

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
ESCOLA DE ENGENHARIA  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA  
CURSO DE GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**GUILHERME OLIVEIRA DA ROSA**

**ALGORITMO PARA  
BINAURALIZAÇÃO BIDIMENSIONAL  
DE SONS PARA FONES DE OUVIDO**

Porto Alegre  
2016

**GUILHERME OLIVEIRA DA ROSA**

**ALGORITMO PARA  
BINAURALIZAÇÃO BIDIMENSIONAL  
DE SONS PARA FONES DE OUVIDO**

Projeto de Diplomação apresentado ao Departamento de Engenharia Elétrica da Universidade Federal do Rio Grande do Sul como parte dos requisitos para a obtenção do título de Engenheiro Eletricista.

**ORIENTADOR: Prof. Dr. Adalberto Schuck Jr.**

Porto Alegre  
2016

**GUILHERME OLIVEIRA DA ROSA**

**ALGORITMO PARA  
BINAURALIZAÇÃO BIDIMENSIONAL  
DE SONS PARA FONES DE OUVIDO**

Este Projeto foi julgado adequado para a obtenção dos créditos da Disciplina Projeto de Diplomação do Departamento de Engenharia Elétrica e aprovado em sua forma final pelo Orientador e pela Banca Examinadora.

Orientador: \_\_\_\_\_  
Prof. Dr. Adalberto Schuck Jr., UFRGS  
Doutor pela Universidade Federal do Rio Grande do Sul –  
Porto Alegre, Brasil

Banca Examinadora:

Prof. Dr. Adalberto Schuck Jr., UFRGS  
Doutor pela Universidade Federal do Rio Grande do Sul – Porto Alegre, Brasil

Prof. Dr. Tiago Roberto Balen, UFRGS  
Doutor pela Universidade Federal do Rio Grande do Sul – Porto Alegre, Brasil

Prof. Dr. Luiz Fernando Ferreira, UFRGS  
Doutor pela Universidade Federal do Rio Grande do Sul – Porto Alegre, Brasil

Chefe do DELET: \_\_\_\_\_  
Prof. Dr. Ály Ferreira Flores Filho

Porto Alegre, dezembro de 2016.

## **DEDICATÓRIA**

Dedico esse trabalho aos meus colegas amigos que muito me ajudaram ao longo da graduação.

## **AGRADECIMENTOS**

Gostaria de agradecer aos familiares mais próximos - Juçara Franco de Oliveira (mãe); Nelson Edi da Rosa (pai); Gustavo Oliveira da Rosa (irmão) - por sempre estarem ao meu lado nos momentos mais difíceis, nos auxiliando de alguma forma nessa longa jornada nomeada projeto de diplomação.

Ao meu orientador e Professor Dr. Adalberto Schuck Jr, por acreditar em minha capacidade e aptidão. Sem ele não seria possível a conclusão do projeto. Obrigado pelos conselhos, auxílios e tempo cedido.

Ao meu fiel e inseparável companheiro de todas as noites: Senhor Café.

E finalmente a todos aqueles, que apesar de não citados aqui contribuíram para a conclusão do projeto de formas indiretas, tornando aquilo que alguns consideravam um sonho, uma realidade.

## RESUMO

Este trabalho tem por objetivo a aplicação de um algoritmo, utilizando a abordagem contemplada pelo modelo de Brown e Duda, que aplique técnicas de binauralização em sinais sonoros para fones de ouvido. Não foram adotadas técnicas para adicionar o sentido de elevação, desse modo o trabalho caracterizou-se apenas em um plano bidimensional. Foram utilizados dois indicadores para a síntese da técnica: ITD e ILD, indicadores de diferença de tempo e intensidade, respectivamente. Para aplicação da ILD, utilizou-se um filtro de primeira ordem, baseado em um modelo de espalhamento de Rayleigh e modelado para o fenômeno de sombra acústica que ocorre em altas frequências. Para o indicador ITD, utilizou o modelo de Woodworth, revisado e corrigido por Aaronson e Hartmann, que adiciona atraso ao sinal em função do ângulo incidente da onda emitida.

Foram utilizadas técnicas de discretização no modelo para implementação de um filtro IIR. A implementação foi realizada com sucesso, visto que as curvas da resposta em frequência obtidas coincidem com o modelo de Brown e Duda. O modelo foi comparado à um catálogo de respostas impulsivas, obtidos através de análise experimental em um manequim do tipo KEMAR. Utilizando o catálogo percebe-se uma melhor distinção sonora, contudo, mais custoso computacionalmente

**Palavras-chave: Binauralização, diferenças de tempo interaurais, diferenças de Intensidade interaurais, modelo de woodworth.**

## ABSTRACT

This work aims at the application of an algorithm, using the approach contemplated by the Brown and Duda model, that applies binauralisation techniques in sound signals to headphones. None techniques were used to add the sense of elevation, so the work has application only in a two-dimensional plane. Two indicators were used for the synthesis of the technique: ITD and ILD, indicators of time difference and intensity, respectively. For the ILD application, a first order filter was used, based on a Rayleigh scattering model and modeled for the sound shadow phenomena that occur at high frequencies. For the ITD indicator, was used the Woodworth model, revised and corrected by Aaronson and Hartmann, which adds delay to the signal as a function of the incident angle of the emitted wave.

Discretization techniques were used in the model to implement an IIR filter. Success was achieved in implementation, since the frequency response curves obtained coincide with the Brown and Duda model. The model was compared to a catalog of impulsive responses, obtained through an experimental analysis on a KEMAR dummy. Using the catalog is perceptible a better sound distinction, however, it is more computationally costly.

**Keywords: Binauralization, Interaural time differences, Interaural level differences, Woodworth Model.**

## LISTA DE ILUSTRAÇÕES

1	Sombra Acústica . . . . .	17
2	Sistema auditivo dividido em 4 partes . . . . .	18
3	Exemplo de uma onda sonora atingindo a cabeça de uma pessoa.. . .	19
4	Diferença de tempo. . . . .	19
5	Azimuth percebido de uma fonte de som virtual. a) Diferenças de intensidade e b) diferenças de tempo.Fonte: (ZOLZER, 2011) . . . .	20
6	Cone delimitando a zona de confusão em um sistema interaural de coordenadas polares. Fonte: (ZOLZER, 2011) . . . . .	21
7	Diferença de tempo interaural dependente da frequência para um ângulo de azimuth $\alpha=30^\circ$ . Fonte: (ZOLZER, 2011) . . . . .	22
8	Diferença de intensidade interaural dependente da frequência para um ângulo de azimuth $\alpha=30^\circ$ . Fonte: (ZOLZER, 2011) . . . . .	22
9	Imagem do som percebida quando escutando um ruído de banda larga com grau de correlação igual a a) 1 e b) 0. . . . .	23
10	Exemplo de um manequim do tipo KEMAR. . . . .	24
11	Comparativo entre convolução e correlação. . . . .	26
12	Respostas em frequência de uma esfera rígida ideal para a incidência de uma onda com ângulo de incidência $\theta$ variando entre $0^\circ$ e $180^\circ$ . Fonte: (BROWN; DUDA, 1998) . . . . .	27
13	Influência do ângulo de incidência $\theta$ no fator $\alpha$ . . . . .	29
14	Respostas em frequência para o modelo de uma cabeça esférica, utilizando um filtro de primeira ordem como aproximação, com uma onda com ângulo de incidência $\theta$ variando entre $0^\circ$ e $180^\circ$ . Fonte: (BROWN; DUDA, 1998) . . . . .	29
15	Atraso no tempo oriundos de reflexões no ombro em função do azimuth ( $\theta$ ) e elevação( $\phi$ ). . . . .	32
16	Carregamento e armazenamento das HRIRs obtidas experimentalmente.	34
17	Leitura e extração dos dados amostrados e taxa de amostragem do arquivo de áudio. . . . .	34
18	Aplicação do método da convolução. . . . .	35
19	Saída do processamento. . . . .	35
20	Atribuição das variáveis de interesse para o aplicação do script teste .	36
21	Método utilizado para apresentar, graficamente, o resultado aplicado pelo algoritmo que utiliza o catálogo de HRIRs. . . . .	36
22	Script utilizado para gerar um diagrama de bode para qualquer ângulo de incidência . . . . .	37



23	Ajuste da variável de interesse $\theta$ de acordo com a orelha para qual se deseja aplicar o filtro. . . . .	38
24	Implementação da posição do zero em função do azimute $\theta$ e atribuição do numerador e denominador da função de transferência . . . . .	38
25	Script teste para a implementação do fenômeno de sombreamento da cabeça. . . . .	38
26	Implementação da diferença de tempo interaural em função de $\theta$ . . . . .	39
27	Algoritmo completo do modelo. . . . .	40
28	Modelo completo desenvolvido. . . . .	40
29	Amplitudes do canal stereo geradas através do MATLAB para um ruído branco. . . . .	41
30	Diagrama de amplitudes para o canal stereo da música clássica utilizada. . . . .	42
31	Diagrama de amplitudes após processadas pelo algoritmo. . . . .	42
32	Diagrama de bode obtido para o modelo proposto por Brown e Duda. Para o primeiro gráfico utilizou-se como $\theta_{min} = 150$ e para o segundo, 180. . . . .	43
33	Implementação do filtro que aplica o efeito de sombreamento da cabeça utilizando um filtro IIR. . . . .	44
34	Filtro IIR para a orelha direita. . . . .	45
35	Filtro IIR para a orelha esquerda . . . . .	45
36	Implementação de atraso (ITD) ao sinal emitido em função do ângulo de incidência $\theta$ . . . . .	46
37	Diagramas de amplitude ao longo do tempo para um ruído branco quando submetido ao sistema completo . . . . .	47
38	Script completo que aplica método da convolução direta com catálogo de HRIRs. . . . .	50
39	Script teste para o algoritmo que aplica o método da convolução com o catálogo de HRIRs. . . . .	51
40	Algoritmo para aplicar fenômeno de sombreamento da cabeça. . . . .	52
41	Script teste para o algoritmo que aplica o filtro de sombreamento da cabeça . . . . .	53
42	Script utilizado para adicionar atraso (ITD) . . . . .	54
43	Script utilizado para apresentar atraso (ITD) gerado pelo algoritmo da Figura 42 . . . . .	54

## **LISTA DE TABELAS**

1	Número de medições e incremento em cada nível de elevação . . . .	25
2	Especificações técnicas do headphone Philips - SHL3565BK/00 . . . .	33

## LISTA DE ABREVIATURAS

ITD	Interaural Time Difference
ILD	Interaural level Difference
HRTF	Head-Related Transfer functions
HRIR	Head-Related Impulse Responses
GUI	Graphical User Interface
FIR	Finite Impulse Response
IIR	Infinite Impulse Response
MATLAB	Matrix LABoratory
MLS	Maximum Length Sequence
FTP	File Transfer Protocol.

## LISTA DE SÍMBOLOS

- $\theta$  Ângulo de incidência em relação ao azimute.
- $\alpha$  Fator de controle da localização do zero na HRTF
- $c$  velocidade do som
- $\lambda$  comprimento de onda
- $\omega$  frequência angular da onda
- $a$  raio da cabeça humana
- $f$  frequência da onda
- $d$  diâmetro da cabeça

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	14
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	16
2.1	Psicoacústica	16
2.2	Princípios da binauralidade	18
2.3	Indicadores de diferenças interaurais	20
2.3.1	Diferença de tempo interaural	20
2.3.2	Diferença de intensidade interaural	22
2.4	Externalização e Correlação	23
2.5	Análise Experimental utilizando um Manequim	24
2.6	Modelos estruturais	25
2.6.1	Modelo esférico de Rayleigh	26
2.6.2	Modelo Cabeça-Esfera Aproximado	27
2.7	Funções de transferência "Head-Related" - HRTF	29
2.7.1	Modelando filtros HRTF	30
<b>3</b>	<b>METODOLOGIA EXPERIMENTAL</b>	33
3.1	Método da convolução utilizando HRIRs obtidas experimentalmente	34
3.2	Sistema Proposto: Algoritmo para um modelo cabeça-esfera	36
3.2.1	Validação do modelo obtido por Brown e Duda	37
3.2.2	Filtro para fenômeno de Sombreamento da Cabeça	37
3.2.3	implementação de delay para o indicador ITD	39
3.2.4	Algoritmo para implementação do modelo completo.	39
<b>4</b>	<b>RESULTADOS E DISCUSSÕES</b>	41
4.1	Utilizando o catálogo de respostas impulsivas HRIRs	41
4.2	Utilizando o sistema proposto	43
4.2.1	Apresentação da validação do modelo de Brown e Duda	43
4.2.2	Apresentação do filtro que aplica o fenômeno de sombreamento da cabeça	44
4.2.3	Apresentação do algoritmo que implementa atraso	46
4.2.4	Apresentação da implementação do sistema completo	46
<b>5</b>	<b>DISCUSSÃO E CONCLUSÕES</b>	48
<b>6</b>	<b>TRABALHOS FUTUROS</b>	49

<b>7</b>	<b>ANEXOS</b>	<b>50</b>
7.1	Script que aplica método da convolução utilizando HRIRs previamente carregados e obtidas experimentalmente	50
7.2	Script teste para o algoritmo ilustrado na seção 7.1	51
7.3	Script que implementa algoritmo para o fenômeno de sombreamento da cabeça	52
7.4	Script teste para o algoritmo ilustrado na seção 7.3	53
7.5	Script para implementar atraso ITD	54
	<b>REFERÊNCIAS</b>	<b>55</b>

# 1 INTRODUÇÃO

Os seres humanos ainda estão um pouco longe dos morcegos quando se está em análise o sentido auditivo. Pois estes mamíferos voadores, além de conseguirem escutar até 80kHz, possuem o que chamamos de "ecolocalização", eles conseguem detectar a localização de objetos através da reflexão de ondas por eles emitidas. Contudo, a raça humana também possui a habilidade de poder localizar fontes sonoras em um espaço de três dimensões, utilizando diversos indicadores e métodos para tal. Uma grande parte da forma como assimilamos o ambiente que nos rodeia depende de nossa capacidade de determinar a localização de distintas fontes sonoras.

Os efeitos físicos da difração de ondas sonoras em uma cabeça humana é bem conhecido pela literatura (CARLIE, 1996). Quando um som produzido atinge nosso sistema auditivo, o cérebro humano processa aquela informação para conseguir identificar, da melhor forma possível, a localização da fonte. Indicadores de intervalo de tempo entre a chegada do som nas orelhas e de intensidade sonora, por exemplo, são utilizados pelo cérebro para conseguir realizar a correta localização de um sinal sonoro que foi emitido. Uma vez que se conheça estes indicadores, é possível realizar-se a modelagem estrutural do sistema que influencia, significativamente, na recepção do som, isto é a cabeça, o torso e a orelha. Deste modo, assumindo-se que o ouvinte estará recebendo o material sonoro via um "headphone" stereo, é possível reproduzir estes indicadores e simular a ação deste sistema cabeça-torso-orelha para que torne-se possível a síntese e injeção deste sinal filtrado e processado diretamente nas orelhas.

Gardner e Martin obtiveram e publicaram em (GARDNER; MARTIN, 1994) um conjunto de respostas impulsivas para a cabeça humana através de experimentos em um manequim. Contudo, por motivos como desempenho computacional e limitação de adequações individuais da cabeça humana, tornou-se deveras importante o estudo para uma nova forma de obter uma função de transferência para a cabeça quando submetida a uma onda sonora. Brown e Duda em (BROWN; DUDA, 1998) utilizam-se dessa motivação para pesquisar e desenvolver este novo modelo. Baseando-se em um modelo de Rayleigh, detalhado em (RAYLEIGH), 1904), eles modelam um novo filtro com o objetivo de ser menos custoso computacionalmente e adequar-se facilmente, às individualidades da cabeça humana das pessoas.

Muito estudo tem sido realizado nessa área devido ao enorme aumento de desenvolvimento de dispositivos com tecnologia imersiva. Jogos com realidade aumentada e áudio 3D são dois excelentes exemplos.. Diversas áreas possuem interesse no avanço dessa tecnologia: Aplicações multimídia, Interfaces Homem-Computador, video games, sistemas de realidade aumentada, áudio espacial para teleconferências (CIPIC, 2012). Empresas como a "Waves Audio", que desenvolvem tecnologias de áudio para conjuntos musicais como U2; Rolling Stones e Madonna, desenvolvem pesquisa em processamento de áudio

e produtos com a finalidade de maior imersão do usuário e causar a sensação de tridimensionalidade do áudio em um headphone.

Muitos pesquisadores tem proposto métodos para substituir esse catálogo de respostas impulsivas (MYERS, 1989) (DUDA, 1993). O objetivo desse trabalho é a implementação de um algoritmo, utilizando a ferramenta computacional MATLAB, que utilize o método de binauralização proposto por Brown e Duda para a manipulação de um sinal sonoro ao longo do azimute que circunda a cabeça humana em um plano bidimensional. O resultados desta implementação serão comparados com o conjunto de respostas impulsivas obtidos por Gardner e Martin (GARDNER; MARTIN, 1994). A comparação será realizada através da audição, diagramas de amplitude e espectros de frequência, também gerados no MATLAB. Toda a análise será realizada utilizando Headphone e não serão adotados técnicas ou métodos para aplicar sentido de elevação, o que adicionaria uma nova dimensão ao sistema. A estrutura do trabalho segue com 4 partes principais:

- **Fundamentação Teórica:** Revisão breve da teoria e conceitos abordados no trabalho. Discussão sobre o material de pesquisa
- **Metodologia Experimental:** Explicação do conjunto de procedimentos adotados no desenvolvimento do trabalho efetuado
- **Resultados e Discussões:** Apresentação clara dos resultados obtidos utilizando-se a metodologia experimental com observações e discussões pertinentes.
- **Conclusões Considerações finais** sobre o trabalho desenvolvimento.



## 2 FUNDAMENTAÇÃO TEÓRICA

### 2.1 Psicoacústica

A sensação de intensidade interpretada pelo aparelho auditivo humano é de, aproximadamente, a raiz cúbica da intensidade real do som. Visto isso, se houver uma variação de 1 bilhão de vezes na energia do som, o sistema auditivo interpretará como uma variação de, aproximadamente, mil vezes. Esta compressão permite ao ser humano distinguir variações de uma faixa muito mais ampla.

Devido a essa capacidade de distinção que o ouvido pode detectar, frequentemente utiliza-se escalas logarítmicas para expressar intensidade sonora. Para o aumento de 10 vezes na intensidade sonora é atribuído-se 1 *bel*, enquanto que 0,1 *bel* chama-se *decibel*. Esta conversão é expressada pela Equação 1:

$$P_{DB} = 10 \log_{10} \frac{P}{P_0} \quad (1)$$

Desse modo, conforme a Equação 1, o aumento de 1dB proporciona um aumento de 1,26 vezes da intensidade sonora em uma escala linear. (STANDARD, 2000) O sistema auditivo utiliza pelo menos três métodos para a detecção da intensidade sonora (GUYTON; JOHN E. HALL, 2011).

- 1 - Conforme aumenta-se a intensidade do som, a amplitude do padrão de vibração da membrana basilar, que é responsável por suportar e acomodar os receptores sensoriais, também aumenta e desse modo, estes receptores (células ciliadas) do sistema auditivo excitam as terminações nervosas com uma frequência, conseqüentemente, maior;
- 2 - Conforme aumenta-se a intensidade do som, faz com que mais células ciliadas, que estejam nos arredores da parte em ressonância na membrana basilar, sejam estimuladas. Desse modo, ocorrerá transmissão através de muitas fibras nervosas e não poucas. Isto é detectável pelo sistema auditivo.
- 3 - As células ciliadas externas não são estimuladas substancialmente até que a vibração, oriunda da membrana basilar, seja alta o suficiente. Desse modo, o estímulo dessas células notifica o sistema nervoso quanto a intensidade sonora exercida sobre o canal auditivo.

O ser humano consegue determinar a localização de uma fonte sonora a partir de dois fatores principais:

- O intervalo de tempo de chegada do sinal sonoro emitido entre cada um dos ouvidos;

- A diferença entre a intensidade de som nos dois ouvidos.

O primeiro método, que atua nas diferenças de tempo, funciona melhor em frequências inferiores a 3000 Hz. Por outro lado, o segundo método, que atua nas diferenças de intensidade, é melhor aproveitado em frequências superiores a 3000 Hz (GUYTON; JOHN E. HALL, 2011).

O primeiro método distingue a localização do som de uma forma mais eficiente que o segundo método para a intensidade. Isto se deve ao fato de que o primeiro método depende, apenas, do intervalo de tempo entre a chegada do som nos ouvidos. A direção em que a pessoa está olhando influi diretamente nisso, visto que conforme a pessoa move a cabeça, a distância percorrida pela onda sonora até os ouvidos muda.

A intensidade do som que percebemos é reduzida quando o sinal emitido ocorre em altas frequências para pontos distantes da orelha. Isto faz com que a cabeça cause o fenômeno de sombra acústica. Deste modo, o segundo método torna-se mais eficaz em altas frequências, pois em baixas o som não é muito atenuado pela cabeça. I

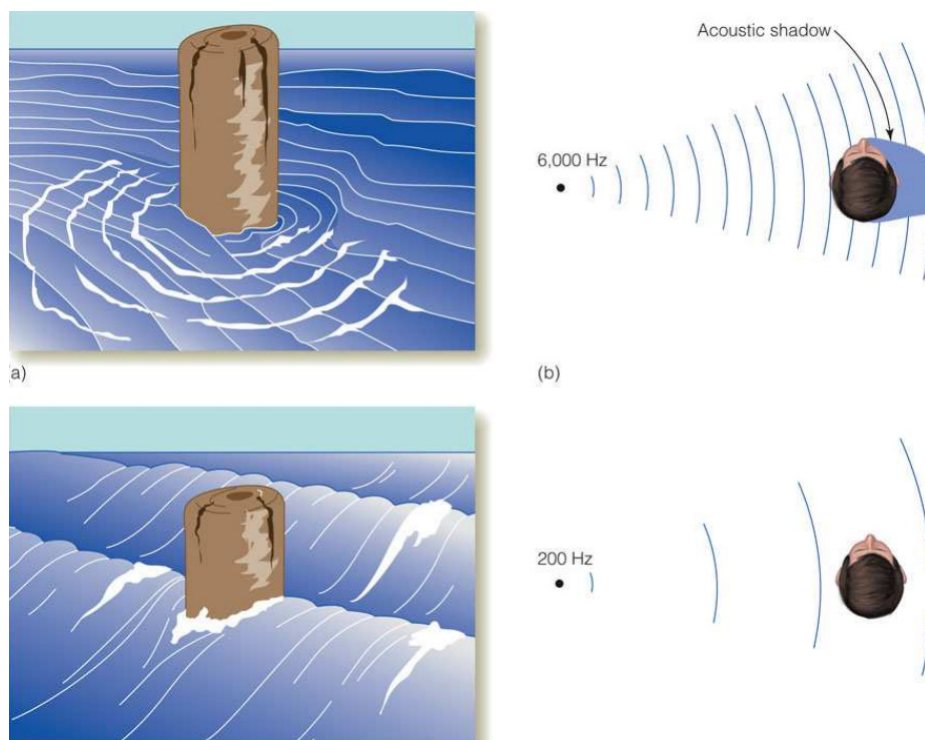


Figura 1: Sombra Acústica

Tanto o método que depende do intervalo de tempo quanto o que atua em função da intensidade da onda sonora, não podem dizer com certeza, se o som é proveniente da região frontal ou traseira da pessoa. Isso nos próximos capítulos será explicado. A esse fenômeno se dá o nome de "Cone da Confusão". Para realizar essa distinção o corpo humano utiliza os chamados *Pavilhões auditivos* dos dois ouvidos, também conhecido como orelha. A orelha muda a qualidade do som dependendo da direção da qual está localizada a fonte sonora. Destes modo, pequenos movimentos da cabeça são, na grande maioria dos casos, mais do que o suficiente para conseguir distinguir quando um som está vindo de frente ou de trás, escapando assim do cone da confusão.

Por possuímos duas orelhas é fácil localizar a direção de uma fonte sonora, visto a diferença entre a chegada do som em cada canal auditivo. Deste modo, torna-se trivial

para nosso cérebro calcular a provável localização do som emitido. Após o som chegar em nosso pavilhão auditivo, o som é conduzido pelo canal auditivo até atingir o tímpano, fazendo-o vibrar. A Figura 2 apresenta este percurso percorrido pelo som.

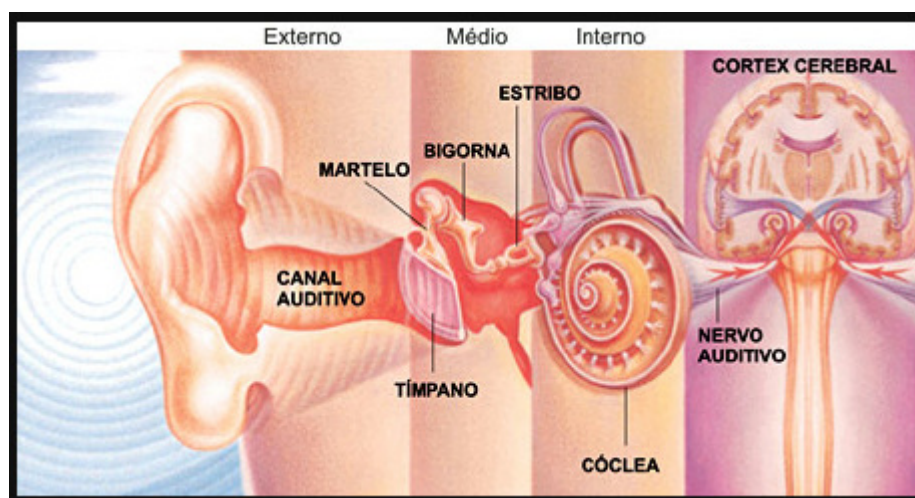


Figura 2: Sistema auditivo dividido em 4 partes

Por consequência dessa vibração, esta será transmitida através dos ossículos: martelo, bigorna e estribo que são responsáveis por conduzir o sinal sonoro e amplificá-lo. Uma vez amplificado, o som chega ao ouvido interno que é a região onde está localizada a cóclea, responsável por abrigar receptores sensoriais, e surgem os impulsos nervosos, o qual levará informação ao cérebro através do nervo auditivo (GUYTON; JOHN E. HALL, 2011).

## 2.2 Princípios da binauralidade

A binauralidade da audição humana permite a descoberta da direção e possível localização de uma fonte sonora. A capacidade de localização da direção de propagação de ondas sonoras é uma das principais características da audição humana. O sistema auditivo humano utiliza diversos indicadores que são aperfeiçoados ao longo da vida para localizar um sinal sonoro em um espaço tridimensional. Para o plano horizontal, utiliza-se as diferenças de sinal (intensidade e tempo) que chegam a ambos os ouvidos. Quando se trata do plano vertical, os indicadores são monoaurais e resultam de modificações do som causadas pelo corpo humano, isto é, fenômenos físicos (como a reflexão) que são causados após o encontro do sinal emitido com o ouvido, cabeça ou tronco do ouvinte (FERNANDES, 2005). Na Figura 3 é mostrado o caminho percorrido por uma onda sonora até os ouvidos.

Pode-se observar que a onda sonora leva mais tempo para percorrer o caminho até o ouvido direito, pois estes estão separados por uma distância "d". A diferença de caminho percorrida pela onda é chamada de  $\Delta l$ .

Sendo  $\alpha$ , o ângulo formado entre uma reta que traça o caminho percorrido pela onda e outra que passa pelo centro da cabeça e segue em direção ao horizonte, é possível calcular a diferença de caminho percorrido  $\Delta l$ . Utilizando-se o conceito de similaridade de triângulos, obtém-se a Equação 2:

$$\Delta l = d \times \text{Sen}(\alpha) \quad (2)$$

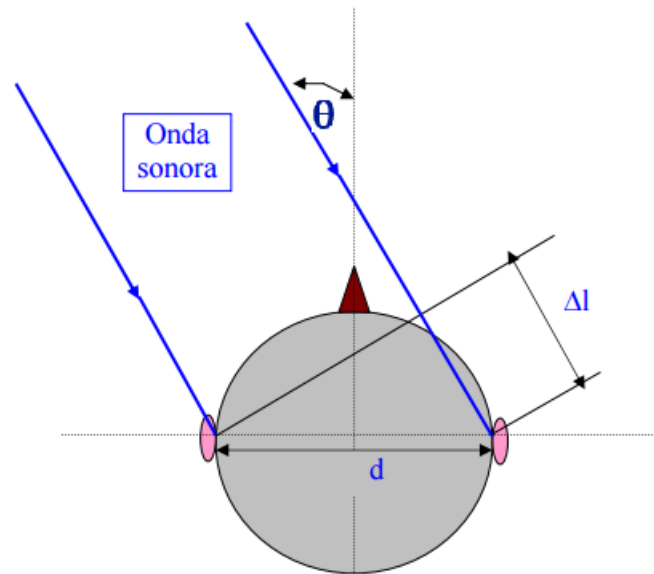


Figura 3: Exemplo de uma onda sonora atingindo a cabeça de uma pessoa..

Considerando a velocidade do som como 344m/s e a distância média entre as orelhas de uma cabeça humana como 0,21m, é possível gerar o gráfico, da Figura 3, em função de  $\alpha$  com o auxílio da ferramenta computacional MATLAB.

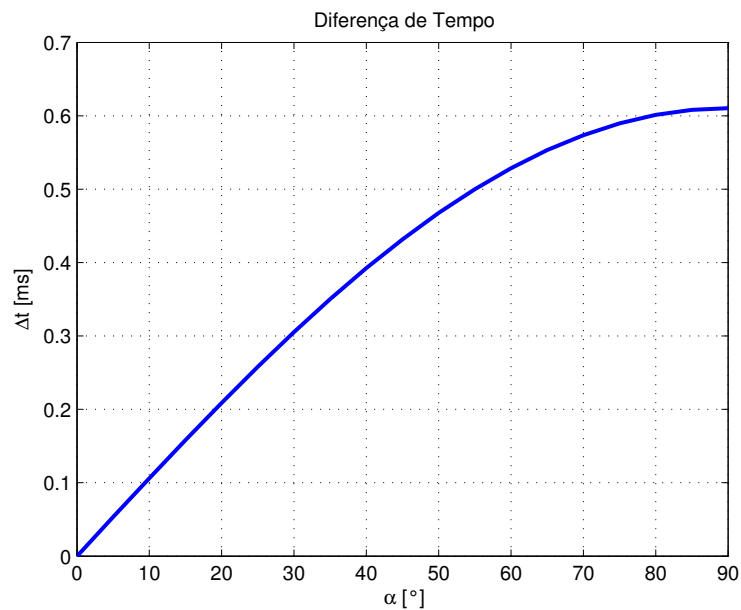


Figura 4: Diferença de tempo.

Têm-se dois casos extremos:

- Para  $\alpha = 0$ ,  $\Delta t = 0$ s;
- Para  $\alpha = 90^\circ$ ,  $\Delta t = 0.6105$ s

Os valores acima correspondem à interpretação geométrica, pois quando a fonte sonora está localizada sobre o eixo das ordenadas da Figura 3, significa que  $\alpha$  é 0 e por

consequente, não haverá diferença de tempo na chegada do som aos ouvidos. Analisando o caso oposto, viu-se que quando a fonte sonora é paralela ao eixo das abscissas e perpendicular a origem (de acordo com a Figura 3), isto é, está alinhada perpendicularmente a um dos ouvidos, o atraso torna-se máximo.

É importante ressaltar que os valores apresentados são aproximados, visto que a velocidade do som depende de vários fatores (temperatura do ar, por exemplo) e que está se utilizando um valor médio para a distância entre as orelhas de uma cabeça humana.

Pequenos movimentos da cabeça, às vezes imperceptíveis pela pessoa, auxiliam fortemente na distinção de fenômenos que poderiam confundir o sistema auditivo. Por exemplo, a incerteza se uma fonte de som central (posicionada sobre o eixo das ordenadas da Figura 3) está a frente ou em uma localização oposta, atrás do ouvinte. Desta forma, indicadores tornam-se fundamentais para a localização de fontes sonoras. Existem dois muito efetivos: ITD (diferença de tempo intra auricular, do inglês *interaural time difference* - *ITD*) e ILD (diferença de intensidade intra auricular, do inglês *interaural level difference* - *ILD*). De acordo com (ZOLZER, 2011), a amplitude e atraso no tempo em cada canal são governadas por curvas similares a Figura 5.

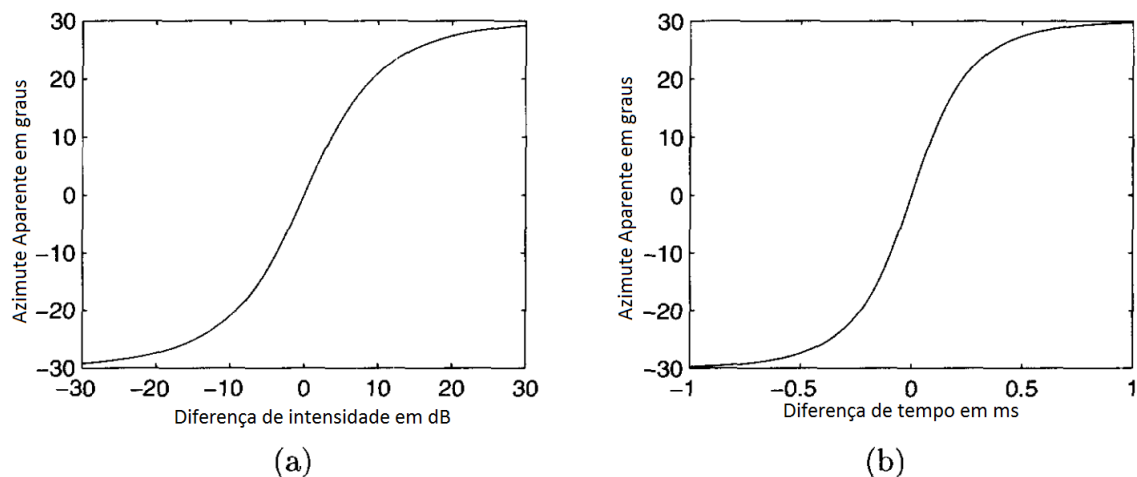


Figura 5: Azimute percebido de uma fonte de som virtual. a) Diferenças de intensidade e b) diferenças de tempo. Fonte: (ZOLZER, 2011)

Estes indicadores, *ITD* e *ILD*, quando provenientes de uma determinada localização, causam uma confusão em nosso sistema auditivo. Para o ouvinte, estes indicadores serão iguais se a fonte sonora estiver em localizações opostas no cone da confusão formado na Figura 6. Contudo, pequenos movimentos da cabeça podem, facilmente, resolver essa situação, fazendo com que o ouvinte consiga determinar a localização aproximada da emissão do som e discernindo, assim, o que seria a zona de confusão.

## 2.3 Indicadores de diferenças interaurais

Conforme explicado anteriormente, existem dois indicadores essenciais para estimular a percepção sonora humana. São eles: *ITD* e *ILD*.

### 2.3.1 Diferença de tempo interaural

A ITD representa a diferença de tempo com que uma determinada frente de onda sonora atinge os ouvidos humanos. Dependendo da localização da fonte sonora, o sinal

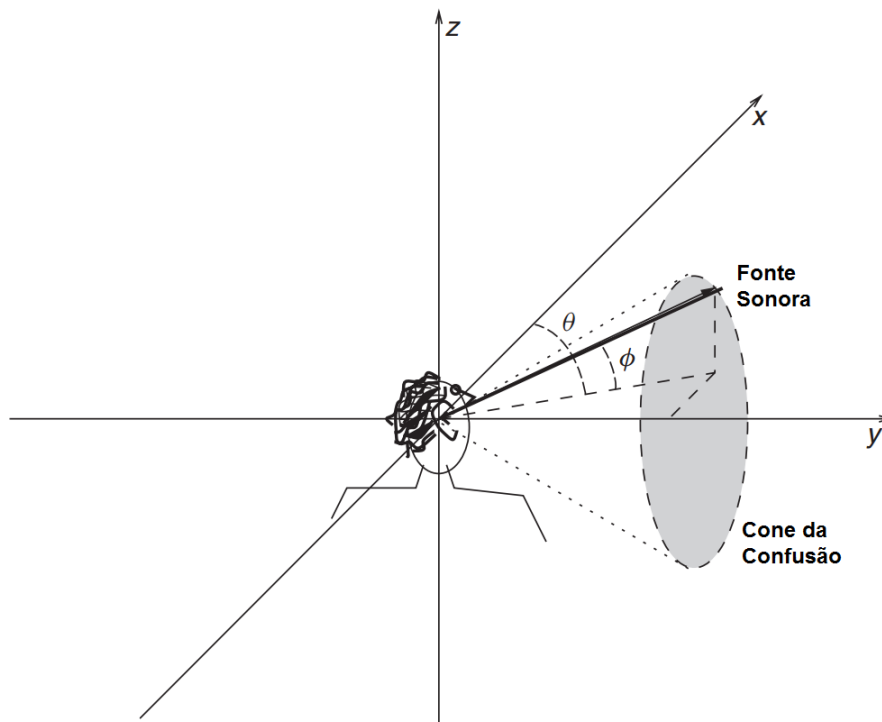


Figura 6: Cone delimitando a zona de confusão em um sistema interaural de coordenadas polares. Fonte: (ZOLZER, 2011)

emitido levará tempos distintos para chegar a cada um dos ouvidos. Este é um indicador deveras importante para estimar a localização da fonte do som. Este é indicador é adequado para se utilizar em frequências inferiores a 1500 Hz (ZOLZER, 2011). Note que se o ângulo  $\alpha$  é 90 (pior caso, o de atraso máximo) é fácil ver que para frequências pouco acima de 1500 Hz, ocorrerá atraso de fase. Comprovando assim, o *ITD* é mais efetivo em frequências pequenas. A Equação 3 mostra a frequência correspondente para um comprimento de onda  $\lambda$  de 0.21m.

$$f = \frac{v}{\lambda} = \frac{344}{0.21} \cong 1640Hz \quad (3)$$

Sabendo que 0.21m é o pior caso de diferença de caminho percorrida pela onda, estabelecido na Figura 3, pode-se constatar que o indicador *ITD* é eficiente, apenas, para frequências pequenas. Pois para frequências mais elevadas, poderá haver atrasos superiores a 360°. Por exemplo, se utilizar-se uma onda sonora com comprimento de onda de 0.15m, e a incidência desta caracterizar o pior caso, a diferença de fase será maior que 360°, visto que a distância a ser percorrida, entre as orelhas, é superior ao comprimento de onda  $\lambda$ .

Cálculos feitos em (ZOLZER, 2011) tornam possível obter uma aproximação, razoável, de curvas *ITD* dependentes da frequência. A Figura 7 apresenta essa relação de dependência com a frequência para um ângulo de 30°.

Note que, na Figura 7, a *ITD* pode ser aproximada por duas constantes, -0.26 e -0.38 para altas e baixas frequências, respectivamente. Este ponto de mudança entre frequências consideradas baixas e altas pode ser obtido através de um ângulo de incidência  $\theta$

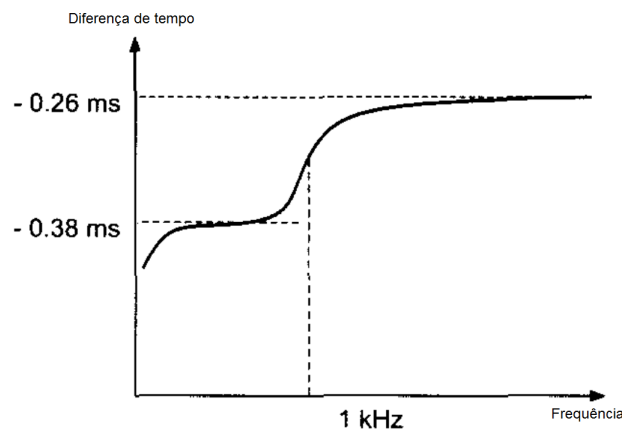


Figura 7: Diferença de tempo interaural dependente da frequência para um ângulo de azimute  $\alpha=30^\circ$ . Fonte: (ZOLZER, 2011)

(ZOLZER, 2011). Desta forma, têm-se a Equação 4.

$$ITD = \frac{1.5d}{c} \times \sin\theta \quad (4)$$

Em que:

- $d$  é a distância entre as orelhas;
- $c$  é a velocidade do som, aproximada por 344m/s

### 2.3.2 Diferença de intensidade interaural

A ILD é um indicador que apresenta a diferença de intensidade do som, oriunda da fonte sonora, que chega aos ouvidos. É apropriada para ondas com frequências acima de 1500 Hz. Este indicador surge, pois a cabeça humana tende a se opôr a passagem dessas ondas, tal fenômeno é chamado de efeito de sombreamento da cabeça (ZOLZER, 2011). A Figura 8 apresenta uma relação de dependência entre a frequência e a diferença de intensidade (ILD).

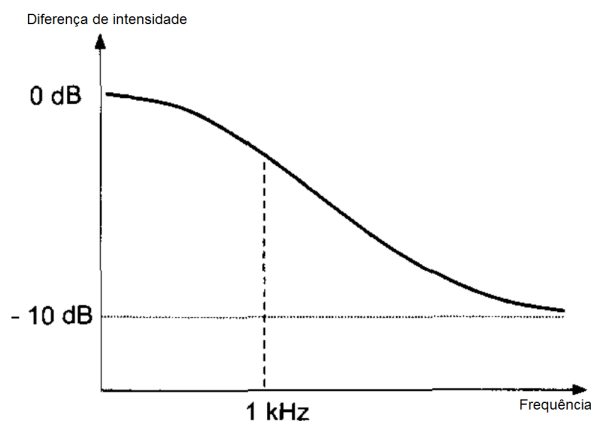


Figura 8: Diferença de intensidade interaural dependente da frequência para um ângulo de azimute  $\alpha=30^\circ$ . Fonte: (ZOLZER, 2011)

## 2.4 Externalização e Correlação

Quando o estímulo total, oriundo de todos modos sensoriais, não pode ser produzido por situações inerentes na natureza envolvendo fontes distantes, parece que os seres humanos tendem a internalizar o som percebido (DURLACH et al., 1992).

Utiliza-se, então, a técnica de decorrelação. Esta técnica é conhecida por ser deveras efetiva para externalizar fontes sonoras ao se utilizar fones de ouvido (KENDALL, 1995). Ela pode ser utilizada para minimizar a correlação cruzada dentro de um conjunto de sinais. Para o caso de interesse, ela decorrelaciona os sinais, provenientes da fonte sonora, que chegam a cada um dos ouvidos.

A técnica de decorrelação pode ser justificada pelo simples fato de que, em situações naturais, o sinal emitido que chega aos ouvidos são, consideravelmente, decorrelacionados (ZOLZER, 2011). Deste modo, para se calcular a correlação entre o sinal que chega a cada um dos canais auditivos, utiliza-se a Equação 5.

$$r(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2Tv} \int_{-T}^T x_L(t)x_R(t+\tau)dt \quad (5)$$

Em que:

- $\tau$  é o atraso de tempo entre os dois canais;
- $x_L$  e  $x_R$  são os sinais que chegaram aos canais da esquerda e direita, respectivamente;
- $v$  é um fator normalizante definido pela Equação 6.

$$v = \frac{1}{2T} \sqrt{\int_{-T}^T x_L^2(t)dt \int_{-T}^T x_R^2(t)dt} \quad (6)$$

Olhando para a Equação 6, pode-se notar que, se os sinais que chegam aos ouvidos forem iguais, os integrandos, também, serão iguais e, desta forma, o resultado de  $v$  será reduzido ao teorema de valor médio quadrático de uma integral definida. O fator de normalização é utilizado para produzir um grau de correlação igual a 1 para quando houver um deslocamento puro e -1 para quando for atrasado em  $180^\circ$ .

Quando um ser humano escuta, via headphone, um ruído de banda larga, o respectivo grau de correlação produz uma imagem espacial que varia do centro da cabeça para as orelhas (ZOLZER, 2011). A Figura 9 apresenta este fenômeno.

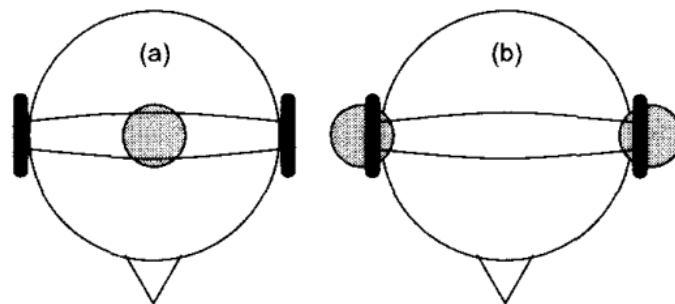


Figura 9: Imagem do som percebida quando escutando um ruído de banda larga com grau de correlação igual a a) 1 e b) 0.



Perceba que se o grau de correlação for 0, isto é, não existir interdependência entre os sinais, a imagem espacial não varia do centro da cabeça para as orelhas. O grau de correlação também é útil para organizar um conjunto de fontes no espaço. Para fontes que possuem graus de correlação semelhantes, o ouvinte assimila apenas imagem espacial do som (ZOLZER, 2011).

Um significativo aumento na externalização pode ser alcançado se o estímulo binaural for variado, dinamicamente, para seguir os movimentos da cabeça do ouvinte. Contudo, esta técnica é muito cara e não muito prática, na maioria dos casos. Portanto, assume-se que o ouvinte deixa sua cabeça parada enquanto escuta o áudio de interesse com seu "headphone".

## 2.5 Análise Experimental utilizando um Manequim

Gardner e Martin, em (GARDNER; MARTIN, 1994) realizaram um conjunto de experimentos em uma manequim do tipo KEMAR (SOUND; VIBRATION, 1994) para a obtenção de uma função de transferência que se aproxime do modelo da cabeça humana. Segue na Figura , um exemplo de um manequim KEMAR.

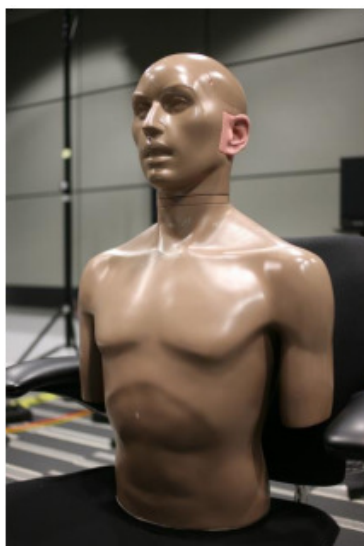


Figura 10: Exemplo de um manequim do tipo KEMAR.

A análise experimental consiste na obtenção das respostas impulsivas da orelha direita e esquerda a partir de um alto falante posicionado cerca de 1.4 metros do manequim. Foi utilizado MLS (sequência de comprimento máximo, do inglês *Maximum Length Sequence*), que é uma sequência binária pseudo aleatória, para obter as respectivas respostas impulsivas em uma taxa de amostragem de 44.1 kHz. Foi utilizado um computador que possuía uma placa de som de 16bits em seus conversores A/D e conversores D/A que operam em uma taxa de 44.1 KHz. A saída do microfone, utilizado para captar os sinais, é pré amplificada e conectada a entrada stereo da placa de áudio. Em (GARDNER; MARTIN, 1994) é descrito como se obter acesso aos dados através um servidor FTP (protocolo de transferência de arquivos, do inglês *File Transfer Protocol*).

As medições foram realizadas em níveis de elevação variando entre  $-40^\circ$  e  $+90^\circ$ , em que o primeiro está abaixo do plano horizontal e o segundo diretamente acima da cabeça. para cada nível de elevação, girou-se a cabeça do manequim em  $360^\circ$  com diferentes

incrementos. A tabela 1 apresenta a quantidade de medições realizadas para cada nível de elevação e o correspondente incremento aplicado para a rotação da cabeça do manequim.

Tabela 1: Número de medições e incremento em cada nível de elevação

<b>Elevação</b>	<b>Número de Medições</b>	<b>Incremento no Azimute</b>
-40	56	6.43
-30	60	6
-20	72	5
-10	72	5
0	72	5
10	72	5
20	72	5
30	60	6
40	56	6.43
50	45	8
60	36	10
70	24	15
80	12	30
90	1	0

As respostas impulsivas foram obtidas através de MLS. Uma descrição detalhada do processo pode ser consultada em (RIFE; VANDERKOOY, 1989). De acordo com (GARDNER; MARTIN, 1994) o comprimento da sequência (MLS) é de  $N=16384$  amostras, o que corresponde a um registrado binário de 14 bits visto que  $2^{14} = 16384$ . Após obtida essa sequência aleatória, replicou-se e concatenou-se a cópia de modo a obter  $2N$  amostras. Simultaneamente foram gravadas  $2N$  amostras na orelha esquerda e direita.

Para a obtenção das respostas impulsivas foram descartadas as primeiras  $N$  amostras de modo que as  $N$  amostras seguintes foram duplicadas para formar novamente uma sequência de comprimento  $2N$ . Para esta nova sequência foi utilizada a técnica de correlação cruzada com as amostras de comprimento original  $N$ , obtidas com MLS, utilizando um algoritmo de FFT e formando um resultado de comprimento  $3N-1$ . É importante ressaltar a diferença entre correlação cruzada e convolução, visto a semelhança entre os dois, a Equação que descreve a operação é descrita na Equação 7.

$$(f * g)[n] = \sum_{m=-M}^M f^*[m]g[n+m] \quad (7)$$

A Figura 11 apresenta um comparativo entre as duas operações. Veja que na correlação cruzada não espelhamento de um dos sinais.

## 2.6 Modelos estruturais

Existem diversos modelos construídos com a finalidade de analisar, entre outros, as consequências da difração em um corpo. O modelo utilizado nesse trabalho se baseia na aproximação da cabeça por uma esfera. Ele utiliza como fundamento o modelo de *Rayleigh* que se baseia na modelagem matemática realizada por Maxwell das equações fundamentais do eletromagnetismo. Ambos serão apresentados a seguir.

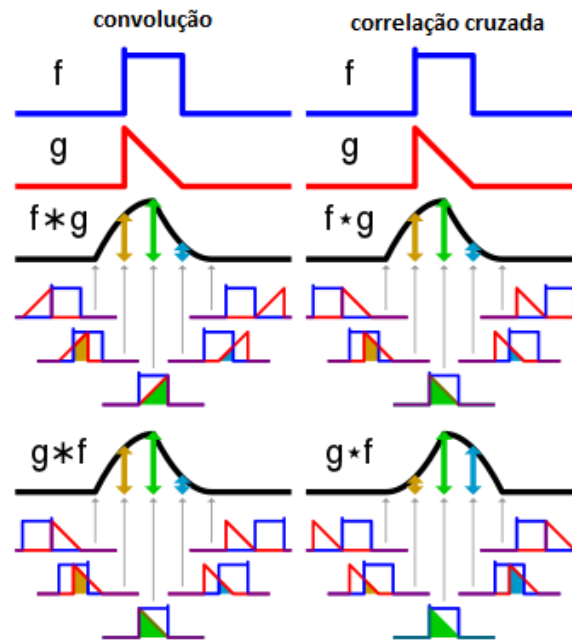


Figura 11: Comparativo entre convolução e correlação.

### 2.6.1 Modelo esférico de Rayleigh

É totalmente possível obter funções de transferência que descrevam o comportamento da cabeça humana, a partir da resolução de equações da onda que possuem como condições de contorno o torso, ombros, cabeça e orelhas. Contudo, *Rayleigh*, obteve e demonstrou em (RAYLEIGH), 1904), uma aproximação simplificada para baixas frequências derivada da solução exata da difração de uma onda plana em uma esfera.

A função de transferência resultante desse modelo é dependente do ângulo de incidência entre o centro da cabeça e a onda incidente, representado por  $\alpha$  na Figura 3 e da frequência da onda emitida.

O diagrama contendo as curvas de respostas em frequências para o modelo proposto por *Rayleigh* é apresentado na Figura 12:

Note que o eixo das abscissas que representa a frequência da onda propagada é normalizada através da Equação 8:

$$\mu = \frac{\omega a}{c} \quad (8)$$

em que,

- $\omega$  representa a frequência angular da onda;
- $c$  é a velocidade do som (aproximadamente 343 [m/s]);
- $a$  é o raio médio de uma cabeça humana. Aproximadamente 8,75 cm de acordo com (KUHN, 1977).

As respostas em frequência obtidas na Figura 12 apontam para o fenômeno de sombreamento da cabeça que é a perda de altas frequências para quando a fonte emissora da onda está distante da cabeça. (BROWN; DUDA, 1998) diz que, para fenômenos binaurais, a resposta de fase é mais importante que a de amplitude e as propriedades de fase para uma esfera são revisadas e aplicadas por (AARONSON; HARTMANN, 2014).

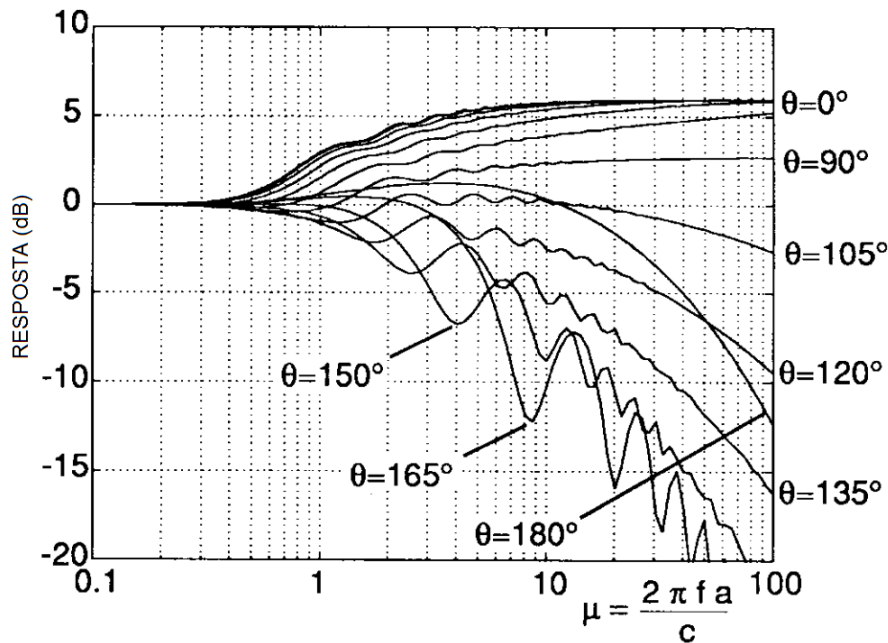


Figura 12: Respostas em frequência de uma esfera rígida ideal para a incidência de uma onda com ângulo de incidência  $\theta$  variando entre  $0^\circ$  e  $180^\circ$ . Fonte: (BROWN; DUDA, 1998)

Para o caso em que  $\mu$  é maior que 1 e conseqüentemente, a frequência da onda for maior que 624 Hz, a diferença  $\Delta T$  entre a localização da fonte emissora e o centro da cabeça (modelado como esfera) é aproximado pela expressão independente de frequência elaborada em (BLAUERT, 1997) e apresentada na Equação 9:

$$\Delta T(\theta) = \begin{cases} -\frac{a}{c}[\theta + \sin(\theta)] & \text{se } 0 \leq |\theta| < \frac{\pi}{2} \\ \frac{a}{c}[\pi - \theta + \sin(\theta)] & \text{se } \frac{\pi}{2} \leq |\theta| \leq \pi \end{cases} \quad (9)$$

Note que a diferença  $\Delta T$  é dependente, exclusivamente, de  $\theta$ , ou seja, apenas do ângulo de incidência da onda emissora. Semelhante ao que foi visto na análise que segue da Figura 3.

O modelo e formula de *Woodworth* para diferenças de tempo interaurais é amplamente adotado como um padrão fisiologia e estudos de psicoacústica envolvendo binauralidade. Este modelo calcula a diferença de tempo interaural para uma cabeça humana, assumindo que a cabeça é uma esfera rígida e que as orelhas são pontos antipodais, isto é, as orelhas estão diametralmente opostas, assim como mostra a Figura 3. Este modelo pode ser considerado válido para casos em que o comprimento de onda do som é muito menor que o raio da cabeça humana. Considerando este raio como 87.5 mm ((KUHN, 1977)), o comprimento de onda torna-se, aproximadamente, 0,26 mm (AARONSON; HARTMANN, 2014).

## 2.6.2 Modelo Cabeça-Esfera Aproximado

Os mais efetivos sistemas de sínteses sonora utilizam catálogos com coeficientes FIR derivados de medições de respostas impulsivas HRIR realizadas em objetos de interesse. Há muito se é desejado um modelo para aproximação de uma HRTF e que substitua esses catálogos de coeficientes. Contudo, há diversas problemas em se desenvolver um modelo

funcional(BROWN; DUDA, 1998):

- É difícil aproximar os efeitos da propagação e difração de uma onda em funções de baixa ordem.
- Não existe critério quantitativo para mensurar o quão próximo o modelo desenvolvido está do ideal. Dessa forma, a avaliação torna-se subjetiva e dependente da audição humana;
- Uma aproximação que funcione bem para um indivíduo, pode não atuar bem para outro. Desta forma, tornaria-se necessário resolver um sistema de identificação repetidas vezes para cada ouvinte.

Segundo (BROWN; DUDA, 1998), é possível modelar uma função de transferência, que descreva o comportamento da cabeça humana, utilizando uma aproximação de primeira ordem através de filtros que fornecem as diferenças de tempo interaurais apresentadas na Equação 9.

O filtro obtido e utilizado por (BROWN; DUDA, 1998) é apresentado na Equação 10:

$$H_{hs}(\omega, \theta) = \frac{1 + j\frac{\alpha\omega}{2\omega_0}}{1 + j\frac{\omega}{2\omega_0}} \quad (10)$$

em que:

- $\omega_0$  representa a frequência relacionada ao raio da cabeça humana. e é dado pela Equação 11
- $\alpha$  é uma função dependente do ângulo de incidência  $\theta$  e que controla a localização do zero no filtro. É apresentado na Equação 12

$$\omega_0 = \frac{c}{a} \quad (11)$$

$$\alpha(\theta) = \left(1 + \frac{\alpha_{min}}{2}\right) + \left(1 - \frac{\alpha_{min}}{2}\right)\cos\left(\frac{\theta}{\theta_{min}}180^\circ\right) \quad (12)$$

Em (BROWN; DUDA, 1998),  $\alpha_{min}$  e  $\theta_{min}$  são utilizados como 0,1 e  $150^\circ$ , respectivamente. Contudo, estes parâmetros podem ser ajustados, individualmente. A dependência do ângulo de incidência para o fator  $\alpha$  pode ser visualizada na Figura 13.

Note que para ângulos superiores a  $78^\circ$ ,  $\alpha$  torna-se inferior a 1, logo haverá um corte para altas frequências. Contudo, se  $\alpha = 2$ , haverá um ganho de 6dB em altas frequências. Isto é ilustrado na Figura 14 que contém um diagrama com as curvas de respostas em frequência para o modelo proposto, utilizando a Equação 10. Assim, como na Figura 12, o eixo das abscissas é normalizado pelo fator  $\mu$ .

De fato, As Equações 10 - 12 fornecem uma boa e simples aproximação para a solução de *Rayleigh* ao obter as respostas da Figura 12 e, assim, uma aproximação para uma HRTF. Elas podem produzir efeitos no azimute convincentes para o ouvinte além de serem de fácil implementação em tempo real (BROWN; DUDA, 1998).

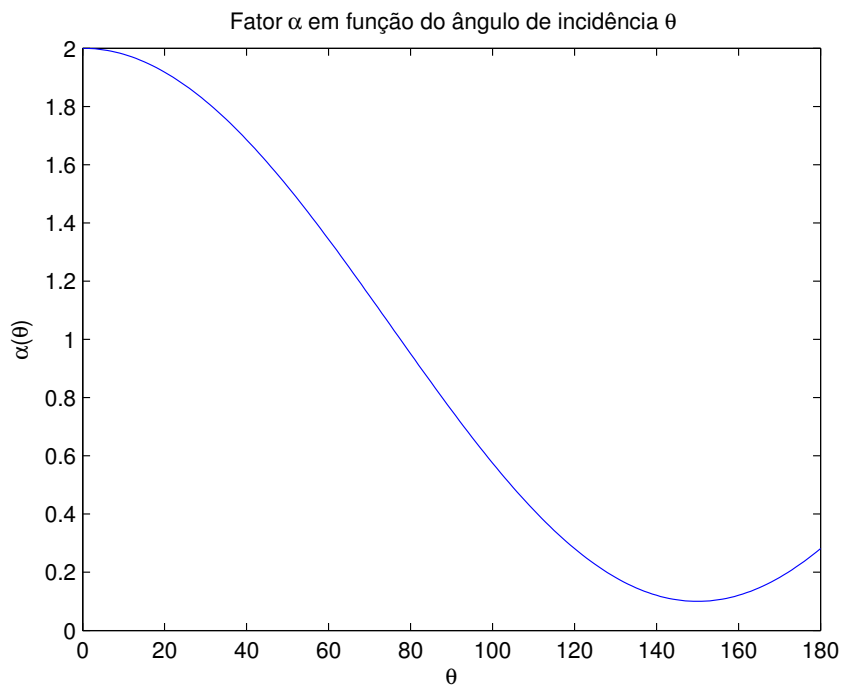


Figura 13: Influência do ângulo de incidência  $\theta$  no fator  $\alpha$

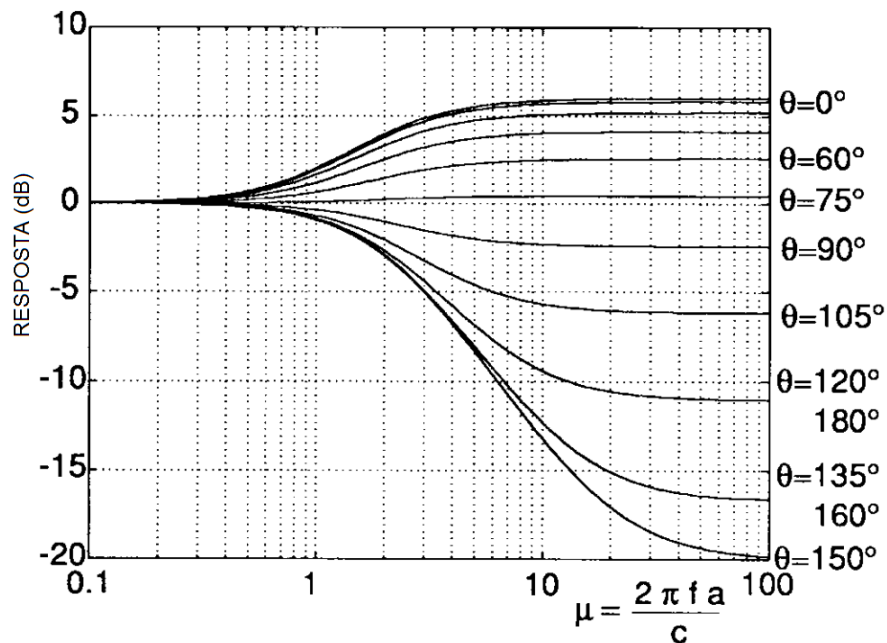


Figura 14: Respostas em frequência para o modelo de uma cabeça esférica, utilizando um filtro de primeira ordem como aproximação, com uma onda com ângulo de incidência  $\theta$  variando entre  $0^\circ$  e  $180^\circ$ . Fonte: (BROWN; DUDA, 1998)

## 2.7 Funções de transferência "Head-Related" - HRTF

Como visto nos subcapítulos anteriores, técnicas binaurais podem ser definidas como métodos cujos quais visam controlar, diretamente, o som no canal auditivo. Isto pode ser

feito por gravações binaurais realizadas em alguma cabeça (seja de uma pessoa ou de um boneco) ou pode ser realizada medindo e modelando funções de transferência.

(GARDNER; MARTIN, 1994) realizaram experimentos em um manequim. Suas medições experimentais foram realizadas em uma câmara anecoica (sem eco) com um alto falante, tocando sinais teste, que é posicionado estrategicamente de diferentes formas para, conseqüentemente, o som emitido, atingir a cabeça do manequim, também, de diferentes formas. O sinal utilizado para a realização destes testes é um ruído pseudo-randômico. O resultado dessas medições é um conjunto de HRIRs que podem ser utilizadas como coeficientes FIR.

Esta coleção de HRIRs e uma convolução direta pode ser uma solução viável para implementar direcionalidade em fontes sonoras utilizando tecnologias atuais (ZOLZER, 2011). Contudo, há uma grande limitação neste processo: o formato das cabeças das pessoas varia largamente. (GARDNER; MARTIN, 1994) obtiveram excelentes resultados utilizando em seus testes um manequim, todavia utilizando HRIRs obtidos individualmente, melhorará drasticamente os resultados obtidos e por conseguinte, a qualidade da localização. Também há o fato de que, inconscientemente, utilizamos pequenos movimentos da cabeça para resolver ambigüidades sonoras oriundas do cone da confusão apresentado na Figura 6.

Dessa forma, por motivos de desempenho, a construção de um modelo do sistema auditivo externo torna-se mais desejável do que uma coleção de HRIRs. Modelando as propriedades estruturais desse sistema auditivo, torna-se possível aplicar contínuas variações para posições de fontes sonoras e para a morfologia do ouvinte (ZOLZER, 2011). A abordagem apresentada em (BROWN; DUDA, 1998) para modelar o sistema auditivo externo contém 3 partes e serão tratadas no próximo sub capítulo.

### 2.7.1 Modelando filtros HRTF

Partindo da abordagem de (BROWN; DUDA, 1998), o modelo utilizado para compôr a HRTF pode ser dividido em três estruturas:

- Fenômeno de sombreamento na cabeça e *ITD*
- Echo pelos ombros
- Reflexões na orelha

Utilizando a aproximação de que a cabeça é uma esfera rígida que difrata uma onda plana, o efeito de sombreamento pode ser aproximado por um sistema de primeira ordem contínuo no tempo. Desta forma, têm-se a Equação 13:

$$s_z = \frac{-2\omega_0}{\alpha(\theta)} \quad (13)$$

e a Equação 14,

$$s_p = -2\omega_0 \quad (14)$$

em que:

- $s_p$  e  $s_z$  representam a posição do polo e do zero, respectivamente, no plano complexo;

- a frequência  $\omega_0$  é definida pela Equação 15 e
- a função  $\alpha$  que depende do azimute  $\theta$  é definida pela Equação 16.

$$\omega_0 = \frac{c}{a} \quad (15)$$

$$\alpha(\theta) = \left(1 + \frac{\alpha_{min}}{2}\right) + \left(1 - \frac{\alpha_{min}}{2}\right) \cos\left(\frac{\theta}{\theta_{min}} 180^\circ\right) \quad (16)$$

O par de polos e zeros podem ser transformados, através de uma transformação bilinear, em um filtro digital IIR estável (MITRA, 2000). O filtro resultante é visto na Equação 21.

Para facilidade e desempenho, em processamento digital de sinais utiliza-se técnicas como a transformada bilinear (também conhecida como método de *Tustin* ou transformada de *Mobius*) para discretizar um sistema contínuo no tempo. Sendo assim, de acordo com (MITRA, 2000), a transformada bilinear é descrita pela Equação 17:

$$s = \frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}} \quad (17)$$

Deste modo, utilizando  $j\omega$  da Equação 10 como  $s$  e aplicando a transformada bilinear da Equação 17, obtém-se:

$$H_{hs} = \frac{1 + \frac{\alpha s}{2\omega_0}}{1 + \frac{s}{2\omega_0}} = \frac{2\omega_0 + \alpha s}{2\omega_0 + s} \quad (18)$$

E substituindo  $s$  pela relação dada em 17, chega-se à:

$$H_{hs} = \frac{2\omega_0 + \alpha \left(\frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}}\right)}{2\omega_0 + \left(\frac{2}{T} \frac{1 - z^{-1}}{1 + z^{-1}}\right)} \quad (19)$$

Sabendo-se que  $T$  é o período da amostragem, utiliza-se a relação de  $1/T \rightarrow F_s$ , em que  $F_s$  é a frequência da amostragem. Deste modo, realizando esta substituição e expandindo os produtos envolvendo  $S$ , obtém-se:

$$H_{hs} = \frac{2\omega_0 + \frac{2F_s \alpha + (-2\alpha F_s z^{-1})}{1 + z^{-1}}}{2\omega_0 + \frac{2F_s + (-2F_s z^{-1})}{1 + z^{-1}}} \quad (20)$$

E, por manipulação algébrica, obtém-se a Equação 21

$$H_{hs} = \frac{(\omega_0 + \alpha F_s) + (\omega_0 - \alpha F_s)^{-1}}{(\omega_0 + F_s) + (\omega_0 - F_s)^{-1}} \quad (21)$$

O indicador *ITD*, pode ser obtido pelo atraso em segundos, como está descrito na Equação 22:

$$\begin{cases} -\frac{a}{c} \cos\left(\theta + \frac{\pi}{2}\right) & \text{se } 0 \leq \left|\theta + \frac{\pi}{2}\right| < \frac{\pi}{2} \\ \frac{a}{c} \left(\left|\theta + \frac{\pi}{2}\right| - \frac{\pi}{2}\right) & \text{se } \frac{\pi}{2} \leq \left|\theta + \frac{\pi}{2}\right| < \pi \end{cases} \quad (22)$$

Apesar de em (BROWN; DUDA, 1998) ser citado que os efeitos do torso e dos ombros são muito menos influentes se comparados ao de sombreamento da cabeça e as reflexões da orelha, pode-se utilizar uma aproximação para que os efeitos do torso e dos ombros



possam ser sintetizados como um simples eco. A partir desse fato e das análises experimentais realizadas por Brown e Duda, obtém-se a Equação 23, que é retratada na Figura 15:

$$\tau_{sh} = 1.2 \frac{180^\circ - \theta}{180^\circ} \left( 1 - 0.00004 \left( (\phi - 80^\circ) \frac{180^\circ}{180^\circ + \theta} \right)^2 \right) [ms] \quad (23)$$

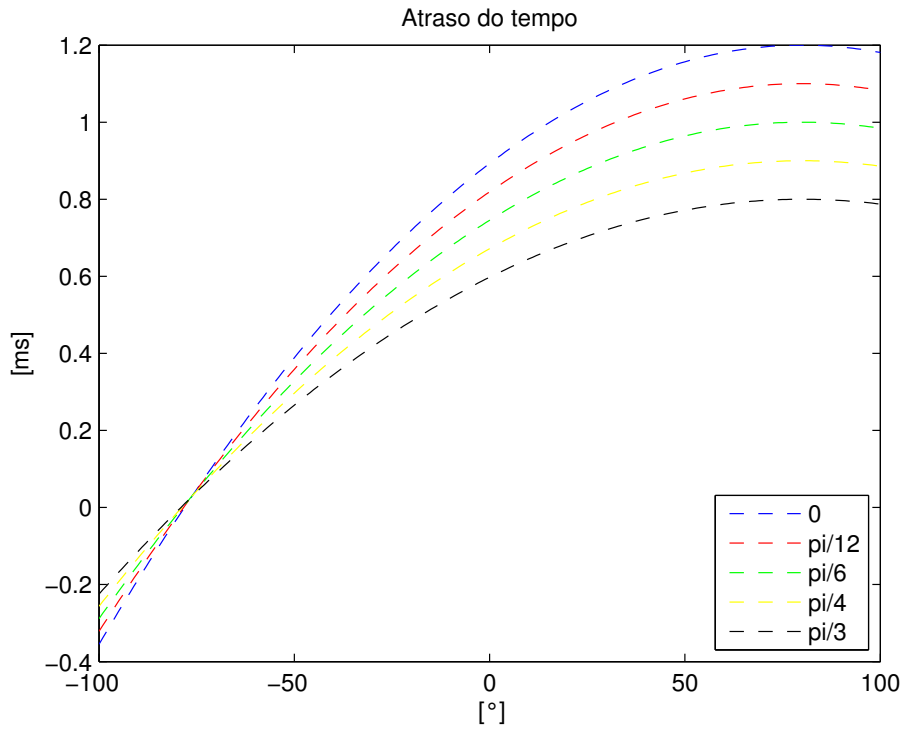


Figura 15: Atraso no tempo oriundos de reflexões no ombro em função do azimuth ( $\theta$ ) e elevação( $\phi$ ).

### 3 METODOLOGIA EXPERIMENTAL

Em toda a metodologia experimental realizada utilizou-se um "headphone" da philips "SHL3565BK/00". Suas especificações técnicas relacionadas a características sonoras são apresentadas na Tabela 2:

Tabela 2: Especificações técnicas do headphone Philips - SHL3565BK/00

Especificações Técnicas	
Sistema acústico	Fechado
Resposta em frequência	7 a 40.000 Hz
Impedância	16 Ohm
Tipo de ímã	Neodímio
Entrada de energia máxima	2200 mW
Sensibilidade	104 dB
Diâmetro da caixa acústica	40 mm
Diafragma	PET

Ao longo dessa metodologia, serão mostrados passo-a-passo como foram desenvolvidos os algoritmos e como esses foram testados, sempre referenciando-os aos respectivos resultados da seção 4. Desse modo, a metodologia aqui abordada seguirá esses passos

Para o desenvolvimento do algoritmo de interesse do trabalho, inicialmente, utilizou-se para fins de estudo e comparação um modelo que não necessita do filtro obtido, utilizando a abordagem encontrada em (BROWN; DUDA, 1998) que modela as propriedades estruturais de uma cabeça humana. Deste modo, utilizou-se um conjunto de coleções de medições, popularmente conhecido, realizadas por Gardner e Martin (GARDNER; MARTIN, 1994). Eles utilizaram um simulador KEMAR para replicar as características estruturais de uma cabeça e um torso (SOUND; VIBRATION, 1994).

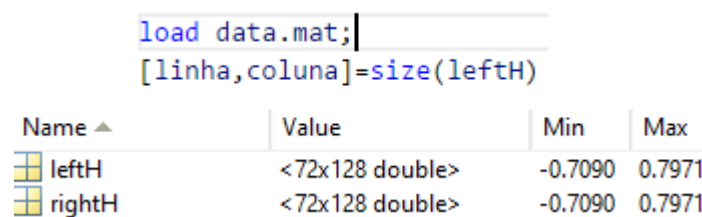
O conjunto de HRIR ( respostas impulsivas "Head-Related", do inglês *Head-Related Impulse Responses* pode ser obtido através de (GARDNER; MARTIN, 1994). Os dados obtidos foram armazenados e compilados em um arquivo chamado "data.mat". É necessário carregá-lo para obter-se acesso as respectivas HRIRs medidas experimentalmente.

### 3.1 Método da convolução utilizando HRIRs obtidas experimentalmente

Para testar este algoritmo utilizou-se dois arquivos de áudio diferentes.:

- 1 - Um arquivo de 14 segundos gerado utilizando a função "rand" do MATLAB para gerar um ruído branco.
- 2 - Um trecho do quarto movimento da primeira sinfonia de Johannes Brahms.

Para carregar-se e armazenar as respostas impulsivas, precisou-se utilizar o "load" no início do script, conforme a Figura 16:



The screenshot shows two lines of MATLAB code in a script editor: `load data.mat;` and `[linha,coluna]=size(leftH)`. Below the code is a variable browser window with the following content:

Name ▲	Value	Min	Max
leftH	<72x128 double>	-0.7090	0.7971
rightH	<72x128 double>	-0.7090	0.7971

Figura 16: Carregamento e armazenamento das HRIRs obtidas experimentalmente.

Note que ao carregar essa coleção de dados, são gerados duas matrizes referentes aos coeficientes das HRIRs, que foram obtidas experimentalmente. "leftH" e "rightH" são abreviações para lado esquerdo e direito da cabeça, respectivamente. A segunda instrução está armazenando a dimensionalidade da matriz "leftH", que será necessária, posteriormente. É indiferente utilizar "leftH" ou "rightH" para esse armazenamento, visto que ambos possuem a mesma dimensionalidade.

Para se realizar a leitura de um arquivo de áudio e sua correta extração de dados, isto é, extrair os dados amostrados e a respectiva taxa de amostragem, utilizou-se o a instrução "audioread". Como mostra a Figura 17:

```
arquivo = uigetfile({'*.mp3'; '*.wav'; '*..*'}, 'Selecione o arquivo de audio');
[pathstr, name, ext] = fileparts(arquivo)
[amostras, Fs] = audioread(arquivo);
```

Figura 17: Leitura e extração dos dados amostrados e taxa de amostragem do arquivo de áudio.

Note que na Figura 17 foram utilizadas 3 funções que fazem parte do conjunto disponibilizado pelas bibliotecas internas do MATLAB.

- "uigetfile": Abre uma "GUI" que solicita que o usuário escolha um arquivo de áudio a ser processado;
- "fileparts": Extrai informações do arquivo que foi recebido. Dividindo-as em caminho do diretório do arquivo ("pathstr"), nome do arquivo ("name") e a extensão do arquivo ("ext"). Estas serão úteis posteriormente.
- "audioread": Realiza a devida extração de dados amostrados e taxa de amostragem do arquivo que foi selecionado pelo usuário.

Para que seja possível aplicar o método da convolução a fim de convoluir amostras de dados com impulsos para se obter a sensação de que o som está se movendo circularmente, em sentido horário, ao redor da cabeça, utilizou-se todos os coeficientes gerados a partir dos dados carregados, isto é, "leftH" e "rightH" a partir de "data.mat", conforme segue na Figura 18:

```

y = [amostras'; amostras'];

N = floor(length(y)/linha);
L = zeros(1,Fs*tempoTotal+127);
R = L;

for i=0:linha-1
    L(N*i+1:(i+1)*N+127) = L(N*i+1:(i+1)*N+127) + conv(leftH(i+1,:),y(1,i*N+1:(i+1)*N));
    R(N*i+1:(i+1)*N+127) = R(N*i+1:(i+1)*N+127) + conv(rightH(i+1,:),y(2,i*N+1:(i+1)*N));
end

```

Figura 18: Aplicação do método da convolução.

Veja que para poder-se efetuar, corretamente a convolução, faz-se necessário alguns "preparativos" para este. A variável "y" está armazenando uma matriz que é uma concatenação da matriz de dados "amostra" transposta. N, L e R são preenchidos de tal forma para serem compatíveis dimensionalmente durante todo o processo de convolução. A convolução, que é o método de maior interesse, é efetuado dentro de um laço iterativo, usando a instrução "conv". Note que a quantidade de iterações que ocorrerão será determinado pelo valor da variável "linha", que representa a quantidade de linhas dos coeficientes obtidos em "leftH", deste modo, verifica-se que utilizará-se todos os coeficientes.

Por fim, para obter-se a saída deste método, com a finalidade de analisá-lo, normaliza-se as respostas referentes ao lado esquerdo e direito, transpondo-as em seguida, conforme a Figura 19:

```

L = L/max(abs(L));
R = R/max(abs(R));
output = [L',R'];
audiowrite( strcat(name,'_saida.wav'),output, Fs);

```

Figura 19: Saída do processamento.

Veja que para armazenar e escutar o áudio que foi processado, utiliza-se a função "audiowrite", que está atribuindo como nome do arquivo uma concatenação entre o nome do arquivo original e a sequência de caracteres "\_saida.wav". O script completo pode ser visualizado na Figura 38, no anexo 7.1.

Para testar e verificar o resultado do algoritmo desenvolvido, elaborou-se um script para teste. Na Figura 20 é apresentado como a função é chamada e a atribuição das variáveis de interesse.

Note que, inicialmente o código recebe os dois valores que a função "circular\_conv" exporta. "output" representa a saída, após o processamento, e "tempoTotal" representa a duração do arquivo de áudio que foi amostrado. As respectivas saídas, referentes aos canal esquerdo e direito são atribuídas as variáveis L e R. Para a variável x é atribuído um array de valores de acordo com a amostragem que foi realizada e o tempo de duração do

```

[output, tempoTotal]=circular_conv();

L=(output(:,1))';
R=(output(:,2))';

x = [0:tempoTotal/(length(L)-1):tempoTotal];

```

Figura 20: Atribuição das variáveis de interesse para o aplicação do script teste

áudio processado. Para poder visualizar os resultados, obtidos utilizou o trecho de código apresentado na Figura 21.

```

subplot(3,1,1);
plot(x,L);
axis tight;
title('Canal Esquerdo');
xlabel('Tempo (s)');
ylabel('Amplitude');

subplot(3,1,2);
plot(x,R, 'r');
axis tight;
title('Canal Direito');
xlabel('Tempo (s)');
ylabel('Amplitude');

subplot(3,1,3);
plot(x,L);
hold on;
plot(x,R, 'r');
axis tight;
title('Canal Esquerdo e Direito');
xlabel('Tempo (s)');
ylabel('Amplitude');

```

Figura 21: Método utilizado para apresentar, graficamente, o resultado aplicado pelo algoritmo que utiliza o catálogo de HRIRs.

Foram geradas três imagens a partir do código acima: uma referente ao auditivo esquerdo, outro para o direito e as duas sobrepostas. As Figuras obtida por este código são apresentadas no sub capítulo 4.1.

## 3.2 Sistema Proposto: Algoritmo para um modelo cabeça-esfera

Conforme visto nos capítulo anteriores, existem motivos para que se prefira aderir à composição de um modelo que estruture o sistema auditivo externo do que ao catálogo de HRIRs. Desse modo, no decorrer desse capítulo serão explicados os métodos e materiais utilizados para desenvolver um algoritmo, utilizando a abordagem de (BROWN; DUDA, 1998).

### 3.2.1 Validação do modelo obtido por Brown e Duda

Inicialmente, para poder-se validar esse modelo, resolveu-se tentar repetir os resultados encontrados. Deste modo, utilizando as Equações 10 - 12, optou-se por obter-se os respectivos diagramas de bode para cada ângulo de incidência  $\theta$ . Na Figura 22 está o código utilizado para gerar os respectivos diagramas de bode.

```
function Hhs(theta)
c=343;
a=0.0875;
w0=c/a;
alfamin=0.1;
thetamin=180;

alfa=(1+(alfamin/2))+(1-(alfamin/2))*cos(theta*pi/thetamin);
H=tf([alfa/(2*w0) 1],[1/(2*w0) 1]);

bode(H);
end
```

Figura 22: Script utilizado para gerar um diagrama de bode para qualquer ângulo de incidência

A função espera um único parâmetro, o ângulo de incidência. "c" e "a" representam, respectivamente, a velocidade do som e o raio da cabeça humana. "alfamin" e "thetamin" são parâmetros que podem ser ajustados individualmente, estes controlam junto com a função "alfa" a localização do zero na função de transferência atribuída a variável "H", através do comando "tf". Este último cria o modelo de uma função de transferência, que espera o numerador e denominador. A função "bode" gera o diagrama de bode (magnitude e fase) da respectiva função de transferência. Esse algoritmo foi utilizado duas vezes para verificar a influência do fator  $\theta_{min}$  no cálculo de  $\alpha$ , na primeira utilizou-se "thetamin" como 180 e na outra, como 150. O resultado dessa validação pode ser verificada na seção de resultados 4.2.1.

### 3.2.2 Filtro para fenômeno de Sombreamento da Cabeça

Baseando-se no modelo das Equações 10 - 12, implementou-se, um algoritmo que aplicasse a transformada bilinear a essas, obtendo assim, um modelo discretizado. Utilizou-se a Equação 21 para a implementação desse filtro. As Figuras 23 e 24 apresentam o algoritmo aplicado.

Note que a função "filtro \_ sombra \_ cabeça" espera três parâmetros em sua assinatura: ângulo de incidência, frequência de amostragem do sinal de entrada e a orelha pra qual se deseja aplicar o filtro. Em (ZOLZER, 2011), é apresentada a ideia de se aplicar o filtro para para duas vezes. Uma para cada orelha, desse modo, a Equação 21 origina as Equações 24 e 25.

$$H_{hsE} = H_{Hs} |_{\theta=\theta+90} \quad (24)$$

$$H_{hsD} = H_{Hs} |_{\theta=\theta-90} \quad (25)$$

```

function [B,A] = filtro_sombra_cabeca(theta, Fs, orelha)
% p/ orelha = 1, o filtro será aplicado sobre a orelha direita
% p/ orelha = -1, o filtro será aplicado sobre a orelha esquerda

theta_orelha = 90 * orelha;
theta=theta - theta_orelha;

```

Figura 23: Ajuste da variável de interesse  $\theta$  de acordo com a orelha para qual se deseja aplicar o filtro.

em que  $H_{hsE}$  e  $H_{hsD}$  representam o filtro aplicado a orelha esquerda e direita, respectivamente.

A localização do zero do filtro foi atribuída conforme a Equação 16 e atribuída a variável "alfa". Os dois coeficientes do numerador e de denominador da função de transferência discretizada são atribuídos a "B" e "A", respectivamente, conforme a Equação 21.

```

% parâmetros para ajuste fino
alfamin=0.1;
thetamin=pi;

% reponsável por ganho e definir localização do zero.
alfa=(1+(alfamin/2))+(1-(alfamin/2))*cos(theta/thetamin*pi);

%numerador da função de transferência
B=[(alfa+w0/Fs)/(1+w0/Fs),(-alfa+w0/Fs)/(1+w0/Fs)];
%denominador da função de transferência
A=[1, -(1-w0/Fs)/(1+w0/Fs)];

```

Figura 24: Implementação da posição do zero em função do azimute  $\theta$  e atribuição do numerador e denominador da função de transferência

Para a validação do filtro aplicado criou-se outro script para gerar as curvas resultantes do filtro aplicado, conforme a Figura 25:

```

figure(1);
hold on;

for theta=0:30:180
    [B,A]=head_shadow(theta,44100,-1);
    [h,w] = freqz(B,A);
    loglog(w/pi,10*log(h),'color',rand(1,3),'LineWidth',2);
end

```

Figura 25: Script teste para a implementação do fenômeno de sombreado da cabeça.

Antes da obtenção das curvas utilizou-se a função "bilinear" da ferramenta MATLAB para comparar se os coeficientes do filtro IIR estavam iguais aos que foram calculados

através da transformação bilinear nas Equações 17 - 21. Para a obtenção das curvas de resposta em frequência utilizou-se o comando "freqz" que calcula a resposta em frequência de um filtro digital, a partir dos coeficientes do polinômio numerador e denominador que são passados por argumento. Utilizou-se um laço iterativo para variar-se o ângulo theta que foi passado para a função que gera o filtro e retorna os coeficientes desse. Os coeficientes do polinômio são transmitidos através das variáveis B e A e a função "freqz" exporta as variáveis "w" e "h", responsáveis pelo vetor de frequência angular e a magnitude da resposta, respectivamente. A função "loglog" é a responsável por gerar o diagrama em escala logarítmica. Este script foi aplicado três vezes, variando-se apenas o parâmetro responsável por decidir para qual orelha será aplicado o filtro:

- 0 para desconsiderar as orelhas e esperar obter respostas semelhantes as curvas do diagrama de bode que foi obtido utilizando o algoritmo aplicado na seção 3.2;
- -1 e 1 para se obter o diagrama de respostas em frequência analisando a orelha esquerda e diretamente, respectivamente

Os diagramas obtidos estão apresentados na seção de resultados 4.2

### 3.2.3 implementação de delay para o indicador ITD

Para aplicar-se o algoritmo utilizou-se o modelo de Woodworth, revisado e corrigido por (AARONSON; HARTMANN, 2014).

O algoritmo para a aplicação desse modelo está ilustrado na Figura 42

```
function [output] = woodworth(theta)
    theta=theta*pi/180;
    c=343;
    a=0.0875;
    if (abs(theta) >= 0 && abs(theta) < pi/2)
        output=(a/c)*(theta+sin(theta));
    end
    if (abs(theta) >= pi/2 && abs(theta) <= pi)
        output=(a/c)*(pi - theta + sin(theta));
    end
end
```

Figura 26: Implementação da diferença de tempo interaural em função de  $\theta$

Note que o algoritmo implementado é função apenas do ângulo de incidência  $\theta$ . Desse modo, espera-se que para um azimute de  $0^\circ$  e  $180^\circ$  a diferença de tempo seja 0 e que seja máxima para uma incidência  $\theta = 90^\circ$ . O resultado obtido para esse delay é apresentado na seção de resultados e o algoritmo completo junto com o que foi utilizado para teste pode ser visualizado no Anexo 7.5.

### 3.2.4 Algoritmo para implementação do modelo completo.

Para o desenvolvimento do modelo final utilizou-se o algoritmo da Figura 27. Note que o algoritmo depende de três argumentos essenciais: ângulo de incidência, sinal de entrada e frequência de amostragem.



```

function [HRTF] = sistema_completo(theta,sinal,Fs)
delayITDe=woodworth(theta-90);
delayITDd=woodworth(theta+90);
[Bd,Ad] = filtro_sombra_cabeca(theta,Fs,1);
[Be,Ae] = filtro_sombra_cabeca(theta,Fs,-1);
outputE = filter(Be,Ae,sinal)
outputD = filter(Bd,Ad,sinal)
HRTF=[outputE*delayITDe outputD*delayITDd]
end

```

Figura 27: Algoritmo completo do modelo.

O algoritmo é composto de funções descritas nas seções anteriores e que, também, são apresentados nos anexos. Ele funciona a partir de um sinal de entrada amostrado. Para realizar a leitura do sinal pode-se utilizar a função "audioread". Inicialmente é calculado o atraso do sinal através da função "woodworth". Note que ela é calculada para "theta-90" e "theta+90". Visto que é necessário ser calculada para ambos canais auditivos. Em seguida, o filtro de sombreado da cabeça é aplicado para as duas orelhas. Utilizou-se a função "filter" para filtrar o sinal de entrada obtido, através dos coeficientes da função de transferência obtida para ambas orelhas. O matlab recomenda para sinais relativamente grandes, o uso da função "fftfilt" (que utiliza o algoritmo de overlap-add (PROAKIS; MANOLAKIS, 1996) (OPPENHEIM; SCHAFER, 1989)), contudo, devido a não perceber diferenças entre o sinal gerado e não perceber diferença significativa no tempo de processamento, continuou-se a usar a função "filter". Por fim, aplicou-se o delay encontrado através da função "woodworth" para as saídas filtradas. O modelo em diagrama de blocos é apresentado na Figura 27.

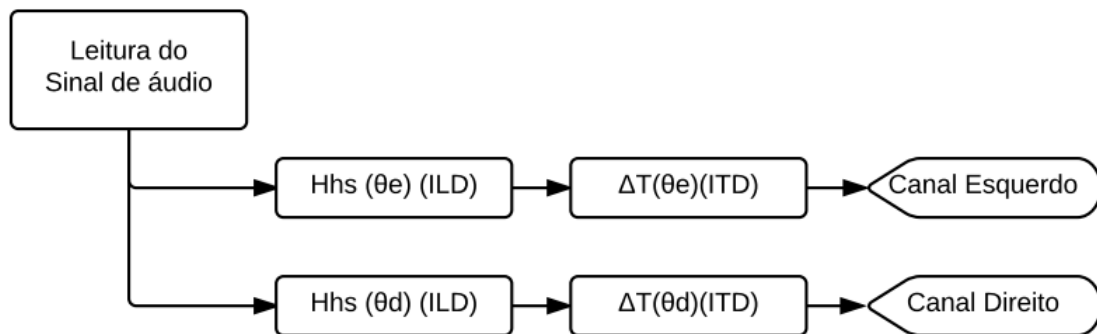


Figura 28: Modelo completo desenvolvido.

## 4 RESULTADOS E DISCUSSÕES

Nesta seção serão apresentados os resultados obtidos a partir da metodologia experimental utilizada. Inicialmente, será contemplada a metodologia que aborda a aplicação do método de convolução para as HRIRs obtidas, experimentalmente. Posteriormente, mostrará-se os resultados utilizando o modelo proposto.

### 4.1 Utilizando o catálogo de respostas impulsivas HRIRs

Ao executar o script que aplica o método da convolução utilizando HRIRs previamente carregadas (Anexo 7.1) em um ruído branco gerado pela ferramenta MATLAB, obteve-se espectro de amplitude da Figura 29 para o canal stereo:

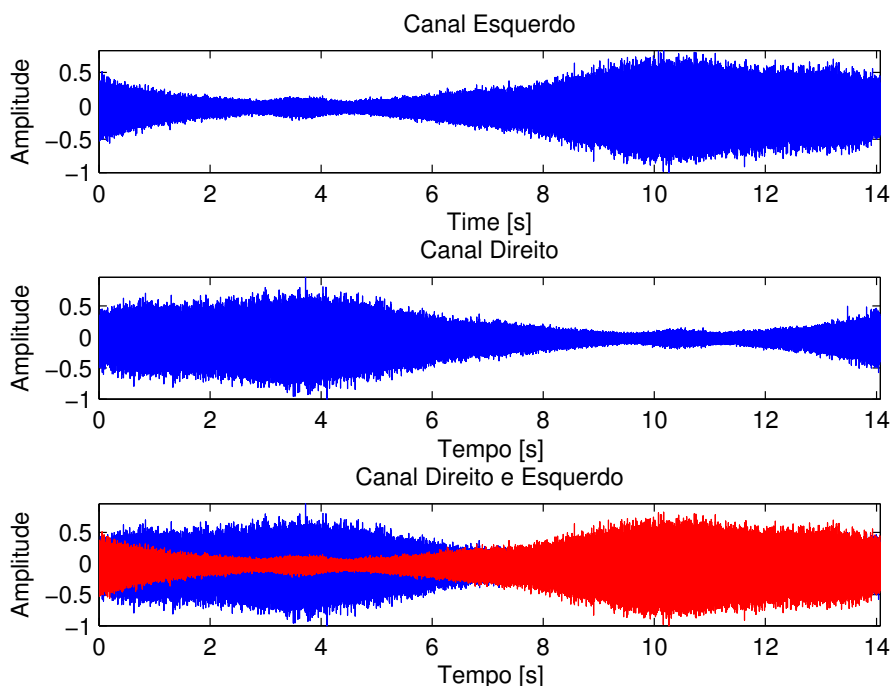


Figura 29: Amplitudes do canal stereo geradas através do MATLAB para um ruído branco.

Note que, de fato, como foi gerado um ruído branco pelo MATLAB, o diagrama de amplitude do canal esquerdo é muito semelhante ao do canal direito, se este for espelhado. Ao observar a Figura 29, verifica-se que apesar de o MATLAB gerar um ruído branco,

o som parece iniciar no canal stereo direito e realizar a transição para o lado esquerdo ao longo dos 14 segundos da amostra de áudio gerada. É fácil afirmar isso, visto que o diagrama de amplitude do canal esquerdo começa os primeiros segundos atenuado ao contrário do direito, sendo que essa situação suavemente se inverte até tornar-se o oposto.

Para o mesmo script, utilizou-se um trecho do quarto movimento da primeira sinfonia de Johannes Brahms. Para este, obteve-se os diagramas de amplitudes da Figura 30 e 31 :

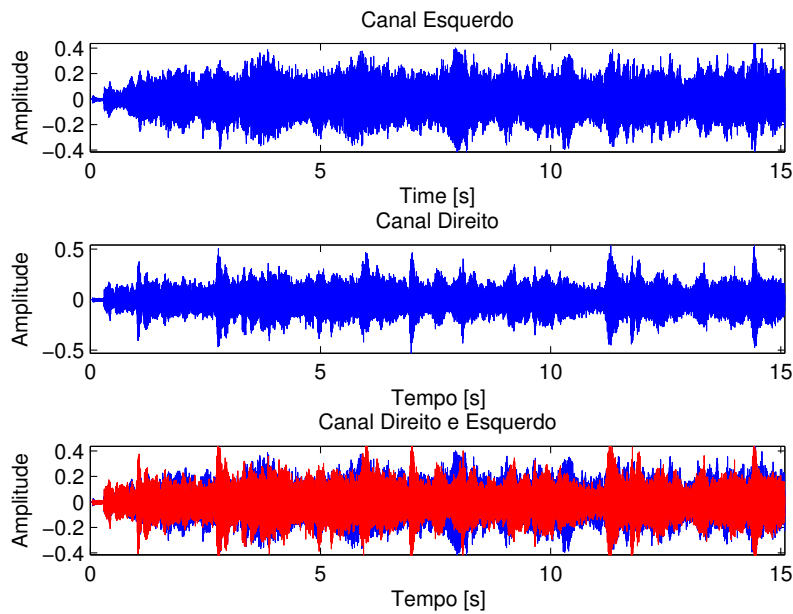


Figura 30: Diagrama de amplitudes para o canal stereo da música clássica utilizada.

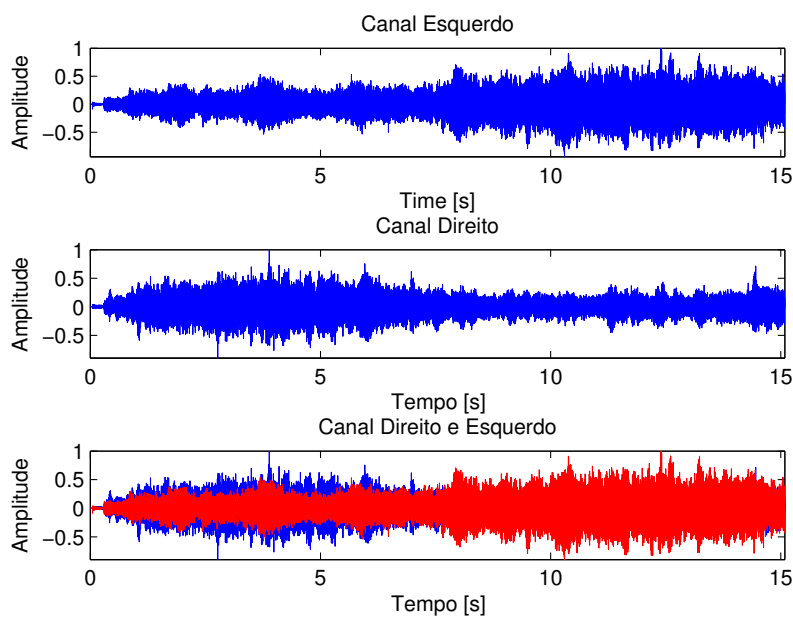


Figura 31: Diagrama de amplitudes após processadas pelo algoritmo.

O diagrama de amplitudes do arquivo de áudio original, isto é, sem processamento é apresentado na Figura 30. Note que não existe uma simetria evidente entre o canal esquerdo e direito, como ocorreu para o ruído branco na Figura 29. Isto se deve ao fato das particularidades do arquivo de áudio escolhido, visto que o ruído branco não possui auto correlação e é temporalmente homogêneo, por óbvio, a simetria será maior. Contudo, é fácil obter o mesmo consentimento que a análise anterior. O canal esquerdo começa atenuado, enquanto que ocorre o oposto ao do lado direito. Essa relação suavemente se inverte até o final do arquivo de áudio.

## 4.2 Utilizando o sistema proposto

Nessa seção serão apresentados e discutidos os resultados referentes ao modelo proposto por Brown e Duda que é o foco do trabalho.

### 4.2.1 Apresentação da validação do modelo de Brown e Duda

Utilizando a metodologia apresentada na seção 3.2.1 e o algoritmo da Figura 22, obteve-se o diagrama de bode (magnitude) apresentado na Figura 32

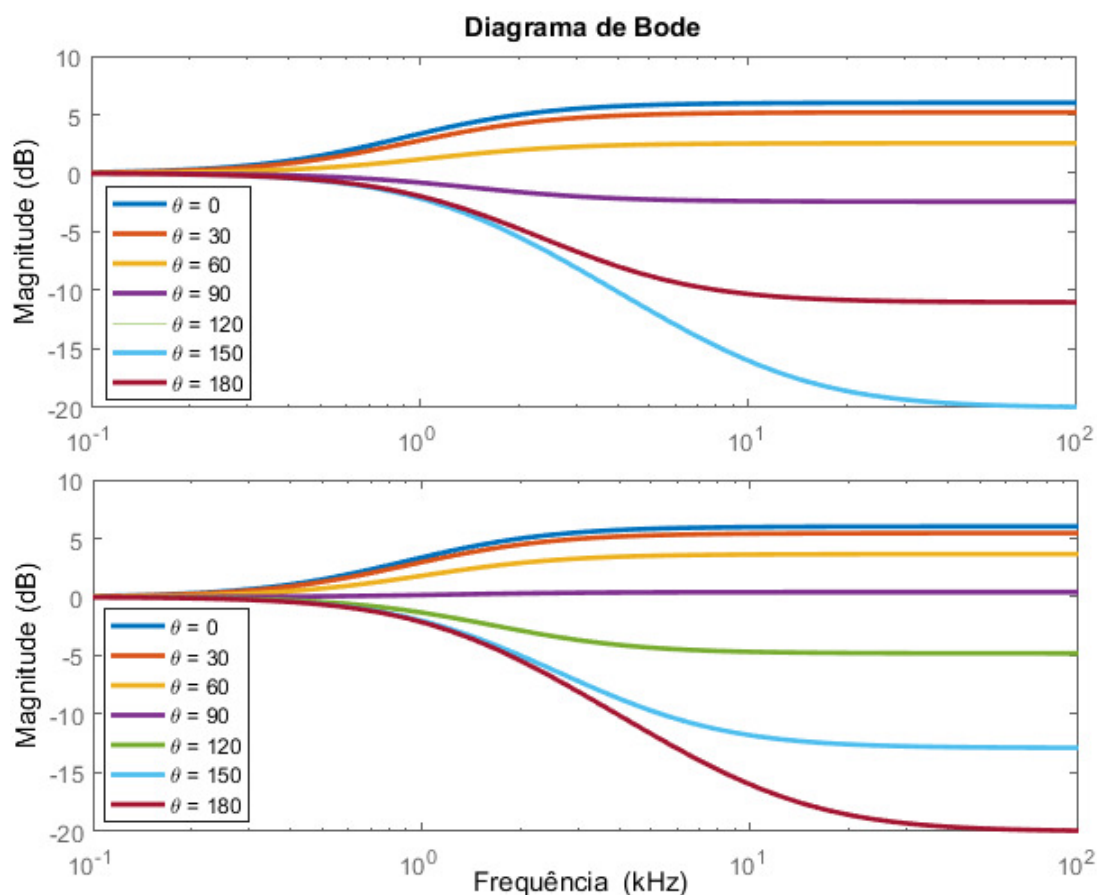


Figura 32: Diagrama de bode obtido para o modelo proposto por Brown e Duda. Para o primeiro gráfico utilizou-se como  $\theta_{min} = 150$  e para o segundo, 180.

Note que conforme o modelo ilustrado por Brown e Duda e suas respectivas curvas de resposta em frequência na Figura 14, a magnitude da resposta, em ambos gráficos varia mais significativamente quando exposto a maiores frequências, comprovando, assim, o

fato de que o indicador de diferenças de intensidade *ILD* é melhor para altas frequências. Note que no primeiro gráfico não está ilustrada a curva referente ao ângulo  $\theta = 120^\circ$ , isso ocorre pois a curva referente ao ângulo de  $\theta = 180^\circ$  está sobreposta a ela. Veja que ao alterar o fator  $\theta_{min}$  ocorre um impacto na simetria e continuidade.

#### 4.2.2 Apresentação do filtro que aplica o fenômeno de sombreamento da cabeça

Utilizando a função "bilinear" do MATLAB obteve-se os mesmos resultados que se esperava encontrar conforme a modelagem realizada nas Equações 17 - 21. Desse modo, considerou-se que a transformação realizada está correta. Aplicando, então, os métodos da seção 3.2.2, obteve-se resultados para o filtro aplicado. A Figura 33 apresenta as curvas da resposta em frequência para  $\theta$  variando entre  $0^\circ$  e  $180^\circ$ .

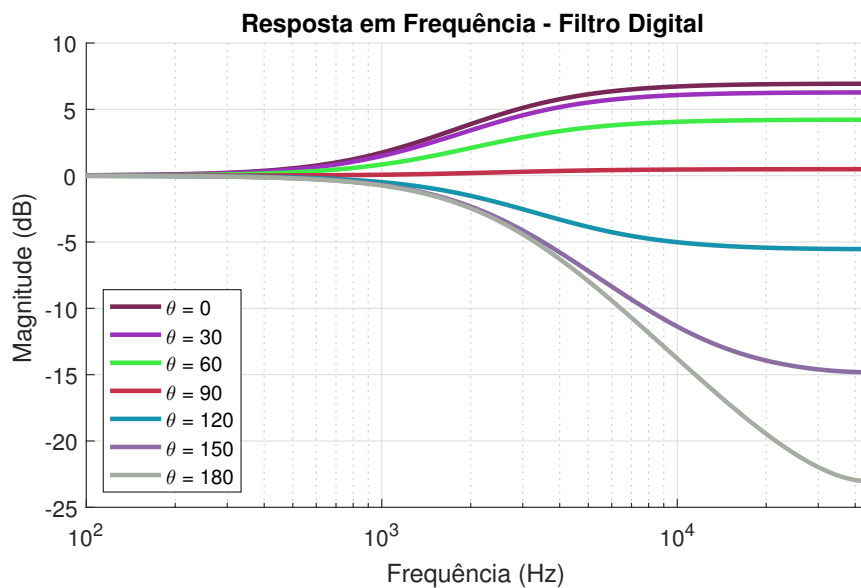


Figura 33: Implementação do filtro que aplica o efeito de sombreamento da cabeça utilizando um filtro IIR.

Perceba que o resultado é semelhante ao encontrado pelo modelo de Brown e Duda, de forma não discreta. Visto que o comportamento das curvas é igual a do modelo, considera-se que a implementação da discretização realizada pela transformada bilinear, obteve sucesso.

As Figuras 34 e 35 apresentam o resultado das respostas em frequência quando o algoritmo é aplicado para as orelhas esquerda e direita. Note que quando é aplicado para a orelha direita, a magnitude da resposta em frequência é sempre superior a 0 dB, visto que o ângulo de incidência  $\theta$  variou entre  $0^\circ$  e  $180^\circ$ , faz sentido obter essas respostas. Note que o ângulo de máxima intensidade se dá quando  $\theta = 90^\circ$ , comprovando, assim, que a implementação do modelo, de fato, funcionou. É importante ressaltar que as curvas referentes ao azimute de  $0^\circ$  e  $180^\circ$  estão sobrepostas, o que é explicável, visto que se o filtro é aplicado para orelha direita, que é perpendicular ao azimute de  $90^\circ$ , quando o ângulo de incidência for  $0^\circ$  e  $180^\circ$ , a magnitude da resposta em frequência deve ser semelhante. O mesmo ocorre para os ângulos de  $30^\circ$  e  $150^\circ$ ;  $60^\circ$  e  $120^\circ$ . Percebe-se uma simetria que espera-se encontrar quando o filtro for aplicado para a outra orelha (esquerda).

A Figura 35 possui simetria idêntica a da Figura 34, conforme o esperado. Note que o para  $\theta = 90^\circ$ , ocorre a menor intensidade. Isto explica-se pelo fato de que para esse

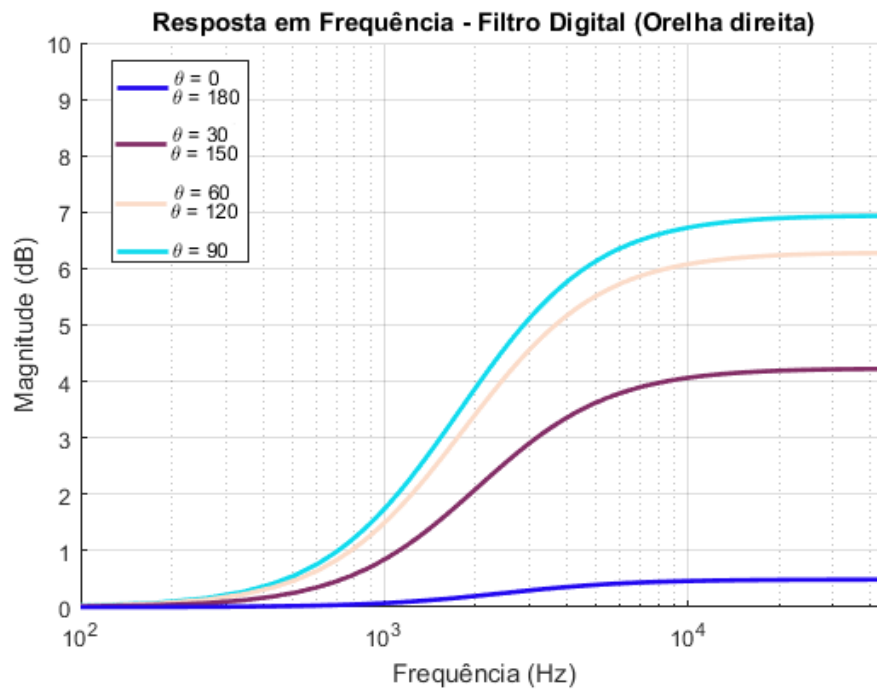


Figura 34: Filtro IIR para a orelha direita.

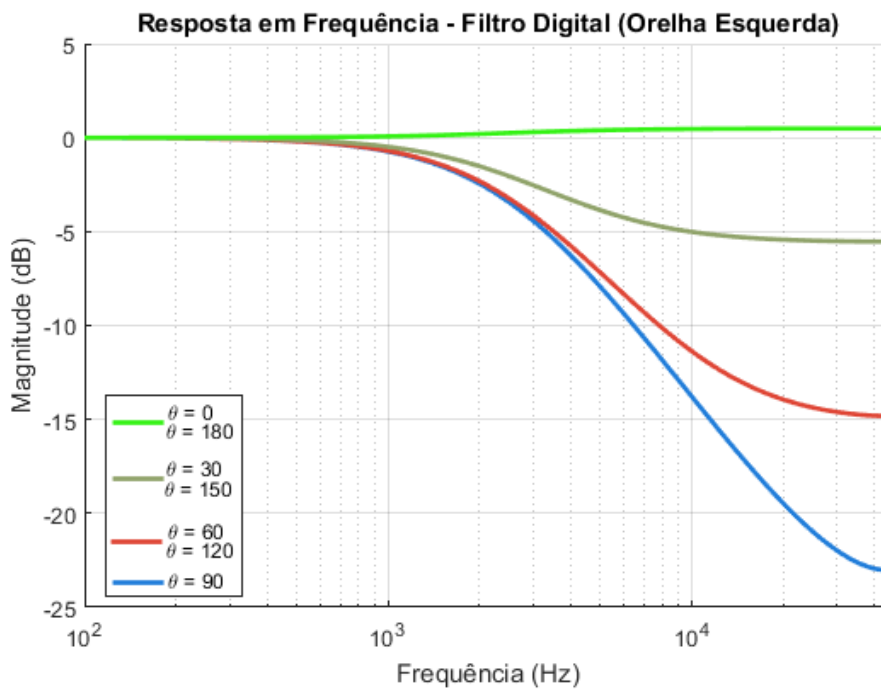


Figura 35: Filtro IIR para a orelha esquerda

azimute a incidência é oposta ao ouvido esquerdo. Desse modo, o som precisa viajar mais, desse modo, sendo atenuado pela distância percorrida e sombreamento da cabeça. Note que, semelhante ao ouvido direito, a sobreposição também ocorre entre os ângulos:  $60^\circ$  e  $120^\circ$ ;  $30^\circ$  e  $150^\circ$ ;  $0^\circ$  e  $180^\circ$ .

Após encontrar resultados, explicáveis e esperado, constata-se que o modelo foi implementado de forma correta, funciona e possui repetibilidade.

### 4.2.3 Apresentação do algoritmo que implementa atraso

Baseado na metodologia da seção 3.2.3 e do algoritmo apresentado no anexo 7.5, gerou-se a respectiva resposta para o modelo de Woodworth.

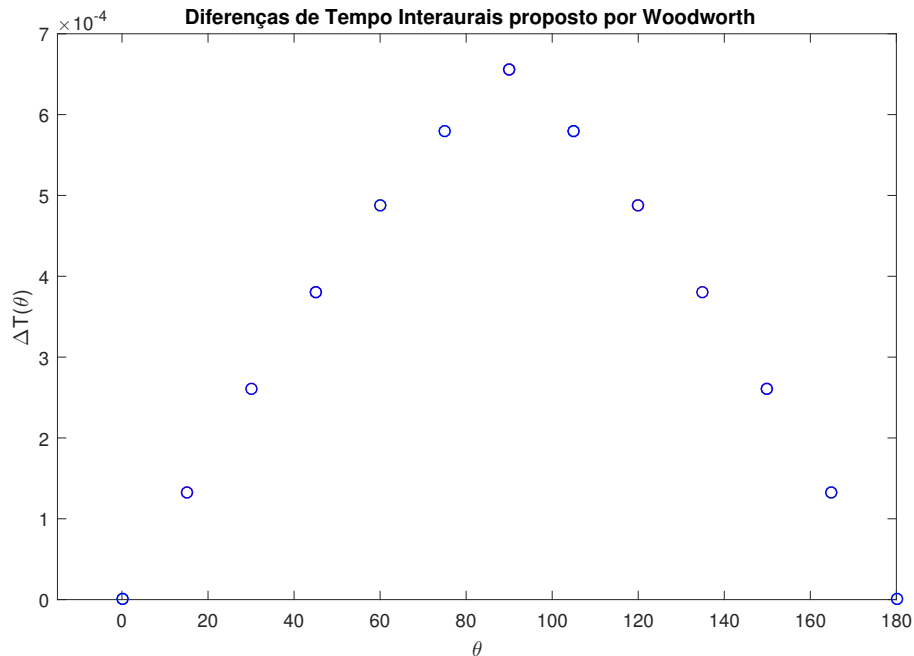


Figura 36: Implementação de atraso (ITD) ao sinal emitido em função do ângulo de incidência  $\theta$

Note que para  $\theta = 0^\circ$  e  $\theta = 180^\circ$  a diferença de tempo obtida é zero, o que é explicável. Isso se deve ao fato de que para esses ângulos o sinal está sendo emitido da região frontal, perpendicular ao nariz e da região traseira, respectivamente. Em ambas situações o sinal chegará ao mesmo tempo nas duas orelhas. No entanto, quando o sinal emitido está sendo emitido de uma localização com um ângulo  $\theta = 90^\circ$ , a diferença de tempo interaural é máxima. O sinal que chegará ao ouvido contrário sofrerá com o atraso e isso causará ao ouvinte a percepção de localização da fonte do som, caracterizando a função do indicador ITD.

### 4.2.4 Apresentação da implementação do sistema completo

Na Figura 37, apresenta-se o diagrama de amplitudes obtido através de um ruído branco injetado no sistema. Utilizou-se a abordagem citada na seção de metodologia 3.2.4.

Note que assim como ocorre para quando utiliza-se o catálogo de respostas impulsivas, está claro a transição que o sinal realiza entre os canais auditivos. Pode-se perceber no diagrama que o sinal começa fraco no ouvido esquerdo e gradualmente vai se intensificando até o término do áudio. O oposto ocorre com o canal direito. Ao sobrepôr os dois diagramas de amplitude, obtém-se a mesma simetria esperada que ocorreu utilizando o catálogo de HRIRs.

Brown e Duda dizem que é difícil mensurar e comparar esses resultados. Pois não existe critério quantitativo para tal (BROWN; DUDA, 1998). Como mensurar, matematicamente, qual o melhor método? Devido a essa dificuldade, optou-se por avaliar-se,

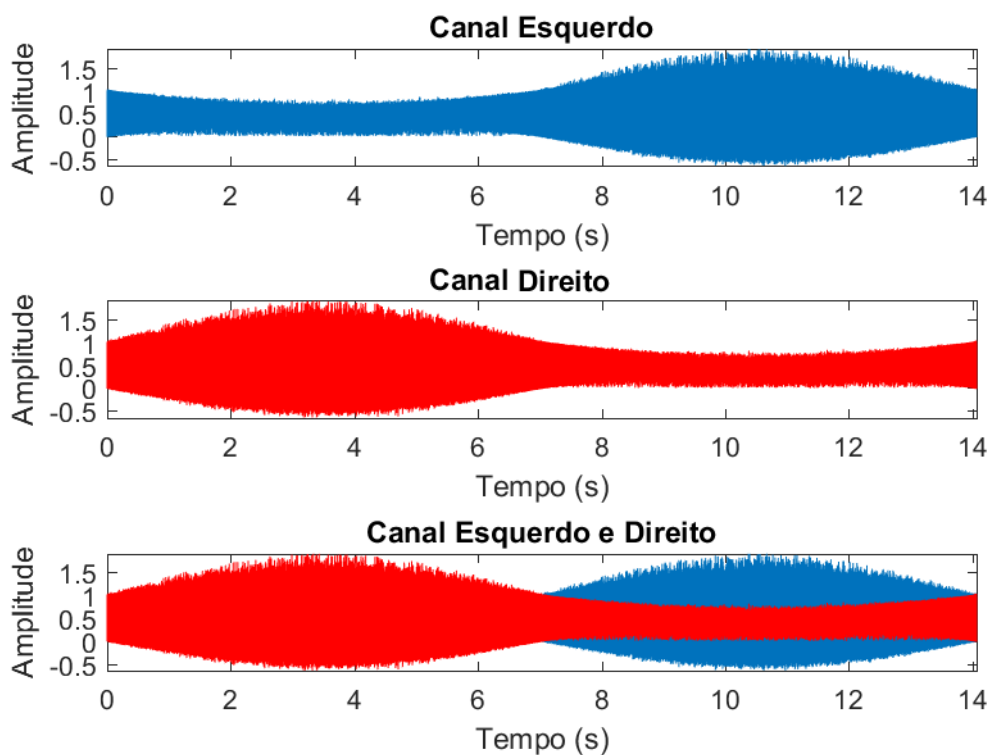


Figura 37: Diagramas de amplitude ao longo do tempo para um ruído branco quando submetido ao sistema completo

também, de forma subjetiva, os métodos. Três pessoas escutaram o sinal processado por ambos os métodos e chegaram a conclusão semelhantes. Em ambos os casos perceberam com boa aproximação a localização da fonte sonora, contudo pequenas alterações do ângulo azimute, passaram despercebidas. Essas pessoas também informaram que quando submetidos ao método que envolve o catálogo de HRIRs, o som processado parecia tornar-se mais claro, fazendo com que a percepção sonora se desse de modo mais trivial.



## 5 DISCUSSÃO E CONCLUSÕES

Apesar da dificuldade em aproximar os efeitos de propagação e difração de ondas, em uma esfera, por funções de baixa ordem, pelos resultados obtidos, é possível perceber que o modelo desenvolvido obteve resultado satisfatório. Pois esse coincide com o proposto, teoricamente, por Rayleigh e conseqüentemente com Brown e Duda, foco desse trabalho. As Equações 9, 16 e 21 conseguem fornecer uma excelente aproximação do modelo tratado teoricamente. Isso abre caminhos para possíveis implementações em tempo real para futuros trabalhos, pois a realização de convoluções diretas utilizando o catálogo de respostas impulsivas é mais custosa computacionalmente. Com o modelo desenvolvido torna-se simples implementar filtros IIR de forma barata sem a necessidade de precisar de hardwares de alto desempenho. Também pode-se ressaltar o fato de que as pessoas possuem diferenças em suas cabeças. Respostas impulsivas obtidas experimentalmente, através de um manequim não podem atender a qualquer indivíduo, o modelo proposto nesse trabalho permite uma fácil configuração. Através de ajustes é possível alterar a localização de um zero do filtro e assim alterar a forma como esse interage com um sinal sonoro de entrada.

O modelo desenvolvido nesse trabalho apresenta características que permitem, através da percepção sonora humana, localizar fontes sonoras em um espaço bidimensional, através de indicadores binaurais. Em (BROWN; DUDA, 1998) é descrito que: *"não existe critério quantitativo para medir a capacidade de uma aproximação para capturar informação direcional que seja perceptualmente relevante"*. Desse modo, torna-se difícil avaliar e comparar resultados. Três pessoas testaram o sistema desenvolvido de forma subjetiva, a opinião delas foi semelhante. As três pessoas conseguiram localizar a direção de onde estava vindo o som, entretanto, essas mesmas informaram que quando utilizado o catálogo de respostas impulsivas obtidas por Gardner e Martin, a percepção e clareza foi maior. Elas conseguiram localizar a posição aproximada da fonte sonora, todavia pequenas variações de ângulos pareceu imperceptível a elas. Isto deve-se ao fato das características individuais que existe entre o sistema composto por cabeça, orelhas, e tronco dos indivíduos.

Por fim, é importante notar que este trabalho pode ser utilizado como para implementações futuras. Visto que a metodologia está detalhada passo a passo. Torna-se possível utilizar o modelo e os resultados apresentados aqui, dado sua repetibilidade, para realizar adaptações e futuras implementações: processamento em tempo real, por exemplo. A análise realizada nesse trabalho, não utilizou-se do sentido de elevação, o que adicionaria uma nova dimensão ao sistema. Visto a complexidade em atuar com reflexões na orelha, essa poderia ser uma continuidade a esse trabalho

## 6 TRABALHOS FUTUROS

É possível listar muitas continuações para este trabalho, visto que essa tecnologia é ascendente e a conferência de efeitos digitais de áudio estimula a publicações de trabalhos na área.

- Realização de mais ensaios com pessoas para obter, assim, uma maior análise dos resultados, mesmo que de forma subjetiva.
- Adicionar os efeitos de reflexões oriundas das orelhas ao modelo para, assim, obter-se o sentido de elevação ao sistema
- Implementar o modelo desenvolvido na *toolbox* Simulink do MATLAB para que seja possível a atuação em tempo real no sistema
- Analisar a complexidade do código para determinar o sistema mínimo necessário para processar os algoritmos desenvolvidos
- Comparação para o modelo desenvolvido, utilizando headphone e fones intra auriculares
- Adicionar um acelerômetro ao fone de ouvido para intensificar a imersão causada em tempo real, de acordo com os movimentos da cabeça da pessoa.

## 7 ANEXOS

### 7.1 Script que aplica método da convolução utilizando HRIRs previamente carregados e obtidas experimentalmente

```

clear all;
close all;
format long;

load data.mat;

[linha,coluna]=size(leftH)

arquivo = uigetfile({'*.mp3'; '*.wav'; '*.*'}, 'Selecione o arquivo de audio');
[pathstr,name,ext] = fileparts(arquivo)
[amostras, Fs] = audioread(arquivo);

tempoTotal = length(amostras)/Fs;

y = [amostras'; amostras'];

N = floor(length(y)/linha);
L = zeros(1,Fs*tempoTotal+127);
R = L;

for i=0:linha-1
    L(N*i+1:(i+1)*N+127) = L(N*i+1:(i+1)*N+127) + conv(leftH(i+1,:),y(1,i*N+1:(i+1)*N));
    R(N*i+1:(i+1)*N+127) = R(N*i+1:(i+1)*N+127) + conv(rightH(i+1,:),y(2,i*N+1:(i+1)*N));
end

L = L/max(abs(L));
R = R/max(abs(R));
output = [L',R'];
audiowrite( strcat(name, '_saida.wav'),output, Fs);

```

Figura 38: Script completo que aplica método da convolução direta com catálogo de HRIRs.

## 7.2 Script teste para o algoritmo ilustrado na seção 7.1

```
[output, tempoTotal]=circular_conv();

L=(output(:,1))';
R=(output(:,2))';

x = [0:tempoTotal/(length(L)-1):tempoTotal];

subplot(3,1,1);
plot(x,L);
axis tight;
title('Canal Esquerdo');
xlabel('Tempo (s)');
ylabel('Amplitude');

subplot(3,1,2);
plot(x,R,'r');
axis tight;
title('Canal Direito');
xlabel('Tempo (s)');
ylabel('Amplitude');

subplot(3,1,3);
plot(x,L);
hold on;
plot(x,R,'r');
axis tight;
title('Canal Esquerdo e Direito');
xlabel('Tempo (s)');
ylabel('Amplitude');
```

Figura 39: Script teste para o algoritmo que aplica o método da convolução com o catálogo de HRIRs.

### 7.3 Script que implementa algoritmo para o fenômeno de sombreamento da cabeça

```

function [B,A] = filtro_sombra_cabeca(theta, Fs, orelha)
% p/ orelha = 1, o filtro será aplicado sobre a orelha direita
% p/ orelha = -1, o filtro será aplicado sobre a orelha esquerda

theta_orelha = 90 * orelha;
theta=theta - theta_orelha;

% velocidade em [m/s]
c=343;
% raio da cabeça humana em
a=0.0875;

w0=c/a;

% parâmetros para ajuste fino
alfamin=0.1;
thetamin=pi;

% responsável por ganho e definir localização do zero.
alfa=(1+(alfamin/2))+(1-(alfamin/2))*cos(theta/thetamin*pi);

%numerador da função de transferência
B=[(alfa+w0/Fs)/(1+w0/Fs),(-alfa+w0/Fs)/(1+w0/Fs)];
%denominador da função de transferência
A=[1, -(1-w0/Fs)/(1+w0/Fs)];

end

```

Figura 40: Algoritmo para aplicar fenômeno de sombreamento da cabeça.

## 7.4 Script teste para o algoritmo ilustrado na seção 7.3

```

figure(1);
hold on;

for theta=0:30:180
    [B,A]=head_shadow(theta,44100, 0);
    [h,w] = freqz(B,A);
    loglog(w/pi,10*log(h),'color',rand(1,3),'LineWidth',2);
end

legend('\theta = 0', '\theta = 30', '\theta = 60', '\theta = 90'...
, '\theta = 120', '\theta = 150', '\theta = 180');

grid on
xlabel('Frequência normalizada (x \pi rad/amostra)')
axis([0 1 -25 10])
ylabel('Magnitude (dB)')
title('Resposta em Frequência - Filtro Digital')

```

Figura 41: Script teste para o algoritmo que aplica o filtro de sombreamento da cabeça

## 7.5 Script para implementar atraso ITD

```
function [output] = woodworth(theta)
    theta=theta*pi/180;
    c=343;
    a=0.0875;
    if (abs(theta) >= 0 && abs(theta) < pi/2)
        output=(a/c)*(theta+sin(theta));
    end
    if (abs(theta) >= pi/2 && abs(theta) <= pi)
        output=(a/c)*(pi - theta + sin(theta));
    end
end
```

Figura 42: Script utilizado para adicionar atraso (ITD)

```
figure(1);

for theta=0:15:180
    plot(theta,woodworth_novo(theta),'b--o');
    hold on;
end
xlabel('\theta')
ylabel('\DeltaT(\theta)')
title('Diferenças de Tempo Interaurais proposto por Woodworth')
```

Figura 43: Script utilizado para apresentar atraso (ITD) gerado pelo algoritmo da Figura 42

## REFERÊNCIAS

- AARONSON, N. L.; HARTMANN, W. M. Testing, correcting, and extending the Woodworth model for interaural time difference. **The Journal of the Acoustical Society of America**, [S.l.], v.2, n.135, p.817–823, Feb. 2014.
- BLAUERT, J. P. (Ed.). **Spatial Hearing**. [S.l.]: Cambridge, MA - MIT Press, 1997.
- BROWN, C. P.; DUDA, R. O. A Structural Model for Binaural Sound Synthesis. **IEEE Transactions on Speech and Audio Processing**, [S.l.], v.6, n.5, p.13, Sept. 1998.
- CARLIE, S. The physical basis and psychophysical basis of sound localization. **Virtual Auditory Space: Generation and Applications**, [S.l.], 1996.
- CIPIC. **The CIPIC Interface Laboratory Home Page**. 2012.
- DUDA, R. O. ModModel head related transfer functions. **Ann. Asilomar Conf. Signals, Systems and Computers**, [S.l.], 1993.
- DURLACH, N. I. et al. On the Externalization of Auditory Images. **Presence: Teleoper. Virtual Environ.**, Cambridge, MA, USA, v.1, n.2, p.251–257, May 1992.
- FERNANDES, J. C. **Acústica e Ruídos**. 2005.
- GARDNER, B.; MARTIN, K. HRTF Measurements of a KEMAR Dummy-Head Microphone. **MIT Media Lab Perceptual Computing - Technical Report 280**, [S.l.], May 1994.
- ELSEVIER (Ed.). **Tratado de fisiologia Medica**. [S.l.]: Elsevier, 2011.
- KENDALL, G. S. The Decorrelation of audio signals and its impacts on spatial imagery. **Computer Music J.**, [S.l.], p.71–87, 1995.
- KUHN, G. F. Model for the interaural time differences in the azimuthal plane. **The Journal of the Acoustical Society of America**, [S.l.], v.62, p.157–167, July 1977.
- MITRA, S. K. K. **Digital Signal Processing: a computer-based approach**. 2<sup>nd</sup>.ed. [S.l.]: McGraw-Hill Higher Education, 2000.
- MYERS, P. **Three-dimensional auditory display apparatus and method utilizing enhanced bionic emulation of human binaural sound localization**. 1989. n.4 8 1 7 149.



OPPENHEIM; SCHAFER (Ed.). **Discrete-Time Signal Processing**. [S.l.]: Prentice-Hall, 1989. pp. 558 - 560.

CLIFFS, E. (Ed.). **Digital Signal Processing**. [S.l.]: Prentice-Hall, 1996. v.3. pp. 430 - 433.

RAYLEIGH), J. W. S. L. On the accoustic shadow of a sphere. **Phil. Trans. R. Soc. London**, [S.l.], v.2, p.87–97, 1904.

RIFE, D. D.; VANDERKOOY, J. Transfer Function Measurements using Maximum Length Sequences. **Journal of the Audio Engineering Society**, [S.l.], p.419–444, June 1989.

SOUND, G.; VIBRATION. **Head & Torso simulators**. 1994.

IEEE (Ed.). **Dicínário de Termos**. [S.l.]: The Institute of Electrical and Electronics Engineering, 2000.

ZOLZER, U. **DAFX Digital Audio Effects**. [S.l.]: John Wiley & Sons Ltd, 2011.