

# Aspectos da Localização de Distúrbios Impulsivos em Sinais de Áudio Usando Técnicas Baseadas em Limiar

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

DEL da EE & PEE do COPPE, Universidade Federal do Rio de Janeiro  
CP 68564, CEP 21945-970, Rio de Janeiro, RJ  
Fone: (021) 260 5010, Fax: (021) 590 0788  
{wagner, diniz, esquef, freeland}@lps.ufrj.br

*Sumário — Este artigo aborda algoritmos clássicos para detecção de ruído impulsivo em áudio, baseados em limiar. Visando a melhorar seu desempenho, propõe a adoção de dois limiares, um para detectar e outro para estimar a duração dos distúrbios; reexamina o cálculo do limiar; adota a fusão de distúrbios adjacentes; e discute critérios para iterar o processamento em cada bloco de sinal. A superioridade dos resultados obtidos após as modificações é confirmada por avaliação humana e por medidas objetivas.*

## 1 INTRODUÇÃO

Um dos maiores problemas associados a gravações musicais históricas é a degradação do sinal de áudio por distúrbios impulsivos. Decorrentes de imperfeições, rugosidades ou material agregado nos sulcos das matrizes disponíveis, os distúrbios impulsivos podem ser caracterizados como descontinuidades localizadas de curta duração (tipicamente menores que 1 ms) que corrompem o sinal de áudio em instantes de tempo aleatoriamente distribuídos ao longo do sinal. Apesar da denominação **ruído impulsivo**, os distúrbios impulsivos que corrompem o sinal de áudio ocorrem em salvas de amostras contíguas, com durações e amplitudes bastante variadas, conforme a extensão física do defeito na superfície da matriz.

Um modelo genérico para esse tipo de degradação consiste em considerar a corrupção por ruído impulsivo como aditiva, ou seja, o sinal corrompido como

$$y(k) = x(k) + d(k), \quad (1)$$

onde  $x(k)$  é o sinal não-corrompido e  $d(k)$ , o ruído impulsivo. O ruído impulsivo também admite um modelo genérico  $d(k) = i(k)n(k)$ , onde  $i(k)$  é um processo aleatório de chaveamento que assume valor 1 quando o distúrbio está presente e 0 em caso contrário. As amplitudes dos distúrbios são regidas pelo processo aleatório  $n(k)$ .

O percentual de corrupção do sinal de áudio pelo ruído impulsivo pode variar bastante de um si-

nal degradado para outro, de acordo com o tipo da matriz e com seu estado de conservação. Por exemplo, um sinal degradado já discretizado com 10% das amostras corrompidas por ruído impulsivo é considerado severamente degradado, se todos os distúrbios forem perceptíveis.

Técnicas de processamento digital visando à supressão de ruído impulsivo em sinais de áudio discretizado podem, em geral, ser divididas em duas etapas: Detecção e Reconstrução. Na etapa de Detecção, os distúrbios impulsivos são localizados temporalmente ao longo do sinal. A etapa de Reconstrução recebe a informação sobre a localização dos distúrbios e reconstrói o sinal de áudio nessas regiões utilizando, por exemplo, um algoritmo de interpolação apropriado para substituição das amostras corrompidas por outras que melhor aproximem o comportamento local do sinal de áudio não-corrompido. A qualidade final do sinal de áudio restaurado depende, obviamente, do bom desempenho de ambas as etapas.

Na Seção 2 é revisada uma técnica para detecção de ruído impulsivo, baseada em limiar, chamada, aqui, de Detecção Convencional. Na Seção 3, são descritas as proposições de modificações na Detecção Convencional. Na Seção 4, o desempenho das Detecções Modificada e Convencional é avaliado através de medidas quantitativas.

## 2 DETECÇÃO POR LIMIAR

No método de Detecção apresentado em [1], aqui denominado de Convencional, o processamento é realizado em blocos de sinal de tamanho fixo  $N$ , nos quais o sinal não-degradado  $x(k)$  pode ser considerado estacionário e bem modelado por um processo Autorregressivo (AR) de ordem  $p$ . Cada amostra do sinal nos instantes  $k = p, p + 1, \dots, N - 1$  é dada por

$$x(k) = \sum_{i=1}^p a(i)x(k-i) + e(k), \quad (2)$$

sendo  $a(i)$  os coeficientes do modelo e  $e(k)$  um sinal com amostras i.i.d., denominado de excitação ou erro de predição.

Para sinais degradados por ruído impulsivo segundo o modelo aditivo da eq. 1, é mostrado em [1] que a razão ruído-sinal na excitação é maior que a no sinal propriamente dito, fato que favorece a detecção dos distúrbios impulsivos na excitação. O sinal degradado é, então, levado à excitação pelo Filtro Inverso

$$H(z) = 1 - \sum_{j=1}^p \hat{a}(j)z^{-j}, \quad (3)$$

que depende de uma estimativa robusta dos coeficientes do modelo AR do sinal não-degradado,  $\hat{a}(j)$ , que pode ser obtida a partir do sinal corrompido  $y(k)$ .

A detecção dos distúrbios impulsivos é realizada comparando-se o módulo de  $e(k)$  com um limiar dado por

$$\lambda = K \hat{\sigma}_{e_x}, \quad (4)$$

onde  $\hat{\sigma}_{e_x}^2$  é uma estimativa da variância da parcela do sinal de excitação correspondente ao sinal não-corrompido e  $K$  é um parâmetro de ajuste do compromisso entre não-deteção e falsa detecção de distúrbios.

Os distúrbios impulsivos num bloco podem ser detectados, basicamente, de duas formas. Na primeira forma [1], os distúrbios são detectados um a um, começando pelo de maior amplitude na excitação, através dos passos descritos a seguir:

- Detecção do distúrbio de maior amplitude na excitação pelo índice  $k$  para o qual  $|e(k)|$  é máximo, com  $|e(k)| > \lambda$ ;
- Arbítrio da duração do distúrbio: são dadas como corrompidas algumas (um número fixo para todos os distúrbios) amostras adjacentes ao distúrbio unitário detectado;
- Mapeamento dos índices das amostras dadas como corrompidas na excitação para os índices correspondentes no sinal propriamente dito;
- Supressão do distúrbio (o que vai minimizar a excitação no trecho corrigido);
- Repetição do procedimento para o próximo distúrbio até que  $|e(k)| < \lambda, \forall k$ .

A segunda forma de detecção de distúrbios [2] consiste em detectar conjuntamente os distúrbios num bloco da excitação, tomando como corrompidas todas as amostras de índice  $k$  tais que  $|e(k)| > \lambda$ . Nessa abordagem, a detecção e a determinação da duração dos distúrbios são realizadas em um único passo. Como na forma anterior, os índices das amostras corrompidas na excitação são mapeados nos índices do sinal propriamente dito, onde a reconstrução é realizada.

Embora existam diversos algoritmos voltados à interpolação de amostras corrompidas em sinais de áudio [3], interpoladores baseados em modelagem

AR para o sinal de áudio têm-se mostrado bastante eficientes na reconstrução de sinais de áudio corrompidos por ruído impulsivo. Neste trabalho, utilizou-se como algoritmo de interpolação, unicamente, o *Least Squares Autorregressive* (LSAR), descrito em [4].

### 3 MODIFICAÇÕES NA DETECÇÃO CONVENCIONAL

O desempenho do método de Detecção Convencional pode ser melhorado através de modificações de motivação empírica em alguns aspectos de sua implementação, objetivando, principalmente, uma maior eficácia na detecção/localização dos distúrbios impulsivos. Inicialmente, o cálculo do limiar é examinado. Em seguida, são propostas a utilização de um segundo limiar para auxiliar na localização dos distúrbios; critérios para união de distúrbios adjacentes; e critérios de iteração do processamento num mesmo bloco de sinal.

#### 3.1 Cálculo do Limiar

Esta seção examina o cálculo do limiar com o intuito de se obter um compromisso satisfatório entre os índices de falsa detecção e não-deteção de distúrbios impulsivos, ao longo de todos os blocos do sinal de áudio, a despeito de sua característica altamente não-estacionária e da diversidade de amplitude e duração dos distúrbios impulsivos corruptores, presentes em situações reais.

Em [1], a adaptação do valor do limiar às características locais de um determinado bloco é realizada fazendo o limiar proporcional a uma estimativa do desvio-padrão do sinal de excitação do bloco, como na eq. (4). A escolha do parâmetro  $K$  — fixo para todos os blocos — é experimental e realiza um ajuste no compromisso entre falsa detecção e não-deteção, como mencionado anteriormente.

A forma mais direta de se obter uma estimativa da variância da excitação correspondente ao sinal não-ruído é calculá-la pela excitação do sinal ruído. Entretanto, essa estimativa está muito sujeita a polarização, devido à presença do ruído impulsivo.

As investigações experimentais realizadas em [5] mostram que a presença dos distúrbios no sinal de excitação tende a elevar o valor do desvio-padrão (conseqüentemente o do limiar) nos trechos de baixa potência, produzindo altos índices de não-deteção nesses trechos. Se, no intuito de compensar esse efeito, o valor de  $K$  for reduzido, obtêm-se altos índices de falsa detecção nos trechos de alta potência ou de transitórios bruscos, visto que, nesses trechos, é menor a influência dos distúrbios impulsivos na estimativa do desvio-padrão da excitação.

A consecução de um melhor compromisso entre falsa detecção e não-deteção passa por uma forma de cálculo do limiar que seja tanto quanto possível

imune aos distúrbios impulsivos presentes na excitação. Nesses termos, uma alternativa que se mostra bastante robusta na estimação da variância da excitação consiste em:

- Armazenar em um vetor o quadrado das amostras da excitação, organizadas em ordem crescente de valor;
- Desprezar um determinado número de amostras de maior valor contidas no final do vetor ordenado, que provavelmente correspondem às amostras do ruído impulsivo presentes na excitação;
- Utilizar como estimativa da variância da excitação do sinal não-ruído a média aritmética do vetor reduzido.

Na realidade, o número real de amostras corrompidas pode variar muito de bloco para bloco. Na impossibilidade de conhecê-las de antemão, arbitra-se um número máximo de amostras corrompidas por bloco — algo em torno de 10% — e descarta-se esse número de amostras do final do vetor ordenado, como descrito anteriormente. Uma opção mais enxuta consiste em tomar o desvio-padrão da excitação como proporcional à mediana do vetor de amostras do módulo da excitação do sinal corrompido. Dessa forma, o valor do limiar fica determinado por

$$\lambda = K \text{Mediana} [|e(k)|]. \quad (5)$$

De fato, se se considerar o sinal de excitação como um ruído branco Gaussiano estacionário, mostra-se em [6] que uma estimativa robusta de seu desvio-padrão é dada por  $\sigma_e = \text{Mediana} [|e(k)|] / 0,6745$ . Apesar de a consideração anterior ser um tanto simplista, a utilização de (5) se mostrou menos suscetível à presença de ruído impulsivo na excitação.

### 3.2 Uso de 2 Limiares

A determinação precisa da duração dos distúrbios é um requisito muito importante no processo de restauração; entretanto, a grande variedade de amplitudes e durações dos distúrbios e o espalhamento produzido pela filtragem inversa nos distúrbios presentes no sinal de excitação [4], onde serão detectados e localizados, tornam essa tarefa bastante difícil de se realizar.

A atribuição de uma duração fixa para todos os distúrbios [1] é uma escolha que não condiz com a variedade de duração dos distúrbios impulsivos observados na realidade. De fato, a escolha de uma duração fixa que contemple distúrbios de maior duração leva, invariavelmente, à correção de um grande número de amostras não corrompidas no entorno dos distúrbios de menor duração. Por sua vez, o critério de detecção e localização em um único passo, dado pela escolha dos  $k$  tais que  $|e(k)| > \lambda$  [4], pode levar a durações subestimadas para alguns distúrbios, principalmente os de menores amplitudes. Essas determinações subestimadas de duração

produzem, em geral, correções insatisfatórias na etapa de Reconstrução do sinal de áudio.

A redução do valor do limiar contribuiria, certamente, para minimizar o problema de determinação de durações subestimadas; entretanto, pode acarretar sérios prejuízos pelo aumento do índice de falsas detecções. O que se propõe, então, é utilizar a detecção e a localização em dois passos como em [1], mas, ao invés de arbitrar uma duração fixa, estimar a duração dos distúrbios com o auxílio de um segundo limiar  $\lambda_L$  (denominado de Limiar de Localização), menor que o limiar original  $\lambda_D$  (agora denominado de Limiar de Detecção) por um fator  $b$ , ou seja,

$$\lambda_L = b \lambda_D. \quad (6)$$

Nesses termos, os distúrbios são detectados um a um, começando pelo de maior amplitude na excitação, através dos passos descritos a seguir:

- Detecção do distúrbio de maior amplitude na excitação pelo índice  $k$  correspondente ao máximo valor de  $|e(k)|$ , com  $|e(k)| > \lambda_D$ ;
- Determinação da duração do distúrbio pelas amostras contíguas ao distúrbio detectado que exibirem módulo (na excitação) maior que o valor do Limiar de Localização  $\lambda_L$ ;
- Emulação da supressão do distúrbio (“zerando” a excitação no trecho considerado corrompido);
- Repetição do procedimento para o próximo distúrbio até que  $|e(k)| < \lambda_D, \forall k$ .
- Mapeamento dos índices das amostras corrompidas na excitação para os índices correspondentes no sinal propriamente dito.

A Fig. 1 ilustra esse procedimento de detecção e localização para um distúrbio impulsivo.

A supressão de cada distúrbio localizado poderia ser feita individualmente, executando-se a etapa de Reconstrução a cada novo distúrbio localizado. Isso se encarregaria de minimizar o trecho da excitação correspondente ao distúrbio e permitiria a detecção dos distúrbios de menores amplitudes. Entretanto, é mais eficiente, computacionalmente, determinar a localização de todos os distúrbios detectáveis em um bloco e promover uma supressão simultânea de todos eles.

### 3.3 União de Distúrbios Adjacentes

O espalhamento dos distúrbios no sinal de excitação, produzido pela filtragem inversa, pode produzir interferência construtiva ou destrutiva entre amostras de um mesmo distúrbio ou de distúrbios próximos [2]. Não é raro o sinal de excitação correspondente a um distúrbio relativamente duradouro exibir amostras intermediárias com módulos menores que o valor do limiar de localização, devido à interferência destrutiva, como mostra a Fig. 2 (b)

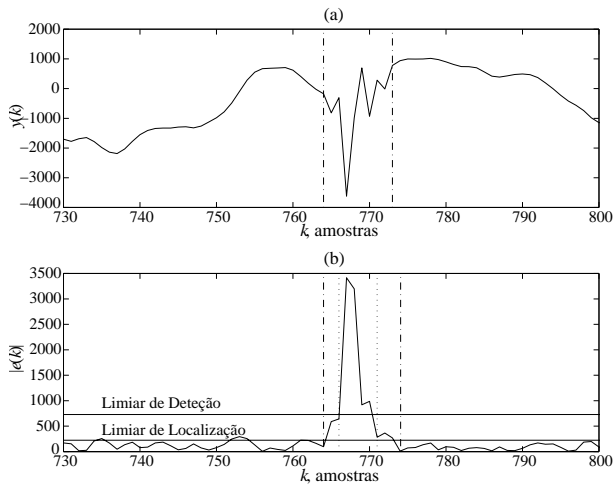


Figura 1: (a) Trecho de sinal de áudio corrompido por distúrbio impulsivo – delimitado pelas linhas mistas verticais. (b) Módulo da excitação e Limiares de Detecção  $\lambda_D$  e Localização  $\lambda_L$  – as linhas mistas verticais delimitam o distúrbio localizado via  $\lambda_L$ , enquanto as linhas pontilhadas delimitam o distúrbio, caso fosse localizado via  $\lambda_D$ .

— observar a magnitude da excitação em torno da posição  $k = 300$ . Mesmo utilizando dois limiares, pode-se obter uma determinação imprecisa da duração do distúrbio ou, o que é pior, sua detecção na forma de dois ou mais distúrbios, separados por amostras consideradas não-corrompidas, mas que, de fato, são corrompidas.

Um critério que se mostra bastante eficiente na solução desses casos é promover uma união forçada de distúrbios adjacentes separados por até  $n$  amostras. Um procedimento possível consiste em incorporar ao distúrbio detectado amostras adjacentes à de maior módulo, até que haja, em ambos os sentidos, pelo menos  $n$  amostras consecutivas da excitação com módulos menores que o limiar de localização. Dessa forma, todo distúrbio localizado é ladeado, garantidamente, por, pelo menos,  $n$  amostras contíguas da excitação com módulos menores que o limiar de localização.

Em geral, unir distúrbios separados por 2 ou 3 amostras, a uma taxa de amostragem de 44100 Hz, parece bastante razoável; entretanto, para distúrbios de maior duração (mais de 30 amostras), esse número pode ser insuficiente. Uma opção é fazer  $n$  dependente da duração  $l$  que se determinaria para o distúrbio se não houvesse qualquer critério de união. O valor variável de  $n$  pode ser ditado, por exemplo, por

$$n = n_{\min} + \left\lfloor \frac{l}{g} \right\rfloor, \quad (7)$$

onde  $n_{\min}$  é o menor valor inteiro arbitrariamente admitido para  $n$ ,  $g$  é um fator de controle do aumento de  $n$  em relação ao aumento de  $l$  e  $\lfloor \cdot \rfloor$  é o operador “maior inteiro menor que ou igual a”.

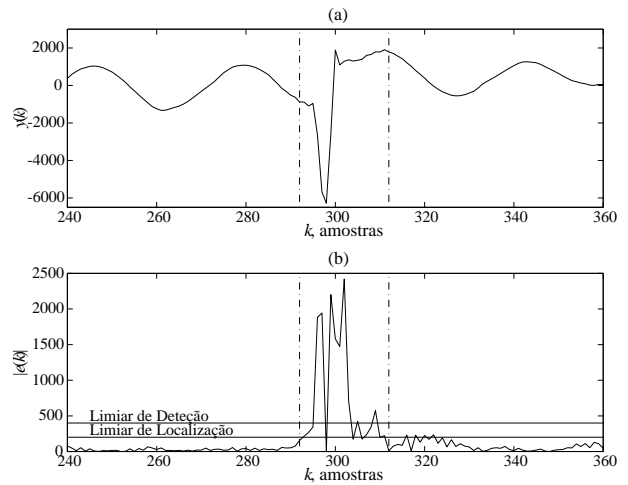


Figura 2: (a) Sinal corrompido – distúrbio impulsivo delimitado, por inspeção visual, pelas linhas mistas verticais. (b) Módulo da excitação e limiares de detecção e localização – o distúrbio localizado usando união de distúrbios adjacentes ( $n = 2$ ) é delimitado pelas linhas mistas verticais. Observe que, sem a união de distúrbios adjacentes, seriam localizados 4 distúrbios distintos.

### 3.4 Iteração de Processamento num mesmo Bloco

Uma vez que as técnicas de detecção descritas nas seções anteriores requerem uma implementação em blocos, é plausível pensar em iterar, num mesmo bloco, as etapas de Detecção e Reconstrução, visto que uma rodada apenas pode não ser suficiente para detectar todos os distúrbios impulsivos presentes, principalmente os de menor amplitude. Sabendo que a etapa de Reconstrução tende a minimizar os trechos do sinal de excitação correspondentes aos distúrbios impulsivos suprimidos, é de se esperar que trechos do sinal já recuperados não sejam mais detectados numa nova etapa de Detecção (utilizando idênticos parâmetros de processamento). É, então, inteiramente justificável adotar a ocorrência  $|e(k)| < \lambda_D, \forall k$ , como critério de interrupção das iterações num mesmo bloco. Entretanto, não há garantia de que tal critério (Critério 1) será sempre satisfeito para todos os blocos a serem processados.

De fato, quando alguns distúrbios, em geral de pequena amplitude, são detectados com duração subestimada, a etapa de reconstrução não produz a desejada minimização do sinal de excitação no trecho correspondente, tornando-os indefinidamente detectáveis. Uma forma de contornar esse problema é fazer com que as iterações sejam interrompidas quando o Critério 1 for satisfeito **ou** quando o número de iterações atingir uma contagem arbitrária  $i_{\max}$ . Esse critério misto será chamado de Critério 2.

Observações experimentais mostraram que, nas primeiras iterações realizadas em um bloco, os distúrbios de maiores duração e amplitude são suprimidos satisfatoriamente, restando por serem suprimidos os de menores amplitudes. Esses distúrbios

têm maiores chances de serem localizados com durações subestimadas e de produzirem rodadas inócuas do processamento. Isso pode ser evitado promovendo-se uma localização mais apurada desses distúrbios, que passa, necessariamente, pela redução do valor do limiar de localização. Essa redução pode ser realizada diminuindo-se a razão  $b$  entre os limiares de localização e detecção conforme o número de iterações  $i$  cresce. Uma forma de implementar essa redução consiste em fazer

$$b[i] = r^{\lfloor \frac{i}{f} \rfloor} \cdot b[1], \quad i = 2 \text{ a } i_{\text{máx}}, \quad (8)$$

onde  $r$  é o fator de redução e  $f$ , um parâmetro que controla a frequência com que  $b$  é alterado. A escolha dos parâmetros  $i_{\text{máx}}$ ,  $f$  e  $r$  é arbitrária e está relacionada a observações experimentais.

A utilização de  $b$  dependente da iteração realizada, além de proporcionar a supressão de distúrbios de menor amplitude, também evita repetições inócuas de processamento, contribuindo, dessa forma, para menores tempos de processamento.

#### 4 COMPARAÇÃO DOS RESULTADOS

A avaliação comparativa dos desempenhos das Detecções Convencional e Modificada foi realizada em um conjunto de sinais de áudio contaminados artificialmente por duas classes de ruído impulsivo. O ruído Classe 1 contém distúrbios impulsivos com amplitudes e durações muito variadas e, quando somado a um sinal de áudio, chega a corromper aproximadamente 0.6% de suas amostras. O ruído Classe 2 contém distúrbios de menores duração e amplitude, entretanto; quando somado a um sinal de áudio, é capaz de corromper aproximadamente 5% de suas amostras. Cada classe foi obtida pela diferença entre um sinal de áudio naturalmente corrompido e uma versão já restaurada sua. As amplitudes dos sinais de ruído também foram ajustadas de acordo com as dos sinais de áudio, de modo a tornar o ruído claramente audível. O método de Reconstrução utilizado em todos os casos foi o LSAR [4].

Para avaliar a eficácia da utilização de dois limiares e do critério de união de distúrbios adjacentes, as duas versões da Etapa de Detecção foram implementadas. A Modificada, denominada de EDM, utiliza os parâmetros e critérios descritos a seguir:

- Processamento em blocos de 1024 amostras;
- Ordem do modelo AR igual a 40;
- $\lambda_D$  calculado por (5), com  $K = 5$ ;
- União de distúrbios separados por  $n$  amostras segundo (7), com  $n_{\text{mín}} = 2$  e  $g = 6$ ;
- Critério 2 para a interrupção das iterações num dado bloco de sinal, com  $i_{\text{máx}} = 7$ ;
- Redução progressiva do valor de  $\lambda_L$  segundo (8), com  $b[1] = r = 0.5$  e  $f = 3$ ;

A Detecção Convencional, denominada de EDC, só difere de EDM pela utilização de limiar único e pela não-utilização do critério de união de distúrbios adjacentes. O limiar único foi obtido estabelecendo  $b = 1$  em (6) e  $r = 1$  em (8), o que é equivalente a fazer  $\lambda_L = \lambda_D$ .

A avaliação qualitativa dos sinais de áudio restaurados é, em última análise, uma questão subjetiva associada aos resultados perceptivos produzidos pelo processamento. Neste trabalho, os sinais originalmente não-corrompidos estão disponíveis para permitir o uso de medidas quantitativas que possam corroborar os resultados perceptivos obtidos. Medidas como as taxas de não-deteção e de falsa deteção são simples e eficazes para esse propósito.

Sejam  $n_o(k)$  e  $n_p(k)$ , respectivamente, o ruído que corrompe os sinais de áudio e o ruído residual após o processamento — obtido pela diferença entre as versões processada e não-ruidosa do sinal de áudio. Como detalhado em [7], as percentagens de não-deteção são calculadas como a razão entre o número de amostras não-nulas em  $n_o(k)$  que permanecem intocadas em  $n_p(k)$  e o número total de amostras não-nulas em  $n_o(k)$ ; as percentagens de falsa deteção são calculadas como a razão entre o número de amostras nulas em  $n_o(k)$  que se tornam não-nulas em  $n_p(k)$  e o número total de amostras nulas em  $n_o(k)$ .

A Tabela 1 mostra os Percentuais de Falsas Detecções (PFD) e Não-deteções (PND) obtidos utilizando-se as versões EDM e EDC da etapa de Detecção. Os Sinais 1 a 4 foram corrompidos pelo sinal de ruído da Classe 1, enquanto que os Sinais 5 e 6 foram corrompidos pelo sinal de ruído da Classe 2. É importante enfatizar que tanto os parâmetros usados em EDM como os usados em EDC foram mantidos fixos para processar os seis sinais.

	PND %		PFD %	
	EDM	EDC	EDM	EDC
Sinal 1	1.559	16.413	3.264	1.165
Sinal 2	2.095	20.388	1.560	0.678
Sinal 3	1.563	10.705	1.831	0.873
Sinal 4	4.210	21.520	2.400	0.930
Sinal 5	1.627	14.864	6.653	3.541
Sinal 6	1.848	12.112	8.002	4.448

Tabela 1: Medidas Comparativas dos percentuais de falsa deteção e não-deteção relativos às versões restauradas usando EDM e EDC.

Todos os sinais restaurados utilizando EDM apresentaram resultados perceptivamente superiores em relação aos restaurados usando EDC, a despeito da diversidade dos sinais de áudio avaliados, das duas classes distintas de ruídos corruptores e da manutenção dos mesmos parâmetros para processar todos os sinais. Isso indica que as modificações propostas aumentam a robustez da etapa de Detecção quanto a variações nas características dos sinais de áudio e de ruído.

As medidas de percentual de não-deteccção associadas à EDM indicam uma redução por um fator entre 5 e 10 no índice de não-deteccção em relação às mesmas medidas associadas à EDC. Já as medidas de percentual de falso alarme associadas à EDM indicam um aumento por um fator entre 2 e 3 no índice de falsa deteção em relação às mesmas medidas associadas à EDC. A desejável redução no índice de não-deteccção é, de fato, responsável pelos resultados perceptivamente superiores dos sinais restaurados utilizando EDM. O aumento do índice de falsa deteção, idealmente indesejável, era esperado, visto que as modificações propostas tendem a superestimar, ligeiramente, a duração de alguns distúrbios. Os resultados perceptivos, entretanto, mostram que tais durações superestimadas não afetam de modo significativo a qualidade do sinal de áudio; ao contrário, são capazes de evitar restaurações insatisfatórias, como a mostrada na Fig. 3 (c).

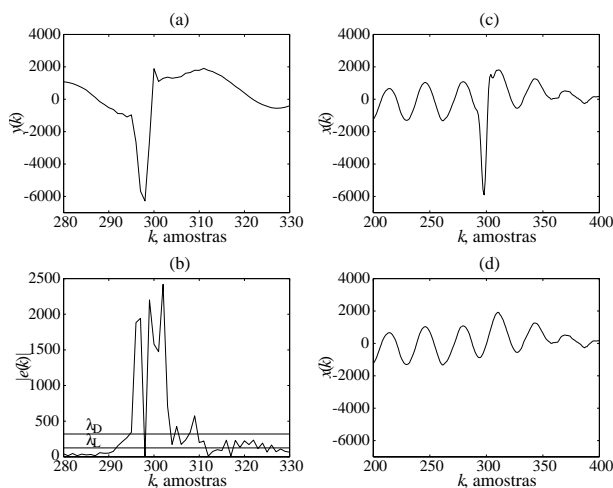


Figura 3: (a) Detalhe do distúrbio impulsivo corrompendo o sinal de áudio. (b) Excitação do sinal mostrado em (a) e limiares de deteção,  $\lambda_D$ , e localização,  $\lambda_L$ . (c) Sinal restaurado usando EDC. (d) Sinal restaurado usando EDM.

Apesar de os resultados apresentados nesta seção serem relativos a sinais de áudio artificialmente corrompidos, as modificações propostas para a etapa de Detecção produzem resultados igualmente satisfatórios na restauração de sinais naturalmente corrompidos por ruído impulsivo.

Todos os exemplos mencionados nesta Seção estão disponíveis em <http://www.lps.ufrj.br/audiobr>.

## 5 CONCLUSÕES

Foram apresentadas neste artigo modificações de motivação empírica no método convencional de deteção de ruído impulsivo em blocos de sinais de áudio, baseado em limiar. Do exame do cálculo do limiar (dependente da variância da excitação do sinal não-corrompido) optou-se pela utilização do estimador Mediana sobre a excitação do bloco de sinal corrompido, por ser essa estimativa pouco polarizável pelos distúrbios impulsivos presentes

na excitação. As modificações propostas incluem a utilização de 2 limiares — um para a deteção e outro (de menor valor) para a localização dos distúrbios —, como forma de evitar que distúrbios fossem localizados com durações subestimadas; um critério de união de distúrbios adjacentes capaz de contornar as localizações imprecisas produzidas por interferências destrutivas na excitação, que mascaram a duração real de alguns distúrbios; e critérios relacionados à iteração do processamento num mesmo bloco de sinal, no qual se propõe a redução gradativa do limiar de localização conforme o avançar da contagem das iterações.

Um sistema completo de restauração de sinais corrompidos por ruído impulsivo, contando com duas versões da etapa de Detecção (Convencional e Modificada), foi implementado e a eficiência na deteção de distúrbios impulsivos foi medida por critérios objetivos (percentuais de falsa deteção e não-deteção). A despeito da variedade de sinais de áudio corrompidos avaliados, a utilização das modificações propostas reduz sensivelmente o percentual de não-deteção ao custo de um pequeno aumento (já esperado) no percentual de falsa deteção. Esse aumento, entretanto, não compromete a qualidade final dos sinais restaurados e, associado à desejável redução do percentual de não-deteção, produz sinais restaurados com qualidade perceptiva superior aos restaurados usando deteção convencional, como comprovam as avaliações auditivas.

## 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] VELDHUIS, R., *Restoration of Lost Samples in Digital Signals*. Prentice-Hall, 1990.
- [4] GODSILL, S. J., *The Restoration of Degraded Audio Signals*. Ph.D. dissertation, University of Cambridge, Cambridge, UK, Dec. 1993.
- [5] ESQUEF, P. A. A., *Implementação e Investigação Comparativa de Técnicas para Restauração Digital de Áudio*. Projeto Final, EE/UFRJ, Rio de Janeiro, RJ, Fev. 1997.
- [6] DONOHO, D. L., *Wavelet Shrinkage and W.V.D.: a 10-Minute Tour*, Report 416, Department of Statistics, Stanford University, Jun. 1993.
- [7] ESQUEF, P. A. A., *Restauração de Sinais de Áudio Degradados por Ruído Impulsivo*. Tese de M.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Mar. 1999.