

Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação

Equalização e Identificação Adaptativas Utilizando Marca d'Água como Sinal de Supervisão

Autor: Mário Uliani Neto

Orientador: Prof. Dr. João Marcos Travassos Romano

Co-orientador: Dr. Leandro de Campos Teixeira Gomes

Dissertação de Mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: **Telecomunicações e Telemática**.
Aprovação em 18/01/2008

Banca Examinadora

Prof. Dr. João Marcos Travassos Romano FEEC/Unicamp
Prof. Dr. Amauri Lopes FEEC/Unicamp
Prof. Dr. Fernando Oscar Runstein Fundação CPqD
Prof. Dr. Jozué Vieira Filho UNESP/Ilha Solteira

Campinas, SP
2008

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

UL3e Uliani Neto, Mário
Equalização e identificação adaptativas utilizando marca
d'água como sinal de supervisão / Mário Uliani Neto.
– Campinas, SP: [s.n.], 2008.

Orientador: João Marcos Travassos Romano.
Dissertação (mestrado) - Universidade Estadual de
Campinas, Faculdade de Engenharia Elétrica e de
Computação.

1. Marca d'água digital. 2. Processamento de sinal
adaptativo. 3. Filtros adaptativos. 4. Equalizadores
(Eletrônica). 5. Identificação de sistemas. 6. Sincronização.
I. Romano, João Marcos Travassos. II. Universidade
Estadual de Campinas. Faculdade de Engenharia Elétrica e
de Computação. III. Título.

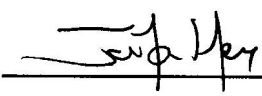
Título em Inglês:	Adaptive equalization and identification using watermark as supervision signal
Palavras-chave em Inglês:	Adaptive filter, Adaptive identification, Adaptive equalization, Digital watermarking, Superimposed training, Resynchronization
Área de concentração:	Telecomunicações e Telemática
Titulação:	Mestre em Engenharia Elétrica
Banca Examinadora:	Amauri Lopes, Fernando Oscar Runstein, Jozué Vieira Filho
Data da defesa:	18/01/2008
Programa de Pós-Graduação:	Engenharia Elétrica

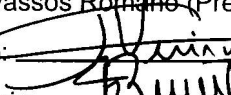
COMISSÃO JULGADORA - TESE DE MESTRADO

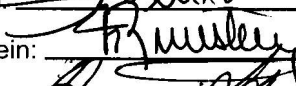
Candidato: Mário Uliani Neto

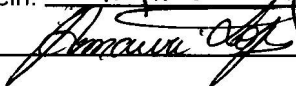
Data da Defesa: 18 de janeiro de 2008

Título da Tese: "Equalização e Identificação Adaptativas Utilizando Marca D'Água como Sinal de Supervisão"

Prof. Dr. João Marcos Travassos Romano (Presidente): 

Prof. Dr. Jozué Vieira Filho: 

Dr. Fernando Oscar Runstein: 

Prof. Dr. Amauri Lopes: 

Resumo

O objetivo deste trabalho é investigar o uso de uma marca d'água digital como sinal de referência na filtragem adaptativa supervisionada, aplicada aos problemas de equalização e identificação de canais. Ao contrário de métodos mais tradicionais, nos quais a comunicação é periodicamente interrompida para a transmissão de seqüências de treinamento, uma marca d'água é transmitida ininterruptamente, juntamente com o sinal de informação. Através da comparação entre o sinal recebido processado pelo filtro equalizador e a marca d'água original, ou entre o sinal recebido e a marca d'água processada pelo filtro identificador, os coeficientes deste filtro são continuamente adaptados para estimar e rastrear as características do canal ao longo do tempo. Tanto a transmissão de informação útil como a adaptação do filtro equalizador/identificador nunca são interrompidas.

Para que o processo de detecção de uma marca d'água digital seja eficiente, é necessária a sincronização da marca no detector. Efeitos de dessincronização podem prejudicar a detecção, potencialmente inviabilizando a extração adequada da informação contida na marca d'água. Esta dissertação apresenta dois métodos de resincronização para sistemas de filtragem empregando marca d'água. Ambos os métodos são baseados no uso de seqüências de treinamento embutidas na marca d'água e reverterem os efeitos de uma ampla classe de ataques de dessincronização.

No caso de sinais que apresentem uma interpretação sensorial (e.g. áudio, imagens, vídeo), a adição da marca d'água não deve causar distorção perceptível no sinal original. Propomos o uso de um modelo psicoacústico em conjunto com um algoritmo de conformação espectral para demonstrar a viabilidade do método quando aplicado a sinais de áudio. Resultados de simulações computacionais são apresentados para evidenciar o desempenho das propostas frente a técnicas tradicionais de filtragem adaptativa supervisionada.

Palavras-chave: Equalização adaptativa, Filtragem adaptativa, identificação adaptativa, marca d'água digital, resincronização, *superimposed training*.

Abstract

The objective of this work is to investigate the use of a digital watermark as a reference signal in supervised adaptive filtering, applied to the channel equalization and identification problems. Contrary to traditional adaptive methods, in which communication is periodically interrupted for the transmission of training sequences, a watermark is transmitted uninterruptedly, along with the information signal. By comparing the received signal and the watermark processed by the identification filter, or the received signal processed by the equalization filter and the original watermark, the coefficients of this filter are continuously adapted to estimate and monitor channel characteristics through time. Both the transmission of useful information and the adaptation of the identification/equalization filter are never interrupted.

For the watermark detection process to be efficient, synchronization of the watermark in the detector is required. Desynchronization effects can reduce detection performance, potentially making impracticable the appropriate extraction of the information contained in the watermark. This dissertation presents two resynchronization methods for filtering systems using a watermark. Both methods are based on training sequences that are embedded in the watermark, and can reverse the effects of a large class of desynchronization attacks.

The embedding of a watermark into signals with sensorial interpretation (e.g. audio, image, video) should not introduce noticeable distortion in the original signal. To achieve this when applying our method to audio signals, we propose the use of a psychoacoustic model with a spectral shaping algorithm. Simulation results are presented to illustrate the performance of the proposed method when compared to traditional supervised adaptive filtering techniques.

Keywords: Adaptive filter, adaptive identification, adaptive equalization, digital watermarking, superimposed training, resynchronization.

À minha mãe Zélia, e irmãos Felipe e Renata

A Suellen pelo amor e ao meu filho Lucas pela alegria

A José Pereira de Miranda Júnior, in memoriam

Agradecimentos

Agradeço

À minha mãe Zélia, e aos meus irmãos, Felipe e Renata, por todo apoio e incentivo em todos os momentos e sentidos.

À minha esposa, Suellen, ao meu filho Lucas, por todo amor e compreensão em todos os momentos.

Ao meu co-orientador, Leandro de C. T. Gomes, pela confiança e apoio, pelas discussões enriquecedoras e sugestões durante todo o desenvolvimento deste trabalho.

Ao meu orientador, Prof. João Marcos T. Romano, pela orientação, apoio e confiança.

Aos meus amigos do Laboratório de Processamento de Sinais para Comunicações (DSPCom).

Aos amigos Prof. Romis R. de F. Attux e Ricardo Suyama, pelas valiosas discussões e sugestões.

Ao amigo Prof. Fernando Oscar Runstein, pelo incentivo constante e apoio técnico.

Aos amigos do CPqD, Edson Nagle, Flávio Simões, Clodocir, José Pedro, Norberto, Ralph Robert e Roberto Petry.

Às minhas avós, Neguita e Santana.

Aos meus tios, José Francisco Miranda e Ivani, pelo apoio.

À minha prima e madrinha Mariane, pelo incentivo.

Aos meus familiares.

Aos meus amigos Renato Gennari e Yara.

Aos amigos da turma de graduação, pelo companheirismo e incentivo.

A todos os meus amigos.

Aos membros da banca.

A todos que puderam comparecer à defesa de tese.

Índice

Lista de Figuras	xvii
Lista de Tabelas	xxi
Lista de Abreviações e Acrônimos	xxiii
Trabalhos Publicados Durante o Mestrado	xxv
1 Introdução	1
2 Filtragem Adaptativa em Equalização e Identificação	5
2.1 Introdução	5
2.2 Estruturas de Filtragem	6
2.2.1 Estruturas Lineares	6
2.2.2 Estruturas Não-lineares	9
2.3 Filtragem Adaptativa	9
2.3.1 Critério de Wiener: Filtragem a Partir de uma Medida de Erro	10
2.3.2 Algoritmo do Gradiente Determinístico	13
2.3.3 Algoritmo <i>Least Mean Square</i>	14
2.3.4 Algoritmo <i>Recursive Least Squares</i>	15
2.4 Equalização	16
2.4.1 Equalização Supervisionada	17
2.5 Identificação	18
2.5.1 Identificação Supervisionada	19
2.6 Introdução aos Métodos Não-Supervisionados	19
2.6.1 O Critério de Decisão Direta	21
2.6.2 O Critério de Godard	22
2.6.3 O Critério do Módulo Constante	23
2.6.4 O Critério de Shalvi-Weinstein	23
2.7 Discussão	24
3 Revisão de Marca d'Água Digital	27
3.1 Introdução	27
3.2 Aplicações	29
3.3 Classificação dos Sistemas de Marca d'Água	32

3.4	Estrutura Geral de Sistemas de Marca d'Água	33
3.4.1	Elementos Básicos do Sistema	33
3.4.2	Marca d'Água como um Sistema de Comunicação	35
3.4.3	Marca d'Água em Áudio	36
3.5	Métodos para Inserção de Marca d'Água	37
3.5.1	Espalhamento Espectral	38
3.5.2	QIM (<i>Quantization Index Modulation</i>)	41
3.5.3	Método da Réplica (<i>Replica Method</i>)	42
3.5.4	Método de Auto-Marcação (<i>Self-Marking Method</i>)	42
3.6	Ataques	43
3.7	Sincronismo	44
3.8	Discussão	45
4	Equalização Supervisionada Utilizando Marca d'Água	47
4.1	Introdução	47
4.2	Equalização Supervisionada Baseada em Marca d'Água: Sistema Proposto	48
4.2.1	Equalizador de Wiener	48
4.2.2	Descrição do Algoritmo Proposto	49
4.3	Equalização de Sinais Descorrelacionados	52
4.3.1	Simulações e Resultados	53
4.4	Equalização de Sinais Correlacionados	70
4.4.1	Simulações e Resultados	72
4.5	Síntese e Conclusões	74
5	Identificação Supervisionada Utilizando Marca d'Água	77
5.1	Introdução	77
5.2	Identificação Supervisionada Baseada em Marca d'Água: Sistema Proposto	78
5.2.1	Identificação de Wiener	78
5.2.2	Descrição do Algoritmo Proposto	79
5.3	Simulações e Resultados	81
5.3.1	Análise de Canais Variantes no Tempo	91
5.4	Síntese e Conclusões	95
6	Sincronismo para Sistemas Adaptativos Utilizando Marca d'Água	97
6.1	Introdução	97
6.2	Dessincronização da Marca d'Água	97
6.3	Sincronismo para os Sistemas Adaptativos Propostos	98
6.4	Métodos de Sincronização	100
6.4.1	Seqüência de Treinamento: Idéia Básica	100
6.4.2	Seqüência de Blocos	101
6.4.3	Seqüência de Símbolos de Codebook	103
6.5	Algoritmos de Sincronização	104

6.5.1	Sincronização por Média de Janelas de Análise	104
6.5.2	Sincronização por Programação Dinâmica	105
6.6	Simulações e Resultados	108
6.6.1	Condições Experimentais	108
6.6.2	Resultados Experimentais	109
6.7	Conclusões	118
7	Equalização e Identificação para Sinais de Áudio Utilizando Marca d'Água	121
7.1	Introdução	121
7.2	Marca d'Água em Áudio Digital	121
7.2.1	Modelos Psicoacústicos	122
7.3	Equalização e Identificação de Sinais de Áudio Utilizando Marca d'Água	125
7.4	Simulações e Resultados	127
7.4.1	Análise do Algoritmo de Marca d'Água para Áudio	128
7.4.2	Sistema de Identificação	136
7.4.3	Sistema de Equalização	137
7.5	Conclusões	138
8	Conclusões e Perspectivas	139
	Referências Bibliográficas	142
A	Fenômenos Auditivos e Modelos Perceptivos	151
A.1	Fenômenos Auditivos	151
A.1.1	Fisiologia do Ouvido Humano	151
A.1.2	Limiar Absoluto de Audibilidade em Silêncio	154
A.1.3	Bandas Críticas	154
A.1.4	Mascaramento	155
A.2	Modelos Perceptivos	160
A.2.1	Modelo de Zwicker para o Cálculo da Sonoridade Percebida	160
A.2.2	Modelos Psicoacústicos em Codificação de Áudio	161

Lista de Figuras

2.1	Esquema de um filtro.	6
2.2	Filtro linear transversal.	7
2.3	Filtro IIR.	8
2.4	Filtro de Wiener.	10
2.5	Função Custo de Wiener.	13
2.6	Modelo simplificado de um processo de equalização.	17
2.7	Esquema clássico de equalização supervisionada.	18
2.8	Esquema clássico de identificação supervisionada.	19
2.9	Esquema clássico de treinamento com o critério DD.	21
2.10	Curvas de nível para função custo DD.	22
2.11	Curvas de nível para função custo CM.	24
3.1	Esquema genérico de um sistema de marca d'água.	34
3.2	Marca d'água como um sistema de comunicação.	35
3.3	Esquema de marca d'água digital em áudio.	36
3.4	Transporte de informação através de um canal acústico.	38
3.5	Esquema de inserção de marca d'água utilizando espalhamento espectral.	39
3.6	Pré-processamento para detecção em um esquema de marca d'água utilizando espalhamento espectral.	41
3.7	Esquema de quantização QIM.	42
4.1	Esquema clássico de equalização supervisionada.	49
4.2	Esquema de equalização utilizando marca d'água.	50
4.3	Função custo e respectivas curvas de nível para o esquema de equalização utilizando marca d'água.	54
4.4	Evolução temporal do algoritmo LMS (SWR = 0 dB, $\mu = 0,0002$).	55
4.5	Evolução temporal do algoritmo LMS para diferentes passos μ (SWR = 10 dB): 4.5(a) $\mu = 0,0002$; 4.5(b) $\mu = 0,0005$	56
4.6	Evolução temporal do algoritmo LMS (SWR = 20 dB, $\mu = 0,0002$).	57
4.7	Evolução temporal do algoritmo RLS (SWR = 0 dB, $\lambda = 1$).	58
4.8	Evolução temporal do algoritmo RLS para diferentes valores de λ (SWR = 10 dB): 4.8(a) $\lambda = 1$; 4.8(b) $\lambda = 0,998$	59
4.9	Evolução temporal do algoritmo RLS (SWR = 20 dB, $\lambda = 1$).	60

4.10	Erro quadrático médio do algoritmo LMS em função da relação de potência sinal-marca d'água.	61
4.11	Convergência do erro quadrático médio do algoritmo LMS para SWR = 20 dB. . . .	62
4.12	Erro quadrático médio do algoritmo RLS em função da relação de potência sinal-marca d'água.	63
4.13	Convergência do erro quadrático médio do algoritmo RLS para SWR = 20 dB.	63
4.14	Coefficientes baseados em funções degrau.	64
4.15	Coefficientes baseados em funções degrau e rampa.	65
4.16	Coefficientes do equalizador obtidos através do algoritmo LMS para um canal variante no tempo baseado em degrau, tendo a solução de Wiener como referência: 4.16(a) w_{eq1} ; 4.16(b) w_{eq2} ; 4.16(c) w_{eq3}	66
4.17	Coefficientes do equalizador obtidos através do algoritmo LMS para um canal variante no tempo baseado em degrau-rampa, tendo a solução de Wiener como referência: 4.17(a) w_{eq1} ; 4.17(b) w_{eq2} ; 4.17(c) w_{eq3}	67
4.18	Coefficientes do equalizador obtidos através do algoritmo RLS para um canal variante no tempo baseado em degrau, tendo a solução de Wiener como referência: 4.18(a) w_{eq1} ; 4.18(b) w_{eq2} ; 4.18(c) w_{eq3}	68
4.19	Coefficientes do equalizador obtidos através do algoritmo RLS para um canal variante no tempo baseado em degrau-rampa, tendo a solução de Wiener como referência: 4.19(a) w_{eq1} ; 4.19(b) w_{eq2} ; 4.19(c) w_{eq3}	69
4.20	Esquema de equalização utilizando marca d'água para sinais correlacionados.	71
4.21	Filtro de erro de predição.	72
4.22	Erro quadrático médio de equalização com o algoritmo RLS em função da relação sinal-marca d'água (sinal de guitarra).	73
4.23	Erro quadrático médio de equalização com o algoritmo RLS em função da relação sinal-marca d'água (sinal de trompete).	73
4.24	Erro quadrático médio de equalização com o algoritmo RLS em função da relação sinal-marca d'água (sinal de rock).	74
5.1	Esquema clássico de identificação supervisionada.	78
5.2	Esquema de identificação utilizando marca d'água.	79
5.3	Função custo e curvas de nível para o esquema de identificação utilizando marca d'água.	82
5.4	Evolução temporal do algoritmo LMS (SWR = 0 dB, $\mu = 0,001$).	83
5.5	Evolução temporal do algoritmo LMS para diferentes passos μ (SWR = 10 dB): 5.5(a) $\mu = 0,001$; 5.5(b) $\mu = 0,005$	84
5.6	Evolução temporal do algoritmo LMS (SWR = 20 dB, $\mu = 0,001$).	85
5.7	Evolução temporal do algoritmo RLS (SWR = 0 dB, $\lambda = 1$).	86
5.8	Evolução temporal do algoritmo RLS para diferentes valores de λ (SWR = 10 dB): 5.8(a) $\lambda = 1$; 5.8(b) $\lambda = 0,999$	87
5.9	Evolução temporal do algoritmo RLS (SWR = 20 dB, $\lambda = 1$).	88
5.10	Evolução temporal do algoritmo RLS (SWR = 23 dB, $\lambda = 1$).	88
5.11	Evolução temporal do algoritmo RLS (SWR = 30 dB, $\lambda = 1$).	89
5.12	Coefficiente w_{id0} do modelo de identificação para um canal variante baseado em degrau.	92
5.13	Coefficiente w_{id1} do modelo de identificação para um canal variante baseado em degrau.	92

5.14	Coeficiente w_{id2} do modelo de identificação para um canal variante baseado em degrau.	93
5.15	Coeficiente w_{id0} do modelo de identificação para um canal variante baseado em degrau-rampa.	93
5.16	Coeficiente w_{id1} do modelo de identificação para um canal variante baseado em degrau-rampa.	94
5.17	Coeficiente w_{id2} do modelo de identificação para um canal variante baseado em degrau-rampa.	94
6.1	Esquema de equalização utilizando marca d'água com sincronização.	99
6.2	Esquema de identificação utilizando marca d'água com sincronismo.	99
6.3	Esquema do detector da marca d'água.	100
6.4	Estrutura de dados para sincronização. O bloco cinza corresponde ao bloco formado por símbolos pseudo-aleatórios; a sucessão de blocos corresponde à marca d'água. . .	101
6.5	Cálculo de correlação para sincronismo.	102
6.6	Deslocamentos de janela deslizante correspondente à máxima correlação para o bloco da marca d'água. Para ilustrar a perda de sincronismo, 5 amostras foram removidas a cada 200 janelas.	103
6.7	Possível caminho ótimo na matriz B	105
6.8	Desempenho dos algoritmos com sinal digital de telecomunicações (SWR = 10 dB) e sem a presença de distorções: 6.8(a) picos espúrios de correlação; 6.8(b) sincronização por média de janelas de análise; 6.8(c) sincronização por programação dinâmica.	111
6.9	Desempenho dos algoritmos com sinal digital de telecomunicações (SWR = 20 dB) e sem a presença de distorções: 6.9(a) picos espúrios de correlação; 6.9(b) sincronização por média de janelas de análise; 6.9(c) sincronização por programação dinâmica.	112
6.10	Desempenho dos algoritmos com sinal de áudio (SWR = 10 dB) e sem a presença de distorções: 6.10(a) picos espúrios de correlação; 6.10(b) sincronização por média de janelas de análise; 6.10(c) sincronização por programação dinâmica.	113
6.11	Desempenho dos algoritmos com sinal de áudio (SWR = 20 dB) e sem a presença de distorções: 6.11(a) picos espúrios de correlação; 6.11(b) sincronização por média de janelas de análise; 6.11(c) sincronização por programação dinâmica.	114
6.12	Desempenho dos algoritmos de sincronismo com sinal digital de telecomunicações (SWR = 20 dB) e presença de sistema hostil: 6.12(a) picos espúrios de correlação; 6.12(b) média de janelas; 6.12(c) programação dinâmica utilizando bloco; 6.12(d) programação dinâmica utilizando codebook.	116
6.13	Desempenho dos algoritmos de sincronismo com sinal de fala (SWR = 20 dB) e presença de sistema hostil: 6.13(a) picos espúrios de correlação; 6.12(b) média de janelas; 6.12(c) programação dinâmica utilizando bloco; 6.12(d) programação dinâmica utilizando codebook.	117
6.14	Sistema completo de identificação utilizando marca d'água.	118
6.15	Sistema completo de equalização utilizando marca d'água.	118

7.1	Densidades espectrais de potência dos sinais mascarador ($U(f)$, linha contínua) e mascarado ($V(f)$, linha tracejada), e limiar de mascaramento ($M_u(f)$). Nas frequências altas, observa-se que o limiar de mascaramento e o sinal mascarado ultrapassam a PSD do sinal mascarador devido à baixa sensibilidade do ouvido humano nesta faixa de frequências.	123
7.2	Exemplo de conformação espectral da marca d'água. A marca audível (linha tracejada) é transformada na marca inaudível (linha contínua) com base no limiar de mascaramento.	124
7.3	Esquema de inserção de marca d'água em áudio digital.	126
7.4	Esquema de equalização utilizando marca d'água em áudio.	127
7.5	Esquema de identificação utilizando marca d'água em áudio.	128
7.6	Tela apresentando as variáveis de saída (MOVs) e o resultado global da avaliação (ODG) do PEAQ básico.	130
7.7	Escala de degradação de 5 graus ITU-R versus ODG PEAQ.	131
7.8	Medida objetiva de qualidade: 7.8(a) <i>dire</i> s; 7.8(b) <i>svega</i> ; 7.8(c) <i>trump</i> ; 7.8(d) <i>symph</i>	133
7.9	Medida objetiva média de qualidade versus SWR usando modelo psicoacústico.	135
7.10	Medida objetiva média de qualidade versus SWR para marca d'água branca.	135
A.1	Corte longitudinal do ouvido, com suas principais estruturas.	152
A.2	Limiar absoluto de audibilidade em silêncio.	154
A.3	Exemplos de sinais tonal e não tonal; ambos os exemplos correspondem a sinais de voz (fones “ó” e “s”, respectivamente).	158
A.4	Padrão de mascaramento causado por ruído de banda estreita.	158
A.5	Representação do mascaramento simultâneo causado por tons a 1 kHz.	159

Lista de Tabelas

5.1	Desvio do RLS para um sinal i.i.d.	90
5.2	Desvio do RLS para um sinal de áudio.	90
7.1	Material de áudio utilizado nas avaliações de qualidade.	129
7.2	Escala ODG com cinco níveis de degradação de áudio.	132
7.3	Desvio dos coeficientes do filtro identificador com uso de modelo psicoacústico.	136
7.4	Desvio dos coeficientes do filtro identificador com marca d'água branca.	136
7.5	Erro quadrático médio do sinal equalizado com uso de modelo psicoacústico.	137
7.6	Erro quadrático médio do sinal equalizado com marca d'água branca.	137
A.1	Bandas Críticas	156

Lista de Abreviações e Acrônimos

AAC - *Advanced Audio Coding*

BPSK - *Binary Phase Shift Keying*

CM - Módulo Constante (*Constant Modulus*)

CMA - Algoritmo do Módulo Constante (*Constant Modulus Algorithm*)

DCT - Transformada Cosseno Discreta (*Discrete Cosine Transform*)

DD - Decisão Direta (*Decision-Directed*)

DFT - Transformada Discreta de Fourier (*Discrete Fourier Transform*)

DTS - *Digital Theater Systems*

DWT - Transformada Discreta Wavelet (*Discrete Wavelet Transform*)

EQM - Erro Quadrático Médio

FIR - Resposta Finita ao Impulso (*Finite Impulse Response*)

HOS - Estatísticas de Ordem Superior (*Higher Order Statistics*)

i.i.d. - independente e identicamente distribuído

ITU - *International Telecommunication Union*

IIR - Resposta Infinita ao Impulso (*Infinite Impulse Response*)

IIS - Interferência intersimbólica

LMS - *Least Mean Square*

MIMO - *Multiple-Input Multiple-Output*

MLS - Sequência de Comprimento Máximo (*Maximum Length Sequence*)

MPEG - *Moving Picture Experts Group*

MOV - *Model Output Variables*

ODG - *Objective Difference Grade*

OFDM - *Orthogonal Frequency Division Multiplexing*

PAC - *Perceptual Audio Codec*

PAM - *Modulação por Amplitude de Pulso (Pulse Amplitude Modulation)*

PEAQ - *Perceptual Evaluation of Audio Quality*

PESQ - *Perceptual Evaluation of Speech Quality*

PSD - *Densidade Espectral de Potência (Power Spectral Density)*

PSK - *Phase Shift Keying*

QIM - *Quantization Index Modulation*

RLS - *Recursive Least Squares*

SDG - *Subjective Difference Grade*

SW - *Shalvi-Weinstein*

SWR - *Relação Sinal-Marca d'Água (Signal-to-Watermark Ratio)*

Trabalhos Publicados Durante o Mestrado

Os trabalhos listados a seguir estão ligados diretamente ao tema desta dissertação.

Artigos Publicados em Anais de Congressos Internacionais

Mário Uliani Neto, Leandro de C. T. Gomes, João Marcos T. Romano, and Madeleine Bonnet. “Égalisation et synchronisation utilisant un tatouage comme signal de référence”. *XXI colloque GRETSI*, 11-14 de setembro de 2007, Troyes, França.

Mário Uliani Neto, Leandro de C. T. Gomes, João Marcos T. Romano, and Madeleine Bonnet. “Adaptive equalization based on watermarking”. *VI International Telecommunications Symposium (ITS 2006)*, 3-6 de setembro de 2006, Fortaleza, Brasil.

Artigos Publicados em Anais de Congressos Nacionais

Mário Uliani Neto, Leandro de C. T. Gomes e João Marcos T. Romano. “Identificação Adaptativa Supervisionada Utilizando Marca d’Água Digital”. *XXV Simpósio Brasileiro de Telecomunicações (SBrT 2007)*, 3-6 de setembro de 2007, Recife, Brasil.

Mário Uliani Neto, Leandro de C. T. Gomes e João Marcos T. Romano. “Identificação e Equalização de Canais FIR Variantes no Tempo Utilizando Marca d’Água”. *XXV Simpósio Brasileiro de Telecomunicações (SBrT 2007)*, 3-6 de setembro de 2007, Recife, Brasil.

Mário Uliani Neto, Leandro de C. T. Gomes e João Marcos T. Romano. “Métodos de Ressincronização para Sistemas de Equalização e Identificação Utilizando Marca d’Água Digital”. *XXV Simpósio Brasileiro de Telecomunicações (SBrT 2007)*, 3-6 de setembro de 2007, Recife, Brasil.

Capítulo 1

Introdução

As primeiras técnicas envolvendo marca d'água (*watermarking* em inglês) foram propostas em meados dos anos 90 como uma solução potencial ao problema da pirataria de documentos multimídia, em particular áudio, imagens e vídeo. Trata-se da inserção de uma marca, a marca d'água, no sinal que se deseja proteger (dito *original* ou *hospedeiro*). Esta marca é um sinal de mesma natureza que o sinal hospedeiro (e.g. um sinal de áudio, uma imagem ou um sinal de vídeo) e que é adicionado a este último. Embora o documento multimídia seja modificado pela adição da marca d'água, sua qualidade sonora e/ou visual deve ser preservada, o que pode ser atingido por meio da exploração de fenômenos psicoacústicos e/ou psicovisuais. O sinal de marca d'água contém dados que podem ser extraídos por um detector. Esses dados podem, por exemplo, informar a equipamentos compatíveis quais são os direitos do usuário sobre o sinal. Em especial, a marca d'água pode indicar se é ou não permitida a duplicação do sinal hospedeiro.

Um sistema baseado em marca d'água pode ser comparado a um sistema de comunicação clássico: o objetivo é transmitir informação de um emissor a um receptor através de um canal. Nesta analogia, o sinal hospedeiro é visto como ruído, enquanto o sinal de marca d'água corresponde ao sinal útil. Do ponto de vista do usuário, no entanto, as relações se invertem: o sinal hospedeiro é o sinal útil e a marca d'água é encarada como uma perturbação.

O domínio de aplicação da marca d'água é hoje muito mais amplo do que a proposta inicial de proteção contra a pirataria, incluindo a inserção de informações complementares em uma gravação e de elementos indicativos de sua origem. Embora na literatura o estudo das técnicas de marca d'água tenha se limitado essencialmente a sinais com uma interpretação perceptiva imediata (tais como sinais de áudio e imagens), não há uma razão fundamental para esta limitação. De fato, qualquer sinal passível de processamento pode ser marcado.

O objetivo deste trabalho é o estudo da aplicação da marca d'água como ferramenta para a melhoria do desempenho de algoritmos de filtragem adaptativa supervisionada, aplicados aos problemas

de equalização e identificação de canais.

Técnicas tradicionais de equalização ou identificação supervisionada baseiam-se geralmente na transmissão periódica de sinais conhecidos (ditos *seqüências de treinamento*). Durante os períodos de treinamento, a transmissão do sinal que transporta informação útil (doravante denominado *sinal de informação*) é interrompida. Em aplicações de tempo real, tais interrupções podem ser inaceitáveis; além disso, as mudanças nas características do canal somente são levadas em conta pelo filtro de equalização ou identificação a partir do período de treinamento subsequente. Uma alternativa ao uso de seqüências de treinamento são os métodos não-supervisionados ou cegos, baseados somente no sinal recebido e em características estatísticas do sinal inicialmente transmitido. Devido à inexistência da seqüência de treinamento, os métodos não-supervisionados utilizam como sinal de referência uma estimativa não-linear do sinal transmitido, introduzindo alguns problemas que não estão presentes no caso supervisionado, como a existência de soluções sub-ótimas (mínimos locais da função custo a ser minimizada).

A abordagem deste trabalho propõe o uso de uma marca d'água em lugar de uma seqüência de treinamento. O sinal de marca d'água é continuamente superposto ao sinal de informação por meio da técnica de espalhamento espectral, que consiste no uso como marca d'água de um sinal pseudo-aleatório uniformemente distribuído em frequência e ortogonal ao sinal de informação, ocupando uma pequena parcela da potência disponível para o sinal transmitido. O sinal resultante desta superposição passa através de um canal e atinge o receptor. Por meio da comparação entre o sinal recebido e a marca d'água original (conhecida no receptor), os coeficientes do filtro são continuamente atualizados. A diferença entre os sinais produz um sinal de erro que é realimentado no próprio filtro. A minimização deste sinal de erro leva à solução ótima de Wiener.

Conforme veremos nos capítulos subsequentes, a principal vantagem desta abordagem em relação a técnicas convencionais supervisionadas é o fato de que o sinal de informação nunca é interrompido. Além disso, a atualização dos coeficientes do filtro adaptativo é contínua, de modo que alterações nas características do sistema têm efeito imediato sobre a filtragem. Em relação às técnicas não-supervisionadas, a vantagem está no fato de que a técnica utilizando marca d'água leva a uma função custo unimodal e, na maioria dos casos, apresenta menor complexidade.

A organização bem como a estrutura e as contribuições desta dissertação estão apresentadas da seguinte maneira:

Capítulo 2 - Filtragem Adaptativa em Equalização e Identificação

Neste capítulo são apresentados os principais conceitos e algoritmos relacionados à filtragem adaptativa, bem como as idéias essenciais relativas à equalização e identificação de canais, sendo apresentados sucintamente os princípios relativos aos métodos supervisionados e não-supervisionados.

Capítulo 3 - Revisão de Marca d'Água Digital

No capítulo 3, apresentamos os conceitos básicos relacionados ao paradigma da marca d'água digital, passando pelo estudo de cada etapa do processo. Em alguns casos, é dada mais ênfase às técnicas de marca d'água aplicadas em áudio digital, pois tais conceitos serão explorados nos capítulos subsequentes.

Capítulo 4 - Equalização Supervisionada Utilizando Marca d'Água

Uma proposta original de equalização adaptativa utilizando marca d'água é o tópico deste capítulo. Após formular o problema de maneira genérica, é feito um desenvolvimento analítico para demonstrar a viabilidade do método. No caso em que o sinal de informação é considerado decorrelacionado, a solução de Wiener corresponde ao ponto que melhor inverte a função de transferência do canal. Já no caso em que o sinal de informação é correlacionado, a solução ótima de Wiener não corresponde à inversão; neste caso, apresentamos uma etapa de branqueamento do sinal de saída do equalizador, por meio de um filtro de erro de predição. Por fim, apresentamos resultados experimentais, buscando tecer comentários e apresentar conclusões.

Capítulo 5 - Identificação Supervisionada Utilizando Marca d'Água

Motivado pelos resultados apresentados no capítulo 4, este capítulo apresenta uma proposta original de identificação supervisionada de canais. Após formular o problema de maneira genérica, é feito um desenvolvimento analítico para demonstrar a viabilidade do método. Por fim, apresentamos resultados experimentais, buscando tecer comentários pertinentes para a obtenção de conclusões.

Capítulo 6 - Sincronismo para Sistemas Adaptativos Utilizando Marca d'Água

No capítulo 6, analisamos dois métodos de sincronismo para marca d'água digital, baseados no uso de sequência de treinamento e correlação. Estes métodos revertem os efeitos de uma grande classe de ataques de dessincronização; além disso, são capazes de resincronizar o sistema continuamente.

Capítulo 7 - Equalização e Identificação para Sinais de Áudio Utilizando Marca d'Água

Para ilustrar a aplicação das propostas anteriormente descritas a sinais que requerem o uso de técnicas perceptivas para garantir transparência da marca d'água, analisamos o desempenho dos métodos de equalização e identificação utilizando marca d'água em sinais discretos de áudio através de técnicas perceptivas que exploram os fenômenos psicoacústicos. Também são apresentados resultados experimentais para demonstrar a viabilidade técnica dos métodos.

Capítulo 8 - Conclusões e Perspectivas

Além das conclusões deste trabalho, este capítulo é dedicado a algumas sugestões para trabalhos

futuros decorrentes do conteúdo da dissertação.

Apêndice A - Fenômenos Auditivos e Modelos Perceptivos

O apêndice A apresenta os principais conceitos relacionados aos fenômenos auditivos do ouvido humano, relevantes para o entendimento dos conceitos aplicados em marca d'água para áudio. Algumas definições referentes a modelos acústicos criados a partir dos fenômenos auditivos, utilizadas no decorrer da dissertação, também são destacadas e discutidas em maiores detalhes neste apêndice.

Capítulo 2

Filtragem Adaptativa em Equalização e Identificação

2.1 Introdução

O homem demonstra ao longo de sua história a necessidade de interagir através da troca de mensagens. Têm-se buscado há séculos formas de codificar e transmitir mensagens de conteúdos diversos. Apesar da aparente simplicidade de tais objetivos é grande a variedade de abordagens possíveis para este problema. A escolha adequada do dispositivo utilizado para levar informação de um ponto a outro depende de diversos fatores, tais como distância de transmissão, características do meio físico, tecnologia disponível etc. Existem diferentes sistemas de comunicação; todos, porém, apresentam três elementos básicos, partindo de um elemento responsável pelo envio da mensagem (transmissor) até aquele que recebe a informação (receptor). Independentemente do mecanismo empregado para trafegar a informação, a separação física entre eles torna inevitável a existência de um canal de comunicação, que pode ocasionar modificações no sinal transmitido e até mesmo tornar a comunicação impraticável.

Para evitar que a comunicação seja impraticável, é necessário que o projetista crie contramedidas que atenuem ou cancelem os efeitos de um canal em um sistema de comunicação, em busca de níveis apropriados de qualidade para transmissão de informação. Neste trabalho serão apresentadas técnicas que viabilizam contramedidas baseadas em processamento de sinais e com foco em soluções adaptativas. Abordaremos dois problemas primordiais: a *equalização* e a *identificação* de um sistema (no nosso caso, um canal de comunicação). É apresentado a seguir uma definição para ambos (Attux, 2005):

- *Identificação* é um processo em que se busca obter um modelo matemático adequado de um

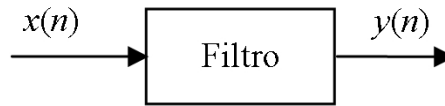


Fig. 2.1: Esquema de um filtro.

sistema ou, mais especificamente, de um dado canal de comunicação.

- *Equalização* é um processo pelo qual os efeitos do sistema são mitigados através da subsequente filtragem por um dispositivo adequadamente projetado para tal fim. A equalização pode então ser vista e por muitas vezes referenciada como um problema de identificação inversa.

Ambos os problemas têm motivações distintas, podendo porém ser empregados de forma hierárquica (e.g. pode-se identificar um canal e de posse desse conhecimento projetar um equalizador para mitigar seus efeitos). Para que seja possível realizar qualquer uma dessas tarefas, é necessário escolher a estrutura de filtragem, ou seja, o filtro utilizado como modelo ou equalizador, bem como o critério de ajuste dos parâmetros da estrutura.

Neste capítulo, analisaremos algumas classificações de estruturas de filtragem, abordaremos os conceitos essenciais relativos à equalização e identificação de sistemas e apresentaremos sucintamente duas grandes vertentes para a solução desses problemas: os métodos *supervisionados* e os *não-supervisionados*.

2.2 Estruturas de Filtragem

Todo sistema capaz de modificar certas componentes de frequência de um sinal, caracterizado por um mapeamento entre as variáveis de entrada e as de saída, é chamado de filtro. A figura 2.1 ilustra essa definição.

Existem duas classes fundamentais de filtros, associadas às respectivas estruturas de implementação. As *estruturas lineares*, cuja resposta a uma combinação linear de entradas é a combinação linear das respostas a cada entrada, e as *estruturas não-lineares*, cujo mapeamento não obedece a este princípio, dito de superposição.

Analisaremos essas duas estruturas de forma mais detalhada.

2.2.1 Estruturas Lineares

Sistemas lineares obedecem a duas relações fundamentais:

$$y(n) = w(n) * x(n) \quad (2.1)$$

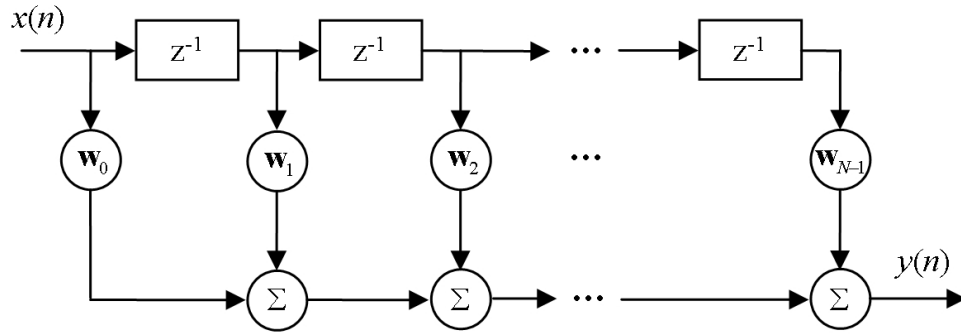


Fig. 2.2: Filtro linear transversal.

onde n em geral corresponde ao índice temporal discreto, $w(n)$ representa a resposta impulsiva do filtro, $x(n)$ o sinal de entrada e $y(n)$ representa o sinal de saída, e:

$$Y(z) = W(z) X(z) \quad (2.2)$$

sendo $W(z)$ a função de transferência do filtro, $X(z)$ a transformada z do sinal de entrada e $Y(z)$ a transformada z do sinal de saída.

Filtros lineares podem ser de resposta impulsiva finita (FIR) ou de resposta impulsiva infinita (IIR).

Filtros FIR

Um filtro FIR pode ser implementado como uma combinação linear das amostras de entrada com diferentes atrasos e ponderações, como mostra a figura 2.2.

Neste caso, a saída do filtro FIR de ordem N pode ser escrita como:

$$y(n) = \sum_{k=0}^{N-1} w(k)x(n-k) = \mathbf{w}^T \mathbf{x}(n) = [w_0 \ w_1 \ \dots \ w_{N-1}] \cdot \begin{bmatrix} x(n) \\ x(n-1) \\ \vdots \\ x(n-N+1) \end{bmatrix} \quad (2.3)$$

A função de transferência de um filtro desse tipo é dada por:

$$W(z) = \sum_{k=0}^{N-1} w(k) z^{-k} \quad (2.4)$$

Como se vê pela expressão na equação 2.4, o filtro FIR é caracterizado apenas pelos zeros de sua

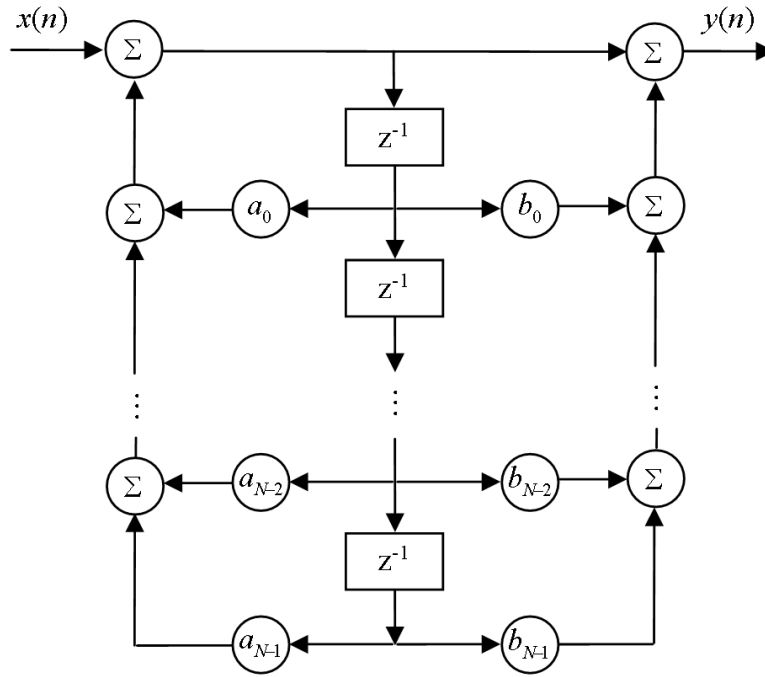


Fig. 2.3: Filtro IIR.

função de transferência, cujas posições são controladas pelos parâmetros do filtro.

Filtros IIR

Os filtros IIR apresentam realimentação, isto é, o valor atual da saída depende de valores atrasados e ponderados da própria saída, como mostra a figura 2.3.

A relação entre saída e entrada pode ser expressa como:

$$y(n) = \sum_{k=0}^{N_a-1} a_k x(n-k) - \sum_{p=1}^{N_b-1} b_p y(n-p) = \mathbf{a}\mathbf{x}(n) - \mathbf{b}\mathbf{y}(n-1) \quad (2.5)$$

A função de transferência desse filtro tem a seguinte forma:

$$W(z) = \frac{\sum_{k=0}^{N_a-1} a_k z^{-k}}{1 + \sum_{p=1}^{N_b-1} b_p z^{-p}} \quad (2.6)$$

Neste caso, o filtro IIR apresenta tanto pólos quanto zeros significativos, cujas posições são controladas pelos parâmetros do filtro.

2.2.2 Estruturas Não-lineares

A análise de sistemas não-lineares é mais trabalhosa do que a dos sistemas lineares, pois sua definição abrange muito mais possibilidades: estruturas não-lineares são aquelas que não obedecem ao princípio da superposição, ou seja, cuja resposta a uma sobreposição de entradas não é a sobreposição das respostas às entradas isoladas.

Duas classes de estruturas não-lineares se destacam pelo seu amplo uso prático (Haykin, 1996): os *filtros polinomiais* e as *redes neurais*.

- *Filtros polinomiais* ou de *Volterra*: a não-linearidade da estrutura é caracterizada como sendo do tipo polinomial.
- *Redes neurais*: a não-linearidade da estrutura é obtida através da conexão de *neurônios*, que podem ser vistos como sistemas de processamento paralelo de informação capazes de *aprender* e de se *auto-organizar*.

No decorrer deste trabalho, lidaremos apenas com estruturas lineares.

2.3 Filtragem Adaptativa

Um filtro pode ser definido como um dispositivo que atua em um conjunto de dados, visando a extração de informação relevante (Haykin, 1996). Na abordagem estatística para a solução do problema de filtragem, assumimos a disponibilidade de certos parâmetros estatísticos do sinal de informação (e.g. média e correlação) e usamos estes parâmetros para projetar um filtro linear de acordo com um critério estabelecido, geralmente valendo-nos de técnicas de otimização. Uma abordagem comumente empregada no problema de filtragem é a minimização do valor quadrático médio de um sinal de erro, definido como a diferença entre a resposta desejada e a saída atual do filtro; a solução resultante é conhecida como *filtro de Wiener*.

O desenvolvimento de um filtro de Wiener requer informações estatísticas prévias dos dados que serão processados. O filtro é considerado ótimo apenas quando as características estatísticas dos dados de entrada se igualam à informação prévia na qual o desenvolvimento do filtro está baseado. Para tal desenvolvimento, é necessário inicialmente estimar os parâmetros estatísticos relevantes dos sinais conhecido previamente e na entrada do filtro, e então igualar os resultados obtidos, utilizando uma formulação não-recursiva, para cálculo dos parâmetros do filtro. Entretanto, em aplicações que trabalham em tempo real, este processo não-recursivo apresenta algumas desvantagens, podendo ser excessivamente elaborado e computacionalmente custoso (Haykin, 1996).

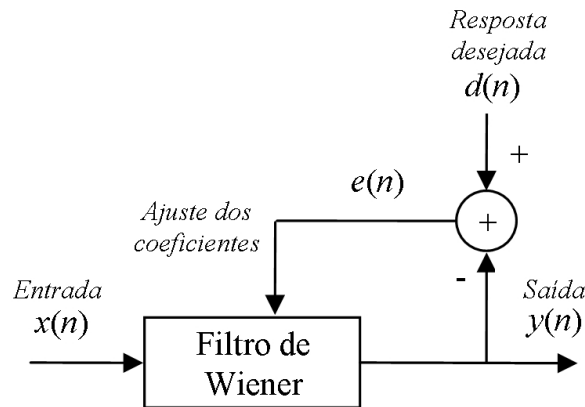


Fig. 2.4: Filtro de Wiener.

Um método mais eficiente para otimização no processo de filtragem é o uso de um *filtro adaptativo*, que possibilita um ajuste dos parâmetros do filtro sem o completo conhecimento prévio das características relevantes do sinal. O algoritmo é inicializado em condições pré-determinadas e, através de um processo iterativo, converge para a solução de Wiener. Na literatura, existe uma grande variedade de algoritmos para a otimização de filtros adaptativos lineares. A escolha do algoritmo mais adequado para um sistema depende de diversos fatores, alguns deles discutidos ao longo desta dissertação.

Para que possamos apresentar uma breve discussão sobre a teoria da filtragem adaptativa, da forma mais conveniente para os assuntos tratados nesta dissertação, teceremos a seguir alguns comentários sobre a teoria da filtragem de Wiener e dos algoritmos LMS e RLS.

2.3.1 Critério de Wiener: Filtragem a Partir de uma Medida de Erro

Ao se projetar um filtro, o passo inicial é a escolha de uma estrutura que seja capaz de resolver os problemas do projetista. Em seguida é necessário ajustar os parâmetros da estrutura em busca de uma configuração ótima segundo algum critério. Uma maneira de se obter uma função custo coerente com os objetivos da equalização ou identificação é realizar uma comparação entre a saída do filtro e um sinal desejado conhecido¹ (aquilo que se esperaria da saída do filtro num caso favorável, conseguido com a presença de um sinal piloto). Este modelo de filtragem, associado ao paradigma de Wiener, é ilustrado na figura 2.4.

A idéia do método consiste em inserir amostras de um processo estocástico $x(n)$ em um filtro com vetor de parâmetros $\mathbf{w}$, produzindo um sinal de saída $y(n)$ tão próximo quanto possível de um sinal desejado $d(n)$, sendo que a diferença entre os sinais dá origem a um erro de estimação $e(n)$. O

¹A existência de um sinal piloto (sinal conhecido aplicado à entrada do canal) caracteriza uma técnica supervisionada.

objetivo do ajuste do filtro é projetar um dispositivo capaz de modificar o sinal $x(n)$ de modo a se ter uma saída semelhante a um modelo de sinal conhecido previamente. O grande desafio é, então, reduzir o erro de estimação a um mínimo.

Uma medida muito comum e efetiva, capaz de tratar indiferentemente erros positivos e negativos e levar em consideração uma minimização no sentido da média e não apenas valores instantâneos, é o erro quadrático médio (EQM). A função custo correspondente, que expressa o critério de Wiener, é definida como²:

$$J = E[e(n) e^*(n)] = E[|e(n)|^2] \quad (2.7)$$

onde E denota o operador estatístico de esperança, $*$ indica complexo conjugado e o sinal de erro $e(n)$ é definido pela diferença:

$$e(n) = d(n) - y(n) \quad (2.8)$$

A saída do filtro é definida como uma convolução linear:

$$y(n) = \sum_{k=0}^q w_k^* x(n-k) \quad (2.9)$$

onde $q+1$ é o número de coeficientes do filtro linear. Definimos o k -ésimo coeficiente do filtro w_k , em termos de partes reais e imaginárias, como segue:

$$w_k = a_k + jb_k \quad (2.10)$$

Substituindo 2.8 e 2.9 em 2.7, e aplicando o operador gradiente, obtemos a seguinte expressão:

$$\nabla_k J = -2E[x(n-k) e^*(n)] \quad (2.11)$$

Para custo mínimo, os elementos do vetor gradiente devem ser iguais a zero:

$$E[x(n-k) e_o^*(n)] = 0, \quad k = 0, 1, 2, \dots, q \quad (2.12)$$

onde e_o é o erro de estimação mínimo (obtido quando o filtro opera na condição ótima). Substituindo 2.8 e 2.9 em 2.12 temos:

$$E\left[x(n-k) \left(d^*(n) - \sum_{i=0}^q w_{oi}^* x(n-i)\right)\right] = 0 \quad (2.13)$$

²A nomenclatura utilizada nesta seção supõe sinais complexos.

onde w_{oi} é o k -ésimo coeficiente na resposta impulsiva do filtro ótimo. Após algumas manipulações, tem-se:

$$\sum_{i=0}^q w_{oi} E[x(n-k)x^*(n-i)] = E[x(n-k)d^*(n)] , \quad k = 0, 1, 2, \dots, q \quad (2.14)$$

O termo $E[x(n-k)x^*(n-i)] = r(i-k)$ é a função de autocorrelação na entrada do filtro para um atraso $i-k$. O termo $E[x(n-k)d^*(n)] = p(-k)$ é a correlação cruzada entre a entrada do filtro e a resposta desejada para um atraso $-k$. Assim, a equação 2.14 pode ser reescrita como:

$$\sum_{i=0}^q w_{oi} r(i-k) = p(-k) , \quad k = 0, 1, 2, \dots, q \quad (2.15)$$

Este sistema de equações (equação 2.15) é chamado de *equações de Wiener-Hopf*.

Podemos reescrever as equações de Wiener-Hopf na forma matricial:

$$\mathbf{R}\mathbf{w}_o = \mathbf{p} \quad (2.16)$$

onde $\mathbf{R} = E[\mathbf{x}(n)\mathbf{x}^H(n)]$ e $\mathbf{p} = E[\mathbf{x}(n)d^*(n)]$, sendo $\mathbf{x}(n)$ definido como um vetor de dimensão $M \times 1$, onde M representa o número de coeficientes do filtro.

Pela inversão da matriz de correlação, obtém-se o valor dos coeficientes ótimos $\mathbf{w}_o$:

$$\mathbf{w}_o = \mathbf{R}^{-1}\mathbf{p} \quad (2.17)$$

Os resultados anteriores que levam à solução de Wiener expressa em 2.17 mostram que há apenas um ponto que anula o gradiente da função, demonstrando que a expressão é *unimodal*. No sentido de obter o valor de mínimo desta função custo, temos:

$$J(\mathbf{w}_o) = E[|e_o(n)|^2] \quad (2.18)$$

Suporemos ainda processos de média nula:

$$\sigma_d^2 = \sigma_{yo}^2 + J(\mathbf{w}_o) \quad (2.19)$$

Portanto, teremos:

$$J_{\min} = J(\mathbf{w}_o) = \sigma_d^2 - \sigma_{yo}^2 \quad (2.20)$$

A equação 2.20 expressa que o valor mínimo da função custo é a diferença entre a variância do sinal desejado e a variância da saída do filtro na condição ótima. Partindo das equações 2.7, 2.8 e 2.9 e, após diversas manipulações, chegamos à expressão geral para a função custo do critério de Wiener:

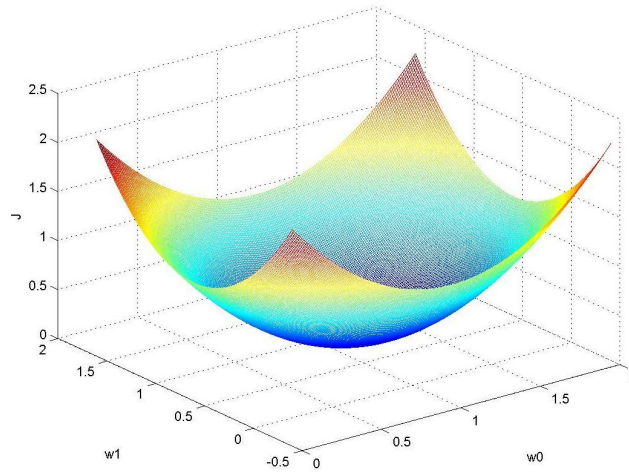


Fig. 2.5: Função Custo de Wiener.

$$J(\mathbf{w}) = \sigma_d^2 - \mathbf{w}^H \mathbf{p} - \mathbf{p}^H \mathbf{w} + \mathbf{w}^H \mathbf{R} \mathbf{w} \quad (2.21)$$

A função custo de Wiener expressa em 2.21 é quadrática e tem a forma de uma parabolóide. Conforme demonstrado, esta parabolóide é caracterizada por um único ponto de mínimo, que representa o valor ótimo do filtro de Wiener. A figura 2.5 apresenta uma típica função custo de Wiener.

2.3.2 Algoritmo do Gradiente Determinístico

O método de adaptação baseado no gradiente emprega uma técnica clássica de otimização conhecida como *steepest descent*. Partindo do fato de que um novo valor da função J pode ser obtido a partir do seu comportamento no ponto $\mathbf{w}(n)$ através de uma expansão em série de Taylor, o método *steepest descent* corresponde a truncar-se esta expansão na derivada de primeira ordem, ou seja, no cálculo do gradiente da função. Intuitivamente, percebe-se que, ao caminhar no sentido contrário ao de maior crescimento do gradiente da função, através de sucessivas correções, é possível buscar um ponto de mínimo, pois trata-se da descida com maior declividade na curva da função custo. Assim, dada uma função custo $J(n)$, a atualização do vetor de parâmetros $\mathbf{w}$ pode ser expressa como:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu [-\nabla J(n)] \quad (2.22)$$

onde μ é denominado *passo de adaptação* e controla o tamanho do passo dado a cada iteração, e $\nabla J(n)$ é o gradiente da função $J(n)$.

Para utilizar este método na busca de um valor que otimiza a função custo de Wiener, precisamos calcular o vetor gradiente de tal função. Aplicando o operador gradiente à equação 2.21, obtemos:

$$\nabla J(n) = -2\mathbf{p} + 2\mathbf{R}\mathbf{w}(n) \quad (2.23)$$

A partir de 2.22 e 2.23, obtém-se a expressão do algoritmo do gradiente estocástico aplicado à busca da solução de Wiener:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + 2\mu [\mathbf{p} - \mathbf{R}\mathbf{w}(n)] \quad (2.24)$$

O fator multiplicativo 2 pode ser incluído no próprio passo de adaptação μ , produzindo a expressão final para o algoritmo do gradiente determinístico:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu [\mathbf{p} - \mathbf{R}\mathbf{w}(n)] \quad (2.25)$$

O algoritmo 2.25 apresenta um único ponto de mínimo para convergência (mínimo global), que corresponde à solução de Wiener. Apesar do algoritmo ser iterativo, pressupõe-se o conhecimento da matriz de autocorrelação $\mathbf{R}$ e do vetor de correlação cruzada $\mathbf{p}$, cujos valores devem ser previamente estimados. Entretanto, esta estimativa pode ser computacionalmente custosa em alguns casos, além de dificultar o processamento em tempo real. Estas dificuldades levaram ao surgimento de algoritmos baseados em aproximação estocástica.

2.3.3 Algoritmo *Least Mean Square*

Conforme discutido anteriormente, o emprego do método *steepest descent* exige uma estimativa prévia da autocorrelação e correlação cruzada do sinal. Porém, tal procedimento é, em alguns casos, indesejável. Torna-se então interessante o uso de um procedimento de menor custo do ponto de vista computacional.

A idéia de aproximação estocástica foi proposta inicialmente por Robbins e Monro (Robbins and Monro, 1951) e constitui a base dos algoritmos de filtragem adaptativa, assim como de áreas como identificação e controle adaptativo, reconhecimento de padrões por redes neurais e outros (Neves, 2001). O algoritmo LMS (*Least Mean Square*) foi proposto por Widrow e Hoff (Haykin, 1996). Este tem por base uma aproximação: a substituição da esperança por valores instantâneos dos sinais envolvidos. Isto dá origem às grandezas aproximadas:

$$\hat{\mathbf{R}} = \mathbf{x}(n) \mathbf{x}^H(n) \quad (2.26)$$

e

$$\hat{\mathbf{p}} = \mathbf{x}(n) d^*(n) \quad (2.27)$$

Substituindo 2.26 e 2.27 em 2.25, obtemos:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu [\mathbf{x}(n) d^*(n) - \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{w}(n)] \quad (2.28)$$

que, após algumas manipulações, resulta em:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \mathbf{x}(n) [d^*(n) - \mathbf{x}^H(n) \mathbf{w}(n)] \quad (2.29)$$

A expressão 2.29 pode ser expressa na forma de três relações básicas:

- saída do Filtro:

$$y(n) = \mathbf{w}^H(n) \mathbf{x}(n) \quad (2.30)$$

- erro de estimação:

$$e(n) = d(n) - y(n) \quad (2.31)$$

- adaptação dos parâmetros do filtro:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \mathbf{x}(n) e^*(n) \quad (2.32)$$

Em 2.32, temos a expressão final do LMS, cujo nome deve-se ao fato de que o algoritmo minimiza o erro quadrático médio de estimação. O fato de o algoritmo basear-se em uma busca estocástica extremamente simples o torna bastante atraente do ponto de vista prático.

As aproximações realizadas no LMS fazem com que seu desempenho tenha diferenças significativas em relação ao *steepest descent*. O LMS não convergirá mais para a solução de Wiener, como no caso do *steepest descent*, mas oscilará em torno da solução ótima. O grau de oscilação depende do passo de adaptação utilizado.

Apesar das variâncias relativamente altas das estimativas empregadas no LMS, o algoritmo opera de forma recursiva realizando uma média das estimativas a cada iteração, o que leva geralmente a um desempenho satisfatório (convergência em torno da solução ótima).

2.3.4 Algoritmo *Recursive Least Squares*

O algoritmo RLS (*Recursive Least Squares*) utiliza uma função custo sem a presença do operador esperança, baseando-se na minimização de uma função custo obtida a partir de uma soma ponderada do valor absoluto do erro quadrático. Porém, as equações nas quais se baseia o RLS são semelhantes às de Wiener, utilizando estimativas temporais da autocorrelação e da correlação cruzada dos sinais envolvidos, atualizadas de forma iterativa.

O RLS apresenta uma complexidade computacional proporcional a M^3 , onde M é o número de coeficientes do filtro, pois necessita da inversão da estimativa da matriz de autocorrelação do sinal de entrada. Com isso, ele se torna computacionalmente custoso. É possível evitar a inversão explícita da matriz de autocorrelação aplicando-se o *lema de inversão de matrizes* (Haykin, 1996). Este lema permite uma estimação recursiva da matriz inversa, reduzindo a complexidade computacional para M^2 . Apesar da redução significativa, esta complexidade ainda é alta se comparada à do algoritmo LMS, que é proporcional a M .

Através do sinal de saída do filtro (2.30) e do erro de estimação (2.31), obtêm-se as equações de atualização do vetor de pesos $\mathbf{w}$ e da estimativa da inversa da matriz de autocorrelação $\hat{\mathbf{R}}^{-1}$:

$$\mathbf{g}(n) = \frac{\lambda^{-1} \hat{\mathbf{R}}^{-1}(n-1) \mathbf{x}(n)}{1 + \lambda^{-1} \mathbf{x}^H(n) \hat{\mathbf{R}}^{-1}(n-1) \mathbf{x}(n)} \quad (2.33)$$

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mathbf{g}(n) e^*(n) \quad (2.34)$$

$$\hat{\mathbf{R}}^{-1}(n) = \lambda^{-1} [1 - \mathbf{g}(n) \mathbf{x}^H(n)] \hat{\mathbf{R}}^{-1}(n-1) \quad (2.35)$$

onde $\mathbf{g}$ é o vetor de ganho, λ é o fator de esquecimento e $\hat{\mathbf{R}}^{-1}(0)$ é igual a $\delta^{-1} \mathbf{I}$, sendo δ uma constante positiva de pequeno valor.

O fator de esquecimento é utilizado como um controle do RLS para que “esqueça” dados mais antigos, ou seja, está relacionado ao fato do algoritmo deixar de usar amostras antigas nos cálculos. Para um fator de esquecimento igual a 1, todas as amostras serão utilizadas no cálculo e, à medida em que o valor diminui, o algoritmo passa a “esquecer” as amostras passadas. Valores superiores a 0,98 são comumente utilizados com sistemas invariantes ou pouco variantes no tempo. Valores menores são utilizados para sistemas variantes, podendo aproximar-se de 0,9 (Panazio, 2001). O RLS proporciona uma solução ótima para a estimativa da matriz $\hat{\mathbf{R}}^{-1}$, e o vetor de ganho sempre aponta para a solução ótima, sendo o RLS independente das características do canal.

2.4 Equalização

Nos modernos sistemas de comunicação, a equalização, realizada por meio de um dispositivo denominado equalizador, tem por objetivo compensar os efeitos de distorção do sinal causados pelo sistema, ou, mais especificamente, um canal de comunicação. O emprego de um equalizador permite assim melhorar a confiabilidade da mensagem recebida. A interferência intersimbólica (IIS, que consiste na extensão de um símbolo transmitido sobre um intervalo de tempo maior do que sua dura-

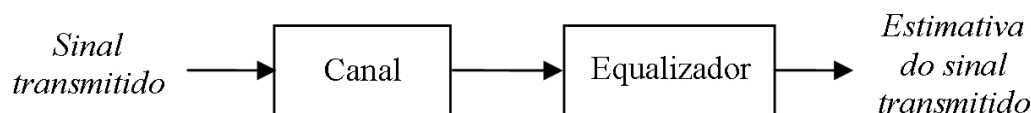


Fig. 2.6: Modelo simplificado de um processo de equalização.

ção original, fazendo com que interfira no próximo símbolo transmitido) e o ruído, são as principais causas de degradação de desempenho em um sistema de comunicação, podendo ser mitigados com o correto uso de um equalizador.

O equalizador é um filtro que, quando localizado no receptor, tenta recuperar, a partir do sinal corrompido, a informação transmitida, conforme ilustrado na figura 2.6. Ele também pode ser utilizado do lado da transmissão, processando o sinal antes mesmo de ser transmitido, de maneira que o sinal, após passar pelo canal, não contenha distorções. Esta tarefa é chamada de pré-equalização.

Esquemas de equalização não são empregados apenas em sistemas de comunicação. Os aparelhos de som, por exemplo, possuem regulagens de graves e agudos que permitem a correção de imperfeições de gravação e reprodução, proporcionando uma melhoria da sensação sonora. O equipamento utilizado em áudio para amplificação de faixas de frequência é chamado de equalizador. Este exemplo de equalização pode ser mais familiar à maioria das pessoas.

Quando se projeta um filtro, tem-se como passo inicial a escolha de uma estrutura, fazendo-se então necessário o ajuste dos seus parâmetros. A equalização pode ser definida como o projeto de um filtro que, atuando sobre o sinal recebido, contribua para uma restauração mais precisa da mensagem original, cancelando os efeitos indesejáveis do canal. Tendo em vista a necessidade de equalizar sistemas desconhecidos ou variantes no tempo, faz-se necessário o uso de equalizadores adaptativos. Estes equalizadores são normalmente adaptados por meio de uma seqüência de treinamento conhecida no receptor. Esta classe de método é conhecida como equalização adaptativa supervisionada.

2.4.1 Equalização Supervisionada

No esquema clássico de equalização supervisionada, associado ao paradigma de Wiener, o objetivo de um equalizador é ajustar os parâmetros do filtro visando à aproximação de sua saída a um sinal desejado conhecido. A figura 2.7 ilustra o esquema deste modelo de filtragem.

Basicamente, tem-se um filtro equalizador em cuja entrada são inseridas amostras de um processo estocástico $x(n)$. A saída $y(n)$ deste filtro é comparada com um sinal de referência $d(n)$, dando origem a um sinal de erro $e(n)$ chamado de erro de estimação.

O esquema tem como objetivo ajustar os parâmetros do equalizador fazendo com que $y(n)$ seja o mais próximo possível de $d(n)$, de modo a reduzir o erro de estimação a um mínimo. Os esforços em torno do problema da equalização se concentram nesta otimização.

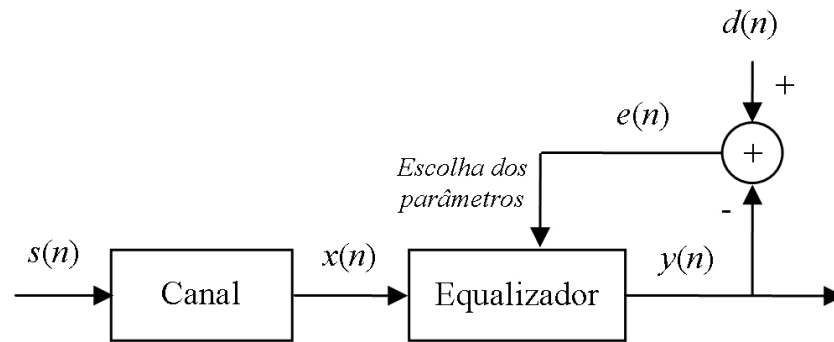


Fig. 2.7: Esquema clássico de equalização supervisionada.

O paradigma da *equalização adaptativa supervisionada* consiste em interromper periodicamente (durante os *períodos de treinamento*) a transmissão de informação útil para transmitir uma sequência de treinamento. Esta sequência é conhecida no receptor, permitindo que os coeficientes do equalizador possam ser ajustados através da comparação entre as sequências de treinamento original e distorcida pelo canal. Com a atuação do equalizador sobre o sinal recebido, obtém-se uma estimativa mais exata do sinal transmitido. Exemplos de algoritmos que tradicionalmente utilizam a abordagem citada são o LMS e o RLS (Haykin, 1996).

Neste cenário, o sinal que carrega informação é interrompido durante a transmissão da sequência de treinamento; com isso, tem-se um gasto de recursos causado pelo envio periódico de dados que não contêm informação. Em aplicações de tempo real, estas interrupções podem ser indesejáveis ou mesmo inaceitáveis. Uma alternativa ao uso de sequências de treinamento é recorrer à teoria da *equalização não-supervisionada* ou *cega*, cujo desafio é definir critérios que indiquem o grau de sucesso na equalização de um sistema sem necessidade de um sinal de referência.

2.5 Identificação

Nos sistemas de comunicação, a identificação consiste em modelar um canal de comunicação que ignoramos, ou seja, obter um modelo adequado e fiel deste canal. Um sistema de identificação é composto de um filtro que, quando localizado no receptor, tenta modelar o canal de comunicação através do qual o sinal foi transmitido. Muitas vezes um sistema de identificação pode ser utilizado de forma hierárquica com um sistema de equalização; em outras palavras, é possível projetar um equalizador a partir do modelo de canal obtido por um processo de identificação.

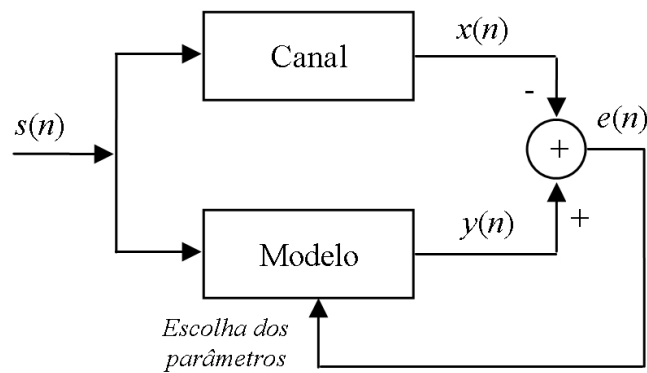


Fig. 2.8: Esquema clássico de identificação supervisionada.

2.5.1 Identificação Supervisionada

O paradigma que norteia a identificação assemelha-se ao da equalização. No esquema clássico de identificação supervisionada associada ao paradigma de Wiener, tem-se como objetivo obter um filtro, utilizado como modelo, que seja tão próximo quanto possível do sistema analisado. Quando um mesmo sinal é aplicado às entradas do canal e do modelo, a saída do modelo deve ser tão próxima quanto possível da saída do canal. A figura 2.8 ilustra o esquema deste modelo de filtragem.

Basicamente, tem-se um filtro em cuja entrada são inseridas amostras de um processo estocástico $s(n)$ (sinal antes de sofrer os efeitos do canal). A saída $y(n)$ deste filtro é comparada à saída $x(n)$ do canal, dando origem a um sinal de erro $e(n)$, chamado de erro de estimação.

O esquema tem como objetivo ajustar os parâmetros do modelo fazendo com que $y(n)$ seja o mais próximo possível de $x(n)$, de modo a minimizar o erro de estimação. Quanto menor o erro, mais próximo o modelo se encontra do canal analisado.

Assim como no caso de equalização, o paradigma da *identificação adaptativa supervisionada* consiste em interromper periodicamente a comunicação de informação útil para transmitir uma sequência de treinamento. Devido esta última excitar tanto o canal quanto o modelo, os coeficientes deste último podem ser ajustados através da comparação entre os sinais de saída de ambos.

2.6 Introdução aos Métodos Não-Supervisionados

Conforme discutimos anteriormente, os métodos supervisionados requerem a presença de uma sequência de treinamento para que seja possível a adaptação dos parâmetros da estrutura de filtragem. Durante o período de treinamento, não há transmissão efetiva de informação. Aplicações onde haja escassez de recursos sistêmicos, ou que trabalhem em tempo real, são exemplos de aplicações nas quais a transmissão de uma sequência de treinamento é custosa, podendo mesmo inviabilizar

sua operação. Tendo isto em mente, iniciou-se a busca por técnicas que não apresentassem essa dependência. Surgiu então um novo paradigma, correspondendo às técnicas *não-supervisionadas* ou *cegas*.

Neste paradigma, a questão central consiste em como projetar um filtro sem precisar de uma sequência de treinamento. De forma geral, as técnicas propostas baseiam-se exclusivamente no sinal recebido. Os critérios não-supervisionados são caracterizados de acordo com a forma como as informações são utilizadas. Com isso, podemos dividir as abordagens existentes em três grandes grupos (Haykin, 1996; Ding and Li, 2001; Suyama, 2003):

- Algoritmos baseados em estatísticas de ordem superior³ (HOS) de forma implícita: são os algoritmos que exploram estatísticas de ordem superior do sinal recebido sem o cálculo explícito de qualquer dessas quantidades.
- Algoritmos baseados explicitamente em HOS: exploram de forma explícita os cumulantes de ordem superior do sinal recebido, ou seus correspondentes no domínio da frequência, os poli-espectros.
- Algoritmos baseados em estatísticas cicloestacionárias: exploram características de cicloestacionariedade do sinal recebido. Utilizando esta propriedade é possível realizar a equalização empregando apenas estatísticas de segunda ordem.

A teoria de estatísticas de ordem superior tem sido muito usada na construção de critérios de equalização e identificação cegas. Tais estatísticas, diferentemente ao que se dá com a autocorrelação, preservam informação de fase do canal; esta última, além da magnitude, faz-se necessária na identificação da função de transferência do canal desconhecido. No entanto, tais técnicas podem apresentar problemas de convergência para mínimos locais e, em alguns casos, custo computacional superior ao das técnicas supervisionadas (Ding and Li, 2001; Neves, 2001).

Dependendo do problema abordado (equalização ou identificação) e do critério estabelecido, vieram à tona resultados que formam uma base genérica da teoria da equalização e identificação não-supervisionada. Apesar de o tema central do nosso trabalho não utilizar critérios não-supervisionados, vale a pena tecer alguns comentários sobre eles, uma vez que, no decorrer desta dissertação, faremos alguns comentários e comparações dos métodos por nós propostos com as técnicas não-supervisionadas.

³Em inglês, *Higher Order Statistics*.

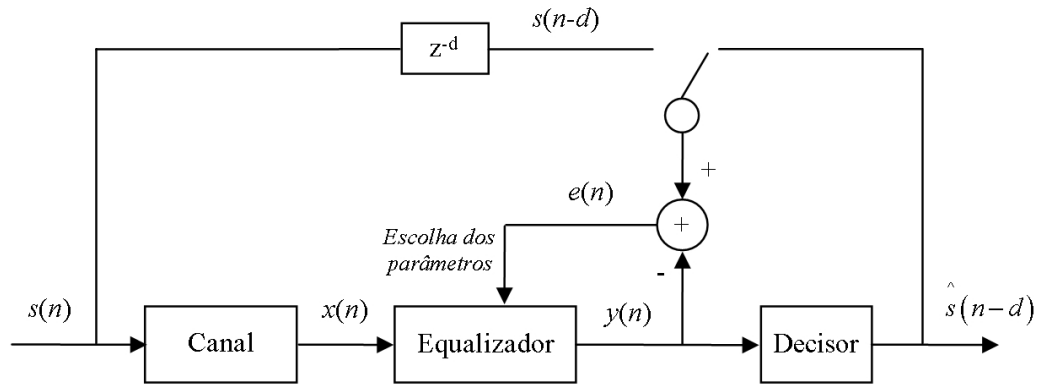


Fig. 2.9: Esquema clássico de treinamento com o critério DD.

2.6.1 O Critério de Decisão Direta

O critério de Decisão Direta (DD) pode ser considerado a primeira abordagem de equalização não-supervisionada, tendo sido proposto inicialmente em 1965 (Lucky, 1965). O critério permite que a informação necessária ao ajuste dos parâmetros do equalizador seja extraída sem interrupção do sinal de informação transmitido.

O algoritmo DD é normalmente utilizado em conjunto com outro algoritmo não-supervisionado ou com um algoritmo supervisionado (e.g. LMS). Ao término da adaptação de um equalizador num sistema supervisionado com duração curta, é de se esperar que a saída do equalizador seja uma estimativa razoável do sinal transmitido. É então possível construir um algoritmo do tipo Wiener chaveando o sistema para o chamado modo DD, ilustrado na figura 2.9. Tem-se a seguinte função custo para o algoritmo DD:

$$J_{DD} = E \left[|\text{sgn}[y(n)] - y(n)|^2 \right] \quad (2.36)$$

onde $\text{sgn}(\cdot)$ é uma função sinal (função *signum*), definida como:

$$\text{sgn}[y(n)] = \begin{cases} +1 & \text{se } y(n) > 0 \\ -1 & \text{se } y(n) < 0 \end{cases} \quad (2.37)$$

e $+1$ ou -1 , com igual probabilidade, se $y(n) = 0$. O algoritmo DD é dado por:

$$\begin{aligned} y(n) &= \mathbf{w}^T(n) \mathbf{x}(n) \\ \mathbf{w}(n+1) &= \mathbf{w}(n) + \mu \{ \text{sgn}[y(n)] - y(n) \} \mathbf{x}(n) \end{aligned} \quad (2.38)$$

Para a utilização do algoritmo, é esperado que o equalizador já se encontre com um ajuste capaz de inverter os efeitos do canal, ou seja, que esteja trabalhando em uma condição de “diagrama de olho aberto”, i.e. com baixa probabilidade de erros de decisão (Lee and Messerschmitt, 1994). Após uma

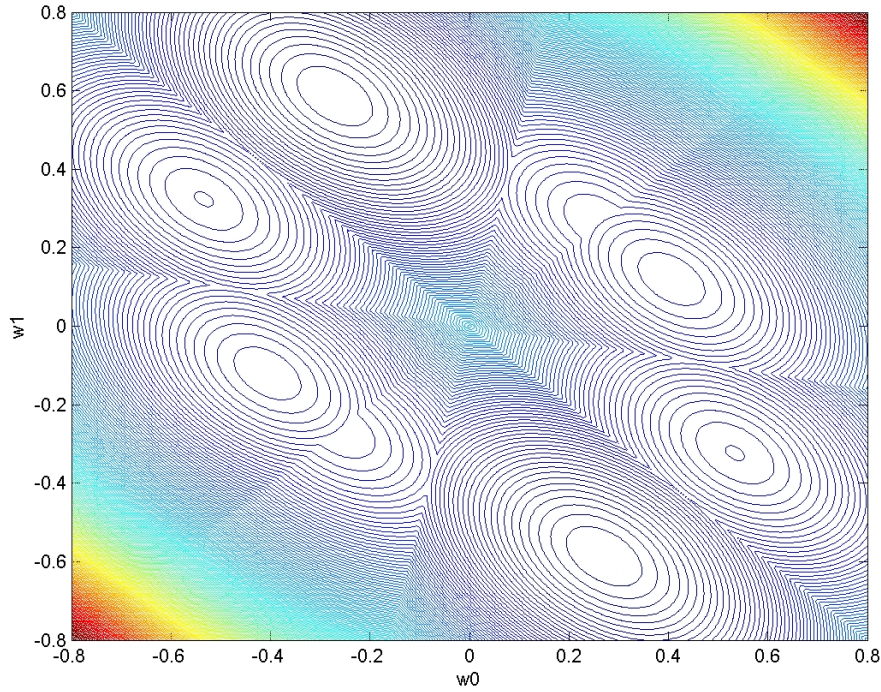


Fig. 2.10: Curvas de nível para função custo DD.

análise mais detalhada, foi demonstrado que o algoritmo DD pode convergir para mínimos espúrios quando utilizado sem o auxílio de um outro algoritmo (Mazo, 1980; Macchi and Eweda, 1984). Isso quer dizer que, exceto em condições muito favoráveis, o algoritmo não é capaz de “abrir o olho do canal”. A figura 2.10 apresenta uma típica função custo DD.

2.6.2 O Critério de Godard

Em outro trabalho representativo (Godard, 1980), foi apresentado uma família de algoritmos autotodatas de módulo constante utilizados com sinais bidimensionais, baseados na seguinte função custo:

$$J_G^{(p)} = E \left[(|y(n)|^p - R_p)^2 \right] \quad (2.39)$$

sendo R_p uma constante definida a partir de estatísticas do sinal transmitido:

$$R_p = \frac{E \left[|s(n)|^{2p} \right]}{E \left[|s(n)|^p \right]} \quad (2.40)$$

Apesar de ter apresentado uma família de algoritmos, Godard limitou-se ao estudo daqueles cor-

respondentes a $p = 1$ e $p = 2$. O segundo, correspondente ao caso mais comum, deu origem ao critério do módulo constante (CM). A função custo destes algoritmos é não-convexa, apresentando mínimos locais não desejáveis; dessa forma, a inicialização do algoritmo é peça chave para uma correta convergência. Apesar desta propriedade desfavorável, observa-se que o algoritmo converge para mínimos próximos ao de Wiener, porém com velocidade de convergência lenta se comparado a métodos supervisionados.

2.6.3 O Critério do Módulo Constante

Apesar de introduzido por Godard (Godard, 1980), o termo *módulo constante* foi consolidado por Treichler e Agee (Treichler and Agee, 1983). A função custo do critério CM é baseada em um caso específico de $J_G^{(p)}$ (2.39), para $p = 2$:

$$J_{CM} = E \left[\left(|y(n)|^2 - R_2 \right)^2 \right] \quad (2.41)$$

onde

$$R_2 = \frac{E \left[|s(n)|^4 \right]}{E \left[|s(n)|^2 \right]} \quad (2.42)$$

O critério do módulo constante deu origem ao correspondente algoritmo CMA (*Constant Modulus Algorithm*):

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu y(n) \left[R_2 - |y(n)|^2 \right] \mathbf{x}(n) \quad (2.43)$$

O critério CM é considerado uma das mais sólidas técnicas não-supervisionadas, tendo merecido muita atenção nos últimos anos (Attux, 2001; Suyama, 2003). A figura 2.11 apresenta uma típica função custo CM.

2.6.4 O Critério de Shalvi-Weinstein

O critério de Shalvi-Weinstein (SW) (Shalvi and Weinstein, 1990) envolve o cálculo explícito de uma estatística de ordem superior, a *curtose*, para determinação dos parâmetros do equalizador. A curtose é calculada a partir de estatísticas do sinal considerado, seguindo a expressão:

$$K[y(n)] = E \left[|y(n)|^4 \right] - 2E \left[|y(n)|^2 \right] - \left| E \left[(y(n))^2 \right] \right|^2 \quad (2.44)$$

Define-se o critério de Shalvi-Weinstein como um problema de maximização:

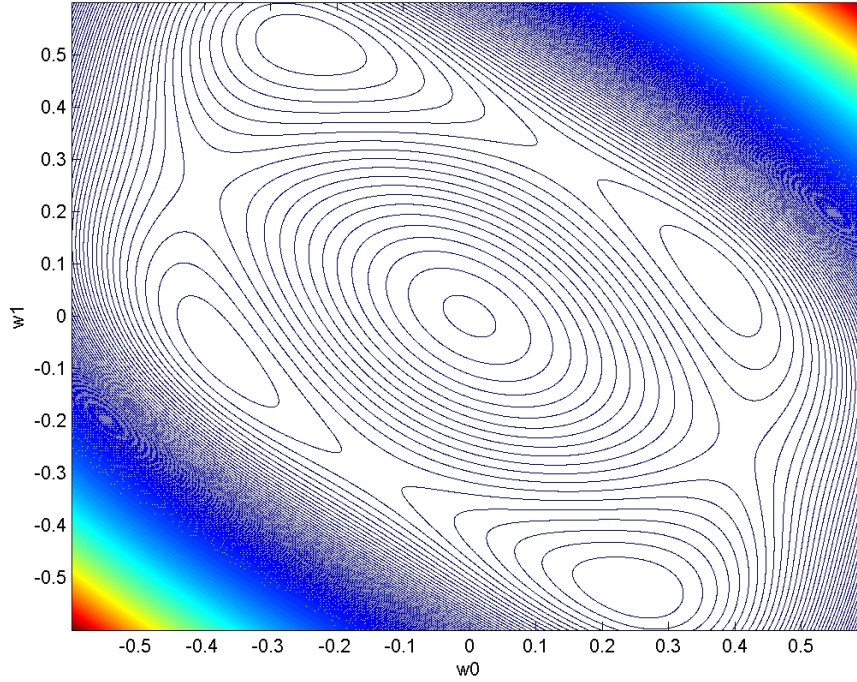


Fig. 2.11: Curvas de nível para função custo CM.

$$\max |K[y(n)]| \quad (2.45)$$

sujeito à restrição:

$$E[|y(n)|^2] = E[|s(n)|^2] \quad (2.46)$$

A maximização da curtose provém de uma propriedade inerente às hipóteses de canal e filtros lineares e da dependência entre símbolos transmitidos. Em um sistema com tal configuração, caso haja igualdade entre as estatísticas de ordem dois da sequência transmitida e equalizada, a curtose do sinal filtrado será menor ou igual à do sinal transmitido. A condição de igualdade somente existirá em caso de equalização perfeita (Shalvi and Weinstein, 1990).

2.7 Discussão

Neste capítulo, foram apresentados os principais conceitos relacionados à filtragem adaptativa e aos problemas de equalização e identificação. Vimos que os métodos clássicos de filtragem super-

visionada possuem as vantagens de levar a uma função custo unimodal e, na maioria dos casos, a uma menor complexidade em relação a técnicas não-supervisionadas. Em contrapartida, apresentam o custo da interrupção periódica da transmissão do sinal de informação, substituindo-o por uma sequência de treinamento. Já os métodos não-supervisionados apresentam a vantagem de não necessitar de uma sequência de treinamento, baseando-se exclusivamente no sinal recebido e viabilizando a transmissão contínua do sinal de informação. No entanto, tais técnicas geralmente apresentam uma função custo multimodal, com vários mínimos associados a custos distintos, além de, na maioria dos casos, apresentar maior complexidade computacional.

Apresentaremos nos próximos capítulos uma proposta para o uso de uma marca d'água em lugar de uma sequência de treinamento para supervisão do sistema adaptativo. A idéia consiste em sobrepor um sinal de marca d'água continuamente ao sinal de informação, sendo ambos transmitidos de forma simultânea. A principal motivação desta proposta é unir as vantagens apresentadas pelas técnicas supervisionadas e não-supervisionadas. Como a marca d'água é continuamente sobreposta ao sinal de informação, não se faz necessária a interrupção deste último para a transmissão de uma sequência de treinamento. Além disso, como nos métodos supervisionados tradicionais, a função custo obtida em nossa proposta é unimodal, e os algoritmos dela decorrentes guardam as propriedades de convergência e de simplicidade de implementação.

Capítulo 3

Revisão de Marca d'Água Digital

3.1 Introdução

A necessidade de garantir integridade, fidelidade e segurança quanto à propriedade e transmissão de informação remonta à antigüidade. O primeiro relato histórico de que temos conhecimento ocorreu em 480 a.C., narrada com a *História de Heródotos* (Moulinand and Koetter, 2005). Histiaeus quis notificar secretamente o regente da cidade grega de Mileto o início de uma revolta contra a ocupação persa. Para tanto, Histiaeus criou um engenhoso método de comunicação: raspou a cabeça de um escravo, tatuou uma mensagem em seu crânio, esperou seu cabelo crescer e, finalmente, enviou-o a Mileto. Chegando lá, o escravo teve a cabeça raspada para revelar a mensagem.

O início do uso de técnicas de marca d'água deu-se no final do século XIII, com o surgimento do papel. O primeiro registro de um papel marcado data de 1292, tendo sua origem na localidade de Fabriano, na Itália (Hartung and Kutter, 1999). Existiam em Fabriano cerca de 40 oficinas para fabricação de papel. Cada oficina fabricava papéis com características específicas. Esse papel era posteriormente tratado por artesãos que poliam e dividiam as folhas. No término do ciclo de produção, o produto era entregue a retalhistas, que o vendiam ao público. A concorrência em cada uma das etapas era grande e, pelo fato de o papel passar por diversas mãos, tornava-se difícil saber sua procedência. Surgiram então as marcas d'água, como um meio de garantir a origem e autenticidade do papel. Logo apareceram tentativas de imitação, falsificação e mesmo anulação das marcas d'água. Ainda hoje, o papel utilizando técnicas de marca d'água é muito empregado em documentos cuja autenticidade deve ser passível de comprovação, tais como documentos de identidade e cédulas monetárias.

Com o crescente uso, no final do século XX, de informação em formato digital, tornou-se simples duplicar documentos e distribuir suas cópias com agilidade (através, por exemplo, da internet), muitas vezes sem respeito à propriedade e aos direitos autorais. Como uma resposta a este problema, surgiram, em meados da década de 90, as primeiras propostas para inserção de marca d'água em

conteúdos digitais (Tirkel et al., 1993), a chamada *marca d'água digital*¹. Desde então, a quantidade de aplicações utilizando marca d'água digital vem crescendo de forma considerável, alavancando o desenvolvimento de novas técnicas e atraindo a atenção tanto de entidades acadêmicas quanto de empresas comerciais. Devido a seu surgimento relativamente recente, esta área de estudo encontra-se permeada por outras disciplinas, porém baseando-se em técnicas de processamento de sinais e criptografia e relacionando-se de forma estreita com as teorias da informação e da codificação.

Já encontramos sedimentada em nossa sociedade a idéia de propriedade de bens físicos, que em geral não é questionada e parece bastante natural. A *propriedade intelectual*² estende o conceito de propriedade a entidades que não pertencem ao mundo físico, tais como uma música ou uma invenção. O conceito de roubo ligado à propriedade intelectual também é diferente: a simples duplicação do conteúdo de um documento sobre o qual incidam direitos autorais já constitui violação de propriedade intelectual. Por ser muito fácil duplicar bens digitais, fazer com que a lei seja cumprida nesses casos é uma tarefa árdua. Pode-se citar um caso recente que causou choque entre detentores de propriedade intelectual: o caso do *Napster* que, permitindo a livre troca de músicas em formato MP3³ entre internautas, foi alvo de um processo judicial por violação de direitos autorais. O *Napster* acabou desativado, mas outros sistemas similares surgiram, e muitas vezes não limitados à troca de informação de áudio. Este cenário motivou diversas tentativas de desenvolver mecanismos de proteção para garantir aos autores o direito de propriedade. No conjunto de técnicas para proteção de direitos autorais, têm merecido muita atenção as baseadas em marca d'água digital.

A marca d'água digital foi proposta inicialmente como uma solução potencial no combate à pirataria. Ela consiste em adicionar uma marca no sinal que se deseja proteger, dito sinal *hospedeiro*. Esta marca não deve degradar a qualidade perceptiva do sinal protegido (propriedade de *transparência*), mas deve ser detectável e indelével. Para garantir transparência, a marca d'água pode utilizar informações perceptivas (psicoacústicas, no caso de um sinal de áudio, ou psicovisuais, no caso de uma imagem). Além disso, a marca d'água deve ser robusta, de modo a garantir sua permanência mesmo quando o sinal sofrer manipulações que de alguma forma o alterem.

O uso de uma marca d'água em um documento multimídia⁴ pode ser comparada a um processo de comunicação, no qual se pretende enviar uma mensagem (a marca d'água) a um receptor através de uma canal sujeito a ruído (o sinal a proteger). Neste sentido, conceitos da teoria da comunicação podem ser estendidos aos sistemas de marca d'água.

Não é certo, nem esperado, que o paradigma da marca d'água digital resolva todos os problemas

¹Em inglês, *digital watermarking*.

²No final do século XIX houve uma ampliação do conceito de propriedade para incluir também a propriedade intelectual.

³MP3 (MPEG 1 - Layer 3): sistema de compressão de áudio em formato digital.

⁴O conceito de multimídia compreende sinais de áudio, voz, imagem, vídeo e texto.

relacionados à proteção dos direitos autorais. No entanto, ele pode dar uma contribuição importante neste sentido, ajudando na identificação de documentos multimídia que foram sujeitos a cópia ou manipulações ilegais. Deve-se salientar que, além de aplicações ligadas à proteção de propriedade intelectual, a marca d'água pode ser utilizada para diversas outras finalidades, pois os dados inseridos através da marca podem ser de qualquer natureza.

O conceito de marca d'água digital pode confundir-se com outros três conceitos relacionados: esteganografia, impressão digital e criptografia. Por isso, faremos uma breve definição destes conceitos:

- *Esteganografia*: inserção de uma mensagem dentro de outra, visando manter secreta a própria existência da mensagem inserida. A mensagem “hospedeira” deve ser tal que não levante suspeitas quanto à existência de uma mensagem oculta.
- *Marca d'água*: inserção de uma mensagem dentro de outra, não sendo imprescindível que a existência da mensagem inserida seja secreta. O que caracteriza a marca d'água é a propriedade de *transparência*, isto é, a inserção da informação secreta não pode causar alterações que degradem perceptivelmente a mensagem hospedeira.
- *Impressão digital*: geração de um código de identificação (impressão digital) a partir de uma mensagem. Este código é uma espécie de *hash* que permite reconhecer a mensagem a partir da qual foi gerado. A impressão digital não causa alterações na mensagem, pois é armazenada separadamente desta.
- *Criptografia*: transformação da informação original para outra forma ilegível, necessitando que seja conhecida uma *chave secreta* para geração e extração da mensagem criptografada. A criptografia torna difícil a leitura da mensagem original por alguém não autorizado.

Apresentaremos neste capítulo os principais conceitos relacionados ao paradigma da marca d'água digital. Daremos ênfase, em alguns casos, às técnicas de marca d'água aplicadas em áudio digital. Tais conceitos serão explorados nos capítulos seguintes deste trabalho.

3.2 Aplicações

O ano de 1993 é considerado o marco que assinala o início das pesquisas em marca d'água digital (Brandão, 2002), embora antes dessa data algumas publicações tenham apresentado a idéia de inserir informação em um sinal multimídia para proteger os direitos de autor (Tanaka et al., 1990a,b). Desde então, a marca d'água digital vem recebendo atenção crescente da comunidade científica, o que tem levado a uma rápida evolução no desempenho das técnicas que envolvem esse conceito.

A motivação inicial para o avanço das técnicas de marca d'água digital foi a proteção de direitos autorais. No entanto, inúmeras outras aplicações têm sido sugeridas. Cada classe de aplicações requer uma marca com diferentes características. Apresentaremos de forma sintética alguns exemplos.

Proteção de Direitos Autorais (*copyright protection*) (Gomes et al., 2003): é a aplicação mais popular da marca d'água. Consiste em adicionar uma assinatura digital secreta a um sinal multimídia, com o objetivo de notificar direitos autorais. A marca deve ser robusta a ataques de adversários que tentem destruí-la. Conforme dito anteriormente, a proteção de direitos autorais alavancou as pesquisas de marca d'água digital no começo da década de 90.

Rastreamento de cópia não autorizada (*traitor tracing*) (Boneh and Shaw, 1998): análogo ao problema de proteção de direitos autorais, com uma diferença: a distribuição da mídia digital é limitada, e uma assinatura distinta é adicionada a cada cópia. Isto possibilita identificar o usuário original caso uma cópia não autorizada seja localizada. Tabelas logarítmicas do século XVII são um exemplo desta aplicação. Erros eram deliberadamente inseridos no último dígito significativo de algumas entradas das tabelas para garantir o direito de propriedade intelectual. Aplicações desta abordagem incluem distribuição de filmes digitalizados para cinemas, distribuição de material audiovisual em redes privadas e distribuição de documentos confidenciais em empresas e governos.

Legenda embutida (*in-band captioning*) (Moulinand and Koetter, 2005): vários tipos de dados podem ser adicionados a programas de áudio e vídeo usando técnicas de marca d'água. Por exemplo: legendas, letras de música, informações financeiras e outros dados quaisquer que possam ser de interesse dos espectadores. Neste caso, a marca d'água adiciona valor ao sinal e não impede a sua reprodução, de modo que não há interesse em apagá-la; portanto, a robustez a ataques intencionais não é exigida. Por outro lado, a marca deve ser robusta a ataques não intencionais (filtragem, modulação, ruído etc).

Autenticação forense (*media forensics*) (Kundur and Hatzinakos, 1999): aplicável sobretudo a perícias criminais. O objetivo é extrair informações sobre qualquer processamento que tenha sido efetuado sobre um sinal. Através de uma marca d'água, métodos forenses podem indicar quais partes de um sinal foram modificadas, identificando trechos introduzidos ou extraídos em um sinal (e.g. uma gravação de áudio).

Anotação de base de dados (*database annotation*) (Moulinand and Koetter, 2005): uma marca d'água pode ser usada em bases de dados audiovisuais que contenham vários tipos de informações associadas (e.g. texto, voz). É possível integrar essas informações com os arquivos audiovisuais inserindo-as diretamente nos próprios arquivos através de uma marca d'água, facilitando manipula-

ções e compartilhamento do conteúdo da base de dados. Informações de partituras e cenas de áudio ou imagens, por exemplo, podem ser adicionadas ao próprio sinal para auxiliar sistemas de busca na localização de informações.

Atualização de sistemas legados (*upgrade of legacy systems*) (Chen and Sundberg, 2000; Barros et al., 2004): é possível atualizar sistemas de transmissão já existentes através da adição de uma *camada* nos dados transmitidos, por meio de uma marca d'água. Exemplos incluem o sinal de crominância adicionado a um sinal de televisão preto e branco e sistemas de áudio digital na faixa de FM (Chen and Sundberg, 2000).

Identificação de conteúdo (*content identification*) (Moulinand and Koetter, 2005): adicionar uma marca d'água que identifique cenas em programas de televisão e rádio possibilita aplicações de monitoramento automático de conteúdo e observação de uso; por exemplo, pode-se avaliar o número de vezes que determinado comercial ou música foram tocados por uma rádio, ou a frequência com que uma propaganda política foi exibida na televisão.

Controle de dispositivos (*device control*) (Cox et al., 2002): vários tipos de informação de sincronismo e controle podem ser adicionados aos sinais de rádio e televisão. Um exemplo é a técnica de redução de ruído Dolby FM, usada por algumas estações de rádio comerciais e que requer o uso de um decodificador apropriado. Um sinal de marca d'água adicionado ao sinal do rádio é usado para *trigger* os receptores Dolby. Essa aplicação também pode ser classificada como atualização de sistemas legados.

Transmissão de informação secreta (Parhi and Nishitani, 1999): é possível estabelecer um processo de comunicação em que a marca d'água contém uma mensagem secreta e o sinal hospedeiro corresponde ao canal de comunicação. Com isso, é possível transmitir um sinal aparentemente inócuo (e.g. uma música) contendo informações secretas presentes na marca d'água. Aplicações deste tipo são comuns em uso militar.

Identificação de radiografias (DeFigueiredo, 2001): um sistema de marca d'água pode criar um vínculo definitivo entre a imagem obtida em uma radiografia e o respectivo paciente, prevenindo os sérios problemas que uma eventual troca poderia causar.

Existe assim um vasto conjunto de aplicações que utilizam marca d'água digital. Dependendo da aplicação, são exigidas diferentes características dos algoritmos de marca d'água. Na seção seguinte, são apresentados critérios para classificação de sistemas de marca d'água conforme essas características.

3.3 Classificação dos Sistemas de Marca d'Água

Diante das aplicações apresentadas e respectivos requisitos, as técnicas de marca d'água podem ser classificadas segundo diferentes critérios. São eles:

- **Transparência ou fidelidade:** a adição da marca d'água não pode causar degradação exagerada no sinal hospedeiro. A informação adicional deve ser invisível em imagens e inaudível em áudio e voz. No caso de sinais que não possuem interpretação sensorial, a potência da marca d'água deve ser tal que não interfira com a finalidade original do sinal hospedeiro; por exemplo, no caso de um sinal digital, a marca d'água não deve prejudicar a detecção desse sinal. Para cada tipo de aplicação existe um nível tolerável de degradação. Para garantirmos a inaudibilidade da marca d'água em sinais de áudio, faz-se necessário o uso de um modelo psicoacústico que explore as características do ouvido humano.
- **Robustez:** refere-se à capacidade do algoritmo de garantir a sobrevivência da marca d'água a operações comuns de processamento de sinais, tais como codificação e compressão (com ou sem perdas), filtragem, inserção de ruído aditivo e transmissão. Os algoritmos são em geral desenvolvidos de modo que a marca sobreviva a diferentes níveis de distorção, conforme os requisitos da aplicação. Se a marca pode ser facilmente removida, através de modificações impostas ao sinal marcado, ela é dita *frágil*. Já no caso em que a marca d'água persiste mesmo após a imposição de modificações significativas ao sinal, ela é dita *robusta*.
- **Segurança:** refere-se à resistência do sistema à decodificação ou remoção não autorizadas da informação contida na marca d'água. Ao atacar uma marca d'água, o objetivo de um pirata é tornar essa marca indetectável introduzindo pouca ou nenhuma degradação perceptível no sinal hospedeiro; num ataque ideal, o pirata recupera o sinal hospedeiro original, apagando qualquer traço da marca d'água. Em analogia com técnicas criptográficas, podem-se definir duas abordagens para algoritmos de marca d'água: sistemas *simétricos* e *assimétricos*. Quando a mesma chave⁵ é utilizada para a inserção e a detecção da marca d'água, o sistema é dito simétrico; caso sejam usadas chaves diferentes em cada etapa, o sistema é dito assimétrico. Sistemas assimétricos tendem a ser mais robustos a ataques intencionais visando a remoção da marca d'água; nestes sistemas, a chave empregada na geração do sinal de marca d'água (denominada *chave privada*) é mantida em segredo, sendo divulgada apenas uma chave (denominada *chave pública*) que permite a detecção da marca mas não contém informação suficiente para a sua remoção.

⁵Por “chave” entende-se um elemento a partir do qual é contruída a marca d'água ou que permite detectá-la. A chave pode ser, por exemplo, a semente de um algoritmo de geração de números pseudo-aleatórios, uma propriedade estatística da marca d'água, ou até mesmo o conjunto completo de amostras do sinal de marca d'água.

- **Detecção:** a detecção de uma marca d'água consiste em determinar se a marca encontra-se ou não presente em um sinal. Podem-se definir as seguintes probabilidades como indicadores de desempenho de um sistema: *probabilidade de detecção correta* é a probabilidade da decisão ser feita a favor da hipótese de que há uma marca quando esta de fato foi inserida, e *probabilidade de falsa detecção* é a probabilidade de ser feita uma decisão de que a marca existe quando na realidade ela não foi inserida. Idealmente, a probabilidade de detecção correta deveria ser igual a 1 e a probabilidade de falsa detecção deveria ser nula; na prática, há um compromisso entre essas duas probabilidades, devendo-se calibrar o sistema conforme as necessidades da aplicação.
- **Payload:** define-se *payload* como sendo o comprimento em bits da marca d'água. Essa característica está relacionada à quantidade de informação inserida no sinal hospedeiro. O *payload* pode variar de acordo com a aplicação, podendo variar de apenas alguns bits (e.g. aplicações para proteção de direitos autorais), até vários megabytes (tal como em aplicações de comunicação secreta).
- **Custo computacional:** como em qualquer tecnologia utilizada comercialmente, o custo computacional para inserção e detecção da marca d'água é importante, tornando-se crítico quando o processamento deve ser efetuado em tempo real. Este fator deve portanto ser levado em conta na escolha da técnica mais apropriada a uma dada aplicação.
- **Uso de informação do sinal hospedeiro:** um sistema de marca d'água pode ou não utilizar informação do sinal hospedeiro na construção da marca. O ajuste da marca às características do sinal hospedeiro permite potencialmente maximizar a quantidade de informação inserida no sinal. Quando não é utilizada informação do sinal hospedeiro, o sistema é dito de detecção cega; caso contrário, o sistema é dito de detecção não-cega.

Existem compromissos entre as características citadas. Grande parte do trabalho de desenvolvimento de um sistema de marca d'água consiste em analisar esses compromissos, identificando requisitos da aplicação e desenvolvendo algoritmos práticos que atendam a esses requisitos.

3.4 Estrutura Geral de Sistemas de Marca d'Água

3.4.1 Elementos Básicos do Sistema

Um sistema genérico de marca d'água consiste de dois blocos básicos: um módulo de *inserção*, que permite gerar a marca d'água a partir de dados que se deseja ocultar e inseri-la no sinal hospedeiro, e um módulo de *detecção*, responsável pela recuperação dos dados ocultos a partir do sinal

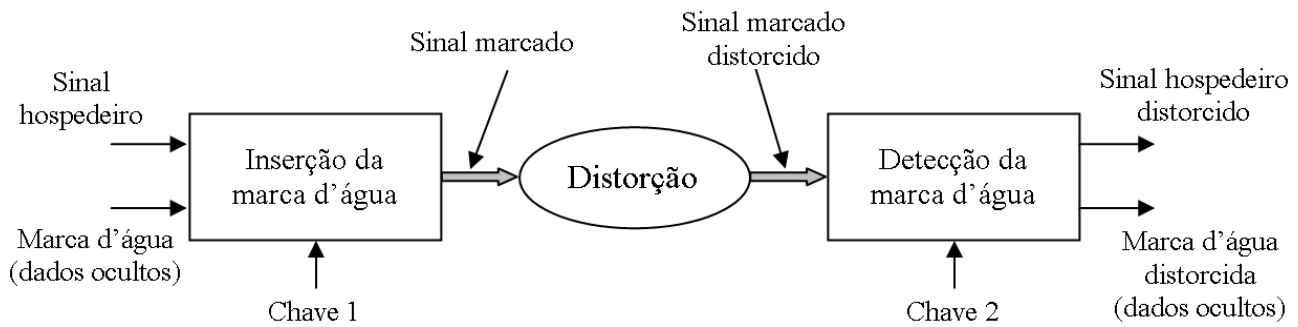


Fig. 3.1: Esquema genérico de um sistema de marca d'água.

hospedeiro marcado. Apesar de o sinal hospedeiro ser modificado com a inserção da marca d'água, esta modificação deve ser tal que não afete a finalidade original do sinal hospedeiro; no caso de sinais com interpretação sensorial (e.g. áudio, imagem), podem ser usados modelos perceptivos para garantir a transparência da marca. Este processo é ilustrado na figura 3.1.

Uma chave (chave 1) é usada para gerar a marca d'água, enquanto uma outra chave (chave 2) é usada para detectá-la. Se as duas chaves são idênticas, o esquema é dito *simétrico*; caso contrário, o esquema é dito *assimétrico* (Gomes et al., 2000).

Um esquema assimétrico admite detecção pública e privada. A detecção privada requer conhecimento da chave 1 (chave privada) e a detecção pública requer conhecimento da chave 2 (chave pública, normalmente um subconjunto da chave privada). Um esquema simétrico somente permite detecção privada.

Na detecção, o sistema pode primeiramente checar a existência da marca d'água (detecção binária). Caso a marca esteja presente, passa-se à extração da informação (dados ocultos) presente na marca d'água. A detecção binária e a extração dos dados ocultos da marca podem ser privadas ou públicas. O sinal marcado pode ainda conter simultaneamente dados privados e públicos, acessados através do uso da chave privada e da chave pública, respectivamente. Devido à chave pública ser usualmente um subconjunto da chave privada, esta última é em geral suficiente para acessar ambas as classes de dados.

Como na esteganografia, uma marca d'água insere uma mensagem no sinal hospedeiro. Porém, no paradigma da marca d'água, diferentemente da esteganografia, não é necessário que a presença de informações ocultas seja secreta; o primordial é a manutenção da qualidade do sinal hospedeiro, que tem valor próprio. Por exemplo, ao inserir uma marca d'água em uma música, é importante que a qualidade sonora não seja afetada, o que é possível com o auxílio de técnicas psicoacústicas.

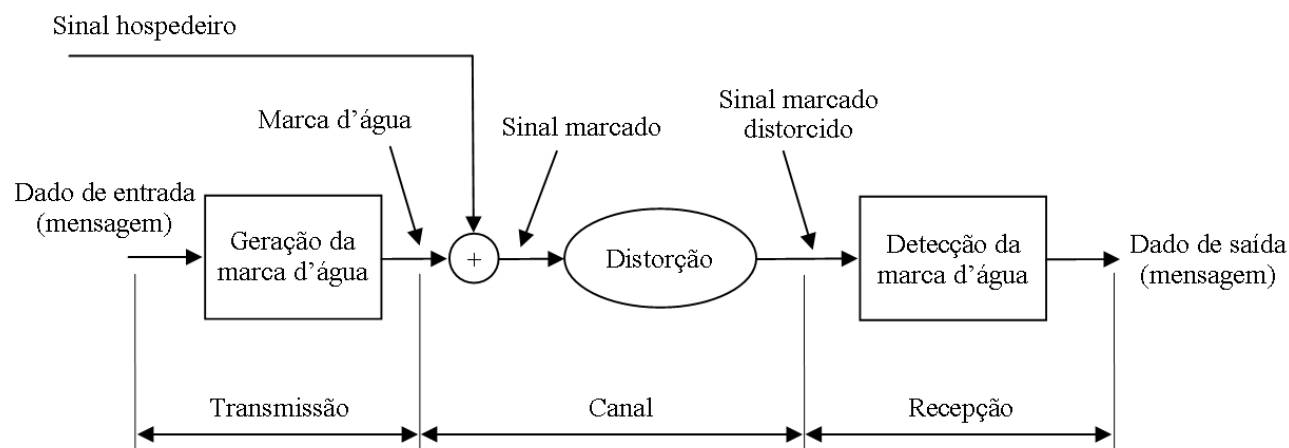


Fig. 3.2: Marca d'água como um sistema de comunicação.

3.4.2 Marca d'Água como um Sistema de Comunicação

Um algoritmo de marca d'água pode ser encarado como um sistema de comunicação (Garcia, 1999): nessa analogia, a marca d'água corresponde ao sinal transmitido, enquanto o sinal hospedeiro (juntamente com as distorções impostas ao sinal marcado) é visto como um ruído que, em geral, tem uma potência muitas vezes maior do que a da própria marca d'água. Esta abordagem é ilustrada na figura 3.2.

Neste modelo, a mensagem a ser transmitida é convertida no sinal de marca d'água por um processo de modulação. Em seguida, a marca d'água é adicionada ao sinal hospedeiro, gerando o sinal marcado. Este processo pode ser modelado como uma adição de ruído ao sinal de marca d'água. O sinal marcado pode então sofrer uma distorção devido aos efeitos do canal, gerando o sinal marcado distorcido. Os processos de adição de ruído e distorção são encarados como um canal de transmissão hostil. Por fim, o sinal marcado distorcido é entregue ao receptor, que tentará detectar a mensagem transmitida.

Em um sistema convencional de comunicação, o sinal transmitido é geralmente mais potente do que o ruído, e este último é muitas vezes assumido gaussiano e branco⁶. Porém, essas considerações não são verdadeiras em um sistema de marca d'água. Para que a marca não degrade demasiadamente o sinal hospedeiro, é essencial que ela seja muitas vezes mais fraca do que este em potência. Além disso, o sinal hospedeiro pode ser não-estacionário e colorido⁷, como no caso de sinais de áudio.

Quando um sistema de marca d'água é visto como um sistema de comunicação, três fontes de ruído podem ser identificadas:

- o sinal hospedeiro, bem mais potente do que a marca d'água; para aplicações de áudio, essa

⁶Sinal dotado de um conteúdo uniforme e de extensão infinita em frequência.

⁷Sinal não dotado de um conteúdo uniforme em frequência.

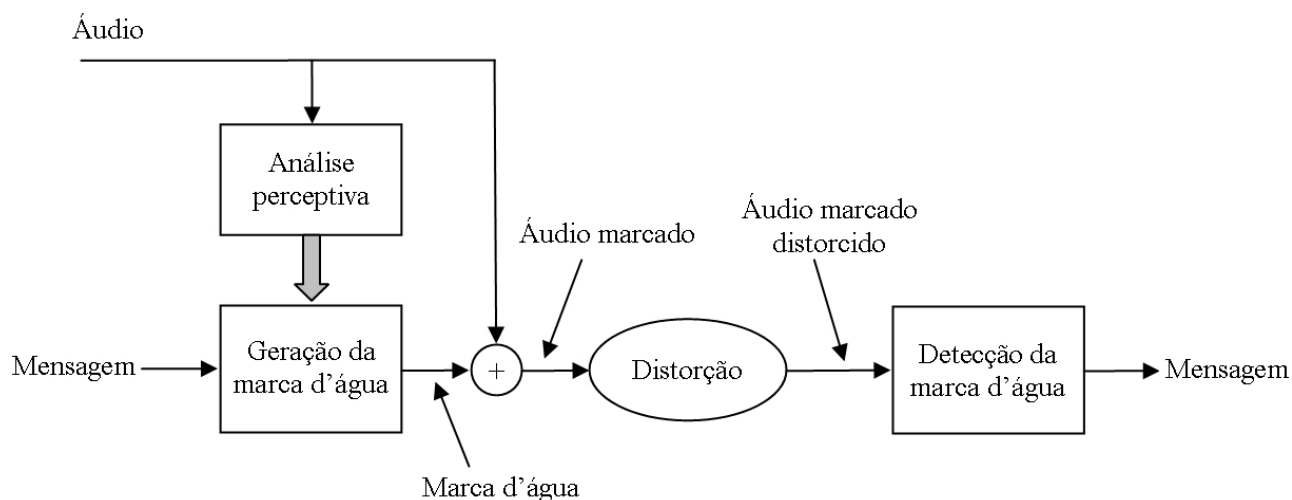


Fig. 3.3: Esquema de marca d'água digital em áudio.

relação é da ordem de 20 dB (Gomes et al., 2000).

- distorções introduzidas por operações lícitas, como compressão/descompressão, filtragem e ruído do canal de transmissão.
- distorções introduzidas por operações ilícitas, tais como ataque de piratas (hackers) que tentam tornar a marca indetectável.

O sinal hospedeiro não marcado em geral é conhecido apenas no transmissor; o receptor normalmente tem acesso apenas ao sinal marcado.

3.4.3 Marca d'Água em Áudio

O primeiro artigo de que temos conhecimento abordando marca d'água digital em áudio foi proposto em 1996 (Boney et al., 1996, 1997). O esquema geral da técnica proposta encontra-se na figura 3.3.

A idéia consiste em adicionar ao sinal de áudio uma seqüência pseudo-aleatória de espectro branco, formada a partir de uma mensagem que se deseja ocultar. Para garantir inaudibilidade e buscando atingir o maior nível possível de potência para a marca d'água, exploram-se conceitos relativos ao sistema auditivo humano. A seqüência pseudo-aleatória passa por um processo de conformação espectral guiado pelo *limiar de mascaramento* do sinal de áudio, obtido por meio de um *modelo psicoacústico*.

Experimentos psicoacústicos mostraram que o sistema auditivo humano não consegue perceber um tom ou sinal de banda estreita na presença de outro sinal suficientemente forte e próximo em

freqüência. Isto é o chamado *efeito de mascaramento* (Moore, 1989; Humes and Jesteadt, 1989). Quando generalizado para sinais de banda larga, o efeito de mascaramento tonal ou de banda estreita é substituído pelo *limiar de mascaramento*, correspondente a uma curva análoga a uma densidade espectral de potência. Caso a densidade espectral de potência de um sinal de áudio esteja abaixo do limiar de mascaramento de outro sinal de áudio para todas as freqüências, o primeiro sinal é *mascarado* pelo segundo, e sua presença não é detectada pelo ouvido humano quando ambos os sinais são reproduzidos simultaneamente. Os algoritmos para cálculo do limiar de mascaramento associado a um sinal de áudio são chamados de *modelos psicoacústicos*.

Modelos psicoacústicos são usados em codificadores de áudio para atingir maiores níveis de compressão sem abrir mão da qualidade de áudio. Contanto que a densidade espectral de potência do ruído de quantização esteja abaixo do limiar de mascaramento, nenhuma distorção será percebida pelo ouvinte. Sinais de áudio são em geral não estacionários; por isso, os modelos psicoacústicos são aplicados a janelas do sinal de áudio, normalmente de algumas dezenas de milissegundos. O mesmo princípio pode ser usado para garantir a inaudibilidade da marca d'água: se, para cada janela, a densidade espectral de potência da marca d'água estiver abaixo do limiar de mascaramento do sinal de áudio, a marca não será percebida na presença desse sinal. Detalhes sobre fenômenos auditivos e modelos perceptivos serão apresentados no apêndice A.

A detecção da marca d'água em sinais de áudio é geralmente realizada através de correlação da marca d'água original com a marca distorcida, produzindo um coeficiente de correlação que é comparado a um limiar específico.

Tomemos como exemplo de marca d'água em áudio um sistema implementado na Universidade Paris V em colaboração com a ENST-Paris (Gomes et al., 2000), ilustrado na figura 3.4. O objetivo é transportar informação entre dois computadores através de um canal acústico, com a informação dissimulada em um sinal de música. No primeiro computador, um texto é convertido em informação binária e adicionado a um arquivo digital de áudio através de uma marca d'água. O arquivo de áudio é reproduzido em um alto-falante, sendo recebido por um segundo computador através de um microfone. É então realizado o processo de detecção da marca d'água, extraindo-se o texto a partir do qual ela foi gerada e exibindo-o na tela do computador. Este exemplo ilustra o potencial de uso da marca d'água em áudio.

3.5 Métodos para Inserção de Marca d'Água

Apresentaremos a seguir uma revisão das técnicas mais comuns de marca d'água digital. Estas técnicas não são específicas para determinados usos, podendo-se aplicar a diversas modalidades de sinais hospedeiros. Além de serem os métodos mais freqüentemente utilizados, eles apresentam

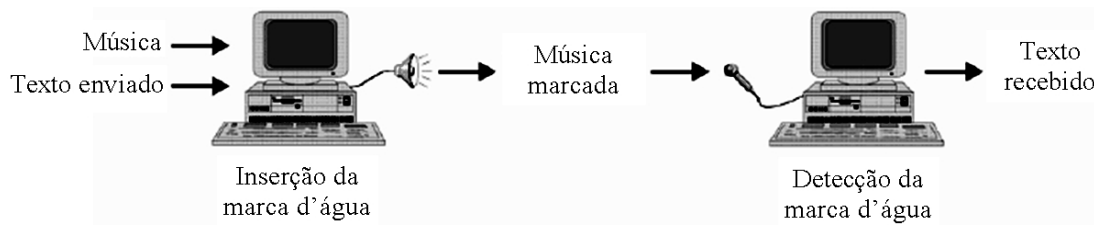


Fig. 3.4: Transporte de informação através de um canal acústico.

características importantes que serão exploradas no decorrer deste trabalho. Apesar de apenas a técnica de espalhamento espectral ser utilizada no escopo deste trabalho, buscaremos dar uma idéia ao leitor de diferentes abordagens encontradas na literatura.

Serão apresentados em detalhe dois métodos de marca d'água: o método de espalhamento espectral, que visa espalhar a marca ao longo de todo o espectro do sinal hospedeiro, e o método de quantização, que quantiza os valores das amostras do sinal hospedeiro em função do sinal de marca d'água. Por fim, serão apresentados brevemente outros métodos encontrados na literatura.

3.5.1 Espalhamento Espectral

O método de marca d'água baseado em espalhamento espectral foi proposto por Cox et al. (Cox et al., 1997). Este é o método de marca d'água de uso mais difundido e vem sendo bastante estudado na literatura (Cox et al., 1997; Swanson et al., 1998; Kim, 2000; Lee and Ho, 2000; Cvejic et al., 2001; Gomes et al., 2003). O método é eficiente e de implementação relativamente simples, mas apresenta a desvantagem de ser suscetível a distorções temporais.

A idéia básica do método de espalhamento espectral consiste em adicionar uma seqüência pseudo-aleatória de banda larga ao sinal hospedeiro. Esta adição pode se dar no domínio do tempo ou no domínio de uma transformada do sinal; neste último caso, são comuns as transformadas DCT (transformada cosseno discreta), DFT (transformada de Fourier discreta) e DWT (transformada wavelet discreta). Um esquema típico de inserção de marca d'água utilizando espalhamento espectral é ilustrado na figura 3.5. A mensagem binária da marca d'água $b(n) = \{-1, 1\}$ é modulada por uma seqüência pseudo-aleatória $r(n)$ gerada através de uma chave secreta. A marca d'água modulada $w(n) = b(n)r(n)$ é escalonada de acordo com a energia do sinal hospedeiro $s(n)$. O fator de escala α controla a relação entre robustez e transparência da marca d'água. A marca modulada $w(n)$ é igual a $r(n)$ ou $-r(n)$, dependendo se $b = 1$ ou $b = -1$. O sinal modulado é adicionado ao sinal hospedeiro para produzir o sinal marcado $x(n)$, seguindo a relação:

$$x(n) = s(n) + \alpha w(n) \quad (3.1)$$

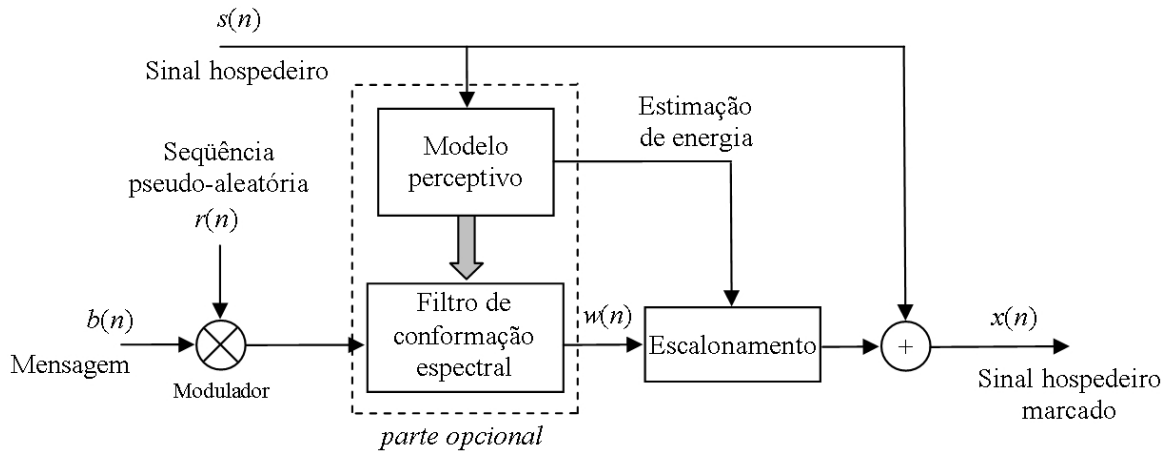


Fig. 3.5: Esquema de inserção de marca d'água utilizando espalhamento espectral.

A sequência pseudo-aleatória deve ter propriedades estatísticas similares às de um sinal realmente aleatório, podendo ser gerada a partir de uma semente ou chave criptográfica. Um algoritmo eficiente para geração de sequências pseudo-aleatórias faz com que duas sequências geradas a partir de sementes diferentes sejam praticamente ortogonais entre si, isto é, o valor da correlação entre elas é muito baixo. Além disso, o valor da autocorrelação de uma sequência deve ser próximo de um impulso.

Uma sequência pseudo-aleatória bastante comum é conhecida como sequência de comprimento máximo⁸ (MLS) ou sequência-m⁹. Esta é uma sequência binária $r(n) = \{-1, 1\}$ com tamanho¹⁰ N . Esta sequência tem propriedades de autocorrelação e correlação cruzada satisfatórias, sendo a autocorrelação dada por:

$$\frac{1}{N} \sum_{i=0}^{N-1} r(i) r(i-k) = \begin{cases} 1 & \text{se } k = 0 \\ -1/N & \text{outro valor} \end{cases} \quad (3.2)$$

O uso de sequências-m apresenta duas desvantagens. Primeiro, o tamanho da sequência-m é limitado e dado por $2^m - 1$. Na prática, esta dependência pode levar a sequências relativamente curtas (e.g. com tamanho $N = 31$ para $m = 5$), que podem apresentar baixa ortogonalidade entre si, uma vez que se afastam das características de um sinal branco. O tamanho normalmente utilizado na literatura e que apresenta bons resultados é 1.023 ou 2.047 (Cvejic et al., 2001). Janelas deste tamanho preservam as características espectrais e temporais do sinal onde a marca d'água será inserida. A segunda desvantagem é o número limitado de diferentes sequências-m quando m é determinado. Isso

⁸Em inglês, *maximum length sequence*.

⁹Em inglês, *M-sequence*.

¹⁰O processo de geração da sequência-m utiliza um registrador de deslocamento com m elementos realimentados. Para um dado elemento de entrada, o registrador de deslocamento produz uma nova saída. O número máximo de elementos produzidos pela sequência-m é $N = 2^m - 1$. Detalhes são encontrados em (Golomb, 1967; Cohn and Lempel, 1977).

demonstra que a seqüência pode ser reproduzida utilizando técnicas de testes exaustivos.

A adição de uma seqüência pseudo-aleatória a um sinal de áudio, por exemplo, pode causar distorções perceptíveis. Apenas reduzir o fator α da seqüência pseudo-aleatória pode não resolver o problema de transparência, ou seja, um ruído, mesmo com pequenas amplitudes, pode ser perceptível. Além disso, baixos valores do fator α tornam o método de espalhamento espectral menos robusto, pois diminui-se o valor da autocorrelação na presença do sinal hospedeiro. Uma solução para se garantir a inaudibilidade da marca d'água é conformá-la espectralmente segundo um limiar de mascaramento produzido por um modelo psicoacústico (Boney et al., 1996; Bassia et al., 2001; Gomes et al., 2003), conforme marcado com uma linha tracejada na figura 3.5. A conformação espectral da marca d'água possibilita aumentar a robustez do sistema, permitindo maximizar o fator de escala α e aproximar do limiar de audibilidade.

Um som afeta a percepção de sons vizinhos no tempo ou em freqüência. Esse fenômeno é conhecido como mascaramento. O modelo psicoacústico analisa o sinal de áudio $s(n)$ para calcular o limiar de mascaramento correspondente. A marca d'água torna-se inaudível através da sua conformação espectral e posterior escalonamento, garantindo que a densidade espectral de potência da marca d'água fique abaixo do limiar de mascaramento do sinal de áudio. Detalhes referentes aos fenômenos auditivos e modelos perceptivos encontram-se no apêndice A.

A detecção no método de espalhamento espectral baseia-se em correlação linear. A seqüência $r(n)$ é conhecida ou pode ser gerada no receptor. A marca d'água é detectada através da correlação entre $x(n)$ e $r(n)$, da seguinte forma:

$$c = \frac{1}{N} \sum_{i=1}^N x(i) r(i) \quad (3.3)$$

onde N denota o tamanho do sinal. A equação 3.3 pode ser expandida na soma de duas componentes:

$$c = \frac{1}{N} \sum_{i=1}^N s(i) r(i) + \frac{1}{N} \sum_{i=1}^N \alpha b(i) r^2(i) \quad (3.4)$$

Assumimos que o primeiro termo na equação 3.4 tem baixas magnitudes, pois os sinais não devem apresentar correlação entre si. Idealmente, este termo deve desaparecer¹¹. O sinal marcado é processado como mostrado na figura 3.6, assumindo válida essa suposição. Outros métodos de detecção utilizam filtragem passa-altas (Hartung and Girod, 1998), codificação preditiva linear (Seok et al., 2002) e técnicas de branqueamento (Kim, 2000) para tornar os sinais $x(n)$ e $r(n)$ independentes entre si e aumentar a robustez do sistema.

¹¹No caso de sinais de áudio, por exemplo, pode ser usada uma técnica de conformação espectral segundo um limiar de mascaramento, o que vai introduzir um certo grau de correlação entre $s(n)$ e $r(n)$; isto será abordado no capítulo 7.

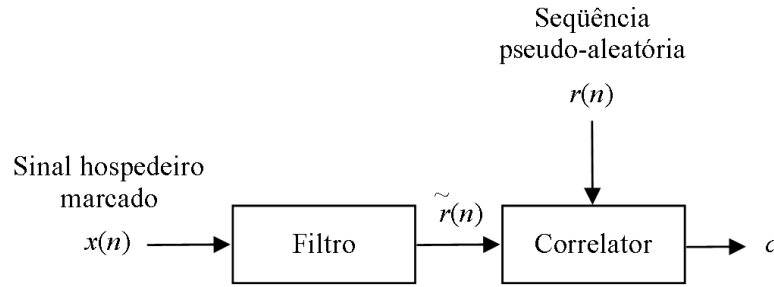


Fig. 3.6: Pré-processamento para detecção em um esquema de marca d'água utilizando espalhamento espectral.

3.5.2 QIM (*Quantization Index Modulation*)

O método QIM (Chen and Wornell, 1999, 2001) é um esquema de quantização escalar. Ele consiste em quantizar o valor x de uma amostra de sinal determinando um valor y baseado na seguinte regra:

$$y = \begin{cases} q(x, D) + \frac{D}{4} & \text{se } b = 1 \\ q(x, D) - \frac{D}{4} & \text{se } b = 0 \end{cases} \quad (3.5)$$

onde $q(\cdot)$ é a função de quantização, D é o passo de quantização, e b é o valor binário da marca d'água que se deseja associar à amostra do sinal. A função de quantização $q(x, D)$ é dada pela seguinte relação:

$$q(x, D) = \left[\frac{x}{D} \right] D \quad (3.6)$$

onde o operador $[\cdot]$ indica o arredondamento para o inteiro mais próximo do argumento. Este esquema de quantização é ilustrado na figura 3.7. O valor x da amostra é inicialmente quantizado para $q(x, D)$, indicado por um círculo preto ($\bullet$) na figura. Se o bit da marca d'água é 1, o valor da amostra é movido para a posição indicada pelo círculo branco ($\circ$); caso contrário, o valor da amostra é movido para a posição indicada pelo símbolo $\times$. Por exemplo, seja $D = 8$ e $x = 81$. Então, $q(81, 8) = 80$. Se $b = 1$, tem-se $y = 82$; se $b = 0$, tem-se $y = 78$. A distância entre as referências é igual a $D/2$.

O processo de detecção da marca d'água pode ser resumido da seguinte forma:

$$b = \begin{cases} 1 & \text{se } 0 < y - q(y, D) < D/4 \\ 0 & \text{se } -D/4 < y - q(y, D) < 0 \end{cases} \quad (3.7)$$

O método é simples de implementar e robusto contra ataques que apresentem uma margem de distorção abaixo de $D/4$. Se a distorção ultrapassar $D/4$, o valor quantizado é perturbado e o detector pode interpretar incorretamente o bit da marca d'água.

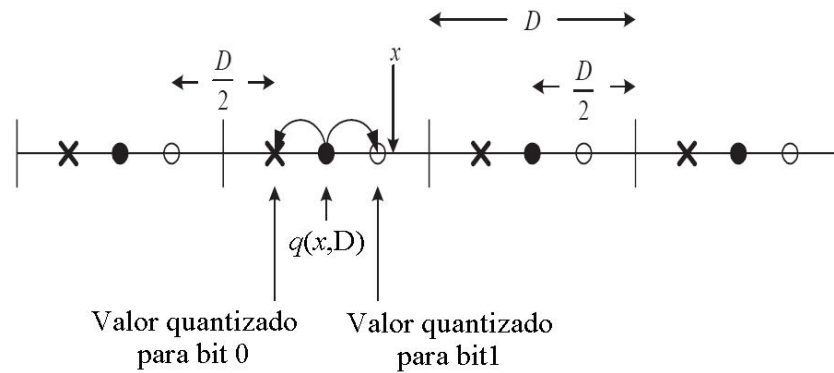


Fig. 3.7: Esquema de quantização QIM.

3.5.3 Método da Réplica (*Replica Method*)

O próprio sinal original pode ser usado como uma marca d'água. A técnica de eco oculto¹² (Bender et al., 1996) adiciona dados no sinal original através de um eco do próprio sinal no domínio do tempo, ou seja, uma cópia de menor amplitude do sinal, defasada no tempo, é adicionada explorando os efeitos do mascaramento temporal. A informação binária da marca d'água é adicionada introduzindo-se ecos com diferentes atrasos. A extração da informação é obtida através da determinação deste atraso, o que pode ser conseguido através do cálculo do cepstrum do sinal marcado. A técnica de modulação da réplica¹³ (Petrovic, 2001) adiciona ao sinal original, no domínio da frequência, parte de si mesmo, explorando os efeitos do mascaramento freqüencial. Por exemplo, cópias de faixas de baixa frequência (e.g. de 1 a 4 kHz) são moduladas (e.g. com deslocamento de 20 Hz e adequado ajuste do fator de escala), e inseridas abaixo do limiar de mascaramento do sinal original (e.g. variando na faixa compreendida entre 1020 e 4020 Hz). No detector, é gerado a mesma réplica a partir do sinal marcado e calculado correlações para verificar a presença da marca d'água.

3.5.4 Método de Auto-Marcação (*Self-Marking Method*)

Métodos de auto-marcação normalmente deixam evidente a presença da marca d'água junto ao sinal original. Uma das técnicas consiste em inserir picos no domínio da frequência, explorando efeitos de mascaramento nas regiões em que a densidade espectral de potência do sinal original apresenta maiores valores (Mansour and Tewfik, 2001). Outra técnica, específica para sinais de áudio, insere distorções temporais, introduzindo amostras no domínio do tempo entre marcas de *pitch*. O intervalo entre dois picos de *pitch* são particionados em N elementos, e amostras são introduzidas com inclinação (derivada) correspondente à informação binária que se deseja adicionar (Wu et al., 2000).

¹²Em inglês, *echo hiding*.

¹³Em inglês, *replica modulation*.

3.6 Ataques

Existem duas categorias de ataques a que um sistema de marca d'água está sujeito. Uma das categorias, chamada *não intencional*, relaciona-se a distorções e ruídos advindos de operações lícitas, tais como compressão, reconstrução, transmissão, modulação e atrasos. Muitos destes ataques são conhecidos previamente e o sistema de marca d'água deve ser robusto a eles se a aplicação assim o exigir. A outra categoria de ataque é chamada de *intencional*; neste caso, existe um hacker (ou pirata) cujo intuito é inutilizar a marca d'água, de forma que o sistema não possa mais detectá-la. Um exemplo de aplicação sujeita a ataques intencionais é a proteção contra pirataria.

Existem várias formas de ataques intencionais que têm por objetivo remover a marca d'água de um sinal. Porém, as distorções provocadas por tais ataques, além de baixarem a probabilidade de detecção correta da informação contida na marca d'água, introduzem degradação no sinal hospedeiro. O objetivo dos ataques é tornar a marca indetectável sem que a degradação resultante inutilize o sinal hospedeiro. Uma técnica de marca d'água robusta a ataques intencionais, por sua vez, propõe-se a introduzir no sinal uma marca que somente se torne indetectável quando o nível de distorção provocado pelo ataque for inaceitável.

Algumas técnicas de ataque tentam confundir o sistema de detecção da marca d'água, provendo a existência de múltiplos sinais da marca d'água original. Outra modalidade de ataque tenta falsificar a existência da marca d'água em sinais não marcados, valendo-se de informações de sinais marcados, sem o conhecimento da chave que permite a inserção da marca d'água. Esses dois exemplos ilustram diferentes técnicas empregadas com o intuito de destruir sistemas que utilizam marca d'água digital. Discutiremos a seguir algumas técnicas de ataque que tentam remover, ou enganar, de forma intencional, um sistema de marca d'água.

- **Distorção:** estes ataques procuram reduzir a probabilidade de detecção da marca d'água através da distorção do sinal marcado. A adição de ruído ao sinal marcado, por exemplo, pode prejudicar a detecção da marca por correlação. Também a filtragem do sinal marcado, colorindo o sinal da marca d'água, pode dificultar a detecção da marca em sistemas baseados em espalhamento espectral.
- **Média estatística:** um pirata pode tentar estimar o sinal de marca d'água e subtraí-lo do sinal marcado, obtendo um sinal próximo do sinal hospedeiro sem marca d'água. Esse tipo de ataque é particularmente perigoso quando é utilizada uma marca d'água que faça pouco ou nenhum uso de informações do sinal hospedeiro. Nesse caso, se o pirata dispuser de uma quantidade suficiente de sinais hospedeiros (independentes entre si) nos quais foi inserida a mesma marca d'água, esta última pode ser estimada calculando-se a média entre esses sinais. Tendo em mãos

a estimativa da marca d'água, é possível extraí-la de qualquer sinal marcado através da mesma técnica de inserção.

- **Colusão:** em algumas aplicações, diferentes marcas d'água podem ser usadas em cópias de um mesmo sinal hospedeiro. Este é o caso, por exemplo, quando um vendedor deseja controlar a distribuição de um produto para diferentes clientes, inserindo números de lote específicos para cada conjunto de cópias. Caso um pirata tenha acesso a cópias do sinal contendo marcas d'água diferentes, ele pode procurar enfraquecer os sinais de marca d'água através do cálculo da média entre essas cópias, eventualmente tornando as marcas indetectáveis.
- **Falsa marca d'água:** essa técnica de ataque tenta falsificar uma marca d'água, imitando a presença da marca no sinal hospedeiro. Esse ataque tenta confundir o sincronismo do sistema, procurando forçá-lo a detectar uma falsa cópia da marca d'água com o intuito de inviabilizar a extração da informação contida na marca.
- **Exploração exaustiva do detector:** quando um pirata tem livre acesso a um detector de marca d'água, ele pode alterar características do sinal marcado à exaustão e observar a decisão do detector, prosseguindo com o procedimento até que a marca não seja mais detectada. Este procedimento pode mesmo ser realizado de forma automatizada.

3.7 Sincronismo

Em geral, o processo de detecção da marca d'água requer o alinhamento da marca no detector, pois a perda de sincronismo pode causar erros ou mesmo inviabilizar a detecção. Modificações aplicadas ao sinal marcado nos domínios de tempo ou frequência podem fazer com que o detector perca o sincronismo. Explorando esse fato, diversos ataques procuram tornar a marca indetectável por meio da sua dessincronização.

Muitos algoritmos de marca d'água assumem que o detector está sincronizado com a marca antes da detecção; neste caso, se o sistema deve ser robusto a perdas de sincronismo, métodos de resincronização eficientes fazem-se necessários. Alguns esquemas de marca d'água, como o método de réplica, são implicitamente mais robustos a ataques de dessincronização. Já os métodos com detecção baseada em correlação, como o espalhamento espectral, são fortemente dependentes do alinhamento correto da marca no detector.

No entanto, desenvolver um esquema de sincronismo eficiente não é uma tarefa trivial. Conforme a aplicação, tal esquema deve ser capaz de combater ataques intencionais de piratas visando à dessincronização da marca; além disso, ataques não intencionais decorrentes de operações lícitas

(codificação, filtragem e outras) também podem provocar perda de sincronismo. Sistemas de sincronismo eficientes devem atuar de forma suficientemente rápida e ser robustos a tais ataques.

Existem dois problemas distintos na sincronização. O primeiro refere-se à localização do ponto de início de um bloco de sinal de marca d'água. O segundo corresponde a modificações nos domínios de tempo e frequência provocadas por ataques intencionais ou não intencionais (Petrovic et al., 1999). Distorções temporais podem ser provocadas, por exemplo, pela adição ou subtração de amostras do sinal marcado (Petitcolas et al., 1998); através dessas modificações, o tamanho do sinal marcado é incrementado ou decrementado.

Uma das técnicas mais comuns de sincronização em comunicação digital baseia-se no uso de seqüências de treinamento, i.e. seqüências de dados que são conhecidas no transmissor e no receptor. As seqüências de treinamento original e recebida são alinhadas no receptor, possibilitando a resincronização do sistema. A mesma idéia pode ser aplicada à marca d'água: partes da marca podem ser compostas por símbolos conhecidos, correspondendo a seqüências de treinamento inseridas periodicamente na marca. Quando o sistema perde o sincronismo, o processo de detecção é prejudicado somente até que a próxima seqüência de treinamento seja localizada.

Outra forma de sincronização em sistemas de marca d'água consiste em espalhar uma seqüência de treinamento ao longo de todo o sinal marcado, habilitando o detector a rastrear continuamente o sincronismo. Devido à seqüência de treinamento estar presente todo o tempo, o sistema é capaz de resincronizar rapidamente em caso de perda de sincronismo. Este esquema pode ser implementado por meio da inserção de uma segunda marca d'água no sinal hospedeiro, usada somente para propósitos de sincronização; a primeira marca, inserida no mesmo sinal hospedeiro, carrega a informação que se deseja de fato transmitir. Para evitar interferência entre as marcas, elas podem ser criadas de forma a serem ortogonais. Outra maneira de espalhar uma seqüência de treinamento ao longo do sinal marcado consiste em utilizar vários *codebooks* para codificar a informação contida na marca d'água. Esses *codebooks* são usados consecutivamente, criando uma seqüência de *codebooks* que corresponde à regra da seqüência de treinamento (Gomes et al., 2001). Esta última forma de sincronização será explorada no capítulo 6.

3.8 Discussão

Neste capítulo, apresentamos os principais conceitos e características relacionados ao paradigma da marca d'água digital. Iniciamos a exposição com um breve histórico da marca d'água e apresentamos suas principais aplicações. Em seguida, vimos as principais características e estruturas de um sistema de marca d'água, ressaltando os requisitos das diversas aplicações e citando algumas das técnicas de marca d'água que atendem a esses requisitos. Por fim, apresentamos alguns tipos de ataques

visando à inutilização da marca d'água e discutimos problemas de sincronismo na detecção da marca. Apresentaremos a seguir algumas conclusões gerais a respeito de sistemas de marca d'água:

- Técnicas robustas para inserção da marca d'água devem ser utilizadas para aumentar o desempenho da comunicação frente a operações lícitas (e.g. filtragem).
- A técnica de espalhamento espectral é a mais popular, embora apresente algumas limitações (e.g. suscetível a modificações no domínio temporal).
- Muitos pesquisadores propõem o desenvolvimento de sistemas práticos, cujo nível de robustez depende da aplicação. Ataques com alto nível de sofisticação podem ocorrer quando existe um adversário inteligente (o pirata) que deseja inutilizar a marca. Porém, muitas aplicações não estão sujeitas a tais ataques; nesse caso o método de marca d'água não necessita de um grau elevado de robustez, embora possa ser necessária resistência a ataques não intencionais (codificação, filtragem etc).
- Um método robusto de sincronismo pode ser obtido com o uso de seqüências de treinamento e detecção por correlação; este método pode ser implementado, por exemplo, através de uma segunda marca d'água inserida no sinal hospedeiro.

Os conceitos introduzidos neste capítulo serão explorados no decorrer da dissertação. Apresentaremos nos próximos capítulos uma proposta de aplicação de marca d'água digital para equalização e identificação de canais explorando a técnica de espalhamento espectral, com ênfase em sinais digitais de telecomunicações e de áudio.

Capítulo 4

Equalização Supervisionada Utilizando Marca d'Água

4.1 Introdução

Será apresentado neste capítulo um método de equalização baseado no uso de uma *marca d'água* como sinal de referência¹. A marca d'água é conhecida no receptor, sendo utilizada para estimar os parâmetros do equalizador. Tanto a marca d'água como o sinal de informação nunca são interrompidos, evitando períodos sem transmissão de informação útil e permitindo que os coeficientes do equalizador sejam continuamente atualizados. Neste esquema, parte da potência do sinal de transmissão deve ser destinada à marca d'água, o que implica em uma redução da potência do sinal de informação. O sucesso desta abordagem depende de uma adequada distribuição de potências, refletida na relação sinal-marca d'água² (SWR).

O esquema que mais se assemelha à nossa proposta é a técnica conhecida como *superimposed pilot training*, que consiste em sobrepor uma seqüência de dados piloto conhecida ao sinal de informação. Esta técnica diferencia-se da nossa proposta por não estar focada na transparência da seqüência de treinamento, que é um dos objetivos básicos de um algoritmo de marca d'água. Além disso, não observamos nenhum trabalho em *superimposed training* que aborde o problema de equalização, estando esta técnica restrita ao caso de identificação até o presente momento. Outro diferencial deste trabalho é a abordagem genérica dada ao método de equalização aqui proposto, demonstrando o seu potencial uso em outras aplicações além da transmissão de dados (e.g. aplicações de áudio e voz).

¹O uso da marca d'água como sinal de referência na técnica aqui proposta assemelha-se ao uso de uma seqüência de treinamento nas técnicas tradicionais.

²Em inglês, *Signal-to-Watermark Ratio*.

Neste capítulo, apresentamos o método de equalização adaptativa proposto e avaliamos seu desempenho por meio de resultados de simulação para canais invariantes e variantes no tempo.

4.2 Equalização Supervisionada Baseada em Marca d'Água: Sistema Proposto

Conforme mencionado no capítulo 2, nos métodos clássicos de equalização supervisionada, o sinal de informação é interrompido periodicamente para a transmissão de uma *seqüência de treinamento*. Este método apresenta duas importantes desvantagens. Durante os períodos de treinamento, a transmissão do sinal de informação é interrompida, o que pode ser inaceitável em diversas aplicações. Além disso, se as características do canal variam significativamente entre períodos de treinamento, a comunicação poderá sofrer com os efeitos de uma má equalização. A solução dessas duas desvantagens é contraditória: a primeira implica em seqüências de treinamento tão espaçadas quanto possível no domínio temporal, enquanto a segunda implica em treinamentos realizados com a maior frequência possível.

A proposta aqui apresentada visa solucionar esses dois problemas simultaneamente. Por um lado, o método utiliza uma marca d'água para supervisão, *transmitida ininterruptamente* através do canal de comunicação. Porém, o sinal de informação *nunca é interrompido*, e os coeficientes do equalizador são continuamente atualizados. De forma a bem situar nossa proposta, apresentamos anteriormente a solução clássica de equalização supervisionada, baseada no paradigma de Wiener.

4.2.1 Equalizador de Wiener

Um esquema clássico de equalização supervisionada a partir de uma medida de erro utilizando seqüência de treinamento é apresentado na figura 4.1. Este método está baseado na comparação da saída $y(n)$ do equalizador com um sinal $d(n)$ de referência, produzindo um sinal de erro $e(n)$. O sinal de referência $d(n)$ é idêntico à seqüência de treinamento $s(n)$ transmitida. Quando o equalizador é adequadamente ajustado, o sinal de erro $e(n)$ é minimizado, o que implica em uma saída do equalizador $y(n)$ a mais próxima possível do sinal $s(n)$ transmitido, de acordo com um critério estabelecido. Esta operação é efetuada periodicamente para rastrear eventuais variações do canal.

O erro quadrático médio, utilizado como função de custo, é dado por:

$$J = E [|e(n)|^2] \quad (4.1)$$

A equação acima expressa o critério de Wiener, cuja solução é obtida através da sua minimização com respeito aos coeficientes $\mathbf{w}$ do filtro de equalização:

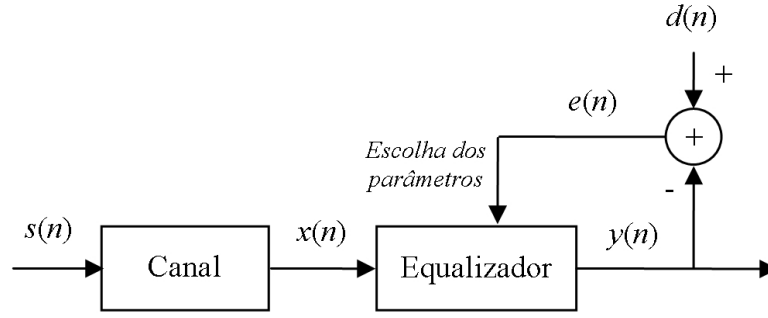


Fig. 4.1: Esquema clássico de equalização supervisionada.

$$J_{\min}(\mathbf{w}) = \min_{\mathbf{w}} E[|e(n)|^2] \quad (4.2)$$

Os coeficientes ótimos $\mathbf{w}_0$ do equalizador são dados por:

$$\mathbf{w}_o = \arg \left(\min_{\mathbf{w}} E[|e(n)|^2] \right) \quad (4.3)$$

A seguir, será apresentada a proposta de um algoritmo utilizando marca d'água.

4.2.2 Descrição do Algoritmo Proposto

Em nosso método, ao invés de substituímos o sinal de informação pela sequência de treinamento, transmite-se continuamente uma marca d'água sobreposta ao sinal de informação, utilizando a técnica de espalhamento espectral para inserção da marca d'água:

$$s(n) = t(n) + \alpha_m m(n) \quad (4.4)$$

onde $t(n)$ é o sinal de informação, $m(n)$ a marca d'água e α_m um fator de escala. O sinal de marca d'água é construído a partir de uma sequência pseudo-aleatória e é decorrelacionado com o sinal de informação $t(n)$. A marca d'água escalonada é usada como sinal de referência:

$$d(n) = \alpha_m m(n) \quad (4.5)$$

Considerando um sistema linear, o sinal recebido $x(n)$ pode ser visto como a soma dos sinais $\tilde{t}(n)$ e $\alpha_m \tilde{m}(n)$, onde $\tilde{t}(n)$ e $\tilde{m}(n)$ são versões distorcidas do sinal de informação e da marca d'água, respectivamente. De forma similar, a saída do equalizador linear $y(n)$ pode ser vista como a soma dos sinais $\hat{t}(n)$ e $\alpha_m \hat{m}(n)$, onde $\hat{t}(n)$ e $\hat{m}(n)$ são versões equalizadas dos sinais de informação e marca d'água recebidos. Este esquema é ilustrado na figura 4.2. O sinal de erro $e(n)$ é expresso como:

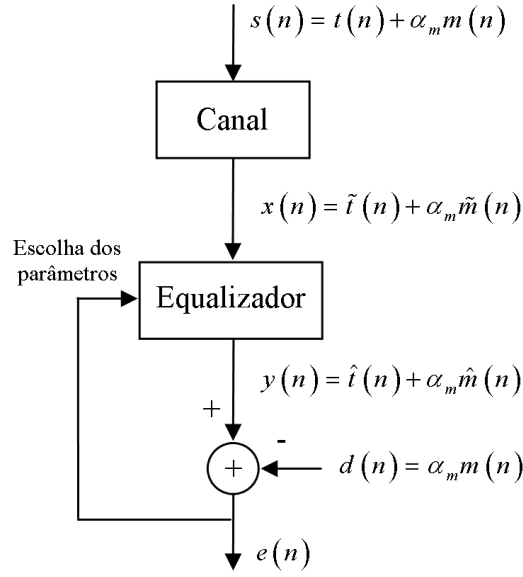


Fig. 4.2: Esquema de equalização utilizando marca d'água.

$$e(n) = y(n) - d(n) = \hat{t}(n) - \alpha_m (m(n) - \hat{m}(n)) \quad (4.6)$$

A função custo J , definida como a esperança do erro quadrático $e^2(n)$, será³:

$$J = E[e^2(n)] = E\left[\left(\hat{t}(n) - \alpha_m (m(n) - \hat{m}(n))\right)^2\right] \quad (4.7)$$

A equação 4.7 pode ser reescrita em notação matricial. Define-se $\mathbf{m}(n)$ e $\mathbf{t}(n)$ como⁴:

$$\mathbf{m}(n) = [m(n), m(n-1), \dots, m(n-M+1)]^T \quad (4.8)$$

$$\mathbf{t}(n) = [t(n), t(n-1), \dots, t(n-M+1)]^T \quad (4.9)$$

Os vetores⁵ $\mathbf{m}(n)$ e $\mathbf{t}(n)$ têm comprimento M , onde M expressa o número de coeficientes do equalizador. Define-se $\mathbf{s}(n)$ como:

$$\mathbf{s}(n) = \mathbf{t}(n) + \alpha_m \mathbf{m}(n) \quad (4.10)$$

com $\mathbf{s}(n)$ apresentando dimensão igual à do equalizador.

³As definições contidas no texto, por simplicidade, têm sempre por base sinais reais.

⁴O tipo negrito é usado para denotar grandezas vetoriais e matriciais.

⁵Quando se lida com sinais complexos, é necessário substituir a transposição matricial simples pelo operador hermitiano. Por simplicidade, estabelecemos a notação tendo em mente sinais reais.

Expandindo-se a função custo definida em 4.1, obtém-se a forma canônica da superfície de erro expressa em forma matricial:

$$J(\mathbf{w}) = \sigma_d^2 - \mathbf{w}^H \mathbf{p} - \mathbf{p}^H \mathbf{w} + \mathbf{w}^H \mathbf{R} \mathbf{w} \quad (4.11)$$

onde $\mathbf{R}$ é a matriz de autocorrelação (dimensão $M \times M$) na entrada do equalizador, $\mathbf{p}$ é o vetor de correlação cruzada (dimensão $M \times 1$) entre a entrada do equalizador e a resposta desejada do mesmo, σ_d^2 é a variância do sinal de referência, e $\mathbf{w}$ (dimensão $M \times 1$) denota os coeficientes do equalizador.

Podemos representar o sinal $\mathbf{x}(n)$ na entrada do equalizador como:

$$\mathbf{x}(n) = \mathbf{H} \mathbf{s}(n) \quad (4.12)$$

onde $\mathbf{s}(n)$ é o sinal transmitido e $\mathbf{H} \in \mathbb{R}^{M \times (M+N-1)}$ é a matriz de convolução do canal, onde M é a ordem do equalizador. Assume-se um canal FIR de ordem N e coeficientes $\mathbf{h} = [h_0 \cdots h_{N-1}]^T$. A matriz de convolução, obtida a partir dos coeficientes $\mathbf{h}$, é uma matriz Toeplitz:

$$\mathbf{H} = \begin{bmatrix} h_0 & \cdots & h_{N-1} & 0 & 0 & \cdots & 0 \\ 0 & h_0 & \cdots & h_{N-1} & 0 & \ddots & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & 0 & \vdots \\ 0 & \ddots & 0 & h_0 & \cdots & h_{N-1} & 0 \\ 0 & \cdots & 0 & 0 & h_0 & \cdots & h_{N-1} \end{bmatrix} \quad (4.13)$$

De posse da relação 4.12, a matriz de autocorrelação $\mathbf{R}$ pode ser escrita da seguinte forma:

$$\mathbf{R} = E [\mathbf{x}(n) \mathbf{x}^T(n)] = \mathbf{H} E [\mathbf{s}(n) \mathbf{s}^T(n)] \mathbf{H}^T \quad (4.14)$$

Expandindo o sinal $\mathbf{s}(n)$ por meio da equação 4.10, obtemos:

$$\mathbf{R} = \alpha_m^2 \mathbf{H} E [\mathbf{m}(n) \mathbf{m}^T(n)] \mathbf{H}^T + \mathbf{H} E [\mathbf{t}(n) \mathbf{t}^T(n)] \mathbf{H}^T \quad (4.15)$$

onde α_m^2 representa a variância da marca d'água. O termo $\mathbf{H} E [\mathbf{m}(n) \mathbf{m}^T(n)] \mathbf{H}^T$ expressa a matriz de autocorrelação do sinal $\tilde{m}(n)$ e $\mathbf{H} E [\mathbf{t}(n) \mathbf{t}^T(n)] \mathbf{H}^T$ a matriz de autocorrelação do sinal $\tilde{t}(n)$, e serão representados por $\mathbf{R}_{\tilde{m}}$ e $\mathbf{R}_{\tilde{t}}$, respectivamente. Temos, com isso:

$$\mathbf{R} = \alpha_m^2 \mathbf{R}_{\tilde{m}} + \mathbf{R}_{\tilde{t}} \quad (4.16)$$

O vetor $\mathbf{p}$ de correlação cruzada entre o sinal na entrada do equalizador e de marca d'água (sinal de referência) pode ser escrito da seguinte forma:

$$\mathbf{p} = E[\mathbf{x}(n) \alpha_m m(n-d)] = \alpha_m \mathbf{H} E[\mathbf{s}(n) m(n-d)] \quad (4.17)$$

onde d é o atraso de equalização. Como $\mathbf{t}(n)$ e $\mathbf{m}(n)$ são independentes, obtemos:

$$\mathbf{p} = \alpha_m^2 \mathbf{H} E[\tilde{\mathbf{m}}(n) m(n-d)] \quad (4.18)$$

O termo $\mathbf{H} E[\tilde{\mathbf{m}}(n) m(n-d)]$ representa a correlação cruzada entre a versão distorcida (na entrada do filtro) e original (sinal desejado de referência) da marca d'água, e será expresso por $\mathbf{p}_{\tilde{m}}$. Com isso:

$$\mathbf{p} = \alpha_m^2 \mathbf{p}_{\tilde{m}} \quad (4.19)$$

Os termos $\mathbf{R}_{\tilde{m}}$ e $\mathbf{R}_{\tilde{t}}$ mostrados em 4.16 representam as matrizes de autocorrelação das versões distorcidas da marca d'água e do sinal de informação, respectivamente. O termo $\mathbf{p}_{\tilde{m}}$ apresentado em 4.19 é a correlação cruzada entre a versão distorcida e original da marca d'água. Substituindo 4.16 e 4.19 na equação 4.11, obtemos:

$$J(\mathbf{w}) = \sigma_m^2 - \alpha_m^2 \mathbf{w}^H \mathbf{p}_{\tilde{m}} - \alpha_m^2 \mathbf{p}_{\tilde{m}}^H \mathbf{w} + \alpha_m^2 \mathbf{w}^H \mathbf{R}_{\tilde{m}} \mathbf{w} + \mathbf{w}^H \mathbf{R}_{\tilde{t}} \mathbf{w} \quad (4.20)$$

O vetor de coeficientes do equalizador, que representa a solução ótima de Wiener, é:

$$\mathbf{w}_0 = \mathbf{R}^{-1} \mathbf{p} = \alpha_m^2 \left(\alpha_m^2 \mathbf{R}_{\tilde{m}} + \mathbf{R}_{\tilde{t}} \right)^{-1} \mathbf{p}_{\tilde{m}} \quad (4.21)$$

A presença da matriz de correlação $\mathbf{R}_{\tilde{t}}$ nas equações 4.20 e 4.21 implica em uma dependência da solução de Wiener com relação ao sinal de informação; portanto, em geral, não basta o conhecimento prévio do sinal de marca d'água para que seja possível o ajuste dos coeficientes do equalizador. Nas próximas seções será mostrado como o esquema de equalização utilizando marca d'água pode ser usado para obter a solução que inverte o canal, apesar da dependência do método com relação às características do sinal de informação.

4.3 Equalização de Sinais Descorrelacionados

Nesta seção, assumiremos que o sinal de informação é descorrelacionado⁶, hipótese mais usual em transmissão digital, devido às técnicas de embaralhamento. Supomos ainda o sinal de marca d'água normalizado e de espectro branco⁷. Com isso, a esperança $E[\mathbf{m}(n) \mathbf{m}^T(n)]$ torna-se uma

⁶Não há correlação entre o sinal e versões atrasadas do mesmo.

⁷Dotado de um conteúdo uniforme e de extensão infinita em frequência, em analogia com a luz branca, que contém todas as cores do espectro.

matriz identidade. Já o termo $E [\mathbf{t}(n) \mathbf{t}^T(n)]$ é uma matriz diagonal, cujo valor é dado pela variância σ_t^2 do sinal $\mathbf{t}(n)$. Assim, a matriz $\mathbf{R}$ representada na equação 4.15 pode ser reescrita como:

$$\mathbf{R} = \alpha_m^2 \mathbf{H}\mathbf{H}^T + \sigma_t^2 \mathbf{H}\mathbf{H}^T = (\alpha_m^2 + \sigma_t^2) \mathbf{H}\mathbf{H}^T \quad (4.22)$$

Portanto, os coeficientes ótimos referentes à solução de Wiener podem ser expressos como:

$$\mathbf{w}_o = \mathbf{R}^{-1} \mathbf{p} = \frac{\alpha_m^2}{\alpha_m^2 + \sigma_t^2} (\mathbf{H}\mathbf{H}^T)^{-1} \mathbf{p}_{\tilde{m}} \quad (4.23)$$

A equação 4.23 demonstra que, exceto por um fator de escala, o equalizador de Wiener tende, neste caso, a inverter a resposta impulsiva do canal. Este fator de escala pode ser facilmente compensado pois a variância α_m^2 da marca d'água é conhecida e, para a maioria das aplicações, estima-se facilmente a variância σ_t^2 do sinal de informação.

Quando o equalizador reverte de maneira ideal os efeitos do canal sobre o sinal transmitido, têm-se $\hat{m}(n) = m(n)$ e $\hat{t}(n) = t(n)$. Combinando este resultado com 4.6, obtemos:

$$e_o(n) = \frac{\alpha_m^2}{\alpha_m^2 + \sigma_t^2} t(n) \quad (4.24)$$

onde $e_o(n)$ é o sinal de erro na condição ótima. Exceto por um fator de escala, este sinal é idêntico ao sinal de informação equalizado.

Na prática, a qualidade da equalização dependerá da estrutura do equalizador. Um equalizador FIR, enquanto previne problemas de estabilidade, não é capaz de inverter exatamente o canal; entretanto, com uma escolha apropriada da ordem do equalizador FIR, a equalização pode ser tão precisa quanto se deseje.

4.3.1 Simulações e Resultados

Para avaliar o método proposto de equalização, inicialmente para sinais descorrelacionados, apresentamos a seguir alguns resultados de simulação. Nelas, foram utilizados sinais com diferentes modulações e canais FIR com diferentes funções de transferência. Além disso, buscamos determinar os limiares de relação sinal-marca d'água (SWR) com os quais o sistema pode trabalhar. A velocidade de convergência e o erro quadrático médio residual dos algoritmos LMS e RLS também foram alvos de análise.

Como exemplo inicial, tomemos um sistema modelado como um filtro linear de fase mínima, passível de inversão perfeita, com a seguinte função de transferência:

$$H_1(z) = 1 + 0,6z^{-1} \quad (4.25)$$

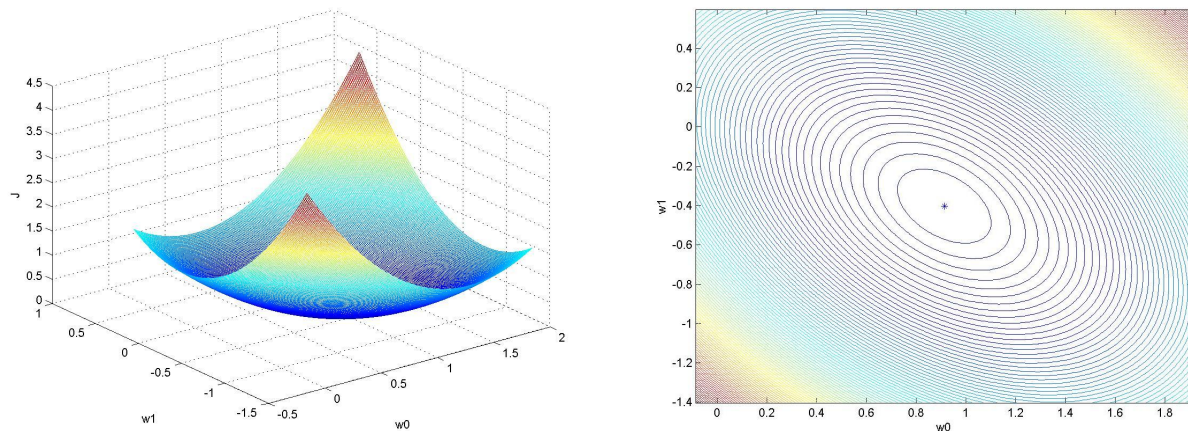


Fig. 4.3: Função custo e respectivas curvas de nível para o esquema de equalização utilizando marca d'água.

O equalizador ótimo correspondente a esta função de transferência tem coeficientes $[0, 913 \ -0, 402]$. Assumimos o sinal de informação formado por amostras i.i.d. (independente e identicamente distribuído) extraídas de uma constelação 2-PAM (modulação por amplitude de pulso) de módulo unitário e obtidas através da função *randsrc* no software de simulação MATLAB, supomos que não há ruído aditivo e que o sistema encontra-se sincronizado.

Na figura 4.3 está traçada a superfície de erro referente à melhor solução para o sistema utilizado no exemplo (para um canal de fase mínima, o menor atraso corresponde à melhor solução). Como seria de se esperar, a função custo para o método de equalização utilizando marca d'água é unimodal e o ponto de mínimo corresponde aos coeficientes do equalizador.

A existência de ruído branco, em tese, não deve modificar a posição do ponto ótimo de Wiener, pois se espera que não haja correlação entre o ruído e o sinal a equalizar. Entretanto, em algoritmos estocásticos como o LMS e o RLS, o ruído interfere no transiente do processo. Do ponto de vista da marca d'água, o sinal transmitido pode ser visto como um forte ruído aditivo, tornando-se fundamental a análise de desempenho dos algoritmos para diferentes valores da relação de potência sinal-marca d'água.

Na figura 4.4, mostramos a evolução temporal dos coeficientes do equalizador supondo dois coeficientes com condição inicial nula, para uma realização do algoritmo LMS. A relação sinal-marca d'água é de 0 dB e passo de adaptação $\mu = 0,0002$. Na ilustração, o regime permanente da evolução dos coeficientes corresponde à solução de Wiener.

A figura 4.5 apresenta duas realizações do LMS para $\text{SWR} = 10$ dB, com passos $\mu = 0,0002$ e $\mu = 0,0005$. Nota-se na figura 4.5(a) que a velocidade de convergência não sofreu influências significativas em relação à realização com $\text{SWR} = 0$ dB (figura 4.4), mantendo-se o valor do passo de

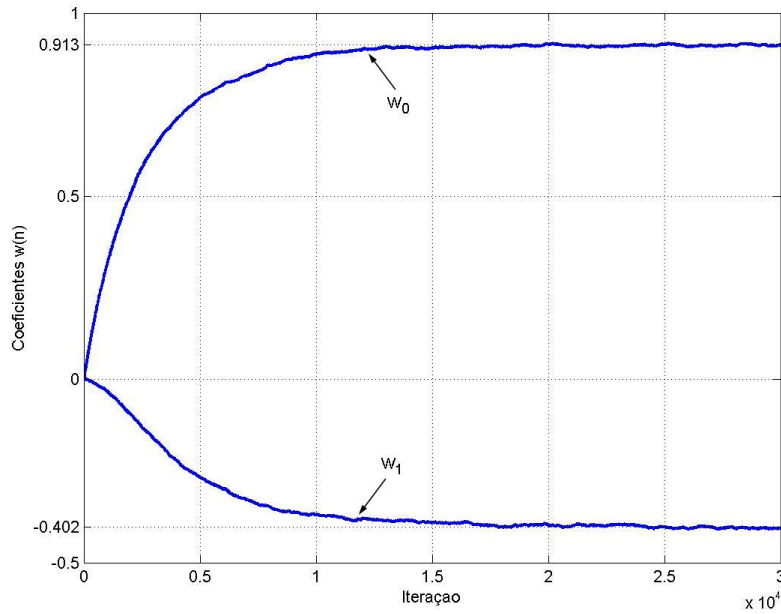


Fig. 4.4: Evolução temporal do algoritmo LMS (SWR = 0 dB, $\mu = 0,0002$).

convergência. A figura 4.5(b) ($\mu = 0,0005$) mostra que um pequeno aumento no passo de adaptação acarretou um grande aumento na velocidade de convergência em relação à realização da figura 4.5(a) ($\mu = 0,0002$), com pouco acréscimo em flutuações estocásticas. Já a figura 4.6 apresenta uma realização do LMS para SWR = 20 dB com passo $\mu = 0,0002$. Neste caso, nota-se um aumento considerável nas flutuações estocásticas e pouca alteração na velocidade de convergência.

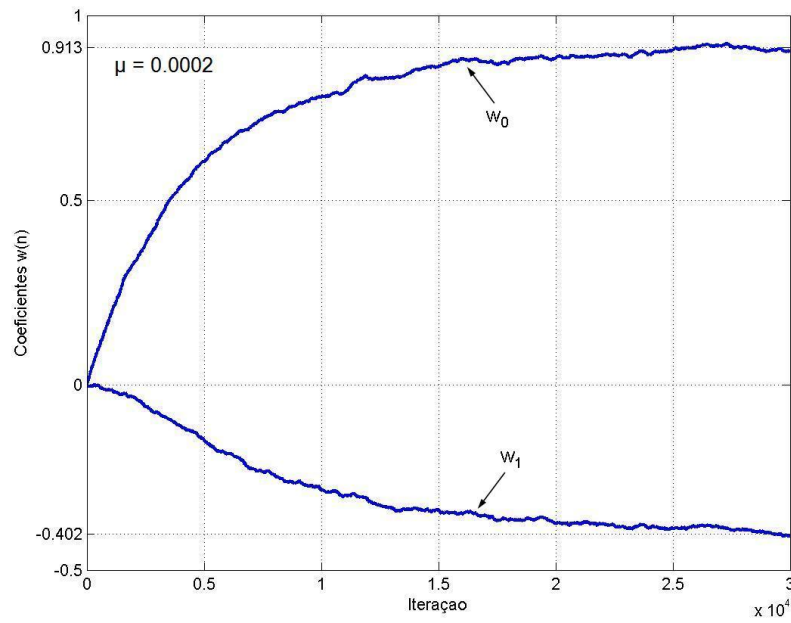
O desempenho apresentado pelo LMS é coerente com a teoria clássica do algoritmo. É possível mostrar que, supondo a curva do algoritmo LMS aproximada por uma exponencial com constante de tempo $(\tau)_{EQM,me}$, pode-se definir a seguinte *constante de tempo média* para o algoritmo (Haykin, 1996):

$$(\tau)_{EQM,me} \approx \frac{1}{2\mu\lambda_{me}} \quad (4.26)$$

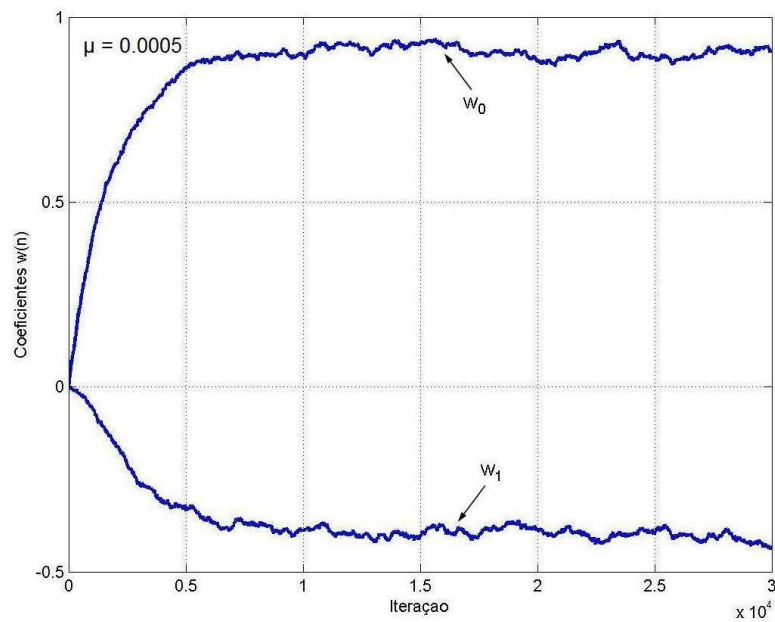
onde μ é o passo de adaptação e λ_{me} o *autovalor médio* da matriz de autocorrelação $\mathbf{R}$ do sinal na entrada do filtro:

$$\lambda_{me} = \frac{1}{M} \sum_{i=1}^M \lambda_i \quad (4.27)$$

sendo λ_i ($i = 1, 2, \dots, M$) os autovalores da matriz de autocorrelação $\mathbf{R}$, com M expressando o comprimento do equalizador.



(a)



(b)

Fig. 4.5: Evolução temporal do algoritmo LMS para diferentes passos μ (SWR = 10 dB): 4.5(a) $\mu = 0,0002$; 4.5(b) $\mu = 0,0005$.

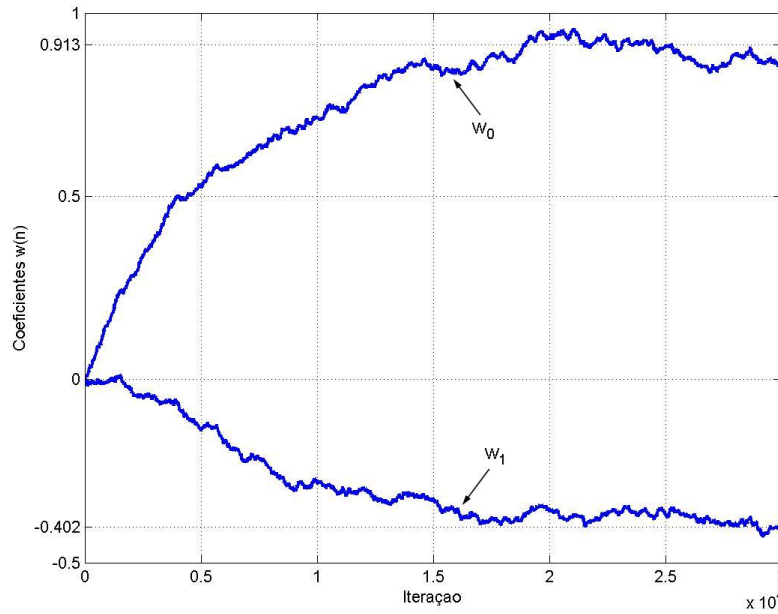


Fig. 4.6: Evolução temporal do algoritmo LMS (SWR = 20 dB, $\mu = 0,0002$).

No esquema de equalização apresentado, tem-se na entrada do equalizador a soma dos sinais de informação e de marca d'água, ambos distorcidos pelo canal; conforme mostrado, a matriz de autocorrelação $\mathbf{R}$ do sinal na entrada do filtro dependerá portanto desses dois sinais. Como a potência da marca d'água em geral é muito menor do que a do sinal de informação, os autovalores da matriz de autocorrelação são poucos alterados à medida que diminuimos a potência da marca. Conseqüentemente, o valor da constante de tempo $(\tau)_{EQM,me}$ do algoritmo LMS sofre pouca alteração. Essa característica justifica o fato de a velocidade de convergência sofrer poucas alterações com variações da SWR. Nota-se claramente que o aumento da SWR tem maior influência nas flutuações estocásticas do algoritmo, afetando diretamente o EQM. Em termos gerais, pode-se concluir que o LMS apresenta desempenho satisfatório, com a vantagem de apresentar custo computacional relativamente baixo.

Passemos agora à avaliação do algoritmo RLS. A figura 4.7 ilustra a evolução temporal dos coeficientes do equalizador para uma realização do algoritmo RLS com relação sinal-marca d'água de 0 dB e fator de esquecimento $\lambda = 1$. Em todos os ensaios utilizando o RLS, observou-se 100% de convergência global.

A figura 4.8 ilustra duas realizações do RLS para uma SWR = 10 dB, com $\lambda = 1$ e $\lambda = 0,998$. Como seria de se esperar, o RLS apresenta uma maior velocidade de convergência e um menor desajuste em relação ao LMS. Com $\lambda = 0,998$, é nítido o acréscimo de pequenas flutuações estocásticas que afetam o EQM, apesar de o algoritmo manter-se bastante estável em torno do ponto ótimo.

A figura 4.9 ilustra uma realização do RLS para SWR = 20 dB, que é a relação sinal-marca d'água

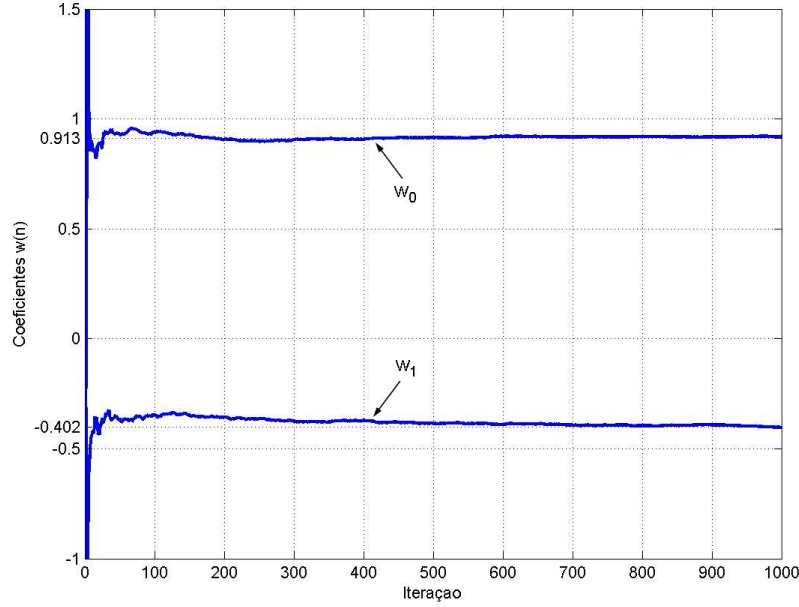


Fig. 4.7: Evolução temporal do algoritmo RLS (SWR = 0 dB, $\lambda = 1$).

normalmente empregada em sistemas de marca d'água de áudio para que seja garantida a transparência da marca utilizando modelos psicoacústicos. A figura mostra um aumento nas flutuações estocásticas e pouca alteração na velocidade de convergência em relação aos ensaios com valores menores de SWR.

O desempenho do algoritmo RLS também pode ser explicado pela teoria clássica do algoritmo. É possível mostrar (Haykin, 2001) que o *desvio quadrático médio* $\mathcal{D}$ do RLS é dado por:

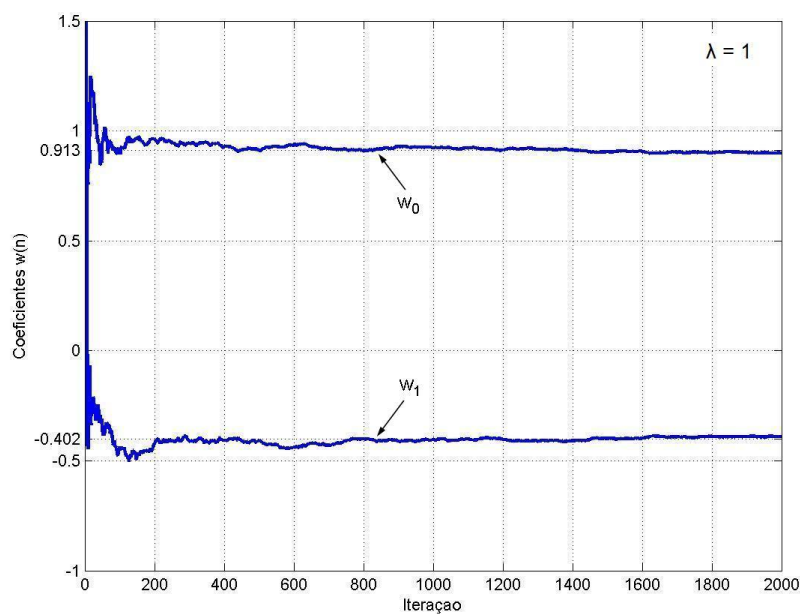
$$\mathcal{D} = E \left[\varepsilon^T(n) \varepsilon(n) \right] \quad (4.28)$$

onde $\varepsilon(n) = \mathbf{w}_o - \hat{\mathbf{w}}(n)$, sendo $\hat{\mathbf{w}}(n)$ a estimativa do vetor de parâmetros e $\mathbf{w}_o$ seu valor ótimo. Esta expressão pode ser dada por:

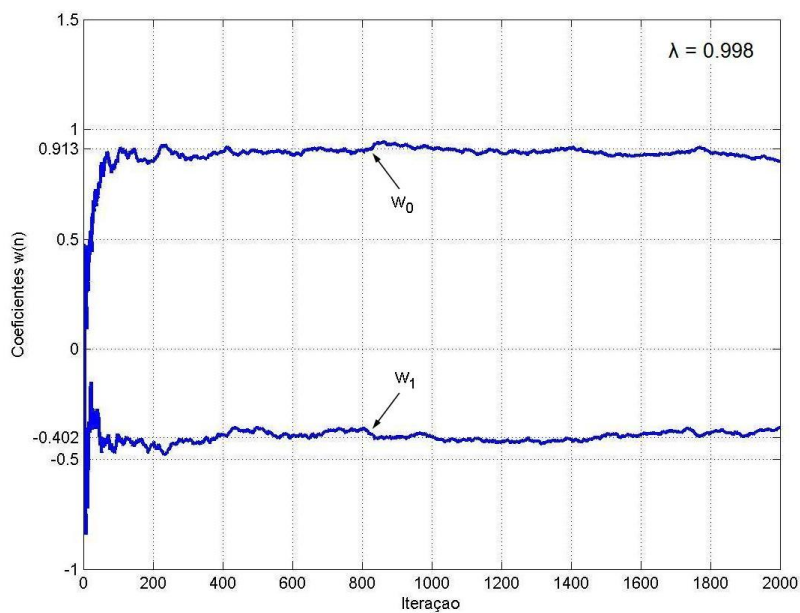
$$\mathcal{D}(n) = \frac{1}{n} \sigma_o^2 \sum_{i=1}^M \frac{1}{\lambda_i}, \quad n > M \quad (4.29)$$

onde λ_i são os autovalores da matriz de autocorrelação $\mathbf{R}$ do sinal na entrada do filtro, σ_o^2 é a variância de $e_o(n)$ (erro na condição ótima, determinado com $\mathbf{w}_o$) e M é o comprimento do equalizador.

No esquema de equalização discutido, os autovalores da matriz de autocorrelação $\mathbf{R}$ dependem dos sinais de informação e de marca d'água. Como a potência da marca é, em geral, muito menor do que a potência do sinal de informação, o desvio do algoritmo RLS é pouco alterado à medida que



(a)



(b)

Fig. 4.8: Evolução temporal do algoritmo RLS para diferentes valores de λ (SWR = 10 dB): 4.8(a) $\lambda = 1$; 4.8(b) $\lambda = 0,998$.

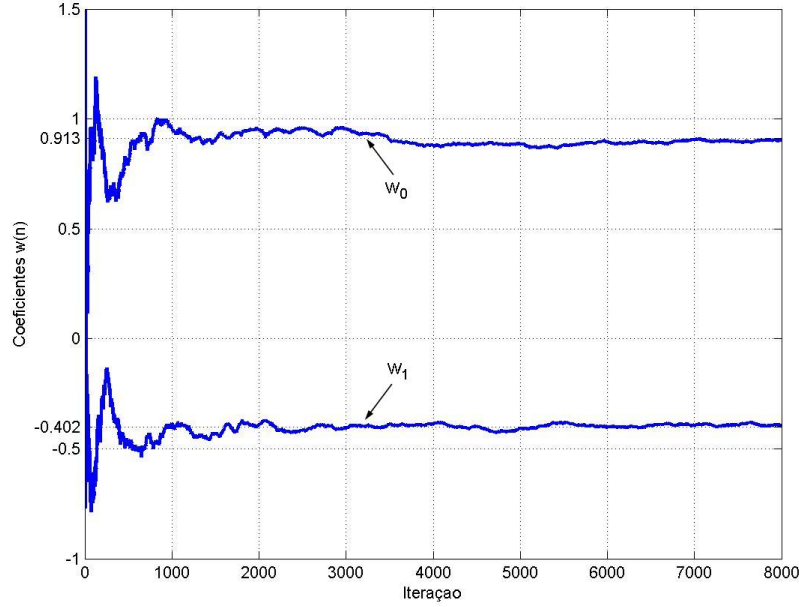


Fig. 4.9: Evolução temporal do algoritmo RLS (SWR = 20 dB, $\lambda = 1$).

diminuímos a potência da marca, justificando o fato de a velocidade de convergência sofrer poucas alterações com as variações da SWR.

Também é possível mostrar que o RLS, diferentemente do LMS (que oscila em torno do ponto ótimo), não apresenta desajuste dos coeficientes para um número de iterações n tendendo a infinito (Haykin, 2001). O valor médio de convergência do RLS pode ser escrito como:

$$E[\hat{\mathbf{w}}(n)] \approx \mathbf{w}_o - \frac{\delta}{n} \mathbf{p}, \quad n > M \quad (4.30)$$

onde $\mathbf{p}$ é o vetor de correlação cruzada entre a entrada do filtro e o sinal desejado e δ uma constante positiva de pequeno valor. O RLS converge para o ponto ótimo de Wiener e é invariante com relação aos autovalores da matriz de autocorrelação $\mathbf{R}$. Para um valor finito de n maior do que M , a estimativa $\hat{\mathbf{w}}$ apresenta um desvio que depende da inicialização do algoritmo; porém, esse desvio decresce a zero com n tendendo ao infinito. Dessa forma, o valor médio de convergência do RLS tende a zero conforme aumenta o número de iterações n .

Conclui-se que, em equalização, a velocidade de convergência sofreu pouca influência com a variação da SWR. Nota-se claramente que a variação da SWR tem maior influência nas flutuações estocásticas do algoritmo, afetando diretamente o EQM.

Adotemos agora um canal de teste com fase não-mínima, passível de inversão perfeita, modelado como um filtro linear com a seguinte função de transferência:

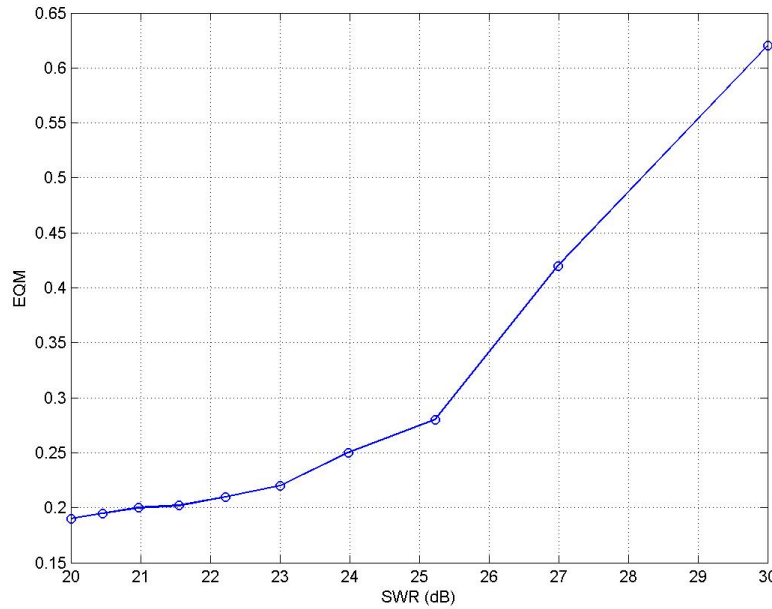


Fig. 4.10: Erro quadrático médio do algoritmo LMS em função da relação de potência sinal-marca d'água.

$$H_2(z) = 1 + 1,2z^{-1} - 0,3z^{-2} + 0,8z^{-3} \quad (4.31)$$

Consideremos um sinal 2-PSK (*Phase Shift Keying*) com amostras i.i.d. obtidas no software de simulação MATLAB, sem ruído, supondo o sistema sincronizado, e um equalizador com cinco coeficientes inicializados na origem⁸. Para avaliação do método de equalização, analisaremos a influência da relação sinal-marca d'água sobre o erro quadrático médio de equalização.

A figura 4.10 apresenta o erro quadrático médio residual do sinal recuperado, após a convergência (regime permanente onde os coeficientes apresentam-se em torno dos valores ótimos), através do algoritmo LMS, versus a razão de potências entre sinal e marca d'água (SWR), com passo $\mu = 0,0005$ e média de 200 realizações. Os resultados de simulações sugerem que o EQM cresce de forma aproximadamente exponencial com a SWR, pois a escala em decibéis é logarítmica. Em escala linear, o EQM cresce de forma aproximadamente linear. No exemplo apresentado na figura 4.10, para $SWR = 20$ dB o EQM residual é próximo do valor ótimo (0,1082 com um equalizador de quinta ordem). Para uma $SWR = 23$ dB, o valor do EQM residual médio é aproximadamente duas vezes o valor ótimo. Já para SWR na faixa de 25-30 dB, o EQM cresce rapidamente. O valor máximo aceitável do EQM dependerá da aplicação.

⁸Todos os coeficientes inicializados com valor 0.

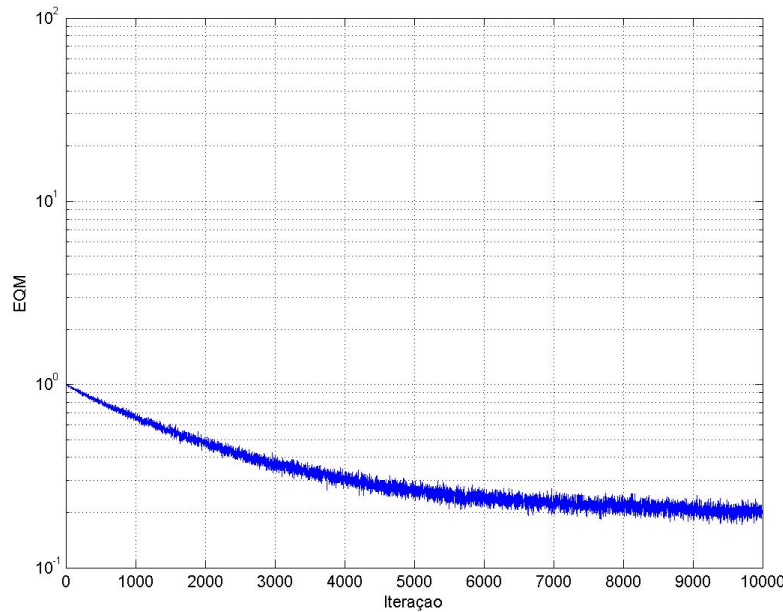


Fig. 4.11: Convergência do erro quadrático médio do algoritmo LMS para SWR = 20 dB.

A figura 4.11 mostra o EQM em função do número de iterações para SWR = 20 dB e uma média de 200 realizações do algoritmo LMS com um passo de 0,0005. Nota-se que o algoritmo atingiu a convergência com menos de 10.000 iterações.

A figura 4.12 apresenta o EQM residual do sinal recuperado, após a convergência (os coeficientes apresentam-se em torno dos valores ótimos), através do algoritmo RLS versus a razão de potências SWR, para $\lambda = 1$ e média de 200 realizações. Os resultados das simulações com este algoritmo foram muito próximos aos obtidos com o algoritmo LMS, sugerindo que o EQM cresce de forma aproximadamente exponencialmente com a SWR. A figura 4.13 mostra o EQM em função do número de iterações para SWR = 20 dB e uma média de 200 realizações do algoritmo RLS com fator de esquecimento $\lambda = 1$. Nota-se que o algoritmo também atingiu a convergência com menos de 10.000 iterações.

Os resultados experimentais atestam assim a viabilidade do uso dos algoritmos LMS e RLS na obtenção da solução ótima para o método de equalização proposto. Na análise dos algoritmos RLS e LMS em regime permanente, notamos que os resultados de ambos são muito parecidos. Porém, vale lembrar que o RLS não apresenta desajuste para um número de iterações tendendo a infinito.

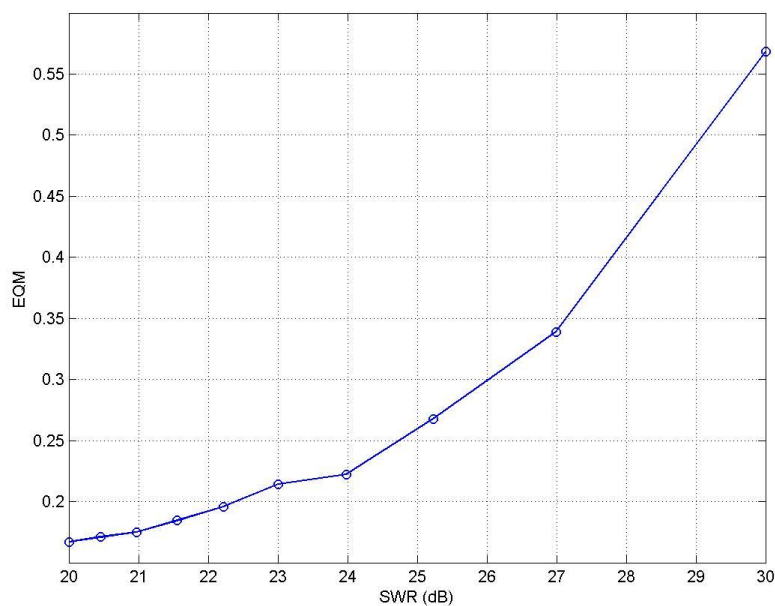


Fig. 4.12: Erro quadrático médio do algoritmo RLS em função da relação de potência sinal-marca d'água.

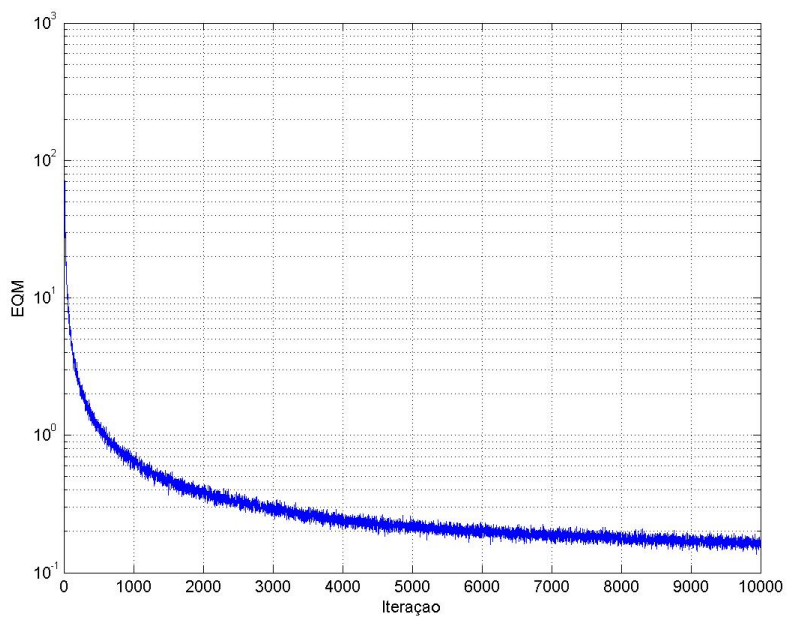


Fig. 4.13: Convergência do erro quadrático médio do algoritmo RLS para SWR = 20 dB.

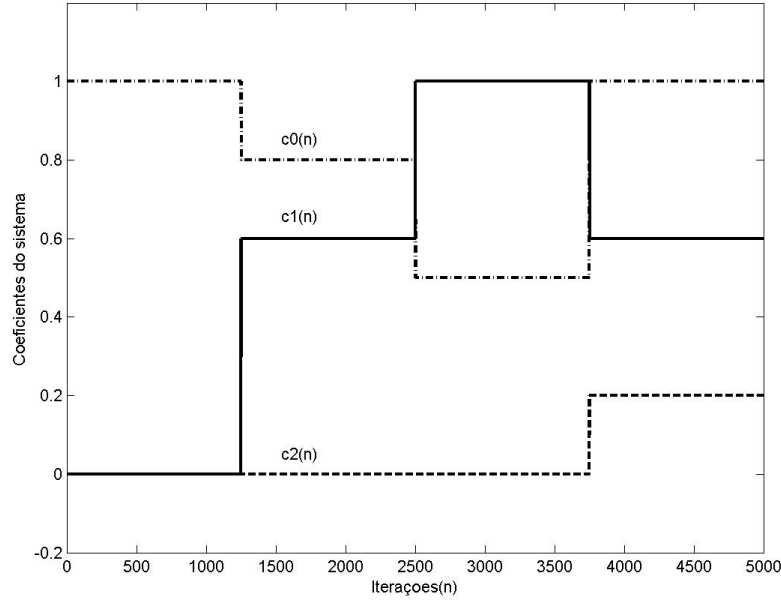


Fig. 4.14: Coeficientes baseados em funções degrau.

Análise de Canais Variantes no Tempo

Para que possamos analisar o desempenho da abordagem proposta quando aplicada a canais variantes no tempo, escolhemos dois cenários onde seus parâmetros apresentam um comportamento dinâmico. Assumimos que as amostras transmitidas são i.i.d. e pertencentes a uma constelação 2-PAM com módulo unitário, obtidas através da função *randsrc* no software de simulação MATLAB. Assumimos também que tanto o canal como o equalizador são representáveis por filtros FIR, sem a presença de ruído aditivo. Além disso, assumimos o sistema sincronizado.

O canal apresenta a seguinte relação entre entrada e saída:

$$x(n) = c_0(n)s(n) + c_1(n)s(n-1) + c_2(n)s(n-2) \quad (4.32)$$

onde $c_0(n)$, $c_1(n)$ e $c_2(n)$ são definidos como ilustrado nas figuras 4.14 e 4.15.

Estes modelos são propostos para a análise de sistemas variantes no tempo (Junqueira et al., 2005). Na figura 4.14, o canal variante é baseado em funções degrau e apresenta as seguintes transições:

- $\begin{bmatrix} c_0(0) & c_1(0) & c_2(0) \end{bmatrix} = \begin{bmatrix} 1, 0 & 0, 0 & 0, 0 \end{bmatrix}$;
- $\begin{bmatrix} c_0(1250) & c_1(1250) & c_2(1250) \end{bmatrix} = \begin{bmatrix} 0, 8 & 0, 6 & 0, 0 \end{bmatrix}$;
- $\begin{bmatrix} c_0(2500) & c_1(2500) & c_2(2500) \end{bmatrix} = \begin{bmatrix} 0, 5 & 1, 0 & 0, 0 \end{bmatrix}$;

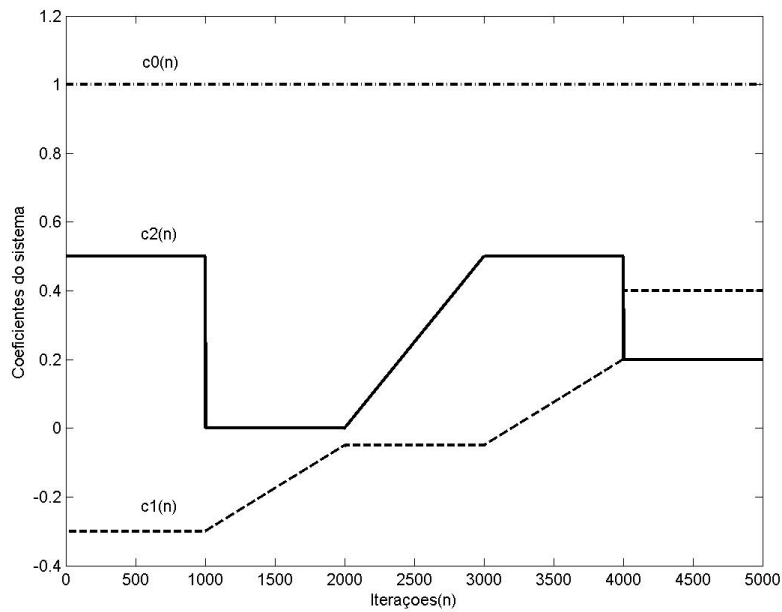


Fig. 4.15: Coeficientes baseados em funções degrau e rampa.

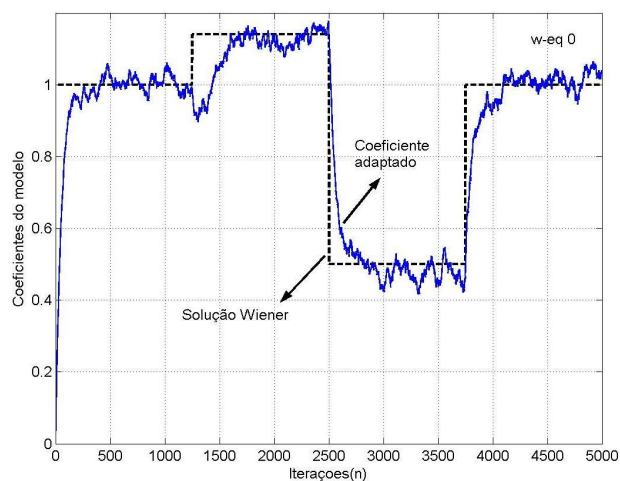
$$\bullet \begin{bmatrix} c_0(3750) & c_1(3750) & c_2(3750) \end{bmatrix} = \begin{bmatrix} 1,0 & 0,6 & 0,2 \end{bmatrix};$$

Na figura 4.15 o canal é baseado em funções degrau e rampa e apresenta as seguintes transições:

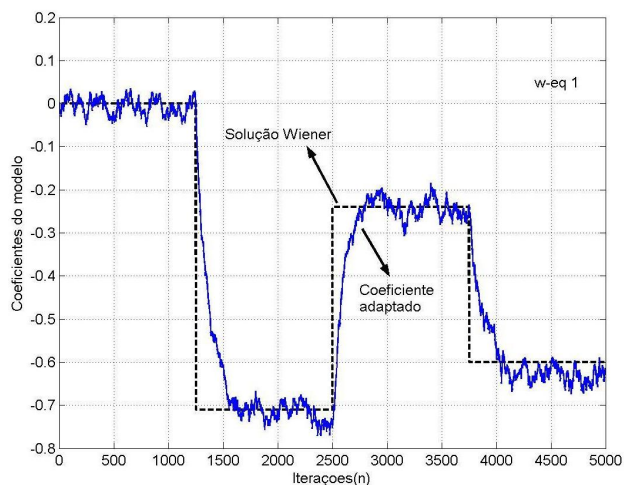
- $\bullet \begin{bmatrix} c_0(0) & c_1(0) & c_2(0) \end{bmatrix} = \begin{bmatrix} 1,0 & -0,3 & 0,5 \end{bmatrix};$
- $\bullet \begin{bmatrix} c_0(1000) & c_1(1000) & c_2(1000) \end{bmatrix} = \begin{bmatrix} 1,0 & -0,3 & 0,0 \end{bmatrix};$
- $\bullet \begin{bmatrix} c_0(2000) & c_1(2000) & c_2(2000) \end{bmatrix} = \begin{bmatrix} 1,0 & -0,05 & 0,0 \end{bmatrix};$
- $\bullet \begin{bmatrix} c_0(3000) & c_1(3000) & c_2(3000) \end{bmatrix} = \begin{bmatrix} 1,0 & -0,05 & 0,5 \end{bmatrix};$
- $\bullet \begin{bmatrix} c_0(4000) & c_1(4000) & c_2(4000) \end{bmatrix} = \begin{bmatrix} 1,0 & 0,4 & 0,2 \end{bmatrix};$

Nas simulações, adotamos um equalizador $[w_{eq1}(0) \ w_{eq2}(0) \ w_{eq3}(0)]$ com três coeficientes. Em todos os experimentos, os coeficientes foram adaptados pelo esquema de equalização proposto e comparados com os coeficientes ideais de Wiener em cada instante.

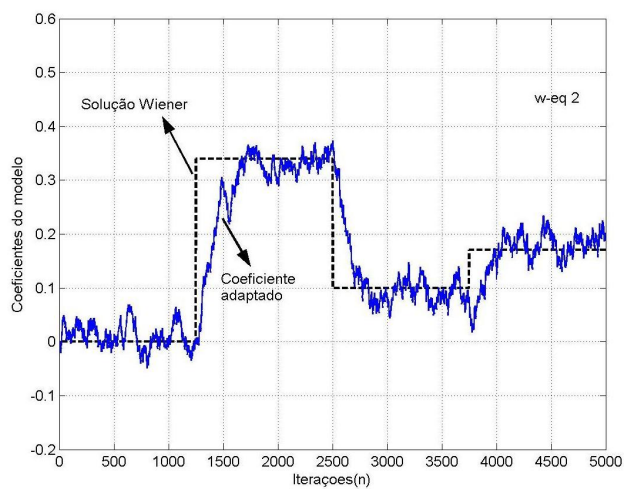
As figuras 4.16 e 4.17 ilustram a estimativa adaptativa obtida pelo equalizador através do algoritmo LMS para os coeficientes dos canais variando segundo funções degrau e degrau-rampa, respectivamente. Nas duas simulações, foi utilizado um passo de adaptação $\mu = 0,02$ e uma média de 200 realizações do algoritmo. Podemos ver nas ilustrações que o algoritmo atinge rapidamente os pontos ótimos de convergência, rastreamento as variações dos canais.



(a)



(b)



(c)

Fig. 4.16: Coeficientes do equalizador obtidos através do algoritmo LMS para um canal variante no tempo baseado em degrau, tendo a solução de Wiener como referência: 4.16(a) w_{eq1} ; 4.16(b) w_{eq2} ; 4.16(c) w_{eq3} .

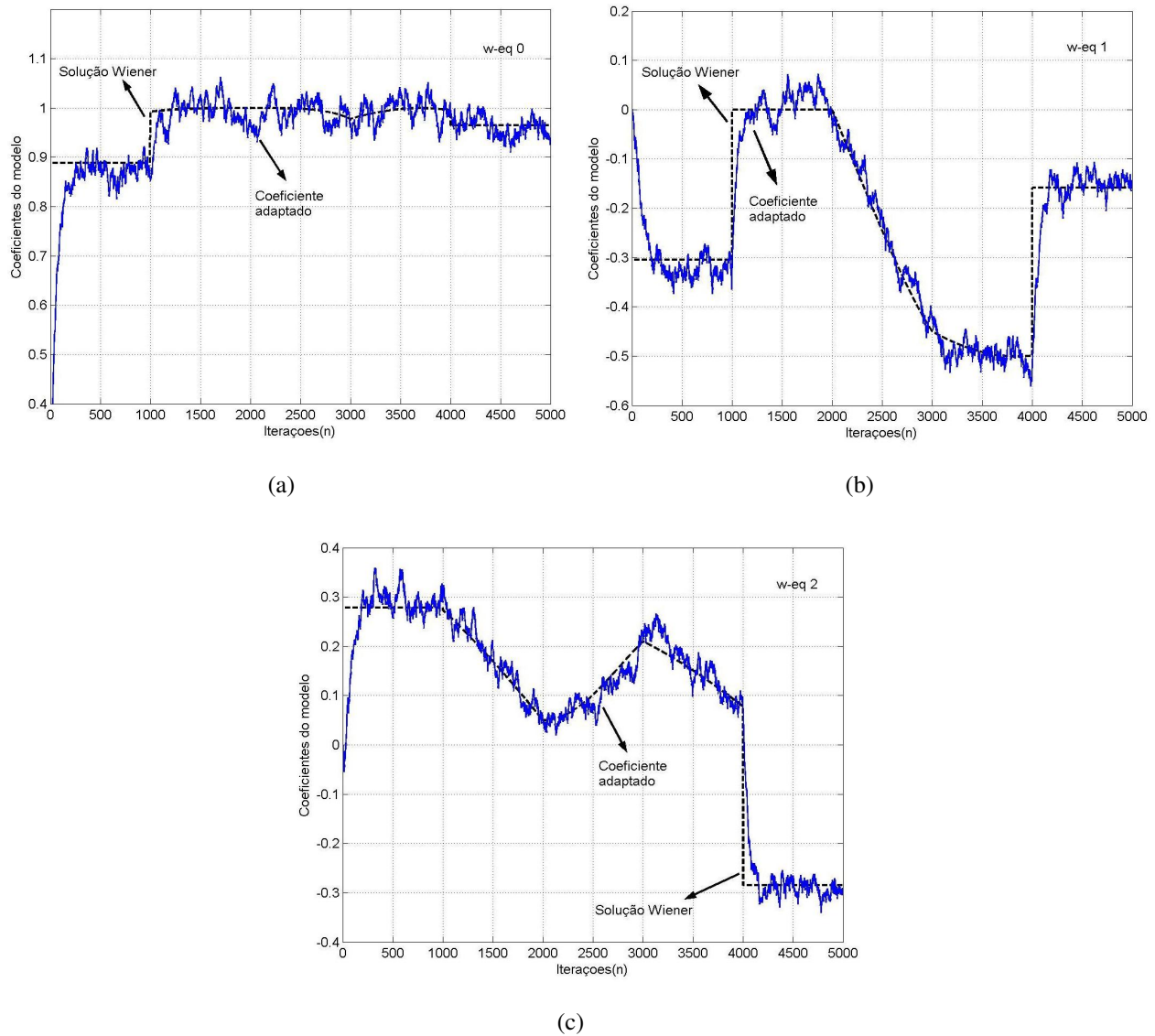
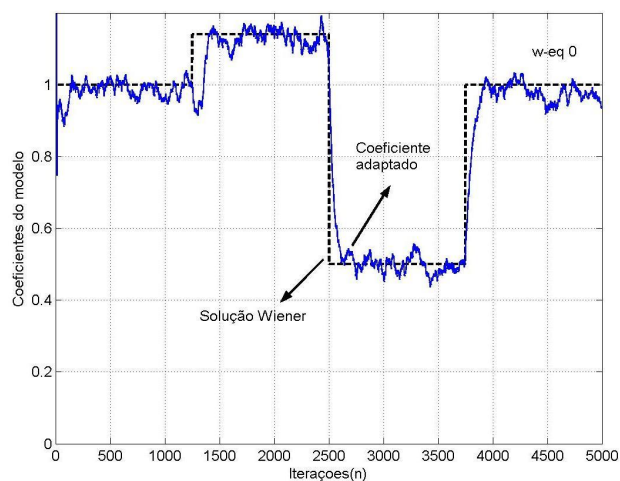
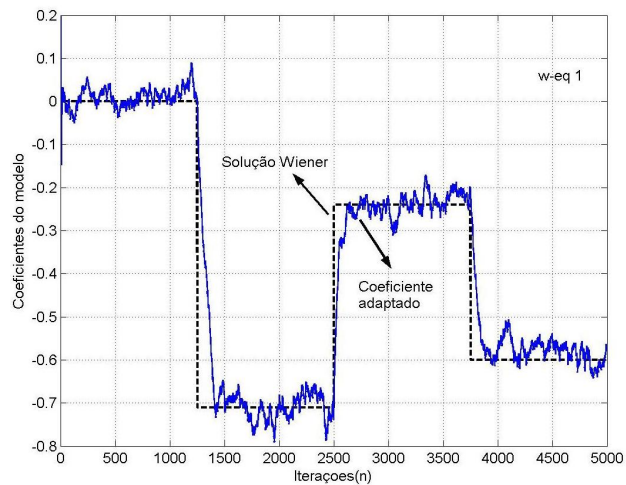


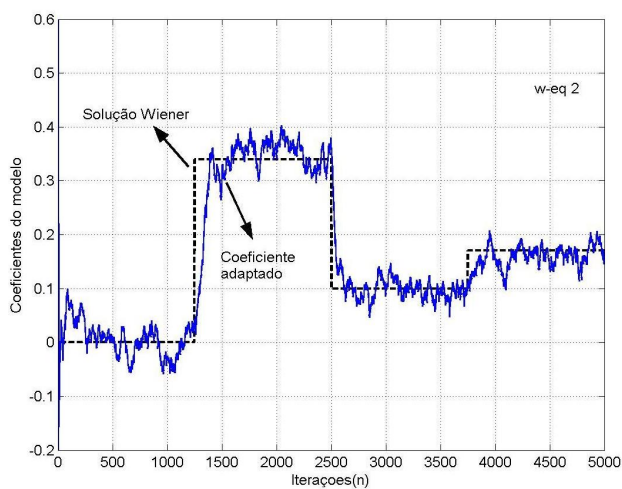
Fig. 4.17: Coeficientes do equalizador obtidos através do algoritmo LMS para um canal variante no tempo baseado em degrau-rampa, tendo a solução de Wiener como referência: 4.17(a) w_{eq1} ; 4.17(b) w_{eq2} ; 4.17(c) w_{eq3} .



(a)



(b)



(c)

Fig. 4.18: Coeficientes do equalizador obtidos através do algoritmo RLS para um canal variante no tempo baseado em degrau, tendo a solução de Wiener como referência: 4.18(a) w_{eq1} ; 4.18(b) w_{eq2} ; 4.18(c) w_{eq3} .

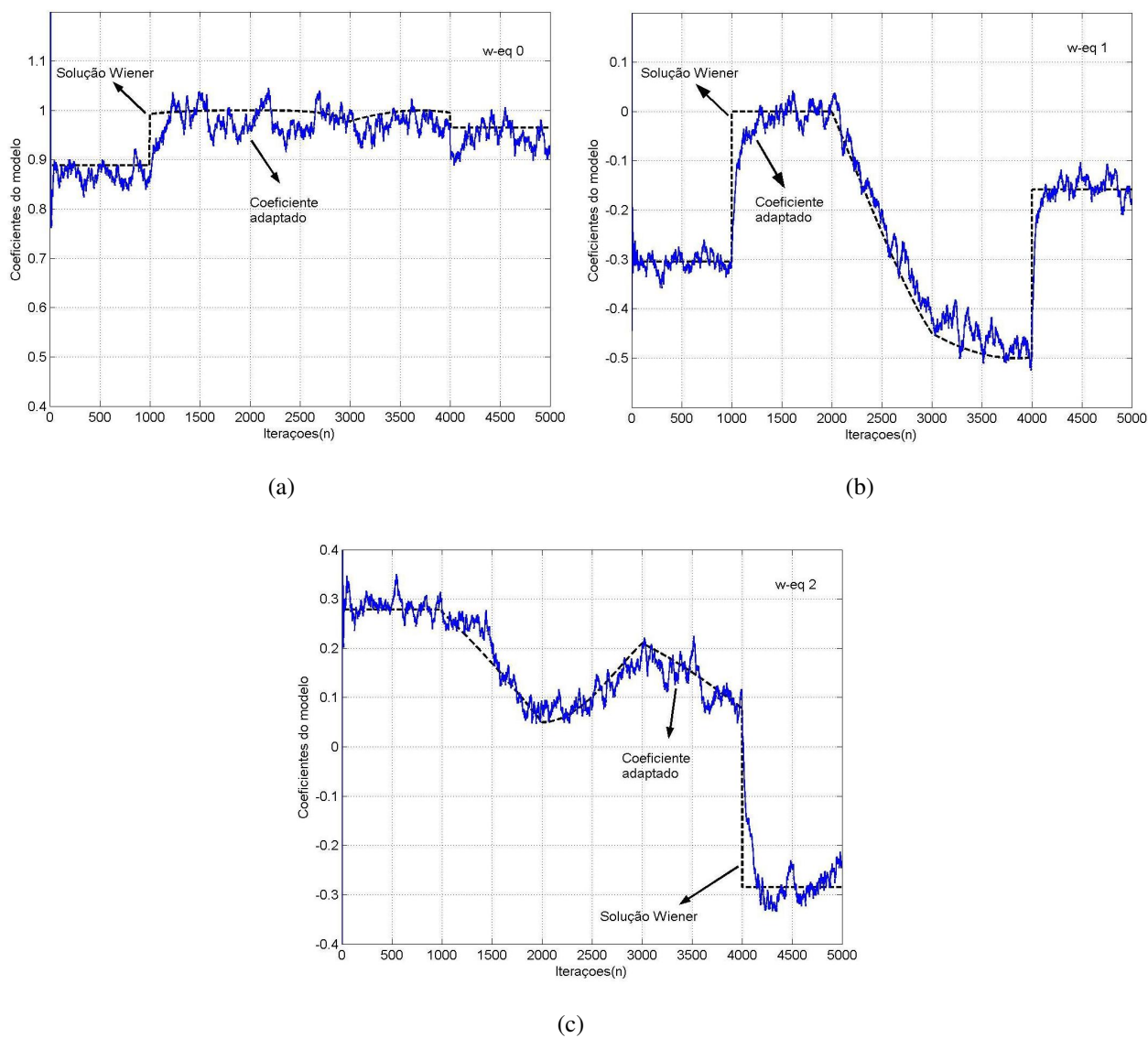


Fig. 4.19: Coeficientes do equalizador obtidos através do algoritmo RLS para um canal variante no tempo baseado em degrau-rampa, tendo a solução de Wiener como referência: 4.19(a) w_{eq1} ; 4.19(b) w_{eq2} ; 4.19(c) w_{eq3} .

As figuras 4.18 e 4.19 ilustram a estimativa adaptativa obtida pelo equalizador através do algoritmo RLS para os coeficientes dos canais baseados em degrau e degrau-rampa, respectivamente. Nas duas simulações, foi utilizado um fator de esquecimento $\lambda = 0,98$ e uma média de 200 realizações do algoritmo, justificado pelo fato de reduzir as oscilações nas figuras, melhorando a visualização. Também neste caso o algoritmo atinge rapidamente os pontos ótimos de convergência, rastreado as variações do canal.

Os desempenhos dos algoritmos adaptativos LMS e RLS, para uma média de 200 realizações, foram muito semelhantes. Ambos atingiram rapidamente a convergência, sendo capazes de acompanhar as variações do canal. No entanto, percebe-se que o RLS atinge o ponto de convergência com um número menor de iterações. Este desempenho depende, obviamente, de uma boa escolha do fator de esquecimento λ . Já o LMS pode ser ajustado alterando-se o passo de adaptação μ ; porém, para valores grandes de μ , aumenta-se a velocidade de convergência às custas de uma maior oscilação estocástica.

Em todas as simulações houve 100% de convergência global e o método de equalização proposto mostrou-se, assim como no caso de identificação, capaz de acompanhar de maneira eficaz as variações dos canais estudados.

4.4 Equalização de Sinais Correlacionados

Conforme discutido anteriormente, o ponto ótimo de convergência do equalizador que inverte a função de transferência do canal somente é atingido quando o sinal de transmissão é descorrelacionado. Demonstramos que, para um sinal desta natureza, a matriz de correlação na entrada do equalizador depende somente da matriz de convolução $\mathbf{H}$ do canal, conforme pode ser visto na equação 4.22.

Para o caso de sinais correlacionados⁹, o termo $E [\mathbf{t}(n) \mathbf{t}^T(n)]$ em 4.15 não será mais uma matriz diagonal. Neste caso, tem-se:

$$\mathbf{R} = \alpha_m^2 \mathbf{H} \mathbf{H}^T + \mathbf{H} E [\mathbf{t}(n) \mathbf{t}^T(n)] \mathbf{H}^T \quad (4.33)$$

Com a matriz de autocorrelação apresentada em 4.33, os coeficientes do equalizador de Wiener não correspondem ao inverso do canal, afastando-se deste à medida que o sinal de informação é mais correlacionado. Uma alternativa para que a solução proposta opere a contento é utilizar um método de branqueamento do sinal $t(n)$.

A figura 4.20 apresenta um esquema de equalização onde um filtro de erro de predição é intro-

⁹Apresenta correlação entre o sinal e versões atrasadas do mesmo.

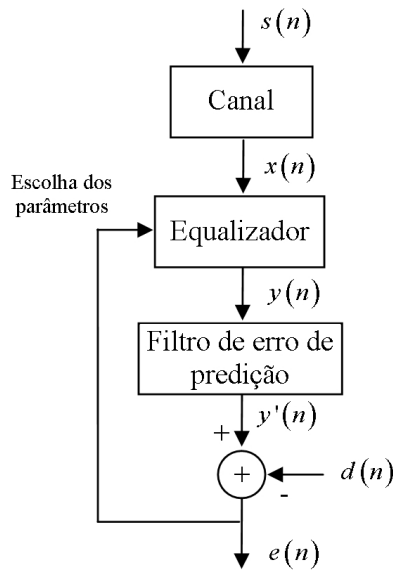


Fig. 4.20: Esquema de equalização utilizando marca d'água para sinais correlacionados.

duzido na saída do equalizador, antes do cálculo do sinal de erro $e(n)$. Este filtro tem como objetivo branquear o sinal $y(n)$ através da filtragem de suas componentes redundantes. A saída do preditor ilustrado na figura 4.21 é composta apenas pela parte correlacionada do sinal $y(n)$. Calculando a diferença entre $y(n)$ e o sinal na saída do preditor, obtemos $y'(n)$, que corresponde às componentes descorrelacionadas na saída do equalizador. Esta operação pode ser expressa da seguinte forma:

$$y'(n) = y(n) - \sum_{k=1}^M w_{f,k} y(n - k - \Delta) \quad (4.34)$$

onde $w_{f,k}$ são os coeficientes do preditor com M elementos de atraso. O atraso das amostras, denotado por Δ , é chamado de *atraso de descorrelação*. O termo

$$\sum_{k=1}^M w_{f,k} y(n - \Delta - k) \quad (4.35)$$

corresponde a uma estimativa de $y(n)$ com base em seus valores passados, ou seja, contém as componentes redundantes do sinal. Quando os parâmetros do equalizador tendem aos ótimos, o sinal de erro $y'(n)$ tende às componentes descorrelacionadas do sinal de informação, mais a marca d'água que, por sua natureza, é um sinal descorrelacionado. Assume-se que o atraso de descorrelação Δ é suficientemente grande para que seja possível remover a autocorrelação do sinal $y(n)$ na saída do equalizador.

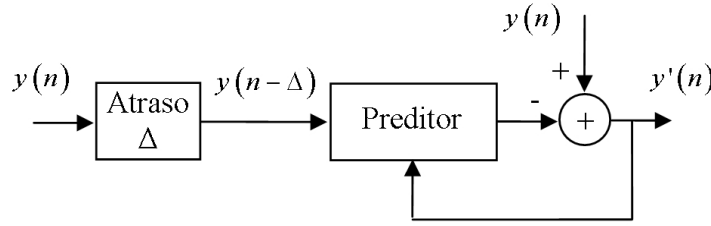


Fig. 4.21: Filtro de erro de predição.

4.4.1 Simulações e Resultados

Nessa seção, apresentamos resultados de simulações para equalização de sinais discretos de áudio que possuem alta correlação, buscando ilustrar a aplicação do processo de branqueamento discutido na seção anterior.

No ensaio, escolhemos o mesmo canal expresso em 4.31 e um equalizador com cinco coeficientes inicializados na origem. Três diferentes amostras de áudio foram utilizadas como sinais de informação (guitarra, trompete e rock, ambas com taxa de amostragem de 8 kHz). O filtro de erro de predição foi modelado como um filtro linear FIR com 750 coeficientes e um atraso Δ de 800 amostras. Este atraso foi determinado supondo que os sinais de áudio apresentam correlação não desprezível consigo mesmos para atrasos de até 100 ms. O equalizador foi otimizado através do algoritmo RLS com fator de esquecimento $\lambda = 1$, e o filtro de erro de predição otimizado através do algoritmo LMS com passo $\mu = 0,0001$. Essa configuração apresentou 100% de convergência nos casos analisados.

O EQM versus SWR, após a convergência, é apresentado nas figuras 4.22, 4.23 e 4.24 para os sinais de guitarra, trompete e rock, respectivamente. Como discutido na seção 4.3.1, os resultados indicam que o EQM residual médio cresce de forma aproximadamente exponencial com a SWR. Para uma SWR na faixa de 20-23 dB, o EQM apresenta um valor relativamente baixo e próximo dos valores ótimos (EQM = 0,0873 para a amostra de guitarra, EQM = 0,0617 para a amostra de trompete e EQM = 0,0712 para a amostra de rock, considerando um equalizador ótimo de quinta ordem). Assim como nas simulações anteriores, o EQM cresce rapidamente para SWR na faixa de 25-30 dB.

Podemos concluir, com base nos resultados apresentados, que a introdução do branqueamento viabiliza o uso do método de equalização baseado em marca d'água com sinais fortemente correlacionados.

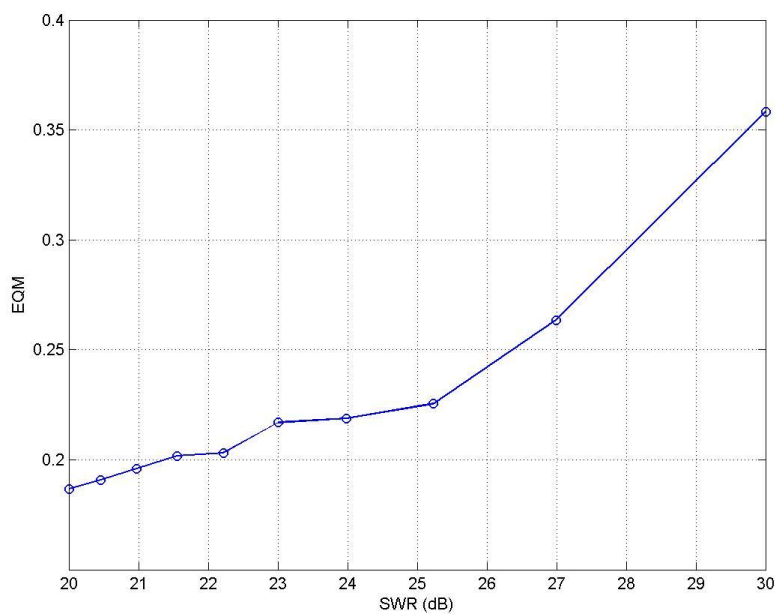


Fig. 4.22: Erro quadrático médio de equalização com o algoritmo RLS em função da relação sinal-marca d'água (sinal de guitarra).

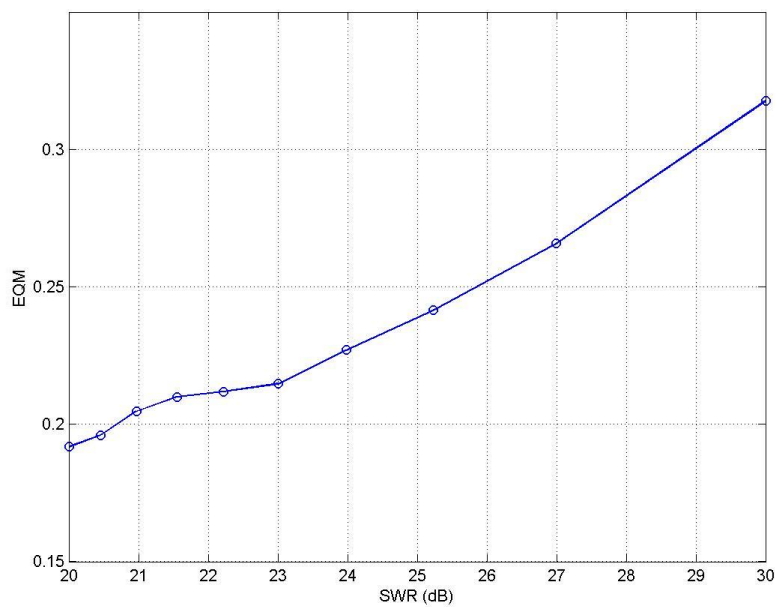


Fig. 4.23: Erro quadrático médio de equalização com o algoritmo RLS em função da relação sinal-marca d'água (sinal de trompete).

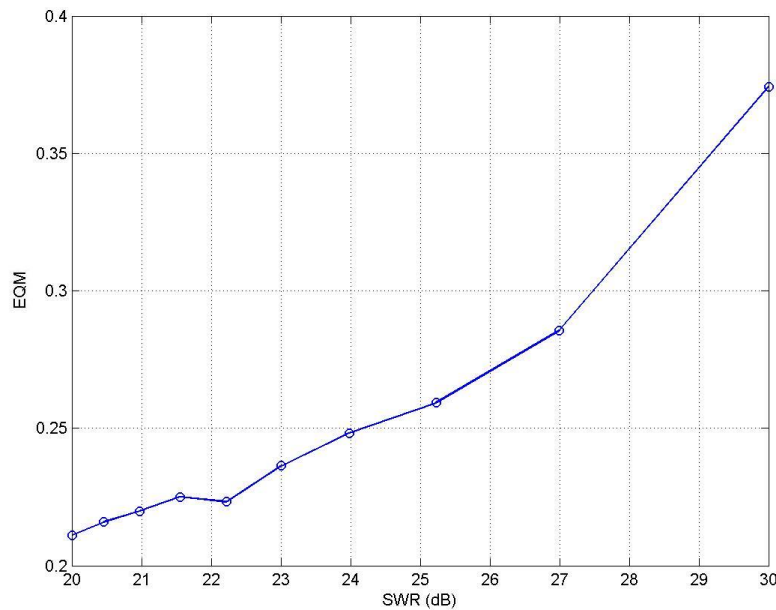


Fig. 4.24: Erro quadrático médio de equalização com o algoritmo RLS em função da relação sinal-marca d'água (sinal de rock).

4.5 Síntese e Conclusões

Neste capítulo, foi proposto um método de equalização supervisionada utilizando uma marca d'água como sinal de referência. O método adiciona uma marca d'água ao sinal de informação que se deseja transmitir e, sem interrupções na transmissão de informação útil, utiliza continuamente o sinal da marca d'água no equalizador para reverter os efeitos do canal de transmissão sobre o sinal transmitido.

Foi demonstrado que, caso o sinal de informação seja branco, é necessário estimar e compensar um fator de escala no equalizador. Já no caso em que o sinal de informação é correlacionado, faz-se necessário o uso de um método de branqueamento em conjunto com o equalizador, para que seja possível a adequada inversão da função de transferência do canal.

O método proposto apresenta vantagens em relação às técnicas tradicionais de equalização adaptativa, destacando-se a ausência de períodos de treinamento nos quais a transmissão do sinal de informação é interrompida, e a possibilidade de rastrear variações do canal, pois o filtro é continuamente adaptado. Por se tratar de um método supervisionado, foi constatado que o algoritmo proposto apresenta menor complexidade computacional em relação à maioria dos critérios não-supervisionados e função custo unimodal. Além disso, a abordagem proposta não se limita à aplicação clássica de transmissão de dados, tendo sido demonstrado o seu uso potencial em outras aplicações (e.g. equalização

de áudio). O método proposto mostrou-se adequado tanto para canais invariantes como variantes no tempo, independente das características do sinal de informação utilizado.

No âmbito da equalização de sinais de áudio, nossa proposta será estendida mais adiante para contemplar o uso de técnicas psicoacústicas, visando garantir que as modificações impostas ao sinal de áudio sejam imperceptíveis, isto é, que a marca d'água seja transparente.

Capítulo 5

Identificação Supervisionada Utilizando Marca d'Água

5.1 Introdução

Apresentaremos neste capítulo um método de identificação de canais baseado no uso de uma marca d'água digital como sinal de referência. A marca d'água, conhecida no receptor, é utilizada para estimar as características do canal, tornando possível associá-lo a um modelo adequado. Ambos os sinais (de marca d'água e de informação) são transmitidos ininterruptamente, de modo que não haja interrupções na comunicação e os coeficientes do modelo de identificação possam ser atualizados continuamente.

Uma técnica similar à aqui proposta é conhecida como *superimposed pilot training* (Tugnait and Luo, 2003; Varma et al., 2004; Tugnait and Meng, 2004; Chen and Zhou, 2004). O paradigma que norteia esta técnica consiste em sobrepor ao sinal de informação, por adição aritmética, uma sequência de dados piloto conhecida; esta sequência é detectada no receptor e usada para identificar o canal de transmissão¹. A técnica *superimposed training* não está focada na transparência da sequência piloto; isso significa que um receptor que não esteja preparado para reconhecer a sequência piloto irá percebê-la como um ruído que pode afetar diretamente o desempenho do sistema. Por outro lado, o paradigma da marca d'água é focado na transparência, assegurando compatibilidade com receptores já existentes que não estejam preparados para reconhecer a marca (ditos *receptores legados*). Embora também encarada como ruído, a marca d'água é idealizada de maneira que sua influência na detecção e percepção do sinal de informação isolado seja insignificante. Outro diferencial em relação à técnica *superimposed training* é a abordagem genérica dada ao método proposto, que pode ser aplicado a

¹Nos trabalhos relativos a *superimposed training* aos quais tivemos acesso, a técnica é sempre aplicada a sistemas de comunicação multiportadora (e.g. OFDM), sendo a sequência piloto inserida no sinal após a etapa de modulação.

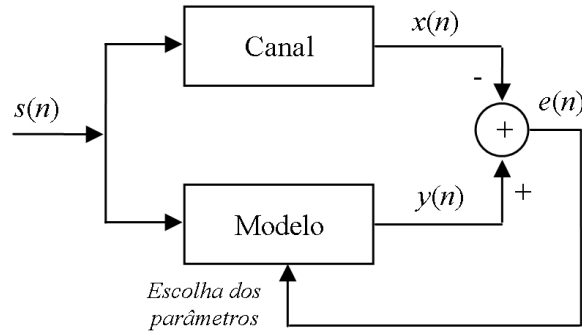


Fig. 5.1: Esquema clássico de identificação supervisionada.

sinais de qualquer natureza, incluindo sinais com interpretação sensorial (e.g. áudio e imagens).

5.2 Identificação Supervisionada Baseada em Marca d'Água: Sistema Proposto

Nos sistemas de comunicação, a identificação consiste em modelar um canal de comunicação que ignoramos, ou seja, obter um modelo adequado e fiel deste canal. Um sistema de identificação é composto de um filtro que, quando localizado no receptor, tenta modelar o canal de comunicação através do qual o sinal foi transmitido. Como já mencionado, muitas vezes um sistema de identificação pode ser utilizado de forma hierárquica com um sistema de equalização; em outras palavras, é possível projetar um equalizador a partir do modelo identificado.

Antes de apresentar nosso método baseado em marca d'água, revisamos a filtragem de Wiener, agora no contexto de identificação, para novamente tê-la como referência para a técnica proposta.

5.2.1 Identificação de Wiener

O esquema clássico de identificação é apresentado na figura 5.1. Quando os coeficientes do modelo são adequadamente ajustados, o sinal de erro $e(n)$ é minimizado, implicando em um sinal $y(n)$ na saída do modelo o mais próximo possível do sinal $x(n)$ na saída do canal, de acordo com o critério estabelecido. Uma medida de erro comumente empregada para a adequação do modelo às características do canal é o erro quadrático médio, que expressa o critério de Wiener:

$$J = E [|e(n)|^2] \quad (5.1)$$

A solução ótima para o problema de identificação é obtida através da minimização da função custo com respeito aos coeficientes $\mathbf{w}$ do filtro-modelo, levando aos coeficientes ótimos $\mathbf{w}_0$ do modelo,

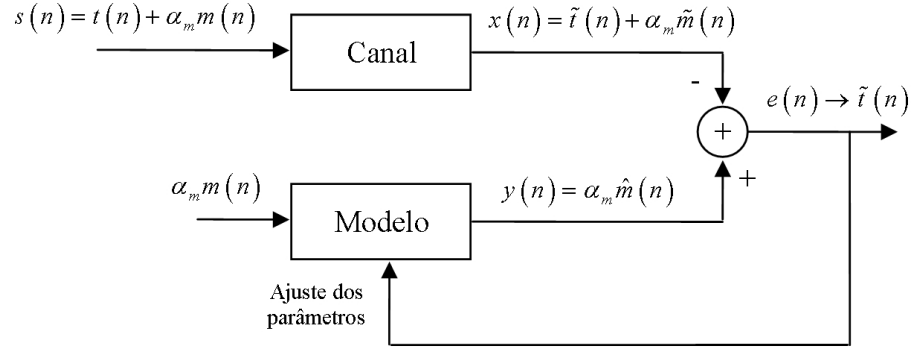


Fig. 5.2: Esquema de identificação utilizando marca d'água.

dados por:

$$\mathbf{w}_o = \arg \left(\min_{\mathbf{w}} E \left[|e(n)|^2 \right] \right) \quad (5.2)$$

5.2.2 Descrição do Algoritmo Proposto

Em nosso método, pretende-se identificar um canal de comunicação transmitindo continuamente uma marca d'água sobreposta ao sinal de informação, utilizando a técnica de espalhamento espectral, conforme já apresentado na seção 4.2.2. Assumindo um canal linear, o sinal na sua saída é:

$$d(n) = x(n) = \tilde{t}(n) + \alpha_m \tilde{m}(n) \quad (5.3)$$

sendo $\tilde{t}(n)$ e $\tilde{m}(n)$ versões distorcidas pelo canal do sinal de informação e da marca d'água, respectivamente. De forma similar, após a convergência do modelo (também assumido linear), a sua saída $y(n)$ pode ser vista como uma estimativa $\hat{m}(n)$ do sinal $\tilde{m}(n)$. Este esquema é ilustrado na figura 5.2. O sinal de erro $e(n)$ é expresso, nesse caso, por:

$$e(n) = y(n) - x(n) = \alpha_m (\hat{m}(n) - \tilde{m}(n)) - \tilde{t}(n) \quad (5.4)$$

A função custo J , definida como a esperança do sinal de erro quadrático $e^2(n)$, é dada por²:

$$J = E \left[e^2(n) \right] = E \left[\left(\alpha_m (\hat{m}(n) - \tilde{m}(n)) - \tilde{t}(n) \right)^2 \right] \quad (5.5)$$

A equação 5.5 pode ser reescrita em notação matricial. Definimos $\mathbf{m}(n)$ e $\mathbf{t}(n)$ nas equações 4.8 e 4.9; porém, M representa agora o número de coeficientes do modelo. Os sinais $x(n)$ e $y(n)$ podem ser representados como a convolução dos sinais nas entradas do canal e do modelo com as respectivas

²Nossas definições, por simplicidade, têm sempre por base sinais reais.

respostas impulsivas:

$$x(n) = \mathbf{h}^T (\mathbf{t}(n) + \alpha_m \mathbf{m}(n)) \quad (5.6)$$

$$y(n) = \alpha_m \mathbf{w}^T \mathbf{m}(n) \quad (5.7)$$

onde $\mathbf{h} = [h_0 \cdots h_{L-1}]$ expressa os coeficientes do canal, suposto de resposta impulsiva finita (FIR) com função de transferência de ordem $L - 1$, e $\mathbf{w} = [w_0 \cdots w_{M-1}]$ representa os coeficientes do modelo FIR de comprimento M , para $M = L$. Com essas definições, o sinal de erro pode ser representado da seguinte forma:

$$e(n) = y(n) - x(n) = \alpha_m \mathbf{w}^T \mathbf{m}(n) - \mathbf{h}^T (\mathbf{t}(n) + \alpha_m \mathbf{m}(n)) \quad (5.8)$$

Ao se determinar a esperança do erro quadrático, chegamos à expressão:

$$\begin{aligned} J(\mathbf{w}) = E[x^2(n)] - \mathbf{w}^T E[\alpha_m^2 \mathbf{m}(n) \mathbf{h}^T \mathbf{m}(n)] - E[\alpha_m^2 \mathbf{h}^T \mathbf{m}(n) \mathbf{m}^T(n)] \mathbf{w} \\ + \mathbf{w}^T E[\alpha_m^2 \mathbf{m}(n) \mathbf{m}^T(n)] \mathbf{w} \end{aligned} \quad (5.9)$$

Em decorrência da técnica de espalhamento espectral empregada na inserção da marca d'água, os sinais $t(n)$ e $\alpha_m m(n)$ são independentes entre si; em decorrência da linearidade, podemos definir a relação:

$$\alpha_m \tilde{m}(n) = \alpha_m \mathbf{h}^T \mathbf{m}(n) \quad (5.10)$$

Analisando os termos da equação 5.9 e de posse da relação expressa em 5.10, podemos reescrevê-los da seguinte forma:

$$\begin{aligned} E[x^2(n)] &= \sigma_x^2 \\ E[\alpha_m^2 \mathbf{m}(n) \mathbf{h}^T \mathbf{m}(n)] &= E[\alpha_m \mathbf{m}(n) \tilde{m}(n)] = \mathbf{p}_m \\ E[\alpha_m^2 \mathbf{m}(n) \mathbf{m}^T(n)] &= \mathbf{R}_m \end{aligned} \quad (5.11)$$

onde assume-se $t(n)$ e $m(n)$ com média nula; os termos $\mathbf{R}_m$ e $\mathbf{p}_m$ expressam, respectivamente, a autocorrelação e a correlação cruzada do sinal da marca d'água.

Substituindo essas relações na equação 5.9, obtemos a superfície de erro do critério de Wiener na forma canônica para o sistema de identificação utilizando marca d'água:

$$J(\mathbf{w}) = \sigma_x^2 - \mathbf{w}^T \mathbf{p}_m - \mathbf{p}_m^T \mathbf{w} + \mathbf{w}^T \mathbf{R}_m \mathbf{w} \quad (5.12)$$

Para obter o ponto de mínimo desta função custo, é necessário minimizá-la com respeito aos coeficientes $\mathbf{w}$ do modelo. Como $J(\mathbf{w})$ expressa um quadrado perfeito, é possível reescrever 5.12 da seguinte forma:

$$J(\mathbf{w}) = \sigma_x^2 - \mathbf{p}_m^T \mathbf{R}_m^{-1} \mathbf{p}_m + \left(\mathbf{w} - \mathbf{R}_m^{-1} \mathbf{p}_m \right)^T \mathbf{R}_m \left(\mathbf{w} - \mathbf{R}_m^{-1} \mathbf{p}_m \right) \quad (5.13)$$

Para atingir o ponto de mínimo na equação 5.13, o termo $\left(\mathbf{w} - \mathbf{R}_m^{-1} \mathbf{p}_m \right)$ deve igualar-se a zero. Tem-se então:

$$\mathbf{w}_0 - \mathbf{R}_m^{-1} \mathbf{p}_m = 0 \quad \Rightarrow \quad \mathbf{w}_0 = \mathbf{R}_m^{-1} \mathbf{p}_m \quad (5.14)$$

Com o ponto de mínimo expresso em 5.14, obtemos o valor mínimo da função custo:

$$\min_{\mathbf{w}} J(\mathbf{w}) = \sigma_x^2 - \mathbf{p}_m^T \mathbf{R}_m^{-1} \mathbf{p}_m \quad (5.15)$$

O ponto de mínimo obtido depende apenas das características do sinal da marca d'água, bastando o conhecimento prévio deste sinal para que seja possível o ajuste dos coeficientes do modelo. Isso significa que, independentemente das características do sinal de informação, o filtro utilizado como modelo converge para um ponto que representa o canal tão bem quanto possível segundo o critério de Wiener. Esta demonstração evidencia a viabilidade do uso da marca d'água para identificação supervisionada.

Nota-se em 5.12 e 5.15 que a influência do sinal de informação sobre a superfície de erro limita-se a um deslocamento no sentido do custo J , representado por σ_x^2 . Como demonstrado, esse deslocamento não influencia o ponto de mínimo da função.

5.3 Simulações e Resultados

Tendo por objetivo avaliar o método aqui proposto, foram realizados diversos ensaios. Foram usados os algoritmos LMS e RLS, sinais com diferentes modulações e canais FIR com diferentes funções de transferência. Além disso, buscamos determinar os limiares da relação sinal-marca d'água (SWR) com os quais o sistema pode trabalhar.

Como primeiro exemplo, consideremos um canal modelado como um filtro linear de fase mínima com a seguinte função de transferência:

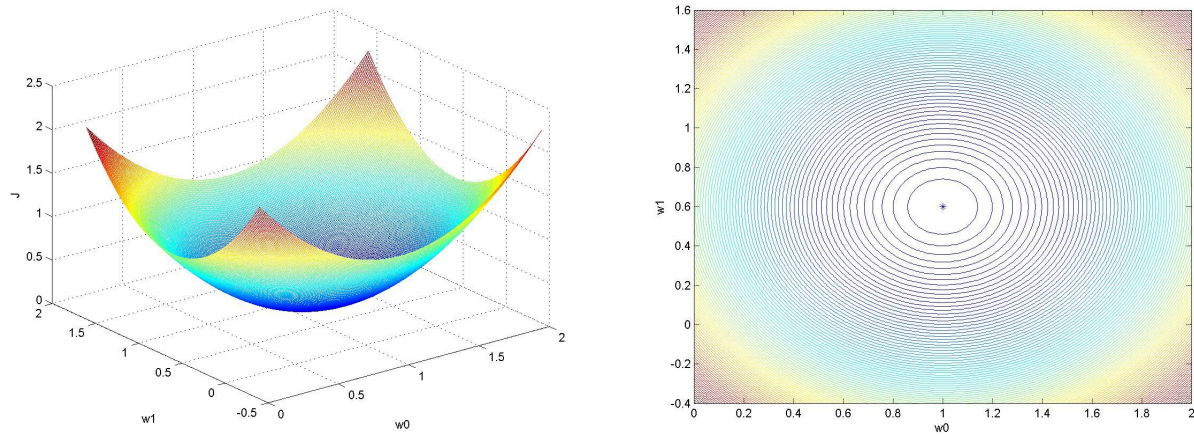


Fig. 5.3: Função custo e curvas de nível para o esquema de identificação utilizando marca d'água.

$$H_1(z) = 1 + 0,6z^{-1} \quad (5.16)$$

Assumamos que as amostras do sinal transmitido através do canal sejam i.i.d. e pertençam a uma constelação binária 2-PSK de módulo unitário e obtidas através da função *randsrc* no software de simulação MATLAB, e suponhamos que não haja ruído aditivo e que o sistema encontra-se sincronizado. Tomemos um modelo de identificação com o mesmo número de parâmetros do canal e coeficientes inicializados na origem.

Na figura 5.3 está traçada a superfície de erro para o modelo do exemplo, obtida pelo critério de Wiener utilizando marca d'água. Como esperado, a função custo é unimodal e o ponto de mínimo corresponde aos coeficientes do canal. O sucesso desta abordagem depende de uma escolha adequada de potências, refletida na relação sinal-marca d'água. A presença de ruído não deve modificar a posição do ponto ótimo de Wiener, pois se espera que não haja correlação entre o ruído e o sinal. Entretanto, para algoritmos estocásticos como o LMS e o RLS, o ruído interfere no transiente do processo. Do ponto de vista da marca d'água, o sinal de informação é encarado como um forte ruído aditivo, sendo necessário analisar o desempenho dos algoritmos para diferentes valores de SWR.

Na figura 5.4, mostramos a evolução temporal dos coeficientes do modelo para uma realização do algoritmo LMS com SWR = 0 dB e $\mu = 0,001$. Nota-se que, em regime permanente, as curvas convergem para a solução de Wiener. Em todos os ensaios utilizando essa SWR, houve 100% de convergência global; porém, na prática, sistemas baseados em marca d'água requerem valores bem maiores de SWR devido à condição de transparência.

A figura 5.5 apresenta duas realizações do LMS para SWR de 10 dB, com $\mu = 0,001$ e $\mu = 0,005$. Com base nestes resultados, podemos destacar duas características do método proposto, sendo pos-

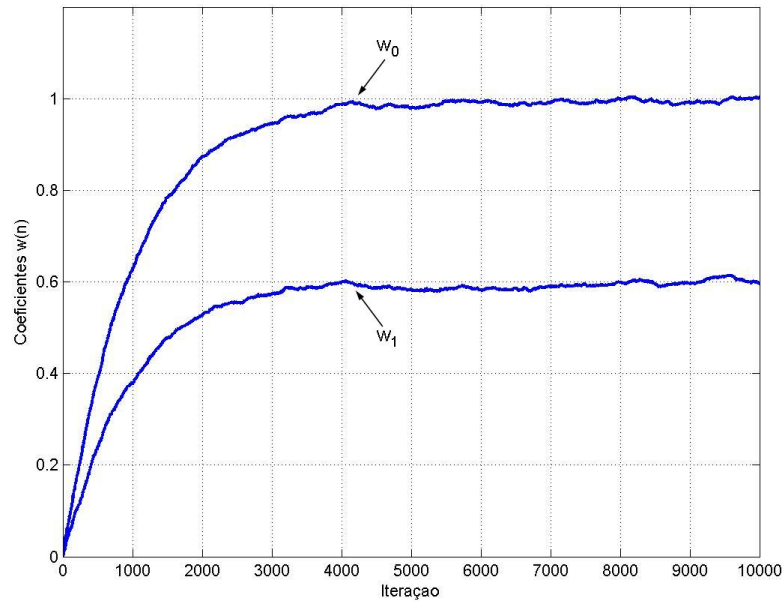
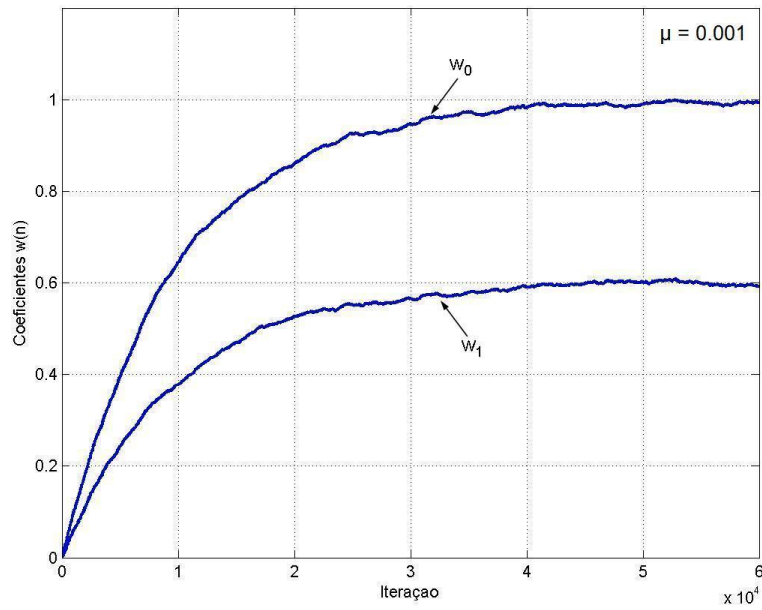


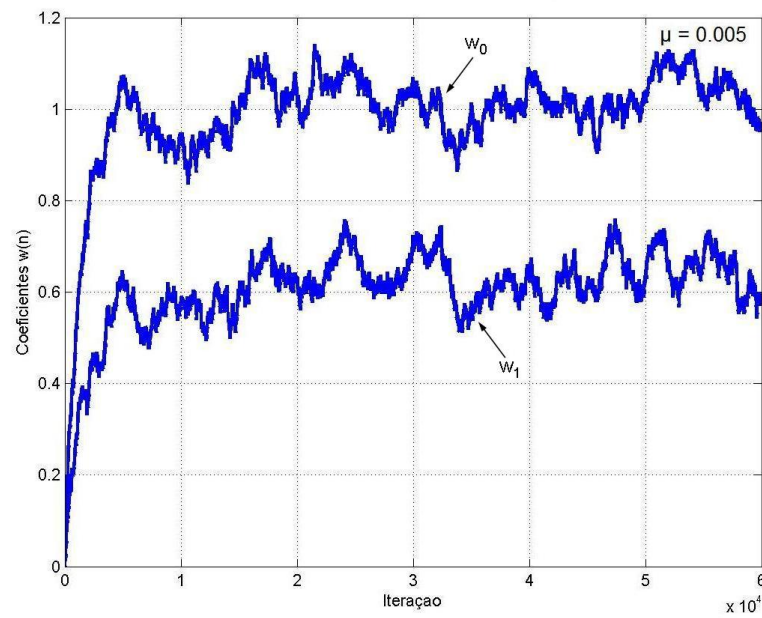
Fig. 5.4: Evolução temporal do algoritmo LMS ($\text{SWR} = 0 \text{ dB}$, $\mu = 0,001$).

sível observar diferenças consideráveis em relação ao comportamento do algoritmo para o caso correspondente em equalização apresentado no capítulo anterior. A primeira característica está relacionada à velocidade de convergência. Nota-se que a convergência foi bem mais lenta que a apresentada na figura 5.4, mesmo mantendo-se o valor do passo de adaptação. Para $\text{SWR} = 0 \text{ dB}$, o ponto de mínimo foi atingido após cerca de 4.000 iterações; já para $\text{SWR} = 10 \text{ dB}$, o número de iterações para convergência foi cerca de dez vezes maior. No caso de equalização, a velocidade de convergência não sofreu forte influência com o aumento da SWR. Seria intuitivo imaginar que, para aumentar a velocidade de convergência do LMS, bastaria aumentar o passo de adaptação, levando a um resultado equivalente para ambos os valores de SWR. Este raciocínio, no entanto, não é totalmente válido, decorrendo daí a segunda característica observada nos resultados: o sinal de informação (que para $\text{SWR} = 10 \text{ dB}$ é 10 vezes mais potente que a marca d'água) representa um forte ruído no LMS, implicando em flutuações estocásticas que tornam a estimativa menos precisa. Pequenos aumentos no passo de adaptação do algoritmo provocam o rápido aumento destas flutuações estocásticas. Na figura 5.5(b), para $\mu = 0,005$, é nítida a presença de flutuações estocásticas, e a convergência foi apenas cinco vezes mais rápida do que para $\mu = 0,001$. Já no caso de equalização (vide figura 4.5), um pequeno aumento no passo de adaptação provocou um considerável ganho de velocidade sem introduzir um grande aumento nas flutuações estocásticas.

Uma relação sinal-marca d'água de 10 dB pode não ser suficiente para garantir a transparência em sistemas de marca d'água digital. Sistemas de marca d'água para áudio, por exemplo, utilizam



(a)



(b)

Fig. 5.5: Evolução temporal do algoritmo LMS para diferentes passos μ (SWR = 10 dB): 5.5(a) $\mu = 0,001$; 5.5(b) $\mu = 0,005$.

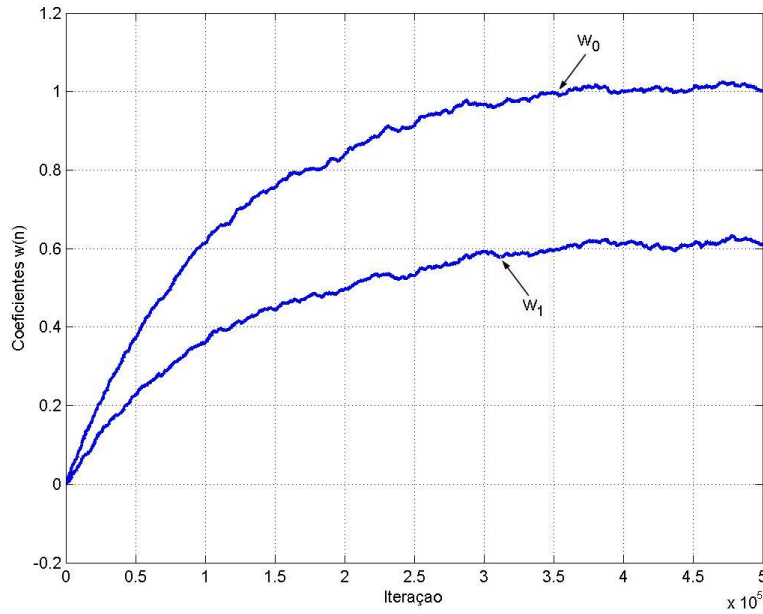


Fig. 5.6: Evolução temporal do algoritmo LMS (SWR = 20 dB, $\mu = 0,001$).

geralmente uma relação sinal-marca d'água da ordem de 20 dB (Gomes et al., 2000, 2001, 2003). A figura 5.6 apresenta uma realização do LMS para SWR = 20 dB e passo $\mu = 0,001$. Nota-se que a convergência é atingida após cerca de 400.000 iterações.

No esquema de identificação apresentado, tem-se na entrada do modelo apenas o sinal da marca d'água $\alpha_m m(n)$; sendo assim, os autovalores da matriz de autocorrelação $\mathbf{R}_m$ dependem apenas do sinal de marca d'água. A constante de tempo $(\tau)_{EQM,me}$ (apresentada na equação 4.26) tem um comportamento inversamente proporcional à potência da marca d'água: à medida que diminui a potência da marca, diminui proporcionalmente o valor de λ_{me} (pois diminuem os autovalores da matriz), fazendo com que a constante de tempo aumente proporcionalmente. Isso justifica o fato de a simulação com SWR = 20 dB ter apresentado velocidade de convergência 100 vezes inferior à simulação com SWR = 0 dB.

Os resultados apresentados permitem concluir que a velocidade de convergência do algoritmo LMS torna pouco prático o seu uso no método de identificação adaptativa utilizando marca d'água; testes realizados com sistemas FIR mais complexos mostraram que a quantidade de amostras necessária para atingir a convergência o torna impraticável. Além disso, pequenas variações no passo de adaptação provocam o rápido aumento das flutuações estocásticas, devido o sinal de informação representar um forte ruído no LMS; esse é o preço pago por o LMS ter por base uma aproximação estocástica.

Passemos agora à análise do algoritmo RLS. A figura 5.7 ilustra a evolução temporal dos coefi-

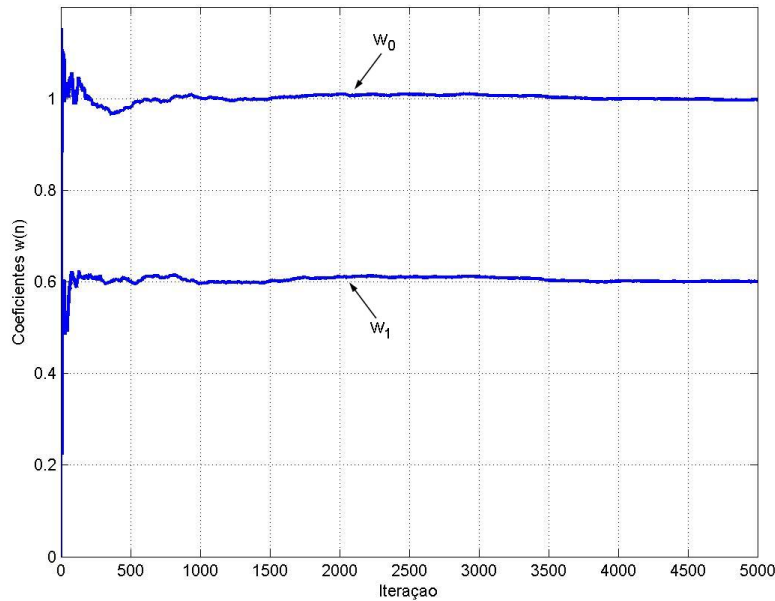
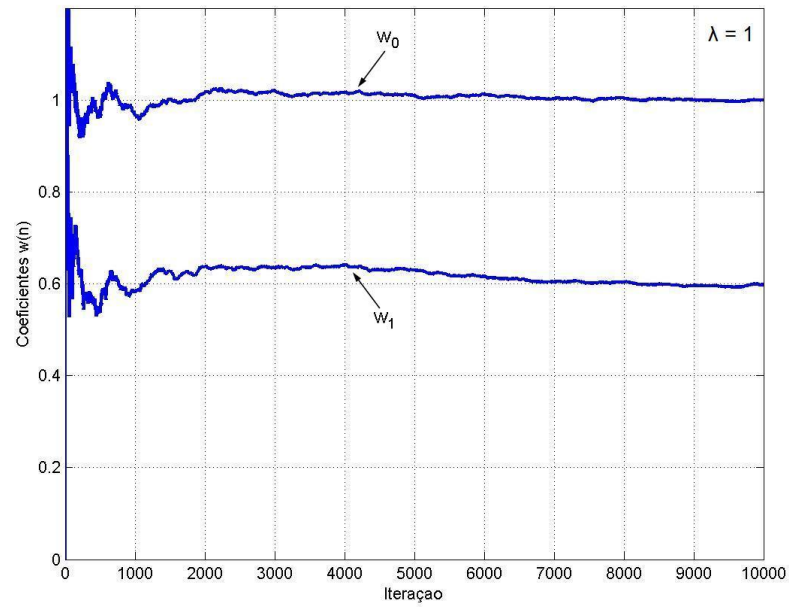


Fig. 5.7: Evolução temporal do algoritmo RLS (SWR = 0 dB, $\lambda = 1$).

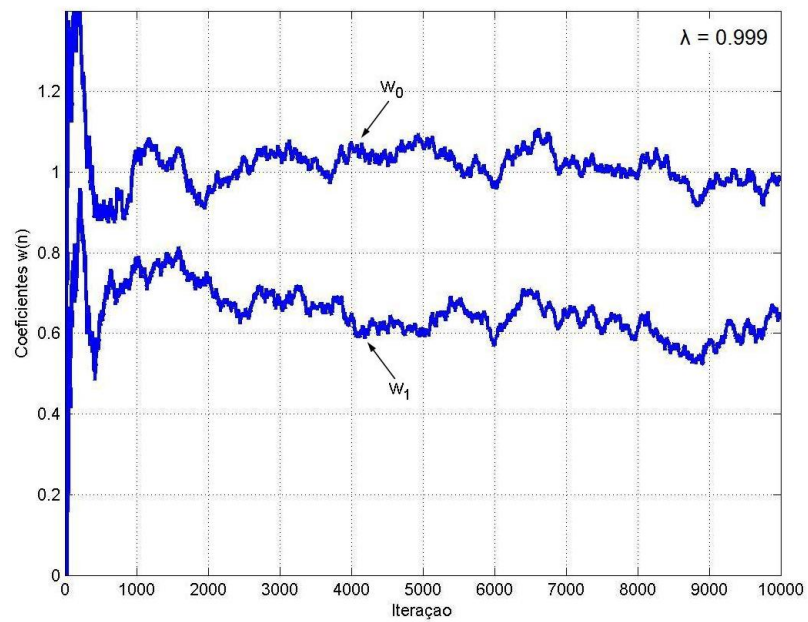
cientes do modelo para uma realização do algoritmo RLS com relação sinal-marca d'água de 0 dB e fator de esquecimento $\lambda = 1$; vê-se que o RLS atingiu o ponto de mínimo após cerca de 1.000 iterações. Em todos os ensaios utilizando essa SWR, observou-se 100% de convergência global. Porém, como dito anteriormente, sistemas baseados em marca d'água geralmente requerem valores bem maiores de SWR para que a propriedade de transparência seja satisfeita. Faremos então uma avaliação da influência deste parâmetro no desempenho do algoritmo.

A figura 5.8 ilustra duas realizações do RLS para SWR = 10 dB, com $\lambda = 1$ e $\lambda = 0,999$. A comparação com a figura 5.7 permite dizer que, com relação à velocidade de convergência, a diferença entre os ensaios com SWR = 0 dB e 10 dB foi significativamente menor do que a verificada para o algoritmo LMS. Para SWR = 10 dB, o RLS atingiu o ponto de mínimo após cerca de 7.000 iterações, valor este muito inferior ao observado com o LMS. Conforme discutido, o sinal de informação representa um forte ruído, implicando em flutuações estocásticas que tornam as estimativas menos precisas. Esse efeito é claramente observado na figura 5.8: para $\lambda = 0,999$, é nítida a presença de ruído. No entanto, apesar do ruído apresentar-se superior que no caso de equalização para as mesmas condições (vide figura 4.8), este mantém-se estável em torno do ponto ótimo. Este ponto será discutido em maior detalhe na seção 5.3.1, referente a sistemas variantes no tempo.

Como os sistemas de marca d'água de áudio geralmente trabalham com uma relação sinal-marca d'água da ordem de 20 dB, as figuras 5.9 e 5.10 ilustram a evolução do RLS para SWR = 20 dB e 23 dB, respectivamente. Em ambos os ensaios, o algoritmo RLS convergiu rapidamente para os



(a)



(b)

Fig. 5.8: Evolução temporal do algoritmo RLS para diferentes valores de λ (SWR = 10 dB): 5.8(a) $\lambda = 1$; 5.8(b) $\lambda = 0,999$.

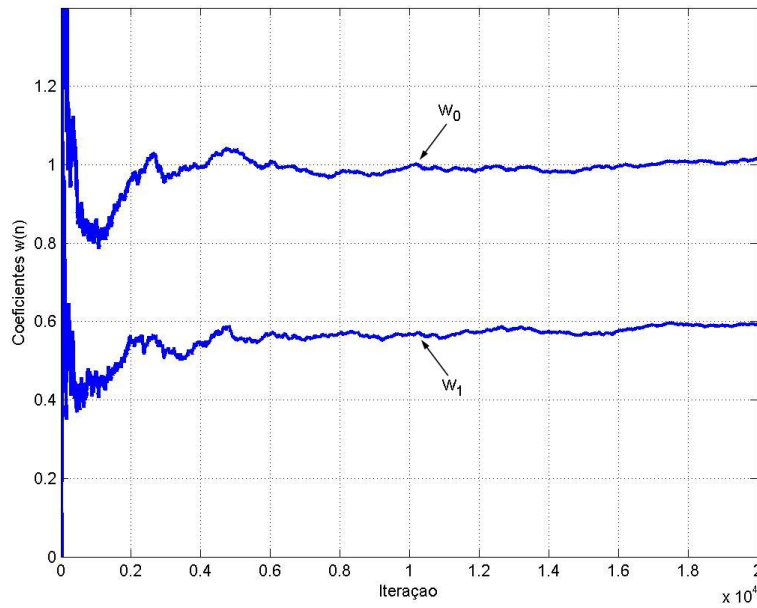


Fig. 5.9: Evolução temporal do algoritmo RLS (SWR = 20 dB, $\lambda = 1$).

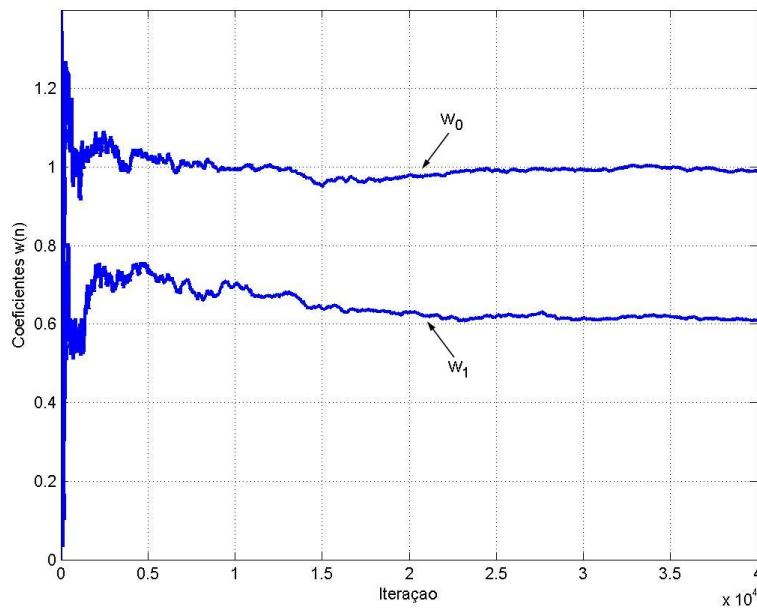


Fig. 5.10: Evolução temporal do algoritmo RLS (SWR = 23 dB, $\lambda = 1$).

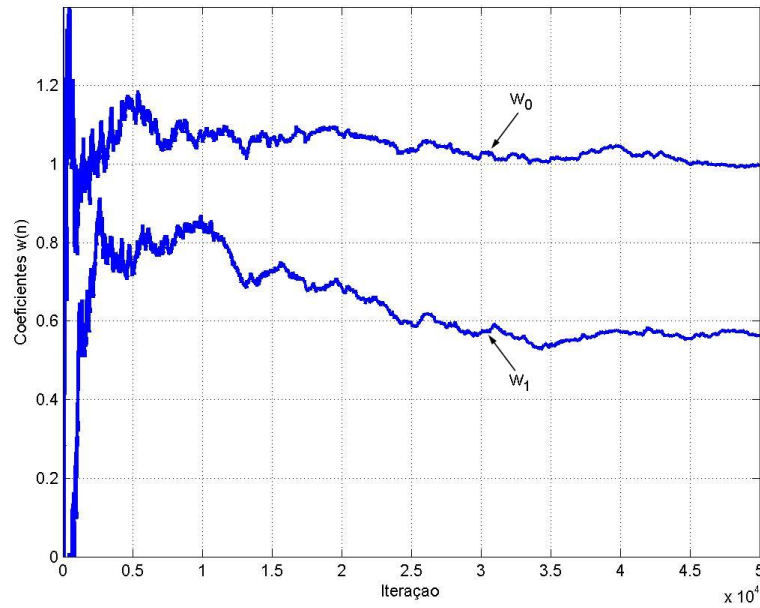


Fig. 5.11: Evolução temporal do algoritmo RLS (SWR = 30 dB, $\lambda = 1$).

pontos de custo mínimo (10.000 iterações para SWR = 20 dB e 20.000 iterações para SWR = 23 dB). Uma relação de 23 dB corresponde a uma marca d'água com potência 500 vezes inferior à do sinal de informação, o que é suficiente para garantir a transparência em sistemas de marca d'água de áudio que exploram características perceptivas (Gomes et al., 2000). Podemos observar na figura 5.11 que, mesmo para SWR = 30 dB, obtém-se uma estimativa razoável do sistema.

No esquema de identificação discutido, os autovalores da matriz de autocorrelação $\mathbf{R}_m$ dependem apenas do sinal de marca d'água. Sendo assim, o desvio quadrático médio do RLS (apresentado na equação 4.28) cresce à medida que diminui a potência da marca d'água, pois diminuem os valores de λ_i ; isto justifica o fato de as oscilações dos coeficientes do modelo aumentarem quando SWR aumenta, evidenciado nas figuras do RLS para baixo número de iterações n . Por outro lado, o desvio quadrático médio $\mathcal{D}$ decai à medida que o número de iterações n aumenta; assim, na média, os coeficientes estimados pelo algoritmo convergem para o vetor de parâmetros $\mathbf{w}_o$ com o tempo.

Conforme podemos concluir da discussão anterior, o algoritmo RLS mostrou-se mais adequado do que o LMS ao método de identificação proposto, diferentemente do caso de equalização no qual os algoritmos LMS e RLS apresentaram desempenho satisfatório; tendo isso em vista, prosseguiremos nossas análises utilizando apenas o algoritmo RLS. Adotemos agora um canal de teste de fase não-mínima, modelado como um filtro linear com a seguinte função de transferência:

$$H_2(z) = 1 + 1,2z^{-1} - 0,3z^{-2} + 0,8z^{-3} \quad (5.17)$$

O ensaio foi dividido em duas etapas. Na primeira, foi utilizado um sinal de informação formado por amostras i.i.d. extraídas de uma constelação 2-PSK de módulo unitário; na segunda, utilizou-se um sinal fortemente correlacionado, formado por amostras de áudio (música *rock*). Em ambos, desconsiderou-se o ruído e empregou-se um modelo de identificação com o mesmo número de parâmetros que o canal, inicializando-os na origem. Para avaliar de maneira mais precisa o desempenho do método neste cenário, definimos a seguinte medida de desvio em relação ao caso ideal (Arenas-García et al., 2003; Attux, 2005):

$$D = \frac{\|\mathbf{h}_{obtido} - \mathbf{h}\|}{\|\mathbf{h}\|} \quad (5.18)$$

sendo $\mathbf{h}_{obtido}$ a estimativa do vetor de parâmetros obtida pelo algoritmo e $\mathbf{h}$ seu valor exato (coeficientes que se deseja modelar); quanto menor o valor de D , melhor é a estimativa do algoritmo (0 para o caso ideal).

Na primeira etapa do ensaio, utilizando amostras i.i.d., obtivemos os valores médios de D , após a convergência, apresentados na tabela 5.1. Para os dois valores de SWR analisados, o coeficiente D revela um grau de precisão do algoritmo satisfatório para uma aplicação prática. A convergência para os dois valores de SWR ocorreu com cerca de 20.000 a 25.000 iterações, próxima à obtida no exemplo anterior (figuras 5.9 e 5.10), utilizando a função de transferência expressa em 5.16.

SWR	D
20 dB	0,0473
23 dB	0,0562

Tab. 5.1: Desvio do RLS para um sinal i.i.d.

Na segunda etapa do ensaio, utilizando amostras discretas de áudio, obtivemos os valores médios de D apresentados na tabela 5.2. Também neste caso, o coeficiente D encontra-se em um patamar que pode ser considerado satisfatório para uma aplicação prática. A velocidade de convergência para os dois valores de SWR também foi muita próxima da obtida no exemplo anterior (cerca de 20.000 a 25.000 iterações, para a função de transferência expressa em 5.16).

SWR	D
20 dB	0,0634
23 dB	0,0982

Tab. 5.2: Desvio do RLS para um sinal de áudio.

Em 100% dos ensaios realizados houve convergência do algoritmo. Além dos dois exemplos apresentados, diversos outros ensaios foram realizados com sinais digitais de telecomunicações e

de áudio, sempre com resultados aceitáveis para aplicações práticas. Em relação ao LMS, o RLS apresenta uma maior velocidade de convergência e um menor desajuste em relação aos coeficientes ótimos. Especialmente no método de identificação proposto, as características de velocidade e desajuste são ainda mais evidenciadas, mostrando-se o RLS mais adequado do que o LMS.

Ao compararmos a convergência de um sistema clássico utilizando seqüências de treinamento com o método utilizando marca d'água, foi percebido que a convergência do primeiro é algumas dezenas de vezes mais rápida. Porém, em sistemas práticos, as seqüências de treinamento são espaçadas no tempo, acarretando em interrupções no ajuste do modelo. Com o uso da marca d'água, o ajuste do modelo de identificação pode ser realizado continuamente, sem interrupções, compensando a adaptação mais lenta quando comparado a sistemas que utilizam seqüências de treinamento.

5.3.1 Análise de Canais Variantes no Tempo

Para analisar o desempenho do algoritmo de identificação proposto quando aplicado a canais variantes no tempo, utilizamos os mesmos canais de teste empregados no caso correspondente de equalização, cuja dinâmica dos coeficientes encontra-se ilustrada nas figuras 4.14 e 4.15. Nas simulações, assumimos que as amostras transmitidas são i.i.d. e pertencentes a uma constelação 2-PAM com módulo unitário obtidas no software de simulação MATLAB, sem a presença de ruído aditivo e com o sistema inicialmente sincronizado. Adotamos um modelo de identificação FIR $[w_{id1}(0) \ w_{id2}(0) \ w_{id3}(0)]$ com três coeficientes. Em todos os experimentos, os coeficientes foram adaptados pelo algoritmo de identificação utilizando marca d'água e comparados com os coeficientes reais do canal (ponto ideal de convergência).

As figuras 5.12, 5.13 e 5.14 ilustram a estimativa obtida para os coeficientes através do algoritmo RLS para o canal com transições em degrau, com $\lambda = 0,98$ e uma média de 200 realizações do algoritmo, justificado pelo fato de reduzir as oscilações nas figuras, melhorando a visualização. Podemos ver nas ilustrações que o algoritmo rastreia rapidamente o ponto ótimo de convergência conforme as variações do canal.

As figuras 5.15, 5.16 e 5.17 ilustram a estimativa obtida para os coeficientes através do algoritmo RLS para o canal degrau-rampa, com $\lambda = 0,98$ e uma média de 200 realizações do algoritmo. Nota-se que, como no caso anterior, o algoritmo é capaz de rastrear rapidamente o ponto ótimo de convergência conforme as variações do canal. Em todas as simulações realizadas houve 100% de convergência global e o método de identificação proposto mostrou-se eficaz com canais variantes.

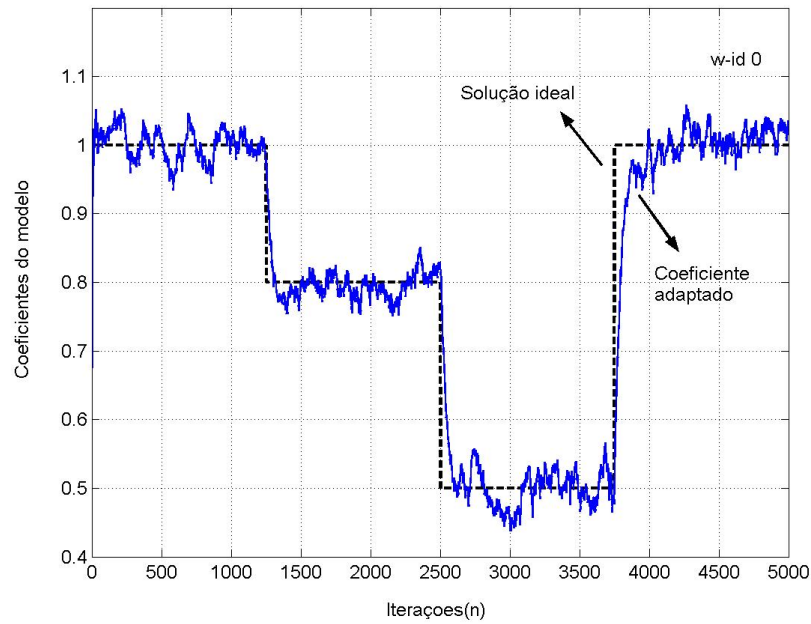


Fig. 5.12: Coeficiente w_{id0} do modelo de identificação para um canal variante baseado em degrau.

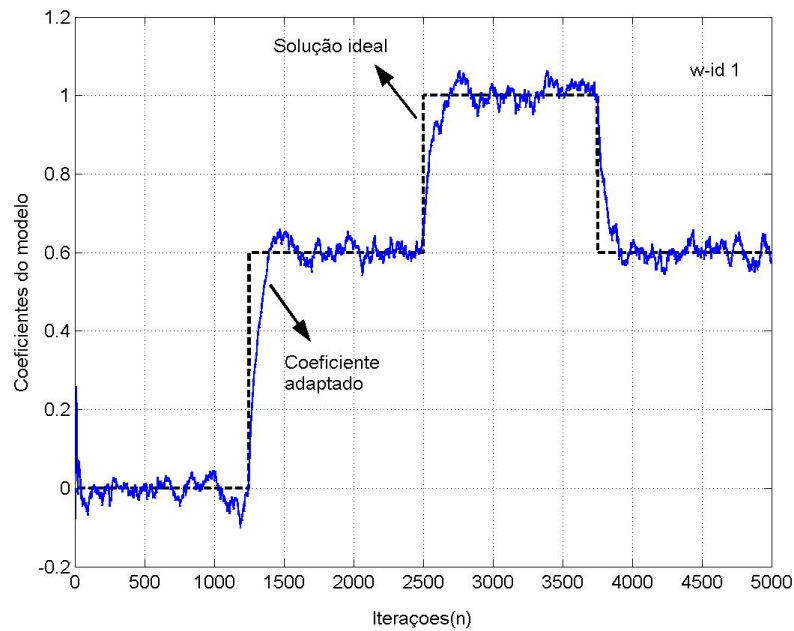


Fig. 5.13: Coeficiente w_{id1} do modelo de identificação para um canal variante baseado em degrau.

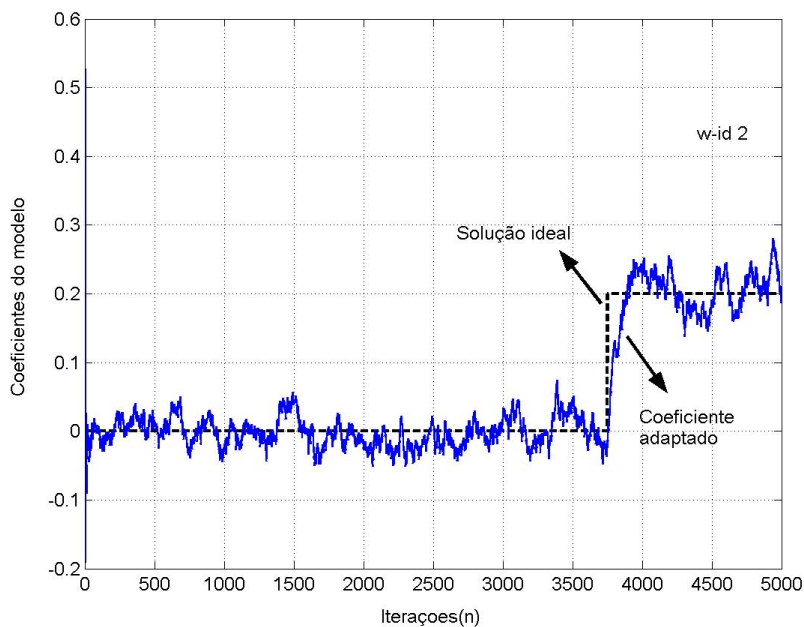


Fig. 5.14: Coeficiente w_{id2} do modelo de identificação para um canal variante baseado em degrau.

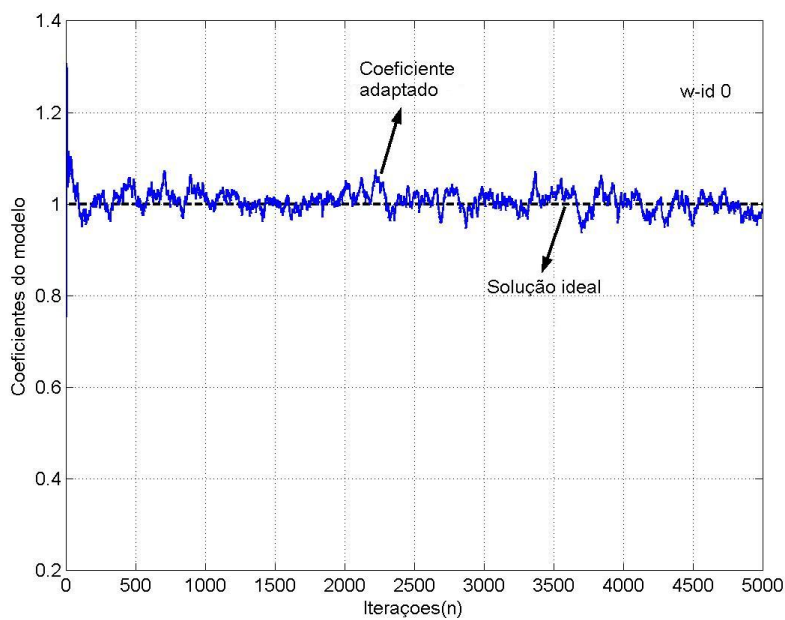


Fig. 5.15: Coeficiente w_{id0} do modelo de identificação para um canal variante baseado em degrau-rampa.

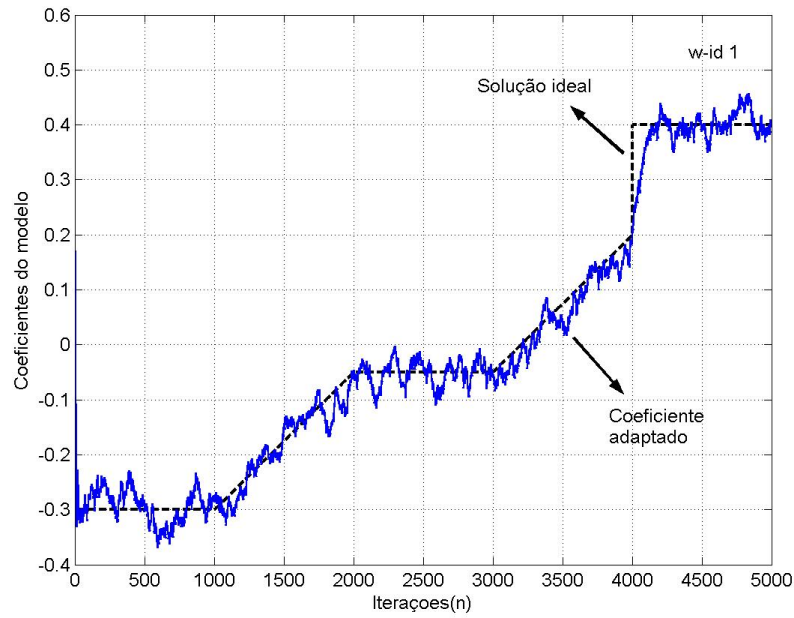


Fig. 5.16: Coeficiente w_{id1} do modelo de identificação para um canal variante baseado em degrau-rampa.

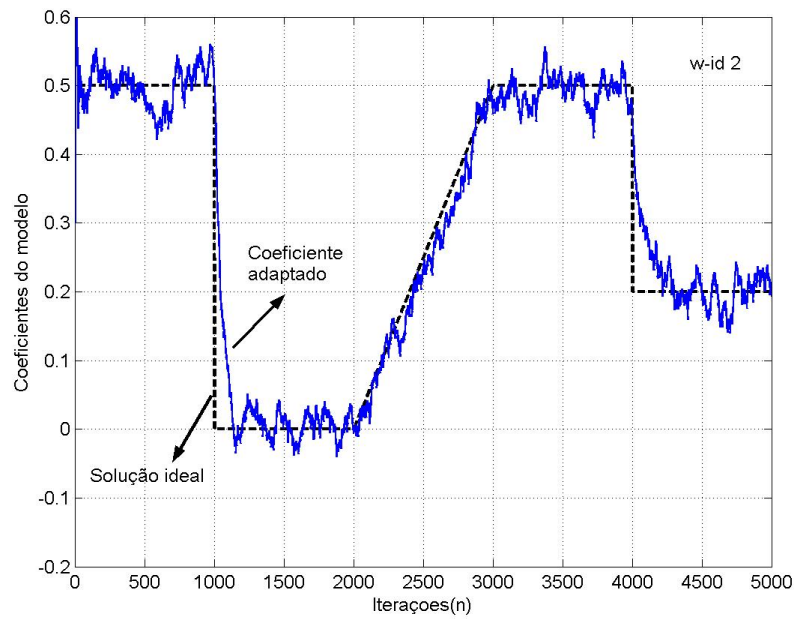


Fig. 5.17: Coeficiente w_{id2} do modelo de identificação para um canal variante baseado em degrau-rampa.

5.4 Síntese e Conclusões

Neste capítulo, foi proposto um método supervisionado de identificação de sistemas utilizando uma marca d'água como sinal de referência. O método adiciona uma marca d'água ao sinal de informação e, sem interrupções na transmissão deste último, utiliza continuamente o sinal da marca d'água no filtro identificador para obter um modelo matemático do canal de transmissão.

O método proposto apresenta vantagens em relação às técnicas de identificação clássicas, podendo-se destacar a ausência de períodos de treinamento, nos quais é interrompida a transmissão do sinal de informação, e a possibilidade de rastrear variações do canal, pois o filtro é continuamente adaptado. Por se tratar de um método supervisionado, foi constatado que o algoritmo proposto apresenta menor complexidade computacional em relação à maioria dos critérios cegos.

Em todas as simulações realizadas, o método funcionou conforme esperado, independentemente do canal ser invariante ou variante no tempo e das características do sinal de informação utilizado. Concluímos assim que o método pode trazer ganhos de desempenho a sistemas de comunicação, sobretudo quando houver escassez de recursos sistêmicos que inviabilize a transmissão periódica de seqüências de treinamento.

Foi demonstrado que o algoritmo RLS é mais adequado do que o LMS ao método de identificação proposto. A velocidade de convergência do algoritmo LMS torna pouco prático o seu uso. Além disso, pequenas variações no passo de adaptação provocam o rápido aumento das flutuações estocásticas.

No âmbito da identificação em sistemas de áudio, o método proposto será estendido mais adiante para contemplar o uso de técnicas psicoacústicas.

Capítulo 6

Sincronismo para Sistemas Adaptativos Utilizando Marca d'Água

6.1 Introdução

Em geral, o processo de detecção de uma marca d'água inicia-se com a sincronização da marca no detector. Dependendo da aplicação, um sistema utilizando marca d'água digital deve ser robusto a diferentes tipos de ataques, intencionais ou não. Ataques que causem dessincronização dificultam a correta recuperação da informação contida na marca d'água pelo detector, e muitas vezes são particularmente difíceis de neutralizar.

Neste capítulo, abordaremos ataques decorrentes de distorções impostas pela transmissão do sinal marcado através de um canal. Consideraremos também ataques caracterizados pela perda e/ou inserção de trechos do sinal transmitido, acarretando modificações no tamanho do sinal na escala temporal. Ambas as classes de ataques analisadas podem ser intencionais ou não.

Analisaremos dois métodos de sincronização para marca d'água, ambos baseados no uso de seqüências de treinamento embutidas na própria marca d'água e com detecção por correlação. O primeiro método utiliza uma média de janelas de análise e o segundo um algoritmo de programação dinâmica. Estes métodos revertem os efeitos de uma ampla classe de ataques de dessincronização, além de serem capazes de ressincronizar o sistema continuamente, ao longo de todo o sinal recebido.

6.2 Dessincronização da Marca d'Água

Conforme a aplicação, a marca d'água deve apresentar um determinado grau de robustez a distorções. Em aplicações envolvendo sistemas de comunicação, a marca d'água deve geralmente resistir a ataques que podem torná-la indetectável, tais como adição de ruído, filtragem e perda ou inserção de

amostras. Além de modificar o sinal de marca d'água transmitido, reduzindo a sua correlação com a marca original, tais ataques podem provocar dessincronização entre a marca transmitida e o sinal de referência no receptor.

Ataques de dessincronização, intencionais ou não, podem ser difíceis de neutralizar. Para se ter acesso à informação contida na marca d'água, o detector deve estar sincronizado com o transmissor; isso significa que o instante de início e de fim de cada um dos símbolos que compõem a marca deve ser conhecido. Na aplicação alvo deste trabalho, na qual a marca d'água é usada como sinal de referência em um filtro adaptativo, o sincronismo é particularmente importante, pois a correlação entre a marca original e o sinal de referência no receptor diminui rapidamente quando a janela de análise se afasta da posição correta.

Várias operações de processamento de sinal podem resultar em dessincronização. Por exemplo, a passagem do sinal marcado através de um canal de transmissão pode introduzir atrasos no sinal. Falhas no processo de comunicação também podem provocar deslocamentos temporais, eliminando ou inserindo amostras no sinal transmitido. Neste capítulo, abordaremos ataques desta natureza, intencionais ou não, que causam desalinhamento entre a marca d'água original e o sinal de referência conhecido no receptor.

6.3 Sincronismo para os Sistemas Adaptativos Propostos

Uma das técnicas mais comuns de sincronização em comunicação digital baseia-se no uso de seqüências de treinamento. O transmissor envia uma seqüência conhecida que atravessa o canal de comunicação; no receptor, a seqüência recebida é alinhada com um sinal de referência idêntico à seqüência original, levando à sincronização do sistema. A mesma idéia pode ser utilizada para sincronizar a marca d'água: nos esquemas de equalização e identificação, a marca é composta por uma sucessão de amostras conhecidas, obtidas a partir de uma seqüência pseudo-aleatória escalonada que pode ser reproduzida no receptor. Esta seqüência de amostras será denominada *bloco da marca d'água*. Para cada janela de análise, o processo de sincronização procura localizar um tal bloco no sinal recebido, o que é feito por meio do cálculo da correlação deste sinal com as amostras que compõem o bloco original. Este método permite que a sincronização seja realizada continuamente ao longo da recepção (Gómez, 2000; Moreau et al., 2000).

Quando um atraso é imposto por um processo de filtragem, ou amostras são extraídas e/ou inseridas no sinal marcado, há um deslocamento no pico da correlação entre janelas do sinal marcado e o bloco da marca d'água. Através dessa medida de correlação, o detector pode estimar o atraso do sinal marcado. Este processo é utilizado tanto no esquema de equalização quanto no de identificação utilizando marca d'água (figuras 6.1 e 6.2, respectivamente).

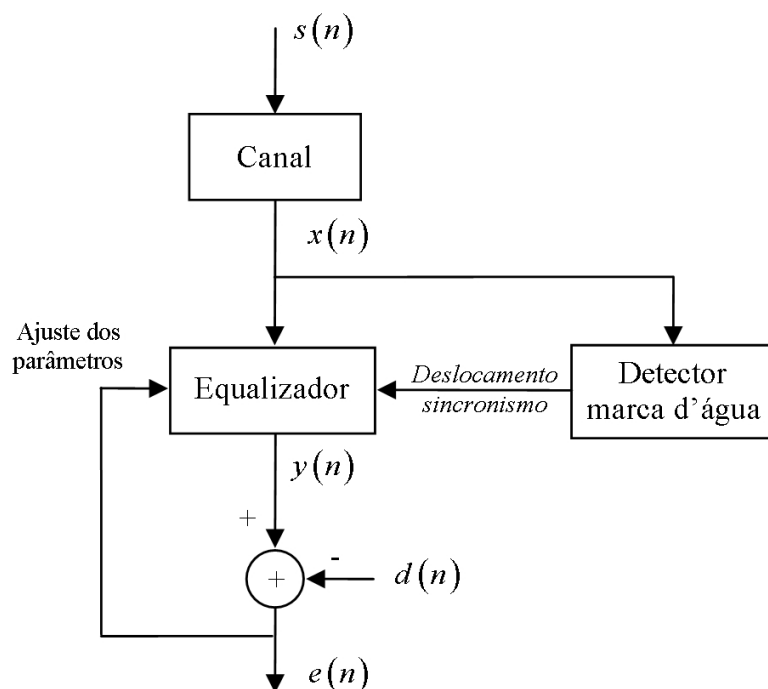


Fig. 6.1: Esquema de equalização utilizando marca d'água com sincronização.

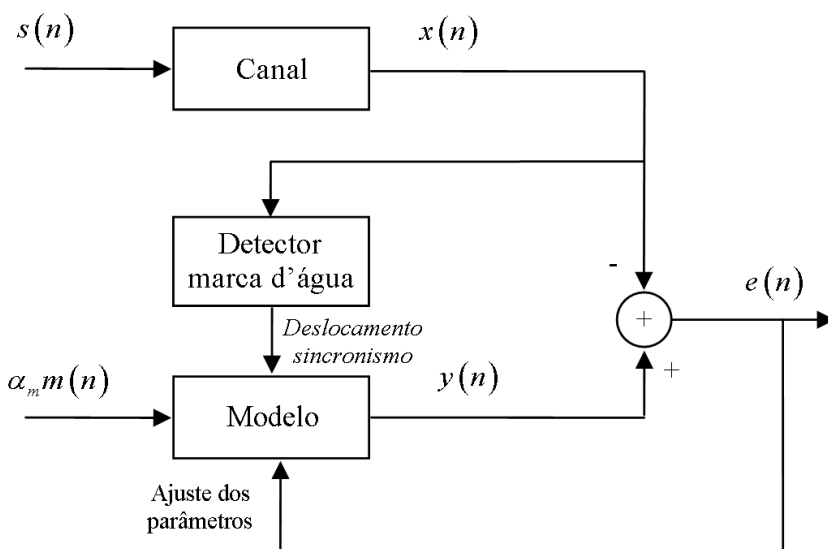


Fig. 6.2: Esquema de identificação utilizando marca d'água com sincronismo.

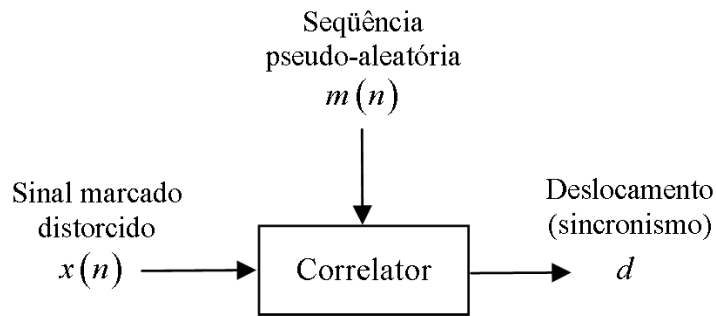


Fig. 6.3: Esquema do detector da marca d'água.

Para cada janela de análise, o algoritmo adaptativo recebe do esquema de sincronismo a informação do atraso ótimo do sistema, expresso por um deslocamento d . A seqüência de treinamento, que corresponde a um bloco da marca d'água, deve ter comprimento suficiente para que a sua autocorrelação se aproxime de uma função impulso¹. Como a marca d'água está presente ao longo de todo o sinal de informação, o sistema é capaz de rastrear o sincronismo de forma contínua, identificando rapidamente deslocamentos da marca caso o sistema perca o sincronismo.

O detector da marca d'água, ilustrado na figura 6.3, opera com base em um processo de correlação linear. O correlator calcula a correlação entre uma janela do sinal marcado distorcido pelo canal $x(n)$ e um bloco da marca d'água original $m(n)$, conhecido no receptor. O método de sincronização mais imediato consiste em adotar como deslocamento ótimo d o deslocamento corresponde ao pico de correlação para cada janela de análise; no entanto, como veremos mais adiante, esse método está sujeito a fortes interferências de ruído, sobretudo no caso em que a marca d'água é muito menos potente do que o sinal de informação. Métodos mais elaborados e que reduzem a interferência do ruído na sincronização serão apresentados nas próximas seções.

6.4 Métodos de Sincronização

6.4.1 Seqüência de Treinamento: Idéia Básica

Para encontrar o deslocamento correspondente ao sincronismo da marca d'água no sinal recebido, considera-se a marca como uma seqüência de treinamento conhecida e presente ao longo de todo o sinal. Analisamos duas abordagens para esse problema. A primeira abordagem baseia-se no uso de uma marca d'água construída através da replicação de um único bloco de sinal obtido a partir de uma seqüência pseudo-aleatória. Como este bloco é conhecido no receptor, é possível calcular a correla-

¹O bloco deve ter características próximas às de um sinal branco, ou seja, sua energia deve ser distribuída de maneira tão uniforme quanto possível em frequência.

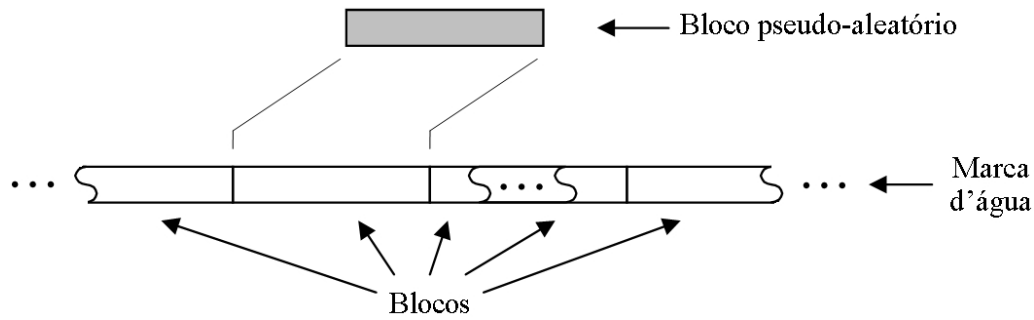


Fig. 6.4: Estrutura de dados para sincronização. O bloco cinza corresponde ao bloco formado por símbolos pseudo-aleatórios; a sucessão de blocos corresponde à marca d'água.

ção entre o sinal recebido e o bloco original por meio de uma janela deslizante de análise; com isso, constrói-se uma matriz contendo os valores de correlação cujas linhas e colunas correspondem ao deslocamento da janela deslizante e ao número da janela de análise, respectivamente. Observando-se os valores máximos de correlação para as janelas, definem-se os deslocamentos ótimos referentes ao sincronismo da detecção da marca. A segunda abordagem utiliza uma marca d'água construída através da concatenação de vetores de um *codebook* em uma ordem determinada (conhecida no receptor). Além da matriz contendo valores de correlação, constrói-se outra matriz contendo o símbolo correspondente ao vetor do codebook que apresenta a maior correlação com a janela deslizante. Para determinação do sincronismo, são levados em conta os valores de correlação e a observância da sequência apropriada de símbolos do codebook.

6.4.2 Seqüência de Blocos

Nesta abordagem, tanto a sincronização inicial como a resincronização contínua da detecção da marca d'água são obtidas por meio de uma sequência de treinamento conhecida pelo receptor, correspondente às amostras que compõem um bloco da marca d'água. Estas amostras são provenientes de uma sequência pseudo-aleatória escalonada. O bloco da marca d'água contém N amostras, apresentando características próximas às de um sinal branco. A marca d'água é construída através da sucessão de múltiplas cópias deste bloco. Esta estrutura é ilustrada na figura 6.4.

O bloco da marca d'água é dado por:

$$\mathbf{m}_\alpha = [m_\alpha(0), m_\alpha(1), \dots, m_\alpha(N-1)] \quad (6.1)$$

onde N expressa a quantidade de amostras no bloco e $\alpha \in \{0, 1\}$ é a estimativa do fator de atenuação para a relação SWR utilizada no sistema. A janela de análise tem o mesmo comprimento do bloco.

O instante em que começa um bloco é inicialmente desconhecido. Na detecção, uma janela desli-

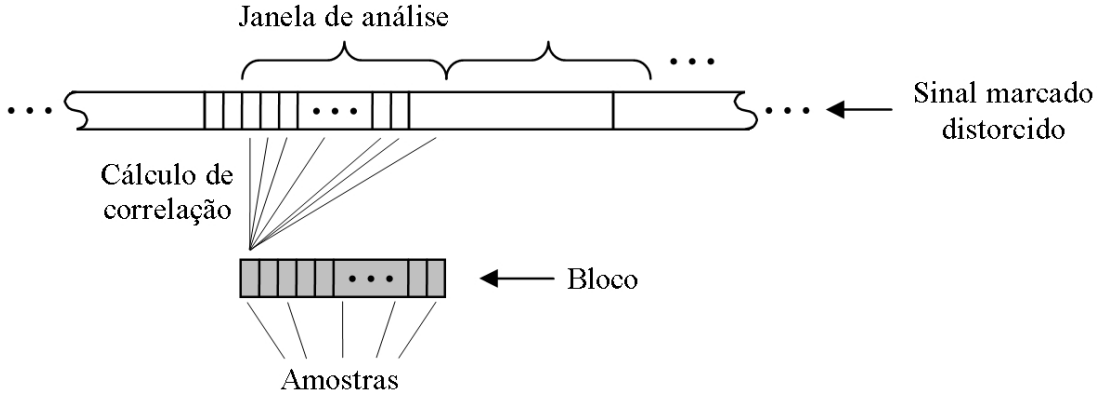


Fig. 6.5: Cálculo de correlação para sincronismo.

zante é utilizada para calcular 2Λ medidas de correlação entre o sinal recebido $x(n)$ e as amostras do bloco da marca d'água $m_\alpha(n)$ para cada janela de análise:

$$r(\lambda, j) = \left| \sum_{n=0}^{N-1} x(jN + n + \lambda) m_\alpha(n) \right| \quad (6.2)$$

onde $\lambda \in [-\Lambda, \Lambda - 1]$ é o deslocamento da janela deslizante e j representa o número da janela de análise. Este processo é ilustrado na figura 6.5. Com isso, podemos construir uma matriz $\mathbf{A} = \{a_{\lambda,j}\}$ que contém as correlações calculadas e cujas linhas e colunas correspondem ao deslocamento λ e à posição j de cada janela de análise, respectivamente. Para que todos os possíveis deslocamentos da marca d'água sejam analisados, deve-se ter $\Lambda = N/2$ (supondo N par). No entanto, caso se tenha segurança de que o sistema se encontra próximo do sincronismo e se deseje apenas refinar esta situação, pode-se utilizar um valor menor para Λ , economizando recursos computacionais.

Pelo fato de a marca d'água apresentar, em geral, potência muito inferior à do sinal de informação, picos espúrios de correlação (não correspondentes ao deslocamento real da marca d'água) podem surgir na matriz $\mathbf{A}$. Por isso, a simples escolha, para cada janela de análise, do deslocamento da janela deslizante que corresponde à máxima correlação pode levar a resultados bastante ruins, como ilustrado na figura 6.6. A interferência do sinal de informação pode ser reduzida por meio de métodos de sincronização que levam em conta várias janelas de sinal simultaneamente, como discutido a seguir.

Apresentaremos nas seções 6.5.1 e 6.5.2 dois algoritmos capazes de determinar um caminho através da matriz $\mathbf{A}$ correspondente aos deslocamentos ótimos da marca d'água, partindo da primeira coluna da matriz ($j = 0$) até a última ($j = J - 1$). O número J de janelas do sinal recebido que devem ser acumuladas para análise simultânea pelos algoritmos de sincronização dependerá de cada aplicação; quanto maior esse número, mais precisa tende a ser a determinação do sincronismo da marca, porém maior será o atraso para que este resultado seja obtido. O primeiro algoritmo apre-

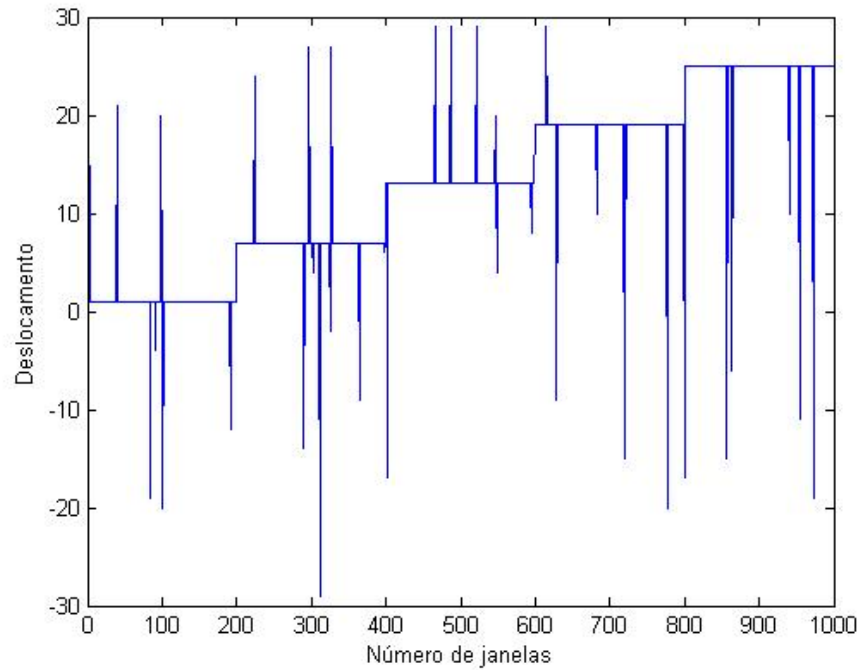


Fig. 6.6: Deslocamentos de janela deslizante correspondente à máxima correlação para o bloco da marca d'água. Para ilustrar a perda de sincronismo, 5 amostras foram removidas a cada 200 janelas.

sentado baseia-se no cálculo da média entre janelas de análise, e o segundo emprega um método de programação dinâmica. Os algoritmos definirão os deslocamentos ótimos correspondentes ao sincronismo da marca d'água, resultando em um conjunto de J valores para λ , $[\hat{\lambda}_0 \dots \hat{\lambda}_{J-1}]$, referentes a cada janela de análise.

6.4.3 Seqüência de Símbolos de Codebook

Uma abordagem para aumentar a robustez na sincronização da marca d'água foi proposta por Gomes et al. (Gomes et al., 2001). Ao invés de um único bloco, como apresentado na seção anterior, um codebook $\mathbf{C}$ é usado para gerar a marca d'água digital $w(n)$. Este codebook contém k vetores $\mathbf{u}_k = [u_k(0) \dots u_k(N-1)]$ ($k \in [0, K-1]$) associados a K símbolos. A seqüência de símbolos $\mathbf{z} = [z_0 \dots z_{M-1}]$ que compõem a marca d'água é obtida de acordo com a regra

$$z_j = j \bmod K \quad (6.3)$$

onde z_j é o j -ésimo símbolo na seqüência da marca d'água ($j \in [0, J-1]$), seqüência esta completamente conhecida no detector. A marca d'água $w(n)$ é construída através da concatenação dos vetores associados aos símbolos na seqüência $\mathbf{z}$.

Na detecção, uma janela deslizante é usada para calcular 2Λ medidas de correlação entre o sinal recebido $x(n)$ e cada um dos K vetores no codebook $\mathbf{C}$, para cada janela de análise (correspondente a cada um dos J símbolos presentes na marca d'água):

$$r(\lambda, k, j) = \left| \sum_{n=0}^{N-1} x(jN + n + \lambda) u_k(n) \right| \quad (6.4)$$

onde $\lambda \in [-\Lambda, \Lambda - 1]$ é o deslocamento da janela deslizante e j é o número da janela de análise. Selecionando os símbolos do codebook que apresentam máxima correlação com o sinal recebido para cada janela de análise, podem-se construir as matrizes $\mathbf{A} = \{a_{\lambda,j}\}$ e $\mathbf{B} = \{b_{\lambda,j}\}$, cujas linhas e colunas correspondem, respectivamente, ao deslocamento λ e à posição j de cada janela de análise:

$$a_{\lambda,j} = \max_k r(\lambda, k, j) \quad (6.5)$$

$$b_{\lambda,j} = \arg \max_k r(\lambda, k, j) \quad (6.6)$$

Dessa forma, $\mathbf{A}$ conterà as maiores medidas de correlação para cada deslocamento da janela deslizante e cada posição na seqüência de símbolos que compõem a marca d'água, e $\mathbf{B}$ conterà o símbolo do codebook $\mathbf{C}$ correspondente à máxima correlação.

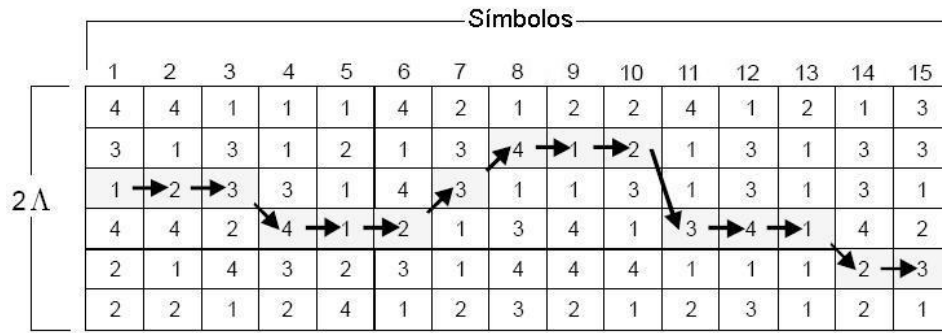
Como na abordagem apresentada na seção anterior, deve-se ter $\Lambda = N/2$ (supondo N par) para que todos os possíveis deslocamentos da marca d'água sejam analisados, podendo-se utilizar valores menores para Λ (com economia de recursos computacionais) caso o sistema já se encontre próximo do sincronismo e se deseje apenas refinar esta situação.

Apresentaremos na seção 6.5.2 um algoritmo capaz de determinar um caminho ótimo nas matrizes $\mathbf{A}$ e $\mathbf{B}$ partindo da primeira coluna ($j = 0$) até a última ($j = J - 1$), levando em conta a ordem esperada de símbolos na marca d'água e as medidas máximas de correlação. Isto resulta em um conjunto de J valores para λ , $[\hat{\lambda}_0 \dots \hat{\lambda}_{J-1}]$, um para cada janela de análise. Como ilustração, um exemplo de possível caminho ótimo na matriz $\mathbf{B}$ é apresentado na figura 6.7.

6.5 Algoritmos de Sincronização

6.5.1 Sincronização por Média de Janelas de Análise

A maneira mais imediata para determinar um caminho através da matriz $\mathbf{A}$, correspondente aos deslocamentos estimados para a marca d'água em cada janela de análise, é calcular uma média linear entre t diferentes colunas do conjunto de janelas J . Para cada janela, é escolhido o deslocamento associado à maior correlação média. Mesmo que haja picos espúrios de correlação, estes tenderão a

Fig. 6.7: Possível caminho ótimo na matriz **B**.

ser neutralizados com o cálculo da média, por apresentarem-se de forma aleatória e não polarizada. O algoritmo de sincronização resultante é descrito a seguir:

Para $j = 0 \dots J - t - 1$

Para $\lambda = -\Lambda \dots \Lambda - 1$

$$I(\lambda) = 0$$

Para $i = 0 \dots t$

$$I(\lambda) = I(\lambda) + a_{\lambda, j+i}$$

$$\hat{\lambda}_j = \arg \max_{\lambda'} I(\lambda')$$

O resultado do algoritmo é um conjunto de deslocamentos $[\hat{\lambda}_0 \dots \hat{\lambda}_{J-t-1}]$ correspondentes ao caminho ótimo segundo o critério adotado.

Este método tende a funcionar bem para valores baixos de SWR, para os quais há relativamente poucos picos espúrios de correlação. Quando o sistema trabalha com valores maiores de SWR, a quantidade de picos espúrios cresce rapidamente; para que a média entre janelas seja capaz de eliminá-los, podem ser necessários valores de t elevados, acarretando em atrasos significativos para o sistema e reduzindo o seu tempo de resposta. Tais atrasos podem ser impraticáveis, sobretudo em sistemas de tempo real que requeiram uma constante verificação de sincronismo. Em particular, este método pode ser interessante para a sincronização inicial do sistema, na qual um certo atraso geralmente pode ser tolerado.

6.5.2 Sincronização por Programação Dinâmica

Função Custo para Sequência de Blocos

Para determinar os deslocamentos da janela deslizando que levam ao melhor alinhamento com o sinal recebido, pode-se utilizar um algoritmo de programação dinâmica (Forney, 1973; Bertsekas,

2000).

Define-se um *estado* do algoritmo de otimização como um par ordenado $[\lambda, j]$ no qual o primeiro elemento corresponde a um deslocamento da janela deslizante e o segundo ao número da janela de análise. Existe uma correspondência direta entre um estado do algoritmo e uma célula da matriz $\mathbf{A} = \{a_{\lambda,j}\}$, definida na seção 6.4.2. O objetivo do algoritmo é minimizar uma função custo calculada em termos da matriz $\mathbf{A}$. Como resultado, obtém-se uma seqüência de estados que corresponde a um caminho ótimo através da matriz $\mathbf{A}$, indicando o deslocamento da marca d'água $([\hat{\lambda}_0 \dots \hat{\lambda}_{J-1}])$ para cada janela de análise.

A função custo $c(\lambda, \lambda', j)$ para passar do estado $[\lambda', j - 1]$ ao estado $[\lambda, j]$ é composta por dois termos:

$$c(\lambda, \lambda', j) = c_1(\lambda, \lambda', j) + c_2(\lambda, j) \quad (6.7)$$

O primeiro termo penaliza mudanças no valor do deslocamento λ :

$$c_1(\lambda, \lambda', j) = \eta_j (\lambda - \lambda')^2 \quad (6.8)$$

O objetivo deste termo é evitar deslocamentos espúrios provocados por picos aleatórios de correlação; em particular, o deslocamento deve permanecer inalterado quando o sistema se encontra sincronizado. O termo quadrático impõe uma penalidade que aumenta rapidamente com o tamanho do salto de λ para λ' , significando que o algoritmo privilegia mudanças suaves de deslocamento em detrimento de mudanças abruptas (menos comuns na prática). O fator η_j é definido como:

$$\eta_j = \begin{cases} \eta_{j-1} + \kappa_1 & \text{se } \lambda \neq \lambda' \\ \max(\eta_{j-1} - \kappa_2, \eta_0) & \text{outro} \end{cases} \quad (6.9)$$

sendo κ_1 , κ_2 e η_0 constantes positivas. Esta definição penaliza caminhos ziguezague, pois η_j tende a crescer em uma tal situação.

O segundo termo na definição da função custo está relacionado às medidas de correlação presentes na matriz $\mathbf{A}$:

$$c_2(\lambda, j) = \rho \left(1 - \frac{a_{\lambda,j}}{\max_{\tilde{\lambda}} a_{\tilde{\lambda},j}} \right) \quad (6.10)$$

onde ρ é uma constante positiva. A expressão entre parênteses assume valores entre 0 (quando o deslocamento λ corresponde ao maior valor de correlação) e 1 (quando a correlação para um deslocamento λ é nula). Esta definição penaliza deslocamentos λ que apresentem baixos valores de correlação.

Função Custo para Seqüência de Símbolos de Codebook

Neste caso, além da matriz $\mathbf{A} = \{a_{\lambda,j}\}$, o processo de otimização leva em conta também a matriz $\mathbf{B} = \{b_{\lambda,j}\}$, ambas definidas na seção 6.4.3. Como na seção anterior, o objetivo é obter uma seqüência de estados do algoritmo que indique o deslocamento da marca d'água $([\hat{\lambda}_0 \dots \hat{\lambda}_{J-1}])$ para cada janela de análise, correspondendo a um caminho ótimo através das matrizes $\mathbf{A}$ e $\mathbf{B}$.

A função custo $c(\lambda, \lambda', j)$ para passar do estado $[\lambda', j-1]$ ao estado $[\lambda, j]$ é composta, neste caso, por três termos:

$$c(\lambda, \lambda', j) = c_1(\lambda, \lambda', j) + c_2(\lambda, j) + c_3(\lambda, \lambda', j) \quad (6.11)$$

Os dois primeiros termos são idênticos aos definidos na função custo da equação 6.7. O terceiro termo, responsável por penalizar a não observância da seqüência de símbolos que compõem a marca d'água, é definido como

$$c_3(\lambda, \lambda', j) = \begin{cases} \psi(b_{\lambda,j} - b_{\lambda',j-1} - 1) & \text{se } b_{\lambda,j} \geq b_{\lambda',j-1} \\ \psi(b_{\lambda,j} - b_{\lambda',j-1} - 1 + K) & \text{outros valores} \end{cases} \quad (6.12)$$

onde ψ é uma constante e K o número de vetores do codebook. Se a seqüência de símbolos da marca d'água é respeitada, este custo é nulo; caso contrário, o custo é proporcional à quantidade de símbolos faltantes em relação à seqüência esperada. Esta definição se justifica pelo fato de que, em geral, a perda de uma pequena quantidade de símbolos é mais comum do que a perda de longos trechos de sinal, devendo-se portanto privilegiar as seqüências de estados correspondentes ao primeiro caso no algoritmo de otimização.

Algoritmo de Otimização

Definamos o custo acumulado $C(\lambda, j)$ como o custo mínimo para atingir o estado $[\lambda, j]$ partindo de um estado na primeira coluna ($j = 0$). Este custo é inicializado em 0 para $j = 0$ e todo λ . O algoritmo de otimização empregado, baseado no algoritmo de Viterbi (Forney, 1973; Bertsekas, 2000), é descrito a seguir:

Para $j = 1 \cdots J - 1$
 Para $\lambda = -\Lambda \cdots \Lambda - 1$
 $\bar{\lambda} = \arg \min_{\lambda'} [C(\lambda', j - 1) + c(\lambda, \lambda', j)]$
 $C(\lambda, j) = C(\bar{\lambda}, j - 1) + c(\lambda, \bar{\lambda}, j)$
 $I(\lambda, j) = \bar{\lambda}$
 $\hat{\lambda}_{J-1} = \arg \min_{\tilde{\lambda}} [C(\tilde{\lambda}, J - 1)]$
 Para $j = J - 2 \cdots 0$
 $\hat{\lambda}_j = I(\lambda_{j+1}, j + 1)$

O resultado do algoritmo é um conjunto de deslocamentos $[\hat{\lambda}_0 \dots \hat{\lambda}_{J-1}]$ correspondentes ao sincronismo da marca d'água para cada janela de análise.

6.6 Simulações e Resultados

6.6.1 Condições Experimentais

Para analisar o desempenho dos algoritmos de sincronização propostos, dois sinais foram utilizados: o primeiro é um sinal digital, descorrelacionado, constituído de amostras i.i.d., pertencente a uma constelação 2-PSK com módulo unitário e obtidas através da função *randsrc* no software de simulação MATLAB; o segundo é um sinal de fala cantada, fortemente correlacionado, “svega”, uma versão capela de “Tom’s diner”, de Suzanne Vega, muito usada como referência em trabalhos de marca d’água de áudio (Gomes et al., 2003, 2000, 2001). Consideramos que não há ruído aditivo no sistema.

Para ilustrar a perda de sincronismo, 5 amostras foram retiradas do sinal a cada 200 janelas, forçando o sistema a resincronizar a detecção da marca d’água. Dois cenários de teste foram utilizados. O primeiro, de baixa complexidade, viabiliza uma análise didática para melhor entendimento dos algoritmos de sincronismo. O segundo cenário, mais complexo, possibilita a análise do desempenho dos algoritmos.

Os dois métodos de sincronização propostos foram testados. O algoritmo de média de janelas utilizou o valor constante $t = 10$. A otimização através de programação dinâmica utilizou os valores constantes $\psi = \Lambda$, $\eta_0 = 1$, $\kappa_1 = 5$, $\kappa_2 = 1$ e $\rho = 10$. Esses parâmetros foram ajustados empiricamente.

Primeiro Ensaio

No primeiro ensaio, assumimos que o sinal marcado não está sujeito a distorções, provocadas por processos de filtragem. Neste caso, a distorção se limita à supressão periódica das amostras do sinal. Com isso, buscamos um ambiente que propicie uma análise didática.

Foram utilizadas janelas com comprimento $N = 60$, com um deslocamento máximo $\Lambda = 30$ para a janela deslizante. Vale frisar que uma janela com esse comprimento pode prejudicar as características de decorrelação do sinal, com influência direta no cálculo da correlação, aumentando a quantidade de picos espúrios na matriz $\mathbf{A}$. Esta escolha viabiliza a análise da influência de janelas curtas no cálculo da correlação.

Segundo Ensaio

Assumimos agora que o sinal marcado está sujeito a um processo de filtragem, modelado como um canal linear invariante no tempo de fase não-mínima, com a seguinte função de transferência:

$$S_2 : H_2(z) = 1 + 1,2z^{-1} - 0,3z^{-2} + 0,8z^{-3} \quad (6.13)$$

Foram utilizadas janelas com comprimento $N = 512$, com um deslocamento máximo $\Lambda = 256$ para a janela deslizante. Uma janela com esse comprimento pode apresentar características próximas às de um sinal branco. Esta escolha viabiliza a análise de janelas longas no cálculo da correlação.

6.6.2 Resultados Experimentais

Apresentamos inicialmente os resultados referentes ao primeiro ensaio, que permite uma análise didática dos algoritmos propostos.

As figuras 6.8 e 6.9 apresentam os resultados para o sinal digital utilizando os algoritmos propostos com relações sinal-marca d'água de 10 e 20 dB, respectivamente. A primeira ilustração de cada figura refere-se aos deslocamentos da janela deslizante que produziram os maiores valores de correlação entre o sinal recebido e a marca d'água original, obtidos a partir da matriz $\mathbf{A}$. Notam-se diversos picos espúrios de correlação que dificultam a determinação dos atrasos correspondentes ao sincronismo. Para $SWR = 10$ dB, apesar da presença de vários picos de correlação espúrios, ainda se pode visualizar de forma relativamente clara os patamares de deslocamento correspondentes ao sincronismo da marca. Já para $SWR = 20$ dB, os picos espúrios tornam impossível determinar por análise visual os deslocamentos correspondentes ao sincronismo da marca. A grande quantidade de picos espúrios deve-se, além da baixa potência apresentada pela marca, ao tamanho relativamente curto apresentado pelas janelas de análise neste ensaio.

As ilustrações 6.8(b) e 6.9(b) correspondem ao resultado do algoritmo de média entre janelas de análise. Nota-se que, apesar de se tratar de um sinal descorrelacionado, o método não foi capaz de determinar com exatidão o sincronismo da marca d'água. Para $SWR = 10$ dB, o algoritmo identificou um deslocamento inexistente no sincronismo em torno da janela 390; já para $SWR = 20$ dB, o resultado mostra vários deslocamentos inexistentes, com um comportamento em zigue-zague que faz com que o sistema altere o deslocamento correspondente ao sincronismo continuamente. As figuras 6.8(c) e 6.9(c) apresentam o resultado do algoritmo de programação dinâmica utilizando sequência de blocos. Neste caso, os resultados mostram que o algoritmo foi capaz de rastrear o deslocamento correspondente ao sincronismo da marca para as relações sinal-marca d'água utilizadas. No entanto, para $SWR = 20$ dB, o algoritmo apresenta problemas nas transições entre deslocamentos, levando algum tempo para ajustar-se a novas condições de sincronismo.

As figuras 6.10 e 6.11 apresentam os resultados para o sinal de fala fortemente correlacionado utilizando os algoritmos propostos com relações sinal-marca d'água de 10 e 20 dB, respectivamente. Como no caso anterior, a primeira ilustração de cada figura refere-se aos deslocamentos da janela deslizante que produziram os maiores valores de correlação entre o sinal recebido e a marca d'água original. No caso deste sinal (assim como na maioria dos sinais de fala e áudio), a potência varia fortemente ao longo do tempo, havendo inclusive trechos de silêncio. Nesses trechos, estabelecemos um limiar mínimo de potência para a marca d'água, obtido empiricamente, que possibilita a sua detecção sem introduzir um ruído demasiadamente incômodo no sinal; como consequência, há uma diminuição da SWR , pois o áudio tem potência muito baixa². Por este motivo, as ilustrações contendo os picos de correlação apresentam alguns trechos nos quais o sincronismo da marca foi detectado com exatidão, inclusive para $SWR = 20$ dB. Com o algoritmo de média entre janelas de análise, o desempenho geral é melhor do que o obtido com a simples maximização da correlação, mas notam-se vários deslocamentos não correspondentes ao sincronismo da marca. Para $SWR = 20$ dB, alguns dos deslocamentos inexistentes são relativamente longos, o que se deve à forte correlação do sinal de fala. As figuras 6.10 e 6.11 também mostram o resultado do algoritmo de programação dinâmica utilizando sequência de blocos. Este algoritmo apresentou bom resultado para ambas SWR .

Passemos agora à análise do segundo ensaio, com os sinais sujeitos a um processo de filtragem linear invariante no tempo de fase não-mínima.

A figura 6.12 apresenta os resultados para o sinal digital de telecomunicações utilizando os algoritmos propostos. Como uma idéia fundamental no paradigma da marca d'água é a transparência, apresentamos a simulação para $SWR = 20$ dB (considerada uma relação suficiente para garantir transparência em sinais de áudio, conforme será apresentado no capítulo 7). A figura 6.12(a) apresenta os

²Para que o limiar mínimo de potência da marca fosse inaudível, sua densidade espectral de potência deveria estar abaixo do limiar absoluto de audibilidade em silêncio. Esta abordagem será detalhada no capítulo 7. Para maiores detalhes sobre fenômenos auditivos, ver apêndice A.

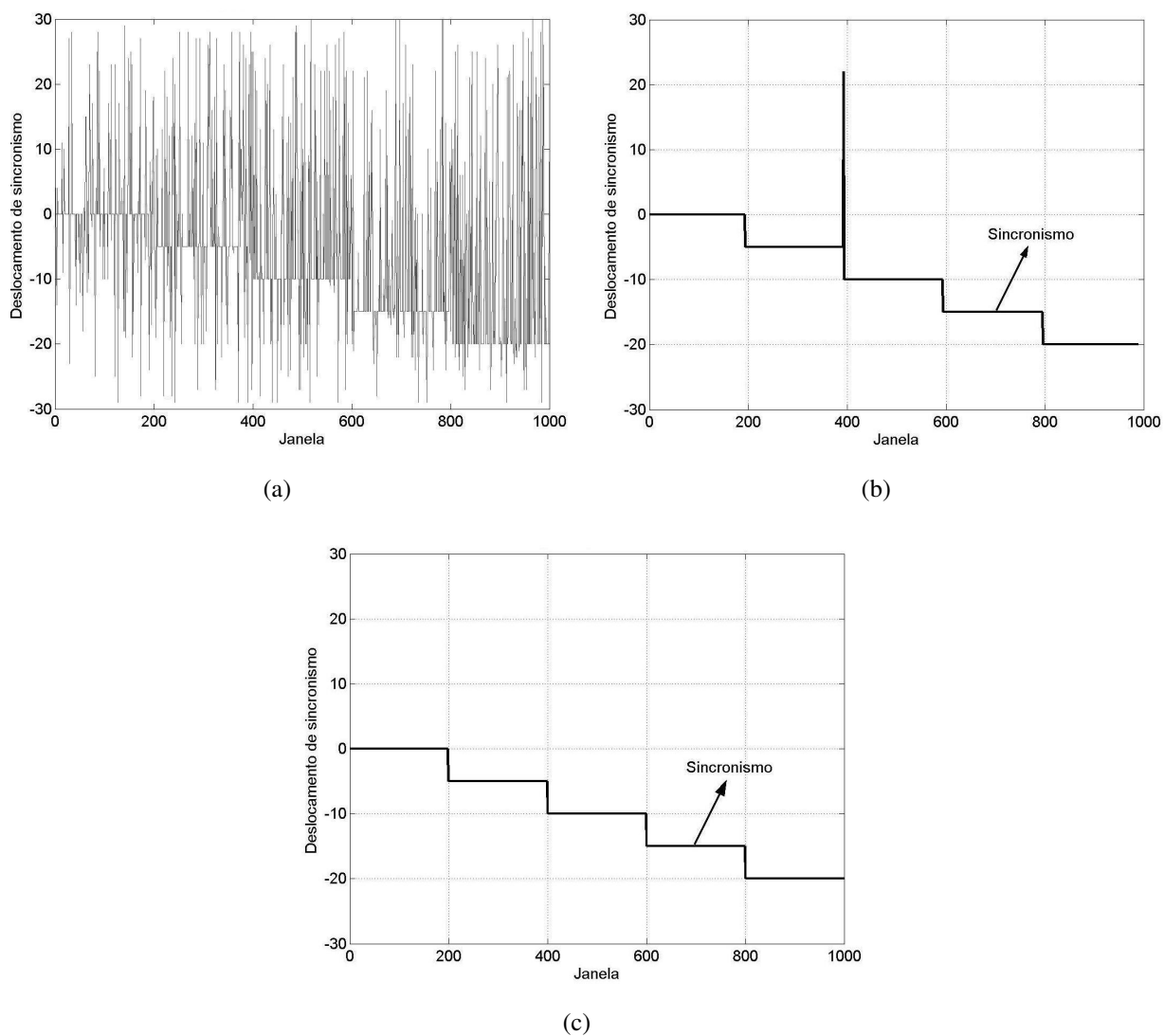


Fig. 6.8: Desempenho dos algoritmos com sinal digital de telecomunicações ($SWR = 10$ dB) e sem a presença de distorções: 6.8(a) picos espúrios de correlação; 6.8(b) sincronização por média de janelas de análise; 6.8(c) sincronização por programação dinâmica.

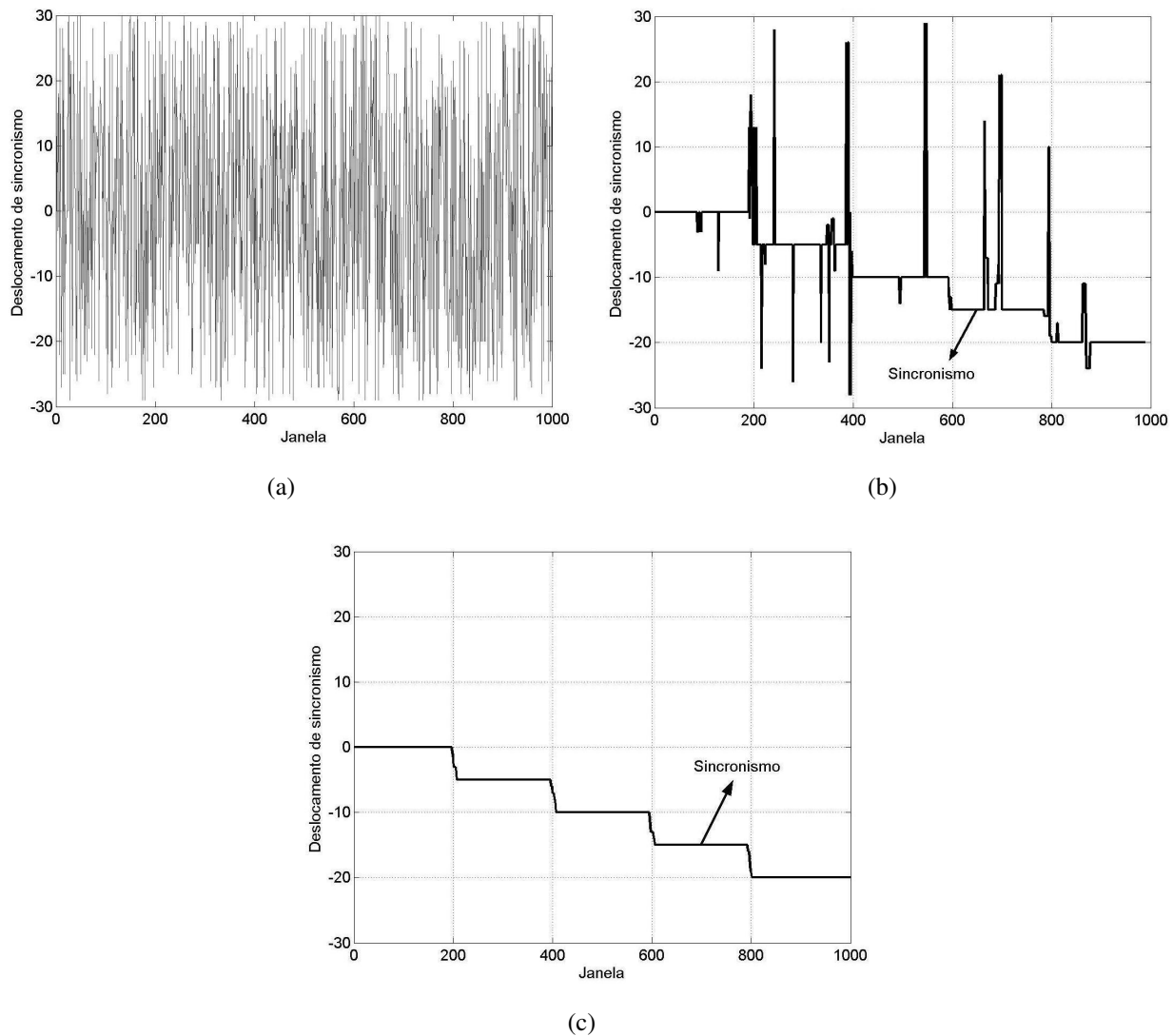


Fig. 6.9: Desempenho dos algoritmos com sinal digital de telecomunicações ($\text{SWR} = 20 \text{ dB}$) e sem a presença de distorções: 6.9(a) picos espúrios de correlação; 6.9(b) sincronização por média de janelas de análise; 6.9(c) sincronização por programação dinâmica.

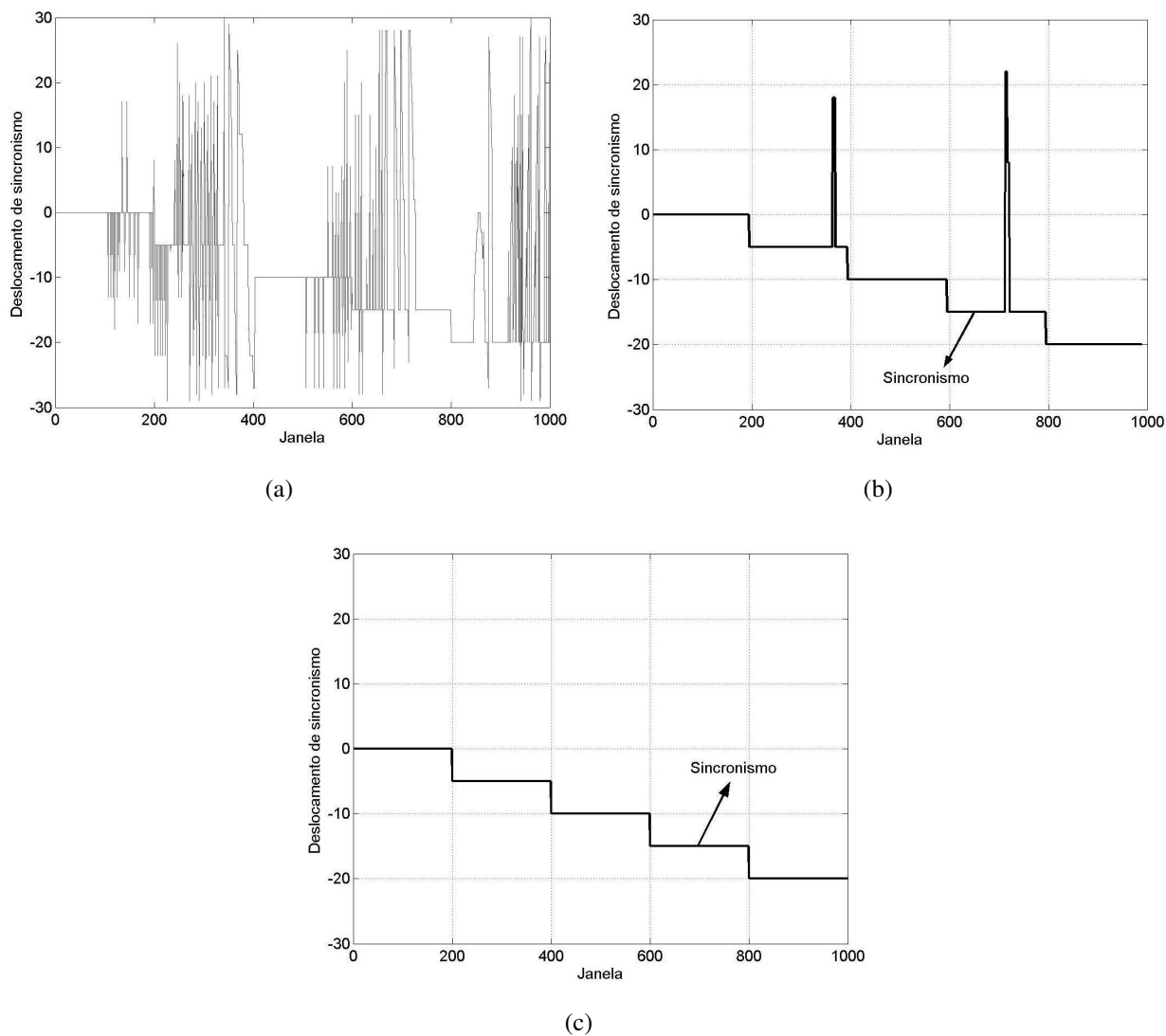


Fig. 6.10: Desempenho dos algoritmos com sinal de áudio ($SWR = 10$ dB) e sem a presença de distorções: 6.10(a) picos espúrios de correlação; 6.10(b) sincronização por média de janelas de análise; 6.10(c) sincronização por programação dinâmica.

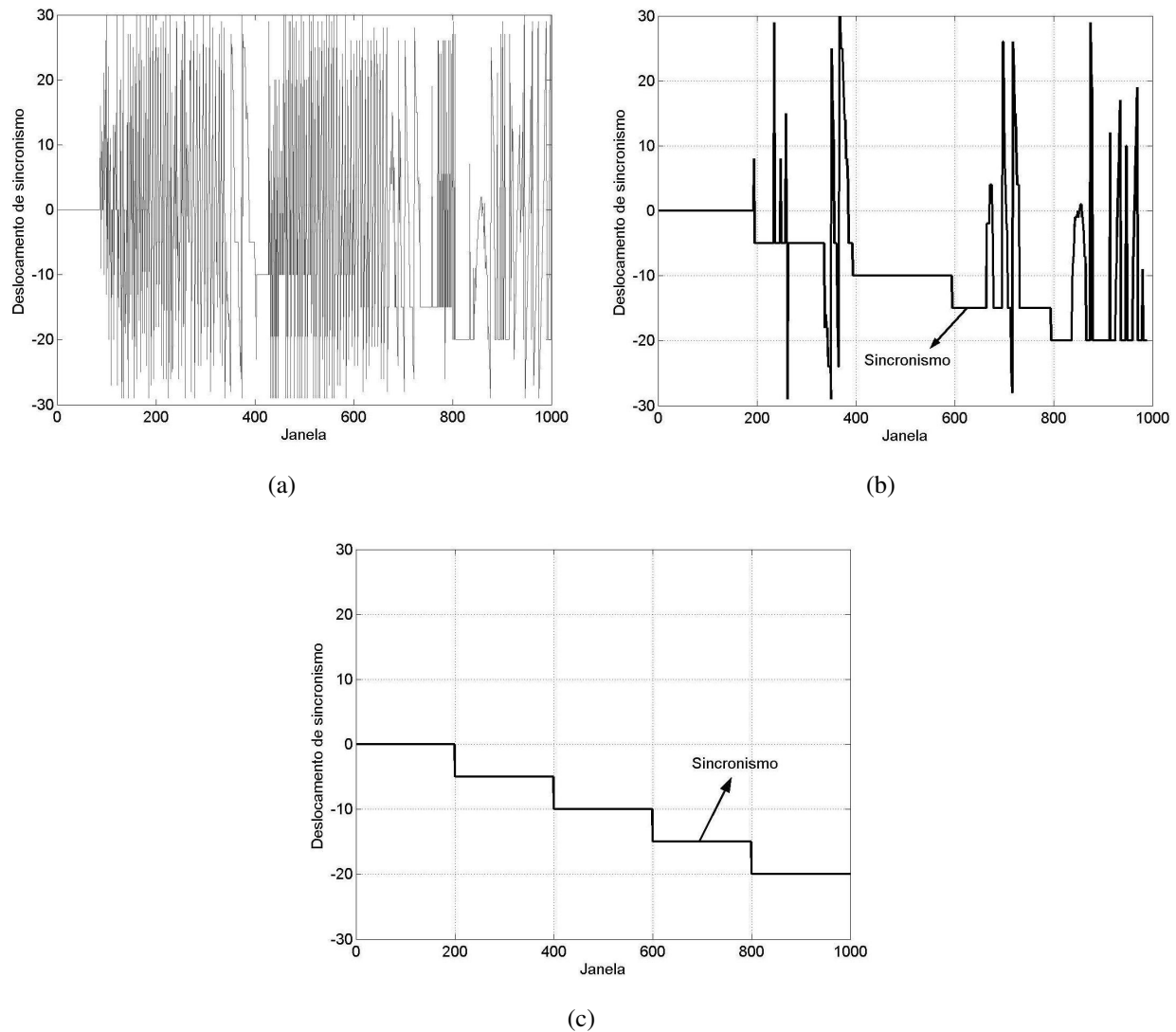


Fig. 6.11: Desempenho dos algoritmos com sinal de áudio (SWR = 20 dB) e sem a presença de distorções: 6.11(a) picos espúrios de correlação; 6.11(b) sincronização por média de janelas de análise; 6.11(c) sincronização por programação dinâmica.

deslocamentos da janela deslizante que produziram os maiores valores de correlação entre o sinal recebido e a marca d'água original, obtidos a partir da matriz **A**. Apesar de ter sido usado $\Lambda = 256$ no cálculo da matriz **A**, apresentamos um zoom da figura entre $\lambda = -25$ e $\lambda = 5$ para facilitar a visualização. Devido à presença do canal linear de fase mista, o deslocamento correspondente ao melhor sincronismo da marca inicia-se em $\lambda = 1$ (posição do coeficiente de maior amplitude do canal), diferentemente do caso anterior, no qual se iniciava em $\lambda = 0$. No entanto, o canal insere atrasos no sinal com soluções válidas para o sincronismo que variam de $\lambda = 0$ a $\lambda = 3$. A ilustração mostra que há uma concentração de máximos de correlação nesta faixa de deslocamentos, embora existam vários picos espúrios. A figura 6.12(b) apresenta o resultado do algoritmo utilizando média entre janelas. Nota-se que o método não foi capaz de determinar com exatidão o melhor sincronismo; entretanto, o deslocamento referente ao sincronismo da marca salta entre as diferentes soluções válidas impostas pelo canal. A figura 6.12(c) mostra que o algoritmo de programação dinâmica utilizando seqüência de blocos foi capaz de determinar os deslocamentos correspondentes ao melhor sincronismo para o sinal digital de telecomunicações, porém apresentou problemas nas transições entre deslocamentos, levando algum tempo para ajustar-se à nova condição de sincronismo. A figura 6.12(d) mostra que o algoritmo de programação dinâmica utilizando seqüência de símbolos de codebook foi capaz de rastrear o melhor sincronismo de forma bastante eficiente, inclusive nos momentos de transição.

A figura 6.13 apresenta os resultados para o sinal de fala utilizando os algoritmos propostos. A figura 6.13(a) apresenta os deslocamentos da janela deslizante que produziram os maiores valores de correlação entre o sinal recebido e a marca d'água original, obtidos a partir da matriz **A**, com $\text{SWR} = 20$ dB. Por facilidade de visualização, foi dado um zoom para deslocamentos λ de -25 a 5 . Diferentemente do caso do sinal digital, a ilustração mostra que os valores máximos de correlação para cada janela não se concentram na faixa de valores ($\lambda = 0$ a $\lambda = 3$) correspondentes aos sincronismos válidos. Isso ocorre devido ao sinal de fala ser altamente correlacionado, levando a uma quantidade de picos espúrios de correlação ainda maior do que no caso anterior. A figura 6.13(b) apresenta o resultado do algoritmo utilizando média entre janelas. Vemos que o método não foi capaz de determinar com exatidão o sincronismo da marca d'água. A ilustração mostra vários deslocamentos inexistentes. A figura 6.13(c) mostra que o algoritmo de programação dinâmica utilizando seqüência de blocos foi capaz de encontrar o deslocamento referente ao melhor sincronismo para o sinal de fala, porém apresentou problemas na transição de deslocamentos, também levando algum tempo para ajustar-se à nova condição de sincronismo. Por fim, a figura 6.13(d) mostra que o algoritmo de programação dinâmica utilizando seqüência de símbolos de codebook foi capaz de rastrear o melhor sincronismo de forma eficiente, mesmo com um sinal fortemente correlacionado.

Conforme ilustrado pelos resultados experimentais, o método de seqüência de símbolos de codebook foi capaz de identificar os deslocamentos correspondentes ao sincronismo da marca d'água

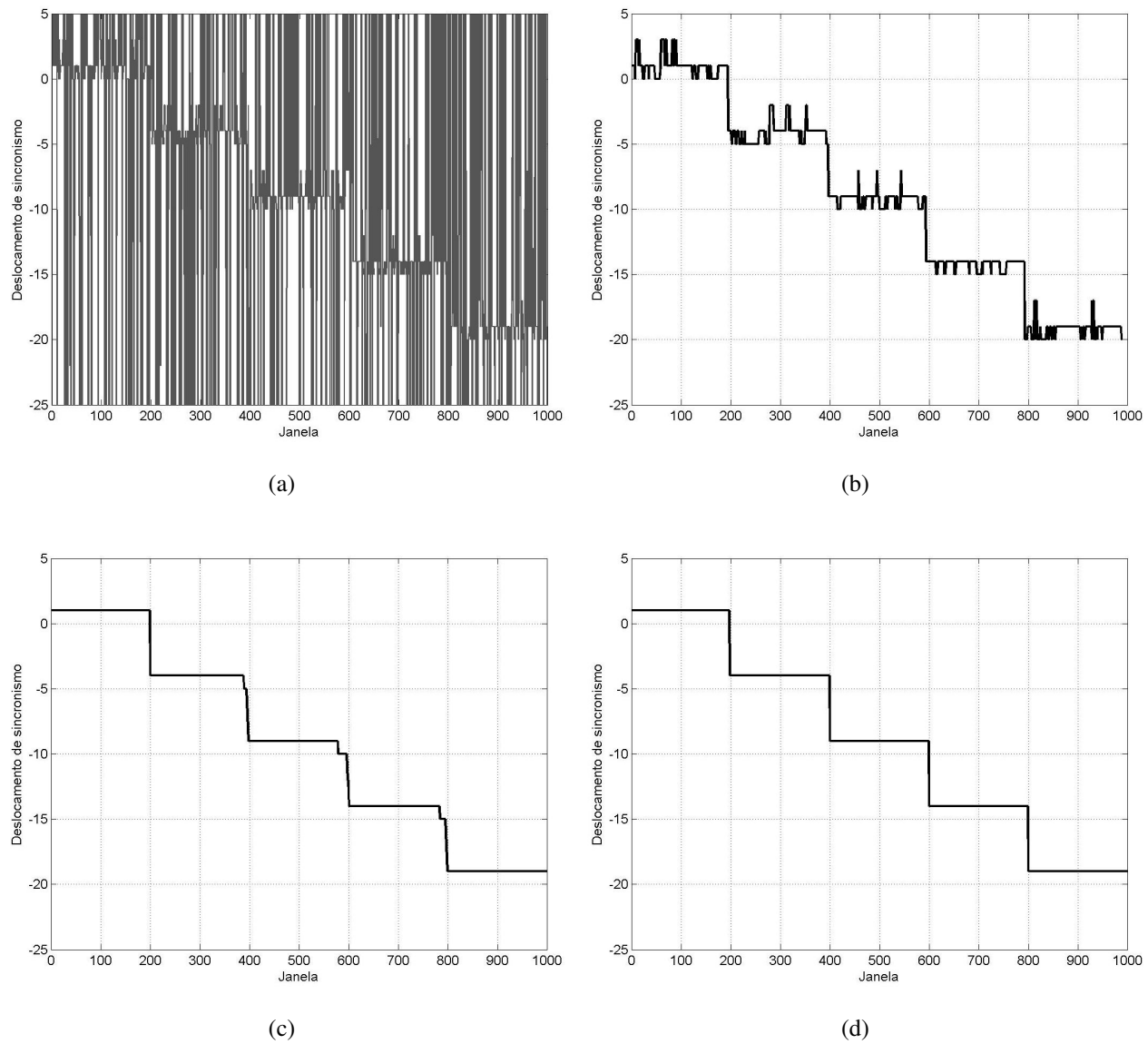


Fig. 6.12: Desempenho dos algoritmos de sincronismo com sinal digital de telecomunicações (SWR = 20 dB) e presença de sistema hostil: 6.12(a) picos espúrios de correlação; 6.12(b) média de janelas; 6.12(c) programação dinâmica utilizando bloco; 6.12(d) programação dinâmica utilizando codebook.

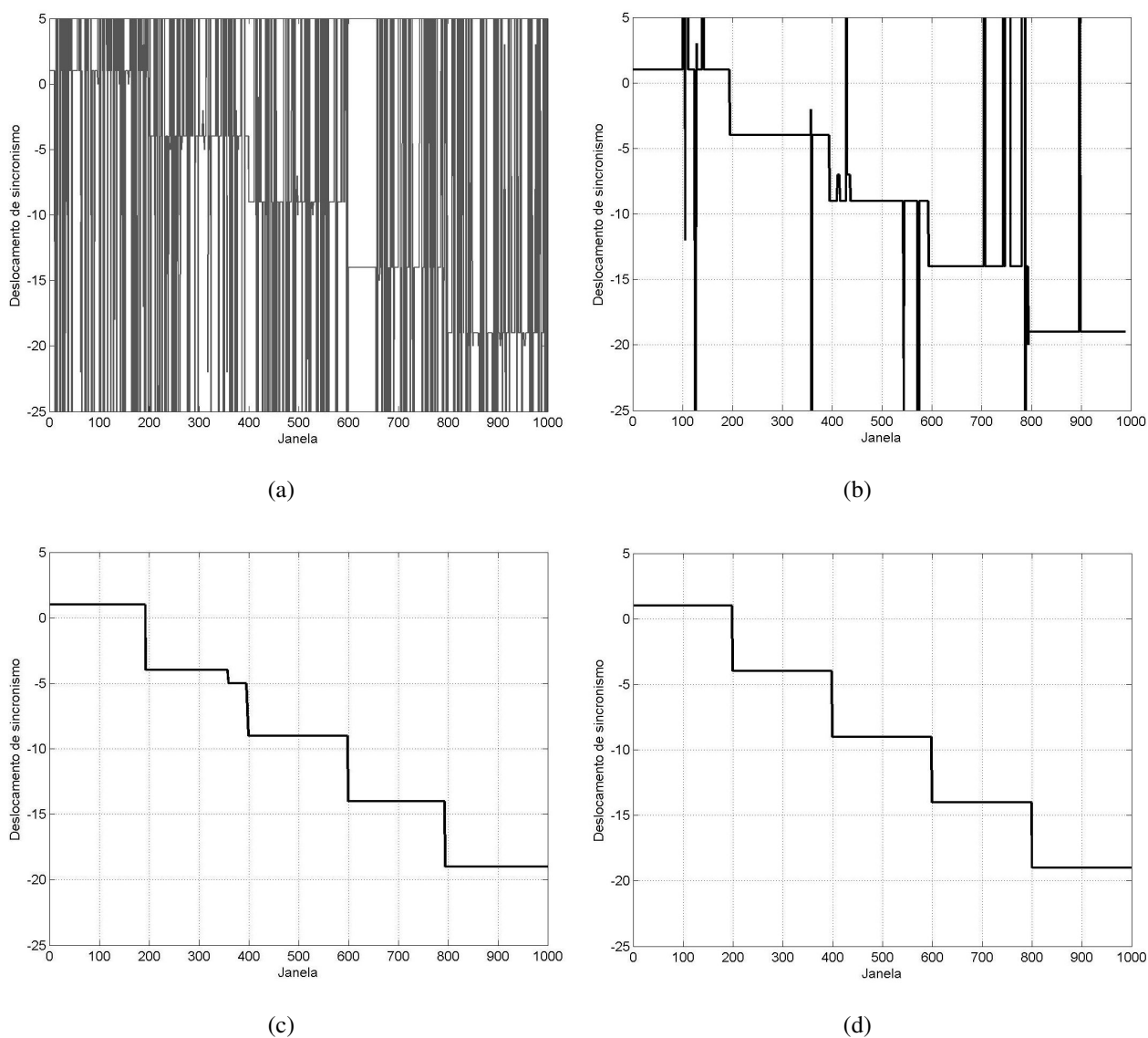


Fig. 6.13: Desempenho dos algoritmos de sincronismo com sinal de fala ($\text{SWR} = 20 \text{ dB}$) e presença de sistema hostil: 6.13(a) picos espúrios de correlação; 6.12(b) média de janelas; 6.12(c) programação dinâmica utilizando bloco; 6.12(d) programação dinâmica utilizando codebook.

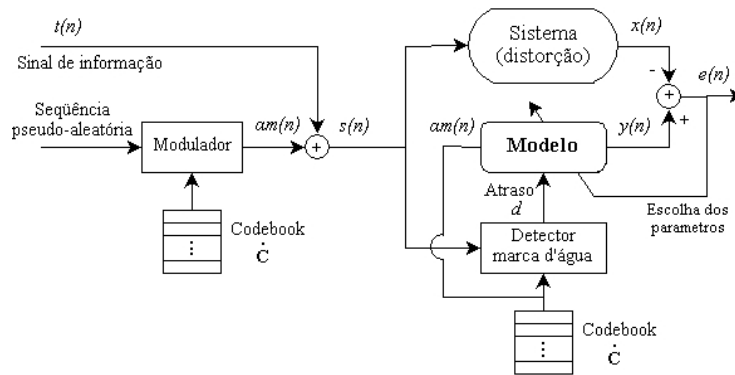


Fig. 6.14: Sistema completo de identificação utilizando marca d'água.

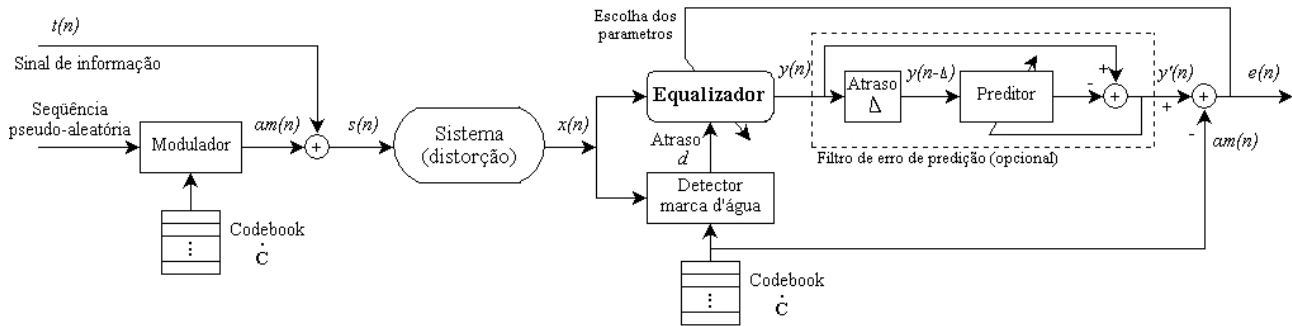


Fig. 6.15: Sistema completo de equalização utilizando marca d'água.

tanto para sinais hospedeiros descorrelacionados como correlacionados. O sistema proposto fornece estes deslocamentos para o filtro de equalização ou identificação, permitindo que este convirja para a melhor solução de Wiener, correspondente ao atraso ótimo do sistema.

As figuras 6.14 e 6.15 apresentam os sistemas completos de equalização e identificação, respectivamente, incluindo o método de sincronização por sequência de símbolos de codebook.

6.7 Conclusões

Neste capítulo, apresentamos dois métodos de sincronização aplicáveis aos sistemas propostos de equalização e identificação utilizando marca d'água. Os métodos são baseados em seqüências de treinamento constituídas pela informação embutida na marca d'água e detectadas por meio de correlação. Estes métodos revertem os efeitos de uma ampla classe de ataques de dessincronização, sendo capazes de monitorar o sincronismo continuamente ao longo de todo o sinal recebido.

Foram propostos dois algoritmos para determinação dos atrasos ótimos do sistema no contexto de equalização e identificação de canais. O primeiro, que é uma contribuição mais específica deste trabalho, baseia-se na média entre janelas de análise do sinal recebido. O segundo, proposto em (Go-

mes et al., 2001), baseia-se em uma técnica de programação dinâmica que leva em conta a correlação entre o sinal recebido e a marca d'água original, e introduz na função custo um termo que verifica se é respeitada uma sequência de símbolos de codebook usada na construção da marca d'água.

Nas simulações, foi abordado o caso de distorções impostas pela transmissão do sinal marcado através de um canal linear. Além disso, foi considerada a situação de perda esporádica de trechos do sinal transmitido. Os resultados experimentais mostram que os métodos foram bem-sucedidos ao encontrar o atraso inicial requerido para o sincronismo da marca d'água. O segundo método apresentado, baseado em programação dinâmica, foi também capaz de ressincronizar continuamente o sinal recebido, corrigindo rapidamente falhas no canal de comunicação que causem perda de sincronismo.

Nas próximas etapas deste trabalho, prevê-se o estudo do comportamento dos métodos propostos frente a ataques mais severos, tais como filtragem não-linear, codificação/decodificação e reamostragem. Além disso, no caso de aplicações de áudio e voz, será considerado o uso de modelos psicoacústicos para aumentar a energia da marca d'água sem introdução de degradações perceptíveis (tema abordado no capítulo 7, porém sem levar em conta questões de sincronismo).

Capítulo 7

Equalização e Identificação para Sinais de Áudio Utilizando Marca d'Água

7.1 Introdução

A adição de uma marca d'água em um sinal de áudio pode causar distorções perceptíveis. Tais distorções poderiam ser minimizadas por meio do escalonamento da marca d'água, reduzindo a sua potência. No entanto, o ouvido humano é muito sensível, percebendo muitas vezes ruídos com pequenas amplitudes. Além disso, pudemos ver nos capítulos 4, 5 e 6 que baixos valores de SWR tornam o esquema de marca d'água menos robusto. O simples escalonamento do sinal de marca d'água não representa portanto uma solução adequada para sinais que possuam interpretação sensorial. Uma solução para garantir a inaudibilidade da marca d'água é empregar um modelo psicoacústico em conjunto com um algoritmo de conformação espectral. A conformação espectral da marca aumenta a robustez do sistema, pois a potência da marca d'água é maximizada, aproximando-a do limiar de mascaramento do sinal de áudio.

Neste capítulo, analisaremos o desempenho dos métodos de equalização e identificação utilizando marca d'água em sinais de áudio através de técnicas perceptivas que exploram os fenômenos psicoacústicos. Além disso, definiremos a faixa de relação sinal-marca d'água necessária para garantir inaudibilidade em áudio.

7.2 Marca d'Água em Áudio Digital

Um esquema de marca d'água pode ser visto como um sistema de comunicação: a marca d'água é o sinal que carrega informação e o sinal de áudio pode ser encarado como ruído. Em sistemas de comunicação tradicionais, a potência do sinal de informação é normalmente maior do que a do ruído,

e este último é muitas vezes assumido como sendo gaussiano e branco. No entanto, este geralmente não é o caso em marca d'água de áudio. Para evitar degradações audíveis, o sinal de marca d'água deve ter, via de regra, uma potência muito inferior à do sinal de áudio; além disso, este último é geralmente não-estacionário e fortemente colorido.

Existem diferentes métodos para inserção de marca d'água em áudio digital. Por exemplo, a marca d'água pode ser adicionada ao sinal de áudio nos domínios do tempo ou da frequência, e podem ser explorados fenômenos de mascaramento temporal e frequencial. Além disso, diferentes chaves podem ser usadas na inserção e na detecção da marca. Ainda, em aplicações de áudio, o sinal marcado pode sofrer diversas operações com potencial de dessincronizar a detecção da marca, tais como transmissão, compressão e conversão A-D/D-A.

O primeiro artigo de que temos conhecimento referente a marca d'água de áudio data de 1996 (Boney et al., 1996). O esquema de marca d'água proposto neste artigo é basicamente uma extensão das técnicas anteriores desenvolvidas para marca d'água em imagem. A idéia consiste em adicionar uma sequência inicialmente branca ao sinal de áudio. Para garantir inaudibilidade ao mesmo tempo em que se maximiza o nível de energia da marca, a sequência passa por um processo de conformação espectral, aproximando-se do limiar de mascaramento do áudio calculado através de um modelo psicoacústico; isto possibilita aumentar a robustez da marca d'água em relação ao que se teria com uma marca branca, cuja potência deveria ser muito mais baixa para que não houvesse distorções perceptíveis no áudio.

7.2.1 Modelos Psicoacústicos

A *psicoacústica* é o estudo da percepção sonora. Através de experimentos psicoacústicos, estabeleceu-se que o ouvido humano apresenta importantes limitações. Em particular, quando dois tons suficientemente próximos em frequência são tocados simultaneamente, pode ocorrer o fenômeno de *mascaramento*: se um dos tons é suficientemente alto, ele irá *mascarar* o outro (Zwicker and Fastl, 1990).

Modelos psicoacústicos generalizam o efeito do mascaramento para sinais não tonais. Dado um sinal de áudio $u(t)$, o modelo calcula uma curva $M_u(f)$ chamada de *limiar de mascaramento*, homogênea a uma densidade espectral de potência¹ (PSD) (Guimarães, 1998). Se a PSD $V(f)$ de um sinal $v(t)$ está abaixo de $M_u(f)$ para todas as frequências, então $v(t)$ é mascarado por $u(t)$. Isto significa que o ouvinte não é capaz de perceber diferença alguma entre $u(t)$ e $u(t) + v(t)$. Este fenômeno é ilustrado na figura 7.1. Tais modelos são usados em diversos algoritmos de compressão de áudio com perdas, como o MP3 e o MPEG-AAC (ISO/IEC-13818-7, 1997), Dolby e DTS. Para tornar o ruído

¹Em inglês, *power spectral density*.

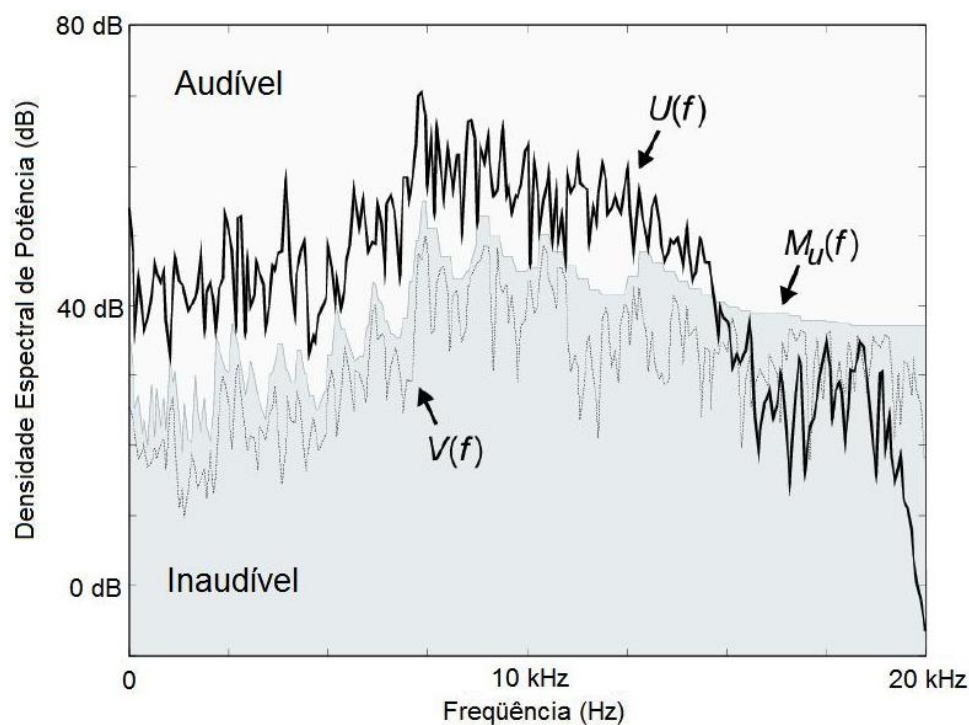


Fig. 7.1: Densidades espectrais de potência dos sinais mascarador ($U(f)$, linha contínua) e mascarado ($V(f)$, linha tracejada), e limiar de mascaramento ($M_u(f)$). Nas frequências altas, observa-se que o limiar de mascaramento e o sinal mascarado ultrapassam a PSD do sinal mascarador devido à baixa sensibilidade do ouvido humano nesta faixa de frequências.

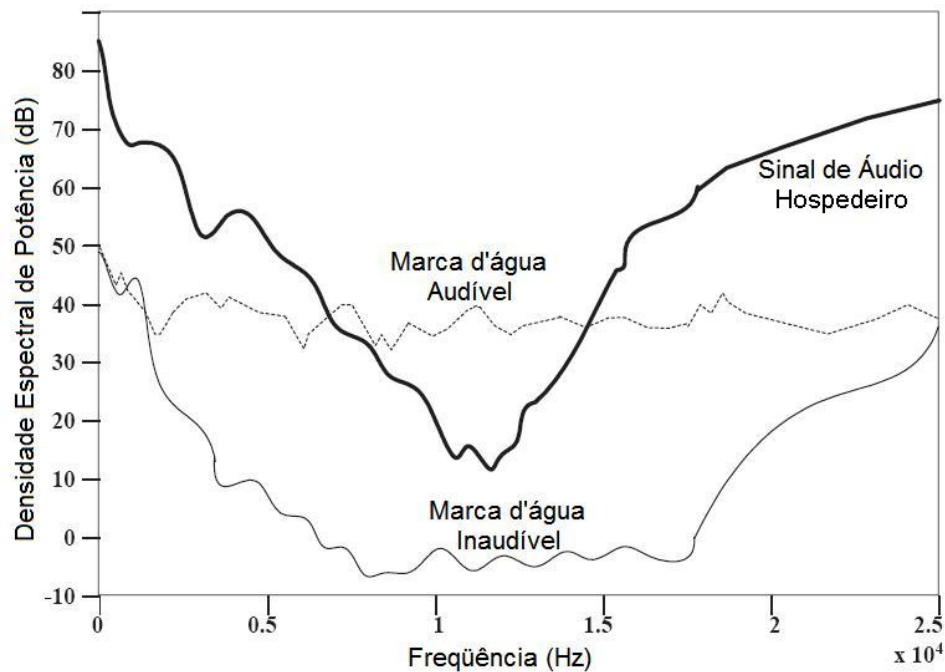


Fig. 7.2: Exemplo de conformação espectral da marca d'água. A marca audível (linha tracejada) é transformada na marca inaudível (linha contínua) com base no limiar de mascaramento.

de quantização inaudível, estes algoritmos utilizam-se do limiar de mascaramento para distribuir os bits disponíveis para codificação entre bandas de frequência.

Em sistemas de marca d'água para áudio, modelos psicoacústicos podem ser usados para estimar o limiar de mascaramento do sinal de áudio e fornecê-lo a um algoritmo de conformação espectral usado para filtrar uma sequência pseudo-aleatória descorrelacionada. Esta operação resulta no sinal de marca d'água, cuja densidade espectral de potência estará abaixo do limiar de mascaramento do áudio e que não poderá ser ouvido na presença deste último. Graças aos modelos psicoacústicos, a inaudibilidade da marca d'água pode ser garantida com SWR de aproximadamente 20 dB (Gomes et al., 2003). Uma marca d'água branca requer valores bem maiores de SWR (de 40 a 60 dB, conforme o áudio (Gomes, 2002)) para evitar distorções audíveis, tornando mais difícil a detecção. A figura 7.2 apresenta as PSDs de sinais de marca d'água audível e inaudível, juntamente com a PSD do áudio hospedeiro. A marca d'água inaudível foi obtida a partir da marca audível (sequência pseudo-aleatória), aplicando-se uma operação de conformação espectral baseada no limiar de mascaramento.

Em linhas gerais, o processo para determinação do limiar de mascaramento em frequência compõe-se dos seguintes passos:

1. Cálculo da densidade espectral de potência do sinal mascarador;
2. Localização de componentes tonais e não tonais no espectro;

3. Dizimação do sinal mascarador para eliminação de informação irrelevante;
4. Cálculo do limiar de mascaramento por banda crítica².

Este procedimento é efetuado para cada janela de áudio, dentro da qual supõe-se que o sinal de áudio é aproximadamente estacionário.

O limiar de mascaramento define a resposta em frequência do filtro de conformação espectral usado para filtrar a marca d'água. A PSD da marca passa a estar abaixo do limiar de mascaramento, embora muito próxima deste, para todas as frequências do sinal. Com isso, a energia da marca d'água é maximizada, aumentando a robustez do sistema.

Caso se necessite, por um motivo qualquer (e.g. maior robustez ou menor distorção), de uma marca d'água com potência diferente da obtida pelo processo descrito no parágrafo anterior, o sinal de marca d'água pode ser escalonado após a conformação espectral. Se a PSD da marca resultante ultrapassar o limiar de mascaramento em uma ou mais bandas do espectro, a marca se tornará audível; no entanto, ainda assim a distorção percebida tenderá a ser significativamente menos acentuada do que a que seria obtida com uma marca de energia uniformemente distribuída em frequência (Gomes, 2002).

O fenômeno de mascaramento também pode ocorrer no domínio do tempo. Se dois sons ocorrerem em instantes suficientemente próximos, e a diferença de potência entre eles é suficientemente alta, o som mais potente pode mascarar o som mais fraco. Caso o som mascarado preceda o som mascarador, o efeito denomina-se *pré-mascaramento*; caso a ordem dos sons seja inversa, tem-se o *pós-mascaramento*. Efeitos de mascaramento temporal são explorados em alguns métodos de compressão com perdas (ISO/IEC-13818-7, 1997), bem como em algumas técnicas de marca d'água; na técnica de *echo-hiding*, por exemplo, a marca d'água é uma versão atrasada e atenuada do sinal de áudio (Bender et al., 1996).

Além das já citadas, várias outras aplicações têm sido propostas para modelos psicoacústicos, tais como o cancelamento de eco, avaliação automática da qualidade de áudio e classificação automática de áudio. Maiores detalhes sobre modelos psicoacústicos podem ser encontrados no apêndice A.

7.3 Equalização e Identificação de Sinais de Áudio Utilizando Marca d'Água

Para garantir a inaudibilidade da marca d'água, a sequência pseudo-aleatória $m(n)$ deve submeter-se a uma conformação espectral, adequando-se a um limiar de mascaramento. Este processo é reali-

²Banda crítica define uma faixa em torno de uma frequência central, a qual está associada a um ponto da membrana basilar.

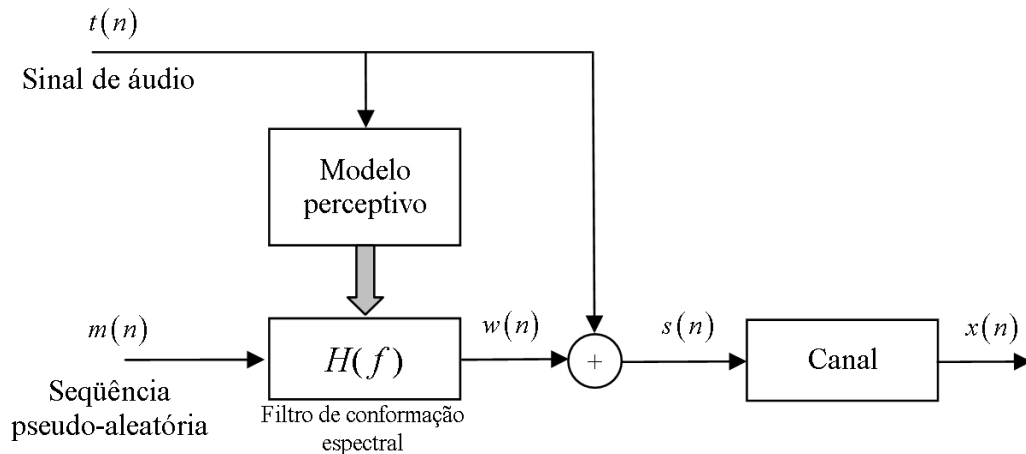


Fig. 7.3: Esquema de inserção de marca d'água em áudio digital.

zado por meio de um filtro $H(f)$, cuja resposta em frequência segue o limiar de mascaramento obtido através do modelo psicoacústico³. O sinal $w(n)$ resultante é simplesmente adicionado ao sinal de áudio $t(n)$, produzindo o sinal marcado $s(n)$. Esta abordagem é ilustrada na figura 7.3.

De forma geral, podemos identificar em sistemas de marca d'água de áudio três fontes principais de ruído:

1. O próprio sinal de áudio, que é aproximadamente 20 dB mais potente do que a marca d'água;
2. Distorções produzidas por operações lícitas, como filtragem e compressão/descompressão, podendo apresentar níveis de energia comparáveis aos da marca d'água se a qualidade sonora não for severamente degradada;
3. Distorções produzidas por operações ilícitas (ataques de piratas com o objetivo de tornar a marca d'água indetectável.)

O ruído introduzido por essas fontes normalmente é colorido; por exemplo, o ruído de quantização introduzido por uma operação de compressão/descompressão com codificadores perceptivos tende a acompanhar o limiar de mascaramento. Abordaremos aqui apenas o caso de distorção imposta por um processo de filtragem FIR linear, que também apresenta a característica de colorir o sinal.

Para que seja possível utilizar a informação da marca d'água como sinal de referência nos filtros identificador ou equalizador, o sinal $x(n)$ (sinal marcado e distorcido pelo canal) é primeiramente filtrado por $G(f)$. Este último é um filtro de conformação espectral com resposta em frequência invertida com relação ao limiar de mascaramento do sinal $x(n)$. O objetivo deste filtro é contrabalançar

³Para esta operação funcionar corretamente, o sinal $m(n)$ deve ser aproximadamente branco e com energia normalizada.

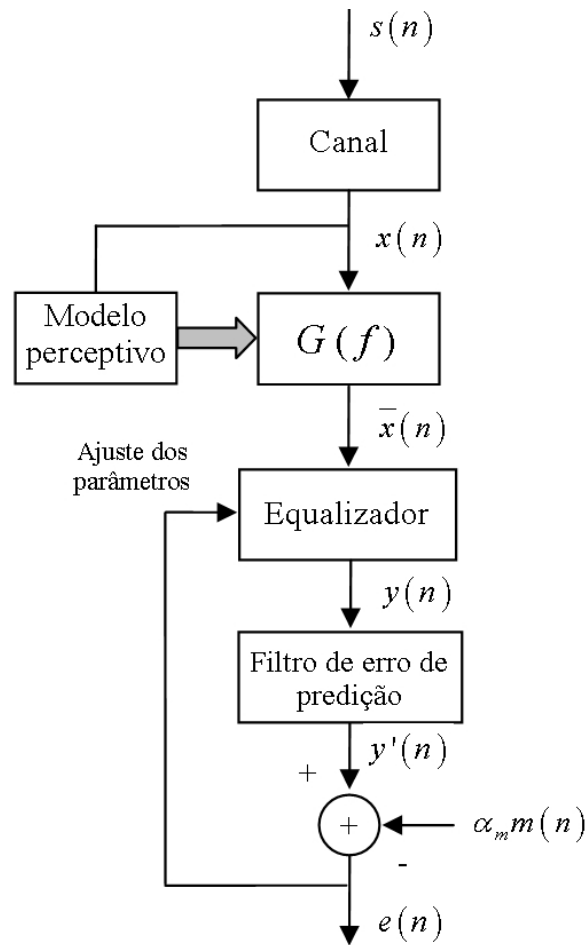


Fig. 7.4: Esquema de equalização utilizando marca d'água em áudio.

a conformação espectral $H(f)$ imposta ao sinal de marca d'água para garantir a sua inaudibilidade. Como o sinal de áudio é geralmente muito mais potente do que a marca d'água, o limiar de mascaramento calculado a partir do sinal marcado tende a ser próximo do obtido para o sinal hospedeiro puro; além disso, supõe-se que a distorção do sistema não será forte o bastante para inutilizar o limiar de mascaramento calculado a partir do sinal recebido. Com isso, o sinal $\bar{x}(n)$ na saída do filtro $G(f)$ contém uma estimativa do sinal $m(n)$ distorcido pelo canal mais o sinal de áudio distorcido pelo canal e filtrado por $G(f)$. O sinal $\bar{x}(n)$ é então utilizado nos processos de equalização (figura 7.4) e identificação (figura 7.5), conforme discutido nos capítulos 4 e 5, respectivamente.

7.4 Simulações e Resultados

Com o objetivo de analisar os métodos de equalização e identificação propostos para sinais de áudio, dividimos as simulações em três fases distintas. A primeira busca avaliar a qualidade do

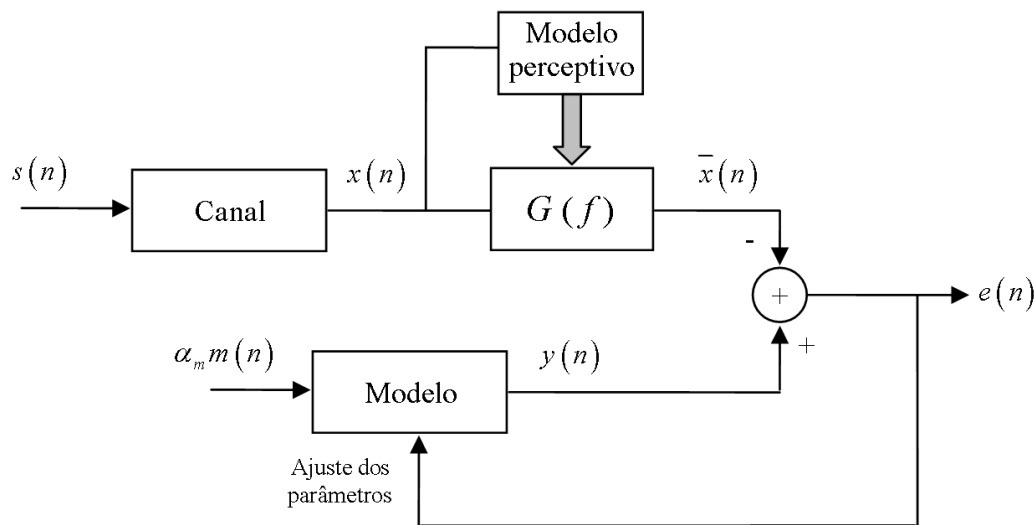


Fig. 7.5: Esquema de identificação utilizando marca d'água em áudio.

método de inserção de marca d'água em áudio. Para isto, é utilizada uma metodologia de avaliação objetiva de qualidade de áudio, buscando-se definir o menor valor da relação sinal-marca d'água que garanta transparência (i.e. para o qual a presença da marca d'água seja imperceptível). As outras duas fases prestam-se à avaliação dos algoritmos de equalização e identificação, respectivamente, em conjunto com um modelo psicoacústico e um algoritmo de conformação espectral.

7.4.1 Análise do Algoritmo de Marca d'Água para Áudio

Nos resultados apresentados nesta seção, foi empregado o modelo psicoacústico número 1 do padrão MPEG-1 (ISO/IEC-11172-3, 1993), disponibilizado pela ISO em implementação de referência. O filtro de conformação espectral é obtido a partir do limiar de mascaramento por meio do algoritmo de Levinson, conforme descrito em (Gomes, 2002).

Duas formas tradicionais para avaliação da qualidade de sinais de fala e áudio são a avaliação subjetiva (inspeção auditiva) e a avaliação objetiva (medida gerada por software). Uma avaliação subjetiva rigorosa deve seguir métodos padronizados para a geração de notas de avaliação de qualidade por avaliadores humanos (P.85, 1994; P.830, 1996; BS.1116-1, 1996), sendo dispendiosa em termos de tempo e requisitos de infra-estrutura. A avaliação objetiva, por sua vez, substitui os avaliadores humanos por algoritmos, também padronizados (P.861, 1998; P.862, 2001), cuja função é modelar o comportamento desses avaliadores, empregando modelos psicoacústicos que estimam as características perceptivas do aparelho auditivo humano. O algoritmo mais empregado para avaliação de sinais de fala é o PESQ (*Perceptual Evaluation of Speech Quality*), padronizado na norma de refe-

rência P.862 da ITU-T⁴ (P.862, 2001); para sinais de áudio genérico, o algoritmo mais empregado é o PEAQ (*Perceptual Evaluation of Audio Quality*), que deu origem à recomendação BS.1387-7 da ITU-R⁵ (BS.1387-1, 1998).

Para avaliação do algoritmo de inserção de marca d'água, foram selecionados materiais de áudio variados e potencialmente sensíveis a degradações. Quatro clipes foram escolhidos, seguindo avaliações realizadas pelo *Communications Research Centre* do Canadá (Soulodre et al., 1998), incluindo fala, música, e fala com música. A tabela 7.1 apresenta o material escolhido para os testes.

Arquivo	Descrição	Duração	Fonte
<i>dirs</i>	<i>Dire Straits</i>	10 s	Warner Bros. CD 7599-25264-2 (trilha 6)
<i>svega</i>	<i>Suzanne Vega (Capela)</i>	10 s	AT&T mix
<i>trump</i>	<i>Solo de trompete</i>	10 s	Gravação original, Universidade de Miami
<i>symp</i>	<i>Orquestra</i>	10 s	EBU SQAM CD (trilha 17)

Tab. 7.1: Material de áudio utilizado nas avaliações de qualidade.

Para a análise dos materiais de áudio selecionados, utilizamos uma implementação do PEAQ contida no software *OPERA* da empresa OPTICOM (Opticom, 2007). De maneira análoga aos testes subjetivos de audição, a qualidade do sinal de áudio sob teste é avaliada em relação a um sinal de referência (o sinal de áudio original). Existem duas versões do PEAQ: a versão *básica*, caracterizada por uma abordagem de baixa complexidade, e a versão *avançada*, que apresenta maior acurácia à custa de uma maior complexidade. A versão *básica* foi utilizada em nossos testes, pois não apresentou diferenças significativas em relação à versão *avançada* para os propósitos desta avaliação. Foram utilizados arquivos de áudio no formato Wave, 16 bits por amostra, mono, com taxa de amostragem de 44,1 kHz (embora a padronização do PEAQ requeira áudio amostrado a 48 kHz, a implementação utilizada possui uma extensão proprietária para amostragem a 44,1 kHz). As diferenças entre os arquivos de áudio são analisadas por um modelo cognitivo nos domínios temporal e freqüencial, extraindo-se características perceptivamente relevantes dos sinais. A norma BS.1387 define um conjunto de variáveis de saída denominadas *Model Output Variables* (MOV). A figura 7.6 mostra a tela de saída do software *OPERA* para as MOVs. Detalhes referentes a todas as MOVs podem ser encontrados na norma ITU-R BS.1387-7 (BS.1387-1, 1998).

As MOVs obtidas são mapeadas em um indicador único de qualidade de áudio denominado *Objective Difference Grade* (ODG). Este indicador apresenta correspondência numérica direta com resultados de avaliações subjetivas de qualidade de áudio medidas através do indicador *Subjective Difference Grade* (SDG), podendo ser interpretado como uma estimativa deste último. A figura 7.7 ilustra a re-

⁴International Telecommunication Union - Telecommunications.

⁵International Telecommunication Union - Radiocommunication.

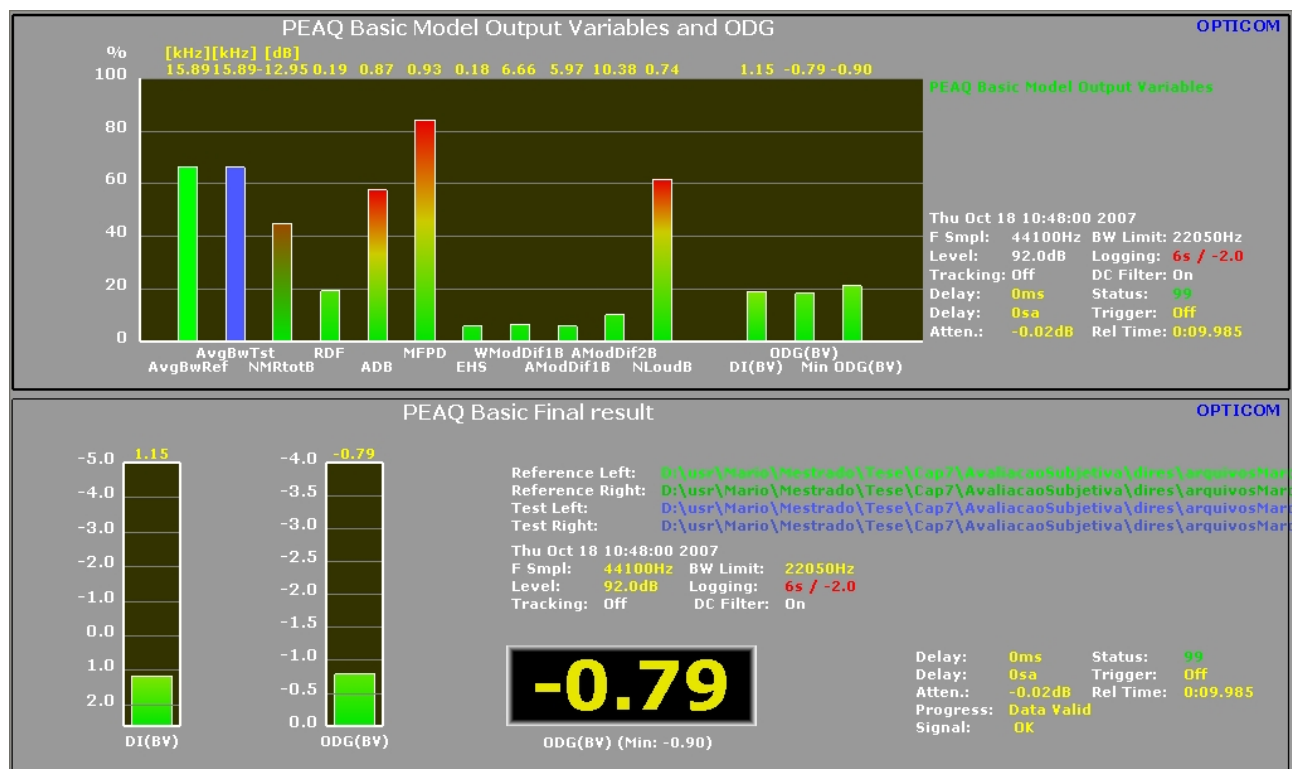


Fig. 7.6: Tela apresentando as variáveis de saída (MOVs) e o resultado global da avaliação (ODG) do PEAQ básico.

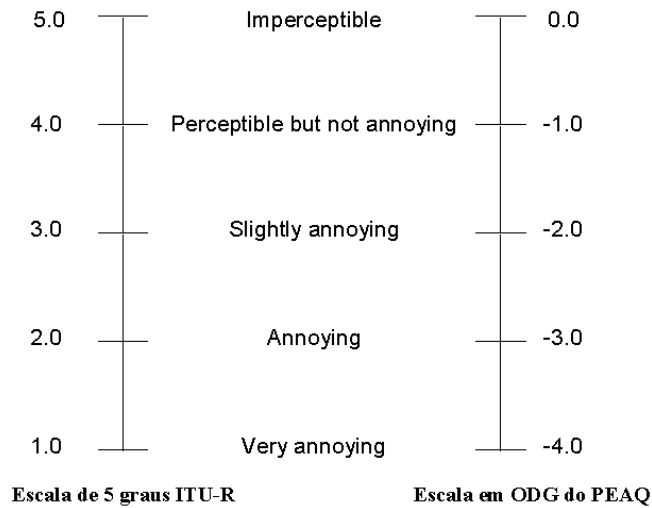


Fig. 7.7: Escala de degradação de 5 graus ITU-R versus ODG PEAQ.

lação entre os valores de ODG (e SDG) e a escala de degradação subjetiva de cinco graus definida na recomendação ITU-R BS.562-3 (BS.562-3, 1990). Esta relação se expressa por:

$$ODG = E_{score_{\text{signal sob teste}}} - E_{score_{\text{signal de referencia}}} \quad (7.1)$$

O indicador ODG é a medida normalmente apresentada nos ensaios de avaliação objetiva de qualidade de áudio. A escala desta medida é contínua, variando de 0,0 (degradação imperceptível) e -4,0 (degradação muito incômoda), com resolução limitada a uma casa decimal. Entre -4,0 e -3,1, a degradação é classificada como *muito incômoda* (*very annoying*); entre -3,0 e -2,1, como *incômoda* (*annoying*); entre -2,0 e -1,1, como *levemente incômoda* (*slightly annoying*); entre -1,0 e -0,1, como *perceptível mas não incômoda* (*perceptible but not annoying*); e, quando igual a 0,0, como *imperceptível* (*imperceptible*). Esta classificação encontra-se resumida na tabela 7.2. Degradações na faixa de -1,0 a -0,1 podem ser consideradas imperceptíveis para ouvintes comuns (i.e. sem treinamento específico para detecção de degradações em áudio), enquanto degradações compreendidas entre -0,5 e -0,1 são em geral imperceptíveis até mesmo para especialistas em áudio, tal como músicos ou engenheiros de áudio (Arnold, 2002; Neubauer and Herre, 1998).

A figura 7.8 apresenta curvas de ODG em função da SWR para o material de áudio selecionado (tabela 7.1), no qual foi inserida uma marca d'água explorando características psicoacústicas através do algoritmo descrito nas seções precedentes. A curva referente ao sinal *diret* (figura 7.8(a)) atinge um valor de ODG de aproximadamente -0,75 para SWR = 20 dB, encontrando-se na faixa que garante transparência da marca d'água para um ouvinte comum. Para SWR = 23 dB, o valor de ODG é de cerca de -0,45, de modo que a marca é transparente até mesmo para especialistas em áudio. Para o sinal *svega*, a figura 7.8(b) mostra que, para SWR = 20 dB, é atingido o limiar de ODG = -1,0;

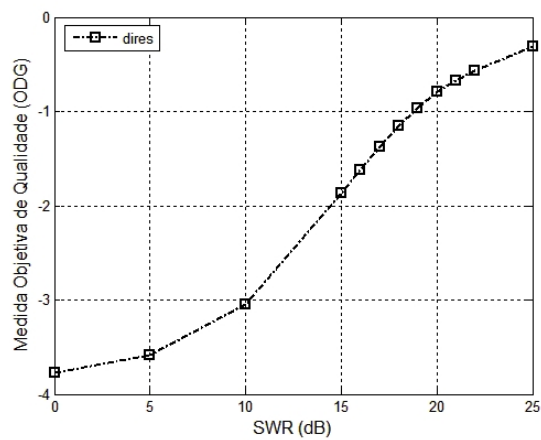
Degradação	ODG
<i>Imperceptível (imperceptible)</i>	0, 0
<i>Perceptível mas não incômoda (perceptible but not annoying)</i>	-1, 0
<i>Levemente incômoda (slightly annoying)</i>	-2, 0
<i>Incômoda (annoying)</i>	-3, 0
<i>Muito incômoda (very annoying)</i>	-4, 0

Tab. 7.2: Escala ODG com cinco níveis de degradação de áudio.

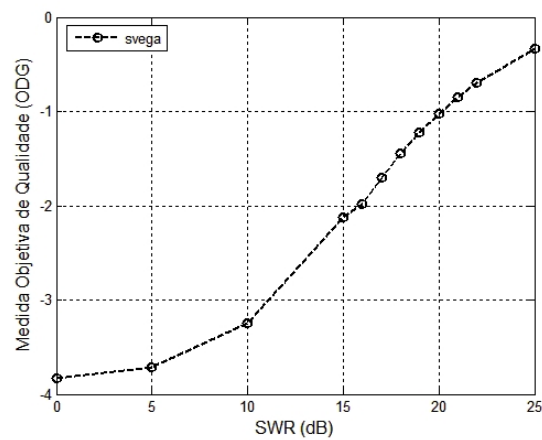
para $SWR = 23$ dB, o valor de ODG encontra-se muito próximo de $-0,5$. Embora satisfatórios, os resultados para este sinal são menos favoráveis do que para o sinal *diret*, o que se explica pelo fato de se tratar de uma versão capela contendo voz sem instrumentos musicais, com energia concentrada em frequências abaixo de 10 kHz; o efeito de mascaramento é, portanto, mais intenso nesta faixa de frequências, levando a energia da marca a concentrar-se também nesta faixa. Como o ouvido humano apresenta maior sensibilidade em baixas frequências (ver apêndice A), torna-se mais difícil evitar a percepção da marca d'água neste caso. A figura 7.8(c) apresenta a curva ODG versus SWR para o sinal *trump*. Observamos um resultado superior ao do sinal *svega*, com o índice ODG em torno de $-0,95$ para $SWR = 20$ dB e de $-0,45$ para $SWR = 23$ dB. O sinal de orquestra (*symph*, figura 7.8(d)) apresenta o melhor resultado dentre as amostras analisadas. Devido ao sinal ser composto por um conjunto de diversos instrumentos musicais, seu espectro contém componentes espalhadas em toda a faixa de frequências (limitada a cerca 20 kHz devido à taxa de amostragem); por esta razão, o limiar de mascaramento e a potência da marca d'água também se encontram distribuídos ao longo de todo o espectro. O índice ODG encontra-se acima do limiar de $-0,5$ para $SWR = 20$ dB e em torno de $-0,15$ para $SWR = 23$ dB.

A figura 7.9 apresenta a curva SWR versus ODG para todas as amostras de áudio analisadas, incluindo também a curva média. Observa-se que, para $SWR = 20$ dB, tem-se um valor ODG médio de $-0,8$, dentro da faixa de imperceptibilidade para um ouvinte comum. Para $SWR = 23$ dB, tem-se um valor ODG médio de $-0,45$, dentro da faixa de imperceptibilidade para especialistas em áudio. Sendo assim, podemos dizer que a faixa de SWR entre 20 e 23 dB é apropriada para sistemas de áudio. O ajuste fino desta relação depende da aplicação em questão.

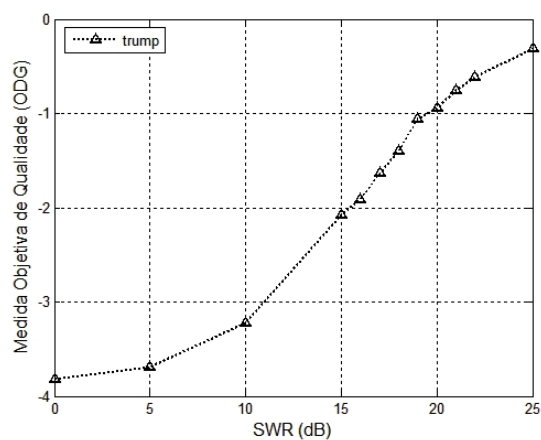
Como referência para avaliação dos resultados obtidos, comparamos a avaliação objetiva média da distorção provocada pela marca d'água, efetuada por meio do PEAQ, com uma avaliação subjetiva formal de codificadores de áudio disponível na literatura, realizada pelo *CRC Signal Processing and Psychoacoustics Audio Perception Lab* ((Soulodre et al., 1998)). Este ensaio testou implementações dos codecs MPEG-1 layer 2 e layer 3, MPEG-2 AAC Main, AC-3 (Dolby) e PAC (*Perceptual Audio Codec*, Lucent), seguindo os procedimentos estabelecidos na recomendação ITU-R BS.1116 ((BS.1116-1, 1996)). O melhor resultado foi obtido com o MPEG-2 AAC, que apresentou uma me-



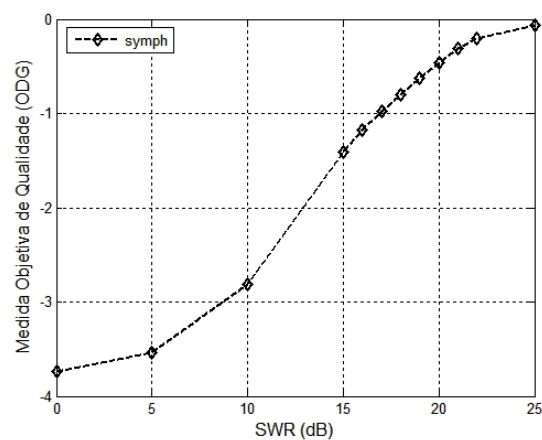
(a)



(b)



(c)



(d)

Fig. 7.8: Medida objetiva de qualidade: 7.8(a) *dires*; 7.8(b) *svega*; 7.8(c) *trump*; 7.8(d) *symph*.

dida SDG de aproximadamente $-0,50$ com uma taxa de saída de 128 kb/s, valor este próximo ao obtido para a marca d'água com SWR média de 23 dB. Outro codec que atingiu SDG próximo de $-0,50$ foi o AC-3 com taxa de 192 kb/s. O codec PAC atingiu SDG máximo de $-0,80$ com taxa de 160 kb/s, equivalente ao algoritmo de marca d'água para $\text{SWR} = 20$ dB. As duas versões do MPEG-1, layer 2 e layer 3, não ultrapassaram o limiar de $-1,0$, independentemente da taxa de bits selecionada.

Com o objetivo de justificar a necessidade do modelo psicoacústico, apresentamos na figura 7.10 a curva SWR versus ODG para as amostras de áudio analisadas, nas quais foi inserida uma marca d'água branca (sem o uso do modelo psicoacústico e do filtro de conformação espectral, semelhante ao método de *superimposed training*). Vê-se que o resultado é muito inferior ao obtido com modelo psicoacústico (figura 7.9). Para SWR entre 0 e 15 dB, os quatro sinais de teste não apresentam diferença significativa entre si, atingindo uma medida ODG máxima de cerca de $-3,80$. Para $\text{SWR} = 20$ dB, o ODG médio está em torno de $-3,60$. Estes valores encontram-se na faixa classificada como degradação muito incômoda, totalmente inaceitável em aplicações práticas de áudio. Para $\text{SWR} = 23$ dB, a medida ODG média está em torno de $-3,25$, também classificada como degradação muito incômoda. Para uma $\text{SWR} = 25$ dB, a medida ODG média aproxima-se de $-3,0$ e apenas o sinal *symph* atinge a faixa de degradação classificada como incômoda, o que também não é aceitável em aplicações práticas. Conclui-se que o uso do modelo perceptivo, em conjunto com um algoritmo de conformação espectral, é imprescindível em sistemas de marca d'água para áudio. Como mostrado nos capítulos anteriores, os algoritmos de equalização e identificação utilizando marca d'água apresentam resultados satisfatórios com SWR entre 20 e 23 dB, faixa em que a marca d'água utilizando o modelo psicoacústico atinge a medida de degradação considerada imperceptível. Essa característica evidencia a diferença entre a técnica proposta e os métodos de *superimposed training*, não focados na transparência do sinal de referência.

Para sinais que não possuem interpretação sensorial, tais como sinais digitais de telecomunicações, o limiar de SWR entre 20 e 23 dB pode ser interpretado, do ponto de vista do sinal de informação (hospedeiro), como um ruído de sistema. Na verdade, sistemas de comunicação típicos toleram níveis de ruído significativamente maiores do que este (Attux, 2005); nestes casos, a marca d'água pode ser considerada transparente. Ruídos adicionais introduzidos no sinal marcado, mesmo que mais potentes do que a marca d'água, em geral não prejudicam a detecção desta última, pois tais ruídos têm tipicamente potência muito inferior à do sinal de informação, este próprio encarado como ruído do ponto de vista da detecção da marca. Este é outro diferencial do método proposto em relação às técnicas de *superimposed training*: enquanto a marca d'água pode trabalhar abaixo do nível de ruído do sistema, no *superimposed training* assume-se que a sequência de treinamento tem energia superior à do ruído (Varma et al., 2004; Tugnait and Luo, 2003).

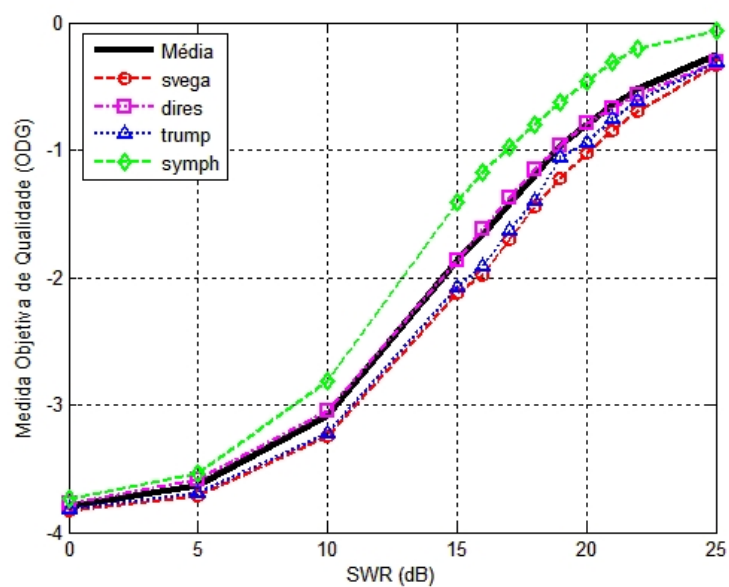


Fig. 7.9: Medida objetiva média de qualidade versus SWR usando modelo psicoacústico.

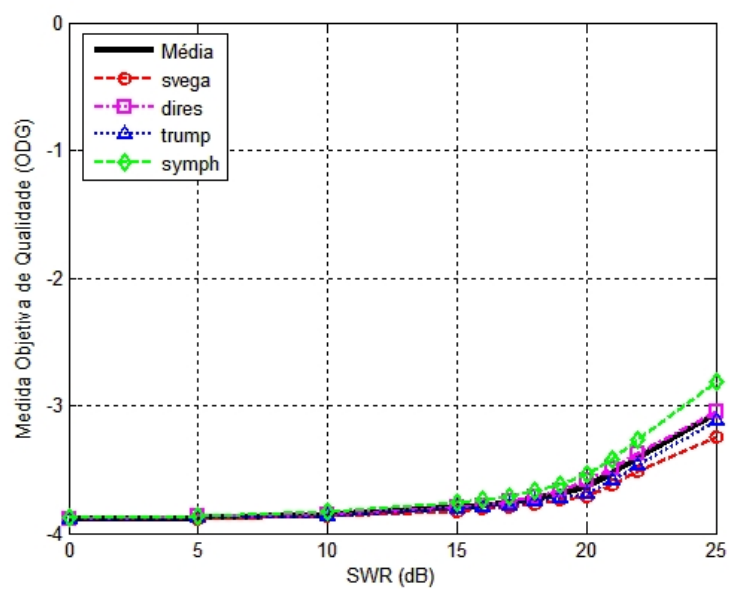


Fig. 7.10: Medida objetiva média de qualidade versus SWR para marca d'água branca.

7.4.2 Sistema de Identificação

Nesta seção, apresentamos resultados de simulação para a proposta de identificação de sinais de áudio utilizando marca d'água. Os sinais de áudio marcados foram filtrados por um canal com quatro coeficientes e fase não-mínima:

$$H(z) = 1 + 1,2z^{-1} - 0,3z^{-2} + 0,8z^{-3} \quad (7.2)$$

Utilizamos um modelo de identificação com o mesmo número de parâmetros do canal (capaz portanto de atingir a condição ideal de identificação), inicializados na origem. O material de áudio utilizado nos testes é o mesmo apresentado na tabela 7.1. O filtro modelo foi otimizado através do algoritmo RLS com fator de esquecimento $\lambda = 1$. O algoritmo de marca d'água utilizou um filtro de conformação espectral com 50 coeficientes.

<i>Sinal</i>	<i>Medida D (20 dB)</i>	<i>Medida D (23 dB)</i>
<i>dire</i> s	0,0586	0,0779
<i>sve</i> ga	0,0464	0,0656
<i>trump</i>	0,0373	0,0513
<i>symp</i> h	0,0428	0,0639

Tab. 7.3: Desvio dos coeficientes do filtro identificador com uso de modelo psicoacústico.

<i>Sinal</i>	<i>Valor D (20 dB)</i>	<i>Valor D (23 dB)</i>
<i>dire</i> s	0,0521	0,0703
<i>sve</i> ga	0,0394	0,0589
<i>trump</i>	0,0318	0,0446
<i>symp</i> h	0,0382	0,0576

Tab. 7.4: Desvio dos coeficientes do filtro identificador com marca d'água branca.

Para avaliar o desempenho do algoritmo de identificação utilizando marca d'água, a tabela 7.3 apresenta a medida de desvio D dos parâmetros do filtro (definida na equação 5.18) em relação ao caso ideal para valores de SWR de 20 e 23 dB. Nestes resultados, foram usados o modelo psicoacústico e o filtro de conformação espectral para minimização da distorção audível. Para uma mesma SWR, os valores obtidos para o desvio D foram muito próximos entre si para todos os sinais analisados. O desvio máximo admissível depende da aplicação em questão; no entanto, os valores obtidos encontram-se dentro de um patamar geralmente considerado satisfatório. Já a tabela 7.4 apresenta medidas similares para uma marca d'água branca, sem o uso do filtro de conformação espectral. Apesar de uma ligeira redução no desvio D, os resultados são praticamente equivalentes aos obtidos

com o uso do modelo psicoacústico. Portanto, a conformação espectral da marca d'água não causou prejuízo notável no desempenho do sistema de identificação.

7.4.3 Sistema de Equalização

Nesta seção, apresentamos resultados de simulação para a proposta de equalização de sinais de áudio utilizando marca d'água. Os sinais de áudio marcados foram filtrados pelo mesmo canal utilizado na seção anterior (equação 7.2). Foi empregado um equalizador com cinco coeficientes inicializados na origem. O material de áudio utilizado nos testes é o mesmo apresentado na tabela 7.1.

Como discutido na seção 4.1, o fato de os sinais de áudio serem correlacionados torna necessário o uso de um método de branqueamento junto ao equalizador. Empregou-se um filtro de erro de predição linear com 750 coeficientes e um atraso Δ de 2.200 amostras; este atraso foi determinado supondo uma correlação significativa dos sinais de áudio numa faixa de até 50 ms, a uma taxa de 44,1 kHz. O equalizador foi otimizado através do algoritmo RLS com fator de esquecimento $\lambda = 1$; já o filtro de erro de predição foi otimizado através do algoritmo LMS com passo $\mu = 0,0001$. Essa combinação apresentou 100% de convergência global.

<i>Sinal</i>	<i>EQM (20 dB)</i>	<i>EQM (23 dB)</i>
<i>dire</i> s	0,179	0,211
<i>sve</i> ga	0,176	0,206
<i>trump</i>	0,161	0,192
<i>symp</i> h	0,169	0,201

Tab. 7.5: Erro quadrático médio do sinal equalizado com uso de modelo psicoacústico.

<i>Sinal</i>	<i>EQM (20 dB)</i>	<i>EQM (23 dB)</i>
<i>dire</i> s	0,173	0,202
<i>sve</i> ga	0,168	0,197
<i>trump</i>	0,154	0,187
<i>symp</i> h	0,165	0,192

Tab. 7.6: Erro quadrático médio do sinal equalizado com marca d'água branca.

Para avaliar o desempenho do algoritmo de equalização utilizando marca d'água, a tabela 7.5 apresenta o EQM entre o sinal original e o sinal equalizado para valores de SWR de 20 e 23 dB. Nestes resultados, foram usados o modelo psicoacústico e o filtro de conformação espectral para minimização da distorção audível. A título de comparação, o equalizador ótimo de cinco coeficientes (calculados analiticamente) fornece os seguintes valores de EQM para os sinais em questão: 0,0823 para *dire*s, 0,0791 para *sve*ga, 0,0595 para *trump* e 0,0796 para *symp*h. Para SWR = 20 dB, o

EQM apresenta valores relativamente baixos e próximos dos valores ótimos (não causa degradação perceptível no áudio). Para $SWR = 23$ dB, o EQM apresenta um ligeiro aumento, porém continua relativamente próximo dos valores ótimos, encontrando-se em um patamar satisfatório para a maioria das aplicações. Já a tabela 7.6 apresenta medidas similares para uma marca d'água branca, sem o uso do filtro de conformação espectral. Apesar de uma ligeira redução no EQM, os resultados são praticamente equivalentes aos obtidos com o uso do modelo psicoacústico. Portanto, a conformação espectral da marca d'água não causou prejuízo notável no desempenho do sistema de equalização.

7.5 Conclusões

Neste capítulo, foi proposta uma extensão para sinais de áudio do método de equalização e identificação supervisionada apresentado nos capítulos anteriores. Para evitar que a marca d'água introduza distorções perceptíveis no sinal de áudio, o sistema emprega um modelo psicoacústico em conjunto com um algoritmo de conformação espectral.

Avaliações de qualidade de áudio mostraram que a marca d'água é imperceptível para SWR acima de um limiar localizado entre 20 e 23 dB. Além disso, as simulações indicaram que a presença do modelo perceptivo não afetou significativamente o desempenho dos sistemas de equalização e identificação.

Como perspectiva de trabalhos futuros, o método proposto pode ser adaptado a outras classes de sinais que apresentem interpretação sensorial, tais como imagens e vídeo. Para o caso de sinais de imagem, faz-se necessário o uso de modelos psicovisuais para maximizar a energia da marca d'água sem introduzir degradação perceptível, além da adaptação do método para trabalhar com sinais bidimensionais.

Capítulo 8

Conclusões e Perspectivas

Neste trabalho, propomos uma nova abordagem de filtragem adaptativa, aplicada aos problemas de equalização e identificação supervisionadas, utilizando uma marca d'água como sinal de referência. O método apresentado adiciona uma marca d'água ao sinal de informação (hospedeiro) e, sem interrupções na transmissão deste último, vale-se continuamente do sinal da marca d'água para obter um modelo matemático do canal, no caso de identificação, ou de seu inverso, no caso de equalização.

O método proposto apresenta importantes vantagens em relação às técnicas tradicionais de equalização e identificação. Dentre elas, podemos destacar a ausência de interrupções periódicas do sinal de informação para transmissão de seqüências de treinamento, ao mesmo tempo em que se evita a complexidade de critérios não-supervisionados. Além disso, a técnica proposta é passível de extensões às aplicações onde se faz uso de sinais de áudio.

No caso de aplicações em áudio, a adição da marca d'água não deve causar distorção perceptível. Neste sentido, propomos o uso de um modelo psicoacústico em conjunto com um algoritmo de conformação espectral, garantindo que a marca inserida no sinal tenha sua densidade espectral de potência limitada pelo limiar de mascaramento do áudio. Resultados experimentais ilustram a viabilidade do método quando aplicado a sinais desta natureza.

No capítulo 4, dedicado ao problema de equalização, foi demonstrado que é necessário compensar um fator de escala no equalizador caso o sinal de informação seja branco. Já no caso em que o sinal de informação é correlacionado, tal como um sinal de áudio, faz-se necessário o uso de um método de branqueamento em conjunto com o equalizador.

Para que o processo de detecção da marca d'água seja eficiente, é necessária a sincronização da marca no detector. Foram apresentados dois métodos de sincronização aplicáveis aos sistemas de equalização e identificação propostos. Os métodos são baseados no uso da própria marca d'água como seqüência de treinamento. Estes métodos revertem os efeitos de uma ampla classe de ataques de dessincronização, sendo capazes de monitorar o sincronismo continuamente ao longo de todo o sinal

recebido. Foram abordados os casos de distorções impostas pela transmissão do sinal marcado através de um sistema linear e de perda de trechos do sinal transmitido. Para determinação do sincronismo do sistema, foram propostos dois algoritmos, um baseado na média de janelas de análise do sinal recebido e outro empregando uma técnica de programação dinâmica.

As simulações realizadas mostraram, independentemente dos sistemas serem invariantes ou variantes no tempo, que o método de filtragem adaptativa supervisionada utilizando uma marca d'água como sinal de referência pode trazer ganhos de desempenho em relação às técnicas tradicionais, principalmente em situações nas quais a escassez de recursos sistêmicos inviabilize o uso de treinamentos periódicos. Para aplicações em áudio, avaliações de qualidade mostraram que o método garante a imperceptibilidade da marca d'água para uma SWR mínima entre 20 e 23 dB. Além disso, simulações mostraram que o uso do modelo perceptivo não introduz degradações significativas no desempenho dos sistemas de equalização e identificação.

Com relação à sincronização, os resultados mostraram que os métodos foram bem-sucedidos ao determinar o sincronismo inicial do sistema. O segundo método apresentado, baseado em programação dinâmica, foi também capaz de ressincronizar continuamente o sinal recebido, contrabalançando rapidamente perdas de sincronismo ao longo da comunicação.

Algumas perspectivas de continuidade deste trabalho incluem:

- análise teórica aprofundada do conjunto equalizador-filtro de erro de predição (utilizado com sinais hospedeiros correlacionados), além do estudo de outras formas de branqueamento do sinal hospedeiro;
- extensão ao caso de canais e filtros (equalizador e identificador) não-lineares;
- extensão ao caso de sistemas MIMO (*Multiple-Input Multiple-Output*);
- estudo do comportamento dos métodos de sincronismo frente a ataques mais severos, tais como filtragem não-linear, codificação/decodificação e reamostragem;
- estudo de outras técnicas de marca d'água, tais como eco oculto e QIM, em particular para áudio;
- análise do desempenho do sistema para outras classes de sinais, além de áudio, que apresentem interpretação sensorial (e.g. imagens, sinais multimídia), empregando-se os respectivos modelos perceptivos para maximizar a energia da marca d'água sem introduzir degradação perceptível.

Outra perspectiva promissora para estudos futuros são aplicações em áudio utilizando os métodos propostos, podendo-se citar:

- *Cancelamento de eco*: o fenômeno do eco em áudio pode ser modelado por meio de um filtro; um equalizador utilizando marca d'água poderia ser projetado para reverter os efeitos deste filtro.
- *Cancelamento dinâmico de reverberação*: poderia ser realizada uma pré-equalização e/ou um controle de potência das diversas fontes sonoras em, e.g. salas de teatro ou cinema, através dos métodos de filtragem adaptativa utilizando marca d'água, levando a uma redução dos efeitos da reverberação.
- *Melhoria do desempenho de algoritmos de compressão de áudio*: codificadores de áudio baseados em modelos perceptivos introduzem distorções no sinal após o processo de codificação/decodificação. O acúmulo de várias distorções desse tipo pode resultar em uma perda de qualidade perceptível. Poderia ser inserida uma marca d'água no áudio original, utilizada ao final de cada etapa de codificação/decodificação com o intuito de minimizar as distorções inseridas pelo codec (Packham and Kurth, 2000). Este procedimento pode ser estendido a codificadores de imagem e vídeo, que também podem se valer de modelos perceptivos.

Referências Bibliográficas

- J. Arenas-García, V. Gómez-Verdejo, Martínez-Ramón, and A. R. Figueiras-Vidal. Separate-variable adaptive combination of lms adaptive filters for plant identification. *IEEE XIII Workshop on Neural Networks for Signal Processing*, 1:239–248, 2003.
- M. Arnold. Subjective and objective quality evaluation of watermarked audio tracks. *Proceedings of the Second International Conference on WEB Delivering of Music (WEDELMUSIC.02)*, 2002.
- R. R. F. Attux. *Novos paradigmas para equalização e identificação de canais baseados em estruturas não-lineares e algoritmos avaliativos*. PhD thesis, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, Brasil, 2005.
- R. R. F. Attux. Sobre dinâmica caótica e convergência em algoritmos para equalização autodidata. Master's thesis, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, Brasil, 2001.
- J. G. A. Barbedo. *Avaliação Objetiva de Qualidade de Sinais de Áudio e Voz*. PhD thesis, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, Brasil, 2004.
- A. Baros, F. Franco, D. Delannay, and B. Macq. Rate-distortion analysis of steganography for conveying stereovision disparity maps. *Proc. SPIE*, 5306:268–273, january 2004.
- P. Bassia, I. Pitas, and N. Nikolaidis. Robust audio watermarking in the time domain. *IEEE Transactions on Multimedia*, 3:232–241, 2001.
- W. Bender, D. Gruhl, N. Morimoto, and A. Lu. Techniques for data hiding. *IBM Systems Journal*, 35:313–336, 1996.
- Dimitri P. Bertsekas. *Dynamic Programming and optimal Control*. Athena Scientific, second edition, 2000.
- R. Bittencourt. Mpeg-audio. <http://www.lsi.usp.br/ricardo/mpeg/>, 1997.

- D. Boneh and J. Shaw. Collusion-secure fingerprinting for digital data. *IEEE Trans. Inf. Theory*, 44(5):1897–1905, 1998.
- L. Boney, A. Tewfik, and K. Hamdy. Digital watermarks for audio signals. *IEEE Int. Conf. on Multimedia Computing Systems*, June 1996.
- L. Boney, A. Tewfik, and K. Hamdy. Digital watermarks for audio signals. *Eurospeech*, September 1997.
- T. G. S. S. Brandão. Assinatura digital de imagens baseada em espalhamento de espectro. Master's thesis, Universidade Técnica de Lisboa, Instituto Superior Técnico, maio 2002.
- ITU-R BS.1116-1. *Methods for the subjective assessment of small impairments in audio systems including*. 1996.
- ITU-R BS.1387-1. *Method for Objective Measurements of Perceived Audio Quality*. 1998.
- ITU-R BS.562-3. *Subjective Assessment of Sound Quality*. 1990.
- B. Chen and C.E. W. Sundberg. Digital audio broadcasting in the fm band by means of contiguous band insertion and precancelling techniques. *IEEE Trans. Commun.*, 48(10):1634–1637, october 2000.
- B. Chen and G. W. Wornell. Quantization index modulation methods: A class of provably good methods for digital watermarking and information embedding. *IEEE Trans. Inf. Theory*, 47(4):1423–1443, May 2001.
- B. Chen and G. W. Wornell. An information-theoretic approach to the design of robust digital watermarking systems. *Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 1999.
- N. Chen and G. T. Zhou. What is the price paid for superimposed training in ofdm? In *Proc. 2004 IEEE International Conf. on Acoustics, Speech, Signal Processing*, volume 4, pages 421–424, Montreal, Canada, May 2004.
- M. Cohn and A. Lempel. On fast m-sequence transforms. *IEEE Transaction Information Theory*, IT-23:135–137, January 1977.
- I. J. Cox, J. Killian, F. T. Leighton, and T. Shamon. Secure spread spectrum watermarking for multimedia. *IEEE Transaction Image Processing*, 6(12):1673–1687, December 1997.
- I. J. Cox, M. L. Miller, and J. A. Bloom. *Digital Watermarking*. CA: Morgan-Kaufmann, San Francisco, 2002.

- N. Cvejic, A. Keskinarkaus, and T. Seppänen. Audio watermarking using m-sequences and temporal masking. *IEEE Workshops on Applications of Signal Processing to Audio and Acoustics*, pages 227–230, New Paltz, New York 2001.
- D. B. DeFigueiredo. Método de inserção de marcas d'água robustas em sinais digitais. Master's thesis, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, Brasil, 2001.
- Z. Ding and Y. Li. *Blind Equalization and Identification*. Marcel Dekker, 2001.
- H. Fletcher. *Speech and Hearing in Communication*. D. Van Nostrand Co., Toronto, 1953.
- G. D. Forney. The Viterbi algorithm. In *Proc. of the IEEE*, volume 61, pages 268–278, March 1973.
- A. J. Fourcin. Speech processing by man and machine. *Group Report*, 1:307–351, 1977.
- R.A. Garcia. Digital watermarking of audio signals using a psychoacoustic auditory model and spread spectrum theory. *107th AES Convention*, september 1999.
- D. N. Godard. Self-recovering equalization and carrier tracking in two-dimensional data communication systems. *IEEE Transactions on Communications*, 28(11):1867–1875, 1980.
- S. Golomb. *Shift Register Sequences*. Holden-Day, San Francisco, 1967.
- L. C. T. Gomes. *Tatouage de signaux audio*. PhD thesis, Université René Descartes (Paris V), Paris, France, 2002.
- L. C. T. Gomes, M. Mboup, M. Bonnet, and N. Moreau. Cyclostationarity-based audio watermarking with private and public hidden data. *108th AES Convention*, 2000.
- L. C. T. Gomes, E. Gómez, and N. Moreau. Resynchronization methods for audio watermarking. *109th AES Convention*, 2001.
- L. C. T. Gomes, P. Cano, E. Gómez, M. Bonnet, and E. Batlle. Audio watermarking and fingerprinting: For which applications? *Journal of New Music Research*, 32:65–81, 2003.
- E. Gómez. Tatouage de signaux de musique (méthodes de synchronisation). *technical report (DEA ATIAM)*, Université de la Méditerranée / ENST, July 2000.
- M. P. Guimarães. *Optimisation de l'allocation des ressources binaires et modélisation psychoacoustique pour le codage audio*. PhD thesis, Université Paris V, 1998.

- F. Hartung and B. Girod. Watermarking of uncompressed and compressed video. *Signal Processing*, 66:283–301, 1998.
- F. Hartung and M. Kutter. Multimedia watermarking techniques. *Proc. IEEE, Special Issue on Identification and Protection of Multimedia Information*, 87(7):1079–1107, julho 1999.
- Simon Haykin. *Adaptive Filter Theory*. Prentice Hall, third edition, 1996.
- Simon Haykin. *Adaptive Filter Theory*. Prentice Hall, fourth edition, 2001.
- L. E. Humes and W. Jesteadt. Models of the additivity of masking. *Journal of the Acoustical Society of America*, 85(3):1285–1294, March 1989.
- ISO/IEC-11172-3. *International Organization for Standardization. ISO/IEC 11172-3 (Information technology – Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s)*, 1993.
- ISO/IEC-13818-7. *International Organization for Standardization. ISO/IEC 13818-7 (MPEG2 Advanced Audio Coding, AAC)*, 1997.
- C. Junqueira, F. O. de França, R. R. F. Attux, R. Suyama, L. N. de Castro F. J. V. Zuben, and J. M. T. Romano. A proposal for blind fir equalization of time-varying channels. *IEEE*, 2005.
- H. Kim. Stochastic model based audio watermark and whitening filter for improved detection. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 4:1971–1974, 2000.
- D. Kundur and D. Hatzinakos. Digital watermarking for telltale tamper proofing and authentication. *Proc. of the IEEE*, 87(7):1167–1180, july 1999.
- E. A. Lee and D. G. Messerschmitt. *Digital communications*. Kluwer Academic Publishers, 2a edition, 1994.
- S. K. Lee and Y. S. Ho. Digital audio watermarking in the cepstrum domain. *IEEE Transactions on Consumer Electronics*, 46(3):744–750, 2000.
- R. Lucky. Automatic equalization for digital communication. *Bell Systems Technical Journal*, 44: 547–588, 1965.
- O. Macchi and E. Eweda. Convergence analysis of self-adaptive equalizers. *IEEE Transactions on Information Theory*, IT-30(2):161–176, 1984.

- M. F. Mansour and A. H. Tewfik. Audio watermarking by time-scale modification. *International Conference on Acoustics, Speech, and Signal Processing*, 3:1353–1356, 2001.
- J. E. Mazo. Analysis of decision-directed equalizer convergence. *The Bell System Technical Journal*, 59(10):1857–1876, 1980.
- B. C. J. Moore. *An Introduction to the Psychology of Hearing*. Academic Press, New York, 1989.
- N. Moreau, P. Dymarski, and L. C.T. Gomes. Tatouage audio : une réponse à une attaque désynchronisante. *CORESA, Poitiers*, October 2000.
- P. Moulinand and R. Koetter. Data-hiding codes. In *Proceedings of the IEEE*, volume 93, No. 12, pages 2083–2126, December 2005.
- C. Neubauer and J. Herre. Digital watermarking and its influence on audio quality. *105th Convention of the Audio Engineering Society (AES'05)*, 46:1041, 1998.
- A. O. Neves. Uma abordagem unificada para algoritmos de equalização autodidata. Master's thesis, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, Brasil, 2001.
- Opticom. <http://www.opticom.de/>, 2007.
- ITU-T P.830. *Subjective performance assessment of telephone-band and wideband digital codecs*. Fevereiro 1996.
- ITU-T P.85. *A method for subjective performance assessment of the quality of speech voice output devices*. Junho 1994.
- ITU-T P.861. *Objective quality measurement of telephone-band (300-3400 Hz) speech codecs*. Fevereiro 1998.
- ITU-T P.862. *Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrow band telephone networks and speech codecs*, Fevereiro 2001.
- N. Packham and F. Kurth. Transport of context-based information in digital audio data. *AES 109th Convention*, pages 22–25, September 2000.
- Cristiano Magalhães Panazio. Utilização conjunta de equalização adaptativa e códigos corretores de erro em processamento espacial e temporal. Master's thesis, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, Brasil, 2001.

- K. K. Parhi and T. Nishitani. *Digital Signal Processing in Multimedia Systems*. Marcell Dekker Inc., 1999.
- F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn. Attacks on copyright marking system. *Information Hiding, Lecture Notes in Computer Science*, 1525:218–238, 1998.
- R. Petrovic. Audio signal watermarking based on replica modulation. *International Conference on Telecommunications in Modern Satellite, Cable, and Broadcasting Service*, 1:227–234, 2001.
- R. Petrovic, J. M. Winograd, K. Jemili, and E. Metois. Data hiding within audio signals. *International Conference on Telecommunications in Modern Satellite, Cable, and Broadcasting Service*, 1:88–95, 1999.
- H. Robbins and S. Monro. A stochastic approximation method. *The Annals of Mathematical Statistics*, 22:400–407, 1951.
- J. Seok, J. Hong, and J. Kim. A novel audio watermarking algorithm for copyright protection of digital audio. *ETRI Journal*, 24:181–189, 2002.
- O. Shalvi and E. Weinstein. New criteria for blind deconvolution of nonminimum phase systems (channels). *IEEE Transactions on Information Theory*, 36(2):312–321, 1990.
- G. A. Soulodre, T. Grusec, M. Lavoie, and L. Thibault. Subjective evaluation of state-of-the-art 2-channel audio codecs. *AES 104th Convention*, 1998.
- R. Suyama. Sobre o critério do módulo constante para equalização não-supervisionada e suas relações com a teoria de wiener. Master's thesis, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, Brasil, 2003.
- M. Swanson, B. Zhu, A. Tewfik, and L. Boney. Robust audio watermarking using perceptual masking. *Signal Processing*, 66:337–355, 1998.
- K. Tanaka, Y. Nakamura, and K. Matsui. Embedding secret information into a dithered multilevel image. *Proc. of the IEEE Military Communications Conference*, september 1990a.
- K. Tanaka, Y. Nakamura, and K. Matsui. Embedding the attribute information into a dithered image. *Systems and Computers in Japan*, 21(7), 1990b.
- E. Terhardt. Calculating virtual pitch. *Hearing Research*, 1:155–182, 1979.
- T. V. Thiede. *Perceptual Audio Quality Assessment Using a Non-Linear Filter Bank*. PhD thesis, Berlin, 1999.

- A. Tirkel, G. Rankin, R. van Schyndel, W. Ho, N. Mee, and C. Osborne. Electric watermark. *Proc. DICTA*, pages 666–672, december 1993.
- J. Treichler and B. Agee. A new approach to multipath correction of constant modulus signals. *IEEE Transactions on Acoustics, Speech and Signal*, 31(2):459–472, 1983.
- J. K. Tugnait and W. Luo. On channel estimation using superimposed training and first-order statistics. *IEEE Communications Letters*, 7(9), September 2003.
- J. K. Tugnait and X. Meng. Synchronization of superimposed training for channel estimation. In *Proc. 2004 IEEE International Conf. on Acoustics, Speech, Signal Processing*, volume 4, pages 853–856, Montreal, Canada, May 2004.
- A. R. Varma, L. L. H. Andrew, C. R. N. Athaudage, and J. H. Manton. Iterative algorithms for channel identification using superimposed pilots. *Australian Communications Theory Workshop*, February 2004.
- C. P. Wu, P. C. Su, and C. C. J. Kuo. Robust and efficient digital audio watermarking using audio content analysis. *Security and Watermarking of Multimedia Contents*, 3971:382–392, 2000.
- E. Zwicker and H. Fastl. *Psychoacoustics, Facts and Models*. Springer Verlag, Berlin, 1990.
- E. Zwicker and R. Feldkeller. *Das Ohr als Nachrichtenempfänger*. Hirzel Verlag, Stuttgart, 1967.

Apêndice A

Fenômenos Auditivos e Modelos Perceptivos

A.1 Fenômenos Auditivos

A seguir, serão apresentados alguns dos principais conceitos provenientes do estudo do comportamento auditivo humano (Guimarães, 1998; Barbedo, 2004). Alguns dos trechos aqui apresentados foram extraídos textualmente de (Barbedo, 2004).

A.1.1 Fisiologia do Ouvido Humano

Apresentamos aqui uma breve descrição dos elementos fisiológicos envolvidos no processo da audição humana. Algumas dessas estruturas podem ser visualizadas na figura A.1, que ilustra o corte longitudinal do ouvido humano. Tradicionalmente, este é dividido em três áreas: o ouvido externo, o ouvido médio e o ouvido interno.

Ouvido Externo

O ouvido externo e a cabeça são componentes de um complexo sistema de recepção acústica, que faz a ligação entre o tímpano e o campo sonoro externo. As duas principais estruturas do ouvido externo são o pavilhão auricular e o canal auditivo externo. O pavilhão, embora seja a parte mais exposta do sistema auditivo, é provavelmente o que desempenha papel menos significativo para a audição. Porém, o ouvido externo protege o tímpano de danos mecânicos e melhora o acoplamento entre este e o campo sonoro, além de contribuir substancialmente para a direcionalidade do sistema, especialmente para altas frequências. A estrutura e localização do pavilhão sugere sua função como refletor de som.

O pavilhão auricular pode ajudar, em certa medida, na localização dos sons, em especial a discriminação entre sons à frente e atrás do ouvinte. As propriedades acústicas do pavilhão auricular

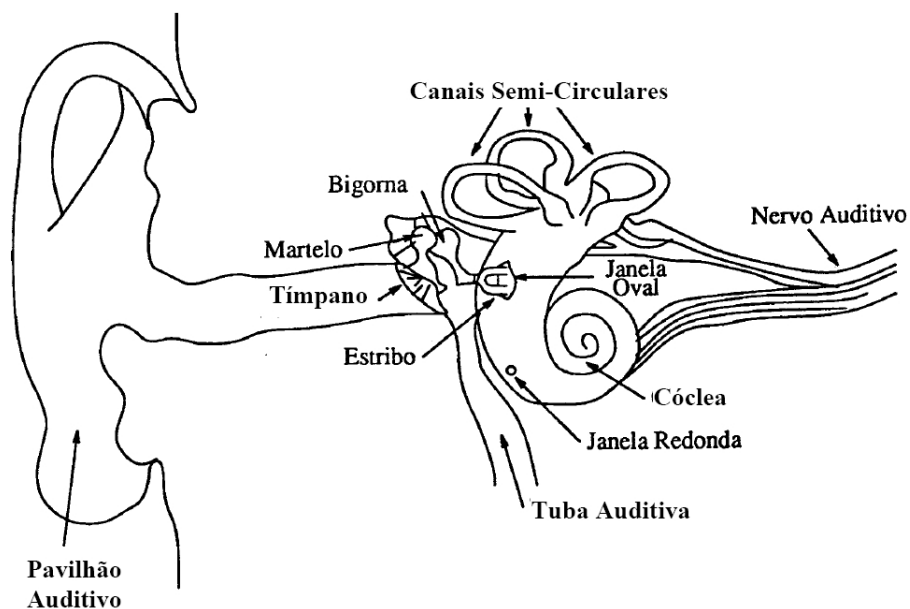


Fig. A.1: Corte longitudinal do ouvido, com suas principais estruturas.

podem afetar a propagação do som no espaço e, conseqüentemente, afetar o sinal acústico que chega ao conduto auditivo, agindo como atenuador ou amplificador de sons de determinada frequência.

O canal auditivo externo consiste de uma cavidade cilíndrica, com aproximadamente três centímetros de comprimento e um diâmetro médio de sete milímetros. A cavidade é aberta do lado externo e fechada do lado interno pelo tímpano. O canal auditivo, através do qual penetram as ondas de pressão, contribui para a sensibilidade máxima a tons entre 2.000 a 5.000 Hz, provocando, juntamente com o pavilhão auricular, um ganho de 10 a 15 dB nessa faixa de frequência.

A membrana timpânica, que fecha a parte média do canal, serve como estrutura separadora entre esta região e o ouvido médio.

Ouvido Médio

As ondas de pressão no canal auditivo provocam movimentos da membrana do tímpano e, com isso, as estruturas do ouvido médio entram em movimento também. A principal função do ouvido médio é melhorar a transmissão sonora entre o ouvido externo e o ouvido interno.

O ouvido médio é, na sua maior parte, uma cavidade cheia de ar, ligada à parte superior da garganta por um canal (tuba auditiva). Esta área também inclui uma outra cavidade menor cheia de ar (antro) e muitos espaços de ar, compreendendo a região mastóide do osso temporal. Uma cadeia de ossos (os ossículos), se estende da membrana até a uma abertura (a janela oval) na região do rochedo do osso temporal. Dois músculos (o tensor do tímpano e o estapédio) são ligados a essa cadeia. Posteriormente, há uma outra abertura (a janela redonda) coberta por uma membrana. O

ouvido interno situa-se aproximadamente entre as janelas oval e redonda, dentro do rochedo do osso temporal.

Os ossículos (martelo, bigorna e estribo) fazem a transmissão do som que é recebido pelo tímpano diretamente à janela oval, além de protegerem a janela redonda do som, que chega a ela com menor intensidade. Se isto não ocorresse, o som chegaria ao mesmo tempo às duas janelas com a mesma magnitude e na mesma fase sonora, pois a distância entre elas é muito pequena, e isto não promoveria a movimentação do fluido do ouvido interno, tornando a audição impossível.

A cavidade do ouvido médio está ligada com a cavidade da boca através da tuba auditiva. Esse arranjo permite igualar a pressão de ambos os lados da membrana do tímpano. Embora as pressões constantes, ou as variações muito lentas de pressão, não sejam eficazes como estímulos auditivos, as diferenças constantes de pressão entre os ouvidos médios e externo alteram a nossa sensibilidade a frequências diferentes.

O processo de transformação do sinal acústico nas ondas do líquido coclear é chamado de função de transferência do ouvido médio. Ele é equivalente a uma filtragem passa baixas com frequência de corte em 5 kHz, com uma sobre-elevação na faixa entre 2.000 e 5.000 Hz e um pico em torno de 3.500 Hz (Bittencourt, 1997).

Ouvido Interno

Os ossos do ouvido médio estimulam a cóclea através da janela oval, fazendo com que seu líquido interno se movimente. A cóclea pode ser modelada como um tubo de aproximadamente 30 mm com duas câmaras separadas por uma estrutura chamada membrana basilar. Na extremidade oposta à janela oval existe um orifício sobre a membrana basilar que comunica essas duas câmaras, chamado de helicotrema. A membrana basilar apresenta uma resistência (mecânica) que varia ao longo de sua extensão: próximo à janela oval ela é mais fina e tensa, ressoando em frequências mais altas, enquanto no seu final (ápice), ela é espessa e flácida, ressoando então para frequências mais baixas. As ondas geradas pelo estribo, em resposta a um sinal senoidal, viajam ao longo da cóclea, fazendo vibrar a membrana basilar na mesma frequência do sinal de entrada (Fletcher, 1953).

Cada ponto da membrana basilar é mais sensível a uma determinada frequência, chamada de frequência característica. Para um ponto específico da membrana basilar, a curva de resposta à frequência de vibração presente na janela oval é equivalente à de um filtro passa-faixa com fator de qualidade aproximadamente constante, resultando numa melhor resolução nas baixas frequências.

Um comportamento similar é obtido ao se traçar a curva de resposta ao longo da membrana basilar para um tom numa frequência específica. Para cada frequência, há um ponto da membrana basilar em que a vibração é máxima. A posição desse ponto é aproximadamente proporcional ao logaritmo da frequência do som. Ao redor desse ponto haverá uma faixa, de cerca de 1,5 mm, onde a vibração

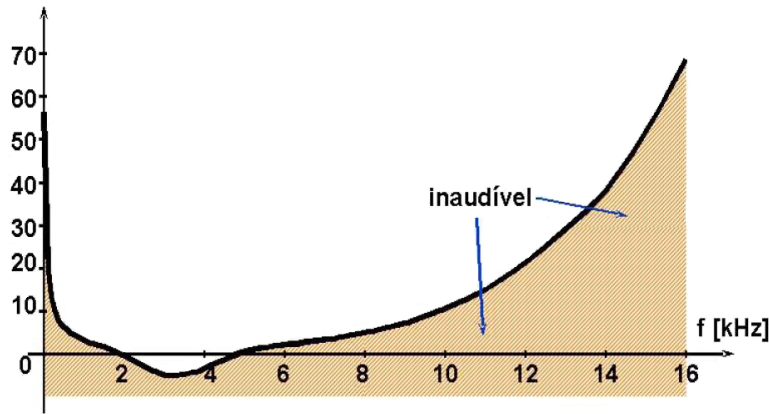


Fig. A.2: Limiar absoluto de audibilidade em silêncio.

estará presente, atenuando-se conforme se afasta do ponto. Tal faixa determina o conceito das bandas críticas, como será visto mais adiante.

A.1.2 Limiar Absoluto de Audibilidade em Silêncio

O limiar absoluto de audibilidade em silêncio é o menor nível, em função da frequência, para o qual um tom se torna audível (Terhardt, 1979). Este limiar, representado na figura A.2, pode ser aproximado pela expressão analítica dada por

$$\text{lim} = 3,64f^{-0,8} - 6,5e^{-0,6(f-3,3)^2} + 10^{-3}f^4 \quad (\text{A.1})$$

onde f é a frequência em kHz. O limiar é dado em dB_{SPL} ¹. Esta aproximação é usada em quase todos os métodos perceptivos. Consiste de três termos: o primeiro descreve a frequência de corte para as baixas frequências; o segundo descreve o aumento de sensibilidade do ouvido para a faixa de frequências em torno de 3 kHz; o último descreve a frequência de corte para as altas frequências.

A.1.3 Bandas Críticas

Alguns dos fenômenos de mascaramento podem ser explicados em termos de faixas de frequências conhecidas como *bandas críticas*, as quais foram determinadas através de experimentos psicoacústicos (Zwicker and Feldkeller, 1967). Uma banda crítica define uma faixa em torno de uma frequência central, a qual está associada a um ponto da membrana basilar, de modo que a cada ponto é possível definir uma banda crítica. Quando dois sinais se situam dentro de uma banda crítica, o

¹Nível de pressão sonora (*Sound Pressure Level*): tem como unidade o decibel SPL (dB_{SPL}), e é dado pela expressão $20\log_{10}(P/P_{\text{ref}})$, onde P é a pressão sonora do sinal que se está medindo e P_{ref} é a pressão sonora de referência.

de maior energia poderá dominar a percepção e mascarar o outro estímulo sonoro. Portanto, dependendo dos níveis, dois tons distintos só serão distinguidos um do outro quando estiverem em bandas críticas diferentes. A resolução para a distinção entre uma frequência e outra varia de 100 Hz, nas frequências mais baixas, a mais de 6.000 Hz, nas frequências mais altas. Além disso, sinais com uma largura de banda suficiente para extrapolar os limites de uma banda crítica sempre proporcionarão uma intensidade perceptiva maior que aqueles cujas componentes espectrais estejam limitadas a uma única banda crítica, ainda que o nível de pressão sonora e a frequência central sejam equivalentes.

Embora exista uma banda crítica ao redor de cada frequência, convencionou-se (com algumas pequenas variações) a adoção dos valores mostrados na tabela A.1 (Fourcin, 1977). Os valores referentes às bandas críticas, apresentados na primeira coluna da tabela, correspondem à escala Bark².

A.1.4 Mascaramento

As limitações do ouvido em termos das resoluções temporal, espectral e de amplitude, em combinação com uma faixa dinâmica também limitada, resultam no chamado fenômeno de mascaramento. Quando dois tons estão suficientemente próximos um do outro, seja no domínio do tempo ou da frequência, o tom mais fraco pode se tornar inaudível devido à presença do tom mais forte. Embora o fenômeno do mascaramento deva ser analisado no plano tempo-frequência, é muito usual considerá-lo como dois efeitos separados, dependendo do domínio que se está considerando. Quando o mascaramento depende unicamente da localização no domínio da frequência, isto é, os sinais mascarado e mascarador são apresentados no mesmo instante de tempo, tem-se o assim denominado mascaramento simultâneo. Se o mascaramento depende primariamente da localização no domínio do tempo, então ele é chamado de mascaramento temporal. Este último pode ser dividido em dois diferentes efeitos: mascaramento progressivo (ou pós-mascaramento) e mascaramento retrógrado (ou pré-mascaramento). No caso do mascaramento progressivo, os componentes do sinal são mascarados após o término do mascarador, e no caso do mascaramento retrógrado, os componentes são mascarados antes do início da execução do mascarador.

O nível de energia abaixo do qual um componente do sinal é mascarado por outros componentes é chamado de limiar de mascaramento. Além de depender da localização dos sinais mascarador e mascarado no plano tempo-frequência, o limiar de mascaramento progressivo também depende da duração do mascarador (Thiede, 1999).

O mascaramento simultâneo se deve basicamente à existência das bandas críticas. Quando dois tons se encontram em uma mesma banda crítica, o de maior amplitude dominará a percepção sonora. Já o mascaramento progressivo ocorre devido ao fato de o ouvido humano necessitar de um certo tempo para se recuperar após um sinal de grande amplitude, uma vez que, nessas condições,

²Uma distância de 1 Bark corresponde à largura de uma banda crítica.

Banda Crítica	Frequência (Hz)		
	Inferior	Superior	Faixa
0	0	100	100
1	100	200	100
2	200	300	100
3	300	400	100
4	400	510	110
5	510	630	120
6	630	770	140
7	770	920	150
8	920	1080	160
9	1080	1270	190
10	1270	1480	210
11	1480	1720	240
12	1720	2000	280
13	2000	2320	320
14	2320	2700	380
15	2700	3150	450
16	3150	3700	550
17	3700	4400	700
18	4400	5300	900
19	5300	6400	1100
20	6400	7700	1300
21	7700	9500	1800
22	9500	12000	2500
23	12000	15500	3500
24	15500	22050	6550

Tab. A.1: Bandas Críticas

os neurônios disparados ficam em um estado refratário que pode durar mais de 100 ms. Enquanto o mascaramento simultâneo (frequencial) e o mascaramento progressivo (temporal) são de fácil entendimento, o mascaramento retrógrado (também temporal) é um fenômeno mais complicado, porque implica em um sinal de grande amplitude mascarar outro sinal antes de o primeiro estar realmente presente. Tal fenômeno é normalmente explicado pela suposição de que o sinal forte é processado mais rapidamente do que o sinal fraco, podendo, portanto, ultrapassar o sinal mascarado durante o processamento dos sinais, ou no nervo auditivo, ou posteriormente, nos níveis mais elevados do sistema auditivo (Thiede, 1999).

Dentre as três categorias de mascaramento, o simultâneo (frequencial) tem sido mais freqüente e detalhadamente analisado. A medição dos mascaramentos temporais é mais difícil do que a medição do mascaramento simultâneo, uma vez que estes são fortemente dependentes dos valores das freqüências dos sinais. Além disso, a determinação do mascaramento temporal requer uma boa resolução tanto no domínio do tempo quanto da freqüência. Porém, isto só é possível até certo ponto. No caso do mascaramento progressivo, tal limitação não é grave porque as constantes de tempo observadas são suficientemente grandes (cerca de 100 ms) para permitir sinais de teste com um espectro suficientemente compacto, sem introduzir demasiada incerteza na estrutura temporal. No caso do mascaramento retrógrado, as constantes de tempo observadas são tão pequenas (entre 1 e 20 ms) que este não pode ser medido de uma maneira confiável para sinais de banda estreita.

Para sons compostos (incluindo componentes tonais e não tonais), os efeitos de mascaramento das várias componentes são cumulativos.

A seguir, serão apresentadas algumas das características inerentes aos diversos tipos de mascaramento.

Limiares de Mascaramento de Tons, Ruídos e Pulsos

Na análise desta seção, consideramos duas classes de sinais: tonais, de maior intensidade e aproximadamente periódicos, e não tonais, de baixa intensidade e com aspecto de ruído. Exemplos dessas classes de sinais encontram-se na figura A.3.

Limiares de mascaramento são normalmente medidos como uma função do tempo ou da freqüência central do sinal mascarado, enquanto o nível e a freqüência central do sinal mascarador são mantidos constantes e são tomados como parâmetros. Tais medidas resultam em curvas de mascaramento como aquelas mostradas na figura A.4, onde o mascaramento é considerado em ambos os domínios, ou como aquelas mostradas na figura A.5, onde apenas o mascaramento simultâneo é analisado.

Na figura A.4, o paralelepípedo em destaque representa um ruído de faixa estreita, o qual faz o papel do mascarador. Tal ruído tem uma duração de 500 ms e uma largura de faixa de cerca de 1 Bark. A curva acima do ruído descreve a função de mascaramento relacionada a esse ruído, representando

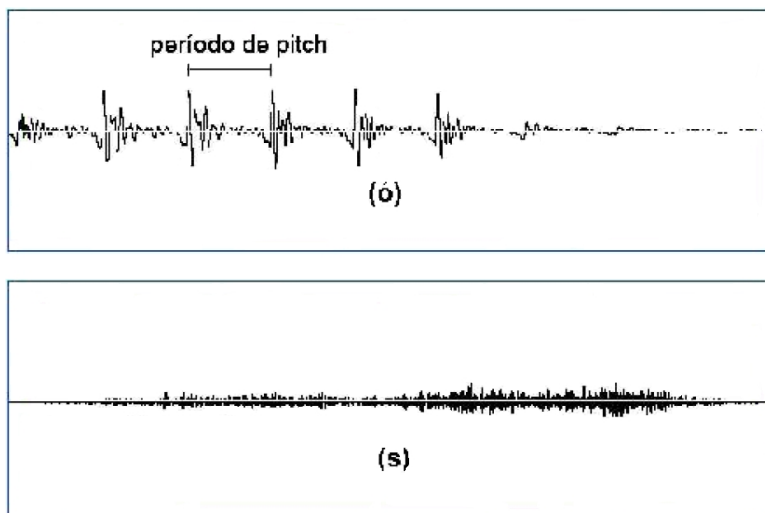


Fig. A.3: Exemplos de sinais tonal e não tonal; ambos os exemplos correspondem a sinais de voz (fones “ó” e “s”, respectivamente).

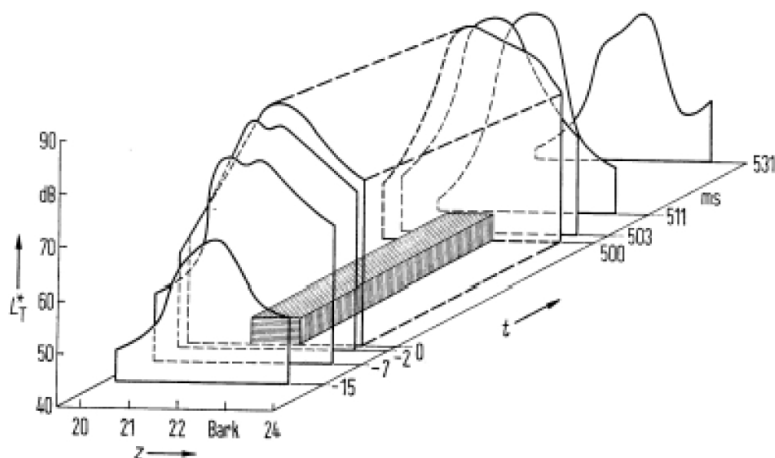


Fig. A.4: Padrão de mascaramento causado por ruído de banda estreita.

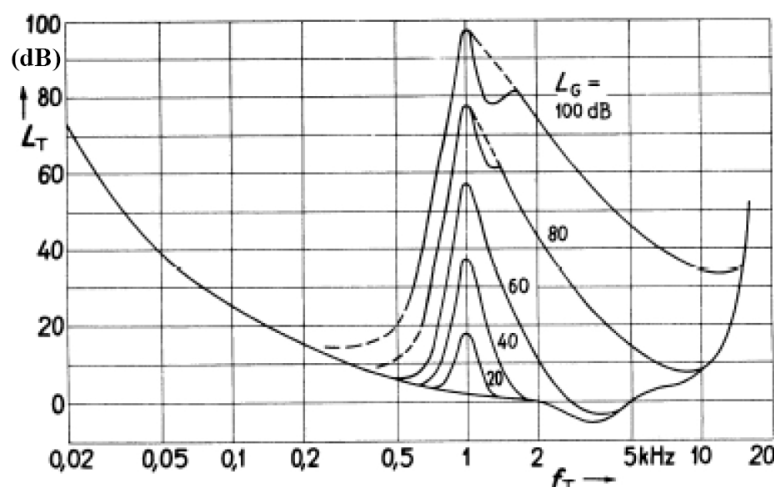


Fig. A.5: Representação do mascaramento simultâneo causado por tons a 1 kHz.

o nível necessário para que um sinal tonal se torne audível na presença do ruído mascarador. O eixo L_T apresenta este nível de forma relativa ao nível do mascarador em uma escala em dB. O eixo f_T permite observar a variação deste nível com o tempo. Aqui, observam-se os fenômenos do mascaramento retrógrado, de menor duração, e do mascaramento progressivo, mais longo. No eixo f_T , representando o domínio espectral perceptivo, percebe-se o comportamento assimétrico da curva de mascaramento, já que a inclinação para as baixas frequências é mais íngreme que a inclinação para as altas frequências. É importante ter em conta que o formato da curva para as frequências superiores é fortemente dependente do nível dos sinais. Para baixos níveis, ela é quase tão inclinada quanto nas baixas frequências, enquanto se torna quase plana para níveis muito elevados do mascarador. Essas curvas de mascaramento podem ser aproximadas por exponenciais de dois lados quando representadas em função de uma escala de frequência perceptiva, como no caso da figura A.4.

A figura A.5 descreve o comportamento do limiar de audibilidade de um sinal na ausência de um mascarador (curva inferior) e na presença de tons mascaradores a 1 kHz e diferentes níveis.

A diferença de nível entre um sinal e o máximo mascaramento por ele produzido (maior nível que outro sinal pode assumir e continuar sendo mascarado) é chamado índice de mascaramento (Moore, 1989). Ele depende da frequência central do sinal mascarador, mas assume-se que seja independente do seu nível. A representação linear do índice de mascaramento em função da frequência é chamada fator de limiar.

A.2 Modelos Perceptivos

Diversos modelos acústicos foram criados a partir dos fenômenos descritos. Esses modelos simulam certas propriedades da audição humana. Como a sonoridade³ é a principal propriedade de um som, modelos da sonoridade percebida têm sido de particular interesse na pesquisa psicoacústica. Muitos métodos perceptivos usam idéias originárias de tais modelos, especialmente do modelo de Zwicker para a computação da sonoridade percebida. Outro modelo da sonoridade percebida foi introduzido por Moore, Glasberg e Baer (Barbedo, 2004). Ambos os modelos serão brevemente descritos a seguir. Também serão brevemente descritos os modelos psicoacústicos 1 e 2 definidos na norma MPEG-2, utilizados em codificadores de áudio baseados em modelos psicoacústicos (Guimarães, 1998). Alguns dos trechos aqui apresentados foram extraídos de (Barbedo, 2004).

A.2.1 Modelo de Zwicker para o Cálculo da Sonoridade Percebida

O modelo de Zwicker (Zwicker and Fastl, 1990) já inclui a maior parte dos passos de processamento que são usados nas medidas perceptivas. Uma versão simplificada deste modelo, que modela a resposta em frequência do ouvido usando filtros, tem sido amplamente usada para a estimação da sonoridade no campo da prevenção de ruído, tornando-se parte de um padrão internacional. No primeiro passo, o sinal de entrada é transformado para o domínio da frequência e agrupado em bandas críticas.

Todos os dados utilizados na elaboração deste modelo foram coletados através de experimentos fisiológicos, os quais consistiram basicamente na observação da reação da membrana basilar a determinadas excitações e na contagem dos neurônios envolvidos em cada situação.

As excitações designadas para as bandas críticas adjacentes são determinadas pela forma das curvas de mascaramento; onde o mascaramento é mais intenso, mais neurônios são envolvidos. Tal procedimento resulta em diversas excitações para cada banda crítica (uma originada pela energia da componente presente naquela banda crítica e outras provenientes das energias de componentes presentes nas bandas críticas adjacentes). A excitação pode ser determinada pelo maior valor entre as excitações parciais devidas a cada componente, ou seja, as excitações são espalhadas de maneira apropriada e, a cada banda, apenas a maior componente presente irá determinar a excitação resultante. Ainda que tal procedimento não respeite a característica de aditividade do mascaramento, sua implementação é consideravelmente mais simples, motivo pelo qual esta estratégia é mais largamente utilizada. Os padrões de excitação resultantes são transformados em uma função densidade, também chamada de sonoridade específica, a qual é definida pela postulação de que a área entre esta função e

³A sonoridade de um sinal é dada na unidade fônon, e corresponde ao nível de pressão sonora em decibéis de um tom puro a 1 kHz que produz a mesma sonoridade percebida para o sinal medido.

o eixo da frequência resulta na sonoridade do estímulo sonoro.

O modelo de Zwicker para o cálculo da sonoridade percebida é a base da maioria dos modelos psicoacústicos usados em codificação de áudio (Zwicker and Fastl, 1990).

A.2.2 Modelos Psicoacústicos em Codificação de Áudio

Codificadores de áudio baseados em modelos psicoacústicos são capazes de obter taxas de compressão bastante significativas com pouca ou nenhuma perda de qualidade perceptível para o ouvido humano.

Estes codificadores partem do princípio de que qualquer ruído adicionado ao sinal de áudio e que cuja potência esteja abaixo do limiar de mascaramento deste sinal, em todas as faixas de frequência do espectro, será inaudível e portanto não introduzirá degradações perceptíveis. O ruído, neste caso, é decorrente da quantização do sinal no processo de codificação.

A idéia básica por trás desses codificadores consiste em identificar, para cada janela de análise, as componentes tonais e não tonais do sinal de áudio presentes em uma gama de frequências, e deduzir, a partir das propriedades de mascaramento do ouvido humano, a quantidade de ruído de quantização que pode ser introduzida em cada faixa de frequência.

A quantidade de bits alocada para a codificação de uma dada faixa de frequência é função da quantidade de ruído de quantização admitida para essa faixa, de forma a minimizar a distorção audível provocada pelo codificador.

Cada codificador utiliza um modelo psicoacústico particular, que difere dos demais pela aproximação que é feita com relação aos resultados teóricos e experimentais descritos nas seções anteriores. Mas a estrutura global desses modelos é similar. Um modelo psicoacústico pode se decompor em três etapas (Guimarães, 1998):

1. Cálculo da excitação
 - Cálculo do espectro basilar
 - Convolução com uma função de espalhamento
2. Cálculo da taxa de mascaramento
 - Estimativa da tonalidade do sinal
 - Ponderação dos valores de taxa de mascaramento nos casos *tonal mascarando ruído* e *ruído mascarando ruído* pela tonalidade do sinal
3. Cálculo do limiar de mascaramento
 - Cálculo do limiar de audição mascarado
 - Cálculo do limiar de mascaramento através de um algoritmo de otimização

Na prática, a essas três etapas deve-se adicionar o cálculo das relações sinal-máscara nas sub-bandas úteis ao codificador. Em um codificador tradicional de áudio, o modelo psicoacústico recebe como entrada uma janela de sinal e retorna uma relação sinal-máscara para cada sub-banda do codificador.

Nas seções seguintes, serão brevemente descritos os modelos psicoacústicos 1 e 2 definidos na norma MPEG-2. Para maiores detalhes, consultar (Guimarães, 1998).

Cálculo da Excitação

A excitação pode ser calculada de duas maneiras. A primeira calcula diretamente a potência na saída da cóclea, considerando-a como um filtro através do qual passa o sinal de áudio. Na segunda, a excitação é calculada a partir do periodograma do sinal através do espectro basilar, isto é, as raiais do espectro do sinal que se encontram dentro de uma banda do espectro basilar são somadas entre si, resultando na potência total de sinal dentro daquela banda.

O nível de excitação é então obtido aplicando-se uma função de espalhamento ao resultado da operação anterior, modelando o efeito da potência em uma banda sobre as bandas adjacentes, e subtraindo-se o nível de pressão acústica correspondente ao limiar de mascaramento absoluto (no silêncio).

Espectro Basilar

O modelo psicoacústico 2 de MPEG-2 calcula a excitação pelo segundo método. É calculado um espectro basilar de 49 sub-bandas basilares, seguindo aproximadamente passos de meio Bark.

O modelo 1 de MPEG também é baseado no periodograma do sinal, mas faz um cálculo mais elaborado. As raiais tonais são extraídas seguindo um procedimento de 3 etapas:

1. São inicialmente selecionadas as raiais que ultrapassam as suas vizinhas imediatas.
2. São eliminadas as raiais tonais que não ultrapassam de ao menos 7 dB um grupo de raiais vizinhas. O grupo de raiais vizinhas é função da taxa de amostragem e da raia tonal. O número de raiais do grupo aumenta com a frequência da raia tonal para levar em conta a resolução frequencial não uniforme do ouvido.
3. Em cada banda crítica é conservada apenas a raia tonal mais forte.

Ao final deste procedimento, restam no máximo 25 raiais tonais selecionadas. A potência de cada tonal é calculada somando-se a raia selecionada com suas duas vizinhas imediatas.

O espectro basilar da parte não tonal do sinal é calculado somando-se para cada banda crítica as raiais do periodograma do sinal de áudio, após a eliminação de todas as raiais selecionadas como tonais e suas vizinhas imediatas.

O espectro é em seguida sub-amostrado. Por exemplo, para a camada 1 com uma taxa de amostragem de 32 kHz, passa-se de um espectro de 256 raiais a 108 raiais. Nas seis primeiras sub-bandas de 500 Hz, todas as componentes são mantidas. Nas seis subbandas seguintes, uma componente a cada duas é mantida. Nas sub-bandas restantes, até 15 kHz, uma raia em cada quatro é mantida.

Os resultados deste processo são as tonais e o espectro basilar da parte não tonal.

Função de Espalhamento

A função de espalhamento determina o quanto da potência presente em uma determinada sub-banda será transferida às sub-bandas vizinhas.

No modelo 2 de MPEG, são definidas funções de espalhamento que não dependem da potência do sinal de áudio, sendo utilizada uma curva muito próxima da obtida experimentalmente para o ouvido humano. Já no modelo 1 de MPEG, a função de espalhamento é representada por uma curva linear por partes e dependente da potência do sinal de áudio.

Somatório da Potência

Os dois modelos MPEG utilizam uma lei de adição linear para calcular a excitação total em cada banda a partir da potência presente na própria banda e nas bandas adjacentes. No caso do modelo 1, são calculadas separadamente as partes tonal e não tonal da excitação.

Cálculo da Taxa de Mascaramento

A taxa de mascaramento é função da frequência e da natureza tonal ou não tonal do som mascarado e mascarador. No caso da codificação de áudio, o som mascarador é o sinal original e o som mascarado é o ruído de quantização introduzido pelo codificador.

O codificador MPEG considera implicitamente que o ruído de quantização é de natureza não tonal. Portanto, os casos considerados para o cálculo da taxa de mascaramento são *ruído mascarando ruído* e *tonal mascarando ruído*. O nível de mascaramento no primeiro caso foi medido experimentalmente, estando entre -2 dB nas baixas frequências e -6 dB nas altas frequências. No caso *tonal mascarando ruído*, é utilizada a fórmula:

$$L_{av}(j) = -(15,5 + j) \quad (\text{A.2})$$

onde L_{av} é o nível de mascaramento e j é a frequência em Bark.

O modelo 2 de MPEG calcula um índice de tonalidade dependente da frequência.

O modelo 1 de MPEG não tem necessidade de estimar a tonalidade do sinal mascarador, uma vez que a parte tonal do sinal já foi separada da não tonal durante o cálculo da excitação.

Ponderação por um Índice de Tonalidade (modelo 2 de MPEG)

O índice de tonalidade no modelo 2 de MPEG é calculado a partir da predição do espectro de uma janela do sinal em função dos espectros das janelas anteriores. Se o espectro pode ser predito com exatidão (sinal completamente tonal), o coeficiente de não capacidade de predição vale idealmente 0; se, ao contrário, a predição é impossível (sinal completamente não tonal), o coeficiente de não capacidade de predição vale idealmente 1. Na prática, este coeficiente resultará em um valor entre 0 e 1, conforme o sinal analisado se aproxime mais de um sinal tonal ou não tonal.

Para o cálculo da taxa de mascaramento, o espectro basilar é calculado como descrito nas seções anteriores, mas com as raias do periodograma do sinal ponderadas pelo coeficiente de não capacidade de predição. As partes tonal e não tonal do espectro basilar são calculadas separadamente e combinadas para produzir a taxa de mascaramento total.

Cálculo Separado Tonais / Não Tonais (modelo 1 de MPEG)

No modelo 1 de MPEG, as componentes tonais e não tonais já são separadas no cálculo da excitação, sendo diretamente convertidas em taxas de mascaramento.

Cálculo do Limiar de Mascaramento

Para obter o limiar de mascaramento, a excitação é multiplicada pela taxa de mascaramento. No caso do modelo 1 de MPEG, no qual a excitação e a taxa de mascaramento são calculadas separadamente para as partes tonal e não tonal do sinal, o limiar de mascaramento global é obtido pela soma dos limiares obtidos para as partes tonal e não tonal.

Os modelos levam ainda em conta o limiar de audição absoluto na construção do limiar de mascaramento do sinal. No modelo 2 de MPEG, se o limiar de mascaramento calculado é inferior ao limiar absoluto para determinadas bandas de frequência, passa-se a considerar o próprio limiar absoluto como limiar de mascaramento do sinal. Essa correção tem efeito sobretudo nas bandas de frequência mais elevada. Já no modelo 1 de MPEG, o limiar absoluto é simplesmente adicionado ao limiar de mascaramento calculado.

Finalmente, é efetuada uma etapa final de conversão do limiar de mascaramento, originalmente obtido nas sub-bandas do modelo psicoacústico, às sub-bandas utilizadas pelo codificador.