LEANDRO FERRARI THOMAZ

**APLICAÇÃO À MÚSICA DE UM SISTEMA DE ESPACIALIZAÇÃO
SONORA BASEADO EM AMBISONICS**

**São Paulo
2007**

LEANDRO FERRARI THOMAZ

**APLICAÇÃO À MÚSICA DE UM SISTEMA DE ESPACIALIZAÇÃO
SONORA BASEADO EM AMBISONICS**

**Dissertação apresentada à Escola
Politécnica da Universidade de São
Paulo para a obtenção do título de
Mestre em Engenharia**

**São Paulo
2007**

LEANDRO FERRARI THOMAZ

**APLICAÇÃO À MÚSICA DE UM SISTEMA DE ESPACIALIZAÇÃO
SONORA BASEADO EM AMBISONICS**

**Dissertação apresentada à Escola
Politécnica da Universidade de São
Paulo para a obtenção do título de
Mestre em Engenharia**

**Área de Concentração:
Engenharia Elétrica
Sistemas Eletrônicos**

**Orientador:
Prof. Dr. Marcelo Knörich Zuffo**

**São Paulo
2007**

Este exemplar foi revisado e alterado em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, 27 de fevereiro de 2007.

Assinatura do autor _____

Assinatura do orientador _____

FICHA CATALOGRÁFICA

Thomaz, Leandro Ferrari

Aplicação à música de um sistema de espacialização sonora baseado em *ambisonics* / L.F. Thomaz. – ed.rev. -- São Paulo, 2007.

148 p.

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Sistemas Eletrônicos.

1.Acústica virtual 2.Áudio 3D 3.Processamento de sinais 4.Som espacial I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Sistemas Eletrônicos II.t.

AGRADECIMENTOS

Primeiramente, agradeço ao Prof. Dr. Regis Rossi Alves Faria, mentor e orientador deste trabalho, pela sua ajuda e constante estímulo transmitido durante todo o mestrado. Também agradeço a todo o pessoal da CAVERNA Digital, pelo suporte acadêmico, e ao Prof. Dr. Marcelo Knörich Zuffo, por possibilitar a realização deste trabalho.

Aos amigos e professores da ECA pelo auxílio musical dado durante a pesquisa e aos músicos que participaram nas gravações realizadas. Agradeço ao Prof. José Augusto Mannis e Pedro Paulo Kohler pela indispensável ajuda durante as gravações, e ao Prof. Dr. Fernando Iazzetta por disponibilizar o necessário para as gravações.

Aos meus grandes amigos Fernando Moya Orsatti e Rodrigo Carareto, pelas conversas acadêmicas, mútuas revisões dos trabalhos e pelas diversas discussões noctívagas, regadas a muitas bebidas, durante os últimos anos.

Agradeço especialmente ao amigo e maestro Otávio Simões pela inestimável ajuda e à Patrícia Carvalho pela revisão do texto.

Deixo também meus agradecimentos aos meus amigos e familiares que me incentivaram, mesmo não contribuindo de forma direta.

Finalmente, à minha mãe, Diva Ferrari Thomaz, que deu suporte e me incentivou ao longo dos anos.

RESUMO

O trabalho envolve a pesquisa de técnicas para espacialização sonora musical e o desenvolvimento de um sistema de espacialização sonora sobre uma arquitetura de áudio multicanal aberta (AUDIENCE) para a auralização imersiva. Utilizando-se a técnica Ambisonics, o sistema construído permite a recriação de um campo sonoro tridimensional. O sistema reproduz virtualmente peças musicais espacializadas, por meio de um arranjo de alto-falantes em torno do ouvinte, onde as fontes sonoras podem assumir posições incomuns e movimentar-se no espaço. Através dos resultados dos experimentos musicais realizados com o sistema, chega-se a conclusão de que o sistema pode ser utilizado por compositores, regentes e produtores para a tarefa de espacialização musical.

Palavras-chave: Acústica Virtual; Áudio 3D; Processamento de Sinais; Som Espacial.

ABSTRACT

This work involves the research of techniques for musical spatialization, and the development of a spatialization system, based on an open multichannel audio architecture (AUDIENCE) for immersive auralization. The constructed system uses the Ambisonics technique, which allows the recreation of a three-dimensional sound field. The system reproduces spatial music through an arrangement of loudspeakers around the listener where sound sources can assume some uncommon positions, and also be put into motion. From the results of experiments with a musical piece it follows that the system can be used by composers, conductors, and producers to the musical spatialization task.

Keywords: Virtual Acoustic; 3D Audio; Signal Processing; Spatial Sound.

LISTA DE FIGURAS

Figura 1 – Posicionamento de fontes sonoras virtuais envolvendo o usuário	14
Figura 2 – Alguns atributos para avaliar a percepção espacial em uma sala de concerto.....	22
Figura 3 – Posicionamento não espacializado, com uma frente sonora única sendo percebida (quarteto de cordas e orquestra).....	24
Figura 4 – Basílica de São Marcos, em Veneza, onde os dois coros se posicionam nos balcões laterais	25
Figura 5 – Trecho de um <i>cori spezzati</i> de G. Gabrieli, <i>Hodie Christus natus est</i> , para 2 coros.....	26
Figura 6 – Ensaio da peça <i>Gruppen</i> , de K. Stockhausen (1957), para três orquestras em torno da audiência	28
Figura 7 – Diagrama com a disposição dos instrumentos na peça <i>Voyage Four</i> , de H. Brant (1964).....	30
Figura 8 – Indicação na partitura para distribuição dos instrumentos entre a audiência na peça <i>Terretektorth</i> , de I. Xenakis (1965-66).....	33
Figura 9 – Pavilhão alemão esférico com orquestra de alto-falantes construído para a apresentação de <i>Spiral</i> , de Stockhausen, na EXPO 70 em Osaka	36
Figura 10 – Esquema de montagem quadrafônica utilizada pelo Pink Floyd no início da década de 70	37
Figura 11 – Sistema de espacialização computacional utilizado para a reprodução de <i>Répons</i> , de Boulez (1985)	38
Figura 12 – Posicionamento das fontes sonoras e movimentação, na peça <i>Répons</i> , de Boulez (1985)	39
Figura 13 – Diversos tipos de sistemas de espacialização sonora e sua generalização	42
Figura 14 – Exemplos de <i>softwares</i> para espacialização sonora.....	44

Figura 15 – Exemplos de mesas de som com espacialização sonora	45
Figura 16 – Disposição dos alto-falantes no formato 5.1 e 7.1, de acordo com a norma ITU-R BS.775.1	47
Figura 17 – Funções de transferências dos ouvidos esquerdo e direito	49
Figura 18 – <i>Dummy Head</i> KU100, da Neumann	50
Figura 19 – Montagem dos sistemas estéreo e dipolo estéreo	52
Figura 20 – Vetores e posicionamento de alto-falantes no SVPA	55
Figura 21 – Montagem Ambiophonics completa, utilizando os 6 canais	58
Figura 22 – Montagem WFS frontal, na qual uma matriz de alto-falantes reproduz a onda original	59
Figura 23 – Montagem WFS circular para gravação e reprodução, com uma matriz densa de microfones e alto-falantes	60
Figura 24 – Sistema Soundfield SPS422B (com detalhe da cápsula interna do microfone), para captura de campo sonoro em B-Format de 1ª ordem.....	62
Figura 25 – Coordenadas esféricas utilizadas nas equações de codificação Ambisonics	63
Figura 26 – Montagem Ambisonics octogonal de alto-falantes	67
Figura 27 – Diagrama de blocos de um decodificador Ambisonics de 1ª ordem com filtros psicoacústicos	68
Figura 28 – Um <i>patch</i> de síntese FM feito no Pure Data	73
Figura 29 – Diagrama UML representando as classes do PD	75
Figura 30 – Arquitetura genérica de camadas do AUDIENCE	77
Figura 31 – Casos de uso do sistema de espacialização sonora.....	79
Figura 32 – Implementação do AUDIENCE baseado em um sistema de espacialização sonora genérico	80
Figura 33 – Implementação da arquitetura AUDIENCE	81
Figura 34 – Diagrama UML das classes desenvolvidas para o AUDIENCE	85
Figura 35 – Componentes gráficos <code>audce_gui_grid</code> e <code>audce_gui_srcfile</code>	86

Figura 36 – Algumas configurações de alto-falantes disponíveis para decodificação no bloco <code>audce_L4_amb_3rd~</code>	89
Figura 37 – Desenho do posicionamento imaginado pelo compositor para seu trio .	91
Figura 38 – Construção da camada 1 do <i>patch</i> exemplo	92
Figura 39 – Construção da camada 2/3 do <i>patch</i> exemplo	92
Figura 40 – Construção da camada 4 do <i>patch</i> exemplo	93
Figura 41 – <i>Patch</i> de segunda ordem em um arranjo quadrado, com gráficos mostrando os canais Ambisonics e a saída dos alto-falantes	95
Figura 42 – Configuração octogonal horizontal de decodificação Ambisonics	96
Figura 43 – Montagem física do sistema cúbico na CAVERNA Digital	97
Figura 44 – Fluxo de montagem do experimento	98
Figura 45 – Ensaio da peça <i>The Unanswered Question</i> , no auditório do departamento de música da ECA-USP	103
Figura 46 – Gravação da peça <i>The Unanswered Question</i> , no estúdio do departamento de música da ECA-USP	104
Figura 47 – Diagrama do estúdio LAMI (ECA-USP) utilizado na gravação com as baias para isolamento acústico	105
Figura 48 – Detalhe da baia de isolamento acústico durante a gravação	106
Figura 49 – Padrão de captura e resposta em frequência dos microfones Sennheiser MD 421 (dinâmico) e AKG C414 (condensador)	107
Figura 50 – <i>Patch</i> para espacialização da peça <i>The Unanswered Question</i> com três fontes sonoras	109
Figura 51 – <i>Patch</i> para espacialização da peça <i>The Unanswered Question</i> com nove fontes sonoras	110
Figura 52 – Posições das fontes para o teste 1, sem acesso à interface gráfica pelo usuário.....	112

LISTA DE TABELAS

Tabela 1 – Possíveis aplicações de um sistema de espacialização sonora.....	16
Tabela 2 – Atributos espaciais	21
Tabela 3 – Algumas composições com espacialização anteriores ao séc. XX	27
Tabela 4 – Composições contemporâneas com espacialização (continua)	31
Tabela 5 – Tecnologias Dolby e suas aplicações.....	54
Tabela 6 – Equações de codificação Ambisonics até a 3ª ordem	64
Tabela 7 – Formatos de armazenamento e transmissão de canais Ambisonics.....	65
Tabela 8 – Manipulações do campo sonoro em Ambisonics de 1ª ordem	66
Tabela 9 – Matriz de ganhos de primeira ordem para uma configuração octogonal ..	68
Tabela 10 – Sugestão de ganhos para os filtros psicoacústicos do Ambisonics.....	69
Tabela 11 – Comparação entre as técnicas de auralização.....	71
Tabela 12 – Blocos desenvolvidos no PD, sobre a arquitetura AUDIENCE.....	82
Tabela 13 – Dados de entrada e saída dos blocos (continua)	83
Tabela 14 – Peças musicais espacializadas para o experimento	100
Tabela 15 – Testes de validação do sistema de espacialização	111
Tabela 16 – Resultados dos testes de validação do sistema de espacialização.....	114

LISTA DE ABREVIATURAS E SIGLAS

AUDIENCE	<i>Audio Immersion Experience by Computer Emulation</i>
CAVE	<i>Cave Automatic Virtual Environment</i>
CMU	Departamento de Música da ECA
DSP	<i>Digital Signal Processing</i>
DII	Diferença de Intensidade Interaural
DTI	Diferença de Tempo Interaural
DTS	<i>Digital Theater System</i>
DVD	<i>Digital Versatile Disc</i>
ECA	Escola de Comunicações e Artes da USP
GUI	<i>Graphical User Interface</i>
HOA	<i>High Order Ambisonics</i>
IR	<i>Impulse Response</i>
IRCAM	<i>Institut de Recherche et Coordination Acoustique/Musique</i>
ITU	<i>International Telecommunication Union</i>
MPEG	<i>Moving Picture Experts Group</i>
OCAM	Orquestra de Câmara da USP
PD	<i>Pure Data</i>
SACD	<i>Super Audio CD</i>
SVPA	Sistema Vetorial De Panorama Por Amplitude
THX	<i>Tomlinson Holman's eXperiment</i>
UML	<i>Unified Modeling Language</i>
WFS	<i>Wave Field Synthesis</i>
VBAP	<i>Vector Based Amplitude Panning</i>

LISTA DE SÍMBOLOS

θ	Ângulo de elevação com relação à origem
ϕ	Ângulo de rotação com relação à origem
S	Sinal de áudio monofônico
$\vec{r}$	Vetor da origem à posição do usuário
α_n	Ganho do canal W para o alto-falante n em Ambisonics 1ª ordem
β_n	Ganho do canal X para o alto-falante n em Ambisonics 1ª ordem
γ_n	Ganho do canal Y para o alto-falante n em Ambisonics 1ª ordem
δ_n	Ganho do canal Z para o alto-falante n em Ambisonics 1ª ordem
$G1$	Ganho do filtro psicoacústico (< 700Hz)
$G2$	Ganho do filtro psicoacústico (> 700Hz)
m	Ordem Ambisonics
N	Número de alto-falantes em uma configuração Ambisonics
$\vec{k}$	Vetor da origem à posição de uma fonte sonora
ψ	Ângulo de rotação de uma fonte sonora com relação à origem
J_m	Função de Bessel de primeira ordem
Y_{00}^1	Função de Bessel de segunda ordem
P_ψ	Pressão sonora de uma fonte real
ϕ_n	Ângulo de rotação de um alto-falante n com relação à origem

SUMÁRIO

1	INTRODUÇÃO	13
1.1	OBJETIVOS	15
1.2	PREMISSAS.....	17
1.3	SUMÁRIO ESTRUTURADO DA DISSERTAÇÃO	19
2	ESPACIALIZAÇÃO SONORA	20
2.1	PERCEPÇÃO DA ESPACIALIZAÇÃO	20
2.2	ESPACIALIZAÇÃO NA MÚSICA	23
2.3	SISTEMAS PARA ESPACIALIZAÇÃO SONORA.....	33
2.4	TENDÊNCIAS	43
3	TÉCNICAS DE AURALIZAÇÃO	48
3.1	BI-AURAL	48
3.2	DIPOLO ESTÉREO.....	51
3.3	DOLBY SURROUND E DTS	53
3.4	SISTEMA VETORIAL DE PANORAMA POR AMPLITUDE.....	55
3.5	AMBIOPHONICS.....	57
3.6	WAVE FIELD SYNTHESIS	58
3.7	AMBISONICS	61
3.8	CONCLUSÃO	70
4	IMPLEMENTAÇÃO DO SISTEMA DE ESPACIALIZAÇÃO.....	72
4.1	PLATAFORMA DE PROGRAMAÇÃO	72
4.1.1	Pure Data	72
4.1.2	AUDIENCE	76

4.2	MODELAGEM DO SISTEMA.....	78
4.3	IMPLEMENTAÇÃO DOS BLOCOS POR CAMADA AUDIENCE.....	86
4.3.1	Camada 1 – Composição de Cena Acústica	86
4.3.2	Camada 2 – Simulador Acústico.....	87
4.3.3	Camada 3 – Codificador de áudio espacial	87
4.3.4	Camada 4 – Decodificador e reprodutor de áudio espacial	88
4.4	BLOCOS DO SISTEMA EM CONJUNTO	90
4.4.1	Construção de um <i>patch</i>	90
4.4.2	Teste dos blocos em conjunto	94
4.5	SISTEMA DE HARDWARE PARA REPRODUÇÃO	96
5	EXPERIMENTO PARA VALIDAÇÃO DO SISTEMA.....	98
5.1	ESCOLHA DA PEÇA MUSICAL	99
5.2	PRODUÇÃO DA PEÇA ESCOLHIDA.....	102
5.3	MONTAGEM DO <i>PATCH</i>	108
5.4	TESTES DE VALIDAÇÃO.....	111
5.5	RESULTADOS	113
6	CONCLUSÕES	116
6.1	CONTRIBUIÇÕES.....	117
6.2	TRABALHOS FUTUROS	118
	REFERÊNCIAS	121
	APÊNDICE A – Fundamentos físicos do Ambisonics	128
	APÊNDICE B – Diagramas UML do sistema.....	134
	APÊNDICE C – Partitura da peça <i>The Unanswered Question</i>	136
	APÊNDICE D – Dados Técnicos da Gravação	145
	APÊNDICE E – Questionário de avaliação do sistema.....	147

1 INTRODUÇÃO

A localização espacial ou percepção espacial se refere à habilidade humana de julgar a direção e a distância de uma fonte sonora. Assim, podemos entender como espacialização sonora o ato de posicionar uma fonte sonora com o intuito de estimular essa capacidade em um ouvinte.

Essa habilidade sempre foi importante para o homem como uma forma de direcionar o nosso sentido mais utilizado e desenvolvido, a visão, para objetos a serem evitados ou buscados (MOORE, 2003). Por exemplo, nos primórdios da civilização, esse era um dos principais auxílios para a caça e, portanto, vital para a sobrevivência. Justamente por isso, a evolução nos favoreceu com um mecanismo eficiente para a localização auditiva. Com a passagem do tempo e o assentamento das comunidades, a caça perdeu a importância, mas a localização sonora tinha sua utilidade nas constantes guerras e lutas. Mesmo atualmente, esta habilidade é útil para escapar de perigos corriqueiros, como atravessar uma rua.

Tendo essa propriedade de chamar a atenção para certo evento, nada mais natural que a espacialização sonora fosse utilizada na música, mas esse caminho demoraria alguns séculos para que fosse seguido conscientemente pelos compositores. Com algumas exceções, somente a partir do século XX é que alguns compositores iriam considerar o espaço como um novo parâmetro a ser explorado na música e, mais de cinquenta anos depois, utilizá-lo sistematicamente na composição das chamadas peças espaciais ou espacializadas (TROCHIMCZYK, 2001).

O grande impulso dado à composição de música espacializada foi a popularização das fitas magnéticas e a possibilidade de reproduzir sons previamente gravados utilizando alto-falantes. Assim, a execução de uma obra não estava mais presa ao músico, permitindo ao compositor dispor de inúmeras fontes sonoras, posicionadas arbitrariamente e interagindo com os músicos.

Embora o interesse pela síntese de campos sonoros bi ou tridimensionais tenha nascido com a capacidade de reprodução sonora, as limitações tecnológicas (restrita a sistemas mono ou estereofônicos) retardaram por meio século os experimentos musicais com espacialização utilizando elementos eletroacústicos.

Com os avanços na capacidade de recriação de campos sonoros, sistemas de espacialização sonora começaram a ser desenvolvidos para serem utilizados por músicos. Atualmente, conjuntos de diversos *softwares* e *hardwares* estão disponíveis e sistemas podem ser montados com grande flexibilidade.

O termo auralização é utilizado neste contexto de forma análoga ao termo visualização, utilizado na computação gráfica, ou seja, é o processo de criar um campo sonoro em um espaço através da modelagem física ou matemática.

A motivação desta pesquisa é criar uma ferramenta computacional que auxilie compositores, regentes e mesmo produtores fonográficos na tarefa de espacializar fontes sonoras. Com um sistema auxiliado por ferramentas computacionais, o usuário do sistema pode fazer experimentos virtuais de posicionamento espacial tridimensional (ver Figura 1).

Figura 1 – Posicionamento de fontes sonoras virtuais envolvendo o usuário

No caso do compositor/regente, o sistema pode auxiliá-los a realizar tecnologicamente suas idéias musicais, mais especificamente com a espacialização dos instrumentos em uma peça musical, e avaliar os resultados instantaneamente, sem a necessidade de uma orquestra a sua disposição. Isso se aplica ao trabalho de modelamento em estúdio e também a apresentações ao vivo.

O sistema é capaz de reproduzir peças espacializadas da forma como o compositor sugeriu, permitindo ainda que o usuário possa manipular e criar novas formas de apresentação espacial da mesma peça.

O trabalho é resultado da primeira implementação da arquitetura de produção, transmissão e reprodução sonora espacial AUDIENCE (FARIA, 2005; FARIA; ZUFFO, 2006). Novos dispositivos de projeção sonora com alta densidade de canais estão previstos no AUDIENCE, sendo esse trabalho importante para investigar o comportamento do sistema como um todo quando se utiliza um número elevado de canais ou fontes sonoras simultaneamente.

Começa-se com uma introdução sobre a percepção da espacialização e seu uso na música. Em seguida, são apresentados sistemas existentes para espacialização sonora, suas capacidades e limitações, inclusive com exemplos comerciais. É feita uma revisão do estado da arte em técnicas para recriação do campo sonoro, ou auralização. A partir daí, descreve-se a proposta de desenvolvimento e a implementação de um sistema de espacialização sonora aplicado ao contexto da música espacializada. Enfim, experimentos de validação do sistema são realizados, chegando-se às conclusões sobre a aplicabilidade do sistema.

1.1 OBJETIVOS

Este trabalho de pesquisa tem os seguintes objetivos principais:

- **Pesquisar sistemas e técnicas para espacialização sonora**

O trabalho envolve o levantamento da bibliografia principal e o estado da arte na área de técnicas, algoritmos e sistemas para espacialização sonora.

Essa abordagem é importante na atual fase tecnológica da engenharia de áudio, na qual sistemas de som multicanal e de áudio imersivo encontram-se em intensa pesquisa e exploração comercial, havendo várias tecnologias para tratamento de som *surround* e mercados para absorver a tecnologia que vão desde o cinema digital até os jogos eletrônicos domésticos (FARIA, 2005).

- **Desenvolver um sistema para a espacialização sonora**

O trabalho tem o objetivo de desenvolver um sistema para a espacialização de fontes sonoras e construção de cenas acústicas virtuais, baseado na arquitetura de codificação, transmissão e decodificação de som multicanal AUDIENCE (FARIA, 2005). A Tabela 1 mostra possíveis aplicações deste sistema.

Tabela 1 – Possíveis aplicações de um sistema de espacialização sonora

<i>Aplicação</i>
Musical - Reprodução de peças musicais espacializadas
Musical - Experimentos com espacialização orquestral pelo compositor/regente
Musical - Produção musical
Musical - Ensaios com orquestra virtual
Som de cinema
Realidade Virtual
Jogos de Computador
Sistemas de simulação e controle em engenharia
Sistemas de teleconferência / comunicação imersiva

A aplicação escolhida para o sistema em construção tem como alvo o problema da espacialização instrumental em música, tanto de pequenas formações, quanto de

orquestras. No sistema em desenvolvimento, não só a localização da fonte sonora é alterável, mas também a movimentação e os efeitos de ambiência são considerados.

Tendo em vista as tendências na incorporação de multifuncionalidades nas plataformas de consumo e *workstations* de áudio, antes restritas a ambientes de editoração fonográfica, tal investigação também permitirá conhecer quais parâmetros de cena acústica são potenciais candidatos para serem editáveis em futuras plataformas de consumo.

Embora seja uma ferramenta voltada para a utilização na composição e execução de música espacial, o sistema também é extensível a outras aplicações como sistemas de teleconferência imersivos, jogos de computador e outros.

- **Realizar experimentos de espacialização musical**

Experimentos de espacialização de instrumentos de orquestra são realizados no sistema durante o desenvolvimento e, ao final, a reprodução de uma peça espacial é executada pelo sistema a fim de validá-lo e colher resultados.

1.2 PREMISSAS

Este trabalho de pesquisa tem as seguintes premissas:

- **Uso de sistemas livres e abertos**

Uma das premissas do trabalho é utilizar plataformas livres e abertas para a implementação do sistema de áudio. A escolha do Pure Data (PD) como plataforma de programação de áudio gráfica deu-se devido à facilidade de uso, código aberto, grande penetração no meio acadêmico, grande comunidade de usuários e a capacidade de ser estendido. Com versões para os sistemas operacionais Windows e MacOS, o PD além de ser mutiplataforma, com poucas alterações necessárias

para a mudança, também é suportado pelo Linux, um sistema operacional gratuito e de código aberto.

- **Uso da arquitetura AUDIENCE**

O sistema de referência no presente trabalho é o AUDIENCE, em desenvolvimento no projeto homônimo AUDIENCE¹ – *Audio Immersion Experience by Computer Emulation* – voltado para a auralização realista e imersão sonora. Um dos objetivos principais é o de investigar e prover soluções flexíveis e escaláveis para imersão sonora multicanal, integradas ou não a ambientes de realidade virtual. Este sistema é utilizado na CAVERNA Digital da Universidade de São Paulo para torná-la um ambiente de realidade virtual imersiva completa.

- **Possível integração com sistemas de realidade virtual imersivos**

O sistema deve ser construído com a possibilidade de ser expandido para se conectar a um sistema de realidade virtual imersiva, como as CAVEs (ZUFFO et al., 2001), de forma que a experiência possa ser aumentada incluindo a visualização dos instrumentos, por exemplo, em uma sala de concerto ou em uma cena complexa.

Poucas CAVEs já vêm incorporando sistemas sonoros mais complexos, como a descrita por Ogi et al. (2003), que utiliza uma configuração maior de alto-falantes e a técnica SVPA de auralização (item 3.4), obtendo uma maior imersão sonora. Mesmo assim, faltam tecnologias de auralização que possam ser acopladas a um sistema de realidade virtual imersivo.

Em muitos casos, o sistema de sonorização das CAVEs restringe-se a um par de alto-falantes e à utilização de alguma biblioteca para programação de áudio 3D com implementações de técnicas simples de auralização, como o estéreo por panorama. Assim, a imersão sonora obtida é muito baixa, sem a capacidade de posicionar fontes sonoras virtualmente com precisão, e não coerente a um sistema visual com alta imersão.

¹ Disponível em: <<http://www.lsi.usp.br/interativos/nem/audience/>>. Acesso em: 11 março 2006.

O sistema de espacialização sonora traz melhorias na integração de áudio e vídeo contribuindo com soluções para sonorizar com precisão vários sistemas de realidade virtual e pode ser extremamente útil para possibilitar uma imersão completa.

1.3 SUMÁRIO ESTRUTURADO DA DISSERTAÇÃO

Capítulo 1 – Introdução: introdução à espacialização na música, problemas e necessidades e definição dos objetivos e aplicações do trabalho desenvolvido;

Capítulo 2 – Espacialização sonora: revisão bibliográfica sobre a percepção da espacialização o sistema sensorial humano, a aplicação da espacialização na música, sistemas de espacialização sonora desenvolvidos, principalmente, para o uso na música espacializada e tendências da área de espacialização.

Capítulo 3 – Estado da arte em técnicas de auralização: revisão bibliográfica do estado da arte das principais técnicas de auralização; detalhamento da técnica de auralização Ambisonics, utilizada como base de codificação espacial no trabalho;

Capítulo 4 – Implementação do sistema de espacialização sonora: descrição da implementação do sistema baseado na arquitetura AUDIENCE, algoritmos utilizados, blocos desenvolvidos;

Capítulo 5 – Experimento para validação do sistema: descreve os experimentos realizados baseado no sistema desenvolvido e seus resultados;

Capítulo 6 – Conclusões: conclusões gerais do trabalho, contribuições e trabalhos futuros.

2 ESPACIALIZAÇÃO SONORA

O texto a seguir apresenta uma breve revisão sobre os mecanismos que o ser humano possui para perceber a localização de uma fonte sonora e alguns parâmetros envolvidos na percepção. Em seguida, é apresentado um resumo histórico de como a espacialização foi aplicada na música ocidental, desde a Idade Média até nossos dias, e uma breve revisão dos sistemas desenvolvidos para permitir essa espacialização. Um sistema de espacialização genérico é apresentado, que será utilizado como base para a avaliação das técnicas de auralização e implementação do sistema. Por fim, algumas tendências na área são identificadas.

2.1 PERCEPÇÃO DA ESPACIALIZAÇÃO

O mecanismo de localização da direção funciona por meio de indicações auditivas (*auditory cues*) que são percebidas pelo nosso sistema auditivo. As principais indicações vêm do fato de termos dois ouvidos posicionados separadamente, fazendo com que o som chegue de forma distinta em cada um.

De acordo com Lord Rayleigh (MOORE, 2003), a diferença de tempo interaural (DTI) e a diferença de intensidade interaural (DII) são as indicações mais utilizadas. A DTI indica a diferença de tempo que a onda sonora leva para chegar a cada ouvido e a DII a diferença da intensidade.

Presume-se que o nosso sistema auditivo, a partir dessas indicações e de outras mais específicas (como a ordem de chegada das ondas e o espectro), considerando um som complexo, determine a posição (direção e distância) da fonte mediante um sistema de votação (GERZON, 1974b). Uma análise detalhada dessas indicações, bem como de outros fatores e avanços na área sobre os mecanismos auditivos de percepção espacial foram descritos por Blauert (1996).

Com o crescente avanço dos sistemas de som envolvente, ou *surround*, a avaliação dos atributos espaciais torna-se importante como uma forma de comparar estes sistemas. Estes atributos tanto podem considerar aspectos simples e objetivos, como a qualidade na determinação da direção e distância da fonte, como podem se aprofundar em aspectos muito mais subjetivos, como a presença e a naturalidade de uma fonte sonora.

Tabela 2 – Atributos espaciais

<i>Atributo</i>	<i>Descrição</i>
Naturalidade	A similaridade do som produzido pelos alto-falantes com a produzida naturalmente.
Presença	A experiência de se encontrar na mesma sala que a fonte sonora.
Preferência	Se o som como um todo agrada.
Conteúdo de Baixa Freqüência	O nível sonoro de baixas freqüências.
Largura do Conjunto	A largura percebida do conjunto de todas as fontes sonoras, o ângulo ocupado pelo conjunto. Não considerar a reverberação.
Largura da Fonte Individual	A largura percebida de uma fonte sonora, o ângulo ocupado pela fonte. Não considerar a reverberação.
Localização	A facilidade de se perceber a posição de uma fonte sonora, ou seja, quão fácil é apontar a direção de uma fonte sonora.
Distância da Fonte	A distância percebida da fonte.
Envoltório da Fonte	A extensão com que uma fonte sonora envolve o ouvinte. Não considera a reverberação.
Largura da Sala	A largura/ângulo ocupado pelos sons vindos das reflexões das fontes sonoras. Não considera o som direto das fontes.
Tamanho da Sala	O tamanho percebido da sala, se é grande ou pequena.
Nível Sonoro da Sala	O nível sonoro gerado na sala como resultado da ação das fontes sonoras. Não considera o som direto.
Envoltório da Sala	A extensão com que o som vindo das reflexões na sala envolve o ouvinte.

Extraído de Berg e Rumsey (2003).

Berg e Rumsey (2003) descrevem uma metodologia para avaliar a qualidade na espacialidade através da definição de atributos e escalas para serem utilizados em testes. Baseado em trabalhos anteriores (citados no artigo) e contribuições próprias,

chegam a um conjunto de atributos capaz de definir a qualidade espacial (descritos na Tabela 2). A Figura 2 ilustra alguns destes atributos em uma cena acústica de um grupo de câmara.

Figura 2 – Alguns atributos para avaliar a percepção espacial em uma sala de concerto

A avaliação desses atributos pode ser utilizada para determinar o nível de imersão sonora proporcionado pelo sistema de reprodução, conforme propõem Faria, M. Zuffo e J. Zuffo (2005). A maior contribuição deste artigo é propor uma escala de percepção da intensidade da imersão sonora medida em sistemas de auralização considerando-se os atributos acima como variáveis de ponderação. Nesta escala, o

nível cinco, mais alto, representa a capacidade de se criar campos sonoros tridimensionais estáveis e com a localização e distância da fonte mais apuradas, enquanto que o nível quatro é similar para campos bidimensionais.

2.2 ESPACIALIZAÇÃO NA MÚSICA

Toda execução de uma peça musical pode ser considerada espacializada, pois o ouvinte utiliza-se de sua percepção espacial para avaliar o posicionamento da fonte. Assim, desde um instrumento solista, um conjunto de alto-falantes até uma orquestra, o ouvinte é capaz de apontar a origem dos sons. Geralmente, esse posicionamento ou ambientação é intrínseco a forma e local da propagação sonora, dependendo mais de questões físicas do que artísticas.

Por exemplo, em uma orquestra, a escolha da posição dos instrumentos deve-se principalmente ao volume sonoro que cada instrumento produz, de forma que todos possam ser ouvidos claramente. Já em formações menores, como quartetos de cordas, o posicionamento das fontes tende a se tornar irrelevante para o resultado sonoro. Nos dois casos, o posicionamento percebido é de uma fonte sonora única, vinda da frente dos ouvintes, como ilustrado na Figura 3.

Contudo, a utilização da espacialização na música vai além dessa simples percepção quando o compositor foge do senso comum e agrega à sua peça a dimensão musical do espaço. Por exemplo, uma orquestra posicionada no palco e um trompete soando atrás da plateia.

Então, aplica-se o termo de música espacializada para toda música na qual, por meio do posicionamento incomum das fontes, da movimentação das fontes e do uso da ambientação, o compositor eleva a espacialização das fontes sonoras a um mesmo patamar de importância que os outros elementos da música como a harmonia, o ritmo e o timbre, entre outros. A seguir, apresenta-se um breve histórico da música espacializada.

Figura 3 – Posicionamento não espacializado, com uma frente sonora única sendo percebida (quarteto de cordas e orquestra)

Na Idade Média, existia um tipo de execução do canto-chão realizada pelos monges em seus conventos e em algumas igrejas chamado Responsório. A formação podia ser de dois coros voltados um para o outro, ou um solista no lugar de um destes coros. Assim, uma frase cantada por um coro ou solista era respondida pelo outro coro (APEL, 1990).

Durante a Renascença, o uso intenso da polifonia permitia ao compositor fazer alguns experimentos com posicionamento de coros. No século XVI, o compositor flamengo Adrian Willaert, criou uma das primeiras peças utilizando o espaço como elemento composicional, utilizando dois coros posicionados em lados opostos da Basílica de São Marcos, em Veneza (ver Figura 4).

Figura 4 – Basílica de São Marcos, em Veneza, onde os dois coros se posicionam nos balcões laterais²

Essa técnica foi chamada de *cori spezzati*, coros espaçados, e foi muito utilizada pelos compositores venezianos, principalmente por Giovanni Gabrieli (a Figura 5 apresenta um trecho de uma partitura para dois coros antifônicos), e posteriormente também na Inglaterra, Alemanha e outros países da Europa (ZVONAR, 1999).

O moteto *Spem in alium nunquam habui*, do compositor inglês Thomas Tallis, é um exemplo grandioso desse tipo de música espacial. Composto para oito corais com

² Figura extraída de: Romanelli, G. **Venice - Art & Architecture**. Konemann, 1999.

cinco vozes cada, totalizando quarenta vozes, esses podem ser dispostos em um formato circular, onde a música circula pelos corais, ou então em um arranjo baseado nos pontos cardeais, criando um efeito antifônico entre os coros (SCHWARTZ, 1998).

The image displays a musical score for two choirs, arranged in two systems. Each system contains five staves, representing the voices of two different choirs. The lyrics are in Portuguese and include the phrase 'al-le-lu-ia' repeated several times, followed by 'Ho-di-e sal-va-tor ap-pa-'. The score is written in a standard musical notation with treble and bass clefs, and the lyrics are written below the staves.

10

ia, al - le-lu - ia, al - le-lu - ia, al - le-lu - ia.

ia, al - le-lu - ia, al - le-lu - ia, al - le-lu - ia.

ia, al - le-lu - ia, al - le-lu - ia, al - le-lu - ia.

8 ia, al - le-lu - ia, al - le-lu - ia, al - le-lu - ia.

ia, al - le-lu - ia, al - le-lu - ia, al - le-lu - ia.

al - le-lu - ia, al - le-lu - ia. Ho - di - e sal - va - tor ap -

8 al - le-lu - ia, al - le-lu - ia. Ho - di - e sal - va - tor ap - pa -

al - le-lu - ia, al - le-lu - ia. Ho - di - e sal - va - tor ap -

al - le-lu - ia, al - le-lu - ia. Ho - di - e sal - va - tor ap -

al - le-lu - ia, al - le-lu - ia. Ho - di - e sal - va - tor ap -

al - le-lu - ia, al - le-lu - ia. Ho - di - e sal - va - tor ap -

Figura 5 – Trecho de um *cori spezzati* de G. Gabrieli, *Hodie Christus natus est*, para 2 coros

Com a criação dos corpos orquestrais e a tendência da homofonia na música, a espacialização esteve praticamente adormecida durante os períodos Barroco, Clássico e Romântico. Ainda que as formações orquestrais fossem mudando ao longo dos anos, como se pode ver ao comparar as posições dos instrumentos em uma orquestra Barroca e em uma Romântica, experimentações e novas sonoridades advindas da espacialização antifonal renascentista foram deixadas de lado.

Algumas exceções são espacializações de instrumentos na qual o compositor visava à criação de algum efeito especial na peça, como se pode observar em peças de George Frideric Handel, Wolfgang Amadeus Mozart e Ludwig Van Beethoven.

Os efeitos utilizados são ambiências diferentes (e.g. concertos ao ar-livre, sobre barcos), ecos, com instrumentos repetindo os outros, e instrumentos posicionados fora do palco. Um efeito muito usado era o posicionamento de metais fora do palco, geralmente dando a impressão de que alguma autoridade ou exército estava se aproximando. A Tabela 3 apresenta alguns exemplos do uso da espacialização nesta época.

Tabela 3 – Algumas composições com espacialização anteriores ao séc. XX

<i>Compositor</i>	<i>Peça</i>	<i>Ano</i>	<i>Detalhes</i>
W. A. Mozart	<i>Noturno para 4 orquestras, K. 286</i>	1777	Efeitos de eco entre as orquestras
W. A. Mozart	<i>Don Giovanni, K. 527</i>	1787	Três grupos tocam separadamente: no buraco da orquestra, no palco e atrás do palco
L. van Beethoven	<i>Fidelio: Abertura Leonora, Op. 72</i>	1805	Trompete tocado fora do palco
H. Berlioz	<i>Requiem: Tuba Mirum, Op. 5</i>	1837	Quatro conjuntos de metais separados, localizados na disposição dos pontos cardeais, anunciam a entrada do coro
G. Verdi	<i>Requiem</i>	1874	Metais fora do palco
G. Mahler	<i>Sinfonia no. 2</i>	1895	Conjunto de metais e percussão fora do palco

A grande revolução musical do fim do século XIX, na qual o sistema tonal se expandiu, abria espaço para novas sonoridades, vindas de novas escalas, novos timbres, novas combinações instrumentais, novas harmonias e, também, nova espacialidade do som (GRIFFITHS, 1994). A tendência das décadas seguintes foi trazer a espacialização para um patamar tão importante quanto de outros parâmetros ou dimensões de uma peça (e.g. harmonia, instrumentação), como pode ser verificado no trabalho de compositores como Charles Ives, Erik Satie e Edgard Varèse (TROCHIMCZYK, 2001).

Essa utilização da espacialidade pode ser observada nas peças de Charles Ives, como *The Unanswered Question*, com grupos de instrumentos distribuídos pela sala de concerto. Edgard Varèse também foi um dos pioneiros em utilizar e dar forma a esse novo conceito de música espacial em peças como *Intégrales* (1925), que utiliza a ambiência do local de execução da peça (MALHAM, 2001). Varèse foi um dos primeiros a utilizar o termo música espacial, ao dar um rótulo para essa peça.

A partir da metade do século XX, a configuração espacial do corpo orquestral começou a ser explorada sistematicamente por compositores e regentes. Nomes proeminentes da música de vanguarda da época, como John Cage, Karlheinz Stockhausen, Iannis Xenakis e Pierre Boulez, entre outros, iniciaram suas experimentações com a espacialização, seja utilizando os instrumentos acústicos ou auxiliados pela revolução tecnológica da música eletroacústica (TROCHIMCZYK, 2001).

Figura 6 – Ensaio da peça *Gruppen*, de K. Stockhausen (1957), para três orquestras em torno da audiência³

³ Figura extraída de (GRIFFITHS, 1994)

A composição espacial eletroacústica aparece com a peça *Pfingstatorium*, de Ernst Krenek (1955) e, a partir daí, Karlheinz Stockhausen viria a compor as peças quadrafônicas, *Gesang der Jünglinge* (1956) e *Kontakte* (1960). A utilização da espacialização do som tornou-se um item importante nas composições de Stockhausen, seja em peças eletroacústicas, como *Spiral* (1970), ou em peças acústicas, como *Gruppen* (1957) e *Carré* (1960). Em *Gruppen*, três orquestras tocam em torno da audiência, criando efeitos como a passagem de acordes entre orquestras (na Figura 6 pode-se ver um exemplo desta espacialização).

Esses experimentos com espacialização acústica foram absorvidos e expandidos por Henry Brant, que conduziu diversos experimentos com posicionamento dos músicos em salas de concerto. Esses experimentos deram origem a um artigo, no qual o autor discorre sobre as possibilidades e efeitos advindos da espacialização (SCHWARTZ, 1998).

Exemplos da espacialização utilizada por Brant podem ser vistos em peças como *Antiphony* (1955), na qual cinco orquestras tocam simultaneamente de locais diferentes, e *Voyage Four* (1964), para orquestra com instrumentos posicionados em diversos locais da sala de concerto, inclusive com sons vindos de dutos de ar do piso. A Figura 7 mostra um diagrama, feito por Brant, da disposição dos instrumentos da peça.

Um outro exemplo do uso da música espacializada acústica pode ser visto na peça *Últimos Ritos*, de John Tavener (1972). A peça, que deve ser executada preferencialmente em uma catedral, tem um significado religioso que é explicitado pelo uso da dimensão espacial. Utilizando as galerias superiores das catedrais, abaixo da cúpula, Taverner representa o céu através de flautas e trompetes tocando nestas posições, enquanto que o resto da orquestra situa-se no plano do chão. Assim, por meio da movimentação vertical da música, Taverner representa a Eucaristia onde Cristo desce para Terra.

No Brasil, diversos compositores incluíram a dimensão espacial em suas obras, como Gilberto Mendes, Mario Ficarelli, Willy Corrêa, Aylton Escobar, Almeida Prado e, mais recentemente, Silvio Ferraz, Flô Menezes e Mauricio de Bonis. Gilberto Mendes, em sua peça *Santos Football Music*, de 1968, utiliza a platéia como fonte sonora, que deve emitir gritos ao longo da execução da música, e alto-falantes posicionados em torno dessa. Outros experimentos com a espacialização foram

feitos por Flô Menezes, em peças como *Parcours de l'Entité* de 1993, para duas flautas, percussão e sons eletroacústicos, na qual flautistas se deslocam pelo espaço cênico durante toda a apresentação (MENEZES, 1999).

Figura 7 – Diagrama com a disposição dos instrumentos na peça *Voyage Four*, de H. Brant (1964)⁴

⁴ Figura extraída de (SCHWARTZ, 1998)

A Tabela 4 contém uma compilação de peças do século XX nas quais se utiliza a espacialização acústica e/ou eletroacústica. Não é o intuito desta tabela ser um catálogo completo de peças espacializadas, mas apenas fornecer exemplos de como a espacialização foi utilizada por compositores contemporâneos.

Tabela 4 – Composições contemporâneas com espacialização (continua)

<i>Compositor</i>	<i>Peça</i>	<i>Ano</i>	<i>Detalhes</i>
Charles Ives	<i>The Unanswered Question</i>	1906	Grupos instrumentais espalhados, com as cordas posicionadas fora do palco de frente para madeiras e metais
Edgard Varèse	<i>Intégrales</i>	1925	Utilização da ambiência proveniente do local de execução
Béla Bartók	<i>Music for Strings, Percussion and Celesta</i>	1936	Plano detalhado para a disposição das cordas aos flancos da percussão e do teclado
Henry Brant	<i>Antiphony</i>	1953	5 orquestras separadas no espaço
Ernst Krenek	<i>Pfingstorianum</i>	1955	Primeira peça eletroacústica estereofônica
Karlheinz Stockhausen	<i>Gesang der Jünglinge</i>	1956	Para 4 alto-falantes dispostos em panorama no palco e 1 alto-falante suspenso em cima da audiência
Karlheinz Stockhausen	<i>Gruppen</i>	1955-57	Para 3 orquestras em torno da audiência
Edgard Varèse	<i>Poème Électronique</i>	1958	Peça eletroacústica para fita utilizada no pavilhão Philips, na qual o som se movimenta através de um arranjo de 425 alto-falantes
Iannis Xenakis	<i>Duel</i>	1959	Para 2 orquestras, que respondem entre si executando partes escolhidas pelos regentes
Karlheinz Stockhausen	<i>Carré</i>	1959-60	Para 4 orquestras e 4 coros dispostos em forma de um quadrado
Karlheinz Stockhausen	<i>Kontakte</i>	1960	Primeira composição quadrafônica, criando o efeito de órbita
Karlheinz Stockhausen	<i>Originale</i>	1961	Peça de teatro baseada em <i>Kontakte</i>
Iannis Xenakis	<i>Stratégie</i>	1962	Similar à peça <i>Duel</i>
Henry Brant	<i>Voyage Four</i>	1963	Para percussão e metais no palco, violinos em um balcão lateral, violas e violoncelos em outro, baixos atrás, madeiras e algumas cordas no balcão traseiro e instrumentistas espalhados pela audiência
Iannis Xenakis	<i>Terretektorth</i>	1965-66	Peça para orquestra, com os músicos espalhados pela audiência

Tabela 4 – Composições contemporâneas com espacialização (conclusão)

<i>Compositor</i>	<i>Peça</i>	<i>Ano</i>	<i>Detalhes</i>
Karlheinz Stockhausen	<i>Stimmung</i>	1968	Para 6 vocalistas amplificados por 6 alto-falantes dispostos em torno da audiência
Luciano Berio	<i>Sinfonia</i>	1968-69	Grupo de violinos localizados atrás da orquestra
Pierre Boulez	<i>Domaines</i>	1968-69	Para clarinete solo e 6 grupos instrumentais, em que o clarinetista se desloca pela cena
Henry Brant	<i>Windjammer</i>	1969	Trompa solo parada e várias madeiras movendo-se em rotas predeterminadas enquanto tocam
Gilberto Mendes	<i>Santos Football Music</i>	1969	Para orquestra, alto-falantes em torno da audiência e participação de sons vindos da platéia
Karlheinz Stockhausen	<i>Spiral</i>	1970	Para solistas, com o som amplificado e manipulado espacialmente em uma projeção esférica
Karlheinz Stockhausen	<i>Sternklang</i>	1971	Para 5 grupos em locais distantes em um parque ao ar livre
John Tavener	<i>Últimos Ritos</i>	1972	Peça para orquestra a ser executada em catedrais, na qual as flautas e trompetes posicionam-se em galerias elevadas
Pierre Boulez	<i>Ritual</i>	1974-75	Para orquestra dividida em 8 grupos
Julio Estrada	<i>Canto Naciente</i>	1975-78	Oito metais em forma de cubo ao redor da audiência
Aylton Escobar	<i>Poética II</i>	1978	Para 4 flautas dispostas em losango em torno da audiência
Pierre Boulez	<i>Répons</i>	1981-88	Seis instrumentos solo, orquestra e equipamento eletroacústico
Luciano Berio	<i>Formazioni</i>	1985-87	Duas harpas posicionadas nos cantos do palco, próximas à audiência
Flô Menezes	<i>Parcours de l'Entité</i>	1993	Para flauta, percussão e sons eletroacústicos, com movimentos do flautista pelo espaço cênico
Mauricio de Bonis	<i>Sogno</i>	2001	Para piano, oboé e trompa, com movimentos dos dois últimos pelo palco

A música popular sempre absorve de alguma forma os aspectos de vanguarda da música erudita e com a espacialização não poderia ser diferente. Assim, experimentos com espacialização quadrafônica em shows começaram a ser feitos pela banda Pink Floyd (PARSONS, 1975), exemplo logo seguido por outras bandas.

2.3 SISTEMAS PARA ESPACIALIZAÇÃO SONORA

Por sistema para espacialização sonora entende-se a cadeia completa de processos para espacializar fontes sonoras, seja em uma peça musical ou em um filme com som *surround*. A reprodução espacial, através de uma configuração de vários alto-falantes ou de fones de ouvido, é apenas uma das tarefas de um sistema de espacialização sonora.

Figura 8 – Indicação na partitura para distribuição dos instrumentos entre a audiência na peça *Terretektorth*, de I. Xenakis (1965-66)

A seguir, através da apresentação de exemplos de aplicações de sistemas de espacialização sonora, do mais simples ao mais complexo, chega-se a um sistema de espacialização genérico.

O uso dos sistemas de espacialização na música acústica é simples, uma vez que só são necessárias as entradas, como a partitura com indicações do posicionamento, e a reprodução é feita diretamente pelo canto ou pelo tocar de um instrumento. Entretanto, eles podem ser extremamente complexos na sua execução, utilizando diversas fontes sonoras, conjuntos separados e exigindo a sincronização entre eles, movimentos e requisitos para o ambiente de apresentação.

Na Figura 8, pode-se ver o sistema utilizado por Iannis Xenakis na peça *Terretektorth* (1966), na qual a orquestra é homogeneamente distribuída entre a platéia, eliminando o conceito de palco, e cada ouvinte experimenta a peça de uma maneira, dependendo de sua posição.

Assim, o sistema necessário para a espacialização de uma peça acústica é composto pela partitura, na qual estão as indicações da execução musical (e.g. notas, instrumentação, posicionamento dos instrumentos), a disposição física dos músicos (de acordo com a partitura) e a execução da música por estes, gerando o som espacializado. Esse tipo de sistema pode ser visto no diagrama da Figura 13, item a.

Em músicas nas quais aparecem fontes eletroacústicas, a complexidade do sistema aumenta, pois se torna necessária a utilização de meios elétricos para a gravação, transmissão e reprodução do som.

Para a reprodução podem ser utilizados alto-falantes que funcionam como fontes primárias, ou seja, que representam uma fonte sonora vinda exatamente da posição onde eles foram colocados, em um mapeamento direto, ou fontes secundárias, que por meio de técnicas avançadas de codificação espacial e conjuntos de alto-falantes, projeta-se uma fonte virtual sem a coincidência de posicionamento com os alto-falantes.

Nesses tipos de sistemas, além da partitura também existem como entrada as gravações. A partir do posicionamento descrito na partitura, a separação dos canais é realizada, bem como o posicionamento dos músicos no caso de peças mistas. A gravação é transmitida para ser reproduzida em uma configuração de alto-falantes,

juntamente com a execução dos músicos. Esse tipo de sistema pode ser visto no diagrama da Figura 13, itens b e c.

Os primeiros experimentos utilizando a espacialização na música eletroacústica foram feitos por Pierre Schaeffer e Pierre Henry na França, no começo da década de 50. O sistema criado por Schaeffer consistia em até cinco reprodutores de fita magnética monofônicos, em que os sinais eram redirecionados para um sistema de quatro alto-falantes, em um arranjo tetraédrico (dianteiro-esquerda, dianteiro-direita, traseiro e acima da cabeça). A distribuição dos sons no espaço era feita com um aparato magnético que captava os movimentos da mão do usuário. John Cage, em sua peça *William Mix* (1952), utilizou oito reprodutores de fita magnética monofônicos para alimentar oito alto-falantes posicionados em torno da audiência.

A EXPO 70, realizada em Osaka, Japão, abrigou grandes experimentos eletroacústicos utilizando a música espacial. O pavilhão Phillips, projetado por Iannis Xenakis, abrigava um complexo agrupamento de 425 alto-falantes, alimentado por 11 canais. Neste local, Edgard Varèse apresentou *Poème Electronique* (1958), uma peça eletroacústica para fita magnética. A fita continha o áudio e também, mediante instruções codificadas nela, controlava o sistema de reprodução, movimentando o som e sincronizando o áudio com efeitos visuais (ZVONAR, 1999).

Um sistema mais avançado foi criado por Karlheinz Stockhausen para a sua peça *Spiral* (1970), executada no pavilhão alemão da EXPO 70, em Osaka. Em uma construção esférica, com capacidade para 600 ouvintes, 55 alto-falantes foram montados e a execução de 20 solistas era controlada pelo próprio compositor, distribuindo o som em caminhos circulares e espiralados através dos alto-falantes (mostrado na Figura 9). Estando totalmente cercada pelos alto-falantes, inclusive por baixo, a platéia pôde experimentar um dos primeiros sistemas de projeção sonora tridimensional (TROCHIMCZYK, 2001).

Já no fim da década de 60, a banda inglesa de rock progressivo Pink Floyd, em busca de novas sonoridades, fez experimentos com som envolvente em seus concertos. Um rudimentar sistema de mixagem foi montado de forma a cercar a audiência com caixas acústicas e, a partir disso, efeitos sonoros controlados por um *joystick* podiam ser jogados de um lado a outro ou descrever movimentos circulares (CUNNINGHAM, 1997).

Figura 9 – Pavilhão alemão esférico com orquestra de alto-falantes construído para a apresentação de *Spiral*, de Stockhausen, na EXPO 70 em Osaka⁵

Posteriormente, o sistema evoluiu para uma montagem de alto-falantes no formato quadrafônico, como mostra a Figura 10. Nesta montagem, o sistema quadrafônico tradicional foi rotacionado em 45 graus, posicionando caixas atrás da banda e atrás e dos lados da audiência. A partir da mesa de mixagem, o engenheiro de som era

⁵ Disponível em: <<http://www.medienkunstnetz.de/works/stockhausen-im-kugelauditorium/>>. Acesso em: 29 julho 2006

capaz de colocar qualquer som microfonado do palco, como o da guitarra, e controlar a sua posição no sistema envolvente (PARSONS, 1975).

Figura 10 – Esquema de montagem quadrafônica utilizada pelo Pink Floyd no início da década de 70⁶

Estes experimentos impulsionaram a gravação de discos no formato quadrafônico que, embora tenha sido um fracasso comercial devido aos problemas na recriação do campo sonoro (GERZON, 1974a), foi uma primeira tentativa de inserir o som envolvente no ambiente doméstico.

Com a chegada da era da computação e do processamento digital de sinais, as possibilidades para manipulação do som se expandiram. O *Institut de Recherche et Coordination Acoustique/Musique* (IRCAM), instituto francês de pesquisa musical e

⁶ Figura extraída de (PARSONS, 1975).

acústica fundado pelo compositor Pierre Boulez, criou em 1980 um sistema para o processamento, geração de ondas e espacialização de fontes sonoras (BOULEZ; GERZSO, 1988).

Na Figura 11, o equipamento da esquerda, chamado de Sogitec 4X e desenvolvido por Giuseppe Di Giugno no IRCAM, é composto por oito processadores digitais de áudio capazes de armazenar, transformar, analisar e sintetizar sinais áudio. O Matrix 32, à direita, controla o tráfego de áudio entre os microfones, o 4X e os alto-falantes. Para controlar o sistema de processamento de sinais, Miller Puckette, também do IRCAM, desenvolveu um *software* chamado Patcher⁷, que mais tarde viria a evoluir nos *softwares* MAX/MSP⁸ e no Pure Data⁹.

Figura 11 – Sistema de espacialização computacional utilizado para a reprodução de *Répons*, de Boulez (1985)¹⁰

⁷ Disponível em: <http://freesoftware.ircam.fr/article.php3?id_article=5>. Acesso em: 9 dez. 2006.

⁸ Disponível em: <<http://www.cycling74.com/products/maxmsp>>. Acesso em: 27 jul. 2006.

⁹ Disponível em: <<http://crca.ucsd.edu/~msp/software.html>>. Acesso em: 9 dez. 2006.

¹⁰ Figura extraída de (BOULEZ, 1985).

Boulez compôs, para a utilização do sistema, a peça *Répons* (1985), na qual o som proveniente dos instrumentos solistas é capturado por microfones, modificado pelo computador e movimentado pela sala de concerto através de alto-falantes posicionados em torno da audiência. A Figura 12 ilustra a disposição dos instrumentos e alto-falantes utilizados nessa peça.

Figura 12 – Posicionamento das fontes sonoras e movimentação, na peça *Répons*, de Boulez (1985)¹¹

Na Itália, o *Centro Tempo Reale*, fundado pelo compositor Luciano Berio com o objetivo de investigar as possibilidades de interação entre sistemas digitais e execuções musicais ao vivo, desenvolveu um sistema chamado SMOV (*Sound Movement*) que possibilitava o movimento de fontes sonoras por meio de um conjunto variável em posição e quantidade de alto-falantes (GIOMI; MEACCI; SCHWOON, 2003).

¹¹ Figura extraída de (BOULEZ, 1985).

Devido a popularização de computadores com poder para processar sinais digitais, a construção destes sistemas deixou de necessitar de equipamentos dedicados e de serem exclusividade de grandes centros de pesquisa. O desenvolvimento de *softwares* capazes de processar sinais de áudio das mais diversas maneiras, incluindo a espacialização, fez com que o desafio deixasse de ser construir um sistema de espacialização, mas sim usar todas as possibilidades disponíveis.

Um exemplo desses *softwares* é o Spat¹² do IRCAM, que é uma aplicação para espacialização sonora com baixa latência, permitindo aos compositores trabalhar com a dimensão espacial na música sem se preocupar com os parâmetros acústicos envolvidos.

O sistema AUDIENCE, descrito anteriormente e objeto deste trabalho, também se presta a implementar a parte de *software* de espacialização como o Spat, mas vai mais além, incluindo a cadeia de espacialização desde a descrição da cena até a reprodução por um conjunto específico de alto-falantes.

Além da aplicação em concertos de música erudita e popular, a espacialização também chegou ao ambiente doméstico devido ao crescente aumento dos sistemas *home theater* e a conseqüente produção de material espacial para estes. Atualmente, um grande número de concertos gravados no padrão DVD apresenta a trilha sonora codificada para a configuração 5.1 de alto-falantes, geralmente utilizando as técnicas Dolby Surround ou DTS (ver item 3.3). Esse tipo sistema pode ser visto no diagrama da Figura 13, item e.

Os jogos eletrônicos também vêm assimilando, como forma de aumentar a sensação de realidade, não somente técnicas de projeção visual tridimensional, mas também técnicas de projeção sonora envolvente. Este tipo de aplicação tem a característica de não ser estática, ou seja, o jogo deve responder aos comandos do usuário e as novas posições assumidas pela personagem no espaço do jogo, portanto a projeção do áudio deve mudar a cada instante. A técnica bi-aural, com o uso de HRTFs (ver item 3.1), apresenta bons resultados para este tipo de aplicação, seja o som projetado via alto-falantes ou fones de ouvido. Esse tipo sistema pode ser visto no diagrama da Figura 13, item d.

¹² Disponível em: <<http://forumnet.ircam.fr/356.html?&L=1>>. Acesso em: 9 dezembro 2006.

Os exemplos descritos acima mostram que os sistemas de espacialização sonora tanto podem ser simples, com músicos executando uma partitura, como podem ser extremamente complexos, como a cadeia completa da produção de uma trilha sonora para filme em formato *surround*. Nesse tipo de aplicação, a partir de gravações de música, diálogo e efeitos especiais, o som é codificado, como no padrão Dolby Surround, armazenado em uma mídia como o DVD e reproduzido em um ambiente doméstico.

Baseado nos exemplos, chega-se a um diagrama de blocos genérico de um sistema de espacialização sonora, ilustrado na Figura 13, item f.

A entrada de um sistema de espacialização sonora é compreendida pela descrição da cena acústica, ou seja, de como os sons serão espacializados e como irão se comportar ao longo da execução, e da descrição dos sons envolvidos na cena, que podem ser partituras, sons gravados ou alguma outra instrução que produza som.

A partir da descrição da cena, o próximo bloco do sistema é responsável por posicionar as fontes sonoras, seja através do posicionamento dos próprios músicos, pela utilização de um *mixer* para colocar um som em um canal de gravação ou um *software* de espacialização.

Em seguida, o som pode ser codificado por meio de alguma técnica de auralização, com o intuito de organizar a informação sonora temporal e espacial. O bloco de decodificação faz o trabalho inverso e reconstrói a informação original. Nestes blocos podem ser empregados algoritmos de compressão de sinais sonoros.

O bloco de transmissão é responsável por levar a informação sonora do local onde ela foi produzida até o local onde ela será reproduzida. Este trabalho pode ser feito por cabos elétricos, mídias (e.g. fitas magnéticas, DVDs, CDs) ou transmissão digital através de redes e protocolos de comunicação (e.g. redes de computadores).

Por fim, a reprodução materializa em ondas sonoras a informação sonora espacializada, da forma como foi descrita na entrada do sistema. Ela pode ser feita, por exemplo, mediante a atuação de músicos ou por alto-falantes.

Como foi visto através dos exemplos, não existe a necessidade de que um sistema apresente todos os blocos. Geralmente, o sistema de espacialização sonora é implementado por um subconjunto destes blocos, e a utilização depende das

necessidades da cena sonora espacializada que foi descrita pelo compositor/regente/produtor.

(a) Peça acústica espacializada

(b) Peça Eletroacústica espacializada

(c) Peça acústica + eletroacústica espacializada

(d) Jogos eletrônicos

(e) Cinema / Home Theater

(f) Sistema de Espacialização Genérico

Figura 13 – Diversos tipos de sistemas de espacialização sonora e sua generalização

2.4 TENDÊNCIAS

A tendência atual dos sistemas *surround* em produções para cinema, sistemas profissionais e audiófilos é de explorar a verticalidade do espaço sonoro. Essa característica é suportada pela configuração de alto-falantes 10.2 (RUMSEY, 2001), com a adição de duas caixas para o efeito de altura, e também na revitalização de formatos como o Ambisonics (ver item 3.7) e na criação de novos, como o Ambiophonics (ver item 3.5).

No caso do setor de *consumer electronics*, a tendência é a popularização da tecnologia *surround* com a criação de sistemas de baixo custo, obtendo o melhor efeito com o mínimo de caixas instaladas e decodificadores baseados em *hardwares* simples ou computadores pessoais. Exemplos de decodificadores de baixo custo, que rodam sobre a plataforma Linux e disponibilizam várias técnicas de auralização, são o AlmusVCU¹³ e o VLC Media¹⁴.

Com a crescente demanda na área da espacialização, nada mais natural que existam cada vez mais ferramentas computacionais que disponibilizem alguma forma de auxiliar o processo de espacialização. Algumas destas ferramentas podem ser vistas na Figura 14.

Atualmente, os estúdios de gravação já trabalham intensamente com o processamento digital e a maioria é capaz de fazer uma produção *surround* 5.1, 6.1, 7.1 e 10.2, muitas vezes associada à produção de vídeos. As ferramentas mais comumente disponíveis para esse trabalho são os *software* de gravação de áudio, como o Digidesign Pro Tools, Steinberg Nuendo e Cubase, Cakewalk Sonar, Apple Logic Pro, Adobe Audition e Sony Vegas.

Uma tendência do mercado é a de divisão das funcionalidades por pequenas tarefas, com a criação de *plug-ins* que podem ser utilizados em diversos *softwares*

¹³ Disponível em: <<http://www.ludd.luth.se/~torger/almusvcu.html>>. Acesso em: 17 julho 2006.

¹⁴ Disponível em: <<http://www.videolan.org/vlc/>>. Acesso em: 17 julho 2006.

de processamento de áudio, comunicando-se através de padrões bem estabelecidos como o Steinberg VST¹⁵, Digidesign RTAS¹⁶ e o Microsoft DirectX¹⁷.

Figura 14 – Exemplos de softwares para espacialização sonora

¹⁵ Disponível em: <<http://www.steinberg.net/331+M52087573ab0.html>>. Acesso em: 22 julho 2006.

¹⁶ Disponível em: <<http://www.digidesign.com/>>. Acesso em: 22 julho 2006.

¹⁷ Disponível em: <<http://www.microsoft.com/windows/directx/>>. Acesso em: 22 julho 2006.

As mesas de sons também vêm equipadas para o processamento do som *surround*, geralmente utilizando a configuração 5.1. Alguns exemplos são as mesas da Yamaha 01V96V2 e DM2000, cujas interfaces gráficas podem ser vistas na Figura 15.

Figura 15 – Exemplos de mesas de som com espacialização sonora

Utilizando programas como o Pure Data e o MAX/MSP, é possível juntar blocos que executam funções específicas, como, por exemplo, um *reverb* ou um espacializador quadrafônico, e criar blocos que não estejam disponíveis. Assim, podem-se obter ferramentas mais complexas, que realizem todos os blocos necessários de um sistema de espacialização.

Um exemplo de *software* que roda sobre essas ferramentas é o Spat, do IRCAM, que pode ser acoplado a um seqüenciador musical para auxiliar o compositor na espacialização de sua obra, além de poder ser utilizado com um sistema de

realidade virtual. Ele é capaz de simular a propagação acústica baseado na geometria e em parâmetros acústicos de uma sala.

Outro *software* é o nSlam¹⁸, da Open Territories, que é um conjunto de blocos para o Pure Data que também tem o objetivo de prover a espacialização sonora, utilizando técnicas de auralização como o panorama por amplitude e o SVPA (ver item 3.4). Ele tem o código aberto e pode ser obtido gratuitamente.

Baseado no diagrama de blocos dos sistemas de espacialização, estas ferramentas são capazes de fornecer as fontes sonoras (através da síntese ou gravação), manipulá-las, codificá-las e armazenar em algum tipo de mídia, ou seja, não executam a cadeia completa. Além disso, como focam apenas uma parte do processo, as ferramentas não avaliam como estão gerando importantes atributos espaciais (ver item 2.1), possivelmente degradando a qualidade espacial final.

A abordagem modular é utilizada neste trabalho, similar aos dois *softwares* dados como exemplo, acrescentando o fato de todo o desenvolvimento estar baseado em uma plataforma genérica para projeção de áudio multicanal (AUDIENCE).

Outra tendência é a criação e uso de normas sobre som envolvente, como a norma de 5.1 canais da ITU e o THX, com o intuito de uniformizar a distribuição de alto-falantes e a qualidade de uma reprodução de som envolvente.

A norma ITU-R BS.775 (ITU-R, 1993) define o posicionamento dos alto-falantes e dos canais de áudio na configuração 5.1 e 7.1, conforme apresentado na Figura 16. Essa norma não define como o áudio é codificado nem transmitido. Desta forma, ela serve apenas como uma indicação do conteúdo que deve ser reproduzido pelos alto-falantes, sendo os três canais frontais utilizados para as fontes sonoras primárias, como diálogo e música, e os dois canais traseiros para efeitos especiais e ambiência.

O THX¹⁹ é um conjunto de normas, criadas por Tomlinson Holman na Lucasfilm, originalmente concebido para cinemas. Ele tem o objetivo de assegurar que o som ouvido durante a apresentação de um filme no cinema seja o mesmo que foi feito durante a mixagem no estúdio (RUMSEY, 2001). Para aperfeiçoar as características

¹⁸ Disponível em: <<http://tot.sat.qc.ca/eng/nslam.html>>. Acesso em: 19 dezembro 2006.

¹⁹ Disponível em: <<http://www.thx.com/>>. Acesso em: 17 julho 2006.

acústicas e níveis de ruído utilizam-se regras para construção da sala e equipamentos a serem usados. Testes regulares são feitos nestas salas para garantir que o sistema esteja dentro dos padrões THX e que a sala possa continuar a usar o selo de garantia THX. Mais recentemente, as normas foram expandidas para considerar *home theaters*, jogos eletrônicos e áudio profissional.

Figura 16 – Disposição dos alto-falantes no formato 5.1 e 7.1, de acordo com a norma ITU-R BS.775.1²⁰

²⁰ Figura extraída de (ITU-R, 1993).

3 TÉCNICAS DE AURALIZAÇÃO

As técnicas de imersão aural ou auralização tem o objetivo de recriar campos sonoros tridimensionais através de modelamentos físicos e matemáticos (FARIA, 2005).

Muitos fatores são considerados no desenvolvimento de cada técnica, como custo computacional, exigência de equipamento (como alto-falantes), número de pessoas inseridas e qualidade da recriação do campo sonoro.

A seguir, faz-se uma breve descrição das técnicas tradicionalmente utilizadas, bem como de técnicas novas e com uma maior capacidade de recriação do campo sonoro. Uma maior ênfase é dada ao Ambisonics, técnica escolhida para a implementação do sistema neste trabalho.

3.1 BI-AURAL

O sistema de projeção sonoro bi-aural baseia-se no fato de o som chegar de forma diferente em cada ouvido, uma vez que esses estão posicionados em lugares distintos. Pode-se dizer que o som que parte de uma mesma fonte é processado por funções de transferências distintas, chamadas de *Head Related Transfer Function* (HRTF), para o ouvido direito e esquerdo. Ele é análogo ao sistema de projeção visual em estéreo, no qual duas imagens, uma para cada olho, são projetadas com o fim de produzir uma imagem tridimensional.

Diferente de outros sistemas que reproduzem um campo sonoro tridimensional, o bi-aural necessita de apenas dois canais de áudio para dar bons resultados audíveis. Para a reprodução, a técnica necessita dos fones de ouvido ou de alto-falantes muito próximos aos ouvidos. Este tipo de sistema é invasivo, uma vez que necessita o uso de fones de ouvido por parte do usuário, além de dificultar o compartilhamento do sistema quando usado por mais de uma pessoa.

Figura 17 – Funções de transferências dos ouvidos esquerdo e direito

A HRTF consiste numa representação matemática da transformação que um som sofre desde a fonte sonora até o ouvido humano. Ela é usualmente expressa segundo uma medição da resposta impulsiva (ou *Impulse Response*, IR) na altura dos ouvidos humanos. Conforme se pode observar na Figura 17, a geração de um som $x(t)$ na posição da fonte sonora é ouvido pela cabeça como $x_d(t)$ no ouvido direito e $x_e(t)$ no ouvido esquerdo, ou seja, existe uma diferença entre esses sons. As respostas impulsivas medidas são relativas a cada ouvido, $h_d(t)$ e $h_e(t)$.

Com essas respostas impulsivas gravadas, h_d e h_e , é possível simular como um som seria ouvido caso fosse gerado naquela dada posição da fonte e recebida naquela posição da cabeça. Isto é possível mediante a aplicação do teorema da convolução (OPPENHEIM; SCHAFER, 1998), pelo qual o som simulado seria obtido através da resposta impulsiva convoluída com um som qualquer.

Para que a reprodução do som seja real, deve-se encontrar estas funções de transferência por meio de microfones posicionados na entrada do canal auditivo, fazendo o levantamento das respostas impulsivas dos ouvidos para cada posição em torno do ouvinte.

A forma mais usual de fazer gravações bi-aurais é com a utilização de uma *Dummy Head*, que é um modelo de uma cabeça humana no qual microfones são

posicionados em cada ouvido e, por meio de estímulos sonoros (estouros ou *sine sweep*), as respostas de cada ouvido artificial são gravadas por esses microfones. Elas podem ser feitas em laboratório ou adquiridas comercialmente, como mostra a Figura 18. Mais recentemente, técnicas para sintetizar HRTF vêm sendo desenvolvidas (RUMSEY, 2001).

Figura 18 – *Dummy Head* KU100, da Neumann²¹

O grande problema deste tipo de abordagem é que não existe um sistema auditivo similar ao outro (e.g. cabeça, pina), assim, o som é percebido de forma distinta para duas pessoas posicionadas no mesmo local. O resultado obtido com o uso de HRTFs genéricas (gravadas com *Dummy Heads*) pode ser pobre, chegando em alguns casos a confundir o sistema auditivo (GLASGAL, 2001). Já HRTFs personalizadas oferecem um resultado muito próximo da realidade. Apesar dos problemas, a técnica é utilizada em gravações comerciais e, devido a sua facilidade de implementação, é muito utilizada em jogos eletrônicos.

Experimentos com o sistema bi-aural, utilizando HRTF em um ambiente de realidade virtual imersiva, foram feitos pelo autor (THOMAZ et al., 2005). Um recente

²¹ Figura obtida no sítio do fabricante: <http://www.fullcompass.com/Products/pages/SKU--27605>.

levantamento de banco de dados de HRTF foi conduzido por Kim (2005) para um melhor modelamento tridimensional das fontes sonoras, utilizando como referência o padrão Dolby e, assim, voltado para o mercado doméstico. Uma aplicação interessante, recentemente desenvolvida por Davis (2005), utiliza HRTF para reproduzir um campo-sonoro remoto, registrado através de uma matriz de microfones em formato esférico, em outra localidade, onde o usuário utiliza fone de ouvidos com um rastreador para posicionar corretamente o campo sonoro. Experimentos para uma melhor modelagem do sistema auditivo humano (orelha, pina) focadas na utilização de HRTF estão sendo conduzidos por Dobrucki (2005), Ramírez e Rodríguez Soria (2005) e Rodríguez Soria (2006).

3.2 DIPOLO ESTÉREO

A técnica estéreo tradicional é constituída por dois alto-falantes, separados por distâncias variáveis, cada qual com o seu canal de reprodução. A divisão de som entre os canais pode ser feita artificialmente, utilizando-se panorama por amplitude, ou através de um arranjo de microfones.

O principal problema do estéreo vem do fato de, devido ao posicionamento de 60° de seus alto-falantes, existir uma grande interferência entre as ondas (*cross-talk*). Uma forma de contornar esse problema é com utilização de filtros anti-interferência, de difícil obtenção.

Kikerby, Nelson e Hamada (1997) demonstraram que, com uma separação de 10° dos alto-falantes, a interferência é menor e a construção dos filtros anti-interferência é simplificada. A Figura 19 mostra as montagens estéreo e dipolo estéreo.

Figura 19 – Montagem dos sistemas estéreo e dipolo estéreo

A técnica do dipolo estéreo funciona a partir de gravações bi-aurais (ver item 3.1) para cada alto-falante, no posicionamento de 10° entre os alto-falantes, na obtenção de filtros anti-interferência e a aplicação através da convolução destes filtros com a gravação bi-aural.

O dipolo estéreo apresenta uma boa reprodução espacial e um bom *sweet spot*²², mesmo com o movimento da cabeça, além de oferecer uma liberdade ao usuário de que o sistema bi-aural (que utiliza fones de ouvido) não dispõe.

²² *Sweet Spot*, neste contexto, refere-se ao local onde a audição é ideal, ou seja, onde o campo sonoro é reconstruído corretamente.

3.3 DOLBY SURROUND E DTS

O sistema Dolby, em suas diversas versões, é o padrão de facto de áudio envolvente atualmente no mercado. Um breve histórico do sistema Dolby é dado a seguir (DOLBY, 1999).

No meio da década de 70, enquanto o padrão quadrafônico não obtinha sucesso no mercado doméstico devido aos seus vários problemas técnicos, a Dolby lançou o Dolby Stereo, um sistema de áudio envolvente para cinema que era embutido nos rolos de filme, codificados em um sinal ótico estéreo, contendo quatro canais de áudio: esquerda (L), centro (C), direita (R) e surround (S). Este padrão é encontrado atualmente em quase todos os cinemas ao redor do mundo.

No início da década de 80, o Dolby Surround foi introduzido, com o objetivo de trazer para o ambiente doméstico as capacidades do sistema Dolby Stereo, mantendo os seus quatro canais codificados. A princípio não era possível decodificar o canal central devido ao alto custo que implicaria no decodificador doméstico, mas, com a chegada do Dolby Pro Logic, este canal passou a ser reproduzido. A codificação feita no Dolby Surround funciona através de um processo de matriciação (DRESSLER, 1993).

O próximo passo da Dolby foi trazer para o universo digital sua codificação de som envolvente, e no final da década de 80 surgiu o Dolby Digital, com o *surround* em estéreo. Esse sistema atende à configuração 5.1 (ITU, 1993) de distribuição de alto-falantes (centro, esquerda, direita, *surround* esquerda, *surround* direita e baixas frequências). Além de ser utilizado em cinemas, o Dolby Digital chegou ao mercado doméstico por meio dos *laser discs*, DVD, TV por cabo, TV digital e outros sistemas multimídia.

A técnica de codificação utilizada no Dolby Digital é o Dolby AC-3 (TODD, 1994). A codificação AC-3 utiliza conceitos psicoacústicos para retirar informação que o ser humano não necessita do sinal de áudio (MOORE, 2003). Um banco de filtros psicoacústico é utilizado para analisar e codificar o sinal de áudio de entrada.

A codificação é feita de forma a se adaptar ao decodificador disponível, chamada de *downmixing*, ou seja, uma trilha em Dolby Digital pode ser decodificada por um

decodificador Pro Logic. A Tabela 5 apresenta uma comparação dos sistemas Dolby disponíveis.

O grande concorrente do sistema Dolby no mercado profissional e doméstico é o DTS, técnica similar de auralização multicanal que também se baseia na norma ITU-R BS.775.1 para distribuição dos alto-falantes. A maioria das mídias vem provida de faixas de áudio nos dois padrões, e os *receivers* atuais também suportam os dois sistemas.

Tabela 5 – Tecnologias Dolby e suas aplicações²³

<i>Sistema</i>	<i>Tipo</i>	<i>Canais</i>	<i>Principal Uso</i>
Dolby Stereo	Analógico	4 (centro, esq., dir, surround)	Cinema
Dolby Surround	Analógico	4 (centro, esq., dir, surround)	Laser disc, TV, VHS
Dolby Digital	Digital	5.1 (centro, esq., dir., esq. surround, dir. surround, baixas freq.)	Laser disc, DVD, TV a cabo, multimídia
Dolby Digital EX	Digital	6.1 ou 7.1	TV, DVD
Dolby Digital Surround EX	Digital	6.1	Cinema
Dolby Headphone	Digital	2 (fones de ouvido)	Entretenimento Doméstico, Jogos, PC

De acordo com Rumsey (2001), os sistemas baseados na norma ITU-R BS.775.1 são na verdade um sistema estéreo de três canais com um canal surround traseiro utilizado para efeitos especiais. Não é possível obter uma imagem estável de 360° em torno do ouvinte, assim têm-se buracos na formação do campo sonoro nos lados e atrás do ouvinte.

O sistema Dolby, embora venha vagarosamente perdendo terreno no mercado devido à introdução de novas técnicas de auralização, ainda tem uma grande aceitação e praticamente toda a produção cinematográfica lança seus títulos com suporte ao formato.

²³ De acordo com (BERG; RUMSEY, 2003).

3.4 SISTEMA VETORIAL DE PANORAMA POR AMPLITUDE

O Sistema Vetorial de Panorama por Amplitude (SVPA), ou em inglês *Vector Based Amplitude Panning* (VBAP), foi desenvolvido por Pulkki (2001) e consiste em um sistema no qual um mesmo sinal sonoro é distribuído com diferentes amplitudes utilizando um arranjo de alto-falantes bi ou tridimensional.

Figura 20 – Vetores e posicionamento de alto-falantes no SVPA²⁴

O funcionamento é similar aos tradicionais sistemas de panorama por amplitude, como o estéreo. No caso do SVPA, existe uma generalização deste método para ser possível a recriação de ambientes sonoros.

²⁴ Figura extraída de (PULKKI, 2001)

A Figura 20 ilustra como o SVPA recria uma fonte sonora virtual, utilizando um trio de alto-falantes e a fonte virtual posicionada entre eles. Os vetores l_k , l_m e l_n indicam a posição dos alto-falantes com relação ao ouvinte. O vetor p , que representa a posição da fonte virtual, é expresso a partir de uma combinação linear dos vetores dos alto-falantes através da aplicação de ganhos (g_k , g_m , g_n) específicos para o sinal de cada alto-falante.

Em situações reais de utilização do sistema, mais de três alto-falantes são necessários para poder posicionar o som em torno do ouvinte. Neste caso, a escolha do trio de alto-falantes utilizado para produzir o som depende de três regras determinadas por Pulkki (1998). Definindo-se um triângulo para cada trio de alto-falantes possível (sem existir triângulos com algum alto-falante dentro), os alto-falantes de um trio não podem estar no mesmo plano do ouvinte, os triângulos não podem se sobrepor e os triângulos devem ter o menor lado possível.

Uma das vantagens do sistema baseado em SVPA é o livre posicionamento e número de alto-falantes. Todavia, como os vetores dos alto-falantes são calculados a partir da posição do ouvinte, tem-se um *sweet spot* menor. Esses vetores poderiam ser redefinidos caso a posição do ouvinte fosse rastreada, melhorando a qualidade da recriação do campo sonoro do sistema.

Através de experimentos, Pulkki (2001) chegou à conclusão que o sistema reproduz a direcionalidade de forma apurada quando o conjunto de alto-falantes encontra-se próximo a um plano mediano, que divide a cabeça do ouvinte em esquerda e direita. Já quando este conjunto é movido para o lado, a direcionalidade tende ao plano horizontal, que divide a cabeça do ouvinte em cima e abaixo. Outro problema do SVPA é que este não leva em conta parâmetros psicoacústicos importantes, como a dependência da amplitude com relação à frequência, diferente de técnicas como o Ambisonics.

A técnica SVPA vem sendo utilizado em várias pesquisas acadêmicas que necessitam de espacialização sonora tridimensional pelo seu baixo custo em equipamento, facilidade de implementação e custo computacional, como por Ogi (2003), Childs (2003), Sheridan (2004) e Schmitz (2004).

O sistema descrito por Ogi (2003) utiliza o SVPA para criar fontes sonoras virtuais em um ambiente de realidade virtual imersiva. O SOUNDSTUDIO4D (2004) utiliza

uma interface de gestos para sintetizar e espacializar campos sonoros, utilizando SVPA com uma montagem de oito alto-falantes. O SAFIR (SCHMITZ, 2004) é uma biblioteca Java para espacializar fontes sonoras de forma interativa também baseada em SVPA, focada em montagens arbitrárias de alto-falantes. Em uma aplicação incomum, Childs (2003) utiliza SVPA para auralizar dados meteorológicos de forma a descobrir padrões em tempestades, afirmando que a utilização de auralização tridimensional possibilita conclusões que não seriam obtidas a partir da visualização tradicional. O sistema de realidade virtual DIVA (PULKKI; LOKKI, 1998) utiliza o SVPA para gerar campos sonoros imersivos.

3.5 AMBIOPHONICS

A técnica de gravação e reprodução *surround* chamada Ambiphonics, descrita por Glasgal (2001), combina o cancelamento da interferência entre um par de alto-falantes (dipolo estéreo), ambiência real da sala e posicionamento de alto-falantes, levando em conta considerações psicoacústicas, para recriar o som próximo do realismo de uma sala de concerto.

O princípio da técnica é separar o som direto, reproduzido por pares de alto-falantes (chamados de Ambiopolos), da ambiência, reproduzido por outro arranjo de alto-falantes. Ela é voltada para o uso doméstico em *home theaters*, uma vez que comporta somente de um a dois usuários. Uma montagem completa é mostrada na Figura 21.

A técnica é escalável e Ambiopolos podem ser adicionados para sons traseiros e elevados, exigindo um maior número de canais, recriando o campo bidimensional e adicionando a sensação de elevação. Ela também pode ser utilizada para reproduzir conteúdo existente estéreo e 5.1, produzindo melhores resultados (GLASGAL, 2003).

Necessitando apenas de seis canais em sua configuração máxima, a técnica pode ser incluída em mídias como o SACD e o DVD-Audio. Glasgal (2003) considera o Ambiphonics o sucessor natural do sistema estéreo e 5.1 para ambientes

doméstico devido sua facilidade e baixo custo de gravação e reprodução. Apesar disso, poucas gravações comerciais já foram feitas e a técnica é usada somente por audiófilos.

Figura 21 – Montagem Ambiophones completa, utilizando os 6 canais

3.6 WAVE FIELD SYNTHESIS

A técnica *Wave Field Synthesis* (WFS), proposta por Berkhout (1988), é baseada na reconstrução da frente de onda original a partir da superposição de várias ondas, conforme o princípio de Huygens. Essa abordagem permite que a reprodução seja muito próxima da onda original, uma vez que as propriedades espaciais e temporais

desta são mantidas. Através de uma matriz densa de alto-falantes, a técnica reconstrói a onda original, como pode ser vista na Figura 22.

O campo sonoro reconstruído é válido para toda a área de reprodução, proporcionando um grande *sweet spot*. Outra vantagem é a possibilidade de poder recriar uma fonte sonora dentro da área de audição.

Figura 22 – Montagem WFS frontal, na qual uma matriz de alto-falantes reproduz a onda original

A gravação espacializada na técnica WFS é feita por meio de uma matriz de pares de microfones que envolvem a fonte sonora a ser gravada. Estes microfones captam a pressão e a velocidade das partículas da onda. A reprodução é feita através de uma matriz idêntica, em que os pares de microfones são substituídos por pares de alto-falantes, sendo um alto-falante um monopolo (onidirecional), para a reprodução da velocidade, e o outro um dipolo, para a reprodução da pressão (DANIEL; NICOL; MOREAU, 2003). A Figura 23 ilustra uma montagem de gravação e reprodução com a técnica WFS, utilizando-se uma matriz densa de alto-falantes.

Como alternativa, pode ser feita uma gravação virtual, na qual a propagação das ondas de uma cena virtual até os microfones é calculada (DANIEL; NICOL; MOREAU, 2003). Este tipo de uso é ideal para aplicações de realidade virtual pois consegue reproduzir com realismo um campo sonoro virtual, suportando várias

peças simultaneamente. Todavia, considerando o número de sinais que devem ser gerados (matriz densa de alto-falantes), ele é muito custoso computacionalmente.

Figura 23 – Montagem WFS circular para gravação e reprodução, com uma matriz densa de microfones e alto-falantes

Pesquisas sobre o uso do WFS vêm sendo conduzidas pelo instituto Fraunhofer (IOSONO²⁵), para uso em cinema, France Telecom R&D (DANIEL; NICOL; MOREAU, 2003) e em projetos como o AUDIENCE (FARIA, 2005).

²⁵ Disponível em: <http://www.idmt.fraunhofer.de/de/projekte_themen/iosono.htm>. Acesso em: 30 ago. 2006.

3.7 AMBISONICS

No início da década de 70, o sistema estéreo já se encontrava difundido nos cinemas e partia para se popularizar na televisão, rádio e sistemas de som domésticos.

Na mesma época, o primeiro sistema envolvente entrou no mercado, o quadrafônico. Diferente do estéreo, que só produz uma imagem frontal, ele é baseado em quatro alto-falantes colocados nos vértices de um quadrado com o ouvinte no meio. O sistema não conseguiu se sustentar devido aos problemas inerentes de sua construção, como a dificuldade em se estabelecer um padrão de codificação (BAMFORD; VANDERKOOY, 1995) e, principalmente, a instabilidade no campo gerado, havendo a criação de muitos buracos acústicos entre os alto-falantes (GERZON, 1974a).

Partindo dos problemas do sistema quadrafônico, mas com uma abordagem completamente diferente, Gerzon (1973) desenvolveu um sistema capaz de gravar e reproduzir campos sonoros, levando em conta que ele deveria ser capaz de reconstruir a onda sonora original e considerar parâmetros psicoacústicos durante a reprodução. Outra novidade do sistema era a possibilidade de recriação do campo sonoro com altura (perifonia). (GERZON, 1973, 1974b, 1980, 1985, 1992).

Embora promissor e muito mais sofisticado conceitualmente que as soluções disponíveis na época, ele não teve sucesso em entrar no mercado por problemas no licenciamento da tecnologia, realizado de forma ineficaz por um órgão de investimento do governo britânico (ELEN, 1991).

Além disso, a mídia para transmissão do conteúdo na época se restringia a um baixo número de canais, mesmo utilizando-se técnicas de matriciação, permitindo apenas a gravação de sistemas de primeira ordem (quatro canais). Assim, o Ambisonics ficou restrito ao ambiente acadêmico e a algumas utilizações em estúdios.

O interesse pelo sistema aumentou na década de 90 e atualmente o Ambisonics, embora ainda restrito ao ambiente de pesquisa e a audiófilos, conta com uma grande comunidade de usuários, pesquisadores e desenvolvedores. O suporte tecnológico para o sistema vem aumentando constantemente, com o

desenvolvimento de codificadores/decodificadores, principalmente em *software*, e sistemas de gravação.

O sistema foi desenvolvido inicialmente como um esquema de gravação, feito a partir de um microfone especial criado por Gerzon (1973). Ele montou um protótipo composto por um microfone onidirecional e três microfones figura de oito, cada um direcionado em um eixo, posicionados praticamente no mesmo ponto.

Posteriormente, um microfone específico foi construído, sem a necessidade da montagem com quatro microfones, que pode ser visto na Figura 24, com detalhe para a cápsula (à direita). Esse microfone é comercializado atualmente pela empresa Soundfield.

Figura 24 – Sistema Soundfield SPS422B (com detalhe da cápsula interna do microfone)²⁶, para captura de campo sonoro em B-Format de 1ª ordem

O grande problema da gravação através de microfone é a inexistência de modelos comerciais que gravem em ordens superiores a primeira, devido à dificuldade de se montar uma cápsula capaz de comportar nove, dezesseis ou mais transdutores.

²⁶ Disponível em: <<http://www.soundfield.com/>>. Acesso em: 26 julho 2006.

Assim, a reconstrução do campo acústico reconstruído não tem a mesma estabilidade e direcionalidade que sistemas de ordens superiores.

Pesquisas vêm sendo conduzidas para a construção de microfones de altas ordens sobre a teoria de microfones de segunda ordem. As técnicas utilizadas para estes microfones são matrizes horizontais circulares (POLETTI, 2000) e matrizes esféricas (MEYER; ELKO, 2002) e (ABHAYAPALA; WARD, 2002). Mais recentemente, experimentos vêm sendo realizados com um microfone de quarta ordem baseado em trinta e duas cápsulas posicionadas sobre uma esfera (DANIEL; MOREAU, 2004).

A codificação em Ambisonics também pode ser feita por meio da síntese, em que um sinal monofônico é posicionado no espaço através de manipulações com ganhos em cada canal. Esses ganhos são provenientes das equações de codificação Ambisonics, que são derivadas diretamente do formalismo matemático/físico da técnica (ver Apêndice A).

Figura 25 – Coordenadas esféricas utilizadas nas equações de codificação Ambisonics

A Tabela 6 sumariza as equações de codificação até a terceira ordem, sendo θ a elevação e ϕ a rotação da fonte relativa aos eixos, como mostra a Figura 25, e S o sinal monofônico ao qual será aplicado o ganho.

As fontes virtuais, através destas equações, são posicionadas sobre uma esfera unitária com centro na origem ($|\vec{r}|=1$), assim, para obter a distância desejada da fonte, deve-se aplicar atenuações ou amplificações ao sinal.

Tabela 6 – Equações de codificação Ambisonics até a 3ª ordem

Ordem (m)	Canal Ambisonics	Equação de Codificação
0	W	$\sqrt{2} \cdot S$
1	X	$\cos \phi \cdot \cos \theta \cdot S$
	Y	$\sin \phi \cdot \cos \theta \cdot S$
	Z	$\sin \theta \cdot S$
2	R	$(3 \sin^2 \theta - 1) / 2 \cdot S$
	S	$\cos \phi \cdot \cos 2\theta \cdot S$
	T	$\sin \phi \cdot \sin 2\theta \cdot S$
	U	$\cos 2\phi \cdot \cos^2 \theta \cdot S$
	V	$\sin 2\phi \cdot \cos^2 \theta \cdot S$
3	K	$\sin^2 \theta \cdot (5 \cdot \sin^2 \theta - 3) / 2 \cdot S$
	L	$8 \cdot \cos \phi \cdot \cos \theta \cdot (5 \cdot \sin^2 \theta - 1) / 11 \cdot S$
	M	$8 \cdot \sin \phi \cdot \cos \theta \cdot (5 \cdot \sin^2 \theta - 1) / 11 \cdot S$
	N	$\cos 2\phi \cdot \sin \theta \cdot \cos^2 \theta \cdot S$
	O	$\sin 2\phi \cdot \sin \theta \cdot \cos^2 \theta \cdot S$
	P	$\cos 3\phi \cdot \cos^3 \theta \cdot S$
	Q	$\sin 3\phi \cdot \cos^3 \theta \cdot S$

O armazenamento e transmissão de sinais Ambisonics dependem principalmente da ordem do sistema, que define o número de canais. Além disso, estes sinais podem sofrer algum tipo de codificação visando à utilização em outros sistemas, como o estéreo e o 5.1. Alguns formatos foram propostos, sumarizados na Tabela 7, sendo

que o B-Format é o mais utilizado nos diversos codificadores/decodificadores Ambisonics.

Tabela 7 – Formatos de armazenamento e transmissão de canais Ambisonics

<i>Formato</i>	<i>Ordem suportada</i>	<i>Número de Canais</i>	<i>Descrição</i>
A-Format	1ª	4	Formato interno utilizado pelo sistema de gravação Soundfield
B-Format	1ª horizontal	3 (WXY)	Formato mais utilizado em codificadores/decodificadores
	1ª 3D	4 (WXYZ)	
	2ª horizontal	5 (WXYUV)	
	2ª horizontal + 1ª vertical	6 (WXYZUV)	
	3ª horizontal	7 (WXYUVPQ)	
	3ª horizontal + 1ª vertical	8 (WXYZUVPQ)	
	2ª 3D	9 (WXYZRSTUV)	
	3ª horizontal + 2ª vertical	11 (WXYZRSTUVPQ)	
	3ª 3D	16 (WXYZRSTUVKLMNOPQ)	
C-Format (UHJ)	1ª horizontal (estéreo)	2 (Σ , Δ)	Codificação em matriz do B-Format, suporta estéreo comum, sendo Σ e Δ , respectivamente, a soma e a diferença dos sinais esquerdo e direito do estéreo convencional
	1ª horizontal	3 (Σ , Δ , T)	
	1ª 3D	4 (Σ , Δ , T, Q)	
G-Format		5.1	Decodificação pré-realizada em estúdio de Ambisonics para a configuração de alto-falantes 5.1

Uma propriedade interessante no uso das equações de codificação, descrita por Malham e Myatt (1995), é a capacidade de manipular o campo sonoro através de operações trigonométricas simples, com baixo custo computacional. Essa

propriedade pode ser utilizada em aplicações de realidade virtual, nas quais é importante executar ações como a rotação e inclinação, tanto do espaço virtual como do usuário. A Tabela 8 apresenta as equações para obtenção dos novos valores de cada canal para o B-Format de 1ª ordem.

Estas operações podem ser aplicadas tanto a campos sonoros previamente gravados com um microfone, ou com campos sintetizados através das equações de codificação. Além disso, combinações dessas operações podem ser feitas, e aplicadas por meio de matrizes de ganho aos sinais de codificados (MALHAM; MYATT, 1995).

Tabela 8 – Manipulações do campo sonoro em Ambisonics de 1ª ordem

<i>Operação</i>	<i>Descrição</i>	<i>Equações</i>
Rotação	Rotação sobre o eixo z	$W' = W$ $X' = X \cos \phi - Y \sin \phi$ $Y' = X \sin \phi + Y \cos \phi$ $Z' = Z$
<i>Tilt</i>	Rotação sobre o eixo y (movimento esq.-dir.)	$W' = W$ $X' = X$ $Y' = Y \cos \theta - Z \sin \theta$ $Z' = Y \sin \theta + Z \cos \theta$
<i>Tumble</i>	Rotação sobre o eixo z (movimento frente-trás)	$W' = W$ $X' = X \cos \theta - Z \sin \theta$ $Y' = Y$ $Z' = X \sin \theta + Z \cos \theta$
Rotação e <i>Tilt</i>	Operações conjuntas	$W' = W$ $X' = X \cos \phi - Y \sin \phi$ $Y' = X \sin \phi \cos \theta + Y \cos \phi \cos \theta - Z \sin \theta$ $Z' = X \sin \phi \sin \theta + Y \cos \phi \sin \theta + Z \cos \theta$

A decodificação dos canais Ambisonics é feita através de matrizes de ganho previamente calculadas. Como o cálculo destes ganhos pode ser extremamente complicado dependendo da configuração de alto-falantes utilizada, adotam-se matrizes pré-calculadas para as configurações mais comuns.

Figura 26 – Montagem Ambisonics octogonal de alto-falantes

Furse (2006) fez os cálculos para canais de primeira e segunda ordem, em configurações como as horizontais quadrada, hexagonal, octogonal, *surround* 5.1, e para as tridimensionais, cúbica e dodecaédrica. A Figura 26 mostra uma montagem octogonal e a Tabela 9 sua respectiva matriz de ganhos de primeira ordem, para cada alto-falante.

Tabela 9 – Matriz de ganhos de primeira ordem para uma configuração octogonal

Alto-Falante (n)	W	X	Y	Z
1	0.1768	0.2500	0.0000	0.0000
2	0.1768	0.1768	0.1768	0.0000
3	0.1768	0.0000	0.2500	0.0000
4	0.1768	-0.1768	0.1768	0.0000
5	0.1768	-0.2500	0.0000	0.0000
6	0.1768	-0.1768	-0.1768	0.0000
7	0.1768	0.0000	-0.2500	0.0000
8	0.1768	0.1768	-0.1768	0.0000

Ganhos calculados por Furse (2006)

A Figura 27 mostra o diagrama de um decodificador Ambisonics de primeira ordem, para uma matriz de N alto-falantes, utilizando os filtros psicoacústicos.

Figura 27 – Diagrama de blocos de um decodificador Ambisonics de 1ª ordem com filtros psicoacústicos

Os ganhos α_n , β_n , γ_n e δ_n são aplicados, respectivamente, aos sinais codificados W, X, Y e Z, para cada enésimo alto-falante. Os ganhos G1 e G2 correspondem ao filtro equalizador (*shelf filter*) utilizado para fazer o ajuste psicoacústico.

Gerzon levou em conta a propriedade do ouvido humano comportar-se diferentemente para sons de várias frequências, principalmente entre as graves, abaixo de 700 Hz (GERZON, 1974b), e acrescentou parâmetros psicoacústicos no sistema Ambisonics para compensar essa diferença. Os parâmetros psicoacústicos em um sistema Ambisonics são ajustados através de filtros equalizadores, mais conhecidos e implementados por *shelf filters*.

Os valores dos *shelf filters* não são padronizados e vários autores e fabricantes de decodificadores utilizam ganhos diferentes em seus sistemas. A Tabela 10 mostra alguns desses valores para configurações bidimensionais (FARINA, 1998), sendo N o número de alto-falantes no sistema.

Tabela 10 – Sugestão de ganhos para os filtros psicoacústicos do Ambisonics

Autor	Aplicação	G1	G2
D. Malham	Sala de concerto Todas as frequências	$\sqrt{\frac{8}{3 \cdot N}}$	$\sqrt{\frac{8}{3 \cdot N}}$
M. Gerzon	Estúdio Altas frequências	$\sqrt{\frac{8}{4 \cdot N}}$	$\sqrt{\frac{8}{2 \cdot N}}$
M. Gerzon	Estúdio Baixas frequências	$\sqrt{\frac{8}{6 \cdot N}}$	$\sqrt{\frac{8 \cdot 2}{3N}}$
M. Gerzon (Patente)	>500 Hz <500 Hz	$\sqrt{2}$ 1	$\sqrt{2}$ $\sqrt{3}$
J. M. Lot	Estúdio Baixíssimas frequências	$\sqrt{\frac{8}{2 \cdot N^2}}$	$\sqrt{\frac{8 \cdot 2}{N^2}}$

Os valores são discretos para as duas faixas de frequência e novos estudos sugerem que deveria existir uma variação contínua entre os dois valores, com uma

transição progressiva entre baixa e alta frequência. Farina (1998) fez uma implementação desta sugestão, em que os valores podem ser alterados graficamente e testes podem ser feitos para obter o melhor resultado psicoacústico do sistema.

O uso de sistemas Ambisonics de primeira ordem apresenta problemas na reconstrução acústica da onda, principalmente no pequeno tamanho do *sweet spot*. A tendência atual é pesquisar sistemas de altas ordens de forma a minimizar o problema do *sweet spot*, aumentando a área de reconstrução acústica (DANIEL; NICOL; MOREAU, 2003; DANIEL; MOREAU, 2004), tanto através do desenvolvimento de microfones de alta ordem como na síntese virtual utilizando-se as equações de codificação.

A tendência atual é a pesquisa de sistemas Ambisonics de altas ordens (DANIEL; NICOL; MOREAU, 2003; DANIEL; MOREAU, 2004), utilização em sistemas de realidade virtual (FARIA, 2005; FARIA et al. 2005; FARIA; ZUFFO, 2006), e sua integração com outros formatos de som envolvente, como transcódificações para sistemas 5.1.

3.8 CONCLUSÃO

A escolha de uma técnica de auralização para uma aplicação de espacialização sonora depende de vários fatores, listados a seguir:

- realismo na recriação, indica o realismo do efeito de espacialização;
- número de usuários, indica qual a capacidade de usuários simultâneos do sistema;
- facilidade de implementação, indica se a montagem do sistema de *hardware* e a programação do *software* são complicadas;
- custo computacional, indica se a técnica exige um grande processamento computacional devido aos cálculos para a reprodução da onda;

- custo de equipamentos, indica qual a necessidade de alto-falantes, amplificadores, computadores e interfaces de áudio para a montagem do sistema.

Através de uma escala de um a cinco, sendo cinco a melhor situação, esses fatores são avaliados para cada técnica de auralização na Tabela 11. Assim, espera-se que ela possa auxiliar na escolha da técnica mais adequada, dados os requisitos de uma aplicação qualquer.

Tabela 11 – Comparação entre as técnicas de auralização

<i>Técnica de Auralização</i>	<i>Realismo</i>	<i>Número de Usuários</i>	<i>Facilidade de Implementação</i>	<i>Custo Computacional</i>	<i>Custo Equipamento</i>
Dipolo Estéreo	2 Sem os problemas de interferência do estéreo comum, mas só reproduz som frontal	1 Suporta apenas um usuário	4 Montagem simples, requer cálculo de filtros	3 Processamento de filtros anti-interferência	4 Necessita de apenas dois alto-falantes
Bi-aural	4 Reconstrução da onda boa se forem utilizadas boas HRTFs	1 Suporta apenas um usuário	4 HRTFs prontas para utilização. Mais complicada no caso de HRTFs personalizadas	3 No caso da síntese, exige um processamento razoável	5 Necessita apenas de fones de ouvido
Dolby / DTS	1 Reconstrução com problemas nas laterais e atrás	2 Suporta poucos usuários	5 Codificadores amplamente disponíveis	4 existem <i>hardwares</i> e <i>softwares</i> específicos	4 Equipamento 5.1 e <i>receiver</i> amplamente disponível
SVPA	3 Panorama por amplitude, é capaz de reproduzir campos tridimensionais	4 Suporta vários usuários simultaneamente	4 Simple implementação computacional, baseada em ganhos	4 Não requer grande processamento	3 Depende do número de alto-falantes utilizado
Ambiophonics	3 Boa reconstrução na montagem completa	2 Suporta até dois usuários	3 Necessita da implementação do dipolo estéreo e do Ambisonics	5 Não requer processamento, gravações prontas	4 Sistema escalável, iniciando em 2 alto-falantes
WFS	5 Reconstrução da onda próxima da realidade	5 É capaz de suportar muitos usuários simultaneamente	1 Montagem complicada	1 Requer grande processamento se for feita a síntese	1 Requer uma matriz densa de alto-falantes
Ambisonics	5 Ótima reconstrução da onda para altas ordens	4 É capaz de suportar vários usuários simultaneamente	3 Montagem depende do número e disposição dos alto-falantes	3 Não requer grande processamento	3 Depende do número de alto-falantes

4 IMPLEMENTAÇÃO DO SISTEMA DE ESPACIALIZAÇÃO

Neste capítulo, é descrita a implementação do sistema de espacialização sonora baseado no AUDIENCE e na plataforma de programação Pure Data. A seguir, a modelagem do sistema é feita utilizando-se diagramas UML (BOOCH, 1999). Primeiramente os casos de uso são explicados e, a partir destes, os blocos são desenvolvidos (diagramas de classes) e sua utilização durante a execução do sistema (diagramas de seqüência). Por fim, a instalação física do sistema (*hardware*) é discutida e sua forma de utilização.

A técnica de auralização Ambisonics foi escolhida para a utilização neste trabalho devido a sua fácil implementação, baixo custo computacional, configuração flexível de alto-falantes, *sweet spot* capaz de suportar algumas pessoas simultaneamente e boa direcionalidade das fontes sonoras virtuais. A técnica também atende plenamente à necessidade de recriação de campos sonoros tridimensionais realistas, e permite a utilização de montagens flexíveis de alto-falantes, que é um requisito para a espacialização de cenas musicais.

4.1 PLATAFORMA DE PROGRAMAÇÃO

4.1.1 Pure Data

Utilizou-se o Pure Data (PD) como plataforma para a construção dos blocos de *software* do sistema e suas conexões. O PD é um ambiente de programação gráfico para aplicações musicais e de áudio. A plataforma foi desenvolvida por Miller Puckette (2006) baseada na concepção do MAX/MSP do IRCAM, de mesma autoria.

A escolha dessa plataforma foi feita por ser uma ferramenta aberta, flexível e com um tempo de reposta com baixa latência para o processamento de áudio, além de

permitir a lógica de ligação entre o subsistema de áudio e o de visualização. O código é aberto, podendo ser compilado em diversas plataformas como Windows, MacOS e Linux.

Figura 28 – Um *patch* de síntese FM feito no Pure Data

A comunidade de usuários do PD vem crescendo com o tempo, e conta com diversas listas de discussão²⁷, tanto para o seu uso como para o seu desenvolvimento. O PD é amplamente utilizado na pesquisa de áudio, como em

²⁷ Listas do PD disponíveis em <<http://lists.puredata.info/listinfo>>. Acesso em: 27 julho 2006.

trabalhos feitos por Noisternig (2003), Frauenberger, Ritsch e Höldrich (2003) e, mais recentemente, por Liuni e Morelli (2006) e Friberg (2006).

A programação no PD é feita através da conexão gráfica de blocos que executam uma tarefa específica, criando um *patch*. Esse termo é muito utilizado no meio musical e de sonorização, e refere-se a um esquema ou diagrama de conexões entre diversos equipamentos. O termo *patch* é muito utilizado também desde os primórdios dos sintetizadores analógicos, construídos por meio de cabos interligando diferentes dispositivos independentes, como osciladores e controladores.

A Figura 28 mostra um *patch* que implementa um sintetizador FM simples, com diversos blocos de funcionalidades simples que se conectam para formar um sistema complexo.

O fluxo de dados entre os blocos funciona através de mensagens que são passadas da saída de um bloco (*outlet*) conectado à entrada de outro (*inlet*). Estas mensagens podem ser de controle (números, listas de números e *strings*), que são passadas quando ocorre algum estímulo no bloco, ou podem ser sinais, que são processados continuamente quando o processamento de áudio está ligado. A ferramenta disponibiliza alguns blocos básicos internos, chamados de *internals*, para o processamento do áudio e controle.

O PD permite que novos blocos sejam criados pelo usuário²⁸, chamados de *externals*. A programação destes blocos é feita na linguagem C, e deve seguir algumas regras para que sejam utilizados em um *patch*. Estas regras incluem a criação de uma estrutura de dados para armazenar os parâmetros do bloco e a implementação de algumas funções de *callback* básicas que são utilizadas pelo programa principal para se comunicar com o bloco.

Existem dois tipos de *externals*: o de controle, chamado *objectclass*; e o de sinal, chamado de *signalclass*. A diferença básica é que o primeiro processa apenas mensagens de controle, enquanto que o segundo pode utilizar ainda sinais de áudio, processados de forma contínua.

²⁸ Um tutorial de como criar *externals* no PD pode ser consultado na Internet. Disponível em: <<http://iem.at/pd/externals-HOWTO/>>. Acesso em: 13 fev. 2007.

A Figura 29, através de um diagrama de classes UML, ilustra a relação dos *patches*, *externals* e *internals* entre eles e com o PD. Três estereótipos²⁹ de classes foram definidos para representar as classes básicas de construção de um *patch* no PD: <<PD Patch>>, um *patch*; <<PD Objectclass>>, um *external* de controle; e <<PD Signalclass>>, um *external* de sinal. Estas definições facilitam posteriormente o modelamento de um sistema baseado no PD.

Figura 29 – Diagrama UML representando as classes do PD

O PD é utilizado no projeto como a ferramenta que liga os diferentes módulos, operando em cada uma das quatro camadas do AUDIENCE, e renderiza o áudio para reprodução final. As funções de cada camada são implementadas no PD como um ou vários *patches*, como um *external* ou como uma combinação destes elementos.

²⁹ Um estereótipo na linguagem UML serve para definir novos tipos de elementos.

4.1.2 AUDIENCE

A arquitetura de produção de som espacial do AUDIENCE, proposta por Faria (2005), está baseada em uma abordagem modular de quatro camadas funcionais, ilustradas na Figura 30. Essa divisão tem o intuito de permitir a utilização de técnicas e ferramentas diferentes na implementação das funções executadas em cada camada (e.g. programas, *plug-ins*, equipamentos externos), ao mesmo tempo em que mantém a comunicação entre elas através de uma interface pré-definida, na qual os sinais são conhecidos ou mesmo padronizados por consenso ou norma. Com essa abordagem, não é necessário refazer o sistema por inteiro no momento da troca das ferramentas e técnicas empregadas em uma das camadas. Isso garante uma maior independência e interoperabilidade entre elas, e uma maior flexibilidade na escolha de algoritmos ou ferramentas de terceiros.

A primeira camada é responsável pela composição da cena acústica, informando à camada seguinte os atributos das fontes sonoras, como a posição de objetos, posição do ouvinte e parâmetros do ambiente (e.g. parâmetros acústicos). Estes dados podem ser fornecidos via uma interface homem-máquina ou por outro sistema que utilize o AUDIENCE, como um navegador visual de realidade virtual.

A simulação acústica, na segunda camada, tem um papel muito importante na produção do realismo sonoro, pois a percepção espacial e sua qualidade estão diretamente associadas a esta. A simulação do campo sonoro calcula a propagação acústica do som proveniente da fonte até chegar ao ouvinte, considerando o meio (ambiente) em que se encontra. Várias técnicas podem ser utilizadas e implementadas a partir de modelos de simulação geométricos (e.g. traçado de raios, radiosidade, fonte-imagem), modelos ondulatórios (e.g. elementos finitos) ou modelos estatísticos.

A terceira camada, de codificação de áudio, codifica o sinal espacializado em algum formato para ser transmitido à próxima camada. Aqui podem ocorrer compressões, *stream* etc., dependendo do meio de transmissão ou da mídia utilizada.

Figura 30 – Arquitetura genérica de camadas do AUDIENCE³⁰

A última camada decodifica o sinal gerado pelo codificador e reproduz o campo sonoro final através de alto-falantes. Ela é sensível à disponibilidade de equipamentos no sistema, como número e disposição de alto-falantes, e, dependendo da escolha do formato de codificação e da disponibilidade de alto-falantes, uma grande variedade de configurações de reprodução pode ser

³⁰ Figura extraída de (FARIA, 2005).

considerada. A mixagem do sinal proveniente das várias fontes sonoras também é feita nesta camada.

Descrições mais completas sobre o AUDIENCE e aplicações deste foram feitas por Faria (2005), Faria et al. (2005), Thomaz et al. (2006a), Thomaz et al. (2006b) e Faria e Zuffo (2006).

4.2 MODELAGEM DO SISTEMA

O objetivo principal do sistema é espacializar fontes sonoras previamente gravadas através do posicionamento virtual dessas fontes e do usuário, podendo o usuário movimentar essas fontes e sua posição durante a execução da peça.

Os casos de uso do sistema são ilustrados na Figura 31. Tanto um usuário humano como outro *software* (como uma aplicação de realidade virtual) é capaz de acionar os casos de uso disponíveis. Visto que o sistema funciona da mesma forma que um reproduzidor de música, existem casos de uso para inicializar e finalizar.

O caso de uso de inicialização do sistema tem o objetivo de preparar o sistema de áudio para a execução, carregar as fontes sonoras e carregar os parâmetros iniciais da peça musical que será executada, como as posições iniciais das fontes e do ouvinte, e o ambiente virtual de execução. Com o sistema inicializado, é possível iniciar a execução da peça musical.

O caso de uso de configurar parâmetros de execução tem as funções de ajustar os parâmetros do ambiente (sala), equalização e volume das fontes sonoras, e quaisquer outros parâmetros disponíveis que não as posições das fontes e do usuário.

Durante a execução, é possível movimentar as fontes sonoras, colocando-as em posições diferentes das previamente programadas pelo compositor. Também se pode alterar a posição do usuário.

O caso de uso para parar a execução pode ser utilizado para pausar ou preparar uma nova execução. Já o caso de uso de finalização do sistema desabilita o áudio e todos os recursos após a utilização do sistema.

Os diagramas de seqüência UML dos casos de uso de iniciar execução do sistema e movimentar fontes sonoras, que ilustram as trocas de mensagens entre os blocos, encontram-se no Apêndice B.

Figura 31 – Casos de uso do sistema de espacialização sonora

O AUDIENCE e o PD, em conjunto, são utilizados para implementar o sistema de espacialização sonora. A Figura 32 mostra como cada camada do AUDIENCE se encaixa nos blocos de um sistema de espacialização genérico. Neste caso, o PD

funciona como o meio de comunicação entre as diversas camadas, tanto na passagem de mensagens de controle como no envio de sinais de áudio.

Figura 32 – Implementação do AUDIENCE baseado em um sistema de espacialização sonora genérico

As funcionalidades esperadas de cada camada foram implementadas através da criação de *patches* e *externals* do PD. Estes blocos podem ser conectados posteriormente para a criação do *patch* utilizado no experimento.

A Figura 33 ilustra uma adaptação, específica para o projeto, da arquitetura de camadas do AUDIENCE, apresentada na Figura 30. Na figura, o índice f representa o número de fontes sonoras, m o número de sinais de áudio codificados e n o número de alto-falantes. Toda a parte de sincronização entre as camadas foi descartada visto que o próprio PD é responsável pela sincronização do fluxo do áudio.

Figura 33 – Implementação da arquitetura AUDIENCE

A arquitetura AUDIENCE descreve os dados que devem ser passados entre as camadas, ou seja, as interfaces a serem seguidas, de forma a respeitar a premissa do AUDIENCE de modularidade. Assim, é possível que sejam desenvolvidos blocos com a mesma funcionalidade (e.g. codificação de áudio), mas com implementações de técnicas diferentes (e.g. B-Format e B-Format com compactação MPEG).

A Figura 33 mostra os dados que são passados entre as camadas do AUDIENCE, que podem ser tanto parâmetros (e.g. posição do ouvinte, respostas impulsivas) como sinais de áudio (e.g. sinal monofônico, sinal em B-Format). A passagem de parâmetros entre os diversos blocos é feita através de mensagens homônimas do PD, enquanto que a passagem de sinais de áudio é feita por meio de ligações exclusivas para esse fim.

A partir da descrição de cada camada, da interface entre elas e dos requisitos específicos do sistema desenvolvido, modelou-se um conjunto de blocos do PD, descritos na Tabela 12. Os blocos foram divididos em seis grupos, identificados pela sua nomenclatura, nos quais L_n indica um bloco que implementa uma camada (sendo n esta camada) do AUDIENCE, *gui* os blocos relacionados com a interface gráfica com o usuário e *aux* os blocos auxiliares.

Tabela 12 – Blocos desenvolvidos no PD, sobre a arquitetura AUDIENCE

<i>Grupo</i>	<i>Nomenclatura</i>	<i>Blocos Desenvolvidos</i>	<i>Tipo</i>	<i>Descrição</i>
Camada 1	audce_L1_*	audce_L1_pd_gui	<i>patch</i>	Controle de parâmetros de entrada baseado em uma interface gráfica
Camada 2	audce_L2_*	audce_L2_allen	<i>external</i>	Simulador Acústico baseado em Fontes Virtuais (FARIA; ZUFFO, 2006)
Camada 3	audce_L3_*	audce_L3_amb_3rd~	<i>external</i>	Codificador Ambisonics de 3ª ordem baseado na posição da fonte
		audce_L3_amb_ir~	<i>patch</i>	Codificador Ambisonics baseado nas respostas impulsivas
Camada 4	audce_L4_*	audce_L4_amb_3rd~	<i>external</i>	Decodificador Ambisonics de 3ª ordem
Interface Gráfica	audce_gui_*	audce_gui_grid	<i>patch</i>	Posicionamento das fontes sonoras e do ouvinte
		audce_gui_srcfile	<i>patch</i>	Controle do arquivo fonte de áudio
		audce_gui_volume	<i>patch</i>	Controla volumes
Auxiliares	audce_aux_*	audce_aux_convolver~	<i>external</i>	Convolucionador baseado na biblioteca FFTW
		audce_aux_mixer	<i>patch</i>	Misturador de sinais
		audce_aux_move	<i>patch</i>	Move a fonte automaticamente a partir de uma trajetória dada
		audce_aux_srcfile	<i>patch</i>	Controla a reprodução de um arquivo de áudio

A Tabela 13 contém todos os dados de entrada e saída de cada bloco implementado, bem como os parâmetros de criação de cada bloco

Tabela 13 – Dados de entrada e saída dos blocos (continua)

<i>Blocos</i>	<i>Dados</i>	<i>Tipo</i>	<i>Descrição</i>
audce_L2_allen	m	P	Ordem de codificação Ambisonics
	npts	P	Número de pontos da resposta impulsiva
	pos_src <x y z>	IN	Posição da fonte na sala
	pos_rcv <x y z>	IN	Posição do ouvinte na sala
	room_dim <width depth height>	IN	Dimensões da sala
	room_coef <c1 c2 c3 c4 c5 c6>	IN	Coefficientes de absorção de cada superfície da sala
	ir[npts]	OUT	Resposta impulsiva em B-Format
audce_L3_amb_3rd~	m	P	Ordem de codificação Ambisonics
	radius	P	Raio da esfera unitária Ambisonics (em metros)
	signal	IN	Sinal monofônico e anecóico
	signal[M]	OUT	Sinais de áudio em B-Format
audce_L3_amb_ir~	ir[npts]	IN	Resposta impulsiva em B-Format
	signal	IN	Sinal monofônico e anecóico
	signal[M]	OUT	Sinais de áudio em B-Format
audce_L4_amb_3rd~	m	P	Ordem de codificação Ambisonics
	speaker_layout	P	Tipo de arranjo dos alto-falantes
	signal[M]	IN	Sinais de áudio em B-Format
	signal[N]	OUT	Sinais de áudio decodificados para o arranjo de alto-falantes
audce_gui_grid	num_src	P	Número de fontes
	height	P	Altura do grid (em pixels)
	widht	P	Largura do grid (em pixels)
	scale	P	Escala do grid (metros/pixel)
	pos_src <num_src x y z>	IN/OUT	Posição da fonte
	pos_rcv <x y z>	IN/OUT	Posição do ouvinte na sala
audce_gui_srcfile	label	P	Nome da fonte sonora
	file	IN	Arquivo de áudio
	signal	OUT	Sinal de áudio do arquivo

Tabela 12 – Dados de entrada e saída dos blocos (conclusão)

<i>Blocos</i>	<i>Dados</i>	<i>Tipo</i>	<i>Descrição</i>
audce_gui_volume	vol	OUT	Volume do áudio
audce_aux_convolver~	ir[npts]	IN	Resposta impulsiva
	signal	IN	Sinal de áudio original
	signal	OUT	Sinal de áudio convoluido
audce_aux_mixer	n	P	Número de fontes
	level <n>	IN	Volume do sinal n de entrada
	signal[n]	IN	Sinais de áudio separados
	signal	OUT	Sinal de áudio mixados
audce_aux_move	move_list	IN	Lista descrevendo os movimentos da fonte no tempo
	pos_src <x y z>	OUT	Posição da fonte
audce_aux_srcfile	file	IN	Arquivo de áudio
	vol	IN	Volume do áudio
	signal	OUT	Sinal de áudio do arquivo

Tipo: P – Parâmetro de criação do bloco / IN – Dado de entrada / OUT – Dado de saída

O diagrama UML da Figura 34 mostra as interfaces de cada camada, bem como os diversos blocos desenvolvidos, com suas operações e atributos, e a relação entre eles.

Definiram-se estereótipos para as operações de cada classe que representam a forma como são utilizadas no PD:

- <<in>>, operação ativada quando o *external* recebe, por um *inlet* de controle, uma mensagem com o nome da operação e os parâmetros necessários;
- <<out>>, operação que, quando ativada, envia pelo outlet de controle uma mensagem com o nome da operação e os valores de retorno;
- <<audio in>>, operação que recebe sinal de áudio através do inlet de sinal;
- <<audio out>>, operação que envia sinal de áudio através do outlet de sinal.

4.3 IMPLEMENTAÇÃO DOS BLOCOS POR CAMADA AUDIENCE

4.3.1 Camada 1 – Composição de Cena Acústica

Sendo a composição e controle da cena feita toda através do PD, os únicos blocos pertencentes a essa camada são os de interface homem-máquina. Além de utilizar blocos internos do PD, ela também usa os blocos do grupo `audce_gui_*`, para a interface gráfica com o usuário.

A montagem da interface é feita no *patch* `audce_L1_pd_gui`, que controla a reprodução e serve como repositório para os outros blocos de interface gráfica. O bloco `audce_gui_grid` controla a posição das fontes e do usuário, e o bloco `audce_gui_srcfile` controla a reprodução dos arquivos de entrada e seus volumes. Esses dois componentes podem ser vistos na Figura 35.

Figura 35 – Componentes gráficos `audce_gui_grid` e `audce_gui_srcfile`

4.3.2 Camada 2 – Simulador Acústico

O bloco `audce_L2_allen` implementa um simulador acústico baseado na técnica de traçado de raios, utilizando uma adaptação do método fonte-imagem descrito por Allen e Berkley (1979). Ele emula o efeito de ambientação, principalmente a reverberação, de uma sala real.

Uma reflexão nesta técnica vem de uma fonte-imagem virtual, localizada atrás da parede, baseada nas leis da ótica geométrica. Desta forma é possível calcular as reflexões (número limitado pela capacidade de processamento) de uma onda sonora e o caminho destas até o ouvinte.

O resultado destes cálculos é uma resposta impulsiva que, aplicada através do teorema da convolução a uma fonte sonora não posicionada, simula o efeito de como o som chegaria a um ouvinte nesta sala, com a fonte na dada posição. A esta resposta impulsiva são aplicadas as equações de codificação do Ambisonics (FARIA; ZUFFO, 2006).

Esse bloco é capaz de calcular as reflexões sonoras de uma sala de geometria simples (paralelepípedo), sem obstruções, com o número previamente definido de reflexões. Os parâmetros de entrada incluem as dimensões da sala, os coeficientes de absorção de cada parede, teto e chão, a posição da fonte e do ouvinte, a ordem Ambisonics e o número de pontos da resposta impulsiva. A saída do bloco fornece vetores de respostas impulsivas para cada canal do Ambisonics.

4.3.3 Camada 3 – Codificador de áudio espacial

O bloco `audce_L3_amb_3rd~` utiliza diretamente as equações de codificação do Ambisonics (descritas na Tabela 6) para codificar o sinal de áudio de entrada, gerando o número de canais necessários para a ordem Ambisonics configurada. Nesse bloco, a fonte sonora é posicionada como se estivesse em um campo aberto

(*free field*), gerando apenas o raio direto da fonte sonora, sem criar o efeito de ambientação.

Utilizando-se as equações de codificação, os sinais de entrada são posicionados em uma esfera em torno do ouvinte, todas a uma mesma distância. Essa distância é especificada pelo parâmetro `radius`, indicando o raio, em metros, da esfera Ambisonics. Assim, pode-se calibrar a intensidade desses sons para serem realistas. A partir daí, a intensidade da fonte é atenuada ou amplificada dependendo da distância do ouvinte, além de considerar o atraso que o som tem devido à distância do ouvinte.

No bloco `audce_L3_amb_ir~` o sinal de áudio monofônico e anecóico é convolucionado com as respostas impulsivas geradas na segunda camada de forma a obter o sinal no formato Ambisonics já espacializado. Esta convolução é feita através do bloco `audce_aux_convolver~`.

O bloco de convolução de sinais `audce_aux_convolver` é implementado utilizando-se um algoritmo baseado no código presente em Press et al. (1992). O método utilizado é o da convolução *overlap-add* (OPPENHEIM; SCHAFER, 1998), usado em sinais de grande comprimento, caso do sinal de áudio. Desta forma, tem-se uma convolução contínua com baixa latência. A biblioteca FFTW³¹ é utilizada para efetuar as transformadas rápidas de Fourier, devido a sua rapidez e fácil integração ao código, tanto no sistema operacional Linux como no Windows.

4.3.4 Camada 4 – Decodificador e reproduzidor de áudio espacial

Cada instância dos blocos descritos anteriormente processa apenas uma fonte sonora. Assim, é necessário que, antes da decodificação, os sinais das diversas fontes sejam misturados em apenas um vetor de sinais no formato Ambisonics. Esta mistura ou mixagem é feita através do bloco `audce_aux_mixer~`.

³¹ Disponível em: <<http://www.fftw.org/>>. Acesso em: 14 fevereiro 2006.

O último bloco, `audce_L4_amb_3rd~`, decodifica o vetor de sinal de áudio Ambisonics para a configuração de alto-falantes dada, reproduzindo o espaço sonoro codificado. Por configuração entende-se o número e o posicionamento dos alto-falantes, como ilustrado na Figura 36.

Figura 36 – Algumas configurações de alto-falantes disponíveis para decodificação no bloco `audce_L4_amb_3rd~`

A decodificação é feita através de uma matriz de ganhos para conjunto ordem/configuração, e foram previamente calculadas por Furse (2006) até a terceira ordem. Os filtros psicoacústicos para um melhor ajuste na reprodução do campo sonoro, descritos por Gerzon (1974) e no item 3.7 desta dissertação, foram implementados por meio de equalizadores paramétricos disponíveis no PD.

4.4 BLOCOS DO SISTEMA EM CONJUNTO

4.4.1 Construção de um *patch*

Cada aplicação tem o seu *patch* específico, que é construído de acordo com os seguintes requerimentos:

- número de fontes sonoras;
- movimento das fontes, ou seja, se alguma fonte necessita de movimentação automática;
- nível de imersão sonora, que influi na ordem Ambisonics e na configuração de alto-falantes.

A seguir, através de um exemplo, mostra-se como construir um *patch*.

Um compositor gostaria de espacializar os instrumentos de sua mais nova composição, um trio para violoncelo, clarinete e flauta. Ele decide que o flautista deverá executar um movimento circular dentro da sala de concerto, enquanto que os outros instrumentistas permanecem fixos em suas posições, sendo que o clarinetista fica à frente da audiência e o violoncelista atrás. A Figura 37 mostra o posicionamento dos instrumentos feito pelo compositor.

Tendo sintetizado as linhas de cada instrumento, ele já possui os arquivos de áudio que serão utilizados como entrada para o sistema. O compositor decide que não existe necessidade de grande imersão, portanto irá utilizar um sistema Ambisonics de primeira ordem, e que não utilizará a simulação acústica da sala, pois ele está interessado apenas na direcionalidade das fontes. Ele também dispõe de apenas quatro caixas, em uma montagem quadrada.

Figura 37 – Desenho do posicionamento imaginado pelo compositor para seu trio

Inicializado o PD, o compositor, partindo de suas premissas e necessidades, começa a construção do *patch* a partir da primeira camada do AUDIENCE, ou seja, a descrição da cena acústica (ver Figura 38).

Uma vez que a peça necessita de três fontes sonoras, ele utiliza três blocos `audce_gui_srcfile`, que irão controlar os arquivos de áudio de entrada. Também adiciona ao *patch* um bloco para o controle gráfico da posição das três fontes, `audce_gui_grid`, que configura para representar uma sala de vinte por vinte metros, e um bloco para controlar o movimento da flauta, `audce_aux_move`.

Adicionalmente, ele cria mensagens no PD que irão passar os nomes dos arquivos (como `open flauta.wav`), que irão controlar a reprodução da peça (`play`, `stop`), e que contém a lista com os movimentos que a flauta irá executar (`move_list flauta.txt`). O bloco `loadbang`, interno do PD, tem a função de enviar as mensagens conectadas a ele no momento em que o *patch* for aberto.

Figura 38 – Construção da camada 1 do *patch* exemplo

Em seguida, ele necessita adicionar os blocos relativos à segunda e terceira camada do AUDIENCE, nas quais o som será posicionado espacialmente (ver Figura 39). Não utilizando da simulação acústica, o único bloco necessário é o codificador Ambisonics, `audce_L3_amb_3rd~`. Cada fonte sonora deve ter um codificador que irá transformar o sinal monofônico em sinais B-Format, assim, o compositor liga cada saída das fontes em um bloco `audce_L3_amb_3rd~`. Ele também liga a saída do *grid*, que indica a posição das fontes e do usuário, aos seus respectivos codificadores.

Figura 39 – Construção da camada 2/3 do *patch* exemplo

Finalmente, a quarta camada do AUDIENCE, de decodificação e reprodução, é adicionada através do bloco `audce_L4_amb_3rd~`, em que é configurada para uma configuração quadrada de alto-falantes (ver Figura 40). Os sinais B-Format de cada fonte sonora devem ser mixados antes de entrarem no bloco de decodificação utilizando-se o bloco `audce_aux_mixer~`, sendo necessário um para cada canal Ambisonics. Enfim, as saídas do decodificador são ligadas ao bloco `dac~` do PD, que representa os alto-falantes conectados ao sistema.

Figura 40 – Construção da camada 4 do *patch* exemplo

Tendo finalizado a construção do *patch*, o compositor agora necessita simplesmente posicionar os instrumentos através da interface gráfica e enviar a mensagem `play` para iniciar a execução de seu trio especializado.

4.4.2 Teste dos blocos em conjunto

A Figura 41 apresenta um *patch* construído para testar a codificação e decodificação Ambisonics de segunda ordem, para reprodução em um arranjo quadrado de alto-falantes.

Dada uma entrada senoidal (no caso, 440 Hz), as formas de onda de cada canal Ambisonics (W, X, Y, Z e V, visto que os outros canais são nulos para este arranjo, pois são utilizadas para o eixo Z) e das saídas dos alto-falantes (frente-esq, frente-dir, trás-dir e trás-esq) são mostradas.

Figura 41 – Patch de segunda ordem em um arranjo quadrado, com gráficos mostrando os canais Ambisonics e a saída dos alto-falantes

4.5 SISTEMA DE HARDWARE PARA REPRODUÇÃO

O bloco de decodificação é ligado diretamente às saídas da interface de áudio. Um ambiente virtual sonoro é reproduzido por uma configuração como a da Figura 42, oito alto-falantes dispostos em uma forma octogonal em torno do ouvinte.

Figura 42 – Configuração octogonal horizontal de decodificação Ambisonics

O sistema também pode ser acoplado a uma CAVE e utilizado em conjunto de aplicações de realidade virtual. O diagrama de montagem da Figura 43 mostra o sistema físico instalado atualmente na CAVERNA Digital. Os alto-falantes são alimentados por dois amplificadores de potência de quatro canais cada, que, por sua vez, recebem o sinal de áudio de uma placa multicanal.

Figura 43 – Montagem física do sistema cúbico na CAVERNA Digital

Descrita a implementação do sistema de espacialização sonora, no próximo capítulo descreve-se o experimento de validação do sistema. Através da execução de uma peça musical espacializada, avalia-se a eficácia do sistema em reproduzir um campo sonoro.

5 EXPERIMENTO PARA VALIDAÇÃO DO SISTEMA

O experimento para validação do sistema consistiu na montagem de uma orquestra virtual para a execução, interativa e de baixa latência, de uma peça musical espacializada. Esse experimento incluiu a gravação das fontes sonoras a partir da partitura, a construção de um *patch* utilizando os blocos descritos anteriormente e baseado nos requerimentos da peça, e a reprodução da peça musical. A partir desta reprodução, testes de validação foram realizados para avaliar a qualidade da espacialização. O diagrama de blocos da Figura 44 ilustra os passos do experimento.

Figura 44 – Fluxo de montagem do experimento

O objetivo final do experimento foi determinar, a partir dos resultados aferidos, se o sistema está apto a ser utilizado para tarefas de espacialização orquestral, bem como para outros usos em tarefas de espacialização (ver Tabela 1).

Experimentos anteriores foram descritos pelo autor (THOMAZ et al., 2006a, 2006b). No primeiro, montou-se um cenário flexível que considera três instrumentos (baixo,

bateria e guitarra) posicionáveis no ambiente sonoro virtual através da interface gráfica do *patch*, manipulada pelo usuário por meio de *slides*. No segundo experimento, realizou-se a espacialização de quatro fontes sonoras, reproduzindo uma formação camerística composta de violino, violoncelo, trompete e flauta.

5.1 ESCOLHA DA PEÇA MUSICAL

Para o experimento de validação do sistema, elaborou-se uma lista de possíveis peças musicais espacializadas dentro das opções apresentadas na Tabela 4, que utilizassem os recursos do sistema como o posicionamento e movimentação das fontes sonoras. Também se acrescentaram peças que, mesmo não sendo espacializadas, poderiam ser utilizadas, e uma espacialização criada.

Embora as peças espacializadas apareçam em grande número na literatura, é difícil encontrar partituras disponíveis, uma vez que são trabalhos geralmente experimentais e de pequena distribuição. As peças mais conhecidas e, por consequência, de maior disponibilidade, são para grandes formações orquestrais e/ou coro, impossibilitando o uso dentro do escopo deste experimento. Além disso, a disponibilidade de músicos e o aparato necessário para gravação é um fator importante na escolha da peça. Para viabilizar a gravação, preferiu-se uma peça com uma pequena formação (três a cinco instrumentos ou grupos).

A Tabela 14 mostra as peças musicais consideradas para o experimento. A metodologia utilizada para a escolha foi dar nota para itens de avaliação e, ao final, selecionar a peça com maior nota, sendo que todos os itens têm o mesmo peso. Os itens de avaliação são:

- disponibilidade, que indica se as partituras e, se for o caso, sons eletroacústicos estão disponíveis;
- complexidade, que indica o esforço necessário para montar a peça;
- execução, que indica qual a facilidade de execução da peça por parte dos músicos;

- envolvimento, que indica o grau de imersão espacial em que o ouvinte está submetido durante a execução;
- movimento, que indica se existe uma movimentação indicada na partitura para as fontes sonoras;
- riqueza de timbres, que indica a variedade de timbres da peça.

Tabela 14 – Peças musicais especializadas para o experimento

<i>Peças Especializadas</i>	<i>Compositor</i>	<i>Espacialização</i>	<i>Disponibilidade</i>	<i>Complexidade</i>	<i>Fácil Execução</i>	<i>Envolvimento</i>	<i>Movimento</i>	<i>Riqueza timbres</i>	<i>Nota Final</i>
<i>Parcour's de L'Entité</i>	F. Menezes	Percussão ao fundo do palco, flautista em várias posições no palco e sons eletroacústicos ao lado da audiência. Flautista percorre 4 diferentes posições durante a execução da peça	10	3	3	5	5	5	5
<i>Poética II</i>	A. Escobar	4 Flautistas dispostos em losango em torno da audiência. Primeiro flautista entra no palco andando, enquanto executa o início da peça.	10	8	8	7	0	0	6
<i>Santos Football Music</i>	G. Mendes	Alto-falantes (1 para cada fita) são posicionados atrás do público, ao lado e atrás da orquestra.	10	4	4	7	0	5	5
<i>Gruppen</i>	K. Stockhausen	Orquestras situadas à frente e aos lados da audiência.	10	0	0	8	0	8	4
<i>The Unanswered Question</i>	C. Ives	Três grupos separados (quarteto de cordas, quarteto de madeiras e trompete). Quarteto de cordas preferencialmente fora do palco e longe dos outros.	10	10	10	5	0	7	7
<i>Noturno em Ré Maior K.286</i>	W. A. Mozart	Quatro orquestras separadas, tocando a mesma música com defasagem, causando um efeito de eco.	10	7	8	5	0	3	6
<i>Quinteto K.452</i>	W. A. Mozart	Sem espacialização.	10	7	6	5	0	7	6
<i>História de um Soldado</i>	I. Stravinsky	Sem espacialização .	10	4	4	8	8	8	7
<i>Omnes Gentes, Plaudite Manibus</i>	G. Gabrielli	4 Corais SATB posicionados em locais distintos.	10	8	8	7	0	5	6

Das duas peças de maior nota, a selecionada foi *The Unanswered Question*, de Charles Ives, por sua riqueza de timbres, com um quarteto de cordas, flautas, oboé,

clarinete e trompete, e pela divisão em três grupos dos instrumentos, facilitando o ensaio e a gravação da peça.

Essa peça foi composta em 1906, e apresentava elementos inéditos para a época, uma marca das composições de Charles Ives. A peça dura pouco mais de cinco minutos, e foi chamada por Ives de um “drama cósmico” (IVES, 1953, p. 2, tradução nossa). Em 1930, Ives revisou a peça e adicionou um texto programático (SWAFFORD, 1996).

Ela é constituída de três camadas, que pouco têm a ver uma com a outra, cada uma com seu estilo, textura, instrumentação e música. Um quarteto de cordas executa uma música coral, aparentemente tonal, que se repete indefinidamente, como uma espécie de moto contínuo. De acordo com Ives, ele representa o “Silêncio dos Druidas, que nada sabem, nada vêem e nada ouvem” (IVES, 1953, p. 2, tradução nossa). Sobre essa música misteriosa, um trompete periodicamente enuncia uma pergunta, “a questão perene da existência” (IVES, 1953, p. 2, tradução nossa). Após cada pergunta, o quarteto de madeiras tenta formular uma resposta, cada vez mais frenética, dissonante e desesperada. Por fim, a última pergunta é feita sem resposta, enquanto o coral das cordas continua indefinidamente.

Ives sugere, no texto que acompanha a partitura, que se dividam os instrumentos em três grupos: o quarteto de cordas, que deve ficar fora do palco ou longe do trompete e das madeiras; o trompete, que executa a pergunta musical; e o quarteto de madeiras, que tenta responder essa pergunta. Assim, têm-se três fontes sonoras distintas, que estão separadas espacialmente.

Baseado nas sugestões dadas por Ives, inúmeras são as possibilidades do arranjo espacial. Henry Brant, em uma execução da peça feita em 1955, posicionou as cordas atrás do palco, as madeiras em um balcão no meio do auditório e o trompete no fundo do auditório no mais alto balcão (SCHWARTZ, 1998).

Pode-se ir mais além das posições sugeridas por Ives, considerando cada instrumento uma fonte sonora, e não cada grupo. Assim, têm-se nove fontes sonoras que podem ser arranjadas espacialmente, criando novos efeitos sonoros, diferentes dos imaginados pelo compositor.

5.2 PRODUÇÃO DA PEÇA ESCOLHIDA

Tendo sido escolhida a peça e conhecida a sua formação instrumental, partiu-se para a seleção dos músicos no Departamento de Música da ECA-USP (CMU), devido à proximidade e disponibilidade destes. Selecionaram-se músicos alunos do departamento e músicos pertencentes à Orquestra de Câmara da USP (OCAM), todos com alguma experiência em música orquestral e acostumados ao processo de ensaio. Além desses, um aluno de regência do último ano do curso foi convidado a organizar e dirigir a parte musical do projeto de gravação.

A edição e revisão da partitura foram feitas baseadas em duas edições obtidas na biblioteca do departamento (IVES, [19--]) e (IVES, 1953). Ambas apresentavam pequenos erros e discordância entre elas, necessitando-se de uma ampla revisão e a confecção de uma nova edição, inclusive das partes separadas para cada músico. A partitura orquestral resultante desta revisão encontra-se no Apêndice C.

Após a distribuição das partituras, decidiu-se organizar um primeiro ensaio para testar a dificuldade de execução da peça por parte dos músicos. A Figura 45 mostra a realização deste ensaio, no auditório do CMU. Sem apresentar dificuldades, partiu-se diretamente para a gravação, sem a necessidade de outro ensaio.

Figura 45 – Ensaio da peça *The Unanswered Question*, no auditório do departamento de música da ECA-USP

Devido à própria natureza da peça composta por Ives, decidiu-se separar a gravação em três sessões, divididas em quarteto de madeiras, quarteto de cordas e trompete. Para estas sessões, reservou-se o estúdio de gravação do CMU pelo período de duas manhãs. A Figura 46 mostra uma visão de dentro da sala de controle do estúdio, com os músicos e o regente ao fundo.

Figura 46 – Gravação da peça *The Unanswered Question*, no estúdio do departamento de música da ECA-USP

Embora, para a execução da peça, pudessem ter sido gravados os instrumentos que compõem cada grupo de forma conjunta, preferiu-se capturar separadamente cada instrumento, para que fosse possível que cada um deles fosse uma fonte sonora no sistema, e não o grupo.

Figura 47 – Diagrama do estúdio LAMI (ECA-USP) utilizado na gravação com as baias para isolamento acústico

Devido à necessidade de que os instrumentistas de cada grupo executassem a peça de forma conjunta, por motivo de afinação e sincronia, estes foram isolados dentro de estúdio através de baias de material absorvente (fibra de vidro), deixando os músicos apenas com a visão do regente (ver Figura 47). A Figura 48 mostra uma das baias de isolamento durante a gravação do quarteto de cordas.

Figura 48 – Detalhe da baia de isolamento acústico durante a gravação

Para cada instrumento, utilizaram-se, na captação, dois microfones, um dinâmico e outro condensador, com o intuito de avaliar qual dos dois permitia um maior isolamento na captura da fonte sonora. O padrão de captura dos microfones foi configurado para cardióide, ou seja, com a maior direcionalidade possível. A Figura 49 mostra os padrões de captura de um microfone dinâmico e de um condensador.

Figura 49 – Padrão de captura e resposta em frequência dos microfones Sennheiser MD 421 (dinâmico) e AKG C414 (condensador)

O microfone dinâmico é menos sensível a sinais de baixa intensidade e apresenta uma resposta menos regular por faixa de frequência. Já o microfone condensador tem uma resposta em frequência mais plana, ou seja, reproduz mais fielmente o espectro do som original, e é mais sensível a sinais de baixa intensidade.

Devido à diferença de sensibilidade, o microfone dinâmico apresentou um maior isolamento acústico na captura do som proveniente do instrumento, enquanto que, na maioria dos casos, o microfone condensador apresentava de forma perceptível o som proveniente de outros instrumentos.

Assim, para a mixagem e masterização final das fontes utilizou-se ambos os sinais, levando em conta as vantagens de cada um: o bom isolamento do microfone dinâmico e a boa resposta em frequência do microfone condensador.

A captura do áudio foi toda feita via um sistema digital baseado em computador, utilizando o sistema de gravação Digidesign Pro Tools|HD³² e a edição, mixagem e masterização final das fontes foram feitas através do Adobe Audition³³.

A ficha técnica da gravação e o material utilizado no estúdio encontram-se no Apêndice D.

5.3 MONTAGEM DO *PATCH*

Para os testes experimentais foram utilizados codificadores / decodificadores Ambisonics de segunda ordem. Preferiu-se não utilizar o simulador acústico (audce_L2_allen) de forma a testar com mais ênfase a direcionalidade das fontes, ao invés da simulação acústica da sala, ou seja, o campo sonoro recriado foi o de um campo aberto.

Dois *patches* foram criados para a espacialização da peça *The Unanswered Question*: o primeiro contando apenas com três fontes sonoras, considerando os três grupos instrumentais da peça, ilustrado na Figura 50; o segundo contando com nove fontes sonoras, em que cada instrumento é espacializado de forma independente, ilustrado na Figura 51.

³² Disponível em: <<http://www.digidesign.com/index.cfm?navid=24&langid=1>>. Acesso em: 12 dez. 2006.

³³ Disponível em: <<http://www.adobe.com/products/audition/>>. Acesso em 12 dez. 2006.

Figura 50 – Patch para espacialização da peça *The Unanswered Question* com três fontes sonoras

Figura 51 – Patch para espacialização da peça *The Unanswered Question* com nove fontes sonoras

5.4 TESTES DE VALIDAÇÃO

Os testes de validação têm o objetivo de determinar o nível de imersão sonora do sistema proposto. Assim, mede-se a qualidade na criação do campo sonoro e, por consequência, da peça musical, e a utilidade do sistema para ser usado por músicos, regentes e produtores.

Foi selecionado um grupo de pessoas com experiência musical, como instrumentistas, compositores e regentes, e outro grupo sem experiência ou de músicos amadores. Cada pessoa foi encaminhada a uma cadeira localizada no centro do arranjo de alto-falantes e submetida individualmente à três execuções da peça musical pelo sistema, ora numa audição cega, ora numa audição interativa, onde podia manipular os parâmetros (posições) na interface. Posteriormente respondia a um questionário para avaliar os itens subjetivos da experiência. A Tabela 15 contém uma lista dos testes realizados.

Tabela 15 – Testes de validação do sistema de espacialização

#	Teste	Parâmetros avaliados	Descrição
1	Localização cega das fontes	Presença Direcionalidade Distância da Fonte Tamanho da Sala	Identificação das fontes sonoras apenas pela audição, sem acesso à interface gráfica do sistema de espacialização, de forma a avaliar a qualidade da recriação do campo sonoro. Executado duas vezes, com três e nove fontes sonoras com posições pré-determinadas.
2	Localização interativa das fontes	Facilidade de uso Aplicabilidade do sistema	Interatividade do usuário com o sistema, de forma a avaliar a aplicabilidade do sistema para músicos/regentes/produtores. Executado uma vez, no qual o usuário podia alterar a posição das fontes sonoras.

Para o teste 1, o usuário fez a audição da peça sem ter acesso à interface gráfica do sistema, utilizando os dois *patches* (de três e nove fontes sonoras). Durante a

audição, ele deveria identificar a posição e a distância das fontes sonoras, preenchendo os quadros presentes no questionário. Desta forma, esperava-se que o usuário não tivesse nenhum auxílio visual da posição das fontes que pudesse influenciar na sua decisão.

Figura 52 – Posições das fontes para o teste 1, sem acesso à interface gráfica pelo usuário

A Figura 52 ilustra as posições dos instrumentos que foram utilizadas no teste 1. No item (a) da figura, utilizando o *patch* com três fontes, os números representam: (1) quarteto de cordas; (2) quarteto de madeiras; (3) trompete. No item (b), utilizando o *patch* com nove fontes, os números representam: (1) violino I; (2) violino II; (3) viola; (4) violoncelo; (5) flauta I; (6) flauta II; (7) clarinete; (8) oboé; (9) trompete.

O teste 2 avalia a facilidade de uso do sistema por pessoas que geralmente não têm afinidade com ambientes computacionais, no caso os músicos e regentes. Assim, a configuração de um *patch* e a manipulação dos parâmetros deve ser acessível para este público. O ouvinte pôde nesse teste, durante a execução da peça, alterar a

posição dos instrumentos, colocando-os em lugares distintos dos originalmente planejados pelo compositor.

O usuário avaliou, durante os testes, a presença, a localização da fonte, a distância da fonte e o tamanho da sala, baseado nos atributos sugeridos por Berg (2003) e apresentados na Tabela 2. A avaliação destes atributos é importante para que se possa saber o nível de imersão sonora obtido pelo sistema e quais modificações devem ser feitas, caso exista deficiência em algum parâmetro. Além disso, também avaliou a facilidade de uso do sistema.

O questionário tem perguntas baseadas nos atributos que se pretende avaliar, e cada uma dessas perguntas tem uma explicação sucinta do parâmetro para melhor avaliação por parte do ouvinte. As respostas são baseadas em uma escala de um a cinco, indicando a qualidade do parâmetro sendo medido, que tem uma breve explicação para cada grau. De acordo com Berg (2003), este tipo de avaliação, em que escalas numéricas são associadas a âncoras verbais, são as mais comumente utilizadas em testes desta natureza com bons resultados. O questionário utilizado no experimento encontra-se no Apêndice E.

5.5 RESULTADOS

A versão atual do simulador acústico (*audce_L2_allen*), em testes separados, apresentou uma boa reprodução dos padrões de reflexão e reverberação da sala (ambiente virtual), possibilitando uma percepção da profundidade do ambiente e da distância da fonte sonora. Ele não apresentou, todavia, uma resposta estável no caso da localização e da distância da fonte.

Já os testes feitos sem o simulador acústico mostraram uma melhora no quesito localização e distância das fontes sonoras. Por outro lado, a ambiência foi perdida, tornando a experiência menos próxima da realidade.

Os resultados apresentados a seguir referem-se ao sistema sem a utilização do simulador acústico, com um sistema Ambisonics de segunda ordem e uma configuração quadrada utilizando quatro alto-falantes.

Nos testes foram avaliadas as respostas de quatro usuários do sistema, sendo três músicos e um não músico. A Tabela 16 resume a respostas dos questionários preenchidos pelos quatro usuários (a coluna de identificação das fontes indica quantas fontes foram corretamente apontadas pelo usuário nos dois casos).

Tabela 16 – Resultados dos testes de validação do sistema de espacialização

<i>Usuário</i>	<i>Experiência Musical</i>	<i>Identificação das Fontes (3 e 9 fontes)</i>	<i>Localização da Fonte</i>	<i>Distância da Fonte</i>	<i>Facilidade de Uso do Sistema</i>
1	sim	3 / 6	4	4	5
2	sim	3 / 6	4	3	4
3	sim	3 / 7	4	4	5
4	não	3 / 5	3	4	5

As notas variam de 1, para a pior avaliação, até 5, para a melhor.

Durante o experimento, constatou-se que os usuários tendem a fixar a visão no alto-falante, como um auxílio para perceber a localização do som. Desta forma, algumas das respostas apontaram a posição do alto-falante mais próximo da fonte sonora virtual. Para melhorar a experiência imersiva, os usuários foram instruídos a fechar os olhos durante a execução do teste 1.

Os resultados apontaram que a noção de proximidade da fonte é muito boa, todavia, quanto mais longe se posiciona a fonte, menos precisa fica a percepção da distância, chegando em um ponto onde já não se faz mais diferença auditiva ao se afastar ou não a fonte. Na localização da fonte, foi apontado que existe uma mudança brusca quando se move uma fonte da esquerda para a direita, ou de frente para trás, passando pela posição do usuário.

Devido à peça não ter uma textura polifônica, ou seja, os instrumentos não apresentam uma linha melódica facilmente destacável dentre eles, a localização individual dos nove instrumentos foi prejudicada durante o teste 1. Mostrou-se difícil identificar a posição das cordas, devido ao estilo coral de sua execução. Da mesma forma, as intervenções relativamente rápidas do quarteto de madeiras (em torno de cinco segundos), dificultou a localização dos instrumentos.

Somente quando o usuário pôde efetivamente mexer na interface gráfica e alterar a posição do instrumento escolhido por ele, é que a identificação da posição foi facilitada (teste 2).

As respostas dos questionários, baseadas nos dois testes, obtiveram para os quesitos de localização e distância da fonte uma média de quatro pontos em cinco. Já para a facilidade de uso do sistema a pontuação obtida foi máxima. Os quesitos presença e tamanho da sala não foram avaliados devido à opção de não utilizar o simulador acústico.

6 CONCLUSÕES

Os objetivos propostos de pesquisar técnicas de espacialização, desenvolver um sistema e executar experimentos de espacialização musical foram atingidos.

O sistema comportou-se de forma estável e sem problemas de desempenho, ou seja, sem atrasos perceptíveis na reprodução do som, mesmo quando auralizava nove fontes sonoras simultaneamente. Todavia, espera-se que com a utilização do simulador acústico e a convolução das respostas impulsivas, a performance deva cair e a utilização de computadores mais poderosos ou aglomerados de computadores (*cluster*) seja necessária.

Os resultados com testes individuais do bloco de simulação acústica indicaram a necessidade de se implementar melhoramentos, especialmente na calibração de escalas e valores das grandezas acústicas como intensidade/pressão nos vetores Ambisonics.

A falta de familiarização com os instrumentos e suas diferenças de timbre de um avaliador prejudicou a localização dos instrumentos, principalmente nas cordas, nas duas flautas e no oboé com o clarinete. Portanto, para a localização da fonte sem o acesso à interface gráfica (teste 1), fez-se necessária a experiência em distinguir timbres e a habilidade de ouvir instrumentos similares executando linhas diferentes.

Os resultados obtidos durante o teste 2, no qual o usuário teve acesso ao sistema, mostraram que a direcionalidade e distância das fontes, baseadas na codificação Ambisonics de segunda ordem, mostrou-se fiel a um campo sonoro real.

Através dos experimentos, principalmente o teste 2, conclui-se que o sistema é promissor para a utilização em aplicações musicais e que pode ser uma ferramenta interessante e de baixo custo para que compositores, regentes e produtores musicais possam fazer experimentos com espacialização. A facilidade de se montar um *patch* e do uso do sistema corrobora com essa conclusão.

Enfim, baseado nos resultados obtidos com os experimentos, classificou-se o sistema de espacialização sonora com um nível de imersão sonora quatro, de

acordo com a escala proposta por Faria, M. Zuffo e J. Zuffo (2005), ou seja, ele é capaz de reproduzir um campo sonoro estável bidimensional.

O sistema está preparado para chegar ao nível cinco, uma vez que a implementação dos blocos permite a adição da verticalidade do som, necessitando apenas de experimentos para comprovar tal nível de imersão.

Os resultados obtidos nos experimentos fornecem pistas para melhoramentos adicionais de forma a aumentar o grau de imersão sonora dos sistemas posteriores, baseados no AUDIENCE.

6.1 CONTRIBUIÇÕES

O trabalho traz contribuições para as áreas de engenharia de áudio e de computação musical, principalmente na área de processamento digital de sinais de áudio, explorando formas de implementar seus algoritmos com alto desempenho.

Outras contribuições estão listadas a seguir:

1. Literatura sobre o Ambisonics. Um grande trabalho foi feito para levantar os artigos e monografias disponíveis sobre o Ambisonics desde a década de 70. O trabalho preenche uma lacuna na literatura disponível em português sobre o Ambisonics, além de fornecer um ponto de partida para um maior estudo sobre a técnica;
2. Primeira implementação completa da arquitetura AUDIENCE. O uso de uma arquitetura genérica de produção de áudio multicanal combinada com a abordagem de programação gráfica do PD, utilizando blocos de funções específicas existentes ou programando-os quando inexistentes, permite a criação de um sistema para espacialização sonora que atende todos os blocos apresentados no item 2.3;
3. Criação de uma ferramenta para espacialização de fontes sonoras, utilizando um sistema gráfico para montagem dos *patches* (PD) de fácil

uso através de blocos previamente construídos, permitindo ao usuário (compositor/regente/produtor/outros) focar-se na aplicação do sistema;

4. Gravação de fontes sonoras em estúdio de forma separada. A natureza do sistema, cujas entradas são fontes sonoras monofônicas, exigia que as gravações de cada instrumento fossem feitas de forma separada. Visto que nem sempre é possível gravar os músicos de forma separada, principalmente na música orquestral e de câmara, utilizou-se um método para que, mesmo com os músicos tocando juntos, cada fonte sonora fosse captada individualmente;
5. Experimentos com a espacialização musical possibilitando a execução e releitura de peças espacializadas;
6. Novas direções para a produção e reprodução de áudio espacial. Através da implementação do AUDIENCE, a área de áudio espacial tem suas fronteiras ampliadas, possibilitando novos estudos a partir deste trabalho;
7. Utilização de uma metodologia para avaliar a qualidade dos atributos espaciais. Desta forma, é possível melhorar a recriação da cena acústica baseado na análise subjetiva de parâmetros espaciais, como a presença, a posição da fonte, entre outros.

6.2 TRABALHOS FUTUROS

A partir das conclusões sobre o trabalho elaborado, alguns trabalhos futuros e possíveis extensões do sistema são descritos a seguir:

1. Realizar uma nova bateria de testes, com um maior número de usuários e baseado em alguma das diversas metodologias de forma a obter

resultados mais sólidos sobre a fidelidade na recriação do campo sonoro, a facilidade de uso e a aplicabilidade do sistema.

2. Desenvolvimento de novos blocos para o sistema de espacialização sonora, como blocos responsáveis por mapear as cenas acústicas (e.g. *parsers* X3D), novas interfaces com o usuário (e.g. *joystick*, *head-tracking*), novos simuladores acústicos (e.g. traçado de raios para poliedros arbitrários (BORISH, 1984), elementos finitos) e novos codificadores/decodificadores (e.g. WFS, HRTF, Dolby/DTS). Este trabalho está dentro do escopo do projeto OpenAUDIENCE³⁴, que vem sendo conduzido como forma de estender a plataforma AUDIENCE por meio da participação da comunidade de pesquisa de áudio *surround*;
3. Experimentos com configurações tridimensionais de alto-falantes, já implementado no bloco de decodificação para arranjos cúbicos e dodecaédricos, para testar a capacidade de espacializar verticalmente o som;
4. Codificadores/decodificadores Ambisonics de altas ordens (HOA), da quarta em diante, a partir de novos trabalhos na área e na crescente capacidade de processamento dos computadores;
5. Decodificação Ambisonics para uma configuração qualquer de alto-falantes, não restrito a configurações regulares (LEE; SUNG, 2003). Desta forma, torna-se mais fácil a adaptação de um sistema Ambisonics em locais onde o posicionamento dos alto-falantes pode ser restrito (e.g. CAVEs, onde o feixe de luz dos projetores não pode ser interrompido).
6. Sistema de calibração automático para os alto-falantes, de forma a normalizar a intensidade de cada um e ajustar o posicionamento físico deste da forma como foi informada ao sistema;
7. Filtros inversos para ajuste dos alto-falantes, de forma a uniformizar um conjunto de alto-falantes de diferentes modelos ou minimizar o efeito

³⁴ Disponível em: <<http://openaudience.incubadora.fapesp.br/portal>>. Acesso em: 18 dez. 2006

difusor de objetos inseridos no local da reconstrução acústica (e.g. telas de projeção em CAVEs).

8. Sistema de espacialização sonora *standalone*, sem a necessidade de rodar sobre o PD. Assim, torna mais fácil a integração do sistema de espacialização a uma aplicação de maior escopo, como um sistema completo de realidade virtual imersiva (CAVEs);

REFERÊNCIAS

ABHAYAPALA, T.; WARD, D. Theory and design of high order sound field microphones using spherical microphone array. In: ICASSP, IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, 2002, Orlando. **Proceedings...** Orlando: IEEE, 2002.

ALLEN, J. B.; BERKLEY, D. A. Image method for efficiently simulating small-room acoustics. **Journal of the Acoustical Society of America**, v.65, n.4, p. 943-950, abr. 1979.

APEL, W. **Gregorian Chant**. Indiana University Press, 1990. 540 p.

BAMFORD, J. S. **An Analysis of Ambisonic Sound System of First and Second Order**. 1995. 270 p. Dissertação (Mestrado) – University of Waterloo. Waterloo, Canada.

BAMFORD, J. S.; VANDERKOOY, J. Ambisonics Sound for Us. In: AES CONVENTION, 99, 1995, New York. **Proceedings...** New York: AES, 1995.

BERKHOUT, A. J. A Holographic Approach to Acoustic Control. **Journal of the Audio Engineering Society**, v. 36, n. 12, p. 977-995, 1988.

BERG, J., RUMSEY, F. Systematic evaluation of perceived spatial quality. In: AES INTERNATIONAL CONFERENCE ON MULTICHANNEL AUDIO, 24., 2003, Banff. **Proceedings...** Banff: AES, 2003. p.184-198.

BLAUERT, J. **Spatial Hearing – The Psychophysics of Human Sound Localization**. MIT Press, 1996.

BOOCH, G.; RUMBAUGH, J.; JACOBSON, I. **The Unified Modeling Language User Guide**. Addison-Wesley, 1999.

BORISH, J. Extension of the image model to arbitrary polyhedra. **Journal of the Acoustical Society of America**, v.75, n.6, p. 1827-1836, abr. 1984.

BOULEZ, P.; GERZSO, A. Computers in Music. **Scientific American**, v. 258, n. 4, abr. 1988.

CHILDS, E.; PULKKI, V. Using multi-channel spatialization in sonification: a case study with meteorological data. In: INTERNATIONAL CONFERENCE ON AUDITORY DISPLAY, 2003, Boston. **Proceedings...** Boston: ICAD, 2003.

CUNNINGHAM, M. 'Welcome to the Machine' The story of Pink Floyd's live sound. **Sound on Stage**, n. 5, mar. 1997.

DANIEL, J. **Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia**. 2001. 319 p. Tese (Doutorado) – l'Université Paris. Paris.

DANIEL, J.; MOREAU, S. Further Study of Sound Field Coding with Higher Order Ambisonics. In: AES CONVENTION, 116., 2004, Berlin. **Proceedings...** Berlin: AES, 2004.

DANIEL, J.; NICOL, R.; MOREAU, S. Further Investigations of High Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging. In: AUDIO ENGINEERING SOCIETY CONVENTION, 114., 2003, Amsterdam. **Proceedings...** Amsterdam: AES, 2003.

DAVIS, S. et al. High Order Spatial Audio Capture and Its Binaural Head-Tracked Playback Over Headphones with HRTF Cues. In: AUDIO ENGINEERING SOCIETY CONVENTION, 119., 2005, New York. **Proceedings...** New York: AES, 2005.

DOBRUCKI, A.; PLASKOTA, P. Numerical Head Model to HRTF Simulation. In: AUDIO ENGINEERING SOCIETY CONVENTION, 118., 2005, Barcelona. **Proceedings...** Barcelona: AES, 2005.

DOLBY LABORATORIES. **Surround Sound** - Past, Present, and Future. 1999. Disponível em: <www.dolby.com/assets/pdf/tech_library/2_surround_Past.Present.pdf>. Acesso em: 15 maio 2005.

DRESSLER, R. **Dolby Surround Pro Logic Decoder Principles of Operation**. Dolby Laboratories Licensing Corporation, 1993.

ELEN, R. Whatever Happened to Ambisonics. **AudioMedia Magazine**, nov. 1991.

FARIA, R. R. A. et al. AUDIENCE - Audio Immersion Experiences in the CAVERNA Digital. In: BRAZILIAN SYMPOSIUM ON COMPUTER MUSIC, 10., 2005, Belo Horizonte. **Proceedings...** Belo Horizonte: SBCM, 2005. p. 106-117.

FARIA, R. R. A. **Auralização em ambientes audiovisuais imersivos**. 2005. 191 p. Tese (Doutorado) – Escola Politécnica da Universidade de São Paulo. São Paulo.

FARIA, R. R. A.; ZUFFO, J. A. An Auralization Engine Adapting a 3-D Image Source Acoustic Model to Ambisonics Coder for Immersive Virtual Reality. In: AES INTERNATIONAL CONFERENCE, 28., 2006, Piteå. **Proceedings...** Piteå: AES, 2006.

FARIA, R. R. A.; ZUFFO, M. K.; ZUFFO, J. A.. Improving spatial perception through sound field simulation in VR. In: IEEE CONFERENCE ON VIRTUAL ENVIRONMENTS, HUMAN-COMPUTER INTERFACES AND MEASUREMENT SYSTEMS, 2005, Giardini Naxos. **Proceedings...** Giardini Naxos: IEEE, 2005. p. 103-108.

FARINA, A. Software Implementation of B-Format Encoding and Decoding. In: AUDIO ENGINEERING SOCIETY CONVENTION, 104., 1998, Amsterdam. **Proceedings...** Amsterdam: AES, 1998.

FRAUENBERGER C., RITSCH, W., HÖLDRICH, R. Internet Archive For Electronic Music IAEM-IARS (Internet Audio Rendering System). In: AES INTERNATIONAL CONFERENCE ON MULTICHANNEL AUDIO, 24., 2003, Banff. **Proceedings...** Banff: AES, 2003.

FRIBERG, A. pDM: An Expressive Sequencer with Real-Time Control of the KTH Music-Performance Rules. **Computer Music Journal**, v. 30, n. 1, p. 37-48, primavera 2006.

FURSE, R. First and Second Order Ambisonic Decoding Equations. Disponível em: <<http://www.muse.demon.co.uk/ref/speakers.html>>. Acesso em: 11 mar. 2006.

GERZON, M. Ambisonics Decoders for HDTV. In: AUDIO ENGINEERING SOCIETY CONVENTION, 92., 1992, Vienna. **Proceedings...** Vienna: AES, 1992.

_____. Ambisonics in Multichannel Broadcasting and Video. **Journal of the Audio Engineering Society**, v. 33, n. 11, p. 859-871, nov. 1985.

GERZON, M. Periphony: With-Height Sound Reproduction. **Journal of the Audio Engineering Society**, v. 21, n. 1, p. 2-10, jan.-fev. 1973.

_____. Practical Periphony: The Reproduction of Full-Sphere Sound. In: AUDIO ENGINEERING SOCIETY CONVENTION, 65., 1980, London. **Proceedings...** London: AES, 1980.

_____. What's wrong with quadraphonics. **Studio Sound**, maio 1974a.

_____. Surround-sound psychoacoustics. **Wireless World**, p. 483-485, dez. 1974b.

GIOMI, F.; MEACCI, D.; SCHWOON, K. *Live Electronics in Luciano Berio's Music*. **Computer Music Journal**, v. 27, n. 2, p. 30-46, verão 2003.

GLASGAL, R. The Ambiophone Derivation Of A Recording Methodology Optimized For Ambiophonic Reproduction. In: AES INTERNATIONAL CONFERENCE ON SURROUND SOUND TECHNIQUES, 19., 2001, Schloss Elmau. **Proceedings...** Schloss Elmau: AES, 2001.

_____. Surround ambiophonic recording and reproduction. In: AES INTERNATIONAL CONFERENCE ON MULTICHANNEL AUDIO, 24., 2003, Banff. **Proceedings...** Banff: AES, 2003.

GRIFFITHS, P. **Modern Music**. World of Art, 1994.

ITU-R. **Recommendation BS. 775.1, Multi-channel stereophonic sound system with or without accompanying picture**. International Telecommunications Union, 1993.

IVES, C. E. **La Pregunta Incontestada**. [S.l: s.n.], [19--]. 1 partitura (7 p.). Quarteto de flautas, trompete e quarteto de cordas.

_____. **The Unanswered Question**. New York: Southern Music Publishing Co. Inc., 1953. 1 partitura (8 p.). Quarteto de flautas, trompete e quarteto de cordas.

KIM, J. et al. New HRTFs (Head Related Transfer Functions) for 3D Audio Applications. In: AUDIO ENGINEERING SOCIETY CONVENTION, 118., 2005, Barcelona. **Proceedings...** Barcelona: AES, 2005.

KIRKEBY, O.; NELSON, P. A.; HAMADA, H. The "Stereo Dipole": Binaural Sound Reproduction using Two Closely Spaced Loudspeakers. In: AUDIO ENGINEERING SOCIETY CONVENTION, 102., 1997, Munich. **Proceedings...** Munich: AES, 1997.

LEE, S.; SUNG, K. Generalized Encoding and Decoding Functions for a Cylindrical Ambisonic Sound System. **IEEE Signal Processing Letters**, v. 10, n. 1, p. 21-23, jan. 2003.

LIUNI, M.; MORELLI, D. *Playing Music*: an installation based on Xenakis' musical games. In: WORKING CONFERENCE ON ADVANCED VISUAL INTERFACES, 2006, Venezia. **Proceedings...** Venezia: ACM, 2006.

MALHAM, D. Towards Reality Equivalence in Spatial Sound Diffusion. **Computer Music Journal**, v. 25, n. 4, p. 31-38, inverno 2001.

MALHAM, D.; MYATT, A. 3-D Sound Spatialization using Ambisonic Techniques. **Computer Music Journal**, v. 19, n. 4, p. 58-70, inverno 1995.

MENEZES, F. **Atualidade Estética da Música Eletroacústica**. São Paulo: Editora Unesp, 1999.

MEYER, J; ELKO, G. A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield. In: ICASSP, IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, 2002, Orlando. **Proceedings...** Orlando: IEEE, 2002.

MOORE, B. **An Introduction to the Psychology of Hearing**. 5 ed., Academic Press, 2003.

NOISTERNIG, M. et al. A 3D Ambisonic Based Bi-aural Sound Reproduction System. In: AES INTERNATIONAL CONFERENCE ON MULTICHANNEL AUDIO, 24., 2003, Banff. **Proceedings...** Banff: AES, 2003.

OGI, T. et al. Immersive sound field simulation in multi-screen projection displays. In: INTERNATIONAL IMMERSIVE PROJECTION TECHNOLOGIES WORKSHOP, 7., EUROGRAPHICS WORKSHOP ON VIRTUAL ENVIRONMENTS, 9., 2003, Zurich. **Proceedings...** Zurich: Eurographics, 2003. p.135-142.

OPPENHEIM, A.; SCHAFER, R. **Discrete-Time Signal Processing**. 2 ed. Prentice Hall, 1998.

PARSONS, A. Four Sides of the Moon. **Studio Sound**, jun. 1975.

POLETTI, M. A Unified Theory of Horizontal Holographic Sound Systems. **Journal of the Audio Engineering Society**, v. 48, n. 12, p. 1115-1182, dez. 2000.

PRESS, W. et al. **Numerical Recipes in C: the Art of Scientific Computing**. 2 ed. Cambridge University Press, 1992.

PUCKETTE, M. Pd Documentation. Disponível em: <http://crca.ucsd.edu/~msp/Pd_documentation/>. Acesso em: 11 mar 2006.

PULKKI, V.; LOKKI, T. Creating auditory displays to multiple loudspeakers using VBAP: A case study with DIVA project. In: INTERNATIONAL CONFERENCE ON AUDITORY DISPLAY, 1998, Glasgow. **Proceedings...** Glasgow: ICAD, 1998.

PULKKI, V. **Spatial Sound Generation and Perception by Amplitude Panning Techniques**. Helsinki: Helsinki University of Technology, 2001. 59 p.

RAMÍREZ, M.A.; RODRÍGUEZ SORIA, S. G. HRTF Individualization by Solving the Least Squares Problem. In: AUDIO ENGINEERING SOCIETY CONVENTION, 118., 2005, Barcelona. **Proceedings...** Barcelona: AES, 2005.

RANDALL, R. **An Introduction to Acoustics**. Dover, 2005.

RODRÍGUEZ SORIA, S. G. **Estudos sobre personalização da função de transferência relativa à cabeça em sistemas binaurais de reprodução acústica virtual**. 2006. 155 p. Dissertação (Mestrado) – Escola Politécnica da Universidade de São Paulo. São Paulo.

RUMSEY, F. **Spatial Audio**. Oxford: Focal Press, 2001.

SCHMITZ, M. SAFIR – Spatial Audio Framework for Instrumented Rooms. In: WORKSHOP ON INVISIBLE AND TRANSPARENT INTERFACES, 2004, Gallipoli. **Proceedings...** Gallipoli: ITI, 2004.

SCHWARTZ, E. S. et al. **Contemporary Composers on Contemporary Music**. Da Capo Press, 1998.

SHERIDAN, J. et al. SOUNDSTUDIO4D - A VR Interface for Gestural Composition of Spatial Soundscapes. In: INTERNATIONAL CONFERENCE OF AUDITORY DISPLAY, 10., 2004, Sydney. **Proceedings.....** Sydney: ICAD, 2004.

SWAFFORD, J. **Charles Ives** – A Life with Music. W.W. Norton, 1996.

THOMAZ, L. F. et al. Experimentações de espacialização orquestral sobre a arquitetura AUDIENCE. In: CONGRESSO DA AES BRASIL, 4., 2006, São Paulo. **Anais...** São Paulo: AES, 2006a.

_____. Orchestration Spacialization Using The AUDIENCE Engine. In: INTERNATIONAL COMPUTER MUSIC CONFERENCE, 2006, New Orleans. **Proceedings.....** New Orleans: ICMC, 2006b. p. 266-269.

_____. Simulação de Som 3D em um Ambiente de Realidade Virtual Imersiva utilizando HRTF. In: CONGRESSO BRASILEIRO DE ENGENHARIA DE ÁUDIO DA AES BRASIL, 3., 2005, São Paulo. **Anais...** São Paulo: AES, 2005. p. 44-48.

TODD, C. et al. AC-3: Flexible Perceptual Coding for Audio Transmission and Storage. In: CONVENTION OF THE AUDIO ENGINEERING SOCIETY, 96., 1994.

TROCHIMCZYK, M. From Circles to Nets: On the Signification of Spatial Sound Imagery in New Music. **Computer Music Journal**, v. 25, n. 4, p. 39-56, inverno 2001.

ZUFFO, J. A. et al. CAVERNA Digital – Sistema de Multiprojeção Estereoscópico Baseado em Aglomerados de PCs para Aplicações Imersivas em Realidade Virtual. In: SYMPOSIUM OF VIRTUAL REALITY, 4., Florianópolis, 2001. **Proceedings...** Florianópolis: SBC, 2001.

ZVONAR, R. A History of Spatial Music. Disponível em: <www.zvonar.com/writing/spatial_music/History.html>. Acesso em: 15 maio 2005.

APÊNDICE A – Fundamentos físicos do Ambisonics

O sistema Ambisonics reproduz um campo sonoro através da reconstrução da onda sonora original. Assim como a técnica WFS, ele se baseia no princípio de Huygens, que diz que cada ponto de uma frente de onda serve como uma pequena fonte esférica, de tal forma que a onda original pode ser descrita como uma soma destas fontes (DANIEL, 2001).

A seguir, demonstra-se como essa reconstrução pode ser feita a partir de um caso em que existe uma fonte sonora real, e como a contribuição de uma matriz de fontes pode recriá-la (BAMFORD, 1995). De forma a tornar os cálculos menos complexos e mais ilustrativos, o equacionamento é feito utilizando apenas as componentes do plano horizontal e considerando as ondas planas, com o ouvinte longe da fonte sonora.

Figura 1 – Campo sonoro gerado por uma fonte na posição $\vec{k}$ e recebido por um ouvinte na posição $\vec{r}$

A Figura 1 mostra uma frente de onda gerada por uma fonte sonora (violino) com a posição dada pelo vetor $\vec{k}$ que forma um ângulo ψ com o eixo x, chegando a um ouvinte na posição dada pelo vetor $\vec{r}$, formando um ângulo ϕ com o eixo x.

Considerando esta onda que parte da fonte, com uma pressão sonora ψ como plana, ela é descrita pela equação (1).

$$S_{\psi} = P_{\psi} e^{i\vec{k} \cdot \vec{r}} = P_{\psi} e^{ikr \cos(\phi - \psi)} \quad (1)$$

A equação de uma onda plana pode ser relacionada a uma série de co-senos e funções cilíndricas de Bessel (J_m) (RANDALL, 2005), chegando-se à equação (2), com sua forma expandida na equação (3), sendo m a ordem Ambisonics.

$$S_{\psi} = P_{\psi} J_0(kr) + 2P_{\psi} \sum_{m=1}^{\infty} i^m J_m(kr) \cos m(\phi - \psi) \quad (2)$$

$$S_{\psi} = P_{\psi} J_0(kr) + P_{\psi} \sum_{m=1}^{\infty} 2i^m J_m(kr) [\cos(m\phi) \cos(m\psi) + \sin(m\phi) \sin(m\psi)] \quad (3)$$

A reconstrução da onda representada pela equação (3) é o objetivo de um sistema Ambisonics. Com um número infinito de fontes seria possível reconstruí-la exatamente, mas essa solução não é factível, pois se tem disponível apenas um número fixo de geradores de onda, no caso um conjunto de alto-falantes.

Figura 2 – Configuração regular octogonal de alto-falantes utilizados para reconstruir a frente de onda em um sistema Ambisonics

Considerando um número N de alto-falantes, dispostos em uma configuração fixa e regular (um exemplo com oito alto-falantes é mostrado na Figura 2), obtém-se a equação de onda proveniente de um alto-falante n que chega ao ouvinte conforme a equação (4) posicionado com um ângulo ϕ_n .

$$S_n = P_n e^{i\vec{k} \cdot \vec{r}} = P_n e^{ikr \cos(\phi - \phi_n)} \quad (4)$$

Somando a contribuição de todos os alto-falantes, chega-se à equação (5), e expandida em (6).

$$S = \sum_{n=1}^N P_n J_0(kr) + \sum_{n=1}^N P_n \sum_{m=1}^{\infty} 2i^m J_m(kr) \cos m(\phi - \phi_n) \quad (5)$$

$$S = \sum_{n=1}^N P_n J_0(kr) + \sum_{n=1}^N P_n \sum_{m=1}^{\infty} 2i^m J_m(kr) [\cos(m\phi) \cos(m\phi_n) + \sin(m\phi) \sin(m\phi_n)] \quad (6)$$

Para reconstruir a onda original, basta igualar a onda S_ψ com a onda gerada pelos alto-falantes S (equação (7)), obtendo assim a equação (8).

$$S_\psi = S \quad (7)$$

$$\begin{aligned} & P_\psi J_0(kr) + P_\psi \sum_{m=1}^{\infty} 2i^m J_m(kr) [\cos(m\phi) \cos(m\psi) + \sin(m\phi) \sin(m\psi)] \\ &= \sum_{n=1}^N P_n J_0(kr) + \sum_{n=1}^N P_n \sum_{m=1}^{\infty} 2i^m J_m(kr) [\cos(m\phi) \cos(m\phi_n) + \sin(m\phi) \sin(m\phi_n)] \end{aligned} \quad (8)$$

Igualando os termos em pares para cada valor de m , chega-se as equações (9), (10) e (11).

$$P_\psi = \sum_{n=1}^N P_n \quad (9)$$

$$P_\psi \cos(m\psi) = \sum_{n=1}^N P_n \cos(m\phi_n) \quad (10)$$

$$P_\psi \sin(m\psi) = \sum_{n=1}^N P_n \sin(m\phi_n) \quad (11)$$

Aqui aparecem duas limitações para a reconstrução exata da frente de onda. Visto que a onda é representada por uma série infinita, em aplicações práticas deve-se fixar um valor máximo para m nas equações e, assim, perder componentes da onda original. Estudos de como a reconstrução se comporta para diferentes valores de m foram realizados por Bamford (1995) e Daniel (2001). Outra limitação é o número de alto-falantes, representado pela letra n .

As equações (9), (10) e (11) definem os sinais Ambisonics, que devem ser obtidos da fonte original, e que serão utilizados na reconstrução. Ou seja, para armazenar um sistema de primeira ordem ($m=1$), necessita-se de três canais (correspondentes a P_ψ , $P_\psi \cos(\psi)$, $P_\psi \sin(\psi)$), para um sistema de segunda ordem ($m=2$), mais dois ($P_\psi \cos(2\psi)$, $P_\psi \sin(2\psi)$) e assim por diante. O número de canais no caso bidimensional é determinado por $(2m+1)$.

Outra forma de decompor a frente de onda original, utilizando coordenadas esféricas, é através das séries de Fourier-Bessel (DANIEL, 2003).

$$P(\vec{r}) = \sum_{m=0}^{\infty} J_m^m(kr) \sum_{m=1}^{\infty} B_{mn}^{\sigma} Y_{mn}^{\sigma}(\theta, \delta) \quad (12)$$

As esferas harmônicas são a parte angular de um conjunto de soluções ortogonais da equação de Laplace, representada em um sistema de coordenadas esféricas (DANIEL, 2001). No contexto da representação do campo sonoro, elas indicam a dependência direcional de uma fonte sonora (BAMFORD, 1995), ilustradas na Figura 3 através de um modelo tridimensional até a terceira ordem.

Figura 3 – Esferas harmônicas até a terceira ordem³⁵

³⁵ Figura extraída de (DANIEL, 2003).

Enquanto a componente Y_{00}^1 representa a intensidade da pressão sonora, as outras componentes representam o gradiente ou velocidade da partícula. Assim tem-se uma indicação do movimento da onda, e, quanto maior a ordem, maior a acuidade na determinação da direção.

No momento da reprodução, para se alimentar os alto-falantes em uma configuração regular, compõe-se o sinal através dos sinais obtidos da onda original (BAMFORD; VANDERKOOY, 1995). Para um alto-falante qualquer n , utiliza-se a equação (13).

$$P_n = \frac{1}{N} (P_\psi + 2 \sum_{m=1}^{\infty} [P_\psi \cos(\phi) \cos(m\phi_n) + P_\psi \sin(m\phi) \sin(m\phi_n)]) \quad (13)$$

Experimentos realizados por Bamford (1995) mostram que a reconstrução do campo sonoro perde estabilidade com o aumento da frequência devido a diferenças de fase geradas pelos alto-falantes. O aumento da ordem do sistema aprimora o casamento de fases, diminuindo o problema e aumentando o tamanho da área da reconstrução da onda (*sweet spot*).

APÊNDICE B – Diagramas UML do sistema

Figura 1 – Diagrama de seqüência para iniciar a execução da peça musical

Figura 2 – Diagrama de seqüência para movimentar uma fonte sonora

APÊNDICE C – Partitura da peça *The Unanswered Question*

The Unanswered Question

Charles E. Ives
(versão revisada, 1934)

Flauta I

Flauta II

Oboé

Clarinete em Sib

Largo molto sempre (para cordas e trompete) (aprox. 50 = ♩)

Trompete em Dó

Violino I *ppp con sordini*

Violino II *ppp con sordini*

Viola *ppp con sordini*

Violoncello *ppp con sordini*

Adagio

The musical score is arranged in two systems. The first system includes parts for Flute I, Flute II, Oboe, and Clarinet in Sib. The second system includes parts for Trumpet in Dó, Violin I, Violin II, Viola, and Violoncello. The tempo is marked 'Adagio'. The music features a variety of notes, including triplets, and dynamics such as 'p' (piano). Slurs are used to group notes across measures. The key signature has one sharp (F#) and the time signature is 3/4.

Fl. I.

Fl. II.

Ob.

Cl. (Sib)

Tpt. (Dó)

VI. I.

VI. II.

Vla.

Vcl.

The first system of the musical score includes parts for Flute I, Flute II, Oboe, and Clarinet in Sib. Each instrument has a melodic line with triplets and a dynamic marking of *p* (piano). The Flute I part includes a triplet of eighth notes. The Flute II part includes a triplet of eighth notes. The Oboe part includes a triplet of eighth notes. The Clarinet in Sib part includes a triplet of eighth notes.

Allegretto

Fl. I. *mf*

Fl. II. *mf*

Ob. *mf*

Cl. (Sib) *mf*

Tpt. (D $\flat$)

Vi. I.

Vi. II.

Vla.

Vcl.

Allegro

Fl. I. *f* *ff*

Fl. II. *f* *ff*

Ob. *f* *ff*

Cl. (Sib) *f* *ff*

Tpt. (D6) *p*

Vl. I. *f*

Vl. II. *f*

Vla. *f*

Vcl. *f*

Allegro molto

This musical score page contains measures 141 through 144. The tempo is marked 'Allegro molto'. The instrumentation includes Flute I, Flute II, Oboe, Clarinet in B-flat, Trumpet in D-flat, Violin I, Violin II, Viola, and Violoncello. The percussion part (Tpt. (Dó)) is shown with a series of rests. The woodwinds and strings play complex patterns, including triplets and sixteenth notes. Dynamics range from *pp* (pianissimo) to *ff* (fortissimo). The score is written in a key with one flat (B-flat) and a common time signature (C).

Fl. I. *f* *sf* *pp*

Fl. II. *ff* *sf* *pp*

Ob. *f* *sf* *pp*

Cl. (Sib) *f* *sf* *pp*

Tpt. (Dó)

Vi. I. *f* *sf* *pp*

Vi. II. *f* *sf* *pp*

Vla. *f* *sf* *pp*

Vcl. *f* *sf* *pp*

Allegro - accel. até Presto

Molto agitando

Fl. I. Fl. II. Ob. Cl. (Sib.) Tpt. (Dó) Vl. I. Vl. II. Vla. Vcl.

[Trompete segura a nota até a entrada das Flautas]

Fl. I.

Fl. II.

Ob.

Cl. (Si**b**)

Tpt. (D**b**)

VI. I.

VI. II.

Vla.

Vcl.

Con fuoco

fff

3

APÊNDICE D – Dados Técnicos da Gravação

Ficha Técnica Da Gravação

Peça	<i>The Unanswered Question</i>	
Compositor	<i>Charles Ives</i>	
Data	1906 / 1934 (revisão)	
Duração	5' 06"	
Local de gravação	Estúdio LAMI – ECA – USP	
Planejamento do experimento	Leandro Ferrari Thomaz	
	Regis Rossi Alves Faria	
Engenheiros de Som	Pedro Paulo Kohler	
	José Augusto Mannis	
Microfonação	José Augusto Mannis	
Mixagem e Masterização	Regis Rossi Alves Faria	
Edição e Revisão da Partitura	Otávio Simões	
	Leandro Ferrari Thomaz	
Direção Musical	Otávio Simões	
	José Augusto Mannis	
Regência	Otávio Simões	
Músicos	Jezreel Silva	(Trompete)
	Felipe Ferreira	(Flauta I)
	Regis Rossi Alves Faria	(Flauta II)
	Diogo Maia	(Clarinete)
	Arthur Ortenblaud	(Oboé)
	Ana Arakawa	(Violino I)
	Carolina Feuerharmel	(Violino II)
	Bruno Menegatti	(Viola)
	Leandro Carlini Tenorio	(Violoncelo)
	Thiago Faria	(Violoncelo)

Equipamento utilizado no estúdio para a gravação

<i>Qtd.</i>	<i>Equipamento</i>	<i>Marca</i>	<i>Modelo</i>	<i>Descrição</i>
1	Microfone	Sennheiser	e835	Dinâmico
2	Microfone	Sennheiser	MD 421	Dinâmico
1	Microfone	Shure	SM-58	Dinâmico
2	Microfone	Neumann	KM 184	Condensador
2	Microfone	AKG	C 414	Condensador
1	Mesa de Mixagem	Yamaha	O2R	
1	Interface de áudio	Digidesign	192	8 canais conectados via interface digital ADAT à mesa de mixagem
2	Monitor de Estúdio	Mackie	HR-824	
1	Computador Pessoal	Apple	Power Mac G4 1 GHz / 768 Mb RAM	
1	Software de gravação	Digidesign	Protools TDM 6.7	

5. Distância da fonte

1	2	3	4	5
Indeterminada				Realística

6. Tamanho da sala

1	2	3	4	5
Campo aberto				Sala esperada

7. Facilidade de uso do sistema

1	2	3	4	5
Difícil				Fácil

8. Comentários