

Detecção de Pulsos Longos em Sinais de Áudio

Paulo A. A. Esquef, Luiz W. P. Biscainho, Paulo S. R. Diniz

PEE do COPPE & DEL da EE, Universidade Federal do Rio de Janeiro

CP 68504, CEP 21945-970, Rio de Janeiro, RJ

Fone: (021) 260 5010, Fax: (021) 590 0788

{esquef, wagner, diniz}@lps.ufrj.br

Sumário — Este trabalho se refere à detecção de distúrbios na forma de pulsos de longa duração presentes em sinais provenientes de matrizes degradadas de áudio. A principal técnica proposta se baseia na FFT aplicada a blocos do sinal discretizado, o que lhe confere simplicidade e velocidade. O emprego auxiliar da filtragem por mediana é responsável por elevada eficácia na detecção, inclusive no caso crítico de distúrbios imersos em sinais de ampla faixa dinâmica e largo espectro de frequências.

1 INTRODUÇÃO

Danos sérios no sulco de uma matriz em disco — originados, por exemplo, por arranhões — excitam impulsivamente o sistema empregado na reprodução da matriz, produzindo distúrbios, em geral de longa duração e alta amplitude, que são um dos tipos mais desagradáveis de degradação em gravações analógicas de áudio. Esses distúrbios serão, doravante, denominados de Pulsos Longos.

Estando em questão apenas técnicas de processamento digital de sinais, ao longo do texto está implícito que todos os sinais estão na forma discreta. Em particular, nos exemplos, a taxa de amostragem empregada é de 44100 Hz, padrão do CD.

O processamento visando à supressão de pulsos longos pode ser dividido nas etapas de Detecção e Reconstrução. Na etapa de Detecção, os distúrbios são localizados temporalmente ao longo do sinal; na etapa de Reconstrução, o sinal de áudio é restaurado nas regiões dadas como corrompidas na etapa anterior. Neste trabalho serão enfatizadas técnicas para a Detecção dos pulsos longos.

Na Seção 2, são discutidas algumas características associadas aos pulsos longos. Na Seção 3, são revistos os algoritmos correntemente empregados no tratamento de pulsos longos. Finalmente, na Seção 4, são apresentados 3 possíveis algoritmos para detecção de pulsos longos em blocos. Entre eles, propõe-se uma abordagem baseada na análise de conteúdo espectral que, com o auxílio da filtragem por mediana, apresenta desempenho superior na detecção dos distúrbios em sinais com conteúdo

espectral rico e dinâmica ampla, quando comparada aos outros 2 algoritmos.

2 CARACTERIZAÇÃO DOS PULSOS LONGOS

Em geral, os pulsos longos podem ser divididos, temporalmente, em duas partes. A parte inicial consiste de uma descontinuidade semelhante a um distúrbio impulsivo de alta amplitude. A segunda parte, que sucede a primeira, é formada por oscilações amortecidas de baixas frequências, de longa duração, associadas à resposta impulsiva, em geral não-linear, do sistema utilizado para reprodução da matriz.

A Fig. 1 mostra um exemplo de pulso longo e ilustra sua divisão nas partes mencionadas.

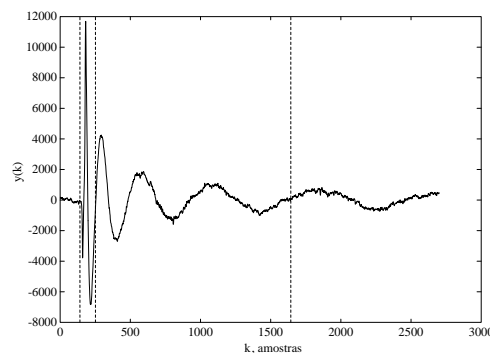


Figura 1: Exemplo de pulso longo: as linhas verticais indicam aproximadamente as regiões da descontinuidade inicial e do transitório de baixa frequência.

A análise do conteúdo espectral dos pulsos longos revela que suas descontinuidades iniciais exibem considerável energia ao longo de todo o espectro. A Fig. 2 mostra o espectrograma de um sinal de áudio originalmente contaminado por pulsos longos. A ocorrência de faixas espectrais largas temporalmente localizadas é bem característica da presença de pulsos longos e pode ser bastante útil na sua detecção, como será visto na Seção 4.

3 MÉTODOS USUAIS DE PROCESSAMENTO

Até recentemente, a única técnica específica de processamento para supressão de pulsos longos men-

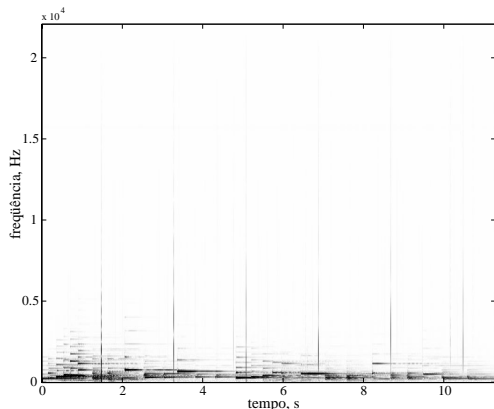


Figura 2: Espectrograma de um sinal de áudio originalmente contaminado por pulsos longos, utilizando blocos e FFTs de 1024 amostras e superposição de 50% entre blocos.

cionada na literatura corrente era a descrita em [1]. Trata-se de um método de detecção e reconstrução de pulsos longos por Similaridade, ou por Pulso-Padrão, e parte da hipótese de que os pulsos longos variam primariamente em amplitude, mantendo praticamente inalterados seus formatos e durações, ditados pela resposta impulsiva do sistema usado na reprodução da matriz.

O processamento propriamente dito consiste em “deslizar” sobre o sinal, amostra a amostra, um pulso-padrão representativo do distúrbio típico, calculando-se o módulo do coeficiente de correlação cruzada correspondente. Picos de correlação acima de um limiar próximo de 1 indicam a possível presença de um pulso a ser suprimido. Isto é feito em grande parte subtraindo-se do sinal uma versão sincronizada do padrão ajustada em polaridade e amplitude por um fator convenientemente calculado. Como resultado, o longo transitório de baixa frequência é praticamente eliminado, restando a descontinuidade inicial, de mais curta duração. Esta pode ser suprimida por técnicas usuais para eliminação de ruído impulsivo [2].

O pulso-padrão pode ser calculado, por exemplo, como a média de prováveis distúrbios encontrados como visto, sucessivamente refinada ao longo do processo.

Numa matriz musical típica, as ocorrências de pulsos longos são pouco numerosas, o que faz a busca pontual dos distúrbios ao longo do sinal soar bastante custosa. Se o responsável pelos pulsos é, por exemplo, um arranhão aproximadamente radial na superfície de um disco, eles aparecem de modo periódico no sinal de áudio, e suas posições podem ser estimadas levando-se em consideração a taxa de amostragem utilizada na digitalização e a velocidade angular de reprodução da matriz. Nesse caso, portanto, é possível realizar a busca de pulsos similares apenas nas regiões em torno das posições estimadas de ocorrência dos pulsos longos. Se, por um lado, o uso desse recurso acelera o processamento, por outro, reduz a generalidade do método em relação à detecção de pulsos em quaisquer posições.

Uma forte limitação desse método está na hipótese de invariabilidade dos pulsos longos, que, em geral, não é sempre satisfeita. De fato, os pulsos longos variam tanto em duração quanto no formato. Uma forma de superar este problema é dispor de uma biblioteca de pulsos-padrões previamente acumulados, que permita a utilização do mais similar a cada distúrbio. Obviamente, a maior generalidade se obtém ao custo do aumento no tempo de processamento pelos múltiplos testes.

Há que se mencionar, ainda, o caso de ocorrência de pulsos longos muito próximos, podendo mesmo aparecer superpostos. Isso dificulta a sua detecção por correlação com um pulso-padrão simples [3].

Com vistas a maior generalidade, em [3] é proposto o método de supressão de pulsos longos por Separação de Processos AR (Autorregressivos — ver Apêndice), que apresenta a vantagem de ser menos dependente da hipótese de invariabilidade dos pulsos longos, permitindo suprimir tanto pulsos com transitórios diferentes como pulsos superpostos. Contudo, o método apresentado se restringe à etapa de Reconstrução, considerando as localizações dos pulsos longos no sinal como informações previamente conhecidas. Isso nos leva novamente ao problema de realizar uma Detecção eficaz.

4 DETECÇÃO DE PULSOS LONGOS

A detecção de pulsos longos consiste em localizar temporalmente as regiões do sinal de áudio corrompidas por pulsos longos. Quando se deseja realizar processamento em blocos, a detecção de pulsos longos tem como objetivo identificar em quais blocos do sinal existem ocorrências de pulsos longos. Nesse sentido, a idéia básica de um sistema de detecção em blocos consiste em extrair informações dos blocos do sinal que permitam distinguir, por algum procedimento seletivo, quais deles estão corrompidos por pulsos longos. Desta forma, o sistema de detecção será tanto melhor quanto maior for o contraste entre as informações extraídas dos blocos corrompidos e não-corrompidos.

A seguir, são abordados 3 métodos para detecção de pulsos longos, usando processamento em blocos. O Método A tem apelo quase trivial; os Métodos B e C são apresentados originalmente neste trabalho.

4.1 Método A

A idéia do Método A é realizar a detecção dos blocos contaminados por pulsos longos comparando-se o módulo do sinal dos blocos com um limiar de valor arbitrariamente adotado. Os blocos em que pelo menos uma amostra do sinal exibir módulo de valor maior que o limiar são, então, selecionados como corrompidos por pulsos longos.

A justificativa para esse método reside nas amplitudes elevadas tipicamente apresentadas pelos pulsos longos, logo no início de suas ocorrências, em relação às amplitudes do restante do sinal. Nesses

termos, a escolha de um limiar de valor bastante alto é conveniente. Essa forma de detecção pode ser problemática quando o sinal de áudio tem faixa dinâmica muito ampla e suas amplitudes podem se tornar comparáveis ao valor elevado adotado para o limiar. Nesse caso, a ocorrência de falsos alarmes pode se tornar muito elevada; em contrapartida, a elevação do valor do limiar pode levar a altos índices de não-deteção.

A Fig. 3 mostra, respectivamente, 2 sinais reais de áudio, denominados, daqui para a frente, de Sinal I e Sinal II. O Sinal I é naturalmente corrompido por pulsos longos; o Sinal II é corrompido artificialmente por pulsos longos reais.

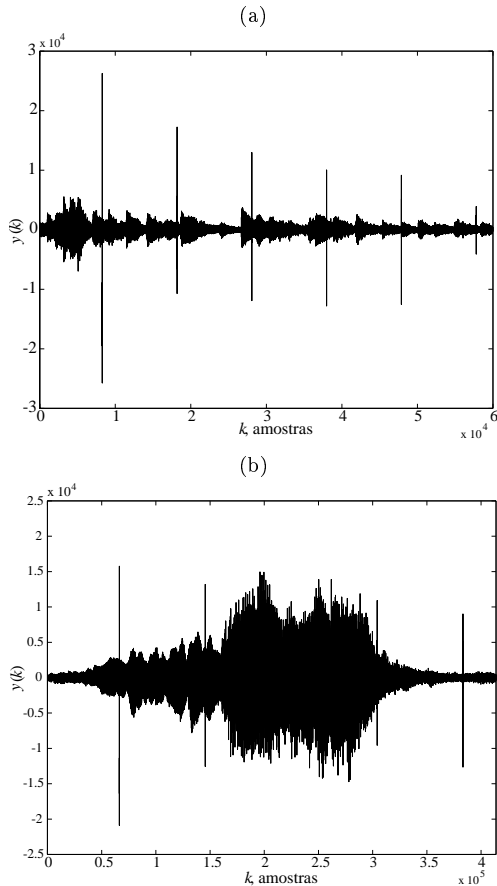


Figura 3: (a) Amplitude do Sinal I (dinâmica restrita). (b) Amplitude do Sinal II (dinâmica ampla).

O Sinal I tem dinâmica restrita, e nele os pulsos longos estão bastante evidenciados, enquanto que no Sinal II, com dinâmica ampla, eles são mascarados pelas amplitudes altas das amostras no início da segunda metade do trecho. O Método A é adequado a detectar pulsos longos em sinais com característica de dinâmica similar à do Sinal I, e inadequado para sinais do tipo do Sinal II.

4.2 Método B

O Método B se baseia nas características espectrais apresentadas pelos pulsos longos, principalmente o fato de apresentarem energia significativa em uma faixa ampla do espectro, como pode ser

observado no espectrograma mostrado na Fig. 2.

O método ora proposto consiste em calcular a Transformada Discreta de Fourier (DFT) [4] de cada bloco, selecionar uma faixa de frequências onde a informação dos pulsos seja mais evidente; computar, nessa faixa, a média aritmética dos módulos dos coeficientes; comparar os valores dessas médias, obtidas para cada bloco, com um limiar arbitrariamente adotado; e selecionar como corrompidos por pulsos longos os blocos cujos valores das médias frequenciais estiverem acima do valor do limiar.

O Método B requer, além da escolha de um limiar, a escolha de uma faixa de frequências na qual será calculada a média dos módulos dos coeficientes da DFT dos blocos.

A média aritmética dos módulos dos coeficientes da DFT na faixa de interesse é então calculada, em um dado bloco b , por

$$\mu(b) = \frac{1}{\beta - \alpha + 1} \sum_{n=\alpha}^{\beta} |Y_b(n)| \quad (1)$$

onde $Y_b(n)$ é a DFT do sinal discreto $y_b(k)$ de N amostras, dada por

$$Y_b(n) = \sum_{k=0}^{N-1} y_b(k) e^{-j \frac{2\pi kn}{N}}, \quad n = 0 \text{ a } N-1 \quad (2)$$

e $0 \leq \alpha, \beta \leq \frac{N}{2}$ são os índices (em n) dos coeficientes correspondentes à faixa de frequências desejada.

Pela inspeção de espectrogramas de sinais de áudio pode-se constatar que a maior concentração de conteúdo espectral ocorre na região de baixas e médias frequências, até 3 kHz aproximadamente. Esse conteúdo contribui para elevar o valor de $\mu(b)$; entretanto, o que se deseja é a obtenção de valores altos de $\mu(b)$ apenas para os blocos contaminados. Por isso, é conveniente desconsiderar no cálculo a faixa de frequências em que a informação do áudio esteja mais presente. Testes experimentais mostraram que a adoção de uma faixa com frequência inferior em torno de 3 kHz é bastante eficaz para acentuar o contraste entre os blocos com e sem contaminação por pulsos longos.

A Fig. 4 mostra uma versão normalizada (por comodidade de representação) de $\mu(b)$ para o Sinal I em 4 faixas distintas do espectro, comparados com um limiar arbitrário de valor 0,5. Como pode ser observado, em (c) e (d) os blocos corrompidos por pulsos longos ficam bastante destacados em relação aos demais, facilitando a seleção daqueles por comparação com um limiar.

O uso do Método B para seleção dos blocos corrompidos por pulsos longos no Sinal II é tão problemático quanto o uso do Método A. Nesse exemplo, o problema reside no fato de o trecho forte do Sinal II ser composto de sons fortes de um naipe de metais, que exibem harmônicos significativos até frequências muito altas. Esse conteúdo em altas frequências produz valores altos de $\mu(b)$ nos blocos

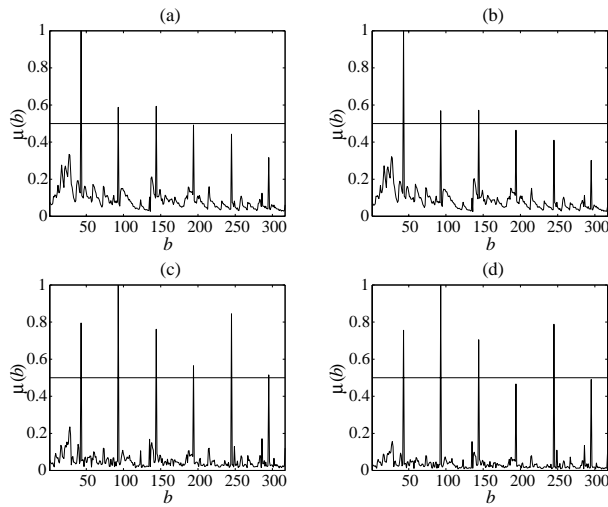


Figura 4: Sinal I, Método B: Limiar de valor 0,5 sobre $\mu(b)$ normalizado calculado entre: (a) 0 e 22,05 kHz; (b) 0 e 10 kHz; (c) 3 e 22,05 kHz e (d) 3 e 10 kHz. Foram utilizados blocos e DFTs de 2048 amostras e superposição de 480 amostras entre blocos.

correspondentes, tornando difícil a seleção dos blocos corrompidos via comparação com um limiar.

A Fig. 5 mostra uma versão normalizada de $\mu(b)$ para o Sinal II em 4 faixas distintas do espectro, comparados com um limiar arbitrário de valor 0,5. Em todos os casos é impossível selecionar os blocos corrompidos por comparação com limiar sem incorrer em um alto índice de falsos alarmes. Resta, assim, solucionar este problema.

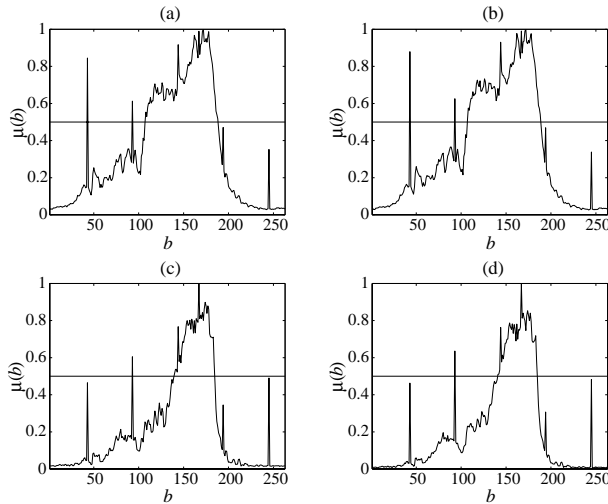


Figura 5: Sinal II, Método B: Limiar de valor 0,5 sobre $\mu(b)$ normalizado calculado entre: (a) 0 e 22,05 kHz; (b) 0 e 10 kHz; (c) 3 e 22,05 kHz e (d) 3 e 10 kHz. Foram utilizados blocos e DFTs de 2048 amostras e superposição de 480 amostras entre blocos.

4.3 Método C

A idéia do Método C é realizar a seleção de blocos contaminados por pulsos longos comparando as potências médias das excitações correspondentes a cada bloco de sinal, modelado como um processo AR (ver Apêndice), com um limiar de valor arbi-

trário. A expectativa é que a ocorrência de valores elevados do sinal de excitação correspondentes aos pulsos longos seja capaz de elevar significativamente a potência do sinal de excitação dos blocos contaminados, destacando-os dos demais.

Para a obtenção do sinal de excitação de cada bloco, é necessário estimar os coeficientes do modelo AR associado ao sinal no bloco e realizar uma filtragem inversa, como exposto no Apêndice.

Obtido o sinal de excitação $e_b(k)$ de um dado bloco b de N amostras, sua potência pode ser calculada por

$$\sigma^2(b) = \frac{1}{N} \sum_{i=0}^{N-1} e_b^2(i). \quad (3)$$

Na Fig. 6, são confrontadas versões normalizadas das seqüências $\sigma^2(b)$ obtidas, respectivamente, para o Sinal I e o Sinal II com limiares arbitrariamente adotados. Embora o Método C tenha sucesso no Sinal I, a seleção dos blocos corrompidos no Sinal II é impossível de ser realizada pela comparação de $\sigma^2(b)$ com um limiar sem incorrer em altos índices de falsos alarmes. A elevação do valor de $\sigma^2(b)$ no trecho forte do Sinal II se deve, em parte, pela ocorrência do fenômeno de subpredição periódica associado a trechos fortemente tonais [3]. Repetese, assim, a dificuldade encontrada no Método B.

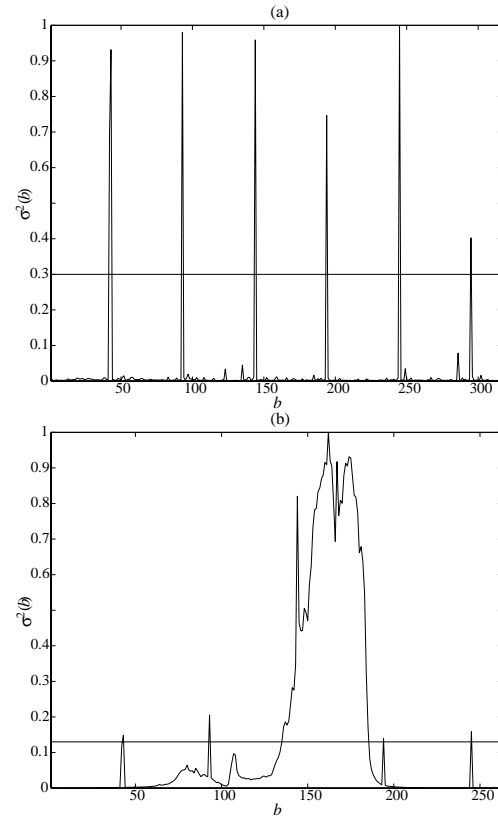


Figura 6: Método C: Limiar λ sobre $\sigma^2(b)$ normalizado. (a) $\sigma^2(b)$ do Sinal I e $\lambda = 0,3$ e (b) $\sigma^2(b)$ do Sinal II e $\lambda = 0,13$. Utilizaram-se blocos de 2048 amostras com superposição de 480 amostras, modelados como processos AR de ordem 40.

4.4 Utilização da Filtragem por Mediana

Para contornar o problema de seleção de blocos com pulsos longos em sinais com dinâmica ampla e conteúdo espectral significativo em altas frequências, propõe-se acrescentar ao Método B (ou C) o seguinte procedimento:

- Obter $\mu(b)$ (ou $\sigma^2(b)$) como já descrito;
- Obter um outro sinal $\mu_m(b)$ (ou $\sigma_m^2(b)$), resultante da filtragem de $\mu(b)$ (ou $\sigma^2(b)$) por mediana com janela de tamanho ímpar igual a c amostras, o que corresponde a substituir cada amostra pela mediana das c amostras nela centradas;
- Obter $\Delta\mu(b) \triangleq \mu(b) - \mu_m(b)$ (ou $\Delta\sigma^2(b) \triangleq \sigma^2(b) - \sigma_m^2(b)$);
- Comparar $\Delta\mu(b)$ (ou $\Delta\sigma^2(b)$) com um limiar λ arbitrariamente adotado e selecionar como corrompidos os blocos de índice b tais que $\Delta\mu(b) \geq \lambda$ (ou $\Delta\sigma^2(b) \geq \lambda$).

A filtragem por mediana tem a propriedade de eliminar ocorrências localizadas de valores atípicos em uma seqüência. Em particular, deve-se observar que a elevação de $\mu(b)$ devido às componentes significativas de alta frequência do sinal, no Método B, é um fenômeno muito “menos localizado” do que a sua elevação devido aos pulsos longos.

A Fig. 7 mostra $\Delta\mu(b)$ e $\Delta\sigma^2(b)$ correspondentes ao Sinal II, ambos comparados com um limiar arbitrário de valor 0,13. Experimentalmente, verificou-se que a filtragem por mediana com $c = 5$ amostras produz resultados bastante satisfatórios, no que diz respeito ao aumento de contraste entre os blocos corrompidos e os blocos não-corrompidos.

4.5 Desempenho dos Métodos de detecção de Pulsos Longos

A avaliação da eficácia dos 3 métodos, de acordo com o exposto anteriormente, aponta os Métodos B e C como os melhores.

Nos Métodos B e C, o fator determinante do sucesso na localização de blocos corrompidos em sinais com dinâmica ampla e conteúdo espectral rico é a utilização do procedimento auxiliar de filtragem por mediana. Nos testes comparativos entre os dois métodos, o tamanho dos blocos, a superposição entre eles e o tamanho da janela do filtro de mediana utilizados foram os mesmos, isto é, respectivamente, 2048, 480 e 5 amostras.

No Método B, o espectro do bloco foi calculado através de DFT de 2048 pontos e a faixa do espectro selecionada foi entre 3 kHz e 10 kHz. Já no Método C, o sinal de cada bloco foi modelado como um processo AR de ordem 40 cujos coeficientes foram calculados por estimativa LS (*Least Squares*). Para essa configuração de processamento e lidando

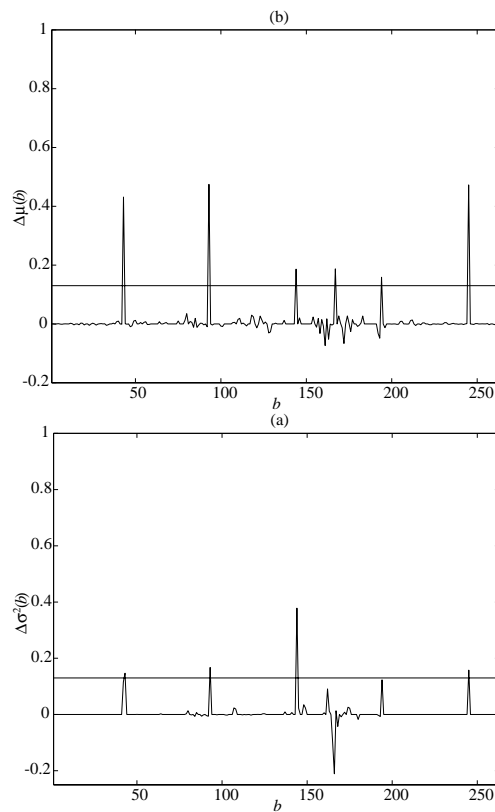


Figura 7: (a) Método B: seqüência $\Delta\mu(b)$ do Sinal II e (b) Método C: seqüência $\Delta\sigma^2(b)$ do Sinal II, obtidas após o emprego auxiliar da Filtragem por Mediana. O limiar de referência, nos dois casos, vale 0,13.

com sinais de dinâmica reduzida e conteúdo espectral pobre, os desempenhos dos Métodos B e C podem ser considerados equivalentes, o que é verdade mesmo sem a filtragem por mediana. Já para sinais com dinâmica ampla e conteúdo espectral rico em altas frequências, como o Sinal II, o Método B tem desempenho superior, como se vê na Fig. 7.

O Método B ainda tem desempenho muito superior em relação ao tempo de processamento, quando comparado ao Método C. Enquanto no Método C é necessário calcular os coeficientes AR do bloco e realizar a filtragem inversa, no Método B basta computar a DFT do bloco. A complexidade computacional para estimar os coeficientes AR de um bloco de N amostras é muito maior que a necessária para computar uma DFT.

Com referência ao Apêndice, para resolver (6) têm que ser computados $\mathbf{X}^T \mathbf{X}$ e $\mathbf{X}^T \mathbf{x}_1$:

- $\mathbf{X}^T \mathbf{X}$ é simétrica de ordem p ; portanto, só há necessidade de calcular $(p^2 + p)/2$ de seus elementos; só para calcular a sua primeira linha, são necessárias $p(N - p)$ multiplicações;
- Para computar $\mathbf{X}^T \mathbf{x}_1$ são necessárias $N - p$ multiplicações.

Para a solução do sistema propriamente dito, são necessárias $p^3/4$ multiplicações [5]. Então, o total subestimado de multiplicações no Método C, até imediatamente antes de se computar a filtragem inversa, é de $p^3/4 - p^2 + p(N - p) + N$, ou seja, para

$p = 40$ e $N = 2048$ são necessárias 98328 multiplicações. A computação da DFT no Método B, por sua vez, pode ser realizada por um algoritmo rápido do tipo *Split-Radix*-FFT para sinais reais, que necessita de $N/2(\log_2(N) - 3) + 2$ multiplicações [6], ou seja, 8194 multiplicações, no caso de $N = 2048$. Isso significa que, para esse exemplo, o Método B é, muito subestimadamente, 12 vezes rápido que o Método C.

É interessante observar que para nenhum dos métodos apresentados foram feitas quaisquer considerações sobre a periodicidade de ocorrência dos pulsos longos, de modo que a ocorrência não periódica dos pulsos, como no Sinal II, é perfeitamente detectada pelo Método B. A similaridade dos pulsos também não é levada em consideração no processo de detecção dos blocos corrompidos, tornando os métodos aqui apresentados mais robustos que o método de detecção apresentado em [1].

5 CONCLUSÕES

Neste trabalho foram discutidos três métodos de detecção de pulsos longos em sinais de áudio em blocos. Foi abordado o problema de detecção de pulsos longos em sinais de dinâmica e conteúdo espectral amplos e proposto um método, baseado em FFT e filtragem por mediana, que explora o contraste entre os conteúdos espectrais dos pulsos e do sinal de áudio. Exemplos práticos ilustram sua eficácia na detecção de pulsos longos sem as restrições de similaridade e espaçamento entre os pulsos.

APÊNDICE – PROCESSO AUTORREGRESSIVO (AR)

Um processo autorregressivo $x(k)$ pode ser caracterizado pela expressão

$$x(k) = \sum_{i=1}^p a(i)x(k-i) + e(k), \quad (4)$$

ou seja, como uma parcela preditível por uma combinação linear das p amostras anteriores ao instante k e uma parcela não-preditível $e(k)$, composta de amostras independentes identicamente distribuídas, chamada de excitação ou erro de predição.

Matricialmente, para um trecho de N amostras, pode-se escrever (4) como

$$\mathbf{e} = \mathbf{x}_1 - \mathbf{X}\mathbf{a}, \quad (5)$$

onde $\mathbf{x}_1 = [x(p) \cdots x(N-1)]^T$, $\mathbf{e} = [e(p) \cdots e(N-1)]^T$, $\mathbf{a} = [a(1) \cdots a(p)]^T$ e

$$\mathbf{X} = \begin{bmatrix} x(p-1) & x(p-2) & \cdots & x(0) \\ x(p) & x(p-1) & \cdots & x(1) \\ \vdots & \vdots & \ddots & \vdots \\ x(N-2) & x(N-3) & \cdots & x(N-p-1) \end{bmatrix}.$$

Uma forma de obter os coeficientes do modelo AR para um trecho de N amostras é por mínimos quadrados (*Least Squares* – LS), que minimiza a função-custo $F = \mathbf{e}^T \mathbf{e}$, resultando na estimativa

$$\hat{\mathbf{a}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{x}_1. \quad (6)$$

Considerando-se conhecidos os coeficientes $a(i)$, deduz-se de (4) que é possível chegar ao sinal de excitação $e(k)$ pela filtragem inversa de $x(k)$ por

$$H(z) = 1 - \sum_{j=1}^p a(j)z^{-j}. \quad (7)$$

O modelamento do sinal de áudio como um processo AR na forma descrita pressupõe sua estacionariedade no sentido amplo por, pelo menos, N amostras. Além disso, o cálculo dos coeficientes do modelo feito diretamente a partir do sinal contaminado (único disponível) considera que os distúrbios presentes no sinal não causam erro significativo. Por fim, a excitação resultante da filtragem inversa aplicada ao sinal contaminado por pulsos longos apresenta versões filtradas dos pulsos, que aparecem bem mais evidenciadas que no sinal original; isto se dá porque onde há distúrbios o modelo falha, causando erro de predição elevado.

REFERÊNCIAS

- [1] VASEGHI, S. V., *Algorithms for Restoration of Archived Gramophone Recordings*. Ph.D. dissertation, University of Cambridge, Cambridge, UK, Feb. 1988.
- [2] GODSILL, S. J., RAYNER, P. J. W., *Digital Audio Restoration*. London, UK, Springer-Verlag, 1998.
- [3] GODSILL, S. J., *The Restoration of Degraded Audio Signals*. Ph.D. dissertation, University of Cambridge, Cambridge, UK, Dec. 1993.
- [4] OPPENHEIM, A., SCHAFER, R., *Discrete-Time Signal Processing*. Englewood Cliffs, NJ, USA, Prentice-Hall, Inc., 1989.
- [5] PRESS, W. H., TEUKOLSKY, S. A., VETTERLING, W. T., *et al.*, *Numerical Recipes in C: The Art of Scientific Computing*. 2 ed. New York, NY, USA, Cambridge University Press, 1992.
- [6] MALVAR, H. S., *Signal Processing with Lapped Transforms*. Norwood, MA, USA, Artech House, 1992.