

Escola das Artes da Universidade Católica Portuguesa
Mestrado em Som e Imagem



Ambientes Sonoros Interativos e Imersivos

Design de Som 2012-2013

Miguel Duarte Gonçalves

Professor Orientador: Prof. Doutor Luís Gustavo Martins

Professor Coorientador: Prof. Vítor Joaquim

Novembro de 2013

Escola das Artes da Universidade Católica Portuguesa
Mestrado em Som e Imagem



Ambientes Sonoros Interativos e Imersivos

Design de Som 2012-2013

Miguel Duarte Gonçalves

Professor Orientador: Prof. Doutor Luís Gustavo Martins

Professor Coorientador: Prof. Vítor Joaquim

Novembro de 2013

Agradecimentos

O trabalho desenvolvido nesta dissertação recebeu o apoio e colaboração de diversas pessoas e instituições. A estas entidades que das mais diversas formas contribuíram para a realização deste trabalho, cumpro o grato dever de expressar os mais sinceros agradecimentos. É devido um agradecimento especial:

Ao Professor Doutor Luís Gustavo Martins da Escola das Artes da Universidade Católica Portuguesa (EA/UCP), meu orientador da dissertação, pela confiança, apoio e conselhos prestados.

Ao Professor Vítor Joaquim da EA/UCP, meu coorientador do projeto, pela orientação e apoio.

Aos Professores João Cordeiro e Vasco Carvalho da EA/UCP, ao Mestre André Perrotta, pelo apoio dado no processo de busca de soluções no âmbito das tarefas de preparação e montagem da cablagem e suporte de altifalantes no Laboratório de *Motion Capture* da EA/UCP. Agradeço também aos colegas Adrian Santos e Vasco Pucarinho, pela ajuda nestas tarefas.

Ao Professor Ricardo Megre da EA/UCP e ao Mestre André Baltazar pela ajuda e tempo disponibilizados no âmbito do sistema de captura de movimentos *Vicon*. Agradeço também aos alunos da turma de Animação por Computador, pelo apoio nesta tarefa.

A Adrian Santos, pela ajuda na captação e edição de uma parte da componente sonora do projeto.

A Luís Freitas, pelo apoio prestado na montagem e resolução de problemas do sistema de jogo.

A Jorge Carvalho e Luís Carvalho pelo apoio no registo fotográfico do projeto.

A Pedro Oliveira pela ajuda no processo de requisição e transporte do material de difusão sonora.

À Escola de Artes da Universidade Católica Portuguesa, pelas instalações, condições e recursos disponibilizados, que permitiram a concretização do projeto.

Resumo

O uso de técnicas de espacialização sonora interativas para a criação de ambientes virtuais sonoros permite atingir uma resolução superior à apresentada pelos ambientes virtuais visuais. Recorrendo a estas ferramentas é possível o desenvolvimento de aplicações interativas e imersivas, como jogos sonoros. Não estando dependentes da componente visual e criando um estímulo à criatividade do utilizador, este género de experiências permite atingir um grau de imersão superior aos jogos tradicionais.

Esta investigação pretendeu analisar os principais métodos de espacialização sonora interativa e o fenómeno de imersão, de modo a permitir reunir o conhecimento necessário à concepção de uma instalação sonora interativa e imersiva.

Foi desenvolvido e testado o jogo sonoro *The Sound of Horror*, do género *first-person shooter* e subgénero *survival horror*, controlado através de uma interface à semelhança de uma arma de grande porte que permite um controlo gestual orgânico para a execução de múltiplas funções, permitindo ao utilizador interagir com alvos sonoros dinâmicos, na forma de criaturas monstruosas e assustadoras.

Palavras chave: Espacialização sonora, Imersão, Interatividade, Psicoacústica, Realidade Virtual.

Índice de Conteúdos

Lista de Figuras	vii
Lista de Tabelas.....	ix
Lista de Acrónimos	xi
1. Introdução.....	1
1.1. Apresentação do tema de investigação e do projeto final que lhe diz respeito	1
1.2. Apresentação da problemática	2
1.3. Metodologia utilizada para a investigação	3
1.4. Descrição da estrutura da dissertação	4
2. Espacialização Sonora Interativa e Imersiva	7
2.1. Espacialização sonora	7
2.1.1. História da tecnologia de espacialização sonora	8
2.1.2. Factores psicoacústicos da espacialização sonora.....	12
2.2. Técnicas de Espacialização Sonora Interativa.....	19
2.2.1. Binaural (Processamento <i>HRTF</i>)	21
2.2.2. <i>Amplitude Panning</i>	23
2.2.3. <i>Ambisonics</i>	29
2.2.4. <i>Wave Field Synthesis</i>	32
2.2.5. <i>Ambiophonics</i>	34
2.2.6. <i>ViMiC</i>	36
2.2.7. Comparação entre as técnicas	37
2.2.8. Interatividade na espacialização sonora	38
2.3. Imersão	40
2.3.1. Modelos de imersão	40
2.3.2. Imersão no som	41
2.3.3. Exemplos de projetos imersivos.....	43
3. <i>The Sound of Horror: Instalação Sonora Interativa e Imersiva</i>	45
3.1. Escolha do projeto	45
3.2. Objectivos	47
3.3. Planeamento	49
3.4. Metodologias	49
3.4.1. Narrativa e estrutura do jogo.....	49
3.4.2. Sistema de espacialização sonora.....	55
3.4.3. Sistema de difusão sonora.....	63
3.4.4. Interface (Arma).....	65
3.4.5. Estrutura para confinar o utilizador.....	69
3.4.6. <i>Tracking</i> de movimentos.....	71
3.4.7. Componente sonora.....	80
3.4.8. Componente de programação.....	90
3.5. Montagem e apresentação	99
3.6. Discussão	106
4. Conclusão	109

5. Referências e Bibliografia	111
5.1. Livros e artigos	111
5.2. Websites	117
5.3. Jogos	118
5.4. Patentes	118
Apêndice 1: Calendarização do Projeto	119
Apêndice 2: DVD com Conteúdos Digitais	121

Lista de Figuras

Figura 1 – Representação dos planos mediano e horizontal.....	13
Figura 2 – Diagrama de representação da ITD.....	14
Figura 3 – Representação da “sombra acústica” (ILD).....	15
Figura 4 – Exemplo de gráficos HRIR e HRTF.....	16
Figura 5 – Representação do cone of confusion.....	18
Figura 6 – Diagrama de um sistema de síntese espacial com base nos princípios binaurais.....	23
Figura 7 – Configuração estereofônica.....	25
Figura 8 – Configuração quadrafônica.....	25
Figura 9 – Configuração 5.1 Surround-sound.....	25
Figura 10 – Representação da técnica de espacialização DBAP.....	26
Figura 11 – Triângulo de altifalantes para a criação de fontes sonoras através do sistema VBAP.....	28
Figura 12 – Formação de triângulos para cinco altifalantes através do sistema VBAP.....	28
Figura 13 – Representação das componentes W, X, Y e Z do B-format.....	30
Figura 14 – Representação do padrão de interferência na reprodução com dois altifalantes.....	33
Figura 15 – Representação da reprodução através de um array de altifalantes.....	33
Figura 16 – Exemplo da reconstrução do campo sonoro utilizando a WFS.....	34
Figura 17 – Representação da captação e reprodução de sinal sonoro num sistema stereo dipole.....	35
Figura 18 – Representação da captação dos impulsos de resposta, convolução multicanal e reprodução da camada surround Ambiophonics.....	36
Figura 19 – Exemplo de utilização do sistema CAVE.....	43
Figura 20 – Representação virtual da estrutura Allosphere.....	44
Figura 21 – Representação neurofisiológica no AlloBrain.....	44
Figura 22 – Fotografia do sistema de jogo sonoro elaborado para a unidade curricular de Interfaces Tangíveis.....	47
Figura 23 – Diagrama simplificado da estrutura do jogo The Sound of Horror.....	51
Figura 24 – Diagrama extenso da estrutura do jogo The Sound of Horror.....	53
Figura 25 – Patch de implementação da técnica DBAP no Max/MSP (1).....	56
Figura 26 – Patch de implementação da técnica DBAP no Max/MSP (2).....	56
Figura 27 – Representação da técnica VBAP no Max/MSP.....	57
Figura 28 – Imagem do módulo de ViMiC do Jamoma no Max/MSP.....	58
Figura 29 – Imagem dos objetos HOALibrary no Max/MSP.....	59
Figura 30 – Exemplo do uso dos objetos Ambisonics do ICST no Max/MSP.....	60
Figura 31 – Imagem do sistema de calibração da distância nos objetos Ambisonics do ICST.....	60
Figura 32 – Exemplo de manipulação da trajetória de conjuntos de fontes sonoras através dos objetos Ambisonics do ICTS.....	61
Figura 33 – Imagem dos objetos Binaural Tools no Max/MSP.....	62
Figura 34 – Fotografia do altifalante Genelec 6010A.....	63
Figura 35 – Fotografia do altifalante Genelec 7070A.....	63

Figura 36 – Representação da distribuição bidimensional de altifalantes, do projeto The Sound of Horror.....	65
Figura 37 – Interface do jogo The Sound of Horror (1).....	66
Figura 38 – Interface do jogo The Sound of Horror (2).....	66
Figura 39 – Interface do jogo The Sound of Horror (3).....	67
Figura 40 – Interface do jogo The Sound of Horror (4).....	67
Figura 41 – Objeto externo ajh.wiiremote para o Max/MSP.....	69
Figura 42 – Construção da estrutura para confinar o utilizador (1).....	70
Figura 43 – Construção da estrutura para confinar o utilizador (2).....	70
Figura 44 – Imagem do sistema de tracking utilizando o dispositivo Kinect desenvolvido por Seldess.....	73
Figura 45 – Fotografia do sistema de leds infravermelhos montados com uma organização semiesférica.....	74
Figura 46 – Fotografia do sistema de leds infravermelhos montados em superfície plana de nível num sistema de contrapeso.....	74
Figura 47 – Fotografia da caixa com iluminação e um marcador fiducial do software ReactIVision.....	75
Figura 48 – Fotografia do Laboratório de Motion Capture da Escola das Artes da UCP.....	76
Figura 49 – Captura de ecrã dos programas Vicon e Vicon2OSC.....	77
Figura 50 – Interface de jogo com marcadores do sistema Vicon.....	78
Figura 51 – Fotografia do reconhecimento e tracking automático do objeto calibrado pelo sistema Vicon.....	79
Figura 52 – Sessão ProTools HD9 dos elementos sonoros relativos ao Cinematic Tutorial.....	88
Figura 53 – Representação da distribuição do sinal sonoro das categorias Ambientes e Cinematics.....	89
Figura 54 – Patch do projeto The Sound of Horror.....	91
Figura 55 – Módulo de programação relativo à interface de jogo.....	91
Figura 56 – Módulo de programação relativo ao tracking de movimentos e cálculo do ângulo de mira.....	92
Figura 57 – Reprodução em loop, com duração e curva de fade-in e fade-out ajustáveis dos elementos sonoros da categoria Ambientes do Menu.....	93
Figura 58 – Processamento dos elementos sonoros relativos aos monstros do Nível 2 com o plugin Waves IR-1.....	94
Figura 59 – Criação, geração da trajetória e cálculo dos “pontos de vida” de um monstro do Nível 1.....	95
Figura 60 – Módulo de associação dos elementos sonoros das categorias Ambientes e Cinematics às etapas de jogo.....	96
Figura 61 – Módulo de motor de jogo.....	97
Figura 62 – Módulo de calibração do sistema de jogo.....	98
Figura 63 – Diagrama do fluxo de informação entre as várias componentes (hardware) do projeto The Sound of Horror.....	100
Figura 64 – Perspectiva do espaço com o sistema de jogo montado (1).....	101
Figura 65 – Perspectiva do espaço com o sistema de jogo montado (2).....	101
Figura 66 – Perspectiva do espaço com o sistema de jogo montado (3).....	102
Figura 67 – Perspectiva do espaço com o sistema de jogo montado (4).....	102
Figura 68 – Perspectiva do espaço com o sistema de jogo montado (5).....	103
Figura 69 – Perspectiva do espaço com o sistema de jogo montado (6).....	103
Figura 70 – Perspectiva do espaço com o sistema de jogo montado (7).....	104
Figura 71 – Perspectiva do espaço com o sistema de jogo montado (8).....	104
Figura 72 – Perspectiva do espaço com o sistema de jogo montado (9).....	105
Figura 73 – Perspectiva do espaço com o sistema de jogo montado (10).....	105

Lista de Tabelas

Tabela 1 – Categorização de aplicações de espacialização de áudio.	8
Tabela 2 – Comparação entre os métodos de espacialização binaural e multicanal.	20
Tabela 3 – Equações de codificação Ambisonics até à 3ª ordem.	31
Tabela 4 – Comparação entre as técnicas de espacialização sonora.	38
Tabela 5 – Quatro domínios da experiência, segundo Pine e Gilmore.	41
Tabela 6 – Lista de etapas presentes no jogo de acordo com o seu grau de interatividade.	52
Tabela 7 – Lista de diálogos do Narrador e Grande Irmão.	54
Tabela 8 – Lista de elementos sonoros do projeto The Sound of Horror.	81
Tabela 9 – Coordenadas cartesianas dos altifalantes no espaço de instalação de acordo com o sistema de captura de movimentos Vicon.	99

Lista de Acrónimos

<i>DBAP</i>	<i>Distance-based Amplitude Panning</i>
<i>e.g.</i>	por exemplo (Latim: <i>exempli gratia</i>)
EA/UCP	Escola das Artes da Universidade Católica Portuguesa
<i>et al.</i>	e outros (Latim: <i>et alteri</i>)
<i>etc.</i>	Etetcetera (Latim: <i>et cetera</i>)
<i>HRIR</i>	<i>Head Related Impulse Response</i>
<i>HRTF</i>	<i>Head Related Transfer Functions</i>
<i>i.e.</i>	isto é (Latim: <i>id est</i>)
<i>ILD</i>	<i>Interaural Level Differences</i>
<i>ITD</i>	<i>Interaural Time Differences</i>
<i>OSC</i>	<i>Open Sound Control</i>
<i>VBAP</i>	<i>Vector Base Amplitude Panning</i>
<i>ViMiC</i>	<i>Virtual Microphone Control</i>
<i>WFS</i>	<i>Wave Field Synthesis</i>

1. Introdução

1.1. Apresentação do tema de investigação e do projeto final que lhe diz respeito

O desenvolvimento de sistemas de reprodução de som com a capacidade de simular a existência ambientes acústicos é um tema que foi alvo de bastante estudo desde o século passado. Através da análise do comportamento das ondas sonoras, da forma como são afectadas pelo seu meio envolvente e das características psicoacústicas humanas foi possível desenvolver sistemas de som com a capacidade de envolver sensorialmente o utilizador, reproduzindo não só ambientes sonoros já existentes com as suas características originais, mas também simulando a existência e o movimento de fontes sonoras. Isto levou a grandes avanços no mundo da engenharia de som, que posteriormente foram transpostos para outras áreas, como a da música e do cinema (Peters, 2010).

Recentemente, com o aumento da capacidade de processamento dos computadores, estes avanços começaram a ser conceptualizados de modo a permitir explorar ao máximo a criação de “realidades virtuais”, aliando a criação de espaços sonoros artificiais com o elemento visual, através de projecções gráficas ou outros sistemas que permitem um envolvimento a este nível. O objectivo deste género de trabalhos é permitir que as informações fornecidas passem a ser tomadas como reais, no decorrer da sua utilização, imergindo assim, o espectador (Wozniowski, 2006).

A imersão, num contexto audiovisual, pode considerar-se como o grau de ligação perceptual, emocional e ligação de concentração relativa a interações com o sistema emocional do utilizador ou espectador à aplicação em causa. De modo a não criar quebras de imersividade é, então, importante criar uma continuidade e coerência no decorrer da ação (Grimshaw, 2007).

No que respeita aos conteúdos audiovisuais, deve haver uma harmonia e lógica nos eventos. Ou seja, com todas as dinâmicas que a peça possa ter, é importante que os componentes expostos preservem um fio condutor que mantenha o utilizador “preso” à experiência. Esse fio é a narrativa apresentada pelo trabalho.

A utilização de controladores que não restrinjam os movimentos do utilizador, agindo de forma intuitiva, permitem, também, uma associação direta com a realidade virtual apresentada, melhorando a experiência do utilizador (Wozniowski, 2006).

Apesar do sentido da visão ter uma resolução muito superior à da audição em termos de localização de fontes no espaço (o que leva a que o ser humano seja considerado por muitos como um ser primariamente visual), este não é essencial para que haja imersão. O som, por outro lado, é indispensável, sendo que providencia informações relativamente ao espaço através do som do ambiente e das suas reverberações e, ao contrário da visão, o campo de escuta é esférico com centro no ouvinte, permitindo receber pistas auditivas de qualquer direção (Moore, 1995).

Dito isto, os objectivos principais desta dissertação são os seguintes:

- Estudo histórico e técnico das ferramentas de espacialização sonora;
- Estudo das capacidades psicoacústicas humanas;
- Estudo da interatividade na espacialização sonora;
- Estudo do conceito de imersão e sua relação com o som;
- Desenvolvimento de uma instalação imersiva e interativa com uma componente sonora preponderante.

1.2. Apresentação da problemática

Tendo em conta que a maioria dos sistemas imersivos, tanto jogos como instalações, contam com dados visuais para uma mais profunda imersão na realidade virtual criada, torna-se pertinente perceber de que forma é possível atingir o mesmo nível de imersão sem fornecer quaisquer dados perceptuais visuais, o que implica tornar a experiência sonora completa o suficiente, de modo a conseguir proporcionar ao utilizador toda a informação que ele necessita para a navegação no espaço da instalação.

Dependendo do método de reprodução do sistema de espacialização sonora escolhido (se este se baseia na utilização de altifalantes ou numa escuta com auscultadores), há um conjunto de vantagens e desvantagens associadas a cada um (Faria, 2005). Em termos de utilização prática no contexto de uma instalação, estes estão essencialmente relacionados com o ponto de escuta ideal. No caso da utilização de colunas, este é um ponto fixo no espaço, o

que impede o utilizador de se deslocar livremente. Na situação dos sistemas de auscultadores, o problema prende-se com captação dos movimentos da cabeça, que geralmente é resolvido através do *tracking* da posição e ângulo desta, de modo a que todo o espaço sonoro se mantenha coerente com os movimentos efectuados. No entanto, há sempre um conjunto de factores que impedem uma reprodução perfeita deste sistema, nomeadamente a latência entre os aparelhos utilizados para a captura destes movimentos, que restringe a velocidade de movimentos do utilizador, bem como uma liberdade total de movimentos, caso para esta detecção seja necessário que o utilizador se mantenha numa inclinação ou ângulo relativamente constante. Pretende-se, então, responder a estas limitações para manter a coerência imersiva e maximizar os resultados.

Em situações onde haja um grande número de elementos com o qual o utilizador possa interagir, e estes se situem num mesmo local do campo sonoro, em contexto de instalações imersivas sem feedback visual, pode tornar-se complicado dar resposta a todos estes componentes mantendo a inteligibilidade de cada um destes, sem gerar confusão para o utilizador. Geralmente, os trabalhos nesta área dependem da componente visual (Cruz-Neira *et al.*, 1992; Cruz-Neira *et al.*, 1993; Ihren & Frisch, 1999; Amatriain *et al.*, 2007; Overholt *et al.*, 2008; Amatriain *et al.*, 2009; Cooper, 2011; “Red Paper Heart”, 2013), de modo a conseguir melhor exprimir a interação entre o sistema e o utilizador, tirando partido das capacidades perceptuais e da forma como cada sentido complementa o outro.

Apesar do uso conjunto destes dois sentidos poder ser uma forma de transmissão de informação que mais se assemelha àquela experienciada no mundo real, não é uma obrigatoriedade o uso de ambas para que seja possível a criação de um ambiente imersivo e com inteligibilidade nos seus elementos.

1.3. Metodologia utilizada para a investigação

A fim de consolidar conceitos e de recolher informação, conhecimento e inspiração, foi efectuado um estudo teórico de modo a permitir o desenvolvimento de uma instalação sonora, interativa e imersiva.

Procedeu-se à leitura de livros e artigos para o desenvolvimento do estudo científico da investigação. Foi possível recorrer ao uso de técnicas em diversas áreas (*i.e.* espacialização sonora, *tracking* de movimentos e reconhecimento gestual) através do acesso a um leque de

ferramentas disponíveis *online*. Através da análise da componente sonora de videogames, foi possível reunir a inspiração necessária para o desenvolvimento do projeto.

O processo de criação do produto final da investigação envolveu o uso de sistemas de espacialização sonora interativa e imersiva como *Ambisonics*, *VBAP*, *ViMiC*, entre outros. Recorreu-se ao uso de *hardwares* para a detecção dos movimentos do utilizador, nomeadamente o comando *Nintendo Wiimote* a câmara *Microsoft Kinect* e o sistema de captura de movimentos *Vicon*. Para a criação e manipulação dos elementos sonoros foi utilizado o *Pro Tools HD9*, em conjunto com a linguagem de programação *Max/MSP* para o processamento em tempo real dos ficheiros de som. Esta linguagem foi utilizada para a integração dos diferentes elementos do projeto, permitindo interligá-los e conceber uma instalação sonora interativa e imersiva.

Este trabalho decorreu segundo o seguinte plano:

- Escolha do tema da investigação e do projeto, em Setembro de 2012;
- Recolha e leitura de artigos, livros e *websites*, bem como a análise de projetos de referência, de Setembro a Dezembro de 2012;
- Planificação da estrutura da dissertação, em Outubro de 2012;
- Preparação e organização do projeto, de Outubro a Dezembro de 2012. No Apêndice 1 apresenta-se um cronograma detalhado das atividades desenvolvidas relativas ao projeto final;
- Escrita do capítulo de revisão do estado da arte de Outubro a Dezembro de 2012 e de Maio a Julho de 2013;
- Desenvolvimento do projeto, de Janeiro a Abril de 2013;
- Apresentação do projeto, de 24 de Abril de 2013;
- Escrita do capítulo relativo ao desenvolvimento do projeto, de Maio a Setembro de 2013;
- Revisão e finalização da dissertação, em Outubro e Novembro de 2013.

1.4. Descrição da estrutura da dissertação

O capítulo 2 apresenta-se subdividido em três secções. Na primeira, é feito um estudo histórico relativamente à tecnologia de espacialização sonora, seguido de uma análise aos

factores psicoacústicos humanos que permitem uma localização de fontes sonoras no espaço. A segunda secção conta com um estudo comparativo relativamente ao modo de funcionamento e aos conceitos subjacentes às principais técnicas de espacialização sonora, em conjunto com uma análise aos aspectos interativos associados à utilização destas ferramentas para síntese de som espacial. Para finalizar, é estudado o fenómeno de imersão de um modo geral e aplicado especificamente ao som, apresentando alguns dos modelos sugeridos no estudo de aplicações de realidade virtual, seguido de uma análise relativamente a alguns projetos de referência na área da imersão.

No capítulo 3 é feito o acompanhamento do processo de criação do projeto *The Sound of Horror*. Em primeiro lugar é relatado o processo de escolha do projeto no contexto da investigação. Seguidamente, são expostos os objetivos pretendidos com este jogo sonoro e é feita a apresentação das várias metodologias testadas e implementadas. Por fim, é feita uma crítica às opções tomadas no projeto, confrontando-as com outros trabalhos.

O capítulo 4 apresenta a conclusão da investigação. É feita uma referência aos temas abordados nos capítulos anteriores e feito um cruzamento da informações obtidas. Finaliza-se a investigação com a apresentação de perspectivas de trabalho futuro.

2. Espacialização Sonora Interativa e Imersiva

2.1. Espacialização sonora

R. Murray Schafer, no seu livro *Soundscapes: Our Sonic Environment and the Tuning of the World* (Schafer, 1993), afirma que o ser humano, desde os primórdios, esteve imerso em ambientes sonoros que influenciaram bastante o seu comportamento e forma de estar, criando relações muito fortes com certos elementos, nomeadamente sons de grande escala de intensidade, como é o caso de temporais, trovões ou o som do próprio mar, que eram impossíveis para o Homem de perceber, mas que aprendeu a temer e a respeitar. Tentou, eventualmente, captar estas características, através da construção e utilização de instrumentos que lhe permitiam replicar estes atributos que temia da Natureza, de modo a conceder este carácter titânico a situações ou indivíduos particulares, como é possível perceber pela utilização de sinos para indicar eventos religiosos ou de canhões para contextos bélicos. Outro exemplo de uma característica de cariz sonoro que prevaleceu desde os exórdios da humanidade foi a conotação negativa associada ao silêncio, estando, inconscientemente, muitas vezes associado a perigo ou morte.

Esta relação forte com os ambientes sonoros que nos rodeia foi sempre um importante objecto de estudo ao longo do tempo, impulsionando investigações ao nível da psicoacústica, para compreender a forma como o som nos chega aos ouvidos e a maneira como interpretamos esses sinais. Eventualmente, com o evoluir das tecnologias surgiu também a necessidade de captar e reproduzir ambientes e cenas sonoras, contendo fontes sonoras em diferentes posições, dando origem ao conceito de realidade virtual e de sistemas capazes de criar espaços artificiais com o intuito de imergir um ou mais espectadores, tirando partido das suas capacidades sensoriais (Jot, 1997).

É possível, através da Tabela 1, adaptada da construída por Francis Rumsey (2002), perceber o tipo de aplicações em que a área da espacialização sonora se insere e o género de precisão sonora necessária para o funcionamento da mesma. É feita uma distinção entre uma espacialização onde apenas é induzida uma sensação de ambiência (associada com fins criativos e artísticos) e um outro género que envolve uma precisão espacial muito mais exata, sendo necessário um *rendering* 3D preciso.

Tabela 1 – Categorização de aplicações de espacialização de áudio (Adaptado de Rumsey, 2002).

Aplicação	Ilusão criativa/artística	Rendering 3D preciso
Música clássica/ao vivo/natural	Muito	Pouco
Música <i>pop</i>	Muito	-
Som de rádio/televisão	Muito	Pouco
Som de cinema	Muito	-
Realidade virtual	Pouco	Muito
Jogos de computador	Muito	Pouco
Sistemas de simulação/controlo	-	Muito
Sistemas de conferência/comunicação	-	Muito

O autor conclui, então, que nas aplicações categorizadas como sendo as mais exigentes a nível de precisão sonora, o objectivo principal do sistema está relacionado com algum género de interação espacial ou interação humana dentro de um ambiente virtual.

Neste capítulo, vai ser feita, inicialmente, uma breve análise histórica das técnicas de espacialização sonora e da sua evolução ao longo do tempo até ao presente. De seguida vai ser feito um estudo com o intuito de perceber que características do campo da psicoacústica são responsáveis pela capacidade de localizar fontes sonoras oriundas de qualquer direcção e distância, para, e seguidamente, um estudo relativamente às técnicas de espacialização sonora existentes, percebendo como elas funcionam e as vantagens e desvantagens associadas à sua utilização. Para finalizar, será feito, um estudo ao nível da Interação e Imersão sonoras, para melhor entender quais os aspectos necessários para a criação de um ambiente sonoro interativo e imersivo.

2.1.1. História da tecnologia de espacialização sonora

Dada a forma como os avanços na área da espacialização começaram, é possível afirmar, antes de mais, que as raízes do seu desenvolvimento se prendem com a invenção de sistemas de captação sonora, rádio e telefonia (Zvonar, 2005).

O primeiro uso registado da utilização de um sistema com múltiplos canais de áudio para criar a sensação de espacialização sonora data de 1881 (Peters, 2010; Rumsey, 2002), através do *Théâtrophone* de Clement Ader, um sistema de distribuição telefónica que permitia

transmitir o sinal sonoro, em tempo real, captado por um conjunto de microfones em peças de teatro ou ópera, através de linhas telefônicas, que eram recebidas posteriormente pelos espectadores, utilizando um *headset* de dois canais, no *Paris Exhibition of Electricity* (Askew, 1981). Apesar desta obra inovadora, não houve nenhum trabalho de relevo nesta área até ao início do século seguinte (Malham *et al.*, 1995), quando, em 1920, Harvey Fletcher, em conjunto com a sua equipa nos *Bell Telephone Laboratories*, desenvolveu um sistema de escuta binaural, tendo patenteado o mesmo uns anos mais tarde (Fletcher & Sivian, 1927), em conjunto com um estudo extenso relativamente às características psicoacústicas e de fala, que, através da utilização de um molde de cabeça humana, visava transmitir um sinal sonoro, emulando a filtragem e os desfasamentos com que os ouvidos humanos captam sons oriundos de diferentes localizações, tentando recriar a mesma sensação de espacialização num ouvinte, com auscultadores (Sanal, 1976), com o objectivo de sair do mundo de comunicações em um só canal (Atkins, 2011). No entanto, esta ideia foi esquecida, tendo sido reavivada mais tarde, na Alemanha, nos anos 70 (Fox, 1982).

Graças à contribuição do engenheiro eléctrico Christopher Wenthe no ano de 1916, que originou os microfones condensadores (Wenthe *et al.*, 1931), foram abertas portas no que toca ao trabalho na área da espacialização de áudio, sendo que com esta invenção, surgiram, alguns anos mais tarde, técnicas de microfonia multicanal, que vieram originar o conceito de *stereo* como o conhecemos hoje (Faria, 2005). Em particular as de microfones coincidentes, em 1931, por Alan Blumlein (1934), que vieram depois a dar origem ao sistema *Ambisonics*, sendo que as ideias inerentes ao conceito por trás desta técnica são fundamentais para este sistema (Farrar, 1979). E as de microfones espaçados, em 1934, desenvolvidas por Steinberg e Snow (Steinberg & Snow, 1934a), que vieram a ser a base para a criação da técnica de reprodução de ambientes sonoros conhecida como *Wave Field Synthesis* (Peters, 2010).

No decorrer dos anos 30, os *Bell Laboratories* desenvolveram um sistema constituído por *arrays* de colunas, que tinham como objectivo recriar um ambiente sonoro, captado por um igual *array* de microfones. No entanto, apesar de não terem obtido resultados satisfatórios com a utilização destes *arrays*, aperceberam-se que, a utilização de apenas três microfones espaçados, ligados a três colunas com uma disposição semelhante, era o mínimo de canais possíveis a utilizar para dar uma sensação de espacialização precisa e aceitável para o efeito (Fox, 1982; Gerzon, 1973; Hull, 1999).

É possível ver a importância destas descobertas nos dias de hoje, através da constituição dos sistemas *surround* atuais, que utilizam estas três colunas numa disposição semelhante

(ITU-R, 2012). Steinberg e Snow tiveram um impacto direto neste trabalho, graças aos seus avanços nesta área (Malham *et al.*, 1995).

O mundo do áudio espacial iria, transitar para o cinema, tendo a sua primeira aparição no filme *Fantasia*, de Walt Disney, em 1939, com a participação do maestro Leopold Stokowski (Goldman *et al.*, 2002), com a utilização de um conjunto de gravadores sincronizados ligados a trinta e três microfones, para a captação sonora da sua orquestra (Hope, 1979). Recorreu, ainda, a um sistema de reprodução desenvolvido especificamente para esta obra, chamado de *Fantasound*, que utilizava dois projetores, sendo que o primeiro continha a informação visual juntamente com um *mono mix* do filme. No segundo projetor, estavam contidas quatro pistas mono, onde a primeira possuía informações de controlo para modelar o volume de algumas colunas no decorrer do espetáculo. As pistas dois, três e quatro continham o áudio para a coluna esquerda, direita e central, respectivamente. Estes canais derivavam ainda para três colunas colocadas atrás dos espectadores, dando nascimento ao primeiro sistema *surround* (Shepherd, 2003).

Nos anos que se seguiram, foi continuado o trabalho nesta área, tendo, em 1942, surgido as primeiras gravações *stereo* feitas em cassete, por Helmut Kruger (Peters, 2010).

Em 1951, Pierre Schaeffer, engenheiro na *Radiodiffusion-Télévision Française* apresentou o seu controlador *Potentiomètre d'espace*, um aparelho que, através da utilização de bobina de indução, controlava dinamicamente os níveis de leitores de vinil ligados a quatro colunas (Holmes, 2008). Apesar da apresentação pública deste sistema ter sido bastante bem sucedida, conseguindo captar a atenção de compositores e entusiastas, a manipulação espacial do som em si não foi bem conseguida, sendo que era muito complicada a sua operação sem automações (Fishman-Johnson, 1993).

Em 1952 surgiu o primeiro vinil estereofónico, de Emory Cook. Sendo seguido, poucos anos mais tarde, em 1957, da primeira unidade de reverberação artificial, o *EMT-140* (Peters, 2010).

A síntese de som espacial e difusão sonora tornou-se possível e popular na música contemporânea graças à integração de microfones, gravadores e colunas na performance musical (Peters, 2010), como é possível observar do trabalho de Stockhausen, *Kontakte* (1958-1960). Esta foi provavelmente a primeira composição feita especificamente para uma reprodução quadrafónica. Stockhausen utilizou um sistema denominado de *Rotation Table*, que era compreendido por um gira-discos com um sistema de colunas rotativas, rodeadas por quatro microfones, permitindo assim captar os sons das colunas que se encontravam a girar,

criando a ilusão de que os sons estavam a orbitar em redor dos espectadores, que se encontravam no centro de uma quadrafonia, que reproduzia o som captado pelos microfones. Stockhausen veio mais tarde a dar continuação a este trabalho na peça *Oktophonie* (1990-1991) (Zvonar, 2005).

Em simultâneo, no ano de 1958, Edgard Varèse apresentou o seu trabalho *Poème électronique*, que consistia no mapeamento de três canais de som para cerca de 350 colunas, através da utilização de *rotaries* de telefone, tendo contado também com projeção de vídeo. (Lombardo *et al.*, 2005), que veio a impulsionar a criação de modelos de colunas chamados de *loudspeaker orchestras*, no início da década de 70 (Deruty, 2012).

Em 1969, foi introduzido o sistema de espacialização sonora Ambisonics, por Michael Gerzon (Gerzon, 1974; Cooper *et al.*, 1998), que viria a ser estendido para a sua versão mais avançada, chamada de *Higher Order Ambisonics* em 2000 (Daniel *et al.*, 2003).

Na *Osaka Internation Exhibition* de 1970, no Japão, Stockhausen apresentou uma obra musical numa configuração multicanal que consistia numa esfera de 50 colunas, posicionadas em sete níveis diferentes, acima e abaixo da audiência, criando uma sensação de imersão (Lynch *et al.*, 2011).

No ano de 1971, John Chowning introduziu o conceito de criação de trajectórias para fontes sonoras (Chowning, 1977), tendo, um ano mais tarde apresentado a sua obra *Turenas*, que utilizava deste sistema (McGee, 2009). Este conceito apresentado por Chowning veio influenciar inúmeros trabalhos, como, por exemplo, o *Spatialisateur*, um software de processamento espacial de áudio que tem vindo a ser desenvolvido desde 1992 pelo IRCAM (*Institut de Recherche et Coordination Acoustique/Musique*) (Jot, 1997).

Chegando à década de 70, no mundo do cinema, os formatos de *film stereo* utilizados (desenvolvidos no início dos anos 50, que compreendiam entre quatro a seis canais, dependendo do utilizado), nomeadamente o *CinemaScope* (35 mm) e o *Todd-AO* (70 mm) começavam a tornar-se pouco eficazes devido aos custos elevados envolvidos, o que impulsionou a *Dolby* a desenvolver o seu próprio formato chamado de *Dolby Stereo* (1975), agora operacional em fita óptica (em contraste com a anterior fita magnética) contendo a informação de quatro canais (Left, Right, Center, Surround) armazenados em apenas dois canais através de uma matriz de inversão de fases (Hull, 1999).

Com o aumento dos recursos ao nível dos computadores, houve (prolongando-se para o presente) uma evolução conceptual e melhoramento das técnicas de posicionamento espacial e de criação de ambientes virtuais sonoros (Pulkki, 1997), onde se destacam o surgimento de

sistemas como o *Ambiophonics* (1986) (Farina *et al.*, 2001) *Wave Field Synthesis* (1988) (Spors *et al.*, 2008), as várias formas de *Dolby Surround* (Peters, 2010) (com a eventual *standardization* do 5.1 em 1991), *Vector base Amplitude Panning* (1997) (Pulkki, 1997) e mais recentemente *Distanced-based Amplitude Panning* (2003) (Lossius *et al.*, 2009) e *ViMiC* (2005) (McGee, 2009).

Esta evolução tem permitido, também, uma sobreposição de matérias do campo da *Human-Computer Interaction* com as da espacialização sonora, possibilitando um controlo em tempo real de características espaciais do som através de interfaces gráficas ou de controlo gestual (Gasteiger, 2010).

2.1.2. Factores psicoacústicos da espacialização sonora

Segundo Dorte Hammershøi (Yôiti, Brungart *et al.*, 2009), o processo de localização sonora consiste na associação do posicionamento de fontes sonoras com a posição do ouvinte, permitindo a criação de um “espaço perceptual”. O autor afirma, ainda, que, em determinados casos, uma localização precisa das fontes sonoras determina o sucesso auditivo, como em situações de navegação num meio onde o sentido principal é o da audição, ou em caso de necessidade de filtrar uma fonte específica em ambientes complexos com um conjunto grande de elementos sonoros (referido originalmente como *Cocktail Party Effect* (Cherry, 1953)). Hammershøi faz ainda uma análise dos testes e experiências feitas na área do estudo das capacidades perceptuais de espacialização auditiva, dividindo-as em três grandes grupos, consoante os objectivos das mesmas, sendo o primeiro de localização absoluta de fontes sonoras no espaço, o segundo de localização relativa (pretendendo averiguar se alterações no ângulo de incidência ou na distância das fontes são interpretadas corretamente) e, por fim, testes de identificação (reconhecimento) de fontes diretas. Esta dissertação incidirá essencialmente no primeiro grupo, pelo que, para isso, neste capítulo serão estudados os factores e indicadores (chamados de *cues* (Faria, 2005)) sonoros psicoacústicos associados a uma localização absoluta no espaço.

2.1.2.1. Indicadores espaciais (*spatial cues*)

Os principais indicadores que permitem uma percepção espacial precisa baseiam-se nas diferenças temporais com que os ouvidos captam os sons e nas alterações ao nível do espectro de frequências que estes sofrem ao colidir com o nosso corpo (Blauert, 1997). Às diferenças temporais dá-se o nome de *Interaural Time Differences (ITD)* e à forma como a cabeça e o tronco “mascaram” o som, atenuando certas frequências denomina-se de *Interaural Level Differences (ILD)*. A estes dois elementos, que juntos permitem uma percepção tridimensional do som, Lord Rayleigh, em 1907, intitulou de *Duplex Theory* (Rayleigh, 1907).

No decorrer do trabalho, serão feitas alusões relativamente à posição, no espaço, dos sons. Deste modo, é importante definir uma representação espacial para referências futuras. Na Figura 1, adaptada de Pulkki (1997), é demonstrado o sistema de representação sugerido por Blauert (1997).

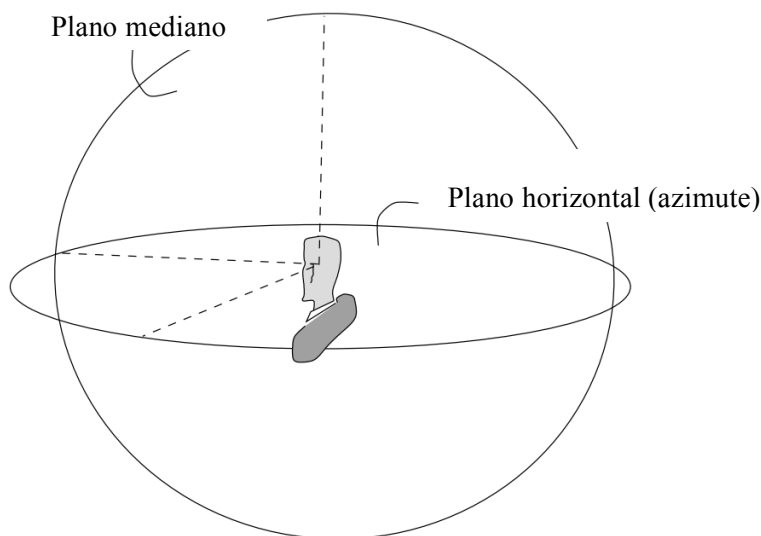


Figura 1 – Representação dos planos mediano e horizontal (adaptado de Pulkki, 1997).

Nesta representação, é assumido como centro a cabeça do ouvinte, de modo a que toda a análise feita tenha em perspectiva a sensação de audição deste ponto de vista. A posição no plano horizontal é correspondente ao ângulo de azimute (θ), onde a referência, o 0 (zero), corresponde à direção exatamente à frente da cabeça. O plano mediano pretende indicar a

elevação angular (φ) da fonte sonora, em que o 0 corresponde à direção à frente da cabeça do ouvinte.

Se for colocada uma fonte sonora com um ângulo 0° no plano mediano e com um ângulo de 90° de azimute (ou no plano horizontal), a onda sonora emitida, irá percorrer o comprimento total da cabeça do ouvinte, onde atingirá o ouvido oposto (Gasteiger, 2010). Considerando como distância entre os dois ouvidos um valor de 20-25 cm e sabendo que a velocidade do som no ar é de 340 m/s, é possível obter o valor máximo possível de *ITD* entre o som recebido nos dois ouvidos, que dá entre 600 a 700 μ s, dependendo do tamanho da cabeça do ouvinte. Como seria de esperar, os valores, mesmo na sua extensão máxima, equivalem a resultados numa escala muito pequena, pelo que estas diferenças temporais são referidas por alguns autores como *Interaural Phase Differences* (Wang *et al.*, 2005).

Na Figura 2, é possível observar a representação da *ITD* de um determinado som ao chegar aos dois ouvidos separadamente, o que permite uma percepção da localização da fonte, para criar o “espaço perceptual”.

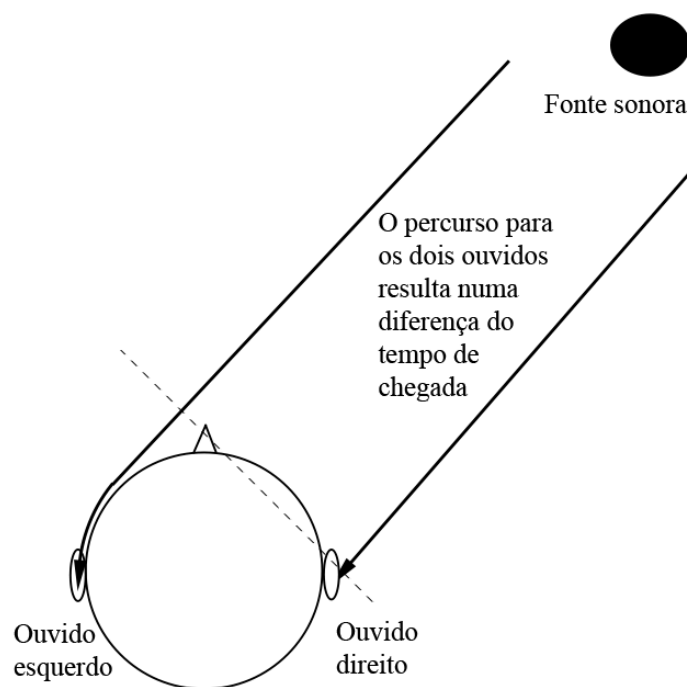


Figura 2 –Diagrama de representação da *ITD* (adaptado de Rumsey, 2002).

Quando os sons recebidos se encontram abaixo dos 1.5 KHz, estes intervalos temporais são percebidos como diferenças de fase de onda. No entanto, quando o valor de frequência se aproxima deste número, torna-se impossível a distinção desta forma, visto que a resolução dos ouvidos não é suficiente para ter uma percepção precisa (Carlile, 2011). Nestes casos, o próprio *envelope* da onda sonora (também chamado de *carrier wave*) é utilizado

perceptualmente de modo a conseguir obter uma a localização da fonte sonora. O *envelope* define-se como sendo a linha que contorna os extremos de amplitude de uma onda, formando um limite contínuo (Burgess, 1992). No entanto a sensibilidade ao *envelope* é perdida quando este se aproxima dos 100-200 Hz (Carlile, 2011; Middlebrooks & Green, 1990).

A própria colisão das ondas sonoras com a cabeça, ombros e tronco influencia a percepção sonora, através de fenómenos de absorção e ressonância, sendo muitas vezes referida como “sombra acústica” (Wang *et al.*, 2005). No entanto, esta *cue* apenas é funcional para frequências acima dos 1.5 KHz, como é possível observar na Figura 3, apesar de este fenómeno estar bastante relacionado com a distância física à fonte (Atkins, 2011).

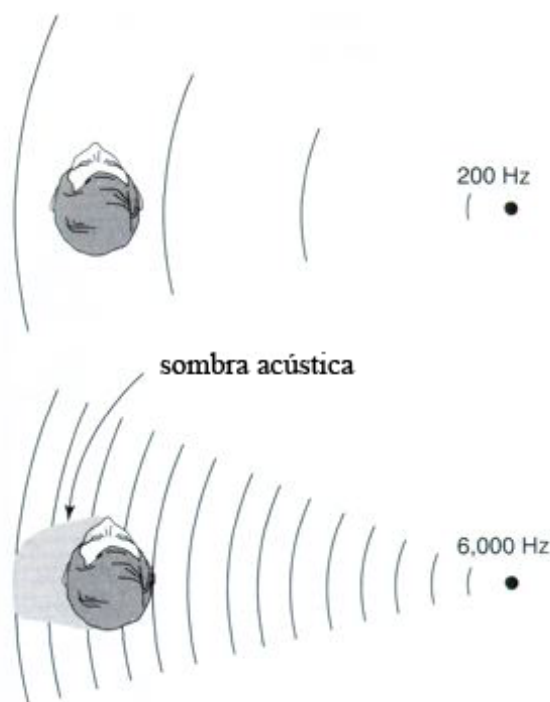


Figura 3 – Representação da “sombra acústica” (ILD) (adaptado de Heeger, 2006).

Alguns autores defendem que a própria forma da orelha serve também como um apoio à percepção elementos sonoros acima dos 7 KHz, afirmando que a sua forma permite filtrar frequências agudas, ajudando assim no processo perceptual (Moore, 1995).

2.1.2.2. HRTF e HRIR

Através dos dados apresentados anteriormente, é possível concluir que a nossa percepção de localização de fontes sonoras baseia-se na interpretação de diferenças, sendo

elas temporais, espectrais ou de ambos os tipos. Estes estudos, que contam com a inexistência de reverberação, podem ser utilizados para formular um conjunto de funções chamadas de *Head Related Transfer Functions (HRTF)*, que podem ser construídas com base em captações de impulsos sonoros chamados de *Head Related Impulse Response (HRIR)*, que têm como principal objectivo captar as *spacial cues* de um determinado indivíduo. Na Figura 4, é possível perceber que, através da análise das respostas obtidas de dois microfones situados nos ouvidos, foi criada uma linha de resposta de frequências oriundas da posição testada. Através de um leque de ensaios desta natureza é, então, possível fazer uma construção da percepção espacial sonora de um qualquer indivíduo (Rumsey, 2002; Wang *et al.*, 2005).

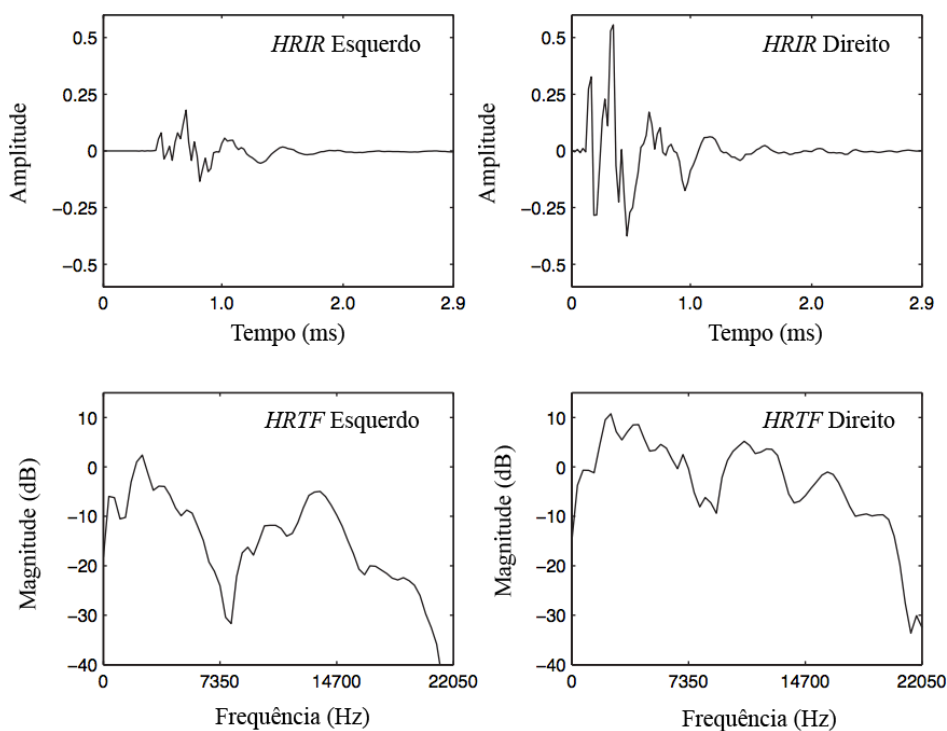


Figura 4 – Exemplo de gráficos HRIR e HRTF (adaptado de Wang *et al.*, 2005).

Através da simples observação do exemplo apresentado anteriormente é possível fazer uma estimativa da posição da fonte sonora relativamente ao ouvinte. A diferença entre o tempo de chegada do impulso aos diferentes ouvidos permite-nos concluir que a fonte se situava mais próxima do ouvido direito. Pode também perceber-se pela diferença de intensidade dos impulsos nos dois ouvidos, em conjunto com a resposta de frequências, que apresenta uma forte atenuação de frequências maioritariamente na banda das agudas e médias agudas, que a fonte se encontrava diretamente à direita do ouvinte.

2.1.2.3. Dimensões do campo sonoro

De acordo com Carlile (2011) é possível definir o espaço perceptual auditivo em três dimensões. A primeira dimensão é a de direção, onde estão compreendidas as capacidades de detectar uma fonte sonora localizada num qualquer ponto do plano horizontal e no plano mediano. A segunda dimensão é a de distância entre a fonte emissora e a cabeça (receptor). E, por fim, a dimensão do espaço acústico, que tem conta todas as absorções e reflexões nele existentes (Moore, 1995; Blauert, 1997).

A localização sonora horizontal (correspondente ao ângulo de azimute) é feita através dos fenómenos psicoacústicos de percepção das *ITD* e *ILD*. No caso da localização vertical (que equivale ao ângulo de elevação), as *cues* utilizadas para esta percepção baseiam-se principalmente no mascaramento de frequências causado pelas orelhas (Gasteiger, 2010). No entanto, há outros factores que intervêm, apesar de o fazerem em muito menor escala, como reflexões ao nível do tronco do ouvinte (Moore, 1995).

Os erros de localização provenientes destes indicadores sonoros prendem-se com a existência de valores de *ITD* ambíguos em diferentes pontos do espaço sonoro, pelo que é possível traçar um cone com centro no eixo interaural (o eixo que atravessa ambos os ouvidos) onde, em qualquer ponto deste, o valor de *ITD* é constante, como é possível observar na Figura 5. Em casos normais, a informação recolhida através das *ILD* serviria como indicador perceptual. No entanto, em situações onde a informação espectral do sinal sonoro não permite uma percepção das *ILD*, surge o chamado *cone of confusion* (Pulkki, 1997). O outro tipo de erros consistem numa percepção das fontes sonoras na proximidade da sua real posição (Carlile, 2011).

Independentemente do motivo da falha perceptual, o sistema psicoacústico humano está preparado para lidar com estas imprecisões. Através da alteração da posição ou rotação da cabeça, que muitas vezes sucede inconscientemente, há mudanças ao nível dos indicadores espaciais sonoros, dependendo da posição da fonte. Estas alterações são também utilizadas como informação de modo a fazer uma localização o mais precisa possível (Pulkki, 1997).

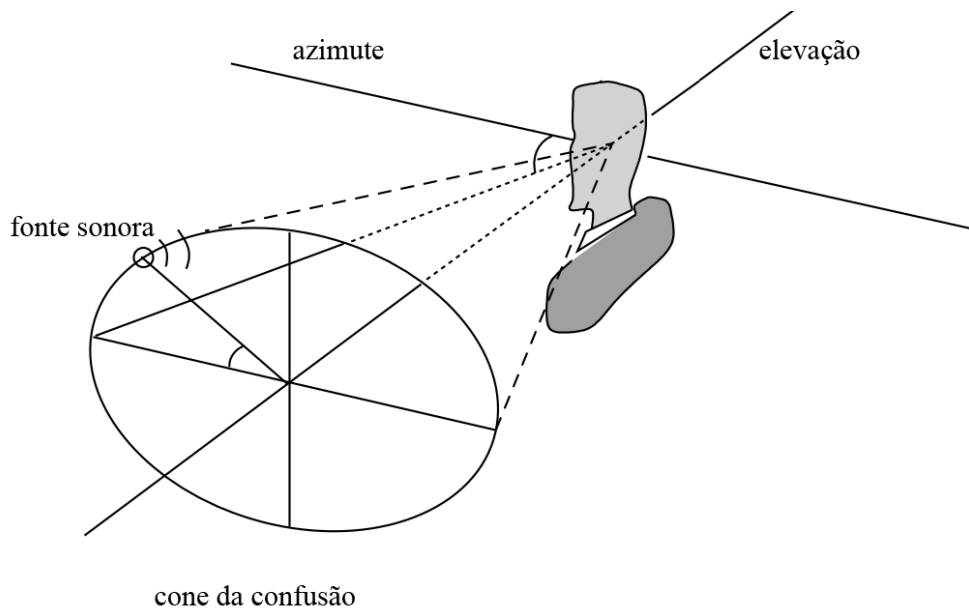


Figura 5 – Representação do cone of confusion (adaptado de Pulkki, 1997).

Segundo Rumsey (2002), Wang e Brown (2005) há um conjunto de indicadores que possibilitam uma percepção de distância da fonte sonora ao ouvinte. O primeiro a considerar será a intensidade da fonte sonora. Esta decresce 6 dB com cada duplicar de distância (num ambiente livre, sem reflexões), pelo que constitui um importante factor na interpretação desta dimensão (Fonseca, 2007).

A atenuação atmosférica consiste na absorção de frequências agudas pelo ar, a qual também é um indicador de distância, pois há uma perda de cerca de 1.6 dB nas frequências acima dos 4 KHz por cada duplicar de distância. No entanto, este indicador só será considerado, se esta perda for notável, o que implica que haja uma familiaridade prévia com o som emitido, no caso de haver um constricção ao nível da distância à fonte.

Outro indicador é o *direct-to-reverberant energy ratio*, um rácio que diminui com o aumento da distância da fonte sonora. Para além de ser influenciado pela perda de intensidade da onda sonora direta (como foi visto no primeiro indicador), é um dos elementos de cálculo deste rácio, conta ainda com as características acústicas de um espaço fechado. Este indicador, no entanto, não se baseia nas alterações das suas características e na distinção perceptual destas, ao contrário dos anteriormente expostos.

Finalmente, o último indicador está relacionado com os valores de *ILD*. Com a proximidade da fonte relativamente ao ouvinte, a diferença do espectro de frequências aumenta drasticamente, mesmo com baixas frequências (Moore, 1995).

Em espaços fechados, a própria natureza das características reverberantes do espaço permitem transmitir uma sensação de largura da fonte, estando relacionado com as reflexões laterais recebidas e com a intensidade das frequências graves do sinal.

Um último aspecto, que permite uma percepção do ambiente, é o chamado *Listener Envelopment*. Este fenómeno está ligado ao campo sonoro reverberante em geral, acentuando-se com reflexões mais tardias, mas com níveis de intensidade elevados (Carlile, 2011).

2.2. Técnicas de Espacialização Sonora Interativa

Segundo Malham (1998), é possível tomar três abordagens distintas no que toca à espacialização sonora. Esta categorização é feita não só relativamente ao conjunto de materiais necessários para a montagem do sistema em causa (*i.e.* altifalantes, auscultadores), o que por si só acarreta um conjunto de vantagens e desvantagens, mas principalmente em relação aos fundamentos por de trás das diferentes técnicas.

O primeiro método de abordagem de espacialização sonora envolve uma simulação da forma como ouvimos os vários sons que nos rodeiam. Na simulação das características psicoacústicas humanas, é enviado som diretamente para os ouvidos (normalmente através do uso de auscultadores) de modo a produzir uma sensação de espaço sonoro no ouvinte.

A abordagem seguinte compreende a utilização de sistemas de colunas, estando em regra associada à reprodução de um campo sonoro previamente “composto”, que pressupõe que haja uma interpretação correta por parte do ouvinte (ao contrário do método anterior, onde a simulação das características psicoacústicas encarrega-se de garantir que esta interpretação seja feita corretamente).

Em último lugar, o próprio espaço de *performance* pode ser utilizado para a espacialização do som. Em contraste com os métodos anteriores, onde a espacialização é processada antes da sua reprodução, nesta abordagem a espacialização é feita pelo espaço acústico. Através da utilização de *loudspeaker orchestras* (Deruty, 2012) e da utilização de mesas de mistura é possível o controlo do envio de sinal para as diferentes colunas presentes no espaço.

Faria (2005) sugere a diferenciação entre os sistemas de espacialização sonora com base no seu objectivo, distinguindo-nos entre sistemas de espacialização ou de auralização. Os

sistemas de espacialização apenas sugerem ambiências e uma sensação de espacialidade. Por outro lado, os sistemas de auralização têm como objetivo principal a simulação de campos sonoros “autênticos”. Neste trabalho não surge esta necessidade de diferenciar estes tipos de sistemas de espacialização sonora, pelo que não é feita esta distinção. Isto permite criar um seguimento mais coerente relativamente às técnicas subjacentes a cada método, como será possível observar mais à frente.

Tendo em conta a natureza deste trabalho, irão ser analisadas as técnicas compreendidas nos dois primeiros métodos mencionados (reprodução através de técnicas binaurais e de sistemas multicanal). No entanto, antes de entrar nos detalhes mais técnicos destas abordagens, é importante perceber as implicações de cada uma. Na Tabela 2 é possível observar características associadas à utilização de cada uma, que, dependendo do contexto em que pretendem ser utilizadas podem ser consideradas como vantagens, desvantagens, ou características a desconsiderar.

Tabela 2 – Comparação entre os métodos de espacialização binaural e multicanal (Adaptado de Faria, 2005).

Binaural	Multicanal
Estabilidade do ouvinte e da audição indiferente à posição dos auscultadores no ambiente	Estabilidade da audição em região limitada no ambiente
Criticidade da posição da cabeça no campo sonoro	Robustez em relação a movimentos da cabeça na área de audição, sem perda da estabilidade da imagem (para várias técnicas)
Liberdade de movimento restrita para auscultadores com fio	Liberdade de movimento assegurada dentro do volume de audição
Baixo custo do ambiente (não necessário câmaras anecoicas ou especialmente tratadas)	Custo elevado, proporcional ao número de canais
Uso individual	Alcance simultâneo a múltiplos ouvintes
Requer uso de auscultadores	Sem necessidade de utilizar auscultadores
Campo sonoro ideal para uma pessoa	Campo sonoro ideal para uma região no ambiente
Instabilidade da medida <i>HRTF</i>	Não restrito ao uso de <i>HRTFs</i>
Requer dois canais para a sonorização final	Requer n canais para a sonorização final ($2 < n < \infty$)
Necessária a equalização dos auscultadores para excelentes resultados	Necessária a equalização do ambiente (quase anecoico) para excelentes resultados
Problemas de localização espacial, principalmente devido à confusão entre frente-trás	Confusão frente-trás reduzida ou eliminada com configurações envolventes
Formação da imagem dentro da cabeça	Formação da imagem externa (no ambiente)
Diferenças sensíveis entre ouvintes para as mesmas <i>HRTFs</i>	Experiência semelhante do campo entre ouvintes diferentes
Menor complexidade na infraestrutura	Infraestrutura complexa

O uso de um sistema de colunas com *subwoofer* permite ainda um aumento do impacto físico produzido pelo sistema, o que em certos casos pode melhorar a experiência (Faria, 2005).

É possível perceber a importância da escolha do método de espacialização relativamente ao contexto da aplicação, sendo que cada um dos sistemas acarreta limitações (*i.e.* liberdade de movimentos, estabilidade na imagem sonora, custos) que poderão colocar em causa toda a experiência sonora.

No decorrer deste capítulo serão apresentadas algumas técnicas de síntese e controlo das características espaciais sonoras.

2.2.1. Binaural (Processamento HRTF)

As técnicas binaurais baseiam-se nos princípios psicoacústicos abordados no capítulo 2.1.2.. Concluiu-se que as funções *HRTF* representam a resposta de frequências de um determinado indivíduo relativamente a um ponto específico no seu campo de escuta, tendo como perspectiva os canais auditivos.

Geralmente, estes testes contam com medições feitas de diversos pontos no espaço, situados a uma distância fixa da cabeça, representados por pares ordenados, onde as coordenadas correspondem aos ângulos de azimute e elevação respectivamente, *e.g.* $(-90^\circ, 0^\circ)$ e $(90^\circ, 0^\circ)$ referem-se às posições exatamente ao lado de cada um dos ouvidos. Desta forma, é possível obter um conjunto de *HRTFs* para vários pontos do campo de escuta dos ouvidos esquerdo e direito do ouvinte, que vão ser “*samples*” das funções tridimensionais que relacionam o valor de intensidade de cada frequência numa determinada posição espacial para ambos os ouvidos, representadas por $HRTF_L(\theta, \varphi, f)$ e $HRTF_R(\theta, \varphi, f)$, onde θ corresponde ao ângulo de azimute, φ corresponde ao ângulo de elevação e f à frequência. Estas funções podem ser interpretadas como a quantidade relativa de energia recebida por ambos ouvidos (L e R) quando recebem um tom de frequência (f) originado de uma determinada localização espacial (θ, φ) (Gardner & Martin, 1994; Cheng, 2001).

Conclui-se então que para caracterizar as *cues* associadas com uma localização espacial são necessários três dados: as medições de intensidade provenientes das *HRTFs* dos ouvidos

esquerdo e direito e as *ITD*. Através destas é então possível filtrar um som monoaural para lhe conceder características binaurais, simulando que é oriundo da posição desejada.

Na Figura 6 é apresentado um diagrama que representa um sistema de espacialização binaural. O sinal sonoro (mono) é duplicado de modo a ser processado individualmente para cada ouvido tendo como referência um conjunto de informações de *ITD* e *HRTF* que irão permitir ao sistema uma reprodução espacial com base nestes dados. Em primeiro lugar é aplicado um *delay* que irá corresponder à diferença temporal que o som demora a atingir os dois ouvidos, seguido de um filtro de convolução (que permite filtrar o som, utilizando a resposta de frequências das funções *HRTF* como base), sendo finalmente enviados para os ouvidos do ouvinte. Tanto a etapa do *delay* como do filtro encontram-se salientadas devido à sua capacidade de dinamismo para o caso da síntese espacial de fontes em movimento. No entanto, sendo que a medição dos impulsos é feita sequencialmente, para cada localização escolhida, surge uma irregularidade das medições ao nível do espaço, mas uma uniformidade relativamente ao espectro de frequências. Isto cria uma incompatibilidade com este tipo de movimentos, que requerem uma variação suave por parte da fonte sonora virtual. Deste modo, torna-se indispensável a aplicação de métodos de interpolação dos valores dos filtros *HRTF* de modo garantir a integridade deste sistema (Hacihabiboglu *et al.*, 2005).

Apesar de conceptualmente bem fundamentada, sendo que se baseia em princípios inerentes às capacidades auditivas humanas, esta técnica de espacialização nem sempre produz os resultados desejados a vários níveis (McAnally & Martin, 2007)

O primeiro conjunto de problemas está relacionado com o género de medição *HRTF*. O facto de estas serem tão características a cada indivíduo, resulta que a utilização de *HRTFs* generalizados nem sempre tenha os efeitos desejados. A própria necessidade de recorrer a mecanismos de interpolação entre os impulsos para o caso de fontes sonoras em movimentos, que nem sempre conseguem simular perfeitamente as variações destes parâmetros de acordo com a escuta humana, revela-se como um problema na criação de movimentos suaves no campo de escuta.

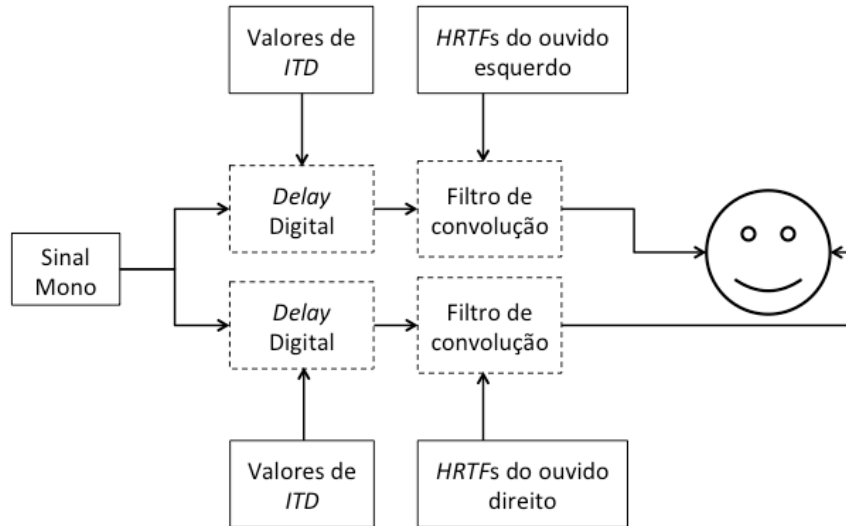


Figura 6 – Diagrama de um sistema de síntese espacial com base nos princípios binaurais (adaptado de Cheng, 2001).

O outro conjunto de problemas nasce da simples incapacidade destas ferramentas simularem de forma perfeita os mecanismos de localização espacial, que dependem de outros factores, como as reflexões do espaço ambiente. Mais especificamente o facto de os sons espacializados no plano mediano, com 0° de azimuth, soarem como se se originassem de dentro da cabeça do ouvinte, ou de ser difícil a síntese de sons com elevação superior a 0° . Há também uma dificuldade em distinguir se certos sons se situam à frente ou atrás do ouvinte (confusão frente-trás).

O facto de esta técnica ter necessidades reduzidas relativamente a materiais para escuta (apenas exige o uso de auscultadores) permite um acesso fácil para a criação de aplicações ou sistemas que envolvam uma componente espacial de áudio. É possível implementar ou recorrer a ferramentas grátis para este tipo de espacialização sonora, para a criação de aplicações interativas ou composição musical (“Binaural tools”, 2009).

2.2.2. Amplitude Panning

A *Amplitude Panning* é a técnica de espacialização sonora mais utilizada, podendo ser formulada da seguinte forma (Pulkki, 2001):

$$x_i(t) = g_i * x(t), i = 1, \dots, N$$

Tendo $x_i(t)$ como o sinal sonoro a ser aplicado à coluna i , g_i o factor de ganho do canal correspondente, t o factor de tempo e N sendo o número de colunas presentes no sistema.

O seu funcionamento baseia-se no princípio de que é possível atribuir uma sensação de direcionalidade aos sons reproduzidos (independentemente do número de colunas utilizado), enviando o mesmo sinal sonoro para todas as colunas, mas concedendo-lhes valores de intensidade diferentes, que irão controlar a direção desta fonte virtual (Rumsey, 2002).

Este conceito foi introduzido em 1931 por Blumlein (1933), relativamente a sistemas estereofónicos, afirmando que se as intensidades dos dois canais forem controladas corretamente, é possível produzir diferenças de amplitude e fase que se assemelham àquelas resultantes de fontes naturais, dando a sensação de imagens virtuais em qualquer posição entre os altifalantes esquerdo e direito.

Rumsey (2002) afirma que uma configuração tradicional estereofónica é o *setup* mais comum para efeitos de som espacial. Esta compreende uma abertura de 60° entre as duas colunas, tendo o ouvinte como referência, como é exemplificado na Figura 7.

Ele acrescenta ainda que há um conjunto de limitações associadas a esta técnica, que não possui qualquer tipo de processamento psicoacústico, sendo limitada na sua capacidade espacial. A própria necessidade de manter as colunas num ângulo de abertura de 60° , com o objectivo de produzir o maior grau de abertura possível com apenas duas colunas sem criar um “buraco” sonoro no meio, restringe o campo de audição a esse espaço entre as elas. Estas condições são aceitáveis para a maioria dos ouvintes, que têm necessidade de se manter focados num ponto à sua frente (*e.g.* televisão). No entanto, a inexistência de uma coluna central levanta também alguns problemas, como o facto de os sons colocados ao centro (*phantom*) sofrerem alterações ao nível das frequências médias (dos 500 Hz até aos 1000 Hz) e também a necessidade de garantir que a posição de escuta é a ideal para que a imagem sonora não colapse num dos lados.

Elen (2001) afirma que o sistema 5.1 *Surround-sound* é um descendente da quadrafonia, ambos baseados na técnica de *amplitude panning*, sendo, portanto, extensões da estereofonia. Sugere também que as próprias configurações, pela abertura entre as colunas, não oferecem uma localização espacial ideal, pois os ângulos de separação dos altifalantes criam zonas de muito baixa qualidade espacial, como é possível observar na Figura 8 e Figura 9.

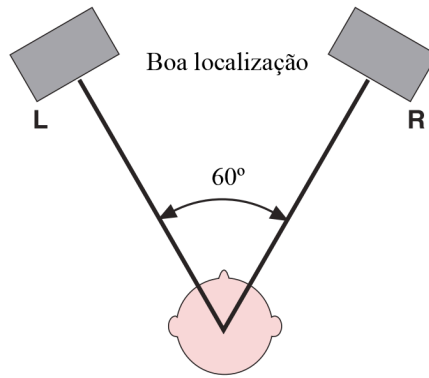


Figura 7 – Configuração estereofônica (adaptado de Elen 2001).

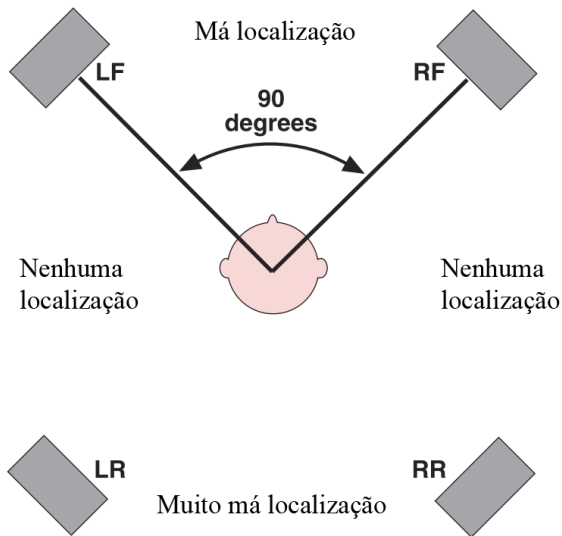


Figura 8 – Configuração quadrafônica (adaptado Elen 2001).

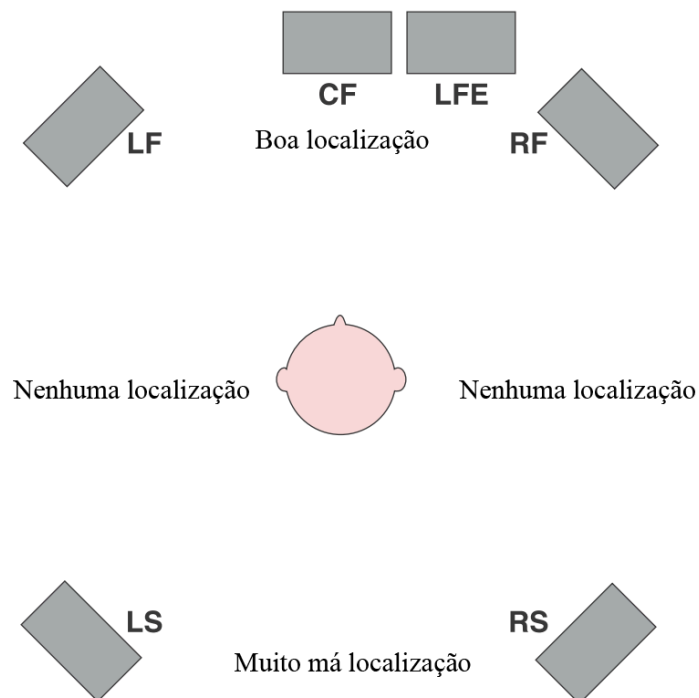


Figura 9 – Configuração 5.1 Surround-sound (adaptado de Elen 2001).

Seguidamente serão apresentadas técnicas de espacialização sonora que têm como base os conceitos inerentes à *amplitude panning*.

2.2.2.1. Distance-based Amplitude Panning

A técnica *Distance-based Amplitude Panning (DBAP)* (Lossius, *et al.* 2009) nasceu pela necessidade de dar resposta a situações onde não é possível (ou desejável) utilizar configurações tradicionais de colunas com o seu habitual *sweet spot*. Mais especificamente em casos de concerto, peças de teatro ou instalações artísticas. A *DBAP* compreende qualquer tipo de posicionamento por parte dos altifalantes a utilizar, podendo estar situados em qualquer ponto do espaço, sendo necessário ter o registo da posição relativa de cada uma das colunas em coordenadas cartesianas (x, y, z) , ou caso apenas sejam necessárias duas dimensões, assumindo que não é tido em conta o eixo z , (x, y) .

É possível perceber o modo de funcionamento desta técnica através do diagrama apresentado na Figura 10.

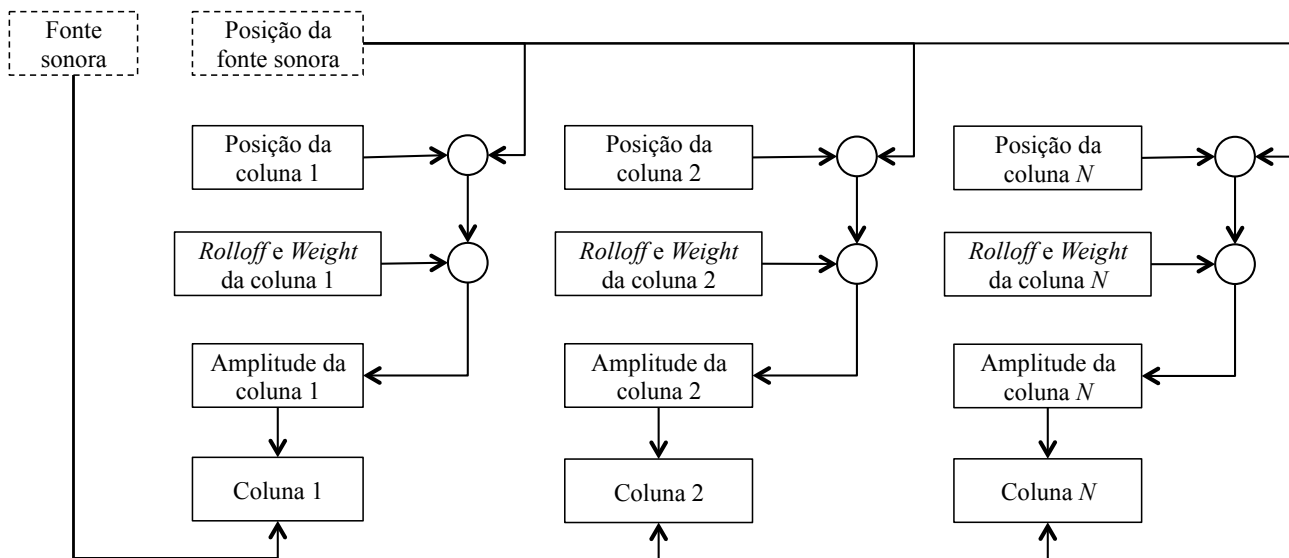


Figura 10 – Representação da técnica de espacialização DBAP.

O valor de amplitude de cada coluna é obtido a partir da distância entre a posição da fonte sonora a espacializar (x_s, y_s, z_s) e posição da própria coluna (x_i, y_i, z_i) , através do cálculo de distância entre dois pontos. Para isso é possível utilizar a seguinte fórmula:

$$D_i = \sqrt{(x_i - x_s)^2 + (y_i - y_s)^2 + (z_i - z_s)^2}$$

Sendo que em aplicações reais as coordenadas do espaço dificilmente teriam uma relação direta com os valores de amplitude das colunas, pois seria necessário escalar os valores obtidos para que os componentes pudessem ser ligados diretamente.

Seguidamente são aplicados os factores *rolloff* e *weight* de cada coluna. O *rolloff* controla a forma como o som reage ao afastar e aproximar de cada coluna (aumentar e diminuir de intensidade). Isto permite definir a quantidade de sobreposição dos sons nos vários altifalantes disponíveis, podendo ter finalidades técnicas (calibração do espaço) ou criativas (definindo sons que irão reagir como ambiências, tomando conta de uma zona alargada do espaço). O *weight* permite atribuir factores de ganho a determinadas colunas, de modo a definir limites de intensidade para certos canais, dando a possibilidade de restringir uma fonte sonora a um conjunto limitado de altifalantes, entre outras abordagens artísticas.

Esta técnica, recorrendo a métodos matemáticos simples, é uma solução básica e de fácil implementação e acesso (Jamoma, s.d.) para uma espacialização com um *setups* irregulares de colunas.

2.2.2.2. *Vector base Amplitude Panning*

A *Vector base Amplitude Panning (VBAP)* (Pulkki, 1997; Pulkki, 2000; Pulkki, 2001) surgiu da mesma premissa que a *Distance-based Amplitude Panning*: a reprodução de ambientes sonoros com características espaciais, sem haver uma dependência a configurações tradicionais de altifalantes. No entanto, os conceitos por detrás desta técnica assemelham-se mais aos do *stereo* tradicional.

Numa configuração densa de altifalantes num plano bidimensional (sem elevação), é possível controlar a direção de um sinal sonoro de modo a que este origine de qualquer ângulo de azimute do campo de escuta. Para isto, é possível recorrer à técnica de *amplitude panning* tradicional, aplicando o sinal sonoro a duas colunas de cada vez. A *VBAP* estende este conceito para um sistema tridimensional (com elevação), também chamado de perifonia (Gerzon, 1973), juntando as colunas em grupos de três, de modo a formar triângulos, como apresentado na Figura 11. Na eventualidade de configurações com mais de três colunas, conjuntos de triângulos (não sobrepostos uns aos outros) são formados (Figura 12). O sistema usa, então, o triângulo necessário para a espacialização desejada.

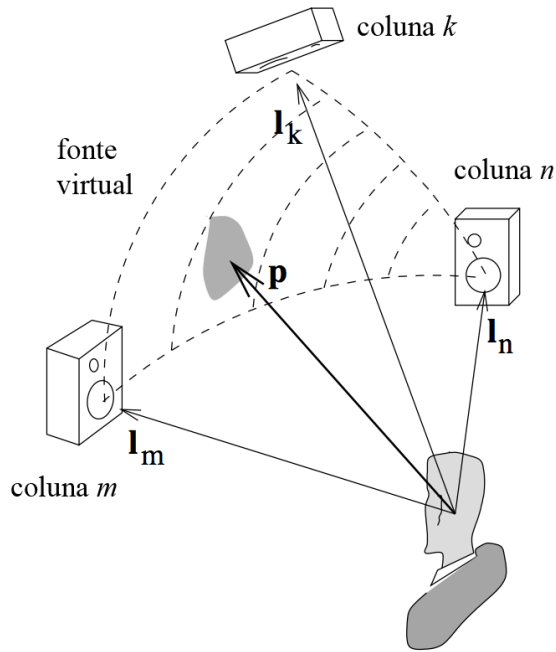


Figura 11 – Triângulo de altifalantes para a criação de fontes sonoras através do sistema VBAP (adaptado de Pulkki, 2001).

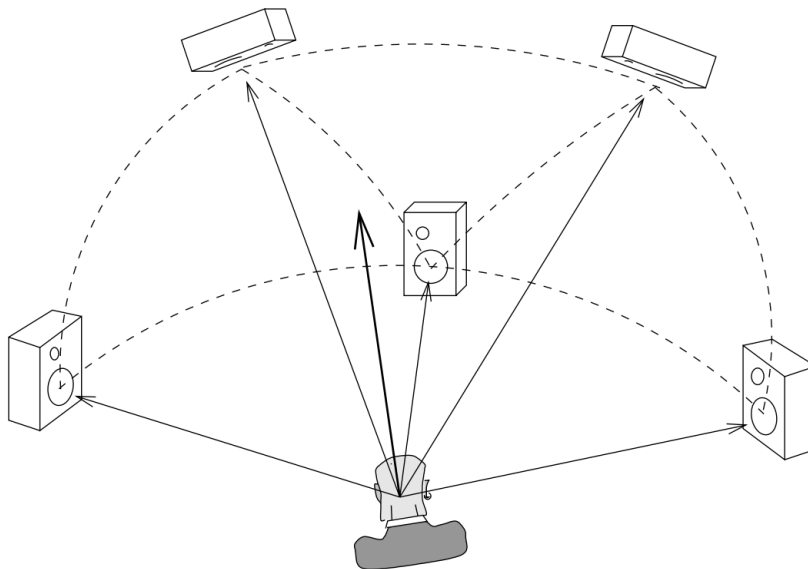


Figura 12 – Formação de triângulos para cinco altifalantes através do sistema VBAP (Pulkki, 2001).

Como o nome sugere, esta técnica envolve o uso de vetores para formular a configuração de altifalantes. Na Figura 11, os vetores l_k , l_m e l_n definem a direção do ponto de escuta para cada um dos altifalantes k , m e n , respectivamente. O vector p é relativo à fonte sonora, que se pode definir como a combinação dos vetores das três colunas, tendo g_k , g_m e g_n como factores de intensidade associados a cada altifalante:

$$p = g_k l_k + g_m l_m + g_n l_n$$

Erros a nível de localização de fontes sonoras estão relacionados com o tamanho da área da região activa (triângulo de colunas). Deste modo, para maximizar o nível de precisão é necessário aumentar o número de altifalantes nessa mesma área.

2.2.3. *Ambisonics*

A técnica *Ambisonics* destaca-se como tendo sido a primeira abordagem deste conjunto de técnicas a focar-se na reprodução de ambientes sonoros completos (Gerzon, 1973), em contraste com as anteriores mencionadas que apenas contam com a simulação de fontes “fantasma”. Na *Ambisonics*, ao contrário da *DBAP* e *VBAP*, na fase de codificação do sinal sonoro, a posição dos altifalantes é desconhecida e desconsiderada. É feita a síntese espacial de modo a que todos os altifalantes “cooperem” no sentido de criar não só uma fonte virtual, mas todo o ambiente sonoro que ela produz (não tendo em conta as reflexões do ambiente).

A *Ambisonics* nasceu como uma técnica de microfonia, tendo como objectivo fazer uma extensão da técnica de gravação *stereo* de Blumlein (1933) para um campo tridimensional, através da utilização de um microfone que incorporava um microfone omnidireccional em conjugação com 3 microfones “figura-oito” e a sua posterior reprodução, conhecido como *Soundfield microphone* (Farrar, 1979).

Na primeira ordem (a forma de *Ambisonics* mais simples) a informação sonora é codificada nos canais W , X , Y e Z , que constituem o chamado *B-format*, o mais utilizado dos formatos *Ambisonic*, representado na Figura 13. O canal W carrega a componente mono do sinal sonoro, não contendo qualquer informação relativa à direcção deste (corresponde ao output de um microfone omnidireccional). Por fim, os canais X , Y e Z possuem as informações de direcção relativas a cada uma das dimensões respectivas (correspondem aos *outpus* de três microfones “figura-oito”).

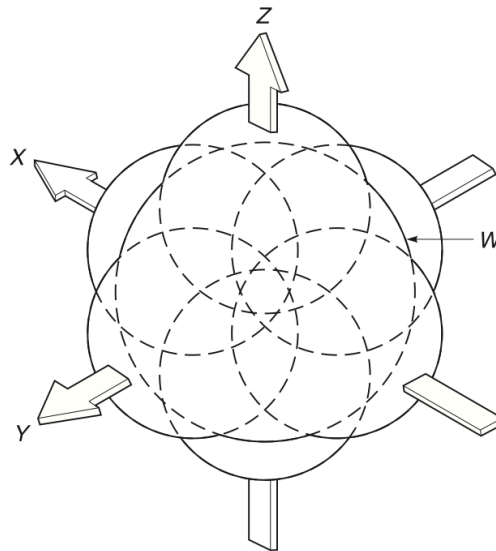


Figura 13 – Representação das componentes W, X, Y e Z do B-format (Rumsey, 2002).

A decomposição do campo sonoro está relacionada com o nível da ordem do sistema, que permite uma maior ou menor representação do número de dimensões do som reproduzido, dependendo da ordem utilizada. Por sua vez, com o aumentar de ordem, torna-se necessário o uso de mais canais para poder descrever corretamente cada uma destas dimensões. Esta decomposição tem como base o conceito de esferas harmônicas (Bamford, 1995).

Apesar de ter surgido como uma sistema de microfonia, é possível proceder à síntese espacial de um sinal mono com esta técnica, através do uso das equações de codificação *Ambisonics*. Thomas (2007) apresenta de uma forma condensada o número de canais necessários para a codificação *Ambisonics* até à 3ª ordem, em conjunto com as respectivas equações de codificação *B-format* (Tabela 3).

Na Tabela 3 são apresentadas as equações relativas às ordens 1 a 3. Estas contam com os canais das ordens anteriores, por exemplo, a codificação de um sinal *Ambisonics* de ordem 2, a qual iria necessitar dos canais *W, X, Y, Z, R, S, T, U, V*. Relativamente às equações de codificação, o sistema *Ambisonics* utiliza coordenadas esféricas para uma orientação espacial, logo θ irá corresponder ao ângulo de azimute e ϕ ao ângulo de elevação. Nesta fase do processo, assume-se que a distância das fontes ao centro é estática e corresponde ao valor de 1. O ajuste da distância é feito através de diminuições ou aumentos do sinal. Por fim, *S* corresponde ao sinal sonoro.

Tabela 3 – Equações de codificação Ambisonics até à 3ª ordem (adaptado de Thomas, 2007).

Ordem (m)	Canal Ambisonics	Equação de Codificação
0	W	$1/\sqrt{2} \cdot S$
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	$\sin 3\phi \cdot \cos^3 \theta \cdot S$

É possível tirar partido do facto de que o *Ambisonics* funciona por etapas de codificação e descodificação com base em equações através da manipulação do campo sonoro codificado. Para isso é apenas necessário proceder a algumas alterações nas equações de codificação, o que permite criar rotações nos diversos eixos (Malham, 1998).

O sinal codificado não alimenta diretamente nenhum altifalante. No entanto, possui toda a informação relativa ao ambiente sonoro captado ou espacialmente sintetizado. Como foi referido anteriormente, isto salienta o facto de que a configuração de colunas é completamente independente do processo de espacialização do som em si. No entanto, o *setup* de altifalantes está relacionado com o processo de descodificação do sinal (Hollerweger, 2008).

Tal como no DBAP e VBAP, é possível utilizar configurações irregulares de altifalantes. No entanto, vários autores aconselham o uso de distribuições de colunas com espaçamentos iguais e em círculo em volta do centro de escuta, de modo a otimizar o grau de desempenho do sistema, maximizando a precisão espacial (Zotter *et al.*, 2009; Zotter *et al.*, 2011).

O processo de descodificação envolve a soma de todos os canais codificados, que são enviados para cada uma das colunas (Gerzon, 1979; Gerzon, 1977). Dependendo da posição

desta, é-lhe atribuído um coeficiente de ganho (*weight*). Há diferentes matrizes de coeficientes que poderão ser aconselhadas para utilizações diferentes (Stitt *et al.*, 2013).

Mais recentemente, o conceito da técnica foi estendido para o chamado *Higher Order Ambisonics* (Daniel *et al.*, 2003) através do aumento de ordem, que apesar das vantagens óbvias ao nível das potencialidades de espacialização sonora, como visto anteriormente, tem requisitos associados à sua utilização. De modo a tirar partido do aumentar de ordem, é necessário aumentar o número de altifalantes utilizados. Para reproduções nas três dimensões do espaço, é possível obter o número de colunas mínimo necessário, de acordo com Hollerweger (2008), pela seguinte equação:

$$N = (m + 1)^2$$

onde N é o número mínimo de altifalantes necessário e m é o número da ordem *Ambisonics* utilizada. No entanto, se for pretendida uma reprodução em apenas duas dimensões, a fórmula de cálculo é a seguinte:

$$N = 2 * m + 1$$

Apesar de não se considerar uma técnica de *Amplitude Panning*, tendo em conta todos os seus fundamentos teóricos, ao aplicar a *Ambisonics* a uma fonte mono, o processamento efectuado ao sinal enviado para os diferentes altifalantes passa unicamente pela alteração ao nível de amplitude dos vários canais, pelo que alguns autores (*e.g.* Pulkki, 2001) consideram que ao ser utilizada para síntese espacial, a *Ambisonics* é de facto uma técnica de *Amplitude Panning*.

2.2.4. Wave Field Synthesis

A *Wave Field Synthesis* baseia-se na ideia de que, conhecendo os valores de pressão sonora e a sua respectiva velocidade dentro de qualquer ponto de um espaço isento de fontes sonoras, é possível fazer a síntese espacial de qualquer ambiente sonoro. Apesar de Steinberg e Snow (1934b) terem feito menções a este tipo de espacialização muito antes do seu surgimento, o conceito foi estabelecido por Berkhout (1993) em 1988, tendo como base teórica o princípio de Huygens. Para uma leitura mais detalhada dos fundamentos

matemáticos inerentes a esta técnica é possível utilizar como referência o trabalho publicado por Berkhout (1993).

A utilização da *Wave Field Synthesis* permite resolver problemas de interferência resultantes dos padrões de direccionalidade dos altifalantes, que resultam na restrição a uma área de audição limitada ou *sweet spot* (Devries, 1999), aumentando, assim, o campo de escuta (Figura 14 e Figura 15), o que por sua vez permite uma escuta por um maior número de ouvintes. No entanto, esta abordagem acarreta grandes custos a nível de processamento, tendo em conta que envolve a modelação física da propagação das ondas sonoras através de grandes matrizes de altifalantes que se comportam como fontes sonoras secundárias. Isto permite a localização espacial de fontes sonoras dentro ou fora do limite definido pelos altifalantes, como representado na Figura 16.

Para que este género de reprodução sonora seja possível são colocadas restrições no posicionamento das colunas, que por sua vez irão limitar a gama de frequências que a *Wave Field Synthesis* consegue sintetizar. Vogel (1993) aborda este tema de forma extensa, concluindo que as colunas devem encontrar-se à distância de meio comprimento de onda entre cada uma, o que resulta no uso de quantidades muito grandes de altifalantes para que seja possível reproduzir frequências agudas corretamente.

Apesar das excelentes capacidades a nível de síntese espacial, a *Wave Field Synthesis* revela-se impraticável para a maioria das situações reais, pelos seus custos a nível computacional e de altifalantes.

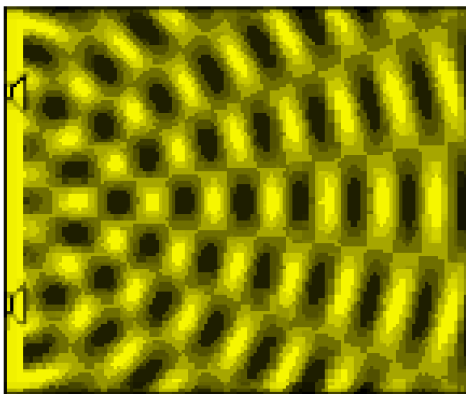


Figura 14 – Representação do padrão de interferência na reprodução com dois altifalantes (De Vries & Boone, 1999).

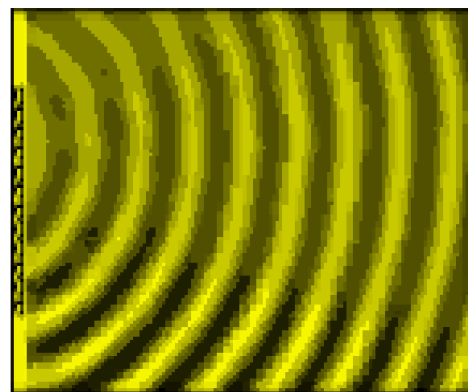


Figura 15 – Representação da reprodução através de um array de altifalantes (De Vries & Boone, 1999).

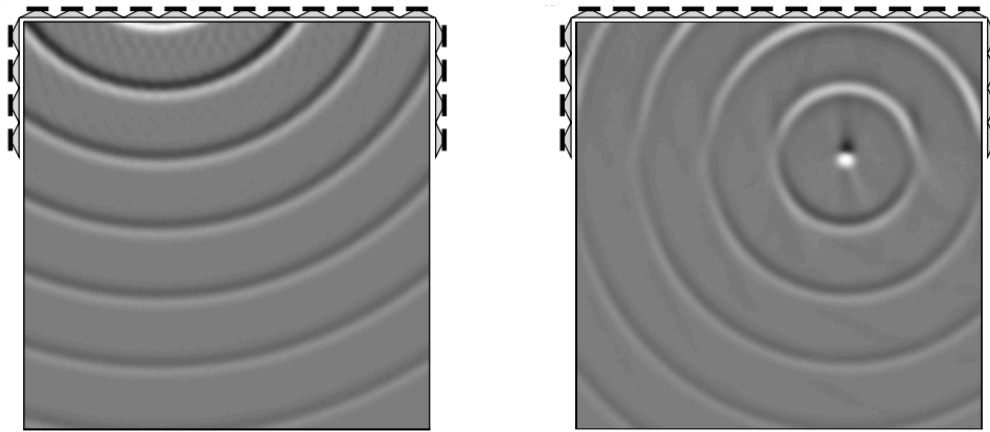


Figura 16 – Exemplo da reconstrução do campo sonoro utilizando a WFS (De Vries & Boone, 1999).

2.2.5. *Ambiophonics*

A *Ambiophonics* (Glasgal, 2003; Glasgal, s.d.) permite a reprodução espacial de ambientes sonoros recorrendo à conjugação de duas técnicas. A primeira consiste na reprodução por altifalantes *stereo* com cancelamento de interferência (*cross-talk*), chamado de *stereo dipole* (Kirkeby *et al.*, 1998), em que um afastamento de 10 a 30 graus permite a construção de imagens sonoras com graus de abertura de quase 180°. A segunda técnica baseia-se na reconstrução da ambiência de um determinado espaço utilizando mecanismos de convolução de impulsos de resposta.

É uma técnica que envolve gravação e reprodução multicanal tendo como base conceitos psicoacústicos, de modo a produzir ambientes e fontes sonoras virtuais. Este formato, utilizando apenas dois canais de áudio consegue níveis de desempenho na síntese das características espaciais e de localização superiores ao formato *stereo* tradicional.

Um *setup Ambiophonics* completo compreende o uso de duas camadas de altifalantes. A primeira é a de colunas na disposição *stereo dipole*, que pode ser estendida ao uso de altifalantes traseiros na mesma disposição dos frontais, o que permite aumentar o grau de abertura da imagem sonora. Estes recebem o som proveniente dos dois canais contidos no formato *Ambiophonics*, diretamente. A segunda camada envolve o uso de pelo menos oito colunas que são colocadas em redor do ouvinte e têm como objectivo transmitir uma sonoridade de ambiência (*surround*).

A Figura 17 representa a primeira parte do sistema, onde é feita a gravação do sinal sonoro a espacializar. Para isso, são colocados dois microfones no local dos ouvidos,

semelhante ao descrito no capítulo 2.1.2. para a captação dos impulsos *HRIR*, não havendo preferência relativamente ao padrão de direccionalidade dos microfones utilizados. Os mecanismos de cancelamento de *cross-talk* envolvem a aplicação de filtros ao sinal.

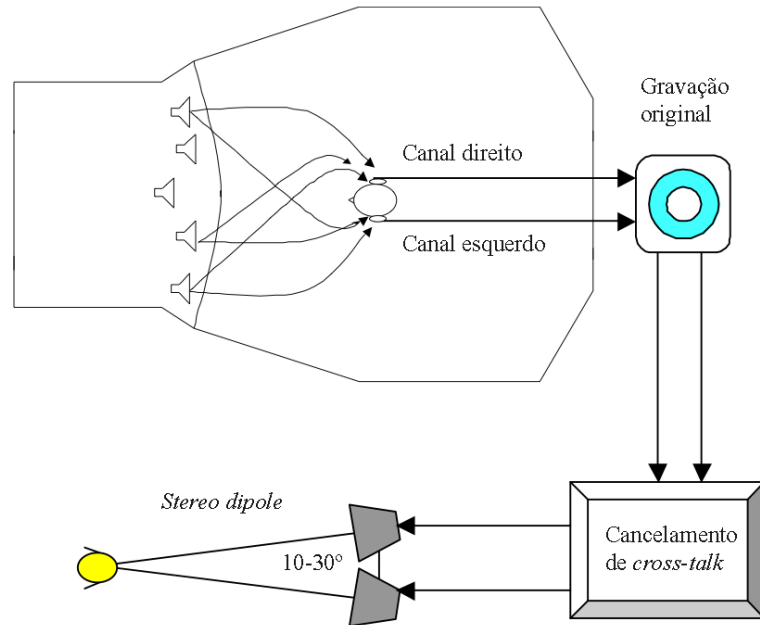


Figura 17 – Representação da captação e reprodução de sinal sonoro num sistema stereo dipole (adaptado de Farina et al., 2001).

A segunda parte (Figura 18) envolve a captação de impulsos relativos às propriedades de reverberação do espaço tridimensionalmente, sendo, para isto utilizado o formato *B-format* do sistema *Ambisonics* (não confundir com *Amphiophonics*) para captação e descodificação do sinal. Estes impulsos, através de mecanismos de convolução em conjunto com as gravações obtidas na fase anterior, irão definir as características de reverberação para cada um dos altifalantes *surround*, que serão dispostos de acordo com o sistema *Ambisonics*.

Tendo em conta a área de escuta reduzida do *stereo dipole*, o *Amphiophonics* não permite a audição por muitos utilizadores simultaneamente.

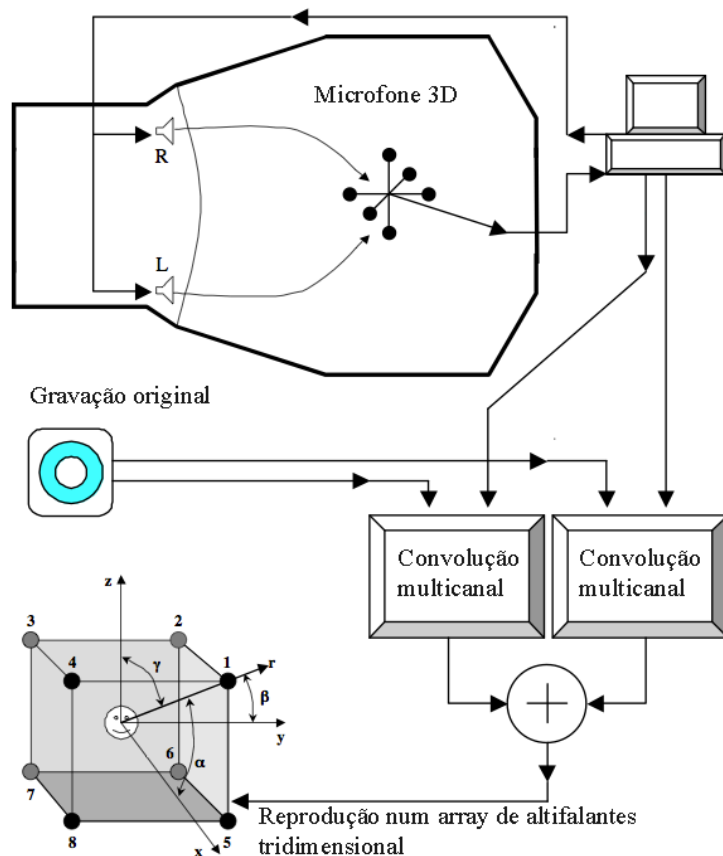


Figura 18 – Representação da captação dos impulsos de resposta, convolução multicanal e reprodução da camada surround Ambiophonics (adaptado de Farina et al., 2001).

2.2.6. ViMiC

A *ViMiC* (*Virtual Microphone Control*) é uma ferramenta de espacialização sonora tridimensional em tempo real, tendo sido desenhada para utilização em situações de concerto ou instalações imersivas. Foi desenvolvida originalmente para o software *Pure Data*, tendo sido adaptada mais recentemente para o *Max/MSP*. (Peters et al., 2008)

Em termos de *software*, baseia-se na utilização de fontes virtuais e microfones virtuais, cuja direccionalidade de captação é ajustável. Os microfones no espaço virtual irão corresponder aos canais de saída para o número de colunas desejado. O *ViMiC* incorpora um sistema de cálculo de ganho e *delay* entre fontes e microfones virtuais através da distância entre cada, uma simulação das reflexões e absorção do espaço ambiente virtual (um paralelepípedo).

Tendo em conta que é feito um processamento das simulações do comportamento físico do som, o *ViMiC* torna-se intensivo ao nível de processamento de computador.

2.2.7. Comparação entre as técnicas

O conceito de holofonia (do inglês *holophony*) foi introduzido por Maurice Jessel em 1973 (Jessel, 1973), sendo descrito como o equivalente acústico da holografia. A holofonia tem como objectivo a construção (ou reconstrução) total de um ambiente sonoro, em vez de uma reprodução local de fontes “fantasma” (Jessel, 1973). Do estudo feito no decorrer deste capítulo, é possível determinar quais das técnicas de espacialização que podem ser consideradas holofónicas, nomeadamente *Ambisonics*, *Wave Field Synthesis*, *Ambiophonics* e *ViMiC*.

Existe um conjunto alargado de técnicas de espacialização sonora que apresentam diferentes abordagens com incidências em diversas áreas de estudo. Por exemplo, a recriação dos mecanismos psicoacústicos humanos (espacialização binaural), a modelação física do percorrer das ondas no espaço (*Wave Field Synthesis*), a aplicação de conceitos matemáticos (*Distance-based Amplitude Panning*) ou a fusão de várias destas matérias (*Ambiophonics*). Estas escolhas trazem limitações e possibilidades, pelo que a escolha do método de espacialização está dependente inteiramente da aplicação em causa. Um exemplo disso é a utilização recorrente de técnicas de *Amplitude Panning* em contextos de realidade virtual, que surge dos baixos requisitos a nível de processamento destes sistemas e da flexibilidade ao nível de configurações de altifalantes que provém do seu uso (Faria, 2005).

Na Tabela 4, adaptada de Thomas (2007), são apresentados os factores envolvidos na escolha de uma técnica de espacialização sonora. É também feita uma comparação entre a eficiência de algumas das técnicas mais comuns de reprodução espacial, sendo atribuída uma cotação, na escala de 1 a 5, a cada parâmetro apresentado, em que a nota 1 representa o pior resultado de eficiência e a nota 5 o melhor.

Para além das características técnicas de cada um destes métodos de espacialização, é importante considerar a forma como os materiais de difusão sonora utilizados se relacionam com o espaço físico do ouvinte. Cohen *et al.* (1999) apresentam uma relação entre o sistema de difusão sonora e o contexto proxémico do ouvinte. É possível recorrer a estes mecanismos de comunicação não-verbal para estimular determinadas sensações no utilizador.

Tabela 4 – Comparação entre as técnicas de espacialização sonora (adaptação de Thomas, 2007).

Técnica	Realismo	Número de utilizadores	Facilidade de implementação	Custos computacionais	Custos de equipamento
Binaural	4 Bom realismo se forem utilizadas boas <i>HRTFs</i>	1 Suporta apenas um utilizador	4 Fácil no caso do uso de <i>HRTFs</i> gerais	3 No caso da síntese, exige um processamento razoável	5 Necessita apenas de auscultadores
<i>VBAP</i>	3 Permite a reprodução de campos tridimensionais	4 Suporta vários utilizadores em simultâneo	4 Simples implementação computacional	4 Não requer grande processamento	3 Depende do número de altifalantes utilizados
<i>Ambisonics</i>	5 Ótima reconstrução da onda para ordens elevadas	4 É capaz de suportar vários utilizadores simultaneamente	3 Montagem depende do número e disposição de altifalantes	3 Não requer grande processamento	3 Depende do número de altifalantes
<i>WFS</i>	5 Reconstrução da onda próxima da real	5 É capaz de suportar muitos utilizadores	1 Montagem complicada	1 Requer grande processamento para a síntese espacial	1 Requer uma matriz densa de altifalantes
<i>Ambiophonics</i>	3 Boa reconstrução para sistemas completos	2 Suporta até dois utilizadores	3 Precisa da implementação do <i>stereo dipole</i> e do <i>Ambisonics</i>	5 Baixos requisitos computacionais	4 Custo escalável

2.2.8. Interatividade na espacialização sonora

Desde que o ser humano nasce, começa a interagir com o mundo que o rodeia, e é a partir dessas interações com outros elementos que aprende determinadas regras, que resultam em associações de ação-percepção que toma por garantidas (Lemaitre *et al.*, 2009).

A própria manipulação sonora é uma ferramenta de diagnóstico natural humana. O som é um indicador temporal dos processos físicos, sincronizando os sentidos visuais e tácteis, confirmando o contacto físico com objetos e dando informações relativamente às suas propriedades físicas. No entanto, o nível de controlo que o utilizador tem dessas ações vai influenciar a forma como interpreta as respostas resultantes. A própria complexidade das

interações e o grau de controlo dos objetos por parte da interação e do utilizador vai resultar em ligações de intimidade de controlo diferentes. (Hermann & Hunt 2011)

O estudo feito por Schafer (1973) na relação entre os sons, o seu meio ambiente e a cultura em que se inserem teve um papel muito importante na interligação das áreas relacionadas com o som que envolvem uma sonificação de dados multidimensionais (Hermann & Hunt., 2011). Nestas situações, a integração de ambientes sonoros é crucial para garantir uma experiência de interação homem-computador credível (Mauney & Walker, 2004).

Hermann e Hunt (2005) definem sonificação interativa como o uso de som dentro de uma interface homem-computador, onde o sinal auditório fornece informação sobre aquilo que está a ser analisado ou sobre a interação em si, que por sua vez ajuda a aperfeiçoar ou refinar essa atividade. A sonificação interativa é caracterizada por dar ao utilizador a capacidade de controlar o processo de sonificação através da associação das ações do utilizador e da resposta sonora do sistema (Hermann & Hunt., 2005).

Na área de espacialização sonora, a componente interativa tem como objetivo suportar todas as alterações físicas presentes no espaço sonoro virtual. Para isto é importante identificar as variáveis existentes. Num espaço, a resposta acústica interpretada pelo ouvinte varia de acordo com a fonte sonora e a posição e orientação do ouvinte. Assim, um modelo de espacialização interativa deve conseguir produzir um *output* que dependa das propriedades dinâmicas da fonte, ouvinte e do ambiente (Savioja *et al.*, 1999). É possível distinguir os seguintes parâmetros interativos de espacialização sonora:

- Geometria do espaço;
- Materiais nas superfícies do espaço;
- Locais e orientações de cada fonte sonora;
- Local e orientação do ouvinte.

Sempre que há uma variação de qualquer um destes parâmetros há alterações no espaço sonoro, pelo que estas mudanças devem transmitir essas alterações. É também importante que o modo de implementação permita que essas mudanças sejam feitas de forma suave de modo a que o ouvinte não seja capaz de distinguir o tempo exato da atualização (Savioja *et al.*, 1999).

2.3. Imersão

O conceito de imersão é utilizado e discutido num conjunto alargado de áreas, no entanto, a sua definição não é completamente clara (Brown & Cairns, 2004).

Janet Murray (1997) afirma que uma narrativa condutora apresentada em qualquer canal comunicativo pode ser experienciada como uma realidade virtual graças à capacidade de abstração do cérebro humano. A esta experiência de estar submerso numa realidade que se apodera de todas as componentes do sistema perceptual dá-se o nome de imersão.

As próprias características de interatividade do canal comunicativo influenciam o estado de imersão. Dovey e Kennedy (2006) salientam a distinção entre *passive media* (e.g. filmes) e *interactive media* (e.g. videojogos), referindo que apesar de poder ocorrer o fenómeno de imersão em diferentes *medias*, apenas quando o participante tem um papel ativo, é possível que haja uma combinação de imersão no ambiente virtual em conjugação com uma concentração intensa na atividade em causa.

2.3.1. Modelos de imersão

O conceito de categorização dos tipos de experiência foi estudado por Pine e Gilmore (1999), através do uso de duas dimensões para fazer esta classificação, a participação e a conexão. A participação tem como objectivo determinar o grau de interação que o utilizador tem com a atividade, variando de passiva a ativa. Por sua vez, a conexão pretende classificar o grau de ligação físico e emocional criado, passando de absorção (prestar atenção à tarefa em causa) até imersão (envolver-se com a experiência, tornando-se física ou virtualmente parte dela). Os autores definem ainda quatro domínios da experiência que podem ser definidos através destas dimensões, apresentados na Tabela 5.

É possível perceber, desta categorização, que o género de atividades com uma participação ativa têm a potencialidade de produzir uma experiência mais envolvente, o que explica a razão pelo qual o tema da imersão nos videojogos e sistemas de realidade virtual é tão estudado (Brown & Cairns, 2004; Ermi & Mäyrä, 2005; Grimshaw, 2007).

Tabela 5 – Quatro domínios da experiência, segundo Pine e Gilmore (1999).

	Conexão	Absorção	Imersão
Participação			
Passiva		Entretenimento	Estético
Activa		Educacional	Escapista

Brown e Cairns (2004) classificam a imersão nos videojogos como o envolvimento de um utilizador com o jogo. Consideram ainda que há três patamares imersivos, interligados, que estão dependentes tanto do utilizador como do próprio jogo. A primeira fase tem o nome de *Engagement*. Esta fase inicial compreende o investimento de tempo, esforço e atenção por parte do utilizador; por sua vez o jogo deverá permitir um controlo e *feedback* que responda às necessidades do jogador. É também necessário que haja uma recompensa do tempo investido pelo utilizador. O jogador começa-se-á a sentir mais abstraído do meio que o rodeia e entra no patamar *Engrossment*. Aqui, o jogo terá de oferecer características que afectem as emoções do jogador, que tem de se submeter a essa envolvência. Por fim, a fase *Total Immersion* envolve uma empatia relativamente às personagens e ao ambiente de jogo. Os elementos do próprio jogo (*i.e.* visuais, auditivos, narrativos) encarregam-se de manter o utilizador concentrado.

Ermi e Mäyrä (2005) apresentam um modelo (*SCI-model*) para descrever a experiência imersiva em videojogos baseando-se na distinção de três dimensões heurísticas. A primeira é a imersão sensorial, provocada pelos estímulos sensoriais do jogo. A escolha do sistema de reprodução de cada elemento sensorial irá influenciar a forma como essa informação se sobrepõe aos estímulos exteriores ao jogo. Em segundo lugar está a imersão baseada em desafios, que se prende com a dificuldade dos desafios presentes no jogo (quer a nível motor como de pensamento estratégico). Finalmente, a imersão imaginativa está relacionada com as relações emocionais que os utilizadores criam com os elementos do jogo (*i.e.* personagens, história).

2.3.2. Imersão no som

Como foi visto anteriormente, o factor de interatividade numa experiência cria potencialidades a nível de imersão, permitindo ao utilizador envolver-se de forma mais íntima

com a experiência. Um concerto musical e uma simulação de treino num ambiente de realidade virtual enfatizam experiências psicoacústicas diferentes, pelo que dependendo do nível de interatividade que a experiência exige, é também imposto ao sistema dar resposta a um conjunto de critérios a nível perceptual (Matthew, 2011).

Os ambientes de realidade virtual gerados a computador, sendo interfaces tridimensionais e multimodais entre computadores e humanos, permitem uma interação em tempo real por parte do utilizador (Vorländer, 2008). Apesar de ser aceite que o realismo a nível gráfico e sonoro não é necessário para que ambientes de realidade virtual tenham a capacidade de produzir uma sensação de imersão nos seus utilizadores (McMahan, 2003), considera-se que a utopia da realidade virtual visual é a criação de um sistema com uma capacidade dinâmica que permita igualar as propriedades da luz, cor e fluidez de movimentos às do mundo real e uma resolução que exceda as capacidades do olho humano, em conjunto com uma reprodução fiel das características da visão humana (*i.e.* visão binocular, *motion parallax*, alteração da distância) (Matthew, 2011). Igualmente, a realidade virtual sonora ideal deverá exceder a gama dinâmica e de resposta de frequências humana e ser capaz de criar e igualar a complexidade da física acústica de um ambiente dinâmico (Matthew, 2011).

Atualmente, dentro dos limites impostos pelos avanços tecnológicos, a criação de ambientes sonoros consegue atingir resultados mais realistas e com uma resolução superior aos ambientes virtuais visuais (Neuhoff, 2011), o que permite o desenvolvimento de aplicações onde a componente sonora assume um papel principal (Papworth, 2010; Hugill, 2012).

Audiogames são jogos onde a componente sonora tem um papel dominante, focando-se maioritariamente nesta criação de ambientes e mundos sonoros virtuais (Friberg & Gärdenfors, 2004; Röbert & Masuch, 2005). Através dela, são transmitidas todas as informações necessárias para a orientação do utilizador no universo de jogo. A ausência, em grande parte, da componente visual destas experiências resulta num maior estímulo da imaginação do utilizador, o que permite atingir níveis mais elevados de imersão em comparação com os jogos tradicionais (Höner, 2011).

2.3.3. Exemplos de projetos imersivos

De seguida, serão apresentados alguns projetos desenhados para proporcionar uma experiência imersiva e interativa.

2.3.3.1. CAVE

CAVE (CAVE Automatic Virtual Environment) é um sistema imersivo de realidade virtual concebido na University of Illinois e demonstrado em 1992 (Cruz-Neira *et al.*, 1992). Consiste numa sala cúbica com cerca de 3 metros de lado, dentro da qual os utilizadores se situam. O sistema envolve a projeção gráfica para quatro a seis das faces do interior do cubo em conjunto com a utilização de *shutter glasses*, que permitem visão tridimensional. É utilizado um sistema multicanal para a difusão sonora para altifalantes situados nos vértices da sala. Todas as perspectivas são calculadas do ponto de vista do utilizador através de um sistema de *tracking* da posição da cabeça e orientação da cabeça (Cruz-Neira *et al.*, 1992; Cruz-Neira *et al.*, 1993). Na Figura 19 é possível ver um exemplo deste género de aplicações.

Atualmente, o sistema *CAVE* é disponibilizado comercialmente pela empresa *Mechdyne*, que disponibiliza uma versão alternativa deste sistema, a *CAVE2* (Mechdyne, s.d.).

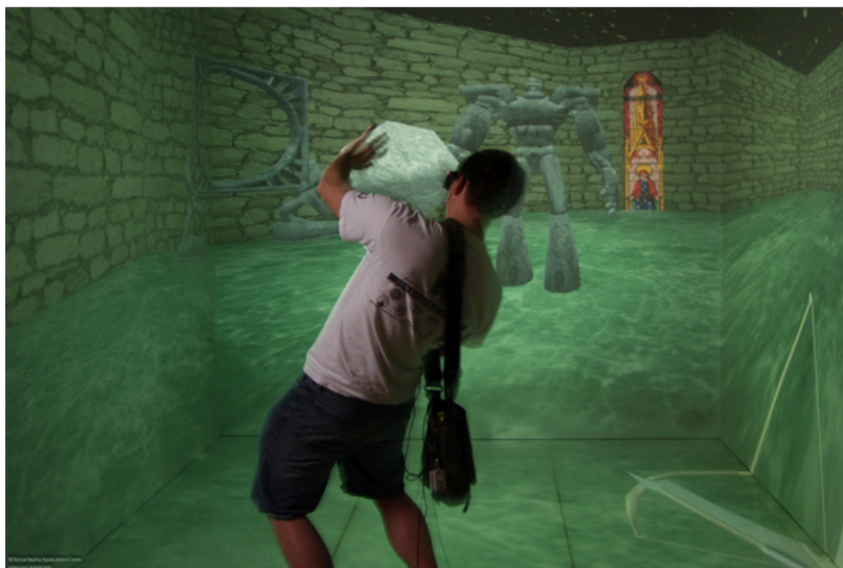


Figura 19 – Exemplo de utilização do sistema *CAVE* (Mechdyne, s.d.).

2.3.3.2. *Allosphere*

Allosphere é uma estrutura esférica com diâmetro de 10 metros situada na University of California em Santa Barbara (Figura 20) (Amatriain *et al.*, 2007). Este projeto pretende ser um palco para a apresentação de conteúdos audiovisuais imersivos gerados a computador, através de uma projeção gráfica envolvente de alta resolução e um sistema sonoro tridimensional, que conta com a utilização de centenas de altifalantes e uma grande versatilidade a nível do processamento da síntese espacial empregue, podendo recorrer a *VBAP*, *Ambisonics* e *Wave Field Synthesis*. Em conjunto são utilizados mecanismos de *tracking* dos utilizadores visam a permitir uma interação com os conteúdos (Amatriain *et al.*, 2007; Amatriain *et al.*, 2009).

AlloBrain, um projeto desenvolvido para o *Allosphere* permite aos utilizadores navegarem virtualmente pela estrutura neurofisiológica humana, captada por técnicas de ressonância magnética e mapeada gráfica e sonoramente (Figura 21) (Overholt *et al.*, 2008).

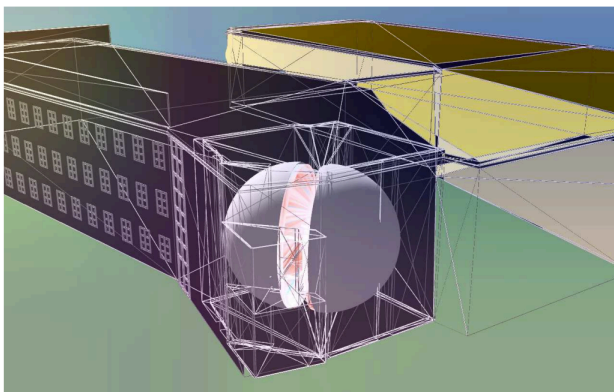


Figura 20 – Representação virtual da estrutura *Allosphere* (Amatriain *et al.*, 2007).

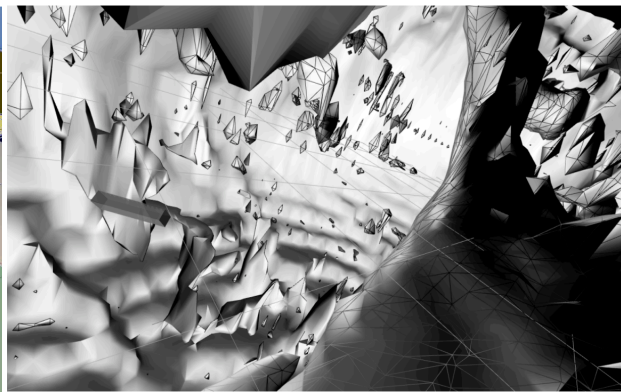


Figura 21 – Representação neurofisiológica no *AlloBrain* (Overholt *et al.*, 2008).

3. *The Sound of Horror*: Instalação Sonora Interativa e Imersiva

De modo a produzir uma instalação sonora interativa e imersiva, foi necessário proceder a um planeamento a nível conceptual e técnico, de modo a conceber uma experiência completa e de grande impacto, capaz de absorver o utilizador no ambiente virtual. Neste capítulo é feito um acompanhamento das várias fases envolvidas no projeto *The Sound of Horror*, apresentado no Laboratório de *Motion Capture* da Escola das Artes da Universidade Católica Portuguesa no dia 24 de Abril de 2013.

3.1. Escolha do projeto

Inicialmente, o projeto foi idealizado como um conjunto de pequenas instalações sonoras, desenvolvidas separadamente, com o intuito de explorar as capacidades interativas do som. Seguidamente serão apresentadas algumas das ideias que surgiram ao longo deste desenvolvimento.

Simulação sonora do *lightsaber* do *Star Wars*, semelhante ao produzido por Matt Jarvis (2009). Recorrendo a um comando *Wiimote* da *Nintendo* para o registo de movimentos, permitindo ao utilizador empunhar virtualmente a espada. Os movimentos executados com o comando iriam definir a forma como a síntese sonora e espacial ocorreria, dando um controlo total ao utilizador.

Construção de uma instalação musical através de uma guitarra equipada com um *pickup* hexafónico, inspirada no trabalho *Smeck* de Miller Puckette (2007). Este captador permite reconhecer o sinal de cada corda do instrumento separadamente, sendo possível a conversão para MIDI através do uso de conversores próprios (Axon, 2008). A ideia deste projeto era permitir aos utilizadores a composição de música ambiental em tempo real, não exigindo conhecimentos teóricos musicais, mas oferecendo uma interface familiar, com a qual poderiam interagir. A conversão do sinal para MIDI permitiria utilizar a guitarra como seis instrumentos monofónicos separados, num só, permitindo um grande nível de controlo da interação por parte do utilizador.

Desenvolvimento de uma instalação interativa que atribuía ao utilizador o papel de “Deus do Trovão”, permitindo-lhe, gestualmente, controlar a geração sonora de trovões e o controlo interativo de um ambiente tempestuoso. Esta ideia contava com a utilização de mecanismos de controlo de pressão do ar, como ventoinhas, que reagiam ao alterar do estado do ambiente sonoro, bem como sistemas de luz com o objectivo tornar a experiência multissensorial.

Criação de um jogo sonoro interativo e imersivo, do subgénero *survival horror*, onde através do uso de uma arma, à semelhança de um *first-person shooter*, o utilizador teria de disparar contra monstros que o tentavam atacar. Esta ideia não contava com a utilização de um *feedback* visual, estando o espaço de jogo num ambiente escuro para ampliar o efeito produzido pelos elementos sonoros e tácteis de modo a aumentar o grau de imersão. A ideia surgiu do projeto final realizado para a unidade curricular de Interfaces Tangíveis, onde foi desenvolvido um sistema muito pouco eficiente que permitia ao utilizador disparar contra alvos sonoros. Na Figura 22 é possível ver um exemplo da utilização deste sistema, que contava com quatro altifalantes numa disposição quadrafónica. Para a captura de movimentos, era utilizada a câmara de um comando *Wii mote* para fazer o *tracking* de *leds* infravermelhos montados numa estrutura com forma de arma, onde estava colocado um segundo comando *Wii mote*, que permitia detectar a operação de disparo da arma e os movimentos efectuados pelo utilizador. Devido às limitações de tempo, materiais e espaço, foi impossível conceber um projeto funcional a todos os níveis, sendo que o *tracking* da posição da arma não estava funcional e o sistema de espacialização sonora tinha muito pouca precisão, devido ao número de altifalantes e à técnica síntese espacial empregue.

Outras ideias menos exploradas envolviam o desenvolvimento de um jogo à semelhança do *Audiosurf* (Valve Corporation, 2008) e do *Tap Tap Revenge* (Tapulous, 2008) o uso de controladores alternativos e uma difusão sonora espacial e imersiva, e o desenvolvimento de uma instalação artística que envolvia um conjunto de pêndulos reativos ao toque ou choque uns com os outros, que produziam sonoridades etéreas com características musicais, em conjunto com ambientes sonoros dinâmicos.

Após um conjunto de reuniões com o orientador, foi decidido que seria mais proveitoso produzir apenas uma instalação, mas torná-la o mais desenvolvida e eficiente possível, de modo a aprofundar o estudo das matérias em causa e tirar melhor partido dos materiais oferecidos pela Universidade Católica Portuguesa. Deste modo, foi decidido desenvolver o projeto do jogo sonoro do género *first-person shooter*, após um estudo relativamente às várias componentes envolvidas.



Figura 22 – Fotografia do sistema de jogo sonoro elaborado para a unidade curricular de Interfaces Tangíveis.

3.2. Objectivos

Tendo em conta as características do projeto, pode considerar-se como uma instalação interativa e imersiva sob a forma de um jogo sonoro, tendo os seguintes objectivos:

- Criar um ambiente virtual imersivo através da utilização de som como elemento principal;
- Gerar fontes sonoras dinâmicas que permitam uma interação com o utilizador;
- Permitir a interação com o sistema através de métodos naturais e orgânicos.

De modo a desenvolver esta experiência interativa e imersiva torna-se necessário dar resposta a um conjunto de factores técnicos.

Em primeiro lugar, é necessário utilizar um sistema sonoro capaz de produzir uma sensação de imersão sensorial no utilizador. Tendo em conta que o elemento sonoro vai ter um papel preponderante, é indispensável o uso de um sistema de difusão sonora que tenha a capacidade de sobrepor o som do jogo a outros sons exteriores, para manter o utilizador

dentro do universo virtual da instalação. Em conjugação com este sistema, é fundamental a implementação de uma configuração de altifalantes e uma técnica de espacialização sonora que permitam maximizar a inteligibilidade espacial dos vários elementos presentes, de modo a que o utilizador consiga perceber a posição no espaço das fontes sonoras sem esforço, não sentindo falta de um apoio visual.

É necessário, dentro de determinados limites, permitir que o utilizador tenha liberdade de movimentos, para que seja possível um grau elevado de interação com a aplicação, nomeadamente, através do uso de um sistema que permita fazer o *tracking* de movimentos do utilizador e da utilização de interfaces que permitam um controlo gestual. Se bem conseguido, isto permite criar, no utilizador, a sensação de uma interação orgânica com o sistema, abstraindo-se do facto de que está realmente a comunicar com um computador. No entanto, isto apenas se torna possível num sistema de latência reduzida, onde os mecanismos de ação-reação se equiparam aos do mundo real.

De modo fazer a ligação entre estes componentes e desenhar o motor de jogo que vai inicializar os diferentes eventos presentes na aplicação é necessário recorrer a uma linguagem de programação que permita interligar os diferentes sistemas e automatizar processos.

É também necessária a criação dos conteúdos sonoros que vão integrar o jogo. É essencial atender ao género de sonoridade pretendida, tendo em conta que este será o principal *output* de informação para o utilizador, pelo que tem de haver uma harmonia entre os diferentes elementos.

Como objetivo específico do trabalho foi escolhida a criação de um jogo individual que pretende colocar o utilizador numa posição defensiva, onde todos os seus esforços se direcionam para a sua própria sobrevivência no mundo virtual. O jogador experiencia o jogo de uma perspectiva de primeiro plano, estando no centro de ação, de modo a perceber o ambiente virtual que o rodeia à semelhança de um jogo do género *first-person shooter*. Todo o ambiente deverá produzir no utilizador uma sensação de medo e insegurança, de modo a que os seus instintos naturais tomem partido na experiência, guiando-o em situações de maior perigo. A instalação conta com um ambiente com pouca ou nenhuma visibilidade e sem qualquer *feedback* visual, sendo o som a componente principal de orientação com o suporte táctil proveniente de um controlador gestual que se assemelha a uma arma, à qual poderá recorrer para se defender dos atacantes. De modo a ampliar ainda mais a sensação de desamparo emocional, é possível a utilização de uma estrutura física, que confina o utilizador a um espaço restrito.

Este projeto tem como público-alvo utilizadores de qualquer idade, sendo apenas necessário que tenham destreza física suficiente para empunhar e manipular a interface de jogo bem como a capacidade auditiva para a orientação na realidade virtual sonora.

3.3. Planeamento

O planeamento estruturado da conceptualização, desenvolvimento e exposição do projeto é apresentado no *Apêndice 1*.

3.4. Metodologias

O trabalho está estruturado de modo a que seja feito um estudo de cada componente, individualmente, de modo a maximizar a sua eficiência. Após definidos os métodos e materiais a utilizar, foi, então, feita a montagem da instalação, através da conjugação dos vários elementos.

3.4.1. Narrativa e estrutura do jogo

A narrativa do trabalho foi bastante influenciada pelo videogame *Portal* da *Valve Corporation* (2007), devido ao seu humor negro e clima tenso e imprevisível. Foi escolhido o nome *The Sound of Horror* de modo a transmitir de forma direta e simples a natureza da experiência, que se destaca não pela inovação a nível de género ou de conteúdos, mas pela forma como estes são apresentados ao utilizador e o modo como a interação é efectuada.

Rouz-Girard (2010) defende que os jogos de terror não criam medo no utilizador apenas pela sua dimensão estética, mas pela sua estrutura narrativa e *gameplay*

O jogo passa-se num ambiente misterioso e obscuro com uma ambiência que se assemelha a um misto de caverna e laboratório subterrâneo, onde o jogador se encontra aprisionado, por motivos desconhecidos. Não lhe é dada nenhuma informação relativamente

ao modo como foi capturado ou em que local se encontra. No entanto, no decorrer da experiência, começa a aperceber-se de que o propósito da sua estadia involuntária se deve ao facto de estar a ser utilizado como cobaia num conjunto de testes que se prendem com a sua sobrevivência face ao ataque de criaturas monstruosas. Para se defender, é-lhe disponibilizada uma arma à semelhança de uma caçadeira com uma baioneta incorporada, permitindo ao jogador disparar contra os atacantes ou esfaqueá-los, se estes se encontrarem a uma distância apropriada para tal.

Tendo em conta que este tipo de experiência interativa é estranha à maioria dos utilizadores, apesar dos objectivos no jogo serem comuns para videojogos, *e.g.* *Left 4 Dead 2* (Valve Corporation, 2009), *The Last of Us* (Naughty Dog, 2013), torna-se necessária a implementação de um percurso tutorial, onde o utilizador possa ter a possibilidade de aprender as funcionalidades da interface e de ficar familiarizado com o método de orientação sonora, sem estar exposto às condicionantes impostas pelo percurso do jogo real.

A Figura 23 apresenta uma representação muito simples da estrutura do jogo. Após o seu início, o utilizador é enviado para um menu de seleção, onde pode escolher entre o modo tutorial ou o modo de jogo. Independentemente do seu percurso em cada um, nomeadamente do tempo que demore a terminar o tutorial ou do seu resultado no modo de jogo, é enviado de volta para o menu principal. Foi escolhida esta abordagem tendo em conta o carácter da experiência de uso individual, salientando o facto de que o jogador apenas tem controlo da arma que empunha, não tendo qualquer influência no iniciar ou terminar da aplicação.

A seleção no menu principal é feita através da arma, que nesta fase desempenha um papel de um comando. O disparo contra alvos que são apresentados ao utilizador, permite a seleção entre o modo tutorial e o de jogo. Estes alvos sonoros são apresentados em pontos distintos do campo de escuta do jogador e são devidamente identificados como pontos de acesso aos dois modos.

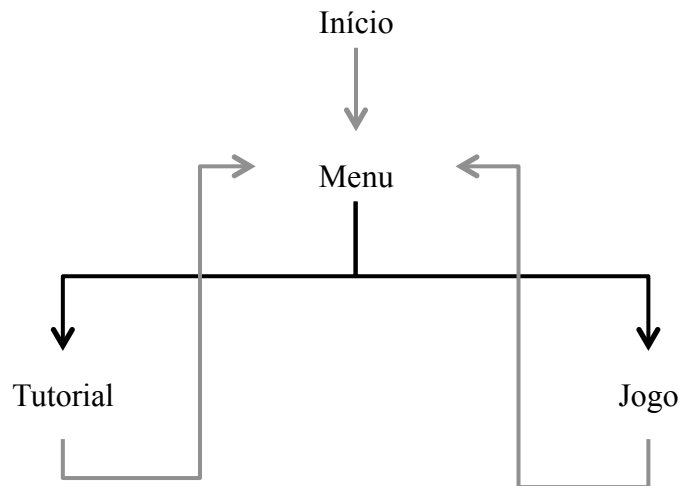


Figura 23 – Diagrama simplificado da estrutura do jogo *The Sound of Horror*.

O modo de jogo é compreendido por três níveis, que são apresentados como sequências de teste impostas ao utilizador. Em cada uma, de forma sequencial, surgem monstros de direções aleatórias do campo de escuta do utilizador, dirigindo-se a ele, com o propósito de o matar. No primeiro nível, estes monstros são grandes e lentos, tendo trajetos mais fáceis de intersectar. Passando ao segundo nível, os monstros tornam-se mais pequenos e ágeis, obrigando o jogador a melhorar o seu tempo de reação. No último nível são apresentadas criaturas mais inteligentes e com trajetos imprevisíveis, ziguezagueando de forma rápida para esquivar os ataques do utilizador. Para passar de um nível para o seguinte, é necessário eliminar todos os monstros sem exceção. Quando algum dos alvos consegue aproximar-se do jogador, ele começa o seu próprio ataque. É imposto um limite curto de tempo dentro do qual o monstro terá que ser eliminado, caso contrário, o jogador é morto, terminando a sequência de teste e enviando o utilizador de volta para o menu principal, sinalizando o perder de jogo.

A arma permite ao utilizador disparar e esfaquear os inimigos, sendo necessário recarregar quando se encontra sem balas. Com a exceção da ação de disparar, que é universalmente definida como o pressionar de um gatilho, o esfaquear e recarregar são funções que requerem competências específicas por parte do utilizador. O modo de tutorial ensina o utilizador a executar cada uma destas ações, sequencialmente, apresentando um ambiente menos tenso que o modo de jogo e alvos sonoros que não produzem uma sensação de medo. São dadas instruções ao utilizador seguido de um período de teste onde tem que aplicar os conhecimentos adquiridos para prosseguir.

Cada mudança de etapa do jogo envolve transições, para que o utilizador perceba que está a haver uma alteração que pode resultar de uma ação tomada ou simplesmente de uma

automatização do jogo. Foram inseridos momentos de transição, ou *Cinematics*, com o propósito de transportar virtualmente o utilizador para a etapa seguinte da experiência. Na Tabela 6 é possível ver a lista completa das várias fases do jogo, organizadas pelas suas características interativas.

Tabela 6 – Lista de etapas presentes no jogo de acordo com o seu grau de interatividade.

Sem interação	Com interação
<i>Cinematic</i> Início	Menu
<i>Cinematic</i> Tutorial	Tutorial
<i>Cinematic</i> Tutorial – Menu	Nível 1
<i>Cinematic</i> Jogo	Nível 2
<i>Cinematic</i> Nível 1 – 2	Nível 3
<i>Cinematic</i> Nível 2 –3	
<i>Cinematic</i> Vitória	
<i>Cinematic</i> Derrota	

As etapas assinaladas como interativas estão condicionadas, em alguma forma, pelo *input* do utilizador. O percurso do jogo nestes momentos está dependente das ações tomadas, o que resulta numa imprevisibilidade a nível temporal. Na Figura 24 é apresentado um diagrama detalhado da estrutura do jogo, onde são especificados cada um dos passos por onde o utilizador passa nas diferentes abordagens possíveis que pode tomar.

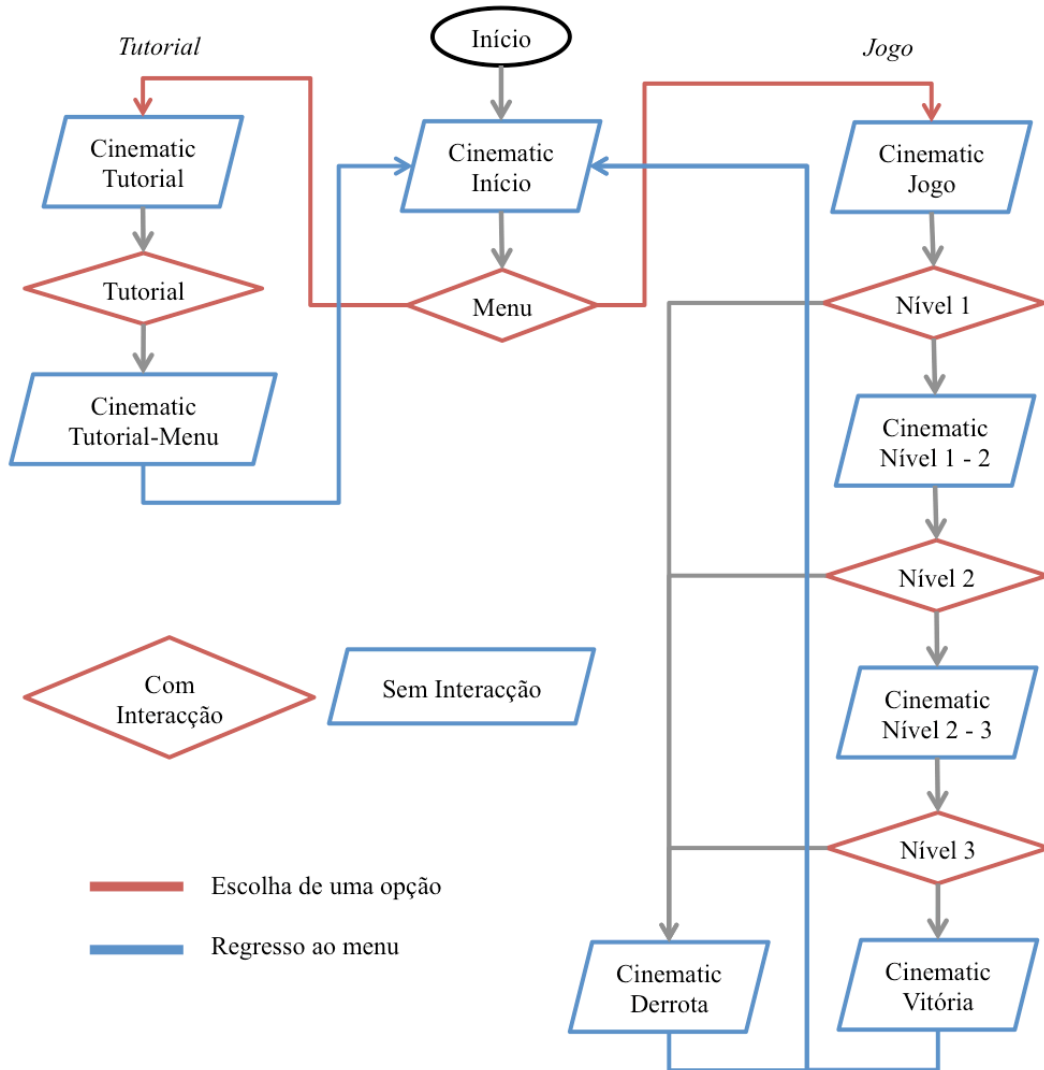


Figura 24 – Diagrama extenso da estrutura do jogo *The Sound of Horror*.

O jogo conta com a presença de duas figuras que se dirigem ao jogador através de diálogos. Estes são o Narrador e o Grande Irmão¹. O Narrador orienta o utilizador no Menu e no Tutorial, ajudando-o a navegar no universo de jogo e a aprender as várias funcionalidades da interface. Por sua vez, o Grande Irmão está envolvido na narrativa do modo de jogo, sendo uma figura fria e isenta de emoções que faz a gestão das várias sequências de teste apresentadas ao utilizador. À semelhança da forma como um cientista encara uma cobaia, independentemente do resultado final do teste, o utilizador é eliminado, tornando toda a experiência questionável de um ponto de vista existencial.

A Tabela 7 apresenta a lista de falas destas duas entidades do jogo. Foi escolhido o uso da língua inglesa de modo a tornar a instalação utilizável por um grupo mais abrangente de utilizadores.

¹ Inspirado em Orwell, G. (1949). 1984. New York: Ed. Erich Fromm.

Tabela 7 – Lista de diálogos do Narrador e Grande Irmão.

Entidade	Etapa	Diálogo
Narrador	Menu	<ul style="list-style-type: none"> • <i>“To enter the Tutorial Mode shoot here.”</i> • <i>“To enter the Game Mode shoot here.”</i> • <i>“Tutorial Mode selected.”</i> • <i>“Game Mode Selected.”</i>
	Tutorial	<ul style="list-style-type: none"> • <i>“Welcome to the Tutorial Mode.”</i> • <i>“To kill targets press the trigger while aiming at the source. Try hitting this.”</i> • <i>“Good job! You can also stab targets if they’re close enough.”</i> • <i>“To stab, thrust the weapon forward in the direction of the source. Try stabbing this.”</i> • <i>“Nice one! Each magazine has five bullets.”</i> • <i>“To reload, point the weapon upwards and then press the trigger. Try reloading.”</i> • <i>“Great! You are now ready to play.”</i>
	Nível 1, 2, 3	<ul style="list-style-type: none"> • <i>“Level Completed.”</i>
Grande Irmão	<i>Cinematic</i> Jogo	<ul style="list-style-type: none"> • <i>“Greetings test subject number 31739. You will find yourself in an enclosed area. These measures were implemented for your own safety. Do not be alarmed. This will all be over soon.”</i>
	Nível 1	<ul style="list-style-type: none"> • <i>“You have been given a weapon. Feel free to use it to defend yourself.”</i>
	Nível 2	<ul style="list-style-type: none"> • <i>“Initiating the second test sequence. Please stand by.”</i>
	Nível 3	<ul style="list-style-type: none"> • <i>“This next test is impossible. Make no attempt to solve it.”</i>
	<i>Cinematic</i> Vitória	<ul style="list-style-type: none"> • <i>“Thank you for participating in this computer-aided enrichment activity. Goodbye. Initiation termination protocol.”</i>

Para além destas, são ainda utilizadas duas falas não pertencentes a nenhuma das entidades previamente mencionadas. No *Cinematic* Início, para introduzir o utilizador ao jogo é usada a frase *“Welcome to The Sound of Horror”*. Para sinalizar a derrota, no *Cinematic* Derrota, é utilizada a expressão *“Game over.”*

3.4.2. Sistema de espacialização sonora

Para a implementação do sistema de espacialização sonora foi utilizada a versão 5.1.9 da linguagem de programação *Max/MSP*, pela facilidade que oferece na manipulação de sinal sonoro, não exigindo conhecimentos profundos a nível dos princípios matemáticos envolvidos no processamento de sinal ou de programação em geral, através de uma linguagem gráfica intuitiva.

Foram feitos testes relativamente a várias técnicas, que têm como base os princípios teóricos abordados no capítulo 2.2., e feita uma análise relativamente às vantagens e desvantagens de cada uma, de modo a encontrar a mais adequada para o projeto em causa.

3.4.2.1. DBAP

Tendo em conta a facilidade de implementação da *Distance-based Amplitude Panning*, foi desenvolvido um *patch* que permite a espacialização sonora de um sinal monofónico para uma distribuição variável de até 24 altifalantes (Figura 25). É feito o cálculo entre a posição da fonte sonora e de cada coluna para determinar o nível de ganho desta e é ainda definida a linha de resposta do *rolloff* (Figura 26). Tendo em conta a natureza deste teste, não foram aplicados os factores de *weight*, apesar da sua implementação ser extremamente simples.

Desta experiência concluiu-se que a baixa precisão espacial da *DBAP* é insuficiente para uma aplicação que está inteiramente dependente da localização precisa de fontes sonoras virtuais. O sistema tem a capacidade de confinar determinados sinais sonoros a zonas do espaço, o que é ideal para a criação de ambiências sonoras com características espaciais, mas a falta de precisão sonora, que em parte advém da simplicidade da técnica, torna-a inviável para o *The Sound of Horror*.

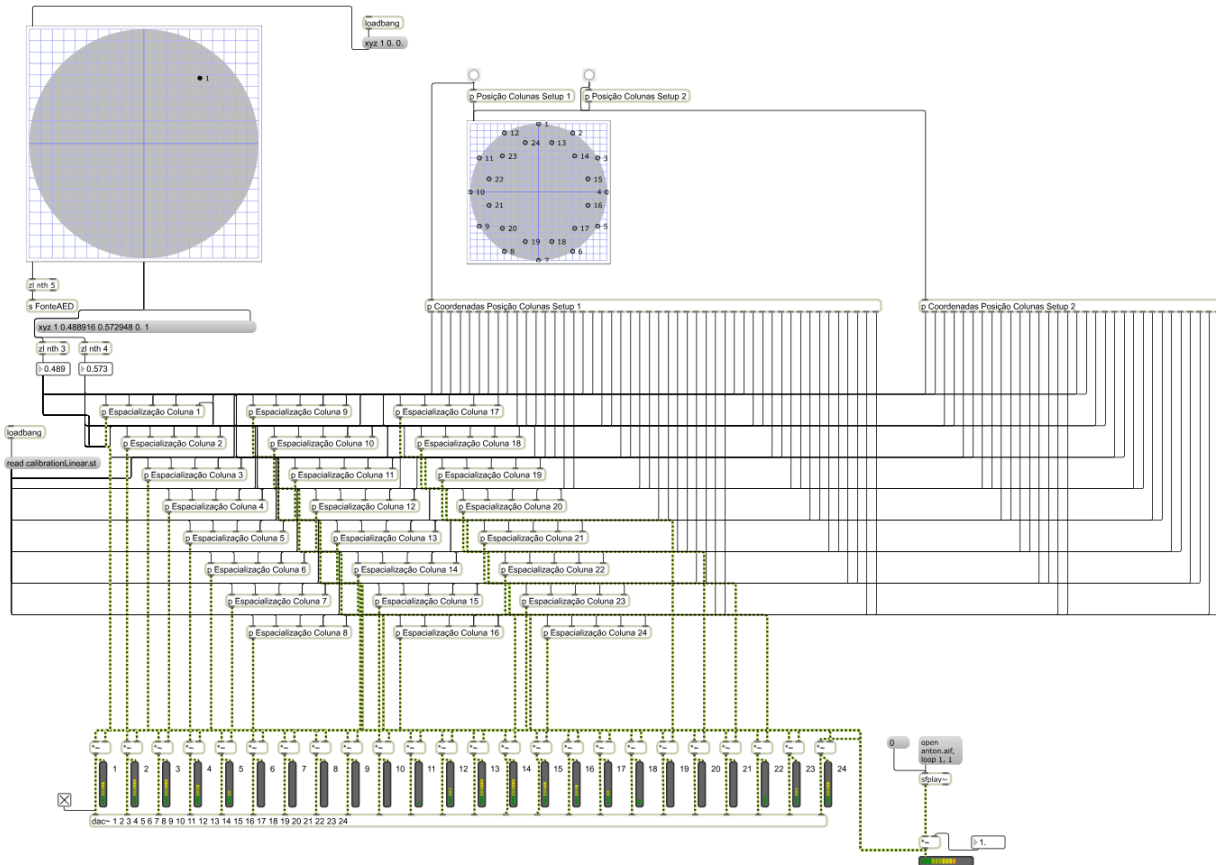


Figura 25 – Patch de implementação da técnica DBAP no Max/MSP (1).

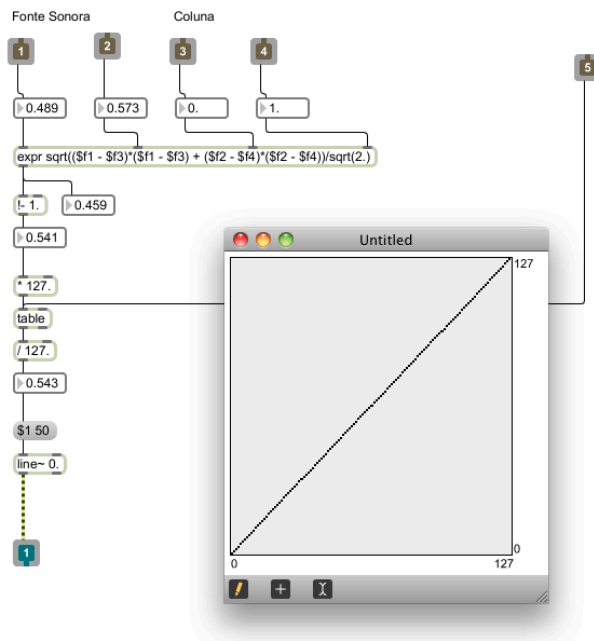


Figura 26 – Patch de implementação da técnica DBAP no Max/MSP (2).

3.4.2.2. VBAP

Em conjunto com o desenvolvimento da *Vector base Amplitude Panning*, Pulkki (2000) implementou este sistema de espacialização sonora nas linguagens *Max/MSP* e *Pure Data*, disponibilizando gratuitamente os objetos externos (TKK Acoustics, 2006). A Figura 27 apresenta uma experiência feita, utilizando esta técnica, na qual foi incorporada uma variável de *Spread*, que permite controlar a quantidade de sinal que se “espalha” para os triângulos de altifalantes mais próximos, permitindo controlar sensação de largura da fonte sonora.

Para o projeto em causa, esta técnica apresenta melhores resultados que a *DBAP* relativamente à precisão espacial das fontes. No entanto, o objecto utilizado não apresenta grande flexibilidade em termos de calibração do sistema de espacialização, como a definição da linha de resposta do *spread*, à semelhança do *rolloff* da *DBAP*.

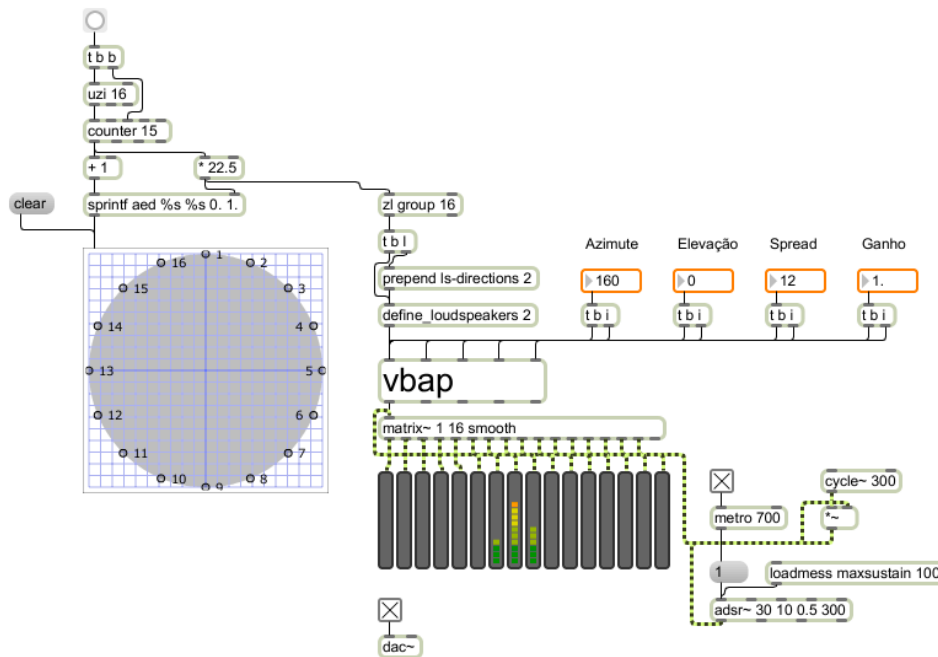


Figura 27 – Representação da técnica VBAP no Max/MSP.

3.4.2.3. ViMiC

A *ViMiC* foi implementada no *Max/MSP* como um módulo da ferramenta *Jamoma* (s.d.), oferecendo um controlo do número, posição e orientação dos microfones virtuais e os seus padrões de direccionalidade, as características de reverberação da sala e da posição e orientação da fonte sonora (Figura 28). Apesar de ter excelentes resultados a nível de criação de ambientes sonoros virtuais, tem um conjunto de limitações. Em primeiro lugar, oferece

apenas a possibilidade de sintetizar espacialmente uma fonte de cada vez, o que, tendo em conta os requisitos computacionais da *ViMiC*, a torna inviável para o projeto. Apresenta, também uma certa latência no controlo em tempo real dos parâmetros, o que pode comprometer a experiência.

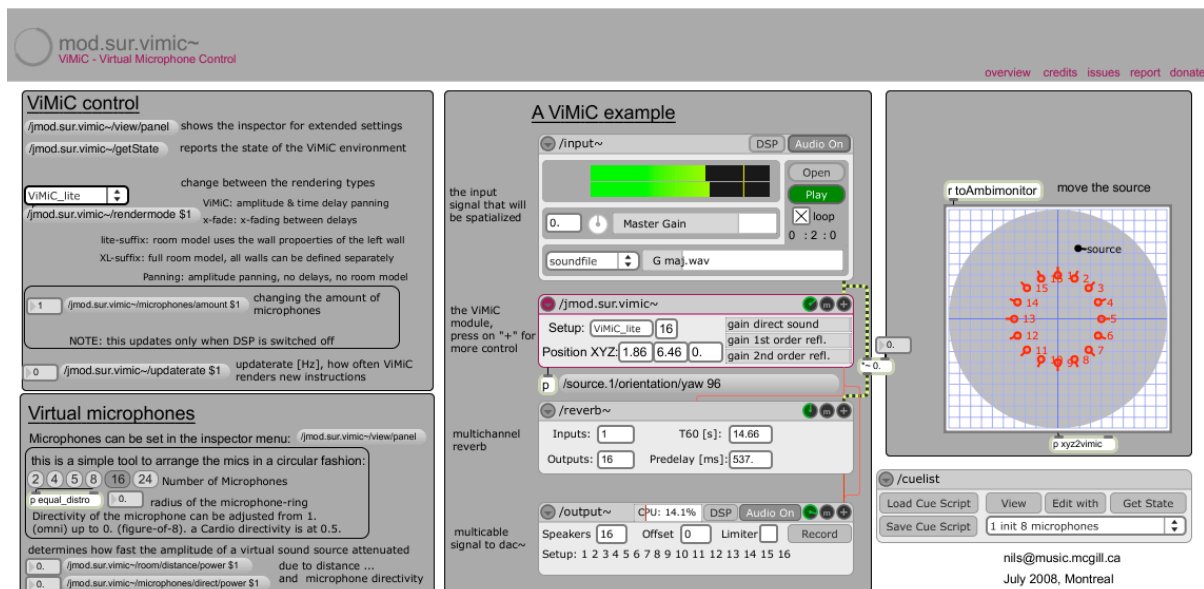


Figura 28 – Imagem do módulo de *ViMiC* do *Jamoma* no *Max/MSP*.

3.4.2.4. *Ambisonics*

Foram testadas duas ferramentas de espacialização *Ambisonics* no *Max/MSP*. Ambas permitem definir a ordem de espacialização, o número de *inputs* de sinal áudio e de *outputs* para altifalantes. Há, no entanto, diferenças significativas nestas duas implementações.

O primeiro teste foi feito com a biblioteca de objectos *HOALibrary* (HoaLibrary, 2012) (desenvolvida pelo *CICM*, *Centre de recherche Informatique et Création Musicale*, em Paris (Figura 29). Apesar dos resultados a nível sonoro serem bastante aceitáveis, esta implementação revelou-se bastante limitada, permitindo apenas configurações em duas dimensões. Não é possível, também, definir a posição de cada altifalante, sendo assumida uma disposição circular, que é a mais eficiente neste sistema. Apesar de ser possível escolher entre três conjuntos de *weights* (coeficientes de ganho) para os canais codificados, não é permitido um controlo de calibração da linha de resposta dos canais ao afastar de distância da fonte. São oferecidos objetos de monitorização que permitem ajudar a compreender os fundamentos teóricos do *Ambisonics*, no entanto para a criação de fontes sonoras virtuais, apenas é possível controlar o ângulo de azimute, não sendo viável para o *The Sound of Horror*.

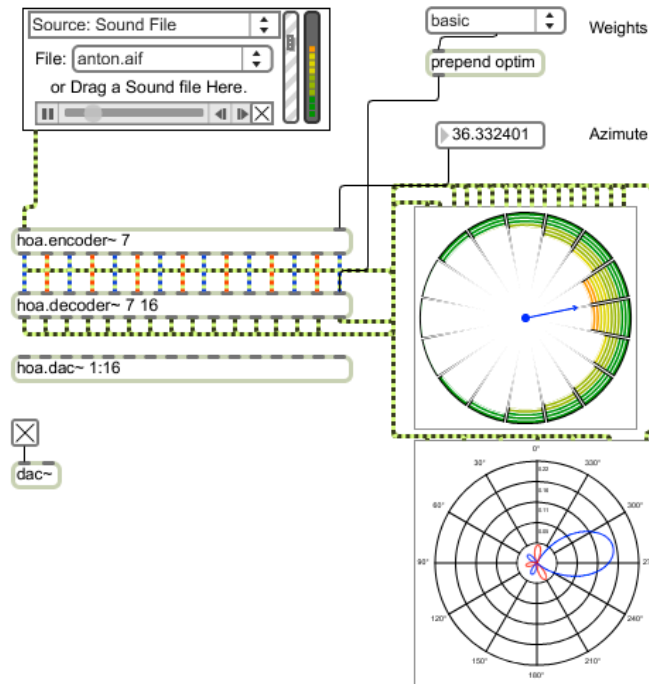


Figura 29 – Imagem dos objetos HOALibrary no Max/MSP.

Foram também testados os objetos externos do *ICST (Institute for Computer Music and Sound Technology)* (ICST, 2012), apresentados na Figura 30. Estes apresentaram muitos melhores resultados que os anteriores, tanto a nível de espacialização, como do controlo dinâmico da posição de fontes e colunas, e de parâmetros de calibração para o espaço de instalação (Figura 31). Oferece, também, um conjunto de operações de controlo do trajeto e movimento das fontes sonoras, como a gravação de trajetórias e o movimento combinado de conjuntos de fontes (Figura 32).

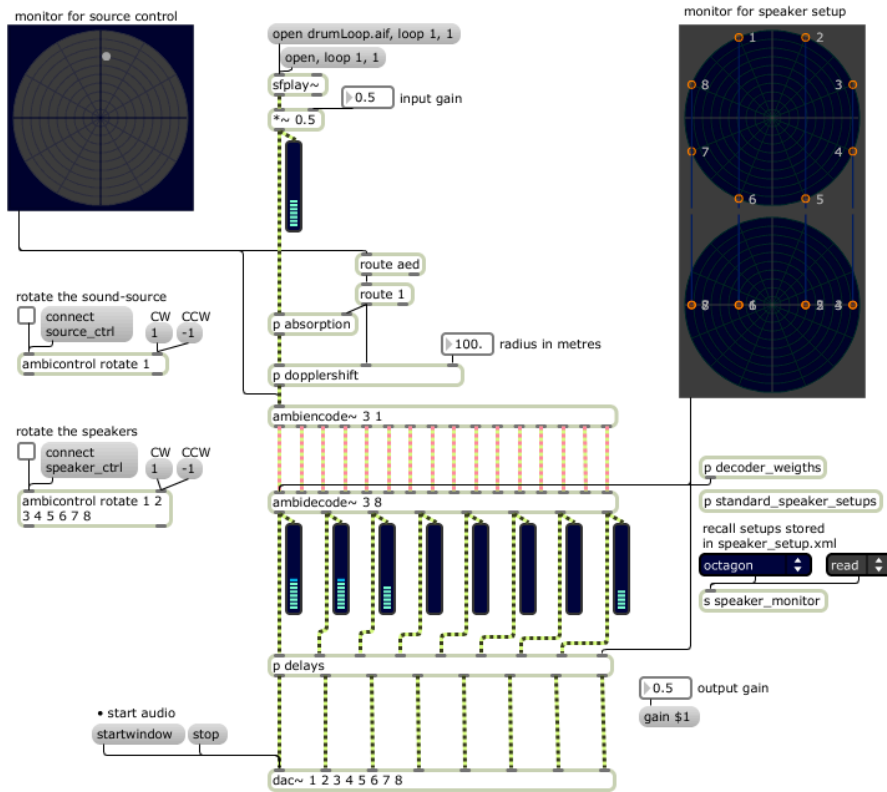


Figura 30 – Exemplo do uso dos objetos Ambisonics do ICST no Max/MSP.

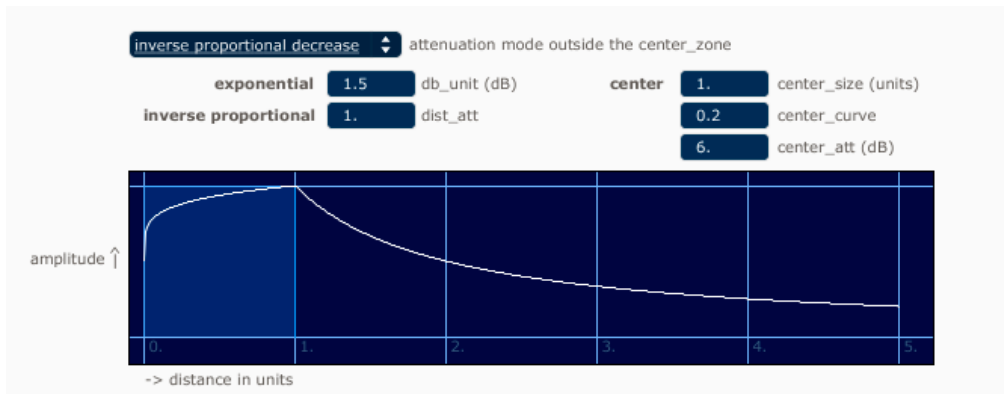


Figura 31 – Imagem do sistema de calibração da distância nos objetos Ambisonics do ICST.

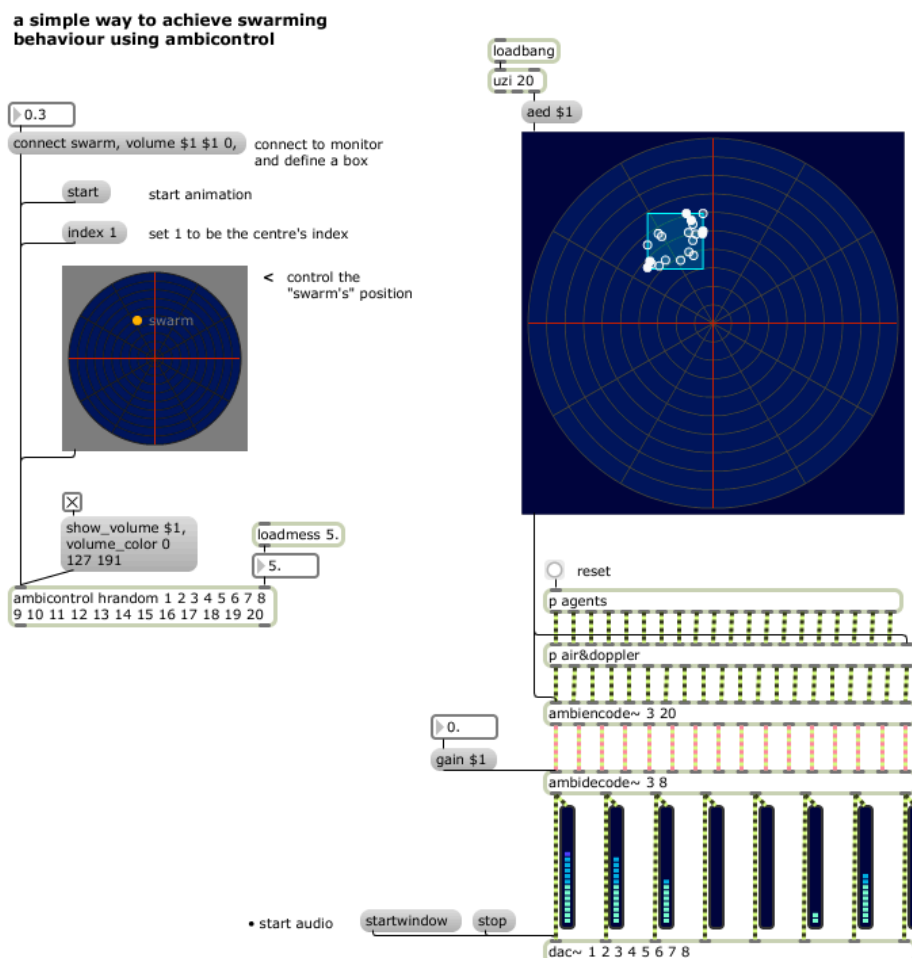


Figura 32 – Exemplo de manipulação da trajetória de conjuntos de fontes sonoras através dos objetos *Ambisonics* do *ICTS*.

3.4.2.5. Binaural

A espacialização sonora interativa com base nos princípios binaurais foi testada através da ferramenta *Binaural Tools* (2009). Esta fornece quatro conjuntos de *HRTFs*, e objetos externos que permitem ao *Max/MSP* fazer o processamento do sinal com base no azimute e elevação escolhidos (Figura 33).

Apesar de ser uma técnica que exige um sistema de difusão menos elaborado e de, teoricamente, permitir a espacialização sonora com um elevado grau de fidelidade, os problemas detectados no capítulo 2.2.1. são confirmados pela experiência feita nesta implementação, que não apresenta uma boa precisão espacial.

Para a implementação desta técnica de espacialização sonora no projeto, seria necessário o uso de um sistema de *tracking* da cabeça do utilizador, para que houvesse uma compensação dos movimentos efectuados, de modo a garantir a estabilidade do campo sonoro. Apesar das várias lacunas e exigências desta abordagem, o uso desta ferramenta não restringe o utilizador a um campo de escuta reduzido.

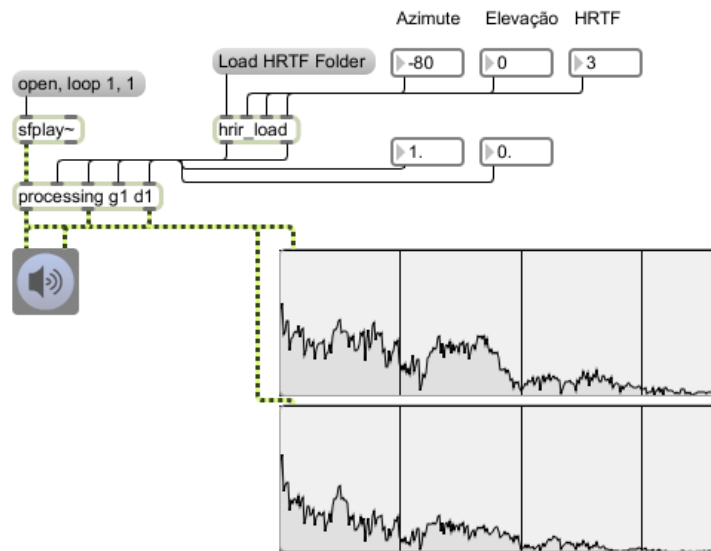


Figura 33 – Imagem dos objetos Binaural Tools no Max/MSP.

3.4.2.6. Conclusões

Os testes e experiências elaborados permitiram concluir que das técnicas de espacialização sonora estudadas, o sistema *Ambisonics* disponibilizado pelo *ICST* apresenta os melhores resultados para o projeto *The Sound of Horror*, permitindo uma grande flexibilidade e precisão na criação de fontes sonoras dinâmicas e dando ferramentas para a calibração do sistema para diferentes espaços de instalação e distribuições de altifalantes.

A lacuna apresentada pelo *Ambisonics* é a restrição a um espaço de escuta relativamente reduzido. No entanto, este problema é resolvido com a implementação de uma estrutura que delimita o espaço de utilização do jogador, à semelhança de uma jaula, tendo também o efeito de ampliar a sensação de medo e pânico produzido pelo atacar dos monstros, fazendo o utilizador aperceber-se, tactilmente, de que está preso à realidade virtual apresentada.

3.4.3. Sistema de difusão sonora

Após definido o método de espacialização sonora, foi necessário fazer um estudo relativamente ao material de difusão sonora disponível na EA/UCP capaz de responder às necessidades impostas pelo projeto. Para a utilização de um sistema de colunas, foi identificado como o material mais eficiente a seguinte lista de material:

- Altifalantes:
 - *Genelec 6010A*, 16 unidades (Figura 34);
 - *Genelec 7070A* (Figura 35).
- Placas de som:
 - *MOTU 828mk3 Hybrid*;
 - *MOTU 828mk3*;
 - *Mackie Onyx 820i*.



Figura 34 – Fotografia do altifalante Genelec 6010A (Imagem retirada de <http://a248.e.akamai.net/origin-cdn.volusion.com/b3o4z.gn3gt/v/vspfiles/photos/9999-09402-2.jpg>).



Figura 35 – Fotografia do altifalante Genelec 7070A (Imagem retirada de http://vintageking.com/media/catalog/product/g/e/genelec_7070a_3.jpg).

Para a ligação deste material, é, ainda, necessária a seguinte cablagem:

- 16 cabos corrente bipolar;
- 1 cabo corrente tripolar
- 16 cabos Jack – RCA
- 1 cabo Jack – XLR macho

É ainda necessária cablagem e fichas eléctricas macho e fêmea para garantir que todas as colunas recebem alimentação.

Como observado no ponto 2.2.3., para tirar a máxima potencialidade do sistema *Ambisonics* é necessário o uso de uma distribuição regular de altifalantes (circular ou esférica). Deste modo, é necessário o uso de um espaço suficientemente grande para suportar uma configuração destas dimensões. O local escolhido para fazer esta instalação foi o Laboratório de *Motion Capture* da EA/UCP.

Foi feito um estudo em conjunto com os professores João Cordeiro, Vitor Joaquim e André Perrotta relativamente aos métodos disponíveis para suporte ou suspensão dos altifalantes, de modo a perceber quais os *setups* de colunas possíveis. As duas soluções encontradas passavam pela suspensão de colunas numa configuração rectangular utilizando cabos de aço, ou pelo uso de suportes específicos dos altifalantes *Genelec 6010A*, com altura e ângulo e orientação ajustáveis, permitindo uma altura máxima de aproximadamente 1,6 metros.

Tendo em conta a importância que o projeto coloca na capacidade de localizar com precisão fontes sonoras, decidiu-se utilizar os suportes das colunas numa configuração bidimensional e circular, representada na Figura 36, de modo a garantir uma distância exata entre o ponto de escuta e cada um dos altifalantes *Genelec 6010A*, bem como um afastamento uniforme entre cada um deles, para criar um campo sonoro consistente. O canal *LFE (Low-frequency Effects)*, reproduzido através do altifalante *Genelec 7070A*, não apresenta a direccionalidade dos restantes 16, devido à gama de frequências que reproduz. Assim sendo, não é necessário atender à sua posição no círculo.

De acordo com a fórmula apresentada no capítulo 2.2.3. para o cálculo do número mínimo de colunas necessário para cada grau de ordem de *Ambisonics*, conclui-se que para o projeto *The Sound of Horror* é necessário o uso de um sistema de 7ª ordem.

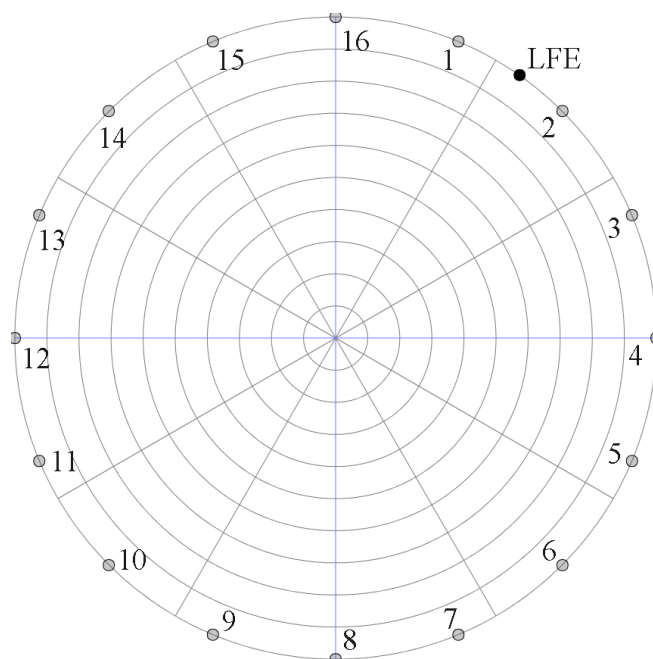


Figura 36 – Representação da distribuição bidimensional de altifalantes, do projeto *The Sound of Horror*.

3.4.4. Interface (Arma)

A interface do jogo deve permitir detectar um conjunto de ações por parte do utilizador, oferecendo um controlador à semelhança de uma arma. Tendo em conta o género de utilização que é lhe é pretendido, tem de possuir as seguintes características:

- Detecção de gestos simples;
- Presença de um gatilho idêntico ao de uma arma de jogo;
- Resistente a movimentos bruscos;
- Relativamente leve;
- *Output táctil* (e.g. vibração)

Dos materiais existentes na Universidade Católica Portuguesa, o comando *Wiimote* da *Nintendo* responde muito bem a todas estas necessidades, oferecendo um controlador leve e de dimensões pequenas, com três acelerómetros incorporados, que permitem a detecção de movimentos nas três dimensões.

Procedeu-se à construção de uma estrutura para suportar o comando e conceder-lhe as dimensões e forma de uma arma. Para isto foi utilizado o acessório *N'PLAY MB-8132* para

segurar o comando e a sua extensão *Nunchuck*, e colado a uma calha de PVC para estender o cano da pistola, assemelhando-a mais a uma arma de grande porte, sem restringir o acesso ao Wiimote. A Figura 37, Figura 38, Figura 39 e Figura 40, apresentam a interface construída.

Para recolher as informações do comando, foi utilizado o objeto externo *ajh.wiiremote* para o *Max/MSP*, desenvolvido por Alexander Harker, apresentado na Figura 41. Este permite fazer a ligação *bluetooth* entre o dispositivo e o computador e identificar o pressionar de qualquer um dos botões, bem como os valores dos três acelerómetros presentes no comando e as posições e tamanhos de informação infravermelha captada pela câmara. Esta ferramenta reconhece também extensões do aparelho, como o *Motion Plus* e o *Nunchuck*.

A partir das informações dos vários botões e valores dos acelerómetros, foi possível definir as ações associadas à arma.

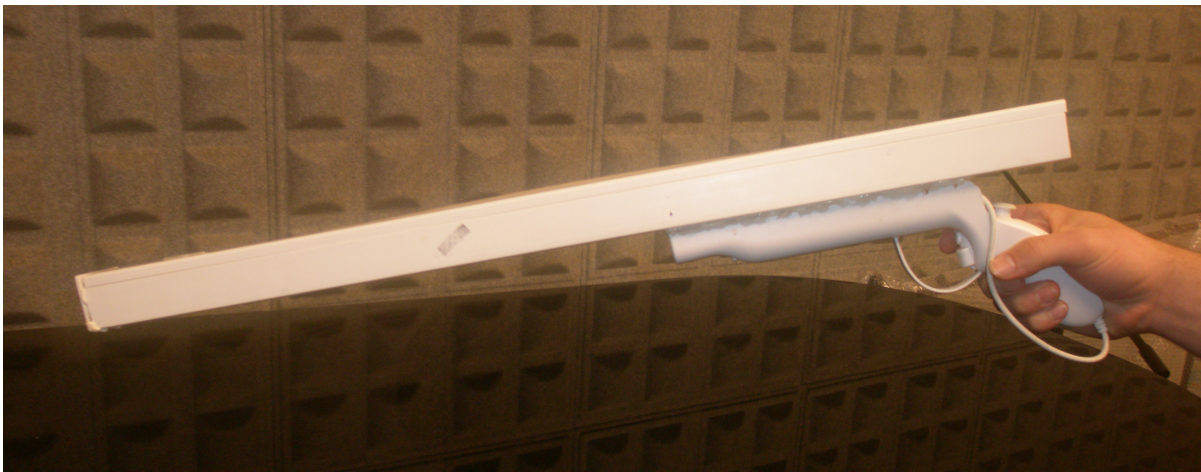


Figura 37 – Interface do jogo *The Sound of Horror* (1).



Figura 38 – Interface do jogo *The Sound of Horror* (2).

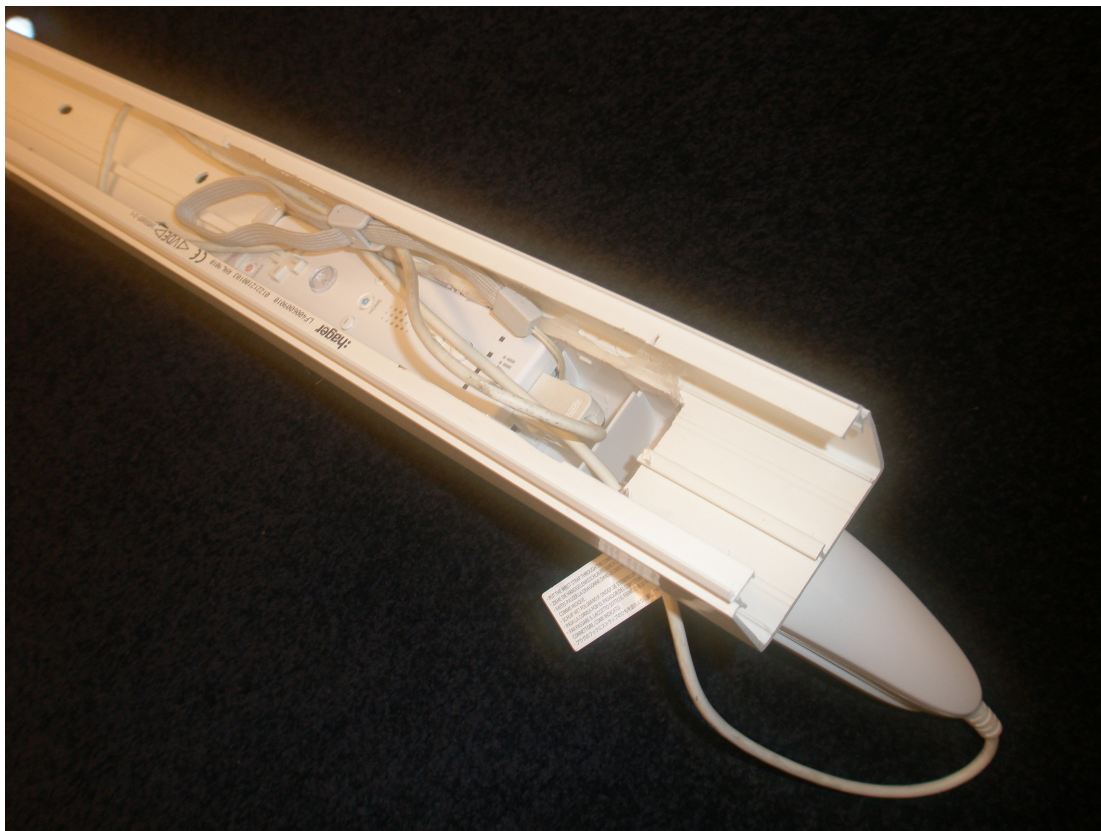


Figura 39 – Interface do jogo The Sound of Horror (3).



Figura 40 – Interface do jogo The Sound of Horror (4).

3.4.4.1. Ações da interface

A arma possui três comandos básicos. O tiro da arma funciona tanto para disparar contra os monstros como para fazer a seleção no Menu. É exigida uma ação semelhante ao tiro de uma arma real, para que seja possível ao utilizador, sem qualquer instrução prévia, selecionar uma opção ao iniciar o jogo. Na interface construída, o botão que mais se assemelha ao gatilho de uma arma, é o *Z* da extensão *Nunchuck*, seguido do *B* do *Wiimote*. Para garantir uma utilização correta foram associados ambos botões ao tiro da interface ambos para garantir uma utilização eficaz.

De modo a evitar que tiros com a arma apontada para cima ou para baixo tivessem o mesmo efeito que aqueles a apontar para os alvos (na horizontal), foi feita uma distinção entre os tiros através dos valores dos acelerómetros *Y* e *Z* do comando no instante em que o gatilho é pressionado.

O gesto escolhido para a ação de esfaquear baseou-se na forma convencional de uso de uma baioneta, impulsionando a arma na direção do alvo, paralelamente ao chão. Foi utilizado o valor do acelerómetro *Y* de modo a identificar este movimento e calibrada a brusquidão necessária para ativar o comportamento.

Foram feitas experiências de modo a encontrar o controlo gestual que mais se adequava à ação de recarregar a arma. O uso de gestos como abanar a arma rapidamente para cima e para baixo revelou-se pouco prático e desconfortável. Concluiu-se que a opção mais viável seria utilizar o gatilho para acionar a ação de recarregar, em conjunto com a arma apontada para cima, tirando partido desta variante do tiro anteriormente definida. Isto permite criar uma sensação de maior controlo sobre a interface, mantendo um controlo gestual simples. Para tornar o jogo mais desafiador definiu-se que após cinco tiros seria exigido ao utilizador recarregar a arma para continuar a disparar.

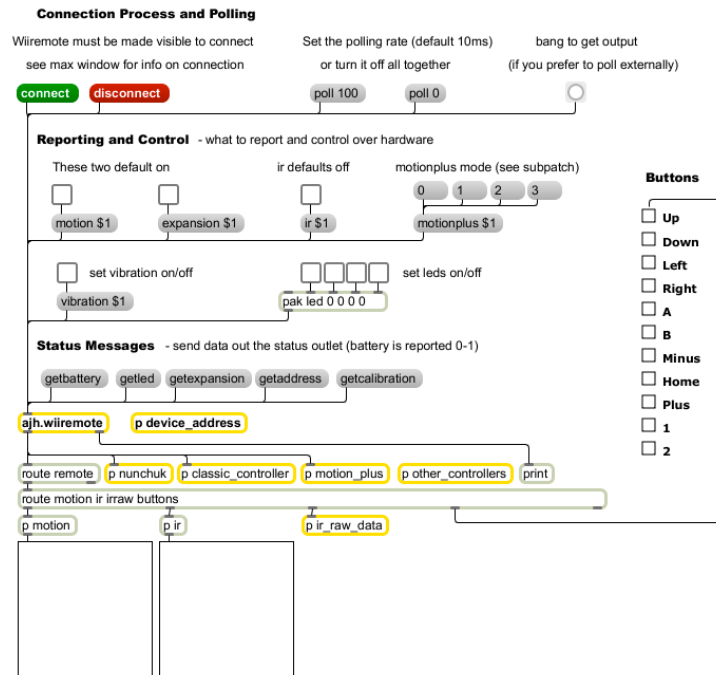


Figura 41 – Objeto externo *ajh.wiiremote* para o Max/MSP.

3.4.5. Estrutura para confinar o utilizador

Para a elaboração da estrutura que pretende delimitar os movimentos do utilizador é necessário ter em conta um conjunto de necessidades que advêm tanto dos objetivos do projeto como da forma como este vai ser utilizado, para garantir uma experiência imersiva e interativa orgânica e completa. A estrutura deve então responder às seguintes características:

- O espaço não deve ser demasiado pequeno, de modo a não restringir os movimentos do utilizador excessivamente;
- Não deve ser demasiado grande, para que o utilizador não saia do *sweet spot* do campo sonoro;
- A estrutura deve ser estável e resistente;
- Deve proporcionar um espaço seguro para o utilizador, na ausência de luz.

A Figura 42 e Figura 43 apresentam o processo de construção desta componente. Procedeu-se à construção de uma base de madeira quadrada com aproximadamente 1,3 metros de lado com uma cerca construída com cantoneiras de ferro com 1 metro de altura. Juntou-se também um círculo de borracha para permitir uma orientação com os pés, e tubos de plástico

a unir os vértices. Para aumentar a segurança para o utilizador, foi aplicada espuma de poliuretano em todas as superfícies rígidas e pontiagudas da cerca.



Figura 42 – Construção da estrutura para confinar o utilizador (1).



Figura 43 – Construção da estrutura para confinar o utilizador (2).

3.4.6. Tracking de movimentos

Esta componente tem como objetivo principal a detecção do ângulo de mira do utilizador, de modo a identificar a direção do tiro e da ação de esfaquear. Na tentativa de encontrar um sistema que permita recolher esta informação, foi feita uma análise a um conjunto de possíveis abordagens.

3.4.6.1. Wiimote

No protótipo desenvolvido na unidade curricular de Interfaces Tangíveis, foi utilizada a câmara de um comando *Wiimote* para detectar *leds* infravermelhos montados no topo da interface. O comando foi pendurado por cima do espaço de utilização e através das posições de cada um dos pontos identificados, foi feito o cálculo do ângulo de mira (Figura 22).

Este método revelou-se ineficaz pela reduzida abertura da câmara, que resultou numa área de utilização demasiado pequena para os movimentos naturais do utilizador. Outra falha encontrada foi a uni direccionalidade que os *leds* infravermelhos apresentam, tornando-se difíceis de detectar quando a arma é inclinada para qualquer um dos lados.

3.4.6.2. Polhemus

A empresa *Polhemus*² desenvolve e disponibiliza comercialmente dispositivos de *motion tracking* que permitem a uma liberdade total por parte do utilizador, através do uso de sensores (com ou sem fios, dependendo do modelo) ligados a um sistema geoestacionário, que permite fazer uma localização com base em coordenadas cartesianas e na orientação do sensor. Os seus produtos são usados em diversas áreas, como realidade virtual, simulações de treino, pesquisa médica, entre outras.

Cheng (2001) faz uso de um sensor *Polhemus* para o *tracking* da cabeça de um utilizador para medições *HRTF*. Há outros registos de uso destes aparelhos para estudos relacionados com a área do som (e.g. Jensenius *et al.*, 2008). Apesar deste sistema dar

² Website da empresa disponível em <http://www.polhemus.com/>

resposta a todas as necessidades do projeto *The Sound of Horror*, o seu custo elevado tornou-o inviável, tendo sido desconsiderado.

3.4.6.3. *Microsoft Kinect*

O dispositivo *Microsoft Kinect* conjuga um emissor de radiação infravermelha com uma câmara de infravermelhos para compor um sensor de profundidade 3D³. Há um conjunto de *Software Development Kits (SDK)* disponíveis (e.g. *Microsoft SDK*, *OpenNI*) que permitem a detecção e *tracking* tridimensional do corpo humano até seis utilizadores simultaneamente.

Tohm Judson (Judson, 2011) publicou um guia para a instalação e ligação do *Kinect* ao *Max/MSP* através do protocolo de *Open Sound Control (OCS)*. Este método utiliza a *SDK OpenNI*, permitindo uma detecção automática do esqueleto humano.

O projeto *The Sound of Horror* exige um sistema de detecção que permita captar o utilizador em qualquer orientação (numa rotação de 0 a 360 graus), pelo que o uso de apenas um *Kinect* não é viável, sendo que quando o jogador se encontra de costas para o dispositivo, os pontos ocultos à câmara não são captados.

Com vista a solucionar este problema, a empresa *iPi Soft*⁴ desenvolveu um *software* de captura de movimentos que conjuga a informação de múltiplos *Kinects*, criando uma imagem tridimensional sem as falhas apresentadas anteriormente.

No entanto, tendo em conta que o ângulo de mira não está dependente das coordenadas do esqueleto, mas da forma como o utilizador empunha a interface, o *tracking* tem de estar relacionado com a arma, de modo a detectar a sua posição e orientação, para proceder ao cálculo do ângulo de mira posteriormente.

Utilizando a *SDK libfreenect*, disponibilizada pela comunidade *OpenKinect*⁵ Zachary Seldess (2012) desenvolveu um sistema em *Max/MSP* para a detecção e *tracking* de *blobs* (aglomerados de informação) captados pelo sensor de radiação infravermelha do *Kinect* (Figura 44). É possível definir os limites de distância e tamanho dos *blobs*, de modo a assegurar a qualidade do *tracking*.

³ Especificações do dispositivo *Microsoft Kinect* disponíveis em <http://support.xbox.com/en-US/xbox-360/kinect/kinect-sensor-components>

⁴ *Website* da empresa disponível em <http://ipisoft.com/>

⁵ *Website* disponível em <http://openkinect.org/>

À semelhança da abordagem de *tracking* testada utilizando o comando *Wiimote*, em teoria, será possível suspender o *Kinect* do tecto e obstruir o emissor de radiação infravermelha, utilizando duas fontes externas (*leds* infravermelhos) colocadas na interface que serão captadas pela câmara de infravermelhos. Estes permitem detectar a posição e ângulo de orientação da arma, em conjunto com o sistema de *blob tracking* de Seldess. Para contornar o problema de unidireccionalidade dos *leds*, procedeu-se a um conjunto de abordagens com o intuito de possibilitar inclinações da interface sem perder a posição de cada um dos *blobs*. A primeira experiência consistiu na montagem de *leds* com uma organização semiesférica, apresentada na Figura 45. Em segundo lugar foi construído um sistema de contrapeso com uma placa embutida com *leds*, apresentado na Figura 46.

Infelizmente, os *leds* infravermelhos não possuem potência suficiente para permitirem a detecção pelo sensor do *Kinect*, o que inviabilizou o uso do dispositivo.

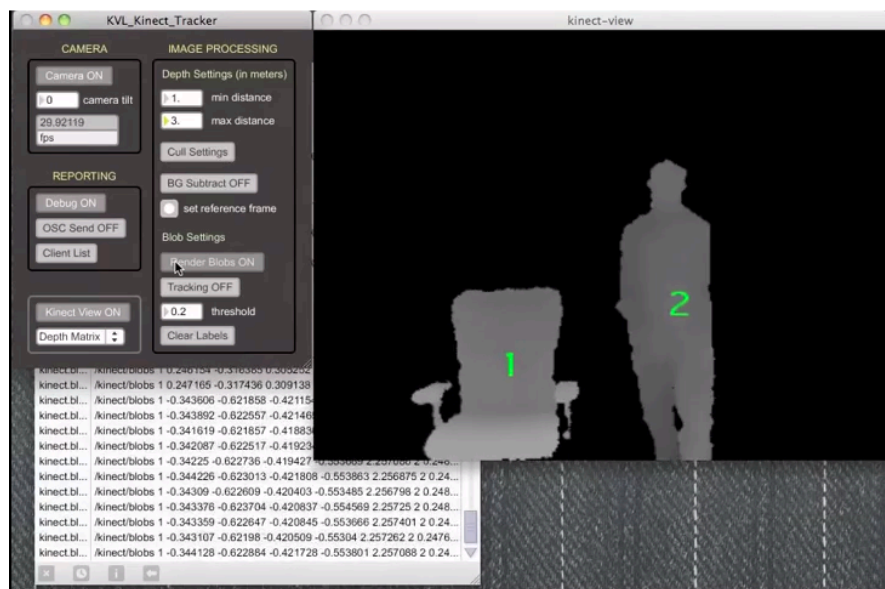


Figura 44 – Imagem do sistema de tracking utilizando o dispositivo Kinect desenvolvido por Seldess (retirado de Seldess, 2012b).



Figura 45 – Fotografia do sistema de leds infravermelhos montados com uma organização semiesférica.

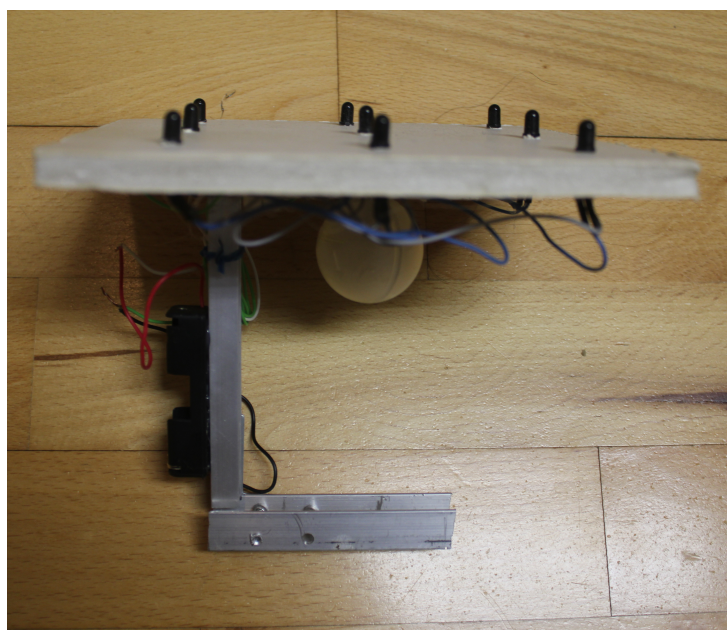


Figura 46 – Fotografia do sistema de leds infravermelhos montados em superfície plana de nível num sistema de contrapeso.

3.4.6.4. *ReactIVision*

ReactIVision (2005) é um *software open-source* de visão computacional. Este permite detecção da posição e orientação de marcadores fiduciais e do *tracking* de dedos, comunicando através do protocolo *Open Sound Control (OSC)*.

Foi testado este *software* através da construção de uma caixa com iluminação e um marcador fiducial, de modo a ser captado por uma câmara às escuras. Foi ainda necessário o uso de uma placa semitransparente por baixo do marcador de modo a atenuar o brilho e melhorar a visibilidade (Figura 47).

De acordo com as experiências efectuadas, concluiu-se que seria necessário o uso de uma câmara de alta definição para permitir uma liberdade de movimentos à semelhança da estrutura para confinar o utilizador.

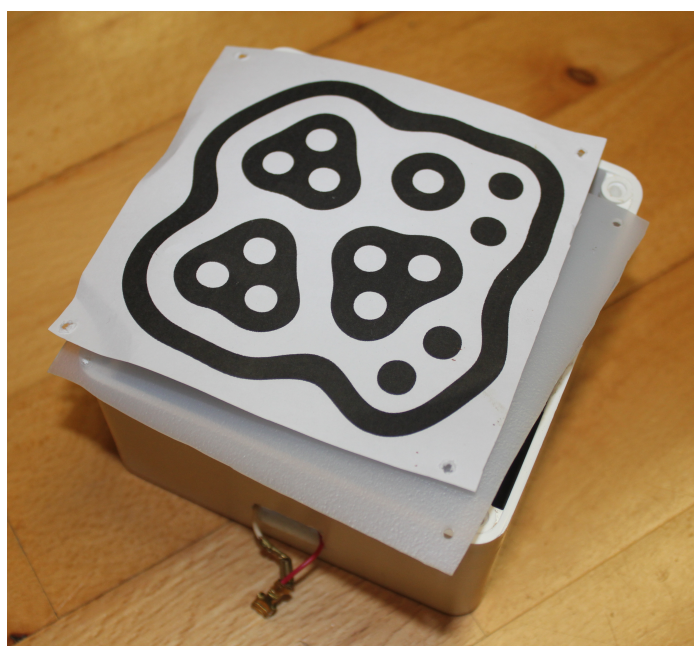


Figura 47 – Fotografia da caixa com iluminação e um marcador fiducial do *software ReactIVision*.

3.4.6.5. *Vicon Motion Capture*

O sistema de captura de movimentos da empresa *Vicon*⁶ utiliza mecanismos de *tracking* de marcadores refletivos. Para a detecção destes marcadores são utilizados focos emissores de

⁶ Website da empresa disponível em <http://www.vicon.com/>.

radiação infravermelha de alta intensidade, em conjunto com câmaras de infravermelhos. A radiação infravermelha é reflectida dos marcadores e captada por múltiplas câmaras, que criam uma imagem tridimensional do espaço captado. O sinal recolhido é filtrado de modo a que o sistema reconheça apenas os marcadores.

A *Vicon* é um dos principais fabricantes de sistemas de *motion capture*, tendo ganho um conjunto de prémios pelos seus sistemas que proporcionam precisão milimétrica na captura 3D.

O sistema *Vicon* encontra-se instalado no Laboratório de *Motion Capture* na EA/UCP (Figura 48), que conta com 10 câmaras *Vicon T-Series* e uma placa *Vicon MX Giganet* que permite a ligação e sincronização das câmaras a um computador para realizar capturas, através do *software Vicon Blade*, versão 2.0.2.



Figura 48 – Fotografia do Laboratório de Motion Capture da Escola das Artes da UCP (Imagem retirada de <http://artes.porto.ucp.pt/VisitaVirtual/EdArtes/Mocap/tour.html>).

O *software Vicon Blade* tem a capacidade de fazer *streaming* das informações dos marcadores captados para um segundo computador ligado por rede à placa *Vicon MX Giganet*, em tempo real, ou através da leitura de capturas previamente feitas, com o objetivo de proceder à animação modelos 3D.

André Baltazar (comunicação pessoal, EA/UCP, Dezembro de 2012 a Fevereiro de 2013) desenvolveu o programa *Vicon2OSC*, que faz a conversão da informação enviada por *streaming* do *Vicon Blade* para o protocolo *OSC*, permitindo definir o endereço IP e porta de

envio, de modo a que os dados sejam corretamente enviados a um outro computador (Figura 49).

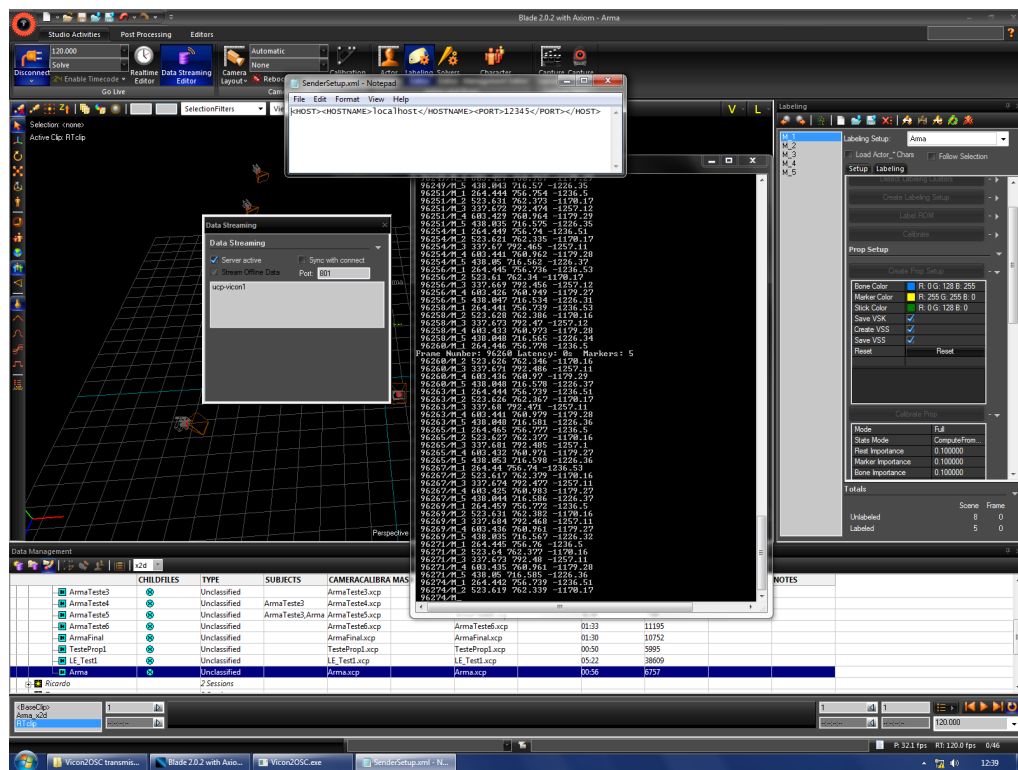


Figura 49 – Captura de ecrã dos programas Vicon e Vicon2OSC.

Tendo em conta a eficiência do sistema *Vicon*, considerou-se que tirar partido deste método de *motion capture* seria extremamente vantajoso para o projeto *The Sound of Horror*, visto que tem a capacidade de tornar a interação com o jogo mais natural e orgânica que a maioria dos outros sistemas testados, que dependem de apenas uma câmara para a detecção da posição da arma.

Graças à ajuda do Professor Ricardo Megre foi possível aprender o modo de funcionamento do *Vicon Blade*. Para a detecção e reconhecimento de um objeto definido por marcadores, é necessário proceder a um conjunto de passos. Foi observado que a colocação de marcadores em posições assimétricas permite melhorar o reconhecimento e *tracking* dos objetos. Com o aumentar do número de marcadores melhora também o seu desempenho. A Figura 50 apresenta a interface do jogo equipada com um conjunto de marcadores.

1. Em primeiro lugar, é necessário calibrar o espaço de captura. Nomeadamente, definir o plano de chão, o ponto de origem e o volume do espaço, para garantir a consistência do *tracking* em qualquer localização

2. Seguidamente, efetua-se uma captura com o objeto que se pretende registar. É necessário garantir que todas as câmaras captam todos os marcadores do objeto;
 3. Utilizando a captura previamente feita, é definido o novo objeto e feita a sua calibração;
- Após estar o objeto calibrado na sessão ativa, ao iniciar uma nova captura, este é automaticamente reconhecido (Figura 51).

Recorrendo à opção de *streaming* e ao *Vicon2OSC*, foi possível proceder ao envio dos valores dos vários marcadores em tempo real para o *Max/MSP*.

Do conjunto de abordagens estudadas para o *tracking* da posição da interface de jogo, o sistema de captura de movimentos *Vicon* revelou-se o único viável para o projeto, proporcionando resultados excelentes a nível de precisão e latência.



Figura 50 – Interface de jogo com marcadores do sistema Vicon.

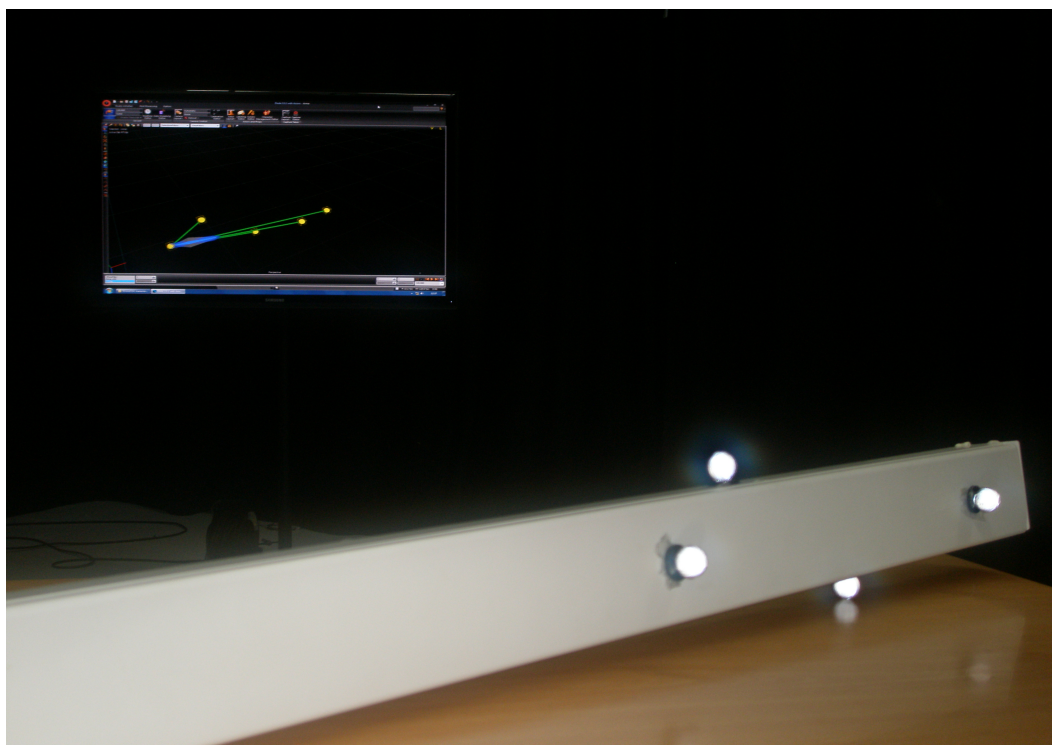


Figura 51 – Fotografia do reconhecimento e tracking automático do objeto calibrado pelo sistema Vicon.

3.4.6.6. Cálculo do ângulo de mira

A partir das coordenadas cartesianas obtidas do sistema *Vicon*, é necessário fazer a conversão destes valores para o ângulo de mira do utilizador, tendo conta os movimentos de rotação e translação que surgem naturalmente da interação com o jogo. Sendo o sistema sonoro bidimensional, não é necessário ter em conta a coordenada z dos marcadores da interface para o cálculo do ângulo de mira. Utilizando as coordenadas de dois marcadores que formam uma linha com a mira da arma, é possível definir uma recta, que vai intersectar o círculo formado pelos altifalantes. Essa intersecção irá ocorrer em dois pontos, sendo um deles o local correto do disparo e o outro o ponto exatamente atrás da arma. Para determinar a posição correta, é possível calcular a distância entre ambos os pontos de intersecção e a posição dos dois marcadores. Aquele que se encontrar mais próximo do marcador frontal irá corresponder ao ponto correto. Finalmente, conversão das coordenadas cartesianas desse ponto para coordenadas esféricas permitem determinar o ângulo de mira.

O utilizador *Chmike* (2009) publicou um algoritmo que permite determinar os pontos de intersecção de uma recta definida por dois pontos $A (A_x, A_y)$ e $B (B_x, B_y)$ com um círculo de centro em (C_x, C_y) e com raio R . Caso haja intersecção em dois pontos, estes são definidos

por $F (F_x, F_y)$ e $G (G_x, G_y)$. Foi feita a implementação deste algoritmo no *Max/MSP*, permitindo definir os valores das coordenadas dos pontos A, B, C e o raio R . A partir o valor do ponto de intersecção F , é calculada a distância entre A e F e a distância entre B e F . Com estes valores de distância, é possível determinar o ponto de intersecção mais próximo do ponto A . Para a conversão das coordenadas cartesianas em coordenadas esféricas, foi utilizado uma componente do sistema *Ambisonics* da *ICST*, e isso, determinar o ângulo de mira.

3.4.7. Componente sonora

Tendo em conta o cariz interativo e à não-linearidade da experiência de jogo, há um conjunto de elementos sonoros que estão dependentes do modo como o utilizador reage e às opções que toma (Collins, 2008). A componente sonora do jogo *The Sound of Horror* foi planeada e desenvolvida de modo a responder a estas necessidades.

O *Apêndice 2* contém o conjunto de elementos sonoros do jogo em conjunto com um *trailer* sonoro, que apesar de não transmitir a experiência completa da instalação, nem apresentar todos os conteúdos sonoros desenvolvidos, permite compreender o ritmo e tipo de sonoridade do projeto.

“Não é apenas pelo jogador ouvir algo com que identifica como um *zombie* que fica assustado, mas por escutar um gemido ou um grunhido, que corresponde aos elementos sonoros que reconhece como símbolos de horror. Desta forma, não é pelo facto da lamentação ser gerada por um *zombie* e por conter frequências graves que é assustadora, mas por, na sua essência, conter uma energia remanescente de uma forma de dor e agonia. Os sons ambiente têm um efeito semelhante, associando qualidades acústicas com situações ou locais desagradáveis e assustadores” (Roux-Girard, 2010).

Não dando respostas a perguntas que surgem naturalmente da interação com o jogo e recorrendo a elementos sonoros ambíguos, também permite aumentar o grau de confusão e medo (Roux-Girard, 2010).

Nas várias etapas do jogo, são necessários conjuntos de sons diferentes com características distintas. A lista completa de sons necessários para o projeto, organizados por categoria e referenciados por etapa de jogo, é apresentada na Tabela 8.

Tabela 8 – Lista de elementos sonoros do projeto *The Sound of Horror*.

Categoria	Nome do elemento sonoro	Número de canais	Loudness (1-5)	Espacialização	Categoria segundo Jørgensen (2010)	Colaboração	Etapa
Diálogos	MenuDiálogos_1	<i>Mono</i>	3	E1	J1	C1	Menu
	MenuDiálogos_2	<i>Mono</i>	3	E1	J1	C1	
	MenuDiálogos_3	<i>Mono</i>	3	E1	J1	C1	
	MenuDiálogos_4	<i>Mono</i>	3	E1	J1	C1	
	TutorialDiálogos_1	<i>Mono</i>	3	E1	J1	C1	Tutorial
	TutorialDiálogos_2	<i>Mono</i>	3	E1	J1	C1	
	TutorialDiálogos_3	<i>Mono</i>	3	E1	J1	C1	
	TutorialDiálogos_4	<i>Mono</i>	3	E1	J1	C1	
	TutorialDiálogos_5	<i>Mono</i>	3	E1	J1	C1	
	TutorialDiálogos_6	<i>Mono</i>	3	E1	J1	C1	
	TutorialDiálogos_7	<i>Mono</i>	3	E1	J1	C1	
	CinematicJogoDiálogos_1	<i>Mono</i>	3	E2	J2	C1	<i>Cinematic</i> Jogo
	Nível1Diálogos_1	<i>Mono</i>	3	E2	J2	C1	Nível 1
	Nível2Diálogos_1	<i>Mono</i>	3	E2	J2	C1	Nível 2
	Nível3Diálogos_1	<i>Mono</i>	3	E2	J2	C1	Nível 3
	VitóriaDiálogos_1	<i>Mono</i>	3	E2	J2	C1	<i>Cinematic</i> Vitória
	GameOver	<i>Stereo</i>	3	E1	J1	C1	<i>Cinematic</i> Derrota
LevelCompleted	<i>Mono</i>	3	E1	J1	C1	Nível 1, 2, 3	
Arma	ArmaTiroMenu_1	<i>Mono</i>	2	E2	J3	C2	Menu
	ArmaTiroMenu_2	<i>Mono</i>	2	E2	J3	C2	
	ArmaTiroMenu_3	<i>Mono</i>	2	E2	J3	C2	
	ArmaTiro_1	<i>Mono</i>	3	E2	J4	Nível 1, 2, 3 e Tutorial	
	ArmaTiro_2	<i>Mono</i>	3	E2	J4		
	ArmaTiro_3	<i>Mono</i>	3	E2	J4		
	ArmaTiro_4	<i>Mono</i>	3	E2	J4		
	ArmaTiro_5	<i>Mono</i>	3	E2	J4		
	ArmaReload_1	<i>Mono</i>	3	E2	J4		
	ArmaReload_2	<i>Mono</i>	3	E2	J4		
	ArmaReload_3	<i>Mono</i>	3	E2	J4		
	ArmaStab_1	<i>Mono</i>	3	E2	J4		
	ArmaStab_2	<i>Mono</i>	3	E2	J4		
	ArmaStab_3	<i>Mono</i>	3	E2	J4		
	ArmaTiroEmpty_1	<i>Mono</i>	3	E2	J4		
ArmaTiroEmpty_2	<i>Mono</i>	3	E2	J4			
ArmaTiroEmpty_3	<i>Mono</i>	3	E2	J4			
Ambientes	Menu_1	<i>Stereo</i>	2	E3	J1		Menu
	Menu_2	<i>Stereo</i>	2	E3	J1		
	Nível1_1	<i>Stereo</i>	2	E3	J2	Nível 1	
	Nível1_2	<i>Stereo</i>	2	E3	J2		
	Nível2_1	<i>Stereo</i>	2	E3	J2	Nível 2	
	Nível2_2	<i>Stereo</i>	2	E3	J2		
	Nível3_1	<i>Stereo</i>	2	E3	J2	Nível 3	
	Nível3_2	<i>Stereo</i>	2	E3	J2		
	Tutorial_1	<i>Stereo</i>	2	E3	J2	Tutorial	
Tutorial_2	<i>Stereo</i>	2	E3	J2			
Cinematics	CinematicInício_1	<i>Stereo</i>	4	E3	J1	<i>Cinematic</i> Início	
	CinematicInício_2	<i>Stereo</i>	4	E3	J1		
	CinematicTutorial_1	<i>Stereo</i>	4	E3	J1	<i>Cinematic</i> Tutorial	

Categoria	Nome do elemento sonoro	Número de canais	Loudness (1-5)	Espacialização	Categoria segundo Jørgensen (2010)	Colaboração	Etapa	
	CinematicTutorial_2	<i>Stereo</i>	4	E3	J1		Cinematic Tutorial - Menu	
	CinematicTutorial-Menu_1	<i>Stereo</i>	4	E3	J1			
	CinematicTutorial-Menu_2	<i>Stereo</i>	4	E3	J1			
		CinematicJogo_1	<i>Stereo</i>	4	E3	J1		Cinematic Jogo
		CinematicJogo_2	<i>Stereo</i>	4	E3	J1		
		Cinematic1-2_1	<i>Stereo</i>	4	E3	J1		Cinematic Nível 1-2
		Cinematic1-2_2	<i>Stereo</i>	4	E3	J1		
		Cinematic2-3_1	<i>Stereo</i>	4	E3	J1		Cinematic Nível 2-3
		Cinematic2-3_2	<i>Stereo</i>	4	E3	J1		
		Vitória_1	<i>Stereo</i>	4	E3	J1		Cinematic Vitória
		Vitória_2	<i>Stereo</i>	4	E3	J1		
		Derrota_1	<i>Stereo</i>	4	E3	J1		Cinematic Derrota
Derrota_2		<i>Stereo</i>	4	E3	J1			
Menu	MenuAlvo1	<i>Mono</i>	3	E2	J3	C2	Menu	
	MenuAlvo1_Hit	<i>Mono</i>	3	E2	J3			
	MenuAlvo2	<i>Mono</i>	3	E2	J3	C2		
	MenuAlvo2_Hit	<i>Mono</i>	3	E2	J3			
Monstros	Monstro1Nível1	<i>Mono</i>	3	E2	J2	C1	Nível 1	
	Monstro1Nível1Ferido	<i>Mono</i>	3	E2	J2	C1		
	Monstro1Nível1Morto	<i>Mono</i>	3	E2	J2	C1		
	Monstro1Nível1Passos	<i>Mono</i>	3	E2	J2			
	Monstro2Nível1	<i>Mono</i>	3	E2	J2	C1		
	Monstro2Nível1Ferido	<i>Mono</i>	3	E2	J2	C1		
	Monstro2Nível1Morto	<i>Mono</i>	3	E2	J2	C1		
	Monstro2Nível1Passos	<i>Mono</i>	3	E2	J2			
	Nível 2	Monstro3Nível1	<i>Mono</i>	3	E2	J2	C1	
		Monstro3Nível1Ferido	<i>Mono</i>	3	E2	J2	C1	
		Monstro3Nível1Morto	<i>Mono</i>	3	E2	J2	C1	
		Monstro3Nível1Passos	<i>Mono</i>	3	E2	J2		
		Monstro1Nível2	<i>Mono</i>	3	E2	J2	C1	
		Monstro1Nível2Ferido	<i>Mono</i>	3	E2	J2	C1	
		Monstro1Nível2Morto	<i>Mono</i>	3	E2	J2	C1	
		Monstro1Nível2Passos	<i>Mono</i>	3	E2	J2		
	Nível 3	Monstro2Nível2	<i>Mono</i>	3	E2	J2	C1	
		Monstro2Nível2Ferido	<i>Mono</i>	3	E2	J2	C1	
		Monstro2Nível2Morto	<i>Mono</i>	3	E2	J2	C1	
		Monstro2Nível2Passos	<i>Mono</i>	3	E2	J2		
		Monstro3Nível2	<i>Mono</i>	3	E2	J2	C1	
		Monstro3Nível2Ferido	<i>Mono</i>	3	E2	J2	C1	
		Monstro3Nível2Morto	<i>Mono</i>	3	E2	J2	C1	
		Monstro3Nível2Passos	<i>Mono</i>	3	E2	J2		
	Monstro1Nível3	<i>Mono</i>	3	E2	J2	C1		
	Monstro1Nível3Ferido	<i>Mono</i>	3	E2	J2	C1		
	Monstro1Nível3Morto	<i>Mono</i>	3	E2	J2	C1		
	Monstro1Nível3Passos	<i>Mono</i>	3	E2	J2			
	Monstro2Nível3	<i>Mono</i>	3	E2	J2	C1		
	Monstro2Nível3Ferido	<i>Mono</i>	3	E2	J2	C1		
	Monstro2Nível3Morto	<i>Mono</i>	3	E2	J2	C1		
	Monstro2Nível3Passos	<i>Mono</i>	3	E2	J2			
	Monstro3Nível3	<i>Mono</i>	3	E2	J2	C1		
	Monstro3Nível3Ferido	<i>Mono</i>	3	E2	J2	C1		
	Monstro3Nível3Morto	<i>Mono</i>	3	E2	J2	C1		

Categoria	Nome do elemento sonoro	Número de canais	Loudness (1-5)	Espacialização	Categoria segundo Jørgensen (2010)	Colaboração	Etapa
	Monstro3Nível3Passos	<i>Mono</i>	3	E2	J2		
	Monstro4Nível3	<i>Mono</i>	3	E2	J2	C1	
	Monstro4Nível3Ferido	<i>Mono</i>	3	E2	J2	C1	
	Monstro4Nível3Morto	<i>Mono</i>	3	E2	J2	C1	
	Monstro4Nível3Passos	<i>Mono</i>	3	E2	J2		
Tutorial	TutorialAlvo1	<i>Mono</i>	3	E2	J2		Tutorial
	TutorialAlvo1Hit	<i>Mono</i>	3	E2	J2		
	TutorialAlvo2	<i>Mono</i>	3	E2	J2		
	TutorialAlvo2Hit	<i>Mono</i>	3	E2	J2		

Loudness determina a intensidade do elemento sonoro na mistura, sendo 1 o mínimo e 5 o máximo (Wilhelmsson & Wallén, 2010).

Espacialização

- E1, Sinal sonoro enviado diretamente para todos os altifalantes;
- E2, Síntese espacial *Ambisonics*;
- E3, Sinal sonoro enviado diretamente para os altifalantes de forma cruzada.

Categoria segundo Jørgensen (2010)

- J1, *Metaphorical Interface Sounds*;
- J2, *Iconic Interface Sounds*;
- J3, *Overlay Interface Sounds*;
- J4, *Emphasized Interface Sounds*;

Colaboração

- C1, Captação de som feita com o apoio de Adrian Santos;
- C2, Som parcialmente produzido por Adrian Santos.

Para o processo de captação, síntese e edição dos vários elementos recorreu-se ao uso dos microfones *Rode NT1-A* e *Rode K2*. A *Digital Audio Workstation* utilizada foi o *software ProTools HD9*. Recorreu-se ainda a *plugins* adicionais, nomeadamente o *Absynth 5*, o *Izotope Iris* e o *Waves IR-1*.

Todos os elementos sonoros foram equalizados, de modo a realçar ou atenuar certas frequências, para se tornarem mais apelativos. Foi também aplicada compressão com o objetivo de controlar a sua gama dinâmica.

3.4.7.1. Diálogos

Para a sonorização dos diálogos do Narrador, foi utilizada uma entoação feminina e fria com características robóticas, à semelhança da personagem *GLaDOS* do jogo *Portal*. No entanto, contrasta com o restante ambiente, sendo o elemento mais humano apresentado ao utilizador. Isto permite criar empatia com esta entidade. Recorreu-se ao uso de ferramentas de

pitch shift e *frequency shift* para atribuir estas qualidades femininas e robóticas às faixas sonoras captadas. No som “LevelCompleted” foi adicionado uma alerta sonoro desenvolvido através da conjugação de sinais audíveis emitidos por computadores e outros aparelhos electrónicos de modo a salientar o terminar de nível.

O Grande Irmão é uma criatura cuja raça é indistinta e que não trespassa sentimentos, tendo um discurso relativamente monocórdico. Esta sonoridade foi obtida através da duplicação da faixa de diálogo e do uso de técnicas de *pitch shift* e *ring modulation*. Para aumentar a sensação de distância entre esta entidade e o utilizador, foi aplicado um efeito de ruído de intercomunicador em cada fala do Grande Irmão, de modo a situá-lo dentro do ambiente de jogo, como se a sua voz surgisse de altifalantes situados no interior do universo virtual.

Apesar de não estar enquadrado na categoria de Diálogos, no som “CinematicInício_1” foi introduzido um diálogo que pretende fazer a apresentação do jogo. À semelhança do “GameOver”, este desenvolvido através da triplicação da faixa sonora e do uso de *pitch shifting* de modo a atribuir-lhe uma sonoridade grotesca e monstruosa, mantendo a inteligibilidade do discurso. Recorreu-se também ao uso de rugidos de feras (e.g. leões, tigres, cães) que, em sincronia com as falas, permitiu ampliar esta dimensão animalésca. No caso do elemento sonoro “GameOver”, em específico, utilizou-se o som de um impacto de grandes dimensões em conjunto com uma sonoridade turbulenta e misteriosa.

3.4.7.2. Arma

Os elementos relativos à categoria Arma contêm um conjunto de variações suaves para cada uma das componentes sonoras. Estas variações têm como objetivo tornar a experiência de jogo mais natural, não alertando o utilizador para o facto de que está a inicializar a leitura de ficheiros sonoros repetidamente. Foi utilizado *pitch shifting* de modo a gerar múltiplas variações suaves mas eficazes, a partir de apenas um ficheiro sonoro. Concluiu-se que seria necessário recorrer a cinco transformações do som do tiro da arma, sendo aquele o mais repetido no decorrer do jogo. Para os restantes, foram aplicadas apenas três.

A base dos sons de tiro no Menu (“ArmaTiroMenu”), construída por Adrian Santos, foi modificada, de modo a conferir-lhe uma sonoridade mais apropriada para uma um seletor à semelhança de um rato de computador, mas respeitando a natureza da interface. Para isso, foi

sobreposto o som de disparo de uma rolha de garrafa, dando-lhe o ataque necessário para tornar a ação de disparo inteligível.

Para a criação dos sons de tiro de arma (“ArmaTiro”), de tiro sem munições (“ArmaTiroEmpty”) e de mudança de carregador (“ArmaReload”), recorreu-se à sobreposição de vários sons de armas reais, de modo atribuir à arma de jogo um misto de sonoridade de espingarda e caçadeira. O esfaquear (“ArmaStab”) foi obtido através de sons de desembainhar de espadas, raspar de facas e de chicotear com uma vara.

3.4.7.3. Ambientes

Atendendo ao facto de que a experiência pretende colocar o utilizador numa posição de medo e desconforto, é necessário que os ambientes de cada etapa do jogo criem uma base sonora que permita prender o utilizador à realidade virtual. Os elementos sonoros desenvolvidos apresentam uma duração aproximada de 1 minuto de modo a não se tornarem repetitivos para o jogador.

Para a etapa de Menu, foi desenvolvido um ambiente grave com alguns elementos agudos. Tendo em conta que esta etapa, apesar de estar dentro do universo de jogo, se assemelha a um menu de seleção de um jogo convencional, torna-se necessário preparar o utilizador para o contexto de jogo.

Os ambientes nos três níveis de jogo pretendem transportar o utilizador para um ambiente à semelhança de uma caverna e um laboratório subterrâneo. Não são dadas muitas pistas auditivas relativamente às características acústicas do espaço de modo a não permitir ao utilizador formar uma imagem definida do local em que se encontra, para aumentar o grau de confusão. Com o passar de nível, os sons ambiente apresentam uma tensão crescente, de modo a salientar o aumento de dificuldade dos níveis.

Para permitir ao utilizador aprender as funcionalidades da interface sem sentir a pressão do modo de jogo, foi desenvolvida uma sonoridade ambiente para a etapa Tutorial com características mais descontraídas que os restantes ambientes desenvolvidos. Para isto foram utilizados sons menos graves e com características metálicas e electrónicas, assemelhando o espaço de aprendizagem a uma área artificial, criada com o propósito da instrução do jogador.

Foram desenvolvidas duas faixas sonoras para cada ambiente, com o objetivo de serem reproduzidas em simultâneo.

3.4.7.4. Cinematics

Os elementos sonoros *Cinematics* têm como objetivo salientar a transição entre etapas de jogo, realçando características que advêm dessa passagem e transmitindo informações ao utilizador relativas ao seu percurso no jogo *The Sound of Horror*. Sendo um jogo sem componente visual, o realce das transições permite ajudar a navegar no jogo, sendo que não há um suporte visual que torne estas mudanças óbvias.

O *Cinematic* Início consiste num acumular progressivo de energia constituído por uma série de diferentes elementos sonoros (*e.g.* gritos humanos, elementos musicais), até atingir um auge, deixando o utilizador numa posição submissa perante a brusquidão e o impacto do elemento sonoro apresentado. De seguida é feita a apresentação do jogo, através da fala “*Welcome to the Sound of Horror.*”

Para as transições entre o Menu e os dois percursos alternativos, foi desenvolvida uma sonoridade que se assemelha ao inicializar de um sistema tecnologicamente avançado, criando uma sensação de transporte virtual para uma outra realidade. A título de exemplo, a Figura 52 apresenta a sessão *ProTools HD9* relativa a estes elementos sonoros.

Revelou-se necessário criar uma transição inteligível entre cada nível de jogo, mantendo uma ligação sonora com ambiente, de forma a não criar quebras imersivas na experiência. Foi criada uma sonoridade que se assemelha a um feixe de energia eléctrica a percorrer o espaço de jogo, como que a recarregar o sistema para o nível seguinte. Gradualmente, esta sonoridade começa a assemelhar-se aos ambientes dos níveis de jogo, permitindo uma transição suave.

Os *Cinematics* que pretendem transportar o utilizador do modo de Jogo e Tutorial de volta para o Menu, assemelham-se ao desligar de um sistema de grandes dimensões, simbolizando o terminar da porção da experiência. Para realçar este efeito, foi utilizado o som invertido do bater de asas de pássaros e de uma turbina de avião a desacelerar para realçar este efeito.

À semelhança da categoria de Ambientes, os *Cinematics* apresentam duas faixas *stereo* por componente, com vista a explorar as capacidades espaciais da distribuição de altifalantes.

3.4.7.5. Narrador

A partir da base sonora criada por Adrian Santos para os dois alvos presentes no Menu, correspondentes às opções de seleção de modo de Jogo ou Tutorial, foi feita a adição de elementos sonoros de forma a proporcionar-lhe a aparência de uma bolha sonora.

A sonorização do disparo sobre estes alvos consistiu na sobreposição do ressoar de um elemento metálico agudo e o som de borbulhar num líquido, para simular o rebentar da bolha.

Para cada nível foram concebidos monstros com características diferentes com o objetivo de tornar a experiência gradualmente mais complexa sem se tornar repetitiva. Foi feito um conjunto de gravações de rugidos recorrendo a instrumentos, como funis e garrafas, com o objetivo de alterar ressonância da voz humana.

No decorrer do jogo, os monstros movem-se são expostos ao ataque do utilizador, sendo necessário criar faixas sonoras separadas para os seus rugidos, passos e sons de ferimento e morte, de modo a permitir ajustar o seu comportamento às ações do utilizador.

O primeiro nível é constituído por três monstros lentos e de grandes dimensões. Para emular este tipo de criaturas, foi triplicada a faixa sonora e foram utilizadas ferramentas de *pitch shift* de modo a criar a tonalidade pretendida. Os monstros do segundo nível de jogo apresentam dimensões mais pequenas, mas uma atitude voraz e uma rapidez de movimentos superior aos anteriores. Foi utilizado o mesmo processo na sonorização destas criaturas. O último nível apresenta um conjunto de monstros mais diverso, contendo criaturas com uma sonoridade semelhante à de seres fantasmagóricos, lagartos e outros monstros. Utilizou-se reverberação invertida para o desenvolvimento de alguns destes sons.

Para os passos dos diferentes monstros, foram aplicados efeitos de *time stretch* e *pitch shift* de modo a assemelhar passos humanos aos das diferentes criaturas.

Os alvos apresentados no Tutorial do jogo assemelham-se a pequenos seres engraçados. Foi escolhido um tipo de alvo com uma sonoridade amigável de modo a não colocar o utilizador numa posição de medo, permitindo-lhe concentrar-se na manipulação da interface de jogo.

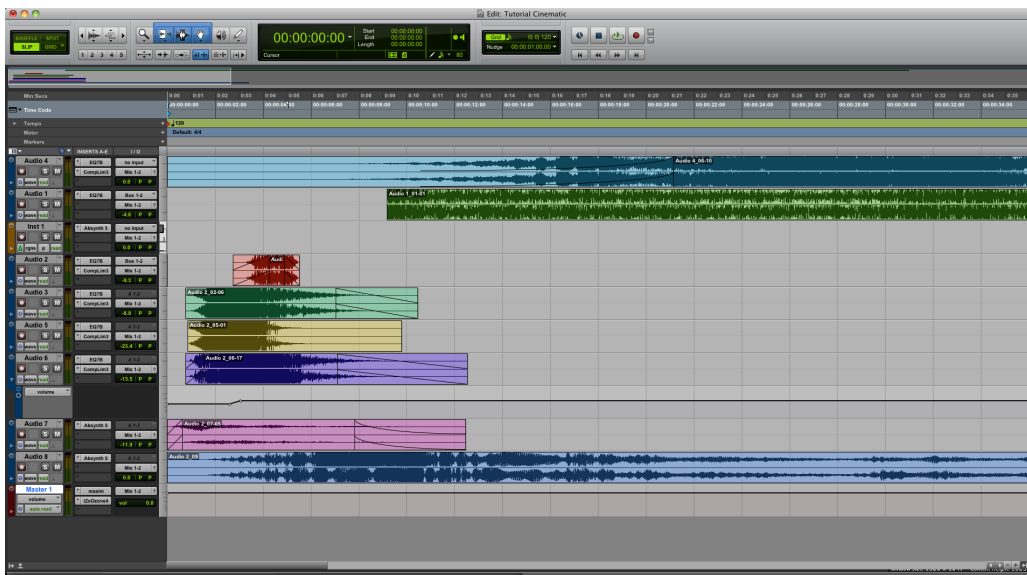


Figura 52 – Sessão ProTools HD9 dos elementos sonoros relativos ao Cinematic Tutorial.

3.4.7.6. Categorização segundo Kristine Jørgensen (2010)

Jørgensen (2010) propõe uma forma de categorização dos elementos sonoros de um videogame com base na sua relação com o utilizador e universo de jogo. Esta categorização advém da incapacidade de aplicar a distinção tradicional entre sons diegéticos e não-diegéticos a um contexto de jogo da mesma forma que a um tipo de média passivo.

As categorias sugeridas são *Metaphorical Interface Sounds*, que se refere a elementos sonoros externos ao universo de ficção do jogo que dão informação de usabilidade ao utilizador. *Overlay Interface Sounds* pretende classificar sons exteriores ao universo de ficção que estão associados à interface de jogo. *Integrated Interface Sounds* engloba os sons que estão ligados a elementos da interface que se encontram dentro do universo de jogo. *Emphasized Interface Sounds* categoriza sons acionados pelo utilizador que se encontram dentro do universo do jogo. Por fim, *Iconic Interface Sounds* pretende identificar sons gerados pelo sistema de jogo, encontrando-se completamente integrados no universo de ficção do jogo.

A Tabela 8 apresenta a categorização dos elementos sonoros do projeto *The Sound of Horror* com base no modelo sugerido por Jørgensen.

3.4.7.7. Espacialização sonora

A espacialização dos elementos sonoros do jogo foi planeada de acordo com o papel desempenhado por cada som, tal como apresentado na Tabela 8. Foi escolhido o sistema *Ambisonics* para a síntese espacial de elementos provenientes de locais específicos campo sonoro (categoria “E2”). Como observado no capítulo 2.2.3., o método de codificação do sinal nesta técnica não permite a criação de fontes sonoras cuja origem seja o centro do campo de escuta. Para estas situações, foi enviado sinal diretamente para os altifalantes, não sendo necessário qualquer processamento espacial (categorias “E1” e “E3”). Apesar de isto poder levantar alguns problemas em situações onde o número de colunas seja inconstante, esta abordagem revelou-se a alternativa mais viável, dada a natureza deste projeto.

De modo a explorar a configuração do sistema de difusão sonora, foi feita uma distribuição do sinal dos elementos (na Tabela 8, categoria “E3”) à semelhança do apresentado na Figura 53. As duas camadas de som, desenhadas para serem reproduzidas simultaneamente, foram sobrepostas de forma cruzada nos altifalantes, permitindo criar uma sonoridade consistente e com características espaciais.

A distribuição e envio de sinal para o canal *LFE* (Figura 36) foi conseguida no *Max/MSP* através da soma dos vários canais e da aplicação de um filtro *crossover* de terceira ordem (18 dB/oitava) aos 80 Hz.

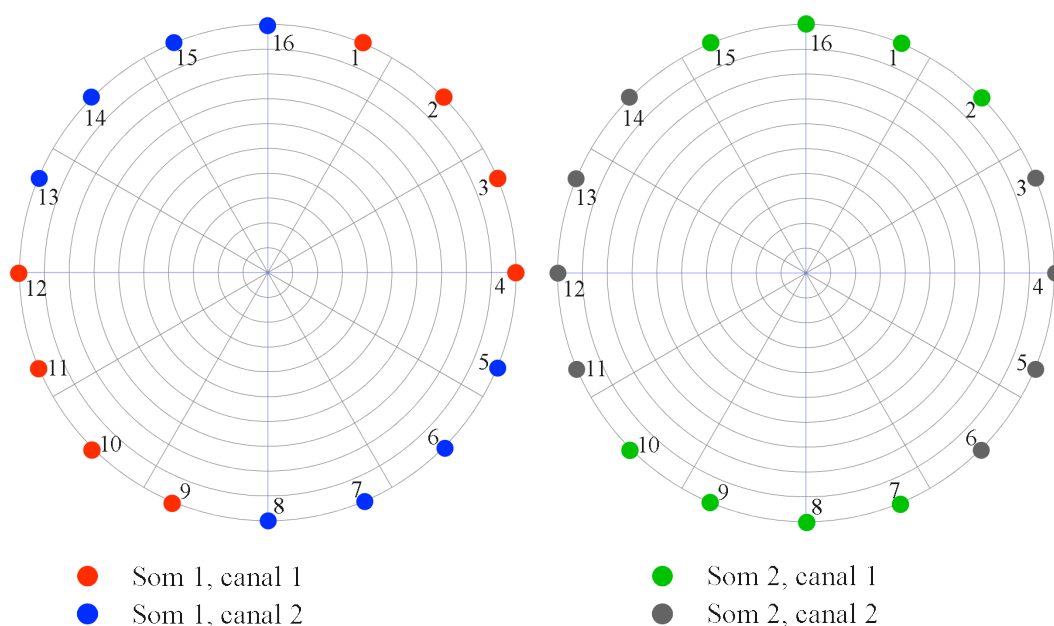


Figura 53 – Representação da distribuição do sinal sonoro das categorias *Ambientes* e *Cinematics*.

De acordo com a *framework* proposta por Wozniowski *et al.* (2006) e com as conclusões obtidas do capítulo 2.1.2., procedeu-se à implementação de uma cadeia de processamento de sinal dos elementos sonoros situados no interior do espaço virtual, concedendo-lhes características espaciais adequadas à realidade virtual sonora. Esta cadeia envolve a atenuação do sinal sonoro com o aumentar de distância (integrado no sistema *Ambisonics*), o aumento das frequências graves causado pelo efeito de proximidade, a atenuação das frequências agudas resultantes da absorção atmosférica, o efeito de *doppler* resultante do movimento de fontes sonoras relativamente ao ouvinte e a reverberação do espaço acústico. Foi feita uma implementação no *Max/MSP* de modo possibilitar o mapeamento destas características à posição das fontes sonoras.

Dependendo do elemento sonoro, determinadas componentes desta cadeia de processamento espacial apresentam-se estáticas. Por exemplo, o tiro da arma, apesar de estar dependente do ângulo de mira do utilizador, encontra-se sempre à mesma distância. Atendendo ao dinamismo de cada elemento, é possível economizar recursos computacionais, aplicando o processamento diretamente na faixa sonora, exteriormente ao jogo.

3.4.8. Componente de programação

Para o desenvolvimento da componente de programação do projeto, foi escolhido o uso da linguagem *Max/MSP*, de modo a unir os restantes elementos do trabalho (*i.e.* sistema de *tracking* de movimentos, interface de jogo, componente sonora) e gerir o fluxo de informação a ser enviado como *output* do sistema.

Para além dos objetos externos relativos ao comando *WiiMote* e ao sistema *Ambisonics* da *ICST*, recorreu-se ao uso do objeto *f0.smooth* desenvolvido por Fredrik Olofsson (2010).

O *Apêndice 2* apresenta o *patch* do jogo juntamente com os objetos externos utilizados.

O trabalho está estruturado por módulos interligados, que operam individualmente. Foi tomada esta abordagem com o objetivo de melhorar a legibilidade do *patch*, para aumentar a eficiência na implementação de novas características e na resolução de erros de programação. O *patch* do projeto *The Sound of Horror* é apresentado na Figura 54.

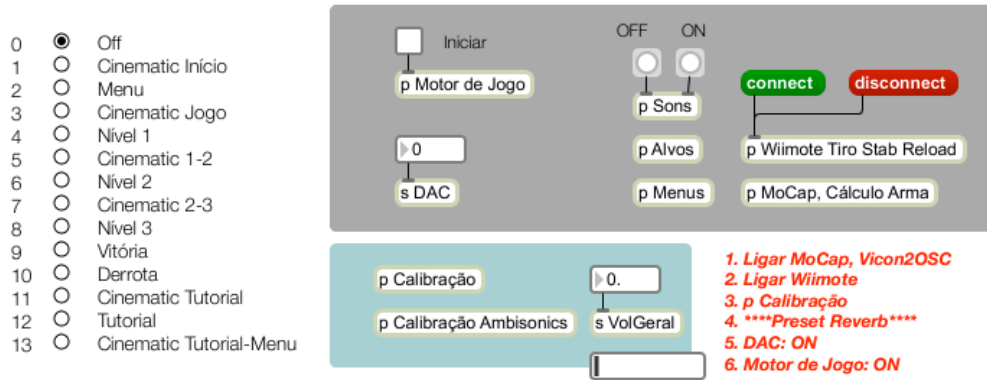


Figura 54 – Patch do projeto *The Sound of Horror*.

3.4.8.1. Interface

O módulo relativo à interface do jogo (Figura 55) contém o sistema de ligação ao comando *Wiimote* contido na arma e o reconhecimento das várias ações associadas à sua utilização. Dependendo das exigências de cada etapa do jogo, são ativadas ou desativadas as várias funcionalidades da interface, como é o caso da necessidade de recarregar a arma, que não se apresenta no Menu.

Foi feita a associação da funcionalidade de vibração do comando ao disparar e ao ataque dos monstros, de modo tornar a experiência multissensorial.

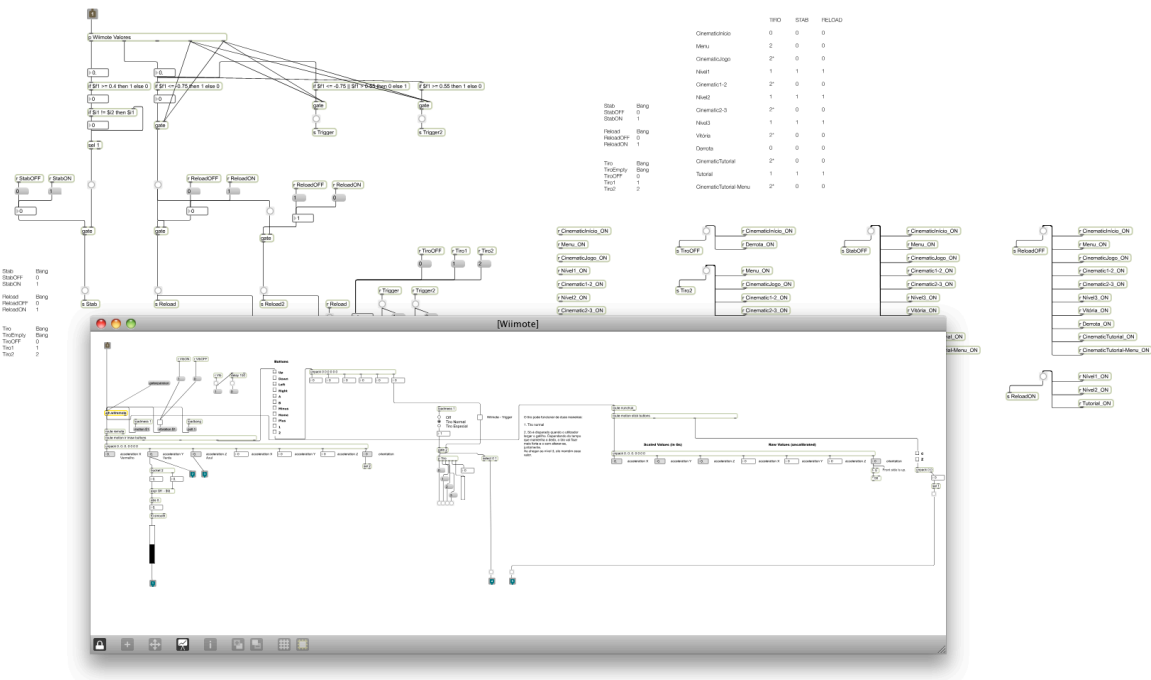


Figura 55 – Módulo de programação relativo à interface de jogo.

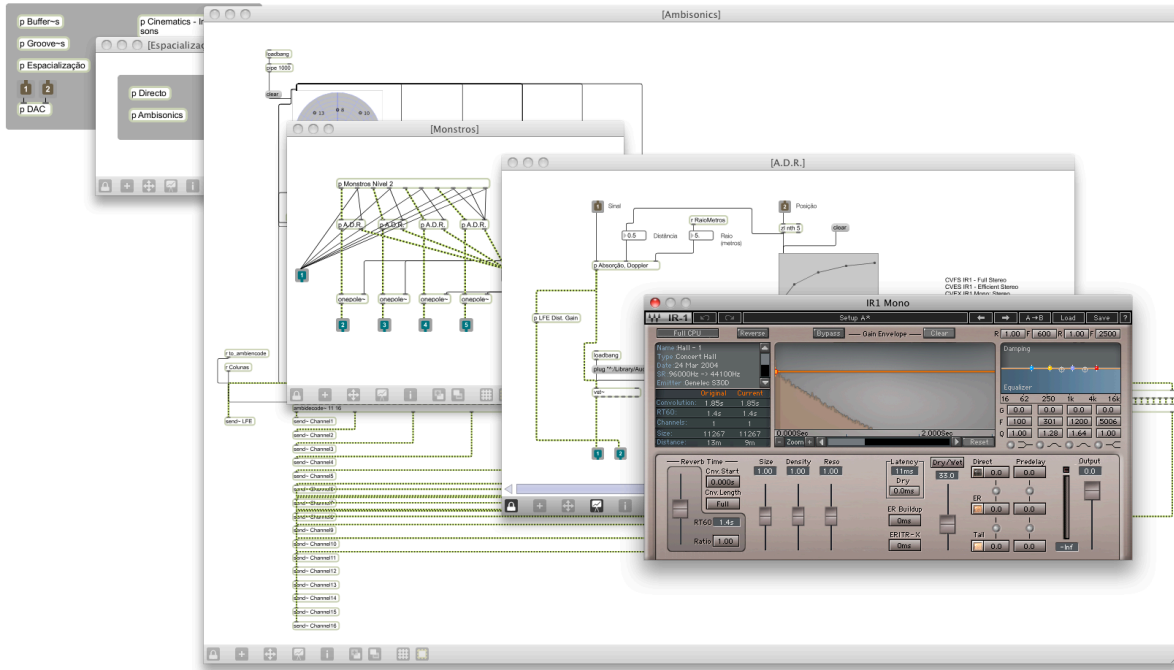


Figura 58 – Processamento dos elementos sonoros relativos aos monstros do Nível 2 com o plugin Waves IR-1.

3.4.8.4. Alvos sonoros

Foi desenvolvido um módulo designado para a criação dos alvos sonoros. Este permite gerar o trajeto e acionar os vários elementos sonoros associados ao alvo. Faz também a ligação entre o componente de interação da arma e a posição dos alvos, determinando se a ação de disparar ou esfaquear atingiu o alvo pretendido.

Os monstros presentes no Nível 1 do jogo apresentam trajetos simples, aproximando-se em linha recta ou circulando o utilizador no decorrer de aproximadamente 10 segundos. No segundo nível, os trajetos são idênticos, mas apresentam um tempo de aproximadamente 6 segundos. No último nível de jogo, os tempos são semelhantes aos do Nível 2, mas apresentam trajetos complexos, zigzagueando à medida que se aproximam do utilizador. Apesar do trajeto dos monstros ser predefinido, o ângulo em que surgem é aleatório, tornando o jogo em parte imprevisível.

É possível definir também o número de “pontos de vida” de cada alvo, ou seja, o número de vezes que é necessário atingi-lo para que seja abatido. Isto permite fazer a distinção entre os alvos do Menu e Tutorial e os monstros do modo de Jogo.

Cada alvo sonoro apresenta uma margem de erro de tiro. Isto permite atribuir aos monstros dimensões físicas, virtualmente, e tornar a experiência mais realista, mapeando a

distância do alvo ao valor de margem de erro de modo a que com o aproximar do monstro, este se torne um alvo mais fácil. A Figura 59 apresenta uma secção deste módulo relativo a um monstro do Nível 1.

Este módulo apresenta um sistema que inicia um contador quando um monstro atinge o utilizador. Após um determinado tempo, caso o monstro não seja eliminado, é dada a informação de que o utilizador foi abatido. Foi implementado um sistema que associa este contador a um *Low Pass Filter*, reduzindo a intensidade das frequências agudas à medida que o utilizador é atacado, de modo a tornar o som menos definido e aumentar o grau de confusão.

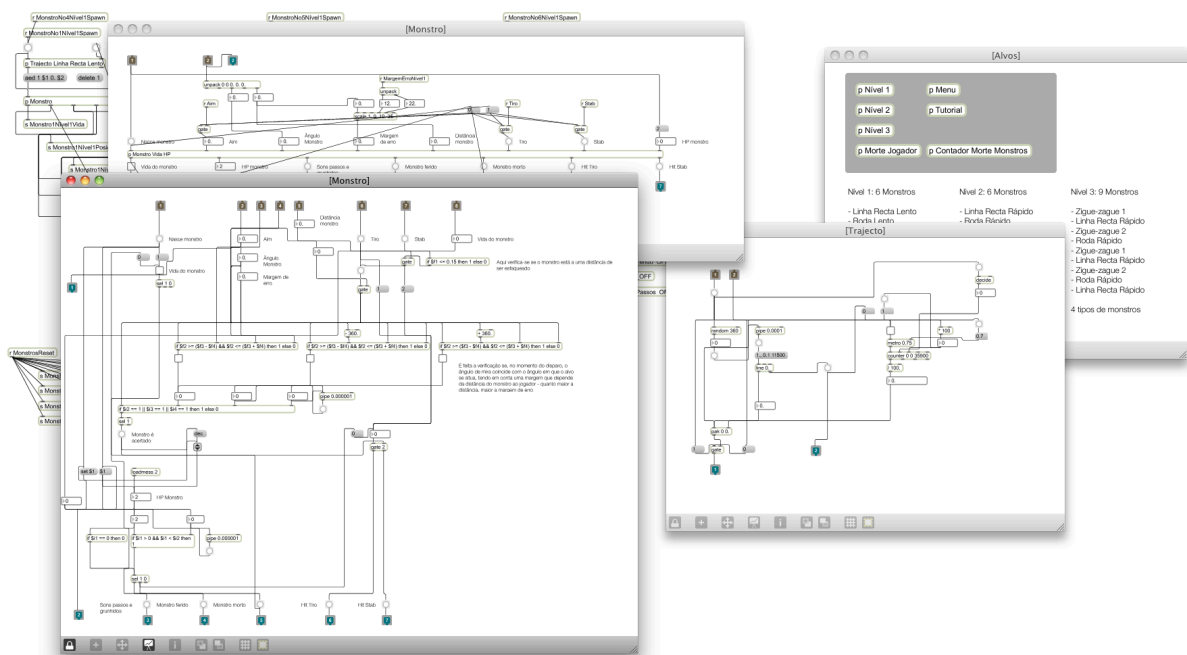


Figura 59 – Criação, geração da trajetória e cálculo dos “pontos de vida” de um monstro do Nível 1.

3.4.8.5. Etapas de jogo

Neste módulo foi feita uma associação das várias etapas de jogo aos elementos sonoros Ambientais e *Cinematics* que lhes dizem respeito, de modo a tornar a inicialização de sons em simultâneo mais simples (Figura 60).

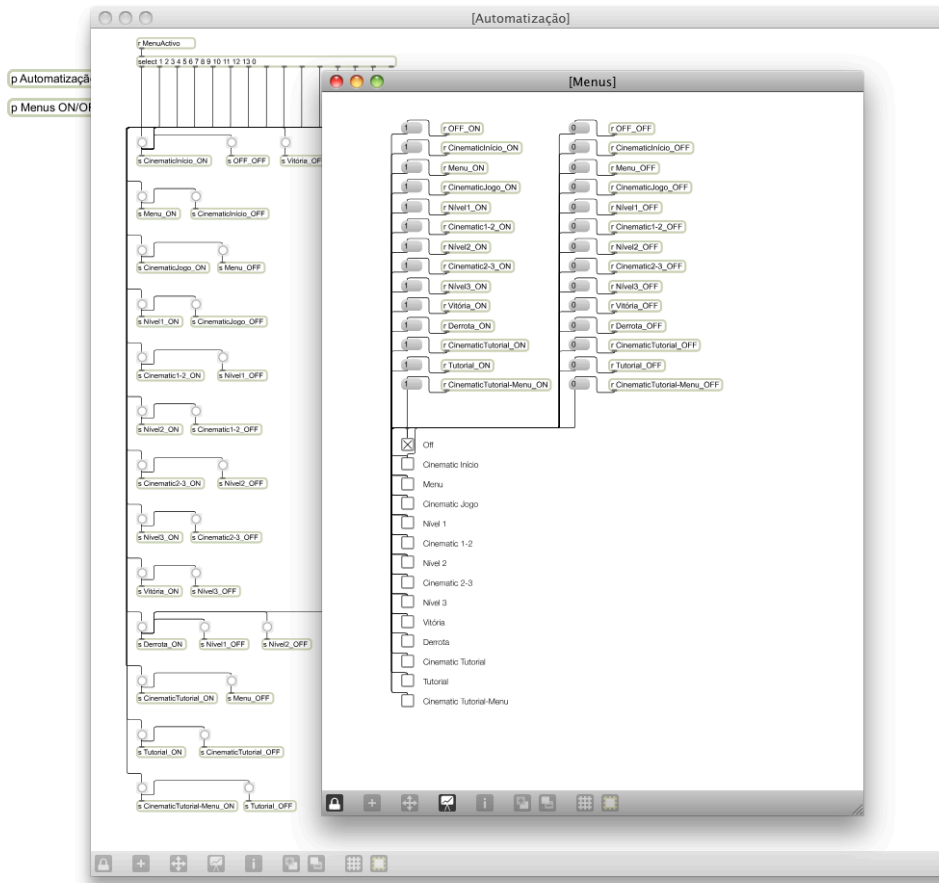


Figura 60 – Módulo de associação dos elementos sonoros das categorias Ambientais e Cinematics às etapas de jogo.

3.4.8.6. Motor de jogo

O motor de jogo é o elemento central do *patch* (Figura 61). Este controla a passagem entre etapas e define os *timings* com que os elementos sonoros vão ser inicializados. Envia também, para o módulo da interface, a informação para ativar ou desativar determinadas funcionalidades.

A partir do motor de jogo é também possível controlar o tempo que decorre entre o surgimento dos monstros, bem como o número de alvos sonoros por nível. Optou-se pelo uso de seis monstros nos Níveis 1 e 2 e nove monstros no Nível 3, sendo necessária a repetição dos elementos sonoros.

Recorreu-se ao uso de *game loops* (Valente *et al.*, 2005) para que o jogo ficasse parado até que receba informação, do utilizador ou do sistema, para prosseguir para a etapa seguinte. No caso do Tutorial, após dadas as instruções relativamente a uma das funcionalidades da interface, é esperada a resposta do utilizador. Caso decorra um período de tempo definido no

qual não é recebida tal informação, é dada instrução para repetir o elemento sonoro relativo ao *game loop*.

A partir deste módulo é possível reiniciar o jogo, revertendo todas as opções tomadas e preparando o sistema para uma nova utilização.

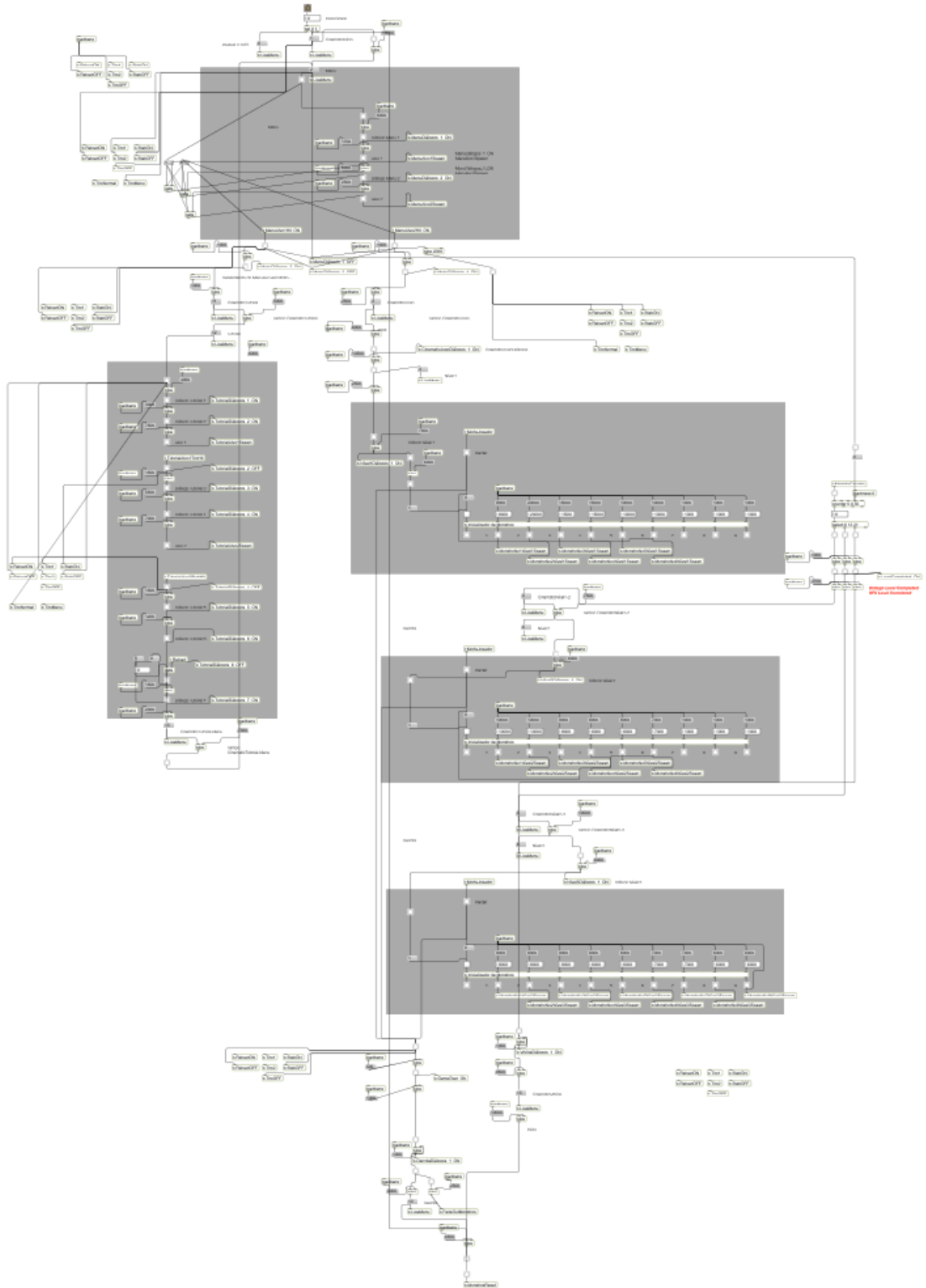


Figura 61 – Módulo de motor de jogo.

3.4.8.7. Calibração

O módulo de calibração (Figura 62) permite definir a posição e o número de altifalantes utilizados, comunicando com o sistema *Ambisonics* para a descodificação do sinal *B-format* e a frequência de *crossover* para o canal *LFE*. Através do sistema de calibração disponibilizado nos ficheiros de apoio das ferramentas *Ambisonics* do *ICST*, é possível definir a atenuação de distância e os valores de *weight* a utilizar no sistema.

O cálculo do sistema de captura de movimentos é feito através de um círculo que virtualmente irá coincidir com o sistema de difusão sonora. Neste módulo é possível definir o valor do raio deste círculo. Para garantir que o ângulo de mira do utilizador corresponde ao ângulo do sistema sonoro foi implementado um sistema de ajuste que unifica estas duas componentes.

É também possível controlar os valores de margem de erro para cada alvo sonoro, controlando o nível de dificuldade deste parâmetro, bem como o aumento de ganho do canal *LFE* com o aproximar do monstro.

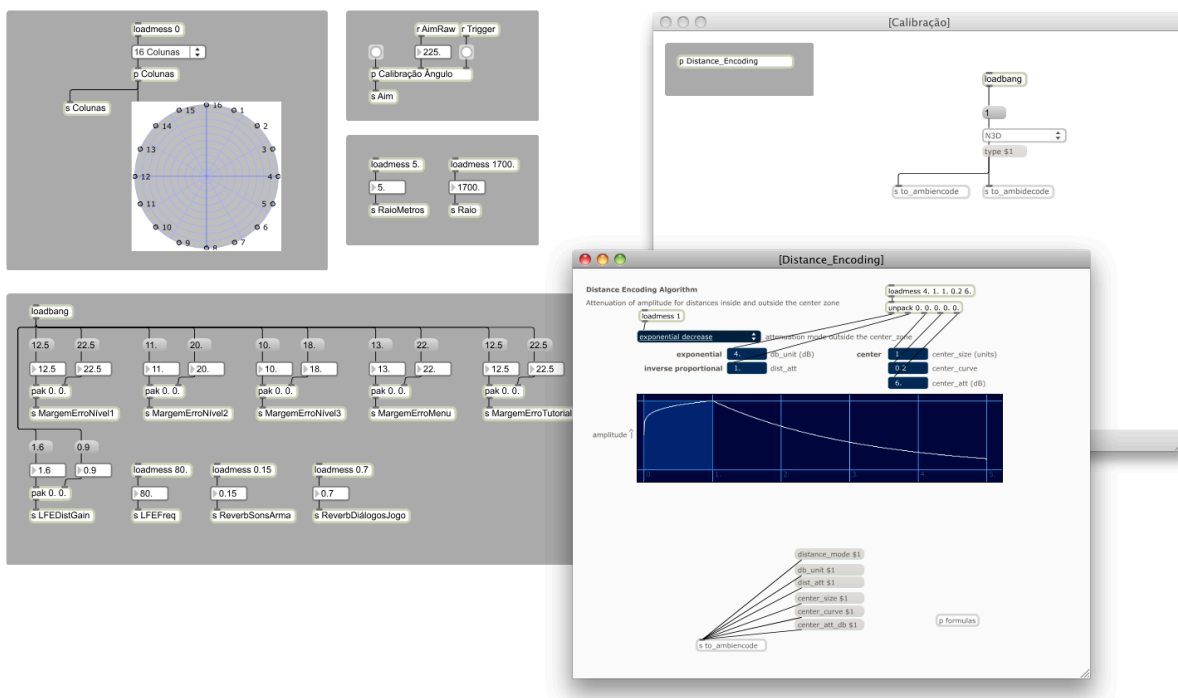


Figura 62 – Módulo de calibração do sistema de jogo.

3.5. Montagem e apresentação

Optou-se pelo uso do sistema de captura de movimentos *Vicon* para o posicionamento dos altifalantes, de modo a garantir uma localização precisa. Assumiu-se como raio da configuração circular, bidimensional, o valor de 2300 unidades do sistema *Vicon* (2,3 metros) e foi feito o cálculo da posição de 16 pontos igualmente distribuídos sobre a circunferência formada à semelhança da Figura 36. A Tabela 9 apresenta as coordenadas de cada altifalante. Foi colocada a estrutura para confinar o utilizador no centro do círculo e ajustada a orientação de cada altifalante para coincidir com o espaço de utilização.

A Figura 63 apresenta o diagrama de funcionamento do projeto, especificando a forma como a comunicação decorre entre cada componente e o modo como o fluxo de informação decorre. A componente sonora é o elemento principal de *output* do sistema, dando ao utilizador todas as informações necessárias para a orientação no ambiente de jogo. Através da vibração do comando *Wiimote* na interface de jogo, é proporcionado um segundo elemento de *output*. Apesar de não ser apresentado ao utilizador de forma óbvia, a interface de jogo apresenta o único elemento de *input* no sistema.

Tabela 9 – Coordenadas cartesianas dos altifalantes no espaço de instalação de acordo com o sistema de captura de movimentos *Vicon*.

Altifalante	X (mm)	Y (mm)
1	880	2125
2	1626	1626
3	2125	880
4	2300	0
5	2125	-880
6	1626	-1626
7	880	-2125
8	0	-2300
9	-880	-2125
10	-1626	-1626
11	-2125	-880
12	-2300	0
13	-2125	880
14	-1626	1626
15	-880	2125
16	0	2300

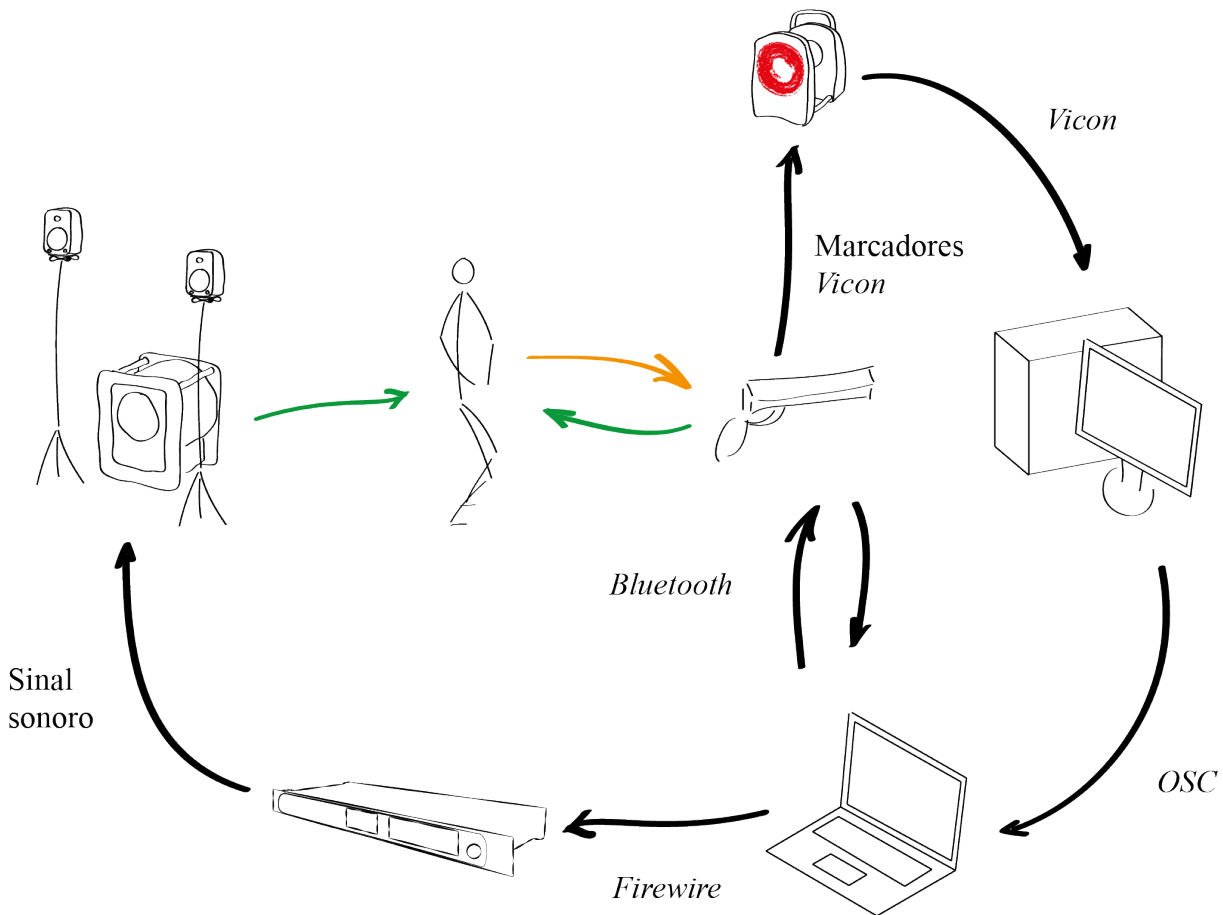
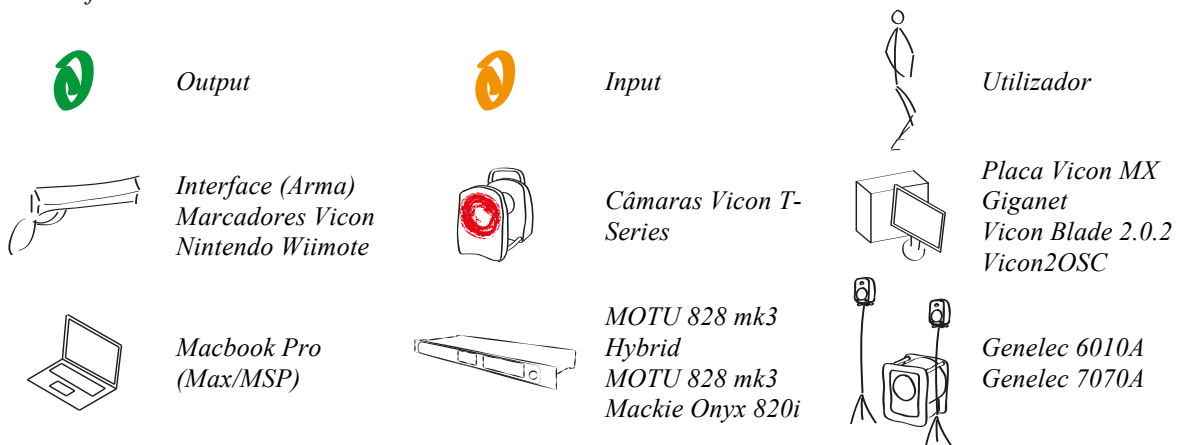


Figura 63 – Diagrama do fluxo de informação entre as várias componentes (hardware) do projeto *The Sound of Horror*.



A apresentação do projecto decorreu no dia 24 de Abril de 2013 no Laboratório de *Motion Capture* da Escola das Artes da Universidade Católica Portuguesa.

Tinha sido concebida a entrada de apenas um utilizador no espaço de cada vez, criando intriga nos restantes utilizadores e tornando a sua experiência mais interessante e única, não lhes apresentando os conteúdos previamente. Por motivos logísticos, isto não se revelou viável, acabando por ser cedida a entrada na sala aos restantes espectadores. Para iniciar a experiência, o utilizador foi guiado para o interior da estrutura de jogo e foi-lhe entregue a

arma. As luzes da sala foram apagadas, tornando o espaço escuro, removendo o estímulo visual da experiência, e foi dada a instrução ao *patch* para inicializar o jogo.

A Figura 64, Figura 65, Figura 66, Figura 67, Figura 68, Figura 69, Figura 70, Figura 71, Figura 72 e Figura 73 apresentam o espaço com o sistema de jogo montado.



Figura 64 – Perspectiva do espaço com o sistema de jogo montado (1).



Figura 65 – Perspectiva do espaço com o sistema de jogo montado (2).



Figura 66 – Perspectiva do espaço com o sistema de jogo montado (3).



Figura 67 – Perspectiva do espaço com o sistema de jogo montado (4).



Figura 68 – Perspectiva do espaço com o sistema de jogo montado (5).



Figura 69 – Perspectiva do espaço com o sistema de jogo montado (6).

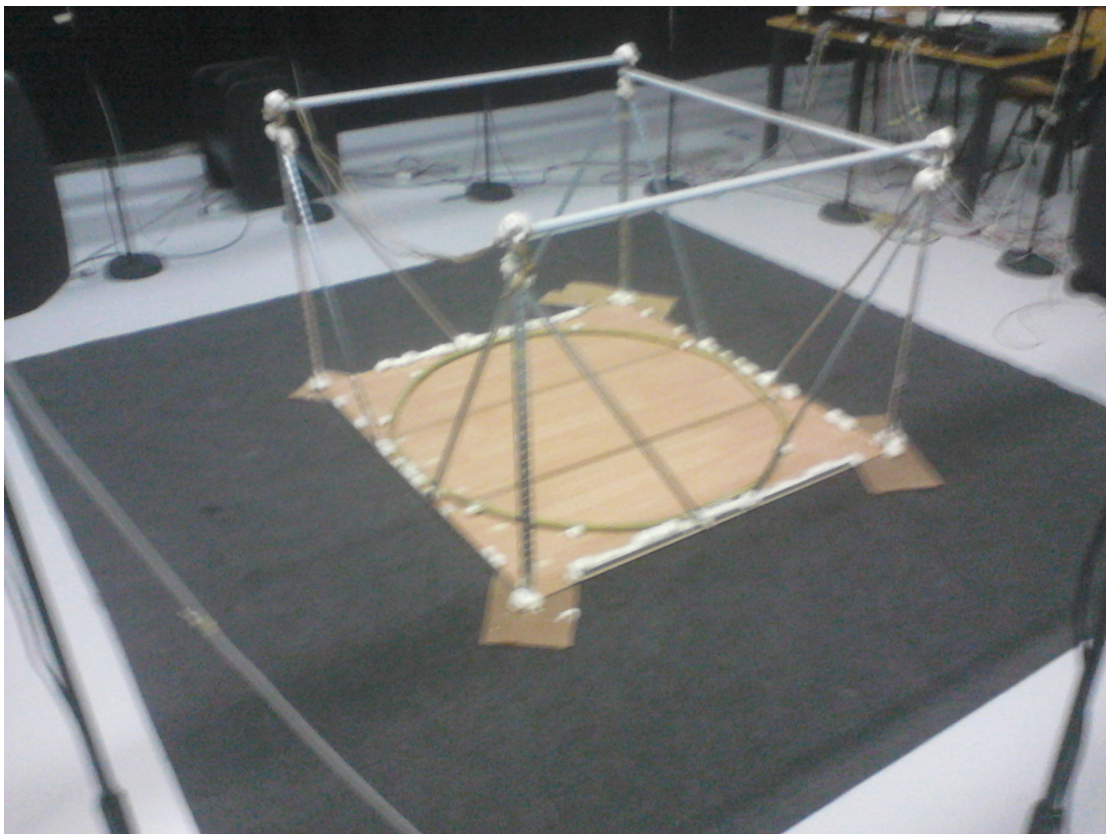


Figura 70 – Perspectiva do espaço com o sistema de jogo montado (7).



Figura 71 – Perspectiva do espaço com o sistema de jogo montado (8).



Figura 72 – Perspectiva do espaço com o sistema de jogo montado (9).

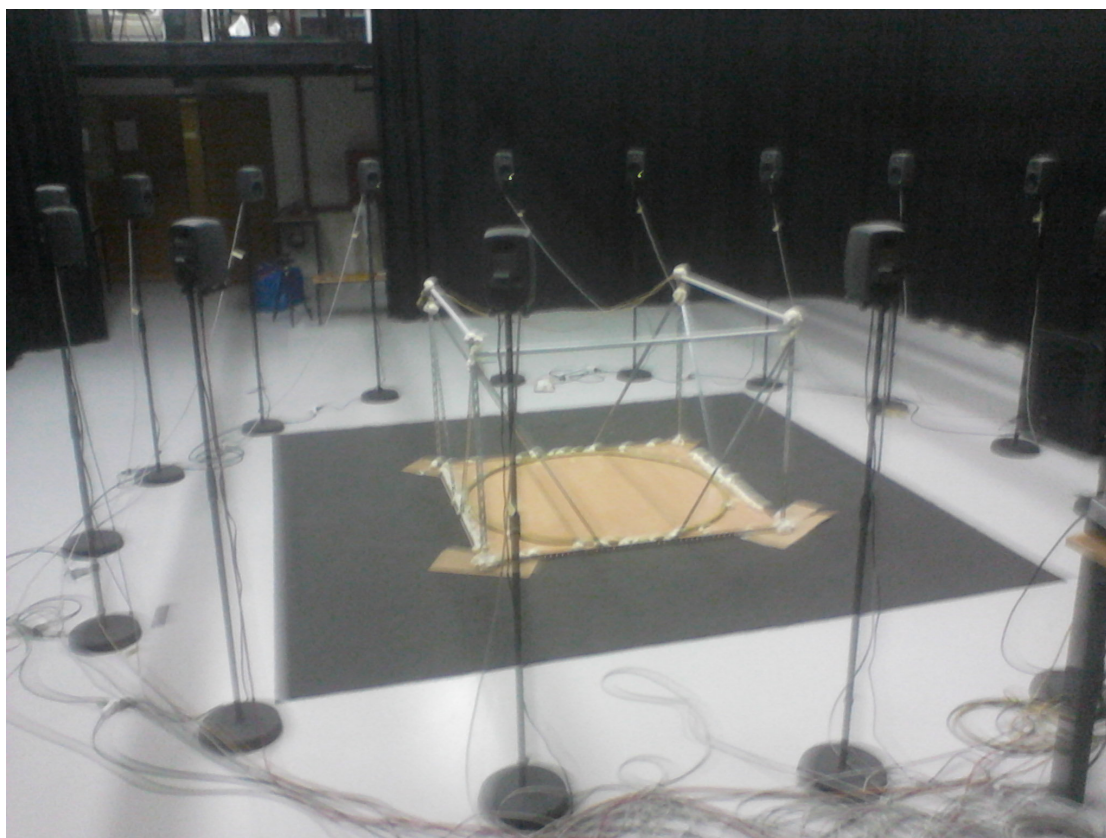


Figura 73 – Perspectiva do espaço com o sistema de jogo montado (10).

3.6. Discussão

O jogo foi conceptualizado como uma experiência individual. Atendendo ao método como a apresentação decorreu, houve uma repetição de conteúdos para os jogadores. Apesar de ter sido implementado um sistema que torna o ângulo de azimute dos alvos sonoros aleatório, os elementos apresentados foram idênticos. Tendo em conta que o *The Sound of Horror* é um jogo do subgénero *survival horror*, a repetição de conteúdos é inevitável. Seria, no entanto, possível a implementação de *power-ups* à semelhança do jogo *Dead Space* (EA Redwood Shores, 2008) para tornar a experiência mais dinâmica.

Apesar dos conteúdos apresentados neste projeto não serem originais relativamente ao estilo de jogo, há inovação no modo como estes são expostos ao utilizador. Apresenta uma forma de interação orgânica e uma realidade virtual sonora que é estranha à maioria das pessoas. Para a tirar partido destes elementos, foi apresentado um ritmo de jogo acessível a utilizadores menos experientes, permitindo assim a apreciação da experiência a um nível estético, sem criar uma sobrecarga de desafios por parte do jogo. Por este motivo, a lentidão do ritmo da experiência foi a crítica principal ao projeto, por parte dos utilizadores, algo que terá sido agravado por terem sido expostos aos conteúdos repetidamente.

Os diálogos do jogo foram sonorizados em inglês de modo a atingir um público maior. No entanto, isto poderá ter sido um obstáculo a pessoas com dificuldades na língua inglesa. Seria possível a implementação bilingue para garantir o discernimento dos diálogos apresentados.

O uso do sistema *Ambisonics* apresentou resultados excelentes, não tendo havido qualquer problema na localização dos alvos sonoros no espaço em qualquer posição.

A incapacidade de recorrer a uma distribuição de altifalantes esférica ou a um espaço acusticamente tratado, à semelhança do *Allosphere* (Amatriain *et al.*, 2007), reduziu bastante a qualidade da experiência sonora.

Numa situação sem condicionantes, seria interessante a implementação do sistema de espacialização *Wave Field Synthesis*, à semelhança do instalado no *Institut für Sprache und Kommunikation*, na *Technische Universität* em Berlim⁷.

⁷ Website do instituto acessível em

<http://www.ak.tu-berlin.de/menue/forschung/forschungsprojekte/wellenfeldsynthese/>

Ao serem guiados para a estrutura, os utilizadores mantiveram-se no centro do espaço sonoro. O uso desta permitiu transmitir de forma imediata a instrução de que não deveriam sair daquela posição. No entanto, e apesar da experiência decorrer às escuras, a nível visual apresentou-se pouco interessante. O uso de uma estrutura mais parecida a uma jaula real teria sido uma alternativa interessante para a narrativa do jogo.

A solução implementada como sistema de captura de movimentos *Vicon* apresentou óptimos resultados, não tendo havido falhas a nível de continuidade de transmissão e obtendo um valor de latência extremamente reduzido (cerca de 5 milissegundos). O uso de uma passadeira rolante omnidireccional à semelhança da utilizada pelo *Gadget Show* (2011) para a construção do simulador do jogo *Battlefield 3* permitiria o movimento do utilizador dentro de um espaço restrito, dando a possibilidade de tirar partido da estabilidade criada pelo sistema de espacialização sonora por colunas, não confinando virtualmente o utilizador a um espaço delimitado.

De acordo com as reações gerais dos utilizadores, os vários elementos sonoros funcionaram em harmonia, produzindo um ambiente característico de um jogo de terror. No entanto, seria benéfica a implementação de um indicador sonoro adicional da categoria *Integrated Interface Sounds* (Jørgensen, 2010) para sinalizar a morte de cada um dos monstros. Isto ajudaria a manter a inteligibilidade dos elementos em situações com múltiplos monstros ativos. A nível de narrativa, integrar-se-ia naturalmente, agindo como uma anotação feita pelo Grande Irmão.

Houve um esforço no sentido de criar um ambiente imersivo, dando resposta às características analisadas no capítulo 2.3. Segundo o *SCI-model*, de Ermi e Mäyrä (2005), a nível sensorial foi desenvolvido um sistema sonoro que abrangeu o utilizador e o expôs a ambiências que permitiram uma abstração dos elementos sensoriais exteriores à experiência de jogo. Em termos de desafios, foi apresentado um controlador gestual, em conjunto com alvos sonoros dinâmicos. Relativamente à dimensão imaginativa da experiência, criou-se um ambiente de terror com personagens e monstros sonoros que permitiram transportar o utilizador para a realidade virtual do jogo.

Num balanço final do trabalho realizado, concluiu-se que os objetivos inicialmente estabelecidos foram atingidos com êxito.

As diversas condicionantes que foram surgindo, com o tempo e recursos disponíveis, foram sucessivamente ultrapassadas através de um esforço de grande interação com os orientadores e colegas. Como resultado da experiência obtida na resolução das diversos

problemas surgidos, está-se convicto de que o desenvolvimento deste projeto foi uma grande oportunidade para obter uma aprendizagem profunda, consistente e coerente. Destaca-se a que foi obtida ao nível da programação computacional e na gestão do trabalho de equipa.

4. Conclusão

Esta investigação surgiu com o propósito da criação de uma instalação interativa e imersiva onde a componente sonora desempenha um papel fundamental. De forma compreender as necessidades que este projeto acarreta, foi necessário o desenvolvimento de um estudo bibliográfico que permitiu usufruir dos conhecimentos teóricos que permitem a aplicação das metodologias indicadas.

Através do enquadramento histórico relativamente à espacialização sonora, bem como do estudo relativamente aos mecanismos psicoacústicos subjacentes à localização de fontes sonoras no espaço, foi possível aprofundar um conjunto de conhecimentos teóricos. Estes permitiram a compreensão do modo de funcionamento das principais técnicas de espacialização sonora e a forma como é possível a sua aplicação para contextos interativos de realidade virtual.

Para a criação deste género de experiência, é necessário garantir que o sistema proporciona ao utilizador os meios necessários para a sua imersão com a realidade virtual apresentada. O estudo desta matéria de uma forma geral, bem como aplicada especificamente ao som permitiu conhecer alguns dos modelos de imersão existentes, alertando para a importância das diferentes componentes envolvidas neste processo.

Os conhecimentos obtidos nestes estudos foram aplicados de forma a criar uma experiência sonora rica e dinâmica. O jogo sonoro *The Sound of Horror* classifica-se como um *first-person shooter* do subgénero *survival horror*, de uso individual. Esta experiência sonora interativa e imersiva contou com o uso de um sistema *Ambisonics* de 7ª ordem e de um sistema multicanal de 17 altifalantes. Através deste foi possível sobrepor o estímulo sensorial do jogo aos exteriores à experiência, permitindo aumentar o grau de imersão na realidade virtual. Não tendo contado com qualquer informação visual, decorreu num espaço escuro, de modo a remover este estímulo e aumentar o medo e terror causados pelo jogo.

Foi desenvolvida uma narrativa com vista a tornar o ambiente sombrio e imprevisível, num modelo de jogo tradicional. Introduziram-se personagens de jogo para ajudar a orientar o utilizador no decorrer da experiência e criados monstros sonoros que obrigaram o utilizador a recorrer aos seus instintos primitivos, defendendo-se como se a sua vida estivesse em risco. Os conteúdos sonoros produzidos foram determinantes para o impacto da experiência,

permitindo ampliar as sensações criadas pelo jogo. O uso de elementos sonoros tensos, desconfortáveis e assustadores, e de todo um contexto de jogo onde o utilizador é colocado numa posição sem qualquer controlo e se apresenta indefeso face aos maus tratos de uma entidade fria e distante, permitiram a criação de uma experiência de terror.

O projeto *The Sound of Horror* contou com uma interação orgânica, através do uso de uma interface à semelhança de uma arma, permitindo um controlo gestual para a execução de múltiplas funções, que permitiram ao utilizador interagir com fontes sonoras dinâmicas.

A combinação de uma realidade virtual sonora, sem componente visual, com um meio de interação orgânico e com liberdade de movimentos tem um grande potencial imersivo. A precisão obtida na criação de ambientes sonoros consegue atingir resultados mais realistas e com uma resolução superior aos ambientes virtuais visuais (Neuhoff, 2011).

Esta investigação permitiu redescobrir a potencialidade do *design* de som. A necessidade de conceber um mundo virtual completamente inteligível unicamente através da sua componente sonora revelou-se um desafio extremamente interessante. Sendo o som um meio distintamente ambíguo, o efeito criado pelo ambiente de jogo é, inconscientemente, diretamente proporcional à capacidade imaginativa do utilizador (Höner, 2011).

Seria interessante expandir esta investigação através da realização de testes e inquéritos aos utilizadores, permitindo validar cientificamente as escolhas tomadas, de modo a obter dados estatísticos.

A adaptação do sistema para uma experiência multijogador permitiria tornar o projeto mais dinâmico, adicionando uma dimensão de trabalho de equipa.

O projeto conta com o uso de componentes de reprodução sonora e captação de movimentos de alta qualidade. Apesar da maioria dos utilizadores não ter acesso a este género de sistemas, seria interessante a adaptação do *The Sound of Horror* para um modelo caseiro e possivelmente comercial.

Apesar do projeto apresentar um modelo de jogo tradicional, o modo como estes conteúdos são apresentados ao utilizador e a forma como a interação com o sistema decorre foram desenvolvidos especificamente com o propósito de satisfazer as necessidades deste género de jogo sonoro. Seria interessante explorar o desenvolvimento ambientes sonoros interativos e imersivos na criação de *audiogames*. A forma como estes são concebidos difere dos jogos audiovisuais tradicionais. Atendendo à potencialidade que este género de projetos apresenta, é necessário criar modelos específicos para tirar partido dos benefícios que advêm da criação de realidades virtuais sonoras.

5. Referências e Bibliografia

5.1. Livros e artigos

- Amatriain, X., Kuchera-Morin, J., Hollerer, T., Pope, S. (2007). Immersive Audio and Music in the Allosphere. *International Computer Music Conference*, 2007. Disponível em: <http://www.cs.ucsb.edu/~holl/pubs/Amatriain-2007-ICMC.pdf>
- Amatriain, X., Kuchera-Morin, J., Hollerer, T., Pope, S. (2009). The AlloSphere: Immersive Multimedia for Scientific Discovery and Artistic Exploration. *IEEE Multimedia*, 16(2), 64-75. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.178.189&rep=rep1&type=pdf>
- Askew, A. (1981, Novembro). The Amazing Clement Ader. *Studio Sound*, 23(11), 100-102. Disponível em <http://www.benmook.com/tech/documents/Schematics 2/Schematics 1/Ader Article.pdf>
- Askew, A. (1981, Outubro). The Amazing Clement Ader. *Studio Sound*, 23(10), 66-68. Disponível em <http://www.benmook.com/tech/documents/Schematics 2/Schematics 1/Ader Article.pdf>
- Askew, A. (1981, Setembro). The Amazing Clement Ader. *Studio Sound*, 23(9), 44-46, 48. Disponível em <http://www.benmook.com/tech/documents/Schematics 2/Schematics 1/Ader Article.pdf>
- Atkins, J. (2011). *Spatial acoustic signal processing for immersive communication*. Baltimore: UMI. Disponível em <http://gradworks.umi.com/3496161.pdf>
- Bamford, J. (1995). *An Analysis of Ambisonic Sound Systems of First and Second Order*. Ontario: University of Waterloo. Disponível em: http://www.engineeringharmonics.com/archive-papers/JSB_thesis.pdf?q=papers/JSB_thesis.pdf
- Berkhout, A.J.; De Vries, D.; Vogel, P. (1993, Maio) Acoustic Control by Wave Field Synthesis, *J.Acoust.Soc.Am.*, 93, 2764–2778.
- Binaural Tools. (2009). College of Engineering - University of California. Disponível em: http://www.ece.ucdavis.edu/binaural/binaural_tools.html
- Blauert, J. (1997). *Spatial hearing, the psychophysics of human sound localization*. (revised ed.). The MIT Press.
- Brown, E. & Cairns, P. (2004). A Grounded Investigation of Game Immersion. *Conference on Human Factors in Computing Systems*, 2004, 1297-1300. Disponível em: [http://complexworld.pbworks.com/f/Brown+and+Cairns+\(2004\).pdf](http://complexworld.pbworks.com/f/Brown+and+Cairns+(2004).pdf)
- Burgess, D. A. (1992). *Real-time audio spatialization with inexpensive hardware*. Georgia: Graphics Visualization and Usability Center, Georgia Institute of Technology. Disponível em <https://smartech.gatech.edu/bitstream/handle/1853/3677/92-22.pdf>
- Carlile, S. (2011). Psychoacoustics. Hermann, T., Hunt, A., & Neuhoff, J. G. (eds.), *The Sonification Handbook*. Berlin: Logos Verlag.
- Cheng, C. (2001). *Visualization, measurement, and interpolation of head related transfer functions (HRTF's) with applications in electroacoustic music*. University of Michigan.
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, 25(5), 975-979. Disponível em <http://www.ee.columbia.edu/~dpwe/papers/Cherry53-cpe.pdf>
- Chowning, J. M. (1977, Junho). The simulation of moving sound sources. *Computer Music Journal*, 1(3), 48-52. Disponível em <http://create.ucsb.edu/~matt/240/chowning-simulation-moving-sound-sources.pdf>
- Cohen, M., Herder, J. & Martens, W. (1999) Cyberspatial Audio Technology. *Journal of the Acoustical Society of Japan*, 20(6), 389-395 <http://web-ext.u-aizu.ac.jp/~mcohen/welcome/publications/JASJ-reviewE.pdf>

- Collins, K. (2008). *Game Sound: An Introduction to the History, Theory and Practice of Video Game Music and Sound*. Cambridge, Massachussets: The MIT Press.
- Cooper, D. (2011, Outubro, 20). *The Gadget Show builds an FPS simulator that shoots back*. Disponível em: <http://www.engadget.com/2011/10/20/the-gadget-show-builds-an-fps-simulator-that-shoots-back-video/>
- Cooper, M., & Taylor, M. E. (1998). *Ambisonic sound in virtual environments and applications for blind people*. Milton Keynes: Institute of Educational Technology/Knowledge Media Institute/Multimedia Enabling Technology Group, Open University. Disponível em http://www.icdvrat.rdg.ac.uk/1998/papers/1998_15.pdf
- Cruz-Neira, C., Sandin, D. & DeFanti, T. (1993). Surround-Screen Projection-based Virtual Reality: The Design and Implementation of the CAVE. *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, 135–142. Disponível em: <http://www.cs.utah.edu/~thompson/vissim-seminar/on-line/CruzNeiraSig93.pdf>
- Cruz-Neira, C., Sandin, D., DeFanti, T., Kenyon, R. & Hart, C. (1992). The CAVE: Audio Visual Experience Automatic Virtual Environment. *Communications of the ACM*, 35(6), 64–72. Disponível em: <http://www.evl.uic.edu/files/pdf/CACM92-CAVE-cruz-neira.pdf>
- Daniel, J., Nicol, R., & Moreau, S. (2003). *Further investigations of high order ambisonics and wavefield synthesis for holophonic sound imaging*. Lannion Cedex: Audio Engineering Society: Convention paper 5788. Disponível em <http://pcfarina.eng.unipr.it/Public/AES-114/00048.pdf>
- De Vries, D. & Boone, M. (1999). Wave field synthesis and analysis using array technology. *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. 1999. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.119.306&rep=rep1&type=pdf>
- Deruty, E. (2012). Loudspeaker orquestras, Non-Standard Multi-Loudspeaker Diffusion Systems. *Sound on Sound*. Disponível em <http://www.soundonsound.com/sos/jan12/articles/nsml.htm>
- Dovey, J. & Kennedy, H. (2006). *Game Cultures: Computer Games as New Media*. Berkshire: Open University Press
- Elen, R. (2001). Ambisonics: the Surround Alternative. *Proceedings of the 3rd Annual Surround Conference and Technology Showcase*. 1-4. Disponível em: <http://www.ambisonic.net/pdf/ambidvd2001.pdf>
- Ermi, L. & Mäyrä, F. (2005). Fundamental Components of the Gameplay Experience: Analysing Immersion. *Proceedings of DiGRA 2005 Conference: Changing Views – Worlds in play*. Disponível em: http://people.uta.fi/~tlilma/gameplay_experience.pdf
- Faria, R. R. A. (2005). *Auralização em ambientes audiovisuais imersivos*. Tese de doutoramento. São Paulo: Universidade de São Paulo.
- Farina, A., Glasgal, R., Armelloni, E., & Torger, A. (2001). *Ambiophonic principles for the recording and reproduction of surround sound for music*. Scholl Elmau: Audio Engineering Society 19th International Conference.
- Farrar, K. (1979, Novembro). Soundfield microphone – 2: Detailed functioning of control unit. *Wireless World*. 99-103. Disponível em: <http://ambisonics.dreamhosters.com/wireless-world-farrar-10-1979.pdf>
- Farrar, K. (1979, Outubro). Soundfield microphone: Design and development of microphone and control unit. *Wireless World*, 85(99), 48-103. Disponível em <http://ambisonics.dreamhosters.com/wireless-world-farrar-10-1979.pdf>
- Farrar, K. (1979, Outubro). Soundfield microphone: Design and development of microphone and control unit. *Wireless World*. 48-50. Disponível em: <http://ambisonics.dreamhosters.com/wireless-world-farrar-10-1979.pdf>
- Fishman-Johnson, E. (1993-1994). The movement of sound in space: An update. *Contemporary Music Forum*, 5-6, 15-21. Disponível em <http://www.bgsu.edu/colleges/music/MACCM/media/cmf/CMF5-6.pdf>
- Fonseca, N. (2007). *Introdução à engenharia de som*. Lisboa: FCA - Editora de Informática.
- Fox, B. (1982, Maio). Early Stereo Recording (Alen Blumlein). *Studio Sound*, 24(5), 36-42. Disponível em <http://www.benmook.com/tech/documents/Schematics%202/Schematics%201/Blumlein%20Article.pdf>

- Gärdenfors, J. & Friberg, D. (2004). Audio games: new perspectives on game audio. *Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology*. Disponível em: <http://www.cms.livjm.ac.uk/library/AAA-GAMES-Conferences/ACM-ACE/ACE2004/FP-18friberg.johnny.audiogames.pdf>
- Gardner, W. & Martin, K. (1994). HRTF measurements of a KEMAR dummy-head microphone. *MIT Media Lab Perceptual Computing Technical Report*, 280, Disponível em <http://www.linux.bucknell.edu/~kozick/elec32007/hrtfdoc.pdf>
- Gasteiger, B. (2010). *A tangible user interface for playing virtual acoustics*. Tese de diploma. Graz: Institute of Electronic Music and Acoustics (IEM), University of Music and Performing Arts Graz. Disponível em <http://iem.kug.ac.at/fileadmin/media/iem/projects/2010/gasteiger.pdf>
- Gerzon, M. (1977, Novembro). Design of Ambisonic Decoders for Multispeaker. *58th Audio Engineering Society Convention*. Disponível em: <http://decoy.iki.fi/dsound/ambisonic/motherlode/source/Design%20of%20Ambisonic%20Decoders%20for%20Multispeaker%20Surround%20Sound%20MAG%201977.pdf>
- Gerzon, M. A. (1973, Janeiro-Fevereiro). Periphony: With-height sound reproduction. *Journal of the Audio Engineering Society*, 21(1), 2-10. Disponível em http://decoy.iki.fi/dsound/ambisonic/motherlode/source/Periphony_With_height_sound_reproduction_Michael%20Gerzon_JAES_Jan_Feb_1973.pdf
- Gerzon, M. A. (1974). Surround-sound psychoacoustics, Criteria for the design of matrix and discrete surround-sound systems. *Wireless World*, 80(12), 483-485. Disponível em http://www.audiosignal.co.uk/Resources/Surround_sound_psychoacoustics_A4.pdf
- Glasgal, R. (2003). Surround ambiophonic recording and reproduction. *Audio Engineering Society International Conference on Multichannel Audio*, 24, 11-20.
- Goldman, D., & Yuval, T. (2002). *The cartoon music book*. Chicago: A Cappella Books.
- Grimshaw, M. (2007). Sound and Immersion in the First-Person Shooter. *Proceedings of CGAMES'2007. 11th International Conference on Computer Games: AI, Animation, Mobile, Educational and Serious Games*. Disponível em: http://wlv.openrepository.com/wlv/bitstream/2436/35995/2/Grimshaw_CGAMES07.pdf
- Grimshaw, M. (2007). *Sound and immersion in the first-person shooter*. Wolverhampton: University of Wolverhampton. Disponível em http://wlv.openrepository.com/wlv/bitstream/2436/35995/2/Grimshaw_CGAMES07.pdf
- Hacihabiboglu, H., Gunel, B., & Zondo, A. (2005, Outubro). *Head-related transfer function filter interpolation by root displacement*. 2005 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, New York. Disponível em: <http://epubs.surrey.ac.uk/2393/1/SRF002503.pdf>
- Heeger, D. (2006). *Perception lecture notes: Auditory pathways and sound localization*. Department of Psychology, New York University. Disponível em <http://www.cns.nyu.edu/~david/courses/perception/lecturenotes/localization/localization.html>
- Hermann, T. & Hunt, A. (2011). Interactive Sonification. Hermann, T., Hunt, A., & Neuhoff, J. G. (eds.), *The Sonification Handbook*. Berlin: Logos Verlag.
- Hermann, T., & Hunt, A. (2005). An Introduction to Interactive Sonification (Guest Editors' Introduction). *IEEE MultiMedia*, 12(2), 20-24. Disponível em: <http://eprints.whiterose.ac.uk/655/1/hunta2.pdf>
- Holmes, T. (2008). *Electronic and experimental music*. (3ª ed.). New York: Routledge.
- Höner, O. (2011). Aiding Movement with Sonification in “Exercise, Play and Sport”. Hermann, T., Hunt, A., & Neuhoff, J. G. (eds.), *The Sonification Handbook*. Berlin: Logos Verlag.
- Hope, A. (1979). Fantasia-Multitracked. *Studio Sound*. 21(8), 29-30
- Hugill, A. (2012). *Towards an analysis of Papa Sangre, an audio-only game for the iPhone/iPad*. Leicester: Institute Of Creative Technologies De Montfort University. Disponível em: http://www.orema.dmu.ac.uk/sites/default/files/analysisfiles/Papa%20Sangre%20analysis_0.pdf
- Ihren, J. & Frisch, K. (1999). The Fully Immersive Cave. *3rd Int'l Immersive Projection Technology Workshop*, 59-63.

- Irie, R. E. (1993). *Robust sound localization: An application of an auditory perception system for a humanoid robot*. Tese de Mestrado. Massachusetts: Massachusetts Institute of Technology Disponível em http://groups.csail.mit.edu/lbr/hrg/1995/irie_ms_thesis.pdf
- ITU-R. (2012). *Recommendation ITU-R BS.775-3, Multichannel stereophonic sound system with and without accompanying picture*. International Telecommunication Union. Disponível em http://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.775-3-201208-I!!PDF-E.pdf
- J. Braasch (2005), A loudspeaker-based 3D sound projection using Virtual Microphone Control (ViMiC). Convention of the Audio Engineering Society, 118.
- Jensenius, A., Nymoen, K. & Godøy, R. (2008). *A multilayered GDIF-based setup for studying coarticulation in the movements of musicians*. Oslo: Department of Musicology, University of Oslo.
- Jessel, M. (1973). *Acoustique Théorique: Propagation et Holophonie*. Paris: Masson.
- Jørgensen, K. (2010). Time for New Terminology? Diegetic and Non-Diegetic Sounds in Computer Games Revisited. Grimshaw, M. (ed.) *Game Sound Technology and Player Interaction: Concepts and Developments*. IGI Global.
- Jot, J. (1997). *Real-time spatial processing of sounds for music, multimedia and interactive human-computer interfaces*. Paris: Institut de Recherche et Coordination Acoustique/Musique.
- Judson, T. (2011, Janeiro) *OpenNI to Max/MSP via OSC*. Disponível em <http://tohmjudson.com/?p=30>. Consultado em Janeiro de 2013.
- Kirkeby, O., Nelson, P. & Hamada, H. (1998). The stereo dipole—A virtual source imaging system using two closely spaced loudspeakers. *Journal of the Audio Engineering Society*, 46, 387–395.
- Lemaitre, G., Houix, O., Franinovic, K., Visell, Y., & Susini, P. (2009, Julho). The Flops glass: a device to study emotional reactions arising from sonic interactions. *Sound and Music Computing*, 2009, 23-25. Disponível em: <http://smc2009.smcnetwork.org/programme/pdfs/281.pdf>
- Lombardo, V., Arghinenti, A., Nunnari, F., Valle, A., Vogel, H. H., Fitch, J., Dobson, R., Padget, J., Tazelaar, K., Weinzierl, S., Benser, S., Kersten, S., Starosolski, R., Borczyk, W., Pytlik, W. and Niedba, X. (2005). *The virtual electronic poem (VEP) project*. Barcelona: Proceedings of the International Computer Music Conference (ICMC). Disponível em <http://www.fonurgia.unito.it/andrea/pub/vep-icmc05.pdf>
- Lord Rayleigh (1907). On our perception of sound direction. *Philosophical Magazine*, 13(74) 214-232
- Lossius, T., Baltazar, P. & Hogue, T. (2009). DBAP - Distance-Based Amplitude Panning. *International Computer Music Conference*. 2009. Disponível em: <http://www.trondlossius.no/system/fileattachments/30/original/icmc2009-dbap-rev1.pdf>
- Lossius, T., Baltazar, P., & Hogue, T. (2009). *DBAP Distanced-based Amplitude Panning*. Bergen: BEK - Bergen Center for Electronic Arts. Disponível em <http://www.trondlossius.no/system/fileattachments/30/original/icmc2009-dbap-rev1.pdf>
- Lynch, H., & Sazdov, R. (2011). *An investigation into compositional techniques utilized for the three-dimensional spatialization of electroacoustic music*. New York: Digital Media and Arts Research Centre (DMARC) Department of Computer Science and Information Systems University of Limerick. Disponível em http://www.ems-network.org/IMG/pdf_EMS11_lynch_sazdov.pdf
- Malham, D. (1998). *Tutorial article: Approaches to spatialisation*. *Organised Sound*, 3(2), 167–177. Cambridge: Cambridge University Press. Disponível em: <https://ccrma.stanford.edu/workshops/gaffta2010/spatialsound/topics/why/materials/malham.pdf>
- Malham, D. G. & Myatt, A. (1995). 3-d Sound Spatialization Using Ambisonic Techniques. *Computer Music Journal*, 19(4), 58-70. Disponível em
- Marije A.J. Baalman (2010). Spatial Composition Techniques and Sound Spatialisation Technologies. *Organised Sound*, 15, 209-218. Disponível em: http://create.ucsb.edu/~matt/240/Baalman-spatial_comp_tech.pdf
- Matthew, N. (2011). *Wave Field Synthesis In Three Dimensions By Multiple Line Arrays*. Florida: University of Miami. Disponível em: <http://www.mattmontag.com/projects/wfs/Montag%20Thesis%202011%20-%20Wave%20Field%20Synthesis%20in%20Three%20Dimensions%20by%20Multiple%20Line%20Arrays.pdf>

- Mauney, B. & Walker, B. (2004). Creating functional and livable soundscapes for peripheral monitoring of dynamic data. *Proceedings of the 10th International Conference on Auditory Display*. Disponível em: <http://sonify.psych.gatech.edu/publications/pdfs/2004ICAD-MauneyWalker.pdf>
- McAnally, K., & Martin, R. (2007). *Interpolation of head-related transfer functions*. Victoria: Air Operations Division DSTO Defence Science and Technology Organisation.
- McGee, R. (2009). *John Chowning: overview, techniques, and compositions*. Santa Barbara: The University of California. Disponível em http://www.lifeorange.com/writing/ChowningAnalysis_McGee.pdf
- McMahan, A. (2003). Immersion, Engagement, and Presence: A New Method for Analyzing 3-D Video Games. J.P. Wolf and B. Perron (eds.), *The Video Game Theory Reader*, New York: Routledge. Disponível em: <http://people.ict.usc.edu/~morie/SupplementalReadings/ch3-McMahanrev.pdf>
- Meinicke, P. & Ritter, H. (2000). Principal curve sonification. *Proceedings of the International Community of Auditory Display*, 2000, 81–86. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.24.3276&rep=rep1&type=pdf>
- Middlebrooks, J. & Green, D. (1990). Directional dependence of interaural envelope delays. *Journal of the Acoustical Society of America*, 87(5) 2149-2162.
- Moore, B. C. J. (1995). *Hearing, handbook of perception and cognition*. (2ª ed.). Cambridge: Academic Press.
- Murray, J. (1997). *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*. Massachusetts: The MIT Press.
- Neuhoff, J. G. (2011). Perception, Cognition and Action in Auditory Displays. Hermann, T., Hunt, A., & Neuhoff, J. G. (eds.), *The Sonification Handbook*. Berlin: Logos Verlag.
- Overholt, D., Wakefield, G., Morin, J. K., Novak, M., Putnam, L., Thompson, J., & Smith, W. (2008). The AlloBrain: an Interactive Stereographic, 3D Audio Immersive Environment. *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*. Association for Computing Machinery. Disponível em: https://mat.ucsb.edu/Publications/08_Wakefield_SIDCHI_AlloBrain.pdf
- Papworth, N. (2010). iSpooks: an audio focused game design. *Proceedings of the 5th Audio Mostly Conference: A Conference on Interaction with Sound*.
- Peters, N. (2010). *Sweet [re]production: Developing sound spatialization tools for musical applications with emphasis on sweet spot and off-center perception*. Montreal: Music Technology Area Department of Music Research Schulich School of Music McGill University. Disponível em <http://www.music.mcgill.ca/~nils/PetersThesis-web.pdf>
- Peters, N., Matthews, T., Braasch, J. & McAdams, S. (2008). *Spatial sound rendering in MAX/MSP with ViMiC*. Proceedings of the International Computer Music Conference (ICMC). Disponível em http://nilspeters.info/papers/ICMC08-VIMIC_final.pdf
- Pine, B. & Gilmore, J. (1999). *The Experience Economy: Work is Theatre & Every Business a Stage*. Boston: Harvard Business School.
- Puckette, M. (2007). Patch for Guitar. *Second Pd Convention*, Montreal.
- Pulkki, V. (1997, Junho). Virtual sound source positioning using vector base amplitude panning. *Journal of the Audio Engineering Society*, 45(6), Disponível em <http://lib.tkk.fi/Diss/2001/isbn9512255324/article1.pdf>
- Pulkki, V. (2000). Generic panning tools for MAX/MSP. *International Computer Music Conference*. 2000. Disponível em: <http://lib.tkk.fi/Diss/2001/isbn9512255324/article4.pdf>
- Pulkki, V. (2001). *Spatial Sound Generation and Perception by Amplitude Panning Techniques*. Helsinki: Helsinki University of Technology. Disponível em: <https://aaltodoc.aalto.fi/bitstream/handle/123456789/2345/isbn9512255324.pdf?sequence=1>
- Röber, N. & Masuch, M. (2005). Playing Audio-only Games: a compendium of interacting with virtual, auditory worlds. *Proceedings of the 2nd DIGRA Gamesconference*. Disponível em: <http://www.sciweavers.org/read/playing-audio-only-games-a-compendium-of-interacting-with-virtual-auditory-worlds-121079>
- Roux-Girard, G. (2010). Listening to Fear: A Study of Sound in Horror Computer Games. Grimshaw, M. (ed.) *Game Sound Technology and Player Interaction: Concepts and Developments*. IGI Global.

- Rumsey, F. (2002). *Spatial audio*. Oxford: Focal Press.
- Sanal, A. J. (1976, Dezembro). Looking backward. *Journal of the Audio Engineering Society*, 24(10), 832.
- Savioja, L., Huopaniemi, J., Lokki, T. & Väänänen, R. (1999). Creating interactive virtual acoustic environments. *Journal of the Audio Engineering Society*, 47(9), 675-705. Disponível em: <http://lib.tkk.fi/Diss/2002/isbn9512261588/article1.pdf>
- Schafer, R. M. (1993). *The soundscape: Our sonic environment and the tuning of the world*. Massachusetts: Destiny Books.
- Shepherd, A. (2003). *Pro tools for video, film and multimedia*. Boston: Muska & Lipman Publishing.
- Spors, S., Rabenstein, R., & Ahrens, J. (2008). *The theory of wave field synthesis revisited*. Amsterdam: Audio Engineering Society 124th Convention. Disponível em http://www.deutsche-telekom-laboratories.de/~spors/publications/2008/AES124_Spors_WFS_Theory.pdf
- Steinberg, J. C., Snow, W. B. (1934a). Auditory perspective – physical factors. *Transactions of the American Institute of Electrical Engineers*, 53(1), 12-17.
- Steinberg, J. C., Snow, W. B. (1934b). Physical Factors. *Bell Systems Technical Journal*, 132, 245-258.
- Stitt, P., Bertet, S., Walstijn, M. (2013). Perceptual investigation of image placement with ambisonics for non-centered listeners. *16th International Conference on Digital Audio Effects*. Disponível em: http://dafx13.nuim.ie/papers/61.dafx2013_submission_32.pdf
- Thomas, L. (2007). Aplicação à música de um sistema de espacialização sonora baseado em Ambisonics. São Paulo: Escola Politécnica da Universidade de São Paulo. Disponível em: http://www.ime.usp.br/~lfthomaz/dissertacao_mestrado.pdf
- Valente, L., Conci, A. & Feijó, B. (2005). Real Time Game Loop Models for Single-Player Computer Games. *Proceedings of the Fourth Brazilian Symposium on Computer Games and Digital Entertainment*. Disponível em https://www.ssugames.org/pluginfile.php/1026/mod_resource/content/1/2005_sbgames.pdf
- Vogel, P. (1993, Dezembro). *Application of wave field synthesis in room acoustics*. Delft: Delft University of Technology. Disponível em: http://repository.tudelft.nl/assets/uuid:4d236099-096e-444c-bf40-b7b163076bf6/as_vogel_19931222.PDF
- Vorländer, M. (2008). *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*. Berlin: Springer.
- Wang, D., & Brown, G. J. (2005). *Computational auditory scene analysis*. (pp. Cap 5, 1-34). New York: John Wiley & Sons, Inc. Disponível em <http://www.cs.cmu.edu/~robust/Papers/SternWangBrownChapter.pdf>
- Wente, E. C. & Thuras, A. L. (1931). Moving-coil telephone receivers and microphones. *Journal of the Audio Engineering Society*, 3(1A), 44-45. Disponível em <http://scitation.aip.org/getpdf/servlet/GetPDFServlet?filetype=pdf&id=JASMAN00000300001A00004400001&idtype=cvips&doi=10.1121/1.1915544&prog=normal&bypassSSO=1>
- Wilhelmsson & Wallén (2010). A Combined Model for the Structuring of Computer Game Audio. Grimshaw, M. (ed.) *Game Sound Technology and Player Interaction: Concepts and Developments*. IGI Global.
- Wozniowski, M., Settel, Z. & Cooperstock, J. (2006). A paradigm for physical interaction with sound in 3-D audio space. *Proceedings of the International Computer Music Conference*, 456-463. Disponível em http://sheefa.net/zack/publications/audioscape_ICMC2006.pdf
- Yôiti, S., Brungart, D., Yukio, I., Kazuhiro, I., Cabrera, D., & Hiroaki, K. (2009). *Principles and applications of spatial hearing*. Zao: World Scientific.
- Zotter, F., Pomberger, H. & Noisternig, M. (2011, Dezembro). Energy-Preserving Ambisonic Decoding. *Acta Acustica United with Acustica*, 98(1), 37-47.
- Zotter, F., Pomberger, H., & M. Frank (2009) An alternative ambisonics formulation: Modal source strength matching and the effect of spatial aliasing. AES 126th Convention, Munich, Germany. Disponível em: <http://iem.kug.ac.at/fileadmin/media/iem/projects/2009/alternative.pdf>

5.2. Websites

- Axon (2008). *AX 50 USB*. Disponível em: <http://axon-technologies.net/modules.php?op=modload&name=News&file=article&sid=36&menu=102>. Consultado em Dezembro de 2012.
- Chmike (2009). *Stackoverflow. Circle line collision detection*. Disponível em <http://stackoverflow.com/questions/1073336/circle-line-collision-detection>. Consultado em Maio de 2012.
- GLASGAL, R. (s.d.) *Ambiophonics*. Disponível em: <http://www.ambiophonics.org/>. Consultado em Agosto de 2013.
- Harker, A. (2012). Disponível em <http://www.alexanderjharker.co.uk/Software.html>. Consultado em Fevereiro de 2012.
- HoaLibrary (2012). *HoaLibrary: High Order Ambisonics Library*. Disponível em <http://www.mshparisnord.fr/hoalibrary/en/>. Consultado em Janeiro de 2013.
- Hollerweger, F. (2008). *An Introduction to Higher Order Ambisonic*. Disponível em: <http://flo.mur.at/writings/HOA-intro.pdf>. Consultado em Novembro de 2012.
- Hull, J. (1999). *Surround sound past, present, and future, a history of multichannel audio from mag stripe to dolby digital*. Disponível em [http://pacificav.com/library/Dolby Surround History.pdf](http://pacificav.com/library/Dolby%20Surround%20History.pdf). Consultado em Novembro de 2012.
- ICST (2012). *Ambisonics Externals for MaxMSP*. Disponível em <http://www.icst.net/research/downloads/ambisonics-externals-for-maxmsp/>. Consultado em Dezembro de 2013.
- Jamoma* (s.d.). Disponível em: <http://jamoma.org/>. Consultado em Agosto de 2013
- Jarvis, M. (2009). *Making a 'Sonic Saber' game in MAX/MSP*. Disponível em: <http://www.mattjarvis.co.uk/creativity/sonic-saber-game-in-maxmsp-its-like-a-light-saber-only-sound-not-light kinda-thang>. Consultado em Novembro de 2012.
- Mechdyne* (s.d.). Disponível em: <http://www.mechdyne.com/>. Consultado em Setembro de 2013
- Olofsson, F. (2010). Disponível em <http://www.fredrikolofsson.com/pages/code-max.html>. Consultado em Março 2012.
- ReacTIVision (2005). Disponível em <http://reactivision.sourceforge.net/#files>. Consultado em Fevereiro 2012.
- Red Paper Heart . (2013). Disponível em: <http://www.redpaperheart.com/>. Consultado em Novembro de 2012.
- Seldess, Z. (2012a). *KVL Kinect Tracker. A depth-sensitive Kinect blob tracker/server*. Disponível em http://www.zacharyseldess.com/KVL_KinectTracker/. Consultado em Janeiro de 2013.
- Seldess, Z. (2012b). *KVL Kinect Tracker – Demo*. Disponível em <http://www.youtube.com/watch?v=bx02WIG7ooU&feature=youtu.be>. Consultado em Janeiro de 2013.
- The Gadget Show (2011). *Up close with the Battlefield 3 simulator!* Disponível em <http://gadgetshow.channel5.com/gadget-show/gadget-news/up-close-with-the-battlefield-3-simulator>. Consultado em Outubro de 2012
- TKK Acoustics (2006). *Vector base amplitude panning*. Disponível em <http://www.acoustics.hut.fi/research/cat/vbap/>. Consultado em Dezembro de 2012.
- Zvonar, R. (2005). *A history of spatial music. Canadian Electroacoustic Community*. Disponível em <https://pantherfile.uwm.edu/kdschlei/www/files/a-history-of-spatial-music.html> Consultado em Novembro de 2012.

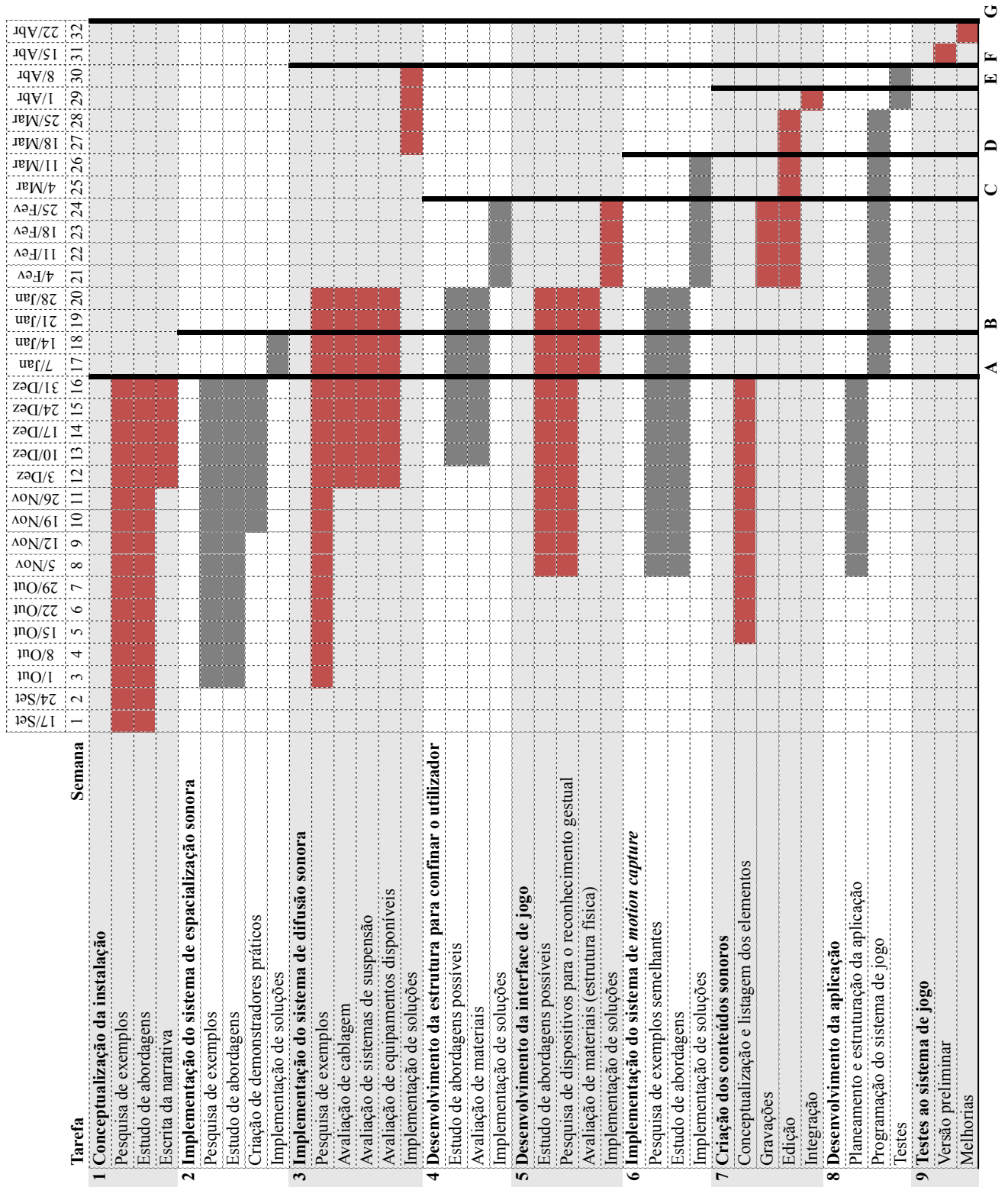
5.3. Jogos

- EA Redwood Shores (2008). *Dead Space*.
- Naughty Dog. (2013). *The Last of Us*.
- Somethin' Else. (2010). *Papa Sangre*.
- Tapulous (2008). *Tap Tap Revenge*.
- Valve Corporation (2007). *Portal*.
- Valve Corporation (2008). *Audiosurf*.
- Valve Corporation (2009). *Left 4 Dead 2*.
- WARP, Inc. (1997). *Real Sound: Kaze no Regret*

5.4. Patentes

- Blumlein, A. (1933). *British Patent No. 394,325*. Reimpresso em *Journal of the Audio Engineering Society*, 6(2), 32-40, 1958. Disponível em: <http://decoy.iki.fi/dsound/ambisonic/motherlode/source/British%20patent%20specification%20394325%20Blumlein%201933.pdf>
- Blumlein, A. (1934). *U.S. Patent No. 1,975,709*. London, England: United States Patent Office. Disponível em: <http://www.google.pt/patents?hl=pt-PT&lr=&vid=USPAT1975709&id=Lv1oAAAAEBAJ&oi=fnd&dq=alan+blumlein+1931&printsec=abstract#v=onepage&q=alan%20blumlein%201931&f=false>
- Fletcher, H. & Sivian, L. J. (1927). *U.S. Patent No. 1,624,486*. New York: United States Patent Office. Disponível em: <http://www.google.pt/patents?hl=pt-PT&lr=&vid=USPAT1624486&id=JetaAAAAEBAJ&oi=fnd&dq=harvey+fletcher+binaural+patent&printsec=abstract#v=onepage&q=harvey%20fletcher%20binaural%20patent&f=false>
- Gerzon, M. (1979). *U.S. Patent No. 4,151,369*. Oxford, England: United States Patent Office. Disponível em: <http://www.google.com/patents/US4151369>

Apêndice 1: Calendarização do Projeto



- A. Finalização da tarefa 1
- B. Finalização da tarefa 2
- C. Finalização das tarefas 4 e 5
- D. Finalização da tarefa 6
- E. Finalização da tarefa 7
- F. Finalização das tarefas 3 e 8
- G. Apresentação do projeto (24/Abril)

Apêndice 2: DVD com Conteúdos Digitais

O DVD contém três pastas com os seguintes conteúdos:

Pasta 1. A presente dissertação em suporte digital, no formato *Adobe Acrobat PDF*:

- *Dissertação – Miguel Duarte Gonçalves 375411004.pdf*

Pasta 2. O patch do projeto *The Sound of Horror*, no formato *Max/MSP*, versão 5.1.9:

- *The Sound of Horror.maxpat*

Os respectivos elementos sonoros, em formato *WAV*, com taxa de amostragem de 48 kHz e 24 *bits/sample*.

- | | | |
|-----------------------------------|---------------------------------------|---------------------------------------|
| • <i>MenuDiálogos_1.wav</i> | • <i>MenuDiálogos_2.wav</i> | • <i>MenuDiálogos_3.wav</i> |
| • <i>MenuDiálogos_4.wav</i> | • <i>TutorialDiálogos_1.wav</i> | • <i>TutorialDiálogos_2.wav</i> |
| • <i>TutorialDiálogos_3.wav</i> | • <i>TutorialDiálogos_4.wav</i> | • <i>TutorialDiálogos_5.wav</i> |
| • <i>TutorialDiálogos_6.wav</i> | • <i>TutorialDiálogos_7.wav</i> | • <i>CinematicJogoDiálogos_1.wav</i> |
| • <i>Nível1Diálogos_1.wav</i> | • <i>Nível2Diálogos_1.wav</i> | • <i>Nível3Diálogos_1.wav</i> |
| • <i>VitóriaDiálogos_1.wav</i> | • <i>GameOver.wav</i> | • <i>LevelCompleted.wav</i> |
| • <i>ArmaTiroMenu_1.wav</i> | • <i>ArmaTiroMenu_2.wav</i> | • <i>ArmaTiroMenu_3.wav</i> |
| • <i>ArmaTiro_1.wav</i> | • <i>ArmaTiro_2.wav</i> | • <i>ArmaTiro_3.wav</i> |
| • <i>ArmaTiro_4.wav</i> | • <i>ArmaTiro_5.wav</i> | • <i>ArmaReload_1.wav</i> |
| • <i>ArmaReload_2.wav</i> | • <i>ArmaReload_3.wav</i> | • <i>ArmaStab_1.wav</i> |
| • <i>ArmaStab_2.wav</i> | • <i>ArmaStab_3.wav</i> | • <i>ArmaTiroEmpty_1.wav</i> |
| • <i>ArmaTiroEmpty_2.wav</i> | • <i>ArmaTiroEmpty_3.wav</i> | • <i>Menu_1.wav</i> |
| • <i>Menu_2.wav</i> | • <i>Nível1_1.wav</i> | • <i>Nível1_2.wav</i> |
| • <i>Nível2_1.wav</i> | • <i>Nível2_2.wav</i> | • <i>Nível3_1.wav</i> |
| • <i>Nível3_2.wav</i> | • <i>Tutorial_1.wav</i> | • <i>Tutorial_2.wav</i> |
| • <i>CinematicInício_1.wav</i> | • <i>CinematicInício_2.wav</i> | • <i>CinematicTutorial_1.wav</i> |
| • <i>CinematicTutorial_2.wav</i> | • <i>CinematicTutorial-Menu_1.wav</i> | • <i>CinematicTutorial-Menu_2.wav</i> |
| • <i>CinematicJogo_1.wav</i> | • <i>CinematicJogo_2.wav</i> | • <i>Cinematic1-2_1.wav</i> |
| • <i>Cinematic1-2_2.wav</i> | • <i>Cinematic2-3_1.wav</i> | • <i>Cinematic2-3_2.wav</i> |
| • <i>Vitória_1.wav</i> | • <i>Vitória_2.wav</i> | • <i>Derrota_1.wav</i> |
| • <i>Derrota_2.wav</i> | • <i>MenuAlvo1.wav</i> | • <i>MenuAlvo1_Hit.wav</i> |
| • <i>MenuAlvo2.wav</i> | • <i>MenuAlvo2_Hit.wav</i> | • <i>Monstro1Nível1.wav</i> |
| • <i>Monstro1Nível1Ferido.wav</i> | • <i>Monstro1Nível1Morto.wav</i> | • <i>Monstro1Nível1Passos.wav</i> |
| • <i>Monstro2Nível1.wav</i> | • <i>Monstro2Nível1Ferido.wav</i> | • <i>Monstro2Nível1Morto.wav</i> |
| • <i>Monstro2Nível1Passos.wav</i> | • <i>Monstro3Nível1.wav</i> | • <i>Monstro3Nível1Ferido.wav</i> |
| • <i>Monstro3Nível1Morto.wav</i> | • <i>Monstro3Nível1Passos.wav</i> | • <i>Monstro1Nível2.wav</i> |
| • <i>Monstro1Nível2Ferido.wav</i> | • <i>Monstro1Nível2Morto.wav</i> | • <i>Monstro1Nível2Passos.wav</i> |
| • <i>Monstro2Nível2.wav</i> | • <i>Monstro2Nível2Ferido.wav</i> | • <i>Monstro2Nível2Morto.wav</i> |
| • <i>Monstro2Nível2Passos.wav</i> | • <i>Monstro3Nível2.wav</i> | • <i>Monstro3Nível2Ferido.wav</i> |
| • <i>Monstro3Nível2Morto.wav</i> | • <i>Monstro3Nível2Passos.wav</i> | • <i>Monstro1Nível3.wav</i> |
| • <i>Monstro1Nível3Ferido.wav</i> | • <i>Monstro1Nível3Morto.wav</i> | • <i>Monstro1Nível3Passos.wav</i> |

- *Monstro2Nível3.wav*
- *Monstro2Nível3Passos.wav*
- *Monstro3Nível3Morto.wav*
- *Monstro4Nível3Ferido.wav*
- *TutorialAlvo1.wav*
- *TutorialAlvo2Hit.wav*
- *Monstro2Nível3Ferido.wav*
- *Monstro3Nível3.wav*
- *Monstro3Nível3Passos.wav*
- *Monstro4Nível3Morto.wav*
- *TutorialAlvo1Hit.wav*
- *Monstro2Nível3Morto.wav*
- *Monstro3Nível3Ferido.wav*
- *Monstro4Nível3.wav*
- *Monstro4Nível3Passos.wav*
- *TutorialAlvo2.wav*

Os objetos externos *Max/MSP* utilizados no projeto:

- *ajh.wiiremote.mxo*
- *ambicontrol.mxo*
- *ambidecode~.mxo*
- *ambiencode~.mxo*
- *ambimonitor.mxo*
- *f0.smooth.mxo*

Pasta 3. O *trailer* sonoro do projeto *The Sound of Horror* no formato *WAV*, com taxa de amostragem de 48 kHz e 24 *bits/sample*.

- *Trailer – The Sound of Horror – Miguel Gonçalves 375411004.wav*