

Mônica Ximenes Carneiro da Cunha

Técnicas Seletivas de Filtragem Adaptativa de
Sinais de Áudio Degradados por Ruído
Impulsivo

Dissertação submetida ao corpo docente da Coordenação dos Cursos de Pós-Graduação em Engenharia Elétrica da Universidade Federal da Paraíba - Campus II como parte dos requisitos necessários para obtenção do grau de Mestre em Engenharia Elétrica.

Benedito Guimarães Aguiar Neto - Dr.-Ing., UFPB
Orientador

Campina Grande, Paraíba, Brasil

©Mônica Ximenes Carneiro da Cunha. 1995



C973t Cunha, Mônica Ximenes Carneiro da.
Técnicas seletivas de filtragem adaptativa de sinais de áudio degradados por ruído impulsivo / Mônica Ximenes Carneiro da Cunha. - Campina Grande, 1995.
144 f.

Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal da Paraíba, Centro de Ciências e Tecnologia.


1. Processamento de Sinais. 2. Sinais de Áudio Degradados. 3. Redução de Ruído. 4. Dissertação - Engenharia Elétrica. I. Aguiar Neto, Benedito Guimarães. II. Universidade Federal da Paraíba - Campina Grande (PB). III. Título

CDU 004.383.3(043)

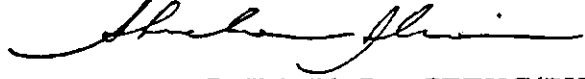
TÉCNICAS SELETIVAS E FILTRAGEM ADAPTATIVA DE SINAIS DE ÁUDIO
DEGRADADOS POR RUÍDO IMPULSIVO

MÔNICA XIMENES CARNEIRO DA CUNHA

Dissertação aprovada em 02.06.1995


BENEDITO GUIMARÃES AGUIAR NETO, Dr.Ing., UFPB
Orientador


JOÃO MARQUES DE CARVALHO, Ph.D., UFPB
Componente da Banca


ABRAHAM ALCAIM, Ph.D., CETUC/PUC-RJ
Componente da Banca

CAMPINA GRANDE - PB
Junho - 1995

Dedico este trabalho a meus pais, Silvio e Wilma,

e ao meu avô Isaac (*in memoriam*).

“O Senhor é o meu rochedo, e o meu
lugar forte, e o meu libertador;
o meu Deus, a minha fortaleza,
em quem confio; o meu escudo,
a força da minha salvação,
e o meu alto refúgio.” Salmo 18:2

Agradecimentos

A Deus, razão do meu viver, amigo fiel e companheiro de todas as horas, por mais uma vitória.

Aos meus pais, Silvio e Wilma, bases do meu caráter e formação, por todo apoio, incentivo, compreensão e dedicação ao longo da minha existência.

Aos meus irmãos, Silvio Filho, Isabel Cristina e Carlos Alberto, pela amizade, carinho e grande apoio em todos os momentos.

Ao orientador e amigo, Benedito Guimarães Aguiar Neto, por todo apoio técnico, sugestões e incentivo durante o desenvolvimento deste trabalho.

Aos meus grandes amigos: Renata Santos, Erika Alves, Giovanni Lyra, Magno Gomes, Alécio Barreto, Glaucio Fontgalland e Jordan Paiva, pelo estímulo e amizade nas horas mais difíceis.

Aos professores Rosângela França, Antônio Marcos, Ângelo Perkusich, Robert Brandt, Elmar Melcher, João Marques e Fátima Turnell pelos ensinamentos, sugestões e estímulo.

Ao amigo e engenheiro do LAPS, Rinaldo Santos, pela presteza e suporte na parte das implementações durante todo o desenvolvimento deste trabalho.

Aos amigos e técnicos do LAPS, Paulo Márcio e José Wallington, pelo constante suporte técnico, que viabilizou imensamente este trabalho.

Aos amigos do LAPS: Joseana Fachine, Paulo Cortez, Leocarlos Lima, Aldenor Falcão, Francisco Madeiro, Silvana Porto e Lírida Naviner pelo companheirismo, ajuda e sugestões.

A Ângela Ribeiro, por sua amizade, eficiência, dedicação e presteza junto à COPELE.

À CAPES, órgão financiador desta pesquisa, e à Universidade Federal da Paraíba - Campus II, pela oportunidade oferecida.

Resumo

Esta dissertação apresenta, implementa e compara algumas técnicas adaptativas seletivas de redução de ruído impulsivo em sinais de áudio degradados. Os sistemas propostos são compostos de duas etapas: detecção e filtragem. A etapa de detecção é baseada no cálculo da derivada da envoltória do sinal e comparação com um limiar pré-estabelecido, determinado a partir das características do sinal. A etapa de filtragem, por sua vez, é baseada na predição linear adaptativa progressiva, regressiva e dupla. Um pós-processamento, baseado na interpolação polinomial de Lagrange, aumenta o desempenho do sistema, proporcionando um maior ganho na relação sinal-ruído (SNR). Os resultados obtidos através do ganhos na SNR do sinal processado e confirmados através de testes de escuta informais mostraram-se bastante satisfatórios, uma vez que apresentam uma melhoria considerável na qualidade do sinal.

Sumário

1	Introdução	1
1.1	Organização da Dissertação	3
2	Técnicas de Supressão de Ruído em Sinais de Áudio Degradados	4
2.1	Introdução	4
2.2	Sistemas Complementares	6
2.3	Sistemas Não Complementares	9
2.3.1	Técnicas Não Seletivas de Redução de Ruídos	12
2.3.2	Técnicas Seletivas de Redução de Ruídos	16
3	Codificação por Predição Linear	19
3.1	Introdução	19
3.2	Princípios Básicos da Predição Linear	20
3.3	Cálculo dos Coeficientes do Preditor	24
4	Descrição do Sistema de Redução de Ruídos Impulsivos	26
4.1	Algoritmo de Detecção	26
4.1.1	Cálculo do Desvio Padrão	27
4.1.2	Cálculo do Limiar	27

4.1.3	Filtro Passa-Altas	28
4.1.4	Detector de Envoltória	32
4.1.5	Derivada da Envoltória e Comparação com Limiar	33
4.2	Filtragem: Algoritmo de Estimação do Sinal	33
4.2.1	Predição Linear Progressiva	34
4.2.2	Predição Linear Regressiva	35
4.2.3	Predição Linear Dupla	36
4.2.4	Pós-processamento do Sinal Estimado: Interpolação Polinomial de Lagrange	37
5	Simulação e Avaliação do Sistema de Filtragem do Ruído Impulsivo	43
5.1	Introdução	43
5.2	Condições Experimentais	44
5.2.1	Digitalização do Sinal	44
5.2.2	Escolha da Ordem do Preditor e do Tamanho do Segmento . . .	44
5.3	Procedimento Experimental	46
5.3.1	Etapa de Detecção	50
5.3.2	Etapa de Filtragem	56
5.4	Avaliação dos Resultados	65
5.4.1	Descrição dos Métodos de Avaliação Utilizados	65
5.4.2	Apresentação dos Resultados	67
5.4.3	Avaliação no Domínio da Frequência	70
6	Conclusão	77
A	Formulário para Avaliação Subjetiva	79

Lista de Tabelas

5.1	Escala de Graduação do MOS	66
5.2	Escala de Graduação do MOS Invertido	67
5.3	Comparação entre as Técnicas de Predição Linear	67
5.4	Resultados do MOS Invertido	68
5.5	Resultados do MOS para o som de cravo	68
5.6	Resultados do MOS para o som de flauta	69
5.7	Resultados do MOS para o som de piano	69
5.8	Resultados do MOS para o som de violão e voz	69

Lista de Figuras

2.1	Sistemas Complementares	5
2.2	Sistemas Não Complementares	5
2.3	Representação Gráfica: (a) Compansão de Baixo Nível; (b) Compansão de Alto Nível	7
2.4	Esquema Básico do Método da Compansão	8
2.5	Supressor de Ruído de Baixo Nível	10
2.6	Supressor Gradual de Ruído: (a) Baixo Nível; (b) Alto Nível	10
2.7	Método de Supressão de Ruído por Subtração Espectral	14
2.8	Sistema Adaptativo de Supressão de Ruídos	15
2.9	Princípio de um Sistema Multicanal de Supressão de Ruídos	16
2.10	Cancelador Adaptativo de Ruído	17
2.11	Filtragem Adaptativa Seletiva do Sinal Degradado pelo Ruído Impulsivo	17
3.1	Estrutura Geral do Processo de Predição Linear	21
4.1	Etapas de Detecção	26
4.2	Características do Filtro Passa-Altas	28
4.3	Filtro $H(s)$ em cascata	28
4.4	Ilustração para um segmento do sinal $x(n)$	29

4.5	Sinal Degradado por Ruído Impulsivo	31
4.6	Sinal na Saída do Filtro Passa-Altas	31
4.7	Sinal na Saída da Envoltória	32
4.8	Representação Analógica do Detector de Envoltória	32
4.9	Sinal na Saída do Estágio Derivador	34
4.10	Estimação Seletiva do Sinal Degradado	34
4.11	Sinal na Saída da Estimação Progressiva	35
4.12	Sinal na Saída da Estimação Regressiva	36
4.13	Sinal na Saída da Estimação Dupla	37
4.14	Representação de pontos no plano cartesiano	38
4.15	Representação de uma seqüência de amostras	39
4.16	Pós-Filtragem após a Estimação Progressiva	41
4.17	Pós-Filtragem após a Estimação Regressiva	41
4.18	Pós-Filtragem após a Estimação Dupla	42
4.19	Sinal Original sem degradação	42
5.1	Curva da RSR para escolha da Ordem do Preditor	45
5.2	Som de cravo digitalizado	46
5.3	Som de flauta digitalizado	47
5.4	Som de piano digitalizado	47
5.5	Som de violão e voz digitalizado	48
5.6	Som de cravo degradado	48
5.7	Som de flauta degradado	49
5.8	Som de piano degradado	49
5.9	Som de violão e voz degradado	50

5.10 Som de cravo na saída do filtro passa-altas	50
5.11 Som de flauta na saída do filtro passa-altas	51
5.12 Som de piano na saída do filtro passa-altas	51
5.13 Som de violão e voz na saída do filtro passa-altas	52
5.14 Som de cravo na saída do detector de envoltória	52
5.15 Som de flauta na saída do detector de envoltória	53
5.16 Som de piano na saída do detector de envoltória	53
5.17 Som de violão e voz na saída do detector de envoltória	54
5.18 Som de cravo na saída do estágio derivador	54
5.19 Som de flauta na saída do estágio derivador	55
5.20 Som de piano na saída do estágio derivador	55
5.21 Som de violão e voz na saída do estágio derivador	56
5.22 Som de cravo filtrado	57
5.23 Som de flauta filtrado	58
5.24 Som de piano filtrado	58
5.25 Som de violão e voz filtrado	59
5.26 Zoom no som de cravo digitalizado	59
5.27 Zoom no som de cravo degradado	60
5.28 Zoom no som de cravo filtrado	60
5.29 Zoom no som de flauta digitalizado	61
5.30 Zoom no som de flauta degradado	61
5.31 Zoom no som de flauta filtrado	62
5.32 Zoom no som de piano digitalizado	62
5.33 Zoom no som de piano degradado	63

5.34	Zoom no som de piano filtrado	63
5.35	Zoom no som de violão e voz digitalizado	64
5.36	Zoom no som de violão e voz degradado	64
5.37	Zoom no som de violão e voz filtrado	65
5.38	Espectro tridimensional do som de cravo digitalizado	70
5.39	Espectro tridimensional do som de cravo degradado	71
5.40	Espectro tridimensional do som de cravo filtrado	71
5.41	Espectro tridimensional do som de flauta digitalizado	72
5.42	Espectro tridimensional do som de flauta degradado	72
5.43	Espectro tridimensional do som de flauta filtrado	73
5.44	Espectro tridimensional do som de piano digitalizado	73
5.45	Espectro tridimensional do som de piano degradado	74
5.46	Espectro tridimensional do som de piano filtrado	74
5.47	Espectro tridimensional do som de violão e voz digitalizado	75
5.48	Espectro tridimensional do som de violão e voz degradado	75
5.49	Espectro tridimensional do som de violão e voz filtrado	76

Capítulo 1

Introdução

Apesar do domínio dos circuitos analógicos em sistemas de áudio profissionais por mais de um século, nos últimos vinte anos eles foram gradativamente substituídos por circuitos digitais [6], que oferecem vantagens que vão desde a eficiência na digitalização e no processamento do sinal até a flexibilidade de programação, evitando a introdução de degradações no sinal original, proporcionando assim um aumento na relação sinal-ruído e, conseqüentemente, um sinal de alta qualidade.

Muito poucas gravações analógicas são completamente satisfatórias, uma vez que a maioria delas sofre degradações, do tipo contínua e/ou impulsiva, produzidas muitas vezes pelo próprio meio de gravação, que provocam uma perda da qualidade e inteligibilidade do sinal. Algumas das principais causas destas degradações são as variações das respostas em frequência e os efeitos de ressonância dos transdutores empregados, as limitações das respostas em frequência dos meios físicos de armazenamento da informação de áudio e os vários tipos de ruídos introduzidos pelo ambiente de gravação ou pelas componentes do sistema [24].

O ruído do tipo contínuo pode ser classificado, de uma forma geral, como **ruído de fundo**, que provoca o efeito do chiado na música e é geralmente produzido pelo próprio equipamento de gravação, e **ruído acústico ambiental**, que é produzido por fontes acústicas externas e adicionado ao sinal no momento da gravação. Este tipo de ruído degrada todas as amostras do sinal e pode ser eliminado, em grande parte, utilizando-se

técnicas adaptativas não seletivas, que permitem uma filtragem contínua do sinal degradado.

Um outro tipo de ruído, conhecido como **ruído impulsivo**, geralmente produzido por defeitos superficiais em discos fonográficos e, mais recentemente, por erros de leitura do sinal digitalizado, degrada apenas uma amostra ou um grupo de amostras do sinal, provocando o aparecimento de descontinuidades na amplitude do sinal.

O sistema ideal para a remoção do ruído impulsivo é o que usa **técnicas adaptativas seletivas** de tratamento do sinal degradado, através das quais apenas as amostras degradadas são filtradas e as demais amostras permanecem na sua forma original. Este sistema possui duas etapas. a de detecção, onde o cálculo de um limiar é usado para decidir se uma amostra ou um grupo de amostras está degradado, e a de restauração, onde a amostra ou o grupo de amostras, após detectado estar degradado, passa por um processo de filtragem.

As exigências quanto a se obter uma alta relação sinal-ruído (RSR), aumentaram a importância do desenvolvimento de métodos especiais para o tratamento de sinais degradados. Esta dissertação apresenta e compara algumas técnicas de processamento digital para a filtragem do ruído do tipo impulsivo em gravações, especialmente as mais antigas, que possuem um grande valor artístico e cultural e não podem ser refeitas.

A etapa de detecção é baseada no cálculo da derivada da envoltória do sinal e comparação com um limiar pré-estabelecido, determinado a partir das características do sinal. Na etapa de filtragem, que se baseia na teoria dos filtros ótimos, foram implementados algoritmos de extrapolação e interpolação seletiva para restauração do sinal degradado, sendo os melhores resultados obtidos através da predição linear regressiva com uma pós-filtragem baseada na interpolação polinomial de Lagrange.

Os resultados obtidos foram avaliados através de critérios objetivos e subjetivos e comparados com a literatura disponível, apresentando uma melhoria substancial na qualidade do sinal que se aproxima bastante do sinal original sem degradações.

1.1 Organização da Dissertação

O capítulo 2 apresenta um resumo da evolução dos sistemas de supressão de ruídos, enfatizando e classificando os mais relevantes.

O capítulo 3 apresenta uma revisão teórica da Codificação por Predição Linear (LPC), que vem a ser a base teórica para todas as técnicas propostas neste trabalho.

O capítulo 4 descreve os sistemas de redução de ruído do sinal de áudio degradado implementados neste trabalho, detalhando, ilustrando e analisando a importância de cada etapa do algoritmo de detecção.

O capítulo 5 apresenta as condições experimentais e os resultados obtidos em cada uma das técnicas propostas e, em seguida, é feita uma avaliação do desempenho de cada uma através do cálculo da relação sinal-ruído e de testes de escuta informais.

O capítulo 6 apresenta as conclusões a partir dos resultados obtidos e sugestões para trabalhos posteriores a serem realizados neste campo de pesquisa.

Capítulo 2

Técnicas de Supressão de Ruído em Sinais de Áudio Degradados

2.1 Introdução

Em um sistema de gravação/reprodução de áudio, o sinal pode sofrer degradações, do tipo contínua e/ou impulsiva, introduzidas ou produzidas pelo próprio equipamento ou meio de comunicação utilizado. Estas degradações provocam, em muitos casos, uma perda da qualidade e inteligibilidade do sinal. Existem, porém, alguns sistemas para reduzir a audibilidade destas degradações que, sem realizar nenhum tipo de correção na fonte, possibilitam um aumento considerável da relação sinal-ruído (RSR) levando, conseqüentemente, a um aumento da qualidade e inteligibilidade do sinal.

As exigências quanto a uma alta relação sinal-ruído, levaram ao desenvolvimento de sistemas especiais para o tratamento de sinais degradados. Em [12] estes sistemas são classificados em dois tipos: sistemas complementares e sistemas não complementares.

Os sistemas complementares, apresentados na Figura 2.1, teoricamente não introduzem alterações no sinal de áudio, pois é realizada uma compressão da faixa dinâmica do sinal de áudio antes da gravação (pré-processamento) e uma expansão complementar (pós-processamento) para restaurar a faixa dinâmica do sinal original. Os sistemas

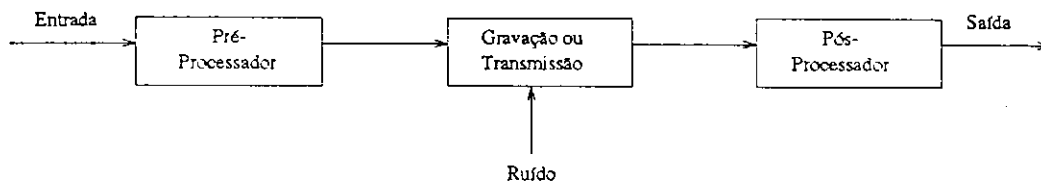


Figura 2.1: Sistemas Complementares

complementares são muito usados na gravação/reprodução de fita cassete.

Os sistemas não complementares, apresentados na Figura 2.2, onde o tratamento do sinal é realizado por pós-processamento, geralmente utilizam técnicas baseadas no mascaramento auditivo ou filtragem adaptativa para reduzir o ruído em sinais de áudio já gravados. Estas técnicas podem ser seletivas ou não seletivas e a escolha, entre uma e outra, é feita em função do tipo de degradação introduzida, ou seja, se a degradação é do tipo contínua ou impulsiva.

Os sistemas não complementares podem ser usados para melhorar a reprodução de discos fonográficos e de fitas cassetes gravadas que apresentam algum tipo de degradação. A maioria dos sistemas não complementares apresentam limitações [23], entretanto, somente este tipo de sistema de redução de ruído pode ser empregado para a restauração de gravações antigas, uma vez que estas gravações não podem ser refeitas.

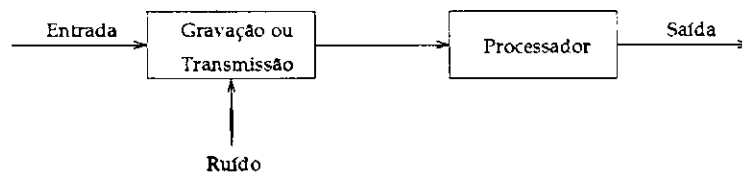


Figura 2.2: Sistemas Não Complementares

A degradação impulsiva se caracteriza por apresentar descontinuidades na amplitude do sinal recebido e é provocada geralmente por fadiga, defeitos superficiais e arranhões do material utilizado na gravação. Esse tipo de degradação não atinge todas as amostras e pode ser detectado e corrigido em sistemas não complementares utilizando técnicas adaptativas seletivas de redução de ruídos [3].

A degradação causada por ruídos acústicos ambientais ou por ruído térmico atinge todas as amostras do sinal, sendo portanto, neste caso, adequado o uso de técnicas adaptativas não seletivas de redução de ruídos, que permitem uma filtragem contínua do sinal degradado [2].

As técnicas adaptativas não seletivas de redução de ruídos são aplicadas nos sistemas conhecidos como **sistemas de cancelamento de ruídos e sistemas de supressão de ruídos** [1]. Nos sistemas de cancelamento de ruídos, parte-se do princípio que, além do sinal degradado, tem-se disponível um outro sinal, descorrelacionado com o sinal original, mas correlacionado com o ruído. Este sinal, normalmente chamado de sinal de referência, é filtrado adaptativamente, de forma a se obter uma estimação do ruído, que por sua vez é subtraída do sinal degradado, fornecendo assim uma estimação do sinal original [27].

Nos sistemas de supressão de ruídos apenas o sinal degradado é necessário; em geral, estes sistemas se baseiam na teoria dos filtros ótimos e/ou na teoria de estimação espectral a curtos intervalos de tempo. A redução do ruído é levada a efeito mediante um balanceamento adaptativo da amplitude espectral do sinal, obtido segundo estimações do espectro do sinal degradado e do espectro do ruído [2].

Os sistemas descritos acima podem ainda ser classificados como **sistemas monocanais** ou **sistemas multicanais** em função do número de entradas disponíveis do sinal [9]. Os sistemas de cancelamento de ruídos são casos típicos de sistemas multicanais, pois necessitam de pelo menos uma entrada de referência para o ruído. Os sistemas de supressão de ruídos podem, por outro lado, ser monocanais ou multicanais. Exemplos de sistemas de supressão de ruídos multicanais podem ser vistos em [10] e [11].

2.2 Sistemas Complementares

Os sistemas complementares mais conhecidos são do tipo analógico, que realizam um processamento do sinal antes da gravação do sinal, normalmente uma compressão, e um processamento equivalente, normalmente expansão, na reprodução. A Figura

2.3(a) apresenta um sistema complementar denominado de baixo nível, com compressão e expansão em sinais de baixa amplitude. A Figura 2.3(b) apresenta um sistema complementar de alto nível, com compressão e expansão em sinais de alta amplitude. O ruído adicionado pelo meio de transmissão após a compressão é reduzido na respectiva expansão. A compansão (compressão e expansão) pode ser realizada utilizando duas técnicas distintas.

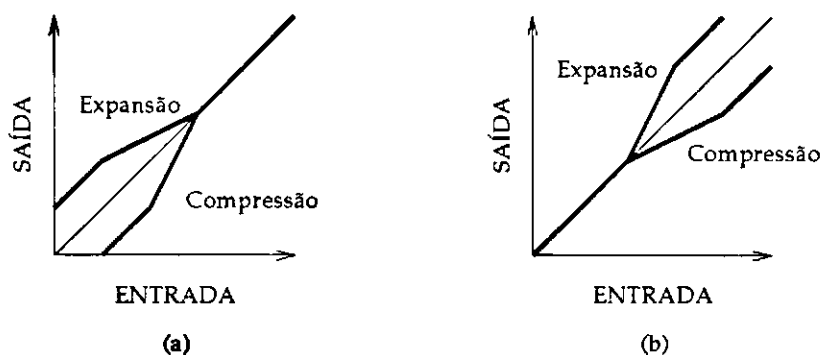


Figura 2.3: Representação Gráfica: (a) Compansão de Baixo Nível; (b) Compansão de Alto Nível

Na primeira técnica, mostrada na Figura 2.4(a), o sinal de controle do compressor é obtido após a passagem do sinal de entrada pelo elemento de ganho variável e o sinal de controle do expansor é obtido antes do sinal ser processado pelo elemento de ganho variável.

Na segunda técnica, mostrada na Figura 2.4(b), o sinal de controle do compressor é obtido antes da passagem do sinal pelo elemento de ganho variável e o sinal de controle do expansor é obtido depois do sinal ser processado pelo elemento de ganho variável [23].

Um sistema complementar bastante conhecido é o **Sistema Dolby** [12], que utiliza a técnica de compressão e expansão dos sinais. A primeira versão desse sistema, o Dolby A, divide o espectro de áudio em quatro faixas de frequência: Faixa 1, passa-baixas limitado em 80Hz; Faixa 2, passa-faixa de 80Hz a 3kHz; Faixa 3, passa-altas limitado em 3kHz; Faixa 4, passa-altas limitado em 9kHz. As faixas 1, 3 e 4 são filtros convencionais de 12 dB por oitava, enquanto que a faixa 2 tem uma resposta em

ao sinal original. Esta contração e/ou expansão automática da largura de faixa tem um efeito desprezível no sinal original, por causa da característica do limiar do ouvido e da distribuição da energia em sinais de áudio.

O sistema de Scott proporcionava uma redução considerável do ruído provocado por estalos e arranhões, ou seja, ruído de alta freqüência, mas em contrapartida alguns detalhes musicais se perdiam junto com o ruído. Com o crescimento da indústria fonográfica, a qualidade do processo de gravação/reprodução do disco melhorou substancialmente e o sistema de Scott caiu em desuso.

O **Filtro Dinâmico de Ruído** desenvolvido por Burwen [8] se baseia nos princípios do supressor dinâmico de ruído de Scott, que consiste de um filtro passa-faixa cuja freqüência de corte é determinada pela energia contida no sinal. Sendo que, neste caso, o filtro passa-faixa é composto pela associação de um filtro passa-altas e um filtro passa-baixas, com freqüências de corte variáveis.

A variação da freqüência de corte do filtro passa-altas é obtida através de um sinal de controle proporcional à energia fornecida por um filtro passa-baixas fixo e, em contrapartida, a variação da freqüência de corte do filtro passa-baixas é obtida através de um sinal de controle proporcional à energia fornecida por um filtro passa-altas fixo. Os limites máximos e mínimos das freqüências de corte são ajustados de tal maneira que as flutuações na largura de faixa devido ao ruído não sejam audíveis.

A redução do ruído de baixa freqüência, em torno de 30 Hz, fica entre 20 e 68 dB e do ruído de alta freqüência, em torno de 10 kHz, fica entre 20 a 57 dB. A redução de todo o ruído atinge apenas 10 a 15 dB, mas é considerada importante especialmente quando se utiliza equalização em freqüências altas e baixas para compensar deficiências em equipamentos (alto-falantes, caixas acústicas, etc) [8].

Um outro sistema, proposto por Ives [15], deriva o sinal de controle do espectro do sinal de áudio com maior confiabilidade e controla com maior precisão o sistema de filtragem dinâmica. O sistema é composto de um filtro passa-baixas com corte abrupto que utiliza o atual espectro de energia do sinal de áudio para posicionar o filtro até à freqüência extrema.

A frequência extrema pode variar de 200 Hz a 20kHz sem provocar nenhuma distorção ou mudança no formato do filtro. Com o filtro situado a 2kHz, a redução do ruído de 4kHz fica em torno de 12 dB e a redução do ruído a 8kHz fica em torno de 24 dB [15].

Este sistema não provoca nenhuma mudança no ganho ou na busca da resposta em frequência porque a filtragem é uma função unicamente do formato do espectro de música, daí o nome filtro espectral.

Mais recentemente, com o avanço da tecnologia digital, muitas técnicas digitais de redução de ruído foram desenvolvidas e aplicadas ao sistema de áudio, o que proporcionou um completo desuso das técnicas analógicas. Essas técnicas podem ser classificadas como seletivas e não seletivas e algumas delas são descritas sucintamente a seguir.

2.3.1 Técnicas Não Seletivas de Redução de Ruídos

As técnicas adaptativas não seletivas de redução de ruídos são aplicadas nos sistemas conhecidos como sistemas de supressão de ruídos e sistemas de cancelamento de ruídos [1].

Os sistemas de supressão de ruídos, cujo princípio geral foi descrito na Seção 2.1, podem ser classificados de várias formas, algumas das quais são descritas a seguir:

Os sistemas de Supressão de Ruído Acústico usando Subtração Espectral, mostrados na Figura 2.7, pertencem a uma classe de sistemas de melhoramento do sinal que explora a noção de que está na magnitude espectral, ao invés da fase, a principal característica da informação para inteligibilidade e qualidade do sinal. Nesta classe de sistemas, o sinal degradado é inicialmente segmentado em curtos intervalos de tempo nos quais se considera que os sinais sejam estacionários e a cada segmento são determinadas estimativas espectrais a curto intervalo de tempo para o sinal degradado e para o ruído que, subtraídas entre si, representam a estimação espectral do sinal [4].

Seja o sinal $x(n) = s(n) + r(n)$ a versão do sinal de áudio $s(n)$ degradada pelo ruído $r(n)$, os quais são descorrelacionados entre si. A técnica de subtração espectral consiste em determinar-se uma estimação da densidade espectral de potência $\hat{S}_{ss}(w)$

(DEP) do sinal original, a partir da subtração da estimação da densidade espectral de potência $\hat{S}_{xx}(w)$ do sinal degradado, da estimação da densidade espectral de potência $\hat{S}_{rr}(w)$ do ruído. Tomando-se essas estimações de densidade espectral de potência em curtos intervalos de tempo, obtém-se para um k -ésimo intervalo o seguinte:

$$\hat{S}_{ss}(w, k) = \hat{S}_{xx}(w, k) - \hat{S}_{rr}(w, k) \quad (2.1)$$

As estimações espectrais na Equação 2.1 para o sinal degradado e o ruído, são obtidas através do quadrado das respectivas amplitudes espectrais $|X(w, k)|$ e $|N(w, k)|$ determinadas para o k -ésimo segmento considerado:

$$\hat{S}_{xx}(w, k) \simeq |X(w, k)|^2 \quad (2.2)$$

e

$$\hat{S}_{rr}(w, k) = \hat{E}[|N(w, k)|^2] \quad (2.3)$$

$\hat{E}[|N(w, k)|^2]$ representa a estimação do ruído através da média das estimações obtidas ao longo dos intervalos de pausas [9].

De forma semelhante, a estimação da densidade espectral de potência a curto intervalo de tempo $\hat{S}_{ss}(w, k)$ do sinal é dada por:

$$\hat{S}_{ss}(w, k) \simeq |S(w, k)|^2 \quad (2.4)$$

O algoritmo de subtração é dado por:

$$\begin{cases} \hat{S}_{ss}(w, k) = |X(w, k)|^2 - E[|N(w, k)|^2] & \text{para } |X(w, k)|^2 > E[|N(w, k)|^2] \\ \hat{S}_{ss}(w, k) = 0 & \text{caso contrário} \end{cases} \quad (2.5)$$

Tem-se, portanto, que a estimação da amplitude espectral a curtos intervalos de tempo do sinal original será

$$|\hat{S}(w, k)| = \{ |X(w, k)|^2 - E[|N(w, k)|^2] \}^{1/2} \quad (2.6)$$

e a estimação do seu espectro de curto intervalo de tempo dado por:

$$\hat{S}(w, k) = |\hat{S}(w, k)| \exp j\theta(w, k) \quad (2.7)$$

onde $\theta(w, k)$ é a fase do espectro do próprio sinal degradado a qual é utilizada para reconstruir o sinal melhorado. Considera-se que a fase do sinal degradado é uma aproximação aproveitável da fase do sinal original, uma vez que o ouvido é relativamente insensível a degradações de fase [14, 4, 2].

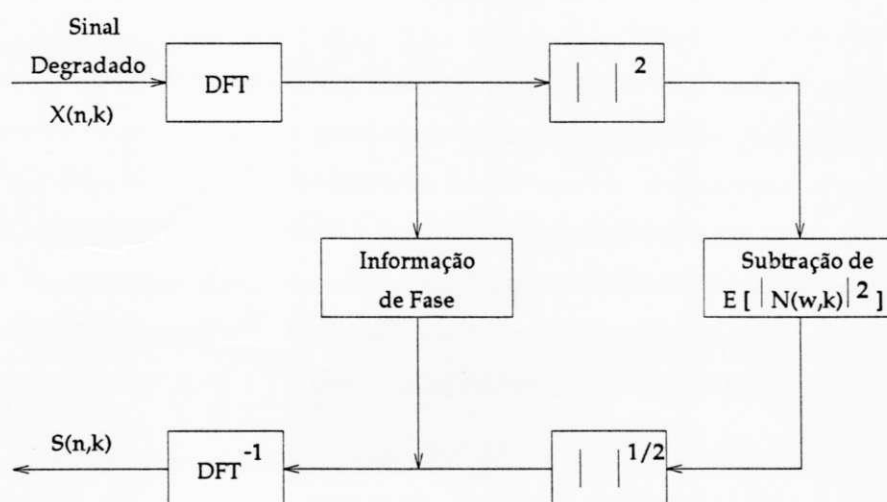


Figura 2.7: Método de Supressão de Ruído por Subtração Espectral

Os **Sistemas Adaptativos de Supressão de Ruídos**, cuja estrutura geral é mostrada na Figura 2.8 utilizam, em geral, algum tipo de filtro adaptativo cuja resposta ao impulso é determinada em função das propriedades estatísticas do sinal a ser melhorado e do ruído [1]. Estas propriedades, em geral, não são disponíveis nestes sistemas. Um pré-reconhecimento delas também não é possível, se os sinais não são estacionários. Por isso, as informações necessárias a respeito da estatística dos sinais devem ser obtidas a partir do sinal degradado e em curtos intervalos de tempo correspondentes a segmentos de igual comprimento. Os parâmetros do filtro são calculados e atualizados a cada novo segmento.

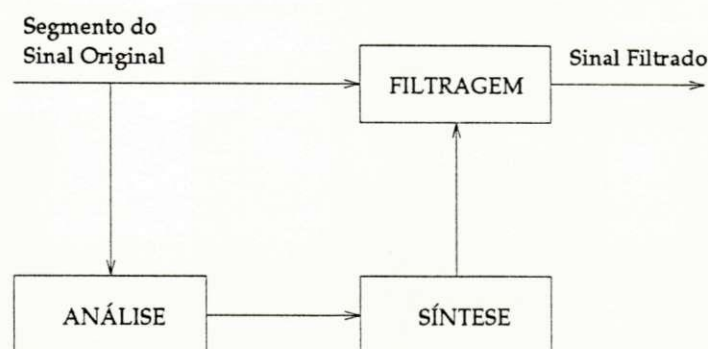


Figura 2.8: Sistema Adaptativo de Supressão de Ruídos

Os **Sistemas de Supressão de Ruídos por Balanceamento Espectral Adaptativo** baseiam-se na estrutura geral dos sistemas adaptativos de supressão de ruídos, mostrada na Figura 2.8. Estes sistemas consistem em se determinar uma função de transferência de um filtro, obtida a partir de estimações do espectro do sinal degradado e do espectro do sinal de ruído, que é utilizada para modificação da amplitude espectral do sinal degradado. Este método é baseado em estimações espectrais a curto intervalo de tempo usando a Transformada Discreta de Fourier (DFT - Discrete Fourier Transform).

Os sistemas adaptativos de supressão de ruídos podem ser classificados como monocanais e multicanais, de acordo com a disponibilidade de mais de uma entrada para o sinal degradado sobre o qual serão efetuadas as estimações para o cálculo do filtro de síntese.

O sistema multicanal se justifica pela não estacionariedade do ruído que torna difícil de se obter uma entrada de referência externa para o mesmo. No entanto, há uma maior facilidade na obtenção de entradas de referência para o sinal de áudio degradado. A Figura 2.9 mostra o princípio de um **Sistema Multicanal de Supressão de Ruídos**, no qual os parâmetros para a síntese do filtro são obtidos explorando-se as características de correlação do sinal de áudio e do ruído. O sistema se baseia no fato de que, com o deslocamento dos sinais no tempo, aumenta a correlação entre o sinal

de áudio presente nas várias entradas e o sinal original (correlação cruzada), ao passo que a correlação entre o ruído presente nestas entradas torna-se menor [11].

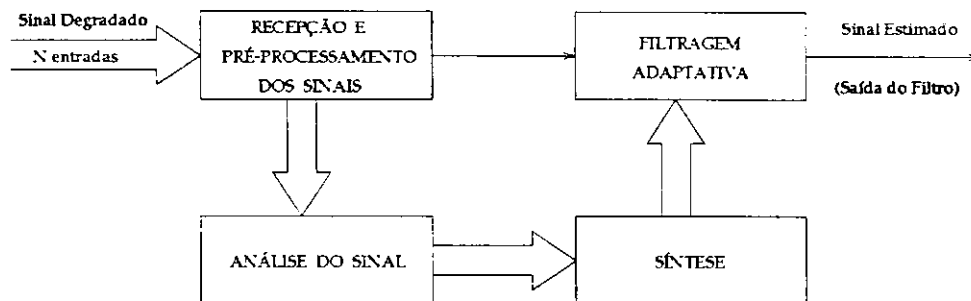


Figura 2.9: Princípio de um Sistema Multicanal de Supressão de Ruídos

Finalmente, nos **Sistemas de Cancelamento de Ruídos**, cujo algoritmo é mostrado na Figura 2.10, parte-se do princípio de que, além do sinal degradado $x(n)$, tem-se disponível um outro sinal $r(n)$, descorrelacionado com o sinal original $s(n)$, mas correlacionado com o ruído $d(n)$. Este sinal, normalmente chamado de sinal de referência, é filtrado adaptativamente, de forma a se obter uma estimação do ruído $\hat{d}(n)$, que por sua vez é subtraída do sinal degradado, fornecendo assim uma estimação do sinal original $\hat{s}(n)$. Os sistemas de cancelamento de ruídos são casos típicos de sistemas multicanais, pois necessitam de pelo menos uma entrada de referência para o ruído [27, 10].

O melhoramento da relação sinal-ruído utilizando a técnica de cancelamento adaptativo de ruídos é geralmente alto, mas é uma técnica limitada na prática, uma vez que a entrada de referência contém tanto o sinal quanto o ruído e, desta forma, o cancelador tenderá a cancelar tanto o sinal quanto o ruído degradante.

2.3.2 Técnicas Seletivas de Redução de Ruídos

Considerando que a degradação do tipo impulsiva não atinge todas as amostras do sinal, é indicado o uso de **técnicas de filtragem adaptativa seletivas** para reduzir o ruído no sinal degradado, que se baseiam em uma busca contínua de erros analisando

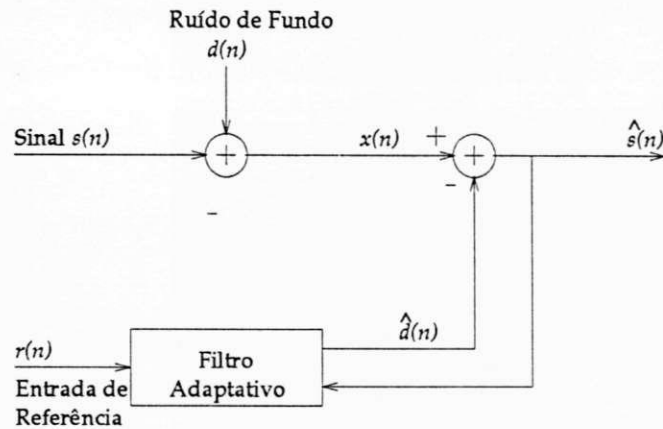


Figura 2.10: Cancelador Adaptativo de Ruído

amostra por amostra e só quando uma amostra é considerada degradada é que se processa a sua correção.

As técnicas seletivas de redução de ruídos se constituem basicamente de duas etapas: detecção e correção, conforme mostra a Figura 2.11.

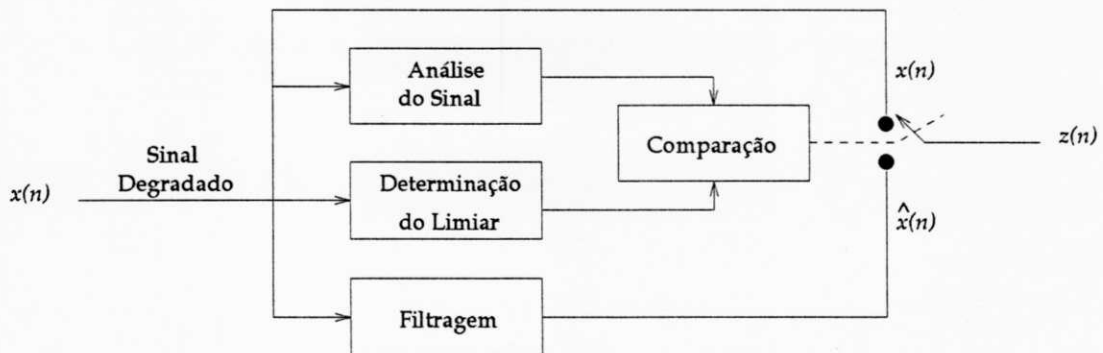


Figura 2.11: Filtragem Adaptativa Seletiva do Sinal Degradado pelo Ruído Impulsivo

Na técnica de redução seletiva de ruídos, proposta por Aguiar Neto [3], a etapa de detecção se baseia no princípio de que, em um sinal correlacionado, a diferença entre duas amostras consecutivas, $x(n) - x(n - 1)$, tende a ser relativamente pequena. O processo de busca de erros consiste em medir essa diferença, a curtos intervalos de tempo, e compará-la com um valor de referência (limiar) determinado a partir das características do sinal. Quando o valor da diferença estiver abaixo do limiar, o sinal é

preservado e o valor na saída $z(n)$ é igual ao da entrada $x(n)$. Em contrapartida, quando o valor da diferença estiver acima do limiar, esta amostra é considerada degradada e é então acionada a etapa de filtragem que efetua uma estimação do valor daquela amostra, $\hat{x}(n)$, e faz a substituição na saída pelo valor estimado pelo processo de filtragem, ou seja, $z(n) = \hat{x}(n)$.

Em uma outra técnica de redução seletiva de ruídos, proposta por Kenia Souza [24], a etapa de detecção se baseia no fato de que os ruídos impulsivos mais comuns em sinais de áudio possuem uma densidade espectral que se situa fora da região de concentração da densidade espectral do sinal de áudio. É utilizado um filtro passa-altas que realça a presença do ruído e um detector de envoltória que produz um sinal cuja derivada indica, com boa resolução, a região do sinal atacada pelo ruído. Após a identificação da região degradada do sinal, as amostras situadas nessa região são substituídas por estimações obtidas por amostras vizinhas não degradadas utilizando interpolação não linear. Esta proposta serviu de base para a etapa de detecção de ruídos no sistema apresentado neste trabalho.

Capítulo 3

Codificação por Predição Linear

3.1 Introdução

Uma das técnicas mais eficientes para a análise do sinal de voz é o método da análise preditiva linear. Este método tem sido predominantemente usado para estimação de parâmetros básicos do sinal de voz, tais como, período fundamental, formantes, espectro, funções na área de trato vocal para representação ou armazenamento do sinal de voz a baixas taxas de transmissão. Sua importância reside na capacidade de fornecer estimativas extremamente corretas dos parâmetros do sinal de voz e na sua relativa velocidade de cálculo.

O fato da técnica de predição linear ter sido inicialmente usada para aplicações em comunicação de voz, especialmente no que se refere ao modelamento do trato vocal em sintetizadores de voz, tornou-a aparentemente inadequada para aplicações em sinais de áudio, onde o nível de qualidade esperado é muito mais alto que em aplicações de sinais de voz. Entretanto, foi constatado que, no contexto de sintetizadores de música e música computadorizada, a flexibilidade do modelo de predição linear o torna um bom candidato para produção e reprodução de sons musicais [7].

A idéia básica do processo de predição linear é que uma amostra do sinal pode ser aproximada por uma combinação linear das amostras anteriores ou posteriores. As

técnicas de predição linear estão disponíveis na literatura de engenharia e têm sido usadas em sistemas de controle, automação, telecomunicações e teoria da informação e codificação.

A predição linear pode ser aplicada utilizando os seguintes métodos:

1. O Método da Covariância;
2. O Método da Autocorrelação;
3. A Formulação do Filtro Inverso;
4. A Formulação da Estimação Espectral;
5. A Formulação da Máxima Verossimilhança;
6. A Formulação do Produto Interno;

dentre outros.

Um estudo dos vários métodos e a comparação entre eles podem ser encontrados em [20]. Neste trabalho será utilizado apenas o Método da Autocorrelação.

Neste capítulo será feita inicialmente uma descrição dos princípios básicos da predição linear. Em seguida, será feita uma abordagem sobre o método da autocorrelação e apresentado o algoritmo para o cálculo dos coeficientes do preditor.

3.2 Princípios Básicos da Predição Linear

O processo de predição linear [4, 7], cuja estrutura geral é mostrada na Figura 3.1, consiste na estimação do valor de uma amostra atual de um sinal $s(n)$ a partir da combinação linear dos valores das amostras passadas $s(n - k)$; $k = 1, 2, 3, \dots, P$:

$$\hat{s}(n) = s(n) * h(n) \quad (3.1)$$

Um preditor linear com coeficientes de predição, a_k , é definido como um sistema cuja saída é:

$$\hat{s}(n) = \sum_{k=1}^P a_k s(n-k) \quad (3.2)$$

onde P representa o número de coeficientes do preditor [4, 16].

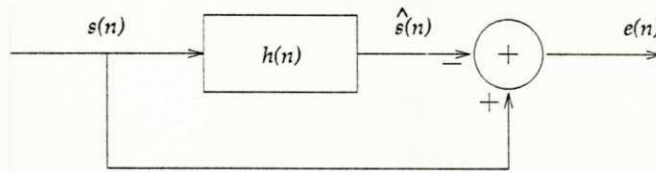


Figura 3.1: Estrutura Geral do Processo de Predição Linear

A resposta ao impulso $h(n)$ do preditor é definida como

$$h(n) = \sum_{k=1}^P a_k \delta(n-k) \quad (3.3)$$

e a correspondente função de transferência é

$$H(z) = \sum_{k=1}^P a_k z^{-k} \quad (3.4)$$

onde a resposta ao impulso do preditor deve ser tal que forneça a melhor estimaco $\hat{s}(n)$, ou seja, o preditor é otimizado de forma a fornecer o menor erro de estimaco [4, 16], $e(n)$, dado por:

$$e(n) = s(n) - \hat{s}(n) = s(n) - \sum_{k=1}^P a_k s(n-k) \quad (3.5)$$

O problema bsico da predico linear consiste na determinaco de um grupo de coeficientes do preditor a_k , a partir das caractersticas do sinal, com o objetivo de se obter uma boa estimativa das suas propriedades espectrais.

Devido a não estacionariedade do sinal, os coeficientes do preditor devem ser estimados a curtos segmentos do sinal, com o objetivo de minimizar o erro médio quadrático de predição. Esse cálculo deve ser executado a curtos intervalos de tempo de forma a seguir a dinâmica do sinal. Como o método utilizado para o cálculo dos coeficientes do preditor é o da autocorrelação, torna-se necessário efetuar um janelamento do sinal original.

No janelamento, o sinal $s(n)$ é inicialmente multiplicado por uma função que define uma janela de tempo finita, $w(n)$, produzindo um sinal $x(n)$ da forma:

$$x(n) = \begin{cases} w(n)s(n), & 0 \leq n \leq N - 1 \\ 0, & \text{caso contrário} \end{cases} \quad (3.6)$$

assumindo que a janela é zero fora do intervalo $0 \leq n \leq N - 1$. A largura da janela, N , geralmente corresponde a um intervalo de 20-30 ms, onde o sinal pode ser considerado estacionário. Os tipos de janelas mais comuns são: Retangular, Hamming e Hanning [7].

Uma justificativa para o uso do erro médio quadrático mínimo de predição como base para estimação dos parâmetros do modelo é que esta aproximação fornece um grupo de equações lineares que podem ser eficientemente resolvidas para obtenção dos parâmetros do preditor. O mais importante é que os parâmetros resultantes compreendem uma representação bastante segura do sinal.

O sinal residual cuja energia deve ser minimizada é obtido pela passagem de $x(n)$ pelo filtro $A(z)$, que é chamado de *filtro do erro de predição*, dado por:

$$A(z) = 1 - \sum_{k=1}^P a_k z^{-k} \quad \text{ou} \quad A(z) = 1 - H(z) \quad (3.7)$$

onde a Equação 3.7 é obtida aplicando-se a transformada z à Equação 3.5. A energia residual ou erro médio de predição em um curto intervalo de tempo é definida como

$$E = \sum_n e_x^2(n) \quad (3.8)$$

$$E = \sum_n [x(n) - \hat{x}(n)]^2 \quad (3.9)$$

$$E = \sum_n [x(n) - \sum_{k=1}^P a_k x(n-k)]^2 \quad (3.10)$$

onde $e_x(n)$ é o residual correspondente ao sinal janelado $x(n)$. O conjunto de coeficientes a_k que minimiza E é obtido a partir de

$$\frac{\partial[E]}{\partial[a_k]} = 0, \quad 1 \leq k \leq P \quad (3.11)$$

Substituindo-se a Equação 3.10 na Equação 3.11 e realizando-se as P derivadas parciais, chega-se ao seguinte sistema de equações lineares:

$$\sum_{k=1}^P a_k R(i-k) = R(i), \quad 1 \leq i \leq P \quad (3.12)$$

onde

$$R(i) = \sum_{n=-\infty}^{\infty} x(n)x(n-i) = \sum_{n=1}^{N-1} x(n)x(n-i), \quad 0 \leq i \leq P \quad (3.13)$$

é a função de autocorrelação a curto intervalo de tempo. A Equação 3.13, pode ser visualizada mais facilmente se colocada na forma matricial:

$$\begin{bmatrix} R(0) & R(1) & R(2) & \cdots & R(P-1) \\ R(1) & R(0) & R(1) & \cdots & R(P-2) \\ R(2) & R(1) & R(0) & \cdots & R(P-3) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ R(P-1) & R(P-2) & R(P-3) & \cdots & R(0) \end{bmatrix} \cdot \begin{bmatrix} a(1) \\ a(2) \\ a(3) \\ \cdots \\ a(P) \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ R(3) \\ \cdots \\ R(P) \end{bmatrix} \quad (3.14)$$

o sistema resultante pode ser resolvido de forma recursiva, utilizando o algoritmo de *Levinson-Durbin* [7, 20, 25], que será descrito detalhadamente na próxima seção.

A energia residual mínima ou erro médio quadrático mínimo E_P é obtida pela substituição da Equação 3.12 na Equação 3.10

$$E_P = R(0) - \sum_{k=1}^P a_k R(k) \quad (3.15)$$

E_P é também conhecido como o *erro quadrático total mínimo* ou, simplesmente, *erro de predição mínimo*. O ganho do filtro $H(z)$ é da forma [7]

$$G^2 = E_P \quad (3.16)$$

3.3 Cálculo dos Coeficientes do Preditor

Devido às propriedades especiais da matriz dos coeficientes, apresentada em (3.14), é possível resolver o sistema de equações de forma bastante eficiente.

A matriz de autocorrelação não é somente simétrica, mas também Toeplitz, ou seja, os elementos ao longo das diagonais são idênticos. Para este caso especial, Levinson e Durbin desenvolveram uma solução recursiva eficiente para o cálculo dos coeficientes ótimos do preditor de ordem sucessiva, cujo algoritmo é como se segue [7]:

$$E_0 = R(0) \quad (3.17)$$

$$k_m = [R(m) - \sum_{k=1}^{m-1} a_{m-1}(k)R(m-k)]/E_{m-1}, \quad 1 \leq m \leq P \quad (3.18)$$

$$a_m(k) = k_m \quad (3.19)$$

$$a_m(k) = a_{m-1}(k) - k_m a_{m-1}(m-k), \quad 1 \leq k \leq m-1 \quad (3.20)$$

$$E_m = (1 - k_m^2)E_{m-1} \quad (3.21)$$

Os coeficientes $a_m(k)$, $1 \leq k \leq m$, são os coeficientes do preditor para um preditor ótimo de m -ésima ordem. A solução final é dada pelos coeficientes $a_P(k)$, $1 \leq k \leq P$.

Da Equação 3.21 pode-se escrever

$$E_m = R(0) \prod_{i=1}^m (1 - k_i^2) \quad (3.22)$$

O erro mínimo E_m , que pode ser monitorado a cada etapa de cálculo, geralmente diminui quando a ordem do preditor aumenta.

Se os coeficientes de autocorrelação $R(i)$ forem substituídos por um grupo de coeficientes de autocorrelação normalizados, ou seja, $r(k) = R(k)/R(0)$, então a solução para a equação matricial permanece inalterada e o erro E_m é agora interpretado como um erro normalizado. Se ele for chamado de V_P , então

$$V_P = \frac{E_P}{R(0)} = \prod_{m=1}^P (1 - k_m^2), \quad 0 \leq V_P \leq 1 \quad (3.23)$$

onde k_m é conhecido como coeficiente de reflexão ou coeficiente de correlação parcial (PARCOR) e

$$|k_m| < 1, \quad 1 \leq m \leq P \quad (3.24)$$

pode ser uma condição suficiente e necessária para que todas as raízes do polinômio $A(z)$ estejam inseridas no círculo unitário, garantindo assim a estabilidade do sistema $H(z)$, que é de extrema importância na estimação de sinais, uma vez que um filtro instável pode degradar o sinal estimado.

Capítulo 4

Descrição do Sistema de Redução de Ruídos Impulsivos

4.1 Algoritmo de Detecção

Em uma filtragem seletiva, a etapa de detecção deve ser capaz de indicar com segurança apenas as amostras que estão degradadas e que, conseqüentemente, devem ser submetidas à etapa de filtragem. Em [24] é proposto um algoritmo de detecção bastante eficiente que foi modificado, no que se refere ao cálculo do limiar e das constantes de tempo do detector de envoltória, e utilizado neste trabalho. As etapas do algoritmo adaptado são mostradas na Figura 4.1 e descritas a seguir.

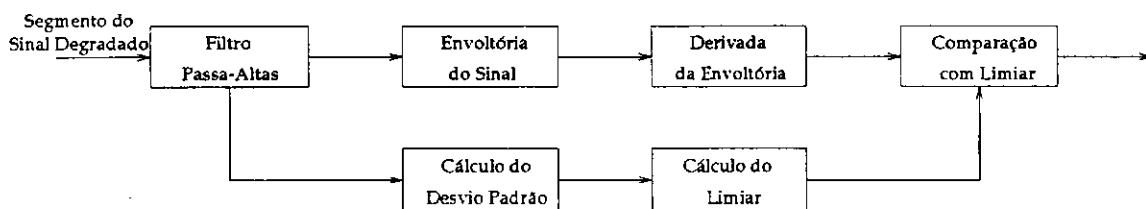


Figura 4.1: Etapas de Detecção

4.1.1 Cálculo do Desvio Padrão

O sinal a ser tratado é dividido em segmentos iguais de L amostras $x(n)$; $n = 1, 2, 3, \dots, L$, tal que para cada segmento é determinado um valor de referência σ , dado pelo desvio padrão,

$$\sigma = \sqrt{\frac{1}{L} \sum_{n=1}^L [x(n) - x(n-1)]^2} \quad (4.1)$$

que representa o valor eficaz da diferença entre as amostras no segmento de comprimento L [3].

Em seguida, é calculado o desvio padrão médio, σ_{med} , definido como,

$$\sigma_{med} = \frac{1}{N} \sum_{m=1}^N \sigma_m \quad (4.2)$$

onde N é o número de segmentos do sinal e σ_m é o desvio padrão referente a cada segmento do sinal.

4.1.2 Cálculo do Limiar

O limiar escolhido é da forma $F \cdot \sigma_{med}$, onde o fator F foi otimizado em função da relação sinal-ruído obtida para o sinal filtrado, e de testes de escuta informais, a fim de possibilitar a detecção dos erros mais relevantes.

Considerando que um pequeno valor para F resulta em uma grande taxa de erros, $F = 0$ significa que todas as amostras serão consideradas degradadas. Em contrapartida, $F \gg 0$ significa que todas as amostras serão consideradas sem erros. O valor ótimo para F depende das características da degradação. Para degradações produzidas por ruído do tipo impulsivo agudo foi observado que $F = 4$ apresenta ótimos resultados; outros tipos de degradações levam a valores diferentes.

4.1.3 Filtro Passa-Altas

Levando em consideração que a maior parte da energia do espectro de música concentra-se em torno de 1kHz, foi utilizado um filtro passa-altas, cujas características estão mostradas na Figura 4.2, que cresce 18dB/oitava abaixo de 1kHz e 12 dB/oitava entre 1kHz e 5kHz.

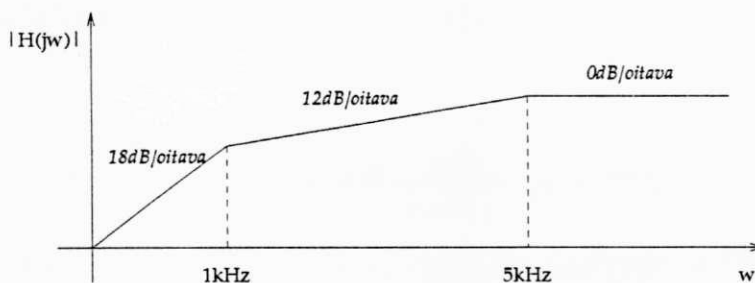


Figura 4.2: Características do Filtro Passa-Altas

A função de transferência do filtro passa-altas é dada por [24]:

$$H(s) = \frac{s^3}{(s + P_1).(s + P_2)^2} \tag{4.3}$$

onde

$$\begin{cases} P_1 = 2\pi.1000 \\ P_2 = 2\pi.5000 \end{cases} \tag{4.4}$$

De acordo com a Equação 4.3, $H(s)$ pode ser representado por três filtros em cascata, como mostra a Figura 4.3.

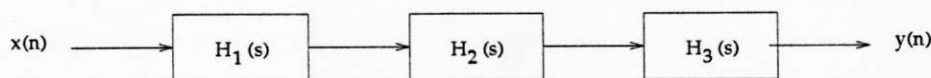


Figura 4.3: Filtro H(s) em cascata

onde

$$\begin{cases} H(s) = H_1(s) \cdot H_2(s) \cdot H_3(s) \\ H_1(s) = \frac{s}{(s+P_1)} \\ H_2(s) = H_3(s) = \frac{s}{(s+P_2)} \end{cases} \quad (4.5)$$

Para implementação deste filtro na forma digital, foram utilizados filtros IIR (resposta impulsiva infinita). A função de transferência do filtro digital foi obtida pela aplicação da transformação bilinear [19]:

$$s = \frac{2 \cdot (1 - z^{-1})}{T \cdot (1 + z^{-1})} \quad (4.6)$$

onde $z = e^{j\theta}$, $\theta = \omega T$ e T é o período de amostragem, ilustrado na Figura 4.4.

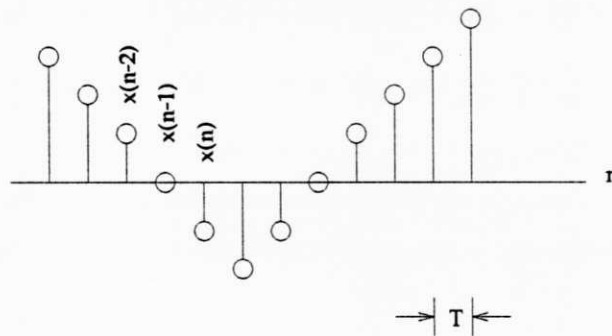


Figura 4.4: Ilustração para um segmento do sinal $x(n)$

Substituindo-se 4.6 em 4.3, obtém-se a função de transferência do filtro no plano z . Assim, para $H_1(s)$, tem-se:

$$H_1(z) = \frac{\frac{2}{T}(1 - z^{-1})}{A_1 - B_1 z^{-1}} \quad (4.7)$$

onde $A_1 = \frac{2}{T} + P_1$ e $B_1 = \frac{-2}{T} + P_1$.

Aplicando-se o operador da transformada z inversa, Z^{-1} , obtém-se:

$$Z^{-1}\{H_1(z)\} = Z^{-1}\{Y(z)/X(z)\} \quad (4.8)$$

$$Z^{-1}\{Y(z).\{A_1 - z^{-1}.B_1\}\} = Z^{-1}\{X(z).\frac{2}{T}.(1 - z^{-1})\} \quad (4.9)$$

A propriedade do deslocamento no tempo da transformada z , estabelece que:

$$X(z).z^{-m} = Z\{[x(n - m)]\} \quad (4.10)$$

onde n é a variável indexada discreta no tempo.

Para calcular a equação de diferença do filtro IIR, usa-se a propriedade da Equação 4.10 na Equação 4.9:

$$y(n) = \frac{1}{A_1} \left\{ \frac{2}{T} [x(n) - x(n - 1)] + y(n - 1).B_1 \right\} \quad (4.11)$$

A representação dos dados variando no tempo é baseada na notação padrão, usada na teoria de filtros digitais [17]. Nesta notação, $x(n)$ representa o valor da amplitude da amostra atual e $x(n - 1)$ o valor da amplitude da amostra anterior. De maneira similar, $y(n)$ é o valor de saída da amostra atual e $y(n - 1)$ o valor de saída da amostra anterior [24].

As equações de diferença para as funções de transferência $H_2(s)$ e $H_3(s)$ são deduzidas de maneira similar a $H_1(s)$. Através destas equações o filtro pode ser implementado em software.

A Figura 4.5 apresenta um trecho de música degradada pelo ruído impulsivo, com duração de aproximadamente 10ms. A Figura 4.6 mostra este mesmo trecho após a passagem pelo filtro passa-altas. Comparando os sinais pode-se ver que a passagem pelo filtro tem como objetivo reforçar o ruído, facilitando assim a sua detecção.

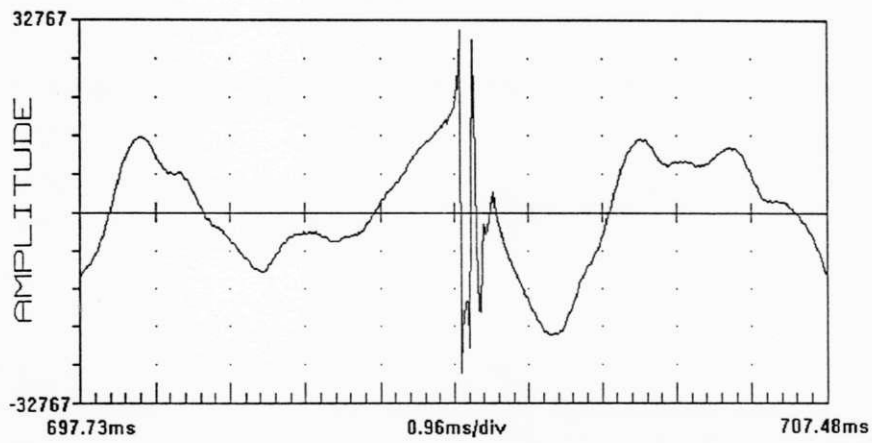


Figura 4.5: Sinal Degradado por Ruído Impulsivo

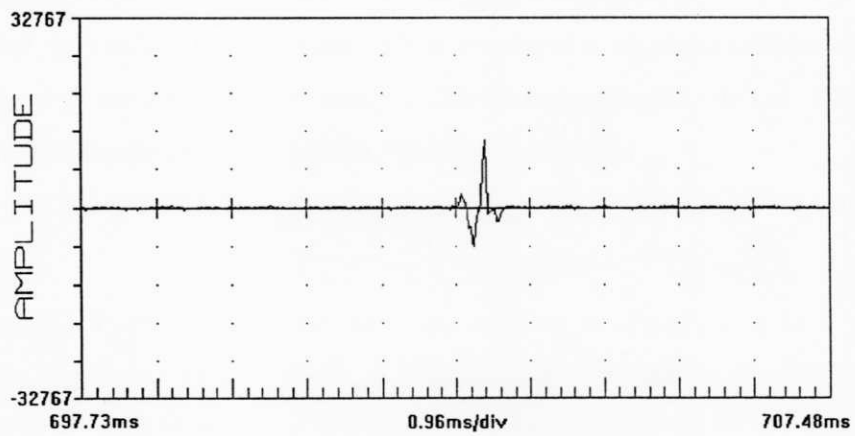


Figura 4.6: Sinal na Saída do Filtro Passa-Altas

4.1.4 Detector de Envoltória

Após o filtro passa-altas é utilizado um detector de envoltória [24], cuja saída é ilustrada na Figura 4.7, onde se pode verificar que o trecho em que está localizado o ruído tem uma amplitude bem evidenciada em relação ao resto do sinal.

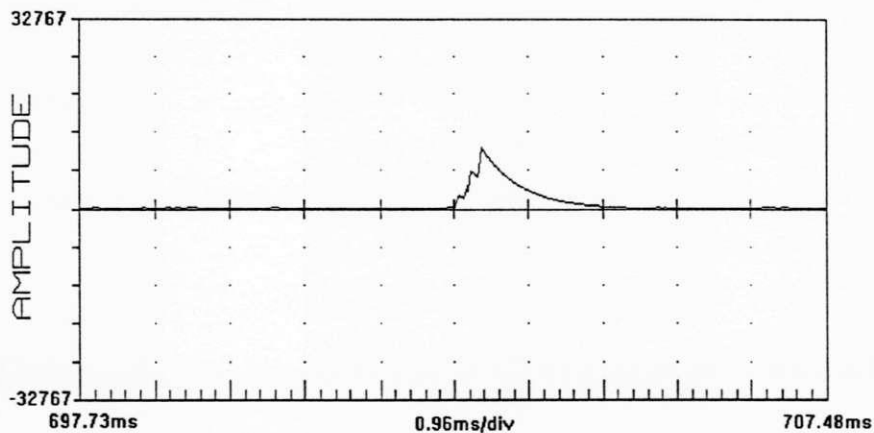


Figura 4.7: Sinal na Saída da Envoltória

O circuito analógico do detector de envoltória é representado na Figura 4.8, onde $y_1(t)$ é o sinal na saída do filtro passa-altas e $y_2(t)$ é o sinal na saída do detector de envoltória, ambos no domínio do tempo. Na implementação digital, como o sinal é discretizado, a comparação se processa da seguinte forma:

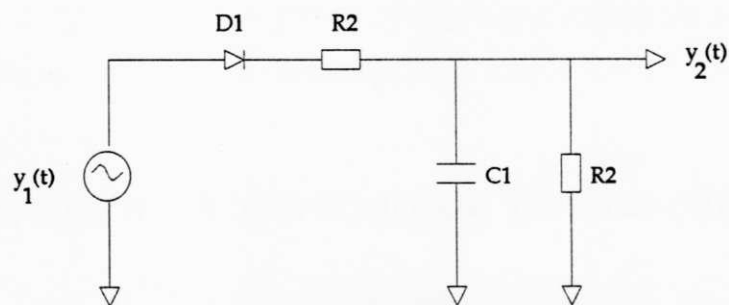


Figura 4.8: Representação Analógica do Detector de Envoltória

- Quando $y_2(n-1) \leq y_1(n)$:

$$y_2(n) = \frac{T}{T+T_1} [y_1(n) + \frac{T_1}{T} y_2(n-1)] \quad (4.12)$$

- Quando $y_2(n-1) > y_1(n)$:

$$y_2(n) = \frac{T}{T+T_2} [\frac{T_2}{T} y_2(n-1)] \quad (4.13)$$

onde n é a variável indexada discreta no tempo, T_1 e T_2 são as constantes de tempo (ataque e relaxamento), escolhidas empiricamente através da análise do sinal de saída do filtro passa-altas, e T é o período de amostragem do sinal, ou seja, $T = \frac{1}{f_A} = 0.0000227s$, onde $f_A = 44.1kHz$ é a frequência de amostragem para o padrão CD (Compact Disk). Os melhores resultados foram obtidos para $T_1 = 5\mu s$ e $T_2 = 0.5ms$.

4.1.5 Derivada da Envoltória e Comparação com Limiar

Uma melhor resolução para o processo de detecção das amostras degradadas é obtida utilizando-se um circuito derivador que atua sobre o sinal obtido na saída do detector de envoltória. A derivada deste sinal é obtida simplesmente pela diferença entre cada duas amostras consecutivas, ou seja, $\Delta y = y_2(n) - y_2(n-1)$, de forma a se obter um sinal mais descorrelacionado, como mostra a Figura 4.9, que possa ser comparado com o limiar de referência pré-estabelecido para uma melhor eficiência na detecção da degradação. No caso de $\Delta y < \text{limiar}$, a amostra é classificada como não degradada; no entanto, se $\Delta y > \text{limiar}$, a amostra é classificada como degradada e é ativada a etapa de filtragem. Como indicado anteriormente, foi utilizado como limiar o valor $F \cdot \sigma_{med}$.

4.2 Filtragem: Algoritmo de Estimação do Sinal

Após a detecção seletiva da degradação, conforme descrito anteriormente, o segmento do sinal considerado degradado é então submetido a um processo de estimação

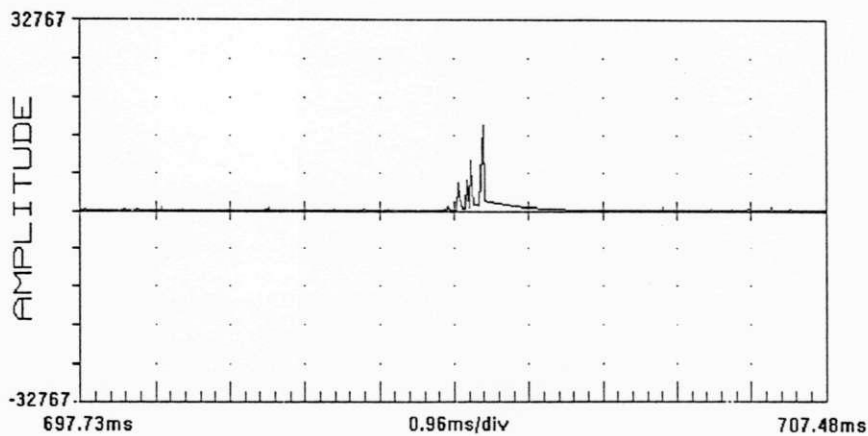


Figura 4.9: Sinal na Saída do Estágio Derivador

por predição linear, onde os valores das amostras degradadas são substituídos por valores estimados em função de amostras do sinal situadas fora do intervalo de degradação do sinal.

A Figura 4.10 apresenta as etapas de filtragem utilizadas no sistema, onde $s(n)$ é o sinal original, $r(n)$ é o ruído, $x(n)$ é o sinal degradado, $z(n)$ é o sinal na saída da predição linear e $f(n)$ é o sinal na saída da interpolação de Lagrange.

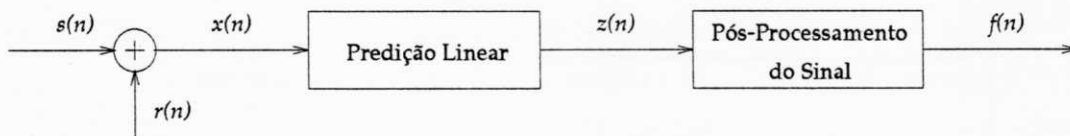


Figura 4.10: Estimativa Seletiva do Sinal Degradado

4.2.1 Predição Linear Progressiva

Uma das técnicas de predição linear utilizada foi a predição progressiva [4, 7], que permite a estimação de amostras através da observação dos valores anteriores. Uma amostra atual do sinal de áudio $x(n)$ pode ser estimada a partir dos valores das amostras passadas $x(n - k)$; $k = 1, 2, 3, \dots, p$, ou seja,

$$z(n) = \sum_{k=1}^p a_k x(n - k) \quad (4.14)$$

onde p representa o número de coeficientes do preditor, a_k representam os coeficientes do preditor que são calculados adaptativamente, ou seja, a curtos intervalos de tempo, de forma a poderem acompanhar a variação da estatística do sinal em função do tempo, e $z(n)$ corresponde ao valor estimado. Para o cálculo dos coeficientes do preditor é utilizado um segmento imediatamente anterior, fora do intervalo de degradação. Uma abordagem detalhada sobre os algoritmos utilizados para o cálculo desses coeficientes já foi apresentada no Capítulo 3. A Figura 4.11 ilustra o processo de estimação progressiva em um trecho de curta duração.

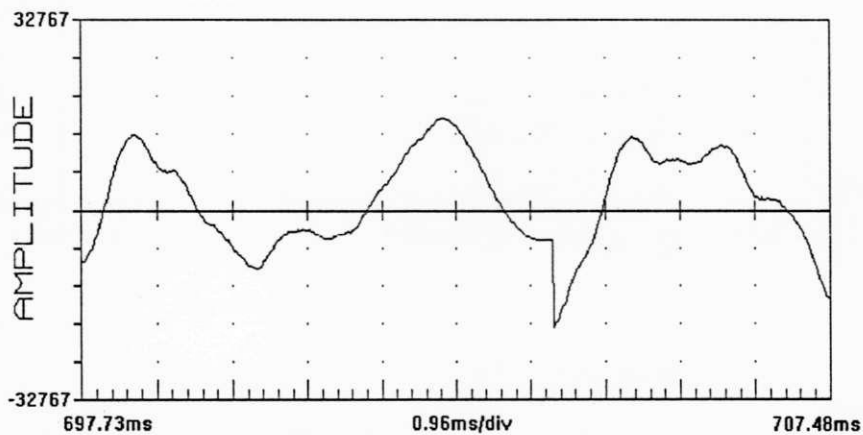


Figura 4.11: Sinal na Saída da Estimação Progressiva

4.2.2 Predição Linear Regressiva

Uma outra técnica de predição linear utilizada foi a predição regressiva, que utiliza amostras “futuras” com relação a um dado instante de referência no processo de predição. Uma amostra atual do sinal de áudio $x(n)$ pode ser estimada a partir dos valores das amostras “futuras” $x(n + k)$, $k = 1, 2, 3, \dots, p$, ou seja:

$$z(n) = \sum_{k=1}^p a_k x(n+k) \quad (4.15)$$

O cálculo dos coeficientes segue o mesmo padrão da predição progressiva, sendo que tomando o segmento imediatamente posterior, fora do intervalo de degradação. A Figura 4.12 ilustra o processo de estimação regressiva no mesmo trecho de curta duração utilizado na predição progressiva.

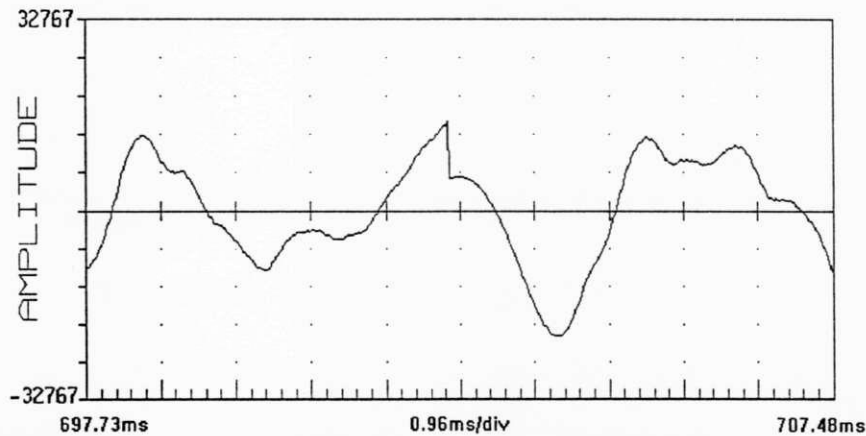


Figura 4.12: Sinal na Saída da Estimação Regressiva

4.2.3 Predição Linear Dupla

A predição dupla de uma amostra em um n-ésimo instante de tempo de referência é obtida calculando a média dos valores de estimação obtidos de amostras passadas (predição progressiva) e amostras futuras (predição regressiva), respectivamente, ou seja:

$$z(n) = \frac{1}{2} \left\{ \sum_{k=1}^p a_{kp} x(n-k) + \sum_{k=1}^p a_{kr} x(n+k) \right\} \quad (4.16)$$

onde a_{kp} são os coeficientes obtidos a partir da predição progressiva e a_{kr} são os coeficientes obtidos a partir da predição regressiva. A Figura 4.13 ilustra o processo

de estimação dupla no mesmo trecho de curta duração utilizado nos outros tipos de estimação mencionados, a fim de proporcionar uma análise comparativa da eficiência deste método.

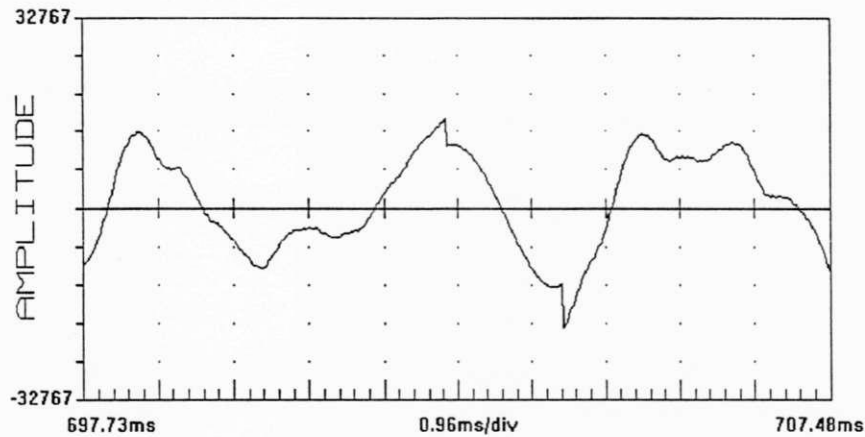


Figura 4.13: Sinal na Saída da Estimação Dupla

4.2.4 Pós-processamento do Sinal Estimado: Interpolação Polinomial de Lagrange

Após o processo de estimação linear, o sinal obtido apresenta um certo ruído residual que pode ser minimizado de maneira bastante eficiente através de uma etapa de pós-processamento do sinal baseada na interpolação polinomial de Lagrange, cujo teorema é descrito a seguir.

Teorema. Sejam (x_i, v_i) , $i = 0, 1, 2, \dots, N, N + 1$ pontos distintos, ou seja, $x_i \neq x_j$ para $i \neq j$, conforme mostra a Figura 4.14. Existe um único polinômio $P(x)$ de grau não maior que N , tal que $P(x_i) = v_i$, para todo i [5].

O polinômio $P(x)$ pode ser escrito na forma:

$$P_N(x) = a_0 + a_1x + a_2x^2 + \dots + a_Nx^N \quad \text{ou} \quad P_N(x) = \sum_{i=0}^N a_i x^i \quad (4.17)$$

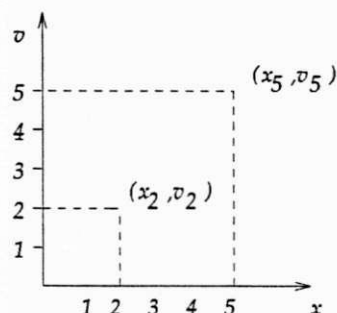


Figura 4.14: Representação de pontos no plano cartesiano

onde $P(x)$ é, no máximo, de grau N , se $a_N \neq 0$ e, para determiná-lo, deve-se conhecer os valores de a_0, a_1, \dots, a_N . Como $P_N(x)$ contém os pontos (x_i, v_i) , $i = 0, 1, \dots, N$, pode-se escrever que $P_N(x_i) = v_i$. Logo,

$$S \begin{cases} a_0 + a_1x_0 + a_2x_0^2 + \dots + a_Nx_0^N = v_0 \\ a_0 + a_1x_1 + a_2x_1^2 + \dots + a_Nx_1^N = v_1 \\ \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \\ a_0 + a_1x_N + a_2x_N^2 + \dots + a_Nx_N^N = v_N \end{cases} \quad (4.18)$$

Resolvendo o sistema S , determina-se o polinômio $P_N(x)$, que pode ser escrito na forma

$$P_N(x) = \sum_{i=0}^N v_i \prod_{j=0, j \neq i}^N \frac{(x - x_j)}{(x_i - x_j)} \quad (4.19)$$

conhecida como o Polinômio Interpolador de Lagrange [5].

Para provar que $P_N(x)$ é único, basta que se mostre que o determinante da matriz A , dos coeficientes das incógnitas do sistema S , é diferente de zero. A matriz A é:

$$A = \begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^N \\ 1 & x_1 & x_1^2 & \dots & x_1^N \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_N & x_N^2 & \dots & x_N^N \end{bmatrix} \quad (4.20)$$

O determinante da matriz A é conhecido como determinante de Vandermonde [13] e seu valor é dado por:

$$\det(A) = \prod_{i>j} (x_i - x_j) \tag{4.21}$$

Para a aplicação do polinômio de Lagrange na etapa de pós-filtragem de uma seqüência de amostras, convém realizar-se a seguinte adaptação:

$$f_N(n) = P_N(n) = \sum_{i=0}^N f_i \prod_{j=0, j \neq i}^N \frac{(n - n_j)}{(n_i - n_j)} \tag{4.22}$$

onde n é a posição da amostra no segmento que se deseja processar e f é o valor da amplitude da amostra, conforme mostra a Figura 4.15. Foram utilizadas quatro amostras próximas à descontinuidade a fim de proporcionar uma interpolação de ordem 3. Duas delas escolhidas de um trecho imediatamente anterior à descontinuidade e as outras duas escolhidas de um trecho imediatamente posterior.

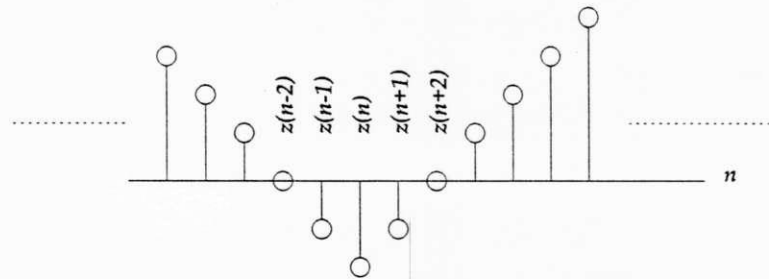


Figura 4.15: Representação de uma seqüência de amostras

A título de ilustração do funcionamento do polinômio de Lagrange, será descrito abaixo um exemplo de uma interpolação de ordem 3.

Exemplo: Dada a tabela abaixo, determinar $f(10)$, ou seja, o valor da décima amostra.

Como pode ser visto na tabela, existem quatro amostras disponíveis para a interpolação, o que permite uma ordem igual a 3. Assim sendo, pode-se aplicar a fórmula do polinômio interpolador de Lagrange:

i	n_i	$f(n_i)$
0	6	200
1	8	230
2	12	180
3	15	150

$$f_3(10) = \sum_{i=0}^3 f_i \prod_{j=0, j \neq i}^3 \frac{(n - n_j)}{(n_i - n_j)} \quad (4.23)$$

$$\begin{aligned} f_3(10) &= f_0 \frac{(n-n_1)(n-n_2)(n-n_3)}{(n_0-n_1)(n_0-n_2)(n_0-n_3)} \\ &+ f_1 \frac{(n-n_0)(n-n_2)(n-n_3)}{(n_1-n_0)(n_1-n_2)(n_1-n_3)} \\ &+ f_2 \frac{(n-n_0)(n-n_1)(n-n_3)}{(n_2-n_0)(n_2-n_1)(n_2-n_3)} \\ &+ f_3 \frac{(n-n_0)(n-n_1)(n-n_2)}{(n_3-n_0)(n_3-n_1)(n_3-n_2)} \end{aligned} \quad (4.24)$$

$$\begin{aligned} f_3(10) &= 200 \frac{(10-8)(10-12)(10-15)}{(6-8)(6-12)(6-15)} \\ &+ 230 \frac{(10-6)(10-12)(10-15)}{(8-6)(8-12)(8-15)} \\ &+ 180 \frac{(10-6)(10-8)(10-15)}{(12-6)(12-8)(12-15)} \\ &+ 150 \frac{(10-6)(10-8)(10-12)}{(15-6)(15-8)(15-12)} \end{aligned} \quad (4.25)$$

$$f_3(10) = 247.9 \quad (4.26)$$

As Figuras 4.16, 4.17 e 4.18 ilustram a etapa de pós-filtragem do segmento de curta duração analisado, com uma interpolação polinomial de Lagrange de ordem 3 após a estimação progressiva, regressiva e dupla, respectivamente. A título de comparação, é

mostrado na Figura 4.19 o segmento de curta duração do sinal original sem degradação.

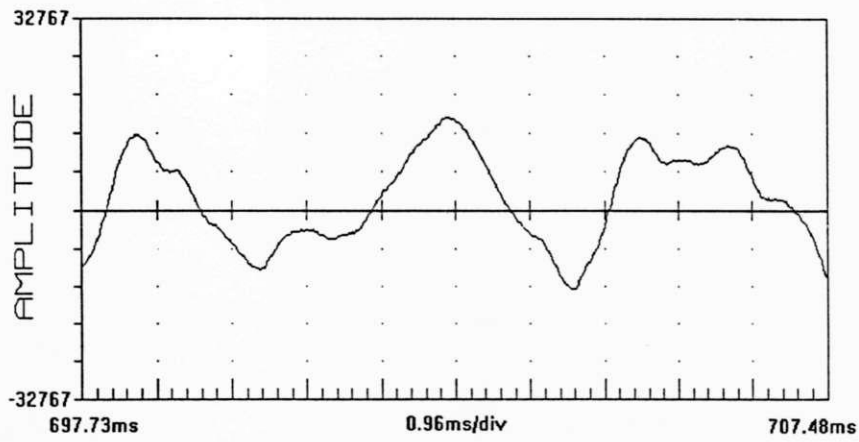


Figura 4.16: Pós-Filtragem após a Estimação Progressiva

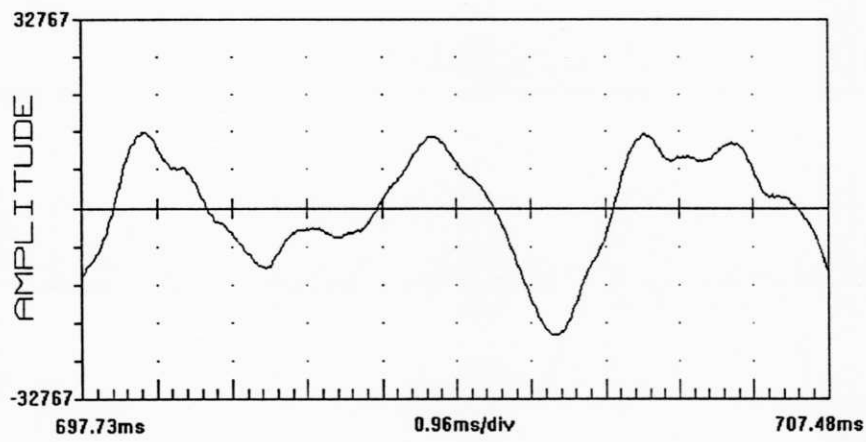


Figura 4.17: Pós-Filtragem após a Estimação Regressiva

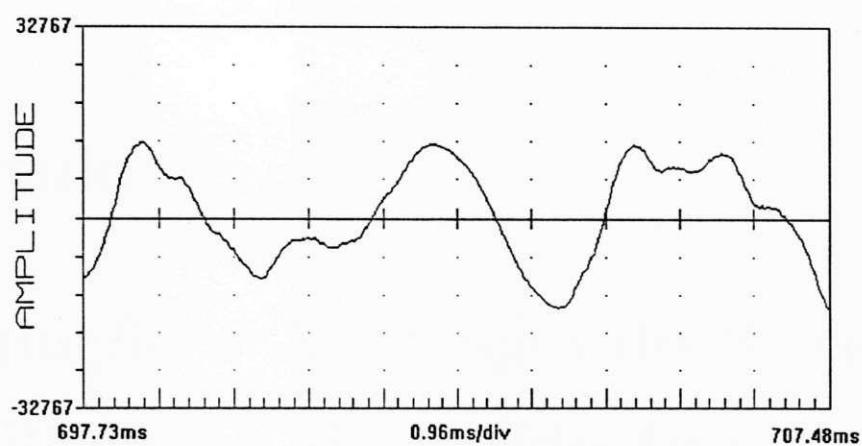


Figura 4.18: Pós-Filtragem após a Estimação Dupla

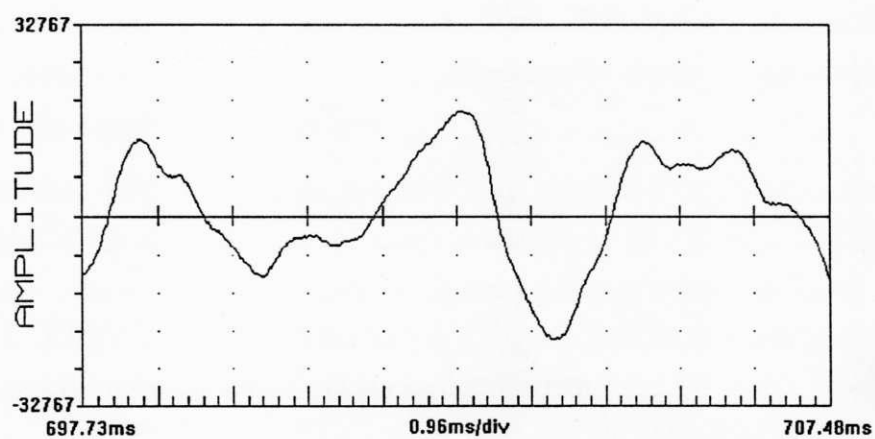


Figura 4.19: Sinal Original sem degradação

Capítulo 5

Simulação e Avaliação do Sistema de Filtragem do Ruído Impulsivo

5.1 Introdução

Devido à sua surpreendente acuracidade, o ouvido humano pode detectar níveis bastante reduzidos de distorção, entretanto, a presença de distorção pode não ser perceptível em muitas situações. Assim sendo, um bom critério para avaliação da qualidade do sinal é baseado no fato de que se o ouvido não pode detectar degradações, pode-se dizer que o som reproduzido é perfeito.

Normalmente o ouvido de uma pessoa é mais sensível à frequências situadas entre 2kHz e 5kHz. Esta sensibilidade cai aproximadamente 20 dB para frequências superiores ou inferiores a esta faixa. Embora algumas pessoas possam detectar frequências em torno de 20kHz, existem evidências para sugerir que muitos ouvintes não podem distinguir se o limite mais alto é 16kHz ou 20kHz [26].

Alguns estudos sobre audibilidade, realizados em 1931, estabeleceram que a faixa compreendida entre 40Hz e 15kHz é suficiente para a reprodução de música. Posteriormente foi constatado que frequências na faixa de 20Hz podem contribuir para uma melhor qualidade do sinal. Daí, chegou-se à conclusão que, para um sistema de áudio,

a largura de faixa ideal deve estar entre 0Hz e 20kHz e uma largura de faixa aceitável pode estar entre 20Hz e 15kHz [6].

Para qualidade CD, a frequência de amostragem de 44.1kHz permite uma largura de faixa de áudio de 20kHz, ou seja, a frequência de amostragem é maior que o dobro da maior componente de frequência do sinal e as amostras são uniformemente quantizadas a 16 bits [18]. É importante enfatizar que o processo de amostragem não elimina nenhum trecho de informação do sinal, não gera componentes de ruído nem degrada a qualidade do sinal.

5.2 Condições Experimentais

5.2.1 Digitalização do Sinal

Para aquisição dos dados foi utilizada a placa de aquisição de dados SOUND BLASTER 16 ASP instalada em um PC-AT 486, que proporcionou a digitalização do sinal de áudio com os seguintes parâmetros de referência: frequência de amostragem de 44.1kHz (qualidade CD), 16 bits por amostra, mono.

5.2.2 Escolha da Ordem do Preditor e do Tamanho do Segmento

O primeiro passo após a digitalização do sinal foi a realização de testes para a escolha da ordem do preditor e do tamanho do segmento para o processamento do sinal, com o objetivo de se obter uma boa relação sinal-ruído (RSR).

Na escolha do tamanho do segmento, foi observada a literatura disponível [7] que indica quadros de 20 a 30 ms. Considerando uma frequência de amostragem de 44.1 kHz, encontra-se um tamanho ideal com 1024 amostras. A título de comparação, foram realizados testes com quadros de 256 e 512 amostras, através dos quais pode-se comprovar que, no que se refere aos resultados numéricos, a diferença entre eles é

pequena. Além disso, a velocidade de processamento do segmento de 1024 amostras é bastante satisfatória.

Quanto à ordem do preditor, em [25] é proposta a seguinte equação para sua determinação:

$$\text{ordem} = \frac{\text{frequência de amostragem}}{1000} \tag{5.1}$$

Sendo a frequência de amostragem igual a 44.1kHz, a ordem indicada seria em torno de 44, que é considerada bastante alta, uma vez que eleva consideravelmente o nível de complexidade dos cálculos dos coeficientes do preditor. Entretanto, os testes comprovaram que, a partir da ordem igual a 8, a tendência é se obter uma RSR quase constante, como mostra a Figura 5.1, o que possibilita uma boa predição com menor volume de cálculos matemáticos.

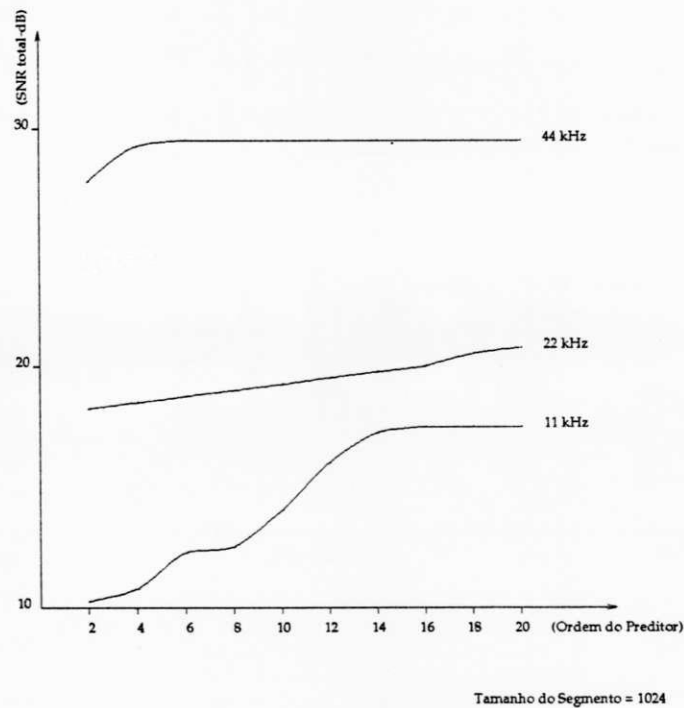


Figura 5.1: Curva da RSR para escolha da Ordem do Preditor

Uma explicação para se ter encontrado uma ordem relativamente baixa, mas, em contrapartida, bastante satisfatória, deve-se ao fato de que, com uma frequência de amostragem mais alta, tem-se amostras bem próximas, que fornecem informações bastante precisas sobre o sinal.

5.3 Procedimento Experimental

Visando a realização de testes do algoritmo, foram gravados alguns trechos variados, de aproximadamente 2, 3 e 4 segundos de duração, de músicas instrumentais com cravo (Figura 5.2), flauta (Figura 5.3) e piano (Figura 5.4), além de violão e voz (Figura 5.5), completamente livres de ruídos, utilizando o aplicativo *Creative WaveStudio*, que grava, reproduz e edita arquivos no formato .WAV. Nas gravações foram obedecidos os parâmetros de referência estabelecidos na Seção 5.2.1.

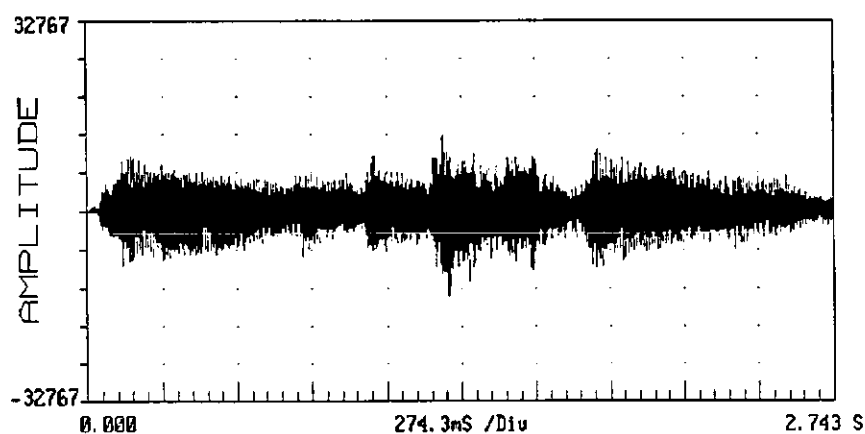


Figura 5.2: Som de cravo digitalizado

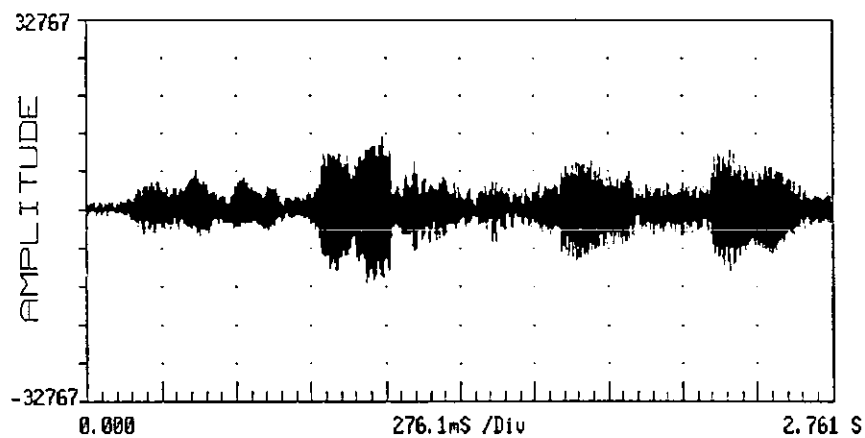


Figura 5.3: Som de flauta digitalizado

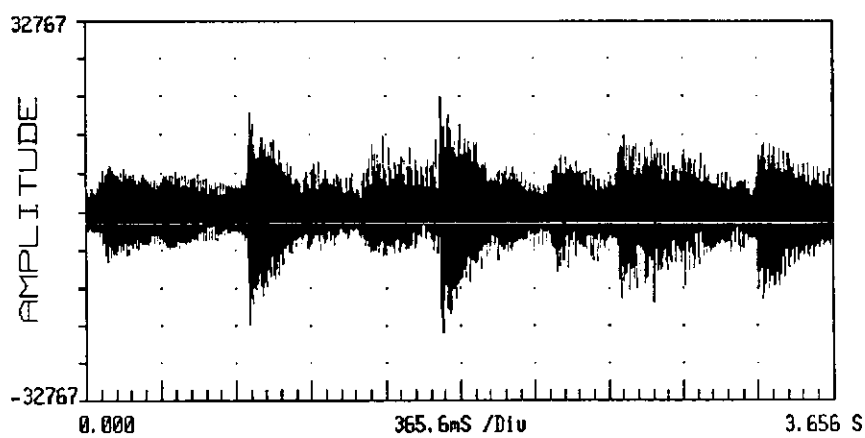


Figura 5.4: Som de piano digitalizado

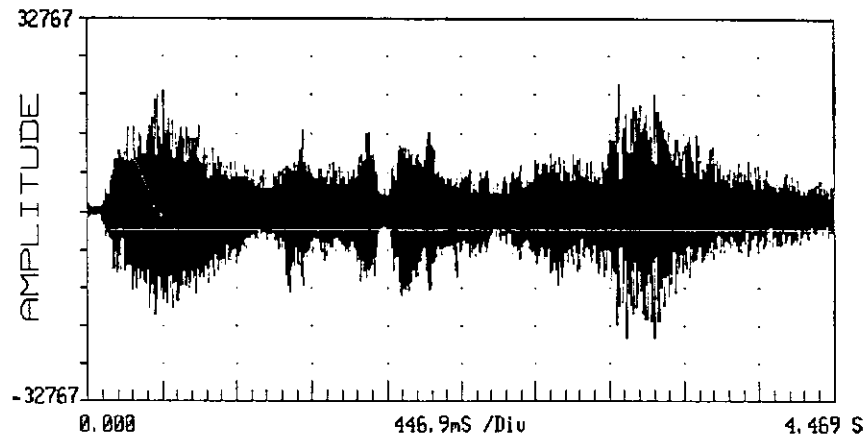


Figura 5.5: Som de violão e voz digitalizado

Em seguida foram adicionados ruídos impulsivos do tipo agudo, com duração aproximada de 0.5ms, escolhidos da biblioteca de ruído construída a partir de ruídos isolados em músicas já degradadas. A degradação foi realizada de forma aleatória em cada uma das músicas originais. O resultado desse processo foram as músicas degradadas apresentadas nas Figuras 5.6, 5.7, 5.8 e 5.9.

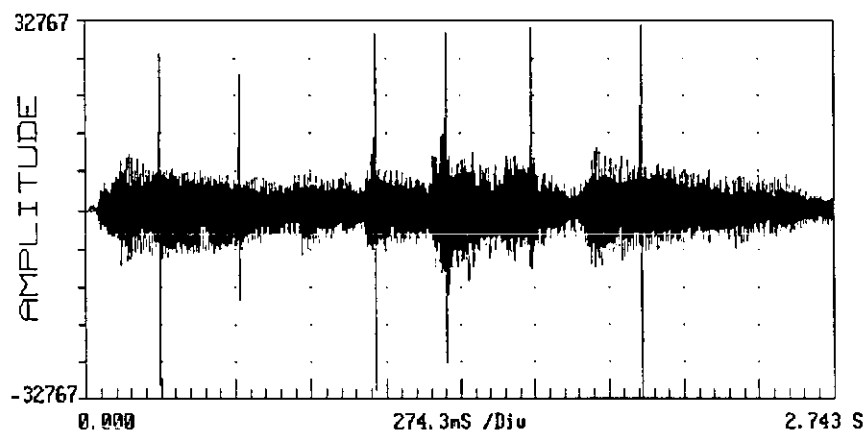


Figura 5.6: Som de cravo degradado

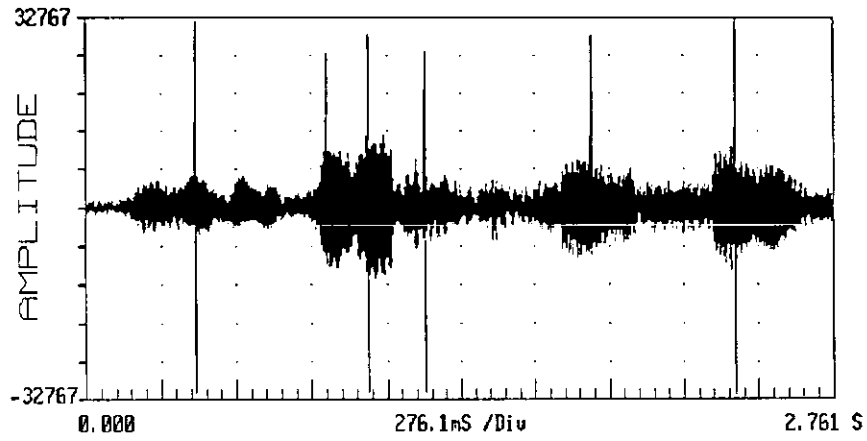


Figura 5.7: Som de flauta degradado

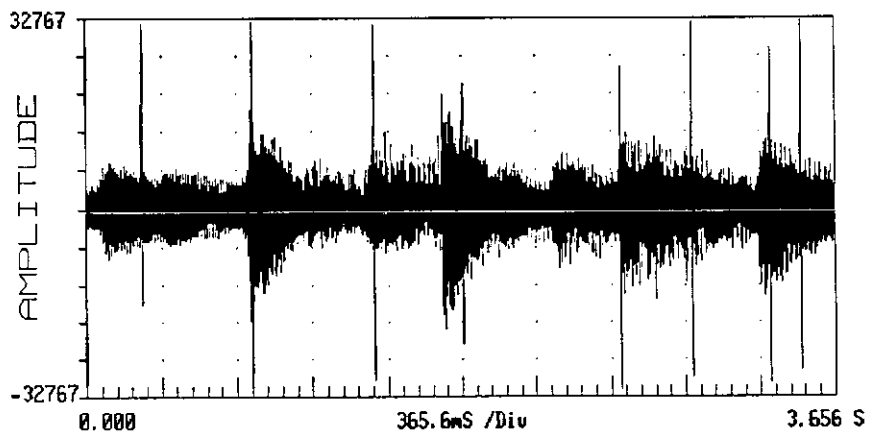


Figura 5.8: Som de piano degradado

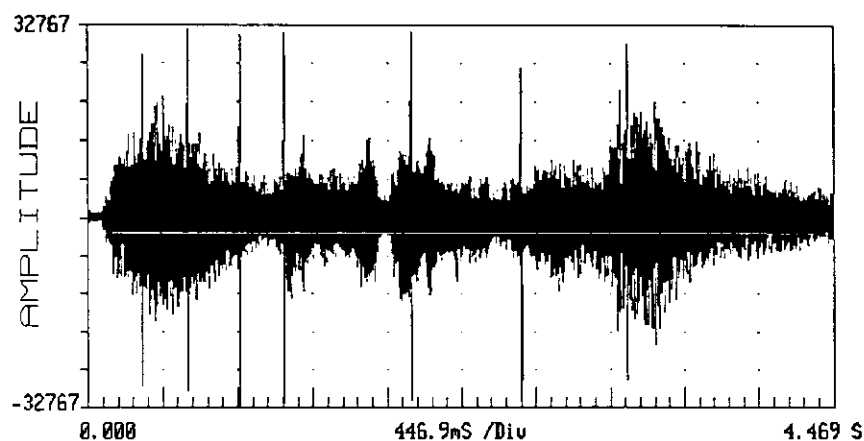


Figura 5.9: Som de violão e voz degradado

5.3.1 Etapa de Detecção

A análise, em tempo não real, de cada um dos sinais degradados apresentados acima, foi feita utilizando os programas implementados na linguagem C, que executam as funções apresentadas no diagrama de blocos da Figura 4.1.

A primeira etapa do processo de detecção do ruído impulsivo é a filtragem do sinal pelo filtro passa-altas, cujo objetivo é reforçar o ruído para facilitar sua detecção, como pode ser visto nas Figuras 5.10, 5.11, 5.12 e 5.13.

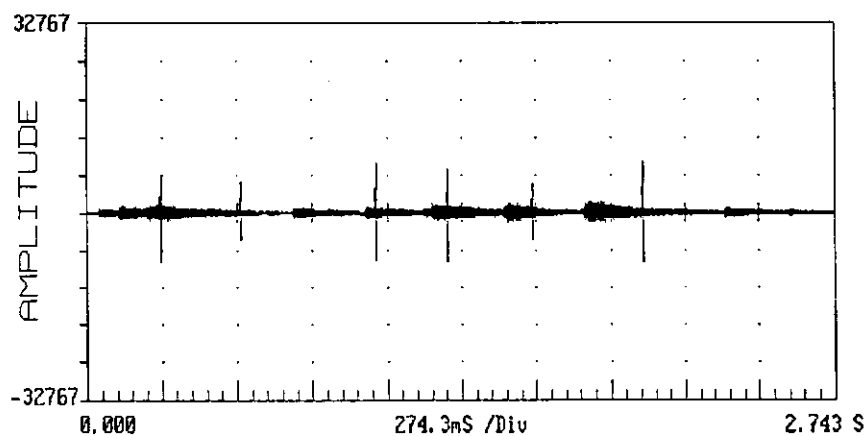


Figura 5.10: Som de cravo na saída do filtro passa-altas

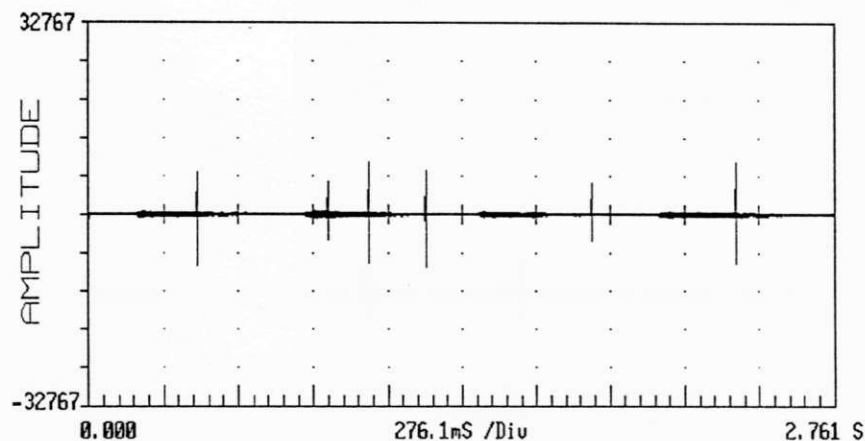


Figura 5.11: Som de flauta na saída do filtro passa-altas

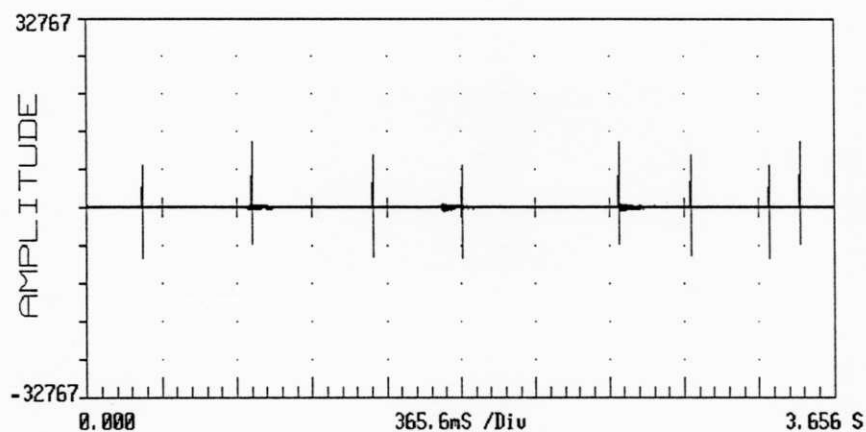


Figura 5.12: Som de piano na saída do filtro passa-altas

Em seguida, é calculada a envoltória de cada um dos sinais utilizados no teste, mostradas nas Figuras 5.14, 5.15, 5.16 e 5.17. Através da envoltória pode-se verificar que o trecho onde está localizado o ruído tem uma amplitude bem evidenciada em relação ao resto do sinal.

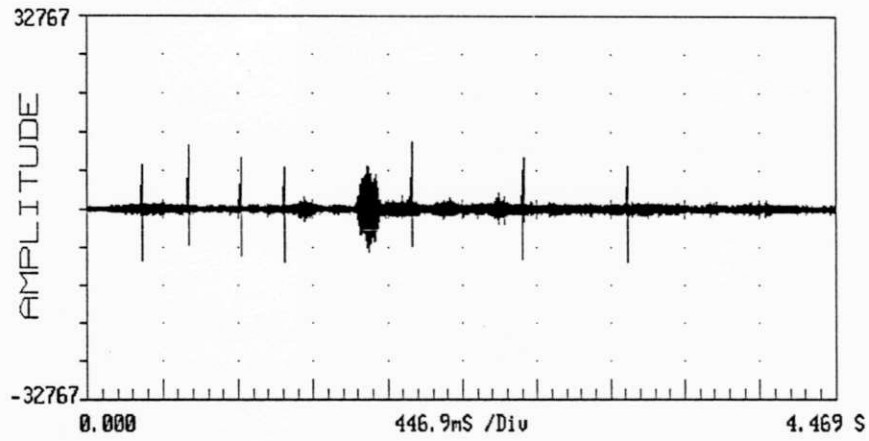


Figura 5.13: Som de violão e voz na saída do filtro passa-altas

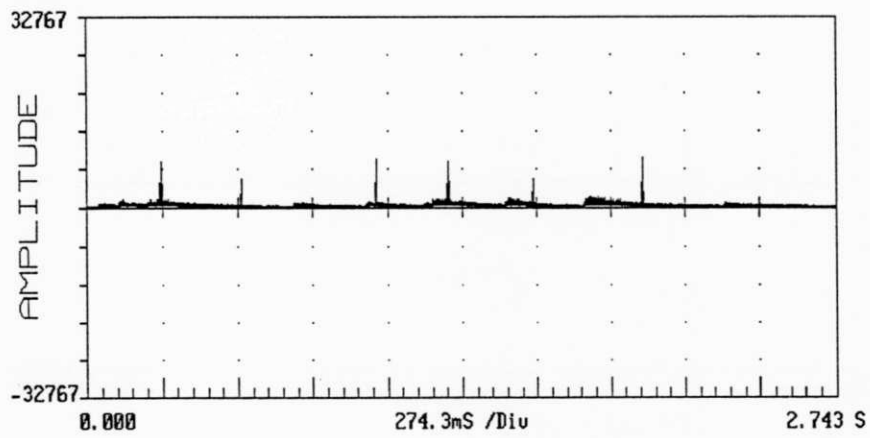


Figura 5.14: Som de cravo na saída do detector de envoltória

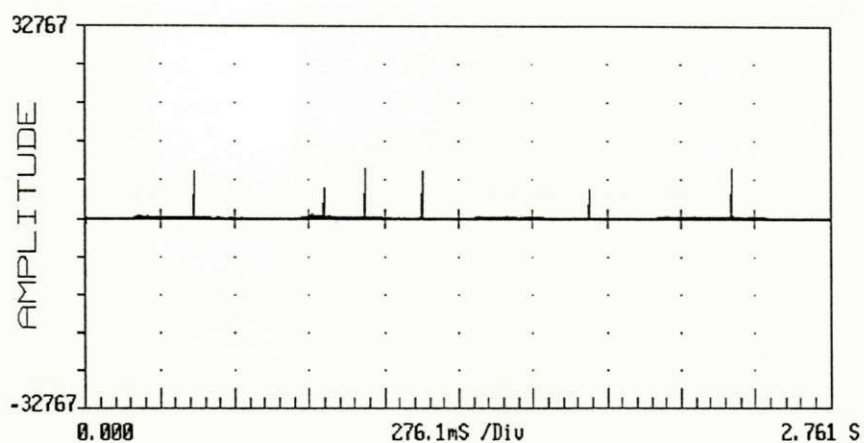


Figura 5.15: Som de flauta na saída do detector de envoltória

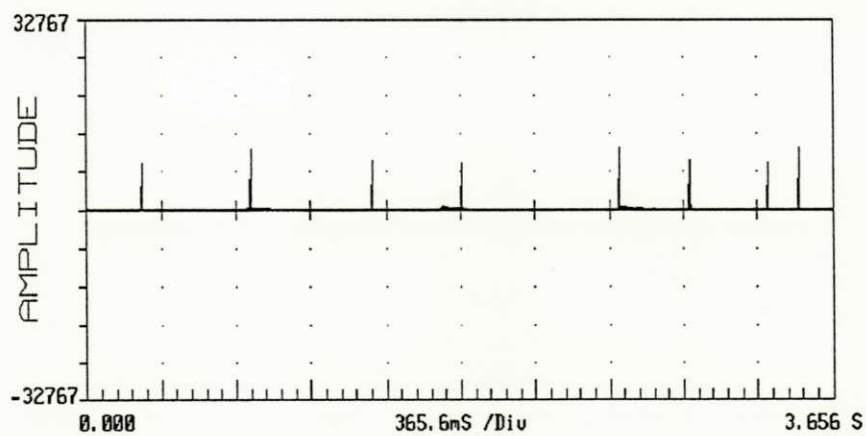


Figura 5.16: Som de piano na saída do detector de envoltória

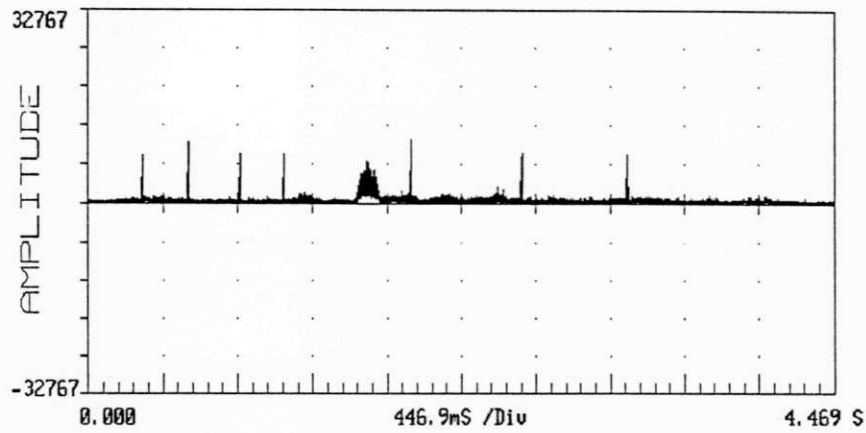


Figura 5.17: Som de violão e voz na saída do detector de envoltória

A detecção do ruído, como foi mencionado anteriormente, é feita através da comparação do limiar com a derivada da envoltória. A saída do estágio derivador de cada um dos trechos utilizados no teste é apresentada nas Figuras 5.18, 5.19, 5.20 e 5.21.

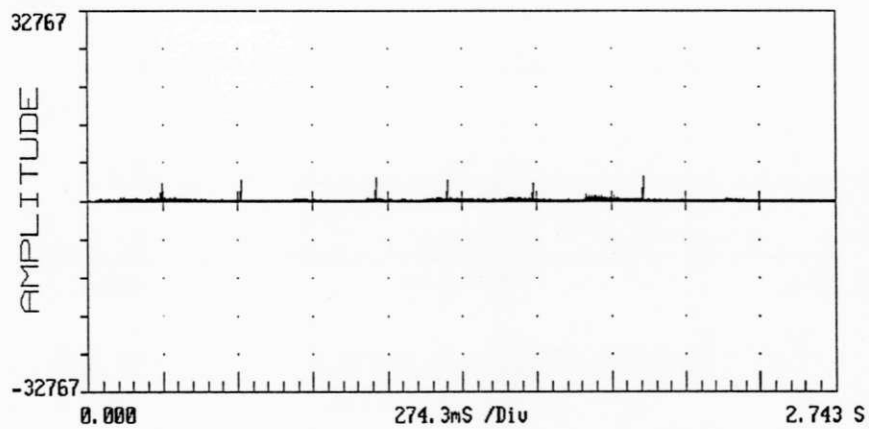


Figura 5.18: Som de cravo na saída do estágio derivador

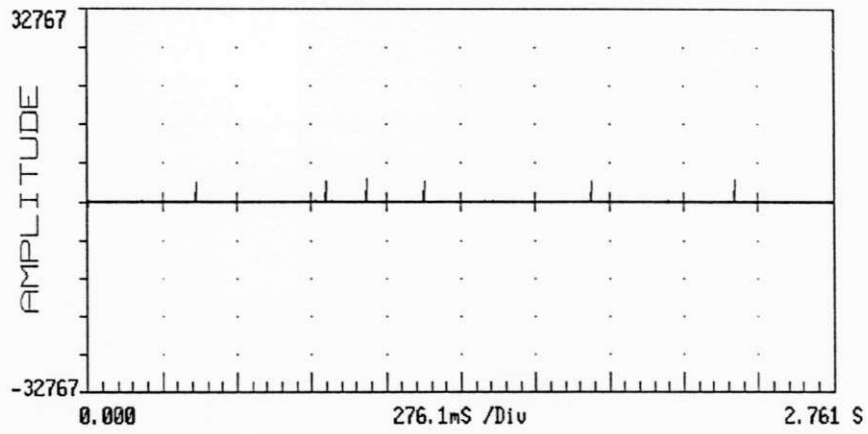


Figura 5.19: Som de flauta na saída do estágio derivador

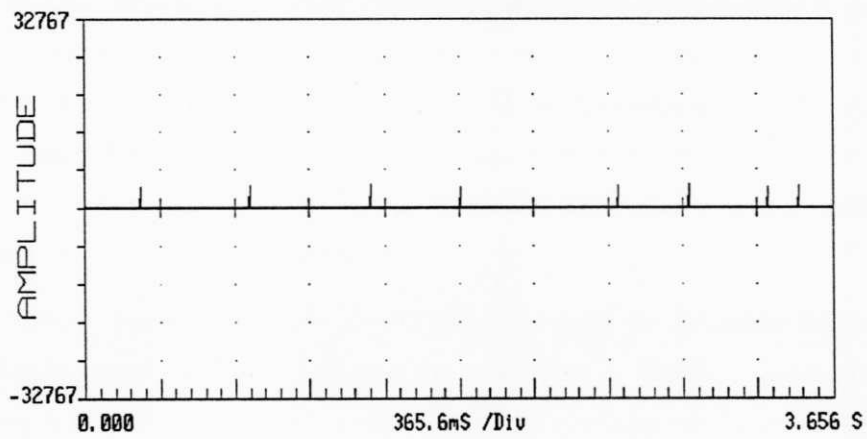


Figura 5.20: Som de piano na saída do estágio derivador

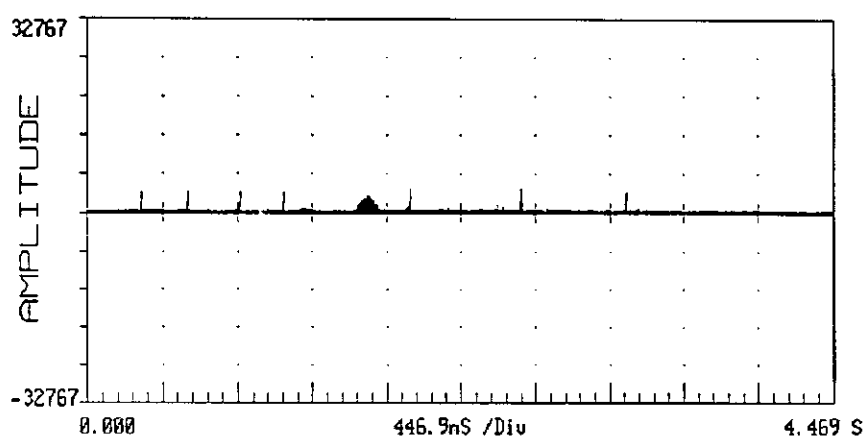


Figura 5.21: Som de violão e voz na saída do estágio derivador

5.3.2 Etapa de Filtragem

Uma vez identificado o ruído, é acionada a etapa de filtragem das amostras degradadas, realizada através do cálculo do valor estimado de cada uma das amostras degradadas utilizando as técnicas de predição linear progressiva, regressiva e dupla.

Na estimação progressiva, após detectada a posição da amostra degradada, é estabelecido, através de testes com a largura do pulso de ruído, o número de amostras, à direita e à esquerda, a serem restauradas. O cálculo adaptativo dos coeficientes do preditor é realizado com o segmento imediatamente anterior ao grupo de amostras degradadas. A primeira amostra do trecho é então estimada e seu valor corrigido, $\hat{x}(n)$, substitui o $x(n)$ nos próximos cálculos.

Na estimação regressiva, após detectada a posição da amostra degradada e, conseqüentemente, estabelecido o número de amostras, à direita e à esquerda, a serem restauradas, é realizado o cálculo adaptativo dos coeficientes do preditor utilizando o segmento imediatamente posterior ao grupo de amostras degradadas. A última amostra do trecho é então estimada e seu valor corrigido, $\hat{x}(n)$, substitui o $x(n)$ nos próximos cálculos, até ser obtido o valor estimado da primeira amostra do trecho degradado.

Finalmente, a estimação dupla é realizada de forma simples, através do cálculo da média dos valores obtidos na estimação progressiva e na estimação regressiva para cada uma das amostras do trecho degradado, conforme indicado na Equação 4.16.

Verifica-se, contudo, que a partir de um determinado número de amostras a predição não oferece tanta segurança, devido ao erro de predição acumulado. Este erro provoca, em geral, uma descontinuidade no final do trecho estimado, no caso de estimação progressiva, no início do trecho estimado, no caso de estimação regressiva, ou no final e no início do trecho estimado, no caso de predição dupla, ambas em relação ao restante do sinal.

A fim de resolver este problema de descontinuidade, foi efetuada uma pós-filtragem utilizando a Interpolação de Lagrange de ordem 3. Para tanto foram escolhidas quatro amostras, relativamente próximas à descontinuidade, que produziram uma excelente suavização do sinal, cuja saída é mostrada nas Figuras 5.22, 5.23, 5.24 e 5.25.

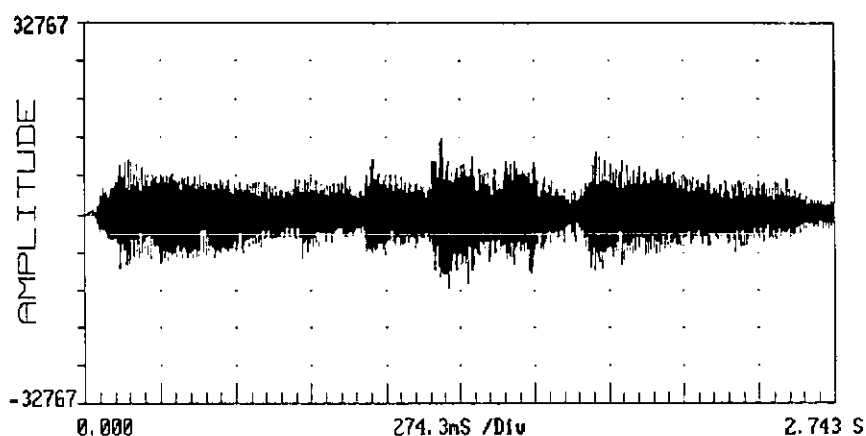


Figura 5.22: Som de cravo filtrado

A fim de proporcionar uma análise mais detalhada dos resultados obtidos, será dado um zoom em um dos trechos de cada um dos segmentos em análise, do sinal original, do sinal degradado e do sinal filtrado. As Figuras 5.26, 5.29, 5.32 e 5.35 mostram um trecho do sinal original sem degradação. As Figuras 5.27, 5.30, 5.33 e 5.36 mostram o mesmo trecho do sinal degradado. E as Figuras 5.28, 5.31, 5.34 e 5.37, por sua vez, mostram o sinal após a estimação regressiva e a interpolação de Lagrange, que foi a técnica que se mostrou mais eficiente.

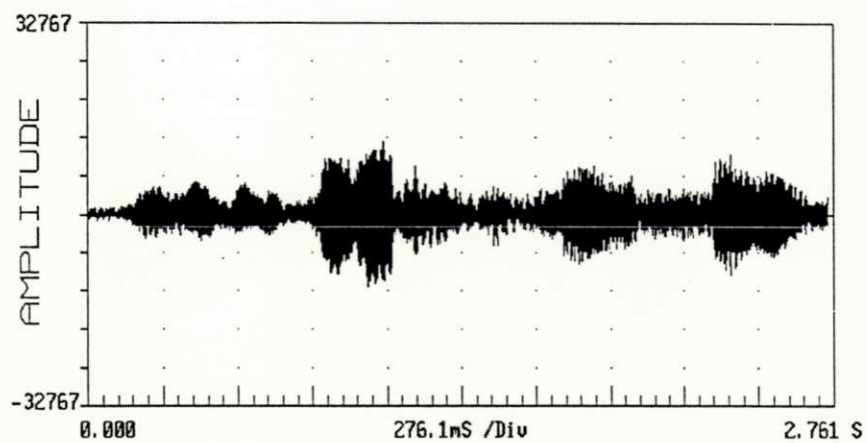


Figura 5.23: Som de flauta filtrado

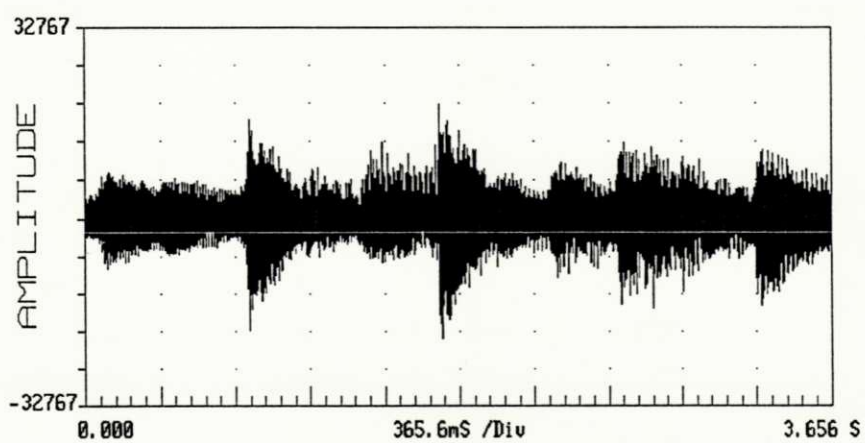


Figura 5.24: Som de piano filtrado

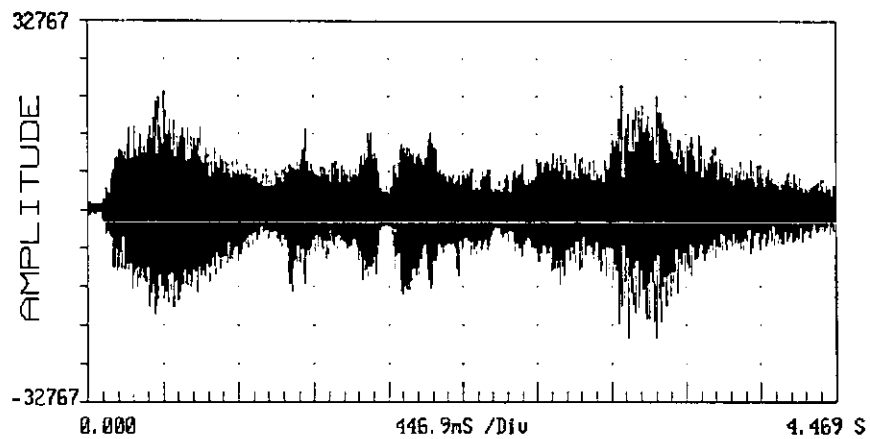


Figura 5.25: Som de violão e voz filtrado

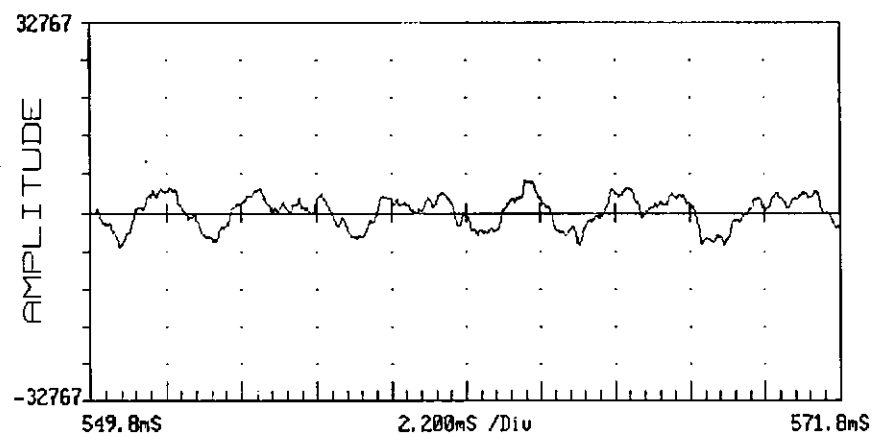


Figura 5.26: Zoom no som de cravo digitalizado

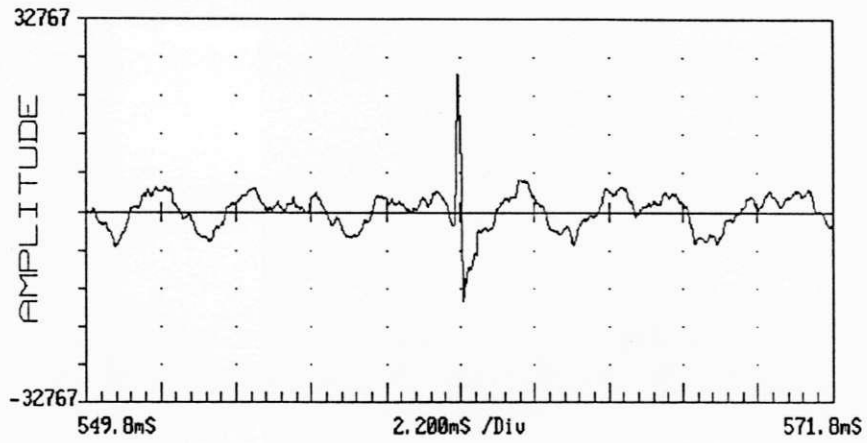


Figura 5.27: Zoom no som de cravo degradado

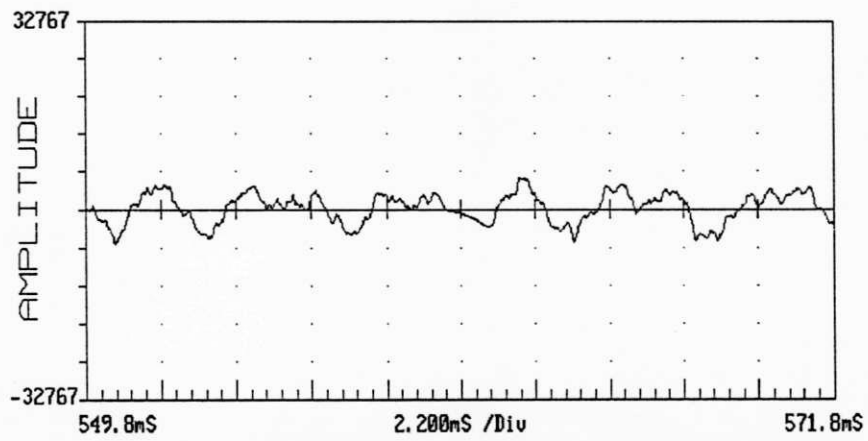


Figura 5.28: Zoom no som de cravo filtrado

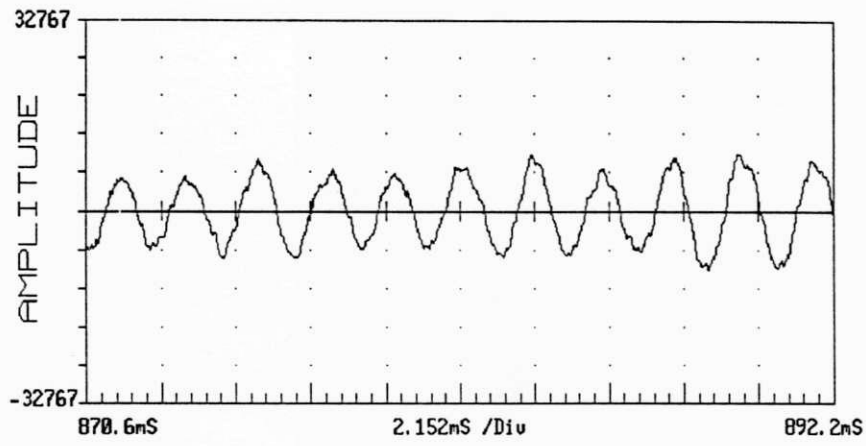


Figura 5.29: Zoom no som de flauta digitalizado

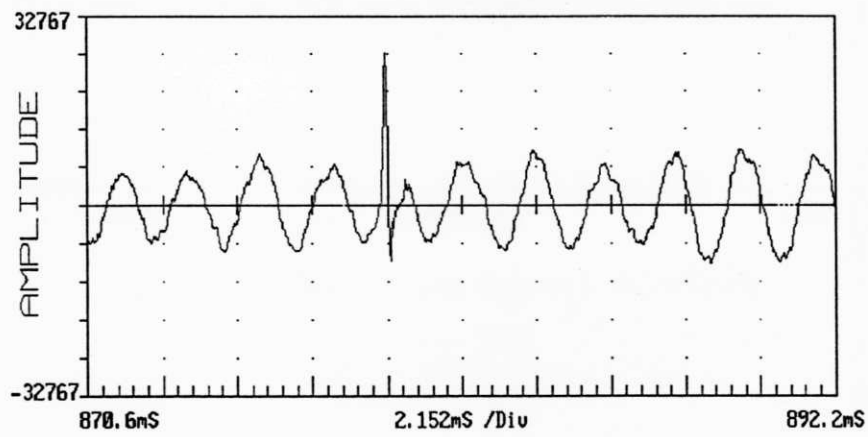


Figura 5.30: Zoom no som de flauta degradado

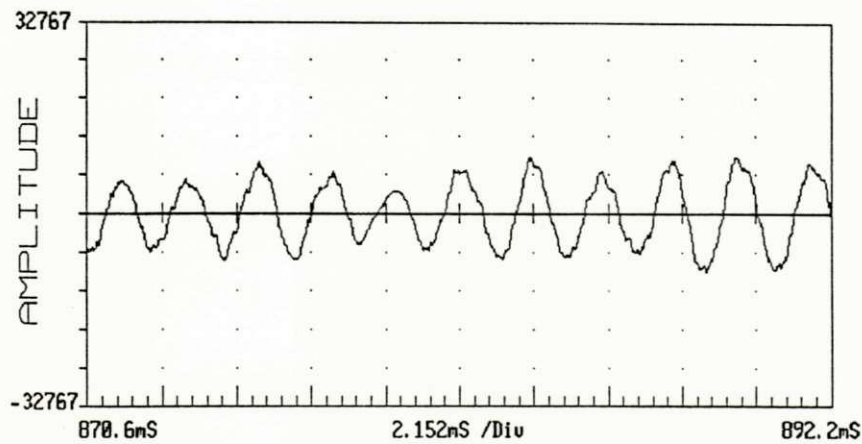


Figura 5.31: Zoom no som de flauta filtrado

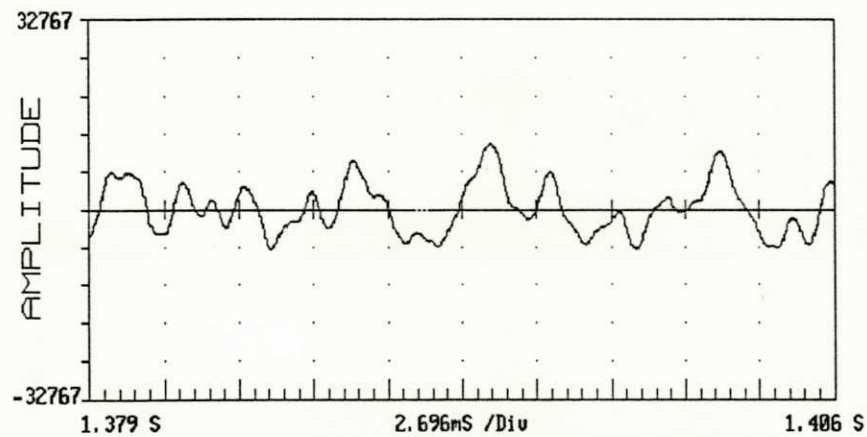


Figura 5.32: Zoom no som de piano digitalizado

A estimação por predição linear mostra-se bastante eficiente, especialmente quando o ruído impulsivo do tipo agudo está presente em trechos de baixa e média amplitude, que praticamente dispensam a pós-filtragem. No caso do ruído estar presente em

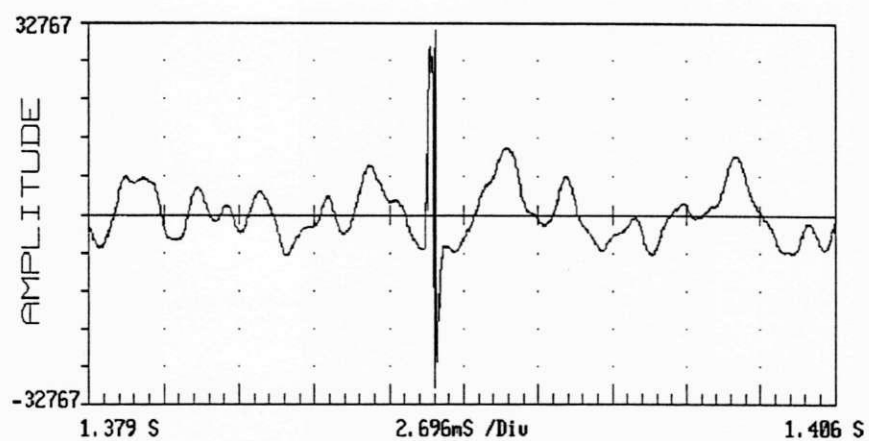


Figura 5.33: Zoom no som de piano degradado

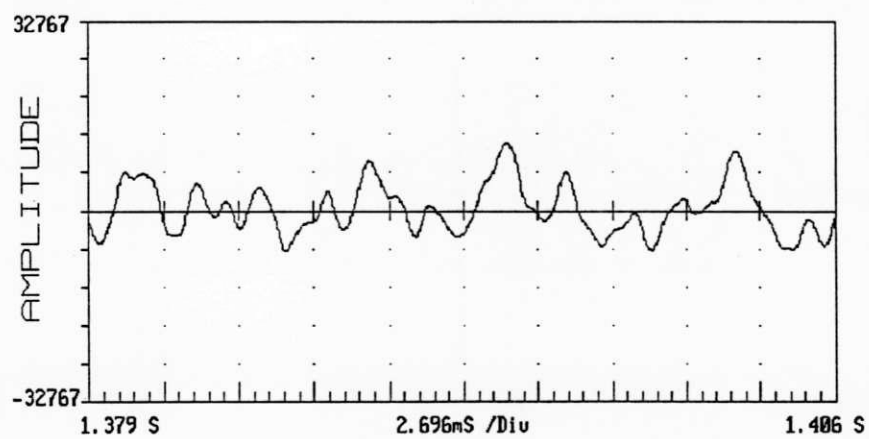


Figura 5.34: Zoom no som de piano filtrado

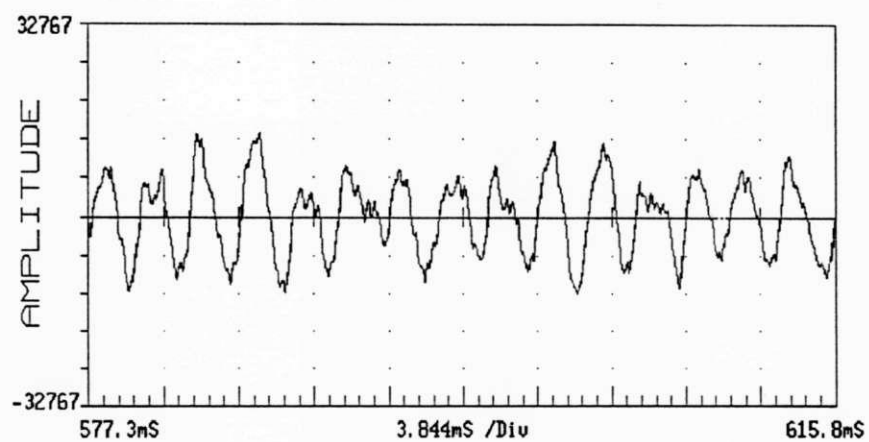


Figura 5.35: Zoom no som de violão e voz digitalizado

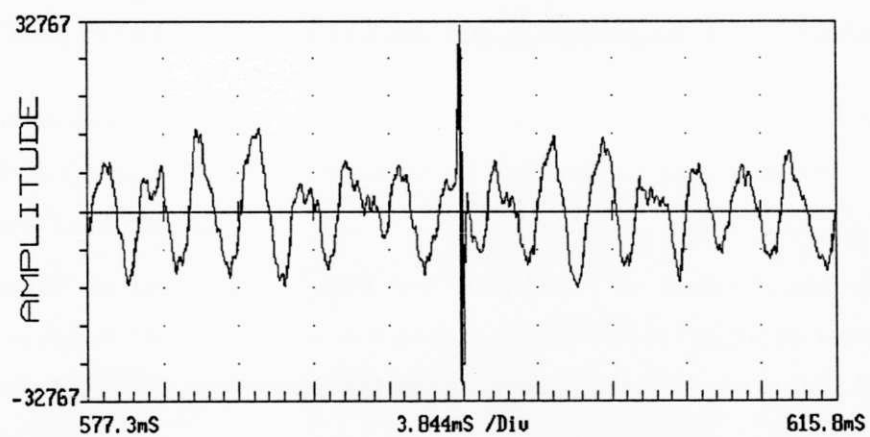


Figura 5.36: Zoom no som de violão e voz degradado

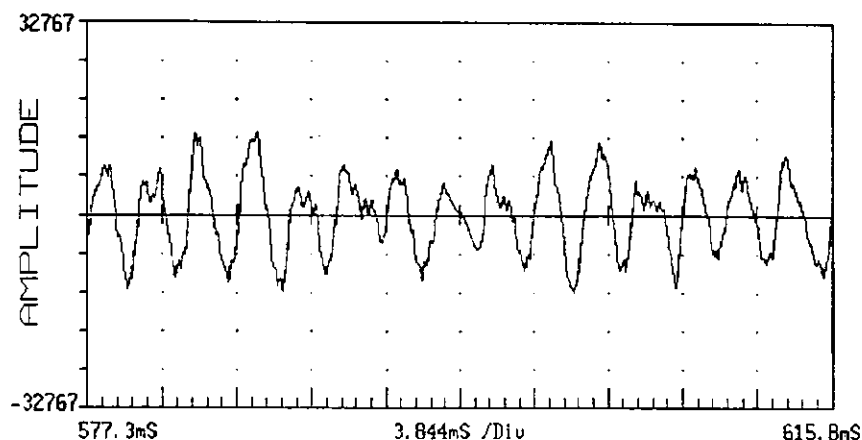


Figura 5.37: Zoom no som de violão e voz filtrado

trechos de alta amplitude, a Interpolação de Lagrange melhora sensivelmente o sinal, evitando descontinuidades abruptas.

5.4 Avaliação dos Resultados

5.4.1 Descrição dos Métodos de Avaliação Utilizados

Em sistemas de redução de ruído em sinais de áudio degradados, a avaliação da qualidade do sinal é realizada através de medidas objetivas na primeira fase e de medidas subjetivas na fase final [4].

Os métodos de avaliação objetiva são mais fáceis de serem implementados, são muito menos dispendiosos em tempo e podem ser repetidos inúmeras vezes durante o desenvolvimento de um algoritmo. Por outro lado, as técnicas de avaliação subjetivas são bastante dispendiosas em tempo e de difícil implementação, pois utilizam normalmente um número relativamente grande de pessoas. Entretanto não é possível avaliar de forma plena o desempenho de um sistema de processamento de áudio com base em critérios unicamente objetivos; é necessário complementar a avaliação com medidas subjetivas [4].

A seguir serão apresentadas, de forma breve, algumas das técnicas de avaliação objetiva e subjetiva da qualidade aplicadas a sistemas de áudio utilizadas neste trabalho.

Na **avaliação objetiva** foi utilizado o *Cálculo da Relação Sinal-Ruído (RSR)*, que indica a relação entre a potência PS do sinal original $s(n)$ e a potência PR do ruído $r(n)$ presente no sinal processado:

$$RSR = \frac{PS}{PR} = 10 \log_{10} \left(\frac{PS}{PR} \right) \quad \text{dB} \quad (5.2)$$

onde o sinal de ruído $r(n)$ é determinado pela diferença entre o sinal original, $s(n)$ e o sinal processado $z(n)$:

$$r(n) = s(n) - z(n) \quad (5.3)$$

Semelhantemente, os métodos de **avaliação subjetiva** utilizados foram:

O *Teste de Qualidade Absoluta: Escore de Opinião Média (MOS)*, que tem por objetivo avaliar a qualidade do sinal processado, no qual cada avaliador atribui uma nota dentre a escala de graduação apresentada na Tabela 5.1 e em seguida é calculada a média aritmética das notas obtidas.

5	A qualidade é excelente
4	A qualidade é boa
3	A qualidade é razoável
2	A qualidade é pobre
1	A qualidade é ruim

Tabela 5.1: Escala de Graduação do MOS

O *Teste de Qualidade Absoluta: Escore de Opinião Média (MOS) Invertido*, cujo objetivo é avaliar o nível de degradação do sinal a ser processado, é realizado de forma que cada avaliador, após a escuta do sinal degradado, atribua uma nota dentre a escala de graduação apresentada na Tabela 5.2. Em seguida é calculada a média aritmética das notas obtidas.

5	A degradação é muito incômoda
4	A degradação é incômoda
3	A degradação é pouco incômoda
2	A degradação é perceptível mas não incomoda
1	A degradação é imperceptível

Tabela 5.2: Escala de Graduação do MOS Invertido

5.4.2 Apresentação dos Resultados

A performance da avaliação objetiva, que apresenta o ganho da relação sinal-ruído (RSR) total, efetuada para cada um dos trechos analisados neste experimento, é mostrada na Tabela 5.3.

GANHO DA RSR-TOTAL (dB)						
Sinal de Música	Predição Linear					
	Progressiva		Regressiva		Dupla	
	s/ Lag	c/ Lag	s/ Lag	c/ Lag	s/ Lag	c/ Lag
cravo	6.93	6.96	8.19	8.37	7.17	7.30
flauta	9.85	9.92	10.01	10.29	9.80	9.87
piano	6.70	6.72	7.38	7.48	5.62	5.70
violão e voz	7.67	7.74	7.97	8.03	6.98	7.13

Tabela 5.3: Comparação entre as Técnicas de Predição Linear

Observando a tabela pode-se notar que os melhores resultados foram obtidos com a Predição Linear Regressiva com a etapa de pós-filtragem baseada na Interpolação de Lagrange, cujos valores da RSR estão em destaque.

A título de comparação, foi implementado o algoritmo proposto em [24], com a etapa de filtragem baseada na interpolação não linear. Os ganhos obtidos ficaram na faixa de 2 a 4 dB, observando as mesmas condições experimentais descritas neste trabalho. Estes resultados comprovam ainda mais a eficiência do método apresentado

neste trabalho, uma vez que os ganhos obtidos ficaram na faixa de 7 a 10 dB, como mostra a Tabela 5.3.

Embora o ganho da RSR após a interpolação de Lagrange não seja tão significativo, na prática ele oferece uma suavização na forma de onda do sinal que é bastante notada pelo ouvido humano, como foi constatado nos resultados dos testes subjetivos apresentados a seguir.

Para a avaliação subjetiva foi elaborado um formulário padrão, cujo modelo é apresentado no Apêndice A.

Foram utilizadas 10 pessoas para avaliar cada um dos trechos apresentados e as médias das notas obtidas por cada um deles na avaliação subjetiva utilizando o MOS Invertido são apresentadas na Tabela 5.4 e na avaliação subjetiva utilizando o MOS são apresentadas nas Tabelas 5.5 (cravo), 5.6 (flauta) e 5.7 (piano) e 5.8 (violão e voz).

Sinal Analisado	Média	Atributo
Som de cravo degradado	4.0	A degradação é incômoda
Som de flauta degradado	4.0	A degradação é incômoda
Som de piano degradado	3.9	A degradação é incômoda
Som de violão e voz degradado	4.3	A degradação é incômoda

Tabela 5.4: Resultados do MOS Invertido

Sinal Analisado	Média	Atributo
Degradado	1.9	A qualidade é pobre
Após Pred. Progressiva s/ Lagrange	3.3	A qualidade é razoável
Após Pred. Regressiva s/ Lagrange	3.4	A qualidade é razoável
Após Pred. Dupla s/ Lagrange	3.3	A qualidade é razoável
Após Pred. Regressiva c/ Lagrange	4.0	A qualidade é boa

Tabela 5.5: Resultados do MOS para o som de cravo

Sinal Analisado	Média	Atributo
Degradado	1.9	A qualidade é pobre
Após Pred. Progressiva s/ Lagrange	3.4	A qualidade é razoável
Após Pred. Regressiva s/ Lagrange	3.0	A qualidade é razoável
Após Pred. Dupla s/ Lagrange	2.9	A qualidade é razoável
Após Pred. Regressiva c/ Lagrange	4.1	A qualidade é boa

Tabela 5.6: Resultados do MOS para o som de flauta

Sinal Analisado	Média	Atributo
Degradado	2.0	A qualidade é pobre
Após Pred. Progressiva s/ Lagrange	3.3	A qualidade é razoável
Após Pred. Regressiva s/ Lagrange	3.1	A qualidade é razoável
Após Pred. Dupla s/ Lagrange	3.2	A qualidade é razoável
Após Pred. Regressiva c/ Lagrange	4.6	A qualidade é boa

Tabela 5.7: Resultados do MOS para o som de piano

Sinal Analisado	Média	Atributo
Degradado	2.2	A qualidade é pobre
Após Pred. Progressiva s/ Lagrange	4.2	A qualidade é boa
Após Pred. Regressiva s/ Lagrange	4.2	A qualidade é boa
Após Pred. Dupla s/ Lagrange	4.0	A qualidade é boa
Após Pred. Regressiva c/ Lagrange	4.4	A qualidade é boa

Tabela 5.8: Resultados do MOS para o som de violão e voz

5.4.3 Avaliação no Domínio da Frequência

Uma última avaliação foi feita através da visualização do sinal original, degradado e filtrado, no domínio da frequência, de cada um dos trechos analisados, *cravo* (Figuras 5.38, 5.39 e 5.40), *flauta* (Figuras 5.41, 5.42 e 5.43), *piano* (Figuras 5.44, 5.45 e 5.46) e *violão e voz* (Figuras 5.47, 5.48 e 5.49). Para tanto, foi utilizado um programa chamado *Hipersignal*, que calcula automaticamente a Transformada Rápida de Fourier (FFT).

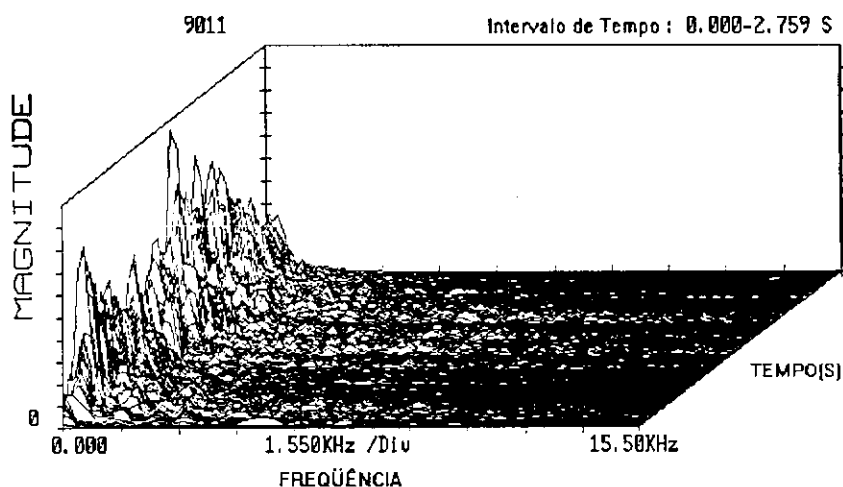


Figura 5.38: Espectro tridimensional do som de cravo digitalizado

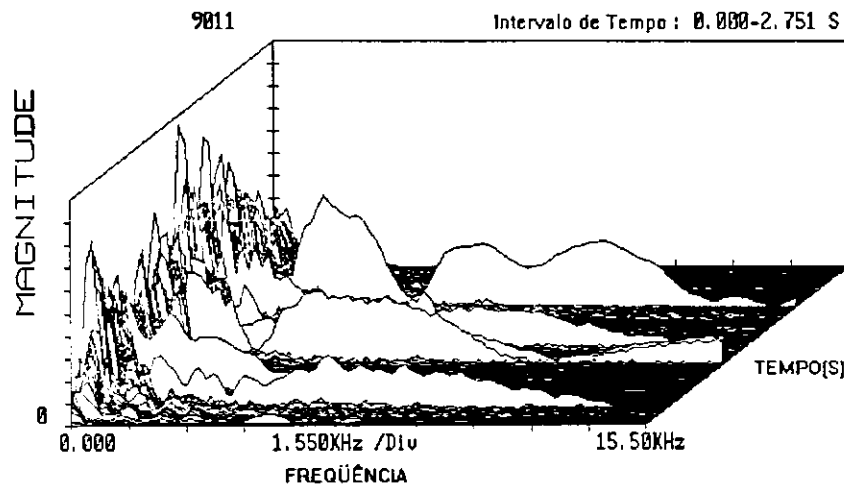


Figura 5.39: Espectro tridimensional do som de cravo degradado

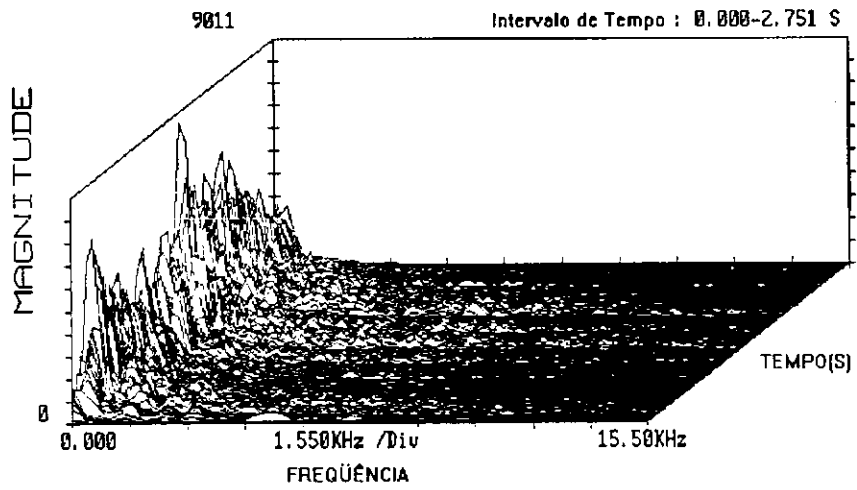


Figura 5.40: Espectro tridimensional do som de cravo filtrado

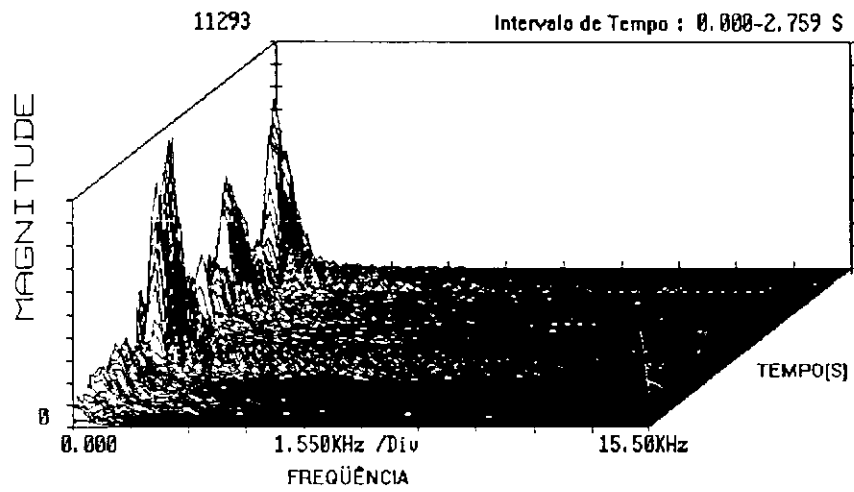


Figura 5.41: Espectro tridimensional do som de flauta digitalizado

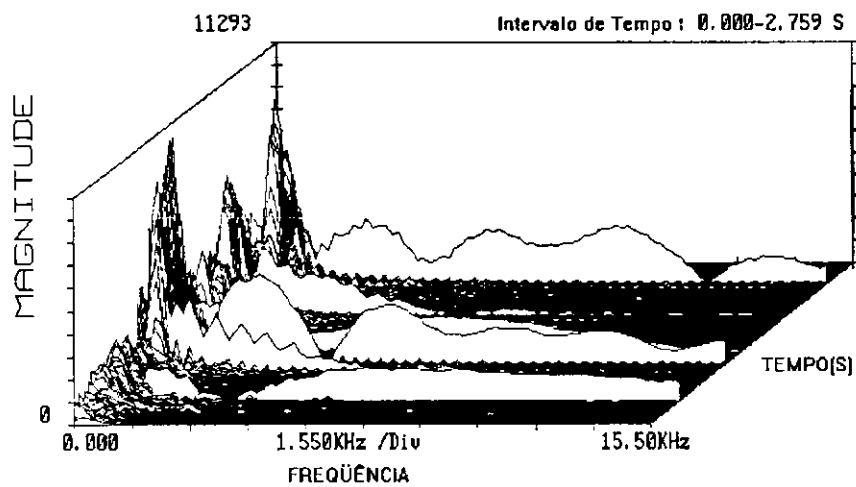


Figura 5.42: Espectro tridimensional do som de flauta degradado

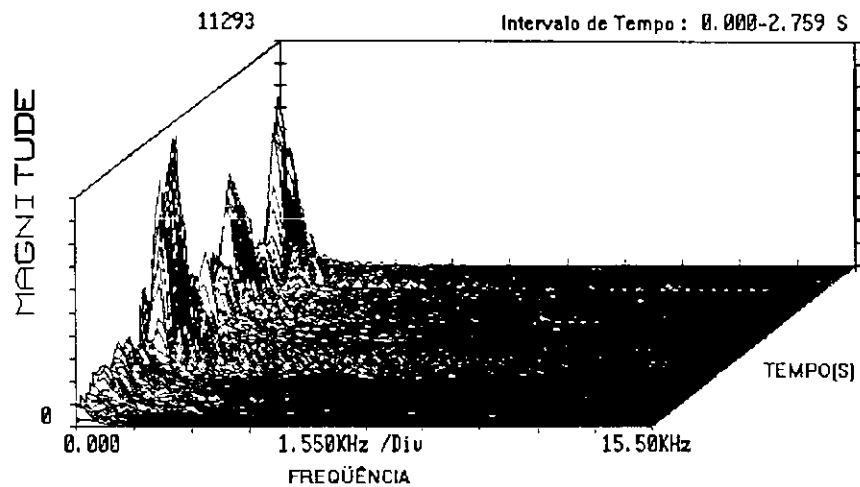


Figura 5.43: Espectro tridimensional do som de flauta filtrado

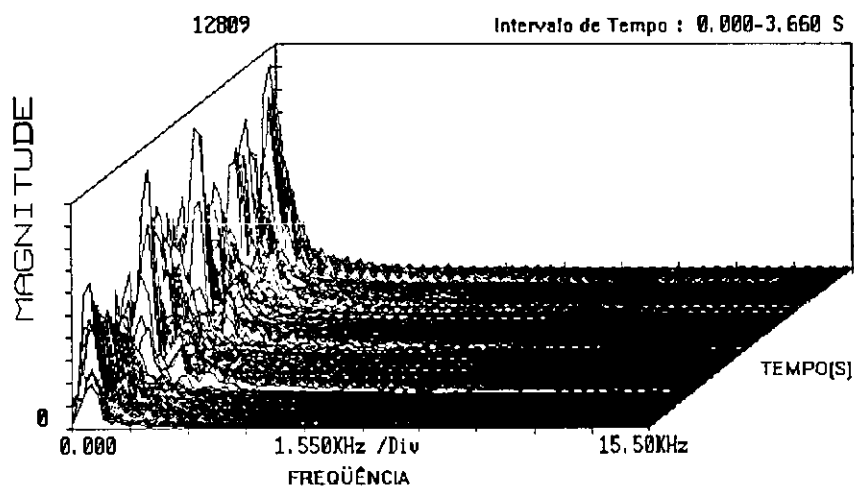


Figura 5.44: Espectro tridimensional do som de piano digitalizado

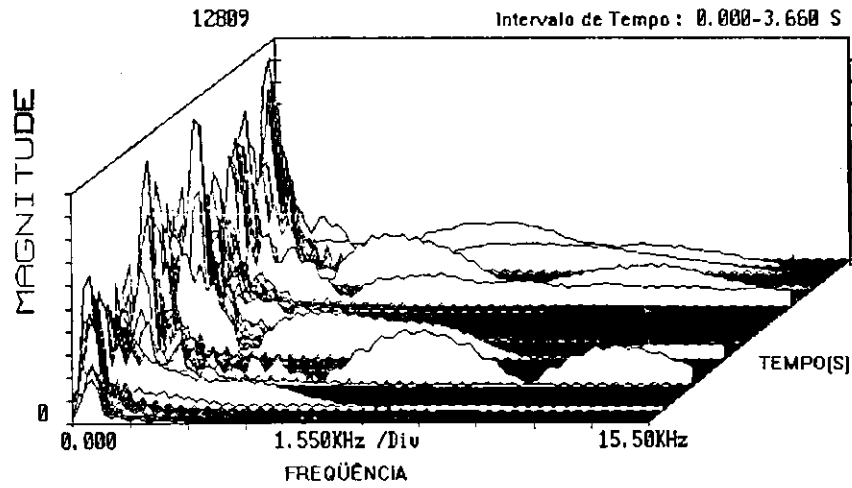


Figura 5.45: Espectro tridimensional do som de piano degradado

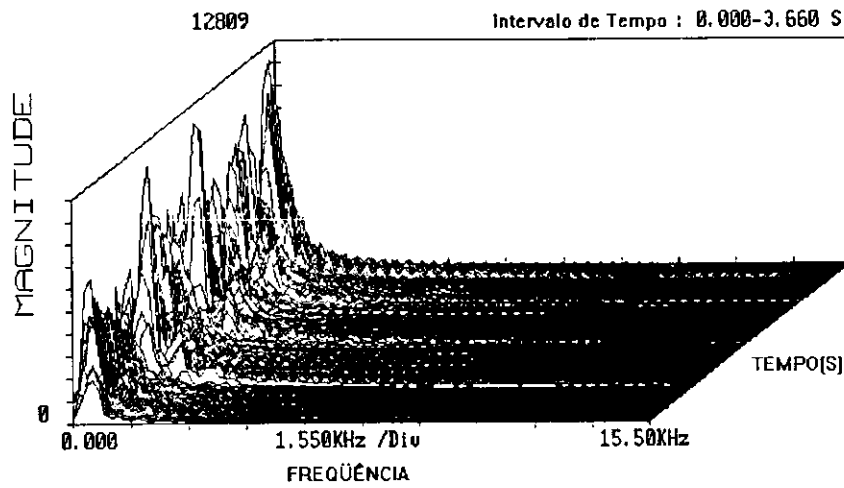


Figura 5.46: Espectro tridimensional do som de piano filtrado

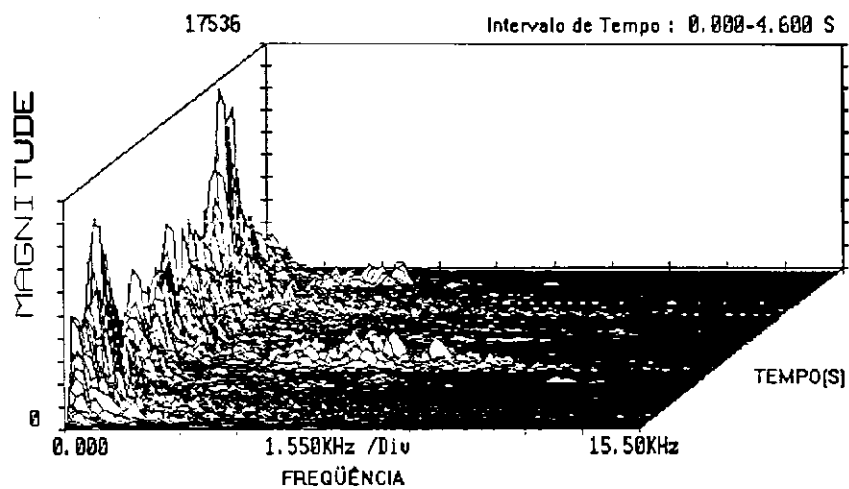


Figura 5.47: Espectro tridimensional do som de violão e voz digitalizado

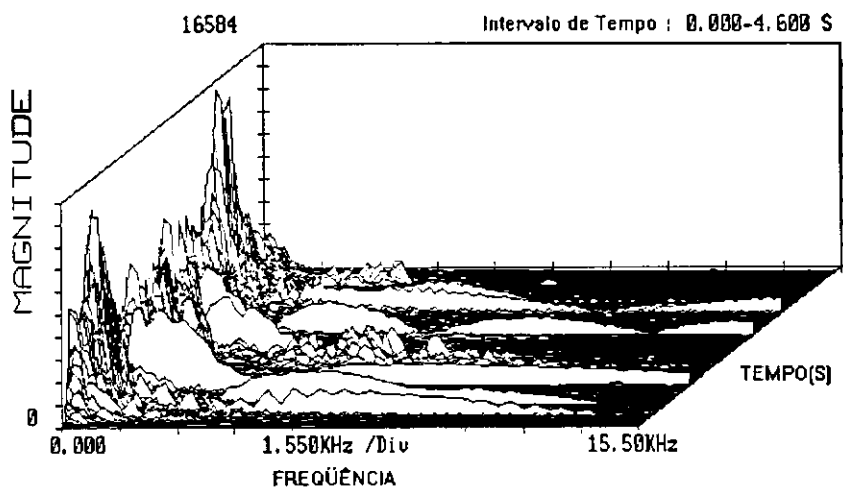


Figura 5.48: Espectro tridimensional do som de violão e voz degradado

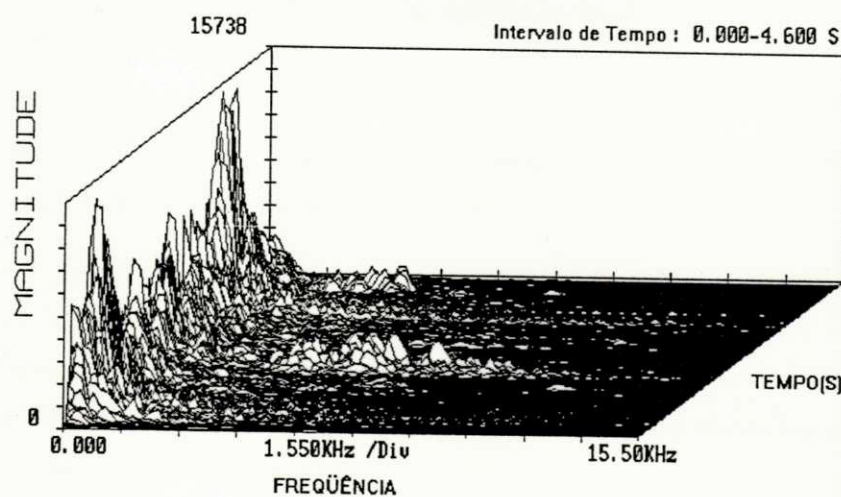


Figura 5.49: Espectro tridimensional do som de violão e voz filtrado

Capítulo 6

Conclusão

Neste trabalho foram implementadas várias técnicas seletivas digitais para a filtragem adaptativa de sinais de áudio degradados por ruído impulsivo, baseadas em sistemas não complementares, ou seja, cujo processamento é realizado apenas na reprodução do sinal de áudio (pós-processamento).

O algoritmo proposto divide-se em duas etapas: detecção e filtragem. A etapa de detecção foi baseada no modelo proposto em [24], com alguns ajustes no que se refere à escolha do limiar de referência, uma vez que este foi determinado a partir das características do sinal de áudio degradado, e à escolha das constantes de ataque e relaxamento do detector de envoltória.

Na etapa de filtragem, foi avaliado o uso da predição linear adaptativa para a estimação dos valores das amostras degradadas do sinal de música, proposta em [1] para sinais de voz, além de uma etapa de pós-filtragem baseada na interpolação polinomial de Lagrange.

As técnicas implementadas para a supressão do ruído impulsivo foram executadas em um tempo relativamente rápido, apresentando-se bastante eficientes, especialmente no que se refere à supressão de impulsos de média e grande energia, que são justamente os mais audíveis. Constatou-se que existe a necessidade de ajustes no limiar de referência para melhorar a detecção dos impulsos de pequena energia.

Cada uma das técnicas foi avaliada através de testes objetivos, como o cálculo da relação sinal-ruído, que acusou ganhos na faixa de 7-10 dB na saída da etapa de pós-filtragem, e através de testes subjetivos, como os testes de escuta informais, que confirmaram a boa qualidade do sinal processado, comparada à qualidade do sinal original sem degradação. De todas as técnicas propostas a que forneceu melhores resultados foi a da predição linear regressiva com a etapa de pós-filtragem baseada na interpolação de Lagrange.

Os resultados obtidos mostraram uma redução significativa de ruídos impulsivos em sinais de áudio, apresentando uma melhoria substancial na qualidade do sinal de áudio em gravações fonográficas antigas. Desta forma, gravações de músicas antigas, com um valor cultural inestimável, poderão ser recuperadas através das técnicas de estimação propostas.

Considerando-se que grande parte do sucesso na supressão da degradação depende da forma como é feita a detecção dos erros, presume-se que melhores resultados podem ser obtidos desde que técnicas mais seletivas de detecção sejam encontradas, especialmente apropriadas para o uso em segmentos de alta energia, onde muitas vezes as transições naturais do sinal são comumente confundidas com o ruído impulsivo.

As principais dificuldades encontradas na etapa de simulação e avaliação das técnicas propostas dizem respeito à qualidade ainda não totalmente satisfatória do equipamento utilizado para a aquisição de dados (SOUND BLASTER 16 ASP), que geralmente introduz ruído térmico no sinal gravado, e à falta de um ambiente com uma acústica adequada. Um outro ponto desfavorável diz respeito ao difícil acesso a uma bibliografia mais atualizada sobre sistemas de supressão de ruídos em sinais de áudio degradados, para eventuais necessidades de comparação do sistema, bem como maiores informações sobre o assunto.

Algumas sugestões para a continuidade deste trabalho ou de trabalhos afins são a utilização de técnicas de análise no domínio da frequência, que permitem uma melhor caracterização espectral do ruído, e a implementação de outras técnicas de detecção utilizando parâmetros baseados nas características espectrais do sinal.

Apêndice A

Formulário para Avaliação Subjetiva

Este formulário padrão será utilizado por cada avaliador na análise subjetiva do sinal degradado por ruído impulsivo (Parte I) e do sinal filtrado (Parte II). As técnicas utilizadas são o Escore de Opinião Média (MOS), para o sinal filtrado, e o Escore de Opinião Média (MOS) Invertido, para o sinal degradado.

Parte I - Teste de degradação

Este teste tem como objetivo a avaliação da audibilidade do ruído impulsivo no sinal degradado. Após a escuta, o avaliador deve atribuir somente uma das seguintes notas:

5	A degradação é muito incômoda
4	A degradação é incômoda
3	A degradação é pouco incômoda
2	A degradação é perceptível mas não incomoda
1	A degradação é imperceptível

Parte II - Teste de Qualidade Absoluta

Este teste tem como objetivo a avaliação do sinal filtrado fornecido por cada algoritmo utilizado. Após a escuta, o avaliador deve atribuir notas, conforme a tabela apresentada.

A) Algoritmo 1

5	A qualidade é excelente
4	A qualidade é boa
3	A qualidade é razoável
2	A qualidade é pobre
1	A qualidade é ruim

B) Algoritmo 2

5	A qualidade é excelente
4	A qualidade é boa
3	A qualidade é razoável
2	A qualidade é pobre
1	A qualidade é ruim

C) Algoritmo 3

5	A qualidade é excelente
4	A qualidade é boa
3	A qualidade é razoável
2	A qualidade é pobre
1	A qualidade é ruim

D) Algoritmo 4

5	A qualidade é excelente
4	A qualidade é boa
3	A qualidade é razoável
2	A qualidade é pobre
1	A qualidade é ruim

Referências

- [1] B. G. Aguiar Neto. *Signalaufbereitung in Digitalen Sprachübertragungssystemen*. Tese de Doutorado, Technischen Universität Berlin, 1987.
- [2] B. G. Aguiar Neto. Melhoria de Voz Degradada por Método baseado em Subtração Espectral Adaptativa. *Anais do VII Simpósio Brasileiro de Telecomunicações*, Setembro 1989.
- [3] B. G. Aguiar Neto. Detecção e Correção de Erros em Transmissão Digital de Voz sem Codificação de Canal. *Anais do IX Simpósio Brasileiro de Telecomunicações*, 1991.
- [4] B. G. Aguiar Neto. Processamento Digital de Sinais de Voz e Imagem. Relatório técnico, Universidade Federal da Paraíba, 1995.
- [5] L. C. Barroso. *Cálculo Numérico*. Editora Harbra, 1987.
- [6] P. J. Bloom. High-Quality Digital Audio in the Entertainment Industry: An Overview of Achievements and Challenges. *IEEE ASSP Magazine*, pp. 2-25, October 1985.
- [7] G. Bristow. *Electronic Speech Synthesis: Techniques, Technology and Applications*. Mc Graw-Hill, 1984.
- [8] R. S. Burwen. A Dynamic Noise Filter. *Journal of the Audio Engineering Society*, 19(2):115-120, February 1971.

- [9] S. L. N. Cunha. Sistemas Multicanal Adaptativo para Supressão de Ruído usando Arranjos de Microfones. Dissertação de Mestrado, Universidade Federal da Paraíba, Setembro 1994.
- [10] S. L. N. Cunha e B. G. Aguiar Neto. Adaptive Multichannel System for Speech Enhancement using Microphone Array. *International Telecommunications Symposium of IEEE*, pp. 152-155, August 1994.
- [11] S. L. N. Cunha e B. G. Aguiar Neto. An Evaluation of an Adaptive Multichannel System for Speech Enhancement with Automatic Phase Alignment. *International Conference on Acoustics, Speech and Signal Processing*, Maio 1995.
- [12] R. M. Dolby. An Audio Noise Reduction System. *Journal of Audio Engineering Society*, 15(4):383-388, October 1967.
- [13] J. L. Boldrini et al. *Álgebra Linear*. Harbra, São Paulo, 3a. edição, 1980.
- [14] G. Stoll G. Theile e M. Link. Low Bit-rate Coding of High-Quality Audio Signals. *EBU Tech. Rev.*, (230):71-94, August 1988.
- [15] F. H. Ives. A Noise Reduction System: Dynamic Spectral Filtering. *Journal of the Audio Engineering Society*, 20(7):558-561, September 1972.
- [16] N. S. Jayant e P. Noll. *Digital Coding of Waveforms*. Prentice Hall, Englewood Cliffs, N. J., 1984.
- [17] A. V. Oppenheim e R. W. Schaffer. *Digital Signal Processing*. Prentice-Hall, New-Jersey, 1975.
- [18] J. B. H. Peek. Communications Aspects of the Compact Disc Digital Audio System. *IEEE Communications Magazine*, 23(2):7-15, February 1985.
- [19] L. R. Rabiner e B. Gold. *Theory and application of Digital Signal Processing*. Prentice-Hall, 1975.
- [20] L. R. Rabiner e R. W. Schafer. *Digital Processing of Speech Signals*. Prentice Hall, 1978.

- [21] H. H. Scott. Dynamic Supression of Phonograph Record Noise. *Electronics*, pp. 92-95, December 1946.
- [22] H. H. Scott. Dynamic Noise Suppressor. *Electronics*, pp. 96-105, December 1947.
- [23] G. Shorter. An Introduction to the Dolby Noise Reduction System. *Wireless World*, pp. 200-205, May 1975.
- [24] K. P. Souza. Restauração de Gravações Fonográficas por meio de Processamento Digital de Sinais. Dissertação de Mestrado, FEE-UNICAMP, 1991.
- [25] M. N. Vieira. Módulo Frontal para um Sistema de Reconhecimento Automático de Voz. Dissertação de Mestrado, FEE - UNICAMP, 1989.
- [26] J. Watkinson. *The Art of Digital Audio*. Focal Press, 1988.
- [27] B. et al Widrow. Adaptive Noise Cancelling: Principles and Applications. *Proceedings of the IEEE*, 63(12):1692-1716, December 1975.
- [28] M. Yamazaki e I. Masuda. A New Automatic Noise Reduction System (ANRS). *Journal of the Audio Engeneering Society*, 21(6):445-449, July/August 1973.