granulares em tempo real, e que resultou na criação do sistema GTSPat – Granular Timbre Spatialization. Como descrito no Capítulo I, um dos objetivos desta pesquisa foi o estudo e implementação de um modelo interativo aplicado à síntese e espacialização sonora no domínio micro-temporal, mais especificamente, espacialização de fluxos sonoros granulares.

O sistema GTSPat foi então concebido e desenvolvido como modelo de síntese e espacialização granular, pelo qual se estabelece uma função entre qualidades timbrísticas e a posição aural de cada grão sonoro na difusão em suporte eletroacústico. O método de síntese granular adotado como referência para este trabalho é conhecido por Granular Sampling (Lippe, 1994), que trata da obtenção de formas de onda a partir da segmentação de amostras de áudio em fragmentos de curta duração (geralmente entre 10 e 100 milissegundos) “janeladas” por uma curva de modulação de amplitude (gaussiana, exponencial, senoidal, dentre outras).

GTSPat foi implementado em Pure Data (PD)⁴⁶ e utiliza a biblioteca *iem_ambi* desenvolvida por Thomas Musil (2008) junto ao Instituto de Música Eletrônica e Acústica (IEM) da Universidade de Graz (KUG) para a espacialização sonora com Ambisonics (Gerzon, 1973). Os descritores acústicos utilizados baseiam-se em algoritmos adaptados da biblioteca *PDescriptors*⁴⁷, desenvolvida por Monteiro (2010) junto ao NICS/UNICAMP.

O sistema foi subdividido em três módulos distintos, sendo o primeiro um módulo de análise e classificação dos segmentos sonoros, o segundo de síntese e espacialização e o terceiro um módulo de interação (Figura 21). O usuário seleciona inicialmente uma amostra sonora e define os parâmetros de análise da mesma. No módulo de análise, a amostra de áudio é então seccionada em pequenos quadros (ou grãos), que são analisados pelo conjunto de descritores. Os dados obtidos dos descritores para cada grão formarão um vetor que será atribuído em função da posição espacial de cada grão. No módulo de interação o usuário define quais grãos serão enviados ao módulo de síntese e espacialização. Os grãos

⁴⁶ www.puredata.info/

⁴⁷ Disponível em <https://sites.google.com/site/pdescriptors/>. Acesso em 26 mar. 2012.

sonoros selecionados, juntamente com sua posição espacial, são endereçados ao módulo de síntese e espacialização, responsável pela codificação do sinal de entrada (monofônico) em relação à sua posição espacial e, em seguida, pela decodificação dos sinais resultantes para a quantidade de saídas de áudio do sistema.

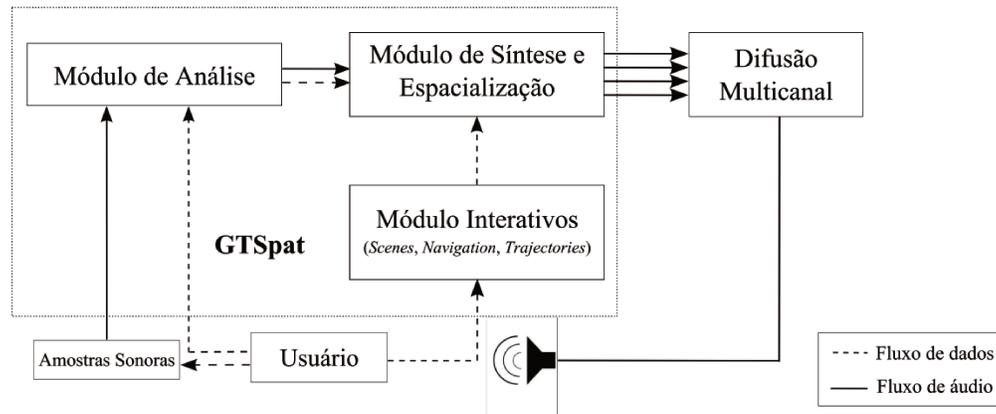


Figura 21: Arquitetura geral do sistema GTSpac

Desenvolvemos três distintos módulos interativos, cada qual explorando diferentes propriedades de síntese e espacialização como elemento de interação. Nos dois primeiros, procuramos explorar a localização espacial dos grãos como elemento de interação enquanto que, no terceiro, utilizamos as propriedades cinéticas na interação entre o sistema e o usuário como parâmetros de controle da síntese e espacialização.

Associado aos módulos principais do sistema GTSpac, construímos ainda dois submódulos adaptativos, pelos quais o sistema pode modificar o posicionamento dos grãos em decorrência da interação para com o usuário. O primeiro submódulo modifica a distribuição espacial dos grãos através do controle automático do vetor de características enquanto que, o segundo o faz por meio de processos de transformação timbrística e subsequente reenvio dos grãos para o módulo de análise.

4.1. Módulo de análise dos segmentos sonoros

A partir da informação unidimensional da amostra sonora digitalizada, podemos obter um enorme conjunto de medidas objetivas que se relacionam às múltiplas dimensões da percepção auditiva humana. Esta metodologia, emergente e relativamente recente no campo da Computação Musical,

aplica-se a diversos fins que vão desde a determinação de padrões através de medidas de similaridade até a classificação automática de gêneros musicais (CASEY et al., 2008; ORIO, 2006; TYPKE et al., 2005).

Em grande parte, os procedimentos encontrados na literatura para extração dos dados do sinal de áudio são padronizados, apresentando algumas variações de algoritmos e parâmetros. Da amostra digitalizada são tomadas sucessivas sequências de pontos, de tamanho invariável e sob uma taxa de sobreposição entre as mesmas⁴⁸. Para cada quadro ou sequência de pontos, ponderados por uma envoltória de amplitude⁴⁹, atribui-se então um *vetor de características*, resultante de um conjunto de medidas extraídas do sinal, as quais serão denominadas por descritores acústicos. (FOOTE, 1997).

Em nosso modelo, o vetor de características é composto por duas categorias de descritores: temporais e espectrais.⁵⁰ A primeira consiste de medidas computadas a partir do sinal representado no domínio do tempo, enquanto que a segunda baseia-se em informações computadas da representação do sinal no domínio das frequências, através da Transformada Discreta de Fourier (*Discrete Fourier Transform* – DFT). Realizamos as Transformadas de Fourier através do objeto (algoritmo) *tab_fft* da biblioteca *iem_tab*, uma versão otimizada da DFT denominada Transformada Rápida de Fourier (*Fast Fourier Transform* – FFT) implementada em Pure Data por Thomas Musil.

No sistema GTSpat foi utilizado um conjunto de seis descritores, sendo dois temporais (energia média e taxa de cruzamento por zero) e quatro espectrais (centroide espectral, variância espectral, obliquidade espectral, curtose espectral). Os conceitos e fórmulas aplicadas são detalhados no

⁴⁸ A taxa de sobreposição entre os quadros de análise é conhecida na literatura como *hop-size*, que se refere à quantidade de pontos tomados a cada início de quadro da Transformada Rápida de Fourier.

⁴⁹ O tipo de envoltória de amplitude aplicada no domínio temporal afetará diretamente a resposta em frequências da Transformada de Fourier. O janelamento do sinal justifica-se na literatura pelo fato de que, dada a natureza discreta do espectro de frequências (*bins*) obtido pela Transformada Discreta de Fourier (TDF), este não é capaz de reconstruir fielmente o espectro contínuo do sinal. Se, por exemplo, determinado pico de frequência encontra-se entre dois pontos discretos da TDF, a energia deste será distribuída aos pontos adjacentes, fenômeno este conhecido como *vazamento espectral*. De acordo com o teorema da amostragem de *Nyquist-Shannon*, a taxa de amostragem de um sinal necessita ser igual a $2F_m$ amostras por segundo, onde F_m é a maior frequência do sinal original. Com isso, a escolha por janelas temporais de menor duração (no presente caso, 10ms, equivalentes a 441 pontos de amostragem por TDF) implicará menor número de frequências e maior resolução temporal, ao passo que janelas com maior duração (200ms, equivalentes a 8820 pontos) implicará maior resolução temporal, porém menos sensível à variações no tempo. O tamanho da janela, como também o início e fim da amostragem são elementos preponderantes do vazamento de espectro. O início e a interrupção repentinos da amostragem de um sinal contínuo no tempo comprometem a resolução de frequência. O vazamento de espectro refere-se ao aparente espalhamento de energia em frequências adjacentes causado pelo início e pela interrupção na aquisição de dados (LATHI, 1998).

⁵⁰ Esta categorização faz referência à classificação apresentada por Peeters (2004).

Apêndice 8.1. Na Figura 22 mostramos um diagrama do procedimento algorítmico aplicado ao módulo de análise do sistema GTSpat.

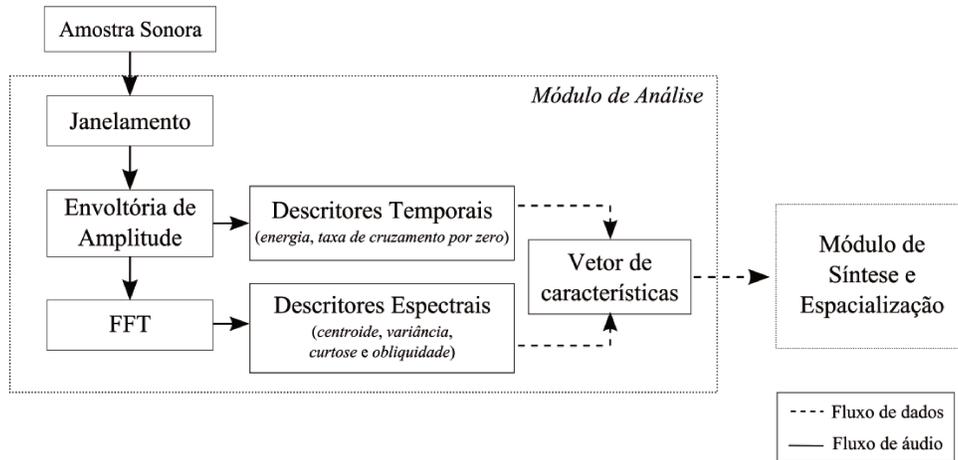


Figura 22: Módulo de análise do sistema GTSpat

A definição do tamanho do quadro de análise e a taxa de sobreposição podem ser previamente definidas pelo usuário. A quantidade de pontos analisados a cada quadro pode variar de 441 a 8820⁵¹, equivalentes a respectivos 10 e 200 milissegundos sob uma taxa de amostragem de 44,1 kHz. A taxa de sobreposição entre quadros pode variar entre 0 a 75 por cento do tamanho da janela temporal de análise. O tamanho dos quadros e a taxa de sobreposição dos mesmos determinarão, conseqüentemente, a quantidade de grãos sonoros obtidos a partir da amostra de áudio.

O usuário pode também definir qual tipo de envoltória de amplitude aplicar aos quadros de análise. A envoltória deve ser estabelecida anteriormente à análise, visto que o tipo de janelamento multiplicado ao sinal modifica a resposta em frequência da análise e pode, portanto, alterar o vetor de características de cada grão. Os tipos de envoltória disponíveis no sistema são: senoidal, gaussiana, exponencial inversa (*expodec*), exponencial (*expodec*), retangular e ADSR, conforme mostrado na Figura 23. Além disto, o usuário pode, a seu critério alterar duas constantes α e β presentes nas funções⁵² de envoltória e que modificam o formato da mesma.

⁵¹ Para os tamanhos de janela não correspondentes a valores de potências de dois, são adicionados zeros até que se complete a quantidade de pontos equivalentes à potência de dois superior mais próxima. Este procedimento é conhecido na literatura como *Zero Padding*.

⁵² A constante β não se aplica unicamente à envoltória senoidal.

Definidos tais parâmetros, cada quadro da amostra sonora é analisado por todos os descritores disponíveis ao sistema e os dados extraídos dos mesmos são armazenados em matrizes numéricas. Para cada grão sonoro analisado, definimos um vetor de características. A localização espacial de cada grão é dependente das entradas deste vetor que são convenientemente normalizadas⁵³ e ponderadas por pesos definidos pelo usuário.

Matematicamente, em nosso modelo, denotamos por n a quantidade de descritores do sistema e m a quantidade de grãos sonoros obtidos a partir de uma amostra de áudio. Seja d_{kj} o valor do descritor acústico k atribuído ao grão sonoro j , com $k = 1, 2, 3, \dots, n$ e $j = 1, 2, 3, \dots, m$ e ainda p_k , q_k e s_k pesos atribuídos ao descritor k , com $0 \leq p_k, q_k, s_k \leq 1$ e normalizados como $\sum_{k=1}^n p_k = \sum_{k=1}^n q_k = \sum_{k=1}^n s_k = 1$. Definimos a posição espacial (x, y, z) para cada grão j como:

$$x_j = \sum_{k=1}^n p_k d_{kj} \quad y_j = \sum_{k=1}^n q_k d_{kj} \quad z_j = \sum_{k=1}^n s_k d_{kj} \quad (4.1)$$

⁵³ Sabendo que os descritores são quantidades dimensionais e que só podemos somar quantidades com a mesma dimensão, optamos por definir os descritores normalizados pela quantidade adimensional d_{kj} , conforme descrita pela fórmula:

$$d_{kj} = \frac{D_{kj} - d_{kMIN}}{d_{kMAX} - d_{kMIN}}$$

onde D_{kj} é o valor absoluto do descritor k para o grão j , d_{kMAX} o valor máximo e d_{kMIN} o valor mínimo do conjunto de dados obtidos do descritor k .

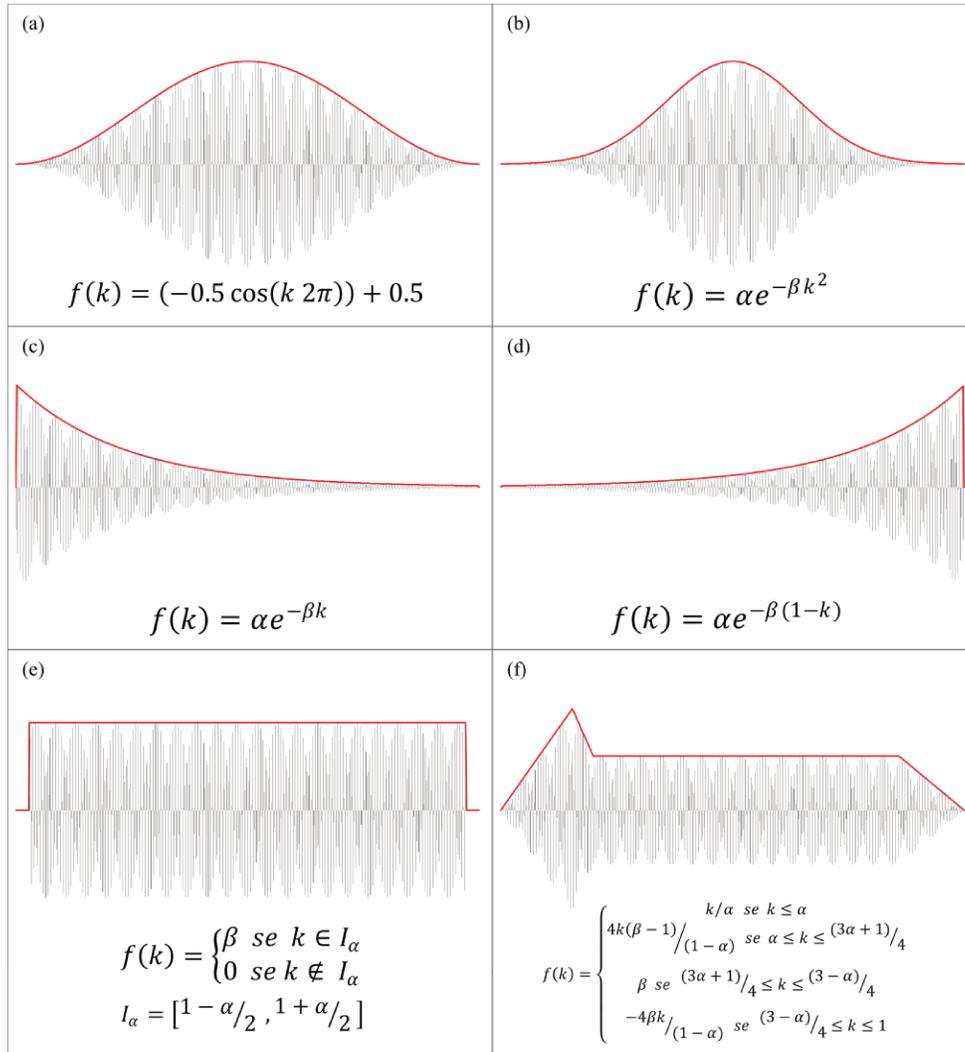


Figura 23: Envoltórias de amplitude do sistema GTSpat (em vermelho), com valores de k , α e β pertencentes ao intervalo $[0, 1]$: senoidal (a), gaussiana (b), rexpodec (c), expodec (d), retângulas (e) e ADSR (f).

O modelo admite distribuições espaciais bi- ou tridimensionais. Para o primeiro caso, adota-se $z_k = 0$, ou seja, a elevação da fonte sonora na espacialização é considerada constante. A Figura 24 ilustra a interface de controle de análise e distribuição espacial dos grãos do sistema GTSpat. Através desta, o usuário determina e pondera quais pontos do vetor de características serão variáveis da localização espacial de cada grão em cada eixo da espacialização sonora. Para cada descriptor, admite-se tanto sua forma normalizada d_k , como também seu dual simétrico \tilde{d}_k . O descriptor dual \tilde{d}_k lê a distribuição dos dados do descriptor k no sentido inverso, ou seja, $\tilde{d}_k = 1 - d_k$.

Na Figura 24 apresentamos a interface do Módulo de análise do sistema GTSpat. Nela, o usuário pode utilizar até quatro amostras sonoras e definir parâmetros de análise independentes para cada uma delas. Os pesos atribuídos a cada um dos eixos da espacialização são aplicados a todas as amostras analisadas.

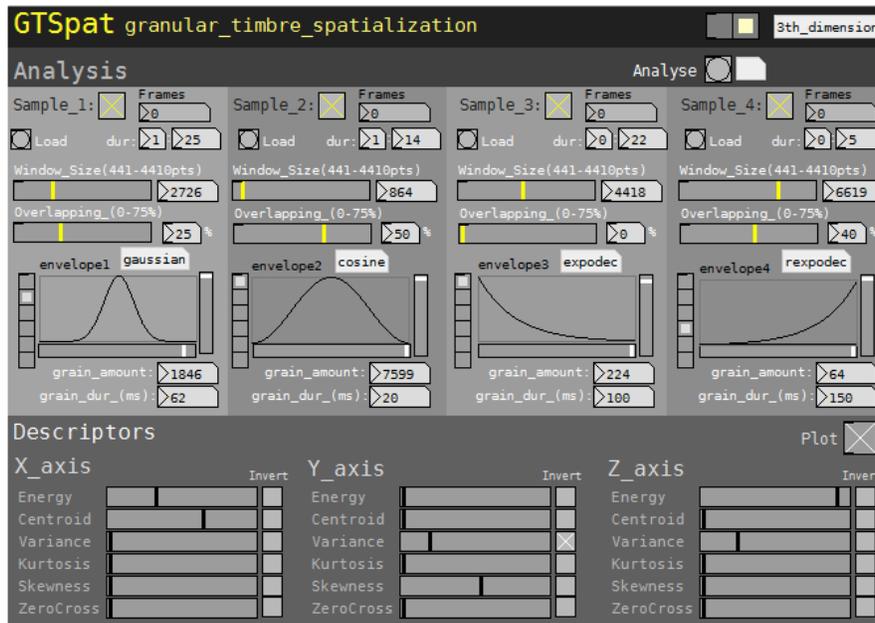


Figura 24: Interface do Módulo de análise do sistema GTSpat

Desenvolvemos ainda um submódulo de análise de amostra sonora gravada em tempo real (Figura 25). Os procedimentos utilizados neste seguem os mesmos aplicados ao Módulo de análise.



Figura 25: Submódulo de análise em tempo real do sistema GTSpat.

4.2. Módulo de Síntese e Espacialização

O módulo de síntese e espacialização é composto por três processos distintos: granulação, codificação e decodificação. O primeiro deles é responsável pela seleção do fragmento do áudio

amostrado e a multiplicação do sinal pela envoltória de amplitude pré-estipulada. O sinal de áudio é multiplicado (modulado) pela mesma envoltória de amplitude aplicada na análise dos descritores e, então, enviado ao codificador.

No codificador, o sinal monofônico enviado pelo granulador é renderizado para uma quantidade de canais que simulam o posicionamento espacial virtual do mesmo. Como referido anteriormente, este posicionamento é determinado pelos valores de coordenada espacial (x, y, z) atribuída a cada grão pelo vetor de características, que são convenientemente convertidas em coordenadas polares (θ, ϑ, r) , adequando-se assim à padronização da técnica Ambisonics. Os sinais renderizados pelo codificador são então enviados ao decodificador, que através de uma matriz de ganhos de amplitude, determina o quanto de cada um destes sinais codificados serão enviados a cada canal de saída do sistema.

A técnica de espacialização utilizada foi o Ambisonics de 2ª ordem. A escolha deste sistema justifica-se principalmente pela fácil portabilidade a diversas configurações de alto-falantes, pelo custo computacional relativamente baixo, como também pelo bom desempenho desta técnica na reconstrução de cenas espaciais tridimensionais.

Conforme descrito no item 2.4.5, esta técnica não contempla tradicionalmente a distância como um parâmetro na codificação espacial. Para isso, utilizamos filtros do tipo IIR (*Infinite Impulse Response*) para simular a atenuação da energia do espectro de frequências em decorrência da distância de cada fonte em relação ao ouvinte (cf. DANIEL, 2003). Desta forma, para cada função de codificação do Ambisonics, conforme Tabela 7, tem-se um fator adicional normalizado que modula o sinal monofônico s_i em função da distancia r_i , isto é, a distância do i -ésimo grão em relação ao ouvinte (*sweet spot*).

Após a codificação de distância, o sinal monofônico é então enviado ao codificador direcional. A codificação direcional do Ambisonics de 2ª ordem constitui-se de nove canais (W, X, Y, Z, R, S, T, U e V) que são por sua vez endereçados ao decodificador. Neste, o sinal de cada um dos canais de codificação é multiplicado por uma matriz de ganho que determina o quanto cada um destes será enviado para cada canal da decodificação. A matriz de ganho varia em relação à configuração de alto-falantes estipulada ao sistema. Em nossos experimentos, adotamos o padrão quadrifônico⁵⁴ para decodificação.

⁵⁴ A configuração padronizada da quadrfonia baseia-se em quatro alto-falantes, dois frontais (LF e RF) e dois posteriores (RB e LB), distribuídos simetricamente a -45° , 45° , 135° e -135° , respectivamente em relação ao centro (*sweet spot*).

A Figura 26 ilustra o processo de síntese e espacialização do sistema GTSpat. O sinal monofônico enviado por cada granulador é filtrado por um codificador de distância e, em seguida, codificado para nove canais em relação à sua informação direcional (θ, φ). Os nove sinais são enviados ao decodificador que redistribui o sinal destes para a quantidade especificada de canais da difusão sonora.

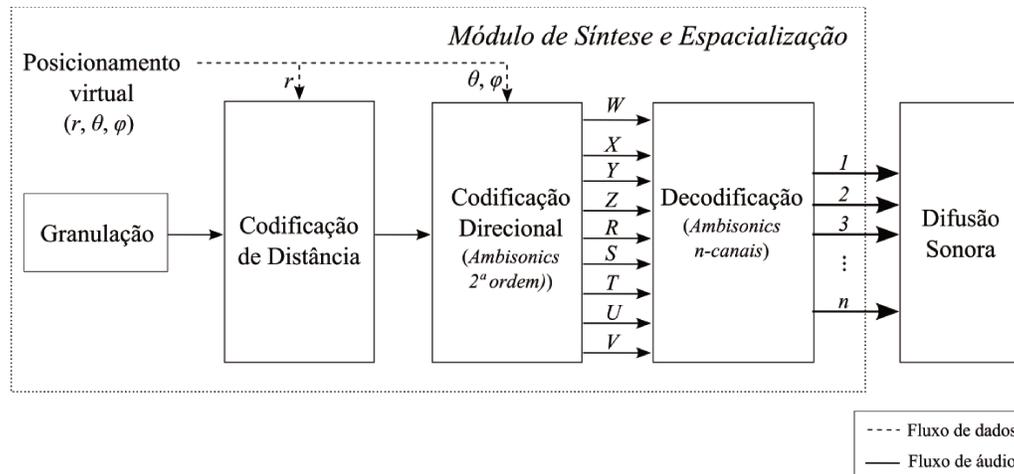


Figura 26: Módulo de Síntese e Espacialização do Sistema GTSpat

4.3. Módulos de interação

Através dos módulos de interação, o usuário pode controlar o sequenciamento do fluxo granular que será enviado ao módulo de síntese e espacialização. A fim de explorar distintas possibilidades de controle da espacialização, construímos diferentes propostas interativas das quais resultam os três módulos de interação: *GTSpat Scenes*, *GTSpat Navigation* e *GTSpat Trajectories*.

Cada um destes módulos explora determinadas propriedades de síntese e espacialização sonora. Nos dois primeiros, procuramos explorar a localização espacial dos grãos como elemento de interação. O primeiro baseia-se na metáfora de múltiplas fontes sonoras fixas no espaço, pelo qual o usuário detém o controle paramétrico de fluxos granulares distribuídos em distintas regiões da espacialização. O segundo baseia-se na metáfora de uma única fonte sonora que se desloca no espaço através do controle gestual do usuário. No terceiro módulo, procuramos salientar as propriedades cinéticas na interação entre o sistema e o usuário, utilizando a diretividade e taxa de deslocamento de movimento como parâmetros de controle da síntese e espacialização.

4.3.1. GTSpat Scenes

Neste módulo, o usuário detém o controle de diversos fluxos granulares espalhados através do espaço de difusão sonora. Tendo distribuído os grãos no espaço bi ou tridimensional de controle da espacialização através de seu vetor de características, o usuário pode então determinar sub-regiões deste espaço as quais definirão cada um dos objetos da cena espacial. Estas sub-regiões são representadas na interface por retângulos⁵⁵ (Figura 27), que podem ser manipulados pelo usuário (posição e área) a fim de se definir qual grupo de grãos formará o referido objeto na cena espacial. O sistema comporta até quatro objetos sonoros simultâneos, cada qual composto por de até dez fluxos granulares síncronos ou quase síncronos. A quantidade de fluxos e intervalo de tempo entre os grãos em cada um dos objetos são definidos pelo usuário, podendo este último variar entre 1 e 10 vezes a duração do grão nesta versão do GTSpat. Cada objeto possui ainda um controle individual de amplitude, que permite que o usuário determine diferentes níveis de intensidade para cada objeto na cena espacial.

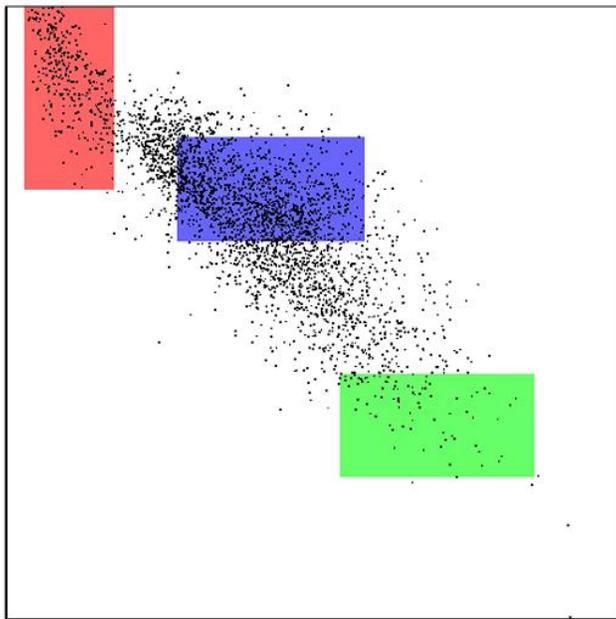


Figura 27: Interface de visualização dos grãos sonoros e objetos da cena espacial no GTSpat Scenes

⁵⁵ Nesta versão do sistema, optamos por uma representação bidimensional da espacialização, considerando a elevação constante das fontes sonoras.

Selecionada a área de um determinado objeto, o granulador de cada fluxo deste objeto pode sortear de maneira aleatória qualquer grão que esteja contido na área determinada para este objeto no espaço dos descritores. Todos os parâmetros definidos pelo usuário podem ser salvos no banco memória e serem recuperados a qualquer momento. O sistema tem capacidade de armazenar até 10 cenas diferentes, sendo ainda possível salvar as predefinições em formato de texto e carregá-lo posteriormente.

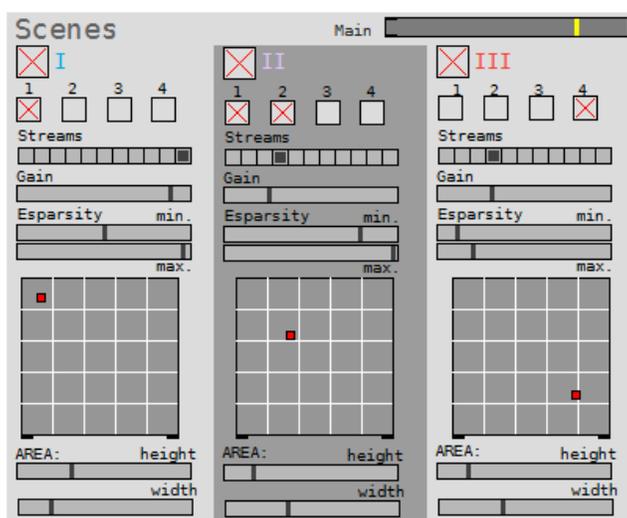


Figura 28: Interface de controle do GTSpat Scenes

Os comandos de controle da interface podem também ser controlados por outros softwares que estabelecem troca de dados via protocolos MIDI e OSC. No Apêndice 8.2.1 encontram-se listados as mensagens de comando aceitas pelo sistema.

4.3.2. GTSpat Navigation

Neste segundo módulo interativo, o sequenciamento temporal dos grãos sonoros como também a sua espacialização são controlados pelo usuário através de um “navegador bidimensional”. Este consiste de um quadrado virtual com tamanho definido pelo usuário e posição indicada através de uma interface de controle. A área demarcada pelo navegador no espaço bi ou tridimensional de controle da espacialização, ao qual denominaremos Espaço de Navegação, define a região alvo do sistema. Os

grãos localizados na região alvo no Espaço de Navegação são, então, enviados ao módulo de síntese e espacialização.

Abordamos dois diferentes tipos de interface de controle. O primeiro consiste em interfaces gestuais capazes de reportar ao sistema informações de posicionamento espacial (*mouse* ou controladores de jogos, por exemplo). O segundo tipo consiste num microfone, pelo qual o sistema recebe informações de áudio que serão processadas e analisadas pelo próprio sistema e, em seguida, transformadas em informação numérica referente à posição do navegador.

O processamento aplicado à entrada de áudio da interface de controle é similar ao aplicado às amostras sonoras no Módulo de Análise, conforme descrito em no item 4.1. O áudio captado pelo microfone é analisado em tempo real pelo mesmo conjunto de descritores do Módulo de análise. A soma do vetor de características resultante, ponderado pelos mesmos pesos atribuídos pelo usuário a cada eixo da espacialização, é então atribuída à posição do navegador no respectivo eixo.

Os dados absolutos obtidos da análise dos descritores são normalizados através de um procedimento de calibração do sistema. Dado certo intervalo temporal estipulado pelo próprio usuário, o sistema registra os valores mínimos e máximos encontrados para cada descritor. Os dados recebidos em subsequente pela interface de controle serão deste modo, normalizados entre $[0, 1]$, equiparando-se à distribuição normalizada dos grãos no Espaço de Navegação. Uma nova calibragem dos dados pode ser requisitada a qualquer momento pelo usuário, em que o sistema registrará novamente os valores mínimos e máximos encontrados e adequará, quando necessário, a função de normalização empregada.

O módulo de síntese para este modelo compreende 24 granuladores, comportando desta forma até 24 grãos sonoros executados simultaneamente⁵⁶. Os comandos de controle deste módulo também podem ser recebidos via mensagens dos protocolos MIDI ou OSC, conforme listado no Apêndice 8.2.3.

⁵⁶ Quando a quantidade grãos enviados num determinado instante excede a esta quantia, o sistema envia os dados excedentes com um atraso equivalente à duração do último grão sonoro da série anterior.

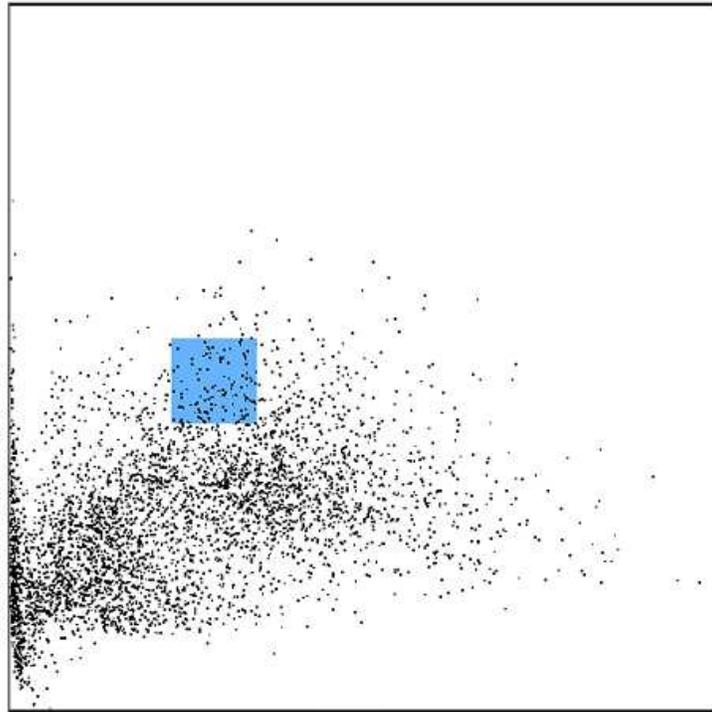


Figura 29: Espaço de Navegação e Navegador (quadrado azul) do sistema GTSpat *Navigation*

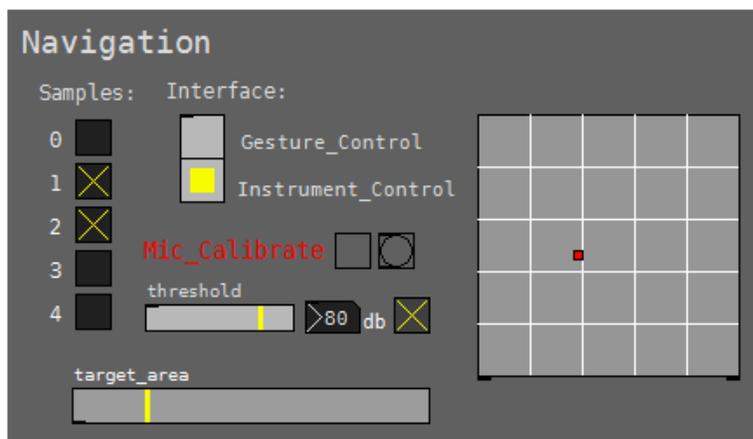


Figura 30: Interface de controle do sistema GTSpat *Navigation*

4.3.3. GTSpat *Trajectories*

Com a criação deste módulo, procuramos estabelecer outras possibilidades de interação do sistema. Utilizamos então medidas de segunda ordem das posições registradas pela navegação do usuário. As coordenadas cartesianas do Navegador, ao serem manipuladas pelo usuário no Espaço de

Navegação, são registradas no sistema como informações de trajetória, em magnitude (tamanho do deslocamento) e ângulo.

A partir desta trajetória, à qual denominaremos por “alvo”, o sistema compara as possíveis trajetórias dentre o conjunto de posições cartesianas atribuídas a cada grão no Espaço de Navegação. O sistema elege então a trajetória, dentre as possibilidades de deslocamento condicionado pela distribuição dos grãos, que mais se assemelhe da trajetória “alvo” segundo a função de comparação. A cada “passo” da trajetória no Espaço de Navegação o sistema compara as possibilidades de deslocamento dos grãos distribuídos no Espaço de Navegação e elege aquele que resultará numa trajetória “próxima” daquela estipulada pelo usuário como “alvo”.

Para a definição das trajetórias resultantes do sequenciamento dos grãos, utilizamos um processo markoviano⁵⁷ de decisão, pelo qual o sistema elege o grão subsequente do fluxo em função da posição espacial do grão atual. Para exemplificarmos, utilizaremos o exemplo representado na Figura 31. Dada a coordenada inicial (x_0, y_0) como sendo a posição espacial de um grão emitido pelo sistema. A partir de uma trajetória “alvo” \vec{T} definida pelo usuário, com magnitude $\|\vec{T}\|$ e ângulo de deslocamento σ_T , o sistema calcula as magnitudes e ângulos de todas as possibilidades de trajetórias em relação ao seu ponto inicial (x_0, y_0) . Dentre as possíveis trajetórias, condicionadas pela quantidade e distribuição dos grãos sonoros no Espaço de Navegação, o sistema escolhe a que mais se assemelhe à trajetória “alvo”. Supondo que no exemplo ilustrado na Figura 31 o sistema eleja o grão localizado em (x_1, y_1) como subsequente, uma nova medida será tomada pelo sistema, comparando as possíveis trajetórias a partir desta nova localização em relação ao “alvo”. No decorrer de sucessivos “passos” do trajeto, o vetor “alvo” pode manter-se fixo ou também alterar-se a cada novo estado.

⁵⁷ As cadeias de Markov são um caso particular de processo estocástico em que os estados anteriores são irrelevantes na predição dos próximos estados, desde que se conheça o estado atual. Num processo discreto de primeira ordem, os estados do sistema são condicionados por uma matriz de transição $P(m, n)$, que indica a probabilidade do sistema sair de um estado m para outro estado n . Em nosso modelo, adotamos $p(k) = 1/K$ para os grãos k de máximo índice l_{jt} (conforme equação $l_{jt} = h \left(1 - \frac{\|\vec{v}_j\| - \|\vec{T}\|}{\sqrt{2}} \right) + \tilde{h} \left(\frac{\sigma_j - \sigma_T}{\pi} \right)$, (4. 2), com K igual a quantidade de grãos que obtiveram tal índice, e $p = 0$ para dos demais grãos.

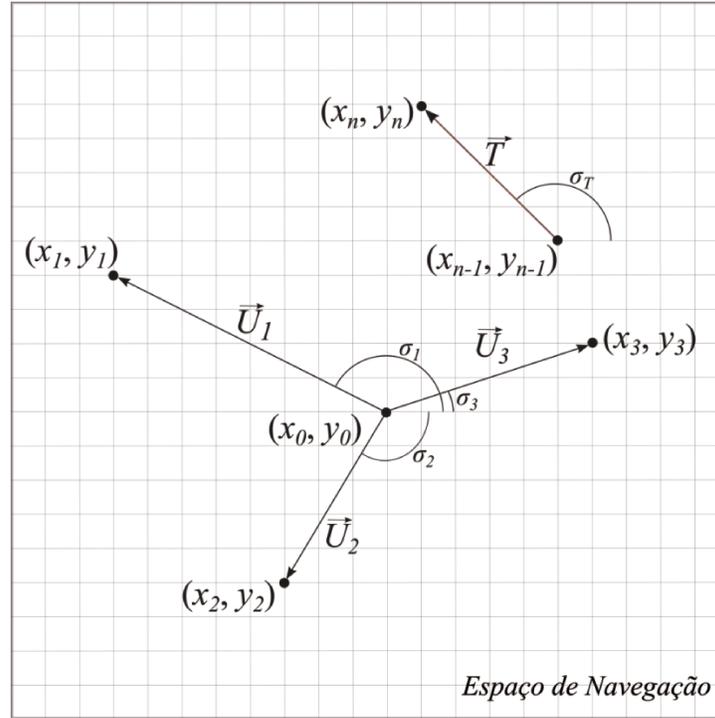


Figura 31: Representação vetorial das possíveis trajetórias a partir do ponto (x_0, y_0) comparadas à trajetória "alvo" (em vermelho).

Em nosso modelo, o sistema sorteia aleatoriamente um grão sonoro inicial de posição definida por (x_0, y_0) para cada fluxo granular (*streams*). Definimos $\vec{U}_j = (x_j - x_0, y_j - y_0)$ como o vetor formado entre a posição (x_0, y_0) do grão inicial emitido e a posição (x_j, y_j) de cada grão j ⁵⁸, com magnitude $\|\vec{U}_j\|$ e ângulo σ_j . Definimos ainda a trajetória "alvo" do sistema pelo vetor $\vec{T}_n = (x_n - x_{n-1}, y_n - y_{n-1})$ formado pelas diferenças sucessivas⁵⁹ de posição (x_t, y_t) do navegador, com magnitude $\|\vec{T}\|$ e ângulo σ_T .

A comparação entre a trajetória alvo e as possíveis trajetórias do sistema é dada pela função de comparação:

$$l_{jt} = h \left(1 - \frac{\|\vec{U}_j\| - \|\vec{T}\|}{\sqrt{2}} \right) + \tilde{h} \left(\frac{\sigma_j - \sigma_t}{\pi} \right), \quad (4.2)$$

⁵⁸ Com $j = \{1, 2, 3 \dots m\}$, referente à quantidade de grãos extraídos da amostra sonora.

⁵⁹ Definimos a taxa de amostragem pela duração temporal equivalente ao *hop-size* aplicado à amostra sonora no Módulo de Análise (cf. item 4.1).

onde h e \tilde{h} são pesos que podem ser definidos pelo próprio usuário, a fim de se atribuir maior ênfase ao tamanho ou à direção do trajeto “alvo”.

Nesta fórmula, o sistema calcula as trajetórias possíveis a partir do ponto (x_0, y_0) , com índices $\|\vec{U}_j\|$ e σ_j , mais se aproxima da trajetória “alvo” de magnitude $\|\vec{T}\|$ e ângulo σ_T indicada pelo usuário.

O sistema seleciona o grão j que obteve maior índice l_{jt} em relação à trajetória alvo e o envia ao granulador. Uma nova métrica é então calculada a partir da posição espacial do grão selecionado, comparando as possíveis trajetórias em relação à trajetória “alvo”. Na Figura 32, apresentamos a interface de controle do módulo GTSpat *Trajectories*.

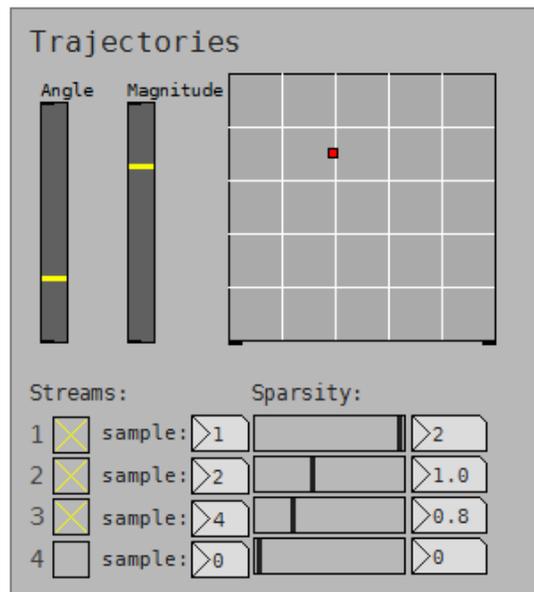


Figura 32: Interface de controle do módulo GTS Trajectories.

No Apêndice 8.2.4 estão listadas as mensagens aceitas por este módulo para controle via OSC ou MIDI.

4.4. Controle adaptativo da distribuição dos grãos

Associado aos módulos de interação do sistema GTSpat, construímos dois submódulos adaptativos. O propósito da criação destes submódulos foi de estabelecer novas instâncias de interação entre o sistema computacional e o usuário. Para isto, procuramos estabelecer estratégias nas quais o

sistema se adapta (ou se auto-organiza) em busca de um estado ótimo (distribuição dos grãos) que mais se aproxime das trajetórias do navegador no Espaço de Navegação. Em ambos os submódulos o sistema procura redistribuir os grãos no Espaço de Navegação a partir desses dados cinemáticos extraídos da navegação do usuário.

No primeiro submódulo, os pesos dos descritores associados a cada um dos eixos da espacialização são redistribuídos de maneira automática, por uma função distância euclidiana entre a distribuição do conjunto de dados de cada descritor e a distribuição dos pontos das trajetórias do navegador.

No segundo modelo, definimos uma função que mede o percentual de disparos de grãos enviados ao módulo de síntese e espacialização. Os grãos que apresentaram baixo percentual de disparos num dado intervalo temporal são submetidos a um processo de ressíntese e, mediante isto, são reenviados ao módulo de análise e redistribuídos no Espaço de Navegação. Nos tópicos a seguir descreveremos mais detalhadamente cada um dos submódulos.

4.4.1. Controle adaptativo do vetor de dados

Neste submódulo, o sistema mensura objetivamente alguns comportamentos do usuário enquanto este navega sobre a interface. A medida destes comportamentos será critério para uma possível redistribuição dos pesos atribuídos a cada descritor e, conseqüentemente, uma reconfiguração da disposição espacial dos grãos em função da interação do usuário. Na Figura 33 apresentamos o fluxograma de transmissão de informações ao primeiro submódulo adaptativo do sistema GTSpat.

O coeficiente de assimetria (As), conhecido como *Coeficiente de Assimetria de Bowley*, é dado pela razão entre a diferença e a soma das distâncias do segundo quartil ou mediana (Q_2) em relação ao primeiro (Q_1) e terceiro quartil (Q_3) da distribuição de pontos de trajetórias do navegador N . Matematicamente, podemos denotar como:

$$As_{Nx} = \frac{Q_{1Nx} + Q_{3Nx} - 2Q_{2Nx}}{Q_{3Nx} - Q_{1Nx}} \quad \text{e} \quad As_{Ny} = \frac{Q_{1Ny} + Q_{3Ny} - 2Q_{2Ny}}{Q_{3Ny} - Q_{1Ny}} \quad (4.5)$$

onde mostramos o coeficiente de assimetria para cada eixo coordenado do Espaço de Navegação. Esta medida representa o quanto a distribuição de pontos é assimétrica em relação à mediana. Valores negativos indicam assimetria para a esquerda (em direção a Q_1) e valores positivos indicam assimetria para a direita (em direção a Q_3).

As mesmas medidas extraídas da navegação do usuário são também realizadas para o conjunto de dados de cada descritor⁶⁰.

Depois de navegar durante certo intervalo de tempo, o usuário pode então solicitar que o sistema se adapte, ou seja, redistribua os grãos de maneira a aproxima-los dos pontos percorridos pelo navegador. Para isto, adotamos uma função distância entre o vetor (μ, Cv, As) de cada eixo do Espaço de Navegação e dos dados de cada descritor.

As distâncias⁶¹ entre o descritor k e a distribuição de pontos de trajetórias do navegador N nos eixos x e y do Espaço de Navegação são definidas por:

$$f(N_x, d_k) = \sqrt{\alpha(\mu_{Nx} - \mu_{dK})^2 + \beta(Cv_{Nx} - Cv_{dK})^2 + \frac{\gamma}{4}(As_{Nx} - As_{dK})^2}$$

$$f(N_y, d_k) = \sqrt{\alpha(\mu_{Ny} - \mu_{dK})^2 + \beta(Cv_{Ny} - Cv_{dK})^2 + \frac{\gamma}{4}(As_{Ny} - As_{dK})^2} \quad (4.6)$$

⁶⁰ Aplicamos a análise estatística aos descritores (em sua forma normalizada), como também para seus duais simétricos \tilde{d} .

⁶¹ Adotamos uma função distância inspirada na distância euclidiana, porém com pesos diferentes para as variáveis estatísticas (coeficientes α , β e γ). Observamos que o fator $1/4$ no terceiro termo é devido ao fato que a variável Coeficiente de Assimetria (As) está definido no intervalo $[-1, 1]$ e, portanto, o termo $(As_N - As_{dK})^2$ encontra-se no intervalo $[0, 4]$. Nesta versão do sistema GTSpat adotamos $\alpha = 3$, $\beta = 1$ e $\gamma = 4$, porém outros coeficientes poderão ser adotados posteriormente.

sendo α , β e γ são coeficientes reais positivos.

Os descritores que obtiveram menor distância em relação a cada um dos eixos das trajetórias percorridas pelo navegador no Espaço de Navegação são então selecionados.

A quantidade padrão de descritores que o sistema selecionará para a distribuição espacial de cada eixo foi padronizada em dois, porém o usuário pode alterar a quantidade de descritores entre 1 e 6 para cada eixo do Espaço de Navegação individualmente.

Dentre os descritores selecionados, o sistema atribuirá um peso entre 0 e 1 inversamente proporcional à distância apresentada por cada um, ou seja, aquele que apresentou menor distância terá portanto maior peso na distribuição espacial dos grãos. É possível ainda ao usuário determinar que a adaptação do sistema aconteça de maneira automática, isto é, ordenar que o sistema calcule novamente a distância a cada intervalo fixo de tempo.

a. Exemplo de aplicação do controle adaptativo do vetor de características

No exemplo abaixo, utilizamos uma amostra de áudio de 18 segundos para a geração de 738 grãos de duração de 50ms e taxa de sobreposição de 50%. Os gráficos da Figura 34 representam a distribuição dos dados extraídos de cada descritor e a

Tabela 9 apresenta a média, o grau de dispersão e grau de assimetria dos respectivos descritores.

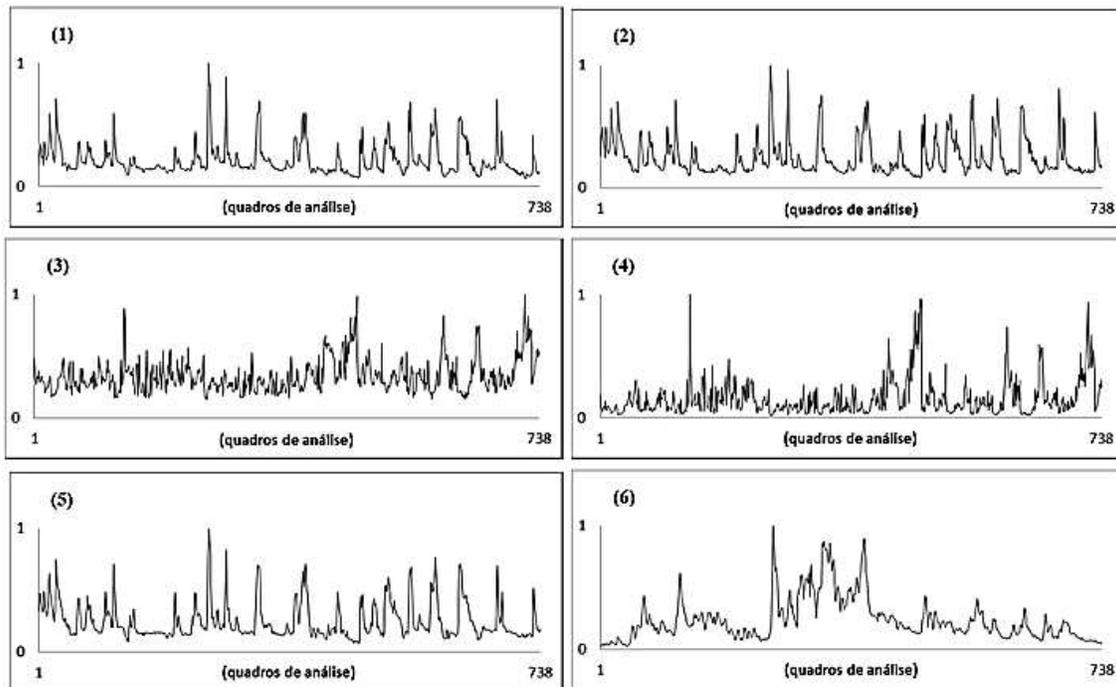


Figura 34: Distribuição normalizada dos dados extraídos dos descritores centroide espectral (1), variância espectral (2), obliquidade espectral (3), curtose espectral (4), taxa de cruzamento por zero (5) e energia média (6).

Tabela 9: Média, grau de dispersão e grau de assimetria na distribuição dos dados extraídos de cada descritor acústico.

	Centroide Espectral		Variância Espectral		Obliquidade Espectral		Curtose Espectral		Taxa de Cruzamento por zero		Energia Média	
	d	\bar{d}	d	\bar{d}	d	\bar{d}	d	\bar{d}	d	\bar{d}	d	\bar{d}
Média	0,2102	0,789	0,2471	0,752	0,3421	0,657	0,1535	0,846	0,249	0,75	0,2379	0,762
Dispersão	0,6163	0,383	0,614	0,385	0,3933	0,606	0,935	0,046	0,5965	0,403	0,7217	0,278
Assimetria	0,3554	-0,3554	0,47	-0,47	0,2327	-0,2327	0,3554	-0,3554	0,5206	-0,5206	0,1703	-0,1703

A atividade de navegação do usuário foi registrada durante um limite temporal e na Figura 35 são representadas as distribuições de posicionamento do navegador para cada um dos dois eixos do espaço dos descritores e na Tabela 10 a média, grau de dispersão e grau de assimetria da navegação.

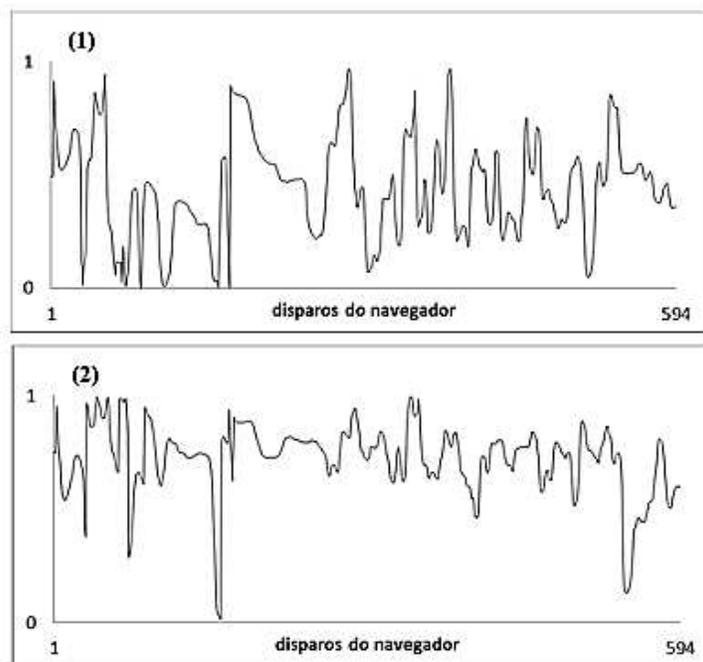


Figura 35: Exemplo de distribuição de posições no eixo X (1) e eixo Y (2) do navegador bidimensional do sistema GTSpac Navigation durante 594 disparos.

Tabela 10: Média, grau de dispersão e grau de assimetria na distribuição de posições nos eixos X e Y do navegador bidimensional do sistema GTSpac

	Eixo X	Eixo Y
Média	0,447	0,7236
Dispersão	0,4886	0,2165
Assimetria	-0,151	-0,139

Tabela 11: Distâncias entre os dados de cada descritor e a dados da trajetória espacial do navegador para cada eixo da espacialização

	Centroide espectral		Variância espectral		Obliquidade espectral		Curtose Espectral		Taxa de Cruzamento por zero		Energia Média	
	<i>d</i>	\bar{d}	<i>d</i>	\bar{d}	<i>d</i>	\bar{d}	<i>d</i>	\bar{d}	<i>d</i>	\bar{d}	<i>d</i>	\bar{d}
Eixo X	0,664	0,636	0,772	0,626	0,255	0,392	0,855	0,845	0,762	0,648	0,537	0,584
Eixo Y	1,093	0,296	1.1	0,375	0,698	0,417	1,327	0,347	1,12	0,427	1,028	0,095

Primeiramente o sistema seleciona a menor dentre as duas distâncias (d e \tilde{d}) aplicadas a cada um dos descritores e, em seguida, as compara selecionando as menores distâncias dentre todos os descritores, de acordo com a quantidade previamente estipulada de descritores a ser selecionados. Para este exemplo, utilizamos a predefinição de três descritores a serem atribuídos ao eixo X e dois ao eixo Y. Conforme os dados apresentados na Tabela 11, obteremos assim os seguintes descritores em função da espacialização: energia média, variância espectral e obliquidade espectral para o eixo X, e energia média e centroide espectral para o eixo Y.

Na Figura 36, demonstramos a distribuição inicial dos grãos no espaço dos descritores (primeiro quadro), a movimentação do navegador no espaço bidimensional da interface (segundo quadro) e, enfim, a redistribuição dos grãos por meio da redistribuição dos pesos dos descritores (terceiro quadro).



Figura 36: Distribuição inicial dos grãos (primeiro quadro), trajeto percorrido pelo navegador bidimensional (segundo quadro) e redistribuição dos grãos através de adaptação do sistema GTSpat (terceiro quadro).

4.4.2. Controle adaptativo através de ressíntese

Neste segundo submódulo adaptativo, utilizamos algumas ferramentas de transformação timbrística aplicada aos grãos sonoros como estratégia para a redistribuição destes no Espaço de Navegação. Diferentemente do primeiro submódulo adaptativo, o vetor de características atribuído para cada eixo da espacialização é mantido fixo, ao passo que os grãos modificados timbristicamente são reanalisados e, possivelmente, reposicionados no Espaço de Navegação⁶². Na Figura 37 apresentamos o

⁶² Dizemos *possivelmente* devido ao grau de imprevisibilidade que se tem na redistribuição dos grãos em função dos descritores a partir de um processo de ressíntese, de forma que não podemos prever quais serão as novas posições ocupadas pelos grãos modificados timbristicamente.

diagrama de interações do sistema GTSpac com o módulo de adaptação através de ressíntese dos grãos sonoros.

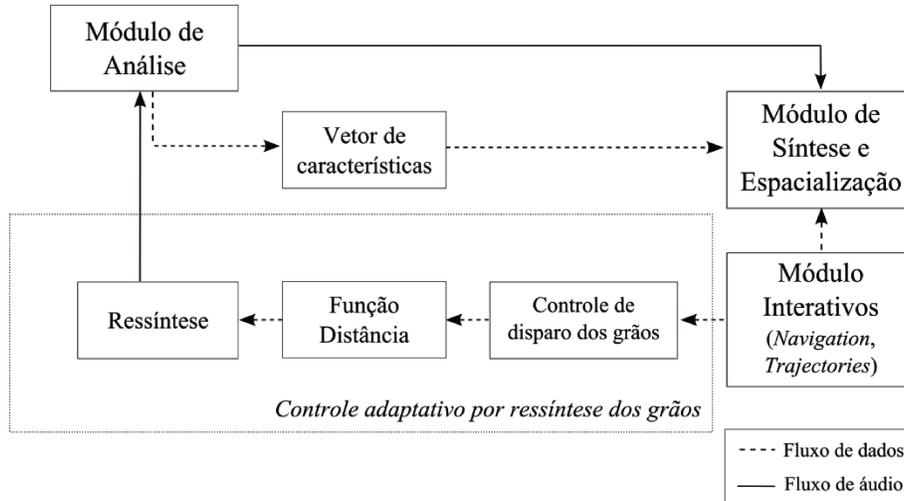


Figura 37: Diagrama de interação do submódulo de adaptação por ressíntese do GSTpat *Navigation*

a. Método de seleção dos grãos para Ressíntese

Como estratégia para determinação dos grãos a serem transformados timbricamente, definimos uma função que mede o percentual de disparos de grãos enviados ao módulo de síntese e espacialização. Após certo período de navegação, o usuário pode solicitar que o sistema se adapte. O sistema calcula, então, o percentual de disparos de cada grão e seleciona para ressíntese aqueles que obtiveram percentuais inferiores a certo limiar previamente estabelecido.

Para o cálculo da função que determinará quais grãos serão submetidos ao processo de ressíntese, normalizamos a quantidade de disparos de cada grão j pelo valor máximo de disparos encontrados. Assim, para o conjunto de valores percentuais de disparos dos grãos no intervalo $[0,1]$, definimos:

$$f_j = \begin{cases} -1, & \text{se } \frac{disp_j}{disp_{MAX}} \leq \alpha \\ 0, & \text{se } \alpha < \frac{disp_j}{disp_{MAX}} < \beta \\ 1, & \text{se } \frac{disp_j}{disp_{MAX}} \geq \beta \end{cases} \quad (4.7)$$

onde α e β são constantes⁶³, satisfazendo $0 \leq \alpha \leq \beta \leq 1$. O sistema selecionará os grãos que apresentaram $f_j = -1$ para ressíntese, convoluindo o sinal sonoro destes ao sinal de outro grão, selecionado dentre os que apresentaram $f_j = 1$.

Para determinar qual dentre os grãos que apresentaram $f_j = 1$ será selecionado para ressíntese de cada grão com $f_j = -1$, utilizamos uma função distância que compara as localizações dos grãos pertencentes a estes dois grupos no Espaço de Navegação. Assim, o sinal sonoro de cada grão com $f_j = -1$ será convoluído ao sinal do grão com $f_j = 1$ que apresentou menor distância em relação ao mesmo no Espaço de Navegação.

No exemplo mostrado na Figura 38, os grãos A e E apresentaram baixo percentual de disparos ($f_A = f_E = -1$) e os grãos B, C e D apresentaram alto percentual de disparos ($f_B = f_C = f_D = 1$). Calculada a função distância entre os grãos A e E, selecionados para ressíntese, e os grãos B, C e D, que obtiveram percentual de disparos superior a X, temos que $d(A, B) < d(A, D) < d(A, C)$ e $d(E, D) < d(E, C) < d(E, B)$, logo, o sistema elegerá o grão B para convolução do sinal do grão A e o grão D para convolução do grão E.

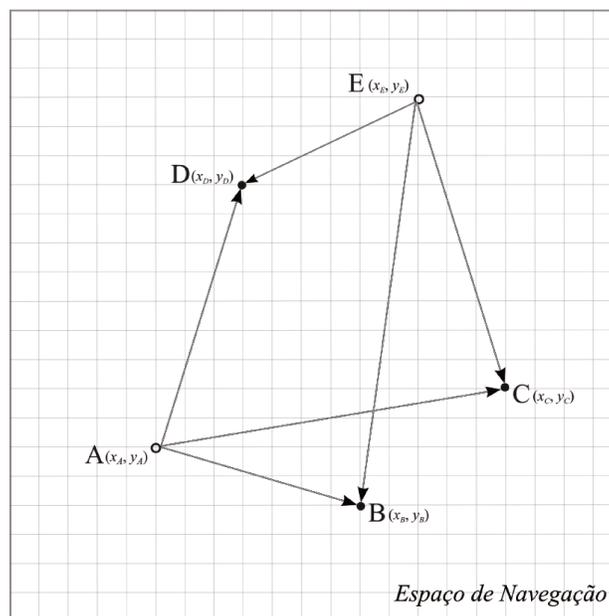


Figura 38: Distância entre os grãos selecionados para ressíntese e grãos com alto percentual de disparo no sistema GTSpat Navigation.

⁶³ As constantes podem ser definidas pelo usuário. Como padrão, adotamos $\alpha = 0,2$ e $\beta = 0,85$.

b. Síntese Cruzada (via Convolução de Fourier)

O processo de transformação timbrística aplicada a este modelo adaptativo foi o método de *Síntese Cruzada* ou *Ressíntese por Convolução de Sinais*. Através deste processo tem-se a convolução do espectro de um determinado sinal com o espectro de outro sinal. Pelas propriedades matemáticas de convolução, a transformada do produto de convolução entre duas funções é igual ao produto da transformada de ambas, ou seja, o produto ponto a ponto das transformadas de Fourier aplicada a cada uma das funções. No processamento de sinais de áudio, a Transformada de Fourier do produto de dois sinais no domínio do tempo é dada pela convolução das Transformadas destes sinais no domínio da frequência e, inversamente, a Transformada da convolução de dois sinais no domínio do tempo é levada no produto ponto a ponto das Transformadas de cada um dos sinais. Em outras palavras, a Transformada de Fourier da convolução entre sinais resulta na multiplicação normalizada das magnitudes de cada parcial harmônica obtida da transformada dos dois sinais analisados. O processo de ressíntese é, portanto, o retorno deste produto ao domínio do tempo, através da Transformada Inversa do sinal (*Inverse Fast Fourier Transform* - IFFT).

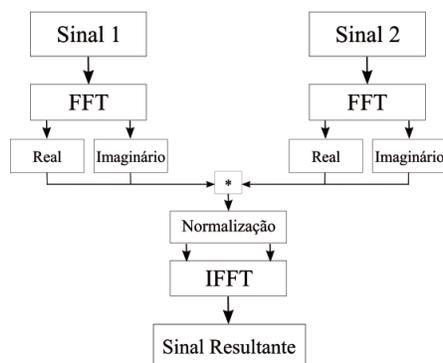


Figura 39: Processo de síntese cruzada do módulo adaptativo do sistema GTSpAt.

Conforme demonstrado na Figura 39, uma FFT é aplicada aos sinais 1 e 2. As coordenadas cartesianas (parte real e parte imaginária) de cada ponto da transformada do primeiro sinal são multiplicadas pelas coordenadas de cada ponto do segundo sinal. O produto das Transformadas é normalizado e, então, um novo sinal é produzido a partir da IFFT do mesmo. Após a ressíntese dos grãos, estes são novamente analisados pelo conjunto de descritores e, então, redistribuídos no Espaço de Navegação em função de seu novo vetor de características.

Capítulo 5

APLICAÇÕES DO SISTEMA GTSPAT

5.1. *interferência* (2012)

A obra *interferência*, para instrumentista e eletrônica em tempo real, foi composta mediante suporte do sistema GTSPat. A obra constitui-se de improvisação dirigida, em que é indicado ao intérprete que tipo de materiais e comportamentos sonoros realizar ao longo da improvisação. A peça é formada por duas camadas sonoras simultâneas. A primeira é formada pelos sons do instrumento acústico e a interação direta dos sons granulares e a segunda pela sequência de fluxos granulares que se modificam ao longo da obra pela leitura de uma partitura algorítmica, acionada indiretamente pela interação com o instrumentista. Utilizamos, respectivamente, os módulos *Navigation* e *Scenes* para a construção dessas duas camadas. Na Figura 40, demonstramos o fluxograma de dados e áudio presentes nesta obra.

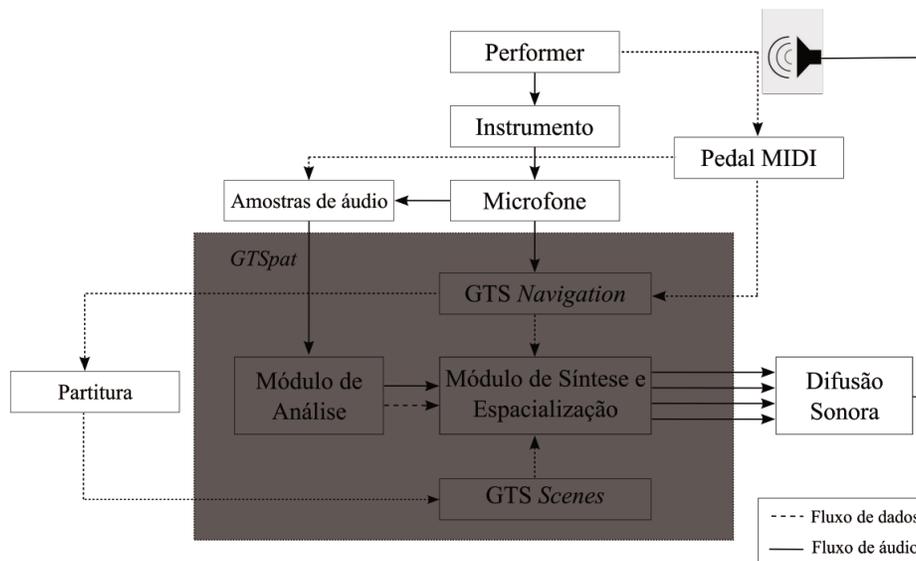


Figura 40: Fluxograma de informações (áudio e dados) da obra *interferência* (2012)

Foram utilizadas quatro amostras sonoras diferentes para granulação. Ao iniciar o *patch*, cada fragmento destas amostras é analisado pelo conjunto de descritores acústicos do sistema e distribuído

no Espaço de Navegação conforme seu vetor de características. Na Tabela 12, apresentamos os parâmetros de análise determinados para cada amostra.

Tabela 12: Definição de parâmetros de análise das amostras sonoras utilizadas em *interferência* (2012)

Amostra	Janela (pts)	Sobreposição (%)	Envoltória de Amplitude	Duração Total	Quantidade de Grãos
1	3528	30	ADSR	1m25s	1527
2	3969	50	Gaussiana	1m14s	1654
3	1304	50	Gaussiana	22s	1513
4	3969	0	Retangular	5s	63

Ao iniciar a improvisação, o som do instrumento, captado através de um microfone, é analisado em tempo real pelo mesmo conjunto de descritores aplicados às amostras selecionadas. Os dados extraídos em tempo real são normalizados a partir de uma calibragem realizada pelo próprio intérprete. O vetor de características normalizado controlará o Navegador, responsável pela concatenação dos grãos localizados no Espaço de Navegação.

Com o auxílio de pedais, o intérprete determina quais das amostras ele deseja utilizar, bem como o tamanho da área total do Navegador. O instrumentista pode ainda amostrar o som captado pelo microfone, através do módulo que analisa e distribui os grãos através do vetor de características extraído do sinal em tempo real. Outros parâmetros também são controlados através do uso de pedais, sendo eles: o início/fim da calibragem do vetor de características e o formato da envoltória de amplitude (controle das variáveis α e β) para cada amostra.

Utilizamos uma interface MIDI, composta por dez pedais de disparo e dois pedais de expressão. Os dez pedais que funcionam ao toque do intérprete e que podem ser programados para gerar até 100 comandos MIDI diferentes. Na Tabela 13, apresentamos os comandos enviados ao sistema GTSpa a partir do Pedal MIDI e na Tabela 14, apresentamos os comandos enviados a partir dos dois pedais de expressão.

Tabela 13: Comandos enviados ao sistema GTSpAt via interface MIDI

Programa Geral	Nº do Pedal	Código MIDI	Ação
00	1	0	Aciona a amostra 0
	2	1	Aciona a amostra 1
	3	2	Aciona a amostra 2
	4	3	Aciona a amostra 3
	5	4	Aciona a amostra 4
	6	5	Inicia nova calibragem
	7	6	Finaliza nova calibragem
	8	7	Inicia gravação de nova amostra 0
	9	8	Finaliza nova gravação de amostra 0

Tabela 14: Comandos enviados ao sistema GTSpAt através dos pedais de expressão da interface MIDI

Programa Geral	Nº do Pedal	Código MIDI	Pedal de Expressão A	Pedal de Expressão B
00	1	0		
	2	1		
	3	2	Tamanho da área do Navegador	Amplitude geral do GTSpAt <i>Scenes</i>
	4	3		
	5	4		
01	1	10	Controla variável α do envelope 0	Controla variável β do envelope 0
	2	11	Controla variável α do envelope 1	Controla variável β do envelope 1
	3	12	Controla variável α do envelope 2	Controla variável β do envelope 2
	4	13	Controla variável α do envelope 3	Controla variável β do envelope 3
	5	14	Controla variável α do envelope 4	Controla variável β do envelope 4

A segunda camada sonora foi criada a partir do módulo GTSpAt *Scenes* e seus parâmetros de controle são automatizados por uma partitura textual, que é “lida” pelo sistema ao longo da improvisação. A leitura dos comandos de cada linha da partitura está condicionado pela quantidade total de disparos de grãos foram enviados pelo GTSpAt *Navigation*, resultando assim numa interação indireta para com o improvisador. Na Tabela 15 apresentamos a Partitura Algorítmica na qual descrevemos os comandos de síntese e espacialização enviados ao módulo GTS *Scenes* em decorrência da quantidade de disparos de grãos.

Tabela 15: Partitura Algorítmica utilizada no módulo GTS *Scenes* em *interferência* (2012).

Etapa	Disparos	Parâmetros de Síntese	Parâmetros de Espacialização
0	0		X: Energy=0.798; Kurtosis = 0.55; ZeroCross = 1; Y: Centroid=1; Skewness=0.568; Invert_Centroid=1;
1	800	II: start=1; sample1=1; streams=8; gain=0.71; esparsity=(1, 3); position=(0, 0.71); size=(1.33, 4); III: start=1; stream=2; sample2=1; gain=0.5; sparsity=(1.34, 2.5); position=(1.56, 2.323); size=(0.38, 0.419);	
2	1400	II: sparsity=(1, 2); III: stream=3; position=(1.409, 2.704); size=(0.457, 0.457);	
3	1800	II: gain=0.2; size=(2.2, 4); III: gain=0.6; position=(0.952, 2.4); size=(1.219, 0.342);	
4	2100	II: start=0; sample1=0; III: gain=0.2; sparsity=(1.5, 6.5);	
5	2700	I: start=1; stream=5; sample2=1; gain=0.38; sparsity=(8.2, 10); position=(-1.79; 0.723); size=(1.142, 1.752); III: gain=0.15;	
6	3000	I: sparsity=(1, 10); position=(-0.11, 0.723); III: start=0; sample2=0;	
7	3200	I: sparsity(1, 5);	
8	3400	I: stream=10;	
9	3700	II: start=1; stream=8; sample2=1; gain=0.14; sparsity=(1, 1.8); position=(1.066, -0.94); size=(1.171, 2.247);	
10	4100	II: gain=0.27;	
11	4400	I: start=0;	
12	4800	II: start=0;	
13	5000	I: start=1; II: start=1;	
14	5300	I: start=0; II: start=0;	
15	5500	I: start=1;	
16	5600	II: start=1;	
17	5700	I: sparsity=(1, 2);	
18	5800	II: sparsity=(1, 1.1);	
19	6000	I: start=0; II: start=0;	
20	6100	III: start=1; stream=5; sample3=1; gain=0.304; sparsity=(5, 10); position=(2.628, 0.495); size=(2.895, 0.608);	X: Energy=1; Kurtosis=0 Invert_energy=1; Y: Invert_centroid=0; Invert_skewness=1;
21	6150	II: start=1; stream=1; sample3=1; gain=0.75; sparsity=(2.2, 2.2); position=(0.419, 0.126); size=(0.228, 0.227);	
22	6250	II: stream=3; gain=1; sparsity=(1, 3.4); position=(0.647, 0.736); size(0.076; 0.113);	

Etapa	Disparos	Parâmetros de Síntese	Parâmetros de Espacialização
23	6300	I: start=1; stream=10; sample3=1; gain=1; sparsity=(1, 4.4); position=(-1.48, -2.62); size(1.104, 0.571); III: stream=10; gain=0.4; sparsity=(1, 7.7); position=(0.723; -1.1); size(0.723, 0.569);	X: Energy=0.183; Centroid=1; Invert_energy=0; ZeroCross = 0; Y: Centroid=0; Sdeviation=0.412; Skewness=0; Zerocross=1; Invert_zerocross=1; Invert_skewness=0;
24	6600	III: start=0; sample3=0;	
25	6800	II: start=0; sample3=0;	X: Invert_centroid=1; Y: Invert_zerocross=0;
26	7000	I: gain=0.5; II: start=1; sample3=1; gain=0.523; stream=4; sparsity=(1, 7); position=(-1.67, 1.638); size=(0.57, 0.120);	
27	7200	I: gain=0.2; III: start=1; sample1=1; sample3=1; gain=0.876; stream=1; sparsity=(6.314, 9.65); position=(3.085, -1.93); size=(0.227, 0.721);	
28	7500	I: start=0; sample3=0; II: gain=0.3; III: sample3=0;	
29	7800	II: stream=2; gain=0.45; size(0.57, 1.217);	
30	8000	II: sparsity=(1, 7); stream=4; III: sample1=0; sample3=1; gain=0.5; stream=5; sparsity=(1, 4); position=(-0.72, -0.94); size=(0.684, 1.217);	
31	8100	III: stream=10; gain=0.65; sparsity=(1, 2.62);	
32	8150	I: start=1; sample2=1; gain=1; stream=10; sparsity=(1, 1.25); position=(0, 0); size=(4, 4); II: stream=10; sparsity=(1, 3); gain=0.8; position=(0.119, 0.046); size=(0.875, 1.217); III: size=(0.875, 1.217);	
33	8250	I: stream=2; gain=0.5; II: start=0; sample3=0; III: stream=2; sample1=1; sample3=0;	
34	8350	I: start=0; sample2=0; II: start=1; sample1=1; gain=1; stream=7; sparsity=(1, 2.457); position=(0.652, 0.723); size=(0.38, 4); III: start=0; sample3=0;	
35	8500		Y: Invert_zerocross=1;
36	8900	III: start=1; sample1=1; gain=1; stream=10; sparsity=(1, 2.36); position=(-1.25, -0.1); size=(0.379, 1.217);	
37	9100	I: start=1; sample2=1; gain=0.095; stream=8; sparsity=(1, 3.2); position=(0, 0); size=(4, 4);	
38	9200	II: start=1; sample1=1; gain=0.35; stream=9; sparsity=(1, 2.8); position=(0, 0); size=(4, 0.25);	

Etapas	Disparos	Parâmetros de Síntese	Parâmetros de Espacialização
39	9500	I: gain=0.45; stream=5; sparsity=(1, 5); II: gain=0.55; stream=6; sparsity=(1, 6);	X: Invert_centroid=0; Y: Invert_zerocross=0;
40	9800	I: gain=0.6; stream=3; sparsity=(1, 10); II: gain=0.75; stream=3; sparsity=(1, 10);	
41	10000	I: stream=1; sparsity=(7, 10); II: sparsity=(5, 10); gain=1; stream=2;	
42	10500	I: sparsity=(9, 10); II: sparsity=(8, 10); gain=1; stream=1;	X: Energy=0.642; Centroid=1; Variance=0.568; Skewness=1; Invert_energy=1; Y: Centroid=1; Sdeviation=0.715; Skewness=0.07; Zerocross=0.66; Invert_Zerocross=1;
43	11000	I: start=0; sample2=0; gain=0; II: start=0; sample1=0; gain=0;	
44	11500	III: start=0; sample1=0; gain=0;	

Na Figura 41, apresentamos o mapa de distribuição espacial dos grãos obtidos as amostras (de 1 a 4), resultante dos seis momentos de automação dos parâmetros de espacialização ao longo da performance (cf. Tabela 15). Salientamos para o contraste estabelecido entre os estágios iniciais (a-b) com maior concentração na região central da espacialização, seguido de estágios intermediários (c-e) com maior concentração espacial nas regiões periféricas da difusão e, por último, o retorno à concentração em regiões centrais da distribuição espacial (f).

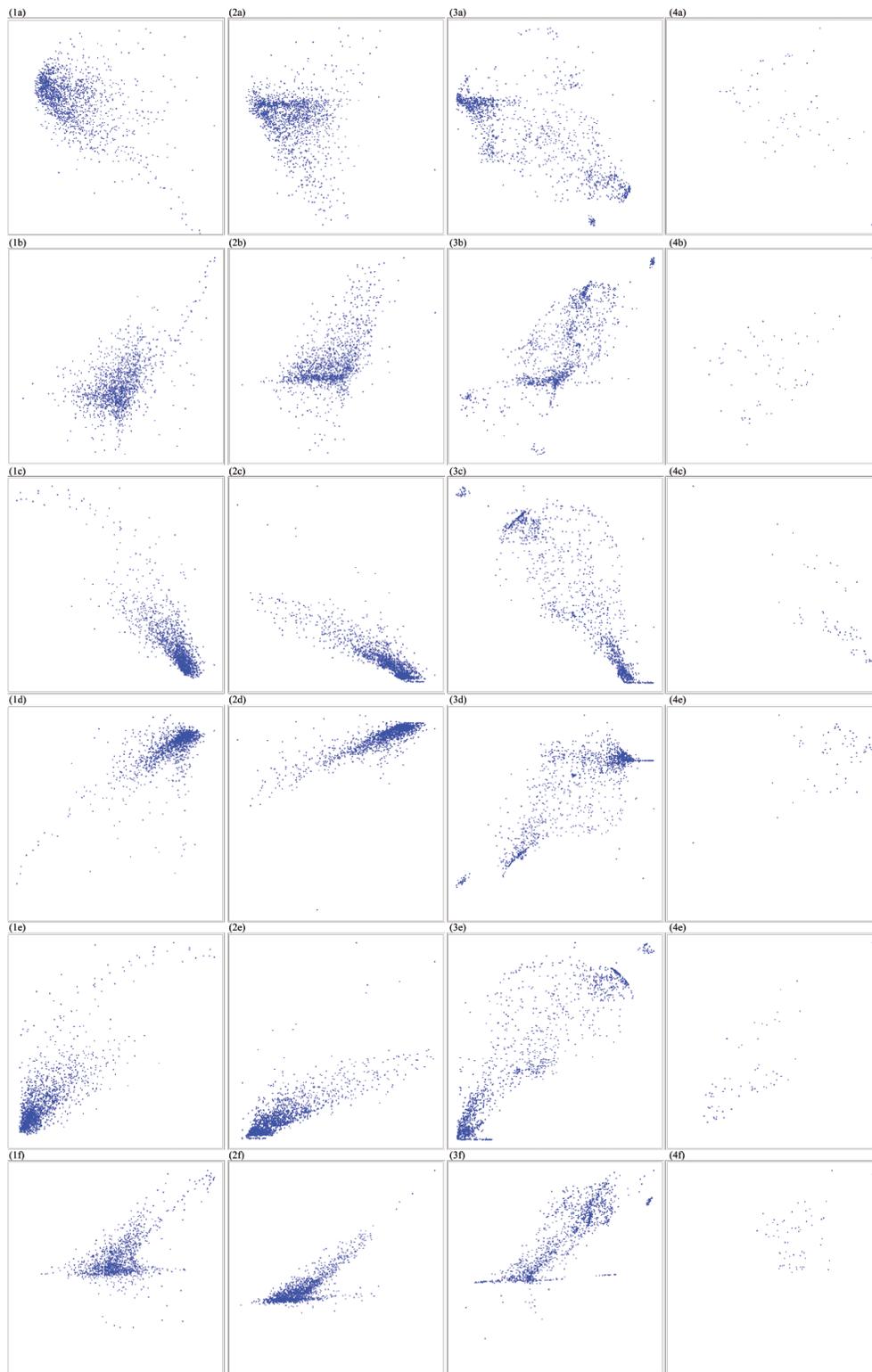


Figura 41: Mapas de distribuição espacial dos grãos sonoros das amostras (1-4) ao longo da peça interfereência (2012), subdividida em seis momentos (a-f).

Durante a improvisação, o intérprete recebe diversas instruções sobre que tipo de materiais e/ou gestos deve utilizar. São também indicadas quais amostras este deverá utilizar para granulação junto ao módulo GTSpat *Navigation*. Na Tabela 16 listamos as instruções dadas ao improvisador durante a performance, que estão relacionadas às etapas descritas anteriormente na Tabela 15.

Tabela 16: Lista de instruções da Partitura Algorítmica dadas ao improvisador em *interferência* (2012)

Etapa	Indicações sobre materiais e/ou Gestos Sonoros	Amostras utilizadas
0	Sons ruidosos de pequena duração	1
4	Contraste entre sons ruidosos e de altura definida	1, 2
6	Gradual crescendo de densidade	
10	Pico de densidade	
12	Baixa densidade, ruptura com seção anterior	
14	Alta densidade,	1, 2, 3
15	Valorização do silêncio / pequenos gestos	
18	Crescendo gradual	3
19	Saturação	
20	Sons percussivos e mais esparsos	2, 4
22		3, 4
23	Sons ruidosos e contínuos	
25	Quebra / ruptura / valorização do silêncio	
26	Decrescendo de densidade	
31	Gestos esparsos / alternância entre nota/ruído	2, 3, 4
35	Pequenos gestos com altura definida	2

A Figura 42 apresenta o espectrograma de uma performance realizada da peça. O áudio relativo a esta gravação pode ser escutado no CD em anexo, no qual constam três arquivos de áudio em estéreo, os dois primeiros relativos aos canais 1-2 e 3-4 e o terceiro contendo uma mixagem dos quatro canais da obra juntamente com o som captado do instrumento. Os traços em vermelho demarcam três comportamentos globais da peça, demarcados respectivamente pelas etapas 0, 20 e 25 da partitura. Na primeira seção, podemos observar um crescente acúmulo de energia na região baixa do espectro, que pela distribuição espacial definida apresenta maior energia no terceiro canal. Na segunda parte, observamos uma concentração da energia em frequências específicas e de aspecto mais contínuo que anterior. Por último, observamos a manutenção de comportamento contínuo da seção anterior nos canais 2 e 4 e sonoridades mais pontuais esparsas pelos quatro canais.

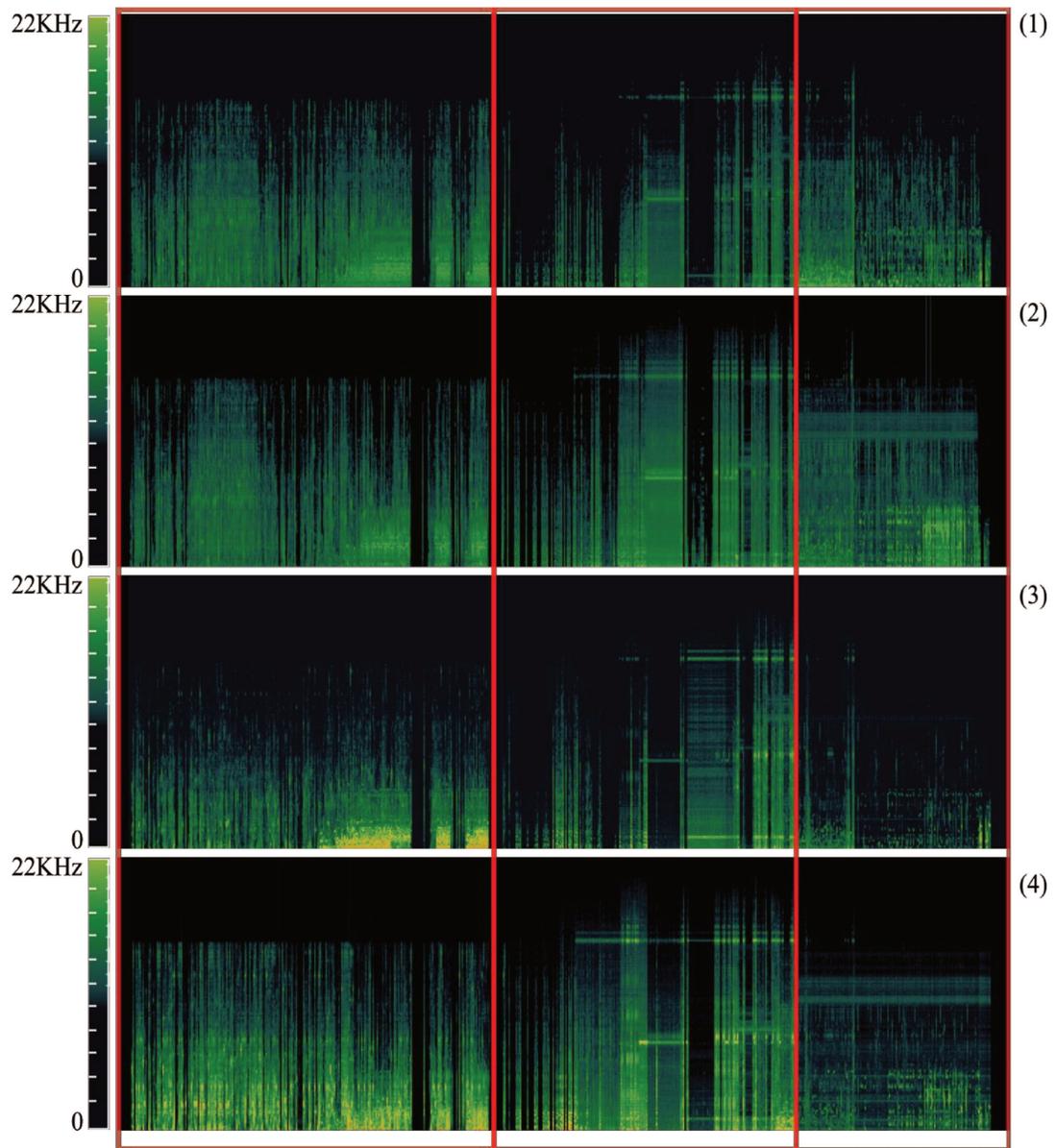


Figura 42: Espectrograma da parte eletrônica (4 canais) de performance de *interferência* (2012)

Capítulo 6

CONSIDERAÇÕES FINAIS

Apresentamos neste trabalho um estudo sobre modelos matemáticos aplicados à síntese granular e espacialização sonora em tempo real. Nos capítulos iniciais, apresentamos uma revisão da literatura sobre a temática relacionada, que fundamenta o desenvolvimento do modelo o qual denominamos GTSpat, um sistema interativo de síntese granular e espacialização sonora.

No Capítulo 2, relatamos inicialmente sobre os principais conceitos da escuta espacial, sob os quais estão baseadas as técnicas de espacialização também descritas neste capítulo. Com base no estudo destas técnicas, definimos a utilização do Ambisonics como plataforma de espacialização para nosso modelo. Neste capítulo também apresentamos um panorama sobre a prática da espacialização nos contextos da música instrumental e eletroacústica, que tiveram grande importância na decisão de quais elementos poderíamos explorar interativamente em nosso sistema.

De modo geral, a implementação computacional do modelo comportou-se de maneira bastante estável e apresentou resultados sonoros satisfatórios na distribuição dos grãos pela difusão sonora. Em trabalhos posteriores, pretendemos aprimorar o algoritmo de espacialização utilizado, sobretudo no que tange à codificação de informação de distância e na simulação de ambientes acústicos. Pelo custo computacional que isto implicava, optamos nesta primeira versão do sistema restringir a espacialização apenas à simulação de posicionamento das fontes sonoras. A codificação de informações relacionadas às primeiras reflexões e reverberação de fontes sonoras virtuais é elemento bastante rico à espacialização e muito contribui para o grau de realismo da cena acústica criada. Posteriormente, pretendemos ainda utilizar codificadores/decodificadores Ambisonics de ordem superiores, bem como aplicar técnicas híbridas de espacialização, que garantam maior definição da auralização de fontes sonoras virtuais. Pesquisas recentes têm apresentado bons resultados com a aplicação de sistemas híbridos, utilizando, por exemplo, o Ambisonics conjugado à técnica Wave Field Synthesis.

Acreditamos que as estratégias interativas por nós adotadas demonstraram-se bastante eficazes e sintetizam um rico conjunto de potencialidades aplicadas à criação musical. Cada módulo interativo desenvolvido explora diferentes características da espacialização, que são manipuladas através de

informações recebidas do usuário. Nos dois primeiros, *Scenes* e *Navigation*, delineamos processos reativos de processamento, enquanto que no terceiro módulo, *Trajectories*, utilizamos mecanismos computacionais de tomada de decisão pautados em dados de segunda ordem das informações recebidas. Através dos módulos adaptativos, estabelecemos ainda estratégias de organização interativa do sistema, que pode alterar-se mediante dados de segunda instância das informações recebidas do usuário.

Com a implementação de sistemas de espacialização mais sofisticados, poderemos ainda explorar interativamente diversos outros atributos espaciais. Dentre os tais, destacamos a manipulação em tempo real das propriedades acústicas da sala e mudanças de perspectivas da auralização pela utilização de *sweet spot* dinâmico.

O escopo teórico levantado no Capítulo 3, no qual discutimos sobre a representação microtemporal do som, foi um alicerce de nosso trabalho, vertendo-se tanto como ferramenta de análise e síntese sonora como também de criação composicional. Neste capítulo apresentamos ainda um estado da arte sobre espacialização no contexto microtemporal.

A extração de características do espectro como parâmetro de comparação entre segmentos sonoros já tem sido amplamente utilizada pelos métodos de síntese concatenativa descritos na literatura. Em nosso modelo, esta metodologia demonstrou-se também bastante eficaz como estratégia para a distribuição espacial em sistemas multicanal. Observamos, no entanto, que o conjunto de descritores utilizado em nosso modelo aborda a questão do timbre ainda de maneira elementar, levando em consideração apenas alguns aspectos objetivos extraídos do sinal de áudio. Com base nos resultados obtidos a partir desta pesquisa, poderemos ampliar o sistema com a utilização de descritores de mais alto nível, contemplando, sobretudo, aspectos psicoacústicos da percepção sonora.

Consideramos ainda que a metodologia de espacialização empregada pode ser ampliada a outros contextos que não propriamente o da síntese granular ou da síntese concatenativa. Procedimentos similares podem, por exemplo, ser utilizados na distribuição espacial de sons de longa duração, que se deslocariam no espaço da difusão em função das medidas extraídas em nível microtemporal.

No Capítulo 5 relatamos a elaboração da obra *interferência*, para instrumentista e eletrônica em tempo real e que teve como suporte de criação alguma das ferramentas desenvolvidas por esta pesquisa. Acreditamos que a obra serviu como uma primeira demonstração da funcionalidade do sistema e validação da metodologia empregada. No entanto, futuras aplicações poderão explorar mais as potencialidades do sistema e esperamos nestas a obtenção de resultados ainda mais satisfatórios.

Observamos que o desenvolvimento da pesquisa aqui reportada atendeu aos objetivos estabelecidos e o sistema GTSpac satisfaz as premissas estabelecidas no item 1.3. Através da disponibilização das ferramentas desenvolvidas, projetamos que o resultado desta pesquisa seja útil para compositores e interessados em criação sonora através de suporte tecnológico.

Em continuidade à nossa pesquisa, pretendemos estender o modelo elaborado pela utilização de diferentes interfaces que, sobretudo, estabeleça um diálogo mais direto entre o espaço físico da difusão e o espaço virtual das fontes sonoras. Outro possível desdobramento deste trabalho é a aplicação de algoritmos genéticos para o controle adaptativo da distribuição espacial. Nestes, a representação no domínio das frequências de cada grão sonoro poderá ser interpretada como informação genotípica, as medidas dos descritores como equivalente fenotípico e as informações de deslocamento espacial do intérprete/usuário como função de adequação do sistema.

BIBLIOGRAFIA

- BAYLE, F. La musique acousmatique, ou l'art des sons projetés. In: BAYLE, F. **musique acousmatique: propositions..positions.** Buchet/Chastel, 1993, p. 66. Paris: Buchet/Chastel, 1993. Texto original publicado em 1984 e revisado em 1993.
- BERG, J.; RUMSEY, F. Systematic Evaluation of Perceived Spatial Quality. **Proceeding of 24th International Conference: Multichannel Audio, The New Reality**, Canadá, 2003.
- BERG, P. PILE. A language for sound synthesis. **Foundation of Computer Music**, Cambridge, Massachussets, p. 160-187, 1985.
- BERKHOUT, A. J.; DE VRIES, D.; VOGEL, P. Acoustic control by wave field synthesis. **Journal of the Acoustical Society of America**, v. 93 (5), p. 2764-2778, 1993.
- BLAUERT, J. **Spatial Hearing: the Psychophysics of Human Sound Localization.** Tradução de John S. Allen. USA-Cambridge: MIT Press, 1996.
- CAGE, J. **Silence: Lectures and Writings.** Middletown: Wesleyan University Press, 1961.
- CHADABE, J. **Electronic Sound: The Past and Promise of Electronic Sound.** Upper Saddle River, NJ: PrenticeHall, Inc., 1997.
- CHION, M. Les deux espaces de la musique concrète. [S.l.]: [s.n.], v. 1, 1988. p. 31-33.
- CHOWNING, J. M. The Simulation of Moving Sound Sources. **Computer Music Journal**, v. 1 (3), p. 48-52, Jun. 1977.
- CHOWNING, J. M. Turenas: the realization of a dream. **Journées d'Informatique Musicale**, Université de Saint-Etienne, 2011.
- DANIEL, J. **Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia.** PhD thesis, Université Paris 6. 2000.
- DANIEL, J. Spatial Sound Encoding Including Near Field Effect: Introducing Distance Coding Filters and a Viable, New Ambisonic Format. **Proceedings of AES 23rd International Conference**, Copenhagen, Denmark, p. 23-25, 2003.
- DE POLI, G.; PICCIALI, A. Pitch-synchronous granular synthesis. In: POLLI, D.; PICDALLI; (ORG.), R. **Representations of Musical Signal.** Cambridge: MIT Press, 1991. p. 187-219.
- DI SCIPIO, A. Composition by exploration of nonlinear dynamic systems. **Proceedings of the International Computer Music Conference**, San Francisco: ICMA, 1990. 324-327.

-
- DI SCIPIO, A. Micro-Time Sound Design and Timbre Formation. **Contemporary Music Review**, Malasya, v. 10 (2), p. 135-148, 1994.
- DI SCIPIO, A. Compositional Models in Xenakis's Electroacoustic Music. **Perspectives of New Music**, v. 36, n. 2, p. 201-243, 1998.
- EMMERSON, S. New Spaces/New Places: a Sound House for the performance of electroacoustic music and sonic art. **Organised Sound**, v. 6 (2), p. 103-105, 2001.
- FURSE, R. First and second order Ambisonic decoding equations, 1999. Disponivel em: <<http://www.muse.demon.co.uk/ref/speakers.html>>. Acesso em: 23 jul. 2012.
- GABOR, D. Acoustical quanta and the theory of hearing. **Nature**, v. 159(4044), p. 591-594, 1947.
- GERZON, M. A. Periphony: With-Height Sound Reproduction. **Journal of Audio Eng. Soc.**, v. 21, p. 2-10, 1973.
- GLASGAL, R. Ambiophonics, 2004. Disponivel em: <<http://www.ambiophonics.org/>>. Acesso em: 12 nov. 2012.
- GOODMAN, N. **Languages of Art**. [S.l.]: Hackett Publishing Company, 1976.
- HARASZY, Z. Personalized Head-Related Transfer Function Measurement for Acoustic Virtual Reality Development. **U.P.B. Sci. Bull.**, v. Series C, Vol. 74, Iss. 3, 2012.
- HARLEY, M. A. **Space and Spatialization in Contemporary Music: History and Analysis, Ideas and Implementations**. Ph.D. diss., McGillUniversity: [s.n.], 1994.
- HARLEY, M. A. An American in Space: Henry Brant's "Spatial Music". **American Music**, v. 15 (1), p. 70-92, 1997.
- HARVEY, J. Mortuos Plango, Vivos Voco: A Realization at IRCAM. **Computer Music Journal**, v. 5 (4), p. 22-24, 1981.
- HARVEY, J. Mortuos Plango, Vivos Voco. **Computer Music Currents** 5, n. Wergo, WER 2005-2, 1990.
- HUSARIK, S. John Cage and LeJaren Hiller: HPSCHD, 1969. **American Music**, v. 1 (2), 1983.
- KENDALL, G. S. Spatial Perception and Cognition in Multichannel Audio for Electroacoustic Music. **Organised Sound**, v. 15 (3), p. 228-238, 2010.
- KENDALL, G. S. Spatial Perception and Cognition in Multichannel Audio for Electroacoustic Music. **Organised Sound**, v. 15 (3), n. Cambridge University Press, p. 228-238, 2010.
- KIM, J. E. A. New HRTFs (Head Related Transfer Functions) for 3D Audio Applications. **Proceeding of 188th Audio Engineering Society Convention**, Barcelona, 2005.

- KIM, S.; KONG, D.; JANG, S. Adaptive Virtual Surround Sound Rendering System for an Arbitrary Listening Position. **Journal of the AES**, v. 56, n. 4, April, 2008, v. 56 (4), April 2008.
- KOENIGSBERG, C. **Stockhausen's new morphology of musical time**. [S.l.]: [s.n.], 1991. Disponível em: <http://www.music.princeton.edu/~ckk/smmt>. Acesso em: 10dez2012.
- LATHI, B. P. **Signal processing and linear systems**. [S.l.]: Berkeley Cambridge Press, 1998.
- MALHAM, D. G. Approaches to spatialisation (tutorial article). **Organised Sound**, 3 (2), 1998. 167-177.
- MANZOLLI, J. Musical applications derived from FracWave sound synthesis method. **Proceedings of 94th Audio Engineering Society Convention, 1993, Berlin.**, 1993.
- MILLS, A. W. Auditory localization. In: TOBIAS, J. V. **Foundations of Modern Auditory Theory**. New York: Academic Press, v. 2, 1972. p. 301-345.
- MOORE, B. **An Introduction to the Psychology of Hearing**. 5 ed. ed. [S.l.]: Academic Press, 2003.
- OTONDO, F. Contemporary trends in the use of space in electroacoustic music. **Organised Sound**, 13(1). United Kingdom: Cambridge University Press, 2008, p. 77-81.
- PLASKOTA, P.; DOBRUCKI, A. The influence of pinna flare angle on Head-Related Transfer Function. **Signal Processing Algorithms, Architectures, Arrangements, and Applications (SPA)**, Sept. 2008. 43-46.
- PULKI, V.; KARJALAINEN, M.; VÄLIMÄKI, V. Localization, coloration and enhancement of amplitude panned. **Proceedings of 16th International Conference: Spatial Sound Reproduction**, Rovaniemi, Finland, 1999.
- PULKKI, V. Localization of Amplitude-Panned Virtual Sources II: three-dimensional panning. **Journal of Audio Eng. Soc.**, v. 49(9), p. 753-767, Sept. 2001.
- ROADS, C. **Microsound**. Cambridge: MIT Press, 2001.
- ROBINDORÉ, B.; ROADS, C. Forays into Uncharted Territories: An Interview with Curtis Roads. **Computer Music Journal**, Cambridge, v. 29 (1), p. 11-20, 2005.
- RUMSEY, F. Spatial Quality Evaluation for Reproduced Sound: Terminology, Meaning, and a Scene-Based Paradigm. **Journal of AES**, v. 50(9), p. 651-666, 2002.
- RUMSEY, F. Spatial Quality Evaluation for Reproduced Sound: Terminology, Meaning, and a Scene-Based Paradigm. **Journal of Audio Eng. Soc.**, v. 50 (9), p. 651-666, Sept. 2002.
- SMALLEY, D. Spatial Experience in Electro-Acoustic Music. In: DHOMONT, F. **L'espace du son II**. Ohain: Musiques et Recherches, 1991. p. 123-126.

-
- SMALLEY, D. Space-form and the acoustic image. **Organised Sound**, v. 12 (1), p. 35-58, 2007.
- STOCKHAUSEN, K. Music in Space. In: _____ **Die Reihe 5**. Tradução de Ruth Koenig. [S.l.]: [s.n.], 1961. Originalmente publicado em “Musik im Raum”, Die Reihe5 (1959).
- TOLONEN, T.; VÄLIMÄKI, V.; KARJALAINEN, M. **Evaluation of Modern Sound Synthesis Methods**. Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology. Espoo, Finlandia, p. 114. 1998.
- TOPPER, D.; BURTNER, M.; SERAFIN, S. Spatio-Operational Spectral (SOS) Synthesis. **Proceedings of the 5th Int. Conf. on Digital Audio Effects (DAFX-02)**, Hamburg, Germany, v. 133-137, Sept. 2002.
- TROCHIMCZYK, M. From Circles to Nets: on the Signification of Spatial Sound Imagery in New Music. **Computer Music Journal**, v. 25 (4), p. 37-54, 2001.
- TROCHIMCZYK, M. Trochimczyk, Maja. “From Circles to Nets: on the Signification of Spatial Sound Imagery in New Music. **Computer Music Journal**, v. 25 (4), p. 37-54, 2001.
- TRUAX, B. Chaotic non-linear systems and digital synthesis. An exploratory study. **Proceedings of of the International Computer Music Conference**, San Francisco: ICMA, 1990a. San Francisco: ICMA.
- TRUAX, B. Truax, B. (1990b) Composing with real-time granular sound. **Perspectives of New Music**, 28 (2), 1990b. 120-134.
- WARREN, R. M. **Auditory Perception: A New Analysis and Synthesis**. Cambridge, UK, and New York: Cambridge University Press, 1999.
- XENAKIS, I. **Musiques formelles**. Paris, Stock: [s.n.], 1963. 232 p. p. ISBN Revue Musicale n253-254. reeditado em 1981.
- XENAKIS, I. **Formalized Music: Thought and Mathematics in Composition**. Revised Edition. ed. Stuyvesant, NY: Pendragon Edition, 1992.
- ZACHAROV, K; KOIVUNIEMI, K. Unravelling the perception of spatial sound reproduction: Techniques and experimental design. **Proceeding of 19th International Conference: Surround Sound - Techniques, Technology, and Perception**, Schloss Elmau, 2001.
- ZACHAROV, N.; KOIVUNIEMI, K. Unravelling the perception of spatial sound reproduction: Techniques and experimental design. **Proceedings of the Audio Engineering Society 19th International Conference on Surround Sound**, Schloss Elmau, 2001.
- ZVONAR, R. A History of Spatial Music. **eContact!**, v. 7 (4), 2010. Disponível em http://cec.concordia.ca/econtact/Multichannel/spatial_music.html. Acesso em 20 ago. 2011.

APÊNDICE

8.1. Conceitos e Fórmulas dos descritores utilizados no sistema GTSpaT

a. Centroide espectral (Ce)

É o baricentro do espectro, o centro da massa do espectro sonoro detectado a cada quadro da Transformada Discreta de Forrier (TDF). O centroide espectral é calculado pela razão entre a soma de magnitude espectral de cada parcial da TDF e a soma total de magnitude espectral.

$$Ce_i = \frac{\sum_{k=0}^{N/2} k \cdot |x_i(k)|}{\sum_{k=0}^{N/2} |x(k)|}$$

Onde $x_i(k)$ representa a magnitude no quadro i para cada banda k de frequências e N é o número de pontos tomados a cada quadro da transformada rápida de Fourier (FFT). Tal parâmetro está relacionado à sensação de “brilho”, de forma que uma maior concentração de energia na região aguda é indicativa de um som mais “brilhante”.

b. Achatamento Espectral (Ae)

É a razão entre a média geométrica e a média aritmética de magnitude do espectro, como demonstra a equação abaixo.

$$Ae_i = \frac{[\prod_{k=0}^{N/2} |x_i(k)|]^{1/N/2}}{\frac{1}{N/2} \sum_{k=0}^{N/2} |x_i(k)|}$$

Quando normalizados, valores próximos de 1 indicam maior homogeneidade na distribuição da energia e qualificam, portanto, espectros mais ruidosos e sem formas claramente definidas. Quando próximos de zero, estes valores indicam heterogeneidade na distribuição e maior clareza da presença de picos de magnitude através do espectro, configurando-se assim de sons tonais ou senoidais.

c. Obliquidade Espectral (Oe)

É a medida de assimetria na distribuição da energia em relação à média. Quando igual a zero, tal índice sinaliza uma distribuição simétrica em relação à banda de frequência de magnitude média. Valores maiores que zero indicam assimetria para a direita, ou seja, maior desvio em relação às frequências superiores à média, ao passo que, valores menores que zero indicam assimetria para a esquerda, ou seja, maior desvio em relação às frequências inferiores à média. A obliquidade espectral é computada a partir do momento central de terceira ordem da distribuição, conforme demonstrado na equação abaixo.

$$Oe_i = \frac{\sum_{k=0}^{N/2} (x_i(k) - \mu_i)^3}{\sigma_i^3}$$

Sendo $x_i(k)$ a magnitude no quadro i para cada banda k de frequência, μ_i e σ_i a média e desvio da magnitude no quadro i .

d. Curtose Espectral (Ke)

É uma medida de dispersão da magnitude do espectro em relação à média que caracteriza o “pico” ou achatamento da curva de distribuição da magnitude. Seu cálculo é determinado pela razão entre o quarto momento central e o quadrado da variância da magnitude, conforme indicado na equação abaixo.

$$Ke_i = \frac{\sum_{k=0}^{N/2} (x_i(k) - \mu_i)^4}{\sigma_i^4}$$

e. Energia Média (E)

É a média quadrática (*Root Mean Square*, RMS) de amplitude um sinal. Para cada quadro i com N amostras tem-se:

$$\bar{E}_i = \sqrt{\frac{1}{N} \sum_{k=0}^{N-1} x_i^2(k)}$$

f. Taxa de Cruzamento por Zero (Z)

Indica o número de cruzamentos por zero do sinal no domínio do tempo. Este é um hábil descritor para diferenciação de fonemas sonoros ou sons fricativos na emissão de sons vocais, por exemplo. Seu cálculo é determinado pela diferença da função sinal na amostra $[n]$ e sua anterior $[n-1]$, conforme demonstra a equação abaixo.

$$Z_i = \frac{1}{2N} \sum_{n=0}^{N-1} |\text{sgn}(x_i[n]) - \text{sgn}(x[n-1])|$$

Sendo que:

$$\text{sgn}(x) = \begin{cases} x < 0, -1 \\ x = 0, 0 \\ x > 0, 1 \end{cases}$$

8.2. Mensagens processadas pelo sistema GTSpac (via comandos MIDI ou OSC)

8.2.1. GTSpac Analysis

Mensagem	Descrição
analysis/#1axis/descriptors/#2 f	Sendo #1 o eixo de espacialização (X, Y ou Z), #2 o descritor relacionado (<i>energy</i> , <i>zerocrossing</i> , <i>centroid</i> , <i>variance</i> , <i>kurtosis</i> e <i>skewness</i>) e f (entre 0 e 1) o peso atribuído ao mesmo.
analysis/#1axis/invert/#2 f	Sendo #1 o eixo de espacialização (X, Y ou Z), #2 o descritor relacionado (<i>energy</i> , <i>zerocrossing</i> , <i>centroid</i> , <i>variance</i> , <i>kurtosis</i> e <i>skewness</i>) e f (0 ou 1) a forma de leitura dos dados (normalizada ou dual, respectivamente)
analysis/envelope/type/#1 f	Sendo #1 a amostra selecionada (0, 1, 2, 3 ou 4) e f o tipo de envelope (de 0 a 5), sendo: senoidal (0), gaussiano (1), expodec (2), rexpodec (3), retangular (4) e ADRS (5).
analysis/envelope/alpha/#1 f	Controla a variável α (f entre 0 e 1) da função de envoltória na amostra #1 (0, 1, 2, 3 ou 4).
analysis/envelope/beta/#1 f	Controla a variável β (f entre 0 e 1) da função de envoltória na amostra #1 (0, 1, 2, 3 ou 4).
analysis/winsize/#1 f	Sendo #1 a amostra (0, 1, 2, 3 ou 4) do referido janelamento de quantidade f (441 a 8820) de pontos de amostragem.
analysis/overlapping/#1 f	Sendo #1 a amostra (0, 1, 2, 3 ou 4) à qual a sobreposição de janelamento f (de 0 a 75) se aplica.
analysis/iniciate bang	Inicia análise das amostras sonoras.

8.2.2. GTSpac Scenes

Mensagem	Descrição
scenes/start/#1 f	Sendo #1 o objeto selecionado (I, II ou III) e f (0/1) o comando de ligar/desligar do mesmo.
scenes/#1/#2 f	Sendo #1 o objeto selecionado (I, II ou III) e f a amostra (1, 2, 3 ou 4) a ser granulada.
scenes/#1/stream f	Sendo #1 o objeto selecionado (I, II ou III) e f a quantidade de fluxos granulares simultâneos (de 0 a 8).
scenes/#1/gain f	Controla a amplitude do sinal do objeto #1 (I, II e III).
scenes/#1/esparsity f1 f2	Sendo #1 o objeto selecionado (I, II ou III), f1 e f2, respectivamente, os valor mínimo e máximo de esparsidade (com distribuição aleatória neste intervalo) entre os grãos.
scenes/#1/position f1 f2	Sendo #1 o objeto selecionado (I, II ou III), f1 e f2, respectivamente, as posições centrais x e y dos retângulos de cada objeto.
scenes/#1/size f1 f2	Sendo #1 o objeto selecionado (I, II ou III), f1 e f2, respectivamente, a altura e largura dos retângulos de cada objeto.
scenes/main f	Controla o ganho geral de amplitude do módulo (f entre 0 e 1)

8.2.3. GTSpac Navigation

Mensagem	Descrição
navigation/start/#1 f	Seleciona a referida amostra #1 (0, 1, 2, 3 ou 4) para granulação.
navigation/mic/calibration f	Inicia e conclui a calibragem do microfone como interface de controle
navigation/target/position f f	Envia posição espacial (x, y) do Navegador.
navigation/target/area f	Envia o tamanho da área da região alvo do sistema com centro determinado por (x, y).

8.2.4. GTSpat Trajectories

Mensagem	Descrição
trajectories/target_position f f	Envia posição espacial (x, y) do Trajetória.
trajectories/stream/#1 f	Determina qual amostra sonora (0, 1, 2, 3 ou 4) será utilizada pelo fluxo granular #1 (1, 2, 3 ou 4)
trajectories/sparsity/1 f	Determina o grau de esparsividade
trajectories/weights/angle f	Determina o peso (entre 0 e 1) atribuído ao ângulo da trajetória alvo na definição das trajetórias
trajectories/weights/magnitude f	Determina o peso (entre 0 e 1) atribuído à magnitude da trajetória alvo na definição das trajetórias