

AURALIZAÇÃO DE SALAS UTILIZANDO WAVELETS PARA A MODELAGEM DAS HRTFs

Julio C. B. Torres Mariane R. Petraglia Roberto A. Tenenbaum
Programa de Eng. Elétrica Programa de Eng. Elétrica Programa de Eng. Mecânica
julio.torres@ufrj.br mariane@lps.ufrj.br roberto@serv.com.ufrj.br
COPPE - Universidade Federal do Rio de Janeiro
Rio de Janeiro, RJ, Brasil,
CP 68503 - CEP 21945-970

Resumo – Neste trabalho é apresentada uma nova estrutura para modelagem das funções de transferência relacionadas à cabeça (HRTFs). A modelagem é realizada utilizando-se um sistema adaptativo multitaxa, implementado através de uma transformada wavelet e filtros esparsos. A seleção da base mais apropriada da wavelet para a modelagem das HRTFs é o principal foco do trabalho. São apresentadas as bases, o critério de seleção e a influência de cada base na redução do número de coeficientes do modelo, visando sua aplicação em um sistema de auralização de salas.

Abstract – In this work is presented a new structure for modeling the Head Related Transfer Functions (HRTFs). The model is obtained by using an adaptive multirate system, implemented through wavelet transforms and sparse filters. The wavelet selection is the focus of this work. We present the wavelet basis, the criteria for transform selection and its influence in the reduction of the coefficients of the model, in order to be implemented in an auralization system.

I. INTRODUÇÃO

Os sistemas de realidade virtual vêm recebendo grande atenção nos últimos anos graças à evolução tecnológica e ao crescente número de aplicações onde a percepção espacial auditiva é necessária. Dentre estas aplicações pode-se citar os jogos eletrônicos para computadores pessoais, videogames, simuladores em parques de diversão, simuladores de vôo e sistemas de som “tridimensional” para salas de exibição de filmes, além de uma ferramenta fundamental de auxílio ao projeto acústico de salas.

Independente do método utilizado para simular a propagação do som pelo ambiente acústico, como por exemplo os métodos dos raios [1], [2], das imagens [3] e da propagação de energia [], qualquer sistema de simulação acústica de salas que possua um módulo de auralização requer uma modelagem do receptor (ouvinte).

A modelagem do receptor é a etapa mais importante no processo de auralização. Ela consiste em simular a influência do ouvinte no ambiente acústico e a percepção espacial do som que chega aos seus ouvidos. Para isso são obtidas as funções de transferência relacionadas à cabeça - HRTFs (“Head Related Transfer Functions”) no domínio da frequência ou HRIRs (“Head Related Impulse Responses”) no domínio do tempo.

As HRTFs são funções direcionais que dependem apenas da direção e da distância entre a fonte sonora e o receptor. Nessas funções estão incluídas as diferenças interaurais de tempo, de nível de pressão sonora e variações espectrais do som percebido pelo ouvinte, em função da direção da fonte. A

obtenção dessas funções pode ser realizada através do uso de uma cabeça artificial [4], [5], que simula um torso humano, ou através da inserção de um par de pequenos microfones nos canais auditivos de uma pessoa [6].

Os principais problemas na utilização direta desse conjunto de dados são a grande quantidade de funções e o comprimento (duração) das suas respostas impulsivas. A audição humana possui uma resolução angular para identificação da posição de uma fonte sonora que varia de 5 a 20 graus [7]. Para manter esta precisão, um conjunto de HRTFs deve possuir em média 700 funções para cada ouvido, com aproximadamente 500 coeficientes cada, a uma taxa de 44.1 KHz.

Apesar de novas tentativas terem sido realizadas no sentido de reduzir o conjunto de dados representativo das HRTFs [8], [9], [10], [11], a utilização de tais modelos em sistemas com auralização ainda possui diversos contratempos. Um deles é a própria implementação da HRTF que, mesmo alcançando um elevado grau de redução dos coeficientes, possui um custo computacional alto. Outro fator negativo é a própria redução dos coeficientes, que pode remover características importantes para uma percepção realística da direção da fonte sonora e, no caso da auralização, do ambiente acústico.

A não redução dos coeficientes torna um sistema de auralização extremamente lento, devido à grande quantidade de raios de chegada ao receptor, pois cada raio precisa ser convolvido com a HRTF da direção correspondente. Caso o sistema de auralização processe o som através da geração e convolução de respostas impulsivas binaurais (BIRs - “Binaural Impulse Responses”) com o sinal da fonte, uma BIR com duração de 2 segundos, possuirá aproximadamente 88.000 coeficientes a uma taxa de amostragem de 44.1 KHz. Portanto são exigidos métodos mais eficientes de implementação da convolução além de processadores extremamente rápidos e sistemas dedicados. Com tais contratempos o sistema de auralização passa a ter um custo elevado, restringindo seu uso a pessoas especializadas na área e dificultando sua popularização.

Neste trabalho apresentaremos um tipo de modelagem cujo objetivo é reduzir o número de coeficientes ao máximo sem que as características das HRTF sejam comprometidas. Para isso propomos a utilização transformadas wavelets, que mantêm as características temporais das HRTFs (ou HRIRs) e possibilitam uma resolução diferenciada para cada faixa de frequência.

II. MODELAGEM ATRAVÉS DE BANCOS DE FILTROS E FILTRAGEM ADAPTATIVA

Uma alternativa às implementações tradicionais é a modelagem das HRTFs através de transformadas wavelets e filtros esparsos [12]. Nessa abordagem a HRTF é vista como um sistema FIR de resposta impulsional desconhecida. O modelo desta HRTF é obtido através de filtragem adaptativa, onde os coeficientes do filtro adaptativo convergem para valores que produzam o menor erro médio quadrático entre os sinais de

saída e desejado do sistema.

A estrutura adaptativa pode ser observada na Figura 1, onde os coeficientes dos filtros adaptativos $G(z)$ são ajustados para obter o modelo do sistema FIR “desconhecido” $P(z)$, neste caso a HRTF. À medida em que o sinal $x(n)$ pas-

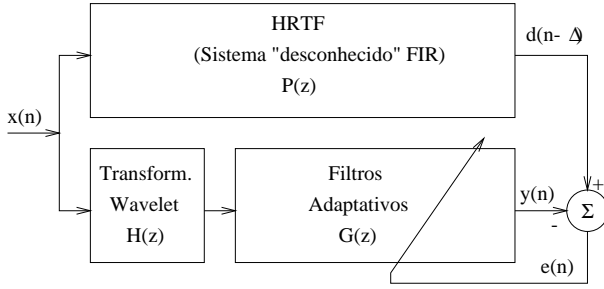


Figura 1. Estrutura para identificação da HRTF

sa através da HRTF, produzindo o sinal desejado $d(n - \Delta)$ (atrasado devido à filtragem), o mesmo sinal $x(n)$ é decomposto, na parte inferior da estrutura, por uma transformada wavelet discreta, implementada através de um banco de filtros $H_m(z)$, onde $m = 0, 1, \dots, M - 1$, sendo M o número de sub-bandas adotado. O sinal transformado passa pelos filtros adaptativos esparsos $G_m(z^{L_m})$, onde L_m é o fator de esparsidade da sub-banda m , produzindo o sinal de saída $y(n)$. Este sinal é comparado com o sinal desejado gerando o sinal de erro $e(n)$, utilizado na convergência dos coeficientes de cada filtro adaptativo.

No nosso interesse está em obter os coeficientes que produzam o menor erro possível. O sinal $x(n)$ utilizado para adaptar os coeficientes de $G_m(z^{L_m})$ é um ruído branco de média zero e variância unitária, o que proporciona a convergência mais rápida e o menor erro médio quadrático (MSE) possível [12].

A wavelet discreta pode ser obtida se escolhermos adequadamente os coeficientes dos filtros protótipos $H^0(z)$ e $H^1(z)$, também chamados de *wavelet mãe*, que ao serem cascateados como na Figura 2, proporcionam uma decomposição em oitavas do sinal de entrada.

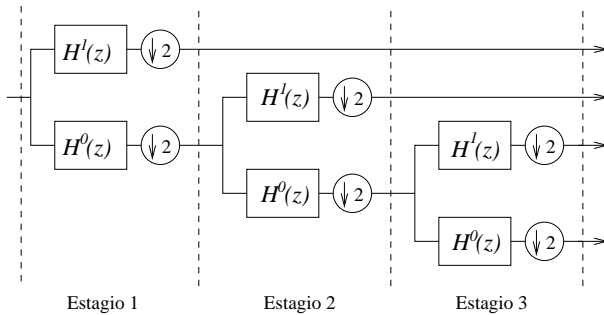


Figura 2. Estrutura em oitavas - implementação da wavelet.

Neste caso, os correspondentes filtros de análise $H_m(z)$ da Figura 3, obtidos do cascateamento de cada ramo da Figura 2, estão relacionados com os protótipos $H^0(z)$ e $H^1(z)$ através das equações (1) e (2):

$$H_0(z) = \prod_{k=0}^{J-1} H^0(z^{2^k}) \quad (1)$$

e

$$H_i(z) = H^1(z^{2^{J-i}}) \prod_{k=0}^{J-1-i} H^0(z^{2^k}), \quad (2)$$

para $i = 1, \dots, M - 1$ e $M = J + 1$, onde J representa o número de estágios da decomposição. O fator de esparsidade L_m dos filtros adaptativos correspondem aos fatores de decimação impostos pela sub-amostragem do sinal na Figura 2, dados por

$$L_m = \begin{cases} 2^{m+1}, & m = 0, 1, \dots, M - 2 \\ 2^m, & m = M - 1. \end{cases} \quad (3)$$

III. GENERALIZAÇÃO DA ESTRUTURA

Para considerarmos o caso de M sub-bandas, vejamos primeiro o caso onde há apenas duas sub-bandas ($M = 2$), na Figura 3.

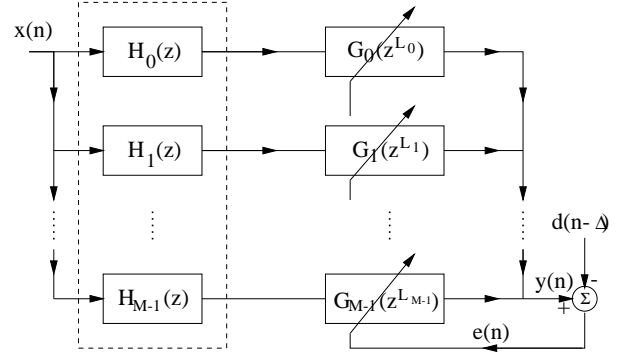


Figura 3. Estrutura adaptativa usando um banco de análise e filtros esparsos.

A função de transferência implementada pela estrutura de dois canais é dada por

$$R(z) = [G_0(z^2) \quad G_1(z^2)] \mathbf{H}_p(z^2) \begin{bmatrix} 1 \\ z^{-1} \end{bmatrix}, \quad (4)$$

onde $\mathbf{H}_p(z) = [H_{i,j}]$ é a matriz polifásica 2×2 do banco de análise [13].

A decomposição polifásica tipo 1 da função de transferência direcional é dada por

$$P(z) = [P_0(z^2) \quad P_1(z^2)] \begin{bmatrix} 1 \\ z^{-1} \end{bmatrix}. \quad (5)$$

Observa-se das equações (4) e (5), que a estrutura em sub-bandas pode modelar exatamente qualquer HRTF $P(z)$, se

$$[G_0(z^2) \quad G_1(z^2)] \mathbf{H}_p(z^2) = [P_0(z^2) \quad P_1(z^2)]. \quad (6)$$

tiver solução. Tal solução, geralmente, não pode ser obtida com filtros causais $G_0(z^2)$ e $G_1(z^2)$.

Porém, se fizermos

$$[G_0(z^2) \quad G_1(z^2)] = [P_0(z^2) \quad P_1(z^2)] \mathbf{F}_p(z^2), \quad (7)$$

com $\mathbf{F}_p(z^2)$ tal que

$$\mathbf{F}_p(z^2) \mathbf{H}_p(z^2) = z^{-\Delta} \mathbf{I}, \quad (8)$$

a estrutura com dois canais passa a ter função de transferência $P(z)z^{-\Delta}$.

As matrizes $\mathbf{H}_p(z)$ e $\mathbf{F}_p(z)$ que satisfazem a equação (8) correspondem, respectivamente, às matrizes polifásicas dos bancos de análise $H_i(z)$ e de síntese $F_i(z)$ de um sistema multitaxa com reconstrução perfeita, sendo $\mathbf{H}_p(z) = [H_{i,j}(z)]$ e $\mathbf{F}_p(z) = [F_{j,i}(z)]$, tal que

$$H_i(z) = H_{i,0}(z^2) + z^{-1} H_{i,1}(z^2) \quad (9)$$

$$F_i(z) = z^{-1}F_{0,i}(z^2) + F_{1,i}(z^2), \quad (10)$$

respectivamente.

Portanto, se utilizarmos um banco de análise que possua reconstrução perfeita a filtros adaptativos cuja ordem satisfaça a equação (7), então a estrutura com duas bandas implementará exatamente qualquer sistema FIR $P(z)$, se permitirmos a introdução de um atraso constante Δ na saída da estrutura.

Na equação (7), o número de coeficientes diferentes de zero em cada sub-filtro adaptativo $G_i(z^2)$ deve ser no mínimo $K_i = \lceil (N_P + N_{F_i})/2 \rceil - 1$, onde N_P e N_{F_i} são os comprimentos do sistema desconhecido $P(z)$ e do i -ésimo filtro de síntese $F_i(z)$, respectivamente.

Os resultados obtidos para a estrutura com duas bandas podem ser estendidos para o caso de M bandas. Desta forma, pode-se obter uma estrutura generalizada para qualquer número de sub-bandas, desde que os requisitos de reconstrução perfeita e número mínimo de coeficientes adaptativos continuem sendo satisfeitos. Como não há restrição às larguras de banda dos filtros do banco de análise, pode-se obter bancos com diferentes faixas de passagem (transformadas não-uniformes [14]).

Para atualizar os coeficientes dos sub-filtros utilizou-se um algoritmo LMS (“Least Mean Square”) normalizado. Definindo o sinal $x_m(n)$ como o sinal na saída do m -ésimo filtro de análise e $g_{m,k}$ como o k -ésimo coeficiente do sub-filtro $G_m(z^{L_m})$, a equação de atualização dos coeficientes dos filtros esparsos que minimiza o erro médio quadrático é dada por

$$g_{m,k}(n) = g_{m,k}(n-1) + \mu_m(n)e(n)x_m(n-L_mk), \quad (11)$$

para $m = 0, 1, \dots, M-1$ e $k = 0, 1, \dots, K_m-1$.

Na equação (11), o sinal de erro $e(n)$ é dado por

$$e(n) = d(n-\Delta) - y(n), \quad (12)$$

onde $d(n)$ é a resposta desejada do sistema desconhecido, $y(n)$ é a saída da estrutura em sub-bandas e Δ é o atraso necessário para compensar a filtragem:

$$\Delta = \frac{N_{H_{\max}} + N_{H_{\min}}}{2} - M + 1, \quad (13)$$

sendo $N_{H_{\max}}$ e $N_{H_{\min}}$ os comprimentos do maior e do menor filtro de análise $H_m(z)$ associados à wavelet, respectivamente. O passo de adaptação de cada sub-filtro é normalizado pela potência de cada sinal transformado $x_m(n)$, ou seja:

$$\mu_m(n) = \mu/p_m(n), \quad (14)$$

onde $p_m(n)$ é uma estimativa da potência de $x_m(n)$.

A resposta impulsional $r(n)$ da estrutura pode ser obtida pela soma das convoluções de cada filtro de análise $h_m(n)$ com o respectivo filtro esparso $g_m(n)$, mostrado a seguir:

$$r(n) = \sum_{m=0}^{M-1} h_m(n) * g_m(n) \quad (15)$$

Após o processo de adaptação, os coeficientes de $G(z)$ convergem para valores que fazem com a resposta impulsional $r(n)$ da estrutura seja praticamente a da HRTF $p(n)$ que se deseja modelar, ou seja, $P(z) = R(z)$ na frequência.

IV. SELEÇÃO DA TRANSFORMADA

Nesta seção nosso objetivo é buscar a melhor base para a wavelet e qual esquema de decomposição dos filtros protótipos que permitam reduzir os coeficientes modelados, sem que haja degradação considerável nas HRTFs.

Investigou-se dois tipos de decomposição em oitavas. Primeiro, a decomposição do protótipo passa-baixas, denominada *estrutura direta* cuja estrutura é mostrada na Figura 2. Depois, a decomposição do protótipo passa-altas, denominada *estrutura invertida*, onde as posições dos filtros $H^0(z)$ e $H^1(z)$ são invertidas na Figura 2. Nesse caso, teremos uma distribuição do espectro diferente da anterior, onde os filtros possuem bandas de passagem menores à medida em que a frequência aumenta.

Se os protótipos forem filtros de “meia-banda”, teremos uma distribuição em oitavas *tradicional* do espectro, no caso da estrutura sem inversão. Porém, as bases wavelet geralmente não possuem uma simetria de espectro, proporcionando um fatiamento não-uniforme.

Essa distribuição não-uniforme do espectro é bastante favorável à modelagem das HRTFs e à redução dos coeficientes, uma vez que as HRTFs requerem um elevado nível de detalhe nas altas frequências e a manutenção das características inter-aurais nas baixas frequências. Um exemplo desta distribuição do espectro pode ser observada na Figura 4, onde o filtro passa-altas do protótipo bio5-3, da tabela I, foi decomposto até o terceiro estágio.

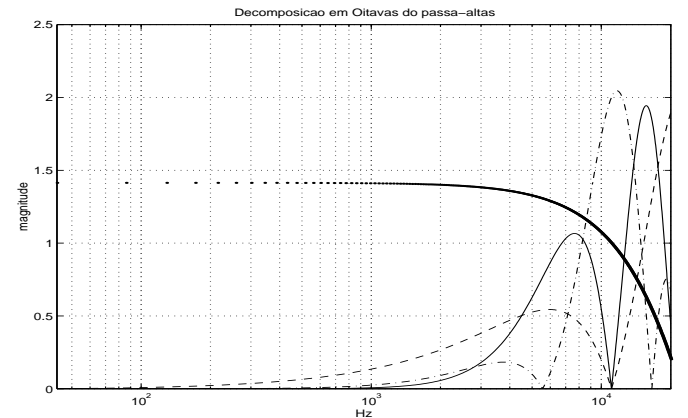


Figura 4. Fatiamento não-uniforme do espectro usando a wavelet bio5-3

As estruturas em oitava *direta* e *invertida* foram investigadas para quatro bases diferentes: duas biortogonais e duas maximamente planas (Daubechies [15]), cujos protótipos são mostrados na tabela I.

Decompondo a estrutura até o terceiro estágio ($M = 4$ sub-bandas), em cada sub-banda o filtro adaptativo terá um comprimento diferente, devido às diferentes esparsidades. As Figuras 5 e 6 mostram os valores finais dos coeficientes para cada filtro adaptativo em cada uma das duas decomposições, direta e invertida, respectivamente, utilizando o protótipo bio5-3.

Como as estruturas são diferentes, os coeficientes adaptados de $G_m(z^m)$ também diferem para cada tipo de decomposição e wavelet. Porém, como era esperado, a resposta impulsional para ambas estruturas é praticamente igual à HRTF que se deseja modelar, como pode ser observado na Figura 7. O erro médio quadrático entre a HRTF original e a modelada é de aproximadamente -60 dB.

Fazendo a modelagem da HRTF com os 488 coeficientes originais, obtemos em cada sub-banda o seguinte número de

| Wavelet | Protótipos |
|---------|--|
| Daub4 | $h^0(n) = \{-0.1294; 0.2241; 0.8365; 0.4829\}$ $h^1(n) = \{-0.4829; 0.8365; -0.2241; -0.1294\}$ |
| Daub8 | $h^0(n) = \{-0.0105; 0.0328; 0.0308; -0.1870$ $-0.0279; 0.6308; 0.7148; 0.2303\}$ $h^1(n) = \{-0.2303; 0.7148; -0.6308; -0.0279$ $0.1870; 0.0308; -0.0328; -0.0105\}$ |
| Bio5-3 | $h^0(n) = \frac{1}{4} \{-1; 2; 6; 2; -1\}$ $h^1(n) = \frac{1}{4} \{-1; 2; -1\}$ |
| Bio9-7 | $h^0(n) = \frac{1}{64} \{1; 0; -8; 16; 46; 16; -8; 0; 1\}$ $h^1(n) = \frac{1}{64} \{1; 0; -9; 16; -9; 0; 1\}$ |

TABELA I
FILTROS PROTÓTIPOS DA WAVELET

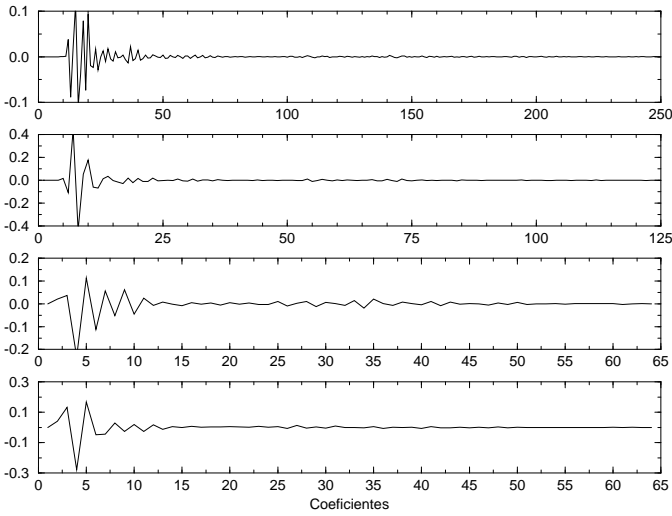


Figura 5. Coeficientes obtidos na decomposição passa-baixas do protótipo bio5-3

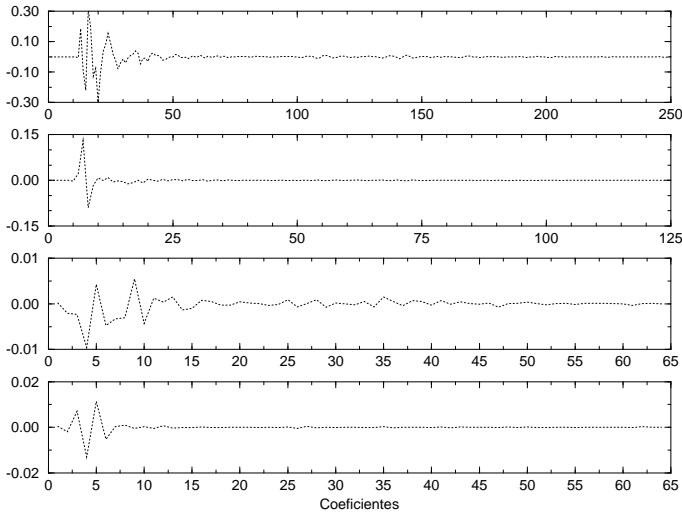


Figura 6. Coeficientes obtidos na decomposição passa-baixas do protótipo bio5-3

coeficientes: 253, 127, 64 e 64, para $m = 0, 1, 2$ e 3, respectivamente. O número total de coeficientes é 508, maior do que os 488 originais. Isto ocorre devido ao atraso Δ introduzido pela estrutura, porém os 28 coeficientes restantes podem ser removidos do final da resposta impulsiva da estrutura pois os

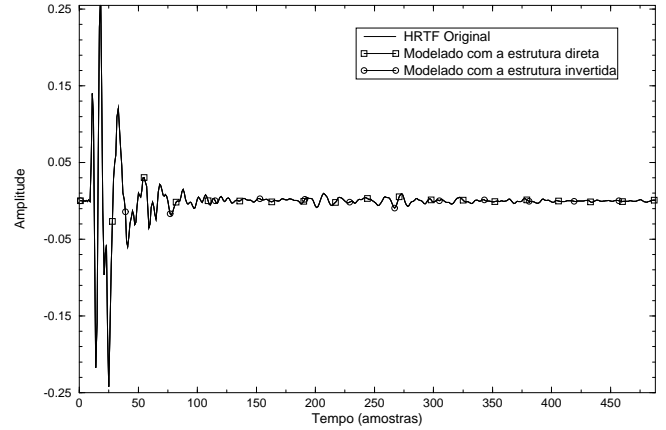


Figura 7. Respostas impulsivas da estrutura

valores tendem a zero. As HRTFs utilizadas nesse trabalho foram medidas pelo MIT [4], cujo comprimento é de 512 coeficientes. Porém, como até a posição 25 os coeficientes são praticamente nulos, foram utilizados apenas os coeficientes de 25 até 512.

Se utilizarmos todos os coeficientes obtidos para o modelo, com qualquer das wavelets mães descritas na tabela I, não há praticamente, variação no erro médio quadrático entre a magnitude das respostas em frequências originais e das modeladas com wavelets. Nas Figuras 8(a) e 8(b), encontram-se as respectivas respostas em frequência (módulo) para as decomposições dos protótipos passa-baixas e passa-altas, na direção de 60° de azimute e 30° de elevação. Para qualquer direção o comportamento é o mesmo.

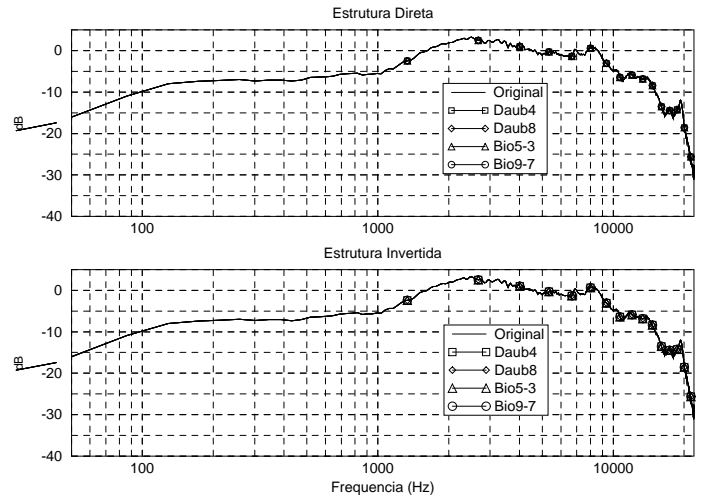


Figura 8. Respostas em frequência para decomposição (a) do passa-baixas e (b) do passa-altas

V. REDUÇÃO DO NÚMERO DE COEFICIENTES

Para que este tipo de modelagem possa ser utilizada em uma sistema de auralização, é necessário reduzir o número de coeficientes sem afetar a HRTF reconstruída a partir da wavelet. Observando as Figuras 5 e 6, pode-se notar, por inspeção, que apenas um conjunto de coeficientes são realmente importantes para a reconstrução da HRTF. O restante do coeficientes pode ser descartado devido a sua baixa amplitude,

quando comparado aos mais significantes.

Como os coeficientes mais relevantes encontram-se no início de cada sub-filtro esparsos, uma primeira idéia seria manter em cada filtro apenas os primeiros \tilde{K}_m coeficientes *efetivos* e após esta posição zerar os coeficientes restantes. Mantendo, por exemplo, os primeiros 40, 20, 10 e 10 coeficientes para as sub-bandas 0, 1, 2, e 3, respectivamente, observa-se, na Figura 9, que a “cauda” da resposta impulsional é forçada a ter coeficientes nulos, introduzindo um erro considerável quando comparada com a resposta original.

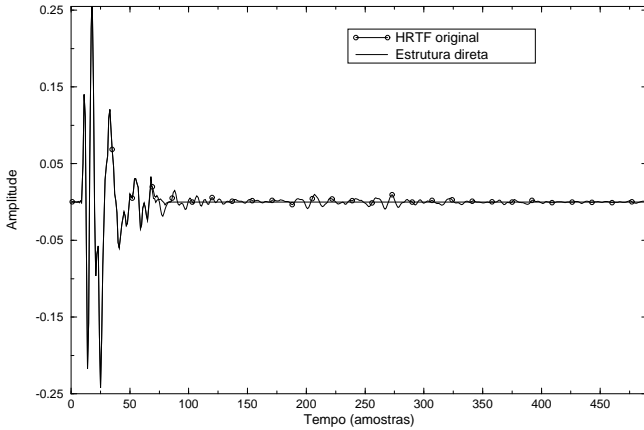


Figura 9. Resposta impulsional reconstruída com redução de coeficientes

Uma alternativa mais eficiente, seria selecionar em cada sub-banda o intervalo no qual os coeficientes são realmente efetivos, eliminando também os coeficientes nulos existentes no início de cada filtro esparsos. Este atraso nos coeficientes de cada banda, é proporcional ao atraso produzido pela diferença interaural, podendo ser calculado e inserido posteriormente. Assim, verificou-se em cada sub-banda, quais coeficientes possuíam valores absolutos maiores do que 90% do valor de pico da sub-banda para definir esses intervalos.

Para todas as direções disponíveis no banco de dados das HRTFs, foram obtidos os coeficientes dos filtros esparsos e determinados quais os intervalos (coeficientes inicial e final), para ambos os ouvidos. Após determinar os limites do intervalo útil em cada sub-banda, encontrou-se a média dos pontos iniciais e finais, cujos valores estão na tabela II para a estrutura direta e a invertida. Um exemplo da aplicação desta metodologia é que, na decomposição passa-baixas, se utilizarmos na primeira sub-banda apenas os coeficientes do intervalo de 10 a 41, a HRTF é modelada com menos erro do que se utilizássemos os primeiros 31 coeficientes desta sub-banda (começando da posição 1).

A resposta em frequência original e a obtida com esta modelagem usando os intervalos descritos na tabela II são mostradas nas Figuras 10 e 11 para ambos os ouvidos na direção com 30° de elevação e 30° de azimute, para decomposição do passa-baixas e passa-altas, respectivamente.

VI. CONCLUSÕES

Neste trabalho foi apresentada uma nova técnica para modelar as HRTFs, utilizando coeficientes obtidos através de transformadas wavelets e filtragem adaptativa. O modelo mostra-se bastante atraente do ponto de vista computacional, pois permite reduzir o número de coeficientes de 500 para 80, uma redução de aproximadamente 85%. A escolha de bases com

| protótipo | estrutura | sub-banda / intervalos | | | | total |
|-----------|-----------|------------------------|-------|------|------|-------|
| | | 0 | 1 | 2 | 3 | |
| bio5-3 | oitavas | 10-41 | 4-26 | 2-16 | 1-15 | 81 |
| | invertida | 11-52 | 5-20 | 1-12 | 1-15 | 81 |
| bio9-7 | oitavas | 23-53 | 10-30 | 4-18 | 3-17 | 78 |
| | invertida | 23-67 | 9-32 | 3-12 | 4-13 | 85 |
| daub4 | oitavas | 10-41 | 4-24 | 2-15 | 1-15 | 78 |
| | invertida | 10-53 | 4-25 | 2-11 | 2-12 | 83 |
| daub8 | oitavas | 22-52 | 10-27 | 4-16 | 2-16 | 73 |
| | invertida | 22-64 | 9-27 | 3-12 | 4-15 | 80 |

TABELA II
INTERVALOS EM CADA SUB-BANDA ONDE EXISTEM COEFICIENTES SIGNIFICANTES PARA ESTRUTURA EM OITAVAS E INVERTIDA

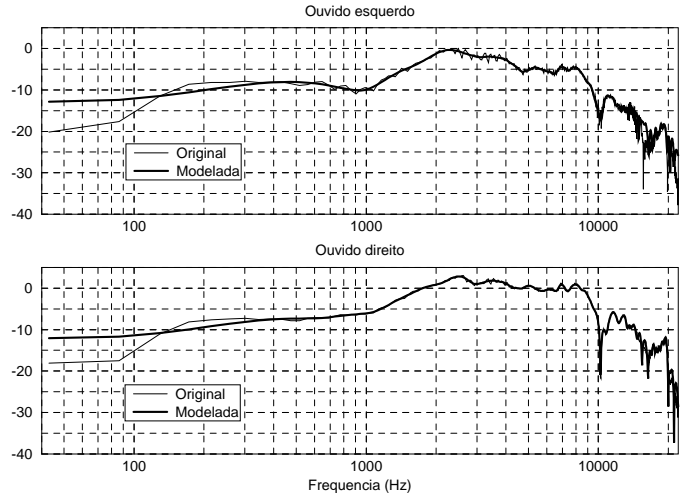


Figura 10. Resposta em frequência para a HRTF original e a modelada pela decomposição do passa-baixas

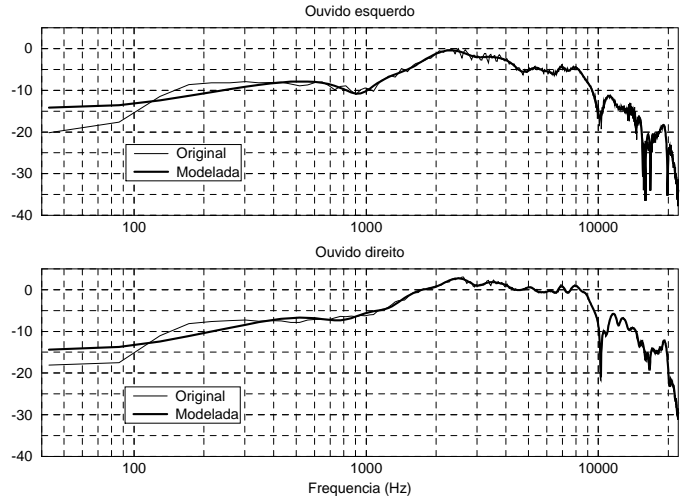


Figura 11. Resposta em frequência para a HRTF original e a modelada pela decomposição do passa-altas

reconstrução perfeita é suficiente para modelar corretamente as HRTFs, além de proporcionarem praticamente a mesma redução no número de coeficientes. Aparentemente, qualquer base wavelet pode ser utilizada sem prejuízo considerável para as HRTFs, desde que as condições descritas na seção 2 sejam satisfeitas. O erro proveniente da redução do número de coe-

ficientes é mais acentuado nas frequências abaixo de 100 Hz, como pode ser observado nas Figuras 10 e 11.

Isto ocorre para todas as direções, tendo seu valor máximo em torno de 10 dB. Para a faixa de baixas frequências, o ouvido humano possui pouca resolução na percepção da direcionalidade da fonte, importando apenas as diferenças interaurais de tempo e pressão [16]. O erro na faixa de frequência abaixo de 100 Hz não terá influência perceptível para a auralização de um ambiente acústico simulado, uma vez que a diferença interaural de nível permanece praticamente inalterada, assim como a diferença interaural de tempo, que é preservada pela wavelet.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] M. Vorlander, "Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm," *Journal of the Acoustical Society of America*, vol. 86, no. 1, pp. 172–179, 1989.
- [2] A. Krokstad, S. Strom, and S. Sorsdal, "Calculating the acoustical room response by use of a ray tracing technique," *Journal of Sound Vibration*, vol. 8, no. 1, pp. 118–125, 19968.
- [3] U. Stephenson, "Comparison of the mirror image source method and the sound particle simulation method," *Applied Acoustics*, vol. 29, pp. 35–72, 1990.
- [4] W. G. Gardner and K. D. Martin, "Hrtf measurements of a kemar," *Journal of the Acoustical Society of America*, vol. 97, no. 6, pp. 3907–3908, 1995, MIT website: <http://sound.media.mit.edu/KEMAR.html>.
- [5] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The cipc hrtf database," in *WASSAP '01 (2001 IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics)*, Oct. 2001, CIPIC website: <http://interface.cipic.ucdavis.edu/>.
- [6] D. J. Kistler and F. L. Wightman, "Headphone simulation of free-field listening. i: Stimulus synthesis," *Journal of the Acoustical Society of America*, vol. 85(2), pp. 858–867, Feb. 1989.
- [7] J. Blauert, *Spatial Hearing*, The MIT Press, 1997.
- [8] D. J. Kistler and F. L. Wightman, "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," *Journal of the Acoustical Society of America*, vol. 91, no. 3, pp. 1637–1647, Mar. 1991.
- [9] Y. Haneda, Y. Kaneda, S. Makino, and N. Kitawaki, "Common-acoustical-pole and zero modeling of head-related transfer functions," *IEEE Trans. on Speech and Audio Proc.*, vol. 7, no. 2, pp. 188–195, Mar. 1999.
- [10] T. Keränen, Huopaniemi, J., and A. Kärkkäinen, "Simplified numerical modeling of individual hrtf," *Proc. International Congress on Acoustics (ICA'2001), Rome, Italy, 2-7 September 2001*.
- [11] P. Georgiou and C. Kyriakakis, "Modeling of head related transfer function for immersive audio using a state-space approach," *Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems and Computers*, 1999.
- [12] J. C. B. Torres, "Estruturas para filtragem adaptativa usando transformadas uniformes, wavelets e bancos de filtros," M.S. thesis, COPPE/UFRJ, 1998.
- [13] G. Strang e T. Nguyem, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, 1997.
- [14] J. C. B. Torres and M. R. Petraglia, "Performance analysis of an adaptive filter employing wavelets and sparse subfilters," in *EUSIPCO 2000*, Sept. 2000, vol. II, pp. 997–1001.
- [15] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inform. Theory*, vol. 36, pp. 961–1005, Sept. 1990.
- [16] D. J. Kistler and F. L. Wightman, "The dominant role of low-frequency interaural time differences in sound localization," *Journal of the Acoustical Society of America*, vol. 91, no. 3, pp. 1648–1661, Mar. 1992.