


Imagem sonora imersiva com arranjos circulares de transdutores espaçados em segmentos iguais

José Augusto Mannis 

Universidade Estadual de Campinas, Instituto de Artes, Depto. de Música, LASom | Campinas, SP, Brasil

Igor Abdo Aguiar 

Universidade Estadual de Campinas, Instituto de Artes | Campinas, SP, Brasil

Amanda Chang Rodrigues 

Universidade Estadual de Campinas, Instituto de Artes | Campinas, SP, Brasil

Resumo: Este artigo apresenta o desenvolvimento de uma pesquisa em gravação e reprodução de áudio imersivo multicanal ocorrida de 2016 a 2020 na Universidade Estadual de Campinas (Brasil), que resultou em arranjos circulares de transdutores (microfones e caixas acústicas) espaçados em segmentos iguais. A mesma tecnologia de captação sonora multicanal imersiva já havia sido formulada em 1991 pelo engenheiro de som Michael Williams. Por caminhos distintos, os dois estudos chegaram a um resultado comum. Reconhecendo o mérito do achado a Williams, detalha-se, neste artigo, o percurso da pesquisa brasileira, com comentários complementares que ilustram a potência e a riqueza dos fenômenos envolvidos na solução do problema. Também são apresentados os desdobramentos do produto inicial, notadamente o desenvolvimento de uma ferramenta de simulação acústica e auralização com aplicações em artes, comunicação, bioacústica, ecologia e monitoramento ambiental. Ao final apresentamos composições musicais em áudio espacial destacando as facilidades expressivas dos novos recursos desenvolvidos.

Palavras-chave: Áudio imersivo, Espacialização sonora, Paisagem sonora, Gravação de som, Sonologia.

Abstract: This article presents the development of research on multichannel immersive audio recording and reproduction from 2016 to 2020 conducted at the Universidade Estadual de Campinas - Unicamp (Brazil), which resulted in circular arrangements of transducers (microphones and speakers) spaced in equal segments. Similar technology was previously developed by sound engineer Michael Williams, who published his findings in 1991. Despite following different paths, both studies reached similar conclusions. Acknowledging Williams' contributions, the article details the Brazilian research journey, highlighting the phenomena involved and the innovations achieved. Key developments include an acoustic simulation and auralization tool applicable to arts, communication, bioacoustics, ecology, and environmental monitoring. Finally, the article features spatial audio musical compositions demonstrating the expressive potential of these new technologies.

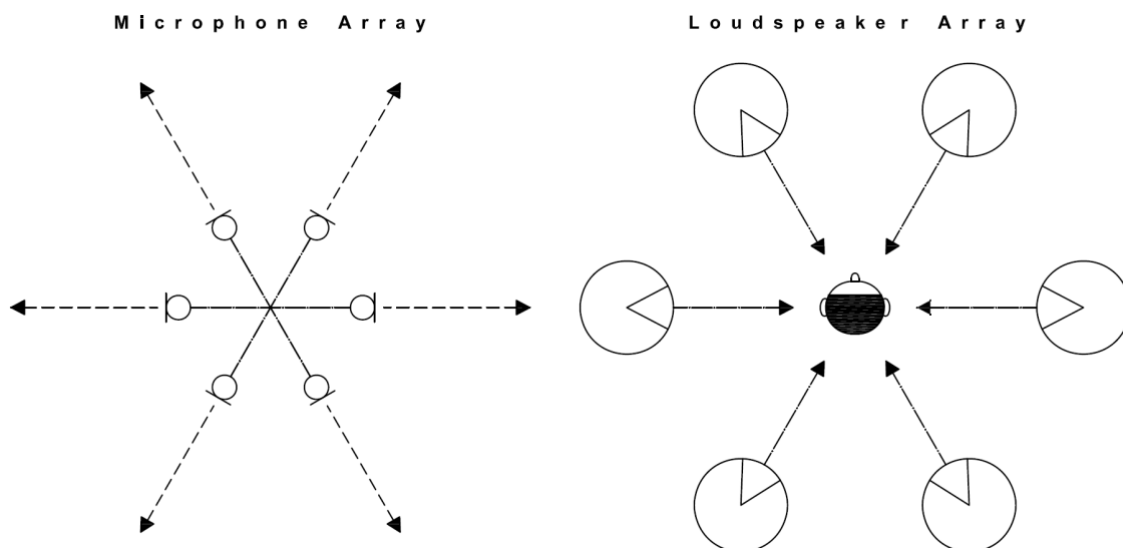
Keywords: Immersive audio, Sound spatialization, Soundscape, Sound recording, Sonology.

Este artigo relata uma pesquisa em áudio imersivo que, após seu desenvolvimento, foi submetida a verificação de anterioridade, resultando na recuperação de uma mesma solução para um mesmo problema, encontrada por Michael Williams, publicada em 1991. Além do aprendizado sobre a importância dessa verificação antes de se iniciar uma pesquisa, se ela foi realizada sem que isso tenha sido feito com maior abrangência, como no nosso caso, entendemos que a prioridade, então, é atender com honestidade à pergunta de pesquisa inicialmente formulada e relatar, também com honestidade e a maior clareza possível, o processo da nossa investigação. Com a sua conclusão, podemos afirmar que nossos resultados corroboram e validam o que foi produzido antes. Segue, assim, um texto que é um misto de relatos experimentais, desdobramentos tecnológicos e aplicações artísticas. Esperamos que nossas observações durante a pesquisa contribuam para a apreensão mais ampla do produto comum que ela reencontrou, sendo, assim, uma base sólida para os desdobramentos que imaginamos.

Michael Williams iniciou sua carreira em áudio profissional na BBC, em Londres, em 1960. Mudou-se para a França em 1965. A partir de 1980, tornou-se consultor e instrutor *freelancer*, atuando nas principais empresas francesas de *broadcast* (rádio e televisão) e em escolas e instituições de formação de áudio profissional (AES Paris 2016 presenter or author, 2016). Em 1991, em Nova York, participou da 91ª Convenção da Audio Engineering Society, apresentando o trabalho *Microphone arrays for natural multiphony* [*Arranjos de microfones para multifonia natural*], no qual postula que “o campo sonoro ao redor de um arranjo [*array*] multicanal de microfones pode ser definido em um *set* de segmentos adjacentes inscritos num plano [horizontal] de referência” (Williams, 1991, p. 1, tradução nossa), e propõe três possíveis soluções. O objetivo era captar o som e reproduzi-lo posteriormente, reconstituindo virtualmente o campo sonoro do instante da gravação, com precisão de localização e mínima distorção geométrica da imagem sonora. Para a reprodução, aplica-se, às caixas acústicas, a mesma configuração dos microfones. As três configurações que ele propôs foram: arranjo de seis segmentos, arranjo de cinco segmentos e arranjo de quatro segmentos. A estrutura prevista sobre cada segmento precisa ser tal que as características de diretividade, o ângulo de orientação e a distância entre microfones contribuam para uma imagem sonora apropriada, tendo cada uma delas um ângulo de largura da imagem estereofônica – *stereo recording angle* (SRA) – que

assegure a continuidade das imagens adjacentes sem lacunas ou sobreposições. A configuração de seis segmentos para microfones e caixas acústicas proposta por Williams é mostrada na Figura 1.

FIGURA 1 – À esquerda, dispositivo de captação sonora, com seis microfones; à direita, dispositivo de escuta, com seis caixas acústicas.



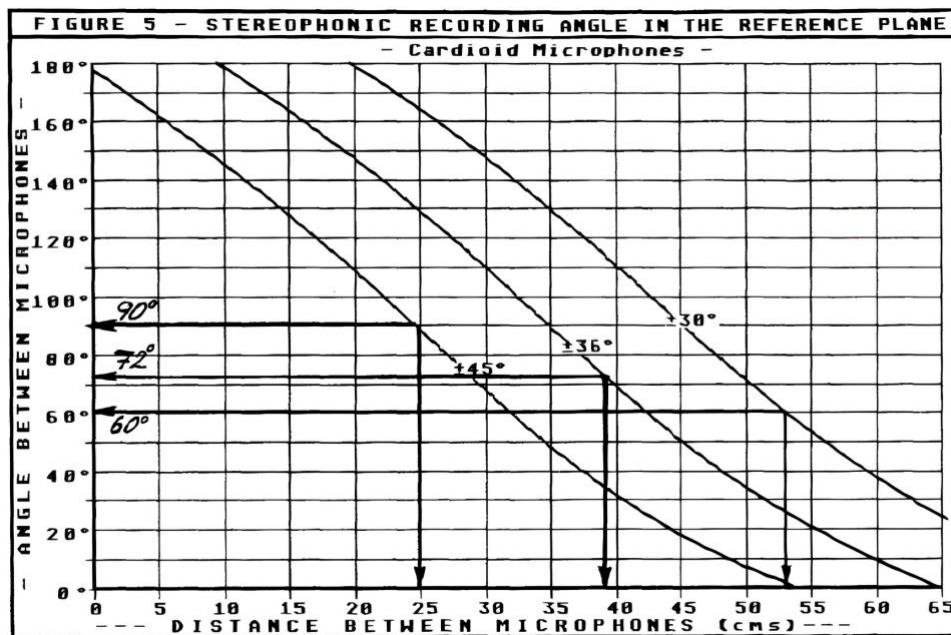
Fonte: Williams (1991, p. 7).

Considerando a diretividade cardioide dos microfones e a angulação de 60° entre os microfones de cada segmento, para a melhor imagem sonora geral de todo o entorno a distância entre eles é dada pelo gráfico a seguir (Figura 2)¹⁻². Como a imagem total do entorno compreende 360° , para a continuidade entre as imagens adjacentes as imagens captadas em cada segmento devem ter largura de $360^\circ/6 = 60^\circ$, ou seja, $\pm 30^\circ$. Para essa condição, a distância entre os microfones deve ser de aproximadamente 53 cm, conforme a Figura 2.

¹ Para microfones com outros diagramas de direcionalidade (hipercardiodes, hipocardiodes etc.) o autor elaborou outros gráficos de correlação.

² A Figura 2 é fac-símile do original realizado em 1991. Não tivemos acesso aos dados levantados empiricamente por Williams para plotar o gráfico em alta definição. Por essa razão, reproduzimos o original com definição melhorada através de AI Upscaler (disponível em <https://imgupscaler.com/>).

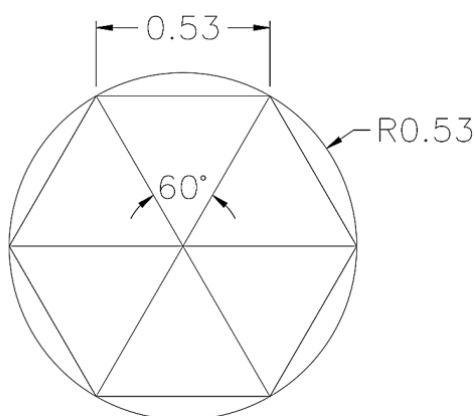
FIGURA 2 – Gráfico que correlaciona o ângulo entre os microfones, o ângulo da abertura da imagem sonora e a distância entre os microfones.



Fonte: Williams (1991, p. 10).

Como o lado do hexágono (Figura 3) é igual ao raio da circunferência na qual se inscreve, temos uma circunferência com 1,06 m de diâmetro. Com exceção da distância de 53 cm entre os microfones, essa configuração foi a mesma que resultou de nossos experimentos.

FIGURA 3 – Traçado de microfones cardioides para imagem virtual contínua de 360°.



A seguir, mostramos como nossa reflexão se iniciou a partir de estudos de estereofonia e de localização e percepção espacial do som em psicoacústica.

1. A estereofonia e o áudio espacial

A estereofonia se caracteriza pela escuta com relevo, ou seja, discernindo três dimensões. Seu precursor foi o francês Clément Ader, em 1881, em Paris, quando criou o *théâtrophone*³, que transmitia, em tempo real, via telefone e em dois canais de áudio, espetáculos musicais e teatrais que estavam acontecendo no Palais Garnier⁴, para ouvintes em suas residências munidos de fones de ouvido (The telephone at the Paris Opera, 1881). Quanto à estereofonia moderna, esta teve início com o britânico Alan Blumlein, na década de 1930, na EMI, que patenteou (GB 394325) vários processos e métodos para a captação e reprodução de som em dois canais (Blumlein, 1931), notadamente o *Blumlein pair* (par Blumlein), com dois microfones figura-8 rotacionados de 90° um em relação ao outro, configuração que se classifica na categoria denominada XY, e o Mid/Side (M/S), com um microfone cardioide e um figura-8, sendo, ambas as configurações, de microfones coincidentes, que produzem imagem sonora clara com movimentos esquerda-direita perceptíveis (Shankleman, 2008). As configurações de microfones coincidentes são aquelas em que todas as cápsulas (membranas) de captação têm diretividade seletiva e se situam (praticamente) num mesmo ponto. Dessa maneira, cada uma recebe ondas sonoras de uma determinada direção. Havendo dois microfones, um orientado à esquerda e outro à direita da cena sonora frontal, esta será separada em dois canais. Ao reproduzir esses sinais individualmente em duas caixas acústicas diante do ouvinte (ITU, 2022), o sinal que está à esquerda soará mais forte no canal esquerdo, e o mesmo acontece com o sinal à direita no canal direito. Assim, pelas diferenças de intensidade entre os canais reproduzidos, a imagem sonora estereofônica se reconstitui à escuta. Trata-se, portanto, de uma estereofonia de intensidade, pois são as diferenças de intensidade (nível sonoro) percebidas pelas duas orelhas⁵ do

³ Comercializado em Paris de 1890 a 1932.

⁴ Conhecido como Ópera de Paris.

⁵ Em 2001, a Sociedade Brasileira de Anatomia (SBA) publicou *Terminologia anatômica*, tradução de *Terminologia anatômica internacional (International anatomical terminology)*, elaborada em 1998 pelo Comitê Federativo de Terminologia Anatômica (FCAT, na sigla em inglês). Desde então, ficou padronizado que, em português brasileiro, a denominação do órgão auditivo é *orelha* (SBA, 2001, p. 179-183, A15.3.00.001): orelha externa, orelha média e orelha interna. Em Portugal, mantém-se a denominação de *ouvido*.

ouvinte que induzem a orientação e a localização das fontes sonoras na cena registrada (Blauert, 1999, p. 82; Dickreiter, 1989, p. 332).

Outra configuração de captação estereofônica é a de microfones não coincidentes, também denominada *spaced microphones* (microfones espaçados) (Ballou, 1991; Slotte, 2004), ou *spaced pair* (par espaçado) (Williams, 1987), cujas cápsulas (preferencialmente omnidirecionais) se posicionam em diferentes pontos no espaço de gravação. Nessa disposição as ondas sonoras incidirão em cada um dos microfones em diferentes instantes, num intervalo na ordem de milésimos de segundo, mas o suficiente para provocar, nas orelhas do ouvinte, a sensação de estarem vindo da esquerda ou da direita da cena sonora gravada. A diferença de tempo de chegada da frente de onda em cada microfone se denomina *time-of-arrival* (Dickreiter, 1989) ou *arrival time* (Blauert, 1999). Essa configuração em dois canais para captação sonora com atraso de incidência da frente de onda em cada microfone se denomina AB.

Há uma configuração híbrida entre a coincidente e a não coincidente, denominada quase coincidente (*near-coincident*) (Slotte, 2004) ou *closely spaced microphones* (Dickreiter, 1989; Rossing, 2014), compreendendo microfones com diretividade seletiva e geralmente espaçados entre 17 e 30 cm. Na configuração AB, a distância entre os microfones varia conforme os autores: de 10 a 48 cm (Brixen, [2024]); de 20 cm a 1,0 m (Dickreiter, 1989, p. 98); de 25 a 50 cm (Fellot, 2007, p. 106); de 90 cm até 3 m (Microphone techniques for recording, 2014); ou, no máximo, até a distância de campo difuso da sala (Slotte, 2004, p. 12). Se as cápsulas dos microfones são omnidirecionais e se ambos estiverem separados pela distância de um a três metros (*AB - spaced pair*), e, ainda, se os sinais gravados em cada canal forem reproduzidos individualmente em duas caixas frontais (ITU, 2022), os acontecimentos na cena sonora gravada estarão separados em cada canal com o mesmo atraso com que a onda sonora incidiu em cada microfone e haverá diferenças de intensidade correspondendo às distâncias entre as fontes sonoras e cada microfone. Se na cena frontal a fonte sonora estiver exatamente sobre a linha mediana entre as cápsulas, o atraso da chegada da frente de onda entre elas será nulo. Quanto mais as fontes sonoras estiverem para os lados, maior o atraso de chegada em cada cápsula. Os atrasos distintos de cada fonte sonora e as variações de intensidade são suficientes para reconstituir, à escuta do ouvinte, a cena sonora gravada com relevo estereofônico.

Segundo Benedict Slotte (2004, p. 1-2), as características subjetivas das mais conhecidas “técnicas convencionais de gravação estéreo podem ser representadas ao longo de um eixo entre dois extremos” (tradução nossa): de um lado a configuração XY (microfones coincidentes cardioides), que proporciona uma imagem na qual as fontes sonoras são bem delineadas e têm localização precisa, mas que carecem de profundidade; de outro a configuração AB larga, com omnidirecionais espaçados⁶, o que resulta em uma imagem com efetivo relevo sonoro e planos de profundidade, mas difusa e menos delineada. No centro do eixo, equilibrando as características de ambas, encontram-se configurações das técnicas quase coincidentes obtidas a partir da configuração AB com microfones cardioides, que são a ORTF, desenvolvida na década de 1960 pelo Serviço de Radiodifusão-Televisão Francesa, e duas variações dela, concebidas na mesma década: a NOS, pela Fundação Holandesa de Radiodifusão, e a DIN, pelo Instituto Alemão de Normalização⁷. Na ORTF os microfones estão orientados, aproximadamente, como numa cabeça humana, separados de 17,5 cm e com abertura horizontal de 110°, captando tanto as variações de intensidade como as diferenças de tempo de chegada da frente de onda (Condamines, 1978, p. 116 e seg.), o que permite uma escuta com destacada sensação de relevo sonoro, profundidade e boa localização. A configuração de microfones quase coincidentes, híbrida de coincidentes (XY) e não coincidentes (AB) proporciona precisão de localização e delineamento das fontes sonoras à profundidade da imagem, com efetiva sensação de relevo sonoro, sendo um bom compromisso entre as qualidades dos extremos (Slotte, 2004, p. 2). As técnicas ORTF, NOS e DIN proporcionam, portanto, uma imagem sonora frontal com boa localização, definição, relevo e profundidade, como uma tela sonora diante de um ouvinte com a cabeça imóvel, voltada para a frente, sentado na plateia com a atenção no palco.

Um impulso semelhante ao que em 1931 levou Blumlein a fazer o som de uma personagem acompanhar seu movimento na tela de cinema, em 1940 motivou engenheiros da então Walt Disney Picture e da RCA a fazerem o som se mover, não somente pela tela, mas por toda a sala de cinema, na

⁶ Quanto mais distantes estiverem os microfones omnidirecionais na configuração AB, maiores serão as diferenças de intensidade

⁷ As siglas ORTF, NOS e DIN se referem às iniciais das instituições onde essas técnicas foram desenvolvidas: respectivamente, Office de Radiodiffusion-Télévision Française, Nederlandse Omroep Stichting e Deutsche Institut für Normung.

projeção de *Fantasia*, que estreou num teatro na Broadway, em Nova York (Torick, 1998). O sistema tomava os sinais sonoros de três pistas ópticas do filme e os distribuía por até oito caixas acústicas ao redor da plateia. Em 1955, Karlheinz Stockhausen performou *O cântico dos adolescentes*, na WDR, em Colônia (Alemanha), em quatro pistas difundidas em quatro caixas acústicas ao redor da plateia (Chang, 2015), e, em 1967, o Pink Floyd fez um grande concerto quadrifônico em Londres (A brief history of surround sound, 2021). O som que envolvia a audiência tornou-se o *surround*, e, em 1987, Dominique Bertrand sonorizou o Moulin Rouge, em Paris, com um dispositivo que logo se tornou referência: o *surround* 5.1, normatizado pela União Internacional de Comunicações (ITU, 2022).

Sistemas de áudio multicanal possibilitam uma projeção sonora com a distribuição de várias caixas de som no espaço de escuta. Para o ouvinte estar acusticamente envolvido, os eventos sonoros devem lhe parecer como pertencentes a um espaço identificado e constituído à sua escuta, e ele deve se sentir imerso no campo sonoro (*surround*). A apreensão sonora do espaço ocorre através da qualidade e da localização daquilo que se escuta. Para caracterizar a escuta do espaço, Jens Blauert (1999) propôs o conceito de audição espacial, mas alertando que não há possibilidade de uma “audição não espacial”, uma vez que tudo o que se escuta está necessariamente em algum lugar do espaço. A qualidade espacial da escuta requer uma identificação e uma aceitação do ouvinte. A familiaridade com o espaço pode ocorrer por verossimilhança (quando há o reconhecimento de uma sensação anteriormente experienciada), mas é dominada pela força, intensidade e tactibilidade do relevo sonoro percebido. Se “a audição espacial abrange relações entre as localizações dos eventos auditivos e outros parâmetros, particularmente entre eventos sonoros (que causam eventos audíveis) e outros relacionados à fisiologia do cérebro” (Blauert, 1999, p. 3, tradução nossa), a imersividade de um ambiente sonoro também abrange essas relações e depende da qualidade daquelas que são percebidas pelo ouvinte entre os diferentes fluxos sonoros pelos canais de difusão, e, evidentemente, de como cada um deles chega às suas orelhas após se propagar no ambiente no qual ocorre a escuta⁸. Portanto, as qualidades do espaço sonoro percebido com sistema *surround* multicanal, entre as quais

⁸ Diversos autores denominam o ambiente de escuta de maneira distinta. Por exemplo, Denis Smalley (2007) o chama de “espaço de escuta” e Fliblio F. de Souza (2023), de “espaço de fruição”.

a imersividade, dependem do ambiente de escuta, da configuração do sistema de projeção sonora e da relação entre os sinais de cada canal, para que haja relevo sonoro suficiente. Contudo, essa sensação é apreendida em uma percepção global:

Um conjunto objetivo de elementos identificados em meio a um todo pode caracterizar uma paisagem. Essa distribuição geométrica fornece, ao observador, uma sensação mais ou menos intensa de relevo. (...) Vamos examinar a sensação de relevo. A questão *não é* saber se tal ou tal elemento está mais ou menos distante, esculpido de maneira mais cinzelada ou mais arredondada. Trata-se de uma percepção global, do relevo percebido como um ‘todo’ psicológico, assim como o amargo, o quente, o apimentado etc. Poder definir (...) a posição, a orientação, o distanciamento relativo entre os elementos é uma simples possibilidade. Essa faculdade é secundária em relação à sensação primária, sintética, de relevo (Condamines, 1978, p. 25, grifo do original, tradução nossa).

Segundo Blauert (1999, p. 164), podemos separar em duas classes os atributos de diferença de sinal interaural, que causa deslocamentos laterais de eventos auditivos: a diferença de nível (sonoro) interaural (*interaural level difference-ILD*) e a diferença de tempo (de chegada da frente de onda) interaural (*interaural time difference-ITD*). Elas remetem, respectivamente, às configurações de microfones coincidentes e microfones não coincidentes, como já vimos anteriormente. O ITD, em razão da distância entre as cápsulas e do comprimento da onda incidente, age sobre diversos parâmetros, como *interaural phase difference* e *interaural arrival-time difference* (Blauert, 1999), e caracteriza a *delay stereophony* ou *time-of-arrival stereophony* (Dickreiter, 1989, p. 82).

As relações entre ILD e ITD são demasiadamente complexas para serem definidas através de algum critério quantitativo (Blauert, 1999, p. 165). Além disso, ambas atuam de maneira distinta no espectro audível e seus efeitos na localização espacial das fontes sonoras variam em função do timbre e de diversas características dos materiais sonoros (Blauert, 1999, p. 204 e seg.). Elas respondem de maneira complementar na localização de fontes sonoras durante audição com duas caixas acústicas frontais na disposição padronizada (ITU, 2022). As configurações de captação sonora quase coincidente são sensíveis a ambas. Nas configurações coincidentes, XY e M/S, próprias à estereofonia de intensidade, os fenômenos de diferença de fase e de diferença de tempo de chegada (atraso) não têm impacto na espacialidade como ocorre nas demais configurações.

Apesar de as configurações citadas apresentarem dois microfones para gravação em dois canais, há configurações estereofônicas com três microfones para gravação estereofônica em dois canais (Dickreiter, 1989; Fellot, 2007; Slotte, 2004), e há, também, a estereofonia multicanal em sistemas *surround* (Baskind et al., 2012; Fellot, 2007; Lyzwa, 2006; Messonnier; Lyzwa; Ling, 2022; Périaux; Ohl; Thévenot, 2015; Slotte, 2004). Salientamos que a qualidade estereofônica se refere especificamente à sensação de percepção espacial com relevo sonoro. As relações de intensidade, de tempo e de fase também precisam ser mantidas e controladas nos sistemas *surround* multicanal, sem o que não há relevo e o resultado sonoro não passaria de uma heterofonia. Por ser determinante da qualidade de sensações de percepção espacial humana, é fundamental que a estereofonia esteja compreendida na estrutura e nos procedimentos dos sistemas de áudio imersivo.

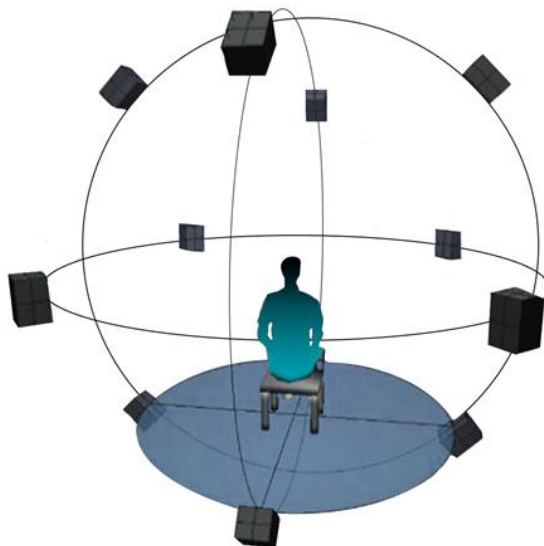
2. Sistema imersivo de captação e reprodução sonora multicanal abrangendo 360° com localização por diferença de intensidade, de tempo (atraso) e livre mobilidade no *sweet spot*

As técnicas de áudio imersivo são:

Ambisonics: sistema de codificação, registro e reprodução sonora (Figura 4) de áudio tridimensional, desenvolvido a partir da década de 1970 na Universidade de Oxford (Reino Unido) por Michael Gerzon e Peter Fellgett. Emprega um *array* de microfones coincidentes (Figura 5) e gravação no formato de áudio *B-format* (Figura 6), sendo, grosso modo, uma extensão do princípio de captação coincidente M/S em estéreo a dois canais (Figura 7), inventado na década de 1930 por Blumlein (What is Ambisonics?, [2024])⁹.

⁹ De 2003 a 2013, o Laboratório de Sistemas Integráveis da Escola Politécnica da USP (LSI-EP-USP) desenvolveu e lançou o sistema Audience, de processamento de áudio espacial (USP-LSI, [2024]), concebido por Régis Rossi Alves Faria (2005), utilizando Ambisonics, para produção, distribuição e reprodução de áudio 2D/3D.

FIGURA 4 – Exemplo de reprodução de áudio imersivo em Ambisonics.



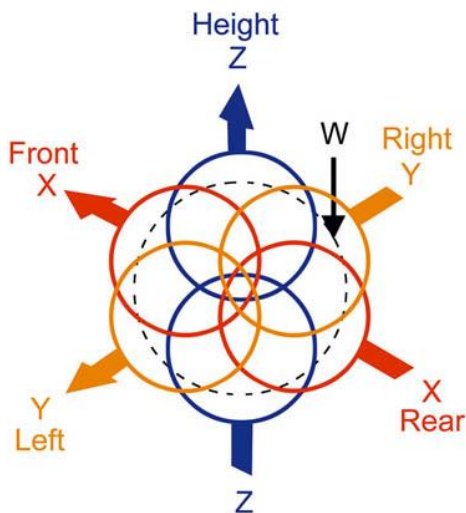
Fonte: Introduction to Ambisonics (2017).

FIGURA 5 – Arranjo de microfones coincidentes Ambisonics de 1ª ordem.



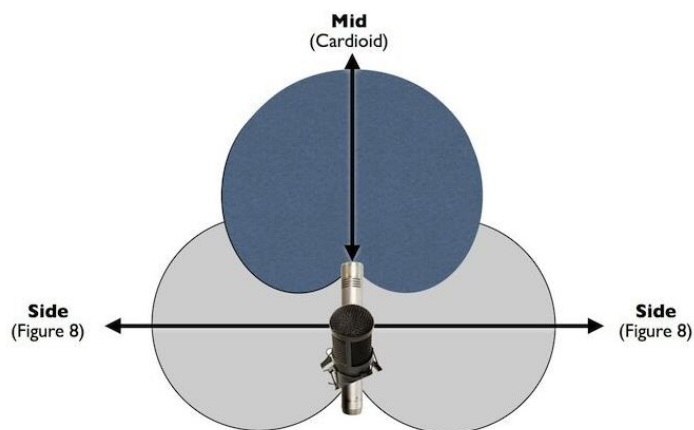
Fonte: Zumer (2021).

FIGURA 6 – *B-format* Ambisonics de 1ª ordem: microfones W (omnidirecional) processado com X (figura-8) como M-S frente-atrás; W é processado com Y (figura-8) como M-S esquerda-direita; W é processado com Z (figura-8) como M-S acima-abaixo.



Fonte: Pickett (2018).

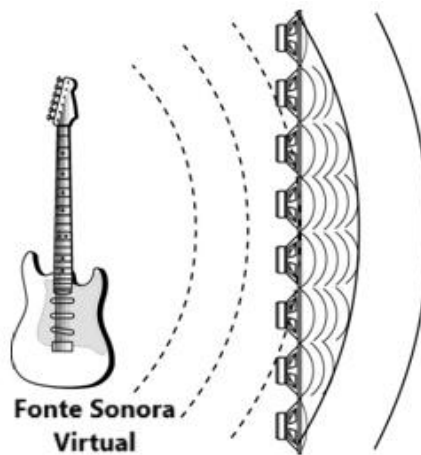
FIGURA 7 – Arranjo de microfones M/S (Mid - cardioide / Side - figura-8) com diagramas de diretividade.



Fonte: Zumer (2021).

Wave field synthesis (WFS): sistema de renderização de áudio espacial que cria campos sonoros que reproduzem frentes de onda artificiais sintetizadas por um grande número de alto-falantes colocados próximos uns dos outros, acionados individualmente (Figura 8). Foi idealizado por A. J. Berkhout (1988).

FIGURA 8 – Princípio de simulação de frente de onda com WFS.



Fonte: López (2007, tradução nossa).

Binaural ou **biauricular**: técnica de gravação com cabeça artificial (*dummy head*) (Figura 9), a ser escutada preferencialmente com fones de ouvido (Paul, 2009).

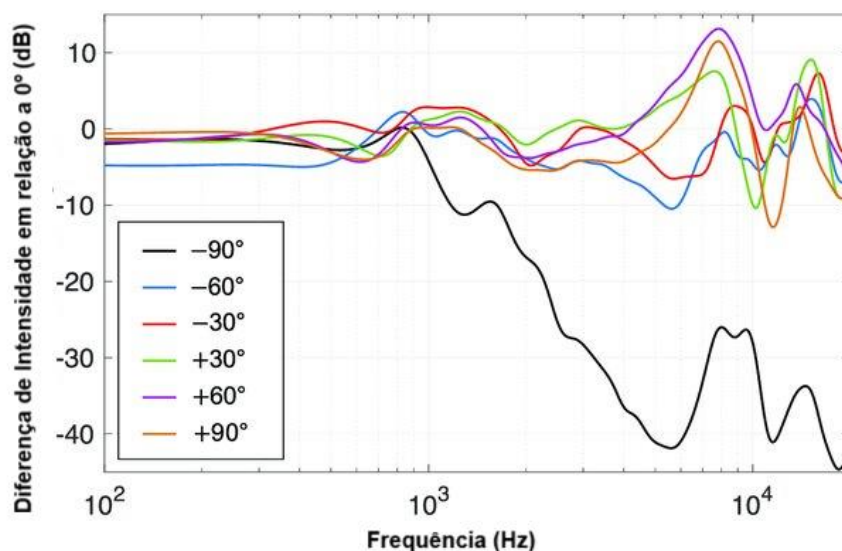
FIGURA 9 – Cabeça artificial (*dummy head*): gravação com dois microfones, um dentro de cada orelha do modelo.



Fonte: Dummy head (2024).

Head-related transfer function (HRTF): processamento de áudio binaural que reproduz, na escuta em fones de ouvido, diferenças de som percebidas em cada uma das orelhas humanas (Figura 10).

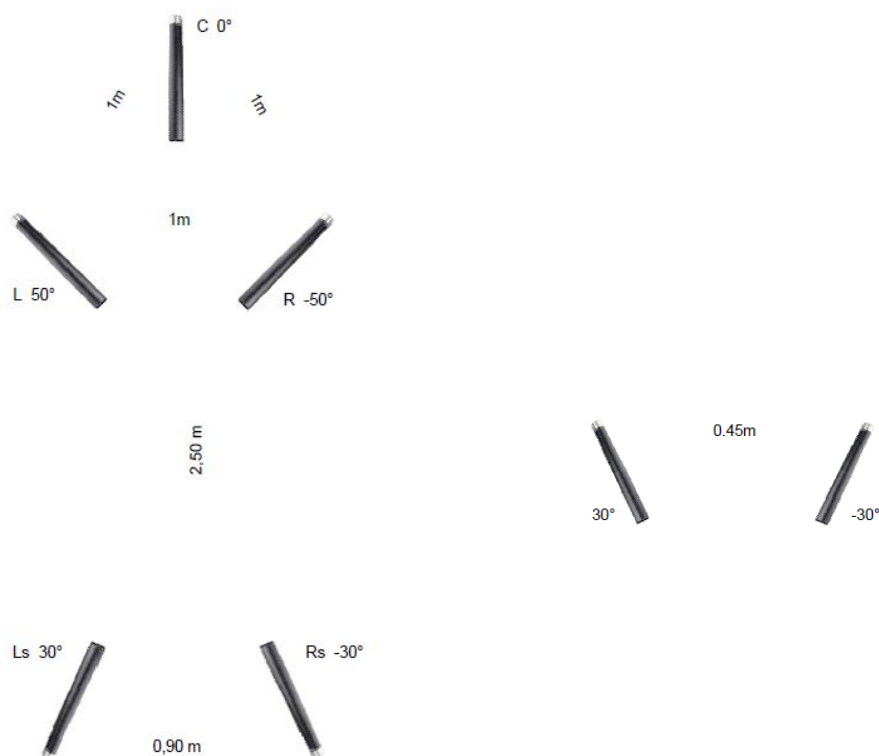
FIGURA 10 – Exemplos de curvas de diferenças de intensidade em função da frequência do som percebido em posições relativas à frontal (0°).



Fonte: Berwick e Lee (2020, p. 5).

Técnicas de espacialização binaural e transaural integradas à captação sonora multicanal e mixagem em *surround*: esse procedimento alia a tradicional mixagem *surround*, usando panorâmico *constant-power*, à tomada de som multicanal em *surround* (quinteto de microfones L-C-R-Ls-Rs) (Figura 11, à esquerda), estando a *Decca Tree* (L-C-R) do quinteto de microfones situada sobre o prosicênio, um par central AB (Figura 11, à direita) recuado em relação à *Decca Tree* e um microfone de proximidade sobre cada instrumento.

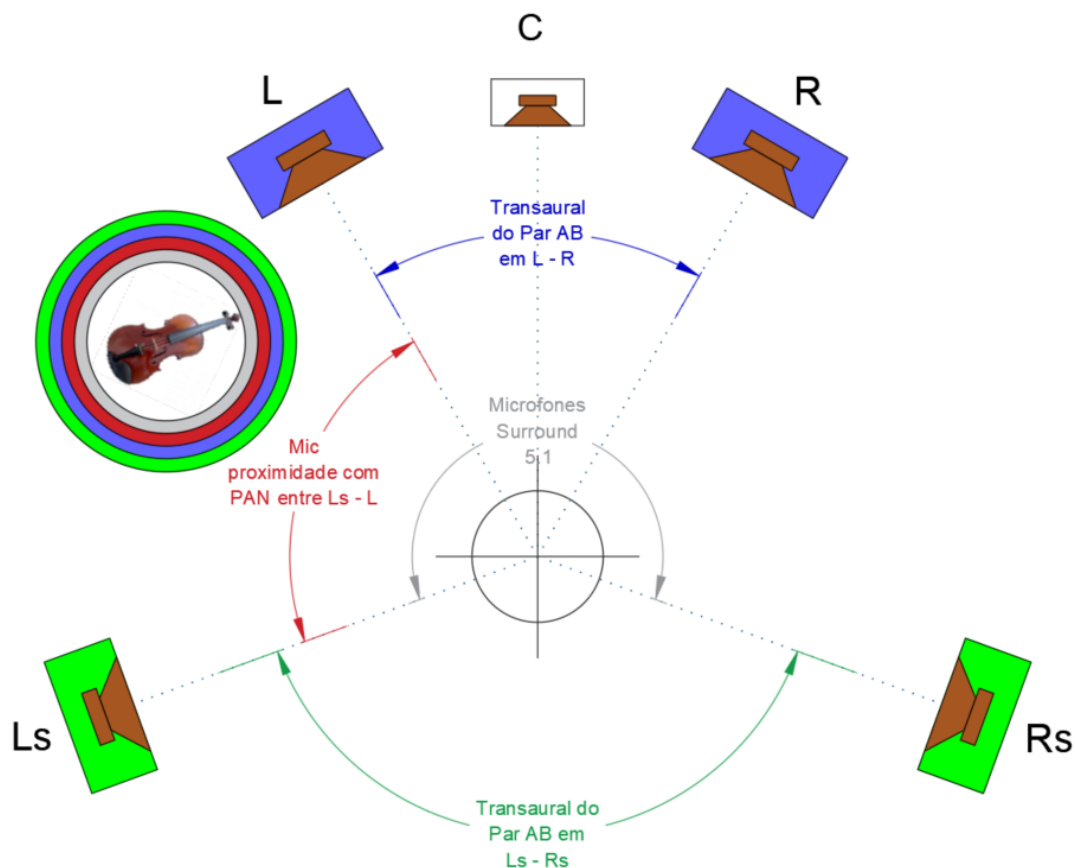
FIGURA 11 – À esquerda, quinteto de microfones *surround* (L-C-R-Ls-Rs): a parte dianteira L-C-R, equivalente a uma *Decca Tree*, está sobre o prosccênio. À direita, par de microfones AB, localizado centralmente e recuado em relação à *Decca Tree*.



Fonte: Lyzwa (2003, p. 15-16).

Posteriormente, acrescentam-se novas camadas de espacialização em formato binaural (HRTF) com processamento transaural (Baskind et al., 2012; Lyzwa, 2006), este último com a função de adaptar sinais binaurais para a reprodução em alto-falantes, compensando a diafonia da emissão sonora de cada alto-falante para cada orelha contralateral, procedimento também conhecido como cancelamento por interferência cruzada (*crosstalk cancellation*) (Atal; Hill; Schroeder, 1962; Masiero; Fels; Vorländer, 2011) ou *crosstalk compensation* (Kuttruff, 2007, p. 431). Esta última etapa do processamento estabiliza definitivamente a localização espacial da fonte sonora, podendo o ouvinte girar a cabeça ou se mover na plateia sem que haja alteração na percepção de localização de cada fonte sonora (Figura 12).

FIGURA 12 – Exemplo de estabilização da localização espacial de um violino em *surround* 5.1: em cinza, posição natural captada pelos cinco microfones *surround*; em magenta, sobreposição do microfone de proximidade do violino com panorâmico ajustado entre os canais L e Ls; em azul, tratamento transaural em HRTF nos dois canais dianteiros (L e R) dos sinais dos microfones AB; em verde, outro tratamento transaural do par AB nos canais posteriores (Ls e Rs).



Fonte: Lyzwa (2003, p. 25; tradução e adaptação de J. A. Mannis).

Trata-se de um procedimento muito eficiente, metódico e complexo, demonstrado publicamente, em 2009, em São Paulo, por Jean-Marc Lyzwa e Alexis Baskind, em sessão especial do 7º Congresso de Engenharia de Áudio da AES-Brasil. Durante a apresentação os ouvintes podiam se mover pela plateia, mas a imagem sonora permanecia espacialmente estável.

Na nossa pesquisa, pretendíamos chegar a um sistema que mantivesse os recursos perceptivos de diferenças de intensidade (*interchannel level difference-ICLD*) e atraso (*interchannel time difference-ICTD*). Além disso, buscamos criar uma solução que fosse estruturalmente simples e de baixo custo. Com financiamento do CNPq, iniciamos a pesquisa em 2016, na Unicamp, em

cooperação com o Laboratório de Sinais, Multimídia e Telecomunicações (SMT)¹⁰ do Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia da UFRJ (COPPE-UFRJ), e realizamos experimentos para uma captação sonora abrangendo 360°. Partimos da configuração AB *spaced pair*, com microfones omnidirecionais distribuídos em cada uma das seis hastes de um guarda-sol de piscina de três metros de diâmetro (na Figura 14, mais adiante, a imagem à esquerda mostra essa disposição em três diâmetros diferentes).

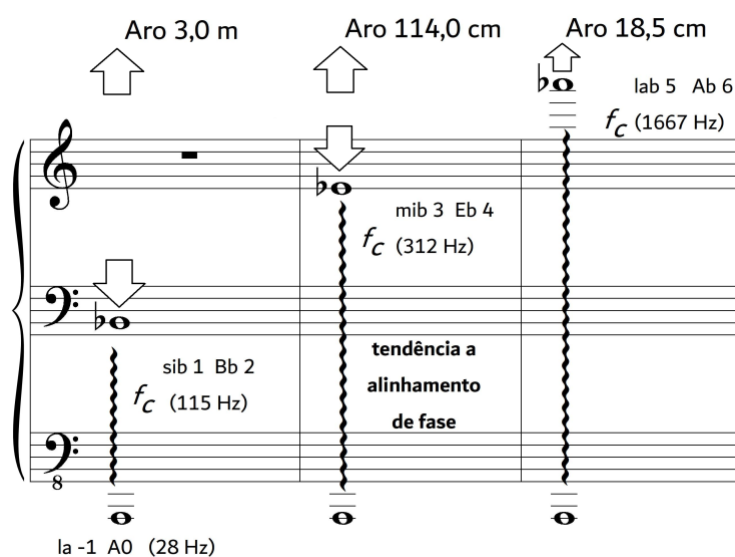
Eram seis pares, com 60° cada um, totalizando 360°. Tínhamos seis imagens sonoras adjacentes, uma terminando onde outra começava (Mannis, 2020, p. 11). Logo na escuta das primeiras gravações decidimos substituir os omnidirecionais por cardioides (Mannis, p. 13-18). Percebemos que, como os microfones eram omnidirecionais, as frentes de onda eram captadas duas vezes: uma quando entravam no círculo delineado pelos microfones e a outra quando saíam pelo lado oposto. Quando os eventos sonoros começavam com ataques claros e definidos, como os sons percussivos, a segunda captação (depois da onda sonora atravessar o círculo) era mascarada na reprodução sonora pelo efeito de precedência (Haas, 1972). Nessa situação, a localização era precisa com os omnidirecionais. Porém, quando os sons eram tônicos e contínuos – uma flauta transversal¹¹ tocando livremente células melódicas nas duas primeiras oitavas com algumas notas longas, por exemplo –, juntamente com o atraso de incidência da onda sonora entre dois microfones ocorriam períodos estáveis e regulares da onda sustentada, e isso era apreendido pela escuta não como atraso, mas como variação de fase, e a localização da fonte sonora se tornava imprecisa. Usando microfones cardioides, que atenuam os sons que incidem por trás, esse efeito foi mitigado. A disposição dos microfones e dos monitores de escuta, todos circunscritos numa circunferência, formava um hexágono regular, com a medida do lado igual ao raio, como na Figura 3.

¹⁰ Coordenado pelo professor Luiz Wagner Biscainho.

¹¹ Foram realizados testes com diversos tipos de instrumentos, tocados livremente enquanto se caminhava ao redor do arranjo de microfones. Nessa etapa específica do trabalho, os microfones ainda eram omnidirecionais. Como os diagramas polares de irradiação da flauta transversal, nas duas primeiras oitavas, são bastante similares (Meyer, 1980, p. 167), entre uma nota tocada e outra não houve mudança suficiente para que a irradiação da flauta pudesse causar a referida imprecisão de escuta.

Comparações entre tomadas de som com microfones circunscritos em diâmetros variáveis (3,0 m; 1,14 m; e 18,5 cm) revelaram que, proporcionalmente ao diâmetro, aumentava a sensação de espacialidade e a doçura dos ataques (Mannis, p. 63 e seg.). Concluímos que isso se devia ao aumento da banda de frequência, em que ocorre descorrelação de fase e atraso entre os canais (Mannis, p. 72). Ou seja, à medida que aumenta o diâmetro, aumenta o âmbito da banda de frequência de descorrelação e a sensação de espacialidade. Quando o diâmetro é maior, os ataques, mesmo quando fortes, tendem a ser percebidos com mais suavidade. Na Figura 13, observa-se como o âmbito de alturas, no qual há sensação de mais especialidade e leveza de ataques, aumenta proporcionalmente ao diâmetro do arranjo circular de microfones.

FIGURA 13 – Na parte superior, indicado pela setas, o âmbito de alturas em que ocorre a descorrelação de fase e atraso; a linha ondulada vertical marca o âmbito no qual há alinhamento de fase entre todas as componentes parciais de cada onda sonora.

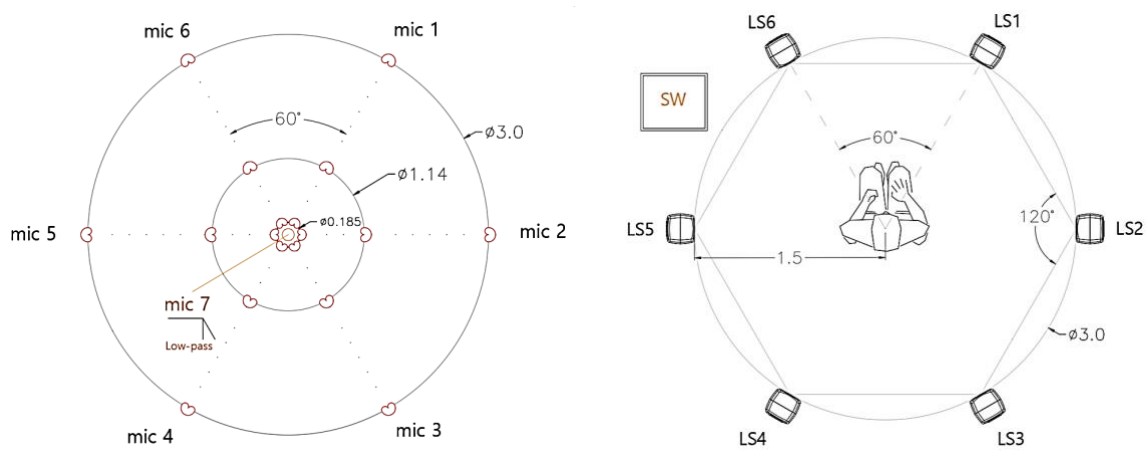


Fonte: Mannis (2020, p. 72).

Observando gravações simultâneas com microfones omnidirecionais e cardioides montados em paralelo, foi possível perceber como a mordida de alguns sons percussivos graves desaparecia devido à fraca resposta, em frequência, dos microfones cardioides nas baixas frequências. Para compensar essa perda, foi introduzido um microfone omnidirecional no centro, que, para a reprodução, é tratado com um filtro *low pass*.

Dos experimentos realizados, resultou um sistema de captação sonora com seis microfones cardioides circunscritos em diâmetro variável entre 18,5 cm e 3,0 m, tendo, de maneira facultativa, um omnidirecional no centro, e um dispositivo de reprodução com seis caixas acústicas dispostas circularmente, com diâmetro de pelo menos 3,0 m e um *subwoofer* (Figura 14; Figura 15). É possível acessar, no Google Drive, um vídeo que apresenta o som imersivo e o sistema de gravação e escuta resultante desta pesquisa¹².

FIGURA 14 – À esquerda, configuração de dispositivo de captação sonora: seis cardioides circunscritos (diâmetro variando entre 18,5 cm e 3,0 m) e um omnidirecional no centro (facultativo); à direita, configuração de dispositivo de escuta: seis monitores circunscritos (diâmetro de pelo menos 3,0 m) e um *subwoofer*.



¹² Ver https://drive.google.com/file/d/1HKmdeLEJ3eZuYOVuKN9bnkPLZ1jOk8k0/view?usp=drive_link.

FIGURA 15 – Duas das versões do dispositivo desenvolvido para captação sonora em campo externo: em cima, ESMA 6.0 com diâmetro de 1,14 m próximo à cabeceira da pista do Aeroporto Santos Dumont, no Rio de Janeiro; embaixo, ESMA 6.1 com diâmetro de 3,0 m no campo do Clube de Aeromodelismo do Rio de Janeiro, na Universidade Federal do Rio de Janeiro (UFRJ), campus da Ilha do Fundão.

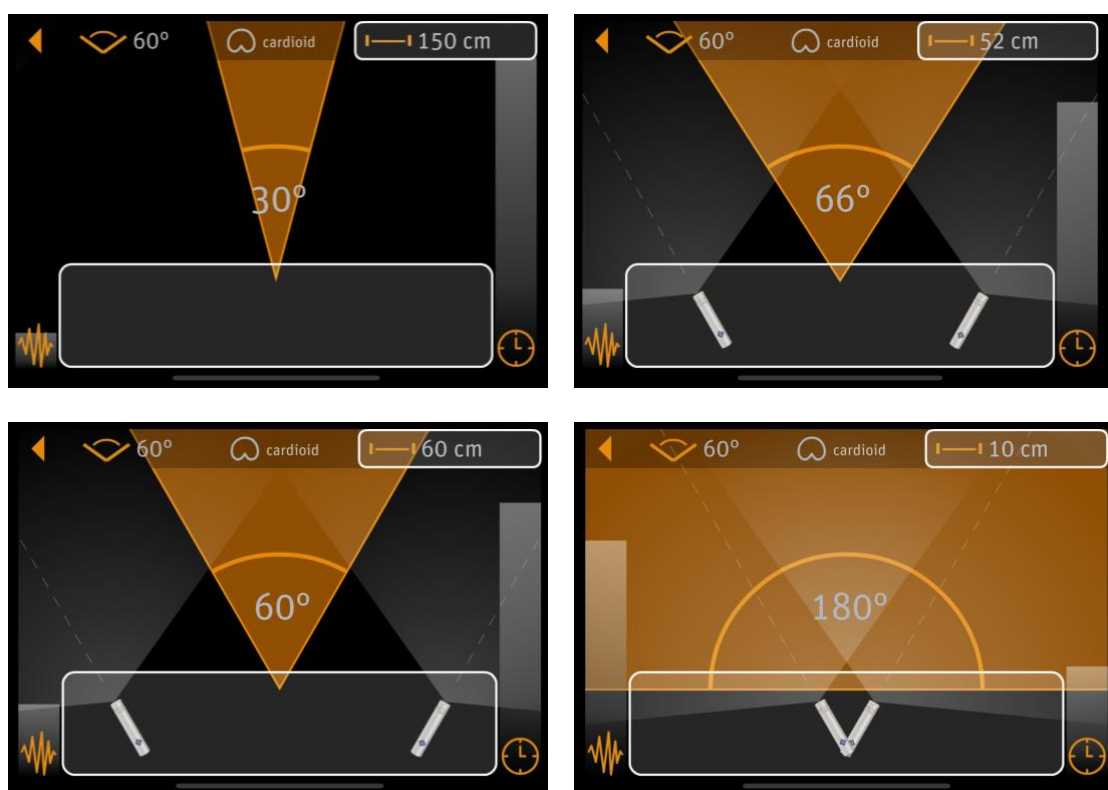


Fonte: J. A. Mannis (2020).

2. Confluência com os trabalhos de Williams

Considerando os gráficos usados para controle da largura da imagem sonora em razão da distância entre os microfones, percebemos que com aro de 3,0 m surgem *gaps* na continuidade da imagem e com aro de 18,5 cm as imagens adjacentes, correspondendo a cada lado do hexágono, se encavalam. A Figura 16 apresenta a largura angular de imagem em função da distância de um par de microfones cardioides entre si, formando um ângulo de 60° (SRA), conforme o simulador Recording Angle Calculator (Neumann, 2018).

FIGURA 16 – Simulação, pelo Stereophonic Recording Angle (SRA), do ângulo de abertura da imagem sonora em função da distância entre os microfones cardioides angulados em 60° pelo Recording Angle Calculator, da Neumann. Em cima, à esquerda, distância de 1,5 m (aro de 3,0 m), imagem sonora com 30° de largura, e, à direita, distância de 52 cm (aro de 1,14 m), imagem sonora com 66° de largura. Embaixo, à esquerda, 60 cm (aro de 1,2 m), imagem sonora com 60° de largura, e, à direita, 10 cm (aro de 20 cm), imagem sonora com 180° de largura. Observe-se que em cima, à esquerda, não aparecem os microfones separados de 150 cm porque eles estão além do campo de plotagem do aplicativo.



A configuração obtida permitiu, ao ouvinte, desde que estivesse dentro do *sweet spot*¹³, uma escuta imersiva durante a qual era possível ele se mover e mudar a direção de orientação da cabeça sem que a imagem sonora deixasse de ser estável e coerente.

De acordo com Williams (1991), as imagens sonoras obtidas com seis microfones arranjos com diâmetro de aproximadamente 1,06 m reproduzidas em seis caixas acústicas dispostas em hexágono regular, tendem a conservar as imagens sonoras originais com menor distorção geométrica e promovem, no *sweet spot*, uma sensação, à escuta, próxima à que se teria *in loco* no momento da gravação da incidência das frentes de onda.

Nas audições em laboratório, tínhamos a impressão de estar no local e no momento da gravação, e podíamos nos mover dentro na área central de escuta sem desestabilizar a percepção espacial da imagem sonora. Podíamos ouvir várias vezes a mesma tomada de som, porém, a cada escuta, nos orientando para uma direção diferente, como que vivendo repetidamente o mesmo instante, mas a cada vez com a atenção para alguma coisa diferente. Poderíamos chamar essas gravações de “fonografias” imersivas, emprestando o termo do designer sonoro francês Louis Dandrel¹⁴.

Reconhecemos a anterioridade da configuração, caracterizada pelo princípio de simetria pela segmentação da circunferência em partes iguais, de Michael Williams, e imaginamos as aplicações dessa técnica de captação e reprodução sonora em música com espacialização, audiovisual imersivo e instalação sonora, o que possibilita criar espaços de escuta em museus, fazer monitoramento ambiental, aprimorar pesquisas em bioacústica, proceder a registro e documentação de ambientes sonoros, e outras mais.

3. Sistemas de captação sonora similares e denominação de configuração

Encontramos referências que mencionam configurações de captação sonora similares em Hyunkook Lee, do Laboratório de Psicoacústica Aplicada da Universidade de Huddersfield (Reino

¹³ Região focal dentro da qual a imagem sonora é escutada de maneira equilibrada entre todos os canais. Quanto maior o diâmetro, maior a área central do *sweet spot*.

¹⁴ Em conversa pessoal com J. A. Mannis no início da década de 1990.

Unido), e em C. Junu Jahana, M. S. Sinith e P. P. Lalu, da Faculdade Governamental de Engenharia Thrissur (Índia). Lee (2019) adotou o acrônimo ESMA, da denominação *equal segment microphone array*, de Williams (2003, p. 6-7), o que nos levou até este autor. Jahana, Sinith e Lalu (2021) chamaram de *circular microphone array* (CMA). A denominação ESMA, que compreende a simetria de distribuição dos pontos de captação, parece-nos compacta e suficiente para um arranjo de microfones igualmente espaçados sobre uma circunferência.

O hexágono de tomada de som desenvolvido nesta pesquisa (Figura 3) corresponde, portanto, à configuração originalmente proposta por Michael Williams em 1991 e a um ESMA 6, ou um ESMA 6.1 (quando há um microfone omnidirecional no centro).

Como desdobramentos tecnológicos desta pesquisa, na medida em que oportunidades se apresentaram, desenvolvemos novos recursos, e em determinado momento imaginamos poder desenvolver um sistema que permitisse acrescentar novos sons a uma gravação com ESMA 6, de maneira que se integrasse naturalmente na imagem sonora registrada.

4. Espacialização sonora multicanal em ambiente imersivo

O problema que tínhamos para resolver, era acrescentar a uma gravação multicanal novos sons de forma coerente e integrada ao espaço e ao ambiente sonoro original. Tal recurso teria como aplicações: a simulação e auralização da presença de uma fonte sonora em um determinado ambiente, de maneira a avaliar o eventual impacto da nova fonte no ambiente, ou vice-versa; a criação artística de intervenção sobre imagens sonoras gravadas; modelar artificialmente uma imagem sonora imersiva, compondo cada um dos seus elementos a partir de fontes sonoras artificiais ou arquivos pré-gravados; auralização histórica de ambientes sonoros (Katz; Poirier-Quinot; Lyzwa, 2021); *design* de paisagens sonoras (Soundscape design, 2024); instalações sonoras imersivas (Nicolas Party, 2018).

A ferramenta que tem esse intuito simula a gravação de uma fonte sonora virtual por um sistema multicanal de microfones e está sendo desenvolvida por Igor A. Aguilar em pesquisa de

mestrado em Música¹⁵.

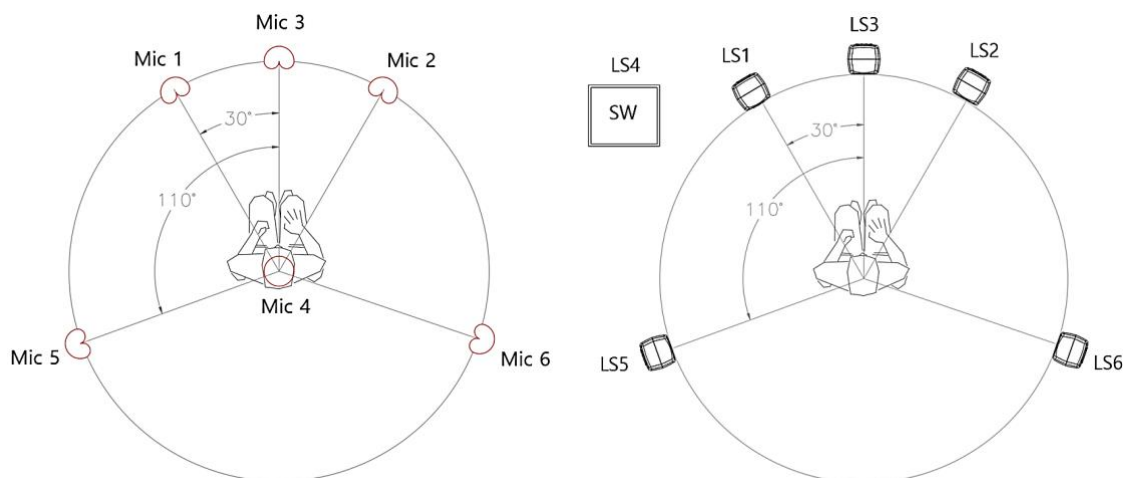
No ambiente simulado, as ondas emitidas se propagam até os pontos de captação (microfones virtuais), chegando em cada um deles em um determinado instante, com perda de energia proporcional à distância percorrida e à direção de incidência. Na captação estéreo em dois canais, as diferenças de tempo e intensidade se denominam ITD e ILD, mas, aqui, em contexto multicanal, passam a ser ICTD e ICLD.

O material sonoro a ser irradiado pela fonte virtual é um sinal sonoro, preferencialmente captado em condições anecoicas, adquirido por uma entrada de áudio ou pela leitura de um arquivo. Após ser virtualmente irradiado, cada canal do sistema de captação sonora virtual (atribuído a um microfone virtual) recebe o sinal da fonte virtual com a respectiva atenuação e atraso, correspondendo à trajetória da frente de onda, desde a fonte sonora até sua captação, simulando o percurso em campo aberto. O sistema virtual de captação – *surround* 5.1 (Figura 17), *surround* 7.1, ESMA 6, ESMA 8 etc. – deve ser configurado no *patch* ASARC (*ambient system and audio recording configurations*; ambiente de configuração de sistema, áudio e gravação)¹⁶ no início do processo (ver, no item 4.2, a Figura 18), bem como a temperatura do ar no ambiente virtual, determinante das condições de propagação da onda sonora.

¹⁵ No Programa de Pós-Graduação em Música da Unicamp, sob orientação de J. A. Mannis e suporte do LASom – Laboratório de Acústica e Artes Sonoras.

¹⁶ Os *patches* (arquivos de código de Pure Data) ASARC e VSSC (adiante citado) foram desenvolvidos por Igor A. Aguilar, na sua pesquisa de mestrado, sob orientação de J. A. Mannis.

FIGURA 17 – Configuração de dispositivo *surround* 5.1: à esquerda, captação sonora com cinco cardioides circunscritos e um omnidirecional no centro (facultativo) (Mic 1-6); à direita, reprodução sonora com cinco monitores circunscritos (LS1-6) e um *subwoofer* (SW).



A simulação e a auralização são processadas em ambientes de programação Pure Data (Puckette, 1996) e Plugdata(x) (Schoen, 2021) com os *patches* ASARC e VSSC (*virtual sound source control*; controle de fonte de som virtual), cada um deles simulando os fenômenos que ocorrem na irradiação e na captação multicanal virtual da fonte sonora. Foram acrescentados outros *patches*, oriundos da biblioteca de domínio público ELSE – EL Locus Solus’ Externals for Pure Data (Porres, 2024). Todos esses recursos reunidos constituem a biblioteca específica de *patches* dessa aplicação.

4.1 Arquitetura do sistema

Os ambientes de programação Pure Data e Plugdata dispõem de interfaces visuais, denominadas *generated user interface* (GUI), que proporcionam uma constante interação visual com o usuário. Através das interfaces, faz-se a configuração do sistema e do contexto simulado, bem como o controle de variação dos parâmetros envolvidos no processamento.

A biblioteca dessa aplicação compreende *patches* que atuam na definição de parâmetros do ambiente, na configuração do sistema de captação sonora multicanal, no posicionamento, na trajetória e na velocidade de rotação das fontes sonoras simuladas, no processamento de sinal e na integração do sistema em Plugdata com uma DAW externa.

Usamos o Pure Data aplicado à simulação de posição, movimentos, e, também, à emissão de som das fontes sonoras e à recepção das ondas pelos diferentes microfones virtuais. Para isso, analisamos, de maneira física e matemática, os fenômenos envolvidos, e os reconstruímos de forma integrada, interativa e coesa usando os objetos de processamento de sinal disponíveis no Pure Data. As dificuldades que surgiram foram superadas ao longo do tempo, porque encontramos alternativas cujas eficiências foram verificadas teórica e empiricamente. Esse processo evoluiu num ambiente coletivo, colaborativo, reflexivo, prático, experimental e flexível, de maneira que, em diversas ocasiões, após inúmeras simulações e análises de resultados, os *patches* foram adequados e aperfeiçoados.

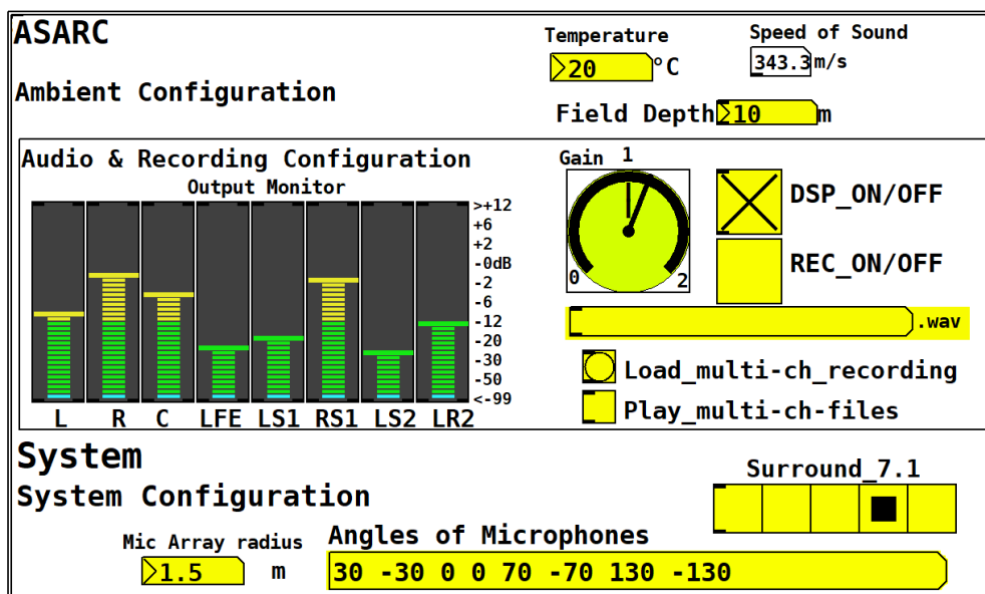
4.2 Interfaces de usuário e sistema desenvolvido

A Figura 18 apresenta a interface visual do *patch* ASARC, que compreende três tipos de controle e está organizada em três setores horizontais: um superior, um central e um inferior. Em cada setor, os itens com fundo amarelo representam elementos controláveis pelo usuário, e os de fundo branco são valores automaticamente gerados pelo sistema.

No setor superior procede-se à configuração do ambiente simulado, na qual são definidas a temperatura¹⁷ (*Temperature*) do ambiente virtual e a profundidade de campo (*Field Depth*: raio máximo), que estabelece a dimensão máxima da área simulada.

¹⁷ A velocidade do som (*speed of sound*) é imediatamente calculada em função da temperatura.

FIGURA 18 – Configuração da interface visual do *patch* ASARC em Pure Data: configuração de parâmetros para caracterizar o ambiente simulado (temperatura do ar e profundidade do campo espacial virtual); configuração do dispositivo virtual de captação sonora (5.1, 7.1, 6.0, 6.1 etc.); monitoramento do nível sonoro multicanal (*Output Monitor*); controle de ganho geral de saída (*Gain*); comando para renderização em arquivo sonoro multicanal após o processamento de espacialização (*REC*); acionamento da DSP; comandos para carregar (*Load*) e disparar (*Play*) gravações multicanais como base sonora de uma espacialização em processamento.



No setor central, onde há um conjunto de *VU meters* para monitoração do nível sonoro das saídas de cada canal processado (*Output Monitor*), procede-se à configuração do áudio e da gravação (*Audio & Recording Configuration*); à regulagem de ganho (*Gain*) de saída do processamento; ao acionamento da DSP¹⁸; ao comando para início de uma gravação multicanal do processamento (marcando o quadrado *trigger*); à definição do nome do arquivo (*.wav) a ser gravado; ao carregamento (*Load*) de arquivos de gravações multicanal; ao início de leitura (*Play*) desse arquivo, de maneira a poder trabalhar sobre uma base imersiva pré-gravada acrescentando-se outras fontes sonoras num espaço imersivo virtual coerente.

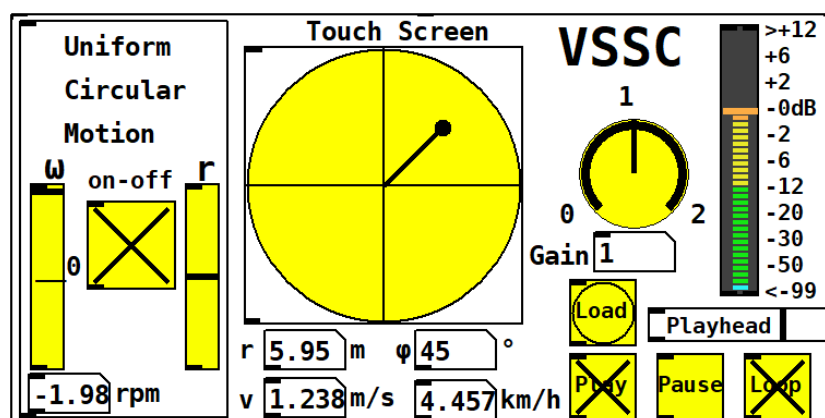
No setor inferior escolhe-se o sistema de configuração de áudio multicanal; o raio da circunferência delineada pelos microfones virtuais (*Mic Array Radius*); a quantidade e a disposição

¹⁸ Digital Signal Processor (DSP) é um microprocessador que processa sinais digitais, como áudio e vídeo, em tempo real ou *off-line*.

desses microfones conforme os formatos ESMA 6.1, *surround* 5.1, ESMA 8 e *surround* 7.1, tendo, na área cinza, a coordenada angular respectiva a cada microfone (*Angles of Microphones*).

A Figura 19 apresenta a interface visual do *patch* VSSC, desenvolvido para controle da fonte sonora virtual.

FIGURA 19 – Interface visual do *patch* VSSC em Pure Data para controle da fonte sonora virtual: ganho (*Gain*), posicionamento [r , ϕ], velocidade linear (v), frequência de rotação (ω), carregamento de arquivo sonoro (*Load*), disparo (*Play*) e acionamento de *Loop*.



Com o VSSC pode-se controlar a posição da fonte sonora no espaço virtual. Na Figura 19, observa-se, no centro, o *Touch Screen*¹⁹, através do qual é possível alterar e monitorar a posição da fonte sonora. Embaixo, observam-se as coordenadas (r , ϕ), que são dadas automaticamente, bem como a velocidade linear (v) da fonte sonora em metros por segundo e quilômetros por hora. Sua posição pode ser deslocada em tempo real arrastando-se, com o *mouse*, o ponto preto sobre o círculo amarelo do *Touch Screen*. Para aplicar uma rotação constante à fonte sonora, acionam-se os comandos do setor *Uniform Circular Motion* (UCM; Movimento Circular Uniforme), situado à esquerda da interface.

Esses acessos estão interligados a um *patch* em Pure Data que contém um oscilador de baixa frequência, normalmente referido como LFO (*low frequency oscillator*), que gera uma onda senoidal da qual a fase²⁰ corresponde à coordenada ϕ da interface *Touch Screen*. O UCM pode ser ligado e

¹⁹ A GUI (v. 4.1) utilizada aqui é o *circle*, da biblioteca de domínio público ELSE - EL Locus Solus' Externals for Pure Data (Porres, 2024).

²⁰ Note-se que a amplitude do LFO não é considerada, pois o raio do movimento é dado pelo parâmetro r gerado pelo *fader* do UCM.

desligado (*On/Off*) através do botão localizado entre os *faders*, que comandam, respectivamente, a frequência de rotação (w) e o raio da trajetória circular (r). É importante notar que o sentido anti-horário corresponde a valores positivos de w e o sentido horário, a valores negativos de w . Todas as funções em amarelo podem ser modificadas durante o funcionamento do *patch*. O ganho²¹ (*Gain*) pode ser alterado rotacionando o potenciômetro à direita do *Touch Screen*. De 0 a 1 há atenuação da intensidade da fonte sonora; em 1 a intensidade original é mantida; de 1 a 2 o sinal é amplificado (+6 dB). À direita do potenciômetro de ganho há um *VU meter* através do qual se pode monitorar o nível sonoro de saída do processamento em decibels. Nos botões amarelos abaixo do *VU meter* procede-se a operações de carregamento (*Load*), acionamento (*Play*), pausa (*Pause*) e repetição contínua (*Loop*) do arquivo de áudio de referência sobre o qual a simulação será feita. Sua reprodução pode ser monitorada através da posição do cursor do *Playhead*, que se move automaticamente da esquerda para a direita conforme progride a leitura do arquivo. Pode-se dizer que o VSSC é equivalente a um algoritmo de automação de movimentos circulares da fonte sonora virtual, podendo, ainda, ter seus parâmetros modificados a qualquer momento por um controlador MIDI interligado ao sistema.

4.3 Controle de dados

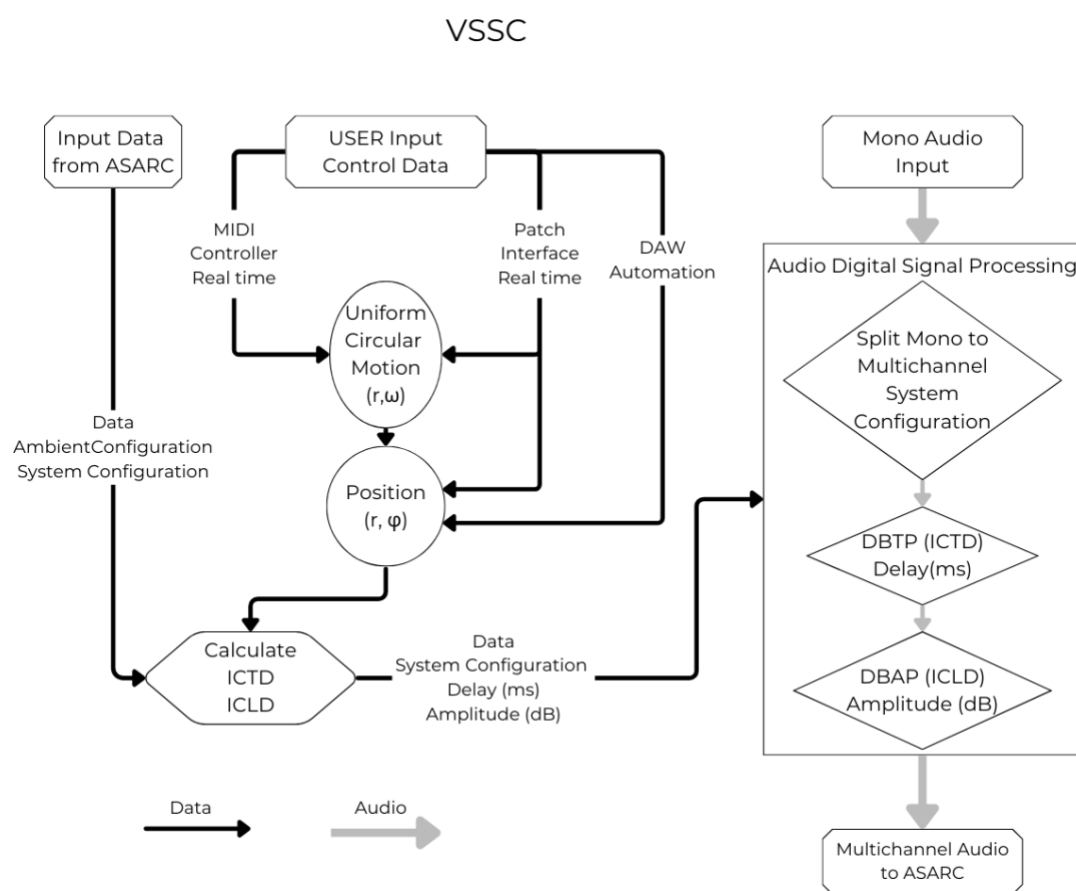
A primeira entrada do *patch* VSSC (Figura 20) recebe dados de variáveis globais do *patch* ASARC. Esses dados são *Ambient Configuration*, referente ao raio máximo e velocidade do som, e *System Configuration*, referente ao posicionamento do arranjo de microfones virtuais e formato de áudio multicanal. A segunda entrada recebe dados de posicionamento da fonte sonora virtual através do módulo UCM que pode ser comandado por um controlador MIDI; bem como pela ação direta do *mouse* no *Touch Screen*; e, ainda, a partir da reprodução de uma automação²² gravada num *track* da DAW. A terceira e última entrada é reservada ao sinal de áudio a ser irradiado pela fonte sonora virtual

²¹ A GUI utilizada aqui é o *knob*, da biblioteca ELSE - EL Locus Solus' Externals for Pure Data (Porres, 2024).

²² A automação de controle de parâmetros diversos é um recurso próprio a diversas DAWs disponíveis na atualidade. No processo de gravação e reprodução de parâmetros de espacialização são usados unicamente os recursos oferecidos por cada aplicativo.

especializada. Esse sinal pode ser proveniente da leitura de um arquivo previamente carregado (*Load*) ou de uma reprodução da DAW em tempo real. O VSSC contém saída de áudio multicanal enviada diretamente para o ASARC.

FIGURA 20 – Diagrama do fluxo de dados e de sinal do *patch* VSSC. As setas de cor preta representam dados e as de cor cinza representam sinal, entrada de dados de controle (configurações de ambiente e de sistema) vindos do *patch* ASARC (*Input Data from ASARC*), entrada do usuário de controle de dados (*User Input Control Data*), entrada de áudio mono (*Mono Audio Input*), movimento circular uniforme (*Uniform Circular Motion*), posição (*Position*), cálculo de atraso (ICTD) e amplitude (ICLD), processamento de sinal digital de áudio (*Audio Digital Signal Processing*), ramificação de sinal mono para multicanal no formato estabelecido no sistema de configuração (*Split Mono to Multichannel System Configuration*).



4.4 Processamento de sinal

Para simular a captação multicanal virtual da fonte sonora, processamos digitalmente o sinal que ela deve emitir, tomando-o a partir de um arquivo de áudio mono previamente carregado ou do

*direct out*²³ de um *track* da DAW. Cada microfone simulado recebe esse sinal após percorrer a distância entre ele e a fonte. Portanto, a cada microfone se atribui um atraso em função do percurso da onda sonora virtual que corresponde ao valor do ICTD. Da mesma maneira, o sinal de cada microfone tem uma atenuação de intensidade que é proporcional à distância percorrida, que corresponde ao valor do ICLD. Esses processos ocorrem com base em objetos dos tipos *distance-based amplitude panning* (DBAP) e *distance-based time panning* (DBTP). O fluxo de dados e de sinal do *patch* VSSC pode ser observado na Figura 20.

4.5 Sistema geral de *patches*

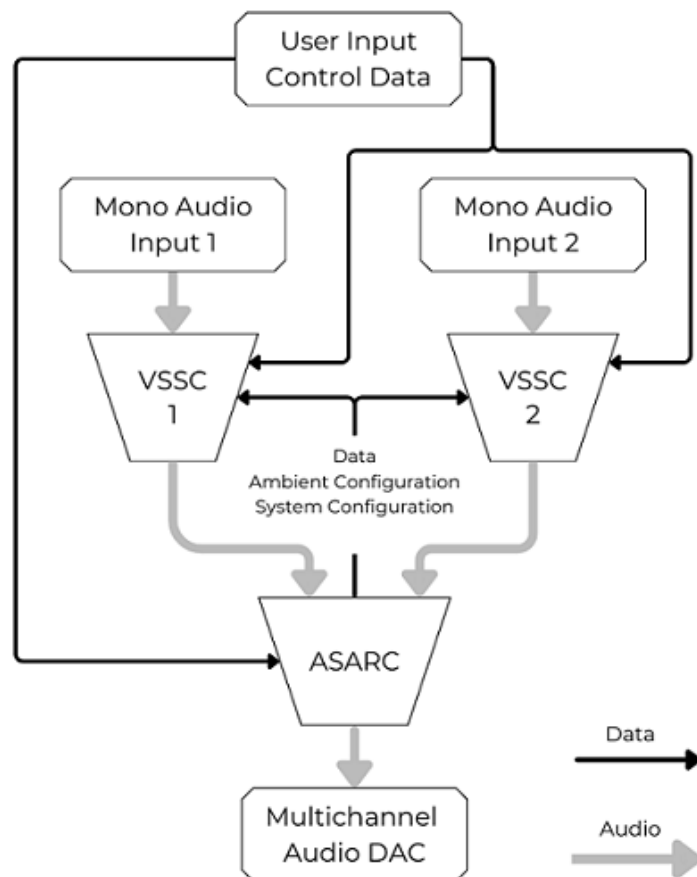
Apresentamos aqui uma visão do sistema geral de *patches* que contém todos os elementos necessários para se realizar uma espacialização em pós-produção ou ao vivo. Há vários *patches* VSSC (conforme a quantidade de fontes sonoras a serem espacializadas) e um *patch* ASARC. No fluxograma do sistema geral de *patches* (Figura 21) são mostrados os processos de fluxo de dados e de sinal de áudio, a visualização das suas respectivas funções e como os *patches* interagem dinamicamente entre si. Nesse sistema, as variáveis globais, ou seja, aquelas que afetam o ambiente e as condições de gravação, estão contidas no ASARC. Já as variáveis locais, como a posição e o ganho de cada fonte sonora, estão presentes em seus respectivos *patches* VSSC. Isso possibilita a espacialização simultânea de várias fontes sonoras de maneira coesa.

Os *patches* em Pure Data também podem ser processados em ambiente Plugdata, o que permite seu uso em uma DAW como VST²⁴. Dessa forma, com o recurso de automação programada nas *tracks* da DAW é possível gravar e sincronizar os parâmetros de posicionamento das fontes sonoras com elementos de outras pistas. Note-se que, nesse caso específico, a sincronização é regida pela DAW (*master*). Os *patches* em Plugdata seguem passivamente (*slave*) a evolução dos parâmetros, cuja leitura evolui conforme o *timecode* da DAW.

²³ Algumas DAWs podem ter outra denominação para uma saída direta do sinal de um *track*.

²⁴ VST se refere a um recurso virtual que atua junto a uma DAW.

FIGURA 21 – Fluxograma do sistema geral de *patches* diferenciando fluxo de dados e de sinal de áudio.



5. Criação artística com espacializador imersivo

Os recursos de espacialização desenvolvidos foram empregados na realização de obras musicais de Igor A. Aguilar e Chang Rodrigues. As aplicações tiveram como objetivos: a) testar o desempenho do sistema; b) analisar como a ferramenta atende e responde a necessidades de composição e performance musicais; c) conhecer como os recursos facilitados contribuem para o processo criativo.

5.1 Nossa natureza (2023), de Igor A. Aguilar

Peça gravada em mídia eletrônica em *surround* 7.1. Dur.: 18'36".

Contexto: sobre uma paisagem sonora imersiva de J. A. Mannis e Janete El Haouli (2019),

gravada com ESMA 6.0 (hexágono) no Parque Ecológico do Gunma (Santa Bárbara, PA), contendo múltiplas vocalizações de pássaros capitão-do-mato (*Lipaugus vociferans*), foram adicionados materiais sintetizados em mono em laboratório e gravações em estéreo (LR) do CD *Cantos das aves brasileiras*, de Jacques Viellard (1995). Os seis canais da gravação de campo foram distribuídos nos canais FL, FR, SL, SR, RL, RR do *surround* 7.1. O material sintetizado e excertos de cantos de aves brasileiras gravados por Viellard foram espacializados diretamente na DAW utilizando o *plugin* em formato VST, que foi configurado para receber o sinal a ser virtualmente irradiado e simular sua captação com microfones virtuais em *surround* 7.1. A cada canto de ave do CD de Viellard acrescentado à gravação original de Mannis e El Haouli foi atribuída uma posição fixa no *surround* 7.1, e os materiais sintetizados foram movimentados em trajetórias circulares ou transversais, programadas diretamente na DAW, comandando os parâmetros do *plugin* em configuração *surround* 7.1.

Representação e recursos criativos: *Nossa natureza* retrata a relação entre o ser humano e a natureza, sobretudo as mudanças que o meio ambiente vem sofrendo devido à ação antrópica. A paisagem sonora de Gunma representa toda a força e exuberância da natureza na floresta Amazônica. Para ilustrar a reação do sistema ecológico, pensamos em uma migração de aves forçada, como se aves de outras regiões estivessem procurando refúgio no parque.

Para isso, usamos as gravações dos pássaros maú (*Perissocephalus tricolor*), espécie amazônica proveniente do norte do rio Amazonas, saci (*Tapera naevia*), que habita áreas abertas de Cerrado e da Mata Atlântica, bem-te-vi (*Pitangus sulphuratus*), que ocupa habitat urbano, jardins e bordas de floresta, e araponga (*Procnia nudicollis*), da Mata Atlântica, simulando sua presença em Gunma. Ironicamente, fizemos os sons de chuva e o coaxar de rãs artificialmente, um efeito “antropogênico”. No ápice do colapso ecológico, os pequenos pássaros se perdem em um pesadelo sonoro. Depois, há uma transição: do pesadelo à volta à paisagem original.

5.2 Ciclos de renascimentos (2021), de Chang Rodrigues

Peça eletrônica com sintetizadores modulares tocada no Projeto Ruínas²⁵ e gravada em três canais. Dur.: 6'48".

Contexto: a peça foi gravada em um cenário desolado, na fazenda Fortaleza de Sant'Anna, em Goianá (MG), onde o abandono das ruínas (que datam do século XIX) contrasta com a presença de uma planta solitária no meio do espaço – uma manifestação de vida e esperança.

Cada um dos canais gravados foi individualmente espacializado em *surround* 5.1 em laboratório. Movimentos circulares e pendulares foram obtidos através do VSSC, estando o ASARC configurado em *surround* 5.1. A espacialização foi feita individualmente a partir de cada um dos canais originalmente gravados enquanto se escutava o movimento sonoro resultante no interior do arranjo circular de caixas acústicas. O resultado sonoro foi simultaneamente gravado na DAW em formato 5.1. Finalmente, os três canais espacializados em 5.1 resultaram em dezoito pistas e foram mixados, na DAW, em uma *master* final em *surround* 5.1.

Representação e recursos criativos: musicalmente, a peça representa a contínua transformação e renovação nos ciclos naturais da vida. A velocidade dos movimentos durante a performance foi controlada manualmente em laboratório, em função da escuta dos elementos sucessivos que entram e saem da cena sonora, podendo a performance realçar um caráter ou um gesto próprio a cada um deles.

²⁵ O Projeto Ruínas é uma proposta artístico-musical integrado ao projeto de *live electronics* denominado Chang Rodrigues (mesmo nome da artista). O objetivo é expressar sentimentos e emoções durante vivências realizadas em ruínas, através de sonoridades produzidas por sintetizadores modulares. Inspirado pelas associações cognitivas e emocionais que as ruínas despertam, o projeto explora essas sensações e nuances como meio de criar uma experiência sonora imersiva. Um curto documentário sobre o Projeto Ruínas está disponível no site da *Fact Magazine* (<https://www.factmag.com/2021/06/01/chang-rodrigues-documentary/>) e no canal dessa revista no YouTube (<https://www.youtube.com/watch?v=fHuhKLB0I4Q>). No canal da artista no YouTube é possível ver a performance *Chang Rodrigues live act Ruínas* (<https://www.youtube.com/watch?v=AWzGHSKHTx4>) e as composições eletroacústicas do Projeto Chang Rodrigues: *Ciclos de renascimentos* (<https://www.youtube.com/watch?v=2NGyMq8Fclg>), *Diálogo com a mãe Terra* (<https://www.youtube.com/watch?v=EGhsZTuBLgk>) e *Velocidade do tempo* (<https://www.youtube.com/watch?v=6TJfjsskZJo>) – também disponíveis em formato binaural. A nova versão dessas obras, em formato 5.1, desenvolvida em parceria com Igor A. Aguilar, pode ser conferida em performances presenciais com áudio imersivo.

5.3 *Diálogo com a mãe Terra* (2021), de Chang Rodrigues

Peça eletrônica gravada em cinco canais durante performance no Projeto Ruínas, em novembro de 2021, utilizando sintetizadores e sons pré-gravados de cigarras na fazenda Fortaleza de Sant'Anna, em Goianá (MG). Dur.: 10'10".

Contexto: a peça é uma composição audiovisual imersiva gravada em um local historicamente significativo, onde a catagem²⁶ de café era realizada. A cena escolhida, marcada pelo renascimento da natureza em um ambiente que mescla vegetação exuberante de Mata Atlântica e ruínas com tijolos expostos, estabelece um pano de fundo poderoso para a obra. O entorno do local abandonado foi tomado pela vegetação e se destaca pela beleza natural e pelo simbolismo de renovação e regeneração, que permeia a narrativa.

Representação e recursos criativos: a espacialização de cada um dos canais gravados na performance foi cuidadosamente desenhada, atribuindo-se, a alguns elementos, posições fixas no espaço *surround*, e a outros, lentos movimentos circulares e pendulares. A composição estabelece um diálogo simbólico entre os sons dos sintetizadores e os elementos naturais, estes representados pelo canto das cigarras.

5.4 *Intuição* (2021), de Chang Rodrigues

Peça eletrônica gravada. Dur.: 3'38".

Contexto: trata-se da fusão de elementos oriundos de três composições eletrônicas que integram o Projeto Ruínas: *Ciclos de renascimentos*, *Diálogo com a mãe Terra* e *Velocidade do tempo*, esta última, também de 2021 e gravada em três canais durante performance. De cada uma dessas peças gravadas em performance foi selecionado apenas um dos canais para ser espacializado em uma nova mixagem em *surround* 5.1.

²⁶ Catagem (ou catação) do café é o processo de separar manualmente os grãos de café defeituosos e outros detritos, e descartá-los.

Representação e recursos criativos: em *Intuição* a espacialização foi realizada por meio de movimentos circulares, tanto lentos quanto rápidos, com controle manual dos movimentos, como nas peças que lhe deram origem. Seu objetivo central foi transformar essas partes em uma única composição sonora que promovesse relaxamento, meditação e a ampliação dos canais perceptivos.

6. Aportes do recurso desenvolvido

A composição de Igor A. Aguilar foi realizada concomitantemente ao desenvolvimento dos recursos VSSC e ASARC. Muitas das necessidades composicionais estimularam a imaginação das facilidades encontradas em ambos, como poder acrescentar sons suplementares a paisagens sonoras gravadas de maneira imersiva, para que soassem com a mesma naturalidade, ou seja, que fossem acusticamente coesas e se integrassem a uma mesma situação sonora. A facilidade em atribuir movimentos circulares regulares à fonte sonora foi aplicada com variação de velocidade angular, como um recurso expressivo gestual presente no pesadelo em forma de vórtice dramático no final de *Nossa natureza*. Como os produtos VSSC e ASARC foram desenvolvidos concomitantemente às necessidades dessa peça, parece redundante analisar como as facilidades desenvolvidas contribuíram para solucionar os problemas que as suscitaram. Portanto, daremos ênfase, aqui, às facilidades que mais contribuíram para o processo criativo das peças de Chang Rodrigues.

Os recursos de espacialização aplicados nas peças de Chang Rodrigues permitiram enfatizar características de cada uma das suas personagens sonoras, dando-lhes vida no ambiente de escuta imersiva e construindo narrativas em que elas pudessem evoluir e interagir entre si. O desenho sonoro se inscreveu na dinâmica da narrativa com movimentos circulares e pendulares, em velocidades rápidas e lentas, por vezes em oposição a posições fixas ou de mobilidade reduzida de cada uma das personagens encenadas no espaço sonoro. A espacialização direta, possibilitada pelos *patches* em Pure Data, permitiu desenhar a trajetória ou controlar a velocidade de cada personagem em tempo real enquanto eram escutadas.

Foi possível observar, nos experimentos de maneira geral, que movimentos diferenciados simultâneos se realçam na escuta, pela constante percepção das posições relativas das fontes sonoras.

Isso também ocorreu quando havia um material sonoro estático e outro em movimento: a sensação de deslocamento espacial na escuta se amplificava.

Em *Ciclos de renascimentos*, os movimentos lentos aplicados à divindade Iansã (senhora dos ventos) conferiram um forte caráter ritualístico à sua presença na cena sonora. Já a personagem Ciclo, que representa ritmos periódicos naturais, modulava-se em um amplo âmbito de velocidades, realçando, nessa agilidade, a dinâmica de sua essência.

Intuição começa com a personagem Oxum em movimento circular no sentido horário. A fluidez e a harmonia do movimento caracterizam a sua força purificadora. A personagem Tempo, com timbre percussivo, também tem trajetória circular no sentido horário. E a personagem Drone oscila entre as posições LS e RS, criando uma base sólida, transmitindo a sensação de estabilidade e continuidade. Na peça, um dos aspectos mais interessantes da espacialização foi o ajuste dos movimentos circulares e pendulares das personagens, testando-se diferentes velocidades e definindo-se, em tempo real, o raio da trajetória simulada. O modo como o som “anda” no espaço e como o ocupa denota os estados imaginários e emocionais das personagens sonoras. Esses movimentos podem sugerir estados de relaxamento, tensão, serenidade ou mesmo transições sutis entre diferentes emoções e atmosferas.

Em *Diálogo com a mãe Terra* a cuidadosa espacialização reforçou as características de cada personagem sonora, dando ênfase ao diálogo simbólico entre os sintetizadores e os elementos naturais, como o canto das cigarras – este, presente em toda a peça, transportando o ouvinte para o coração da natureza. Ao atribuir posições fixas a certos elementos e aplicar movimentos lentos, circulares e pendulares a outros, foi possível criar a sensação de coabitação e interação entre as personagens sonoras. A personagem Coração surge do lado esquerdo do *sweet spot* e aí permanece até se transformar, progressivamente, na personagem Terra, caracterizada por sonoridades graves que se deslocam lentamente no espaço frontal, simbolizando o crescimento e a unidade com o ambiente natural. Ao se mover no espaço, por vezes os sons se ocultam uns atrás dos outros, ressurgindo em seguida, resultando em texturas sonoras complexas e vivas.

A ferramenta de espacialização fomentou a imaginação sonora na composição dessas peças de Chang Rodrigues, permitindo que cada personagem, inicialmente constituída de altura, intensidade,

duração e timbre, fosse “vestida” de espaço, tornando a escuta de cada peça uma experiência auditiva rica e multidimensional.

7. Considerações finais

Apresentamos o nosso percurso na busca por um som imersivo alternativo aos recursos atualmente em uso. A solução para a captura e reprodução sonora à qual chegamos entre 2016 e 2020 já havia sido formulada, e publicada em 1991, por Michael Williams, sendo uma técnica que preserva as diferenças de intensidade e tempo de chegada da frente de onda entre os microfones do arranjo, reproduzindo-as, posteriormente, na escuta multicanal utilizando o mesmo arranjo para as caixas acústicas.

Optamos pela denominação da configuração como proposto por Lee (2019), que a chama de ESMA (*equal segment microphone array*), uma vez que compreende, de maneira coesa e suficiente, os recursos imaginados originalmente por Williams, evitando, assim, a proliferação infértil de denominações. Williams definiu com precisão a melhor distância de separação entre microfones cardioides num hexágono, orientados de modo a formar um ângulo de 60° (SRA), estabelecendo o raio de 53 cm da circunferência do arranjo circular de segmentos onde se situam os microfones, o que conserva, com menor distorção geométrica, a imagem sonora apreendida.

Por outro lado, quando comparamos os resultados obtidos de captações sonoras variando o raio dos arranjos de microfones, destacaram-se as qualidades resultantes e percebemos a mudança da largura da banda de frequência na qual ocorre variação de fase e atraso entre os microfones, correlacionada à dimensão do raio de cada arranjo.

Os resultados práticos obtidos da aplicação da configuração de captação e reprodução sonora ESMA soaram, à nossa escuta, de forma tão impactante que nos dedicamos a simulá-la virtualmente, com os objetivos de intervir posteriormente, de maneira coerente, sobre as gravações realizadas e de ter um recurso rico de espacialização aplicável em produção multicanal, permitindo um resultado sonoro com relevo e profundidade naturais.

Chegamos a um produto efetivamente operacional e que demanda pouco volume de processamento, o que tornou possível sua utilização como VST em uma DAW sem sobrecarregar a DSP. Sua utilização em criação musical revelou-se bastante significativa tanto para a concepção artística preliminarmente a uma ação quanto para explorá-la gestualmente durante a escuta do resultado em tempo real.

Constatamos, ainda, que gestos de espacialização gravados independentemente com a ajuda desses recursos, quando mixados, agregam ao todo uma organicidade natural. As aplicações decorrentes desse trabalho são múltiplas, e os próximos passos compreendem aperfeiçoamentos operacionais da ferramenta. Vislumbramos frentes de trabalho possíveis nos domínios de ampliar os recursos de captação, reprodução sonora e simulação em 3D na dimensão da elevação (zênite), bem como na localização precisa de fontes sonoras identificadas em sinais sonoros em contexto multicanal.

AGRADECIMENTO

Agradecemos a Michael Williams, pela força de sua imaginação acústica espacial.

REFERÊNCIAS

A BRIEF history of surround sound. **KEF**. Disponível em: <<https://us.kef.com/blogs/news/a-brief-history-of-surround-sound?srltid=AfmBOooYRpY4tM6UwMiRjzlm86HVnAzyk91h5FbP3CpBeJy03t4MbBVx>>. Acesso em: 18 ago. 2024.

AES PARIS 2016 presenter or author: Michael Williams. **140th International AES Convention**. 2016. Disponível em: <<https://www.aes.org/events/140/presenters/?ID=4376>>. Acesso em: 22 set. 2024.

ATAL, Bishnu S.; HILL, Murray; SCHROEDER, Manfred R. **Apparent sound source translator**. USA United States Patent Office, 1962. Disponível em: <<https://patents.google.com/patent/US3236949A/en>>. Acesso em: 25 ago. 2024.

BALLOU, Glen. M. **Handbook for sound engineers**: the new audio cyclopedia. 2nd ed. Carmel - Indiana (USA): SAMS, 1991.

BASKIND, Alexis et al. Binaural and transaural spatialization techniques in multichannel 5.1 production. In: TONMEISTERTAGUNG – VDT INTERNATIONAL CONVENTION, 27, 2012, Cologne. **Anais**. Cologne: Verband Deutscher Tonmeister e.V., 2012. p. 1-10. Disponível em: <<https://hal.science/hal-01247648/document>>. Acesso em: 25 ago. 2024.

BERKHOUT, Augustinus J. A holographic approach to acoustic control. **Journal of the Audio Engineering Society**, v. 36, n. 12, p. 977-995, 1988. Disponível em: <<https://secure.aes.org/forum/pubs/journal/?elib=5117>>. Acesso em: 5 nov. 2024.

BERWICK, Nathan; LEE, Hyunkook. Spatial unmasking effect on speech reception threshold in the median plane. **Applied Sciences**. v. 10, n. 15, article 5257, 2020. Disponível em: <https://www.researchgate.net/publication/343324613_Spatial_Unmasking_Effect_on_Speech_Reception_Threshold_in_the_Median_Plane/references#fullTextFileContent>. Acesso em: 28 out. 2024.

BLAUERT, Jens. **Spatial hearing: the psychophysics of human sound localization**. 2nd ed. Cambridge, MA: MIT Press, 1999.

BLUMLEIN, Alan Dower. Improvements in and relating to sound-transmission, sound-recording and sound-reproducing systems. **United Kingdom Intellectual Property Office**, 14 dec. 1931. Disponível em: <<https://patents.google.com/patent/GB394325A/en>>. Acesso em: 26 ago. 2024.

BRIXEN, Eddy B. **Stereo recording techniques and setups**. [2024]. Disponível em: <<https://www.dpamicrophones.com/mic-university/stereo-recording-techniques-and-setups>>. Acesso em: 24 ago. 2024.

CHANG, Ed. Gesang der jüinglinge. **Stockhausen: sounds in space**, n. 8. Jan. 2015. Disponível em: <<https://stockhausenspace.blogspot.com/2015/01/opus-8-gesang-der-junglinge.html>>. Acesso em: 26 ago. 2024.

CONDAMINES, Roland. **Stéréophonie: cours de relief sonore théorique et pratique**. Paris: Masson et CNET-ENST, 1978.

DICKREITER, Michael. **Tonmeister technology: recording environments, sound sources and microphone techniques**. New York: Gemmer Enterprises Inc., 1989.

DUMMY head. **Neumann.Berlin**. 2024. Disponível em: <<https://www.neumann.com/en-us/products/microphones/ku-100>>. Acesso em: 5 nov. 2024.

FARIA, Régis Rossi Alves. **Auralização em ambientes audiovisuais imersivos**. 2005. Tese (Doutorado em Engenharia) - Universidade de São Paulo, São Paulo, 2005.

FELLOT, Dominique. **Précis d'électro-acoustique**: prise de son et reproduction. Les Ulis (France): EDP Sciences, 2007.

HAAS, Helmut. The influence of a single echo on the audibility of speech. **Journal of the Audio Engineering Society**, v. 20, n. 2, p. 146-159, 1972. Disponível em: <<https://www.freestompboxes.org/viewtopic.php?t=31792>>. Acesso em: 16 set. 2024.

INTRODUCTION to Ambisonics – 360 degree audio. **PathPartner Technology**. Apr. 4, 2017. Disponível em: <<https://medium.com/@PathPartner/introduction-to-ambisonics-360-degree-audio-d92c5fe20a97>>. Acesso em: 5 nov. 2024.

ITU - INTERNATIONAL TELECOMMUNICATION UNION. **Recommendation ITU-R BS.775-4**: multichannel stereophonic sound system with and without accompanying picture. Geneva, 2022. Disponível em: <https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.775-4-202212-I!!PDF-E.pdf>. Acesso em: 23 ago. 2024.

JAHANA, C. Junu; SINITH, M. S.; LALU, P. P. Direction of arrival estimation using microphone array. In: INTERNATIONAL CONFERENCE ON MICROELECTRONICS, SIGNALS AND SYSTEMS - ICMSS, 4, 2021, Kollan (India). **Anais**. New York: IEEE, 2021. p. 1-6. Disponível em: <<https://ieeexplore.ieee.org/document/9673617>>. Acesso em: 10 nov. 2024.

KATZ, Brian F. G.; POIRIER-QUINOT, David; LYZWA, Jean-Marc. La Vierge 2020: reconstructing a virtual concert performance through historic auralisation of Notre-Dame Cathedral. In: INTERNATIONAL CONFERENCE 3D AUDIO (I3DA), 2021, Bologna. **Anais**. New York: IEEE, 2021. p. 1-10. Disponível em: <[https://hal.science/hal-03371681/file/I3DA_2021_paper_2%20\(2\).pdf](https://hal.science/hal-03371681/file/I3DA_2021_paper_2%20(2).pdf)>. Acesso em: 10 nov. 2024.

KUTTRUFF, Heinrich. **Acoustics**: an introduction. London: Taylor & Francis, 2007.

LEE, Hyunkook. Capturing 360° audio using an equal segment microphone array (ESMA). **Journal of the Audio Engineering Society**, v. 67, n. 1/2, p. 13-26, jan./feb. 2019. Disponível em: <<https://aes2.org/publications/elibrary-page/?id=19883>>. Acesso em: 5 nov. 2024.

LÓPEZ, Jose J. **WFS fundamentals**. 2007. Disponível em: <<https://personales.upv.es/jjlopez/WFS.html>>. Acesso em: 5 nov. 2024.

LYZWA, Jean-Marc. **Réflexion sur la prise de son et la post-production en 5.1**. Paris: Service audiovisuel-CNSMDP, 2006. Disponível em: <https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://www.conseilratoiredeparis.fr/sites/default/files/Recherche-Editions/CNSMDP_post-prod-multicanal5-1.pdf&ved=2ahUKEwiS0oiylrmIAxVARZUCHQNtFNcQFnoECBIQAQ&usq=AOvVaw2mxC>

[mrTOsx_-DTHDXRMhh>](#). Acesso em: 15 set. 2024.

LYZWA, Jean-Marc. **Prise de son et restitution multicanal en 5.1**. Problématique d'une œuvre spatialisée: Répons - Pierre Boulez. Paris: Conservatoire National Supérieur de Musique et de Danse de Paris, 2003. Disponível em:

<https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://www.conservatoiredeparis.fr/sites/default/files/Recherche-Editions/CNSMDP_repons-boulez.pdf&ved=2ahUKEwiAjbvNntSJAxUnq5UCHUnkLfEQFnoECBQQAQ&usg=AOvVaw2sUdddouqpDYblETMRCOBQ>. Acesso em: 15 set. 2024.

MANNIS, José Augusto. **Sistema de gravação e reprodução de áudio multicanal com imagem sonora horizontal em 360º aplicado a música, artes sonoras e bioacústica**. Brasília: CNPq, 2020. Disponível em:

<https://drive.google.com/file/d/1UpJJaROH1_HQZPh67dQaJAQP7ZOHNXiN/view?usp=drive_link>. Acesso em: 26 ago. 2024.

MANNIS, José Augusto; EL HAOULI, Janete. **Paisagem sonora do Parque Ecológico de Gunma**, gravação de som. Santa Bárbara, PA: out. 2019.

MASIERO, Bruno; FELS, Janina; VORLÄNDER, Michael. Review of the crosstalk cancellation filter technique. In: INTERNATIONAL CONFERENCE ON SPATIAL AUDIO - ICSA, 2011, Detmold (Germany). **Anais**. Köln: Verband Deutscher Tonmeister (VDT), 2011. p. 1-5.

Disponível em: <https://www2.users.ak.tu-berlin.de/akgroup/ak_pub/seacen/2011/Masiero_2011_P2_Review_of_the_crosstalk_cancellation_filter_technique.pdf>. Acesso em: 10 nov. 2024.

MESSONNIER, Jean-Christophe; LYZWA, Jean-Marc; LING, Alexis. Méthodes de prise de son multicanal: objets. **Cahier Louis** Lumière, Saint-Denis, n. 15, p. 114-125, sep. 2022. Disponível em: <<https://www.ens-louis-lumiere.fr/la-recherche/publications-cahiers-louis-lumiere/74-2/>>. Acesso em: 25 ago. 2024.

MEYER, Jürgen. **Akustik und musikalische aufführungspraxis**. Frankfurt: Verlag das Musikinstrument, 1980.

MICROPHONES techniques for recording. **Shure**. 2014. Disponível em:

<https://www.shure.com/damfiles/default/global/documents/publications/en/performance-production/microphone_techniques_for_recording_english.pdf-bb0469316afdb6118691d2f3f5e3ff01.pdf>. Acesso em: 24 ago. 2024.

NEUMANN. **Production expert**: Neumann app to check SRA - Recording Angle Calculator. Neumann, 8 mar. 2018. Disponível em: <<https://apps.apple.com/us/app/recording-tools/id576702914>>. Acesso em: 10 nov. 2024.

NICOLAS Party: talking heads – immersive sound installation. **University Oxford**. Feb. 2018. Disponível em: <<https://empres.music.ox.ac.uk/nicloas-party-talking-heads-immersive-sound-installation>>. Acesso em: 16 set. 2024.

PAUL, Stephan. Binaural recording technology: a historical review and possible future developments. **Acta Acustica United with Acustica**, v. 95, n. 5, p. 767-788, set. 2009. Disponível em: <https://www.researchgate.net/publication/233582452_Binaural_Recording_Technology_A_Historical_Review_and_Possible_Future_Developments>. Acesso em: 10 nov. 2024.

PÉRIAUX, Bergame; OHL, Jean-Luc; THÉVENOT, Patrick. **Le son multicanal**: de la production à la diffusion du son 5.1, 3D et binaural. Paris: Dunod, 2015.

PICKETT, David. **Pantophonic recording**. Dec. 3, 2018. Disponível em: <<https://www.fugato.com/pickett/pantoph.shtml>>. Acesso em: 5 nov. 2024.

PORRES, Alexandre Torres. **ELSE - EL locus solus' externals for Pure Data**. 2024. Disponível em: <<https://github.com/porres/pd-else>>. Acesso em: 10 abr. 2024.

PUCKETTE, Miller. Pure Data. **Los Angeles Pure Data Community**. 1996. Disponível em: <<https://puredata.info>>. Acesso em: 15 set. 2024.

ROSSING, Thomas D. (ed.). **Springer handbook of acoustics**. 2nd ed. Berlin: Springer, 2014.

SBA – SOCIEDADE BRASILEIRA DE ANATOMIA. **Terminologia anatômica**. São Paulo : Ed. Manole, 2001.

SCHOEN, Timothy. **Plugdata**: a visual programming environment for audio experimentation, prototyping and education. Disponível em: <<https://plugdata.org/>>. Acesso em: 15 set. 2024.

SHANKLEMAN, Marin. Early stereo recordings restored. **BBC News**. 2008. Disponível em: <<http://news.bbc.co.uk/2/hi/technology/7537782.stm>>. Acesso em: 26 ago. 2024.

SLOTTE, Benedict. Surround recording using coincident and spaced microphones combined with 2-to-3 upmixing. In: JOINT BALTIC-NORDIC ACOUSTICS MEETING, 2004, Mariehamn, Åland (Finland) **Anais**. Espoo (Finland): Acoustical Society of Finland, 2004. Disponível em: <<http://legacy.spa.aalto.fi/bnam04/webprosari/papers/o48.pdf>>. Acesso em: 5 jun. 2024.

SMALLEY, Denis. Space-form and the acousmatic image. **Organised Sound**, v. 12, n. 1, p. 35-58, Apr. 2007. Disponível em: <<https://beyondlistening.myblog.arts.ac.uk/files/2023/03/Space-form-and-the-acousmatic-image.pdf>>. Acesso em: 10 nov. 2024

SOUNDSCAPE design. **Hydrock**, 2024. Disponível em: <<https://www.hydrock.com/what-we-do/soundscape-design>>. Acesso em: 15 set. 2024.

SOUZA, Flíbio Ferreira de. **A percepção e a espacialização da música eletroacústica no cinema, em surround 5.1: um estudo de caso**. 2023. Tese (Doutorado em Música) - Universidade Estadual de Campinas, Campinas, 2023.

THE TELEPHONE at the Paris Opera. **Scientific American**, v. 45, n. 27, p. 422-423, 1881. Disponível em: <<https://babel.hathitrust.org/cgi/pt?id=mdp.39015024538491&seq=5>>. Acesso em: 10 nov. 2024.

TORICK, Emil. Highlights in the history of multichannel sound. **Journal of the Audio Engineering Society**, v. 46, n. 1/2, p. 27-31, 1 feb. 1998. Disponível em: <<https://issuu.com/disclord/docs/multichannel-history-torrick/1>>. Acesso em: 10 nov. 2024.

USP-LSI - UNIVERSIDADE DE SÃO PAULO - LABORATÓRIO DE SISTEMAS INTEGRÁVEIS. **O projeto AUDIENCE**. Disponível em: <<https://www.lsi.usp.br/interativos/neac/audience/audience.html#0>>. Acesso em: 10 set. 2024.

VELLIARD, Jacques. **Cantos de aves do Brasil**. Campinas: SP Sociedade Brasileira de Ornitologia/ Unicamp, 1995. 1 CD.

WHAT is Ambisonics?. **University Oxford**. [2024]. Disponível em: <<https://intothesoundfield.music.ox.ac.uk/what-is-ambisonics>>. Acesso em: 20 ago. 2024.

WILLIAMS, Michael. **Unified theory of microphone systems for stereophonic sound recording**. AES Preprint 2466. 1987. 41 p. Disponível em: <<https://www.williamsmad.com/Papers/Papers.html>>. Acesso em: 13 set. 2024.

WILLIAMS, Michael. **Microphone arrays for natural multiphony**. AES Preprint 3157. 1991. 13 p. 1991. Disponível em: <<https://www.williamsmad.com/Papers/Papers.html>>. Acesso em: 13 set. 2024.

WILLIAMS, Michael. **Multichannel sound recording practice using microphone arrays**. In: AES INTERNATIONAL CONFERENCE ON MULTICHANNEL AUDIO, 24, 2003, Banff (Canada). New York: AES, 2003. 16 p. Disponível em: <<https://aes2.org/publications/elibrary-page/?id=12287>>. Acesso em: 2 nov. 2024.

ZUMER, Javier. Figuring out: ambisonics. **Javier Zumer Sound Design**. Aug. 30, 2021. Disponível em: <<https://javierzumer.com/blog/2021/8/30/figuring-out-ambisonics>>. Acesso em: 5 nov. 2024.

SOBRE OS AUTORES

Jose Augusto Mannis: Compositor, performer eletroacústico, produtor de rádio, pesquisador em documentação musical, acústica e engenharia de som. É professor do Departamento de Música/IA/Unicamp. Suas composições abrangem música instrumental, eletroacústica, trilhas para vídeo, cinema, teatro, rádio arte. Atuou no Ensemble de L'Itinéraire (Paris), Institut National de l'Audiovisuel – INA/GRM – Groupe de Recherches Musicales, La Grande Fabrique (Rouen), Espace Musical (Puce Muse) (Paris), Núcleo Hespérides-Música das Américas (São Paulo). Desde 2010 tem desenvolvido projetos de rádio arte e ecologia sonora com Janete El Haouli. Atualmente coordena o Laboratório de Acústica e Artes Sonoras – LASom/IA/Unicamp, e o Grupo de Estudos do Som e Processos Criativos junto ao Instituto de Estudos Avançados-IdEA/Unicamp e ao CNPq. ORCID: <https://orcid.org/0000-0002-9484-5674>. E-mail: jamannis@unicamp.br

Igor Abdo Aguilar: Compositor, performer eletroacústico, designer sonoro de jogos digitais, graduado em Música na Unicamp. Atualmente, faz mestrado no Programa de Pós-Graduação em Música/IA/Unicamp, atuando em pesquisas em sonologia, acústica e engenharia de som junto ao Laboratório de Acústica e Artes Sonoras - LASom/IA/Unicamp. Suas composições abrangem música instrumental, eletroacústica, trilhas para jogos digitais. É membro do Grupo de Estudos do Som e Processos Criativos do Instituto de Estudos Avançados-IdEA/Unicamp e CNPq. ORCID: <https://orcid.org/0000-0001-7170-9306>. E-mail: igor.ocarina.dsp@gmail.com

[Amanda] Chang Rodrigues: Compositora, live performer e sintetista, graduada em Comunicação Social pela Universidade Salgado de Oliveira, com MBA em Gestão Empresarial pela UFRJ. Especialista em síntese sonora aplicada a sintetizadores modulares. Atualmente é mestranda na Unicamp sob orientação de José Augusto Mannis, com pesquisa focada em narrativa sonora e movimento espacial. Reconhecida internacionalmente, apresentou-se como artista no Rock in Rio e ministrou workshops sobre live electronics no Amsterdam Dance Event (ADE). Também colabora com a plataforma Echio, dedicada a mentorias artísticas. ORCID: <https://orcid.org/0009-0004-0253-1478>. E-mail: changrodriguesliveact@gmail.com

TAXONOMIA CREDIT

J. A. Mannis			
X	Conceptualization	X	Resources
X	Data curation		Software
X	Formal Analysis	X	Supervision
X	Funding acquisition	X	Validation
X	Investigation	X	Visualization
X	Methodology	X	Writing – original draft
X	Project administration	X	Writing – review & editing

<https://credit.niso.org/>

I. A. Aguilar			
	Conceptualization	X	Resources
X	Data curation	X	Software
X	Formal Analysis		Supervision
	Funding acquisition	X	Validation
X	Investigation	X	Visualization
X	Methodology	X	Writing – original draft
	Project administration		Writing – review & editing

<https://credit.niso.org/>

A. C. Rodrigues			
	Conceptualization	X	Resources
	Data curation		Software
X	Formal Analysis		Supervision
	Funding acquisition		Validation
X	Investigation		Visualization
	Methodology	X	Writing – original draft
	Project administration		Writing – review & editing

<https://credit.niso.org/>

DADOS ABERTOS

Os dados utilizados neste estudo estão disponíveis em **Repositório de Dados de Pesquisa da Unicamp** e podem ser acessados através dos links:

MANNIS, José Augusto, 2024, “Imagem sonora contínua surround a partir de arranjos de microfones e caixas acústicas distribuídas em segmentos iguais de circunferência”, <https://doi.org/10.25824/redu/INTSA5>, Repositório de Dados de Pesquisa da Unicamp,

MANNIS, José Augusto; EL HAOULI, Janete, 2024, “Gunma Belém Brazil 2019 Capitão do Mato - sound rec ESMA6 ch1 to 6”, <https://doi.org/10.25824/redu/1TMGZH>, Repositório de Dados de Pesquisa da Unicamp.

MANNIS, José Augusto; AGUILAR, Igor Abdo, 2024, “Criação musical com especialização em áudio imersivo a partir de simulação de tomada de som multicanal”, <https://doi.org/10.25824/redu/BS3QOE>, Repositório de Dados de Pesquisa da Unicamp.

MANNIS, José Augusto; RODRIGUES, A. Chang, 2024, “Difusão sonora nas Ruínas: estudo de sonologia e composição musical espacial”, <https://doi.org/10.25824/redu/4SNATL>, Repositório de Dados de Pesquisa da Unicamp.