

TÉCNICAS DE GERAÇÃO DE SOM TRIDIMENSIONAL

Fábio Pacheco Freeland

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Aprovada por:

Prof. Paulo Sergio Ramirez Diniz, Ph.D.

Prof. Luiz Wagner Pereira Biscainho, D.Sc.

Prof. Marcio Nogueira de Souza, D.Sc.

Prof. Alexandre Santos de la Vega, D.Sc.

RIO DE JANEIRO, RJ - BRASIL

JULHO DE 2001

FREELAND, FÁBIO PACHECO

Técnicas de Geração de Som Tridimensional [Rio de Janeiro] 2001

IX, 90 pp., 29,7 cm (COPPE/UFRJ, M.Sc., Engenharia Elétrica, 2001)

Tese - Universidade Federal do Rio de Janeiro, COPPE

1.Som Tridimensional 2.Aproximação de Funções de Transferência 3.Filtros IIR

I.COPPE/UFRJ II.Título (série)

Agradecimentos

Gostaria de agradecer, primeiramente, a meus pais, Victor e Inês, meus irmãos, Renata, Miriam e Eduardo, minha noiva Juliana e meus amigos que, de uma forma ou de outra, formaram uma base bastante forte que possibilitou o desenvolvimento deste trabalho.

Novamente à Juliana pela paciência, compreensão e companherismo durante todo o decorrer desse trabalho.

Agradeço, também, a meus orientadores, Paulo Sergio Ramirez Diniz e Luiz Wagner Pereira Biscainho, pelo suporte e idéias que fizeram com que esse trabalho fosse viável e trouxesse alguma contribuição para o grupo de Processamento de Sinais, bem como para as pessoas que trabalham na área de áudio. E, ainda, pela tranqüilidade essencial que me passaram.

Ao amigo e colega Cristiano N. dos Santos, pela ajuda e idéias no decorrer dessa tese, ao amigo Paulo Antônio A. Esquef, que, mesmo estando longe, ajudou-me muito com a obtenção de material utilizado nesse trabalho, aos amigos e colegas Ailton D. Santana Jr., Alexandre G. Ciancio, André C. Vliese, Augusto H. Dantas, Cassio B. Ribeiro, Cassio G. G. Duarte, Charles B. do Prado, Felipe R. Aquino, Lara Christiana R. L. Feio, Lisandro Lovisolo, Maria Heveline V. Duarte, Mauro F. de Carvalho, Ranniery da S. Maia e Rogério Caetano pelo companherismo e ajuda no ultrapassar dos obstáculos físicos e psicológicos.

Também, aos professores, secretárias e funcionários, que proporcionaram um ambiente profissional bastante agradável.

A todas estas pessoas, muito obrigado, esperando que possa retribuir à altura tudo o que fizeram por mim.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

TÉCNICAS DE GERAÇÃO DE SOM TRIDIMENSIONAL

Fábio Pacheco Freeland

Julho/2001

Orientadores: Paulo Sérgio Ramirez Diniz

Luiz Wagner Pereira Biscaíno

Programa: Engenharia Elétrica

O som tridimensional tem sido utilizado em várias áreas, destacando-se entre elas as de realidade virtual e afins, como a cinematográfica, a de simuladores e a de jogos eletrônicos.

Esta tese tem como objetivo principal o estudo de técnicas de geração de som tridimensional, principalmente do ponto de vista da localização e movimentação do som.

Nesse contexto, é proposta uma nova forma de interpolação com vistas à diminuição da complexidade computacional requerida pelo sistema de realidade virtual mantendo a sensação de naturalidade para o ouvinte.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

TRIDIMENSIONAL SOUND GENERATION TECHNIQUE

Fábio Pacheco Freeland

July/2001

Advisors: Paulo Sérgio Ramirez Diniz

Luiz Wagner Pereira Biscainho

Department: Electrical Engineering

Tridimensional sound has been used in many areas, such as virtual reality and those related to it, like movies, simulators and electronic games.

The main objective of this thesis is to study techniques for generation of tridimensional sound, specially from localization and sound motion points of view.

In this context, it is proposed a new form of interpolation aiming to decrease the computational complexity required by the virtual reality systems while preserving the auditory sensations natural.

Sumário

1	Introdução	1
1.1	Motivação	1
1.2	Aspectos Gerais de Implementação	2
1.2.1	Localização e Movimentação	3
1.2.2	Ambientação	5
1.3	Objetivo desta Tese	7
2	Funções de Transferência Relativas à Cabeça (HRTFs)	8
2.1	Aquisição das HRTFs	9
2.1.1	Configuração do Sistema de Medição	9
2.1.2	Cabeça Artificial	12
2.1.3	Compensação de Funções de Transferência Externas à Cabeça Artificial	13
2.2	Características das HRTFs	15
2.2.1	Fatores Envolvidos na Percepção	15
2.2.2	Importância dos Fatores na Percepção	18
2.3	Avaliação das HRTFs	25
2.4	Utilização das HRTFs para Geração de Som Tridimensional	26
3	Funções de Transferência Interposicionais (IPTFs)	29
3.1	Idéia Principal	30
3.2	Funções de Transferência Interposicionais	31
3.2.1	Geração das IPTFs	33
3.3	Interpolação Utilizando as IPTFs	35
3.3.1	Avaliação da Interpolação com as IPTFs	40

3.4	Comentários	42
4	Simplificação das IPTFs	44
4.1	Métodos para Simplificação das IPTFs	45
4.1.1	Método BMR	45
4.1.2	Identificação Adaptativa por <i>Equation Error</i>	49
4.1.3	Métodos de Suavização de Espectro	52
4.1.4	Métodos de Aproximação Complexa	55
4.2	Técnicas de Simplificação	57
4.2.1	Simplificação das HRTFs Utilizadas na Razão (IPTF)	57
4.2.2	Simplificação das IPTFs	58
4.2.3	Simplificação de Funções Suavizadas	64
4.3	Conclusões e Solução Adotada	68
5	Movimento da Fonte Virtual	71
5.1	Algoritmo de Movimento	72
5.1.1	Testes com IPTFs sem Simplificação	74
5.2	Movimento com IPTFs Simplificadas–Solução Adotada	75
6	Conclusão	82
	Referências Bibliográficas	85
A	Método de Medição de Funções de Transferência	88

Lista de Figuras

1.1	Geração de som tridimensional com sistema multi-canal	4
1.2	Comparação entre as direções do som direto e do som refletido.	6
2.1	Esquema do sistema de medição das HRTFs.	10
2.2	Obtenção do par de HRIRs para cada posição.	13
2.3	Diagrama em blocos do sistema de medição.	14
2.4	Verificação da diferença de tempo interaural (ITD).	16
2.5	ITD e ILD ao se variar a elevação da fonte.	17
2.6	Região de importância dos fatores ITD e ILD (I).	19
2.7	Região de importância dos fatores ITD e ILD (II).	20
2.8	Região de importância dos fatores ITD e ILD (III).	21
2.9	Região de importância dos fatores relacionados com as reflexões na pina e no tronco (I).	22
2.10	Região de importância dos fatores relacionados com as reflexões na pina e no tronco (II).	23
2.11	Região de importância dos fatores relacionados com as reflexões na pina e no tronco (III).	24
2.12	Método de interpolação bilinear.	28
3.1	Função de transferência interaural (ITFs).	30
3.2	Modelo das HRTFs através das ITFs.	31
3.3	Comparação das dinâmicas das funções ipsi- e contralateral com a ITF.	32
3.4	Função de Transferência Interposicional.	33
3.5	Aproximações de fase mínima.	35
3.6	VBAP interpretado com as HRTFs.	36

3.7	Limitação da região de interpolação para o conjunto de funções escolhido.	37
3.8	Distâncias angulares absolutas utilizadas no cálculo dos parâmetros α , β e γ	38
3.9	Diagrama em blocos da interpolação utilizando as IPTFs.	40
3.10	Análise gráfica do teste feito para a interpolação.	41
3.11	Exemplo do teste da interpolação (I).	42
3.12	Exemplo do teste da interpolação (II).	43
4.1	Identificação adaptativa IIR.	50
4.2	Identificação de sistemas por <i>equation-error</i>	51
4.3	Aproximação aplicando o método BMR sobre as HRTFs.	59
4.4	Aproximação aplicando o método BMR sobre as IPTFs.	61
4.5	Exemplo de resultado do método BMR sobre as IPTFs representadas por filtros IIR.	62
4.6	Redução de ordem aplicando o método <i>equation error</i>	64
4.7	Efeito das grandes variações nas altas frequências sobre o método de redução de ordem das IPTFs.	65
4.8	Exemplo de simplificação das IPTFs suavizadas com a técnica BMR com aproximação FIR.	67
4.9	Atuação da suavização em funções com “picos” elevados.	69
4.10	Exemplo de suavização utilizando a média geométrica.	70
5.1	Diagrama em blocos do algoritmo de movimento.	73
5.2	Exemplo de mudança de parâmetros durante o movimento.	74
5.3	Efeito do transitório dos filtros.	76
5.4	Exemplo de diagrama de pólos e zeros das IPTFs.	77
5.5	Deslocamento dos pólos das IPTFs em direção à origem.	78
5.6	Efeito do transitório dos filtros com redução dos módulos dos pólos próximos a -1	79
5.7	Exemplo de sinal com a supressão do transitório.	80

Capítulo 1

Introdução

1.1 Motivação

O som é um importante elemento na percepção do ambiente em nossa volta. É com auxílio dele que pessoas com deficiência visual podem ter uma vida praticamente normal. Isso se torna possível porque no som está embutida grande parte da informação do ambiente em que se está imerso.

Além dessa característica, o som tem a capacidade de transmitir/criar emoções. Essa propriedade se mostra claramente quando se toma um susto ao se ouvir um barulho um pouco mais forte ou quando se sente a emoção transmitida pelo autor de uma peça musical.

Essas características do som têm sido largamente utilizadas pela indústria cinematográfica e em jogos eletrônicos para criar sensações cada vez mais realísticas e emocionantes.

Sistemas de realidade virtual devem gerar um som que seja o mais real possível para uma situação determinada ao sistema; por isso, necessitam de técnicas de geração de efeitos sonoros mais precisas.

Algumas técnicas que vêm sendo vastamente utilizadas nessas aplicações são as de geração de som tridimensional. Estas técnicas ganharam força quando o sistema digital de reprodução multicanal começou a ser utilizado em salas de cinema, pois as equipes de efeitos sonoros, pensando nesse tipo de sistema de reprodução, começaram a utilizar efeitos de localização e movimento, criando um som mais envolvente.

Essas técnicas se dividem em dois principais grupos, do ponto de vista da reprodução: as que geram o som tridimensional em caixas acústicas e as que geram em fones de ouvido. As relacionadas com caixas acústicas podem-se dividir ainda naquelas que geram o efeito utilizando dois canais e naquelas que geram o efeito utilizando mais de dois canais.

A técnica a ser utilizada depende da aplicação. Em sistemas de realidade virtual, a utilização de fones de ouvido torna-se mais adequada, tendo em vista que esse sistema é utilizado por apenas um usuário; e, mesmo que seja utilizado por mais de um usuário, o que é ouvido por um não pode ser escutado pelo outro da mesma maneira. Já em sistemas de projeção de filmes, a utilização de fones de ouvido se torna pouco prática, pois, além de o sistema ser muito mais sujeito a problemas técnicos, a utilização de fones parece um pouco desconfortável para o espectador. Nesse caso, o sistema multicanal se torna mais adequado. No caso dos jogos eletrônicos, a forma de reprodução é escolhida mais por razões econômica e prática do que qualquer outra coisa. Na maioria dos casos, o sistema de reprodução encontrado em computadores pessoais e videogames tem como terminais duas caixas acústicas. E, como um jogo não deve exigir equipamentos especiais para seu funcionamento (isso pode acarretar a não aceitação do jogo pelo mercado), os efeitos sonoros devem estar voltados para a reprodução em caixas. No máximo, a utilização de um par de fones de ouvido pode ser sugerida.

Para cada uma dessas técnicas de geração de sons tridimensionais, diferentes aspectos são levados em conta na implementação. Alguns deles serão discutidos na próxima seção.

1.2 Aspectos Gerais de Implementação

As técnicas utilizadas em um sistema de geração de som tridimensional podem ser divididas em dois principais grupos: o responsável pela localização espacial do som e o responsável pela sua ambientação.

1.2.1 Localização e Movimentação

A separação em duas partes pode ser feita porque a sensação de localização tem como principal causa a geometria da cabeça humana. A frente de onda gerada por uma fonte em uma determinada posição percorre dois caminhos principais em direção aos dois ouvidos, respectivamente. Nesse percurso, algumas modificações, diferentes para cada um dos caminhos, são atribuídas às frentes de onda. Essas alterações atribuídas ao som são o resultado de reflexões no tronco e na pina (ouvido externo) e da difração causada pela cabeça (chamada também de sombra da cabeça).

Dependendo da forma de reprodução, a criação da sensação de localização se dá de maneiras bem distintas. No caso da utilização de caixas acústicas, a preocupação está em gerar um campo acústico que dê a impressão de localização, já que as modificações relacionadas ao conjunto cabeça-tronco-pina serão atribuídas pelo próprio conjunto do ouvinte. Esse campo acústico é gerado pela composição das frentes de onda referentes a cada caixa acústica.

No caso dos sistemas multicanal, a posição de cada caixa deve ser determinada e cada uma é excitada de forma a compor o campo desejado. Nessa forma de reprodução, que é chamada de Base de Vetores Balanceada em Amplitude (*VBAP-Vector Base Amplitude Panning* [1]), cada caixa é representada por um vetor com origem na posição do ouvinte, e o módulo desses vetores é tal que a soma vetorial deles tem módulo 1 na direção desejada para a fonte virtual (veja na Figura 1.1).

Para o caso estéreo com caixas acústicas, outras alternativas como inserção de atraso e diferenças na fase entre os canais [2] devem ser utilizadas. Mas problemas como interferência entre canais acabam por prejudicar a eficiência do sistema de localização.

A utilização de caixas acústicas torna o problema da interferência bastante relevante [3], por isso devem ser utilizadas caixas bem direcionais e a sala de reprodução deve ser preparada acusticamente para ter poucas reflexões. Pode-se notar também que, ao se utilizarem apenas dois canais, não é possível, de uma forma simples, conseguir posicionar a fonte virtual fora do plano formado pelos pontos correspondentes às posições das caixas e do ouvinte. No caso multicanal, isso se torna possível, mas devem existir caixas acima e abaixo do plano horizontal. Na Figura 1.1, vê-se a fonte virtual acima do plano horizontal. Para que seja possível

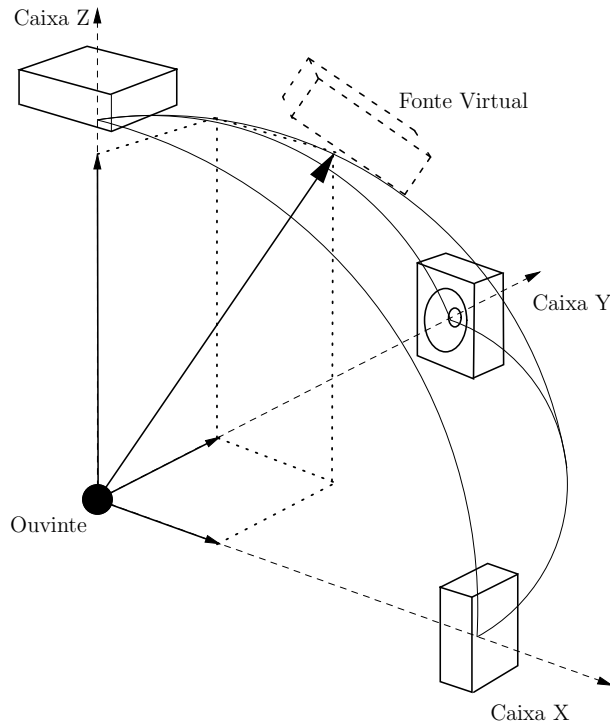


Figura 1.1: Esquema de geração de som tridimensional com sistema multicanal.

posicionar a fonte virtual abaixo do plano horizontal, é necessário mais um canal com a caixa a ele relacionada posicionada abaixo desse plano (não necessariamente no eixo vertical).

Já no caso estéreo utilizando fones de ouvido, a mesma sensação obtida com o campo acústico deve ser reproduzida. Para isso, é necessário saber como o caminho modifica o som; isso significa que é preciso levar em consideração as alterações ocasionadas pelo conjunto cabeça-tronco-pina. Essa informação de como o conjunto cabeça-tronco-pina altera o som pode ser representada por funções de transferência entre o ponto onde está a fonte virtual e cada um dos ouvidos. Essas funções de transferência, chamadas de Funções de Transferência Relativas à Cabeça (*HRTFs—Head-Related Transfer Functions*) são medidas com o auxílio de uma cabeça artificial.

A cabeça artificial consiste de uma espécie de cabeça de manequim oca com orifícios no lugar das orelhas. Por dentro da cabeça, 2 microfones muito sensíveis e calibrados são colocados nos orifícios. Por fora da cabeça, no entorno desses orifícios são colocados orelhas moldadas em material especial (parecido com uma borracha), de forma que a cabeça possa ser considerada um bom modelo para a cabeça humana

média. Em uma sala anecóica, repete-se para vários pontos diferentes a medição da função de transferência entre uma fonte sonora real e os ouvidos da cabeça artificial. Essas medições são feitas em pontos pertencentes a uma esfera de raio r que envolve a cabeça. Ao final das medições, obtêm-se duas funções de transferência para cada ponto medido, uma referente a cada ouvido.

Em [4], Gardner e Martin executaram tais medições, que depois disponibilizaram na *Internet*. Essas funções medidas por eles são utilizadas em todo este trabalho.

Como não é possível obter as HRTFs para todas as direções possíveis, torna-se necessária uma forma de interpolação para que seja possível posicionar a fonte virtual em todos os pontos em volta do ouvinte. Em [5], é utilizada uma interpolação bilinear. Essa interpolação consiste de uma ponderação das HRTFs mais próximas que delimitam a região no espaço onde se encontra a posição desejada para a fonte virtual.

A interpolação se torna mais necessária ainda quando se precisa dar noção de movimento ao som. O movimento exige a interpolação para que não se gere um som se movimentando “em saltos” e para que o número de HRTFs medidas não aumente muito.

1.2.2 Ambientação

O som tridimensional gerado somente a partir das HRTFs tem a informação de localização. Essa informação, no entanto, não é suficiente, tendo em vista que o som que estamos acostumados a ouvir carrega, também, bastante informação do ambiente em que foi gravado. Isso implica voltar um pouco da atenção para a criação da ambientação adequada à posição virtual desejada.

A ambientação é construída por modificadores externos ao sistema cabeça-tronco-pina mencionado anteriormente: reflexões em anteparos, absorção atmosférica e absorção pelos anteparos são efeitos atribuídos à ambientação. Tais modificadores não são totalmente independentes da localização, pois para cada nova frente de onda criada por reflexões está implícita uma localização. Por exemplo, uma frente de onda refletida em um anteparo chega aos ouvidos percorrendo um caminho diferente do som direto (que parte da fonte para os ouvidos sem encontrar nenhum

anteparo). Os efeitos da reflexão são aplicados ao som no momento da reflexão, e o de localização, ao chegar ao ouvinte. Portanto, esse som refletido tem a ele atribuída uma posição diferente da posição da fonte.

Isso pode ser percebido melhor na Figura 1.2. Nessa Figura, pode-se notar a diferença de direção entre o som direto e o refletido. Nota-se também que, no caso dos sons refletidos, o sistema de localização deve atribuir a cada reflexão uma direção diferente, de forma dar uma sensação mais real do ambiente em questão.

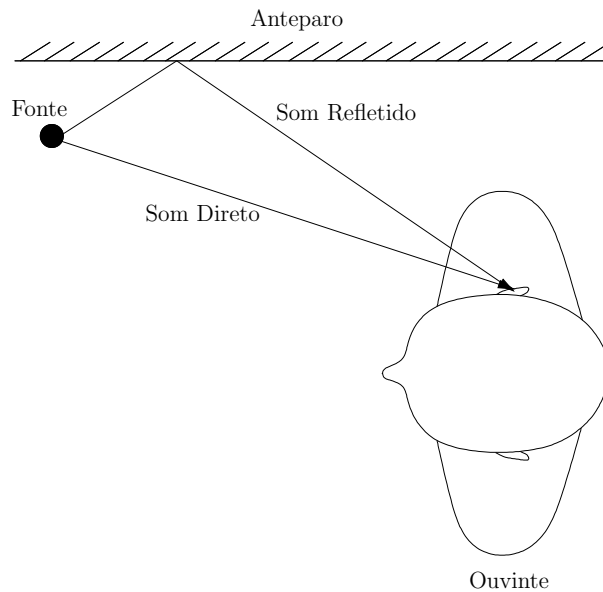


Figura 1.2: Comparação entre as direções do som direto e do som refletido.

Além das reflexões, a ambientação artificial também carrega em si a informação da distância à fonte (já que as HRTFs são medidas a uma distância fixa da cabeça artificial), das absorções atmosféricas e daquelas ocorridas quando das reflexões em anteparos.

Quanto à distância da fonte, pode-se utilizar a lei do inverso do quadrado da distância [2], que dá a amplitude do sinal como proporcional ao inverso da distância do ouvinte à fonte. Isso pode ser feito porque o modelo adotado para a fonte, nesse caso, é o de uma fonte omnidirecional e, assim, essa redução da energia com a distância é causada simplesmente pelo fato de a energia se distribuir igualmente em todas as direções.

Existe um efeito que é desconsiderado pela parcela da ambientação mostrada acima: a absorção atmosférica. O efeito da absorção depende não só da distância como também da frequência. Em [6], é mostrada uma forma analítica de se calcular

a atenuação imposta pela distância para cada frequência como consequência dessa absorção sem levar em conta a atenuação modelada pela lei do inverso da distância, que atua em todas as frequências igualmente. Se esse efeito for, então, utilizado em conjunto com a lei do inverso da distância, resolve-se o caso das atenuações dependentes da distância [5, 7].

Para o caso das atenuações causadas em reflexões, pode-se utilizar um fator de absorção inerente a cada superfície. Esse modelo não é absolutamente rigoroso, pois é sabido que para a reflexão ser perfeita os comprimentos de onda das frequências que compõem o som devem ser menores que o menor comprimento da superfície do anteparo. Portanto, para sons com comprimentos de onda maiores (frequências baixas), pode ocorrer uma parcela importante de reflexão difusa que os atenuará, pelo fato de distribuir a energia da onda incidente em várias direções. Essa variação do fator de absorção com a frequência normalmente não é modelada como tal, mas o efeito da difusão pode ser considerada.

1.3 Objetivo desta Tese

Esta tese se ocupa da implementação de um sistema de geração de som tridimensional utilizando fones de ouvido.

Tem como objetivo principal o estudo das técnicas de geração de som tridimensional, principalmente do ponto de vista da localização e movimentação do som, para servir de base a futuros estudos.

Além disso, é proposta uma nova forma de interpolação das HRTFs visando à redução da complexidade computacional na geração do movimento, diminuindo o número de multiplicações por amostra calculada.

Tem-se em vista a aplicação dessa técnica em sistemas de realidade virtual, jogos eletrônicos e quaisquer aplicações que se utilizem de fones de ouvido e requeiram do som uma maior sensação de realidade.

Capítulo 2

Funções de Transferência Relativas à Cabeça (HRTFs)

Todos os sistemas de geração de som tridimensional com vistas à reprodução utilizando fones de ouvido têm, de alguma forma, que levar em consideração as modificações atribuídas às frentes de onda sonora pelo conjunto receptor que consiste da cabeça, da pina e do tronco.

Essas modificações podem ser representadas por funções de transferência dependentes da direção de que está vindo o som. Essas funções de transferência são chamadas de Funções de Transferência Referentes à Cabeça (*Head-Related Transfer Functions*–HRTFs).

Essas funções, como já foi mencionado, são medidas de várias direções diferentes em condições especiais. Um conjunto delas foi obtido por Gardner e Martin em [4] e, na Seção 2.1, será mostrada a descrição do procedimento de medição utilizado naquela referência para se ter uma visão mais precisa do processo de geração de som tridimensional.

Após a Seção 2.1–**Aquisição das HRTFs**, nesse capítulo discutem-se algumas características importantes das HRTFs na Seção 2.2–**Características das HRTFs**; tecem-se alguns comentários sobre os métodos de avaliação das HRTFs medidas na Seção 2.3–**Avaliação das HRTFs**; e mostra-se como utilizar as HRTFs para geração do som tridimensional na Seção 2.4–**Utilização das HRTFs para Geração de Som Tridimensional**.

2.1 Aquisição das HRTFs

Um dos pontos principais na geração de som tridimensional em fones de ouvido é a obtenção das HRTFs para várias direções. Como nesta tese foram utilizadas HRTFs medidas em [4], será feita nesta seção uma breve explanação de como se devem realizar as medições, mencionando-se o sistema utilizado e como é a estrutura da cabeça artificial com as quais foram medidas as HRTFs.

Além disso, serão vistas também quais as funções de transferência auxiliares que devem ser medidas além das HRTFs para que se possa ter um conjunto de HRTFs que só digam respeito à função de transferência entre o ponto onde está a fonte e os ouvidos. Afinal, a resposta em frequência do sistema de reprodução utilizado na medição não deve interferir na obtenção das HRTFs.

2.1.1 Configuração do Sistema de Medição

O sistema de medição utilizado por Gardner e Martin é baseado em um computador equipado com uma placa de aquisição/reprodução de áudio que, em conjunto com um sistema de caixas acústicas e uma cabeça artificial dentro de uma sala anecóica, obtém a função de transferência na forma da resposta impulsiva para cada direção.

Para se poder medir as HRTFs em várias direções diferentes, a cabeça artificial é montada sobre uma base giratória, controlada pelo computador, que pode posicionar a frente da cabeça, com bastante precisão (centésimo de grau), em um ângulo de azimute θ relativo ao azimute da caixa acústica (referência zero). A caixa, por sua vez, é fixada a um trilho que a deixa a uma distância fixa de $1,4m$ do centro da cabeça, deixando mudar somente sua posição angular de elevação ϕ . Na Figura 2.1, pode-se ver um esquema do sistema que deve ser montado na sala anecóica.

Nota-se, nessa figura, que o trilho onde é afixada a caixa não está sendo mostrado. Somente suas possíveis posições é que estão representadas. O trilho não só mantém fixa a distância da caixa ao centro da cabeça como também mantém o azimute da caixa igual a zero. Outro detalhe são os ângulos mínimo e máximo de elevação: só são medidas as HRTFs com elevação entre -40° e 90° .

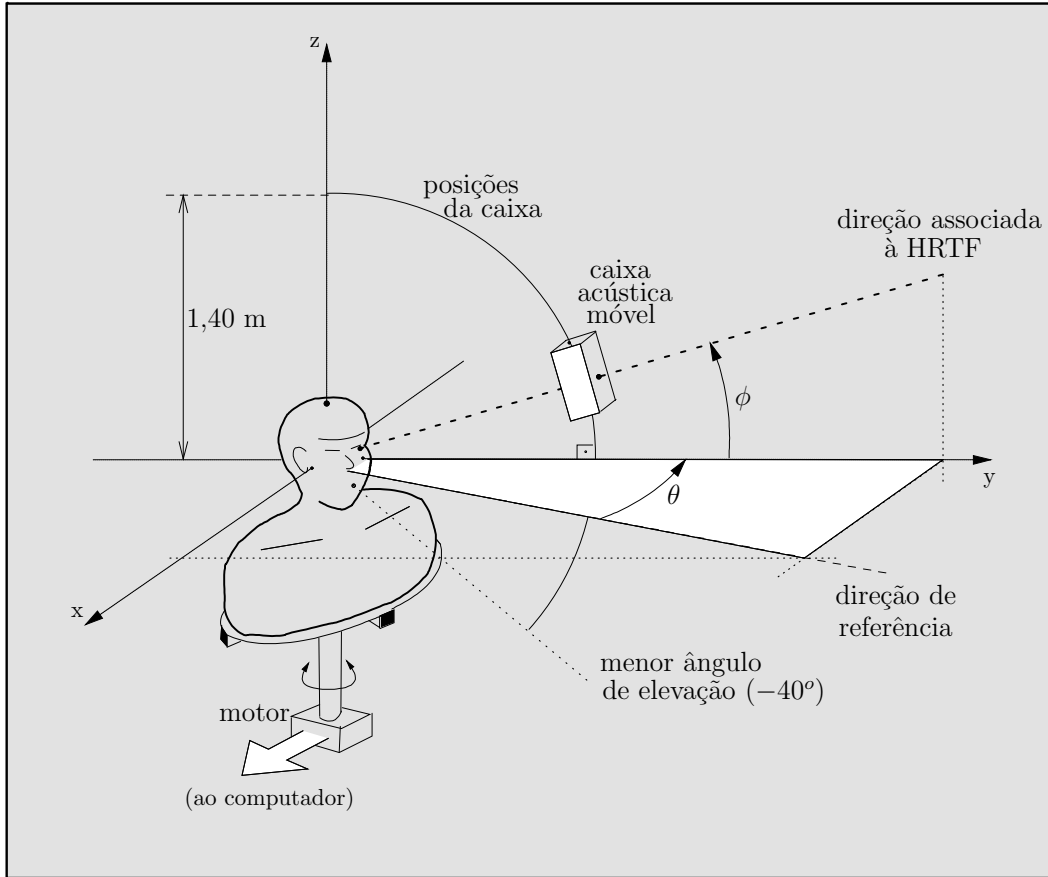


Figura 2.1: Esquema do sistema de medição das HRTFs utilizado por Gardner e Martin em [4]: θ é o ângulo de azimute e ϕ é o ângulo de elevação.

Com a configuração mostrada, efetuam-se as medidas¹ reproduzindo-se uma seqüência pseudo-aleatória chamada de ML (*Maximum Length*) [8] e, simultaneamente, gravando-se o sinal captado pelos microfones da cabeça artificial. O sinal de medida é formado por $2N$ amostras resultantes da duplicação das N amostras da seqüência ML. Das $2N$ amostras medidas em cada canal, as N primeiras amostras são descartadas e as N restantes são novamente duplicadas. Com essas $2N$ amostras, calcula-se a correlação cruzada com a seqüência original de N amostras, gerando um resultado de $3N - 1$ amostras. A resposta impulsiva de N amostras, chamada de HRIR (*Head-Related Impulse Response*), relativa à HRTF que está sendo medida pode ser obtida extraindo-se as amostras do resultado anterior a partir da amostra $N - 1$. Nessas medidas, Martin e Gardner utilizaram $N = 16383$. Esse procedimento de medida é repetido para cada uma das posições desejadas. Em [4], foram medidas

¹No apêndice A, tem-se uma explicação do procedimento utilizado na medição das HRTFs.

Tabela 2.1: Número de HRTFs medidas e passo em azimute para cada elevação.

Elevação (ϕ)	Número de medidas	Passo em azimute ($\Delta\theta$)
-40	56	6,43
-30	60	6,00
-20	72	5,00
-10	72	5,00
0	72	5,00
10	72	5,00
20	72	5,00
30	60	6,00
40	56	6,43
50	45	8,00
60	36	10,00
70	24	15,00
80	12	30,00
90	1	—

respostas com elevações entre -40° e 90° , com passo de 10° , e, nos azimutes de 0° a 360° , com passo dependente da elevação (veja Tabela 2.1).

Cada uma das medidas gera, para cada pina, uma resposta impulsiva com 16383 amostras a uma taxa de 44100 Hz. Delas, observou-se que as primeiras 230 amostras eram relativas ao atraso causado pela propagação no ar no percurso entre a caixa e a cabeça em conjunto com o atraso entre as operações *play/record*. Além disso, Martin e Gardner notaram que após algumas centenas de amostras, além do efeito da resposta à cabeça artificial ser muito pequeno, aparecem reflexões causadas pelo equipamento montado na câmara anecóica, como, por exemplo, da mesa onde está montada a cabeça artificial. De forma a reduzir a base de dados relativos às HRTFs sem que nada de interesse fosse perdido, das 16383 amostras da resposta impulsiva eliminaram-se as 200 amostras iniciais (atrasos relativos à propagação no ar e o atraso entre a operação *play/record*) e extraíram-se como HRIR as 512 amostras seguintes (pois após essas amostras a resposta da cabeça atingia valores muito pequenos).

A essas HRIRs de 512 amostras, foram incorporadas as respostas em frequência do sistema amplificador/caixa acústica utilizado na medição (a discussão de funções de transferências que interferem nas medidas será feita na Subseção 2.1.3). Ao incorporar a resposta do sistema de medição, a amplitude das HRIRs atingem valores baixos em um tempo mais curto, então, pode-se reduzir o número de amostras de cada HRIR para 128. Sendo assim, criou-se um conjunto de HRIRs com 128 amostras que já levam em consideração (compensam) a resposta em frequência do sistema da medição. Essas foram as funções de transferência utilizadas nesta tese.

2.1.2 Cabeça Artificial

A cabeça artificial consiste de um par de microfones de precisão montados na parte interna da cabeça de um manequim. Nesse manequim mostrado na Figura 2.1, com tronco superior (torso), pescoço e cabeça, existem dois orifícios no lugar dos ouvidos, onde os microfones são colocados para captar o que possivelmente os tímpanos de uma pessoa receberiam.

A cabeça artificial utilizada por Martin e Gardner é uma KEMAR–*Knowles Electronics Mannequin for Acoustic Research* (modelo DB-4004) equipada com dois modelos diferentes de pina: uma de tamanho médio para o ouvido esquerdo (modelo DB-061) e uma de tamanho um pouco maior para o ouvido direito (modelo DB-065). O conjunto microfone/preamplificador é do modelo Etymotic ER-11. As saídas do pré-amplificador são ligadas às entradas do conversor A/D de uma placa Audiomedia, que equipa o computador que controla as medições.

Com a cabeça artificial montada dessa forma, obtêm-se para os dois tipos de pina respostas impulsivas diferentes em uma mesma elevação e em azimutes simétricos, com isso, em uma mesma elevação, a resposta ao ouvido direito no azimute θ é diferente da resposta ao ouvido esquerdo no azimute $-\theta$. Assim, com as medidas realizadas para os ouvidos direito e esquerdo, geram-se os pares de HRIRs que representam a resposta de uma cabeça artificial simétrica com pina de tamanho médio e pouco acima da média, respectivamente.

Ao final, um par de HRIRs relativo a uma determinada posição (azimute θ e elevação ϕ) é formado por duas respostas do mesmo ouvido: uma no azimute θ e outra de azimute $-\theta$ como ilustrado na Figura 2.2. O par referente à pina de tamanho

médio é obtido agrupando as HRTFs $H_{DB-061}(\theta)$ e $H_{DB-061}(-\theta)$ para representar os ouvidos esquerdo e direito, respectivamente. Já o outro par, $H_{DB-065}(-\theta)$ e $H_{DB-065}(\theta)$, deve representar os ouvidos esquerdo e direito, respectivamente, para a pina de tamanho acima da média. Por exemplo, na elevação de 20° o par de HRIRs relativo ao azimute 30° com o tamanho de orelha médio é formado pelas HRIRs medidas pelo microfone montado no ouvido esquerdo (lado da pina média) que tem elevação 20° com azimutes 30° e -30° para representar as respostas dos ouvidos esquerdo e direito, respectivamente. No caso da resposta da cabeça artificial com a pina de tamanho um pouco maior, as HRIRs são as medidas com o microfone direito, na mesma elevação mas com os azimutes trocados: -30° para representar o ouvido esquerdo e 30° para o direito.

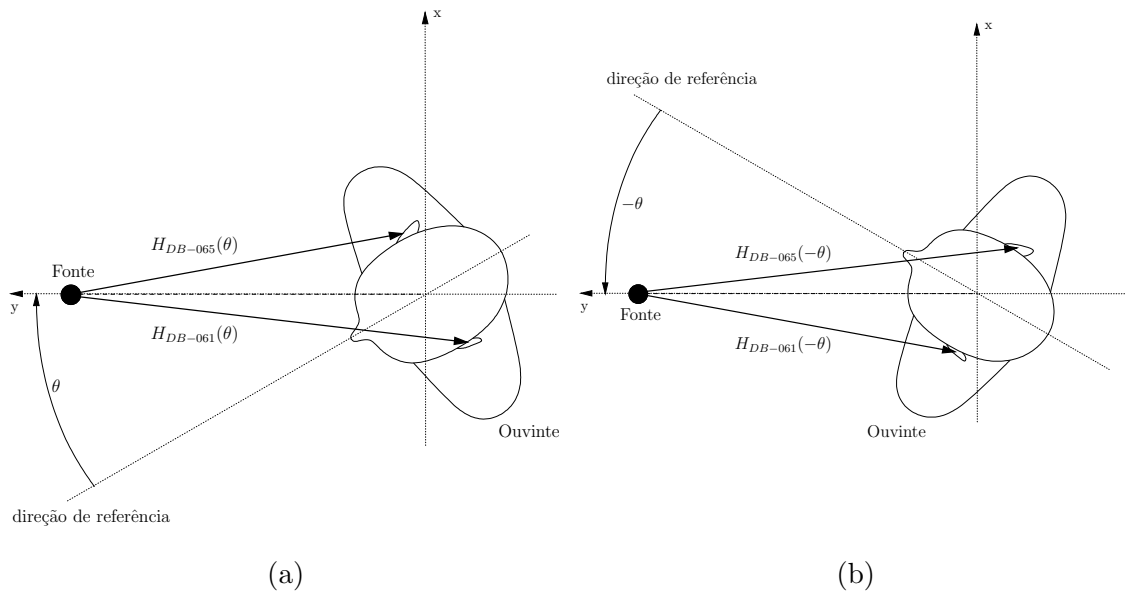


Figura 2.2: Obtenção do par de HRIRs para cada posição. Em (a), com a fonte localizada no azimute de θ e em (b) com a fonte localizada no azimute de $-\theta$.

2.1.3 Compensação de Funções de Transferência Externas à Cabeça Artificial

Como em qualquer procedimento de medida de função de transferência, é necessário ter o cuidado de isolar completamente a função de transferência que se quer medir da influência dos instrumentos utilizados em tal medida.

No caso das medidas das HRIRs, o sistema utilizado para a medição pode

influir no resultado da medida. As respostas do amplificador, da caixa acústica, da sala onde é feita a medida e do microfone devem ser devidamente compensadas para que nenhuma dessas partes do sistema de medida venha interferir na medição da função de transferência desejada.

O esquema da Figura 2.3 mostra as funções de transferência envolvidas na medição das HRTFs. Nessa figura, vê-se o caminho percorrido pelo sinal de medição

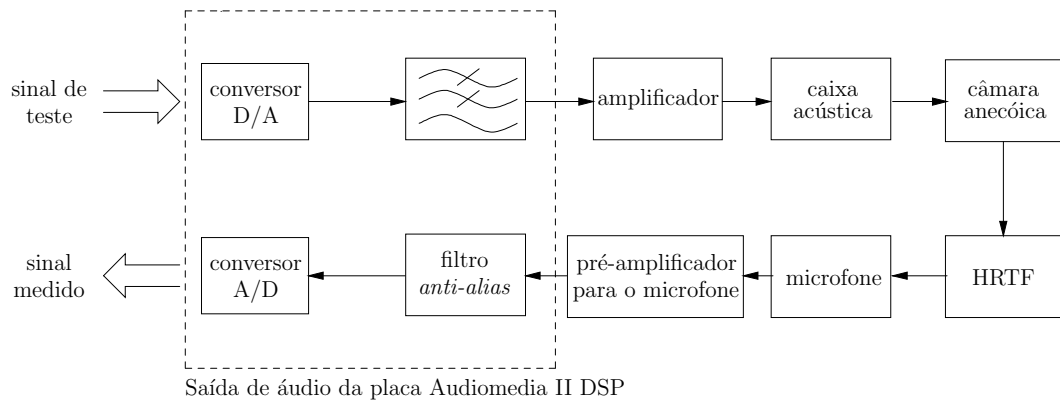


Figura 2.3: Diagrama em blocos do sistema de medição. Funções de transferência envolvidas na medição.

das HRIRs. Se o amplificador for considerado de alta qualidade (com resposta em frequência praticamente plana na região de interesse—20Hz a 20kHz) e que o par microfone/pré-amplificador tem também resposta bastante plana, as funções de transferência que devem ser compensadas são as da câmara anecóica e da caixa acústica.

A resposta da câmara anecóica é formada pelas reflexões provenientes dos equipamentos montados no interior da mesma, bem como pelo eco residual das próprias paredes. Esses dois efeitos ocorrem normalmente bem depois da resposta da cabeça ao sinal direto. Por isso, não é necessária uma compensação para essas funções de transferência.

Já a resposta da caixa acústica, que quase sempre não é nada plana, deve ser compensada. Para isso, faz-se a medida da sua resposta impulsiva separadamente utilizando-se a mesma técnica de medida das HRIRs, mas com um microfone colocado no lugar da cabeça, e implementa-se o filtro inverso correspondente. Gardner e Martin fizeram tal medição e, como já foi mencionado, incluíram nas HRIRs utilizadas nesta tese a compensação da caixa acústica.

Além da resposta à caixa acústica, uma outra função de transferência envolvida, presente no momento da reprodução do sinal tridimensional, é a relativa ao fone de ouvido. Essa deve ser também compensada. Gardner e Martin fizeram algumas medidas de respostas para alguns tipos de fones de ouvido. No entanto, neste trabalho não foram utilizadas, porque os fones utilizados nos testes não eram do mesmo modelo dos que foram medidos. Além disso, há normalmente uma discrepância muito grande entre as funções de transferência de unidades diferentes de um mesmo tipo de fone ou até mesmo entre as de um mesmo fone simplesmente ao tirar e colocar o fone na cabeça artificial no momento da medição de sua resposta.

2.2 Características das HRTFs

As HRTFs variam muito de indivíduo para indivíduo. Essas variações se devem às diferenças da geometria do conjunto cabeça-tronco-pina de cada um. Isso torna também variável a percepção pelos indivíduos dos efeitos gerados com as HRTFs medidas com a cabeça artificial, a qual representa um indivíduo médio. Assim sendo, a adequação individual às sensações transmitidas pelo som gerado artificialmente se faz necessária.

Podem-se, no entanto, destacar características importantes, do ponto de vista da percepção da correta localização do som, verificando-se a diferença existente entre HRTFs adjacentes. Dois tipos de características podem ser obtidos: os fatores responsáveis pelas diferenças na percepção da localização e em quais regiões a variação da posição afeta mais um determinado fator. É o que será discutido a seguir.

2.2.1 Fatores Envolvidos na Percepção

A resposta da cabeça artificial, por sua geometria, sugere alguns pontos importantes na percepção da localização do som. Como já foi dito, as diferenças entre as respostas dos ouvidos são as responsáveis pela localização da fonte sonora. Alguns fatores, que podem ser observados pela comparação de HRIRs, mostram-se importantes na percepção dessas diferenças. São eles a diferença de tempo interaural (ITD—*Interaural Time Difference*), a diferença de nível interaural (ILD—*Interaural Level Difference*) e as reflexões no tronco e na pina. Cada um desses fatores atua

diferentemente em cada região e, ainda, é responsável pela percepção das variações de posição.

A ITD é o resultado do atraso relativo entre as respostas de cada ouvido. Esse atraso resulta da diferença de caminho entre a fonte e cada um dos ouvidos. Essa diferença de caminho é causada pela distância entre os ouvidos e pela difração causada pela cabeça. Na Figura 2.4, pode-se perceber que a diferença de atrasos existente

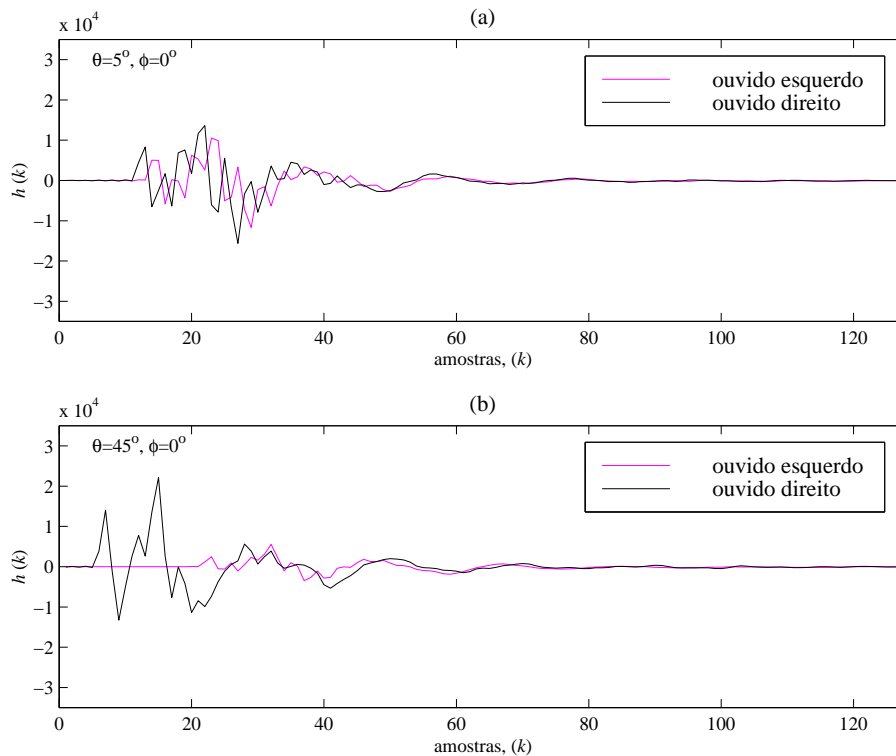


Figura 2.4: Verificação da diferença de tempo interaural (ITD). HRIRs do canal direito e esquerdo na elevação de zero grau e azimutes de 5° (a) e 45° (b).

entre as HRIRs do ouvido direito e esquerdo para o azimuth de 5° (Figura 2.4a) é bem menor que a diferença para o azimuth de 45° (Figura 2.4b). Os dois casos mostrados estão na elevação de zero grau. Note que a ITD está muito associada a variações de azimuth, já que para uma mesma elevação e azimutes diferentes, a diferença de atraso muda muito.

A ILD se caracteriza pela diferença de amplitude do som percebido por cada um dos ouvidos. Essa diferença é coerente com a ITD, sendo as duas em conjunto um indicativo do ângulo de azimuth onde se encontra a fonte sonora. Na Figura 2.4b, mostrada anteriormente, pode-se notar a existência da ILD. Comparando-a com a Figura 2.4a, pode-se perceber que, à medida que a fonte vai se deslocando para um

dos lados da cabeça, a amplitude percebida pelo ouvido do lado oposto diminui. Esse é o efeito chamado de “sombra da cabeça”.

Como se pôde perceber, a ITD e ILD carregam bastante informação sobre o azimute em que se encontra a fonte sonora. No entanto, ao se variar a elevação, essas diferenças não mudam muito. Isso pode ser notado ao se compararem duas HRIRs referentes ao azimute zero com elevações diferentes. Nesse azimute, as HRIRs medidas para o ouvido direito e esquerdo são iguais, por causa da simetria adotada e, assim, fica fácil comparar as variações devido somente à elevação. Sendo assim, na Figura 2.5a, onde se vêem duas das HRIRs relativas ao azimute zero em elevações

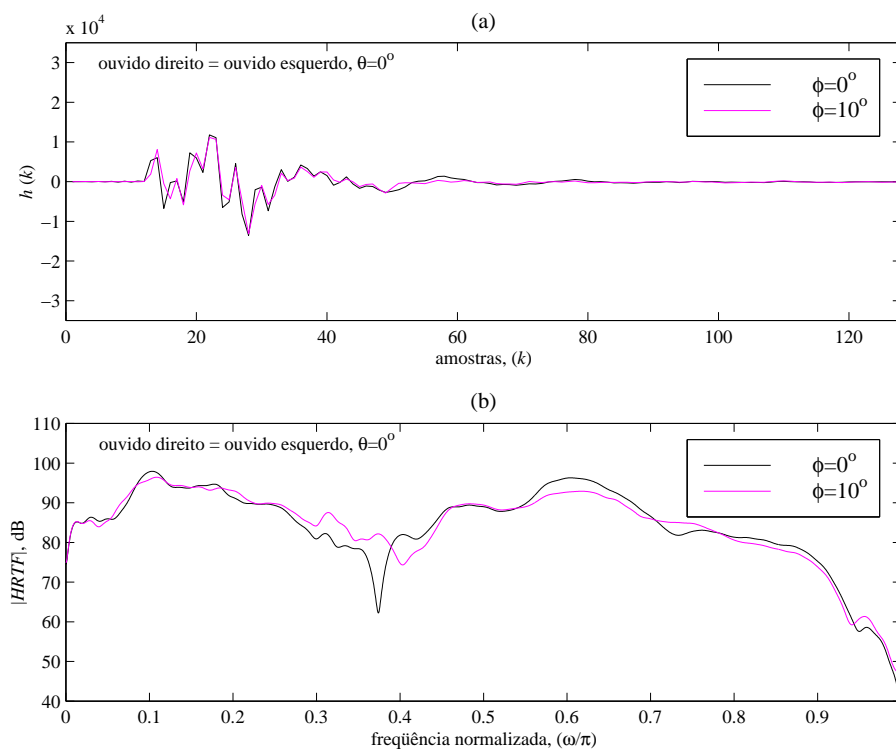


Figura 2.5: ITD e ILD ao se variar a elevação da fonte. Respostas impulsivas relativas ao azimute zero grau (a) e suas respectivas respostas em frequência (b). Essas HRTFs são iguais para o ouvido direito e esquerdo em uma mesma elevação.

diferentes, a diferença entre as posições das duas funções não se mostra através da ITD e da ILD, já que elas não existem nesse caso. Como se pode notar na Figura 2.5b, as diferenças são muito mais evidentes no módulo da resposta em frequência. Isso indica que as mudanças das características das HRTFs relacionadas com mudanças de elevação são mais visíveis na resposta à geometria da cabeça artificial, e assim, não só a distância entre os ouvidos (que gera a ITD) e a sombra

da cabeça (que gera a ILD e modifica um pouco a ITD), mas também as reflexões no tronco e na pina alteram significativamente a HRTF. É possível notar também a importância dos vales e picos para se distinguir em qual elevação se encontra a fonte sonora.

2.2.2 Importância dos Fatores na Percepção

Como foi visto na subseção anterior, existem alguns fatores, observados na comparação de HRIRs, que podem ser utilizados na distinção da posição de uma fonte sonora.

Cada um desses fatores contém um tipo de informação responsável pela distinção e, como consequência dessa diferenciação, certos fatores são mais importantes do que outros em determinadas regiões do espaço, como pode ser visto a seguir.

A ITD e a ILD são fatores influenciados pela distância entre os ouvidos (incluindo a difração causada pela cabeça) e pelo efeito da sombra da cabeça, respectivamente. Intuitivamente, pode-se perceber que as variações desses fatores são maiores à medida que a fonte sonora tem seu ângulo de azimute alterado e, mais ainda, quando o ângulo de azimute está mais próximo dos azimutes zero grau e 180° . Então, se variarmos o ângulo de azimute no entorno do ângulo de 90° , a ITD e a ILD quase não se alteram, diferentemente do caso em que se altera o azimute no entorno dos ângulos de zero grau e 180° (veja Figuras 2.6, 2.7 e 2.8).

Nas Figuras 2.6, 2.7 e 2.8 pode-se notar que a ITD e a ILD são mais sensíveis à variação de azimute próximo dos ângulos de zero grau (Figuras 2.6a e 2.6b) e 180° (Figuras 2.8a e 2.8b) do que próximo de 90° (Figuras 2.7a e 2.7b). Nessas figuras, para efeito de verificação, tem-se também a comparação das respostas em frequência de cada ouvido (Figuras 2.6c, 2.6d, 2.7c, 2.7d, 2.8c, 2.8d). Na comparação, pode-se perceber também que não há variações muito bruscas entre as respostas de um azimute e de outro, tanto próximo aos ângulos de zero grau e 180° quanto do ângulo de 90° .

A pequena e gradativa variação das respostas em frequência para cada ouvido ao se variar o ângulo de azimute da fonte sonora não se mantém quando, ao invés de se mudar o ângulo de azimute, se muda o ângulo de elevação (veja nas Figuras 2.9, 2.10 e 2.11). Nessas figuras, pode-se notar que o módulo da resposta em

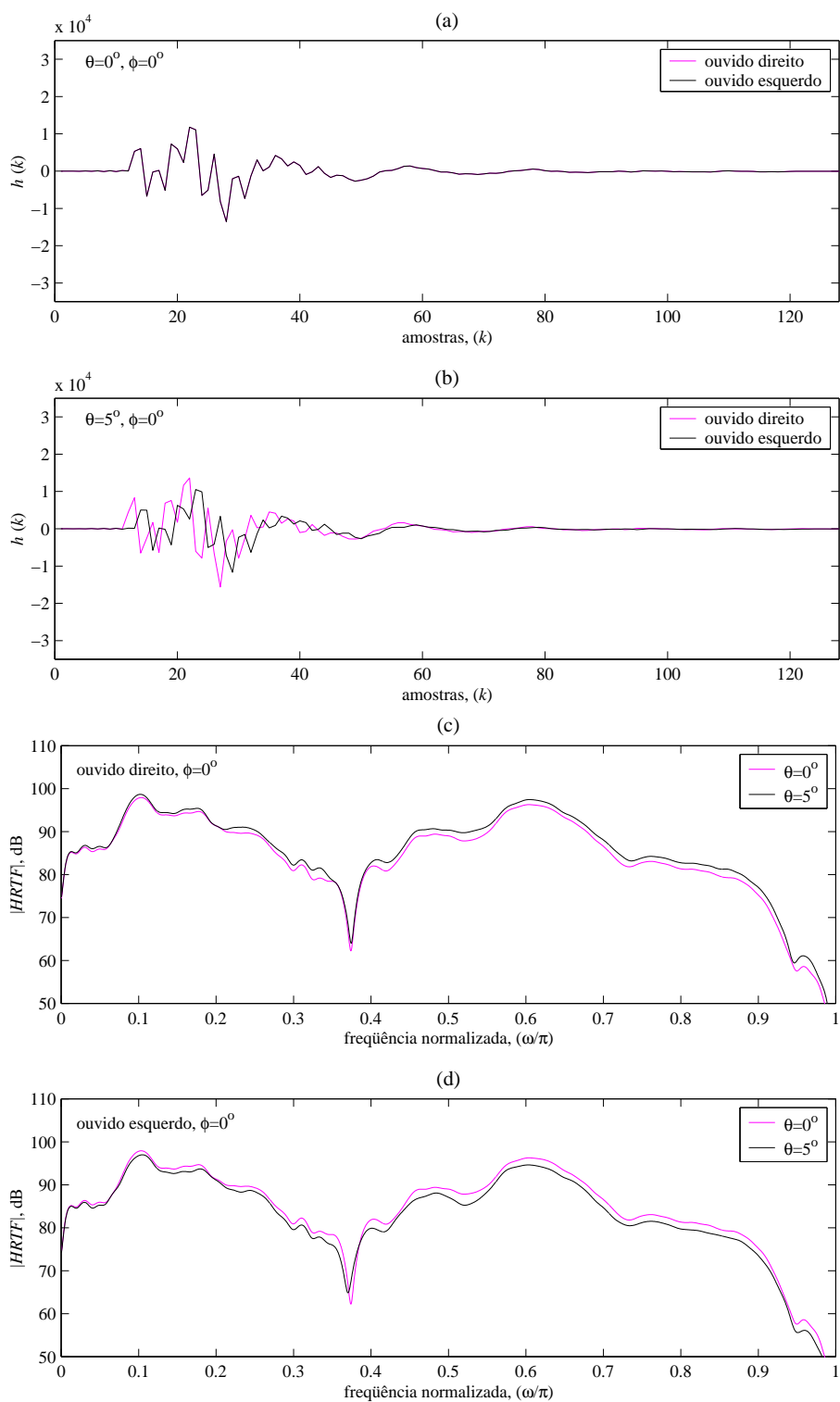


Figura 2.6: Região de importância dos fatores ITD e ILD (I). Comparação das respostas impulsivas (a, b) e dos módulos das respostas em frequência (c, d) das HRTFs com as das Figuras 2.7 e 2.8.

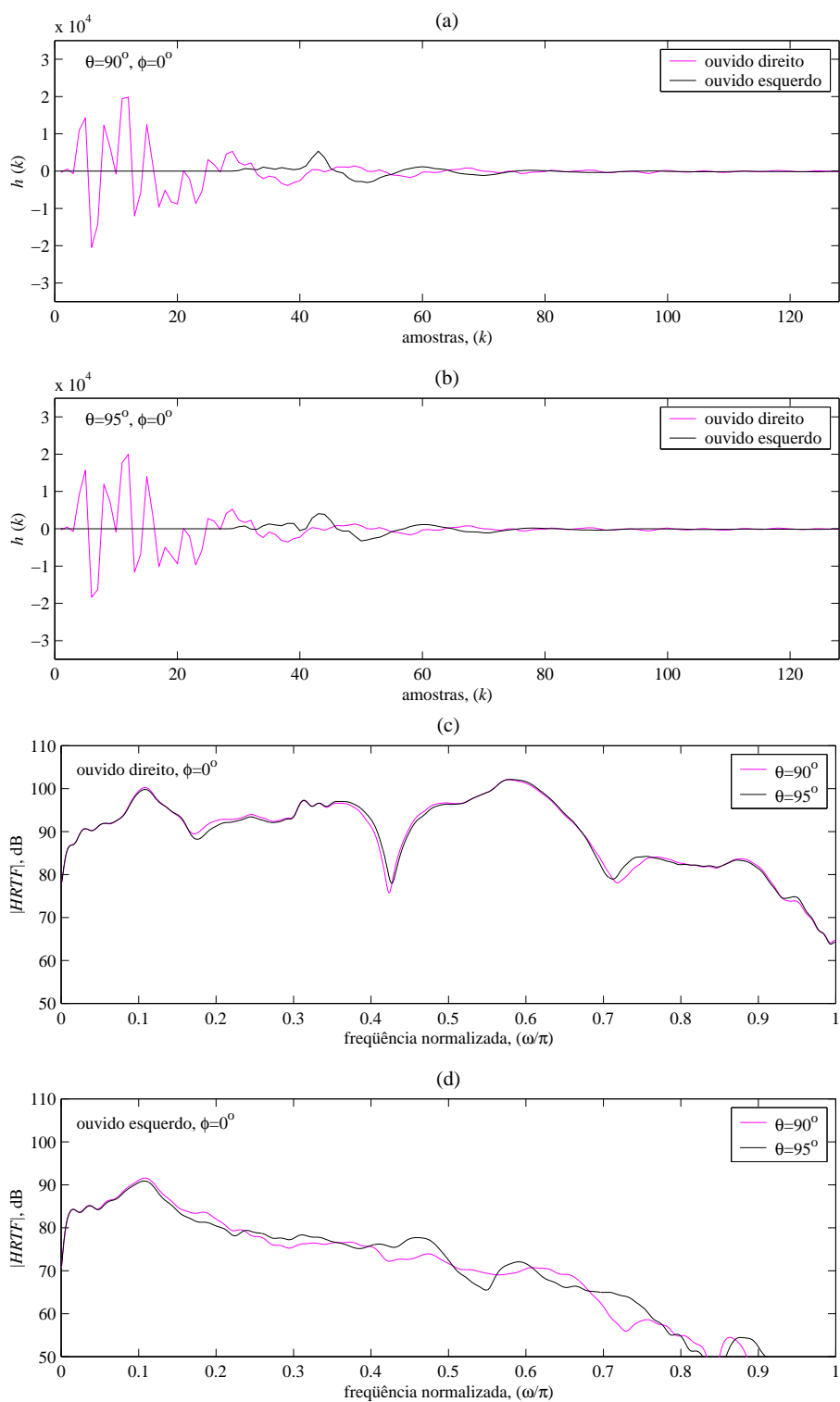


Figura 2.7: Região de importância dos fatores ITD e ILD (II). Comparação das respostas impulsivas (a, b) e dos módulos das respostas em frequência (c, d) das HRTFs com as das Figuras 2.6 e 2.8.

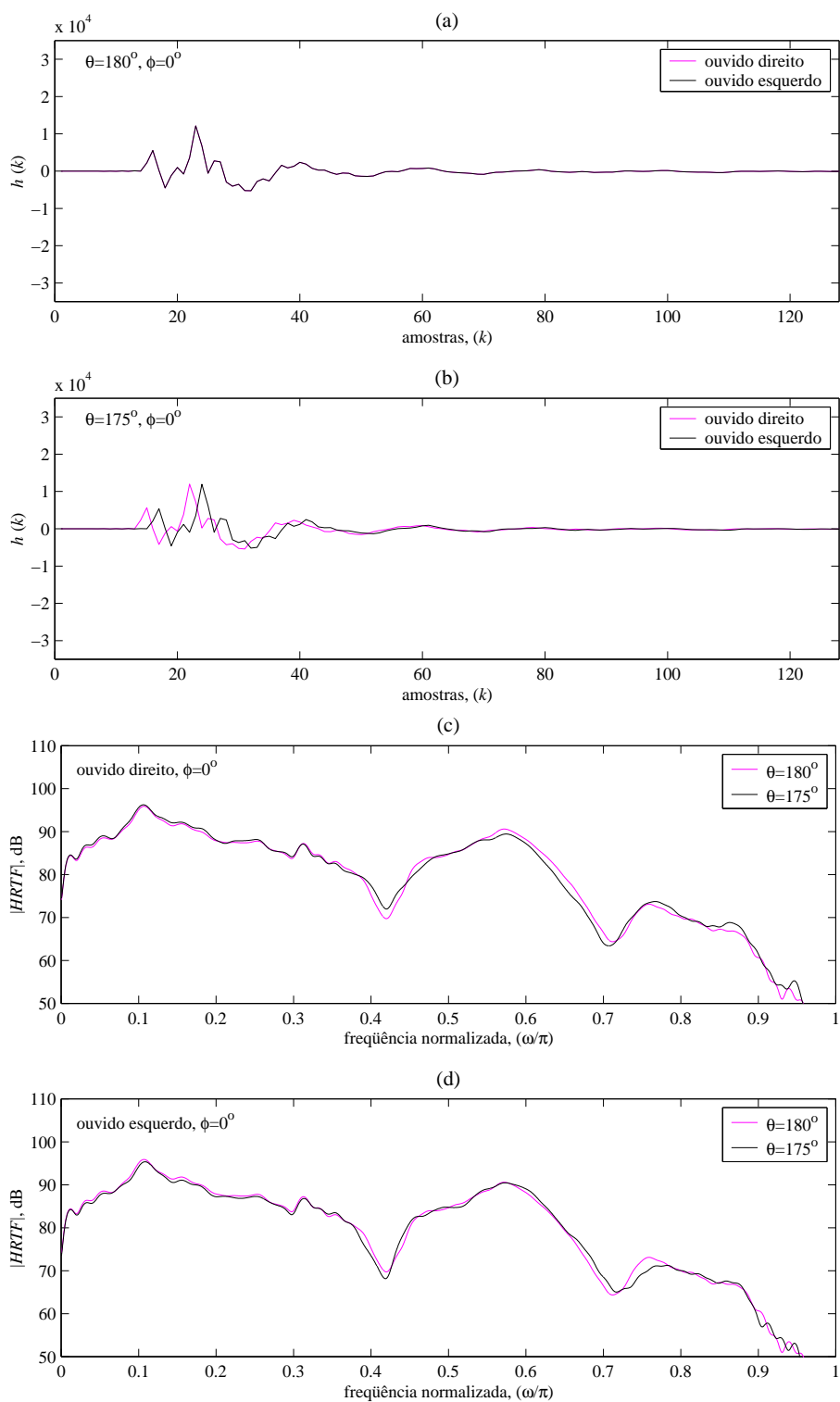


Figura 2.8: Região de importância dos fatores ITD e ILD (III). Comparação das respostas impulsivas (a, b) e dos módulos das respostas em frequência (c, d) das HRTFs com as das Figuras 2.6 e 2.7.

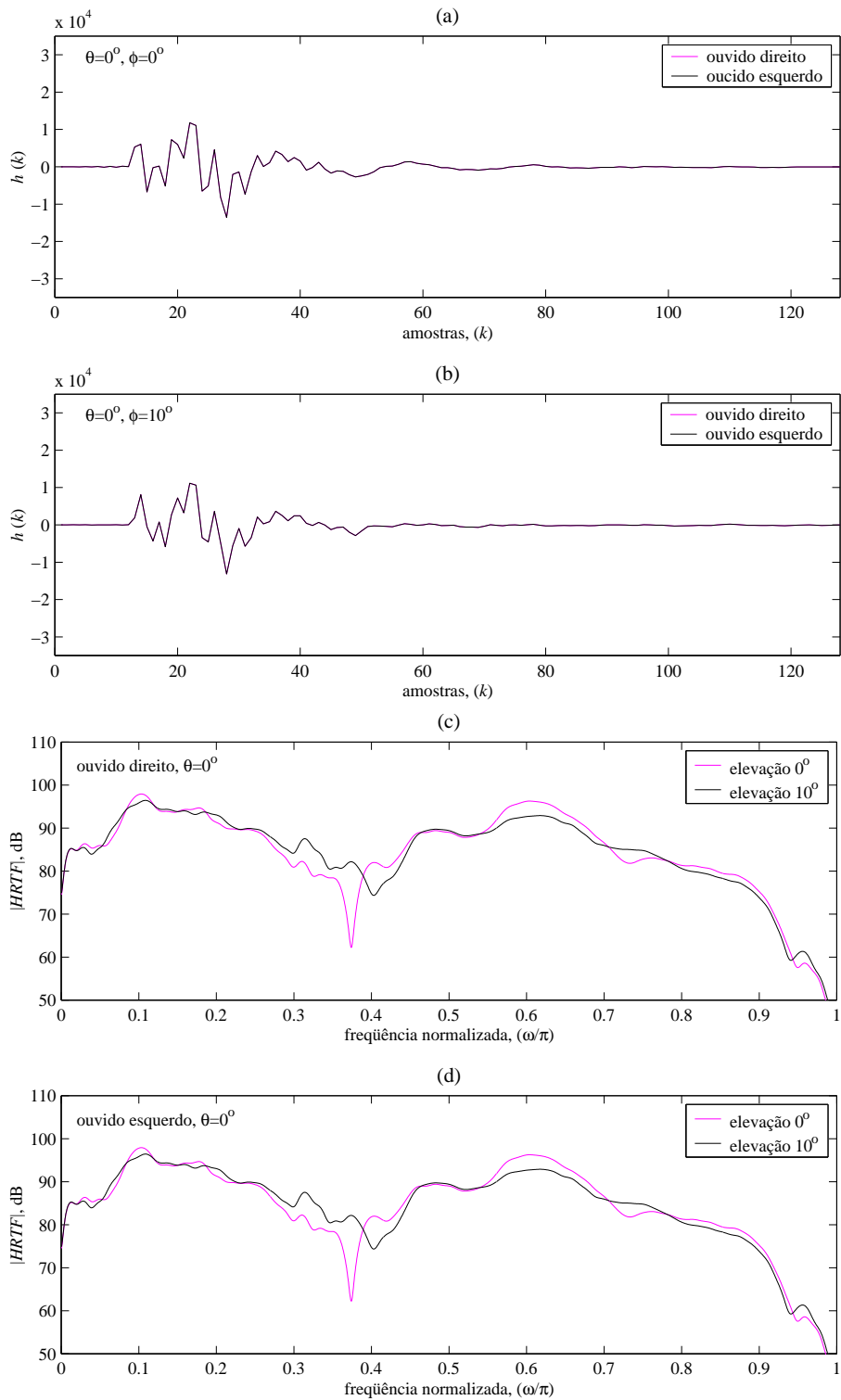


Figura 2.9: Região de importância dos fatores relacionados com as reflexões na pina e no tronco (I). Comparação das HRIRs (a, b) e seus respectivos módulos das respostas em frequência (c, d) com as das Figuras 2.10 e 2.11.

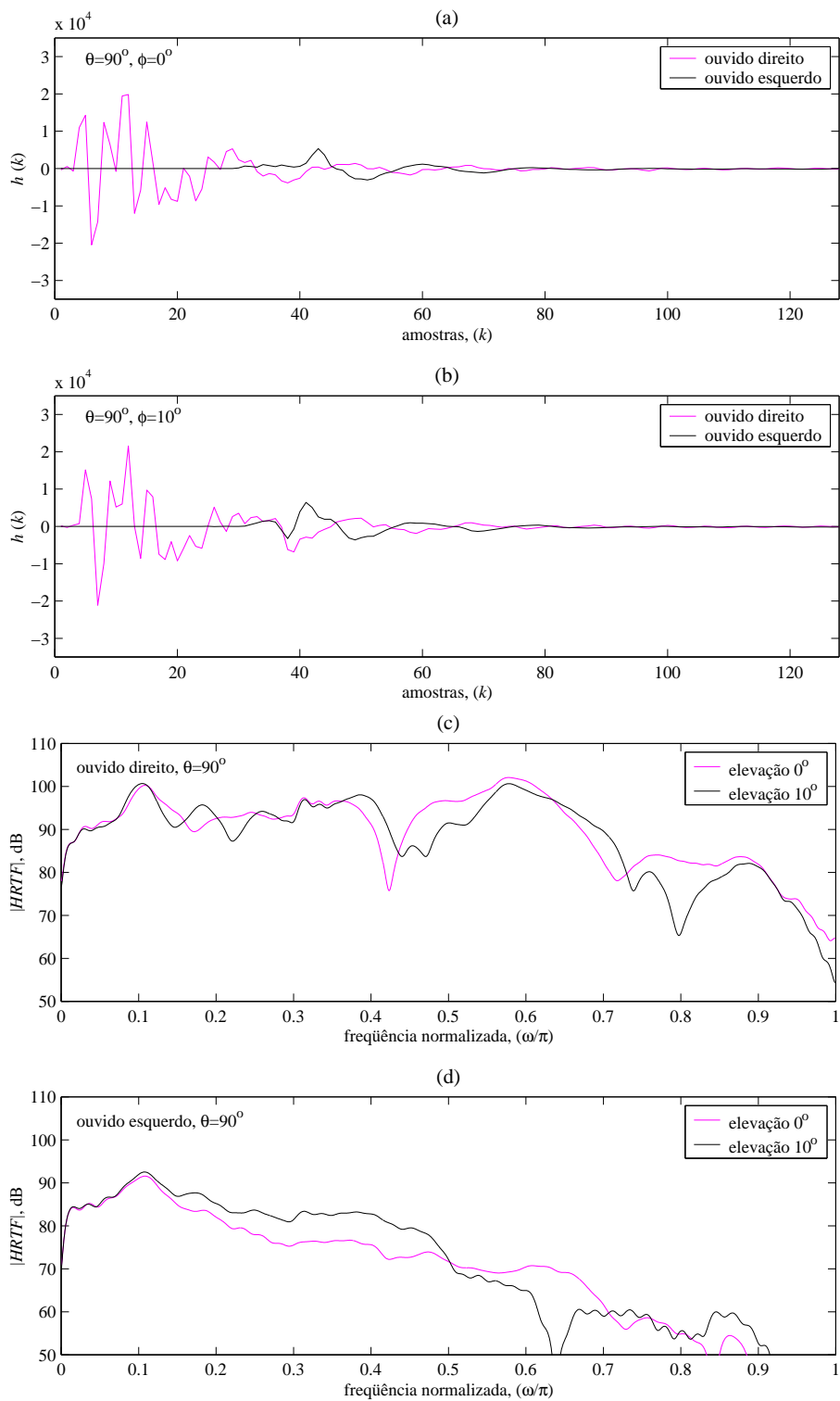


Figura 2.10: Região de importância dos fatores relacionados com as reflexões na pina e no tronco (II). Comparação das HRIRs (a, b) e seus respectivos módulos das respostas em frequência (c, d) com as das Figuras 2.9 e 2.11.

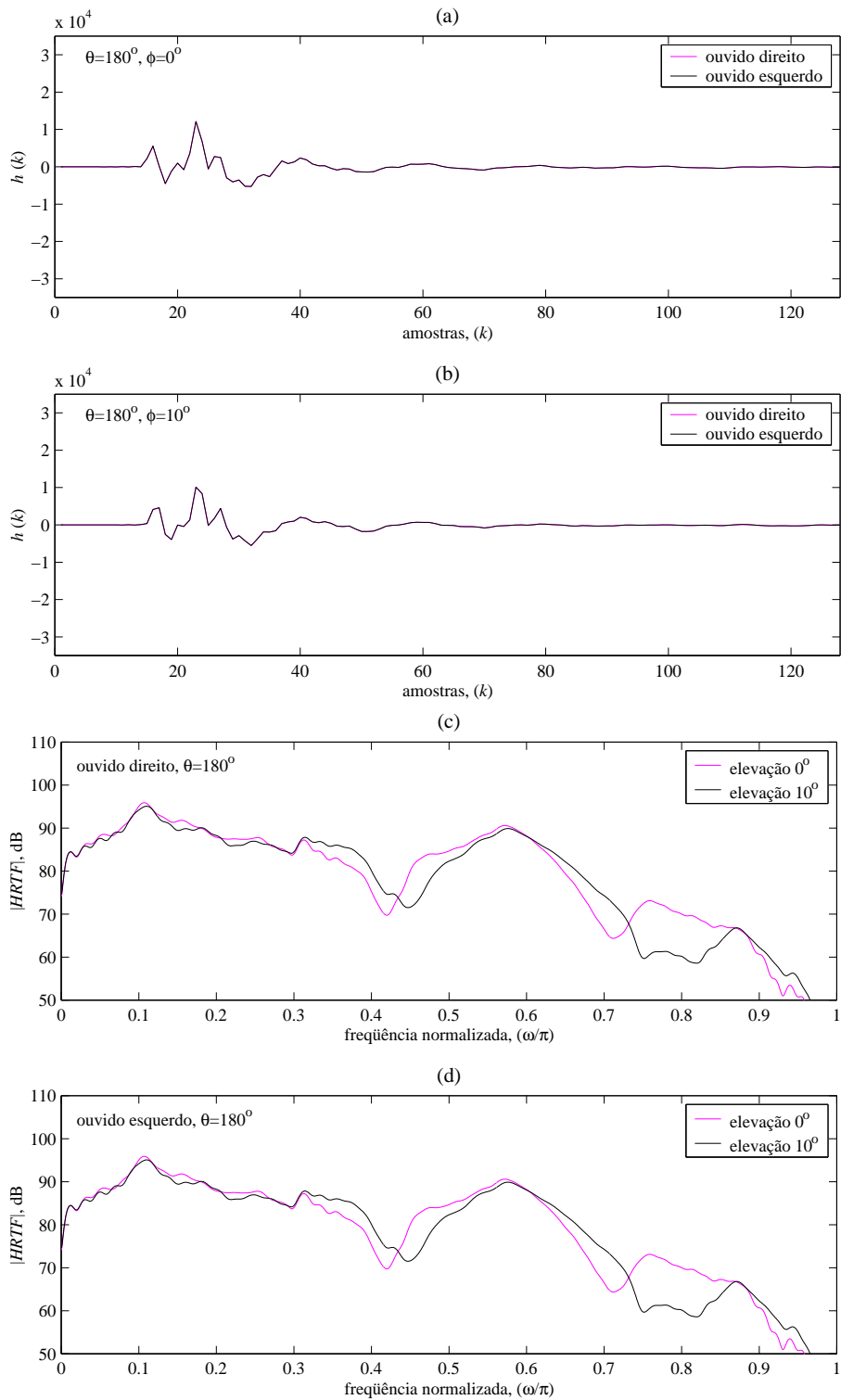


Figura 2.11: Região de importância dos fatores relacionados com as reflexões na pina e no tronco (III). Comparação das HRIRs (a, b) e seus respectivos módulos das respostas em frequência (c, d) com as das Figuras 2.9 e 2.10.

freqüência varia mais para mudanças de elevação próximas do azimute de 90° (Figuras 2.10c e 2.10d) do que para mudanças de elevação próximas dos azimutes de zero grau (Figuras 2.9c e 2.9d) e 180° (Figuras 2.11c e 2.11d). Pode-se notar, também, que os fatores ITD e ILD são pouco sensíveis à variação de elevação (Figuras 2.9a, 2.9b, 2.10a, 2.10b e 2.11a, 2.11b).

Dessa forma, pode-se concluir que nas regiões frontal e traseira podem-se perceber melhor variações em azimute, já que nessas regiões os fatores ITD e ILD mudam muito com pequenas variações de azimute. Por outro lado, nas regiões laterais, tem-se maior sensibilidade às variações de elevação, pois, nessas regiões, as variações na resposta em freqüência (causadas pelas diferentes formas de reflexão na pina) são mais evidentes do que variações na ITD e/ou na ILD.

2.3 Avaliação das HRTFs

A avaliação das HRTFs ou de filtros que as representam normalmente é feita segundo critérios subjetivos [9, 10, 11, 12, 13, 14], mas recentemente tem-se tentado criar métodos objetivos que consigam avaliar corretamente se as respostas impulsivas dão a impressão correta da posição desejada para a fonte.

Nesta tese, não foi utilizado nenhum critério subjetivo formal para a avaliação dos resultados. Apenas testes auditivos informais foram utilizados para se definir se a função em questão está bem representada ou não. Apesar disso, considerou-se interessante tecer alguns comentários sobre esses métodos de avaliação encontrados na literatura.

A maioria dos testes subjetivos encontrados na literatura são feitos reproduzindo-se o sinal de teste com fones de ouvido, posicionando-se a fonte virtual aleatoriamente. As pessoas submetidas ao teste indicam em qual direção está a fonte virtual para cada sinal ouvido e, assim, gera-se uma estatística de quão boas são as HRIRs medidas ou modeladas, no caso de se tratar de HRTFs obtidas de uma outra forma que não por medidas diretas.

Em alguns casos, certos procedimentos são utilizados para melhor avaliar as HRIRs ou seus modelos. Em [12], por exemplo, onde se propõe um realce à resposta da pina, menciona-se o fato de que o treinamento do ouvinte melhora

a percepção da localização da fonte. Esse treinamento consiste em apresentar ao ouvinte, primeiramente, os sinais de teste em todas as direções (no caso, o ângulo de elevação é igual a zero e reproduzem-se os sinais variando-se o azimute para a direita começando do azimute zero) com o prévio conhecimento por parte do ouvinte. Isso faz com que o ouvinte se prepare psicologicamente para ouvir os sinais de teste e, assim, conseguir se transportar melhor para a situação virtual do teste em questão. A influência visual também deve ser eliminada, instruindo-se os ouvintes a fecharem os olhos durante o teste.

Os métodos objetivos têm como base a estimação dos fatores, como por exemplo a ITD e a ILD, para com eles se conseguir avaliar as funções de transferência modeladas ou medidas. Em [14], fez-se uma investigação sobre vários métodos de modelagem das HRTFs comparando-se os resultados com sistemas objetivos de avaliação, obtendo coerência com as avaliações subjetivas. Nesse sistema de avaliação objetiva, desenvolvido por Pulkki *et al.* em [15], estimam-se os fatores ITD e ILD através da correlação cruzada interaural e da estimação do espectro do nível de audibilidade (*loudness*) dos canais direito e esquerdo, respectivamente.

2.4 Utilização das HRTFs para Geração de Som Tridimensional

No contexto de som tridimensional, as HRTFs são utilizadas, através das suas respostas impulsivas (HRIRs), como filtros que acrescentam a informação de localização a sinais originalmente provenientes de um canal. Cada um desses sinais dá origem a outros dois alterados por filtros relativos às respostas para os ouvidos esquerdo e direito, respectivamente. Após a filtragem, esse sinal estéreo é reproduzido com fones de ouvido.

O procedimento acima indica a utilização de um único par de HRIRs que tenha sido medido. Assim, para um ponto cujo par de HRIRs seja conhecido, gera-se um som como se a fonte virtual estivesse parada na posição relacionada àquele par.

Para que se pudesse gerar o som em todas as direções possíveis, dever-se-iam medir todos os pares de HRIRs existentes. Mas como isso é impraticável, adota-

se, na medição, um determinado espaçamento (grade) entre as HRIRs medidas. Assim, para se posicionar a fonte virtual em qualquer posição que não tenha a respectiva HRIR medida, já que existe esse espaçamento entre os pontos em cujas posições foram medidas as HRIRs, vê-se necessária a utilização de alguma forma de interpolação.

O método de interpolação mais simples, mencionado em [5], é o bilinear. Esse método se utiliza dos quatro pares de HRTFs mais próximos do ponto em que se deseja posicionar a fonte virtual para obter uma estimacão do par relativo a tal posição. Isso pode ser descrito da seguinte forma:

$$\hat{h}(k) = (1 - c_\theta)(1 - c_\phi)h_a(k) + c_\theta(1 - c_\phi)h_b(k) + c_\theta c_\phi h_c(k) + (1 - c_\theta)c_\phi h_d(k), \quad (2.1)$$

onde $h_a(k)$, $h_b(k)$, $h_c(k)$ e $h_d(k)$ são os pares de respostas impulsivas mais próximas da posição (θ, ϕ) (veja Figura 2.12) e c_θ e c_ϕ são os coeficientes da interpolação, que são dados por:

$$\begin{cases} c_\theta = \frac{(\theta \bmod \theta_{grid})}{\theta_{grid}} = \frac{C_\theta}{\theta_{grid}} \\ c_\phi = \frac{(\phi \bmod \phi_{grid})}{\phi_{grid}} = \frac{C_\phi}{\phi_{grid}} \end{cases}, \quad (2.2)$$

onde θ_{grid} e ϕ_{grid} são as precisões de azimute e elevação da medição das HRTFs e, ainda, C_θ e C_ϕ dão a posição em graus referente à $h_a(k)$.

Outros métodos de interpolação são utilizados e, com isso, a comparação entre eles se torna necessária. Geralmente, toma-se como base a interpolação bilinear. Em [13], faz-se a comparação entre o método bilinear (nesse artigo chamado de método de ponderação com o inverso da distância—*method of inverse-distance weighting*) e o método de interpolação esférica (*spherical splines*), cada um deles implementado no domínio do tempo e da frequência.

No domínio do tempo, a interpolação bilinear é implementada segundo a equação (2.1), ponderando-se as HRIRs mais próximas. Já no caso da interpolação esférica, todas as HRIRs medidas são consideradas no cálculo da resposta para qualquer ponto para o qual não foi medido um par de HRIRs.

No domínio da frequência, as ponderações são feitas sobre o módulo, em dB, e sobre a fase da resposta em frequência das HRIRs, obtidas pela transformada discreta de Fourier.

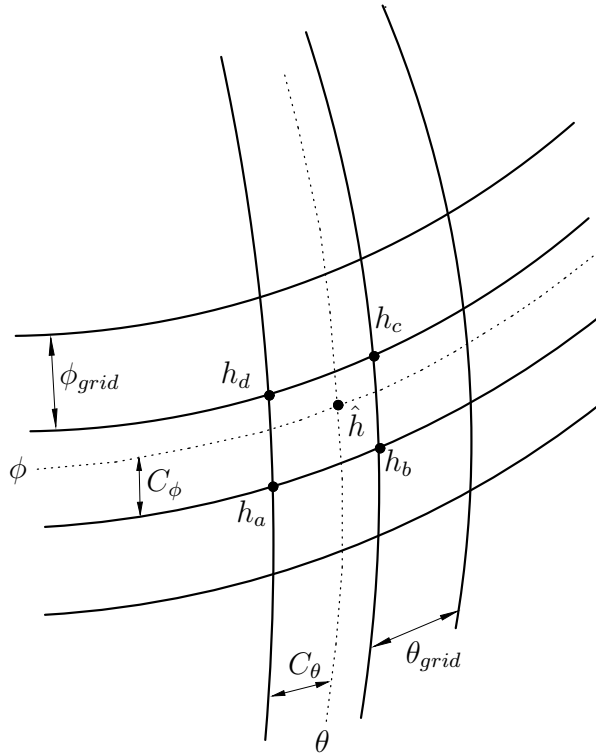


Figura 2.12: Método de interpolação bilinear.

A criação da sensação de movimento utiliza essas formas de interpolação para que, com a evolução do tempo, tenha-se a fonte virtual em pontos diferentes de um determinado caminho.

Muitos trabalhos [7, 9, 10, 14], então, vêm sendo desenvolvidos no intuito de obter a redução da complexidade computacional pela redução da ordem dos filtros envolvidos na geração do som tridimensional. No Capítulo 3, uma nova função de transferência auxiliar será definida e a forma de utilizá-la em uma nova técnica de interpolação será discutida. Já a questão da redução de ordem dos filtros será discutida mais adiante, no Capítulo 4, onde se abordará a complexidade computacional.

Capítulo 3

Funções de Transferência Interposicionais (IPTFs)

Como foi visto no Capítulo 2, para se posicionar a fonte virtual em qualquer elevação e azimute, é necessária a obtenção da HRTF relativa ao ponto desejado através de algum tipo de interpolação. Naquele capítulo, foram também mencionadas algumas formas de interpolação que, utilizando as HRTFs medidas, conseguem aproximar as HRTFs relativas a outros pontos para os quais as medidas não estão disponíveis.

Este capítulo trata da definição e da interpretação de uma nova função de transferência: a função de transferência interposicional (IPTF–*Interpositional Transfer Function*), que é a base para uma nova técnica de interpolação que também será descrita neste capítulo. Essa forma de interpolação é utilizada, no Capítulo 5, no algoritmo que dá movimento ao som.

Neste capítulo será mostrada a idéia de onde se partiu para chegar até a IPTF na Seção 3.1–**Idéia Principal**. Em seguida, na Seção 3.2–**Funções de Transferência Interposicionais**, apresenta-se o que é a IPTF e como ela é gerada a partir das HRTFs. A forma de interpolação que utiliza as IPTFs está descrita na Seção 3.3–**Interpolação Utilizando as IPTFs** e na Seção 3.4–**Comentários**, tecem-se algumas observações sobre a complexidade computacional dessa nova técnica de interpolação.

3.1 Idéia Principal

A idéia inicial para criação dessa nova função de transferência surgiu ao se encontrar em [9] uma proposta de redução dos modelos das HRTFs.

A proposição daquele artigo objetivava reduzir a dificuldade encontrada no modelamento das funções de transferência relativas ao ouvido localizado no lado oposto ao da fonte sonora (chamadas de funções contralaterais—veja na Figura 3.1). Essa dificuldade está no fato de o efeito da “sombra da cabeça” atenuar muito as frequências mais altas e, com isso, tornar mais difícil a identificação dos modos dessas frequências, podendo criar a necessidade de modelos de ordem mais elevada.

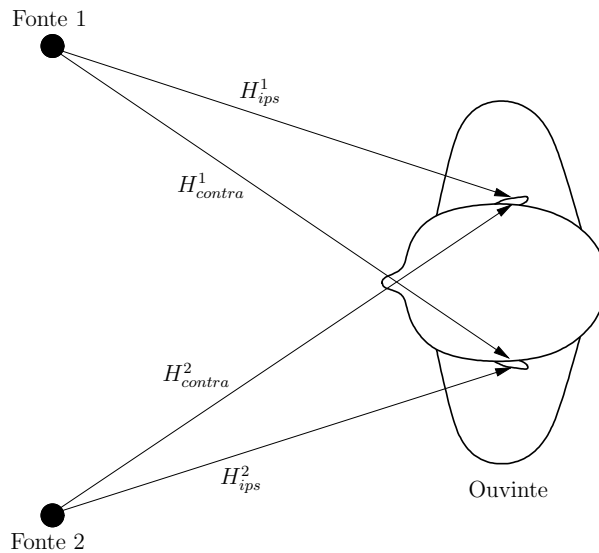


Figura 3.1: Função de transferência interaural (ITFs). H^m_{ips} e H^m_{contra} são as funções ipsi- e contralateral relativas à m -ésima fonte.

Então, tendo em vista que as funções relativas ao ouvido do mesmo lado em que se encontra a fonte sonora (chamadas de funções ipsilaterais) não sofrem este tipo de atenuação, foi sugerida, naquele trabalho, uma pequena modificação na estrutura de implementação, que pode ser vista na Figura 3.2. Essa modificação leva a um diagrama equivalente ao anterior, só que os blocos que representam as funções de transferência a serem modeladas passam a ser a função ipsilateral e a razão entre as funções contra- e ipsilateral. Essa função-razão é denominada por aqueles autores de Função de Transferência Interaural (ITF—*Interaural Transfer Function*), pelo fato de representar a transferência entre o ouvido do lado oposto e o do mesmo lado da

fonte.

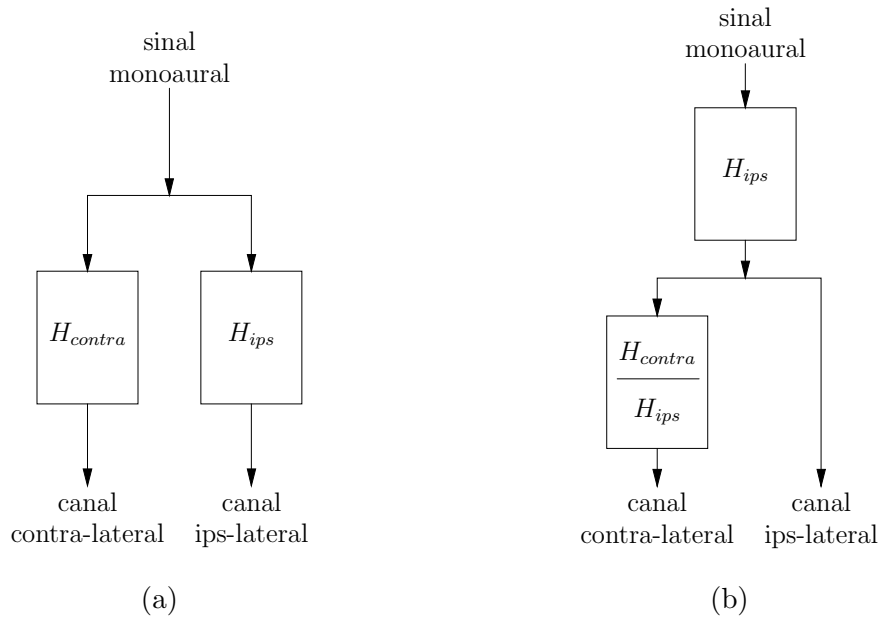


Figura 3.2: Modelo das HRTFs através das ITFs. H_{ips} e H_{contra} são as funções ipsi- e contralateral: (a) estrutura que emprega as funções ipsi- e contralateral separadamente e (b) estrutura que emprega a função ipsilateral e a ITF.

A questão da dinâmica das respostas em frequência também pode ser importante sob o ponto de vista de simplificação do modelo. Na Figura 3.3a, tem-se a resposta em frequência da ITF relativa às HRTFs vistas na Figura 3.3b. Nota-se que a resposta em frequência da ITF tende a ser mais plana que as respostas das HRTFs que a originaram.

Essa idéia de redução do modelo através da razão entre duas funções de transferência foi utilizada no presente trabalho na criação da IPTF, que tem como objetivo fazer com que a interpolação seja mais simples, como será melhor explicado na Seção 3.2.

3.2 Funções de Transferência Interposicionais

A idéia mostrada na Seção 3.1 atinge maior simplicidade com o modelo criado a partir da ITF (razão de duas HRTFs). Apesar disso, percebe-se que as ITFs tendem a ficar mais difíceis de serem modeladas próximo ao azimute de 90° , porque as diferenças entre as funções da razão se tornam maiores. Esse problema, no

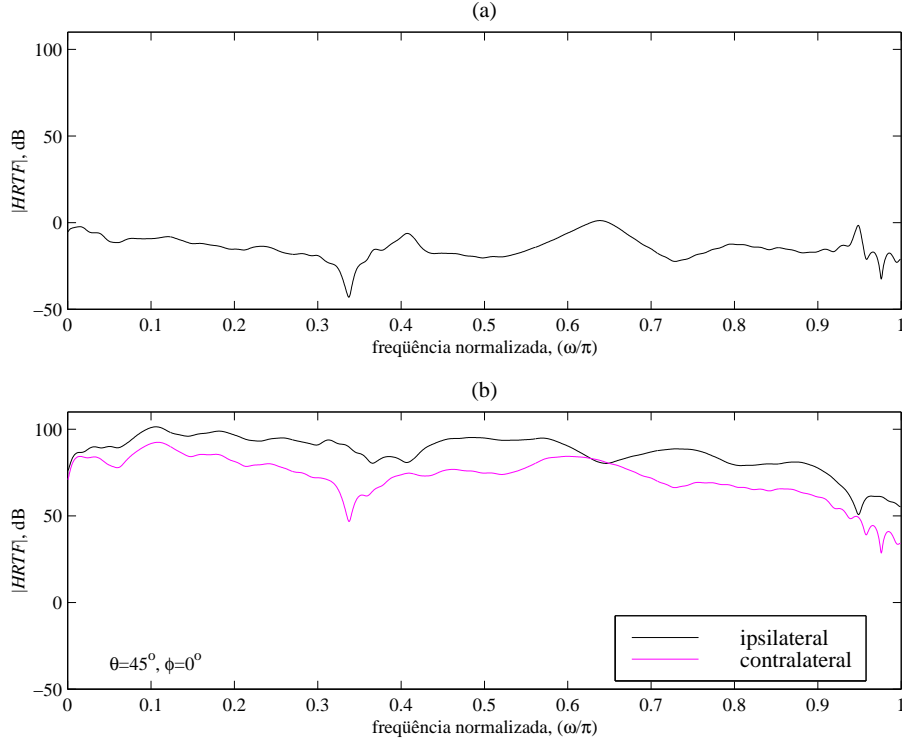


Figura 3.3: Comparação das dinâmicas das funções ipsi- e contralateral com a ITF. Módulo da resposta em frequência da ITF (a) e das funções ipsi- e contralateral (b) que a originaram.

entanto, não parece fácil de ser eliminado representando-se a transferência relativa a um único ponto sem considerar a sua dependência das transferências relativas a pontos vizinhos, como é com a ITF.

Como o principal objetivo deste trabalho é reduzir a complexidade da interpolação, pensou-se em formar uma função-razão que fosse gerada a partir de duas funções o mais possível parecidas; especificamente, utilizar duas HRTFs relativas a um mesmo lado da cabeça e a pontos próximos entre si (veja a Figura 3.4). Nessa figura, pode-se notar que a $IPTF_{i,f}$ é a função de transferência que representa a passagem de um ponto i a outro ponto adjacente f , e pode ser descrita da seguinte forma:

$$IPTF_{i,f} = \frac{HRTF_f}{HRTF_i}, \quad (3.1)$$

onde $HRTF_i$ e $HRTF_f$ são as HRTFs relativas aos pontos inicial e final, respectivamente.

Algumas considerações importantes devem ser feitas para que a Equação (3.1) seja válida. Por exemplo, a função $HRTF_i$ deve ser de fase mínima e os atrasos

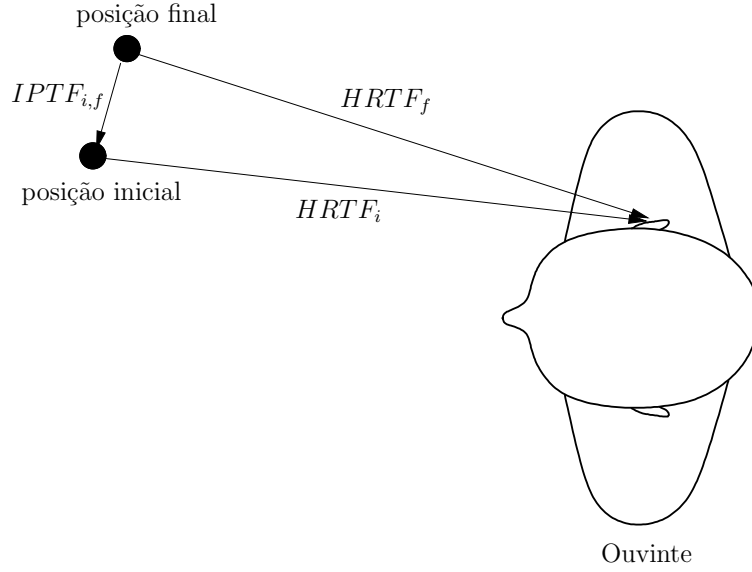


Figura 3.4: Função de Transferência Interposicional.

inerentes à função $HRTF_f$ devem ser iguais aos da IPTF gerada somados aos da $HRTF_i$, para que a ITD (ver Seção 2.2.1) não seja alterada. Outro fato bastante importante é que as IPTFs pressupõem uma função para cada ouvido, isto é, cada ouvido tem uma IPTF para cada par de pontos inicial e final. Todas essas considerações e detalhes de criação das IPTFs serão discutidos na Subseção 3.2.1.

3.2.1 Geração das IPTFs

Como foi visto na primeira parte dessa seção, as IPTFs são funções cuja resposta em frequência é o resultado da razão de duas HRTFs adjacentes. Para cada dois pares de HRTFs, relativas aos dois ouvidos e a duas posições adjacentes, existe um par de IPTFs, cada uma relativa a um dos ouvidos. Isso significa que a Equação (3.1) pode ser reescrita da seguinte forma:

$$IPTF_{i,f}^d = \frac{HRTF_f^d}{HRTF_i^d} \quad (3.2)$$

ou

$$IPTF_{i,f}^e = \frac{HRTF_f^e}{HRTF_i^e}, \quad (3.3)$$

onde $IPTF_{i,f}^d$ e $IPTF_{i,f}^e$ são as IPTFs relativas aos ouvidos direito e esquerdo, respectivamente; $HRTF_i^d$ e $HRTF_f^d$ são as HRTFs para o ouvido direito relativas aos pontos inicial e final, respectivamente; e $HRTF_i^e$ e $HRTF_f^e$ são as HRTFs para o ouvido esquerdo relativas aos pontos inicial e final, respectivamente.

Há, no entanto, um problema que deve ser solucionado antes de se implementarem tais razões: a estabilidade da IPTF deve ser garantida. Em [7, 9, 14], utilizam-se aproximações de fase mínima para facilitar o processo de redução de ordem do modelo para as HRTFs. A utilização de tal recurso se baseia em trabalhos [16, 17, 18] que mostram a insensibilidade à fase das representações se o atraso inicial relativo à ITD for modelado por uma linha de atrasos z^{-D} , onde D é o número inteiro de atrasos. E, mesmo que no domínio do tempo esse recurso mude muito a resposta impulsiva (veja Figura 3.5), não se percebe diferença auditiva entre os sinais modificados pelas funções sem ou com fase mínima ao se gerar o som tridimensional.

Então, sabendo-se que se pode utilizar a versão de fase mínima das HRTFs (tomando apenas o cuidado com os atrasos), antes de se implementar a razão, calculou-se para cada HRTF sua versão de fase mínima. Para calcular as versões de fase mínima, para cada uma das HRTFs estimou-se o número de atrasos existentes por inspeção visual, examinando-se a forma de onda de cada HRTF (pelo número de amostras com valores bem baixos de amplitude que precediam a subida até o primeiro pico). Assim, foram eliminados da i -ésima HRTF as primeiras D_i amostras e, com o restante das 128 amostras, calculou-se a versão de fase mínima. Um exemplo da estimação do número de atrasos de uma HRTF e da função de fase mínima obtida com os atrasos retirados já reincorporados podem ser vistos nas Figuras 3.5a e 3.5b, respectivamente.

Outro ponto importante na implementação da IPTF é o fato de elas poderem ser não-causais. A não-causalidade ocorre naturalmente, pois, a não ser que os números de atrasos das HRTFs utilizadas na razão sejam iguais, o número de atrasos relativo entre essas HRTFs pode ser negativo ($HRTF_f$ ter menos atrasos que a $HRTF_i$). Isso fica mais claro se a Equação (3.1) for reescrita na forma da sua transformada \mathcal{Z} :

$$IPTF_{i,f}(z) = \frac{HRTF_f(z)}{HRTF_i(z)} = \frac{b_0 + b_1z^{-1} + b_2z^{-2} + \dots + b_Nz^{-N}}{a_0 + a_1z^{-1} + a_2z^{-2} + \dots + a_Nz^{-N}}, \quad (3.4)$$

onde b_n e a_n são os valores das HRIRs em cada instante. Nessa equação, pode-se notar que, se os coeficientes da IPTF de b_0 a b_k e de a_0 a a_l forem iguais a zero (atrasos de cada uma das HRTFs), com $l > k$, a função de transferência interposicional irá necessitar de amostras anteriores a partir da amostra $n = k - l + 1$ da entrada,

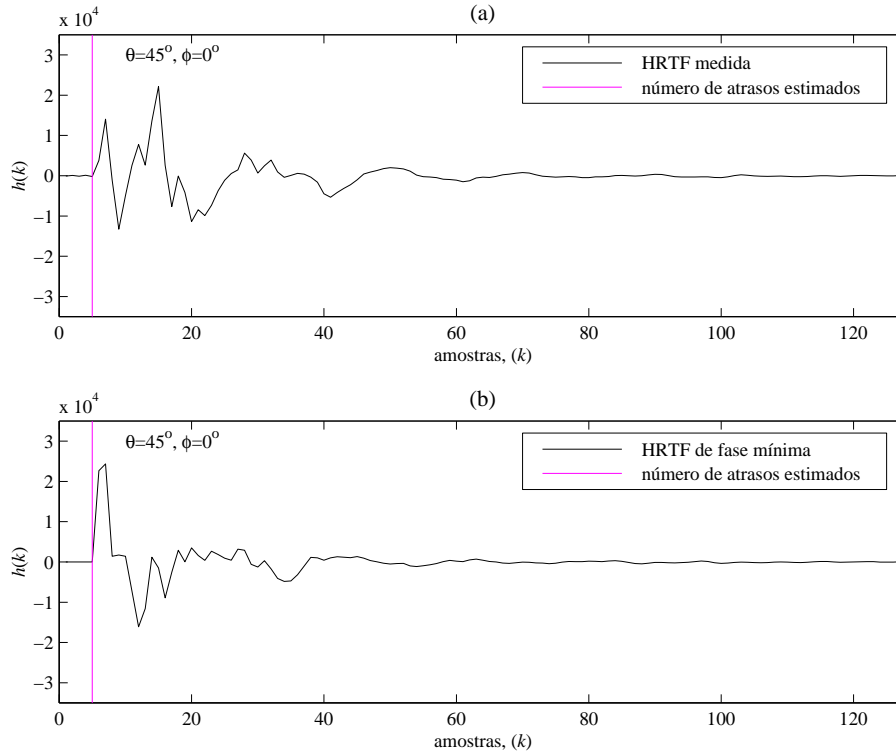


Figura 3.5: Aproximações de fase mínima. Estimação do número de atrasos por inspeção (a) e HRIR de fase mínima calculada.

tornando-se não-causal. Como a IPTF pressupõe que a $HRTF_i$ seja utilizada em cascata com ela (veja na Seção 3.3), esse problema da não-causalidade pode ser solucionado acrescentando-se às IPTFs o atraso relativo à HRTF. Isso torna as IPTFs sempre causais e como o “avanço” (nos casos não-causais) relativo à IPTF é sempre menor do que o atraso inerente à HRTF cascadeada, o sistema resultante se torna causal.

3.3 Interpolação Utilizando as IPTFs

No Capítulo 1, onde se discutiram as várias formas de criação do som tridimensional, mencionou-se o método VBAP [1] que, utilizando um sistema multicanal, faz a localização da fonte virtual depender da direção de um vetor resultante da soma ponderada de cada um dos vetores associados a cada caixa acústica—veja no Capítulo 1 a Figura 1.1. Redesenhando-se aquela figura associando a cada vetor ouvinte-caixa uma HRTF caixa-ouvinte para cada ouvido, ter-se-ia uma ponderação das HRTFs de cada posição formando a transferência global da fonte virtual para o

ouvinte (veja a Figura 3.6).

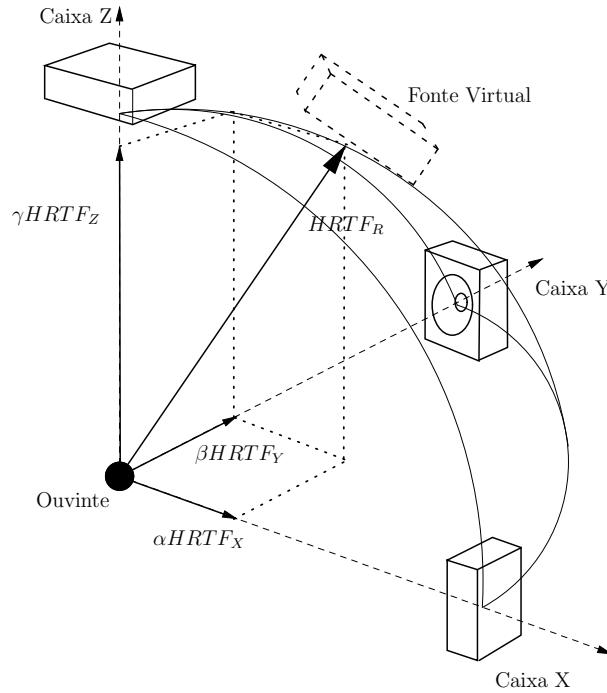


Figura 3.6: VBAP[1] interpretado com as HRTFs. HRTFs X, Y e Z ponderadas, respectivamente, pelos parâmetros α , β e γ , que, compostas, geram a $HRTF_R$ (resultante).

Essa forma de interpretação do sistema VBAP utilizando as HRTFs dá uma idéia da validade do método de interpolação bilinear [5], que gera a transferência de um determinado ponto a partir das HRTFs de pontos adjacentes.

Pode-se interpretar de uma forma similar a interpolação utilizando as IPTFs. Primeiramente, considere a $HRTF_I$, relativa a um determinado ponto I , e as HRTFs interpoladas com as IPTFs para dois de seus pontos vizinhos 1 e 2 (veja Figura 3.7). Essas funções de transferência podem ser combinadas, como na interpolação bilinear, para gerar a transferência de um determinado ponto F , $HRTF_F$, pertencente à região triangular compreendida pelos três pontos, considerando que o ponto mais próximo de F seja o I . Para isso, faz-se com que os ponderadores para cada uma das transferências dependam do inverso da distância angular do ponto F aos pontos a elas relacionados.

Assim, como pode ser visto na Figura 3.8, utilizou-se essa forma de interpolação e, para um dado um ponto F que se quer interpolar:

1. descobre-se qual o ponto I mais próximo para o qual a HRTF foi medida;

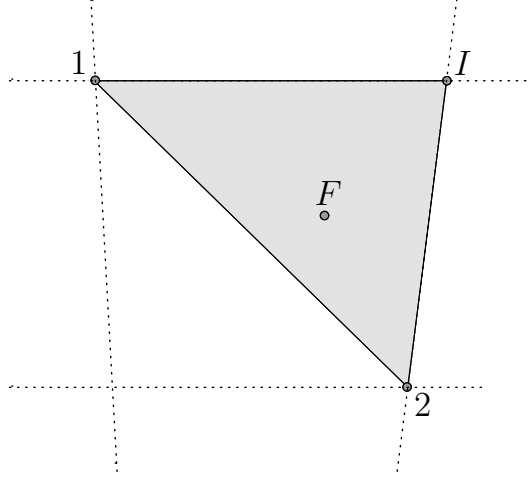


Figura 3.7: Limitação da região de interpolação para o conjunto de funções escolhido.

2. caso dois ou mais pontos sejam equidistantes do ponto F , o ponto escolhido como ponto I é o de menores azimute e elevação;
3. determinam-se quais duas IPTFs, das quatro que têm como $HRTF_i$ a HRTF do ponto I , compreendem a região que contém o ponto F ;
4. e gera-se a transferência do ponto F ao ouvinte através da equação:

$$HRTF_F = \frac{HRTF_I(\alpha + \beta IPTF_{I,1} + \gamma IPTF_{I,2})}{\alpha + \beta + \gamma}, \quad (3.5)$$

onde α , β e γ são dados por

$$\alpha = (1 - c_{\theta_I})(1 - c_{\phi}), \quad (3.6)$$

$$\beta = c_{\theta_I}(1 - c_{\phi}) \quad (3.7)$$

e

$$\gamma = (1 - c_{\theta_2})c_{\phi}, \quad (3.8)$$

sendo c_{θ_I} , c_{ϕ} e c_{θ_2} , respectivamente, as distâncias angulares relativas de azimute do ponto I ao ponto F , de elevação do ponto I ao ponto F e de azimute do ponto 2 ao ponto F , que podem ser descritas por

$$c_{\theta_I} = \frac{|\theta_F - \theta_I|}{\theta_I^{grid}} = \frac{C_{\theta_I}}{\theta_I^{grid}}, \quad (3.9)$$

$$c_{\theta_2} = \frac{|\theta_F - \theta_2|}{\theta_2^{grid}} = \frac{C_{\theta_2}}{\theta_2^{grid}} \quad (3.10)$$

e

$$c_\phi = \frac{|\phi_F - \phi_I|}{\phi^{grid}} = \frac{C_{\phi_I}}{\phi^{grid}}. \quad (3.11)$$

Nessas equações, θ_I^{grid} , θ_2^{grid} e ϕ^{grid} são os passos da grade no azimute (θ) e na elevação (ϕ), sendo que os passos θ_I^{grid} e θ_2^{grid} são os utilizados na elevação dos pontos I e 2 , respectivamente; os pares de coordenadas (θ_F, ϕ_F) e (θ_I, ϕ_I) são relativos aos pontos F e I , respectivamente; θ_2 é a coordenada de azimute do ponto 2 ; e C_{θ_I} , C_{θ_2} e C_ϕ são as distâncias angulares absolutas, que podem ser vistas na Figura 3.8.

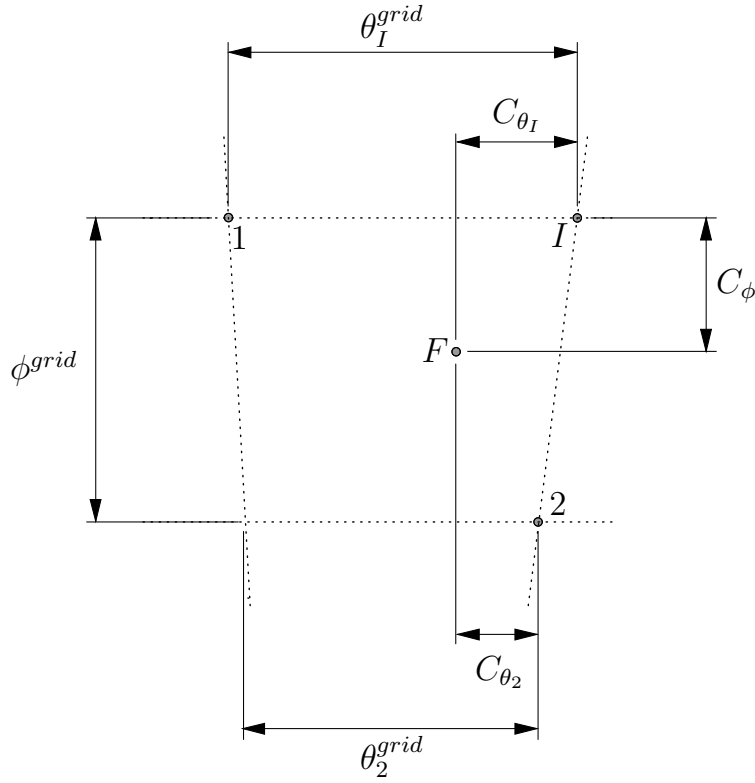


Figura 3.8: Distâncias angulares absolutas utilizadas no cálculo dos parâmetros α , β e γ .

Examinando-se todas essas equações para cálculo dos ponderadores, tem-se uma impressão de que essa forma de interpolação é bem mais custosa, já que são calculados três ponderadores e, além disso, três distâncias angulares relativas. Deve-se notar, no entanto, que nas Equações (2.1) e (2.2), que descrevem a interpolação bilinear, está-se considerando que a grade de funções de transferência medidas tem sempre o mesmo número de funções por elevação, o que transforma as equações de cálculo das distâncias angulares relativas em apenas duas equações. Note que a

operação “mod” foi substituída por uma subtração, pois não se pode garantir que o azimute do ponto 2 é menor que o do ponto F , já que a grade não é quadrada; isso resultaria num valor errado para a distância relativa de azimute, se “mod” fosse utilizada.

Como as funções utilizadas nesse trabalho não foram medidas utilizando-se uma grade quadrada, tanto no caso da interpolação bilinear quanto no caso da que utiliza a IPTF, o número de multiplicações/divisões para gerar os ponderadores seriam iguais a:

$$N_b = 3 + 4 = 7 \quad (3.12)$$

e

$$N_{IPTF} = 3 + 3 = 6. \quad (3.13)$$

No primeiro caso, são 3 operações para calcular as distâncias angulares relativas mais as 4 operações para calcular os 4 ponderadores. No segundo, são 3 operações efetuadas para as distâncias angulares relativas mais 3 para o cálculo dos ponderadores, deixando-se a divisão para normalizar a saída para ser feita no momento do cálculo da saída. Isso mostra que o número de multiplicações/divisões no cálculo dos ponderadores é praticamente o mesmo. A interpolação com as IPTFs tem apenas uma operação a menos que a forma bilinear.

Na Figura 3.9, tem-se o diagrama de blocos do método de interpolação proposto. Nessa figura, o processo de interpolação é mostrado de forma completa, isto é, para os dois canais. Os atrasos D_i^d e D_i^e são relativos a $HRTF_i^d$ e $HRTF_i^e$, respectivamente, enquanto os atrasos $d_{i,1}^d$, $d_{i,2}^d$, $d_{i,1}^e$ e $d_{i,2}^e$ são relativos a $IPTF_{i,1}^d$, $IPTF_{i,2}^d$, $IPTF_{i,1}^e$ e $IPTF_{i,2}^e$, respectivamente. Note que, se qualquer uma das IPTFs for não-causal, isto é, algum dos atrasos $d_{i,1}^d$, $d_{i,2}^d$, $d_{i,1}^e$ e $d_{i,2}^e$ for negativo, ainda assim o sistema permanecerá causal, já que

$$|D_i^d| \geq |d_{i,1}^d|, \quad (3.14)$$

$$|D_i^d| \geq |d_{i,2}^d|, \quad (3.15)$$

$$|D_i^e| \geq |d_{i,1}^e|, \quad (3.16)$$

e

$$|D_i^e| \geq |d_{i,2}^e|. \quad (3.17)$$

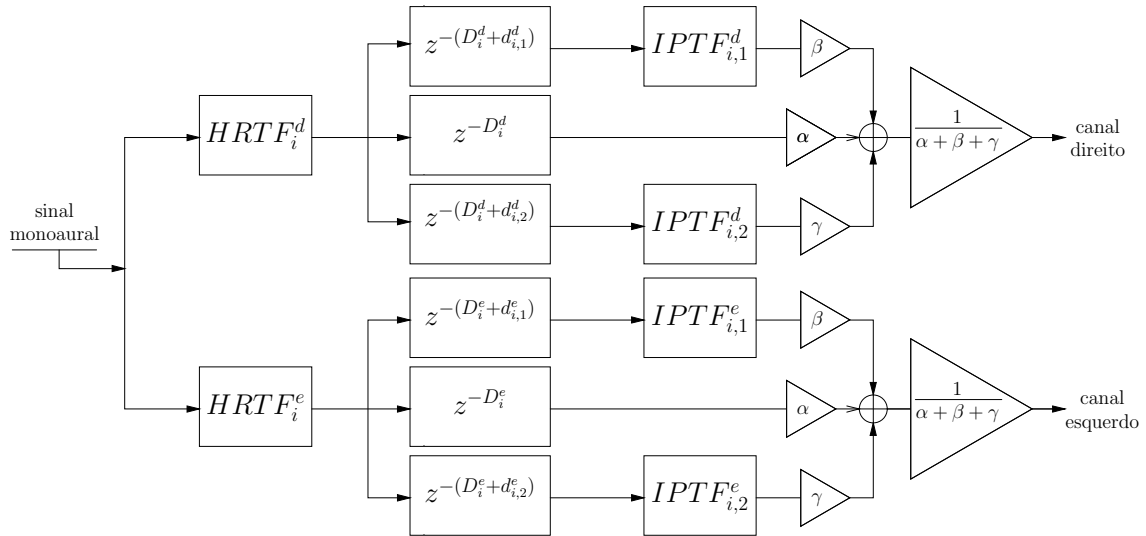


Figura 3.9: Diagrama em blocos da interpolação utilizando as IPTFs.

3.3.1 Avaliação da Interpolação com as IPTFs

A avaliação da forma de interpolação que utiliza as IPTFs foi feita através de testes auditivos informais. Além da comparação visual do módulo das respostas em frequência das funções interpoladas e as originais, os testes auditivos consistiram da simples comparação entre os sinais modificados pelas HRTFs originais e os sinais modificados pelo sistema da Figura 3.9 com α , β e γ tais que fosse possível gerar a HRTF interpolada para o ponto relativo à HRTF original utilizada no teste.

Como se está querendo achar a interpolação para um ponto cuja HRTF foi medida, é necessário o cálculo das IPTFs cuja distância angular seja o dobro do passo da grade em cada sentido (pontos 1' e 2'—veja a Figura 3.10). Nessa figura, é possível notar tal necessidade, já que, com a utilização de uma HRTF medida e duas interpoladas para os pontos 1 e 2, só se podem gerar interpolações no interior do triângulo escuro. Essa limitação espacial é similar à encontrada no sistema VBAP, mencionado no início desse capítulo: apenas com variação de intensidade não se podem gerar fontes virtuais fora da região limitada pelas retas que unem as caixas acústicas.

Nas Figuras 3.11 e 3.12, podem-se ver exemplos da comparação visual efetuada entre as respostas em frequência da $HRTF_F$ e das HRTFs obtidas através das interpolações bilinear e com as IPTFs. Pode-se notar nas Figuras 3.11a e 3.11b que a resposta interpolada com as IPTFs acompanha melhor a resposta original do que

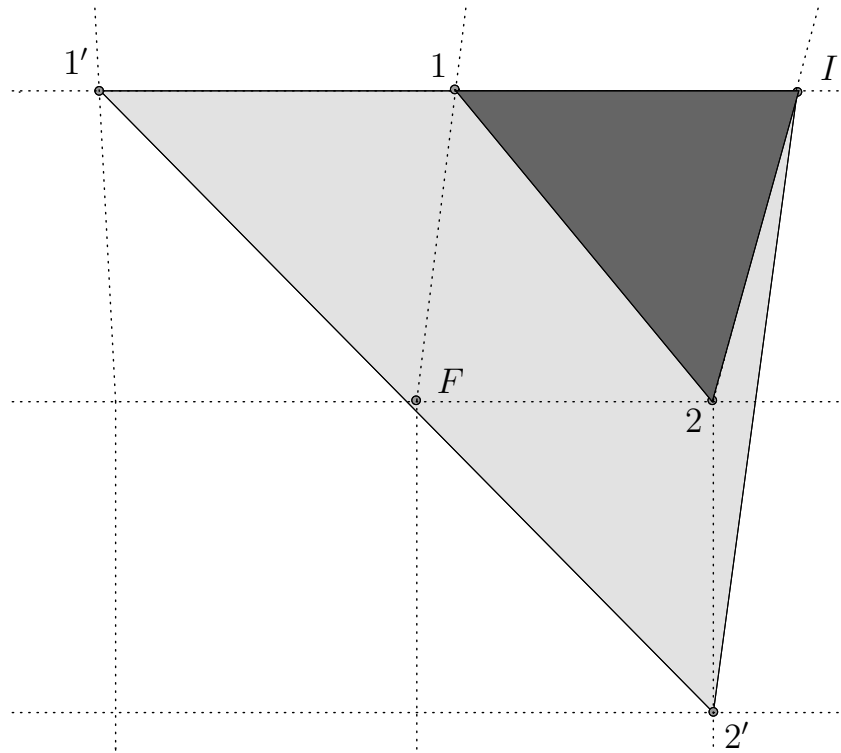


Figura 3.10: Análise gráfica do teste feito para a interpolação.

a resposta resultante da interpolação bilinear, apesar de essas respostas não serem perfeitas, como pode ser observado na resposta da Figura 3.12b. Observando-se grande quantidade de respostas, pode-se verificar que na maioria dos casos a interpolação que utiliza as IPTFs tem resposta mais parecida com a original. É bem verdade que em alguns casos, como o da Figura 3.12a, são observados vales indesejáveis na interpolação feita com as IPTFs; contudo estes também são observados na interpolação bilinear.

Comparando-se os resultados auditivos referentes a uma mesma posição para os dois tipos de interpolação, percebe-se por vezes uma clara mudança de timbre que, em última análise, pode interferir na sensação de localização. A própria forma de compor as funções envolvidas pode resultar num efeito suavizador. Como a interpolação com as IPTFs se mostrou, em muitos casos, melhor que a bilinear, mesmo com essa variação do timbre, os resultados obtidos foram considerados satisfatórios. Deve ser notado que estes testes foram realizados com uma grade de funções mais larga e, ainda, para pontos onde ocorrem os piores casos (pontos mais afastados).

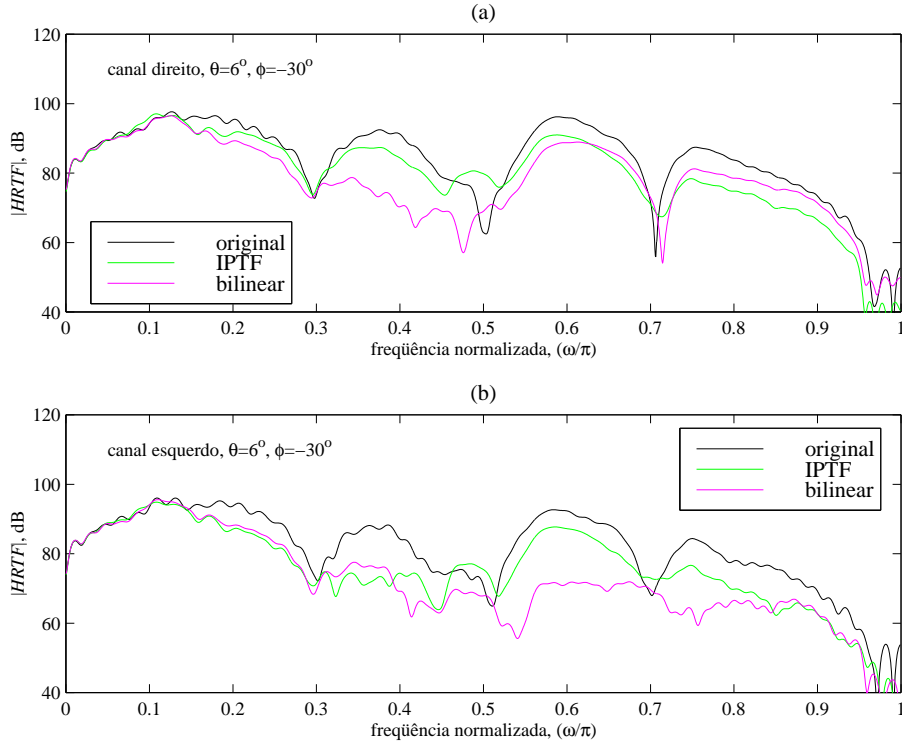


Figura 3.11: Exemplo do teste da interpolação (I). Módulo das respostas em frequência das HRTFs original e interpoladas: (a) canal direito e (b) canal esquerdo.

3.4 Comentários

O intuito desse tipo de interpolação é ter um custo computacional menor do que os que vêm sendo utilizados [13]. A complexidade depende do número de coeficientes dos filtros utilizados por cada um dos sistemas.

Em [14], os autores mostram que o número de coeficientes necessários para um filtro IIR conseguir aproximar bem uma HRTF é 33 (ordem 16). Isso implica que, se fossem utilizadas essas HRTFs numa interpolação bilinear, o número de multiplicações envolvidas no cálculo de uma amostra de saída por canal seria 143 (4 HRTFs mais 4 ponderadores mais as 7 necessárias para calcular os ponderadores—Equação 3.12). No caso da interpolação esférica (*spherical splines*) o número de multiplicações cresce demais, já que todas as HRTFs são utilizadas no cálculo.

Então, é possível estimar a maior ordem que a interpolação feita pelo sistema da Figura 3.9 para que seja vantajosa a sua utilização. O custo computacional associado à interpolação com as IPTFs pode ser calculado somando-se os coeficientes da $HRTF_i$ (33 se utilizado o modelo mostrado em [14]) com os coeficientes da

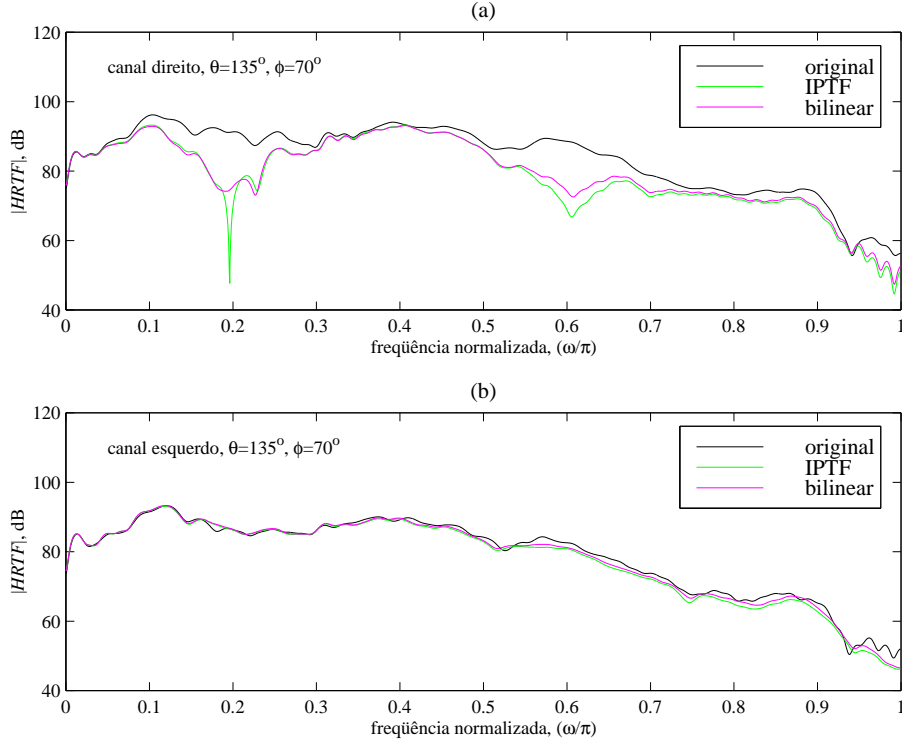


Figura 3.12: Exemplo do teste da interpolação (II). Módulo das respostas em frequência das HRTFs original e interpoladas: (a) canal direito e (b) canal esquerdo.

$IPTF_{i,1}$ e da $IPTF_{i,2}$, os três ponderadores α , β e γ e, ainda, os 6 necessários para o cálculo dos ponderadores mais o normalizador final. Isso significa que o maior número de coeficientes que deveria ser utilizado é dado por

$$NC_{max}^{IPTF} = \left\lfloor \frac{143 - 33 - 3 - 6 - 1}{2} \right\rfloor = 50, \quad (3.18)$$

onde $\lfloor \cdot \rfloor$ é o menor inteiro mais próximo. Dessa forma, se a ordem das IPTFs utilizadas for menor que 24 (ou 49 coeficientes—número de coeficientes menor que NC_{max}^{IPTF}), a complexidade computacional dessa forma de interpolação será menor que a da forma bilinear.

No Capítulo 4, serão discutidas algumas formas de simplificação de modelo que podem ser utilizadas para reduzir a ordem das IPTFs obtidas com a simples razão das HRTFs.

Capítulo 4

Simplificação das IPTFs

Na Seção 3.4, fez-se uma estimativa da ordem máxima que pode ser atribuída às IPTFs para que a utilização da interpolação com as IPTFs ainda se justifique.

Da forma como são geradas, as IPTFs são representadas por filtros IIR de ordem no máximo igual ao número de amostras de cada HRTF menos 1 (se não forem levados em conta os atrasos). Isto significa que a utilização da interpolação sem a simplificação das IPTFs torna-se mais custosa que a forma bilinear se nela forem utilizadas as HRIRs simplificadas obtidas em [14].

Neste capítulo, será feita uma explanação sobre as técnicas utilizadas para se chegar à versão simplificada das IPTFs. Os problemas encontrados e as soluções adotadas serão discutidos. Além disso, os métodos utilizados por cada uma das técnicas serão explicados.

Na Seção 4.1–**Métodos Utilizados na Simplificação das IPTFs**, serão discutidos os métodos que são utilizados na simplificação. Algumas técnicas de simplificação serão mostradas na Seção 4.2–**Técnicas de Simplificação**, onde se terá uma discussão sobre os problemas/limitações de cada uma delas. Na Seção 4.3–**Conclusões e Solução Adotada**, serão feitos alguns comentários sobre os problemas e limitações encontrados e se indicará a solução adotada para tornar possível a simplificação.

4.1 Métodos para Simplificação das IPTFs

Nessa tese, foram implementadas várias técnicas de simplificação que utilizam métodos capazes de reduzir os modelos de funções de transferência ou de tornar melhores os resultados dessas técnicas de redução de modelo.

Esses métodos são utilizados de várias formas diferentes, combinados ou separados, dependendo da técnica utilizada. As técnicas serão descritas na Seção 4.2.

Os métodos que serão descritos nesta seção são:

1. Método de redução de modelo balanceado [7, 19] (BMR–*Balanced Model Reduction*).
2. Identificação adaptativa por *Equation Error* [20, 21].
3. Suavização de espectro [14, 22].
4. Aproximação complexa.

4.1.1 Método BMR

A representação por equações de estado que tem os gramianos de suas matrizes de controlabilidade e observabilidade iguais e diagonais são chamados de modelos balanceados. Esses modelos têm como característica principal permitirem a ordenação dos estados pela sua “importância” na representação do sistema em questão.

Essa característica é bastante importante para os casos de redução de ordem dos modelos, pois, com os estados ordenados, pode-se truncar o modelo, eliminando estados menos importantes, e obter uma aproximação de ordem menor para o sistema.

Para tal, é necessário, primeiramente, tornar balanceado o sistema dado por

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) \end{aligned}, \quad (4.1)$$

onde \mathbf{A} , \mathbf{B} , \mathbf{C} e \mathbf{D} são as matrizes da representação de estados, $\mathbf{x}(k)$ contém os estados, $\mathbf{u}(k)$ contém as entradas e $\mathbf{y}(k)$ contém as saídas. Para isso, calcula-se a

matriz de Hankel dada por

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 & \mathbf{H}_3 & \cdots \\ \mathbf{H}_2 & \mathbf{H}_3 & \cdots & \\ \mathbf{H}_3 & \vdots & \ddots & \\ \vdots & & & \end{bmatrix}, \quad (4.2)$$

onde $\mathbf{H}_k = \mathbf{C}\mathbf{A}^{k-1}\mathbf{B}$ para $k = 1, 2, 3, \dots$, que são escalares, no caso de sistemas de uma única entrada $u(k)$ e de uma única saída $y(k)$. Essa matriz pode ser escrita como

$$\mathbf{H} = \mathbf{W}_o \mathbf{W}_c, \quad (4.3)$$

onde \mathbf{W}_o e \mathbf{W}_c são, respectivamente, as matrizes de observabilidade e controlabilidade do sistema, dadas por

$$\mathbf{W}_o = \begin{bmatrix} \mathbf{C}^T & \mathbf{A}^T \mathbf{C}^T & \cdots & (\mathbf{A}^T)^k \mathbf{C}^T & \cdots \end{bmatrix}^T \quad (4.4)$$

e

$$\mathbf{W}_c = \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \cdots & \mathbf{A}^k \mathbf{B} & \cdots \end{bmatrix}. \quad (4.5)$$

Em função de \mathbf{W}_o e \mathbf{W}_c , as matrizes gramianas do sistema são

$$\mathbf{P} = \mathbf{W}_c \mathbf{W}_c^T \quad (4.6)$$

e

$$\mathbf{Q} = \mathbf{W}_o^T \mathbf{W}_o. \quad (4.7)$$

Como

$$\text{rank}(\mathbf{H}) = \text{rank}(\mathbf{A}) = \mathcal{K}, \quad (4.8)$$

é possível, então, truncando-se as matrizes \mathbf{W}_o e \mathbf{W}_c de forma a manter somente as linhas e colunas LI, reescrever a matriz de Hankel da Equação (4.2), para um sistema mono-entrada mono-saída, como

$$\mathbf{H} = \begin{bmatrix} h_1 & h_2 & \cdots & h_K \\ h_2 & h_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ h_K & 0 & \cdots & 0 \end{bmatrix}, \quad (4.9)$$

onde h_i são escalares.

Essa matriz \mathbf{H} ainda pode ser decomposta na forma

$$\mathbf{H} = \mathbf{V}\Lambda\mathbf{V}^T, \quad (4.10)$$

onde

$$\mathbf{V}\mathbf{V}^T = \mathbf{I}, \quad (4.11)$$

Λ é a matriz diagonal contendo os autovalores $\lambda_1, \dots, \lambda_K$ de \mathbf{H} ordenados decrescentemente com as linhas e \mathbf{I} é a matriz identidade.

Nesse ponto, o objetivo é obter uma transformação \mathbf{T} capaz de tornar balanceado o sistema. Em [19], mostra-se que, para isso, pode-se decompor a matriz \mathbf{Q} na forma

$$\mathbf{Q} = \mathbf{S}^T\mathbf{S} \quad (4.12)$$

e, assim, utilizando a matriz Λ da Equação (4.10), resolver o sistema de equações

$$\mathbf{S}\mathbf{P}\mathbf{S}^T = \mathbf{U}\Lambda\mathbf{U}^T, \quad (4.13)$$

onde

$$\mathbf{U}\mathbf{U}^T = \mathbf{I}, \quad (4.14)$$

obtendo a matriz \mathbf{U} .

Tendo, agora, as matrizes \mathbf{S} , \mathbf{U} e Λ , pode-se definir a transformação

$$\mathbf{T} = \mathbf{S}^{-1}\mathbf{U}\Lambda^{1/4}, \quad (4.15)$$

que leva o sistema formado pelas matrizes \mathbf{A} , \mathbf{B} , \mathbf{C} e \mathbf{D} a um sistema balanceado, fazendo

$$\mathbf{A}_b = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}, \quad \mathbf{B}_b = \mathbf{T}^{-1}\mathbf{B}, \quad \mathbf{C}_b = \mathbf{C}\mathbf{T}. \quad (4.16)$$

Dessa forma [19], as matrizes gramianas do sistema balanceado \mathbf{P}_b e \mathbf{Q}_b acabam por ficar na forma

$$\mathbf{P}_b = \mathbf{Q}_b = \Lambda^{1/2}, \quad (4.17)$$

onde $\Lambda^{1/2}$ é

$$\begin{aligned} \Lambda^{1/2} &= \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_K}) \\ &= \text{diag}(\sigma_1, \dots, \sigma_K), \end{aligned} \quad (4.18)$$

sendo σ_i os valores singulares do sistema.

Se a matriz Λ for decomposta em duas partes de forma que

$$\Lambda = \begin{bmatrix} \Lambda_1 & \mathbf{0} \\ \mathbf{0} & \Lambda_2 \end{bmatrix}, \quad (4.19)$$

onde as matrizes $\mathbf{0}$ são matrizes de zeros,

$$\Lambda_1 = \text{diag}(\lambda_1, \dots, \lambda_k) \quad (4.20)$$

e

$$\Lambda_2 = \text{diag}(\lambda_{k+1}, \dots, \lambda_K), \quad (4.21)$$

as matrizes do sistema balanceado poderão ser também decompostas na forma

$$\mathbf{A}_b = \begin{bmatrix} \mathbf{A}_{1,1} & \mathbf{A}_{1,2} \\ \mathbf{A}_{2,1} & \mathbf{A}_{2,2} \end{bmatrix}, \quad \mathbf{B}_b = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} \quad \text{e} \quad \mathbf{C}_b = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix},$$

onde $(\mathbf{A}_{1,1}, \mathbf{B}_1$ e $\mathbf{C}_1)$ são as matrizes do subsistema que representam o sistema de ordem reduzida k .

Essa é a forma de se obter um sistema de ordem reduzida partindo das equações de estado que o representam. Se, ao invés de se partir de qualquer tipo de sistema, partir-se de um filtro FIR, podem-se aproveitar determinadas características para facilitar o cálculo do sistema de ordem reduzida, diminuindo os problemas numéricos.

4.1.1.1 Método BMR Partindo de Filtros FIR

Em [7], para reduzir a complexidade computacional dos filtros FIR que representam as HRTF, foi utilizado o método BMR partindo de filtros FIR [19].

Para esse método as matrizes de estado da Equação (4.1) ficam da forma

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ & & \cdots & & \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} c_1 & c_2 & c_3 & \cdots & c_n \end{bmatrix}, \quad \mathbf{D} = c_0.$$

Assim, a matriz de Hankel é finita e da forma

$$\mathbf{H} = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \\ c_2 & c_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c_n & 0 & \cdots & 0 \end{bmatrix}, \quad (4.22)$$

onde c_i é o i -ésimo coeficiente do filtro FIR.

Com essa matriz de Hankel e mais algumas propriedades da representação de estados de filtros FIR, chega-se à transformação \mathbf{T} pela simples decomposição da matriz \mathbf{H} na forma da Equação (4.10). Naquela equação, para esse caso, $\mathbf{T} = \mathbf{V}$.

Assim, chega-se no sistema balanceado pelas equações

$$\mathbf{A}_b = \mathbf{V}_k^T \mathbf{A} \mathbf{V}_k, \quad \mathbf{B}_b = \mathbf{V}_k^T \mathbf{B}, \quad \mathbf{C}_b = \mathbf{C} \mathbf{V}_k, \quad (4.23)$$

onde \mathbf{V}_k é composta das primeiras k colunas da matriz \mathbf{V} , de forma que

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_k & \mathbf{V}_{n-k} \end{bmatrix}, \quad (4.24)$$

sendo n a ordem do filtro FIR e k a ordem desejada para filtro IIR final.

O algoritmo final para o caso de redução de um determinado filtro FIR cujos coeficientes são dados fica da forma:

1. Criar a matriz de Hankel \mathbf{H} como indica a Equação (4.22);
2. decompor a matriz \mathbf{H} para obter \mathbf{V} e Λ ;
3. definir a ordem final k desejada ou escolher a partir de qual autovalor a matriz \mathbf{V} pode ser truncada;
4. calcular as matrizes \mathbf{A}_k , \mathbf{B}_k e \mathbf{C}_k ; e
5. converter o sistema \mathbf{A}_k , \mathbf{B}_k , \mathbf{C}_k e $\mathbf{D} = c_0$ na forma de função de transferência.

4.1.2 Identificação Adaptativa por *Equation Error*

Para o caso da redução de modelos, os métodos adaptativos [20] de identificação de sistemas devem ser tais que o número de coeficientes do filtro adaptativo seja menor que o número de coeficientes do modelo original.

Para um sistema genérico a identificar, os métodos adaptativos IIR têm a característica de necessitarem de uma menor ordem para atingir uma boa identificação, em comparação com os métodos FIR. Assim, pode-se adotar este tipo de método para que o sistema final tenha a menor ordem possível.

As formas de identificação por métodos adaptativos baseiam-se na minimização de uma determinada função-objetivo. Essa função, para os casos de identificação, é dependente do erro entre as saídas do filtro adaptativo e do sistema que se quer identificar. O esquema da Figura 4.1 ilustra o cálculo do erro utilizado na função-objetivo, descrito pela equação

$$e(k) = d(k) - y(k), \quad (4.25)$$

onde $y(k)$ é a saída do filtro adaptativo dada por

$$y(k) = \sum_{n=1}^N a_n(k)y(k-n) + \sum_{m=0}^M b_m(k)x(k-m), \quad (4.26)$$

onde $d(k)$ é o sinal desejado (saída do sistema que se quer identificar, no caso de identificação de sistemas), N é a ordem do filtro adaptativo e M é a ordem do numerador da função de transferência do filtro adaptativo.

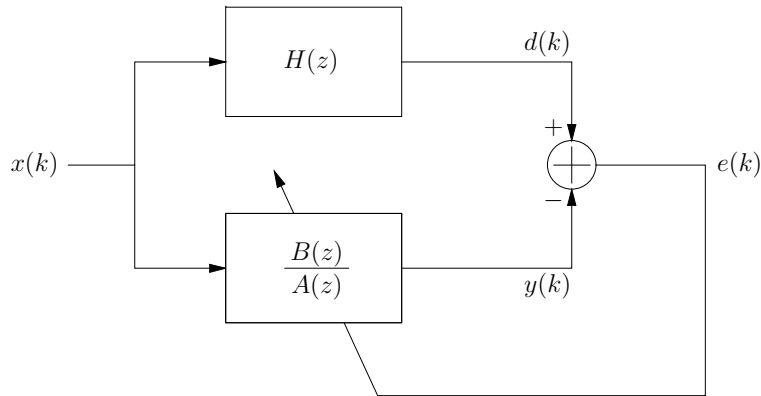


Figure 4.1: Identificação adaptativa IIR. Cálculo do erro.

As funções-objetivo mais freqüentemente adotadas são MSE (*Mean-Square Error*), estocástica, que pode ser descrita da forma

$$J = E[e(k)^2] \quad (4.27)$$

e o LS (*Least Squares*), determinística, dada por

$$J = \sum_{i=0}^k \lambda^i e(i)^2. \quad (4.28)$$

A principal característica dos filtros IIR é o fato de o sistema ter realimentação. Isso faz com que a função-objetivo associada ao filtro adaptativo não seja quadrática, podendo assumir múltiplos pontos de mínimo. Como consequência, o processo de adaptação pode estacionar em um mínimo local, que pode estar bem aquém do mínimo global. Por outro lado, o fato de os coeficientes da realimentação serem também adaptados pode tornar a convergência mais demorada.

Uma forma de reduzir esses problemas da adaptação dos filtros é utilizar a formulação chamada *equation error* [20, 21]. Essa formulação modifica ligeiramente a definição do erro presente na função-objetivo reescrevendo as Equações (4.25) e (4.26) da forma

$$e_e(k) = d(k) - y_e(k), \quad (4.29)$$

onde $y_e(k)$ é dado por

$$y_e(k) = - \sum_{n=1}^N a_n(k)d(k-n) + \sum_{m=0}^M b_m(k)x(k-m). \quad (4.30)$$

Isso torna a minimização da função-objetivo mais simples, já que o valor da saída do filtro adaptativo não entra no cálculo dos coeficientes a cada passo, sendo substituído pelo valor de saída desejado.

Na Figura 4.2, tem-se o esquema utilizado para o cálculo do erro da equa-

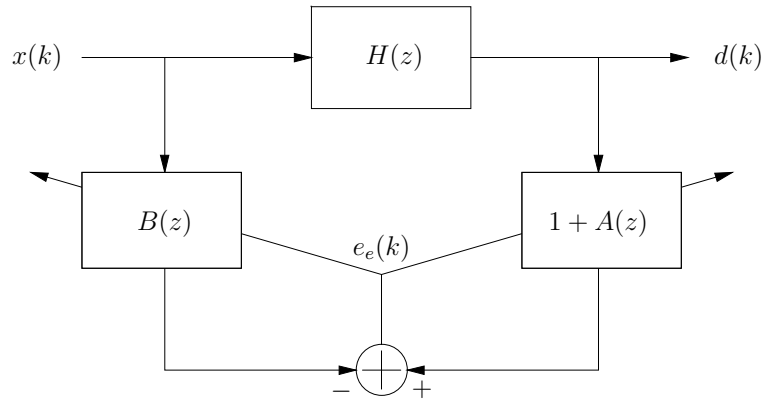


Figure 4.2: Identificação de sistemas por *equation-error*.

ção (4.29). Nessa figura, podem-se ver o sistema $H(z)$ que se quer identificar e os filtros adaptativos $B(z)$ e $A(z)$ (respostas do sistema de identificação), cujos coeficientes formarão a função de transferência de ordem reduzida (no caso da aplicação

desse sistema em redução de ordem) de forma que

$$\hat{H}(z) = \frac{B(z)}{1 + A(z)}. \quad (4.31)$$

Definindo-se os respectivos vetores de coeficientes como

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} \quad (4.32)$$

e

$$\mathbf{b} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_M \end{bmatrix}. \quad (4.33)$$

a atualização dos coeficientes a cada iteração pode ser dada por

$$\mathbf{a}_{l+1} = \mathbf{a}_l + \Delta_{\mathbf{a}} \quad (4.34)$$

e

$$\mathbf{b}_{l+1} = \mathbf{b}_l + \Delta_{\mathbf{b}}, \quad (4.35)$$

onde os vetores de variação de coeficientes do denominador ($\Delta_{\mathbf{a}}$) e do numerador ($\Delta_{\mathbf{b}}$), são calculados no sentido da redução do valor da função-objetivo J .

4.1.3 Métodos de Suavização de Espectro

Em alguns casos, implementar filtros a partir de suas especificações pode ser muito difícil. No caso das HRTFs ou das IPTFs, algumas características não muito importantes podem fazer com que a ordem da aproximação cresça muito sem necessidade.

Os métodos de suavização são utilizados para reduzir as variações de espectro existentes na resposta em frequência desejada, de forma que as variações menos importantes sejam suavizadas e as mais importantes não sejam alteradas.

Nos métodos implementados, são consideradas importantes as regiões de mais baixa frequência. A idéia é fazer com que o valor da resposta para uma determinada frequência seja alterado segundo o valor da resposta da região no seu entorno, região

essa tanto maior quanto maior a frequência em questão. Tal atuação sobre a resposta é justificada pela forma como o som é percebido.

Nessa tese, alguns métodos foram implementados para tentar reduzir os problemas que ocorrem nas aproximações das funções de transferência. São eles:

1. Suavização sobre o quadrado do módulo;
2. suavização sobre o módulo; e
3. suavização sobre a resposta em frequência complexa;

Os três métodos partem do mesmo princípio: calcular o valor da resposta para cada frequência fazendo algum tipo de média dos valores da resposta no entorno. Nas subseções seguintes será mostrada a forma com que cada método suaviza a resposta em frequência.

4.1.3.1 Suavização sobre o Quadrado do Módulo

Esse tipo de suavização faz a média quadrática do módulo da resposta em frequência [14], como pode ser visto na equação

$$|H_s(f)| = \sqrt{\frac{1}{f_2 - f_1 + 1} \sum_{l=f_1}^{f_2} |H_o(l)|^2}, \quad (4.36)$$

onde $H_o(f)$ e $H_s(f)$ são as respostas em frequência que se quer suavizar e suavizada, respectivamente,

$$f_1 = \left[\frac{f}{\sqrt{K}} \right] \quad (4.37)$$

e

$$f_2 = \left[\frac{f}{\sqrt{K}} \right], \quad (4.38)$$

sendo K a razão entre as frequências f_1 e f_2 , e $[\cdot]$ e $\lfloor \cdot \rfloor$ as notações para o maior inteiro mais próximo e para o menor inteiro mais próximo, respectivamente.

Com essa forma de suavização obtém-se somente o módulo da resposta em frequência suavizada. Se for preciso ter a resposta em frequência complexa, pode-se utilizar a fase de $H_o(f)$.

4.1.3.2 Suavização sobre o Módulo

No método mostrado na Subseção 4.1.3.1, o fato de se elevar ao quadrado o módulo da resposta em frequência faz com que as variações fiquem maiores e, com isso, mais difíceis de serem eliminadas.

Uma forma de reduzir essa característica foi testada: fez-se a suavização com a média aritmética do módulo, como na equação

$$|H_s(f)| = \frac{1}{f_2 - f_1 + 1} \sum_{l=f_1}^{f_2} |H_o(l)|, \quad (4.39)$$

onde $H_o(f)$ e $H_s(f)$ são as respostas em frequência que se quer suavizar e suavizada, respectivamente,

$$f_1 = \left\lceil \frac{f}{\sqrt{K}} \right\rceil \quad (4.40)$$

e

$$f_2 = \left\lfloor \frac{f}{\sqrt{K}} \right\rfloor, \quad (4.41)$$

sendo K a razão entre as frequências f_1 e f_2 , e $\lceil \cdot \rceil$ e $\lfloor \cdot \rfloor$ as notações para o maior inteiro mais próximo e para o menor inteiro mais próximo, respectivamente.

4.1.3.3 Suavização sobre a Resposta em Frequência Complexa

Além das formas de suavização descritas pelas Equações (4.36) e (4.39), que utilizam o módulo da resposta em frequência, existem outros métodos que utilizam a resposta complexa.

Em [22], os autores descrevem um método de suavização que, utilizando a resposta complexa, consegue melhorar o desempenho da suavização.

A utilização da resposta em frequência complexa na suavização torna a fase da resposta compatível com o módulo da resposta suavizada, e, se a forma de cálculo do filtro suavizado leva em consideração a fase, fazer a suavização complexa é mais vantajoso.

Então, testou-se fazer a suavização fazendo uma média, como nas Equações (4.36) e (4.39), só que sobre a resposta em frequência complexa, de forma que

$$H_s(f) = \frac{1}{f_2 - f_1 + 1} \sum_{l=f_1}^{f_2} H_o(l), \quad (4.42)$$

onde $H_o(f)$ e $H_s(f)$ são as respostas complexas em frequência que se quer suavizar e suavizada, respectivamente,

$$f_1 = \left\lceil \frac{f}{\sqrt{K}} \right\rceil \quad (4.43)$$

e

$$f_2 = \left\lfloor \frac{f}{\sqrt{K}} \right\rfloor, \quad (4.44)$$

sendo K a razão entre as frequências f_1 e f_2 , e $\lceil \cdot \rceil$ e $\lfloor \cdot \rfloor$ as notações para o maior inteiro mais próximo e para o menor inteiro mais próximo, respectivamente.

Com essa forma obtiveram-se melhores resultados, como será mostrado na Subseção 4.2.3.

4.1.4 Métodos de Aproximação Complexa

Em algumas técnicas, que serão mostradas na Seção 4.2, é necessário calcular os coeficientes de um filtro cuja resposta em frequência se aproxime de especificações dadas.

Foi utilizada uma função `invfreqz` que faz parte da *toolbox* de processamento de sinais do MATLAB [23]. Esta função, dados o vetor complexo da resposta em frequência, o vetor das respectivas frequências e as ordens dos polinômios do numerador e denominador, fornece como resposta os coeficientes do filtro.

O algoritmo utilizado na função `invfreqz` para o cálculo dos coeficientes do filtro faz uma primeira estimativa desses coeficientes através da minimização do erro quadrático. Como a superfície do erro não é quadrática, o que dificulta chegar ao erro mínimo global por esse método, essa resposta parcial é utilizada como ponto de partida para a outra parte do algoritmo, que consiste em uma otimização pelo método Gauss-Newton [20].

Esse algoritmo, de uma forma simplificada, pode ser visto a seguir:

- **Passo 1:** Obtém-se, primeiramente, uma estimativa através do método dos mínimos quadrados, minimizando o erro dado por:

$$\|e\|^2 = \mathbf{e}^H \mathbf{e} = (\mathbf{H}_c - \mathbf{H})^H (\mathbf{H}_c - \mathbf{H}) \quad (4.45)$$

onde \mathbf{H}_c é a resposta em frequência calculada com os coeficientes, \mathbf{H} é a resposta em frequência desejada e $\{\cdot\}^H$ é o Hermitiano de uma matriz, definido como $\mathbf{X}^H = \mathbf{X}^{*T}$.

- **Passo 2:** Sobre os coeficientes calculados no passo **1** (ou, no caso das iterações seguintes, no passo **4**), aplica-se o método Gauss-Newton, que, primeiramente, calcula o gradiente da função-objetivo fazendo

$$\nabla = \frac{\partial}{\partial \mathbf{t}} \{ \mathbf{e}^H \mathbf{e} \} = 2 \times \left[\frac{\partial \mathbf{e}}{\partial \mathbf{t}} \right] \times \mathbf{e} = 2\Delta^H \mathbf{e}, \quad (4.46)$$

onde Δ é a matriz de derivadas parciais que tem o mesmo número de linhas de \mathbf{e} e o número de colunas igual ao número de linhas de \mathbf{t} , sendo este o vetor de coeficientes dado por

$$\mathbf{t} = \begin{bmatrix} a_1 \\ \vdots \\ a_N \\ b_0 \\ \vdots \\ b_M \end{bmatrix}. \quad (4.47)$$

- **Passo 3:** Utilizando-se esse gradiente, que expressa a direção de máxima variação, calcula-se a direção de Gauss-Newton (\mathcal{GN}) aproximando a matriz Hessiana da forma

$$\begin{aligned} \mathcal{H} &= \frac{\partial^2}{\partial \mathbf{t} \partial \mathbf{t}} \{ \mathbf{e}^H \mathbf{e} \} \\ &= \frac{\partial \nabla}{\partial \mathbf{t}} \\ &= \frac{\partial}{\partial \mathbf{t}} [2\Delta^H \mathbf{e}] \\ &= 2 \left[\frac{\partial \Delta}{\partial \mathbf{t}} \right]^H \mathbf{e} + 2\Delta^H \frac{\partial \mathbf{e}}{\partial \mathbf{t}} \\ &\approx 2\Delta^H \frac{\partial \mathbf{e}}{\partial \mathbf{t}} = 2\Delta^H \Delta \end{aligned} \quad (4.48)$$

e, assim,

$$\mathcal{GN} = \mathcal{H}^{-1} \nabla. \quad (4.49)$$

- **item 4:** Com a direção \mathcal{GN} calculada, atualizam-se os coeficientes \mathbf{t} iterativamente da forma

$$\mathbf{t}_{i+1} = \mathbf{t}_i - \mu_i \mathcal{GN}, \quad (4.50)$$

de forma que $\mu_i = 1, 1/2, 1/4, \dots$, até que a norma de \mathbf{e} seja menor que a norma do erro da Equação (4.45) com os coeficientes calculados no passo **1** (ou no passo **4** na iteração anterior).

- **Passo 5:** Quando o passo 4 está concluído, se a norma de \mathcal{GN} for menor que a tolerância especificada (no caso 0,01), finaliza-se o algoritmo; caso contrário, retorna-se para o passo 2, mas, agora, utilizando os coeficientes calculados ao final do passo 4.

4.2 Técnicas de Simplificação

Na Seção 4.1, foram descritos alguns métodos que são utilizados na redução da ordem das IPTFs. Esses métodos podem ser utilizados em algumas técnicas de diferentes formas, dependendo da ordem de utilização e sobre quais funções de transferência estão sendo aplicados. Além disso, podem-se diferenciar as técnicas pela utilização ou não dos métodos responsáveis por melhorias da técnica, como, por exemplo, o que trata da suavização de espectro.

As técnicas utilizadas nessa tese, que serão descritas nas subseções a seguir, são as seguintes:

1. Simplificação das HRTFs utilizadas na razão (IPTF);
2. simplificação das IPTFs; e
3. simplificação de funções suavizadas.

4.2.1 Simplificação das HRTFs Utilizadas na Razão (IPTF)

O método BMR [19], como foi mencionado na Subseção 4.1.1.1, calcula o filtro IIR de ordem reduzida que aproxima o filtro FIR dado.

Como esse método foi utilizado em [7] para reduzir a ordem das HRIRs (filtros FIR), para obter os filtros que representam as IPTFs de ordem reduzida pensou-se, primeiramente, em reduzir a ordem das HRIRs e, posteriormente, obter as IPTFs fazendo a razão na frequência das respostas simplificadas e, em seguida, obter o filtro IIR que melhor aproxima a resposta dessa razão.

O algoritmo utilizado por essa técnica pode ser descrito pelos seguintes passos:

- **Passo 1:** Para cada par de HRTFs utilizado na composição da IPTF (um para o ouvido direito e outro para o esquerdo), aplica-se o método BMR, obtendo

para cada HRTF um filtro IIR de ordem $N/2$, onde N é a ordem desejada para o filtro IIR que representa a IPTF final;

- **Passo 2:** Calcula-se a resposta em frequência de cada um desses filtros calculados no passo 1;
- **Passo 3:** Com as respostas em frequência obtidas no passo 2, calculam-se as respostas em frequência relativas à IPTF para cada ouvido, que são dadas pelas Equações (3.2) e (3.3);
- **Passo 4:** Com a $IPTF_{i,f}^d$ e a $IPTF_{i,f}^e$ calculadas, calcula-se o filtro IIR de ordem N que tem resposta em frequência que mais se aproxima delas, utilizando o método descrito na Subseção 4.1.4 (`invfreqz`).

No entanto, um problema foi encontrado ao se aplicar o método BMR às HRTFs: a redução de ordem das HRTFs causa um erro na resposta em frequência que se transmite às IPTFs por elas formadas, prejudicando sua utilização na interpolação. Na Figura 4.3a, têm-se os módulos das respostas em frequência de uma determinada HRTF e de sua versão simplificada. Note que os “vales” da HRTF medida não são bem acompanhados pela HRTF simplificada. Os erros gerados na simplificação das HRTFs se propagam para as IPTFs finais, que podem levar a funções de transferência interpoladas que não remetem à posição escolhida para a fonte virtual. Na Figura 4.3b, tem-se um exemplo do erro na função interpolada com as HRTFs de ordem reduzida.

Como a faixa dinâmica da HRTF é um dos motivos desse problema, tentou-se solucioná-lo fazendo a redução de ordem pela aplicação do método BMR diretamente sobre as IPTFs. Note que esse problema é agravado pelo fato de a ordem para qual se deve reduzir o modelo da HRTF ter de ser a metade da ordem requerida para a IPTF final, como foi dito no passo 1.

4.2.2 Simplificação das IPTFs

Na Seção 4.2.1, mostrou-se a técnica que simplifica as HRTFs que serão utilizadas na razão. O problema encontrado com a dinâmica da HRTF pode ser reduzido ao se fazer a redução sobre a IPTF, já que esta tem a dinâmica menor do que aquela outra.

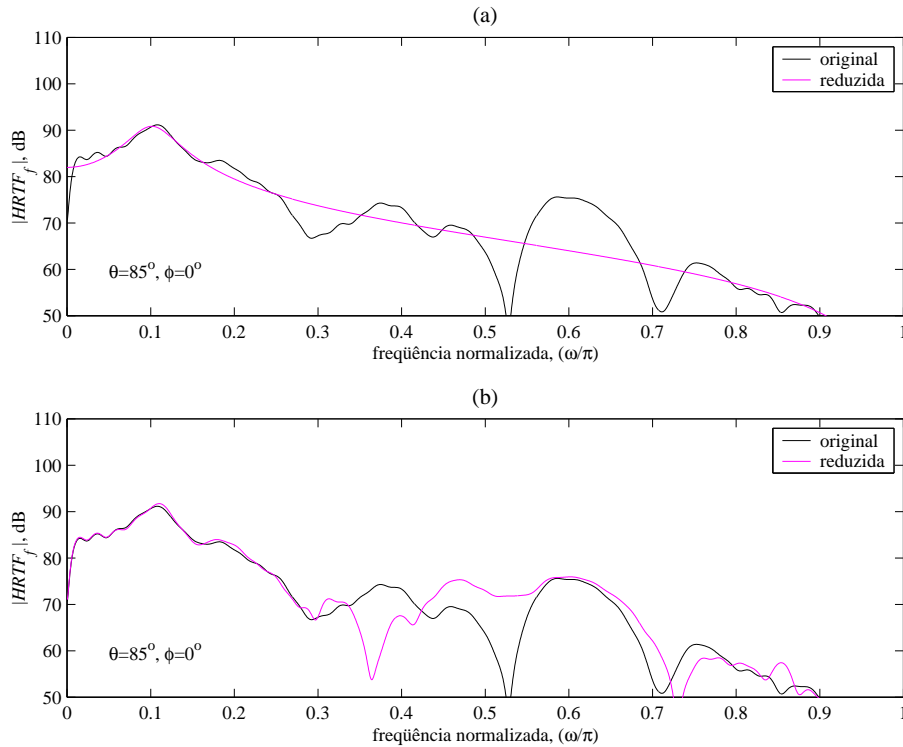


Figure 4.3: Aproximação aplicando o método BMR sobre as HRTFs:(a) comparação entre uma HRTF medida e sua versão simplificada; (b) comparação entre a $HRTF_f$ e o resultado da composição da $HRTF_i$ com a $IPTF_{i,f}$ simplificada.

Existe um pequeno problema ao se abordar a redução sobre as IPTFs: as IPTFs são representadas por uma razão de HRTFs, o que significa tentar reduzir filtros IIR.

Há duas alternativas para se fazer a redução: uma que necessita de uma aproximação FIR (técnica com aproximação FIR) e outra utilizando o método de redução de modelo diretamente sobre o filtro IIR que representa a IPTF (técnica sem aproximação).

4.2.2.1 Técnica com Aproximação FIR

A técnica de redução de ordem da IPTF com aproximação FIR consiste na utilização do método BMR descrito em [19] sobre a aproximação FIR obtida através do método descrito na Subseção 4.1.4.

O algoritmo utilizado nessa técnica segue os seguintes passos:

- **Passo 1:** Calcula-se a resposta em frequência da IPTF utilizando como

polinômios do numerador e do denominador da função de transferência as HRTFs correspondentes;

- **Passo 2:** Calcula-se a aproximação FIR de ordem alta que melhor aproxima a resposta em frequência, calculada no passo 1, através do método descrito na Subseção 4.1.4 (`invfreqz`);
- **Passo 3:** Sobre o filtro FIR calculado no passo 2, aplica-se o método BMR obtendo-se o filtro IIR de ordem reduzida relativo à IPTF.

Essa técnica de redução funciona razoavelmente bem na maioria dos casos. O grande problema, no entanto, é que a aproximação FIR pode ter que ser de ordem muito grande para conseguir aproximar corretamente a IPTF. Assim, como não se pode aumentar indefinidamente a ordem da aproximação FIR (não só pela computação da razão mas também por causa da computação da aproximação pelo método BMR), nos casos em que ocorre o problema a aproximação não acompanha a resposta da IPTF e o método BMR não consegue aproximá-la bem.

Na Figura 4.4a, tem-se um exemplo em que a redução com o método BMR aplicado à IPTF funciona bem. Já na Figura 4.4b, é possível notar o problema ocorrendo na aproximação com FIR. Note que as duas aproximações FIR foram feitas com ordem 400, que pode ser considerada grande comparada à ordem 127 das HRIRs medidas. O problema de se aumentar mais ainda a ordem da aproximação FIR é que o custo computacional do cálculo dos autovalores e autovetores da matriz \mathbf{H} , pela Equação (4.10), se torna muito elevado, já que a ordem da matriz \mathbf{H} é igual a ordem da aproximação FIR. Isso impossibilita, em alguns casos, a obtenção do filtro de ordem reduzida.

4.2.2.2 Técnicas sem Aproximação FIR

Em vista do erro que ocorre ao se tentar aproximar as IPTFs por filtros FIR, pensou-se em utilizar métodos que não precisassem necessariamente partir de filtros FIR, abdicando, assim, da redução da complexidade para o cálculo das funções de ordem reduzida conseguida com o método BMR que parte de filtros FIR, mencionada na Subseção 4.1.1.1. Dois métodos foram testados: o método BMR que parte de

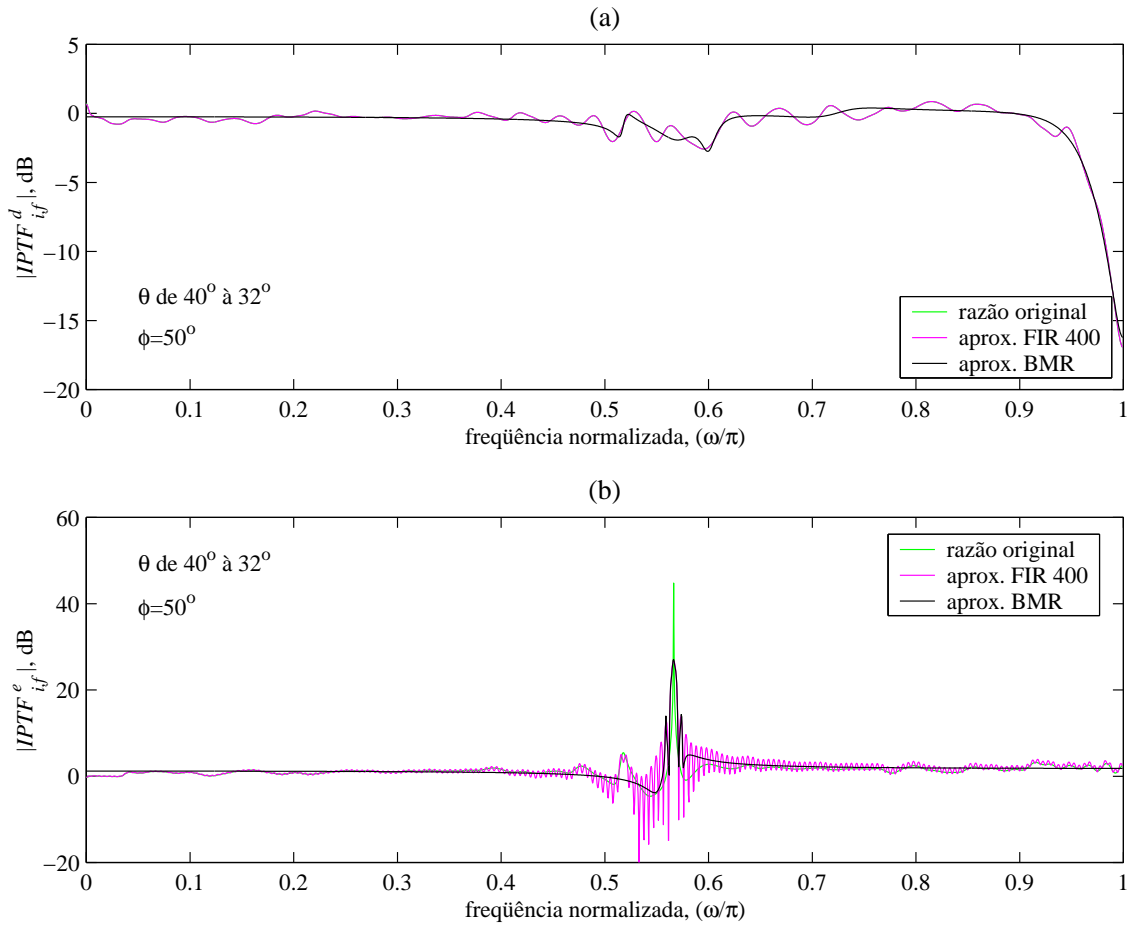


Figure 4.4: Aproximação aplicando o método BMR sobre as IPTFs. IPTFs obtidas com as HRTFs medidas comparadas com suas versões reduzidas: (a) sem erro na aproximação FIR; (b) com erro da aproximação FIR.

filtros IIR de ordem alta (Subseção 4.1.1) e o método adaptativo *Equation Error* (Subseção 4.1.2).

Técnica BMR partindo de filtro IIR

Essa técnica é utilizada de forma similar àquela mostrada na subseção 4.2.2.1. A diferença é que, neste caso, não é necessário calcular a aproximação FIR referente à IPTF, reduzindo o problema visto na Figura 4.4.

Modificando-se o algoritmo da Subseção 4.2.2.1, tem-se:

- **Passo 1:** Forma-se a IPTF com os coeficientes dos filtros FIR relativos às HRTFs medidas;
- **Passo 2:** Aplica-se o método mencionado na Subseção 4.1.1 (BMR a partir

do IIR) diretamente sobre cada IPTF, obtendo-se os filtros IIR de ordem reduzida N .

Utilizando esta técnica, ao se observarem todas as aproximações das IPTFs, percebeu-se uma redução significativa do número de funções com problemas. Isso pode ser verificado na Figura 4.5, onde é mostrado que o problema mencionado na Figura 4.4b não ocorre.

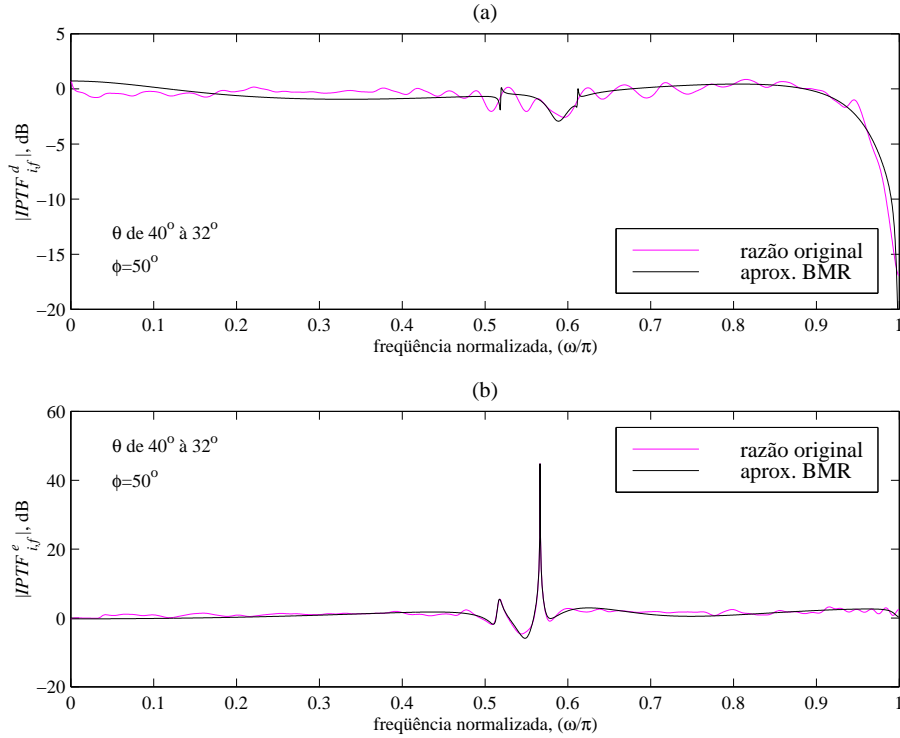


Figure 4.5: Exemplo de resultado do método BMR sobre as IPTFs representadas por filtros IIR. IPTFs obtidas com as HRTFs medidas comparadas com sua versão reduzida: (a) canal direito; (b) canal esquerdo.

No entanto, devido ao fato de a técnica BMR que parte do filtro IIR envolver cálculos mais críticos, como o cálculo de matrizes inversas de ordem relativamente alta, essa técnica fica mais propensa a erros numéricos, que podem tornar o sistema instável, não-controlável ou não-observável. Apesar disso, somente poucas IPTFs se apresentaram em tal situação, impossibilitando, assim, a obtenção de suas respectivas aproximações. Alguns métodos [24] numericamente mais eficazes podem ser utilizados para se tentar resolver tal impossibilidade. Nessa tese, no entanto, eles não foram implementados. Tentou-se, então, resolver tal problema de duas maneiras: primeiramente, verificando-se se as funções $IPTF_{i,f}$ prob-

lemáticas poderiam ser substituídas pelas inversas das $IPTF_{f,i}$ (bastando para isso que essas últimas tivessem sido calculadas) de outra forma, utilizando os métodos de suavização de espectro mencionados na Subseção 4.1.3, como será mostrado na Subseção 4.2.3. O primeiro procedimento obteve êxito, restando sem solução apenas um canal de uma das funções. Isso foi resolvido pelo segundo procedimento, que, utilizando a suavização conseguiu gerar a função que não foi simplificada pelo primeiro procedimento. Assim, conseguiu-se obter um conjunto completo de IPTFs simplificadas.

Técnica Adaptativa por *Equation Error*

A técnica adaptativa por *equation error* utiliza o método de mesmo nome descrito na Subseção 4.1.2 para calcular o filtro IIR de ordem reduzida.

O algoritmo para essa técnica fica da seguinte forma:

- **Passo 1:** Como na técnica mostrada anteriormente (técnica BMR a partir de filtros IIR), forma-se o filtro IIR representante da IPTF que se quer aproximar com os coeficientes das HRIRs medidas a ela relacionadas;
- **Passo 2:** A IPTF é colocada como o sistema a se identificar (como descrito na Subseção 4.1.2) e, após a função-objetivo ter sido minimizada, tem-se o filtro IIR de ordem reduzida.

Essa técnica funciona de forma um pouco melhor que a do método BMR com aproximação FIR. Realmente os problemas com a aproximação FIR são eliminados, só que, como mostrado na Subseção 4.2.1, as variações encontradas nas IPTFs também aparecem muito suavizadas. Nesse caso, o erro está presente em toda a resposta, o que faz que essa técnica tenha um comportamento um pouco pior que a daquela seção (onde os erros apareciam mais nos “vales” da resposta em frequência). Na Figura 4.6, tem-se um exemplo do erro que pode ocorrer quando se utiliza essa técnica. Pode-se notar que o tipo de erro devido à aproximação FIR mostrada na Figura 4.4 não mais aparece aqui. No entanto, nota-se que a resposta aparece muito suavizada, e até os “picos” mais evidentes não são bem modelados.

É possível que a atribuição de um custo para cada frequência, de forma que os “vales” e “picos” fossem priorizados pelo algoritmo de minimização do erro

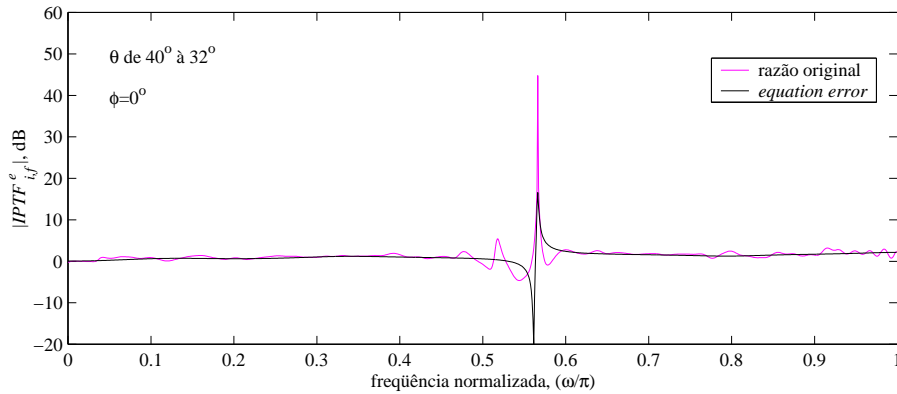


Figure 4.6: Redução de ordem aplicando o método *equation error*. IPTF obtida com as HRIRs medidas comparada com sua versão reduzida.

resultasse em erros menores nas aproximações das IPTFs. Porém, como as IPTFs têm suas respostas com muitas variações, o que acarretaria muitos “vales” e “picos” importantes, esse tipo de artifício pode não se mostrar tão eficaz.

Outro fato importante é que o método *equation-error* é um artifício para se fazer, de forma mais eficaz, a identificação de sistemas IIR com ordem suficiente. No entanto, aqui foi utilizado como um método de redução de ordem. Não está provado que esse método obtém resultados ótimos com ordem insuficiente, o que significa que, apesar de a função-objetivo do método *equation error* ser quadrática, seu mínimo global não é necessariamente o mesmo mínimo conseguido ao se utilizar do erro de saída [20]. Mesmo assim, fez-se a experiência e constatou-se que a minimização da função-objetivo do método *equation error* dá, como solução da identificação, uma função de transferência aceitável embora não tenha a mesma interpretação que a da solução através da minimização do erro de saída.

Um teste que ainda pode ser feito com a filosofia do *equation-error* é utilizar o método Steiglitz-McBride [20], que, ao modificar um pouco a função do erro e, assim, a função-objetivo a ser minimizada, consegue chegar mais próximo da função que seria identificada pelo método que utiliza o erro de saída. Claro que o problema da ordem reduzida persiste, e não se pode garantir que o resultado seja ótimo.

4.2.3 Simplificação de Funções Suavizadas

Nas técnicas mostradas nas Subseções 4.2.1 e 4.2.2, em alguns casos, ocorrem problemas na simplificação das funções, coincidindo com a ocorrência de variações

bruscas, normalmente nas altas frequências, que poderiam ser atribuídas a prováveis erros de medida, já que estão em uma região mais crítica para o equipamento de medição.

No caso da redução de ordem sendo feita sobre as HRTFs, em poucos casos os problemas encontrados foram devido a tais variações, mesmo que essas variações tornassem mais difícil a simplificação das HRTFs. Isso pelo fato de que a própria simplificação prioriza as regiões de ganhos mais elevados (veja na Subseção 4.2.1 a Figura 4.3).

Já no caso da redução feita sobre as IPTFs, esse problema cria variações muito bruscas em altas frequências. Essas variações acabam sendo interpretadas pelo método de simplificação como parte essencial das IPTFs e fazem com que partes mais importantes (até mesmo mais audíveis) sejam desconsideradas para se conseguir menor erro de aproximação nas altas frequências. Na Figura 4.7, pode-se ver um exemplo de IPTF com o problema descrito acima e o resultado da redução de ordem.

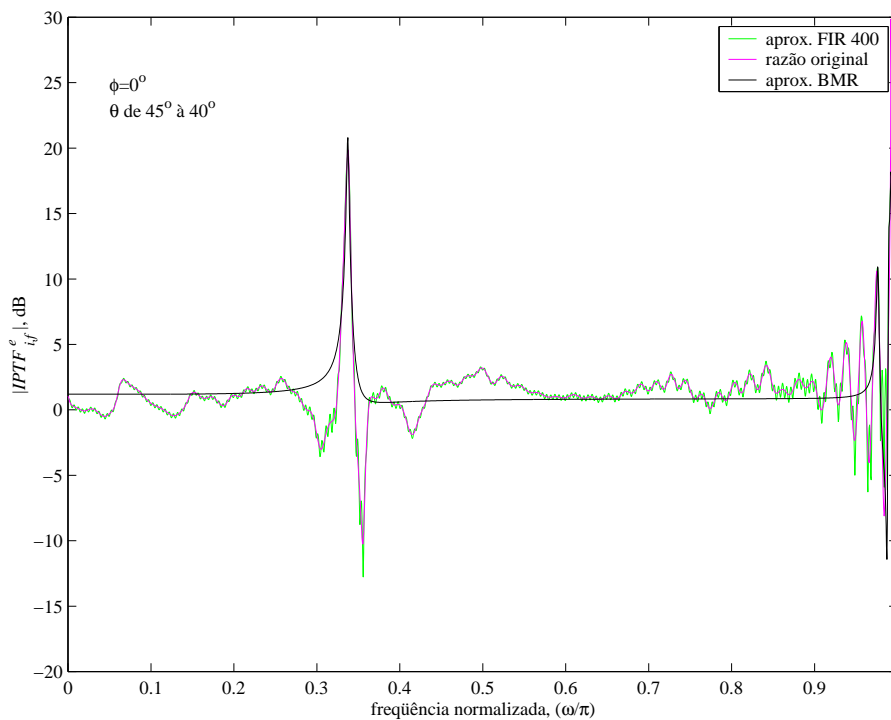


Figure 4.7: Efeito das grandes variações nas altas frequências sobre o método de redução de ordem das IPTFs.

Para tentar minimizar esse tipo de problema e, assim, facilitar a redução de

ordem dos filtros, podem-se utilizar métodos de suavização de espectro [14] descritos na Subseção 4.1.3.

Os métodos de suavização são utilizados sobre as respostas em frequência complexa ou seus módulos e, ainda, sobre as HRTFs ou sobre as IPTFs, dependendo da técnica utilizada.

Quando da técnica de simplificação utilizando a suavização sobre as HRTFs, o algoritmo da Subseção 4.2.1 é modificado de forma que, antes de se simplificar as HRTFs (passo 1), é necessário:

- **Passo 1:** Calcular a resposta em frequência de cada uma das HRTFs;
- **Passo 2:** Aplicar uma das formas de suavização descritas na Subseção 4.1.3 a cada HRTF;
- **Passo 3:** Calcular a função de transferência de cada HRTF suavizada; e
- **Passo 4:** Com a função de transferência calculada, seguir os passos do algoritmo da Subseção 4.2.1.

Nessa técnica, no entanto, ocorre o mesmo mostrado na Figura 4.3: os “vales” não são bem definidos pelo modelo reduzido. Isso já era esperado, pois, se os erros das HRTFs simplificadas já atrapalhavam a obtenção das IPTFs, com as HRTFs suavizadas as variações bruscas são reduzidas, mas as variações importantes, que já eram prejudicadas, ficam mais reduzidas ainda.

Já no caso das IPTFs, a suavização deve ser aplicada de forma diferente, dependendo da técnica escolhida dentre as descritas na Subseção 4.2.2 (técnica com ou sem a aproximação FIR e técnica *equation error*).

No caso da técnica com aproximação FIR (Subseção 4.2.2.1), só é necessário inserir o passo de suavização entre os passos 1 e 2 daquele algoritmo pois, no passo 1, já é calculada a resposta em frequência que será suavizada.

Nessa técnica, a simplificação das funções com as variações mostradas na Figura 4.7 são melhores que a simplificação sem suavização, no entanto, em alguns casos ainda aparecem os problemas da aproximação FIR, como pode ser visto na Figura 4.8. Nessa figura, são mostradas a resposta em frequência suavizada de uma IPTF antes de se calcular a aproximação FIR (chamada **original**) e dois resultados

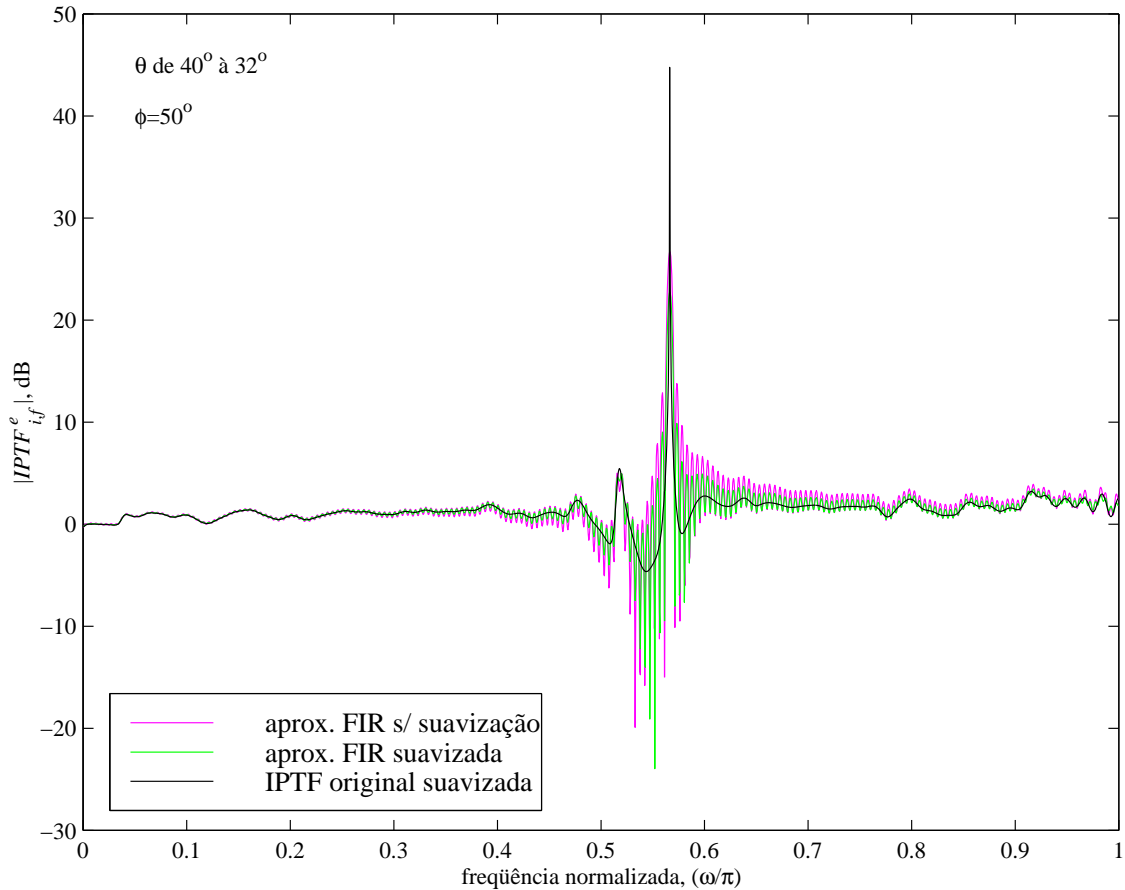


Figure 4.8: Exemplo de simplificação das IPTFs suavizadas com a técnica BMR com aproximação FIR.

da aproximação FIR: aquela que é obtida sem suavização; e a que é obtida com a suavização. Note que há uma pequena redução na variação gerada pela aproximação FIR, mas a variação não é eliminada nesse exemplo.

Já no caso das técnicas sem aproximação FIR, como o processo de redução de ordem é feito diretamente sobre a IPTF original, é necessário o cálculo da resposta em frequência a ser suavizada antes de se fazer a simplificação. Assim, antes de seguir um dos algoritmos descritos na Subseção 4.2.2.2, deve-se executar o seguinte procedimento:

- **Passo 1:** Calcular a resposta em frequência da IPTF que se quer simplificar;
- **Passo 2:** Aplicar um dos métodos de suavização na resposta em frequência;
- **Passo 3:** Calcular o filtro IIR de mesma ordem que a IPTF com o método de aproximação complexa (`invfreqz`);

- **Passo 4:** Aplicar o método que parte do filtro IIR (Subseção 4.2.2.2).

Com a técnica de *equation error*, a suavização não surtiu quase efeito algum. Já ao se aplicar a técnica BMR que parte de filtros IIR, que parecia ser a mais promissora, problemas numéricos na obtenção da função de transferência suavizada (passo **3**) fizeram muitas IPTFs se apresentarem como instáveis, não-controláveis ou não-observáveis. Com isso, um conjunto de IPTFs não pôde ser obtido utilizando esta técnica sobre as funções-razão.

Verificando-se se a inversa da $IPTF_{f,i}$ poderia ser calculada e, assim, ser a substituta da $IPTF_{i,f}$ que não pôde ser calculada, reduziu-se o número de casos problemáticos. Mesmo assim, o número de funções que não puderam ser determinadas, ainda era elevado.

Um dos fatos causadores da não-eficiência da suavização, nesse último caso, pode ser o de a suavização ser feita através de uma média que gera no entorno de “picos” muito elevados uma resposta em frequência não muito suave (veja Figura 4.9). Essa não suavidade torna a função de transferência mais difícil de ser simplificada pelo algoritmo mostrado na seção 4.2.2.2.

Uma solução que está sendo estudada, podendo vir a ser incorporada ao algoritmo proposto, é a da utilização da média geométrica sobre o módulo das funções de transferência, que pode ser descrita da forma:

$$|H_s(f)| = \frac{1}{f_2 - f_1 + 1} 10^{\sum_{l=f_1}^{f_2} \log |H_o(l)|}. \quad (4.51)$$

Assim, a suavização das respostas em frequência atuam de forma mais parecida sobre os “picos” e “vales”. O mesmo exemplo de suavização da Figura 4.9 pode ser visto na Figura 4.10, agora, suavizado com a média geométrica. Neste exemplo, pode-se verificar um melhor resultado, comparando-o com o da Figura 4.9.

4.3 Conclusões e Solução Adotada

Neste capítulo, descreveram-se as formas de simplificação das IPTFs. Foram encontrados alguns problemas relativos às formas de simplificação utilizadas. Os mais importantes foram:

1. erros de simplificação devido à priorização das regiões das respostas em frequência de mais alto ganho, quando da simplificação feita sobre as HRTFs;

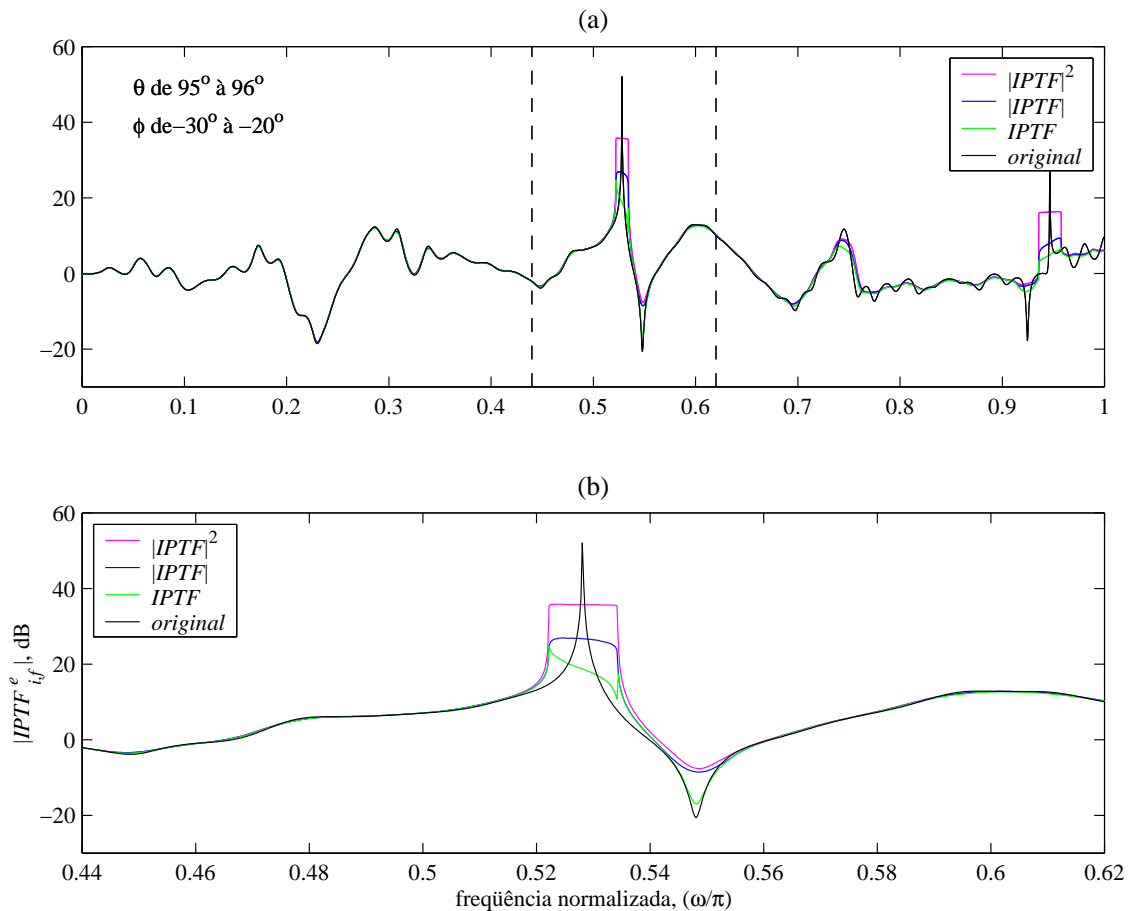


Figure 4.9: Atuação da suavização em funções com “picos” elevados. Exemplos com os três tipos de suavização (a) e detalhe da região delimitada pela linha tracejada (b).

2. erros relativos à aproximação FIR necessária para a implementação do método de redução de modelos balanceados (BMR) que parte de filtros FIR;
3. problemas numéricos devido à complexidade computacional (inversão de matrizes mal-condicionadas de ordem alta) do método de simplificação BMR sobre as IPTFs representadas por filtros IIR; e
4. problemas numéricos de mesma natureza que os do item anterior, só que provocados pelos métodos de suavização.

A técnica que se mostrou mais eficaz foi a que utiliza o método BMR aplicado sobre filtros IIR. Com essa técnica conseguiu-se gerar um conjunto completo de IPTFs simplificadas, sendo que apenas uma delas teve que ser feita com a suavização (que para aquela determinada função apresentava solução).

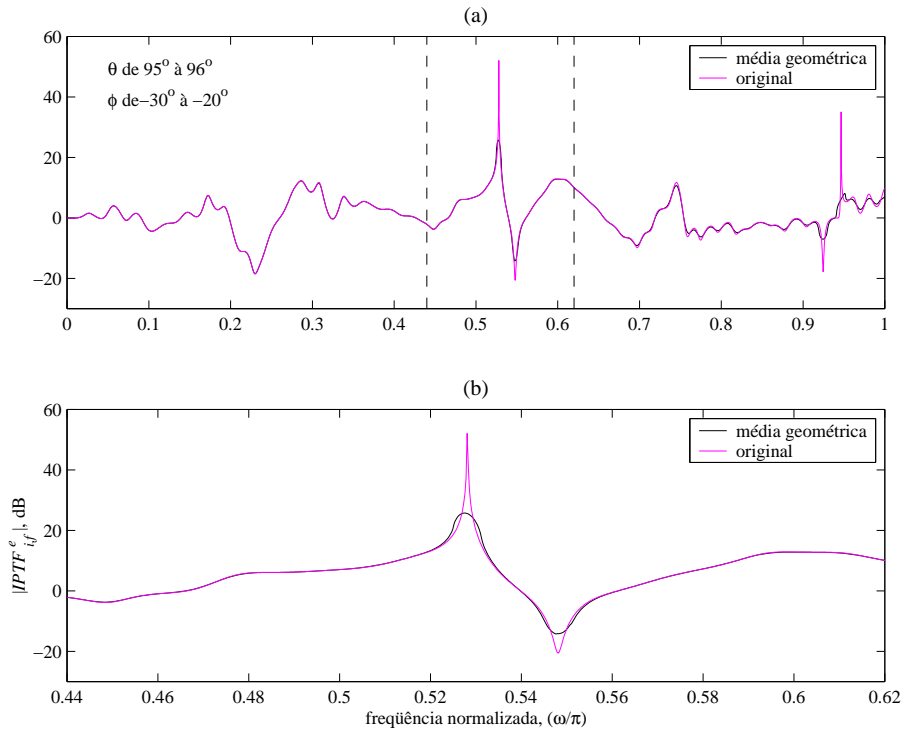


Figure 4.10: Exemplo de suavização utilizando a média geométrica. Em (a), comparação do módulo da função de transferência original e suavizada. Em (b), detalhe da região delimitada pela linha tracejada de (a).

As funções simplificadas são filtros IIR de ordem 10 e a suavização, no caso em que foi utilizada, foi feita sobre a resposta em frequência complexa.

Com os testes implementados, pôde-se notar que a suavização tem grande potencial, pois uma grande parte das IPTFs apresenta as grandes variações em altas frequências mostradas na Figura 4.7. No entanto, deve ser procurada uma forma de suavização mais eficaz.

No próximo capítulo, será mostrado como se pode implementar um sistema de geração de movimento utilizando as IPTFs, sejam elas simplificadas ou não.

Chapter 5

Movimento da Fonte Virtual

O objetivo final do modelamento das HRTFs é criar uma situação auditiva virtual o mais possível parecida com a situação real a que remete. Um efeito bastante interessante é o de variação de posição da fonte virtual.

Esse efeito pode ser conseguido fazendo-se a interpolação das HRTFs para a posição ocupada pela fonte virtual a cada instante. E quanto mais precisa a interpolação referente ao ponto em questão, mais real será a sensação de movimento obtida.

Foi implementado o algoritmo de movimento utilizando a forma de interpolação apresentada no Capítulo 3, que utiliza o par de HRIRs medidas relativas ao ponto mais próximo do qual se quer calcular a HRIR e dois pares de IPTFs que usam as HRTFs associadas àquelas HRIRs como $HRTF_i$, no cálculo das HRIRs de cada instante.

Nesse capítulo, será mostrado na Seção 5.1–**Algoritmo de Movimento**, o algoritmo utilizado para gerar o efeito de movimento sobre sinais mono-aurais utilizando as IPTFs. Por fim, as modificações realizadas nesse algoritmo para se utilizarem as IPTFs simplificadas mencionadas no Capítulo 4, bem como a solução final adotada para o sistema de geração de som tridimensional, serão mostradas na Seção 5.2–**Movimento com IPTFs Simplificadas–Solução Adotada**.

5.1 Algoritmo de Movimento

Para conseguir um efeito de movimento da fonte virtual, o primeiro passo a ser dado é decidir de que forma a posição será atualizada. Essa atualização pode ser feita em tempo real ou de forma *off-line*.

Uma das maneiras de realizar a atualização em tempo real é, por exemplo, verificar e modificar a posição (gerar as amostras do sinal binaural) a cada amostra do sinal mono-aural, isto é, a cada amostra é necessário calcular a HRTF da nova posição e gerar as amostras que seriam provenientes da fonte virtual.

A implementação do sistema em tempo-real necessita de equipamentos especiais, como, por exemplo, um detector de direção que indica ao sistema em que direção deve estar a fonte virtual a cada amostra do sinal mono-aural (detectando em que direção o sensor estaria sendo apontado), além de uma capacidade computacional maior. Para que não fosse necessário ter esse equipamento e já que, ao serem dadas as posições percorridas pela fonte virtual, o sistema *off-line* executa os mesmos procedimentos, implementou-se o sistema *off-line*, utilizando-se, primeiramente, as funções de transferência completas (sem simplificação). Mais tarde, modificou-se um pouco o algoritmo para utilizar as IPTFs simplificadas.

Criou-se, então, um sistema *off-line* pensando no sistema de tempo-real, isto é, a partir de um vetor contendo as posições por onde a fonte virtual deve passar durante o movimento, o sistema calcula as amostra de saída por canal, como se as posições fossem adquiridas em tempo-real.

Assim, implementou-se o sistema da Figura 3.9, só que com os parâmetros (coeficientes dos filtros, seus respectivos atrasos e os multiplicadores α , β e γ) variáveis, como pode ser visto na Figura 5.1.

O cálculo dos parâmetros α , β e γ é descrito pelas Equações (3.6), (3.7) e (3.8); já a determinação das funções de transferência que serão utilizadas segue os passos **1** e **2** do algoritmo da Seção 3.3.

O sistema mostrado nessa figura indica que todos os parâmetros mudam ao longo de um determinado caminho, mas essas mudanças não são feitas necessariamente na mesma taxa, isto é, dependendo da transição de posição efetuada, todos ou alguns parâmetros são modificados.

A forma de mudança dos parâmetros pode ser ilustrada pela Figura 5.2.

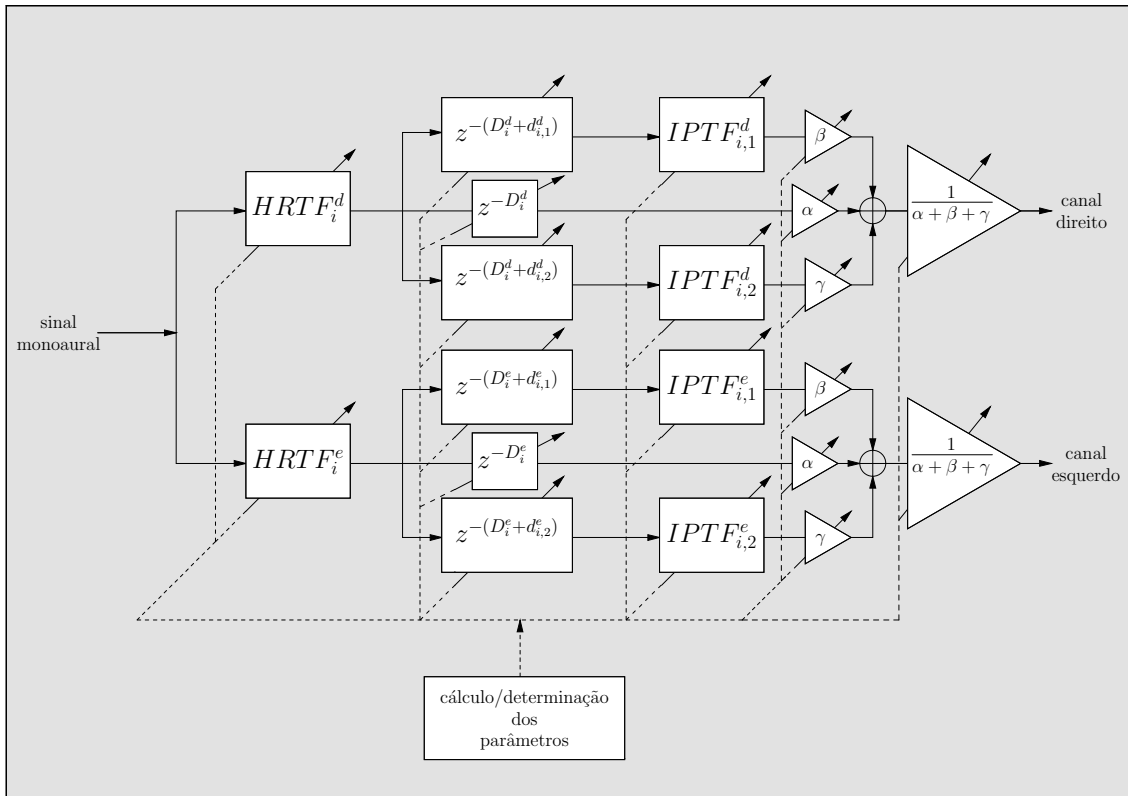


Figure 5.1: Diagrama em bloco do algoritmo de movimento.

Nessa figura, os pontos destacados são aqueles em que é necessário o cálculo dos parâmetros, isto é, são as posições da fonte virtual para a qual se deve obter uma amostra de saída para cada canal. Para esse exemplo, as distâncias entre essas posições estão um pouco exageradas, já que a taxa de amostragem espacial é igual à taxa de amostragem do sinal de áudio, o que resultaria em uma densidade maior de pontos, dependendo da velocidade de deslocamento da fonte.

Olhando para essa figura, pode-se perceber que na transição entre cada dois pontos existe uma variação dos parâmetros α , β e γ , já que a contribuição de cada IPTF está intimamente ligada à posição do ponto para o qual se quer calcular a HRTF final. A mudança desses parâmetros é feita sempre a cada nova posição. Já as mudanças dos coeficientes dos filtros, que acompanham as mudanças das funções de transferência utilizadas na interpolação, não ocorrem necessariamente a cada ponto.

Alguns exemplos de mudanças de parâmetros podem ser vistos na Figura 5.2, ao se verificarem as transições que ocorrem entre os pontos rotulados, pertencentes ao caminho da fonte virtual.

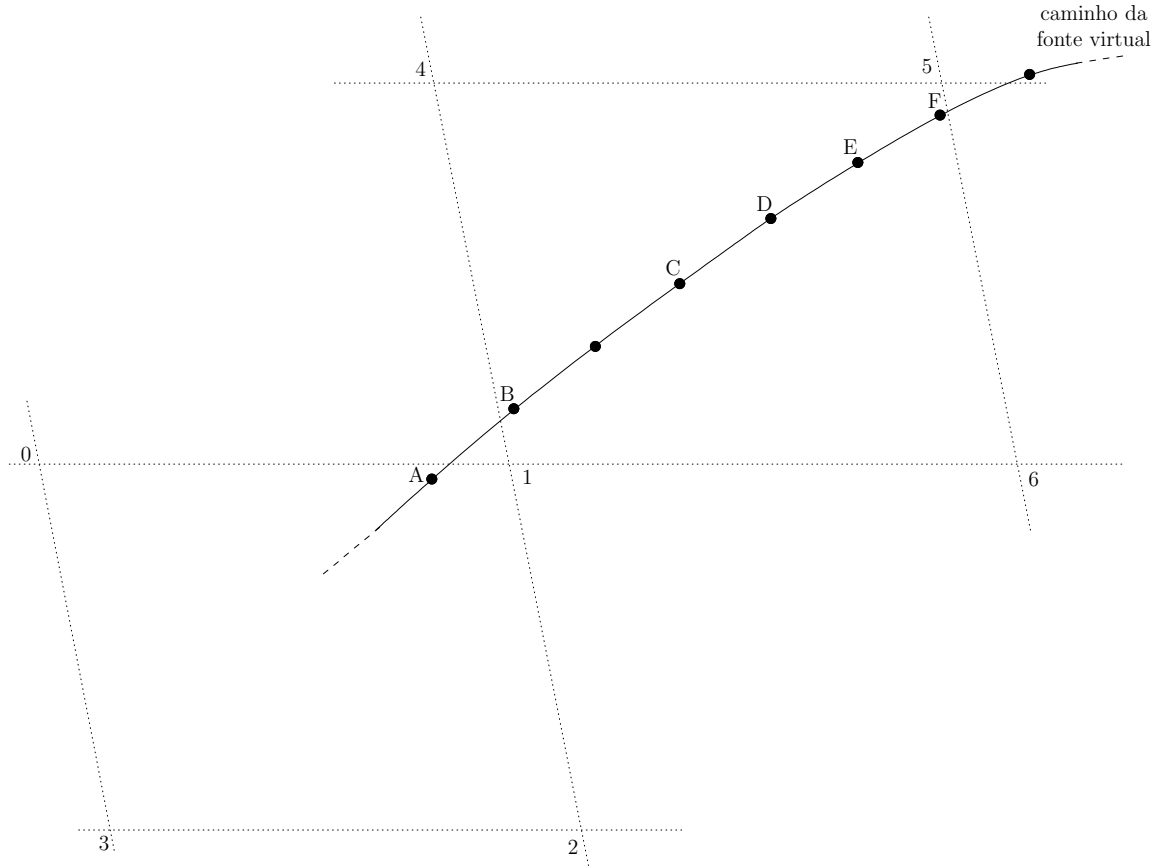


Figure 5.2: Exemplo de mudança de parâmetros durante o movimento. Os pontos numerados são aqueles para os quais foram medidas as HRIRs. Já os pontos sobre o caminho da fonte virtual são aqueles para os quais se quer aproximar suas HRTFs.

Na transição do ponto A para o ponto B, existe a mudança das duas IPTFs utilizadas para o cálculo da HRTF de A ($IPTF_{1,0}$ e $IPTF_{1,2}$) para as IPTFs utilizadas em B ($IPTF_{1,6}$ e $IPTF_{1,4}$). Do ponto C para o ponto D, a mudança é mais radical, pois tanto as IPTFs quanto a HRTF utilizadas mudam de $IPTF_{1,6}$, $IPTF_{1,4}$ e $HRTF_1$ para $IPTF_{5,4}$, $IPTF_{5,6}$ e $HRTF_5$, respectivamente. Já entre os pontos D e E e entre os pontos E e F, a única mudança que ocorre é a dos parâmetros α , β e γ . Note, também, que, sempre que houver mudança de uma das funções de transferência utilizadas, os atrasos relativos a elas e entre elas podem ser alterados.

5.1.1 Testes com IPTFs sem Simplificação

Alguns testes foram feitos para verificação do funcionamento do sistema de geração de movimento da fonte sonora virtual com as IPTFs sem simplificação. Esses testes foram realizados sobre dois arquivos de áudio do formato *WAVE*, um

que consiste no som de um helicóptero e outro com o som de uma abelha.

Esses arquivos foram adquiridos junto com a placa de som da Turtle Beach Montego II Quadzilla. Eles têm a duração de 2s e 0,5s, respectivamente, e, ao serem ouvidos, tem-se a sensação de que o helicóptero e a abelha estão parados (sem variação de intensidade ao longo do tempo) e posicionados no interior da cabeça do ouvinte (característica dos sinais mono-aurais ouvidos com fone).

Na realização dos testes, o caminho a ser percorrido pela fonte virtual foi criado a partir das coordenadas de alguns pares de **pontos-chave** (pares de início e final de um arco pertencente à esfera sobre o qual a fonte virtual se move a velocidade constante) e pelo tempo de percurso entre cada par de pontos-chave, que, juntamente com a taxa de amostragem (44,1kHz, para os sinais em questão), determinam as coordenadas da posição para cada instante.

Como os sinais de teste disponíveis têm duração fixa, caso o tempo de percurso para o caminho desejado seja maior, cria-se uma versão do sinal com repetição periódica até que o tempo de percurso seja atingido. Em caso contrário, isto é, tempo de percurso menor que o tempo do sinal utilizado, trunca-se o arquivo no tempo desejado.

Os testes mostraram o bom funcionamento do método de interpolação utilizando as IPTFs sem simplificação, obtendo-se sinais com vários tipos de movimento, desde os mais simples em que a fonte virtual dá uma volta no entorno da cabeça, até os mais complexos que fazem movimentos de subir e descer enquanto realizam, também, movimentos de rotação.

Esse fato mostrou que se pode utilizar a interpolação com as IPTFs e que, obtendo-se uma simplificação adequada, pode-se reduzir a complexidade computacional sem perda de qualidade.

5.2 Movimento com IPTFs Simplificadas—Solução Adotada

A geração do movimento da fonte virtual utilizando as IPTFs simplificadas pode ser feita, praticamente, da mesma forma que com as IPTFs sem simplificação. No entanto, alguns cuidados devem ser tomados para que as mudanças de parâmetros

não criem um transitório muito grande.

A simplificação deve aproximar bem as funções de transferência e, ainda, as aproximações devem preservar a semelhança entre funções relativas a pontos próximos. Deve-se também cuidar das variações dos atrasos do sistema de interpolação para que essa variação seja igual à do sistema gerado sem a simplificação para um mesmo caminho, garantindo o funcionamento do sistema.

Muitos problemas foram encontrados ao se fazer a simplificação das funções de transferência, como foi mostrado no Capítulo 4; após sua solução, conseguiu-se gerar um conjunto completo de IPTFs simplificadas que possibilitou implementar o sistema de geração de movimento com essas funções. Porém, durante os testes, encontraram-se transitórios indesejáveis, que apareciam no instante de chaveamento dos filtros (veja Figura 5.3). Vale dizer que ao se utilizarem funções sem simplificação

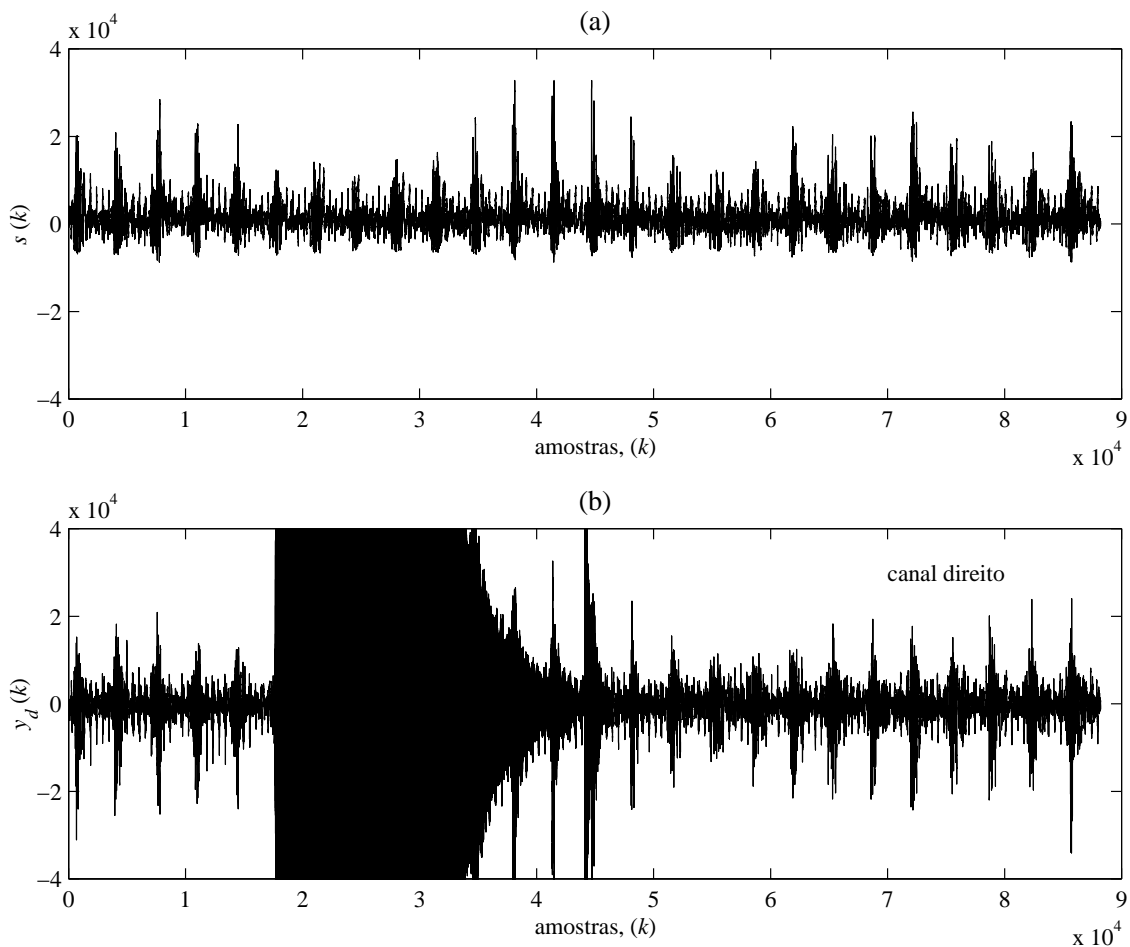


Figure 5.3: Efeito do transitório dos filtros: sinal mono-aural (a) e canal direito sem supressão do transitório (b).

não foram notados esses transitórios. Esse fato se deve à coerência natural entre as HRTFs medidas de pontos adjacentes. A similaridade dos filtros correspondentes resulta em estados muito parecidos, não favorecendo a ocorrência de transitórios nas transições de coeficientes dos filtros. Essa relação entre HRTFs adjacentes é ignorada quando a interpolação é feita com as IPTFs simplificadas, já que as IPTFs são geradas independentemente umas das outras.

Esses transitórios ficavam ainda mais evidentes quando a IPTF simplificada tinha pólos com parte real muito próxima a -1 (veja Figuras 5.3 e 5.4), pois demoravam muito a desaparecer. A Figura 5.4 mostra onde os pólos foram posicionados

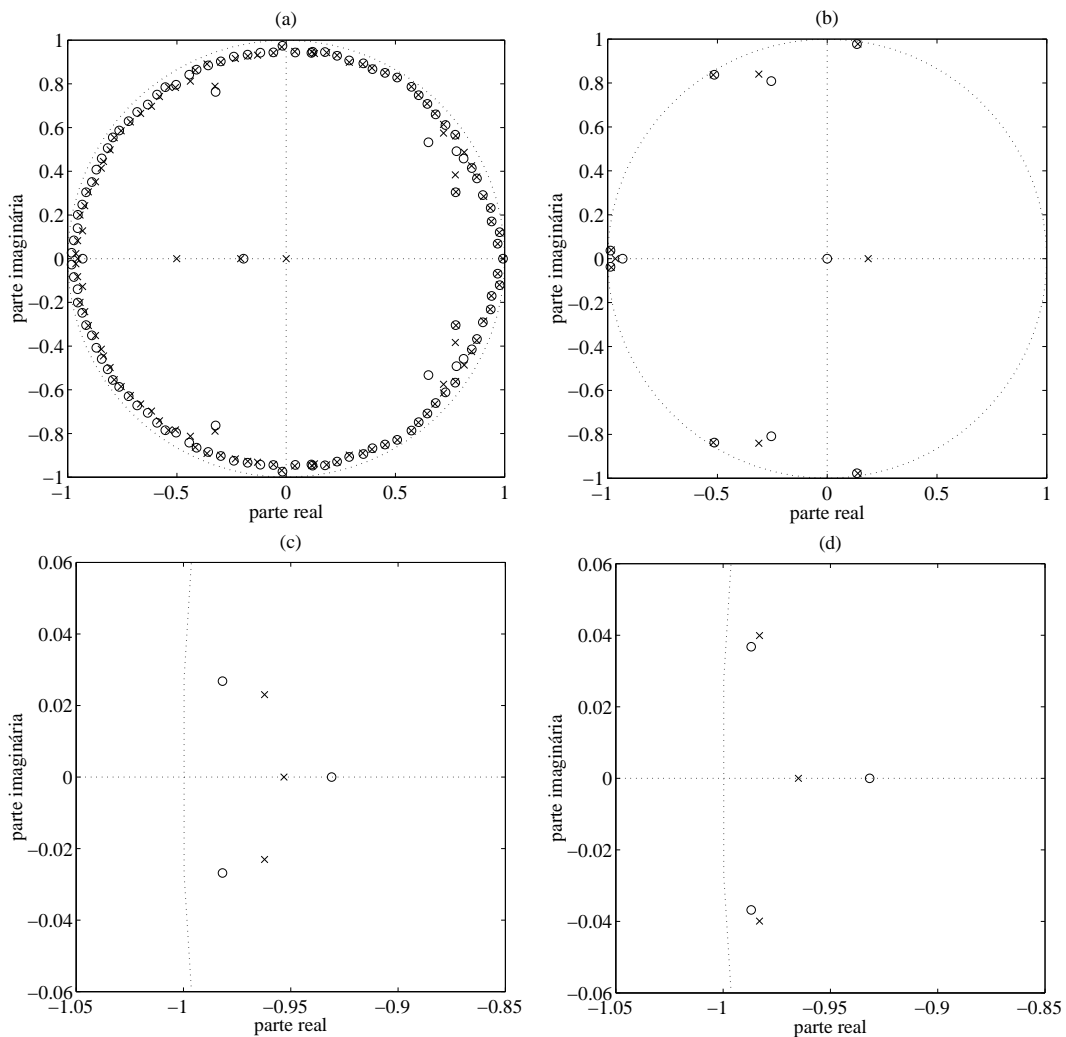


Figure 5.4: Exemplo de diagrama de pólos e zeros das IPTFs: IPTF sem simplificação (a) e com simplificação (b). Em (c) e (d), ampliação da região de (a) e (b), respectivamente, em que as singularidades têm parte real próxima a -1 .

ao se realizar a simplificação. Pode-se notar que os pólos ficaram mais próximos do

círculo unitário, o que tornou o sistema mais lento. Além disso, o pólo real próximo a -1 é responsável pelo comportamento oscilatório do sistema; este comportamento, aliado à lentidão do sistema, acabou por gerar os transitórios observados quando do chaveamento dos filtros. O diagrama de pólos e zeros mostrado na Figura 5.4b corresponde ao filtro que cria o maior transitório da Figura 5.3b.

Pode-se perceber que esses pólos próximos a -1 apareceram no modelo de ordem reduzida por causa das variações bruscas presentes na região das altas frequências no modelo original: já que não foram utilizadas quaisquer formas de suavização de espectro na obtenção do conjunto final de IPTFs, o modelo reduzido tentou preservá-las. Entretanto, como esses pólos não são indispensáveis do ponto de vista do que se quer modelar, para se obter um transitório mais rápido pensou-se em reduzir a influência dos pólos próximos a -1 , delocando-os para mais próximo da origem diminuindo seus módulos. Nas Figuras 5.5a e 5.5b, têm-se os diagramas de pólos e zeros sem e com os pólos deslocados em direção à origem, respectivamente.

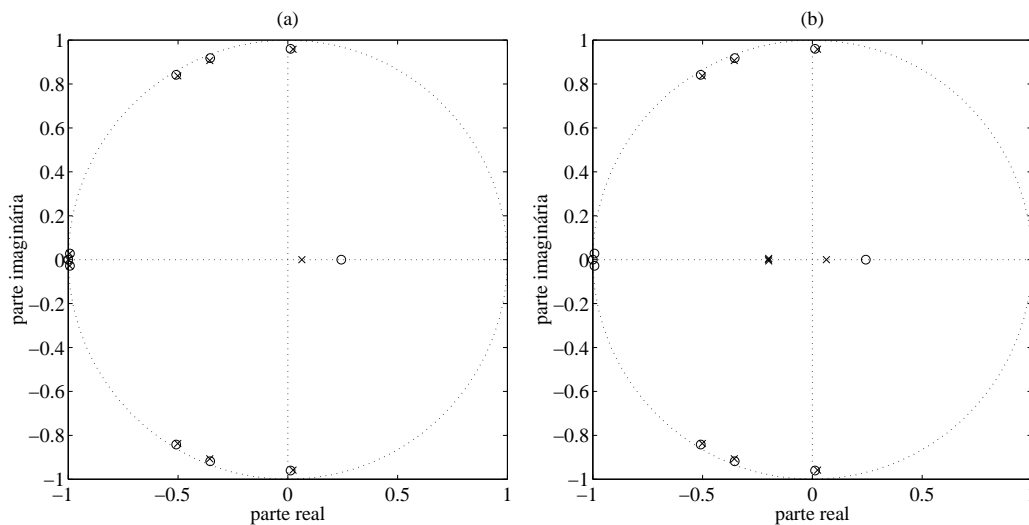


Figure 5.5: Deslocamento dos pólos das IPTFs em direção à origem: modelo original (a) e com a redução dos módulos dos pólos com parte real próxima a -1 (b).

Ao se utilizar esse procedimento, os problemas com os transitórios são reduzidos, mas não em alguns casos. Na Figura 5.6, pode-se verificar esse fato. O instante em que se encontra o transitório é o mesmo em que há a troca dos coeficientes para os relativos ao filtro com a constelação de pólos e zeros da Figura 5.5b. Note que o transitório na Figura 5.3b é bem mais evidente e lento que o da Figura 5.6b. Apesar

disso, essa não é a solução para esse problema.

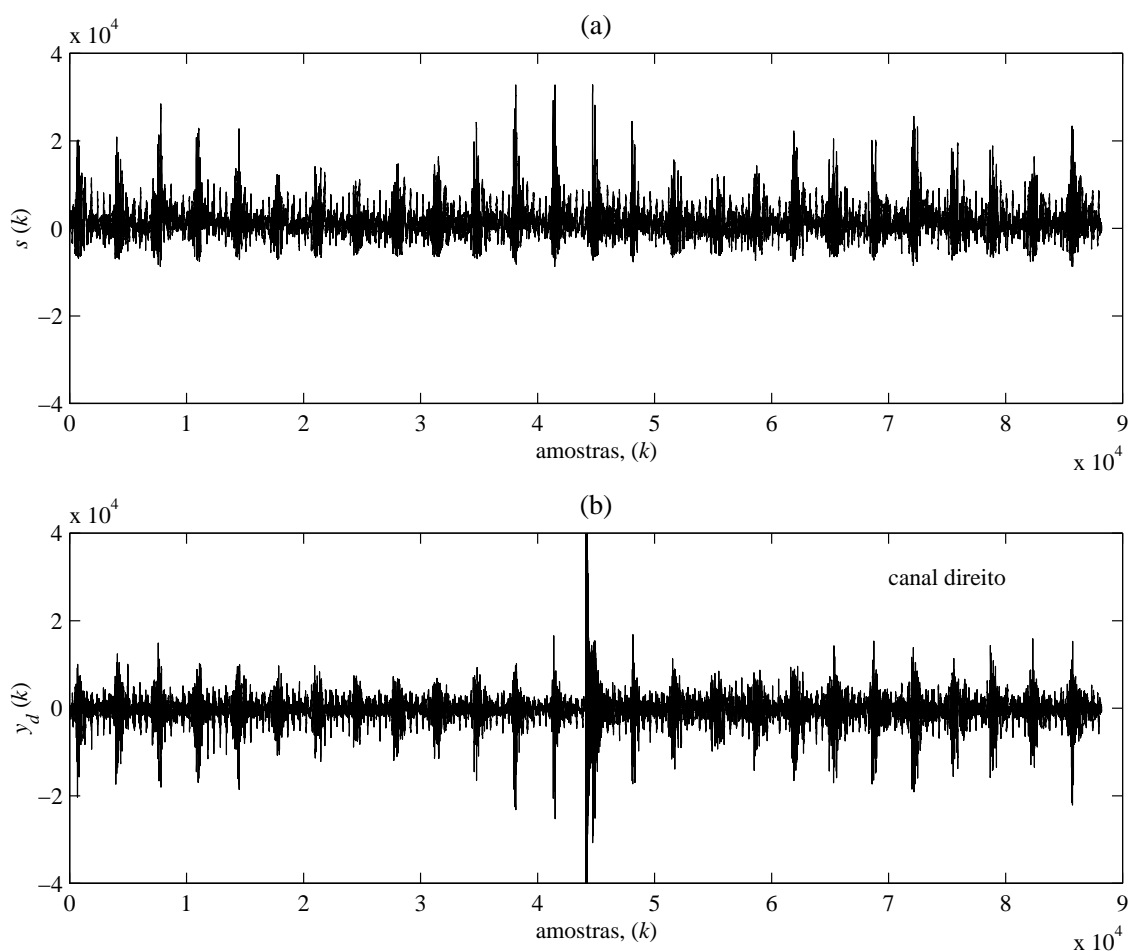


Figure 5.6: Efeito do transitório dos filtros com redução dos módulos dos pólos próximos a -1 : sinal mono-aural (a) e canal direito fazendo-se o deslocamento dos pólos (b).

É bastante conhecido o problema de se chavear entre os coeficientes de dois filtros, principalmente daqueles que necessitam das saídas anteriores no cálculo da próxima amostra de saída (IIR), como é o caso das IPTFs simplificadas. Esse problema é mais grave quando os coeficientes do filtro novo não são muito parecidos com os do filtro anterior.

Uma forma simples de se reduzir o transitório criado na mudança dos coeficientes é o método de *chaveamento de saída*, mencionado em [25], que consiste em se calcularem as saídas anteriores para o filtro IIR novo e, no instante da mudança, atualizar também o estado da saída¹. Esse cálculo pode ser feito a cada amostra,

¹Chamou-se de estado de saída as saídas anteriores necessárias ao cálculo da amostra de saída

devendo-se, para isso, ter o conhecimento dos coeficientes do próximo filtro, ou no momento da troca, sendo então necessário somente um determinado número de entradas anteriores ao instante da troca.

Como, no sistema *on-line* de geração de som tridimensional, saber a posição futura pode inserir um atraso do som com relação a posição (a não ser que se faça uma estimativa da próxima posição), optou-se pelo cálculo das saídas anteriores no momento da troca dos coeficientes. E para eliminar o transitório, mostrou-se suficiente utilizar o número de entradas anteriores igual à ordem das IPTFs simplificadas. Na Figura 5.7, tem-se o resultado obtido com esse procedimento para o canal direito (Figura 5.7b) realizando-se o mesmo movimento que gerou os transitórios da Figura 5.3b.

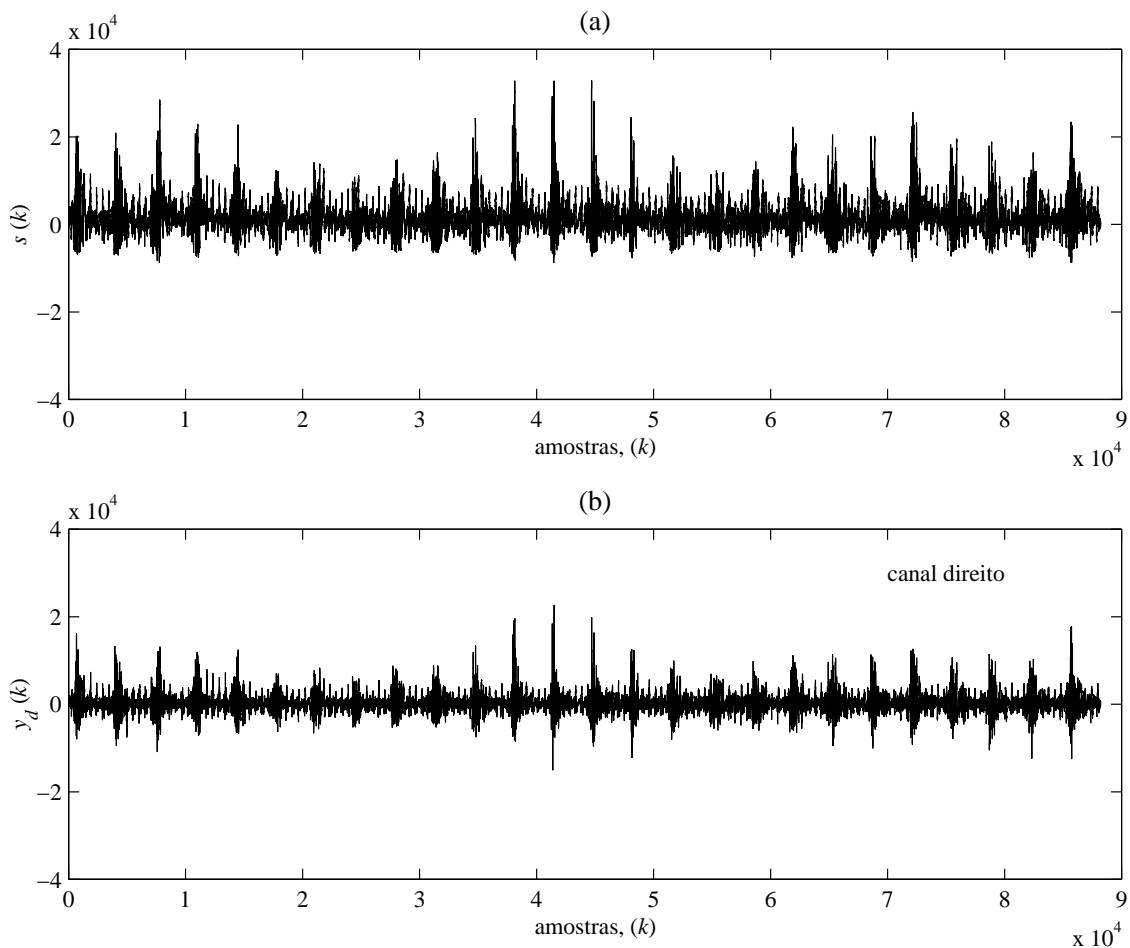


Figure 5.7: Exemplo de sinal com a supressão do transitório: sinal mono-aural (a) e canal direito com supressão dos transitórios (b).

atual do filtro IIR

Os mesmos caminhos realizados com as IPTFs sem simplificação foram feitos com as simplificadas fazendo-se a supressão dos transitórios. O resultado obtido com IPTFs de ordem 10 não apresentaram diferenças com relação ao obtido com as IPTFs sem simplificação. Isso significa que com uma forma de simplificação eficiente pode-se reduzir bastante a complexidade computacional do sistema de geração de som tridimensional realizando a interpolação com as IPTFs.

Chapter 6

Conclusão

Nessa tese, foi desenvolvido um estudo de técnicas de geração de áudio tridimensional que culminou na proposição da técnica de interpolação das HRTFs utilizando modelos de ordem reduzida para as IPTFs. Assim, conseguiu-se uma forma de interpolação que torna possível a redução da complexidade computacional dos sistemas de geração de movimento.

Utilizaram-se algumas técnicas de redução de ordem de modelos para gerar um conjunto de IPTFs contendo todas as IPTFs necessárias ao sistema de geração do movimento. Muitos problemas foram encontrados pelas limitações dos próprios métodos de simplificação, como foi exposto no Capítulo 4. Por fim, conseguiu-se gerar com sucesso um conjunto completo de IPTFs através da aplicação da técnica de redução de modelos balanceados IIR, descrita na Subseção 4.1.1, à razão de HRTFs (filtro IIR) que compõe cada IPTF.

Apesar de essa técnica ter levado quase sempre a uma solução melhor que as outras para os modelos simplificados, algumas funções que não puderam ser simplificadas tiveram que ser substituídas pelo inverso do seu modelo inverso reduzido. No único caso em que as duas soluções não eram possíveis, utilizou-se um dos métodos de suavização complexa do espectro, mencionado na Subseção 4.1.3.3.

Para esse conjunto de IPTFs, foi possível gerar vários exemplos de movimento com dois tipos de sons: o som de helicóptero e o som de uma abelha.

Alguns pontos podem ser revisitados futuramente para se tentar uma melhoria na redução dos modelos das IPTFs:

- a suavização utilizando a média geométrica (veja Subseção 4.2.3) pode facilitar

o cálculo do modelo de ordem reduzida, além de eliminar o efeito dos pólos reais próximos de -1 que aparecem no modelo e deixam o transitório lento;

- pode-se utilizar o método de projeto de filtros através da deformação da escala de frequências (*frequency warping*), mencionado em [14], para obter as HRTFs de ordem reduzida e, ainda, verificar se esse método pode ser estendido para as IPTFs.

Essas modificações podem trazer bastante benefício à simplificação das funções de transferência. A suavização utilizando a média geométrica, por exemplo, pode vir a ser bastante importante, já que não se conseguiu um conjunto de IPTFs com ordem reduzida completo ao se realizar a redução de ordem sobre as IPTFs suavizadas e o procedimento de suavização faz o modelo representar as características mais importantes das funções de transferência. Além disso, como o projeto de filtros com *frequency warping* é considerado em [14] como a melhor forma de modelar as HRTFs (menor ordem e atingindo melhor qualidade), essa técnica pode melhorar bastante os modelos das IPTFs.

Atualmente, tem sido bastante incentivado pelos órgãos internacionais a pesquisa sobre assuntos ligados ao áudio tridimensional. Um bom exemplo disso são os encontros anuais da *Audio Engineering Society* (AES), alguns dos quais têm tido como tema principal o som espacial (*surround* e tridimensional). Esses eventos têm-se voltado a um amplo campo de estudo, que envolve muitos problemas de implementação de sistemas de gravação/reprodução de som espacial. Por exemplo, ainda há muita dificuldade em realizar a gravação do som no sistema de 5.1 canais.

Outro campo que está começando a aparecer é o que trata da interação do vídeo com o áudio. Neste campo, muita coisa ainda pode ser feita, principalmente no sentido da sincronização do som com efeitos de tridimensionalidade e a imagem. As salas de projeção de filmes tridimensionais são um exemplo de aplicação prática envolvida neste campo.

Alguns pontos ligados ao conteúdo desse trabalho que podem vir a ser explorados futuramente:

- a modelagem de ambiente virtual, com o intuito de se ter, além da sensação de localização, a sensação do ambiente em que o som estaria sendo ouvido;

- separação e remixagem de instrumentos, localizando-os de forma a dar uma sensação, por exemplo, de que um determinado músico está se deslocando de um lado para outro no palco;
- remasterização do som de um filme de forma que, através da informação de posição de um determinado objeto da cena em questão, possa-se criar o movimento do som emitido por aquele objeto em movimento.

Bibliography

- [1] PULKKI, V., “Virtual Sound Source Positioning Using Vector Base Amplitude Panning”, *Journal of the Audio Eng. Soc.*, v. 45, n. 6, pp. 456–466, June 1997.
- [2] EARGLE, J., *Sound Recording*. Van Nostrand Reinhold Company, 1976.
- [3] KIRKEBY, O., NELSON, P. A., HAMAD, H., “The ‘Stereo Dipole’-A Virtual Source Imaging System Using Two Closely Spaced Loudspeakers”, *Journal of the Audio Eng. Soc.*, v. 46, n. 5, pp. 387–395, May 1998.
- [4] GARDNER, B., MARTIN, K., *HRTF Measurements of a KEMAR Dummy-Head Microphone*, Technical Report 280, MIT Media Lab., Cambridge, MA, May 1994.
- [5] SAVIOJA, L., *Modeling Techniques for Virtual Acoustics*. Ph.d. thesis, Helsinki University of Technology, Department of Computer Science and Engineering, Telecommunications Software and Multimedia Laboratory Espoo, Finland, December 1999.
- [6] BASS, H. E., BAUER, H. J., EVANS, L. B., “Atmospheric Absorption of Sound: Analytical Expression”, *The Journal of the Acoustical Society of America*, v. 52, n. 3, pp. 821–825, February 1972. Part 2.
- [7] MACKENZIE, J., HUOPANIEMI, J., VALIMÄKI, V., *et al.*, “Low-Order Modeling of Head-Related Transfer Functions using Balanced Model Truncation”, *IEEE Signal Processing Letters*, v. 4, n. 2, pp. 39–41, February 1997.
- [8] HAYKIN, S., *Communication Systems*. John Wiley & Sons, Inc., 1994.
- [9] LORHO, G., HUOPANIEMI, J., ZACHAROV, N., *et al.*, “Efficient HRTF Synthesis using an Interaural Transfer Function Model”. In: *Proceedings of*

- X European Signal Processing Conference (EUSIPICO'2000)*, pp. 2213–2216, EUROSIP, Tampere, Finland, September 2000.
- [10] BROWN, C. P., DUDA, R. O., “A Structural Model for Binaural Sound Synthesis”, *IEEE Transactions on Speech and Audio Processin*, v. 6, n. 5, pp. 476–488, September 1998.
- [11] TAN, C.-J., GAN, W.-S., “Direct Concha Excitation for the Introduction of Individualized Hearing Cues”, *Journal of the Audio Eng. Society*, v. 48, n. 7/8, pp. 642–653, July/August 2000.
- [12] ZHANG, M., TAN, K.-C., ER, M. H., “Three-Dimensional Sound Synthesis Based on Head-Related Transfer Functions”, *Journal of the Audio Eng. Soc.*, v. 46, n. 10, pp. 836–844, October 1998.
- [13] HARTUNG, K., BRAASCH, J., STERBING, S. J., “Comparision of Different methods for the Interpolation of Head-Related Transfer Function”. In: *The Proceedings of the AES 16th International Conference*, pp. 319–329, AES, Rovaniemi, Finland, April 1999.
- [14] HUOPANIEMI, J., ZACHAROV, N., KARJALAINEN, M., “Objective and Subjective Evaluation of Head-related Transfer Function Filter Design”, *Journal of the Audio Eng. Soc.*, v. 47, n. 4, pp. 218–239, April 1999.
- [15] PULKKI, V., KARJALAINEN, M., HUOPANIEMI, J., “Analyzing Virtual Sound Sources Attributes Using a Binaural Auditory Model”, *Journal of the Audio Eng. Soc.*, v. 47, n. 4, pp. 203–216, April 1999.
- [16] KISTLER, D. J., WIGHTMAN, F. L., “A Model of Head-Related Transfer Functions Based on a Principal Components Analysis and Minimum Phase Reconstruction”, *Journal of the Acoustic Society of America*, v. 91, n. 3, pp. 1637–1647, March 1992.
- [17] KULKARNI, A., ISABELLE, S. K., COLBURN, H. S., “On The Minimum-Phase Approximation of Head-Related Transfer Functions”. In: *IEEE Workshop on Applications of the Signal Processing to Audio and Acoustics*, IEEE, New Paltz, New York, October 1995.

- [18] KULKARNI, A., ISABELLE, S. K., COLBURN, H. S., “Sensitivity of Human Subjects to Head-Related Transfer-Function Phase Spectra”, *Journal of the Acoustic Society of America*, , Submitted for publication (1997).
- [19] BELICZYNSKY, B., KALE, I., CAIN, G. D., “Approximation of FIR by IIR Digital Filters: An Algorithm Based on Balanced Model Reduction”, *IEEE Transactions on Signal Processing*, v. 40, n. 3, pp. 532–542, March 1992.
- [20] DINIZ, P. S. R., *Adaptive Filtering: Algorithms and Practical Implementation*. Kluwer Academic Publishers, 1997.
- [21] GOOCH, R. P., *Adaptive Pole-Zero Filtering: The Equation-Error Approach*. Ph.D. thesis, Stanford University, June 1983.
- [22] HATZIANTONIOU, P. D., MOURJOPOULOS, J. N., “Generalized Fractional-Octave Smoothing of Audio and Acoustic Response”, *Journal of the Audio Eng. Soc.*, v. 48, n. 4, pp. 259–280, April 2000.
- [23] MATLAB, “Signal Processing Toolbox”, Math Works Inc.
- [24] LAUB, A. J., HEATH, M. T., PAIGE, C. C., *et al.*, “Computation of System Balancing Transformations and Other Applications of Simultaneous Diagonalization Algorithms”, *IEEE Transactions on Automatic Control*, v. AC-32, pp. 115–122, 1987.
- [25] VALIMÄKI, V., LAAKSO, T. I., “Suppression of transients in Time-Varying recursive Filters for Audio Signals”. In: *in Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP’98)*, v. 6, pp. 3569–3572, IEEE, Seattle, Washington, May 1998.

Apêndice A

Método de Medição de Funções de Transferência

Sendo $x(k)$ uma seqüência pseudo-aleatória de período N e com amostras independentes, pode-se obter a resposta impulsiva $h(k)$ de um determinado sistema, primeiramente, aplicando-se à entrada desse sistema o sinal $x'(k)$ definido por:

$$x'(k) = \begin{cases} x(k+N), & \text{se } k = -N, -N+2, \dots, -1 \\ x(k), & \text{se } k = 0, 1, \dots, N-1 \\ 0, & \text{para outros valores de } k, \end{cases} \quad (\text{A.1})$$

o que equivale a dois períodos da seqüência $x(k)$. Como resultado desse procedimento, obtém-se uma seqüência de $2N$ amostras, das quais as amostras de $k = -N$ a $k = -1$ são descartadas com o intuito de eliminar o transitório. Assim, pode-se dizer que

$$y(k) = h(k) * x'(k), \quad (\text{A.2})$$

com $k = 0, 1, \dots, N-1$, é a seqüência de saída do sistema.

Definindo-se a seqüência $y'(k)$ como sendo a duplicação da seqüência $y(k)$ da forma

$$y'(k) = \begin{cases} y(k+N), & \text{se } k = -N, -N+2, \dots, -1 \\ y(k), & \text{se } k = 0, 1, \dots, N-1 \\ 0, & \text{para outros valores de } k, \end{cases} \quad (\text{A.3})$$

calculando-se a correlação entre $y'(k)$ e $x(k)$ e descartando-se as amostras de $k = -N$ a $k = -1$ dessa correlação, tem-se

$$h'(k) = h(k) * x'(k) * x(-k), \quad (\text{A.4})$$

com $k = 0, 1, \dots, 2N - 1$.

Como a seqüência $x(k)$ aproxima um ruído branco, pode-se dizer que

$$h'(k) = h(k) * \sigma_x^2 \delta(k) = \sigma_x^2 h(k), \quad (\text{A.5})$$

onde $k = 0, 1, \dots, 2N - 1$ e $\delta(k)$ é o impulso unitário definido por

$$\delta(k) = \begin{cases} 1, & \text{se } k = 0 \\ 0, & \text{se } k \neq 0 \end{cases}. \quad (\text{A.6})$$

Dessa forma, pode-se dizer que

$$h'(k) \propto h(k). \quad (\text{A.7})$$