

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA MECÂNICA**

Sergio Luiz Aguirre

**Implementação e avaliação de um sistema de
virtualização de fontes sonoras**

Florianópolis

2017

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA MECÂNICA**

Sergio Luiz Aguirre

**IMPLEMENTAÇÃO E AVALIAÇÃO DE UM SISTEMA DE
VIRTUALIZAÇÃO DE FONTES SONORAS**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia Mecânica da Universidade Federal de Santa Catarina para a obtenção do grau de Mestre em Engenharia Mecânica .

Orientador:

Prof. Dr. Eng. Stephan Paul

Coorientador:

Prof. Dr. Eng. Julio A. Cordioli

Coorientador:

Prof. Ph. D. Bruno S. Masiero

Florianópolis

2017

Aguirre, Sergio Luiz

Implementação e avaliação de um sistema de virtualização de fontes sonoras / Sergio Luiz Aguirre; orientador, Stephan Paul ; coorientador, Júlio Apolinário Cordioli ; coorientador, Bruno Sanches Masiero. - Florianópolis, SC, 2017. 160 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia Mecânica.

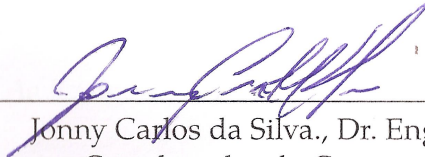
Inclui referências

1. Engenharia Mecânica. 2. Acústica Virtual. 3. Virtualização. 4. Processamento de sinais. 5. Ensaio Subjetivo. I. Paul, Stephan II. Cordioli, Julio Apolinário. III. Masieiro, Bruno Sanches. IV. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia Mecânica. V. Título.

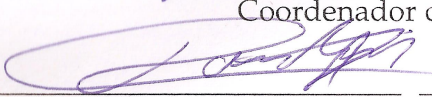
Sergio Luiz Aguirre

IMPLEMENTAÇÃO E AVALIAÇÃO DE UM SISTEMA DE VIRTUALIZAÇÃO DE FONTES SONORAS

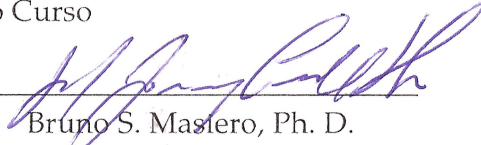
Esta Dissertação foi julgada adequada para obtenção do Título de Mestre em Engenharia Mecânica e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Mecânica da Universidade Federal de Santa Catarina, Florianópolis, 07 de dezembro de 2017.



Jonny Carlos da Silva., Dr. Eng.
Coordenador do Curso



Julio A. Cordioli, Dr. Eng.
Coorientador



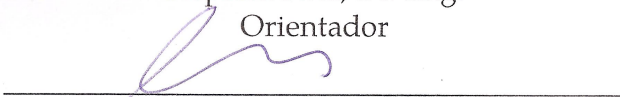
Bruno S. Masiero, Ph. D.
Coorientador

(Participou por videoconferência)

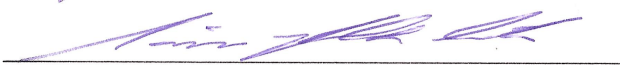
Banca Examinadora:



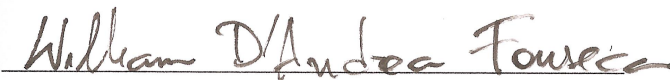
Stephan Paul, Dr. Eng.
Orientador



Andrey Ricardo da Silva, Ph. D.



Márcio Holsbach Costa, Ph. D.



William D'Andrea Fonseca, Dr. Eng.

Ao Ice, inesquecível.

Agradecimentos

À minha esposa, incentivadora e amada Lilian Daminelli Aguirre por suportar-me, compreender e tolerar meus defeitos e minhas ausências principalmente na etapa final.

Ao G.E. pela inspiração, apoio e confiança.

Ao meu orientador Stephan Paul e meus coorientadores Júlio Cordioli e Bruno Masiero pela orientação, dedicação, pelos conselhos, pelos conhecimentos e pela oportunidade oferecida de trabalhar em um tema tão interessante para uma causa tão nobre. Ainda, por todo o empenho e dedicação para que este trabalho fosse realizado da melhor forma possível, realmente obrigado.

Ao Professor Arcanjo Lenzi que com sua paciência infinita esclareceu tantas dúvidas e indicou ótimos caminhos. Ao Professor da Engenharia Acústica da UFSM e amigo William D'Andrea Fonseca pelo contínuo apoio desde a graduação, as inúmeras conversas, dúvidas esclarecidas e o esforço para tornar este documento o melhor possível. Aos Professores Eric Brandão e Paulo Henrique Mareze da Engenharia Acústica da UFSM por toda base teórica, indispensável para a realização do trabalho e pelo incentivo.

A todos que participaram dos ensaios subjetivos ou colaboraram de alguma maneira para a realização deles. À equipe de apoio do LVA, os técnicos Rodrigo e Rafael, em especial agradeço os conselhos e as longas e prazerosas conversas com a dona Sônia. Ao meu amigo Matheus Pereira da UC-PT e aos amigos que fiz no LVA em todos os grupos, obrigado pelas correções, conversas, auxílio nos experimentos e ideias.

Em especial aos membros do x-yri: Matheus Veloso, Henrique Alende da Silveira, Wagner Souza, Guilherme Zanotelli, Guilherme da Silva, Andressa Otto, Danilo Braga, Gleidson Martins, José Pedro, Pedro Nunes, Camilo Arrubla (*in memorian*), Juan Martin Aponte e Rebecca Bugarin. Vocês foram essenciais para o meu crescimento e para a conclusão desse trabalho. Agradeço ao engenheiro acústico e amigo Gil Felix Greco pelo auxílio no desenvolvimento e avaliação do *setup*. Agradecimento especial ao engenheiro acústico Bruno Fujita pelo auxílio nos ensaios e aos engenheiros acústicos Bernardo Murta e Lucas Lobato pela paciência e pelos auxílios e correções.

Ao PosMEC e a todos os professores do LVA, pela excelente formação dada.

Ao CNPq pelo apoio financeiro.

A todas as outras pessoas que direta ou indiretamente colaboraram com este trabalho.

*“ma non esiste prospettiva senza due punti di
vista.”*

Federico Lucia 

Resumo

Próteses auditivas como aparelhos de amplificação sonora individual (AASI) e implantes cocleares (IC) são importantes coadjuvantes na sociabilização de um indivíduo com perda auditiva, por exemplo restaurando a capacidade de compreender a fala. Os testes de inteligibilidade da fala utilizados atualmente para avaliar candidatos a uma prótese auditiva não consideram a distribuição espacial de fontes sonoras e de ruído, podendo levar a indicações incorretas ou excluindo indicações que seriam necessárias. Assim, testes que consideram a distribuição espacial das fontes sonoras e ruído seriam desejáveis e precisam ser desenvolvidos e implementados tendo em vista ainda as condições técnicas existentes nos centros de avaliação. O presente trabalho apresenta uma avaliação do Erro de Localização com a utilização de um sistema octogonal de alto-falantes por meio de duas técnicas de virtualização de fontes sonoras: (1) Sistema Vetorial de Panorama por Amplitude (SVPA) e (2) Sistema Vetorial de Panorama por Intensidade (SVPI), do inglês *Vector Base Amplitude/Intensity Panning* - (VBAP) e (VBIP), respectivamente. O Erro de Localização foi mensurado objetivamente por meio de uma técnica de triangulação utilizando a diferença no tempo de chegada da frente de onda sonora nos sensores. O Erro de Localização foi mensurado também por meio de ensaios subjetivos e denominado Erro de Localização Perceptivo (ELP). O ELP apresenta a resolução da percepção espacial em cada técnica para o número proposto de alto-falantes. A análise de repetibilidade aponta resolução de $\approx 3^\circ$ para os sinais analisados subjetivamente e erros máximos na ordem de $\approx 1,5^\circ$ nas medições objetivas. Foram obtidos valores esperados, com fator de abrangência de 95 %, de $\approx 4^\circ$ nos ensaios subjetivos. Assim as técnicas apresentaram-se equivalentes e comparáveis aos

ensaios que utilizam fontes sonoras reais e às simulações revisadas na literatura.

Palavras-chave: virtualização, VBAP, VBIP, áudio virtual, processamento de sinais.

Abstract

Hearing prostheses such as hearing aids and cochlear implants are of outmost importance for deaf people, providing the ability to understand speech. Actually speech-in-noise recognition tests used to identify candidates for hearing aids and cochlear implants do not take into account that speech and noise are usually spatially distributed. This may result in unnecessary indications but also might exclude patients that should receive a hearing aid or cochlear implant. It is therefore desirable to have test procedures and set-ups that take spatial distributions of noise and speech signal into account and can be implemented in existing facilities. The present work presents an evaluation of the Location Error (LE) with the use of an octagonal system of loudspeakers through two sound sources virtualization techniques: (1) Vector Base Amplitude Panning (VBAP) and (2) Vector Base Intensity Panning (VBIP). The Localization Error was measured objectively by means of a triangulation technique using the difference in the arrival time of the sound wavefront in the sensors. The Localization Error was also measured through subjective tests and called Perceptual Location Error (ELP). The Perceptual Location Error (PLE) delimits the spatial perception resolution in each technique for the proposed number of loudspeakers. Expected values of $\approx 4^\circ$ were obtained in the subjective tests. The repeatability analysis indicates a resolution of approximately 3° for the subjectively analyzed signals and maximum errors on the order of 1.5° in the objective measurements, so the techniques were equivalent and comparable to the tests that use real sound sources and the simulations reviewed in the literature.

Keywords: virtualization, VBAP, VBIP, virtual audio, signal processing.

Sumário

Nomenclatura	25
1 Introdução	29
1.1 Motivação da pesquisa	30
1.2 Objetivo.....	31
1.3 Estrutura do trabalho.....	32
2 Revisão bibliográfica	33
2.1 Tecnologias assistivas para audição	33
2.2 Protocolos de avaliação da inteligibilidade da fala	34
2.3 Audição biauricular e localização sonora	36
2.3.1 Diferença de tempo de chegada e de nível do som entre as orelhas	37
2.3.2 Redução do mascaramento devido à espacialidade.....	39
2.3.3 Erro de localização perceptivo	39
2.4 Virtualização de fontes sonoras.....	42
2.4.1 Sistema Vetorial de Panorama de Amplitude	43
2.4.2 Sistema Vetorial de Panorama de Intensidade	46
3 Sistema para ensaios de localização sonora	47
3.1 Sistema e ambiente de ensaios	48
3.1.1 Reflexões	52
3.1.2 Resposta em frequência da sala.....	53

3.2	Dispositivos de reprodução	53
3.3	Sinais utilizados na avaliação	54
3.4	Nível de pressão sonora de reprodução	55
4	Avaliação do sistema	59
4.1	Avaliação objetiva.....	61
4.1.1	Descrição do experimento	62
4.1.2	Resultados experimentais da medição objetiva ...	65
4.1.2.1	Resultados arranjo em linha.....	65
4.1.2.2	Resultados arranjo em triângulo.....	68
4.2	Avaliação subjetiva	73
4.2.1	Ensaio subjetivo	74
4.2.2	Ensaio piloto	75
4.2.2.1	Método e procedimento de ensaio piloto..	75
4.2.2.2	Instruções e treinamento.....	77
4.2.2.3	Impressões dos jurados	78
4.2.2.4	Análise da influência da duração do ensaio nos resultados do ensaio subjetivo piloto	79
4.2.2.5	Modificações realizadas para o ensaio final	80
4.2.3	Ensaio subjetivo principal	81
4.2.3.1	Erro médio no tempo.....	83
4.2.3.2	Análise de repetibilidade intra indivíduo.	84
4.2.3.3	Análise espacial do erro de localização perceptivo	88
5	Conclusões e trabalhos futuros	99
5.1	Conclusões	99
5.2	Trabalhos futuros	102

Referências bibliográficas	107
Apêndices	117
A RI no domínio do tempo em dB	119
B RI no domínio do tempo	123
Anexo	129
A Monitor de áudio de referência. Genelec 8030C	131
B HEAD <i>Measurement System HMS III</i>	135
C <i>Hammerfall DSP System HDSP 9652 - RME</i>	139
D Conversor AD/DA SQuadriga II - 3320	143
E Datasheet G.R.A.S. 46AE 1/2"	147
F <i>Datasheet G.R.A.S. 46AQ 1/2"</i>	151
G <i>Datasheet Calibrador B&K Tipo 4231 hardware.</i>	155
H <i>Datasheet G.R.A.S. 40AH 1".</i>	157
Índice remissivo	159

Lista de Figuras

1.1	Esquema geral de processamento de sinais e avaliações no âmbito do projeto IMPCoclear.	31
2.1	Representação dos elementos de um implante coclear	34
2.2	Representação das coordenadas esféricas	38
2.3	Fluxograma do procedimento para obtenção da localização sonora a partir dos sinais de 3 sensores. .	42
2.4	Configuração estereofônica formulada com vetores.	44
2.5	Esquema representativo do posicionamento dos alto-falantes na técnica VBAP	45
3.1	Imagem da montagem do sistema na câmara semianecoica.	49
3.2	Fluxograma processamento das técnicas VBAP e VBIP.....	49
3.3	NPS do ruído de fundo e Tempo de reverberação (T20) na câmara semianecoica do LVA-UFSC frente aos requisitos da norma ITU-R BS. 1116-3.	52
3.4	FRFs para as 8 posições de excitação no ponto de audição.	53
3.5	Representação do posicionamento dos alto-falantes e numeração relativa ao ângulo.	56
3.6	Ajuste lateral (rotação) dos alto-falantes.	57

4.1	Representações dos métodos propostos para avaliação do sistema.	59
4.2	Fluxograma: Etapas do trabalho.	61
4.3	Representação do arranjo de microfones para avaliação objetiva.	62
4.4	Fluxograma do conjunto de medição utilizada para aquisição de sinais com 3 sensores.	63
4.5	Arranjos de microfones posicionados.	64
4.6	Arranjo de três microfones em linha.....	64
4.7	Ângulo calculado a partir da medição do tempo de chegada na configuração do arranjo em linha e ângulo da fonte virtualizada - VBAP.	66
4.8	Erro calculado a partir da medição do tempo de chegada na configuração do arranjo em linha - VBAP.	67
4.9	Erro calculado a partir da medição do tempo de chegada na configuração do arranjo em linha - VBIP.	67
4.10	Ângulo calculado a partir da medição do tempo de chegada na configuração do arranjo em triângulo e ângulo da fonte virtualizada - VBAP.	69
4.11	Erro calculado a partir da medição do tempo de chegada na configuração do arranjo em triângulo - VBAP.....	69
4.12	Erro calculado a partir da medição do tempo de chegada na configuração do arranjo em triângulo - VBAP.....	70
4.13	Cadeia de medição utilizada para aquisição de sinais biauriculares.	72
4.14	Simulador de tronco e cabeça posicionado no centro do octógono formado pelos alto falantes na câmara semi anecoica do LVA.	72
4.15	ILDs obtida com simulador de tronco e cabeça para 8 posições.	73
4.16	Fluxograma da fase de treinamento do ensaio subjetivo aplicado.	76

4.17	Sujeito posicionado para o ensaio de localização de fontes sonoras (dentro da câmara semianecoica).	77
4.18	Diferenças entre o ângulo apresentado e o ângulo percebido pelo sujeito na sequência de apresentação.	80
4.19	Sujeito posicionado para o ensaio de localização de fontes sonoras, detalhes destacados	81
4.20	Fluxograma do ensaio subjetivo aplicado.....	82
4.21	Média de desvio absoluto em função do número da apresentação da fonte sonora durante o ensaio ...	83
4.22	Erro quadrático médio entre as apresentações com ângulos repetidos - ruído branco contínuo WN.	85
4.23	Erro quadrático médio entre as apresentações com ângulos repetidos - Fala	86
4.24	Erro quadrático médio entre as apresentações com ângulos repetidos ruído branco intermitente WN_i	87
4.25	VBAP Erro quadrático médio de localização - grupos de 5 graus.....	89
4.26	VBIP Erro quadrático médio de localização - grupos de 5 graus.....	90
4.27	ELP absoluto - sinal: WN Técnicas VBIP e VBAP.	91
4.28	ELP absoluto - sinal: Fala. Técnicas VBIP e VBAP. ...	92
4.29	ELP absoluto - sinal: ruído branco interrompido. Técnicas VBIP e VBAP.....	93
4.30	Erro quadrático médio de localização perceptivo para os ângulos entre dois alto-falantes de um par (grupos de 45°).....	94
4.31	Teste de Tukey para comparação das médias.....	95
A.1	Diferença em dB do nível de pressão sonora do som direto e das primeiras reflexões nas posições 1 e 2 de fonte sonora (0° e 45°).....	119
A.2	Diferença em dB do nível de pressão sonora do som direto e das primeiras reflexões nas posições 3 e 4 de fonte sonora (90° e 135°).....	120

A.3	Diferença em dB do nível de pressão sonora do som direto e das primeiras reflexões nas posições 5 e 6 de fonte sonora (180° e 225°).....	120
A.4	Diferença em dB do nível de pressão sonora do som direto e das primeiras reflexões nas posições 7 e 8 de fonte sonora (270° e 315°).....	121
B.1	Resposta ao impulso na posição de audição com alto falante número 1 (equivalente ao ângulo 0°). ...	123
B.2	Resposta ao impulso na posição de audição com alto falante número 2 (equivalente ao ângulo 45°). .	124
B.3	Resposta ao impulso na posição de audição com alto falante número 3 (equivalente ao ângulo 90°). .	124
B.4	Resposta ao impulso na posição de audição com alto falante número 4 (equivalente ao ângulo 135°). .	125
B.5	Resposta ao impulso na posição de audição com alto falante número 5 (equivalente ao ângulo 180°). .	125
B.6	Resposta ao impulso na posição de audição com alto falante número 6 (equivalente ao ângulo 225°). .	126
B.7	Resposta ao impulso na posição de audição com alto falante número 7 (equivalente ao ângulo 270°). .	126
B.8	Resposta ao impulso na posição de audição com alto falante número 8 (equivalente ao ângulo 315°). .	127

Lista de Tabelas

3.1	Diferenças de amplitude entre som direto e primeiras reflexões para as oito posições de alto-falantes. ...	52
4.1	Valores esperados com fator de abrangência de $\approx 95\%$ $k_p = 2$ na medição objetiva.	70
4.2	Mediana do Erro quadrático médio na localização de fontes sonoras repetidas.	87
4.3	Análise de variância (ANOVA) das medianas de erro absoluto de repetição do ensaio subjetivo final para as técnicas VBAP e VBIP.	88
4.4	Análise de variância (ANOVA) entre as médias de erro absoluto do ensaio subjetivo para grupos de 45°	95
4.5	Valores esperados do ELP absoluto com fator de abrangência de $\approx 95\%$ ($k_p = 2$) na medição subjetiva.	96
4.6	Análise de variância (ANOVA) entre a média de erro absoluto do ensaio subjetivo nas técnicas VBAP e VBIP.	97

Nomenclatura

Símbolos gerais

c	Velocidade de propagação do som no ar, Eq. (4.1), pág. 65.
d_1	Distância entre a frente de onda e o sensor 1, Eq. (4.1), pág. 65.
d_2	Distância entre a frente de onda e o sensor 2, Eq. (4.1), pág. 65.
g	Matriz de ganhos, pág. 45.
k_p	Fator de abrangência que produz um intervalo correspondente a um nível especificado da confiança p , pág. 70.
l_1	Vetor do centro ao canal 1, Eq. (2.1), pág. 44.
l_2	Vetor do centro ao canal 2, Eq. (2.1), pág. 44.
L_{12}	Matriz de posições dos alto-falantes (canais), pág. 45.
M_i	Sensores $i = 1, 2, 3$, pág. 65.
p	Vetor do centro à fonte virtual, Eq. (2.1), pág. 44.

T20	Tempo de reverberação calculado a partir do decaimento de 20 dB extrapolando para 60 dB, pág. 51.
x	Distância entre a fonte sonora e o centro do arranjo de alto-falantes, Eq. (4.2.a), pág. 65.

Símbolos gregos

θ_{DOA}	Ângulo computado por meio de medição objetiva (<i>Direction-of-arrival</i>), pág. 65.
θ_{fonte}	Ângulo relativo à posição da fonte sonora emitida, Eq. (4.3), pág. 82.
θ_r	Ângulo relativo à posição da fonte sonora emitida pelo <i>setup</i> , pág. 66.
θ_{suj}	Ângulo relativo à posição da fonte sonora percebido pelo sujeito, Eq. (4.3), pág. 82.
θ_x	Ângulo relativo ao posicionamento da fonte sonora no círculo, Eq. (4.2), pág. 65.
τ_{12}	Diferença de tempo de chegada da onda sonora entre os sensores M_1 e M_2 , Eq. (4.1), pág. 65.
τ_{32}	Diferença de tempo de chegada da onda sonora entre os sensores M_3 e M_2 , Eq. (4.1), pág. 65.
μ	Média de um conjunto de dados, Eq. (2.0), pág. 41.
μ_z	Esperança de uma grandeza z com distribuição normal, pág. 96.
σ	Desvio padrão, pág. 41.

Acrônimos e abreviaturas

2D	Representação em duas dimensões, pág. 43.
3D	Representação em três dimensões, pág. 44.
AASI	Aparelho de amplificação sonora individual, pág. 33.
AD	Conversor Analógico Digital, pág. 63.
CA	Canal Auditivo, pág. 56.
CAT	Comitê de Ajudas Técnicas, pág. 33.
DA	Conversor Digital Analógico, pág. 63.
DOA	<i>Direction-of-Arrival</i> , pág. 41.
EA	Erros aleatórios, pág. 41.
EC	Erros constantes (sistemáticos), pág. 41.
EDC	<i>Energy Decay Curve</i> , pág. 51.
EL	<i>Erro de Localização</i> , pág. 41.
ELP	Erro de localização perceptivo, pág. 40.
GCC	<i>Generalized Cross Correlation</i> , pág. 42.
IC	Implante Coclear, pág. 33.
iFFT	Transformada Rápida de Fourier Inversa (<i>Inverse Fast Fourier Transform</i>), pág. 51.
LDL	<i>Loudness discomfort level</i> , pág. 36.
LVA	Laboratório de Vibrações e Acústica, pág. 31.
MAD	<i>Mean Absolute Deviation</i> , pág. 83.

MCL	<i>Most Comfort Level</i> , pág. 36.
NPS	Nível de pressão sonora [dB] ref. 20 μ Pa, pág. 35.
RIs	Resposta impulsiva, pág. 51.
SNR	Relação sinal-ruído, pág. 35.
SRM	<i>Spatial release from Masking</i> , pág. 39.
SRT	<i>speech recognition threshold</i> , pág. 35.
SVPA	Sistema Vetorial de Panorama de Amplitude, pág. 33.
SVPI	Sistema Vetorial de Panorama de Intensidade, pág. 33.
TAs	Tecnologias assistivas, pág. 33.
TDOA	<i>Time-Difference-of-Arrival</i> , pág. 41.
TR	Tempo de reverberação, pág. 51.
UCL	<i>Uncomfortable listening level</i> , pág. 36.
UFSC	Universidade Federal de Santa Catarina, pág. 30.
VBAP	<i>Vector Based Amplitude Panning</i> , pág. 33.
VBIP	<i>Vector Based Intensity Panning</i> , pág. 33.
WN	Ruído branco (<i>White Noise</i>), pág. 82.
WN _i	Ruído branco (<i>White Noise</i>) intermitente, pág. 82.

1 Introdução

Perdas de capacidade auditiva podem ser um fator limitante na vida de um indivíduo podendo reduzir a capacidade de comunicação e acarretando na diminuição da interação social em ambientes de conversa em grupo, conversas com estranhos e até mesmo com família e amigos. No escopo das questões emocionais a perda auditiva sem tratamento pode causar uma série de sentimentos negativos: desde a frustração por não compreender a comunicação, ansiedade, entre outros, podendo resultar em depressão. Entre as alternativas disponíveis para problemas de perda auditiva existem tecnologias assistivas para a audição, chamadas próteses auditivas, tais como aparelhos de amplificação sonora individual (AASI), indicados para amenizar os efeitos em perdas leves a severas e implantes cocleares (IC) para perdas de grau severo a profundo.

A partir dos critérios definidos por portarias do Ministério da Saúde brasileiro (atualizados em 2014 [1]) elevou-se o número de pacientes que, por meio de avaliação baseada nestes critérios, têm direito ao implante coclear. Contudo, testes de inteligibilidade da fala considerando a distribuição espacial de fontes sonoras e ruído não fazem parte da rotina clínica convencional de avaliação no Brasil [2,3]. A avaliação que não considera a distribuição espacial de fontes sonoras não leva em conta fenômenos como o melhoramento da inteligibilidade da fala por redução de mascaramento dado pelo posicionamento das fontes sonoras, que, quando separadas no espaço, resultam em uma maior relação sinal ruído para o ser humano. Pode-se assim levar a indicações incorretas ao implante coclear no caso de avaliações que apresentam apenas uma posição

ruído e conteúdo de fala simultaneamente pelo mesmo alto-falante em que o sujeito pode ter um desempenho pior de inteligibilidade de fala do que teria em uma condição com fontes separadas, como acontece em situações reais.

1.1 Motivação da pesquisa

Diante desse cenário, verifica-se a necessidade de propor novos protocolos de avaliação de inteligibilidade da fala, que considerem a distribuição espacial das fontes sonoras, aproximando-se das condições acústicas de ambientes reais. Esse escopo atende parte dos objetivos do projeto “Desenvolvimento de Transdutor Implantável e Algoritmos Avançados para Implantes Cocleares com Objetivo de Adaptação à Realidade Brasileira” IMPCoclear, parceria entre a Financiadora de Estudos e Projeto (FINEP), Universidade Federal de Santa Catarina (UFSC), instituto Otovida e a Wavetech Soluções Tecnológicas. À medida que a estrutura física a ser montada bem como os protocolos, *software* e demais ferramentas desenvolvidas contribuirão para o desenvolvimento e a execução de avaliações dos algoritmos de processamento de ICs que estão em desenvolvimento, tanto com pacientes já implantados quanto com ouvintes normais. Há também a possibilidade da criação e de verificação de métricas objetivas de avaliação desses algoritmos, o que facilita a pesquisa e implementação de novas tecnologias no processamento de sinais de implantes cocleares.

As avaliações subjetivas são parte do projeto IMPCoclear, (ver Figura 1.1). A partir da avaliação e validação do sistema para pessoas com condições normais de audição para as fontes sonoras reais e virtuais¹, ensaios subjetivos poderão ser desenvolvidos com pessoas portadoras de deficiência auditiva e usuários de IC ou AASI. Nesses futuros ensaios a performance de localização pode ser avaliada tendo como base os resultados obtidos neste trabalho.

¹ Fontes sonoras reproduzidas (virtualizadas) por um sistema de virtualização de fontes sonoras onde não há uma fonte sonora real (alto-falante).

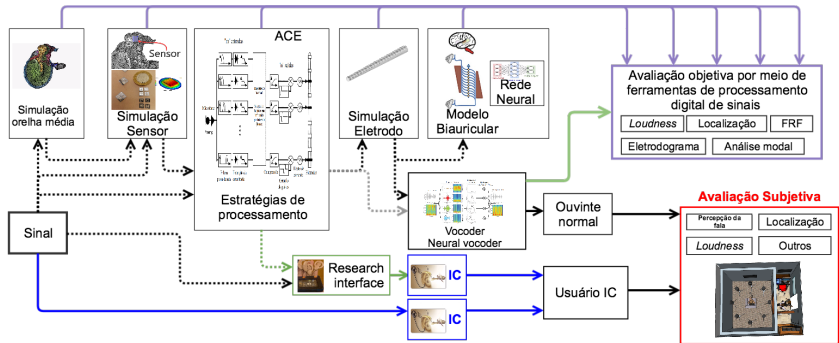


Figura 1.1.: Esquema geral de processamento de sinais e avaliações no âmbito do projeto IMPCoclear.

Além disso, é possível aproveitar o sistema proposto neste trabalho para aprimorar ou desenvolver outras atividades e pesquisas relacionadas às demais tecnologias assistivas para audição e fonação.

1.2 Objetivo

Este trabalho tem como objetivo geral a avaliação objetiva e subjetiva por meio do Erro de Localização (EL) da implementação de um sistema de fontes sonoras reais espacialmente distribuídas e um sistema de fontes sonoras virtuais utilizando duas técnicas de virtualização de fontes sonoras no Laboratório de Vibrações e Acústica (LVA) na UFSC.

Objetivos específicos

Para tal os objetivos específicos do projeto são:

- Implementar um sistema de fontes sonoras virtuais controladas por algoritmos de Matlab;
- Elaborar ensaios subjetivos de localização espacial de fontes sonoras com sinais diferentes, incluindo vozes gravadas em condição anecoica;

- Realizar ensaios subjetivos com normo-ouvintes para validar a localização das fontes sonoras virtuais e reais produzidas pelo sistema;
- Mensurar objetivamente a localização das fontes virtuais.

1.3 Estrutura do trabalho

- O Capítulo 2 apresenta uma revisão bibliográfica sobre o tema, abordando conceitos de tecnologias assistivas para audição, protocolos de avaliação da inteligibilidade da fala (procedimento importante na avaliação de candidatos ao implante coclear), conceitos gerais de audição biauricular, localização sonora e virtualização de fontes sonoras;
- No Capítulo 3 é realizada uma descrição do sistema de virtualização de fontes sonoras abordando as técnicas utilizadas, o ambiente de ensaios e suas condições, os dispositivos de reprodução utilizados, os sinais utilizados nos ensaios, o nível de pressão sonora de reprodução e a calibração em frequência;
- No Capítulo 4 a avaliação objetiva do Erro de Localização a partir de duas configurações de arranjos de sensores é apresentada. Ainda, a avaliação subjetiva do Erro de Localização Perceptivo é descrita e os resultados apresentados;
- O Capítulo 5 apresenta as conclusões do trabalho e tópicos sugeridos para trabalhos futuros.

2 Revisão bibliográfica

A revisão bibliográfica apresentada compreende os tópicos importantes no âmbito da audição biauricular, da virtualização de fontes sonoras utilizando a técnica Sistema Vetorial de Panorama de Amplitude (SVPA) e Intensidade (SVPI), mais conhecidos na literatura como *Vector Based Amplitude Panning* (VBAP) e *Vector Based Intensity Panning* (VBIP), além de conceitos da avaliação da deficiência auditiva. Assim, o objetivo é facilitar o entendimento do texto, bem como consolidar uma base teórica para a pesquisa.

2.1 Tecnologias assistivas para audição

Tecnologias assistivas (TAs) [4, 5] são conceituadas no Brasil pelo Comitê de Ajudas Técnicas (CAT) a partir de 2007 [6]. Para o contexto da audição, os principais produtos de TA são o aparelho de amplificação sonora individual (AASI) e o Implante Coclear (IC).

Implantes cocleares são por definição uma substituição artificial para a orelha interna, sendo apropriados para pessoas que recebem o mínimo ou nenhum benefício com aparelhos auditivos convencionais, os AASI, por terem perda de audição de graus severo (alteração do limiar: 61 a 80 dB) ou profundo (alteração do limiar: maior que 81 dB) [7], devido à doença, lesão ou mau desenvolvimento.

O implante coclear de múltiplos canais, também chamado de *orelha biônica* busca restaurar a audição útil [8] e é atualmente considerado a melhor interface cérebro-máquina já produzida pelo homem [9]. Uma parte do dispositivo, o elemento interno, é implantado cirurgicamente no osso temporal e um pequeno arranjo

de eletrodos é inserido na cóclea que estimula eletricamente o nervo auditivo próximo às células ciliadas, o que leva à ocorrência de sensações auditivas. Outra parte do dispositivo é denominada elemento externo, disposta externamente, possui um microfone, um processador de sinais, os cabos de ligação e uma antena de comunicação com o elemento interno [9, 10]. A Figura 2.1 apresenta um esquema dos elementos citados.

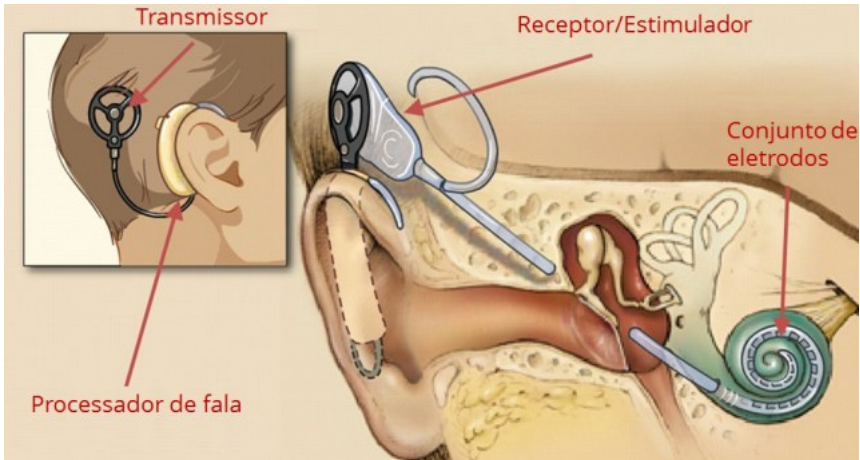


Figura 2.1.: Representação dos elementos de um implante coclear.

Fonte: Biosom [11].

2.2 Protocolos de avaliação da inteligibilidade da fala

A avaliação fonoaudiológica compreende algumas etapas tais como anamnese, avaliação do aproveitamento auditivo e de percepção/inteligibilidade residual da fala por meio de audiometria¹, avaliação da comunicação (fala e voz), avaliação e orientação de expectativas e motivação para ouvir [14]. A avaliação da audição residual pode ser realizada de forma objetiva por meio de avaliação eletrofisiológica da via auditiva ou com avaliações subjetivas por meio de resposta comportamental do indivíduo sujeito a

¹ A audiometria da fala é um conjunto de testes de índice de reconhecimento de fala, limiar de detecção de fala e limiar de reconhecimento de fala [12], também chamados de testes de reconhecimento ou inteligibilidade da fala [13].

estímulos auditivos que podem ser tons puros, ruídos filtrados, sons ambientais, sons verbais (fonemas, palavras ou sentenças) *clicks* ou *tone bursts* [15].

Entre as etapas da avaliação audiológica, a avaliação do aproveitamento auditivo e de reconhecimento/inteligibilidade da fala consiste em testes com apresentação de palavras, sílabas ou fonemas ao paciente em ambiente acusticamente tratado buscando mensurar o limiar de reconhecimento da fala de forma a complementar a audiometria tonal². Contudo, não é encontrado na literatura um consenso quanto às especificações dos ensaios, tais como: tipos de estímulos, o modo de sua apresentação, relação sinal-ruído (SNR) utilizadas e distribuição espacial das fontes do sinal de fala e de ruído.

Nos testes, o reconhecimento da fala é quantificado por meio do *limiar de reconhecimento da fala* (SRT)³, que é uma medida de inteligibilidade da fala que considera uma relação entre limiares de tom puro em determinadas frequências e o nível de pressão sonora mínimo de apresentação do sinal na qual o indivíduo repete o estímulo corretamente em 50% das apresentações, sendo geralmente utilizada a média dos limiares tonais em 500 Hz, 1 kHz e 2 kHz [15, 16].

Existem dois paradigmas na audiologia clínica quanto ao modo de apresentação da fala em testes de avaliação da inteligibilidade da fala: (1) sinais gravados e (2) sinais produzidos por viva voz [17]. Geralmente a prática clínica utiliza a apresentação a viva voz, e em atividades de pesquisa é mais comum a apresentação de estímulos gravados [10, 17].

Uma questão importante, no contexto de audiometria da fala, é o *nível de pressão sonora* (NPS) de reprodução do sinal de fala, ou do sinal de fala combinado com ruído competitivo. Nos testes de audiometria da fala pode ser abordada de algumas maneiras [18]:

² Exame que avalia a capacidade auditiva do indivíduo por meio de estímulos acústicos de tom puro. Esses estímulos são transmitidos por um fone de ouvido, alto-falante ou vibrador ósseo.

³ Sigla em inglês para *speech recognition threshold*.

- O NPS da apresentação pode ser fixado para aquele valor no qual os ouvintes normais alcançam máximo reconhecimento nos testes;
- O NPS da apresentação pode também ser determinado individualmente a partir daquele NPS no qual há o maior conforto para o paciente (MCL)⁴;
- O NPS pode ser definido em 5 dB abaixo do nível de desconforto para o indivíduo (UCL)⁵;

Não há um consenso acerca do valor do nível de pressão sonora adequado para a reprodução em testes de audiometria da fala. Historicamente o nível de reprodução tem sido fixado em 70 dB. Contudo, esse nível não é mais considerado apropriado para representar a conversação, sendo 60 dB(A)⁶ o nível de pressão sonora tido como apropriado para testes de seleção de candidatos a implante coclear [19]. Níveis entre 50 dB e 60 dB são apontados na literatura por Niparko [10], e no português em torno de 60 a 65 dB [15], o que apresenta uma dissonância de valores de referência e conceitos utilizados entre os autores analisados. Uma boa prática é a apresentação em múltiplos níveis de pressão sonora, apesar de estenderem o tempo, pode ser considerada dependendo do tipo do teste [18].

2.3 Audição biauricular e localização sonora

O ato ou processo de identificação da direção de incidência de uma fonte sonora com base no som recebido pelo ouvinte é definido como localização sonora [20]. Para pesquisas em acústica é importante a consideração do receptor, em geral o ser humano. O mecanismo de audição do ser humano tem como principal

⁴ Sigla em inglês para *Most Comfort Level*.

⁵ Sigla em inglês para *Uncomfortable listening level*, também chamado de *Loudness discomfort level* (LDL)

⁶ A notação dB(A) indica que ponderação da curva A é aplicada segundo o referido autor.

característica anatômica ser um sistema biauricular, ou seja, há dois pontos de recepção de sinal (orelhas externas posicionadas em lados opostos da cabeça), no qual todo o conjunto (tronco, cabeça, pavilhões auditivos e outras partes) modifica o sinal que chega às duas membranas timpânicas [21, 22]. A audição biauricular do ser humano e os efeitos associados foram revisados amplamente no livro texto de Blauert [23].

A localização de fontes sonoras em mamíferos é baseada em análises comparativas da energia sonora nas bandas de frequência em cada orelha e de diferença de fase dos sinais nos tímpanos, ou seja, de tempo de chegada entre as orelhas [23, 24], notadamente expressas na literatura como *Interaural Level Difference* (ILD) e *Interaural Time Difference*⁷ (ITD) [25].

O sistema auditivo central, além da análise da localização espacial de fontes sonoras também extrai dos sinais sonoros informações em tempo real do ambiente acústico, tais como geometria e propriedades físicas [22]. Outro benefício é a possibilidade de interpretação de misturas de sons, principalmente oriundas de fontes de diferentes direções [25, 26].

2.3.1 Diferença de tempo de chegada (ITD) e de nível do som entre as orelhas (ILD)

Pelo fato da velocidade de propagação sonora no ar⁸ ser finita e aproximadamente constante, frentes de ondas sonoras percorrem caminhos diferentes até as orelhas, o que acontece quando a incidência não é frontal⁹ ou traseira¹⁰, chegam às orelhas externas em tempos diferentes. Essa pista acústica¹¹ é muito importante para localização de fontes sonoras que emitem sons de baixa frequência [22, 27]. Uma representação dos planos em coordenadas

⁷ Para sinais de tons puros contínuos e outros sinais periódicos utiliza-se o tempo *Interaural Phase Difference* (IPD) [20].

⁸ Propriedade considerada para meios aproximadamente não dispersivos.

⁹ Incidência frontal - ângulo de incidência azimutal igual a 0°.

¹⁰ Incidência traseira - ângulo de incidência azimutal igual a 180°.

¹¹ Em inglês chamadas de *cue*.

esféricas é apresentada na Figura 2.2.

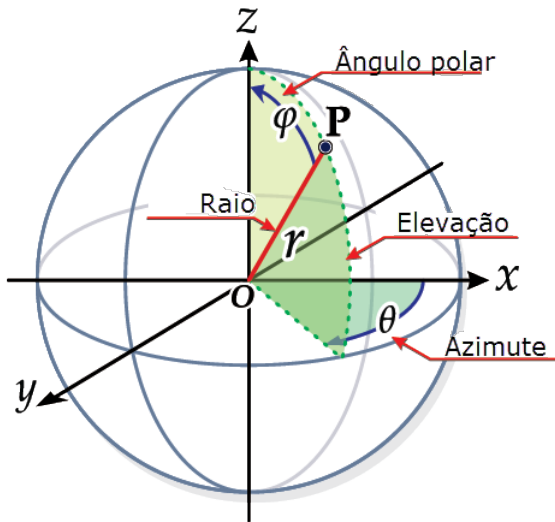


Figura 2.2.: Representação das coordenadas esféricas (adaptado de Fonseca [28]).

A partir de ≈ 1 kHz a cóclea não é mais capaz de codificar a fase de sinais e perde-se então a relação explícita de fase (ou tempo relativo) entre os sinais nas duas orelhas. Assim, quando o comprimento do caminho entre as orelhas excede o comprimento de onda dos sinais provenientes de uma fonte de som lateral, o ITD não é efetivo para resolver laterização. Com o sistema auditivo dividido em duas entradas em lados opostos da cabeça, acontecem atenuações, difrações e absorção sonora diferentes em cada um dos caminhos de propagação sonora até as duas orelhas. Contudo para baixas frequências o tamanho dos obstáculos entre os sinais (elementos anatômicos) não é significativo¹², contando assim com uma baixa influência na diferença de nível do sinal entre as orelhas. Para frequências acima de 200 Hz a forma das orelhas externas, o torso e o formato da cabeça modificam o sinal de maneira significativa e causam diferença de nível entre os sinais que chegam em cada uma das orelhas (ILD) [27].

¹² Significando que a cabeça e o torso são acústicamente transparentes.

2.3.2 Redução do mascaramento devido à espacialidade

A impressão espacial do som percebida por seres humanos depende principalmente da pressão sonora na orelha do ouvinte e da psicoacústica de audição espacial [29]. A grande maioria dos estímulos sonoros ao qual uma pessoa é submetida no dia a dia tem origem em um conjunto de fontes espacialmente separadas [30]. Uma das maiores dificuldades para deficientes auditivos é compreender a fala em ambientes ruidosos. Quando o ruído mascarador é separado espacialmente da fonte sonora existe um aumento da inteligibilidade da fala no ruído, pois os sinais da fala e de ruído chegam às membranas timpânicas em tempos diferentes (ITD) e com amplitudes diferentes (ILD), possibilitando o aumento do foco psicoacústico.

A diferença de tempo de chegada para cada orelha e para cada sinal produz um melhoramento efetivo da SNR [30]. Por isso, sinais acústicos em situações com ruído mascarador são detectados mais facilmente utilizando (quando existe) a informação direcional [31]. Além disso, a detecção de sinais é melhor em campo livre separando espacialmente o sinal e o ruído mascarador; sendo este o fenômeno conhecido como “Redução do Mascaramento devido a Espacialidade”, cujo termo original em inglês é *Spatial release from Masking* (SRM) [32, 33].

Assim, a importância do SRM na comunicação fica evidente, e para uma avaliação adequada da audição e da audição residual em deficientes auditivos o fenômeno SRM precisa ser considerado nos procedimentos de avaliação visando diagnósticos corretos e resultando em indicações coerentes para AASI ou implante coclear.

2.3.3 Erro de localização perceptivo (ELP)

A avaliação da habilidade de percepção humana da localização sonora (ato de localizar a posição verdadeira da fonte sonora no espaço) tem uma quantidade de incerteza associada inerente, bem como um viés operacional que pode resultar em erros de estimativa experimental. Tais erros são dependentes das propriedades do som

emitido, características do ambiente, do tipo de questionamento ou tarefa solicitada ao sujeito e de suas habilidades [20]. A habilidade de localização pode ser separada em habilidade de localização horizontal, vertical e de distâncias [34].

A repetição do ato de localizar fontes sonoras em testes de laboratório transforma o resultado dessa tarefa em uma variável estatística. Para testes de localização nos quais existem valores angulares superiores a 90° no plano azimutal (em que não há movimento da cabeça, ou movimento reduzido), a descrição dos dados é realizada por meio de estatística esférica dada a natureza dos valores angulares. Caso os movimentos da cabeça e tronco sejam livres (apenas avaliação por meio do ângulo frontal do indivíduo) a distribuição assume característica linear¹³ [20]. Testes subjetivos utilizando fontes sonoras reais com ouvintes sem reclamações prévias de perda auditiva (Kerber e Seeber [35]) apresentaram resultados de erro de localização absoluto de fontes sonoras em condições de alta relação sinal-ruído entre $2,3^\circ$ e $6,6^\circ$; com a mediana do erro quadrático médio de $2,8^\circ$. Grohn [36] apresenta a mediana do erro de localização no plano azimutal para fontes sonoras virtuais em movimento de $6,6^\circ$ com ruído branco contínuo. Ainda Romigh *et al.* [37] apresentam uma média de erro de localização com fontes reais para sinais contínuos de $\approx 5^\circ$ e intermitentes $\approx 3^\circ$, apontando assim para uma diretriz para a pesquisa da localização utilizando sinais de ruído branco intermitentes e contínuos. Simulações computacionais realizadas para avaliar diversos aspectos de diferentes técnicas de virtualização de fontes sonoras em diferentes tipos de processamento de aparelhos auditivos Grimm *et al.* [38] apresentam um resultado com uma configuração de oito alto falantes utilizando a técnica VBAP em condição de alta SNR de aproximadamente 5° de Erro de Localização Perceptivo, a partir de algoritmos de predição de localização perceptiva.

Em geral, descreve-se a maior parte das decisões dos seres

¹³ É necessário analisar separadamente os erros em casos de confusão frente-costas [20].

humanos como uma distribuição normal (Gaussiana) na qual, para localização sonora, apresenta a tendência (idealmente) da localização apontar para a direção correta do som e uma distribuição simétrica de erros de localização em torno da média [20]. A distribuição normal é completamente caracterizada a partir da média μ e do desvio padrão σ .

Vieses de erros gerados nas respostas dos sujeitos podem causar desalinhamento dentre o local percebido e a real posição da fonte sonora [20], eles são denominados Erros Constantes (EC) dados pela diferença entre a posição real e a média das observações realizadas. Erros causados por incertezas e imprecisões na avaliação do sujeito, ou mudanças no ambiente de escuta, são denominados Erros Aleatórios (EA) e são caracterizados pelo desvio padrão.

A localização da fonte sonora (e o consequente cálculo de Erro de Localização (EL)) também pode ser realizada por máquinas por meio de algoritmos que comparam sinais obtidos por sensores posicionados em diferentes locais do campo sonoro. Pesquisas na área de localização de fontes sonoras com arranjos de microfones são frequentes na literatura, buscando aprimorar a precisão, a exatidão e aumentar a velocidade de processamento de diferentes métodos [39–44]. Uma das maneiras mais utilizadas é computar a estimativa de *direção de incidência*, em inglês *direction-of-arrival* (DOA), a partir do atraso de tempo de chegada¹⁴ do som entre pares de sensores. A partir disso, conhecendo a posição dos sensores, a localização da fonte é calculada por triangulação [45]. A Figura 2.3 mostra o fluxograma do procedimento para um arranjo triangular na qual s representa a fonte sonora, M1 a M3 representam os sensores, d representa a distância entre os sensores, τ_{21} e τ_{23} representam a estimativa de tempo de chegada da fonte sonora entre os sensores.

¹⁴ Em inglês *Time-Difference-of-Arrival* (TDOA).

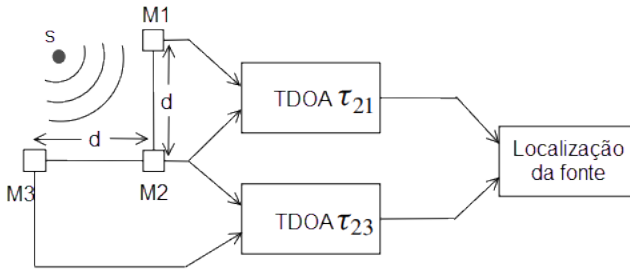


Figura 2.3.: Fluxograma do procedimento para obtenção da localização sonora a partir dos sinais de 3 sensores.

Fonte: Adaptado de [45].

É possível calcular o tempo de atraso em pares de sensores com diversas técnicas, uma das mais utilizadas é chamada *Generalized Cross Correlation* (GCC). Para cada par de sensores é computada a correlação cruzada dos dois sinais. O atraso em que a função de correlação cruzada tem seu máximo é definido como a diferença de tempo de chegada entre os dois sinais [45, 46]. Erros de localização obtidos em medições de (ângulo emitido *vs.* ângulo obtido com medição) por Atmoko *et al.* [47] apresentam valores entre $0,00^\circ$ e $0,75^\circ$. Khaddour [48] apresenta erros utilizando arranjos de três sensores com magnitudes entre $0,0^\circ$ e $2,8^\circ$ com mediana em torno de $1,2^\circ$.

2.4 Virtualização de fontes sonoras

A virtualização de fontes sonoras compreende uma série de técnicas para a criação de eventos sonoros em um ambiente de som virtual utilizando transdutores eletroacústicos (alto-falantes) [49]. Uma visão geral acerca da evolução das técnicas incluindo gravação e reprodução sonora espacial pode ser encontrada em Masi-ero [50]. A aplicação mais generalizada é a reprodução sonora para entretenimento de ouvintes humanos. Também é aplicada para reprodução espacialmente distribuída de ruído (motores, fontes sonoras móveis, interiores de automóveis *etc.*) e vozes em laboratório.

Existem dois conceitos gerais para a virtualização de fontes sonoras por meio de alto-falantes utilizando arranjos distribuídos no espaço. Eles diferem na questão de discretização do campo sonoro existente no ambiente e a consequente criação do campo sonoro para um ponto ou área. O paradigma das técnicas centradas em panorama é a criação de diferenças de tempo de chegada e nível de pressão sonora entre as orelhas em torno de um ponto ótimo (*sweet spot*), de forma que o ouvinte tenha uma impressão de espacialidade naquele ponto. Já o conceito de síntese de campo sonoro procura recriar um campo sonoro completo representativo da área na qual foi gravado [50, 51]. As diversas técnicas de virtualização por meio de alto-falantes têm suas limitações, desde o número de alto-falantes necessários até o posicionamento e o custo computacional.

As maiores diferenças de nível de pressão sonora ILD e de tempo de chegada ITD entre as orelhas são produzidas no plano azimutal Figura 2.2 [38, 52], por isso a maior parte das pesquisas de localização de fontes sonoras e mascaramento de sinais concentra o arranjo de alto-falantes no plano azimutal.

2.4.1 Sistema Vetorial de Panorama de Amplitude (VBAP)

O Sistema Vetorial de Panorama de Amplitude é uma aproximação de primeira ordem da composição dos sinais emitidos, podendo criar fontes virtuais [53]. Considerando uma variação de amplitude entre os alto-falantes a virtualização da fonte sonora pode utilizar o método de Panorama de Amplitude em duas dimensões (2D), o qual é fundamentado na Lei dos Senos e Lei das Tangentes. A derivação de ambas as leis pode ser encontrada em Benesty *et al.* [29]. A principal hipótese do VBAP considera alto-falantes dispostos simetricamente, equidistantes em relação ao ouvinte e sem elevação (mesmo plano horizontal).

A partir do Panorama de Amplitude 2D, o Sistema Vetorial de Panorama por Amplitude, que originalmente em inglês é chamado *Vector Based Amplitude Panning* (VBAP), não limita a quantidade de alto-falantes utilizáveis, contudo utiliza no máximo

3 simultaneamente. Os alto-falantes são arranjados em um círculo de referência (caso 2D) ou esfera (caso 3D), e a fonte virtual criada não pode ser criada fora dessa região [54].

A formulação da técnica VBAP [53] para duas dimensões parte da configuração estereofônica de dois canais (ver Figura 2.4). Reformulada para uma base vetorial, formado por vetores de comprimento unitário $l_1 = [l_{11} l_{12}]^T$ e $l_2 = [l_{21} l_{22}]^T$ que apontam para os alto-falantes e o vetor de comprimento unitário $p = [p_1 p_2]^T$ que aponta para a fonte virtual e apresenta-se como uma combinação linear dos vetores l_1 e l_2 . Utiliza-se aqui a notação T para identificar a transposição

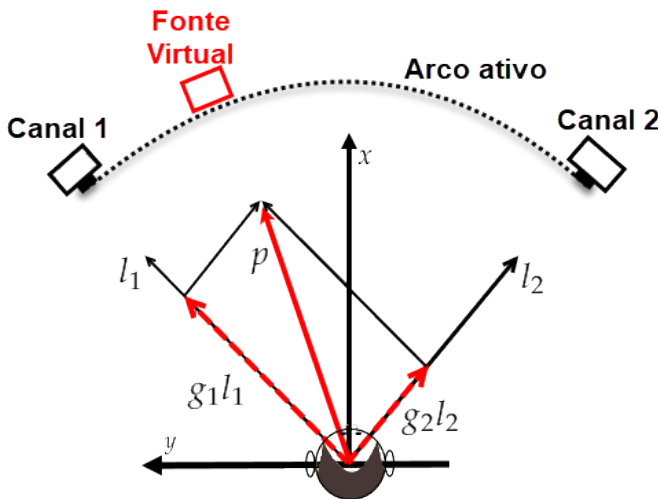


Figura 2.4.: Configuração estereofônica formulada com vetores.
Fonte: Adaptado de [53].

Portanto tem-se o vetor p :

$$p = g_1 l_1 + g_2 l_2, \quad (2.1)$$

sendo g_1 e g_2 (escalares) os fatores de ganho a serem calculados para o posicionamento do vetor relativo à fonte virtual. Em forma

matricial tem-se

$$\mathbf{p}^T = \mathbf{g}L_{12}, \quad (2.2)$$

sendo $\mathbf{g} = [g_1 \ g_2]$ e $L_{12} = [l_1 \ l_2]^T$. Caso L_{12}^{-1} exista, os ganhos podem ser calculados por

$$\mathbf{g} = \mathbf{p}^T L_{12}^{-1} = [p_1 \ p_2] \begin{bmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{bmatrix}^{-1}. \quad (2.3)$$

Para bases ortonormais entre os alto-falante é necessário normalizar os fatores de ganho. A derivação pormenorizada pode ser consultada em [53]. A derivação pode utilizar triângulos e o sistema composto por três dimensões. A Figura 2.5 apresenta um exemplo de distribuição de fontes sonoras para virtualização de uma fonte virtual P utilizando VBAP em três dimensões.

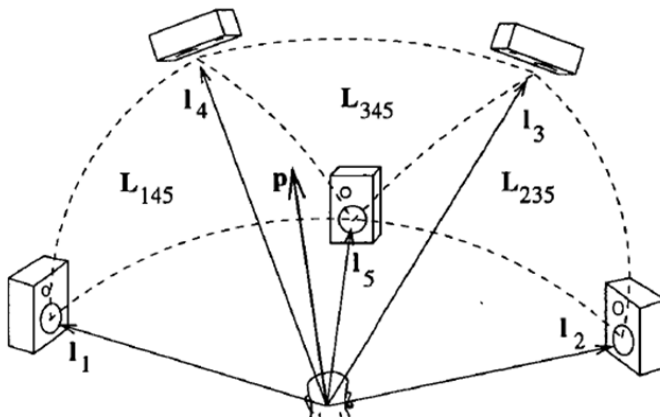


Figura 2.5.: Esquema representativo do posicionamento dos alto-falantes na técnica VBAP.

Fonte: Adaptado de [53].

Alguns fatores colaboram para que métodos baseados em Panorama de Amplitude sejam bastante utilizados em aplicações de áudio virtual, tais como o baixo custo computacional, a flexibilidade

do posicionamento dos alto-falantes e o equacionamento matricial que reduz o custo computacional.

2.4.2 Sistema Vetorial de Panorama de Intensidade (VBIP)

A natureza do sinal de áudio assumida na técnica VBAP é de banda larga desconsiderando fatores dependentes da frequência, em uma combinação linear de fontes sonoras coerentes. Devido a isso, o ajuste dos ganhos do sinal pode trazer problemas em sinais com menores comprimentos de onda devido a interferências causadas pela cabeça do ouvinte [55]. No método VBIP a derivação do panorama é a mesma do VBAP, com exceção da normalização no cálculo dos ganhos que considera que a combinação linear das fontes sonoras é uma soma de sinais incoerentes [56].

Os fatores de ganho, considerando os ângulos azimutais dos alto-falantes como θ_1 e θ_2 , são calculados por [55, 56]

$$g^2 = \mathbf{p}^T \mathbf{L}_{12}^{-1} = [p_1 p_2] \begin{bmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \end{bmatrix}^{-1}, \quad (2.4)$$

a partir $g_1^2 + g_2^2 = 1$, os fatores de ganho g_1 e g_2 são obtidos por meio de

$$g_1 = \begin{cases} \frac{\sqrt{g_1^2}}{\sqrt{g_1^2 + g_2^2}} \\ 1 \text{ se } g_2 \leq 0 \\ 0 \text{ se } g_1 \leq 0 \end{cases} \quad (2.5.a)$$

e

$$g_2 = \begin{cases} \frac{\sqrt{g_2^2}}{\sqrt{g_1^2 + g_2^2}} \\ 1 \text{ se } g_1 \leq 0 \\ 0 \text{ se } g_2 \leq 0 \end{cases}. \quad (2.5.b)$$

3 Sistema para ensaios de localização sonora

Avaliações do aproveitamento auditivo e de percepção ou inteligibilidade da fala podem ser mais acuradas com a inclusão de testes que considerem a distribuição espacial das fontes sonoras e de sons ou cenas acústicas mais realísticas. Assim, é proposto neste trabalho um sistema para a virtualização de fontes sonoras a partir das técnicas VBAP e VBIP utilizando uma configuração octogonal de alto-falantes. A utilização do sistema de virtualização oferece vantagem pela redução de custos e de adaptação às condições clínicas atuais tendo em vista a necessidade de realizar testes mais confiáveis sem a necessidade de um grande número de alto-falantes.

Após da validação do sistema, realizada por meio de uma medição objetiva do erro de localização e por meio de validação por meio de ensaios subjetivos com pessoas normo-ouvintes, observando a resolução obtida, é possível utilizar esse sistema para avaliações de localização de fontes sonoras por pessoas com perda auditiva e usuários de implante coclear ou AASI e futuramente aprimorar o protocolo de avaliação incluindo diversos outros testes. Avaliações utilizando esse sistema podem servir para aprimorar o protocolo de avaliação utilizado para indicação ao IC e ao AASI, bem como servir de instrumento auxiliar no acompanhamento e tratamento de usuários de AASI e IC.

As etapas de realização desta pesquisa foram:

- Implementação de algoritmos de áudio espacial descritos na literatura;
- Implementação de modelo de localização espacial de fontes sonoras;
- Realização de ensaios subjetivos;
- Realização de ensaios objetivos;
- Avaliação de resultados.

3.1 Sistema e ambiente de ensaios

Um arranjo de alto-falantes foi disposto em um formato octogonal em torno do centro da câmara semi-anecoica da LVA. O raio do centro até a base dos alto-falantes foi de 1,27 metros, distância escolhida para tornar-se comparável com o ambiente de audição crítica que está sendo projetado [57] para futuros ensaios no LVA.

Um arco de aço circular foi posicionado em frente aos alto-falantes e uma camada dupla de tecido não tecido (TNT), foi suspensa no arco para que a posição dos alto-falantes não fosse visível para os sujeitos participantes do ensaio com o objetivo de evitar viés de localização visual (quando o sujeito tende a apontar a posição do alto-falante ao invés da posição da fonte virtual induzido pela imagem). A Figura 3.1 apresenta uma imagem da montagem do sistema.



Figura 3.1.: Imagem da montagem do sistema na câmara semianecoica.

As técnicas de virtualização de fontes sonoras utilizadas foram de panorama (VBAP e VBIP) utilizando algoritmos da implementação de Archontis [58]. A Figura 3.2 apresenta o fluxograma das duas técnicas utilizadas.

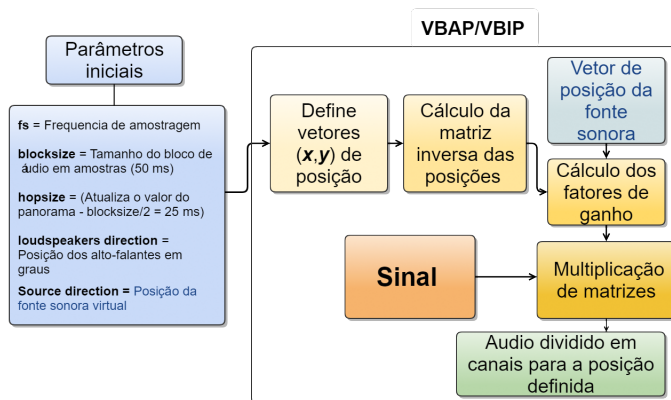


Figura 3.2.: Fluxograma processamento das técnicas VBAP e VBIP.

Diversos ensaios e experimentos podem ser realizados utilizando esse sistema. Para tal, o ambiente de escuta crítica¹ deve ter a mínima influência no campo sonoro. A caracterização acústica do ambiente, ao referir-se ao sistema de áudio, foi realizada conforme a recomendação ITU-R. BS 1116-3 *Methods for the subjective assessment of small impairments in audio systems* [34], sendo essa a mais reconhecida recomendação para ambientes críticos de escuta [59, 60], ela define além das propriedades da sala, as especificações e parâmetros de configuração dos alto-falantes.

O ruído de fundo do ambiente de ensaios foi gravado utilizando um microfone G.R.A.S de 1" modelo 40AF (*datasheet* no Anexo H), calibrado indiretamente com um calibrador B&K modelo 4231 (*datasheet* no Anexo G) e o sistema de gravação HEAD Recorder com o módulo de aquisição SQuadriga 2 (*datasheet* no Anexo D), ambos da HEAD Acoustics. Para o processamento de sinais o *software* Artemis Suite v7.0 foi utilizado.

Para aquisição da resposta impulsiva do ambiente e o cálculo do tempo de reverberação o mesmo microfone foi utilizado com o mesmo sistema de aquisição. O sinal gerado foi um *sweep* exponencial na faixa de frequência de 20 Hz a 20 kHz com duração de ≈ 21 segundos utilizando o ITA-Toolbox [61] no Matlab. Esse sinal foi convertido para o formato ".HDF" no *software* Artemis e posteriormente importado para o sistema do SQuadriga 2. A partir de então utilizou-se a função *play and rec* e as gravações foram armazenadas.

No passo seguinte foram importadas as gravações no *software* Matlab e as funções de transferência entre o sinal de excitação e o sinal adquirido foram obtidas no domínio da frequência e transformados para o domínio do tempo por meio da Transformada Rápida de Fourier Inversa (iFFT) para obter assim as respostas impulsivas (RIs) da sala. Assim, o tempo de reverberação (T20) foi calculado utilizando o módulo *roomacoustics* do ITA-Toolbox [62] a

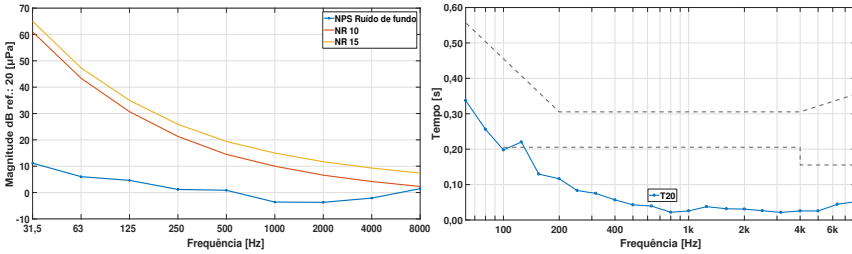
¹ Ambiente de escuta crítica é um termo utilizado para o ambiente acústica e visualmente preparado para avaliações subjetivas visando a mínima influência no ensaio.

partir das RI. O método configurado para o cálculo da curva de decaimento “*edc method*” foi o “*noCut*”².

Os ensaios foram realizados na câmara semianecoica do LVA. A utilização da câmara semianecoica do LVA é justificada por ser um ambiente com características acústicas controladas e pela disponibilidade da estrutura. Nele é possível quantificar os erros associados à utilização do *setup* e compará-lo futuramente, em termos de resposta, com as características do ambiente projetado para abrigar os experimentos posteriores.

Requisitos como o NPS do ruído de fundo e o tempo de reverberação (TR) foram observados e são mostrados na Figura 3.3. Contudo o TR ficou abaixo do valor mínimo requisitado pela norma. Visto que o ambiente tem um campo acústico extremamente controlado é possível utilizá-lo como referência de melhor condição de escuta (maior relação sinal-ruído) para seres humanos em ensaios de localização de sinais no silêncio no plano horizontal, ainda que fora dos limites da recomendação ITU-R BS. 1116-3 [34]. Ambientes dentro dos limites de reverberação (relação sinal-ruído um pouco menor) podem melhorar a resposta de sistemas com fontes virtuais levando a resultados ainda melhores dos que foram obtidos neste trabalho, dado que a reverberação pode auxiliar na naturalidade do campo sonoro para o sujeito.

² Para fornecer a resposta ao impulso integrada com compensação do ruído de fundo, a resposta impulsiva é truncada no tempo de interseção e é aplicada uma correção para truncamento no cálculo da curva de decaimento de energia, em inglês *Energy Decay Curve* (EDC) e o vetor é completado com zeros (*zero padding*) no método *noCut* [62].



(a) NPS do ruído de fundo em bandas de oitava câmara semianecoica LVA (linha azul) curva NR 15 (linha amarela) curva NR 10 (linha laranja). (b) Tempo de reverberação T20 da câmara semianecoica LVA (linha azul) limites estipulados na recomendação ITU-R 1116-3 (linhas tracejadas).

Figura 3.3.: NPS do ruído de fundo e Tempo de reverberação T20 na câmara semianecoica do LVA-UFSC frente aos requisitos da norma ITU-R BS. 1116-3 [34].

3.1.1 Reflexões

Para garantir que não haja influência do ambiente a Recomendação ITU-R 1116-3 [34] determina que a magnitude das primeiras reflexões devem estar ao menos 10 dB abaixo da magnitude do som direto $\Delta NPS \geq 10$ dB. O Apêndice A apresenta os gráficos das respostas impulsivas no domínio do tempo, normalizadas e em com a amplitude em escala logarítmica, para facilitar a visualização da diferença de NPS do som direto e do NPS das primeiras reflexões. As diferenças do NPS determinados no ambiente de ensaios deste trabalho atendem a esse quesito e são sumarizadas e apresentadas na Tabela 3.1.

Tabela 3.1.: Diferenças de amplitude entre som direto e primeiras reflexões para as oito posições de alto-falantes.

N°	1	2	3	4	5	6	7	8
Posição	0°	45°	90°	135°	180°	225°	270°	315°
ΔNPS [dB]	14,64	14,70	14,37	14,35	14,35	16,23	14,57	16,28

O Apêndice B apresenta a resposta ao impulso no domínio do tempo para cada alto-falante, medida na posição de audição dos

avaliadores com um microfone no centro do octógono a 1,4 m de altura com a sala configurada na forma em que foi utilizada no ensaio (incluindo mobiliário), conforme especificação da ITU-R BS. 1116-3 [34].

3.1.2 Resposta em frequência da sala

As curvas de resposta em frequência da sala, obtidas de acordo com a ITU-R BS. 1116-3 [34] para a faixa de frequência entre 50 Hz e 16 kHz no ponto de audição estão dentro dos limites estipulados e podem ser visualizadas na Figura 3.4.

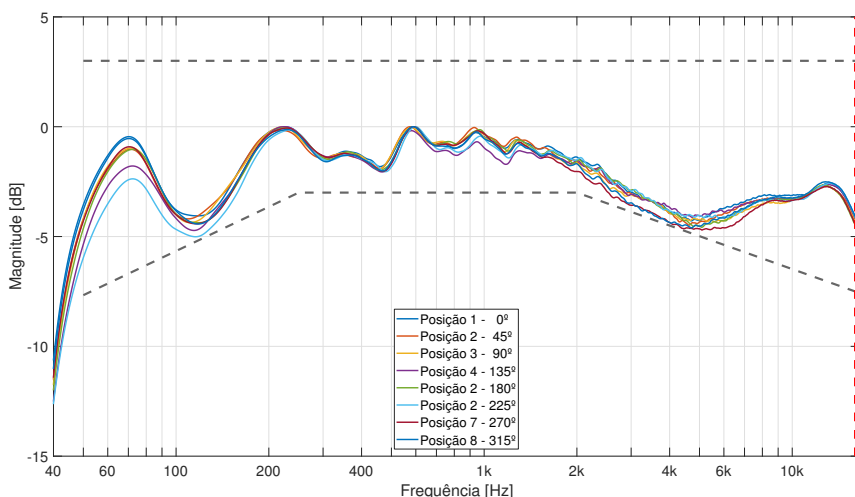


Figura 3.4.: Curvas de resposta em frequência para as 8 posições de excitação no ponto de audição. Linhas tracejadas representam os limites de tolerância conforme ITU-R BS. 1116-3 [34].

3.2 Dispositivos de reprodução

De acordo com a Recomendação ITU-R .1116-3 [34] os dispositivos de reprodução são denominados “alto-falante do monitor de referência”, que significa equipamento de audição de estúdio de alta qualidade, que inclui uma unidade integrada de sistemas de alto-falantes em caixa especificamente dimensionada, combinada com

sonora equalização especial, amplificadores de potência de alta qualidade e redes de cruzamento apropriadas. Assim o termo alto-falante aqui aplica-se ao conjunto de caixa, alto-falante e sistema de amplificação.

Os dispositivos de reprodução utilizados Genelec 8030C, dastheet no Anexo A, atendem às especificações eletroacústicas requeridas pela ITU-R BS. 1116-3 [34].

3.3 Sinais utilizados na avaliação

Para os ensaios de avaliação subjetiva de localização sonora foram escolhidos três sinais distintos utilizados na literatura [35–38], sendo:

- Ruído branco contínuo (WN^3);
- Ruído branco intermitente (WN_i);
- Sinal de fala.

Os sinais escolhidos têm foco em pistas auditivas diferentes, para o sinal de ruído branco contínuo o foco é a pista auditiva de diferença de nível entre as orelhas. Já os sinais de fala e ruído branco intermitente foram escolhidos para avaliar prioritariamente a pista acústica de diferença de tempo de chegada entre as orelhas. Para uma representação do que os sujeitos reconhecem como fala no dia a dia, uma voz feminina foi escolhida, enquanto o ruído branco intermitente é comparável em termos de espectro ao ruído branco contínuo. Todos os sinais foram filtrados mantendo o conteúdo espectral entre 100 Hz e 20.000 Hz. O tempo de ataque e *release* adicionado aos sinais foi de 0,2 segundos utilizando uma janela Hanning. Tal operação (*fade*) é realizada para procurar manter a integridade dos alto-falantes.

Ambos os sinais de ruído branco foram gerados com o AM-Toolbox v0.9.8 [63] e posteriormente os sinais de fala e ruído branco intermitente foram ajustados para o mesmo *loudness* do ruído

³ Abreviação para ruído branco *White Noise* do inglês, mais comumente utilizado.

branco contínuo utilizando a aplicação de psicoacústica do ITA-Toolbox

3.4 Nível de pressão sonora de reprodução

O nível de pressão sonora de preferência para os ensaios de localização realizados é denominado como nível de audição de referência pela ITU-R BS. 1116-3 [34]. O valor foi estipulado em $65 \pm 0,25$ dB(A). O ajuste desse nível de pressão sonora de reprodução entre os 8 alto-falantes foi realizado utilizando ruído rosa e tempo de integração *slow* (1 segundo).

Os sinais foram ajustados por meio de um filtro digital implementado para corrigir a função de resposta em frequência utilizando uma inversão da magnitude das curvas apresentadas na Figura 3.4. Para o ajuste utilizou-se a FRF invertida de cada alto-falante nos 22° imediatamente anteriores e nos 22° posteriores ao ângulo de referência do alto-falante. Exemplo: FRF do alto falante número 2 (posição de 45°) foi invertida e utilizada para filtrar sinais do ângulo 23° ao 67° (ver Figura 3.5).

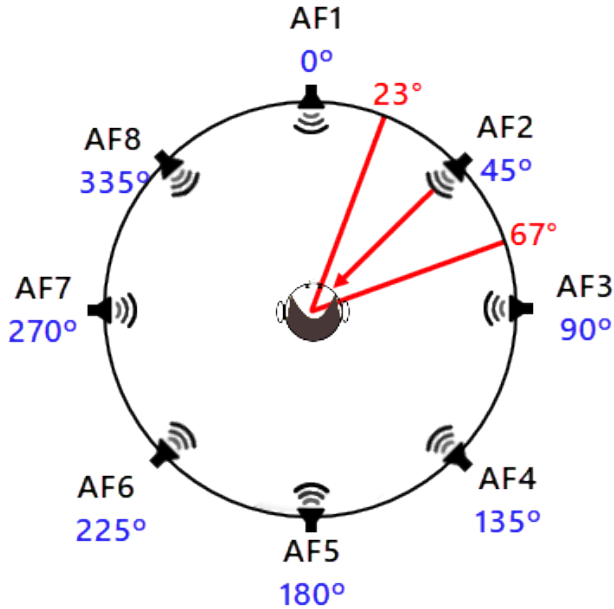
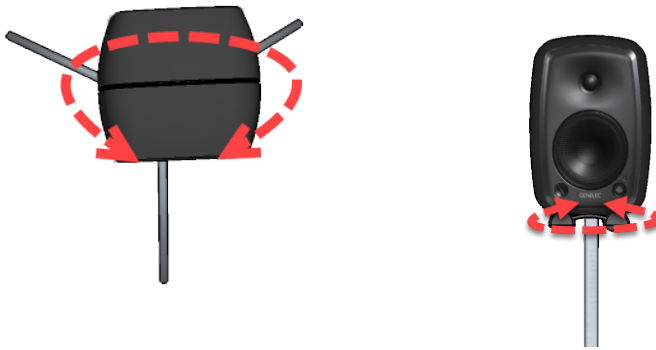


Figura 3.5.: Representação do posicionamento dos alto-falantes e numeração relativa ao ângulo (detalhe de segmentação relativa a inversão do filtro utilizando FRFs).

Para posicionar e alinhar os alto falantes no espaço, o simulador de tronco e cabeça HMS III (*datasheet* no Anexo B) foi utilizado para o ajuste fino. As entradas dos canais auditivos (CAs) do simulador foram posicionadas na altura do centro acústico dos alto-falantes em 1,40 m. Com os alto falantes posicionados na altura (1,40 m) e distância (1,27 m) correta do centro do octógono. O ajuste lateral foi realizado rotacionando os alto-falantes em torno do seu eixo vertical (ver esquema na Figura 3.6) e seguindo o procedimento descrito a seguir.



(a) Vista superior.

(b) Vista frontal.

Figura 3.6.: Ajuste lateral (rotação) dos alto-falantes.

1. Um tom puro na frequência de 250 Hz foi reproduzido pelo alto-falante posicionado em 0 graus (Alto-falante número 1: AF1);
2. O sinal capturado nos dois microfones do simulador de tronco e cabeça com o *software* Head Recorder v7.0 apresentando 1 comprimento de onda foi obtido;
3. Centralizou-se o simulador para obter o mesmo nível de pressão sonora e tempo de chegada nos dois microfones (orelhas);
4. O simulador foi mantido na posição e o sinal então foi reproduzido pelo alto-falante oposto (180° alto-falante número 5);
5. O alto-falante foi ajustado (rotacionado em torno do seu eixo vertical) para obter-se o mesmo nível de pressão sonora e tempo de chegada nos sinais das duas orelhas do simulador;
6. O procedimento foi repetido para os outros três pares de alto-falantes (AF2-AF6; AF3-AF7 e AF4-AF8), sempre ajustando a posição do simulador.

4 Avaliação do sistema

Para analisar a implementação do sistema será avaliado o erro de localização utilizando dois métodos: ensaios subjetivos¹ com normo-ouvintes² (avaliação do Erro de Localização Perceptivo, ELP Figura 4.1 (a)) e avaliação objetiva por meio de um modelo de localização de fontes sonoras que utiliza a diferença no tempo de chegada dos sinais entre os sensores de um arranjo (ver Figura 4.1 (b)), conforme descrito na literatura.

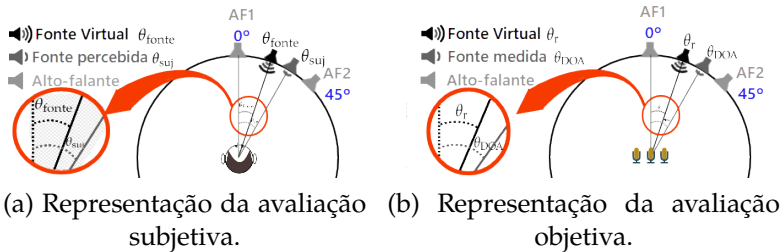


Figura 4.1.: Representações dos métodos propostos para avaliação do sistema.

As etapas para realização da avaliação do sistema foram:

- Preparação da avaliação subjetiva;
- Preparação de algoritmo utilizado nas avaliações subjetivas e análise de resultados;

¹ Ensaios subjetivos utilizam métricas objetivas para avaliar a percepção subjetiva dos sujeitos à alguma variável [64].

² Como critério para classificação de normo-ouvintes a pesquisa considerou pessoas sem reclamações de deficiência auditiva prévia.

- Execução da avaliação subjetiva;
- Preparação da análise objetiva, equipamentos e definição de algoritmos para avaliação;
- Preparação de algoritmo de localização espacial de fontes sonoras;
- Gravação dos sinais em diferentes arranjos de microfones;
- Avaliação de resultados.

Os experimentos objetivos (gravações) aqui apresentados foram realizados após a realização dos ensaios subjetivos finais³ visando à sustentação das características do ambiente de ensaios. Assim é importante notar que todo o aparato de gravação dos sinais utilizou a mesma configuração do *setup* de fontes sonoras utilizado na avaliação subjetiva. A ordem de apresentação dos ensaios e seus resultados (objetivos precedendo subjetivos) foi escolhida em função dos resultados, nos quais os resultados objetivos, apresentando valores menores de erro, qualificam os resultados subjetivos como será apresentado. Um esquema do trabalho pode ser visto na Figura 4.2.

³ O ensaio subjetivo teve uma fase piloto na qual foram avaliadas as modificações necessárias para aumentar a fidedignidade do ensaio como instrumento de medida.

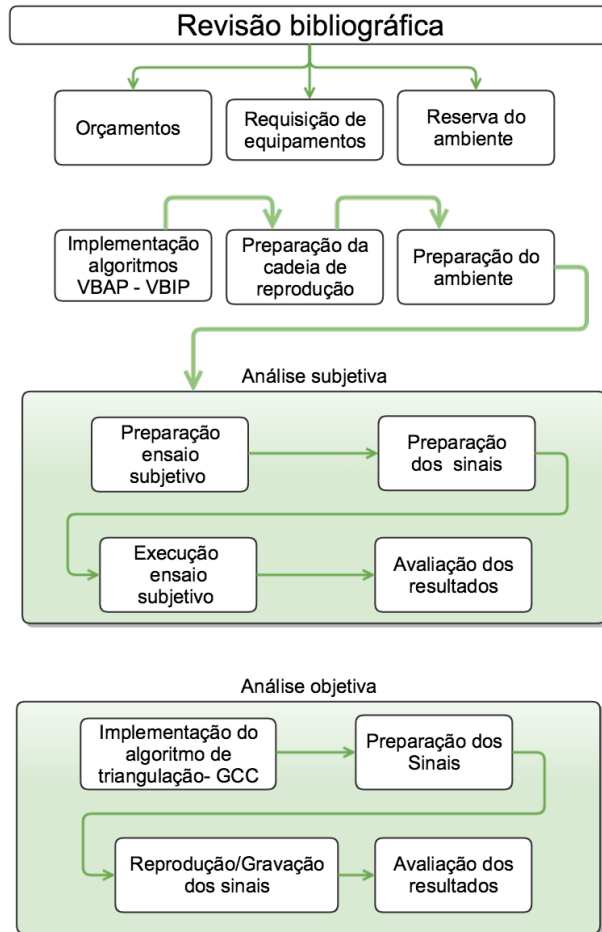


Figura 4.2.: Fluxograma: etapas do trabalho.

4.1 Avaliação objetiva

Nesta seção, o experimento projetado para obtenção da medida objetiva de Erro de Localização (EL) é apresentado. A obtenção de uma medida objetiva permite comparar esse dado e avaliar a exatidão na habilidade de localização sonora dos indivíduos testados posteriormente nos ensaios subjetivos.

4.1.1 Descrição do experimento

O sistema foi testado gravando os sinais computados com as técnicas VBAP e VBIP. O experimento consiste em utilizar com a técnica de triangulação a partir do tempo de chegada do som entre o arranjo de três microfones [47]. Dois tipos de posicionamento dos sensores do arranjo foram utilizadas para gravar os sinais, um em linha e outro em formato triangular. A representação do posicionamento dos sensores dos arranjos é apresentada nas Figuras 4.3 (a) e 4.3 (b).

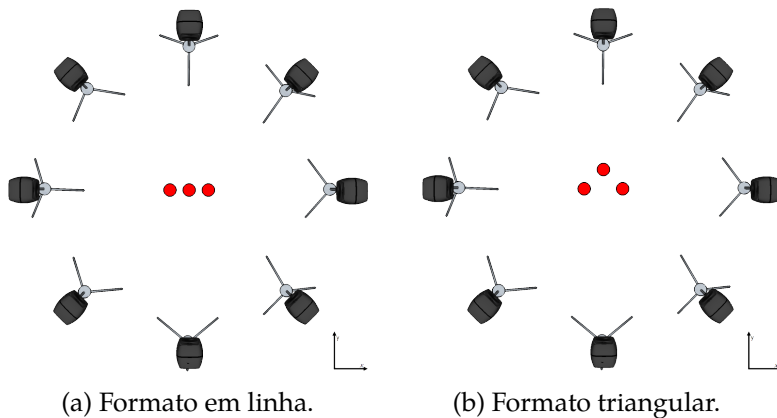


Figura 4.3.: Representação do arranjo de microfones para avaliação objetiva.

Foram utilizados 3 microfones omnidirecionais de $1/2''$ com equalização de campo difuso (*Datasheets* podem ser visualizados no Anexo F), calibrados indiretamente. O sinal reproduzido foi um ruído branco filtrado entre 100 Hz e 20 kHz d e 12 segundos de duração. O sinal foi reproduzido para cada ângulo do sistema (0° ao 359°).

A cadeia de medição é composta, além desses microfones, por um conversor digital analógico (DA) RME HDSP 9256 (*Datasheet* apresentado no Anexo C), 8 alto-falantes ativos - monitores de referência Genelec 8030C (*Datasheet* apresentado no Anexo A) e um conversor analógico digital (AD) SQuadriga II - 3320 (*Datasheet*

apresentado no Anexo D).

O *software* Head Recorder foi utilizado para aquisição de sinais. Os sinais reproduzidos nas fones sonoras foram gerados utilizando o *software* Matlab e reproduzidos com por meio do ITA-Toolbox [61]. A Figura 4.4 ilustra a cadeia de medição e a Figura 4.5 fotos da montagem.

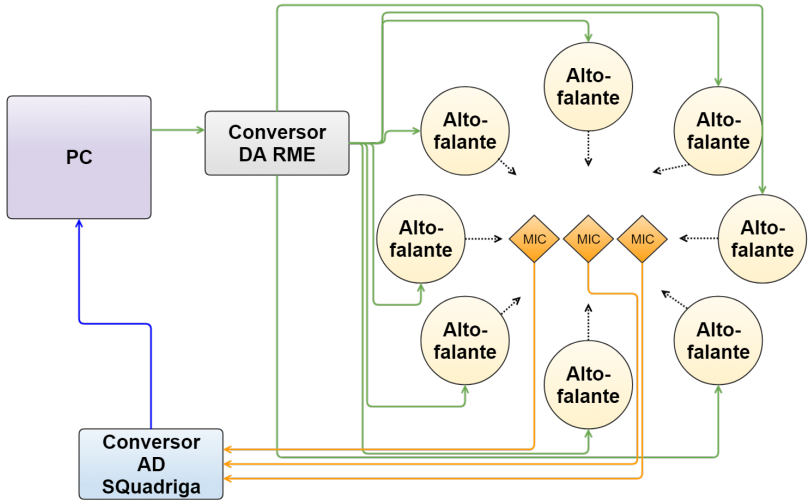
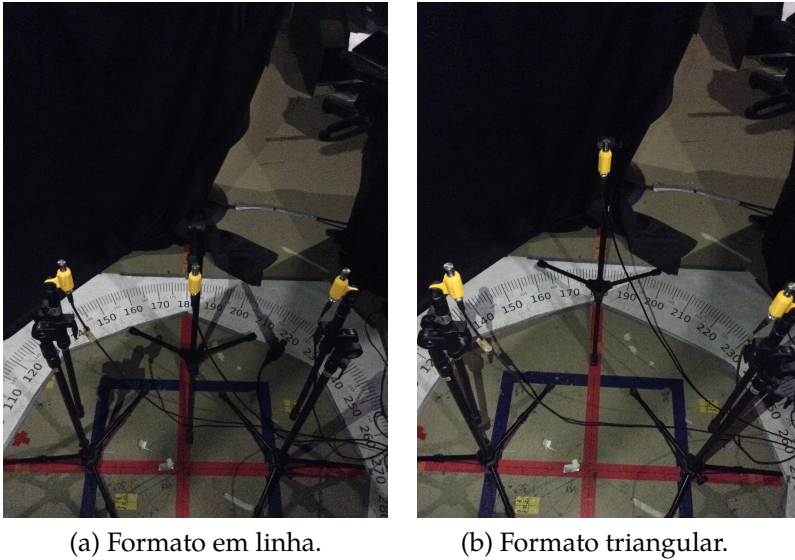


Figura 4.4.: Cadeia de medição utilizada para aquisição de sinais com três sensores.



(a) Formato em linha.

(b) Formato triangular.

Figura 4.5.: Arranjos de microfones posicionados no local da medição (câmara semianecoica).

A Figura 4.6 apresenta uma visualização para o arranjo de três sensores fixos (M_i , com $i = 1, 2, 3$) posicionados no plano x - y ,

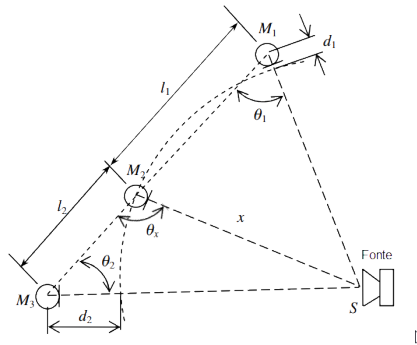


Figura 4.6.: Arranjo de três microfones em linha.
Fonte: Adaptado de [47].

A diferença de tempo de chegada dos sinais entre os sensores M_1 e M_2 , denotada como $\tau_{1,2}$, é utilizada no cálculo da distância d_1 entre o sensor M_1 e a fonte sonora S . A diferença $\tau_{3,2}$, por sua vez

é equivalente à diferença de tempo de chegada dos sinais entre os sensores M_3 e M_2 , e é utilizada para calcular a distância d_2 [47, 65]:

$$d_1 = c\tau_{12}, \quad (4.1.a)$$

$$d_2 = c\tau_{32}. \quad (4.1.b)$$

Logo, conhecendo a distância x entre a fonte sonora S e o sensor M_2 é possível calcular os ângulos θ_1 , θ_2 e θ_x , tem-se então

$$\theta_1 = \cos^{-1} \frac{l_1^2 + (x + d_1)^2 - x^2}{2l_1(x + d_1)}, \quad (4.2.a)$$

$$\theta_2 = \cos^{-1} \frac{l_2^2 + (x + d_2)^2 - x^2}{2l_2(x + d_2)} \quad (4.2.b)$$

e

$$\theta_x = 90 - \frac{\theta_1}{2} + \frac{\theta_2}{2}. \quad (4.2.c)$$

4.1.2 Resultados experimentais da medição objetiva

Nesta seção são apresentados os resultados obtidos com o experimento de localização que utilizou a configuração descrita na Seção 4.1.1.

4.1.2.1 Resultados arranjo em linha

Para a gravação dos sinais utilizando a formação em linha o arranjo utilizou espaçamento de 0,165 m entre os sensores. O microfone de referência foi posicionado no centro do plano (ver Figura 4.3 (a)) na mesma altura em que estava posicionada o centro acústico dos alto-falantes (1,40 m). Foi utilizado o *toolbox Phased Array Antena* do Matlab para computar o ângulo por meio das gravações (θ_{DOA}) do arranjo em linha.

A Figura 4.7 apresenta a relação direta entre os ângulos de incidência do sinal computado θ_{DOA} e o ângulo relativo ao sinal emitido θ_r para o ensaio com os sinais computados por VBAP⁴.

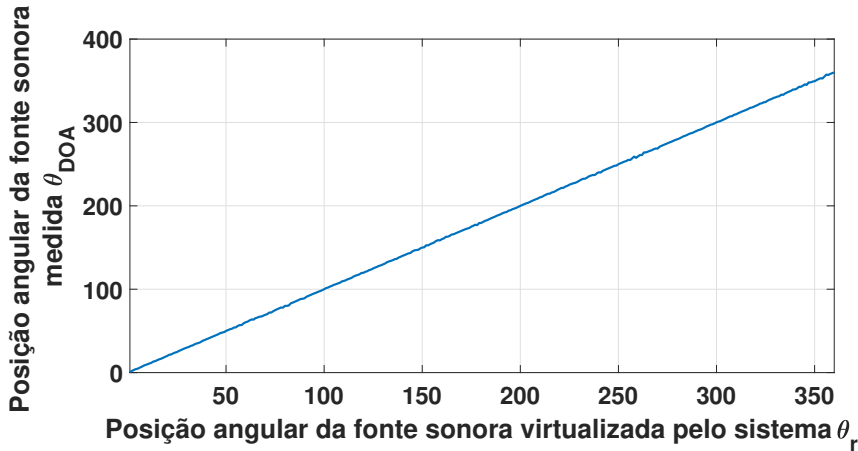


Figura 4.7.: Ângulo calculado a partir da medição do tempo de chegada na configuração do arranjo em linha e ângulo da fonte virtualizada - VBAP.

A Figura 4.8 apresenta o erro obtido entre o ângulo da fonte sonora simulada pela técnica VBAP por meio do *setup* e o ângulo estimado pelo *array* linear de sensores.

⁴ Para uma melhor fluência na leitura, pela similaridade, apenas o gráfico da técnica VBAP relacionando as posições angulares das fontes sonoras virtualizadas e medidas é mostrado.

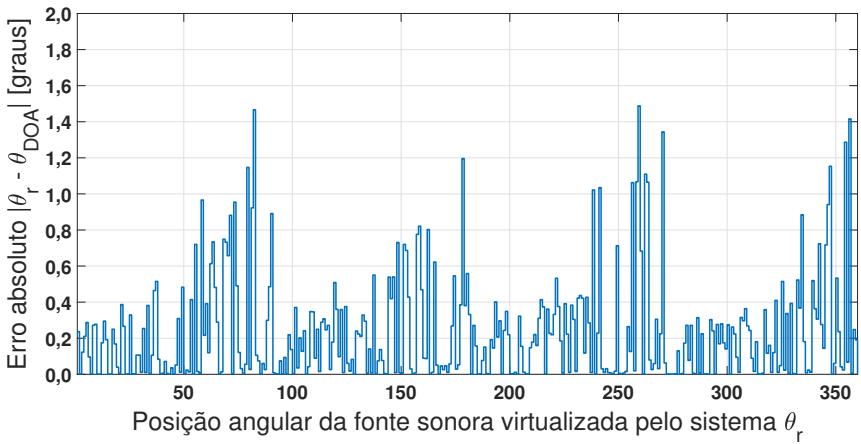


Figura 4.8.: Erro calculado a partir da medição do tempo de chegada na configuração do arranjo em linha - VBAP.

A média de erro entre todos os ângulos foi de 0,234 graus, com desvio padrão de 0,001 graus.

O erro por ângulo, para as gravações dos sinais utilizando o sistema com a técnica VBIP, pode ser visualizado na Figura 4.9.

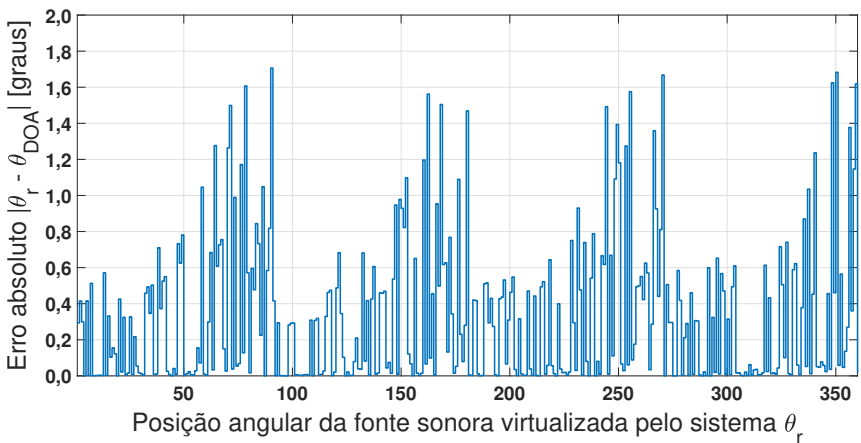


Figura 4.9.: Erro calculado a partir da medição do tempo de chegada na configuração do arranjo em linha - VBIP.

O erro médio entre todos os ângulos foi de 0,2400 graus e o desvio padrão de 0,0007 graus. .

Ambas as técnicas, se avaliadas com o arranjo posicionado em linha, tiveram erros mais elevados em torno dos ângulos 0° , 90° , 180° e 270° . O posicionamento dos sensores com a altura exata e a distância exata pode ter sido uma fonte de erro, dada a dificuldade de posicionamento no espaço por meio dos tripés disponíveis. Apesar dos erros mencionados a média dos erros e o valor esperado (média com apresentação da incerteza) estão de acordo com resultados reportados na literatura para sistemas com fontes reais [47]. A técnica de arranjo em linha pode apresentar problemas por dificuldade em avaliar ângulos normais e perpendiculares ao eixo x por haverem distâncias simétricas. Este fenômeno é explicado pela teoria de arranjos de sensores [28]. Na tentativa de investigar tal hipótese o arranjo em triângulo foi proposto e os resultados serão apresentados a seguir.

4.1.2.2 Resultados arranjo em triângulo

A formação triangular tem por objetivo diminuir o efeito de diminuição de foco colocando a posição de referência ($x = 0, y = 0$) fora da linha média do eixo x do plano.

Para a gravação dos sinais utilizando a formação em triangular o arranjo teve espaçamento de 0,165 m entre os microfones com o microfone de referência no centro do eixo x e deslocado 0,165 m no eixo y (ver Figura 4.3 (b)). O posicionamento da altura dos microfones foi em 1,40 m. Foi utilizado o *toolbox Phased Array Antena* para computar o ângulo por meio das gravações do arranjo em triângulo. A Figura 4.10 apresenta a relação entre os ângulos.

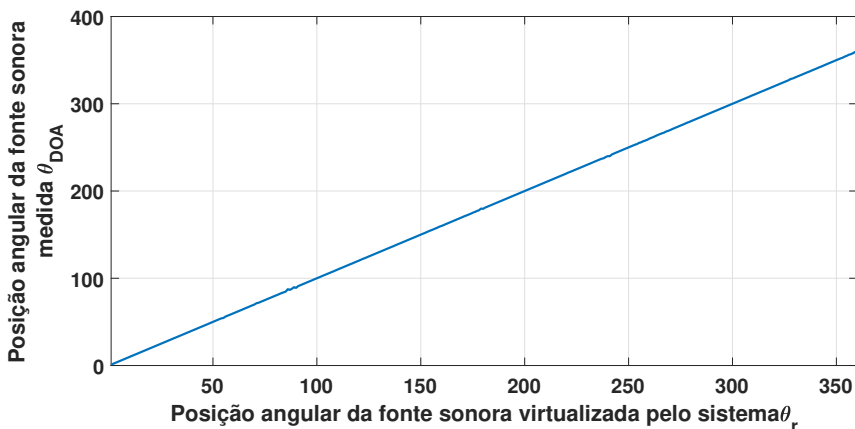


Figura 4.10.: Ângulo calculado a partir da medição do tempo de chegada na configuração do arranjo em triângulo e ângulo da fonte virtualizada - VBAP.

A Figura 4.11 apresenta o erro entre o ângulo relativo a posição da fonte sonora emitida pelo *setup* e o ângulo estimado por meio do procedimento descrito para a técnica VBAP.

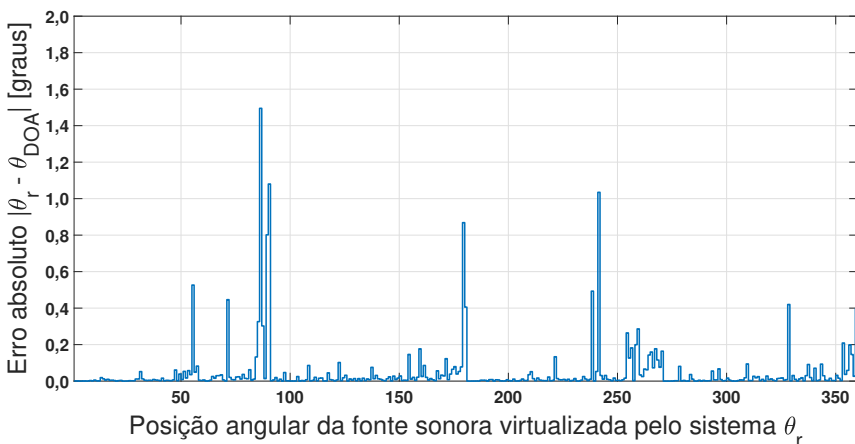


Figura 4.11.: Erro calculado a partir da medição do tempo de chegada na configuração do arranjo em triângulo - VBAP.

Estimou-se assim uma média de erro igual a 0,0459 graus e um desvio padrão igual a 0,0004 graus. Isso acontece pois há diferença de distância em x , criando um meio para ajudar o arranjo

a diferenciar os lados opostos (0° e 180°).

Para as gravações dos sinais computados com a técnica VBIP com o arranjo em triângulo a média foi 0,0972 e o desvio padrão 0,0005. O erro por ângulo pode ser visualizado na Figura 4.12.

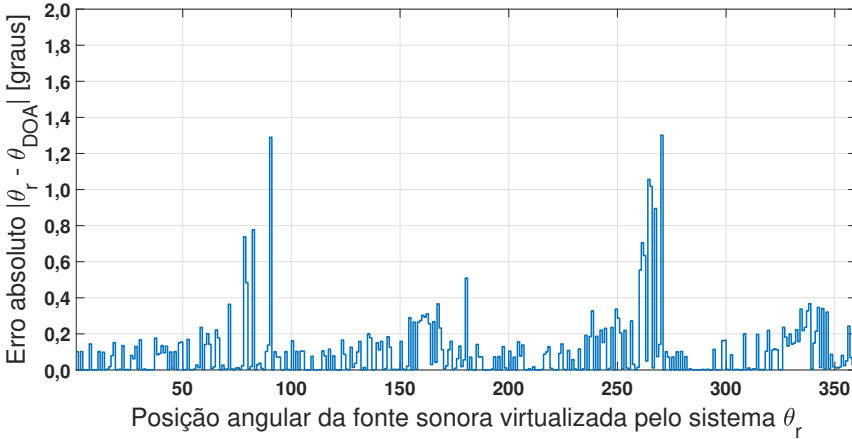


Figura 4.12.: Erro calculado a partir da medição do tempo de chegada na configuração do arranjo em triângulo - VBAP.

A Tabela 4.1 sumariza os resultados dos valores esperados com um fator de abrangência de aproximadamente 95% para as técnicas VBAP e VBIP de virtualização de fontes sonoras.

Tabela 4.1.: Valores esperados com fator de abrangência de $\approx 95\%$ $k_p = 2$ na medição objetiva.

$k_p = 2$		
Técnica	Linha [$^\circ$]	Triângulo [$^\circ$]
VBAP	$0,234 \pm 0,002$	$0,0459 \pm 0,0004$
VBIP	$0,240 \pm 0,002$	$0,0972 \pm 0,0005$

É possível notar nos resultados do arranjo triangular, tanto para sistema reproduzindo por meio do VBAP quanto do VBIP, que a magnitude dos erros diminui em relação aos erros aferidos com o *array* linear, inclusive nos ângulos 0° e 180° , indicando uma dificuldade no arranjo linear em identificar corretamente os ângulos devido as distâncias serem redundantes, o que é

corroborado na teoria de arranjos [28]. Em torno dos ângulos 90° e 270° ainda ocorrem erros maiores.

Apesar de existirem incertezas associadas à montagem do *setup* de medição, o erro obtido é coerente nos dois métodos (arranjos) de gravação e aponta para uma implementação adequada das técnicas de virtualização de fontes sonoras, da instrumentação e da montagem utilizada.

Além da medição do erro de localização utilizando gravações de sinais por meio de arranjos de microfones, outras gravações utilizando uma configuração semelhante foram realizadas. Nestas medições um simulador de tronco e cabeça (*Datasheet* apresentado no Anexo B) configurado para equalização de campo livre foi utilizado para gravar os sinais provenientes de todas as fontes reais e virtuais posicionadas nos 360 graus circundantes. Essa medição tem por objetivo fornecer dados que possam ser utilizados em futuras montagens do sistema, sejam no mesmo ambiente ou em ambientes distintos.

A Figura 4.13 apresenta a cadeia de medição e a imagem apresentada na Figura 4.14 mostra uma vista do torso posicionado para o experimento, sem o tecido para uma melhor visualização. Contudo todas as medições foram realizadas utilizando a cortina dupla de TNT a frente dos alto-falantes e uma cortina simples atrás.

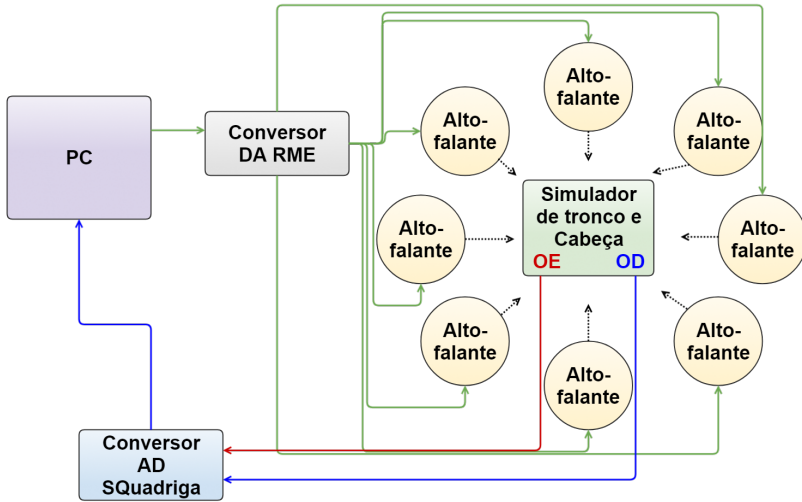


Figura 4.13.: Cadeia de medição utilizada para aquisição de sinais biauriculares.



Figura 4.14.: Simulador de tronco e cabeça posicionado no centro do octógono formado pelos alto falantes na câmara semi anecoica do LVA.

O sinal utilizado no experimento foi o ruído branco com 12 segundos de duração. Os resultados de ILD obtidos para os oito ângulos (ângulos centrais dos cones de 45°) nos quais os alto-falantes estão posicionados foram obtidos com o AM-Toolbox

v0.9.8 [63] e são apresentados na Figura 4.15.

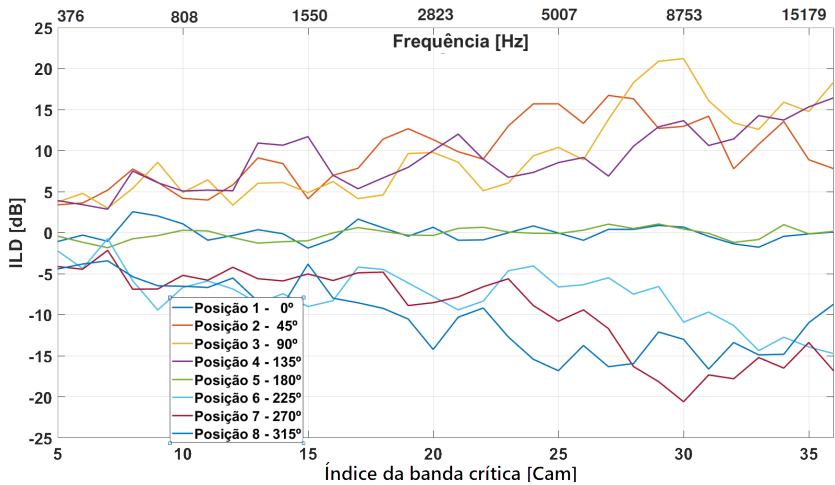


Figura 4.15.: ILDs obtida com simulador de tronco e cabeça para 8 posições.

As curvas de ILD apresentam um comportamento típico, sendo que em incidência frontal (0°) e traseira (180°) a diferença de nível de pressão sonora entre os sinais captados nas orelhas do simulador de tronco e cabeça são pequenas para toda a faixa de frequência. As curvas de ILD para os demais ângulos deslocados dessas posições apresentam valores maiores na medida que a frequência cresce, pois a cabeça atenua e reflete os componentes da onda sonora com comprimento de onda de tamanho menores ou equiparáveis à sua dimensão característica. Para 90° e 270° as maiores diferenças de nível são constatadas (lados opostos da cabeça artificial). Para frequências baixas, dado o comprimento de onda grande, a diferença de nível de pressão sonora entre as orelhas tende a ser menor (a cabeça se torna acusticamente transparente).

4.2 Avaliação subjetiva

O procedimento experimental foi realizado no Laboratório de Vibrações e Acústica da Universidade Federal de Santa Cata-

rina no período de 20 agosto de 2017 a 1 de setembro de 2017. Este estudo foi autorizado pelo comitê de ética em pesquisa CAAE: 56838816.7.0000.0121 parecer número 1.997.931.

Atualmente na prática clínica os testes⁵ para a avaliação auditiva utilizados como critério para indicações às tecnologias assistivas como implante coclear e ao AASI precisam de uma padronização além de maior acurácia.

A proposta do projeto IMPCoclear para o desenvolvimento de um protocolo de testes engloba a utilização de fontes sonoras virtuais nos ensaios de percepção e inteligibilidade da fala com distribuição espacial de fontes sonoras⁶.

Para tal o primeiro passo é a implementação e a mensuração da exatidão dos métodos de virtualização de fontes sonoras.

Nesta seção o ensaio subjetivo de localização de fontes sonoras será descrito, bem como a qualificação dos sujeitos, os resultados e as conclusões preliminares.

4.2.1 Ensaio subjetivo

A exatidão angular da virtualização de fontes sonoras no plano horizontal das técnicas VBAP e VBIP com reprodução por meio de 8 alto-falantes montados em formato octogonal na câmara semianecoica do LVA é o objeto de Ensaio em questão. Neste trabalho o foco é a validação do sistema com pessoas sem perda auditiva declarada.

Esta etapa é considerada parte da validação da discretização angular do sistema juntamente com a avaliação objetiva. A avaliação por meio de ensaios subjetivos procura quantificar a percepção do ângulo de incidência das fontes sonoras reais e virtualizadas e comparar os resultados aos obtidos nos ensaios objetivos e na literatura. A partir disso, ensaios de localização poderão ser aplicados em pessoas com ou sem deficiência auditiva valendo-se dos resultados obtidos no presente trabalho.

⁵ Testes também podem ser utilizados no acompanhamento do tratamento.

⁶ E também testes com e sem adição de ruídos mascaradores.

Uma fase de ensaios piloto foi realizada para o desenvolvimento da ferramenta de avaliação (ensaio subjetivo). A partir dos resultados e comentários dos sujeitos que participaram dessa etapa piloto o ensaio foi remodelado e aplicado aos demais sujeitos em sua versão final. O ensaio visou à conformidade com a recomendação ITU-R 1116-3 [34] (*Methods for the subjective assessment of small impairments in audio systems*).

4.2.2 Ensaio piloto

O primeiro passo para avaliar o protocolo de ensaio (no caso o instrumento de medida) foi uma rodada com 5 sujeitos que já haviam participado de outros ensaios subjetivos, tinham conhecimento na área de acústica e receberam treinamento para execução deste ensaio piloto. As impressões dos sujeitos foram coletadas e as modificações realizadas na estrutura do ensaio, em decorrência das observações feitas, são apresentadas.

4.2.2.1 Método e procedimento de ensaio piloto

O ensaio foi baseado na tarefa de localização de fontes sonoras no plano horizontal por um grupo de 5 indivíduos. As fontes sonoras virtuais foram criadas a partir de dois sistemas de virtualização de fontes sonoras, VBAP e VBIP.

Neste ensaio o sujeito informa verbalmente o ângulo do qual ele percebe a incidência dos sons⁷ (localização da fonte sonora) com o auxílio de um *laser* e uma marcação angular no piso do ambiente de ensaios, vide Figura 4.17. A interface de pesquisa é um questionário no qual o sujeito informa oralmente o valor do ângulo em que ele percebe a incidência do sinal sonoro ao avaliador.

Ao iniciar o ensaio o sujeito participante era posicionado no centro do círculo e tinha completa mobilidade horizontal girando no eixo da cadeira⁸ para identificação da posição da fonte sonora

⁷ Os sinais utilizados são descritos na (Seção 3.3).

⁸ Movimentos verticais da cabeça também foram permitidos, desde que o sujeito permanecesse sentado.

(ângulo de incidência), real ou virtual, com movimentos da cabeça e tronco. A orientação quanto a sua posição foi de que permanecesse com as costas no encosto. Um sinal de ruído branco com as mesmas características dos sinais utilizados no ensaio foi apresentado para verificação do conforto do sujeito em relação ao nível de pressão sonora e pôde ser ajustado conforme indicação do sujeito. O nível de pressão sonora indicado pela ITU-R BS. 1116-3 [34] é de $78,00 \pm 0,25$ dB, contudo não foi considerado como um nível confortável no ambiente anecoico e o ajuste permitido pela recomendação ITU-R BS. 1116-3 [34] foi realizado. O nível de reprodução foi mantido para todos os sujeitos em $65,00 \pm 0,25$ dB. Para cada apresentação foi solicitado aos sujeitos se gostariam de alterar o NPS, no entanto nenhum dos sujeitos solicitou alteração.

O sinal foi gerado no Matlab e reproduzido através do par de alto-falantes selecionado. A seleção dos ângulos nos quais foram criadas as fontes sonoras apresentados para cada sujeito foi aleatória. A primeira etapa no ensaio subjetivo piloto foi o treinamento do sujeito. A Figura 4.16 apresenta o fluxograma da fase de treinamento do ensaio subjetivo.

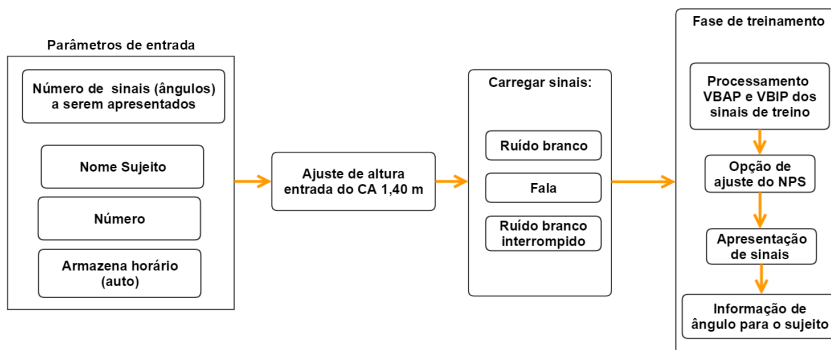


Figura 4.16.: Fluxograma da fase de treinamento do ensaio subjetivo aplicado.

Em relação à diferença de altura dos indivíduos que realizaram o ensaio, foi solicitado para que cada sujeito ajustasse a elevação da

cadeira para que a altura até a base da entrada do canal auditivo fosse igual a 1,40 metros, que foi a altura do centro acústico dos alto-falantes do arranjo.

Para as duas técnicas analisadas (VBAP e VBIP) foram escolhidos três sinais e para cada sinal foram geradas fontes sonoras (reais e virtuais) em 20 ângulos. No total 120 sinais (detalhados na Seção 3.3) foram apresentados a cada sujeito.

4.2.2.2 Instruções e treinamento

Ao iniciar o ensaio os sujeitos foram posicionados no centro do círculo em uma cadeira como mostrado na Figura 4.17. Os sujeitos foram orientados a utilizar a marcação em graus disposta no chão auxiliados pelo *laser* fixado na cadeira para apontar o ângulo percebido relativo à posição da fonte sonora. Foi esclarecido aos participantes que eles poderiam responder antes do sinal encerrar ou solicitar que o sinal fosse repetido.



Figura 4.17.: Sujeito posicionado para o ensaio de localização de fontes sonoras (dentro da câmara semianecoica).

O treinamento foi realizado individualmente com os jurados a fim de familiarizá-los com o ensaio e conseguir uma resposta com o mínimo viés acerca do entendimento das tarefas práticas do ensaio. As etapas do treinamento foram:

- Explicação da motivação da pesquisa e do ensaio;
- Inserção dos parâmetros de entrada;
- Posicionamento do sujeito;
- Sessão de dúvidas e ajuste do NPS;
- Treinamento indicando os ângulos ao sujeito;
- Treinamento solicitando os ângulos ao sujeito.

4.2.2.3 Impressões dos jurados

O ensaio piloto resultou nas seguintes observações:

Jurado 1 Observações:

- nenhum para repetir a apresentação do sinal pelo ângulo apresentado
- Faltaram números/marcações intermediárias de ângulos no piso da câmara
- Fala mais fácil de identificar
- Usou movimentos da cabeça com olhos fechados

Jurado 2 Observações:

- 11 pedidos para repetir a apresentação do sinal pelo ângulo apresentado
- Faltaram números/marcações intermediárias de ângulos no piso da câmara
- Laser deveria ir até ponto mais alto do tecido
- Usou movimentos da cabeça com olhos fechados
- Orientação pelas linhas do tecido

Jurado 3 Observações:

- 2 pedidos para repetir a apresentação do sinal pelo ângulo apresentado

- Movimento com a cabeça
- Mais iluminação no quarto quadrante seria melhor
- Usou movimentos da cabeça com olhos fechados

Jurado 4 Observações:

- 5 pedidos para repetir a apresentação do sinal pelo ângulo apresentado
- Usou movimento da cabeça com olhos fechados
- Sentiu-se tranquilo quanto ao número de apresentações
- Preferia que o conteúdo fala fosse diferente durante as apresentações
- Relatou que sentiu dificuldade em alguns locais por estar mal iluminado

Jurado 5 Observações:

- 6 pedidos para repetir a apresentação do sinal pelo ângulo apresentado
- Preferia que a fala fosse diferente durante as apresentações
- Teve mais dificuldade no segundo quadrante, não soube explicar o motivo.

4.2.2.4 Análise da influência da duração do ensaio nos resultados do ensaio subjetivo piloto

Para avaliar tendências de aprendizado ou cansaço ao longo das apresentações tendo em vista a influência da duração do ensaio foi proposto uma análise da mediana⁹ do erro absoluto para cada apresentação (ângulo) dos sujeitos.

A utilização da mediana como métrica é baseada no pequeno número de sujeitos que participaram no ensaio piloto, logo é necessário complementar a análise com o teste de Wilcoxon [66]. O resultado não apresentou tendência de aumento do erro como pode ser visualizado na Figura 4.18. Assim tanto o tamanho dos

⁹ Por haver apenas 5 valores a mediana é mais representativa.

sinais quanto o número de ângulos e de sinais apresentados foram mantidos.

O teste de Wilcoxon não rejeitou a hipótese nula (H_0) de que dados no vetor provêm de uma distribuição contínua e simétrica cuja mediana é zero no nível de significância de 5%. O resultado para o valor p no teste é de 0,75 demonstrando robustez no resultado do teste.

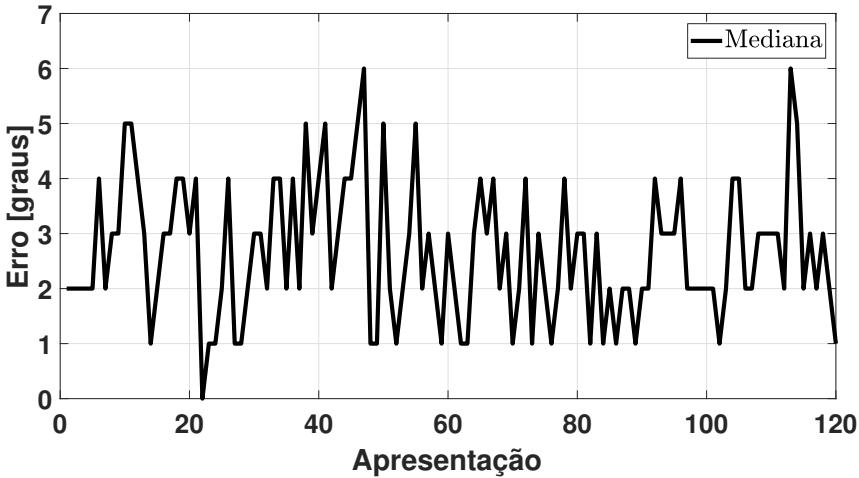


Figura 4.18.: Diferenças entre o ângulo apresentado e o ângulo percebido pelo sujeito na sequência de apresentação (Mediana ELP).

4.2.2.5 Modificações realizadas para o ensaio final

A partir da fase de ensaios piloto o início do protocolo de avaliação foi modificado. Adicionou-se à etapa de instrução a explicação da intenção do ensaio de utilizar o jurado para avaliar o sistema, deixando claro que o sistema estava sendo avaliado e não o jurado.

Números intermediários (5, 15, 25...) foram adicionados na marcação angular posicionada no piso da câmara semianecoica e aumentou-se a iluminação com *leds* direcionais possibilitando uma visualização mais ágil e com maior exatidão.

A marcação proveniente do nível¹⁰ com apontador *laser* (Figura 4.19) foi ajustada na cadeira para que a ponta do feixe de *laser* estivesse na mesma altura que o centro acústico dos alto-falantes.

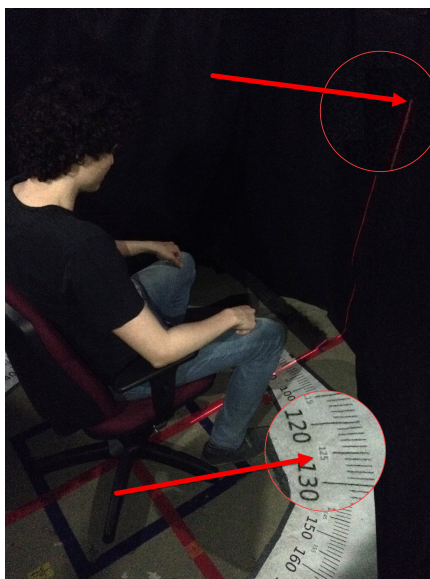


Figura 4.19.: Sujeito posicionado para o ensaio de localização de fontes sonoras, detalhes destacados: posição do *laser* e inclusão de valor intermediário na marcação do piso.

A parte traseira dos alto-falantes foi coberta com uma cortina de TNT para não dar pistas aos sujeitos acerca da posição das fontes reais (alto-falantes) ao entrar no ambiente de ensaios.

4.2.3 Ensaio subjetivo principal

A partir do dia 21 de agosto de 2017 até o dia 1 de setembro de 2017 foram realizados 30 novos ensaios subjetivos. Os ensaios foram realizados no mesmo local, utilizando a mesma configuração do *setup* do ensaio piloto, adicionando-se as modificações descritas na Seção 4.2.2.5. Participaram 6 mulheres e 24 homens com idade

¹⁰ O nível à *laser* proporciona um traço de *laser* que estende-se do chão ao tecido posicionado em frente aos alto-falantes.

entre 20 e 41 anos, sendo estudantes de mestrado, de doutorado e professores no Laboratório de Vibrações e Acústica (LVA) e também produtores de áudio. A duração média do ensaio foi de 27,6 minutos.

A Figura 4.20 mostra o fluxograma do ensaio subjetivo desde o sinal ao armazenamento do valor informado pelo sujeito. Os sujeitos informaram oralmente ao avaliador o ângulo correspondente à localização da fonte sonora. O pesquisador não conheceu o ângulo antes da resposta do sujeito ser computada.

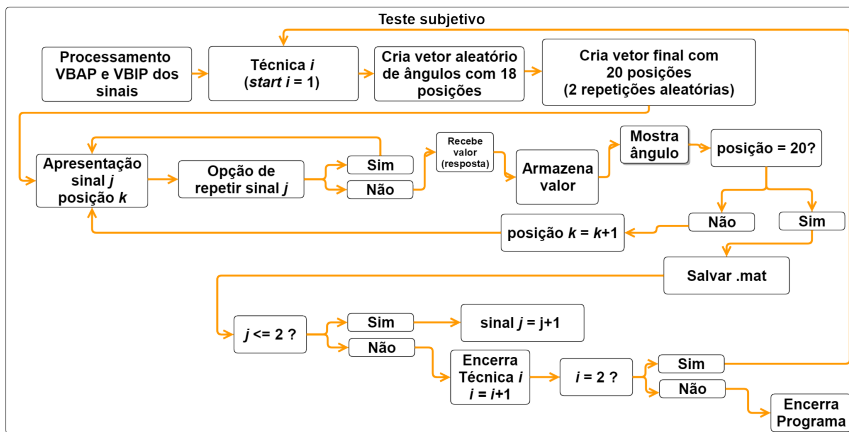


Figura 4.20.: Fluxograma do ensaio subjetivo aplicado.

O valor informado pelo sujeito é adicionado à sua matriz de respostas (ELP), o erro é calculado por meio da diferença entre o ângulo definido no programa θ_{fonte} e a resposta dada pelo sujeito θ_{suj} conforme

$$ELP(i, j, k) = \theta_{\text{fonte}} - \theta_{\text{suj}}, \quad (4.3)$$

na qual $i = 1, 2$, sendo o $i = 1$ para a técnica VBAP e $i = 2$ para VBIP; $j = 1, 2, 3$, sendo referente aos sinais utilizados WN, Fala e WN_i respectivamente, e $k = 1, \dots, 20$ referindo-se ao número da sequência de apresentação do ângulo.

4.2.3.1 Erro médio no tempo

O erro de localização médio (\overline{ELP}) para cada ângulo apresentado para os 30 sujeitos na sequência é calculado por meio da Média de Desvio Absoluto (MAD)¹¹, uma medida eficiente como medida de dispersão [20] sendo calculado como

$$MAD = \frac{1}{n} \sum_{i=1}^n |\theta_{\text{fonte}}(i) - \theta_{\text{suje}}(i)|, \quad (4.4)$$

no qual n é o número de sujeitos.

O MAD de todas as sequências de apresentação pode ser visualizado na Figura 4.21.

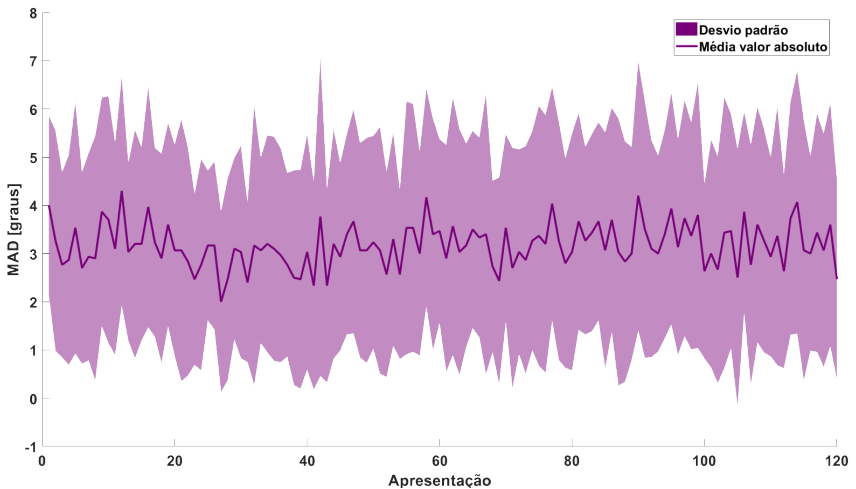


Figura 4.21.: Média de desvio absoluto em função do número da apresentação da fonte sonora durante o ensaio.

É possível notar que não há uma tendência de elevação ou diminuição do ELP ao longo do ensaio, o que seria um indicativo de aumento no cansaço ou aprendizado, respectivamente.

¹¹ Acrônimo do inglês *Mean Absolute Deviation*.

4.2.3.2 Análise de repetibilidade intra indivíduo

A análise de repetibilidade das respostas intra indivíduo foi realizada e casos de confusão frente-costas foram estipulados como critério de exclusão dos sujeitos, ou seja, caso o indivíduo localizasse a fonte sonora no lado oposto (erros de aproximadamente 180°) em qualquer apresentação, suas respostas seriam descartadas da análise.

Ainda, para a análise de repetibilidade, cada bateria de apresentação de sinais foi composta por 20 sinais, desses sinais 18 eram criados aleatoriamente e 2 sinais eram repetições. As repetições também foram executadas de forma aleatória na ordem de apresentação.

A medida de repetibilidade utilizada foi o erro quadrático médio calculado como

$$\text{EQM} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\theta_{i,\text{fonte}} - \theta_{i,\text{suj}})^2}, \quad (4.5)$$

na qual $\theta_{i,\text{fonte}}$ é o ângulo determinado para ser reproduzido pelo sistema e $\theta_{i,\text{suj}}$ é a resposta do indivíduo e $N = 4$ para cada bateria observando que são duas repetições.

O EQM expressa a variância do erro abrangendo erros aleatórios e erros sistemáticos [20], o que permite a análise da resolução espacial para a direção frontal de cada sujeito.

A repetibilidade por indivíduo, em cada técnica é apresentada nas Figuras 4.22, 4.23 e 4.24.

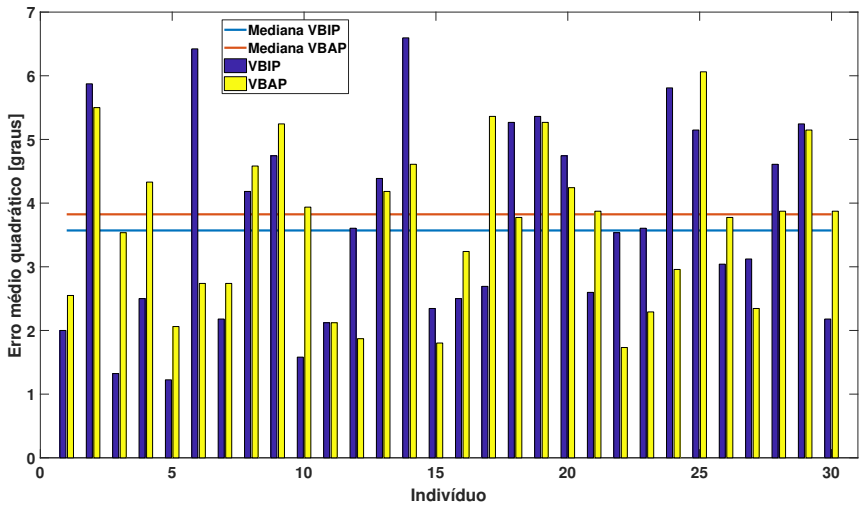


Figura 4.22.: Erro quadrático médio entre as apresentações com ângulos repetidos dentro da bateria com sinal de ruído branco contínuo WN.

Na avaliação da repetibilidade da localização da fonte sonora com o sinal ruído branco contínuo (WN) apresentada na Figura 4.22 é possível notar que não houve erros com magnitudes que representem confusão frente costas. Também é apresentada a mediana do erro quadrático médio para as duas técnicas. A mediana para os 30 sujeitos apresenta a resolução média do conjunto de sujeitos para o sinal de ruído branco inferior a 5° , visto que o EQM calculado para cada indivíduo representa a variância da resolução espacial¹² no plano horizontal.

Para o sinal de fala (Figura 4.23), também não ocorreram erros de localização que levariam à exclusão de sujeitos, indicando que os sujeitos conseguiram localizar a direção de incidência do som. A mediana do EQM de ambas as técnicas (VBAP e VBIP) está inferior a 4 graus para o sinal de fala.

¹² De localização de fontes sonoras.

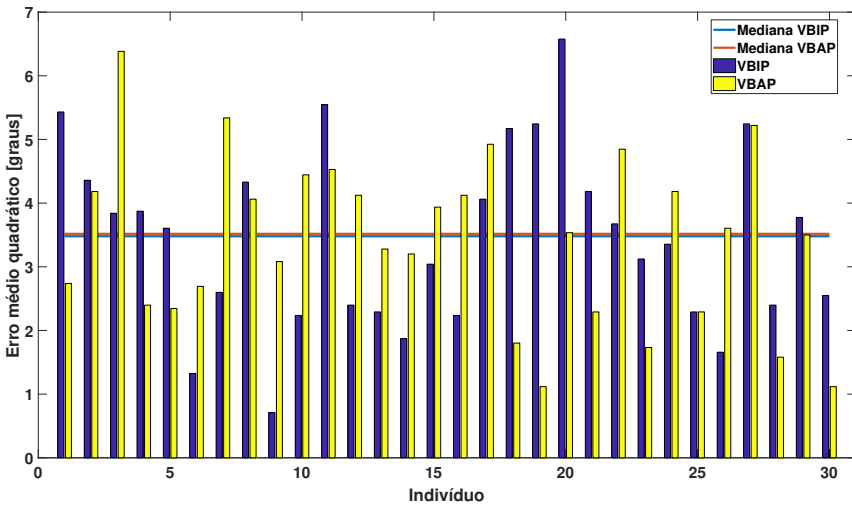


Figura 4.23.: Erro quadrático médio entre as apresentações com ângulos repetidos dentro da bateria com sinal de Fala.

Seguindo a mesma tendência dos ensaios anteriores, a avaliação da repetibilidade de localização de fontes sonoras que emitem ruído branco intermitente (WN_i) também resultou em medianas de erro inferiores a 4 graus para ambas as técnicas, indicando assim a resolução média dos sujeitos para ambas as técnicas.

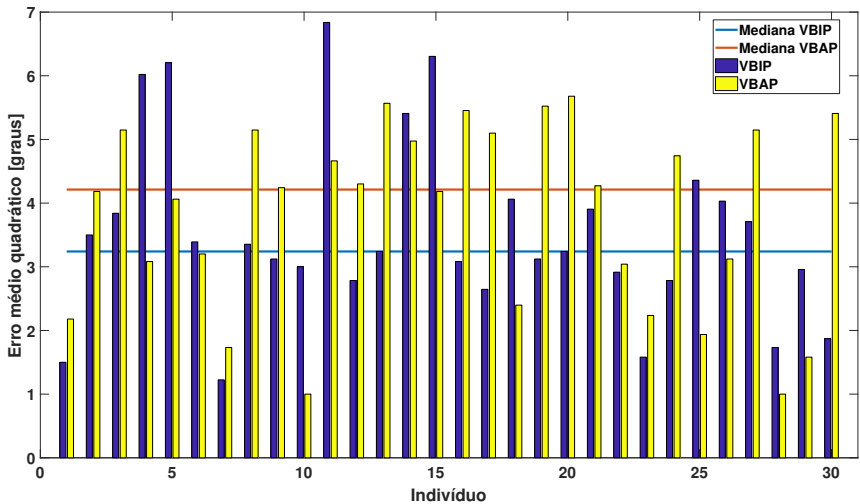


Figura 4.24.: Erro quadrático médio entre as apresentações com ângulos repetidos dentro da bateria com sinal de ruído branco intermitente WN_i .

A Tabela 4.2 sumariza os resultados das medianas apresentados nas Figuras 4.22 a 4.24.

Tabela 4.2.: Mediana do Erro quadrático médio na localização de fontes sonoras repetidas.

Mediana		
Sinal	VBAP [°]	VBIP [°]
WN	4,2	3,2
Fala	3,5	3,5
WN_i	3,8	3,6

Dada a pequena quantidade de repetições por indivíduo a mediana é estatisticamente mais representativa [66]. As medianas de erro de localização foram condizentes com os apresentados na revisão de literatura.

Entretanto, os erros de localização aferidos pelo método subjetivo são superiores ao erro mensurado pelo método objetivo (ver Tabela 4.1). Assim, é possível associar o resultado à resolução para localização de fontes sonoras no plano horizontal¹³ do conjunto de

¹³ Resolução espacial em conjunto aos erros constantes.

sujeitos testados e para os sinais apresentados utilizando o sistema apresentado. A diferença entre as medianas do erro de localização das técnicas VBAP e VBIP não apresentou significância estatística aplicando a análise de variância (estatística F obtida inferior ao F tabelado). O número de graus de liberdade nesta análise foi de 39 com 40 observações (o valor tabelado para Distribuição F de Snedecor a 5% F é de 1,99). O resultado separado por sinais é apresentado na Tabela 4.3.

Tabela 4.3.: Análise de variância (ANOVA) das medianas de erro absoluto de repetição do ensaio subjetivo final para as técnicas VBAP e VBIP.

Sinal	WN	Fala	WN _i
Estatística F	1,04	0,92	1,03

4.2.3.3 Análise espacial do erro de localização perceptivo

A avaliação do erro expresso em termos de localização azimutal é apresentado nesta seção. Duas representações do erro no espaço foram utilizadas, uma representação por grupos de ângulos a cada 5 graus, visando uma discretização maior na avaliação do erro de localização das fontes sonoras do sistema e outra representação com grupos contendo ângulos referentes a conjuntos de 45 graus, o que representa os ângulos no espaço entre os alto-falantes. Em se tratando de avaliação subjetiva, observando especificamente o erro de localização, o Erro de Localização Perceptivo é um conjunto de erros devido à performance do sujeito ao realizar a tarefa associado aos erros devido ao procedimento/sistema de virtualização e medição. A escolha de agrupar os erros em *clusters* de 5 e 45 graus foi realizada em função da natureza aleatória¹⁴ dos ângulos em que foram virtualizadas as fontes sonoras, sendo que nem todos os ângulos foram avaliados.

¹⁴ O algoritmo do ensaio subjetivo escolheu aleatoriamente os ângulos nos quais as fontes sonoras foram reproduzidas para cada sujeito.

Análise do ELP em grupos de 5 graus

O erro quadrático médio para blocos de 5 graus pode ser visto na Figura 4.25 para a técnica VBAP, com os sinais WN (Figura 4.25 (a)), Fala (Figura 4.25 (b)) e WN_i (Figura 4.25 (c)).

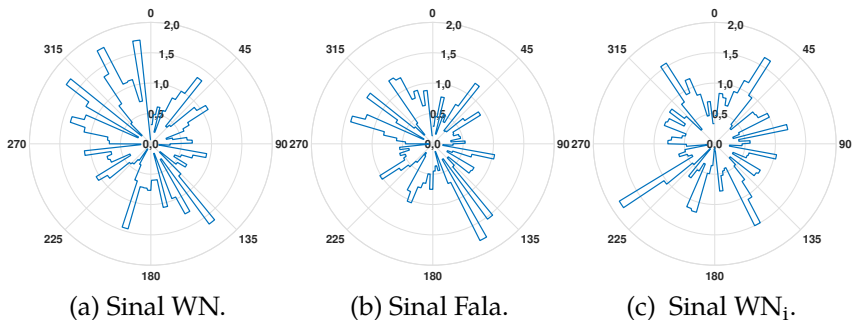


Figura 4.25.: VBAP Erro quadrático médio de localização - grupos de 5 graus.

O agrupamento dos erros a cada 5 graus apresentado não demonstra viés de erro dependente da localização da fonte sonora para nenhum dos sinais com a técnica VBAP. Comprovou-se que não existe um erro sistemático entre os ângulos 180° e 225° , por onde os sujeitos entraram no círculo e, assim, poderiam ter uma tendência maior de não apontar para ângulos entre esses limites por saber que não haviam alto-falantes nessas posições.

É possível notar que nos grupos em que há uma fonte real (alto-falante) há uma tendência menor de erro, indicados nas figuras com os ângulos $n * 45^\circ$ ($n = 0$ até 7).

A Figura 4.26 apresenta o erro quadrático médio de localização em blocos de 5 graus para a técnica VBIP e os sinais WN (Figura 4.26 (a)), Fala (Figura 4.26 (b)) e WN_i (Figura 4.26 (c)).

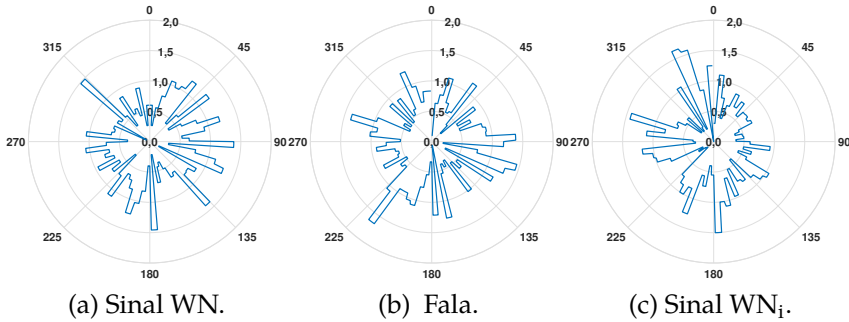


Figura 4.26.: VBIP Erro quadrático médio de localização - grupos de 5 graus.

Na técnica VBIP, ao repetir-se a análise do erro quadrático médio de localização em grupos de 5 graus, assim como na técnica anterior, nenhum conjunto de ângulos destacou-se de forma a ser considerado um viés de localização associado ao espaço na virtualização das fontes sonoras.

Análise do ELP em grupos de 45 graus

O ELP atribuído aos pares de alto-falantes utilizados no *setup* foi calculado a partir da média do ELP de cada cone de 45° . Foram escolhidas duas maneiras de apresentação do ELP, diagramas de caixas e diagrama polar.

A visualização do erro considerando o espalhamento para esses grupos de 45° , separados por tipo dos sinais, pode ser visualizada no diagrama de caixas (*boxplot*) apresentado nas Figuras 4.27, 4.28 e 4.29.

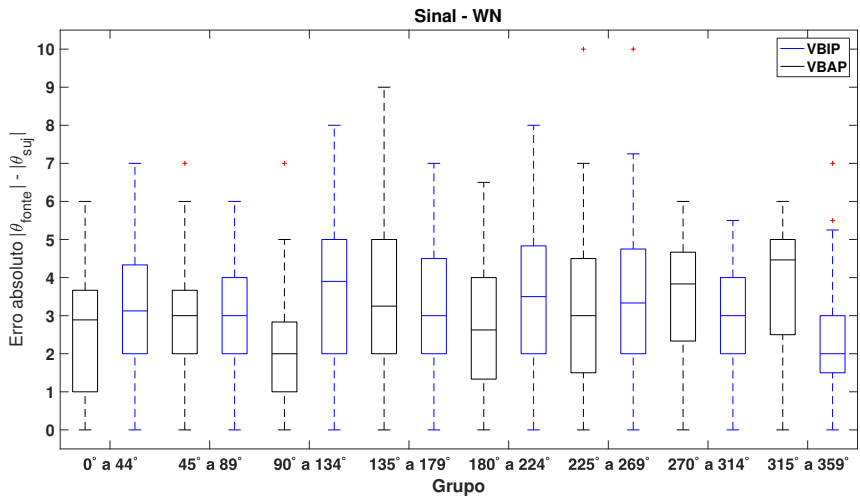


Figura 4.27.: ELP absoluto - sinal: WN Técnicas VBIP e VBAP.

Para o ruído branco contínuo (Figura 4.27) o conjunto de dados resultou em medianas do ELP para todos os grupos de ângulos inferiores a 5 graus e valores interquartis (referentes a 75% dos valores observados) de no máximo 3 graus. Com apenas 6 valores discrepantes¹⁵, sendo 3 em cada técnica, o resultado é consistente e permite uma visão equiparada quanto ao ELP horizontal a partir dos pares de alto-falantes em ambas as técnicas. A avaliação indica uma correta implementação da técnica de virtualização de fontes sonoras para os sinais de ruído branco contínuo.

¹⁵ Valores discrepantes são considerados além do intervalo de 95% dos dados em torno da mediana.

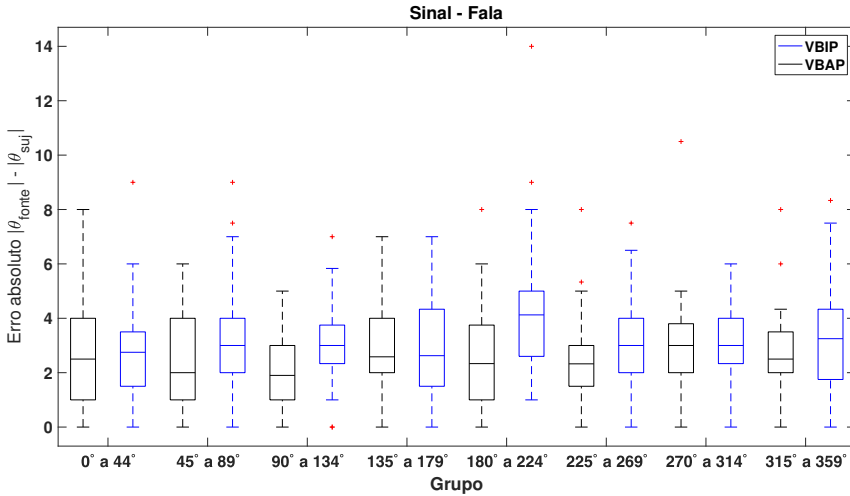


Figura 4.28.: ELP absoluto - sinal: Fala. Técnicas VBIP e VBAP.

O sinal de fala apresentou 8 valores discrepantes na utilização da técnica VBIP para reprodução e 6 valores discrepantes utilizando VBAP. Para todos os grupos de ângulo a mediana do erro foi ligeiramente menor utilizando a técnica VBAP do que a técnica VBIP apontando uma tendência de melhor ajuste ao tipo de sinal focado em diferença de tempo entre as orelhas como pista principal à técnica VBAP.

A mediana de todos os grupos para o sinal WN_i foi inferior a 4 graus (Figura 4.29), contudo com valores de mediana próximos e com melhor desempenho para cada técnica em 4 de 8 grupos. Assim, não há como inferir a tendência de melhor aplicação a partir da pista principal do sinal (ITD) para o sinal de ruído branco intermitente.

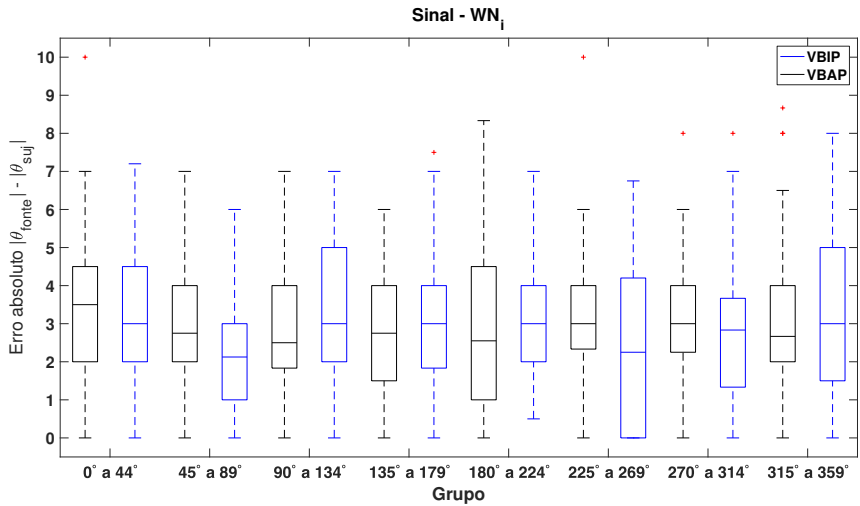


Figura 4.29.: ELP absoluto - sinal: ruído branco interrompido.
Técnicas VBIP e VBAP.

No diagrama polar apresentado na Figura 4.30 é possível visualizar o EQM de localização associado ao intervalo de ângulos entre dois alto-falantes de cada par de alto-falantes e para cada tipo de sinal contabilizando as 20 apresentações de cada um dos 30 sujeitos que realizaram o ensaio.

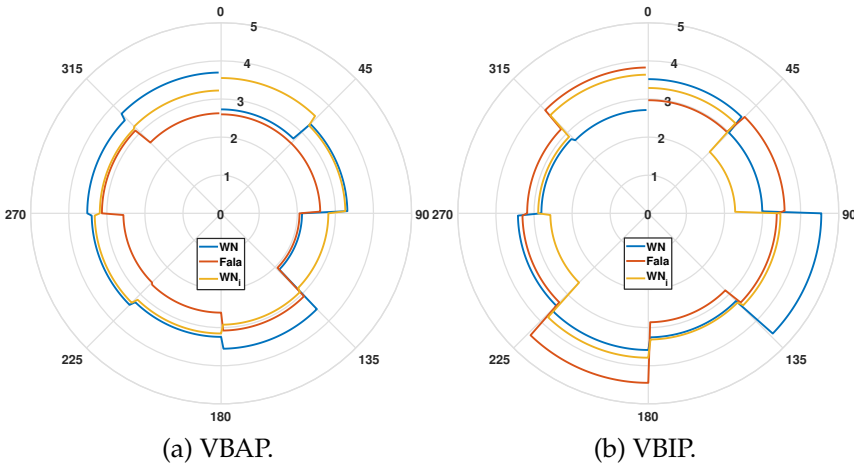


Figura 4.30.: Erro quadrático médio de localização perceptivo para os ângulos entre dois alto-falantes de um par (grupos de 45°).

O erro quadrático médio de localização perceptivo correspondente aos pares de alto-falantes para os três sinais utilizando a técnica VBAP (ver Figura 4.30 (a)) apresenta uma tendência de redução independente do sinal entre os três sinais entre os ângulos de 90° a 135°. Na técnica VBIP (ver Figura 4.30 (b)) há elevação do erro quadrático médio de localização perceptivo independente do sinal apenas no intervalo entre 180° e 225°.

Apesar das tendências nos intervalos apontados a análise não é conclusiva dada a magnitude do erro ser da mesma ordem do valor encontrado como viés na análise das repetições dos indivíduos (ver Seção 4.2.3.2).

Uma análise de variância ANOVA da média dos erros absolutos entre os grupos de ângulos foi proposta. Para a distribuição com 7 graus de liberdade e um número de observações igual a 30 o valor tabelado da Distribuição F de Snedecor a 5% ($p = 0,05$) é igual a 2,33. Assim valores superiores ao F tabelado negam a hipótese nula de que não há diferença significativa entre as médias do erro absoluto dos grupos de ângulos $H_0 : \mu_i = \mu_j$. A partir da análise, os resultados da estatística F (apresentados na Tabela 4.4) apenas para as fontes virtualizadas utilizando VBAP e reproduzindo ruído

branco a H_0 é rejeitada e a hipótese alternativa $H_1 : \mu_i \neq \mu_j$ é aceita.

Tabela 4.4.: Análise de variância (ANOVA) entre as médias de erro absoluto do ensaio subjetivo para grupos de 45° .

Sinal	Estatística F	
	VBAP	VBIP
WN	2,38	1,51
Fala	0,99	1,51
WN _i	0,48	1,28

Para identificar em quais conjuntos de média a discrepância é significativa estatisticamente o teste de comparações múltiplas de Tukey foi realizado e o resultado é apresentado na Figura 4.31.

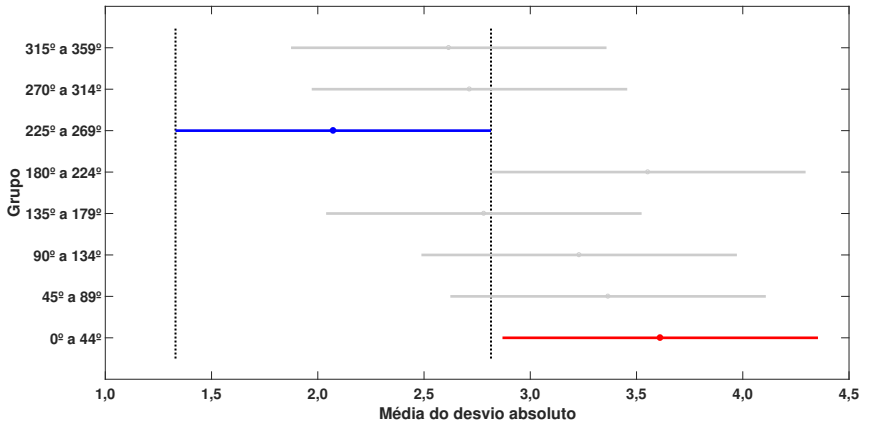


Figura 4.31.: Teste de Tukey para comparação das médias.
Técnica: VBAP, sinal: WN

Houve diferença estatisticamente significativa, segundo o ensaio, apenas entre o grupo 1 (0° a 44°) e o grupo 6 (225° a 269°). Considerando, no entanto, que tanto o intervalo de diferença mínima significativa quanto as médias dos grupos são inferiores a $4,5^\circ$, esse desvio da resposta média pode ser considerado um erro aleatório, dado ainda que apenas ocorreu em uma técnica das duas técnicas e com um dos três sinais.

Valor esperado

A média do erro combinado com o desvio padrão¹⁶, resultando no valor esperado conforme Equação (4.6) [67], é apresentado na Tabelas 4.5 para um intervalo de confiança de aproximadamente 95% ($k_p = 2$) para os três sinais com as duas técnicas utilizadas.

$$\mu_z \pm k_p \sigma, \quad (4.6)$$

na qual,

- μ_z é a esperança de uma grandeza z com distribuição normal,
- k_p é o fator de abrangência que produz um intervalo correspondente a um nível especificado da confiança p ,
- σ é desvio padrão do conjunto.

Tabela 4.5.: Valores esperados do ELP absoluto com fator de abrangência de $\approx 95\%$ ($k_p = 2$) na medição subjetiva.

Sinal	VBAP [°]	VBIP [°]
WN	$3,3 \pm 0,9$	$3,2 \pm 0,7$
Fala	$2,8 \pm 0,7$	$3,4 \pm 0,8$
WN _i	$3,2 \pm 0,9$	$3,2 \pm 0,9$

Fez-se então uma investigação sobre a diferença nas implementações para o erro geral. Utilizando-se a análise de variância da média dos erros absolutos de todas as apresentações de cada sinal nas duas técnicas, com o número de graus de liberdade nesta análise igual a 39 com 40 observações (20 para cada técnica). O resultado é apresentado na Tabela 4.6 e indica que não há diferença significativa entre as médias dado que o valor tabelado para Distribuição F de Snedecor a 5% F é de 1,99.

¹⁶ Para garantir a resiliência da análise o desvio padrão do conjunto foi utilizado.

Tabela 4.6.: Análise de variância (ANOVA) entre a média de erro absoluto do ensaio subjetivo nas técnicas VBAP e VBIP.

Sinal	WN	Fala	WN _i
Estatística <i>F</i>	1,01	1,25	1,15

Assim as duas técnicas apresentam resultados coerentes e podem ser utilizadas em ensaios subjetivos.

5 Conclusões e trabalhos futuros

Neste capítulo serão apresentadas as conclusões e sugeridos trabalhos futuros.

5.1 Conclusões

O trabalho com fontes sonoras virtuais proporciona a possibilidade de otimização de testes de percepção e inteligibilidade da fala, requisito importante na avaliação de candidatos ao implante coclear. Esse trabalho implementou satisfatoriamente duas técnicas (VBAP e VBIP) de virtualização de fontes sonoras utilizando um arranjo de oito alto-falantes posicionados com diferença equiangular polar no plano horizontal.

Ambas as técnicas foram avaliadas em termos objetivos e em termos subjetivos por meio do erro de localização perceptivo. A partir de ensaios nos quais sujeitos localizaram as fontes virtuais reproduzindo diferentes sinais computou-se os erros de localização perceptivo ELP associados a cada posição da fonte sonora (virtual ou real). Assim, concluiu-se que as duas técnicas selecionadas e implementadas nesse trabalho (VBAP e VBIP) são adequadas para virtualização de fontes sonoras no contexto de ensaios subjetivos de localização.

O instrumento de medida (ensaio subjetivo) apresentou-se adequado às necessidades com pequenas mudanças a partir do ensaio piloto. Os erros relativos à confusão frente e costas eram rapidamente resolvidos pelos sujeitos utilizando movimentos da cabeça. Assim, tanto no ensaio subjetivo piloto quanto no principal, os sujeitos informaram os ângulos percebidos sem esse viés e a

análise estatística foi simplificada [22]. Os ensaios piloto e principal tiveram duração média de 26 minutos e não foi constatado indícios de cansaço ou aprendizado.

A análise acerca repetibilidade dos indivíduos aponta uma resolução espacial de localização de fontes sonoras somado ao erro constante entre $3,2^\circ$ e $4,2^\circ$ para os ensaios realizados no trabalho, medida consistente com a mediana ($\approx 3^\circ$) encontrada por meio de ensaios subjetivos com pessoas utilizando apenas fontes sonoras reais [35] (sem virtualização do campo sonoro) para o ângulo frontal e alta relação sinal-ruído. Também é comparável às simulações com técnicas de virtualização de campo sonoro apresentadas por Grimm *et al.* [38] de $\approx 5^\circ$ para a técnica VBAP.

A análise de variância ANOVA aponta que não há diferença estatística significativa entre as medianas dos erros de repetição¹ dos indivíduos para as técnicas implementadas. O ensaio principal foi realizado com 30 sujeitos, estudantes de mestrado, doutorado e professores do Laboratório de Vibrações e Acústica e produtores de áudio. Foi realizado com o mesmo *setup* do ensaio piloto incluindo as modificações oriundas das observações dos sujeitos participantes do ensaio piloto.

Outra análise de variância ANOVA apontou uma distribuição coerente das médias do ELP associadas aos intervalos de ângulos que pertencem a um par de AF (intervalos de 45°) para as técnicas VBAP e VBIP, apontando que não há uma diferença estatisticamente significativa entre elas. O padrão espacial do erro oscila em torno de $4,5^\circ$, indicando a ocorrência de erro aleatório típico de ensaios subjetivos. Nas análises dos erros de localização perceptivo agrupados em intervalos de 5° houve uma tendência da média absoluta do erro diminuir no caso da técnica VBAP nos ângulos nos quais havia o alto-falante (posição de fonte real). Contudo, as médias de erro computadas para todos os intervalos de ângulos de 5° estão inferiores a 2° e como existem variações entre as magnitudes dos erros desses grupos de ângulos onde existem as fontes reais (alto-falantes) não é possível afirmar com certeza a

¹ Avaliados por meio do EQM.

validade de tal tendência.

A implementação das técnicas e da montagem do sistema apresenta-se coerente analisando a distribuição espacial das médias dos ELP absolutos (Figura 4.26) também para a técnica VBIP. Padrões diferentes de erro com magnitudes semelhantes relativos a todos os sinais ficaram abaixo de 2° de erro médio para a média dos grupos de 5° analisados.

Em outra análise, agrupando as médias dos ELP a cada 45°, intervalo de ângulos associado a cada par de alto-falantes, mostra uma tendência homogênea caracterizada pelo mesmo comportamento de aumento ou diminuição do erro médio quadrático dentre os sinais para a mesma técnica.

Uma pequena tendência de diminuição do erro entre 90° e 135° foi observada para a técnica VBAP e de aumento do erro entre 180° e 225°, porém a magnitude de tais diferenças é inferior a 1°.

Assim, os ensaios subjetivos indicaram que ambas as técnicas de virtualização de fontes sonoras são equivalentes para virtualizar fontes sonoras no plano horizontal com os sinais propostos e para serem utilizados em ensaios nesse plano. A média do erro perceptivo de localização obtida é comparável aos erros observados em ensaios com fontes reais e simulações reportadas na literatura indicando uma correta implementação da técnica e da montagem do sistema.

As diferentes configurações de arranjo de sensores para a gravação e posterior aferição objetiva do sistema de virtualização dos sinais e posteriores cálculos dos ângulos de radiação das fontes sonoras demonstraram que um arranjo triangular de sensores consegue identificar com mais exatidão se comparado ao arranjo em linha, um efeito esperado na teoria de arranjos de sensores [28]. Há uma dificuldade inerente no posicionamento dos microfones no espaço que resulta em erros sistemáticos de montagem. Contudo, em ambas as configurações o erro observado é inferior às médias de ELP obtidas nos ensaios subjetivos.

A avaliação do sistema com medidas objetivas resultou em erros máximos na ordem de 1,5°, enquanto as avaliações subjetivas

incorreram em medianas do erro ELP da ordem de 5° na análise de repetição. Assim, o erro nos ensaios subjetivos é associado à resolução espacial dos sujeitos e erros sistemáticos. Isso indica uma adequada implementação das técnicas de virtualização que podem ser utilizadas como referência em ensaios subjetivos posteriores.

A avaliação do sistema proposta desempenhou satisfatoriamente o papel de validação da percepção humana quanto a localização dos sinais executados com as técnicas mencionadas.

5.2 Trabalhos futuros

Dado o curto período de tempo no qual o trabalho foi desenvolvido, métodos, ensaios e aprimoramentos não tiveram oportunidade de ser devidamente aprofundados. Assim, o registro de tais tópicos, bem como a apresentação de ideias para novos trabalhos é importante.

Estudos sobre a eficiência do sistema de fontes virtuais em função do tipo ou do grau da perda auditiva precisam ser conduzidos visando à validação de um protocolo final (completo) de seleção dos candidatos ao implante coclear, bem como estudos de percepção angular desses indivíduos e estudos sobre o mascaramento envolvendo tipos de ruídos e posição espacial do ruído mascarador em pessoas com audição normal, pacientes com perda auditiva e usuários de implante coclear.

Algumas tarefas associadas à ensaios de localização de fontes sonoras, partindo dos ensaios realizados neste trabalho podem envolver

- Ensaios subjetivos de localização de fontes sonoras relativos a distância percebida;
- Ensaios de localização de fontes sonoras com diferentes estímulos (ruídos de banda larga, fala, *clicks* e tonais);
- Criação de ensaios para avaliar mascaramento a partir de diferentes relações sinal-ruído, posicionamento das fontes no

espaço, similaridade dos sinais sonoros e ainda relacionados ao auxílio visual;

- Criação de jogos educativos com tarefas de localização de fontes sonoras ou identificação lúdica do som para crianças;
- Implementação de uma trilha de *leds* à frente do arranjo de alto-falantes para identificação, com cores, da localização da fonte sonora, podendo auxiliar, por exemplo, nos jogos educativos;
- Implementação de dispositivo mecânico de controle angular associado ao conjunto de *leds* para facilitar os ensaios e aumentar a confiabilidade, reduzindo erros constantes (sistemáticos), dos resultados e também reduzindo erros associados a visualizações erradas ou perda de informação entre sujeito e avaliador;
- Desenvolvimento de solução para uniformidade do tecido que fica à frente das caixas para controlar as pistas visuais de acordo com a necessidade do ensaio;
- Avaliação da influência do tecido nas funções de transferência do sistema;
- Gravações com diferentes arranjos para localização do ângulo de radiação da fonte sonora;
- Utilização de sinais tonais de baixas frequências para avaliação objetiva.

Algumas tarefas associadas à criação de um protocolo padronizado, partindo dos ensaios realizados neste trabalho podem envolver

- Ensaios subjetivos de identificação de tamanho ou largura de fontes sonoras;

- Desenvolvimento/Implementação de controle ativo dos modos acústicos do ambiente de escuta crítica a ser construído em que serão realizados os ensaios no Laboratório de Vibrações e Acústica;
- Definição e realização dos ensaios subjetivos de localização com deficientes auditivos;
- Definição e realização dos ensaios subjetivos para avaliação de percepção da fala e inteligibilidade da fala;
- Definição e realização dos ensaios subjetivos para avaliação de mascaramento espacial e outros tipos de mascaramento.

Ainda, os ensaios podem ter uma fase com alunos(as) do curso de fonoaudiologia, os quais têm fácil acesso às medidas de qualidade auditiva e podem auxiliar no desenvolvimento dos protocolos e necessidades dos ensaios com pacientes com implante coclear ou candidatos a implante.

Outras análises podem utilizar o sistema como por exemplo simulações de localização por meio da medição de HRTF (*Head Related Transfer Function*). É possível comparar resultados de localização com HRTFs relativas ao plano horizontal de diferentes simuladores de tronco e cabeça e também HRTFs individualizadas. Pode-se criar métricas objetivas e facilitar implementações de novos algoritmos de áudio espacial ou alterações nos atuais.

Como sugestão para o artigo a ser desenvolvido em decorrência do trabalho o foco é a possibilidade de utilização do sistema de ensaio de localização de fontes sonoras por deficientes auditivos e usuários de Implante Coclear e AASI. O tema de avaliação perceptiva em ouvintes sem perda auditiva foi explorado nesse trabalho com resultados, subjetivos e objetivos, comparáveis aos resultados de simulações e de ensaios subjetivos que utilizaram fontes reais obtidos por diferentes pesquisadores. A utilização das técnicas de virtualização para ensaios de localização de fontes sonoras pode ser amplamente explorada e, a partir de resultados positivos em ensaios com pessoas portadores de deficiência

auditiva, pode resultar em uma nova e poderosa ferramenta de avaliação na prática clínica fonoaudiológica.

Referências bibliográficas

- [1] BRASIL. Portaria do ministério da saúde no. 2.276 de 18 de dezembro de 2014. Diário Oficial da União, Brasília, n. 246,183. (Citado na página 29)
- [2] Mondelli, M. F. C. G.; Santos, M. d. M. d. e José, M. R. Speech perception in noise in unilateral hearing loss. *Brazilian Journal of Otorhinolaryngology*, 82:427 – 432, 08 2016. (Citado na página 29)
- [3] Couto Melo, R. *HINT: um estudo sobre os critérios de julgamento no Brasil*. Mestrado em saúde da comunicação humana, Universidade Federal de Pernambuco, 2015. (Citado na página 29)
- [4] Lynn Rosalina Alves, P. R. Tecnologia assistiva – uma revisão do tema. *HOLOS*, 6(0):170–180, 2014. (Citado na página 33)
- [5] Bersch, R. Introdução à tecnologia assistiva. http://www.assistiva.com.br/Introducao_Tecnologia_Assistiva.pdf, 2013. (Citado na página 33)
- [6] CAT. Ata da reunião vii de 13 e 14 de dezembro de 2007 comitê de ajudas técnicas secretaria especial dos direitos humanos da presidência da república (corde/sedh/pr), 2007. (Citado na página 33)
- [7] Organização Mundial da Saúde. Classificação dos graus de perda auditiva. http://www.who.int/pbd/deafness/hearing_impairment_grades/en/, 2014.

(Citado na página 33)

- [8] Clark, G. *Cochlear Implants Fundamentals and Applications*. : Springer, 2003.
(Citado na página 33)
- [9] Wouters, J.; McDermott, H. J. e Francart, T. Sound coding in cochlear implants: From electric pulses to hearing. *Signal Processing Magazine, IEEE*, 32(2):67–80, 2015.
(Citado nas páginas 33 e 34)
- [10] Niparko, J. *Cochlear Implants: Principles and Practices*. : LWW, 2009.
(Citado nas páginas 34, 35 e 36)
- [11] BIOSOM. *Implante coclear para pessoas com deficiência auditiva*, 2015.
(Citado na página 34)
- [12] Kruger, B. e Kruger, F. Speech audiometry in usa. Em Martin, M, editor, *Speech audiometry, Exc Business And Economy*, page 89. Wiley, Michigan University, 1997.
(Citado na página 34)
- [13] Penrod, J. Logoaudiometria. Em Kantz, J, editor, *Tratado de audiologia clínica 4. ed*, pages 146–162. Manole, São Paulo, 1999.
(Citado na página 34)
- [14] Gomez, M. V. S. G.; Guedes, M. C.; Sant. Anna, S. B. G.; Peralta, C. G. O.; Tsuji, R. K.; Castilho, A. M.; Neto, R. V. B. e Bento, R. F. Critérios de seleção e avaliação médica e audiológica dos candidatos ao implante coclear: Protocolo HC-FMUSP. *Arquivos Internacionais de Otorrinolaringologia*, 8 (4):295–317, 2004.
(Citado na página 34)
- [15] Gamma, M. R. *Desenvolvimento e estudo comparativo de listas de palavras para uso na medida do limiar de reconhecimento de fala em crianças de 5 a 7 anos de idade*. Doutorado, Instituto de Psicologia

- Universidade de São Paulo Neurociências e Comportamento, 2004.
(Citado nas páginas 35 e 36)
- [16] Caporali, S. A. e Silva, J. Reconhecimento de fala no ruído em jovens e idosos com perda auditiva. *Revista Brasileira de Otorrinolaringologia*, 70:525 – 532, 08 2004.
(Citado na página 35)
- [17] Miranda, J. S. e Agra, S. E. R. Logoaudiometria: o uso do mascaramento na avaliação do reconhecimento de fala em indivíduos com deficiência auditiva sensorioneural. *Salusvita*, 27(3):329–339, 2008.
(Citado na página 35)
- [18] Lawson, G. e Peterson, M. *Speech Audiometry*. : Plural Publishing Incorporated, 2011.
(Citado nas páginas 35 e 36)
- [19] Young, N. M. *Pediatric Cochlear Implantation Learning and the Brain*. : Springer New York, 2016.
(Citado na página 36)
- [20] Strumillo, P. *Advances in Sound Localization*. : InTech, 2011.
(Citado nas páginas 36, 37, 40, 41, 83 e 84)
- [21] Paul, S. A fisiologia da audição como base para fenômenos auditivos. Em *Proceedings of the 12th AES Brazil Conference*, São Paulo, SP, 13-15 maio 2014.
(Citado na página 37)
- [22] Letowski, T. e Letowski, S. Localization error accuracy and precision of auditory localization. Em Strumillo, P, editor, *Advances in Sound Localization*, chapter 4, pages 55–78. Intech, Oxford, 2011.
(Citado nas páginas 37 e 100)
- [23] Blauert, J. *Spatial Hearing: The Psychophysics of Human Sound Localization*. : MIT Press, 1997.
(Citado na página 37)

- [24] Grothe, B.; Pecka, M. e McAlpine, D. Mechanisms of sound localization in mammals. *Physiological reviews*, 90(3): 983—1012, July 2010.
(Citado na página 37)
- [25] Pulkki, V. e Karjalainen, M. *Communication Acoustics: An Introduction to Speech, Audio and Psychoacoustics.* : Wiley, 2015.
(Citado na página 37)
- [26] Lyon, R. F. *Human and Machine Hearing Extracting Meaning from Sound.* : Cambridge University Press, 2017.
(Citado na página 37)
- [27] Blauert, J. *Communication Acoustics.* : Springer, 2005.
(Citado nas páginas 37 e 38)
- [28] Fonseca, W. D. *Beamforming Considering Acoustic Diffraction over Cylindrical Surfaces.* Ph.d. dissertation, Federal University of Santa Catarina, Florianópolis, SC, Brazil, May 2013.
(Citado nas páginas 38, 68, 71 e 101)
- [29] Benesty, J.; Sondhi, M. e Huang, Y. *Springer Handbook of Speech Processing.* : Springer-Verlag Berlin Heidelberg, 2008.
(Citado nas páginas 39 e 43)
- [30] Best, V.; Marrone, N.; Mason, C. R. e Kidd, G. The influence of non-spatial factors on measures of spatial release from masking. *The Journal of the Acoustical Society of America*, 131 (4):3103–3110, 2012.
(Citado na página 39)
- [31] Ebata, M.; Sone, T. e Nimura, T. Improvement of hearing ability by directional information. *The Journal of the Acoustical Society of America*, 43(2):289–297, 1968.
(Citado na página 39)
- [32] Suzuki, Y.; Brungart, D.; Iwaya, Y.; Iida, K.; Cabreraa, D. e Kato, H. *Principles and Applications of Spatial Hearing.* : World Scientific Publishing Company, 2009.
(Citado na página 39)

- [33] Ihlefeld, A. e Shinn-Cunningham, B. Spatial release from energetic and informational masking in a selective speech identification task. *The Journal of the Acoustical Society of America*, 123(6):4369–4379, 2008.
(Citado na página 39)
- [34] União Internacional de Telecomunicações, Setor de Radiocomunicação (ITU-R). *Recommendation ITU-R BS.1116-3*, 2015.
(Citado nas páginas 40, 50, 51, 52, 53, 54, 55, 75, 76, 119 e 123)
- [35] Kerber, S. e Seeber, B. Sound localization in noise by normal-hearing listeners and cochlear implant users. *Ear and Hearing*, 33(4):445–457, 2012.
(Citado nas páginas 40, 54 e 100)
- [36] Grohn, M. Localization of a moving virtual sound source in a virtual room: The effect of a distracting auditory stimulus. Em *Stimulus, in International Conference on Auditory Display*, 2002.
(Citado nas páginas 40 e 54)
- [37] Romigh, G. D.; Brungart, D. S. e Simpson, B. D. Free-field localization performance with a head-tracked virtual auditory display. *IEEE Journal of Selected Topics in Signal Processing*, 9(5): 943–954, Aug 2015.
(Citado nas páginas 40 e 54)
- [38] Grimm, G.; Ewert, S. D. e Hohmann, V. Evaluation of spatial audio reproduction schemes for application in hearing aid research. *CoRR*, abs/1503.00586, 2015.
(Citado nas páginas 40, 43, 54 e 100)
- [39] Alameda-Pineda, X. e Horaud, R. A geometric approach to sound source localization from time-delay estimates. *IEEE/ACM Trans. Audio, Speech and Lang. Proc.*, 22(6):1082–1095, June 2014.
(Citado na página 41)
- [40] Kwok, N. M.; Buchholz, J.; Fang, G. e Gal, J. Sound source localization: microphone array design and evolutionary

- estimation. Em *2005 IEEE International Conference on Industrial Technology*, pages 281–286, Dec 2005.
(Citado na página 41)
- [41] Vryzas, N.; Dimoulas, C. A. e Papanikolaou, G. V. Embedding sound localization and spatial audio interaction through coincident microphones arrays. Em *Proceedings of the Audio Mostly 2015 on Interaction With Sound, AM '15*, pages 36:1–36:8, New York, NY, USA, 2015. ACM.
(Citado na página 41)
- [42] Keyrouz, F. e Diepold, K. Binaural source localization and spatial audio reproduction for telepresence applications. *Presence*, 16(5):509–522, Oct 2007.
(Citado na página 41)
- [43] Wu, K. e Khong, A. W. H. Sound source localization and tracking. Em Magnenat-Thalmann, N.; Yuan, J.; Thalmann, D. e You, B.-J., editores, *Context Aware Human-Robot and Human-Agent Interaction*, pages 55–78. Springer International Publishing, Cham, 2016.
(Citado na página 41)
- [44] Dewhirst, M.; Jackson, P.; Rumsey, F. e Zielinski, S. K. Objective assesment of spatial localisation attributes of surroung-sound reproduction systems. Em *Audio Engineering Society Convention 118*, May 2005.
(Citado na página 41)
- [45] Tellakula, A. K. *Acoustic Source Localization Using Time Delay Estimation*. Master of science (engineering), Supercomputer Education and Research Centre Indian Institute of Science Bangalore - INDIA, 2007.
(Citado nas páginas 41 e 42)
- [46] Knapp, C. e Carter, G. The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 24(4):320–327, Aug 1976.
(Citado na página 42)

- [47] Atmoko, H.; Tan, D. C.; Tian, G. Y. e Fazenda, B. Accurate sound source localization in a reverberant environment using multiple acoustic sensors. *Measurement Science and Technology*, 19(2):024003, 2008.
(Citado nas páginas 42, 62, 64, 65 e 68)
- [48] Khaddour, H. *Localization and Rendering of Sound Sources in Acoustic Fields*. Tese de doutorado, FAKULTA ELEKTROTECHNIKY A KOMUNIKACN ICH TECHNOLOGI I, 2015.
(Citado na página 42)
- [49] Fereczkowski, M.; Marschall, M. e Applied Hearing Research, D. T. U. C. *Capturing and Reproducing Realistic Acoustic Scenes for Hearing Research: PhD Thesis*. : Centre for Applied Hearing Research, Technical University of Denmark, 2014.
(Citado na página 42)
- [50] Masiero, B. Qual o futuro do mp3? audio espacial e codificação orientada a objetos. Em *Proceedings of the 14th AES Brazil Conference*, São Paulo, SP, 17-19 maio 2016.
(Citado nas páginas 42 e 43)
- [51] Vorländer, M. *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*. : Springer Berlin Heidelberg, 2007.
(Citado na página 43)
- [52] L. Wightman, F. e J. Kistler, D. The dominant role of low-frequency interaural time differences in sound localisation. *The Journal of the Acoustical Society of America*, 91:1648–61, 04 1992.
(Citado na página 43)
- [53] Pulkki, V. Virtual sound source positioning using vector base amplitude panning. *Journal of the Audio Engineering Society*, 45 (6), 1997.
(Citado nas páginas 43, 44 e 45)
- [54] Faria, R. R. A. *Auralização em ambientes audiovisuais imersivos*. Tese de doutorado, Escola Politécnica da Universidade de São

- Paulo. Departamento de Engenharia de Sistemas Eletrônicos, 6 2005.
(Citado na página 44)
- [55] Galvez, M. F. S.; Menzies, D.; Mason, R. e Fazi, F. M. Object-based audio reproduction using a listener-position adaptive stereo system. *J. Audio Eng. Soc.*, 64(10):740–751, 2016.
(Citado na página 46)
- [56] Pernaux, J.; Boussard, P. e Jot, J.-M. Virtual sound source positioning and mixing in 5.1 implementation on the real-time system genesis. Em *In Proc. Conf. Digital Audio Effects (DAFx-98)*, pages 76–80, 1998.
(Citado na página 46)
- [57] Aguirre, S. L.; Paul, S.; Cordioli, J. A. e Masiero, B. S. Proposta de sistema para medição do limiar de reconhecimento da fala com fontes sonoras e ruído distribuídas espacialmente. Em *Anais do XVII Encontro da SOBRAC*, Brasília, 2017. SOBRAC.
(Citado na página 48)
- [58] Politis, A. Microphone array processing for parametric spatial audio techniques, 2016.
(Citado na página 49)
- [59] Silzle, A.; Geyersberger, S.; Brohasga, G.; Weninger, D. e Leistner, M. Vision and technique behind the new studios and listening rooms of the fraunhofer iis audio laboratory. Em *Audio Engineering Society Convention 126*, May 2009.
(Citado na página 50)
- [60] Silzle, A.; Kosmidis, D.; Greco, Gil Felix, B. D. e Bentz, L. The influence of microphone directivity on the level calibration and equalization of 3d loudspeakers setups. Em *Proceedings of 29th TONMEISTERTAGUNG – VDT INTERNATIONAL CONVENTION*, Köln, Deutschland, November 2016.
(Citado na página 50)
- [61] Berzborn, M.; Bomhardt, R.; Klein, J.; Richter, J.-G. e Vorländer, M. The ITA-Toolbox: An Open Source MATLAB Toolbox

- for Acoustic Measurements and Signal Processing. Em *Proceedings 43th Annual German Congress on Acoustics*. 43th Annual German Congress on Acoustics, Kiel (Germany), 6 Mar 2017 -9 Mar 2017, Mar 2017.
(Citado nas páginas 50 e 63)
- [62] Dietrich, P.; Guski, M.; Klein, J.; Müller-Trapet, M.; Pollow, M.; Scharrer, R. e Vorländer, M. Measurements and room acoustic analysis with the ita-toolbox for matlab. Em *40th Italian (AIA) Annual Conference on Acoustics and the 39th German Annual Conference on Acoustics (DAGA)*, 2013.
(Citado nas páginas 50 e 51)
- [63] Blauert, J. *The Technology of Binaural Listening*. : Springer Berlin Heidelberg, 2013.
(Citado nas páginas 54 e 73)
- [64] Paul, S. Disciplina de acústica ambiental – EAC 1019 - Ensaios subjetivos. Notas de aula - Engenharia Acústica, 2014.
(Citado na página 59)
- [65] Kim, D. H. e Chung, Y. Accurate position detection of sound source by labview. *Proceedings of 20th International Congress on Acoustics ICA*, 2010.
(Citado na página 65)
- [66] Siegel, S. e Castellan, N. *Estatística não-Paramétrica Para Ciências do Comportamento*. : McGraw-Hill, 1975.
(Citado nas páginas 79 e 87)
- [67] Joint Committee for Guides in Metrology. JCGM 100 *Evaluation of Measurement Data Guide to the Expression of Uncertainty in Measurement*. Technical report, JCGM, 2008.
(Citado na página 96)

Apêndices

A RI no domínio do tempo em dB

Como mencionado no Capítulo 4.2, a Recomendação da ITU-R 1116-3 [34] determina um mínimo de 10 dB entre o som direto e as primeiras reflexões para que o ambiente seja apto.

Para melhor legibilidade as respostas foram agrupadas duas a duas, os valores apresentados no segundo pico são referentes à amplitude do som das primeiras reflexões, sumarizados na Tabela 3.1.

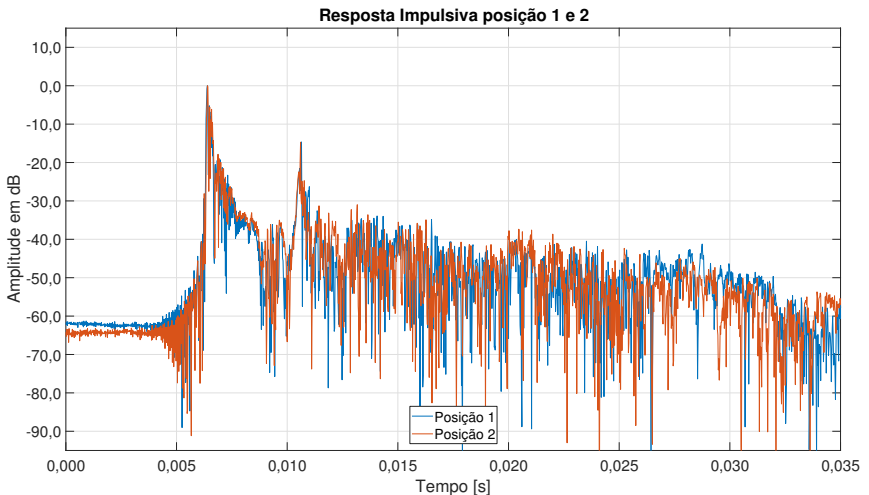


Figura A.1.: Diferença em dB do nível de pressão sonora do som direto e das primeiras reflexões nas posições 1 e 2 de fonte sonora (0° e 45°).

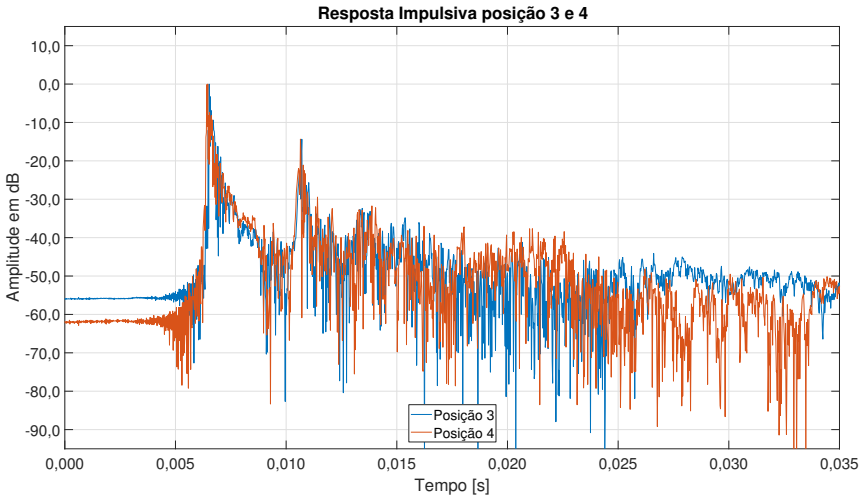


Figura A.2.: Diferença em dB do nível de pressão sonora do som direto e das primeiras reflexões nas posições 3 e 4 de fonte sonora (90° e 135°).

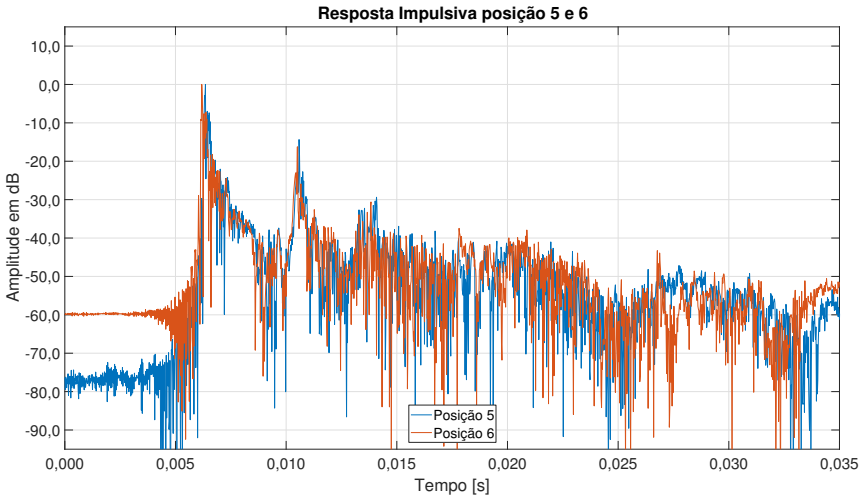


Figura A.3.: Diferença em dB do nível de pressão sonora do som direto e das primeiras reflexões nas posições 5 e 6 de fonte sonora (180° e 225°).

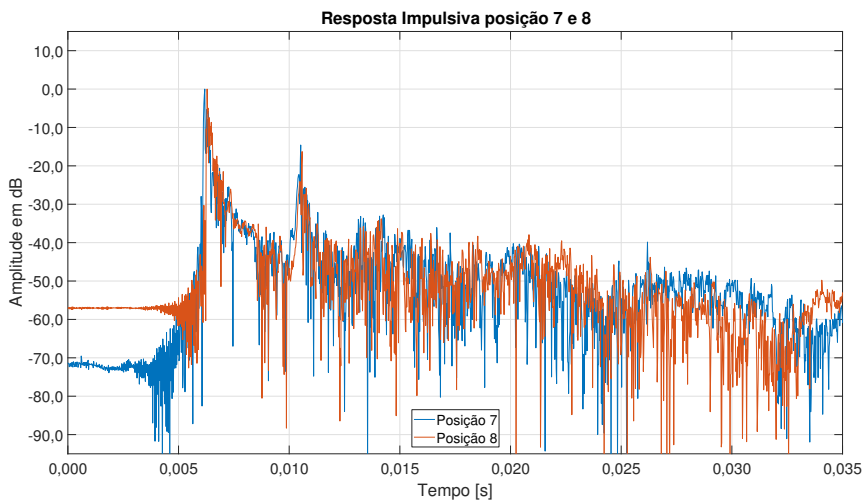


Figura A.4.: Diferença em dB do nível de pressão sonora do som direto e das primeiras reflexões nas posições 7 e 8 de fonte sonora (270° e 315°).

B RI no domínio do tempo

Como mencionado na Seção 3.1.1, a Recomendação ITU-R 1116-3 [34] determina que a resposta impulsiva de cada alto-falante em cada posição de audição seja plotada no domínio do tempo e incluída como apêndice. As Figuras B.1 a B.8 apresentam os dados requisitados.

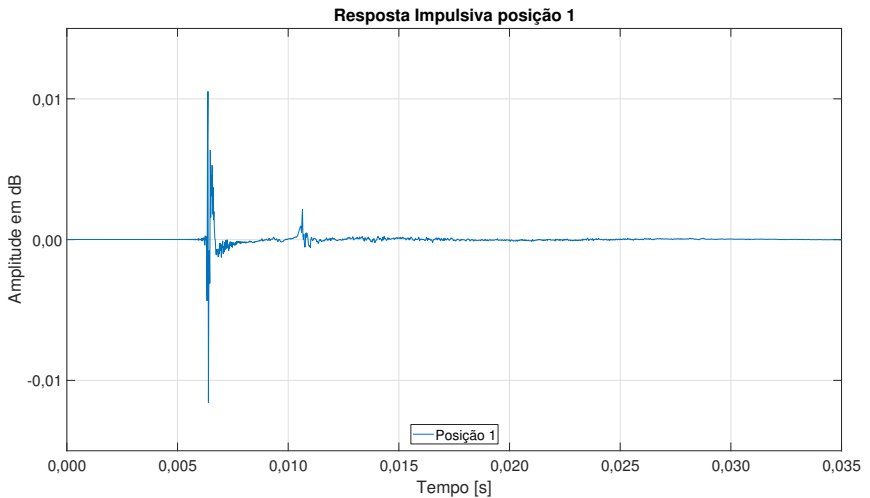


Figura B.1.: Resposta ao impulso na posição de audição com alto falante número 1 (equivalente ao ângulo 0°).

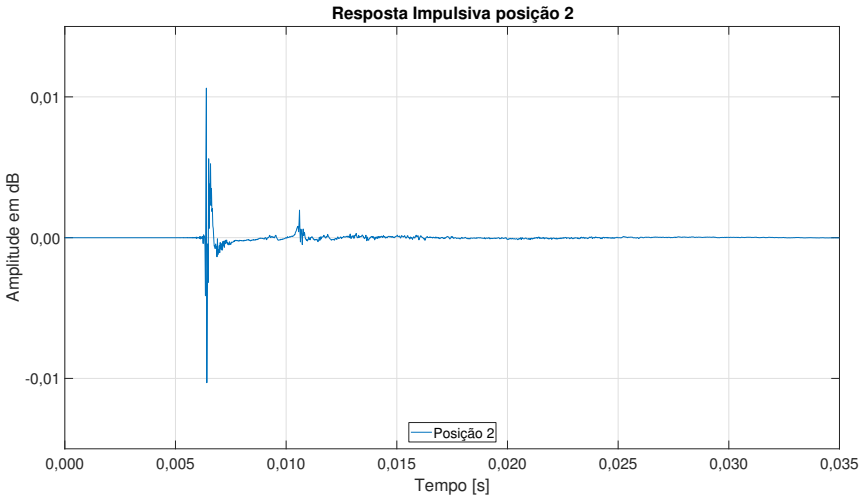


Figura B.2.: Resposta ao impulso na posição de audição com alto falante número 2 (equivalente ao ângulo 45°).

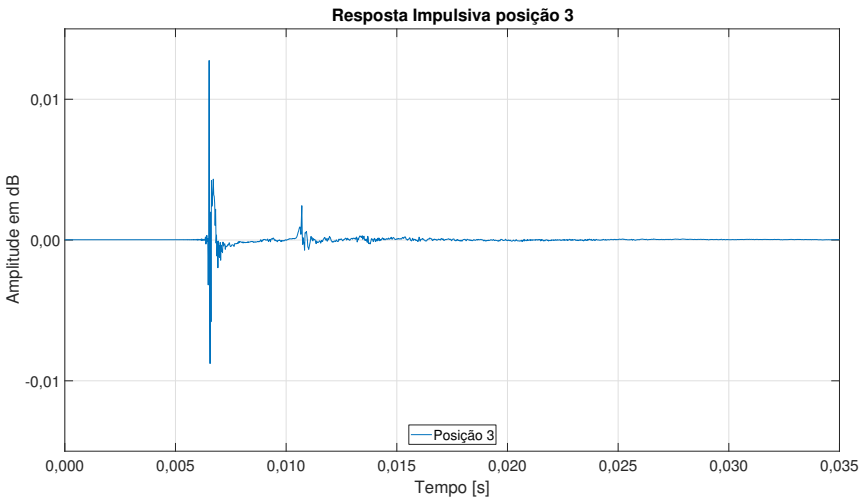


Figura B.3.: Resposta ao impulso na posição de audição com alto falante número 3 (equivalente ao ângulo 90°).

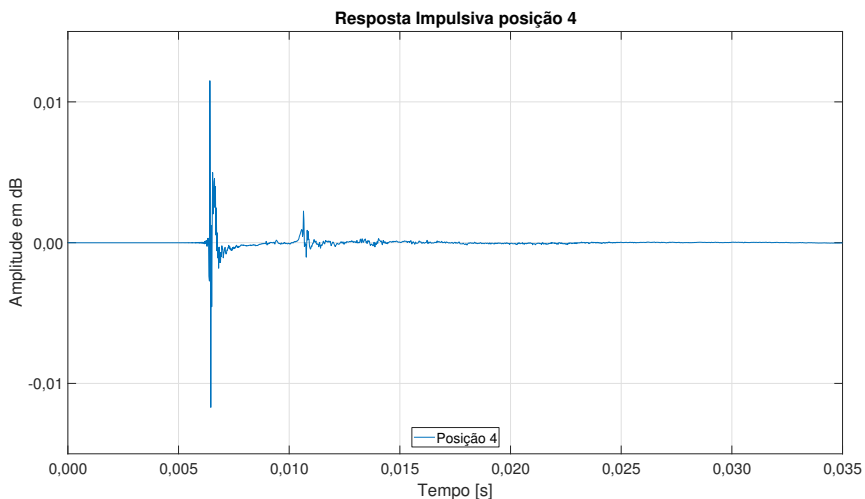


Figura B.4.: Resposta ao impulso na posição de audição com alto falante número 4 (equivalente ao ângulo 135°).

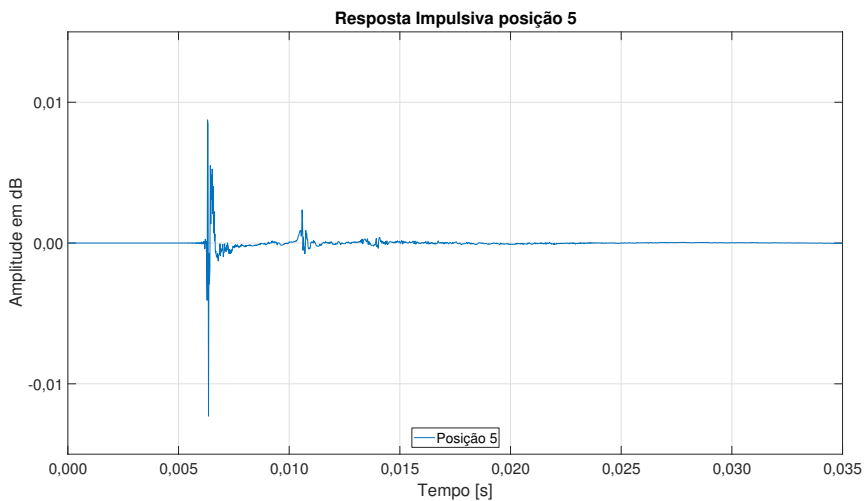


Figura B.5.: Resposta ao impulso na posição de audição com alto falante número 5 (equivalente ao ângulo 180°).

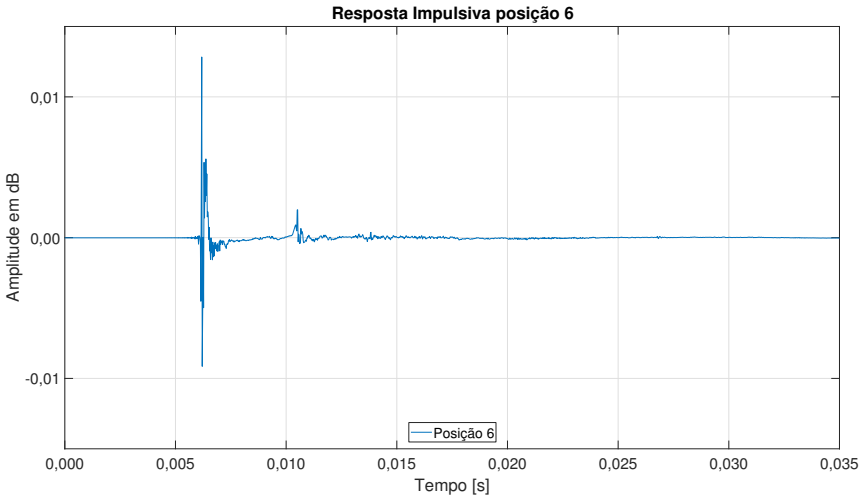


Figura B.6.: Resposta ao impulso na posição de audição com alto falante número 6 (equivalente ao ângulo 225°).

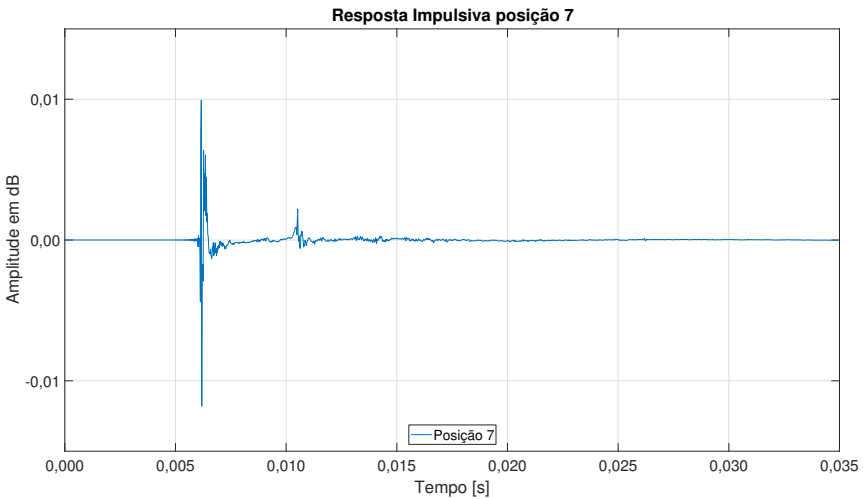


Figura B.7.: Resposta ao impulso na posição de audição com alto falante número 7 (equivalente ao ângulo 270°).

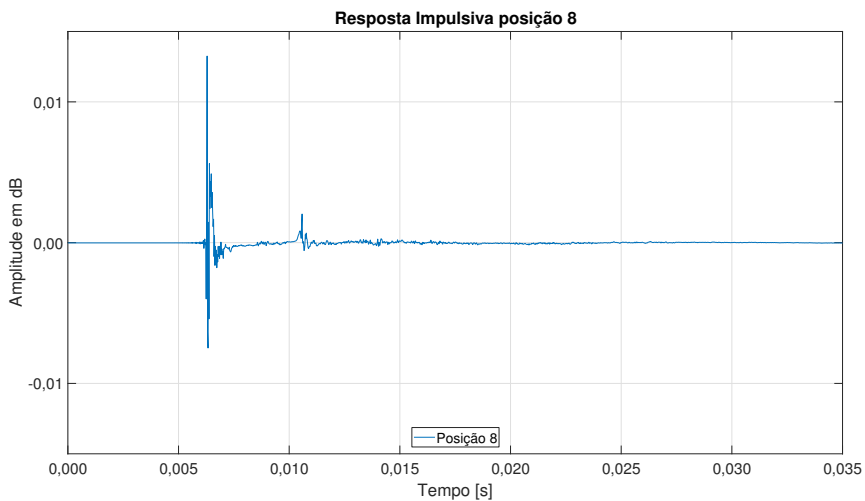


Figura B.8.: Resposta ao impulso na posição de audição com alto falante número 8 (equivalente ao ângulo 315°).

Anexo

**A Monitor de áudio de referência.
Genelec 8030C**

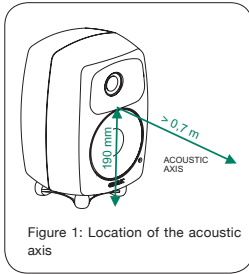


Figure 1: Location of the acoustic axis

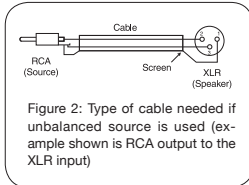


Figure 2: Type of cable needed if unbalanced source is used (example shown is RCA output to the XLR input)

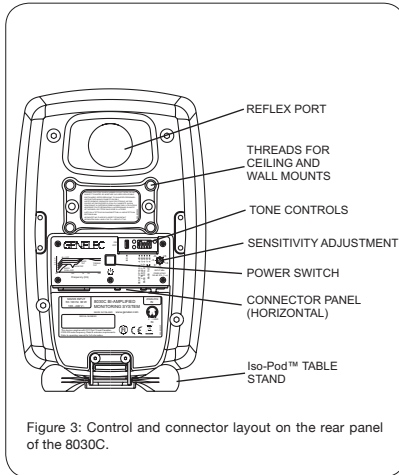


Figure 3: Control and connector layout on the rear panel of the 8030C.

Setting the tone controls

The frequency response of the 8030C can be adjusted to match the acoustic environment by setting the tone control switches on the rear panel. The controls are "Desktop", "Treble Tilt", "Bass Tilt" and "Bass Roll-Off". An acoustic measuring system is recommended for analyzing the effects of the adjustments, however, careful listening with suitable test recordings can also lead to good results if a test system is not available. Table 1 shows some typical settings in various situations. Figure 4 shows the effect of the controls on the anechoic response.

Desktop

The desktop low frequency control (Switch 1) attenuates the bass frequencies around 160 Hz by

4 dB. This feature is designed to compensate for the boost often occurring at this frequency range when the loudspeaker is placed upon a meter bridge, table or similar reflective surface.

Treble Tilt

Treble Tilt (switch 2) attenuates the treble response above 4 kHz by 2 dB, which can be used for smoothening down an excessively bright sounding system.

Bass Roll-Off

Bass Roll-Off (switch 3) activates a -4 dB filter to the lowest bass frequencies (55 Hz). This can be used for compensating excessively heavy bass reproduction typically caused by loudspeaker placement near room boundaries.

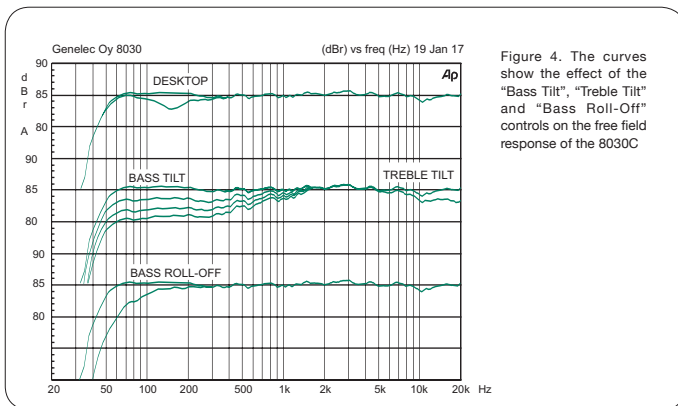


Figure 4. The curves show the effect of the "Bass Tilt", "Treble Tilt" and "Bass Roll-Off" controls on the free field response of the 8030C

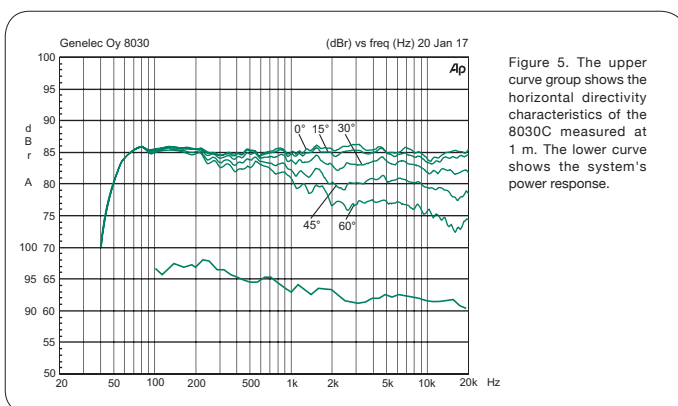


Figure 5. The upper curve group shows the horizontal directivity characteristics of the 8030C measured at 1 m. The lower curve shows the system's power response.

SYSTEM SPECIFICATIONS

Lower cut-off frequency, -6 dB: ≤ 47 Hz

Upper cut-off frequency, -6 dB: ≥ 25 kHz

Accuracy of frequency response:
54 Hz – 20 kHz (± 2.0 dB)

Maximum short term sine wave acoustic output on axis
in half space, averaged from 100 Hz to 3 kHz:
@ 1 m ≥ 104 dB SPL

Maximum long term RMS acoustic output in same
conditions with IEC weighted noise (limited by driver unit
protection circuit):
@ 1 m ≥ 96 dB SPL

Maximum peak acoustic output per pair @ 1 m distance
with music material:
 ≥ 110 dB

Self generated noise level in free field @ 1 m on axis:
 ≤ 5 dB (A-weighted)

Harmonic distortion at 85 dB SPL @ 1 m on axis:
Freq: 50...100 Hz < 2 %
>100 Hz < 0.5 %

Drivers: Bass 130 mm (5 in) cone
Treble 19 mm (7/8 in) metal dome

Weight: 5.0 kg (11 lb)

Dimensions: Height 299 mm (11^{3/8} in)
(including Iso-Pod™ table stand)

Height 285 mm (11^{1/4} in)
(without Iso-Pod™ table stand)

Width 189 mm (7^{1/2} in)

Depth 178 mm (7 in)

CROSSOVER SECTION

Connector: Input: XLR female, balanced 10 kOhm,
pin 1 gnd, pin 2 +, pin 3 -

Input level for 100 dB SPL output at 1 m: -6 dBu

Input sensitivity control range relative to max output:
-12 dB (Constantly variable)

Crossover frequency, Bass/Treble: 3.0 kHz

Treble Tilt control operating range:
0 to -2 dB @ 15 kHz

Bass Roll-Off control: -4 dB step @ 55 Hz

Bass Tilt control operating range in -2 dB steps:
0 to -6 dB @ 100 Hz

Desktop control: -4 dB @ 160 Hz

The 'CAL' position is with all tone controls set to 'off'
and the input sensitivity control to maximum (fully
clockwise).

AMPLIFIER SECTION

Bass amplifier output power: 50 W
Treble amplifier output power: 50 W

Long term output power is limited by driver unit
protection circuitry.

Amplifier system distortion at nominal output:
THD ≤ 0.05 %

Mains voltage: 100-240 V AC 50-60 Hz

Voltage operating range: ± 10 %

Power consumption:
Idle 3 W
Standby in ISS mode <0.5 W
Full output 60 W

B HEAD *Measurement System HMS III*

HEAD acoustics**Quality Certificate**

The system **HEAD Measurement System HMS III.0**
with **HDM I.0**
with serial number **13001475**
was checked and calibrated on **February 12, 2007**
The customer is **Fundacao de Ensino e Engenharia de Santa Catarina**

The results of the calibration are noted on the following page.
The system meets the requirements of HEAD acoustics GmbH.
The next calibration is recommended for February 2008.

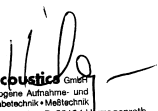
The reference equipment for the calibration was:

Pistonphone: Type 4228 Brüel & Kjaer, with serial number 1742019
The last calibration was done on June 6, 2006
by the Calibration Laboratory of Norsonic Tippkemper GmbH.
Calibration mark: 1404 DKD-K-18901 06-06

Analysis System: Analysis System ACQUA of HEAD acoustics GmbH
with MFE VI.1 of HEAD acoustics GmbH
with serial number 64626000.
The last calibration was done on March 2, 2006
by the Calibration Laboratory of HEAD acoustics GmbH.

The reference equipment is calibrated annually.

Herzogenrath, February 12, 2007



HEAD acoustics GmbH
Kopfbezogene Aufnahme- und
Wiedergabetechnik + Messtechnik
Ebertstraße 30a • D-52134 Herzogenrath
Tel. 02407 / 577-0 • Fax: 02407 / 577-99

Calibration Laboratory

HEADacoustics

Sheet 2/2

Calibration protocol

Calibration of microphone adjustment

During the calibration the air pressure amounted to	1010 hPa
The correction for sound pressure level is	-0.03 dB
The nominal sound pressure level of the pistonphone is	124.05 dB
The corrected sound pressure level of the pistonphone is	124.02 dB

	nominal value	actual value
Sound pressure level left (microphone MK221 32146)	124.02 dB ± 0.1 dB	124.02 dB
Sound pressure level right (microphone MK221 33735)	124.02 dB ± 0.1 dB	124.02 dB

Crosstalk attenuation (minimum {L,R})

measurement frequency	minimum value	actual value
250 Hz	115 dB	> 116 dB
1 kHz	115 dB	> 125 dB
5 kHz	115 dB	> 121 dB

Harmonic distortion (maximum {L,R})

measurement frequency	maximum value	actual value
250 Hz	0.006 %	< 0.004 %
1 kHz	0.006 %	< 0.004 %
5 kHz	0.006%	< 0.005 %

Noise (maximum {L,R})

	maximum value	actual value
Position 94 dB, 48 kHz LIN	16.5 dB _A	15.3 dB _A

Highpass 1 and 2

All variations are in the range of tolerances.

Equalizations ID, FF, DF

All variations are in the range of tolerances.

C *Hammerfall DSP System HDSP 9652 -*
RME

29. Technical Specifications

29.1 Digital Inputs

SPDIF - AES/EBU

- 1 x RCA, transformer-balanced, galvanically isolated, according to AES3-1992
- High-sensitivity input stage (< 0.3 Vpp)
- SPDIF compatible (IEC 60958)
- Accepts Consumer and Professional format, copy protection will be ignored
- Lock range: 28 kHz – 103 kHz
- Jitter when synced to input signal: < 3 ns

ADAT Optical

- 3 x TOSLINK, format according to Alesis specification
- Standard: 24 channels 24 bit, up to 48 kHz
- Double Speed (S/MUX): 12 channels 24 bit 96 kHz
- Bitclock PLL ensures perfect synchronisation even in varispeed operation
- Lock range: 31.5 kHz – 50 kHz
- Jitter when synced to input signal: < 3 ns

Word Clock

- BNC, not terminated (10 kOhm)
- Automatic Double Speed detection and internal conversion to Single Speed
- Not affected by DC-offsets within the network
- Overvoltage protection
- Level range: 1.0 Vpp – 5.6 Vpp
- Lock Range: 28 kHz – 103 kHz
- Jitter when synced to input signal: < 3 ns

29.2 Digital Outputs

SPDIF - AES/EBU

- 1 x RCA, transformer-balanced, galvanically isolated, according to AES3-1992
- Output level Professional 2.3 Vpp, Consumer 1.0 Vpp
- Format Professional according to AES3-1992 Amendment 4
- Format Consumer (SPDIF) according to IEC 60958
- Single Wire mode, sample rate 28 kHz up to 103 kHz

ADAT

- 3 x TOSLINK
- Standard: 24 channels 24 bit, up to 48 kHz
- Double Speed (S/MUX): 12 channels 24 bit 96 kHz

Word Clock

- BNC, max. output voltage: 5 Vpp
- Output voltage @ 75 Ohm termination: 4.0 Vpp
- Output impedance: 10 Ohm
- Frequency range: 27 kHz – 103 kHz

29.3 Digital

- Low Jitter Design: < 3 ns external clock, < 1 ns internal clock
- Internal sample rates: 32 / 44.1 / 48 / 88.2 / 96 kHz
- Supported sample rates through word clock: 28 kHz - 103 kHz
- Internal resolution: 24 bit
- Input PLL ensures zero dropout, even at more than 40 ns jitter
- Bitclock PLL for trouble-free varispeed ADAT operation
- Ground-free digital inputs and outputs

29.4 MIDI

- 2 x MIDI I/O via 5-pin DIN jacks
- Galvanically isolated by optocoupled input
- Hi-speed mode: Jitter and response time typically below 1 ms
- Separate 128 byte FIFOs for input and output

D Conversor AD/DA SQuadriga II - 3320



Features

- Operating modes:
 - stand-alone mode
 - front-end mode (via USB)
- USB 2.0 highspeed interface
 - SQuadriga II as USB mass storage device
- Recording, analysis, playback
- Recording format: 16 or 24 bit
- Sampling rates: 8; 16; 32; 44.1; 48; 51.2; 64; 88.2; 96 kHz
- Recording triggers (start/stop triggers from A/D signals, pulses, CAN or FlexRay signals, etc.)
- Removable SDHC memory card (4 GB, other capacities are optional)
- Large color graphics display respectively touchscreen (7.3 cm/2.83")
- FFT based real-time analyses (not in record mode): FFT, Octave, 3rd Octave, Time Signal, Level, Loudness, Spec. Loudness, Sharpness, Articulation Index
- Adjustable IIR filters (playback and monitoring - not in record mode)
- Display versus time of CAN, FlexRay, RPM and GPS quantities
- Auto Range (channel by channel or for all active channels)
- Limiter (up to 51.2 kHz)
- Real-time clock
- User documentation (ArtemiS SUITE compatible)
- Virtual SQuadriga configuration via the SQuadriga II Simulator
- Power supply via power adapter, car adapter, built-in re-chargeable battery (incl. battery level indicator) or replaceable standard batteries
 - Up to six hours of stand-alone operation capacity

Direct connectors

- USB 2.0
 - PC/notebook/tablet PC (Windows) with ArtemiS SUITE Data Acquisition Module ASM 04 (HEAD Recorder)
- BHS
 - Binaural headset BHS II (aurally accurate recordings, monitoring and playback)
- 6 x BNC
 - ICP/LINE
 - Analog In/Out
 - AC/DC
- 2 x Pulse (SMB; SMB > BNC adapters are included)
- Phones (equalized headphone output e.g. for the headphone HD IV.1)

Connections via adapters or adapter cables

- CAN/OBD-2
- FlexRay
- Artificial head HMS IV
- HEADlab modules labV6/labVF6, labM6
- Connecting a second SQuadriga II device
- Front end BEQ II
- GPS
- Pulse (pulse conditioning)
- High-impedance voltage sources (impedance converter)
- Remote control RC X

DATA SHEET

SQuadriga II (Code 3320)

Mobile recording and playback system - suitable as stand-alone system or USB front end

Overview

SQuadriga II is a 24 bit recording and playback system for a wide range of applications. Its compact dimensions and low weight, its built-in battery and its variety of connection possibilities provide for a high degree of mobility and functional versatility.

SQuadriga II saves time domain data on a removable SDHC card in stand-alone operation, whereas it saves directly to PC or notebook in front-end mode via USB. SQuadriga II can be conveniently operated via its function buttons and its touchscreen display, as well as via the HEAD Recorder software running on a connected computer.

ICP microphones and acceleration sensors, a calibrateable BHS II headset, pulse sensors, CAN, OBD-2 and FlexRay bus systems, an artificial head or other sensors can be connected directly or via adapters or adapter cables. SQuadriga II allows both the channel configuration and the sensor setup to be configured and saved on the device.

Without an external power source, the built-in rechargeable battery allows the extremely compact SQuadriga II unit to run for up to six hours. With additional standard batteries (AA type) inserted in the separate battery compartment, the operating time in stand-alone mode can be further extended.

SQuadriga II operates silently and is ready for operation immediately after turning it on.

Technical Data SQuadriga II

General

Number of channels (direct connections):	10 6 x Line /ICP In, BHS In (2-channel), 2 x Pulse In
Connections via adapters or adapter cables:	CAN/OBD-2, FlexRay, HMS III/HMS IV, HSU/BHM, Analog Out, BEQ II/PEQ V (ADAT In/Out), HEADiob (IobV6/IobV6), 2. SQuadriga II, Pulse Out, pulse conditioning for Pulse In, BHS via BNC, GPS, remote control RC X.1-V1, high-impedance voltage sources
Interfaces:	6 x BNC, 3 x LEMO 14-pin, 2 x SMB, 1 x USB, 1 x jack 3.5 mm
Resolution:	24 bit $\Delta\Sigma$ audio A/D and D/A converter
Input voltage:	5 V DC (+/-5 %), inverse-polarity protection
Power consumption	
Quick charging and operation:	12 W (max.)
Quick charging (device: off state):	5 W (max.)
Power supply via USB:	500 mA (max.), no charging, operation as bus-powered-device
Sampling frequencies (digital)	
Internal or external AES:	32; 44.1; 48; 51.2; 64; 88.2; 96 kHz (different sampling rates adjustable: each with $\frac{1}{2}$ and $\frac{1}{4}$ of f_s)
HEADiink:	48 kHz
External ADAT (opt):	48 kHz
TFT touchscreen:	Colour display 7.2 cm (2.83" / 43.2 x 57.4 mm), TFT, QVGA: 320 x 240 pixel
Memory (SDHC card HSD II.x):	4 GB; optional: 8 GB, 16 GB (file system FAT32)
Cooling:	Convection, no fan
Battery:	LiPo battery, 3.7 V, 5600 mAh
Charging time (ext. power supply):	7 h (max.), with 1000 mA
Operation time	
battery:	Typ. 6 h (stand-alone mode, 6 x ICP, writing on SDHC card)
battery and USB:	Typ. 16 h (500 mA via USB)
External battery:	4 x AA (Mignon), NiMH, Alkali, 1.25 V to 1.5 V, 2 h additional running time
Dimensions (rubber pads and BNC connectors, incl.):	148 x 182 x 45 mm (5.9" x 7.2" x 1.8") (WxDxH)
Weight (without batteries, with stylus and SD card):	970 g (2.14 lb)
Operating temperature:	-20 °C to 50 °C (-4 °F to 122 °F) (0 to 90 % r.h., non-condensing)
Storage temperature:	-20 °C to 70 °C (-4 °F to 158 °F)

BNC Inputs

Interfaces:	6 x BNC, switchable separately as analog outputs
Input impedance:	20 K Ω
Electric strength:	30 V _{eff} (max.)
ICP supply (switchable individually):	18 V, 2 mA (+/-10 %)
Ranges:	14 dB(M) 4 dB(M) -6 dB(M) -16 dB(M) -26 dB(M) -36 dB(M)
Level FS (V _{eff}):	10 V _{eff} 3.162 V _{eff} 1 V _{eff} 316 mV _{eff} 100 mV _{eff} 31.6 mV _{eff}
Level FS (V ₅₅):	28.3 V ₅₅ 8.94 V ₅₅ 2.83 V ₅₅ 894 mV ₅₅ 283 mV ₅₅ 89 mV ₅₅
S/N _{FS} :	97 dB 96 dB 96 dB 94 dB 87 dB 79 dB
	99 dB(A) 99 dB(A) 99 dB(A) 96 dB(A) 90 dB(A) 81 dB(A)
THD+N, at 1 kHz, -6 dB _{FS} :	-80 dB -88 dB -89 dB -88 dB -81 dB -72 dB
	0.001 % 0.004 % 0.004 % 0.004 % 0.009 % 0.003 %
Crosstalk, (sinus 1 kHz, same range, adj. channels):	110 dB 110 dB 110 dB 100 dB 100 dB 100 dB
Linearity:	117 dB 117 dB 118 dB 114 dB 103 dB 98 dB
Frequency range to 20 kHz:	+0.1 dB +0.1 dB +0.1 dB +0.1 dB +0.1 dB +0.1 dB
	-0.3 dB -0.3 dB -0.3 dB -0.3 dB -0.3 dB -0.5 dB
Analog filters	
HP filter (switchable):	DC / 2 Hz / 22 Hz (1st. order, +/-10 %)
LP filter:	30 kHz (anti-aliasing, 1st. order, oversampling: 256 times)
Usable bandwidth:	23.5 kHz at 48 kHz; 47.5 kHz at 96 kHz

Pulse Inputs

Interfaces:	2 x SMB
Galvanical isolation:	yes
Pulse frequency, max.:	600 kHz (at $f_s = 48$ kHz); 1 MHz (at $f_s = 96$ kHz)
Input voltage range	
Low-level:	0 V to +0,8 V
High-level:	+2,5 V to +5 V (max.)
Input impedance:	830 Ω for $U_{IH} = 2,5$ V 400 Ω for $U_{IH} = 5$ V

BHS Headset Inputs

Interfaces:	2 x LEMO 14-pin
Equivalent noise level with BHS II:	30 dB(A) _{SPL} (ID equalization)
Equalization:	ID
Ranges:	134 dB _{SPL} 124 dB _{SPL} 114 dB _{SPL} 104 dB _{SPL} 94 dB _{SPL}
S/N _{ref} , electrical:	99 dB(A) 99 dB(A) 96 dB(A) 90 dB(A) 81 dB(A)
Noise, SPL, electrical:	40 dB _{SPL} (A) 30 dB _{SPL} (A) 24 dB _{SPL} (A) 20 dB _{SPL} (A) 19 dB _{SPL} (A)
THD+N, at 1 kHz, -6 dB _{SPL} :	-82 dB -82 dB -83 dB -81 dB -72 dB 0.008 % 0.008 % 0.007 % 0.009 % 0,03 %
Crosstalk (sinus 1 kHz, same range, adj. channels):	>100 dB >100 dB >100 dB >100 dB >100 dB
Analog highpass filter (switchable):	2 Hz / 36 Hz
Frequency range up to 20 kHz:	+0.1 dB +0.1 dB +0.1 dB +0.1 dB +0.1 dB +0.1 dB

BNC Outputs

Interfaces:	6 x BNC
Level (max.):	-10 dB(V) + 6 dB headroom
Output impedance:	250 Ω
S/N:	96 dB / 99 dB(A)
THD+N:	-82 dB, at sinus 1 kHz, -1 dB _{SPL}
Crosstalk (DA>DA) :	>110 dB, at sinus 1 kHz
Frequency range:	10 Hz to 35 kHz (+0,1 dB / -0,2 dB) at $f_s = 96$ kHz (no DC)
Highpass filter	
analog:	1 Hz (1st. order)
digital:	higher order, scaled with f_s

BHS Headset Outputs

Interfaces:	2 x LEMO 14-pin
Nominal level (frequency dependent):	110 dB _{SPL} with BHS II
THD+N, electrical:	-57 dB at playback sinus 104 dB _{SPL} , 1 kHz at $R_L = 110 \Omega$
Frequency range:	-1 dB at 6 Hz; -1 dB at 20 kHz für $R_L = 110 \Omega$
Equalizations:	FF, ID, DF, LIN (no equalization)

Headphones Output

Interfaces:	1 x Jack 3,5 mm
Nominal level, (frequency dependent):	108 dB _{SPL} with HD IV.1
THD+N, electrical:	-57 dB at playback sinus 104 dB _{SPL} , 1 kHz at $R_L = 110 \Omega$
Frequency range:	-1 dB at 6 Hz; -1 dB at 20 kHz to $R_L = 110 \Omega$
Equalizations:	FF, ID, DF, LIN (no equalization)

USB Interface (USB 2.0 Highspeed)

Data transfer:	480 Mbits/s
----------------	-------------

AUX 1/ AUX 2 (extensions for SQuadriga II)

Interfaces:	2 x LEMO 10-pin
Connecting adapters and adapter cables:	SVA II.0, CLD VII.1, CLF VII, CLX VII.1, CLD VII.8, CLG VII, CLD VII.6, CLL VII.1, CLA VII.5

E Datasheet G.R.A.S. 46AE 1/2"

Conjunto padrão de microfone equalização de campo livre.

G.R.A.S. 46AE 1/2" CCP Free-field Standard Microphone Set Date 30-01-2015. Page 3 of 8

For proper sensitivity calibration we recommend using a reference sound source like the [G.R.A.S. 42AP](#) Intelligent Pistonphone.

Service

Should you by mistake damage the diaphragm on a G.R.A.S. microphone we will in most cases be able to exchange it at a very reasonable cost and short turn-around time. This not only protects your investment but also meets your quality assurance department since you do not have to worry about new serial numbers etc.

Calibration

When leaving the factory, all G.R.A.S. microphones have been calibrated in a controlled laboratory environment using traceable calibration equipment. Depending on the use, measurement environment and internal quality control programs we recommend that the microphone is recalibrated at least once a year.

Quality and warranty

G.R.A.S. microphone sets are made of components from our proven standard portfolio and are all manufactured of high-quality material and branded parts that were chosen and processed to ensure life-long stability and robustness.

All parts are manufactured and assembled at the factory in Denmark by skilled and dedicated operators in a verified clean-room environment. The microphone diaphragm, body and unique protection grid are made of high-grade stainless steel and make the microphone set resistant to physical damage as well as corrosion caused by aggressive air or gasses.

This, together with the enforced gold-plated microphone terminal guarantees a highly reliable connection. Thanks to the high quality, our warranty against defective materials and workmanship is 5 years.

Specifications

Frequency range (± 1 dB)	Hz	5 to 10 k
Frequency range (± 2 dB)	Hz	3.15 to 20 k
Dynamic range lower limit with G.R.A.S. preamplifier	dB(A)	17
Dynamic range upper limit with G.R.A.S. CCP preamplifier	dB	138
Set sensitivity @ 250 Hz (± 2 dB)	mV/Pa	50
Set sensitivity @ 250 Hz (± 2 dB)	dB re 1V/Pa	-26
Output impedance	Ω	< 50

G.R.A.S.
SOUND & VIBRATION

Conjunto padrão de microfone equalização de campo livre.

G. R. A. S. 46AE 1/2" CCP Free-field Standard Microphone Set
Date 30-01-2015. Page 4 of 8

Output Voltage Swing, min. @ 24-28 V CCP voltage supply	Vp	8
Power supply min. to max.	mA	2 to 20
DC bias voltage, typ.	V	12
Microphone venting		Rear
IEC 61094-4 Compliance		WS2P
Temperature range, operation	°C / °F	-30 to 70 / -22 to 158
Temperature range, storage	°C / °F	-40 to 85 / -40 to 185
Temperature coefficient @250 Hz	dB/°C / dB/°F	-0.01 / -0.006
Static pressure coefficient @250 Hz	dB/kPa	-0.007
Humidity range non condensing	% RH	0 to 95
Humidity coefficient @250 Hz	dB/% RH	-0.001
Influence of axial vibration @1 m/s ²	dB re 20 µPa	66
TEDS UTID (IEEE 1451.4)		27 v. 1.0
Connector type		BNC
CE/RoHS compliant/WEEE registered		Yes/Yes/Yes
Weight	g / oz	33 / 1.164

Test conditions

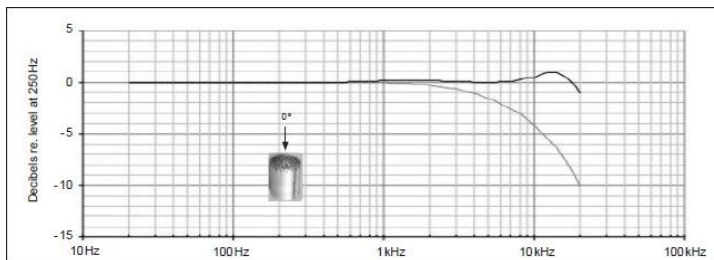
Temperature: 20C

Pressure: 760mm

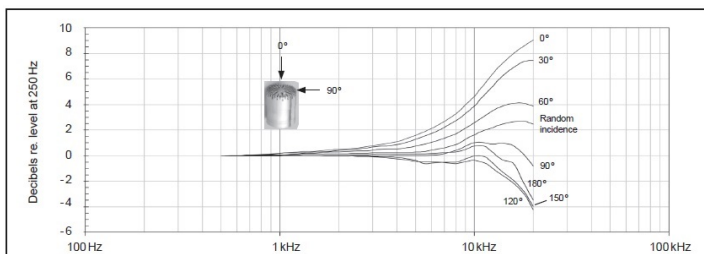
GRAS
SOUND & VIBRATION

Conjunto padrão de microfone equalização de campo livre.

G.R.A.S. 46AE 1/2" CCP Free-field Standard Microphone Set
Date 30-01-2015. Page 5 of 8



Typical frequency response. Upper curve shows free-field response for 0°, lower curve shows pressure response.



Free-field corrections for different angles of incidence

Dimensions

GR.A.S.
SOUND & VIBRATION

F *Datasheet* G.R.A.S. 46AQ 1/2"

Conjunto padrão de microfone equalização de campo difuso.

G.R.A.S. 46AQ 1/2" CCP Random Incidence Standard Microphone Set
Date 24-04-2014. Page 3 of 5

Before leaving the factory, all G.R.A.S. microphone sets are calibrated as a unit in a controlled laboratory environment using traceable calibration equipment. The sets are delivered with calibration charts including sensitivity values and frequency response graphs for the complete set. You can use the sensitivity value directly in your system setup.

Depending on the use, measurement environment, and internal quality control programs, we recommend recalibrating the microphone set at least every second year.

Learn more about G.R.A.S.' calibration services.

Specifications

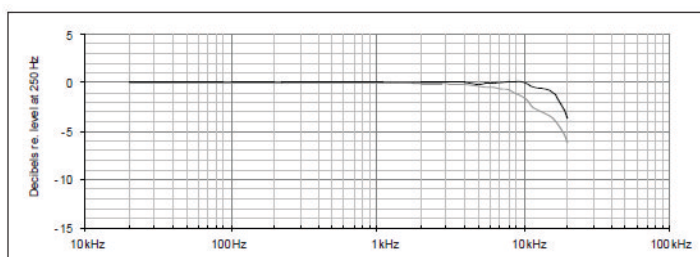
Frequency range (± 1 dB)	Hz	12.5 to 8 k
Frequency range (± 2 dB)	Hz	3.15 to 12.5 k
Frequency range (± 3 dB)	Hz	3.15 to 16 k
Dynamic range lower limit with G.R.A.S. preamplifier	dB(A)	17
Dynamic range upper limit with G.R.A.S. CCP preamplifier	dB	138
Set sensitivity @ 250 Hz (± 2 dB)	mV/Pa	50
Set sensitivity @ 250 Hz (± 2 dB)	dB re 1V/Pa	-26
Output impedance	Ω	< 50
Output Voltage Swing, min. @ 24-28 V CCP voltage supply	Vp	± 8
Power supply min. to max.	mA	2 to 20
Microphone venting		Rear
IEC 61094 Compliance		WS2P/D
Temperature range, operation	$^{\circ}\text{C} / ^{\circ}\text{F}$	-30 to 60 / -22 to 140
Temperature range, storage	$^{\circ}\text{C} / ^{\circ}\text{F}$	-40 to 85 / -40 to 185
Temperature coefficient @250 Hz	dB/ $^{\circ}\text{C}$ / dB/ $^{\circ}\text{F}$	-0,01
Static pressure coefficient @250 Hz	dB/kPa	-0,011
Humidity range non condensing	% RH	0 to 100
Humidity coefficient @250 Hz	dB/% RH	-0,001

GRAS
SOUND & VIBRATION

Conjunto padrão de microfone equalização de campo difuso.

G. R. A. S. 46AQ 1/2" CCP Random Incidence Standard Microphone Set
Date 24-04-2014. Page 4 of 5

Influence of axial vibration @1 m/s ²	dB re 20 µPa	62
TEDS UTID (IEEE 1451.4)		27 v. 1.0
Connector type		BNC
CE/RoHS compliant/WEEE registered		Yes/Yes/Yes
Weight	g / oz	33 / 1.164



Typical frequency response. Upper curve shows response in a diffuse sound field (random incidence), lower curve shows pressure response.


Ordering info

Pistonphone	A2AA
Intelligent pistonphone	A2AP
1-Channel CCP Power Module with A-Weighting filter	L2AL
2-Channel CCP/LEMO Power Module with Signal Conditioning and Computer Interface	L2AQ
3 m BNC-to-BNC cable	AA0035
10 m BNC-to-BNC cable	AA0037

GRAS
SOUND & VIBRATION

G *Datasheet* Calibrador B&K Tipo 4231
hardware.

Compliance with Standards

	<p>The CE marking is the manufacturer's declaration that the product meets the requirements of the applicable EU directives RCM mark indicates compliance with applicable ACMA technical standards – that is, for telecommunications, radio communications, EMC and EME China RoHS mark indicates compliance with administrative measures on the control of pollution caused by electronic information products according to the Ministry of Information Industries of the People's Republic of China WEEE mark indicates compliance with the EU WEEE Directive</p>
<p>Safety</p>	<p>EN/IEC 61010-1: Safety requirements for electrical equipment for measurement, control and laboratory use. ANSI/UL 61010-1: Safety requirements for electrical equipment for measurement, control and laboratory use.</p>
<p>EMC Emission</p>	<p>EN/IEC 61000-6-3: Generic emission standard for residential, commercial and light industrial environments. EN/IEC 61000-6-4: Generic emission standard for industrial environments. CISPR 22: Radio disturbance characteristics of information technology equipment. Class B Limits. FCC Rules, Part 15: Complies with the limits for a Class B digital device. EN/IEC 60942: Instrumentation Standard – Electroacoustics – Sound Calibrators.</p>
<p>EMC Immunity</p>	<p>EN/IEC 61000-6-1: Generic standards – Immunity for residential, commercial and light industrial environments. EN/IEC 61000-6-2: Generic standards – Immunity for industrial environments. EN/IEC 61326: Electrical equipment for measurement, control and laboratory use – EMC requirements. EN/IEC 60942: Instrumentation Standard – Electroacoustics – Sound Calibrators. Note: The above is only guaranteed using accessories listed in this Product Data sheet.</p>
<p>Temperature</p>	<p>IEC 60068-2-1 & IEC 60068-2-2: Environmental Testing. Cold and Dry Heat. Operating Temperature: -10 to +50°C (14 to 122°F) Storage Temperature: -25 to +70°C (-13 to +158°F)</p>
<p>Humidity</p>	<p>IEC 60068-2-78: Damp Heat: 90% RH (non-condensing at 40°C (104°F)).</p>
<p>Mechanical</p>	<p>Non-operating: IEC 60068-2-6: Vibration: 0.3 mm (10 to 58 Hz), 20 m/s² (58–500 Hz) IEC 60068-2-27: Shock: 1000 m/s² IEC 60068-2-29: Bump: 3000 bumps at 400 m/s²</p>
<p>Enclosure</p>	<p>IEC 60529: Protection provided by enclosures: IP 50 with leather protection case.</p>

Specifications – Sound Calibrator Type 4231

STANDARDS SATISFIED

EN/IEC 60942 (2003), Class LS and Class 1, Sound Calibrators
ANSI S1.40 – 2006, Specification for Acoustic Calibrators Class LS and Class 1

SOUND PRESSURE LEVELS

94.0 dB ±0.2 dB (Principal SPL) or
114.0 dB ±0.2 dB re 20 µPa at reference conditions

FREQUENCY

1 kHz ±0.1%

SPECIFIED MICROPHONE

Size according to IEC 61094-4:

- 1" without adaptor
- 1/2" with adaptor UC 0210 (supplied)
- 1/4" with adaptor DP 0775 (optional)
- 1/8" with adaptor DP 0774 (optional)

EQUIVALENT FREE-FIELD LEVEL

(0° incidence, re Nominal Sound Pressure Level)
-0.15 dB for 1/2" Brüel & Kjær microphones. See the Type 4231 User Manual for other microphones

EQUIVALENT RANDOM INCIDENCE LEVEL

(re Nominal Sound Pressure Level)
+0.0 dB for 1", 1/2", 1/4" and 1/8" Brüel & Kjær microphones

NOMINAL EFFECTIVE COUPLER VOLUME

> 200 cm³ at reference conditions

DISTORTION

< 1%

LEVEL STABILITY

Short-term: Better than 0.02 dB (as specified in IEC 60942)
One Year: Better than 0.05 dB (σ = 96%)
Stabilization Time: < 5 s

REFERENCE CONDITIONS

Temperature: 23°C ±3°C (73° ±5°F)
Pressure: 101 ±4 kPa
Humidity: 50%, -10%+15% RH
Effective Load Volume: 0.25 cm³

ENVIRONMENTAL CONDITIONS

Pressure: 65 to 108 kPa
Humidity: 10 to 90% RH (non-condensing)
Effective Load Volume: 0 to 1.5 cm³

INFLUENCE OF ENVIRONMENTAL CONDITIONS (Typical)

Temperature Coefficient: ±0.0015 dB/°C
Pressure Coefficient: +8 × 10⁻⁴ dB/kPa
Humidity Coefficient: 0.001 dB/% RH

POWER SUPPLY

Batteries: 2 × 1.5 V IEC Type LR6 ("AA" size)
Lifetime: Typically 200 hours continuous operation with alkaline batteries at 23°C (73°F)
Battery Check: When Type 4231 stops working continuously, and only operates when the On/Off button is held in, the batteries should be replaced

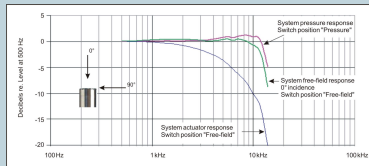
H *Datasheet* G.R.A.S. 40AH 1".

Specifications

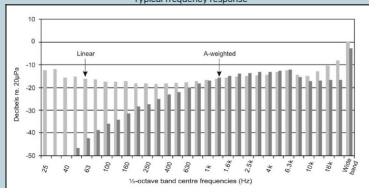
GRAS 40HF 1” Lownoise Microphone System

Page: 3

Frequency range (-1 dB)	Hz	12.5 to 4k
Frequency range (+2 dB)	Hz	10 to 10k
Frequency range (+3 dB)	Hz	6 to 12.5k
Dynamic range lower limit (microphone thermal noise)	dB(A)	-2
Dynamic range upper limit	dB	110
Set sensitivity @ 250 Hz (±2 dB)	mV/Pa	1100
Polarization voltage	V	200 V / Traditional
Microphone venting		Rear
Output impedance		47
Temperature range, operation	°C / °F	-20 to 60 / -4 to 140
Temperature range, storage	°C / °F	-40 to 85 / -40 to 185
Connector type		BNC
CE/RoHS compliant/WEEE registered		Yes/Yes/Yes
Weight	g / oz	620 / 22



Specification Conditions



Typical noise floor shown in 1/3-octave bands for both the linear and A-weighted cases. GRAS Sound & Vibration reserves the right to change specifications and accessories without notice.

Índice Remissivo

- Loudness*, 36
- Abrangência, 70
- Acústica, Biauricular, 33
- Anamnese, 35
- Arranjo, 41
- Audiometria, 35
- Células ciliadas, 34
- Cóclea, 34
- Correlação, 42
- Decaimento, 51
- Eletrofisiológica, 35
- Ensaio
 - Piloto, 75
 - Principal, 82
 - Subjetivo, 74, 82
- Erro
 - Aleatório, 41
 - Constante, 41
 - Localização, 41
- Estereofonia, 44
- Fala
 - Índice, 35
 - Inteligibilidade, 35
- Limiar, 35
- Percepção, 35
- Filtro
 - Digital, 55
- Fontes
 - Reais, 42
 - Virtuais, 43
- Frequências, 73
- Função
 - Resposta em frequência, 51
 - Transferência, 51
- Grau perda auditiva
 - Profundo, 34
 - Severo, 34
- Implante
 - Coclear, 34
- Impulso, 51
- Laterização, 38
- Localização, 37
- Mascarador, 39
- Nível
 - Desconforto, 36

- Pressão sonora, 36
- Nervo auditivo, 34
- Osso temporal, 34
- Panorama
 - Amplitude, 43
 - Intensidade, 46
- Psicoacústica, 55
- Reflexões, 52
- Repetibilidade, 84
- Reverberação, 51
- Ruído
 - Branco, 40
 - de fundo, 51
- Síntese, 43
- Semi-anecoica, 48
- Subjetivo(a), 74
- Tecnologias assistivas, 33
- Transdutores, 42
- Treinamento, 78
- Triangulação, 41
- Virtualização, 42, 74

