

LUIZ W. P. BISCAINHO, PAULO S. R. DINIZ, PAULO A. A. ESQUEF

Laboratório de Processamento de Sinais, DEL-EE & PEE-COPPE, UFRJ
Caixa Postal 68504 - CEP 21945-970 - Rio de Janeiro, RJ, BRASIL
e-mails: {wagner,diniz,esquef}@lps.ufrj.br

Resumo— Este artigo trata da análise em sub-bandas de um processo aleatório autorregressivo com média móvel (ARMA) por um banco de filtros multitaxa. Inicialmente, obtém-se um modelo ARMA em taxa reduzida, equivalente quanto à densidade espectral de potência, para o sinal em cada sub-banda. Com destaque para o caso especial de um processo aleatório autorregressivo (AR) dividido em oitavas por um banco de filtros estruturado em árvore binária, discutem-se a validade do modelo proposto e o mapeamento dos zeros e pólos dos filtros geradores correspondentes ao longo das sub-bandas. Com vistas à utilização do modelo proposto, tecem-se considerações sobre o processamento de sinais baseado em modelos realizado em sub-bandas, em particular aplicado à restauração de sinais de áudio.

Abstract— This paper addresses the subband analysis of an autoregressive-moving average (ARMA) random process by a multirate filter bank. First, an ARMA model at reduced rate, equivalent in the power spectral density sense, is obtained for each subband signal. Focusing the special case of an autoregressive (AR) random process divided in octaves by a binary tree-structured filter bank, the validity of the proposed model is discussed, as well as the zero-pole mapping of the corresponding generator filters along the subbands. Aiming at the proposed model utilization, some considerations are made on model-based signal processing in subbands, specially when applied to audio restoration.

Key Words— ARMA models; Discrete-time signal processing; Filter banks; Multirate; Power spectral density.

1 Introdução

O processamento multitaxa de sinais discretos (Vaidyanathan, 1993) tem sido extensamente aplicado nas comunicações, na indústria de áudio e vídeo e em outras áreas, realizando a codificação de sinais, o processamento adaptativo ou estatístico de sinais etc. Os processos aleatórios discretos (Papoulis, 1991) obtidos pela passagem de ruído branco por um filtro digital linear, genericamente chamados autorregressivos com média móvel (ARMA), são comumente empregados como modelos de sinais no processamento de voz e áudio, particularmente em restauração de áudio (Godsill e Rayner, 1998). O processamento de sinais baseado em modelos realizado em sub-bandas poderia beneficiar-se de uma descrição similar a essa para os sinais nas sub-bandas, se isso fosse possível; é essa a motivação para este trabalho.

Considerando a análise em sub-bandas de um processo ARMA por um banco de filtros com decimação, este trabalho emprega a decomposição polifásica de sistemas (Vaidyanathan, 1993) para chegar a um modelo ARMA na taxa reduzida, equivalente quanto à densidade espectral de potência (Papoulis, 1991), para o sinal em cada sub-banda. O caso particular de um processo autorregressivo (AR) dividido sucessivamente em oitavas por um banco de filtros em árvore binária, de fácil interpretação, é detalhado. Discute-se a validade do modelo proposto e investiga-se a possibilidade de redução da ordem dos modelos pelo mapeamento dos zeros e pólos dos filtros geradores

correspondentes a cada sub-banda. Por fim, dão-se indicações do potencial do modelamento como ferramenta na restauração de áudio.

Após esta Introdução, a Seção 2 apresenta o caso geral do processo ARMA dividido em N bandas e detalha o caso do processo AR dividido em oitavas; na Seção 3, discutem-se a validade do modelo e o mapeamento de zeros e pólos dos filtros geradores; a Seção 4 comenta a potencial aplicação do modelo na restauração de áudio; as Conclusões estão na Seção 5.

2 Obtenção do Modelo

2.1 Processo ARMA em N Bandas

Considere-se o sistema multitaxa em N bandas mostrado na Fig. 1, que analisa o processo ARMA definido por

$$s(k) = \sum_{i=1}^I a(i)s(k-i) + \sum_{j=0}^J b(j)e(k-j),$$

sendo $e(k)$ ruído branco de variância σ_e^2 .

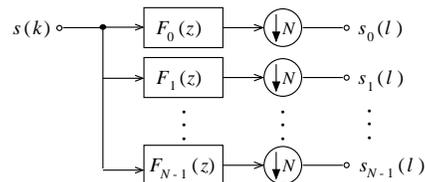


Figura 1. Processo ARMA analisado por banco de filtros em N bandas com decimação.

Para uma sub-banda qualquer n (Fig. 2a), pode-se agregar o filtro gerador de $s(k)$ ao filtro de análise $F_n(z)$ (Fig. 2b), obter a representação polifásica para o filtro resultante $\hat{F}_n(z)$ (Fig. 2c) e chegar a sua versão modificada (Fig. 2d). Para cada polifase p , pode-se posicionar o decimador antes do filtro $\hat{F}_{np}(z)$ e representar a respectiva componente de $e(k)$ como $e_p(l)$ (Fig. 2e). É fácil concluir que $e_p(l)$, para $p = 0, 1, \dots, N-1$, são processos de ruído branco mutuamente não-correlacionados, com variâncias

$$E[e_p^2(l)] = E[e^2(NL + p)] = E[e^2(k)] = \sigma_e^2.$$

Como os sinais $s_{np}(l)$ são versões coloridas de $e_p(l)$ para $p = 0, 1, \dots, N-1$, respectivamente, são também mutuamente não-correlacionados. Assim, a autocorrelação de $s_n(l)$ pode ser escrita como

$$R_{s_n s_n}(\Delta) = \sum_{p=0}^{N-1} R_{s_{np} s_{np}}(\Delta),$$

levando à densidade espectral de potência

$$\begin{aligned} \mathcal{S}_{s_n s_n}(z) &= \sum_{p=0}^{N-1} \mathcal{S}_{s_{np} s_{np}}(z) = \\ &= \sum_{p=0}^{N-1} \hat{F}_{np}(z) \hat{F}_{np}(z^{-1}) \sigma_e^2 = \\ &= \mathcal{S}_{s_n s_n}(-z). \end{aligned}$$

A fatoração da expressão acima na forma

$$\mathcal{S}_{s_n s_n}(z) \triangleq \hat{F}_{d_n}(z) \hat{F}_{d_n}(z^{-1}) \sigma_e^2 \quad (1)$$

permite obter uma representação para $\tilde{s}_n(l)$, equivalente a $s_n(l)$ quanto à densidade espectral de potência, na forma do processo ARMA que resulta da passagem do ruído branco $e_{d_n}(l)$ com variância σ_e^2 pelo filtro de fase mínima $\hat{F}_{d_n}(z)$ (Fig. 2f).

O modelo equivalente completo consistiria, então, de N subsistemas similares em paralelo; mas sua descrição deve incluir, ainda, a correlação mútua entre diferentes sub-bandas. Sejam, então, duas sub-bandas distintas quaisquer n_1 e n_2 , cada uma representada como na Fig. 2e. A correlação cruzada entre seus sinais é dada por

$$\begin{aligned} R_{s_{n_1} s_{n_2}}(\Delta) &= \sum_{p=0}^{N-1} E[s_{n_1 p}(l) s_{n_2 p}(l + \Delta)] = \\ &= \sum_{p=0}^{N-1} E[\hat{f}_{n_1 p} * e_p(l) \hat{f}_{n_2 p} * e_p(l + \Delta)], \end{aligned}$$

resultando na densidade espectral de potência cruzada

$$\mathcal{S}_{s_{n_1} s_{n_2}}(z) = \left[\sum_{p=0}^{N-1} \hat{F}_{n_1 p}(z) \hat{F}_{n_2 p}(z^{-1}) \right] \sigma_e^2.$$

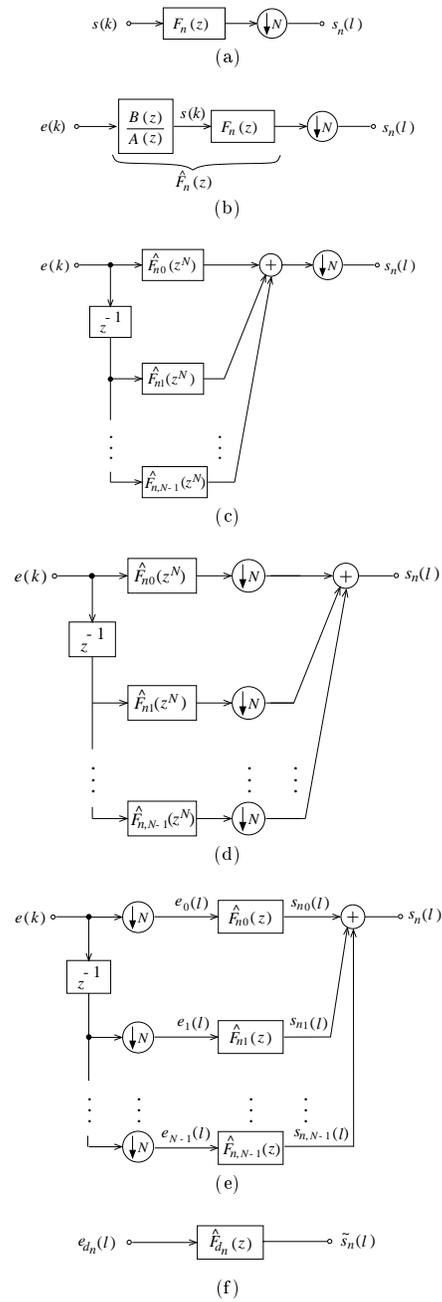


Figura 2. Geração de $s_n(l)$: da representação multitaxa ao modelo em taxa reduzida.

2.2 Processo AR em Oitavas

Seja o processo AR de ordem I definido por

$$s(k) = \sum_{i=1}^I a(i) s(k-i),$$

sendo $e(k)$ ruído branco de variância σ_e^2 . Considere-se, agora, o banco de filtros em árvore binária mostrado na Fig. 3, utilizado para analisar $s(k)$ em oitavas. A cada estágio, os filtros de análise são o passa-altas $F_H(z)$ e o passa-baixas $F_L(z)$, FIR de ordem Q . Os resultados da Subseção 2.1 podem facilmente ser adaptados a essa situação, já que processos AR são um caso particular de processos ARMA e os modelos para as

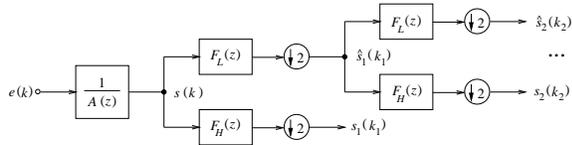


Figura 3. Análise de um processo AR em oitavas por um banco de filtro com decimação.

duas sub-bandas resultantes de cada subdivisão do espectro de frequências podem ser computados recursivamente pela Eq. (1).

A passagem do estágio $m - 1$ para o estágio m é ilustrada na Fig. 4. Na notação modificada, agora as sub-bandas são indexadas pelo estágio m de divisão do espectro e o circunflexo ($\hat{\cdot}$) se relaciona à sua metade inferior. Para cada polifase

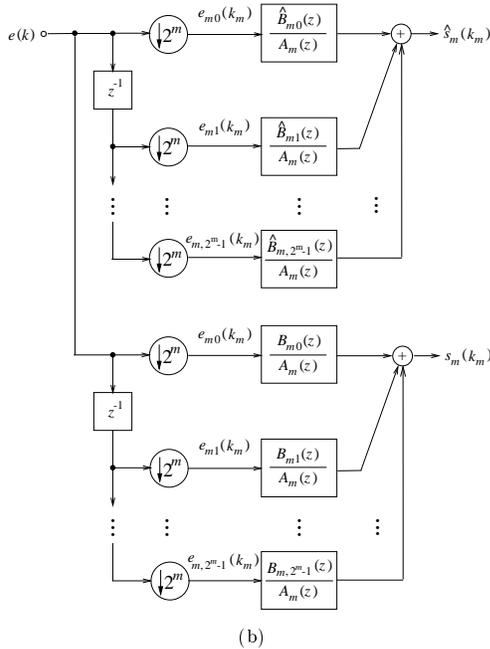
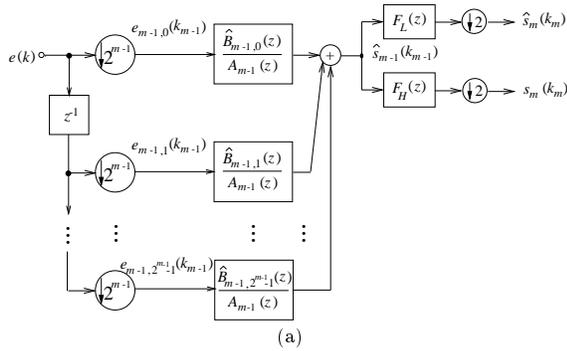


Figura 4. Geração de $\hat{s}_m(k_m)$ e $s_m(k_m)$.

$0 \leq p \leq 2^{m-1} - 1$, as transformações são regidas pelas equações

$$\begin{aligned} F_H(z) \frac{\hat{B}_{m-1,p}(z)}{A_{m-1}(z)} &= F_H(z) \frac{\hat{B}_{m-1,p}(z) A_{m-1}(-z)}{A_{m-1}(z) A_{m-1}(-z)} \triangleq \\ &\triangleq \frac{B_{m-1,p,H}(z)}{A_m(z^2)} \triangleq \frac{B_{m,2p}(z^2) + z^{-1} B_{m,2p+1}(z^2)}{A_m(z^2)} \end{aligned}$$

e

$$\begin{aligned} F_L(z) \frac{\hat{B}_{m-1,p}(z)}{A_{m-1}(z)} &= F_L(z) \frac{\hat{B}_{m-1,p}(z) A_{m-1}(-z)}{A_{m-1}(z) A_{m-1}(-z)} \triangleq \\ &\triangleq \frac{B_{m-1,p,L}(z)}{A_m(z^2)} \triangleq \frac{\hat{B}_{m,2p}(z^2) + z^{-1} B_{m,2p+1}(z^2)}{A_m(z^2)}. \end{aligned}$$

As recursões se iniciam por $m = 0$, com $e_{00}(k_0) = e(k)$, $\hat{B}_{00}(z) = 1$ e $A_0(z) = A(z)$.

Pela mesma argumentação da Subseção 2.1, a densidade espectral de potência de $s_m(k_m)$ é

$$\begin{aligned} S_{s_m s_m}(z) &= \frac{\sum_{p=0}^{2^m-1} B_{mp}(z) B_{mp}(z^{-1})}{A_m(z) A_m(z^{-1})} \sigma_e^2 \triangleq \\ &\triangleq \frac{B_m(z) B_m(z^{-1})}{A_m(z) A_m(z^{-1})} \sigma_e^2. \end{aligned} \quad (2)$$

Em suma, conforme a Fig. 5, descreveu-se $\tilde{s}_m(k_m)$, equivalente a $s_m(k_m)$ quanto à densidade espectral de potência, como o processo ARMA que resulta da passagem do ruído branco $e_{d_m}(k_m)$ de variância σ_e^2 pelo filtro de fase mínima

$$\mathcal{M}_m(z) \triangleq \frac{B_m(z)}{A_m(z)}.$$

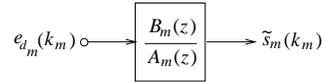


Figura 5. Modelo para o sinal s_m .

A obtenção do modelo para $\hat{s}_m(k_m)$ é análoga à que se viu para $s_m(k_m)$.

Completando o modelo, a densidade espectral de potência cruzada entre sub-bandas de mesma taxa pode ser escrita como

$$S_{\hat{s}_m s_m}(z) = \frac{[\sum_{p=0}^{2^m-1} \hat{B}_{mp}(z) B_{mp}(z^{-1})]}{A_m(z) A_m(z^{-1})} \sigma_e^2.$$

3 Discussão

Por facilidade de interpretação, esta Seção se restringe ao caso particular do processo AR dividido em oitavas, exposto na Subseção 2.2, frequentemente associado ao processamento de áudio.

3.1 Validade do Modelo

Supondo, inicialmente, que modelos ARMA são capazes de representar adequadamente os sinais nas sub-bandas, a resposta em frequência dos filtros $\mathcal{M}_m(z)$ derivados anteriormente corresponde à dos filtros que seriam obtidos por estimação direta dos coeficientes?

Em caso positivo, os modelos ARMA poderiam ser substituídos por modelos AR sem acréscimo de complexidade e ainda manter, com boa aproximação, a resposta em frequência?

Por fim, os filtros geradores correspondentes aos modelos ARMA nas sub-bandas, invertidos, branqueiam os sinais nas sub-bandas?

Um exemplo prático pode ilustrar as três respostas. Um processo AR foi formado com os pólos obtidos pela combinação dos módulos 0,5 e 0,99 com as fases 0 , $\frac{\pi}{30}$, $\frac{\pi}{3}$, $\frac{\pi}{2}$, $2\frac{\pi}{3}$ e $29\frac{\pi}{30}$ e analisado em sucessivos estágios de bancos de filtros Haar (Vaidyanathan, 1993).

Para o sinal em cada sub-banda, obteve-se o filtro gerador correspondente: ao modelo ARMA descrito pela Eq. (2); a um modelo ARMA de mesma ordem por estimação direta de seus coeficientes, em simulação; e a um modelo AR, com denominador de mesma ordem, por estimação direta de seus coeficientes, em simulação. A Fig. 6 compara as três respostas em frequência de módulo e fase correspondentes, para os sinais resultantes de 2 estágios de divisão. As respostas em frequência dos modelos ARMA foram praticamente indistinguíveis, validando o modelo teórico. A possibilidade de utilização de modelos AR de complexidade menor também se evidenciou, embora a ordem escolhida não tenha sido capaz de conformar perfeitamente a resposta em muito baixas frequências de \mathcal{M}_1 .

No mesmo exemplo, realizou-se a filtragem inversa do sinal em cada sub-banda, em simulação, pelo filtro gerador previsto pela Eq. (2) e calculou-se a autocorrelação temporal do sinal resultante. A Fig. 7 mostra as autocorrelações normalizadas para os sinais resultantes de 2 estágios de divisão. Os sinais são, para todos os efeitos, ruídos brancos, confirmando a adequação do modelo.

3.2 Mapeamento de Zeros e Pólos ao longo das Sub-bandas

Suponha-se inicialmente que $F_L(z)$ e $F_H(z)$ na Fig. 3 são filtros passa-baixas e passa-altas de meia banda ideais. Um cosseno de frequência ω analisado sucessivamente em oitavas por esse banco de filtros aparecerá em $\hat{s}_m(k_m)$ com sua frequência dobrada após cada estágio m , até ser “separada” no sinal $s_{m\omega}(k_m)$, após o estágio

$$m_\omega = \left\lceil \log_2 \frac{\pi}{2\omega} \right\rceil + 1,$$

agora com frequência $2\pi - 2^{m_\omega}\omega$. Um processo AR $s(k)$ genérico com pólos z_{p_i} , após 1 estágio de sua análise em oitavas, resultaria em dois processos ARMA com pólos $z_{p_i}^2$ (isto é, com as fases originais dobradas e os módulos originais quadrados) e os zeros necessários para dividir perfeitamente ao meio o espectro de potência, estes também ao quadrado. O efeito se repete ao longo dos estágios. Portanto, os pólos originais não podem ser perfeitamente “separados” pelas sub-bandas, exceto no caso limite de módulo unitário (o cosseno).

Para bancos de filtros realizáveis, a propriedade de quadramento dos pólos (que decorre da decimação por 2) se preserva e o mapeamento dos zeros continua difícil de prever sem cálculos, mas algumas especulações podem ser feitas.

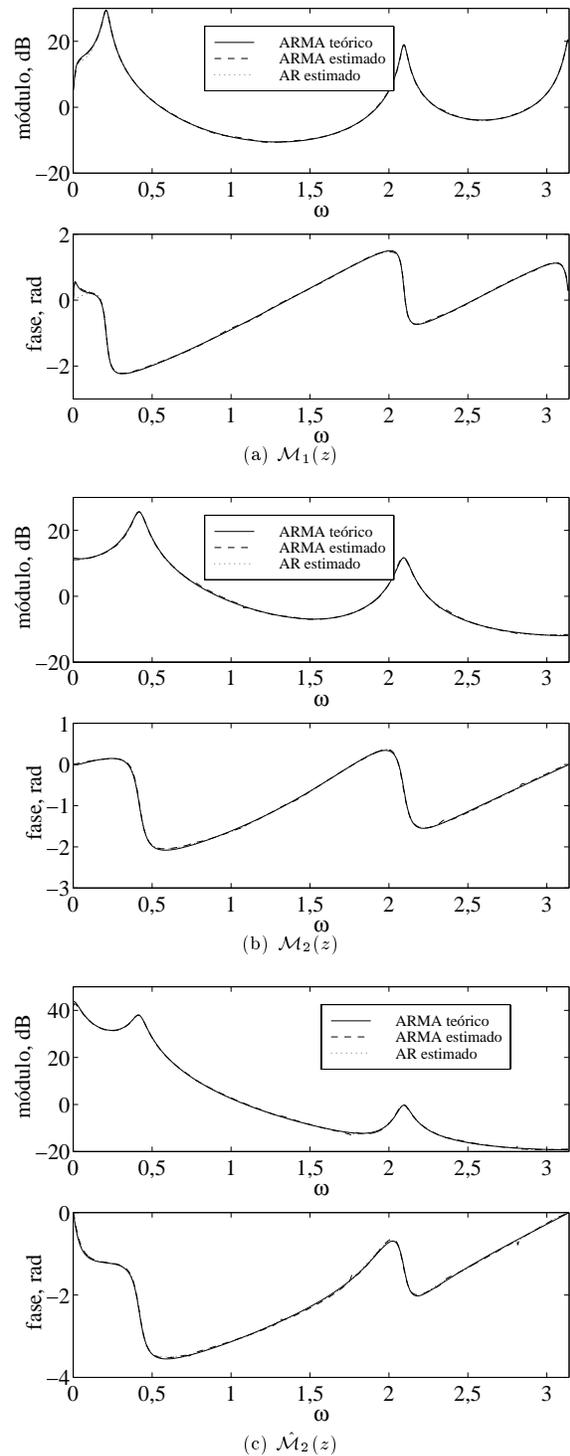


Figura 6. Respostas em frequência de módulo e fase: modelo ARMA teórico e modelos ARMA e AR estimados.

É possível deduzir que as ordens do numerador e do denominador de $\hat{\mathcal{M}}_m(z)$ e $\mathcal{M}_m(z)$ são

$$\begin{cases} O_{\text{num}}(m) = \left\lceil \frac{O_{\text{num}}(m-1) + I}{2} \right\rceil, & m \geq 2 \\ O_{\text{num}}(1) = \left\lceil \frac{Q+I}{2} \right\rceil \end{cases}$$

e $O_{\text{den}}(m) = I$, respectivamente, a menos de possíveis cancelamentos, onde I e Q foram definidos na Seção 2.

Ignorando por ora a decimação e retornando ao banco de filtros ideais, a largura de banda do

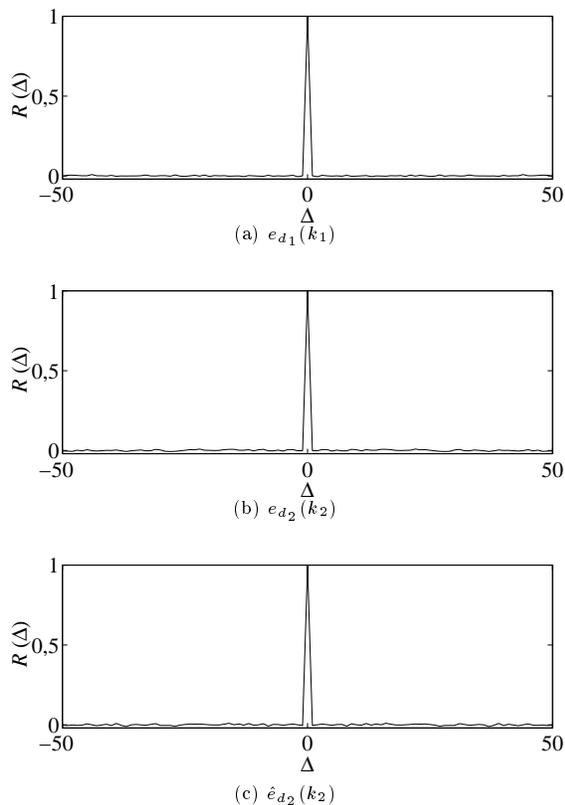


Figura 7. Autocorrelações dos sinais obtidos por filtragem inversa dos sinais nas sub-bandas.

sinal original $s(k)$ é progressivamente reduzida ao longo dos estágios de análise. Num estágio qualquer, o efeito individual de qualquer pólo do filtro gerador de $s(k)$ se distribui entre a banda passante e a banda de rejeição do filtro de análise equivalente; se, por exemplo, tal efeito se encontra praticamente todo na banda de rejeição, sua quase eliminação pode ser conseguida pelo cancelamento desse pólo por um zero praticamente na mesma localização.

Retomando, agora, a decimação e os bancos de filtros não-ideais, é preciso lembrar que, após o m -ésimo estágio, cada pólo é representado por sua versão m vezes quadrada, mas a idéia geral se mantém. Portanto, espera-se que ao longo dos estágios os filtros geradores equivalentes dos modelos tenham um número crescente de pólos cancelados por zeros. Deve-se notar que, embora os bancos de filtros sejam fixos, os cálculos envolvidos na obtenção dos numeradores de $\hat{\mathcal{M}}_m(z)$ e $\mathcal{M}_m(z)$ combinam os pólos do filtro gerador de $s(k)$ com os filtros $F_L(z)$ e $F_H(z)$ de tal forma que permite que os zeros “persegam” os pólos.

Esse fato, associado à progressiva interiorização no círculo unitário, sugere a possibilidade de se reduzirem gradualmente as ordens dos modelos ao longo dos estágios sem prejuízo da acurácia do modelamento global.

Como exemplo, procedeu-se à análise em oitavas de um processo AR de ordem 40 (extraído do modelo para um trecho de gravação musical)

por um sistema como o da Fig. 3 com bancos de filtros Haar. A Fig 8 mostra os diagramas de zeros e pólos calculados para os filtros geradores correspondentes a quadro estágios de divisão, onde se podem observar as migrações e os cancelamentos previstos. Após o cálculo dos modelos completos, como teste descartou-se qualquer zero, pólo ou par zero-pólo cuja resposta em frequência individual se desviasse menos que 10% do módulo unitário e menos de $\frac{\pi}{10}$ da fase linear, considerando-se desprezível seu efeito sobre a função de transferência global. A Tabela 1 mostra as ordens correspondentes aos modelos completos e reduzidos segundo esse critério. A simplificação atingida é considerável. Pode-se, ainda, pensar na substituição dos modelos ARMA de ordem reduzida por modelos AR, de estimação mais simples, com ordem suficiente para assegurar a equivalência.

Tabela 1. Processo AR de ordem 40 analisado em oitavas: número de zeros e pólos dos filtros geradores dos modelos.

filtro	completo		reduzido	
	#zeros	#pólos	#zeros	#pólos
$\mathcal{M}_1(z)$	20	40	6	26
$\mathcal{M}_2(z)$	30	40	2	12
$\mathcal{M}_3(z)$	35	40	3	6
$\mathcal{M}_4(z)$	38	40	2	4
$\hat{\mathcal{M}}_4(z)$	38	40	1	4

4 Potencial Aplicação: Restauração de Áudio

A restauração de áudio proveniente de gravações deterioradas (Godsill e Rayner, 1998) lida com problemas que podem ser genericamente classificados em localizados ou globais, dos quais os exemplos mais comuns são, respectivamente, *clicks* (ruído impulsivo) e *hiss* (ruído de fundo de banda larga). Métodos de supressão de ruído impulsivo geralmente usam modelos autorregressivos e métodos de redução de ruído de fundo se baseiam freqüentemente em sistemas em sub-bandas.

O modelamento proposto anteriormente tem como alvo a realização em sub-bandas do processamento digital de sinais baseado em modelo. Então, tendo em vista particularmente sua aplicação ao processamento de áudio, examinam-se a seguir as possíveis vantagens da realização da restauração de áudio em sub-bandas.

Da mesma forma que a codificação e a compressão de áudio aplicam princípios de Psicoacústica para determinar “quanto o áudio pode ser piorado sem que se perceba”, a restauração de áudio pode utilizá-los para determinar “quanto o áudio precisa ser melhorado que ainda se perceba”, a fim de reduzir esforço computacional e evitar perda de qualidade por pro-

cessamento desnecessário. Um princípio fundamental é o Mascaramento (Jeffress, 1970), propriedade de um som inibir a percepção de outro próximo no tempo ou na frequência. Sabe-se que só uma estreita faixa de frequências, chamada Banda Crítica (Scharf, 1970), em torno de um tom contribui para o mascaramento deste. Um conjunto fixo de bandas críticas é convencionalmente adotado para modelar a percepção sonora, levando a uma configuração correspondente em sub-bandas.

Um sistema multitaxa adequado pode favorecer a supressão de *clicks* baseada em modelo simultaneamente à redução de *hiss* em sub-bandas.

Os distúrbios podem, às vezes, se encontrar preferencialmente em determinadas faixas de frequência; então, o processamento em sub-bandas pode possibilitar redução de processamento ao dar mais “atenção” às sub-bandas mais afetadas.

Pode-se conseguir ainda maior redução de esforço computacional se a equivalência das operações mais complexas realizadas nos sistemas em banda completa e em sub-bandas for favorável a este.

5 Conclusões

Foi proposto e validado um modelo para um processo ARMA dividido em sub-bandas, na forma de um processo ARMA em taxa reduzida para representar o sinal em cada sub-banda. O mapeamento de zeros e pólos dos filtros geradores do modelo indicou a possibilidade de substancial simplificação de suas ordens. Diversas vantagens decorrentes da realização da restauração de áudio em sub-bandas indicam a potencial aplicação do modelamento proposto a esse problema.

Referências Bibliográficas

- Godsill, S. J. e Rayner, P. J. W. (1998). *Digital Audio Restoration*, Springer, London, UK.
- Jeffress, L. A. (1970). Masking, in J. V. Tobias (ed.), *Foundations of Modern Auditory Theory*, Academy Press, New York, NY, USA, chapter 3, pp. 85–114.
- Papoulis, A. (1991). *Probability, Random Variables and Stochastic Processes*, 3 edn, McGraw-Hill, New York, NY, USA, section 10.4, pp. 332–336.
- Scharf, B. (1970). Critical bands, in J. V. Tobias (ed.), *Foundations of Modern Auditory Theory*, Academy Press, New York, NY, USA, chapter 5, pp. 157–202.
- Vaidyanathan, P. P. (1993). *Multirate Systems and Filter Banks*, Prentice-Hall, Englewood Cliffs, NJ, USA.

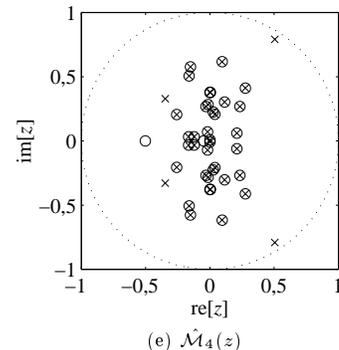
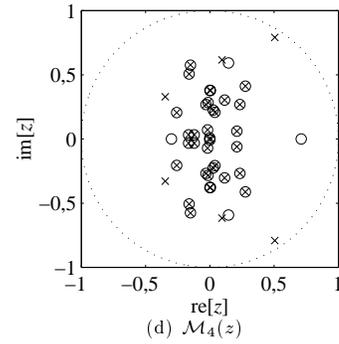
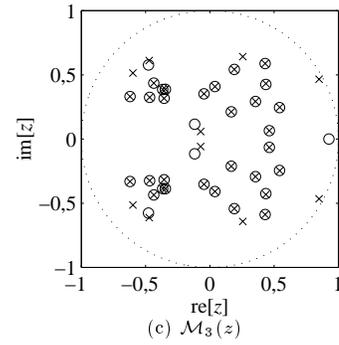
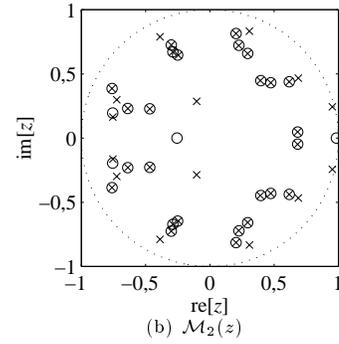
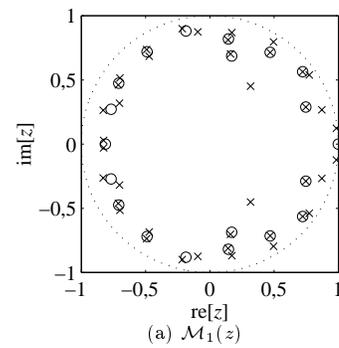


Figura 8. Análise de processo AR em oitavas: diagramas de pólos e zeros.