

Universidade Federal da Paraíba- UFPB
Centro de Ciências e Tecnologia- CCT
Departamento de Engenharia Elétrica- DEE

Relatório de Estágio

Período: Outubro de 1999 a Fevereiro de 2000

Bolsista: Madhavan S. Vajapeyam

Orientador: Prof. Benedito G. Aguiar Neto, Dr. Ing.

Supervisor: Francisco Madeiro B. Júnior, MSc.

ESTAGIÁRIO: MADHAVAN SRINIVASAN VAJAPYAM

MATRÍCULA: 9511172-6

LOCAL DO ESTÁGIO: LAPS- LABORATÓRIO DE AUTOMAÇÃO E
PROCESSAMENTO DE SINAIS

PROFESSOR ORIENTADOR: BENEDITO GUIMARÃES AGUIAR NETO

SUPERVISOR: FRANCISCO MADEIRO BERNADINO JÚNIOR

TIPO DE ESTÁGIO: SUPERVISIONADO

PERÍODO DE ESTÁGIO: 01 DE OUTUBRO DE 1999 A 29 DE FEVEREIRO DE 2000



Biblioteca Setorial do CDSA. Fevereiro de 2021.

Sumé - PB

Agradecimentos

Gostaria de prestar meus mais sinceros agradecimentos:

A Deus, por ter me dado forças para concluir mais uma etapa da minha vida.

Aos professores Benedito Guimarães Aguiar Neto, Bruno Barbosa Albert e Rômulo Raimundo Maranhão do Valle, pela atenção e oportunidade de aprender.

Ao doutorando Francisco Madeiro Bernadino Júnior, pelos incentivos, sugestões e por ter despertado em mim o interesse pela área de Processamento de Sinais.

A meus pais e meu irmão Murali por todo apoio dado até hoje.

Aos colegas da turma concluinte de Engenharia Elétrica, período 99.2, pelo companherismo e amizade demonstrados a todo momento.

A todos os amigos do Laboratório de Automação e Processamento de Sinais e do PET Elétrica, cuja colaboração foi fundamental para a realização deste trabalho.

“Se pude enxergar mais adiante é porque me apoiei sobre ombros de gigantes”

Isaac Newton

SUMÁRIO

1. INTRODUÇÃO.....	2
2. WAVELETS: CARACTERÍSTICAS GERAIS	3
3. MULTIRESOLUÇÃO E FUNÇÕES DE ESCALONAMENTO	8
3.1 ESPAÇOS DE SINAIS	8
3.2 A FUNÇÃO DE ESCALONAMENTO	9
3.3 ANÁLISE MULTIRESOLUCIONAL	10
3.4 AS FUNÇÕES WAVELET	12
3.5 A WAVELET DE HAAR	16
3.6 A TRANSFORMADA WAVELET DISCRETA	19
3.7 TEOREMA DE PARSEVAL	21
4. BANCOS DE FILTROS APLICADOS À OBTENÇÃO DA DWT	22
4.1 ANÁLISE – DA MAIOR PARA A MENOR RESOLUÇÃO	22
4.1.1 Filtragem e Subamostragem (ou Decimação)	23
4.2 SÍNTESE – DA MENOR RESOLUÇÃO PARA A MAIOR	27
4.2.1 Filtragem e Interpolação.....	27
5. QUANTIZAÇÃO VETORIAL	30
5.1 COMPRESSÃO DIGITAL DE SINAIS	30
5.2 QUANTIZAÇÃO VETORIAL	31
5.3 PROJETO DE DICIONÁRIOS PARA QUANTIZAÇÃO VETORIAL	33
5.3.1 Algoritmo LBG	34
5.3.2 Algoritmo SSC.....	35
5.4 AVALIAÇÃO DE DESEMPENHO	35
5.4.1 Medidas Objetivas	36
5.4.2 Medidas Subjetivas	38
6. SIMULAÇÕES REALIZADAS E RESULTADOS.....	40
6.1 SIMULAÇÕES ENVOLVENDO SINAIS DE VOZ	40
6.1.1 Resultados.....	42
6.2 SIMULAÇÕES ENVOLVENDO IMAGENS	47
6.2.1 Resultados.....	47
7. CONCLUSÕES.....	50

1. Introdução

Este trabalho é um relato das atividades do aluno Madhavan Srinivasan Vajapeyam, do curso de Engenharia Elétrica da Universidade Federal da Paraíba – Campus II, durante o período de 01 de Outubro de 1999 a 29 de Fevereiro de 2000, integralizando uma carga horária total de 120 horas.

O estágio foi realizado no Laboratório de Automação e Processamento de Sinais, do Departamento de Engenharia Elétrica, UFPB, Campus II. Nesse laboratório são desenvolvidos trabalhos em várias áreas, incluindo Processamento de Voz e de Imagens, Compressão de Sinais e Aplicações de Redes Neurais e Wavelets, envolvendo alunos de pós-graduação, iniciação científica e estagiários, sob a orientação de professores responsáveis.

O presente trabalho está disposto em capítulos, em que cada um se refere a uma etapa do estágio, e organizados em ordem cronológica de realização.

O plano de estágio consiste nas seguintes etapas:

- Etapa 1: Revisão Bibliográfica
 - Transformada Wavelet ;
 - Quantização Vetorial.
- Etapa 2: Aplicações
 - Compressão de Voz ;
 - Compressão de Imagens.
- Etapa 3: Elaboração do relatório de descrição das atividades realizadas.

O Capítulo 2 deste relatório faz uma introdução às wavelets, ressaltando suas principais características e vantagens. No Capítulo 3 são explorados e discutidos os conceitos de *multiresolução* e *funções de escalonamento*, fundamentais na teoria de wavelets. No Capítulo 4 é mostrado como são feitas as implementações práticas das operações de decomposição e síntese, utilizando filtros discretos. A revisão bibliográfica é concluída no Capítulo 5, que trata da Quantização Vetorial como um eficiente método para compressão de sinais. A etapa prática do estágio, correspondente às aplicações, é descrita no Capítulo 6, em que são também apresentados e discutidos os resultados obtidos.

2. Wavelets: Características Gerais

Uma onda (*wave*) pode ser vista uma função oscilatória no tempo ou espaço, como por exemplo uma senóide [1]. Através da análise de Fourier, um sinal pode ser decomposto como uma soma de senóides (ou exponenciais complexas). A série de Fourier, (bem como a Transformada de Fourier) apresenta-se como uma ferramenta bastante útil para análise e desenvolvimento de aplicações de Engenharia, especialmente no estudo de fenômenos periódicos, invariantes no tempo ou estacionários.

Uma *wavelet* é uma “pequena onda”, cuja energia está concentrada no tempo. A utilização de wavelets constitui uma alternativa bastante adequada para análise de fenômenos transitórios ou não-estacionários. Ela também apresenta caráter oscilatório, mas a decomposição de um sinal como soma de wavelets permite análise simultânea no domínio do tempo e da frequência. A Figura 2.1 ilustra um segmento de uma senóide e wavelet. A senóide (em virtude de oscilar entre $-\infty \leq t \leq \infty$), apresenta energia infinita; a *wavelet*, por sua vez, apresenta energia finita.

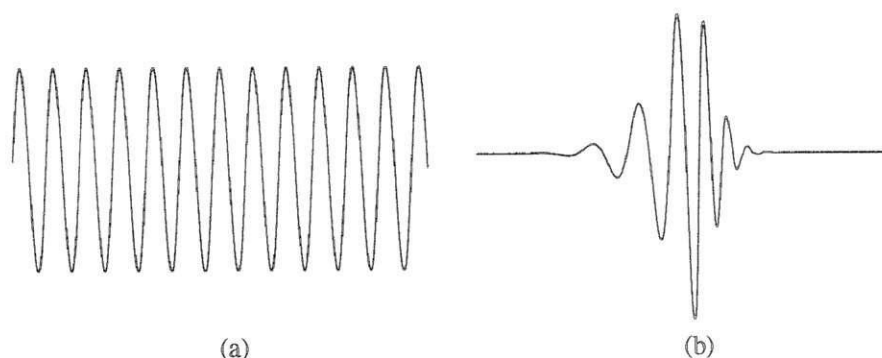


Figura 2.1: Senóide (a) e wavelet (b).

As *wavelets* são utilizadas numa expansão em série, da mesma forma como é feito numa série de Fourier. Como já se conhece, um sinal $f(t)$ pode ser melhor analisado se expresso através de uma decomposição linear por [2]:

$$f(t) = \sum_I a_I \psi_I(t), \quad (2.1)$$

em que a_l são coeficientes reais de expansão e $\psi_l(t)$ formam um conjunto de funções reais de t denominado de conjunto de expansão. Se a expansão (2.1) for única, esse conjunto é denominado *base* da classe de funções que ele representa. Se a base é ortogonal, então os coeficientes podem ser calculados pelo produto interno:

$$a_k = \langle f(t), \psi(t) \rangle = \int f(t) \psi_k(t) dt \quad (2.2)$$

já que o produto interno entre quaisquer duas funções de uma base ortogonal é zero. Para a série de Fourier as bases ortogonais $\psi_k(t)$ são $\sin(k\omega_0 t)$ e $\cos(k\omega_0 t)$, e ω_0 é a frequência angular fundamental de oscilação. Para uma expansão *wavelet*, constrói-se um sistema com dois parâmetros, de modo que (2.1) fica:

$$f(t) = \sum_k \sum_j a_{j,k} \Psi_{j,k}(t) \quad (2.3)$$

em que $\Psi_{j,k}(t)$ são as funções da expansão *wavelet* e geralmente formam uma base ortogonal. O conjunto de coeficientes de expansão $a_{j,k}$ é chamado de *transformada wavelet discreta* (DWT) de $f(t)$ e (2.3) é a transformada inversa.

O conjunto de expansão *wavelet* não é único. Existem muitos sistemas que podem ser utilizados eficientemente, mas quase todos apresentam as seguintes características genéricas [2]:

- 1- Um sistema *wavelet* é um conjunto de blocos para construir ou representar um sinal ou função. É um conjunto de expansão bidimensional (em geral forma uma base) para alguma classe de sinais de uma ou mais dimensões.
- 2- A expansão *wavelet* dá uma localização do sinal em tempo e em frequência. Além disso, a maior parte da energia do sinal é bem representada por poucos coeficientes de expansão $a_{j,k}$.
- 3- O cálculo dos coeficientes a partir do sinal pode ser feito eficientemente. Em geral, o número de multiplicações e adições presentes no cálculo de muitas

transformadas wavelets cresce linearmente com o tamanho do sinal (N). Na FFT, por exemplo, esse número é maior, da ordem de $N \cdot \log N$.

Enquanto que a série de Fourier mapeia uma função unidimensional em uma seqüência unidimensional de coeficientes, a expansão *wavelet* mapeia em um conjunto bidimensional de coeficientes. É essa representação bidimensional que permite dar uma localização do sinal em tempo e freqüência simultaneamente.

Existem três características [2,3] que são mais específicas às expansões wavelet .

- 1- Os sistemas wavelet de primeira geração são gerados a partir de uma única função de escalonamento ou wavelet por simples translação e escalonamento. A parametrização bidimensional é feita a partir da função $\psi(t)$ (chamada de wavelet-mãe) por:

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k) \quad j, k \in \mathbb{Z}, \quad (2.4)$$

em que o fator $2^{j/2}$ mantém uma norma (energia) constante independentemente da escala j .

- 2- Praticamente todos os sistemas úteis satisfazem às condições de *multiresolução*. Isso significa que se um conjunto de sinais pode ser representado por uma soma ponderada de $\psi(t - k)$, então um conjunto maior (incluindo o original) pode ser representado por uma soma ponderada de $\psi(2t - k)$. Em outras palavras, se os sinais básicos de expansão e seus passos de translação forem estreitados pela metade, eles poderão representar uma classe maior de sinais de maneira exata ou dar uma melhor representação de qualquer sinal.
- 3- Os coeficientes de menor resolução podem ser calculados a partir dos de maior resolução através de um algoritmo em árvore, denominado banco de filtros. Isso permite um cálculo eficiente dos coeficientes de expansão (também conhecidos como transformada wavelet discreta).

A Figura 2.2 ilustra a translação e o escalonamento de uma wavelet-mãe, como mostrado em (2.4). À medida que k muda, a localização da wavelet muda ao longo do eixo horizontal. Isso permite a representação explícita de eventos no tempo ou espaço. À medida que j se altera, a wavelet muda de escala. Isso permite uma representação diversificada de detalhe ou resolução. É importante notar que enquanto a escala se torna mais fina (j maior), os passos de translação são menores. De modo a manter a clareza da Figura no eixo horizontal, apenas as wavelets para $k = 1, 5, 9, \dots$ são mostradas.

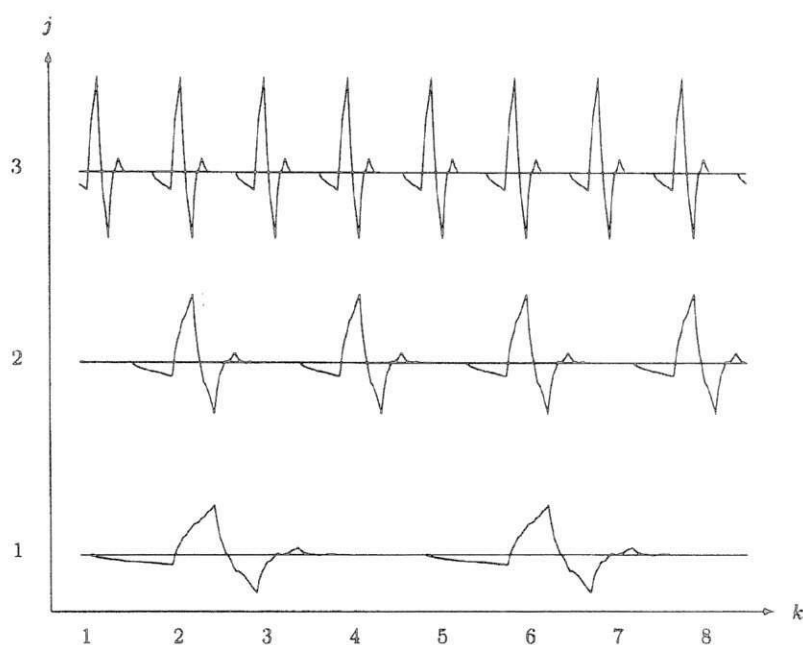


Figura 2.2: Translação e escalonamento da wavelet Daubechies' 4.

A forma da wavelet-mãe também pode ser modificada, e isso é feito durante o projeto de um sistema wavelet e permite que um conjunto represente bem uma certa classe de sinais.

Para a maioria dos sistemas de expansão de sinais (como a série e a transformada de Fourier), a base de expansão é escolhida, e então as propriedades da transformada resultante são analisadas. Para um sistema wavelet, essas propriedades são definidas inicialmente e então as funções de base são obtidas visando atender aos requisitos pré-estabelecidos. Dessa forma, um sistema wavelet pode ser projetado atender às necessidades particulares de uma determinada aplicação.

A análise wavelet é bastante adequada para sinais transientes, enquanto que a análise de Fourier é mais adequada para sinais periódicos ou sinais estacionários. Algumas

propriedades que justificam a eficiência e eficácia da análise wavelet para uma grande variedade de sinais e fenômenos são [1]:

- 1- O valor dos coeficientes de expansão diminui rapidamente com j e k para uma vasta classe de sinais. Isso significa que a base wavelet é uma *base incondicional* e esta propriedade torna a DWT adequada para aplicações envolvendo compressão de sinais, remoção de ruído e detecção.
- 2- A expansão wavelet permite uma descrição local mais exata do sinal e uma separação de suas características. Um coeficiente de Fourier representa uma componente de duração infinita, de forma que eventos temporários têm que ser descritos através de uma característica de fase que permita cancelamentos ou reforços. Um coeficiente wavelet, por outro lado, representa uma componente que já é por si só local e mais fácil de ser compreendida.
- 3- Wavelets são ajustáveis e adaptáveis, devido ao fato de não existir apenas uma wavelet. Dessa forma, elas podem ser projetadas para aplicações específicas, como já mencionado anteriormente.
- 4- Os computadores digitais são bastante adequados ao cálculo da DWT, cuja complexidade é menor que na FFT .

3. Multiresolução e Funções de Escalonamento

Tanto as interpretações matemáticas quanto as práticas das wavelets são melhor compreendidas através do conceito de resolução. Para isso, começa-se o estudo com uma *função de escalonamento* $\varphi(t)$, também representada por $\phi(t)$, e em seguida define-se a wavelet $\psi(t)$. A formulação multiresolucional é útil para se representar sinais em que uma descrição em tempo-frequência ou tempo-escala é desejada.

Neste capítulo serão abordados inicialmente alguns conceitos básicos de análise funcional e, em seguida, será mostrado como funções podem ser representadas através de funções de escalonamento e wavelets. Em seguida será mostrado um exemplo simples de representação em termos da função de Haar, que é uma wavelet bastante simples. O capítulo é finalizado com uma introdução à Transformada Wavelet Discreta e com o Teorema de Parseval.

3.1 Espaços de sinais

De modo a entender-se a idéia de funções ou sinais sendo representados por somas de funções de escalonamento e/ou wavelets, serão apresentadas alguns conceitos e terminologias da análise funcional [1].

Um *espaço de funções* é um espaço vetorial linear (de dimensão finita ou infinita) em que os vetores são funções, os escalares são números reais ou complexos, e a multiplicação escalar e a adição de vetores são similares às operações em (2.1). O *produto interno* é um escalar a obtido através de dois vetores $f(t)$ e $g(t)$:

$$a = \langle f(t), g(t) \rangle = \int f^*(t)g(t)dt, \quad (3.1)$$

com um intervalo de integração dependente da classe do sinal em consideração. Dois sinais são *ortogonais* se seu produto interno for zero.

Um espaço importante em processamento de sinais é $L^2(\mathbb{R})$. Este é o espaço de todas as funções $f(t)$ de quadrado integrável. Embora a maioria das definições apresentadas neste

capítulo se limite a que pertençam a $L^2(\mathbb{R})$, muitos dos resultados valem para uma classe maior de sinais.

Num espaço vetorial de sinais S , se qualquer $f(t) \in S$ puder ser expresso como $f(t) = \sum_k a_k \varphi_k(t)$, o conjunto de funções $\varphi_k(t)$ é chamado de conjunto de expansão para o espaço S . Se a representação é única, o conjunto é uma base. Alternativamente, pode-se partir do conjunto de expansão ou base e definir o espaço S como o conjunto de todas as funções que podem ser expressas por $f(t) = \sum_k a_k \varphi_k(t)$. Esse é o *espaço gerado* pela base.

3.2 A Função de Escalonamento

De modo a usar a idéia de multiresolução, será inicialmente definida a função de escalonamento e a seguir a wavelet em termos desta. Um conjunto de funções de escalonamento é definido em termos de translações inteiras da função de escalonamento básica como

$$\varphi_k(t) = \varphi(t - k) \quad k \in \mathbb{Z} \quad \varphi \in L^2. \quad (3.2)$$

O subespaço de $L^2(\mathbb{R})$ gerado por essas funções é definido como

$$\mathcal{V}_0 = \overline{\text{Span}_k \{ \varphi_k(t) \}}, \quad (3.3)$$

para todos os inteiros k . Isso significa que

$$f(t) = \sum_k a_k \varphi_k(t) \quad \text{para qualquer } f(t) \in \mathcal{V}_0. \quad (3.4)$$

Em geral, pode-se aumentar o tamanho do subespaço gerado mudando a escala de tempo das funções de escalonamento. Uma família bidimensional é gerada a partir da função de escalonamento básica por translação e escalonamento:

$$\varphi_{j,k}(t) = 2^{j/2} \varphi(2^j t - k). \quad (3.5)$$

Desse modo,

$$\mathcal{V}_j = \overline{\text{Span}_k \{ \varphi_{j,k}(t) \}} \quad (3.6)$$

para todos os inteiros $k \in \mathbb{Z}$. Isto significa que se $f(t) \in \mathcal{V}_j$, então pode ser expresso como

$$f(t) = \sum_k a_k \varphi(2^j t + k). \quad (3.7)$$

Para $j > 0$, o subespaço gerado é maior pois $\varphi_{j,k}(t)$ é mais estreito e é transladado a passos menores, podendo, portanto, representar detalhes mais finos. Tem-se, portanto, uma maior resolução do sinal representado. Para $j < 0$, $\varphi_{j,k}(t)$ é mais largo e é transladado a passos maiores. Essas funções mais largas podem representar apenas informações mais grosseiras. Neste caso, o espaço gerado e a resolução são menores.

3.3 Análise Multiresolucional

O requisito básico para a análise multiresolucional [4] é um aninhamento dos espaços gerados, de modo que

$$\dots \subset \mathcal{V}_2 \subset \mathcal{V}_1 \subset \mathcal{V}_0 \subset \mathcal{V}_1 \subset \mathcal{V}_2 \subset \dots \subset L^2, \quad (3.8)$$

ou

$$\mathcal{V}_j \subset \mathcal{V}_{j+1} \text{ para todo } j \in \mathbb{Z}, \quad (3.9)$$

com

$$\mathcal{V}_{-\infty} = \{0\}, \quad \mathcal{V}_{\infty} = L^2. \quad (3.10)$$

O espaço que contém sinais de alta resolução conterá aqueles de mais baixa resolução também. Em virtude da definição de \mathcal{V}_j , os espaços têm de satisfazer a condição natural de escalonamento

$$f(t) \in \mathcal{V}_j \Leftrightarrow f(2t) \in \mathcal{V}_{j+1}, \quad (3.11)$$

que assegura que elementos de um espaço são simplesmente versões escalonadas dos elementos do espaço seguinte. A relação entre os espaços gerados está ilustrada na Figura 3.1.

Em consequência da condição (3.11), se $\varphi(t)$ está contido em \mathcal{V}_0 , estará também em \mathcal{V}_1 , o espaço gerado por $\varphi(2t)$. Isso significa que $\varphi(t)$ pode ser expresso em termos de uma soma ponderada de deslocamentos de $\varphi(2t)$ como

$$\varphi(t) = \sum_n h(n)\sqrt{2}\varphi(2t-n), \quad n \in \mathbb{Z}, \quad (3.12)$$

em que os coeficientes $h(n)$ são uma seqüência de números reais ou complexos denominados de coeficientes da função de escalonamento (ou filtro de escalonamento ou vetor de escalonamento) e o fator $\sqrt{2}$ mantém a norma (energia) da função de escalonamento constante, ou seja, $\int |\varphi(t)|^2 dt = \int |\sqrt{2}\varphi(t)|^2 dt$. Essa Equação recursiva é fundamental para a teoria de funções de escalonamento e é conhecida por Equação de refinamento ou Equação de análise multiresolucional.

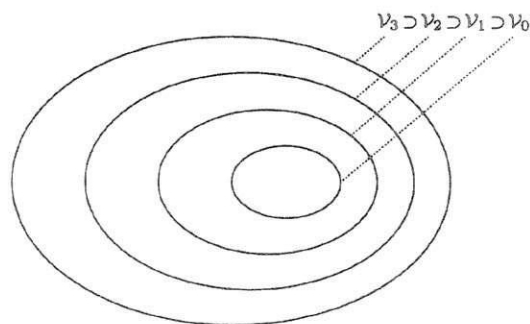


Figura 3.1: Espaços vetoriais aninhados gerados pelas funções de escalonamento.

Um exemplo de função de escalonamento é a de Haar, que é simplesmente um pulso $\varphi(t)$ de largura e altura unitárias, como mostrado na Figura 3.2.

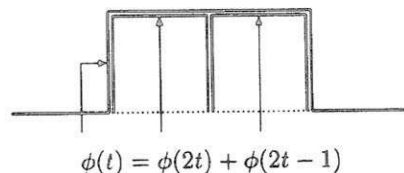


Figura 3.2 : Função de escalonamento de Haar.

Pode-se observar facilmente que $\varphi(2t)$ pode ser usado para construir $\varphi(t)$ fazendo

$$\varphi(t) = \varphi(2t) + \varphi(2t-1). \quad (3.13)$$

Isso significa que (3.12) é satisfeita com os coeficientes $h(0) = 1/\sqrt{2}$, $h(1) = 1/\sqrt{2}$. Já a função de escalonamento Daubechies'4 mostrada na Figura 3.3 satisfaz (3.12) para $h = \{0,4830 \ 0,8365 \ 0,2241 \ -0,1294\}$.

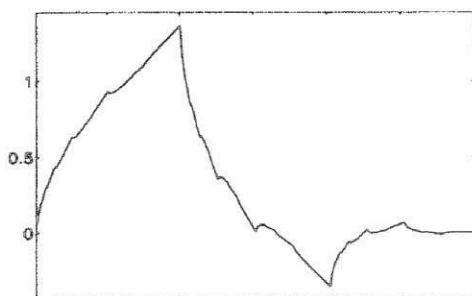


Figura 3.3: Função de escalonamento Daubechies'4.

3.4 As Funções Wavelet

Em vez de serem utilizadas as funções $\varphi_{j,k}(t)$ e aumentar-se j de modo tornar maior o subespaço gerado, pode-se utilizar um conjunto um pouco diferente de funções $\psi_{j,k}(t)$ que gere as diferenças entre subespaços criados através de várias escalas da função de escalonamento [1]. Essas funções são as wavelets.

Há várias vantagens em se requerer que as funções de escalonamento e as wavelets sejam ortogonais. Funções de base ortogonais permitem cálculos simples dos coeficientes de expansão e, uma vez satisfazendo o teorema de Parseval, permitem um particionamento adequado da energia do sinal no domínio da transformada wavelet. O complemento ortogonal de \mathcal{V}_j em \mathcal{V}_{j+1} é definido como W_j . Isto significa que todos os membros de \mathcal{V}_j são ortogonais a todos os membros de W_j . Portanto

$$\langle \varphi_{j,k}(t), \psi_{j,l}(t) \rangle = \int \varphi_{j,k}(t) \psi_{j,l}(t) = 0, \quad (3.14)$$

para todos os valores apropriados de $j, k, l \in \mathbb{Z}$.

As relações entre os vários subespaços podem ser vistas pelas expressões a seguir. De acordo com (3.8) percebe-se que se pode começar a partir de qualquer \mathcal{V}_j , por exemplo em $j = 0$, e então

$$\mathcal{V}_0 \subset \mathcal{V}_1 \subset \mathcal{V}_2 \subset \dots \subset L^2. \quad (3.15)$$

Define-se agora o espaço W_0 gerado pelas wavelets de tal forma que

$$\mathcal{V}_1 = \mathcal{V}_0 \oplus W_0, \quad (3.16)$$

que pode ser estendido a

$$\mathcal{V}_2 = \mathcal{V}_0 \oplus W_0 \oplus W_1. \quad (3.17)$$

De forma geral, tem-se

$$L^2 = \mathcal{V}_0 \oplus W_0 \oplus W_1 \oplus \dots \quad (3.18)$$

quando \mathcal{V}_0 for o espaço inicial gerado pela função de escalonamento $\varphi(t-k)$. A Figura 3.4 mostra o aninhamento dos espaços \mathcal{V}_j das funções de escalonamento para diferentes escalas. Também está ilustrado o fato de os espaços wavelet serem as diferenças disjuntas (exceto pelo elemento zero) ou complementos ortogonais dos espaços \mathcal{V}_j .

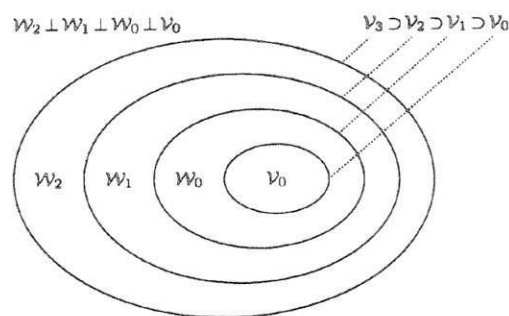


Figura 3.4: Espaços vetoriais gerados pelas funções de escalonamento e wavelets.

A escala do espaço inicial é arbitrária e pode ser escolhida numa resolução maior, por exemplo em $j = 10$, resultando em

$$L^2 = \mathcal{V}_{10} \oplus W_{10} \oplus W_{11} \oplus \dots \quad (3.19)$$

ou numa resolução menor, como $j = 5$, resultando em

$$L^2 = \mathcal{V}_{-5} \oplus W_{-5} \oplus W_{-4} \oplus \dots \quad (3.20)$$

ou até eliminando-se o espaço da função de escalonamento, fazendo $j = -\infty$, de modo que

$$L^2 = \dots \oplus W_{-2} \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus W_2 \oplus \dots \quad (3.21)$$

Uma outra maneira de se descrever a relação de \mathcal{V}_0 com os espaços wavelet é notando-se que

$$W_{-\infty} \oplus \dots \oplus W_{-1} = \mathcal{V}_0, \quad (3.22)$$

que mostra que a escala do espaço de escalonamento pode ser escolhida arbitrariamente. Na prática, ela é normalmente escolhida para se representar o detalhe de interesse mais grosseiro do sinal.

Como as wavelets pertencem ao espaço gerado pela próxima função de escalonamento mais estreita, $\mathcal{V}_0 \subset W_1$, elas podem ser representadas por uma soma ponderada da função de escalonamento $\varphi(2t)$ deslocada, segundo (3.12):

$$\psi(t) = \sum_n h_1(n) \sqrt{2} \varphi(2t - n), \quad n \in \mathbb{Z}, \quad (3.23)$$

em que $h_1(n)$ é algum conjunto de coeficientes. Utilizando-se a condição que as wavelets gerem a “diferença” ou complemento ortogonal dos espaços, e que translações inteiras das wavelets são também ortogonais, pode-se mostrar [1] que os coeficientes wavelet estão relacionados aos coeficientes da função de escalonamento por

$$h_1(n) = (-1)^n h(1-n). \quad (3.24)$$

A função gerada por (3.23) fornece o protótipo ou wavelet-mãe $\psi(t)$ para uma classe de funções de expansão da forma

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k) \quad (3.25)$$

A wavelet de Haar associada à função mostrada na Figura 3.2 está mostrada na Figura 3.5. Para essa wavelet os coeficientes em (3.23) são $h_1(0) = 1/\sqrt{2}$ e $h_1(1) = -1/\sqrt{2}$, que satisfazem (3.24).

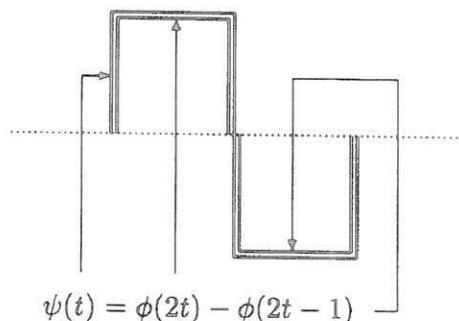


Figura 3.5: Wavelet de Haar.

Na Figura 3.6 está mostrada a wavelet Daubechies'4, e os coeficientes para este caso são $h_1 = \{0,1294 \ 0,2241 \ -0,8365 \ 0,4830\}$.

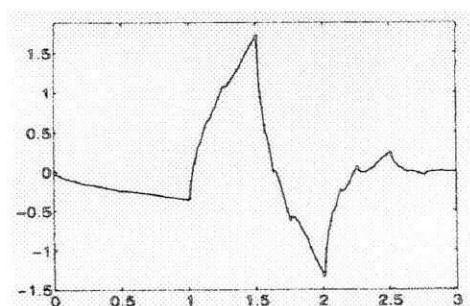


Figura 3.6: Wavelet Daubechies'4.

Dispõe-se agora de um conjunto de funções $\varphi_k(t)$ e $\psi_{j,k}(t)$ que pode gerar todo o $L^2(\mathbb{R})$. De acordo com (3.18) qualquer função $g(t) \in L^2(\mathbb{R})$ pode ser escrita como

$$g(t) = \sum_{k=-\infty}^{\infty} c(k)\varphi_k(t) + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} d(j,k)\psi_{j,k}(t), \quad (3.26)$$

ou seja, como uma expansão em série em termos da função de escalonamento e das wavelets. O primeiro somatório de (3.26) resulta em uma função que é uma aproximação grosseira (de baixa resolução) de $g(t)$. À medida que índices j maiores são incluídos no segundo somatório, uma resolução mais alta é obtida, aumentando os detalhes. Esse

comportamento é similar ao da série de Fourier, em que termos de maior frequência contêm os detalhes finos do sinal.

Caso as funções de expansão $\varphi_k(t)$ e $\psi_{j,k}(t)$ formem uma base ortonormal, os coeficientes podem ser calculados através de

$$c(k) = c_0(k) = \langle g(t), \varphi_k(t) \rangle = \int g(t) \varphi_k(t) dt \quad (3.27)$$

e

$$d_j(k) = d(j, k) = \langle g(t), \psi_{j,k}(t) \rangle = \int g(t) \psi_{j,k}(t) dt. \quad (3.28)$$

Deve-se observar que em (3.26) foi utilizado $j=0$ para a família de funções de escalonamento, mas ela poderia ser definida para qualquer j , conforme mostrado em (3.18), (3.19) e (3.20), por exemplo.

3.5 A Wavelet de Haar

Em 1910, Haar mostrou que certas funções compostas de pulsos quadráticos poderiam ser transladadas e escalonadas de modo a formar uma base para L^2 [1]. Anos mais tarde descobriu-se que o sistema de Haar era de fato um sistema wavelet.

Se a função de escalonamento for escolhida impondo-se a condição de que ela seja diferente de zero apenas no intervalo $0 \leq t \leq 1$, então uma solução simples para (3.12) é a função retangular

$$\varphi(t) = \begin{cases} 1 & \text{se } 0 < t < 1 \\ 0 & \text{caso contrário} \end{cases}, \quad (3.29)$$

com apenas dois coeficientes diferentes de zero, $h(0) = h(1) = 1/\sqrt{2}$. De modo a satisfazer (3.23) e (3.24) a wavelet requerida é

$$\psi(t) = \begin{cases} 1 & \text{se } 0 < t < 0,5 \\ -1 & \text{se } 0,5 < t < 1 \\ 0 & \text{caso contrário} \end{cases}, \quad (3.30)$$

com apenas dois coeficientes diferentes de zero, $h_1(0) = 1/\sqrt{2}$ e $h_1(1) = -1/\sqrt{2}$.

O espaço gerado por $\varphi(t-k)$ é \mathcal{V}_0 . O próximo espaço em ordem crescente de resolução é \mathcal{V}_1 , gerado por $\varphi(2t-k)$, e que inclui \mathcal{V}_0 . À medida que se aumenta o valor de j , o espaço \mathcal{V}_j gerado por $\varphi(2^j t - k)$ torna-se mais adequado para aproximar funções ou sinais por funções constantes cada vez mais estreitas. Dessa forma, Haar demonstrou que, quando $j \rightarrow \infty$, $\mathcal{V}_j \rightarrow L^2$.

As funções de Haar estão ilustradas na Figura 3.7: a primeira coluna contém o pulso de largura unitária que gera \mathcal{V}_0 ; a segunda contém o pulso de largura igual à metade e sua respectiva translação, necessários para gerar \mathcal{V}_1 ; a terceira coluna contém quatro translações de um pulso de largura igual a um quarto, e assim por diante. Isso mostra claramente como o aumento da escala permite que uma representação de maior resolução de um sinal seja obtida.

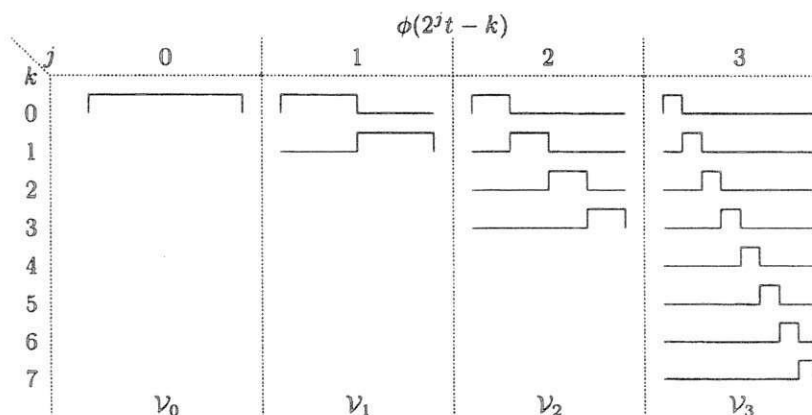


Figura 3.7: Funções de Haar de escalonamento que geram \mathcal{V}_j .

De modo a separar as componentes de alta e baixa resolução, é conveniente usar wavelets. Usando a decomposição ortogonal em \mathcal{V}_3

$$\mathcal{V}_3 = \mathcal{V}_2 \oplus W_2. \quad (3.31)$$

Ou seja, existem dois conjuntos de bases ortogonais que geram o espaço \mathcal{V}_3 : o primeiro, é composto pelas 8 funções de escalonamento para $j = 3$, mostrado na quarta coluna da Figura 3.7 e na primeira coluna da Figura 3.8; o segundo, apresentando as quatro funções de escalonamento para $j = 2$ e mais os detalhes contidos nas quatro wavelets para $j = 2$, totalizando também 8 funções de expansão, ilustradas na segunda coluna da Figura 3.8.

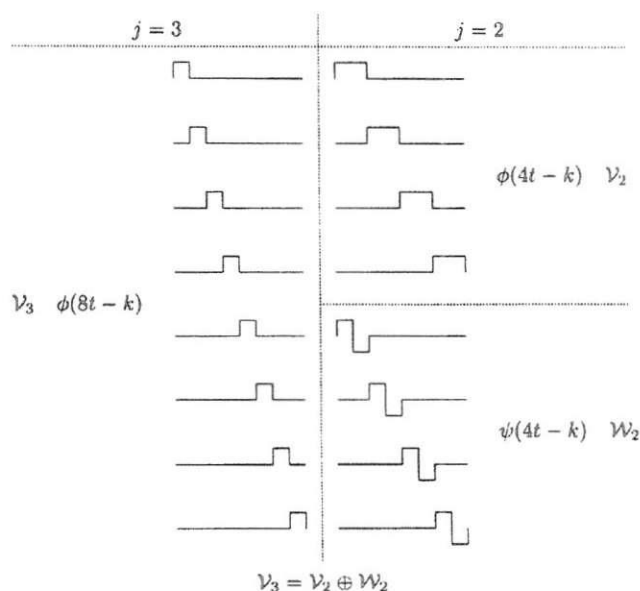


Figura 3.8: Funções de Haar de escalonamento e decomposição wavelet para \mathcal{V}_3 .

O espaço \mathcal{V}_2 pode ser também decomposto em

$$\mathcal{V}_2 = \mathcal{V}_1 \oplus W_1. \quad (3.32)$$

Por fim, \mathcal{V}_1 é decomposto em

$$\mathcal{V}_1 = \mathcal{V}_0 \oplus W_0. \quad (3.33)$$

Em outras palavras, decompõe-se seguidamente o espaço gerado pela função de escalonamento até que a função seja uma constante, a decomposição completa de \mathcal{V}_3 é obtida. Isso é mostrado simbolicamente na Figura 3.9.

O caráter multiresolucional da função de escalonamento e do sistema wavelet pode ser visto facilmente a partir da Figura 3.8, em que um sinal pertencente a \mathcal{V}_3 pode ser expresso em termos de uma soma de oito funções de escalonamento deslocadas na escala $j = 3$ ou uma soma de quatro funções de escalonamento e quatro wavelets deslocadas na escala $j = 2$. No segundo caso, as quatro funções de escalonamento fornecem uma aproximação grosseira para o sinal na escala $j = 2$ e as quatro wavelets contribuem para os “detalhes” mais finos na mesma escala, de modo que o conjunto equivale a uma representação

grosseira na escala $j=3$. As quatro funções de escalonamento na escala $j=2$ podem também ser decompostas em funções de mais baixa resolução, como mostrado em (3.32) e (3.33).

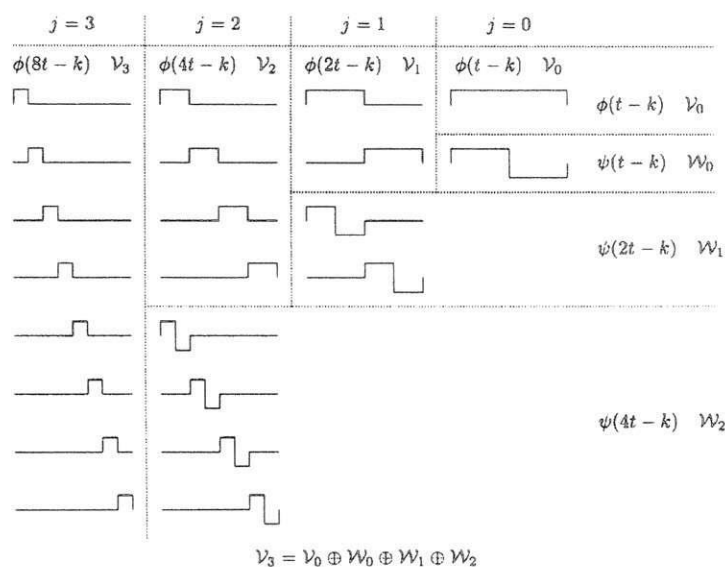


Figura 3.9: Funções de Haar de escalonamento e decomposição em wavelets para \mathcal{V}_3 .

A Figura 3.10 ilustra um exemplo de aproximação de uma função de teste em várias resoluções. O sinal de teste é composto de uma onda senoidal pura, que possui uma representação no domínio de Fourier perfeitamente localizada, e duas descontinuidades que são completamente localizadas no domínio do tempo. A aproximação em \mathcal{V}_0 é simplesmente a média da função de teste. À medida que se aumenta a escala, a aproximação torna-se cada vez mais próxima do sinal original.

3.6 A Transformada Wavelet Discreta

Conforme abordado anteriormente

$$L^2 = \mathcal{V}_{j_0} \oplus W_{j_0} \oplus W_{j_0+1} \oplus \dots \quad (3.34)$$

Usando-se (3.5) e (3.25), uma forma mais genérica da expansão (3.26) pode ser dada por

$$g(t) = \sum_k c_{j_0}(k) 2^{j_0/2} \phi_k(2^{j_0/2} t - k) + \sum_k \sum_{j=j_0}^{\infty} d_j(k) 2^{j/2} \psi_k(2^j t - k) \quad (3.35)$$

ou

$$g(t) = \sum_k c_{j_0}(k) \varphi_{j_0,k}(t) + \sum_k \sum_{j=j_0}^{\infty} d_j(k) \psi_{j,k}(k), \quad (3.36)$$

em que a escolha de j_0 determina a escala mais grosseira, cujo espaço é gerado por $\varphi_{j_0,k}$. O restante de $L^2(\mathbb{R})$ é gerado pelas wavelets, que fornecem os detalhes de resolução acima de $j=j_0$ do sinal [1]. Na prática, quando se possui apenas amostras do sinal e não o sinal em si, existe uma resolução máxima determinada pela taxa de amostragem.

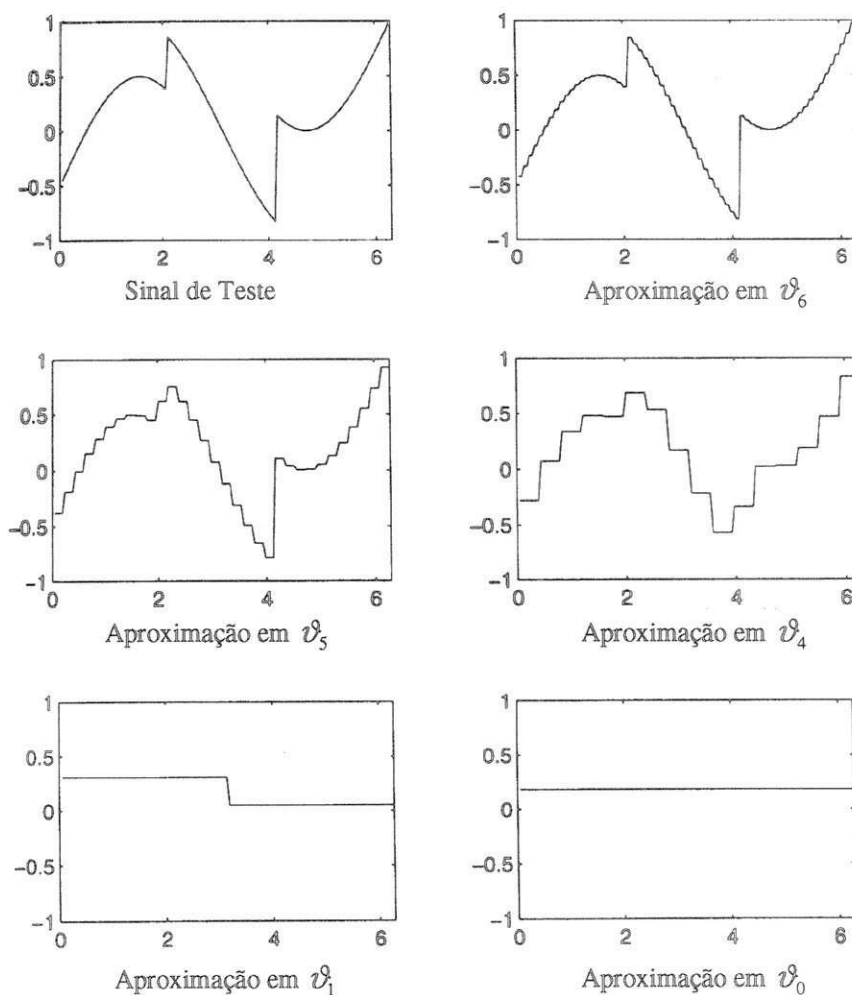


Figura 3.10: Aproximações em v_j por funções de Haar.

Os coeficientes da expansão (3.36) são denominados de *transformada wavelet discreta* (DWT, *Discrete Wavelet Transform*) do sinal $g(t)$. Se certas condições forem satisfeitas,

estes coeficientes descrevem completamente o sinal original e podem ser usados de maneira similar aos coeficientes da série de Fourier para análise, descrição, aproximação e filtragem. Se o sistema wavelet for ortogonal, os coeficientes podem ser calculados pelas expressões de produto interno (3.27) e (3.28).

Se a função de escalonamento for bem comportada, então, numa escala alta, o escalonamento é similar a uma função delta de Dirac e o produto interno simplesmente amostra a função. Em outras palavras, em resoluções suficientemente altas, as amostras do sinal são bastante próximas dos coeficientes de escalonamento.

A DWT é similar à série de Fourier mas é bem mais flexível e informativa. Ela pode ser periódica como a série de Fourier, de modo a representar sinais periódicos eficientemente. No entanto, ao contrário da série de Fourier, a DWT pode ser usada diretamente em sinais não-periódicos transientes com excelentes resultados.

3.7 Teorema de Parseval

Se as funções de escalonamento e wavelets formarem uma base ortonormal, existe um teorema de Parseval que relaciona a energia do sinal $g(t)$ com a energia em cada um de seus componentes e coeficientes wavelet [1]. O teorema afirma que

$$\int |g(t)|^2 = \sum_{l=-\infty}^{\infty} |c(l)|^2 + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} |d_j(k)|^2, \quad (3.37)$$

com a energia no domínio de expansão particionada no tempo por k e em escala por j . De fato, é esse particionamento no plano tempo-escala que descreve a DWT.

Já foi demonstrado [3,5] que é possível que a função de escalonamento e as wavelets apresentem suporte compacto (sejam diferentes de zero apenas sobre uma região finita) e sejam ortonormais. Isso é uma propriedade interessante, pois permite uma melhor localização no tempo das propriedades. Sabe-se que para a transformada de Fourier ortogonalidade e boa resolução tempo-frequência são incompatíveis [6], o que não é o caso da transformada wavelet. Outra propriedade importante é que para a grande maioria dos sinais, os coeficientes de expansão wavelet decaem rapidamente, e, dessa forma, o sinal pode ser representado eficientemente por uma pequena quantidade deles.

4. Bancos de Filtros Aplicados à Obtenção da DWT

Nesse capítulo será explicado como é feita a obtenção dos coeficientes $c(k)$ e $d_j(k)$ da DWT a partir de bancos de filtros discretos, que são uma ferramenta bastante conhecida e utilizada na área de processamento de sinais. Primeiramente, será mostrado como obter os coeficientes de menor resolução a partir dos de maior resolução (análise) e, em seguida, será feito o contrário (síntese). Algumas considerações sobre os coeficientes de entrada encerram o capítulo.

4.1 Análise – Da Maior para a Menor Resolução

Nesta seção, será derivada a relação entre os coeficientes de expansão numa menor escala em termos daqueles numa maior escala [1]. Partindo-se da Equação básica de recursão

$$\varphi(t) = \sum_n h(n)\sqrt{2}\varphi(2t-n), \quad (4.1)$$

e assumindo que exista uma solução única, escalona-se e desloca-se a variável de tempo, resultando em

$$\varphi(2^j t - k) = \sum_n h(n)\sqrt{2}\varphi(2(2^j t - k) - n) = \sum_n h(n)\sqrt{2}\varphi(2^{j+1} t - 2k - n). \quad (4.2)$$

Após a mudança de variável $m = 2k + n$ tem-se

$$\varphi(2^j t - k) = \sum_m h(m-2k)\sqrt{2}\varphi(2^{j+1} t - m). \quad (4.3)$$

Se ϑ_j for denotado por

$$\vartheta_j = \overline{\text{Span}_k \{2^{j/2} \varphi_k(2^j t - k)\}}. \quad (4.4)$$

então

$$f(t) \in \vartheta_{j+1} \Rightarrow f(t) = \sum_k c_{j+1}(k)2^{(j+1)/2}\varphi(2^{j+1} t - k) \quad (4.5)$$

pode ser expressa numa escala de $j+1$ apenas com funções de escalonamento e sem uso de wavelets. Para a resolução a uma escala menor, as wavelets são necessárias para o “detalhe” não contido na escala de j [1]. Dessa forma

$$f(t) = \sum_k c_j(k) 2^{j/2} \varphi(2^j t - k) + \sum_k d_j(k) 2^{j/2} \psi(2^j t - k), \quad (4.6)$$

em que os termos $2^{j/2}$ mantêm a norma unitária das funções da base nas várias escalas. Se $\varphi_{j,k}(t)$ e $\psi_{j,k}(t)$ são ortonormais, os coeficientes de escalonamento no nível j são calculados pelo produto interno

$$c_j(k) = \langle f(t), \varphi_{j,k}(t) \rangle = \int f(t) 2^{j/2} \varphi(2^j t - k) dt \quad (4.7)$$

que, usando (4.3) e trocando a ordem da integração e do somatório, pode ser escrita como

$$c_j(k) = \sum_m h(m-2k) \int f(t) 2^{(j+1)/2} \varphi(2^{j+1} t - m) dt. \quad (4.8)$$

Mas a integral é o produto interno com a função de escalonamento na escala $j+1$, de modo que

$$c_j(k) = \sum_m h(m-2k) c_{j+1}(m). \quad (4.9)$$

A relação correspondente para os coeficientes wavelet é

$$d_j(k) = \sum_m h_1(m-2k) c_{j+1}(m). \quad (4.10)$$

4.1.1 Filtragem e Subamostragem (ou Decimação)

No contexto de processamento digital de sinais, a “filtragem” de uma seqüência de números (o sinal de entrada) é feita pela convolução da seqüência com outro conjunto de números chamados de coeficientes do filtro, pesos ou resposta ao impulso [1]. Para uma seqüência de entrada $x(n)$ e resposta ao impulso $h(n)$, a seqüência de saída $y(n)$ é dada por

$$y(n) = \sum_{k=0}^{N-1} h(k) x(n-k). \quad (4.11)$$

Se o número de coeficientes N é finito, o filtro é denominado de filtro de Resposta ao Impulso Finita (FIR). Caso contrário, é denominado filtro de Resposta ao Impulso Infinita (IIR).

Duas operações básicas usadas em filtros digitais são a decimação (ou sub-amostragem) e a interpolação (ou super-amostragem). O decimador recebe o sinal $x(n)$ como entrada e produz como saída o sinal $y(n) = x(2n)$, como mostrado simbolicamente na Figura 4.1.

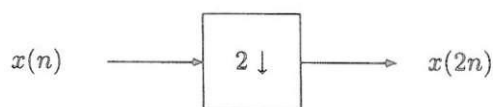


Figura 4.1 : Bloco Decimador.

Na decimação há a possibilidade de se perder informação pois metade dos dados é eliminada. No domínio da frequência, em se tratando da Transformada de Fourier, pode ser observado o efeito conhecido como *aliasing*, que diz respeito à mistura de componentes de frequência. Apenas se o sinal original for limitado em faixa é que não haverá perda de informação causada por decimação, satisfazendo as condições do Teorema da Amostragem.

As Equações (4.9) e (4.10) mostram que os coeficientes wavelet e de escalonamento em diferentes níveis de escala podem ser obtidos pela convolução dos coeficientes de expansão numa escala j pelos coeficientes de recursão invertidos no tempo $h(-n)$ e $h_1(-n)$, seguida da decimação do sinal resultante, de modo a fornecer os coeficientes de expansão para a escala $j-1$. Em outras palavras, os coeficientes na escala j são filtrados por dois filtros digitais FIR com coeficientes $h(-n)$ e $h_1(-n)$, e depois disso a decimação fornece os próximos coeficientes wavelet e de escalonamento de menor resolução [1].

O método descrito para o cálculo dos coeficientes é denominado de algoritmo de Mallat [4,7]. A implementação de (4.9) e (4.10) está ilustrada na Figura 4.2 em que as setas para baixo denotam decimação por dois, e os outros blocos correspondem à filtragem FIR, isto é, convolução com $h(-n)$ ou $h_1(-n)$. Para facilitar a notação, é utilizado $h_0(-n)$ ao invés de $h(-n)$ para denotar os coeficientes da função de escalonamento.

Pode ser mostrado [1] que o filtro implementado por $h(-n)$ é um filtro passa-baixas, enquanto que o implementado por $h_1(-n)$ é um passa-altas. Deve-se notar que o número médio de pontos (amostras de um sinal) na saída desse sistema é o mesmo que na entrada. O número é dobrado pela presença de dois filtros e então é reduzido à metade pela decimação, voltando à quantidade original. Isso significa que há a possibilidade de que não seja perdida informação no processo, de modo que o sinal seja completamente recuperado. Isto é, de fato, o que acontece. O *aliasing* existente no bloco superior pode ser desfeito ou cancelado usando o sinal do bloco inferior.

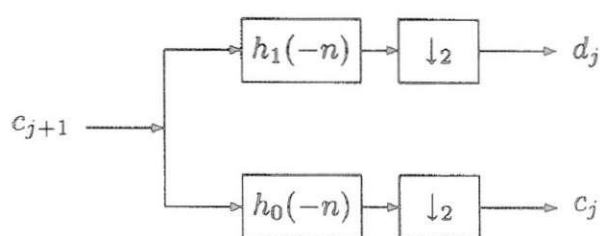


Figura 4.2: Banco de análise em duas bandas.

Esse processo de filtragem e decimação pode ser repetido nos coeficientes de escalonamento de modo a fornecer uma estrutura em duas escalas. A repetição do processo nos coeficientes de escalonamento é denominada “iteração do banco de filtros”. A Figura 4.3 ilustra uma decomposição em 3 níveis (3 estágios), resultante da iteração do banco de filtros.

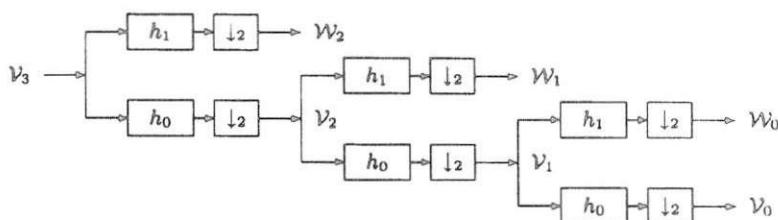


Figura 4.3 : Árvore de análise em duas bandas com 3 estágios.

A resposta em frequência de um filtro digital é a transformada de Fourier discreta no tempo da sua resposta ao impulso $h(n)$, dada por

$$H(\omega) = \sum_{n=-\infty}^{\infty} h(n)e^{i\omega n}. \quad (4.12)$$

O primeiro estágio do banco de dois filtros divide o espectro de $c_{j+1}(k)$ em uma banda passa-baixas e outra passa-altas, resultando nos coeficientes wavelet e nos coeficientes de escalonamento em uma escala menor, denotados, respectivamente, por $d_j(k)$ e $c_j(k)$. O segundo estágio então divide a banda passa-baixas em outra passa-baixas e uma passa-faixas. O primeiro estágio divide o espectro em duas partes iguais. O segundo divide a metade inferior em quartis e assim por diante. O resultado é um conjunto logarítmico de bandas como ilustrado na Figura 4.4.

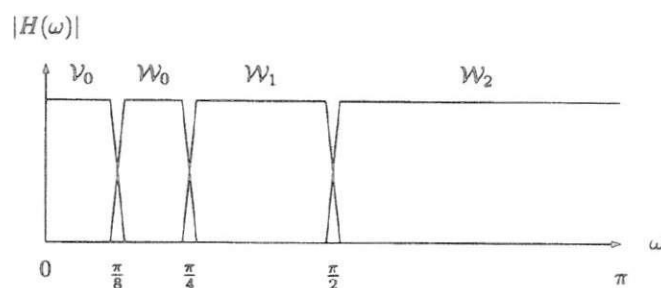


Figura 4.4 Bandas de frequência para a árvore de análise.

Para qualquer sinal limitado em faixa, existe uma escala superior $j = J$, acima da qual os coeficientes wavelet, $d_j(k)$, assumem valores muito pequenos [8]. Começando-se com uma descrição de alta resolução do sinal em termos dos coeficientes de escalonamento c_j , a árvore de análise calcula a DWT até a resolução $j = j_0$, através de $J - j_0$ estágios. Então, para $f(t) \in \mathcal{V}_J$, tem-se que

$$f(t) = \sum_k c_J(k) \varphi_{J,k}(t). \quad (4.13)$$

Fazendo a decomposição em termos de coeficientes de escalonamento na escala $j = j_0$ e dos coeficientes wavelet [1]:

$$f(t) = \sum_k c_{j_0}(k) \varphi_{j_0,k}(t) + \sum_k \sum_{j=j_0}^{J-1} d_j(k) \psi_{j,k}(k), \quad (4.14)$$

que é uma versão de (3.36) com número finito de escalas.

4.2 Síntese – Da Menor Resolução para a Maior

A reconstrução dos coeficientes originais em resolução fina do sinal pode ser feita a partir de uma combinação dos coeficientes wavelet e da função de escalonamento numa resolução mais baixa [1]. Considerando um sinal $f(t) \in \mathcal{V}_{j+1}$ (ou seja, definido no espaço de funções de escalonamento na escala $j+1$) pode ser representado em termos da função de escalonamento por:

$$f(t) = \sum_k c_{j+1}(k) 2^{(j+1)/2} \varphi(2^{j+1}t - k) \quad (4.15)$$

ou, em termos da escala j (o que requer wavelets),

$$f(t) = \sum_k c_j(k) 2^{j/2} \varphi(2^j t - k) + \sum_k d_j(k) 2^{j/2} \psi(2^j t - k). \quad (4.16)$$

A substituição de (4.1) e (3.23) em (4.16) resulta em

$$f(t) = \sum_k c_j(k) \sum_n h(n) 2^{(j+1)/2} \varphi(2^{j+1}t - 2k - n) + \sum_k d_j(k) \sum_n h_1(n) 2^{(j+1)/2} \psi(2^{j+1}t - 2k - n). \quad (4.17)$$

Como todas essas funções são ortonormais, multiplicando (4.15) e (4.17) por $\varphi(2^{j+1}t - k')$ e integrando, obtém-se os coeficientes:

$$c_{j+1}(k) = \sum_m c_j(m) h(k - 2m) + \sum_m d_j(m) h_1(k - 2m). \quad (4.18)$$

4.2.1 Filtragem e Interpolação

De modo contrário ao que é feito na operação de análise, a síntese no banco de filtros é feita com interpolação seguida de filtragem. Isso significa que a entrada para o filtro apresenta um zero inserido entre cada amostra do sinal a ser filtrado. Em outras palavras,

$$y(2n) = x(n) \text{ e } y(2n+1) = 0. \quad (4.19)$$

Como pode ser observado, a interpolação não resulta em perda de informação. Se for feita uma interpolação seguida de uma decimação, volta-se à situação inicial. Isso em geral não ocorre, porém, se a decimação for feita antes da interpolação.

A Expressão (4.18) mostra que a síntese é realizada mediante interpolação seguida de filtragem. Realiza-se interpolação da seqüência de coeficientes $c_j(k)$ na escala j , o que significa dobrar seu comprimento e inserir zeros entre cada termo. Em seguida, é processada a convolução com o vetor de escalonamento $h(n)$. O mesmo é feito com a seqüência de coeficientes wavelet na escala j , e os resultados são somados de modo a se obterem os coeficientes da função de escalonamento na escala $j+1$. Essa estrutura está mostrada na Figura 4.5, em que $g_0(n) = h(n)$ e $g_1(n) = h_1(n)$. É utilizada a notação $\uparrow 2$ para indicar uma interpolação.

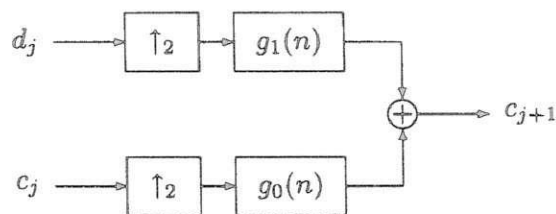


Figura 4.5: Banco de síntese para duas bandas.

Esse processo de combinação pode ser continuado até qualquer escala. A árvore resultante para o caso de duas escalas está mostrada na Figura 4.6.

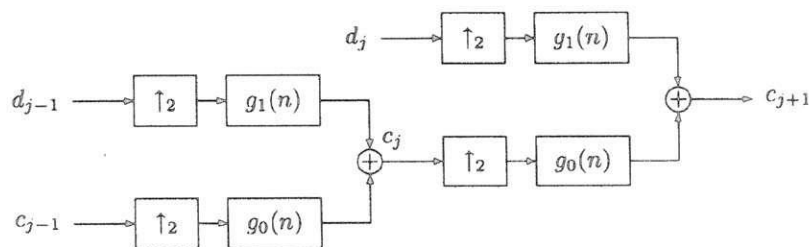


Figura 4.6: Árvore de síntese para duas bandas em dois estágios.

Sabe-se agora como obter os coeficientes de menor resolução da DWT a partir dos coeficientes de escalonamento de maior resolução. O problema agora é saber como o conjunto de entrada de coeficientes de escalonamento c_{j+1} (que irá gerar os demais

coeficientes) é obtido a partir do sinal de entrada para ser usado no sistema da Figura 4.2. Para uma escala alta o suficiente, as funções de escalonamento agem como “funções delta” cujo produto interno com o sinal de entrada (para calcular os coeficientes de alta escala) corresponde simplesmente à amostragem de $f(t)$. Se as amostras forem tomadas numa taxa acima da de Nyquist, elas são boas aproximações para os coeficientes de escalonamento naquela escala. Além disso, numa representação em escala alta, já foi visto que não são necessárias wavelets para representar o sinal.

O conjunto de operações de análise e síntese explicado nesse capítulo é conhecido como algoritmo de Mallat [4,7]. O banco de filtros de análise calcula eficientemente a DWT usando bancos de filtros discretos e decimadores, e o banco de filtros de síntese calcula a DWT inversa para reconstruir o sinal a partir da sua transformada.

Embora uma expansão wavelet pura seja possível, as propriedades da wavelet são melhor desenvolvidas e entendidas através da função de escalonamento. Isso é verdade especialmente se a função de escalonamento apresenta suporte compacto porque nesse caso a wavelet é composta de uma soma finita de funções de escalonamento, de acordo com (3.23).

Em situações práticas, em que a expansão ou a transformada wavelet é usada como ferramenta em processamento de sinais, pode-se tornar a expansão finita. Se as funções da base apresentam suporte compacto, apenas um número finito de adições em k é necessário. Se a função de escalonamento for incluída, como indicado em (3.26) ou (4.6), o limite inferior do somatório em j é finito. Se o sinal é essencialmente limitado em faixa, então há uma escala acima da qual existe pouca ou nenhuma energia e o limite superior em j torna-se finito também.

5. Quantização Vetorial

A Quantização Vetorial (QV) é uma técnica que vem demonstrando ser bastante adequada à compressão de sinais [10, 11]. Neste capítulo serão abordados, inicialmente, alguns conceitos básicos de compressão de sinais e justificativas para o uso da Quantização Vetorial dentro desse contexto. A seguir, é apresentada a formulação matemática da QV, bem como duas técnicas de projeto de dicionários: LBG e SSC. Por fim, serão discutidas algumas técnicas de avaliação de desempenho de sinais de voz e imagens submetidos à compressão.

5.1 Compressão Digital de Sinais

Num sistema de comunicações digitais, uma característica de grande interesse é a sua taxa de transmissão I . Ela é definida como o produto da taxa de amostragem f_s pelo número de bits por amostra do sinal, R . Então [9]:

$$I = f_s R \text{ bits por segundo (bit/s ou bps)} \quad (5.1)$$

Sabe-se que quanto maior for a taxa I , maiores serão os requisitos de largura de faixa do sistema. Este fato, aliado à necessidade de se diminuir os custos com memória para armazenamento, motivou o desenvolvimento de diversas técnicas de codificação, que buscam diminuir a quantidade de bits por amostra (R) necessária para representação das amplitudes.

O principal objetivo das técnicas de compressão de sinais, portanto, é a minimização dos requisitos de largura de faixa e de memória de armazenamento. Algumas das aplicações importantes que justificam a compressão de dados são [9, 10]: os sinais em canais de multiplexação em sistemas de faixa larga, a televisão de alta definição (HDTV), a telefonia móvel celular, redes digitais de serviços integrados (ISDN, *Integrated Services Digital Networks*), entre outras.

A Quantização Vetorial (QV) [10, 11] é uma técnica utilizada para compressão de sinais de voz e imagens que vem despertando bastante interesse dos pesquisadores, devido aos seguintes motivos:

- a Teoria da Distorção Versus Taxa, de Shannon, segundo a qual pode-se sempre obter um melhor desempenho codificando-se um vetor de amostras ao invés de cada amostra individualmente;
- possibilidade de implementação dos sistemas de codificação e decodificação em VLSI (*Very Large Scale Integration*);
- necessidade de sistemas flexíveis e eficientes para a representação dos mais variados tipos de informação, desde símbolos e textos até sons e imagens.

5.2 Quantização Vetorial

Matematicamente, a quantização vetorial consiste em um mapeamento Q de um espaço vetorial k -dimensional R^k em um subconjunto finito W de R^k [12]:

$$Q: R^k \rightarrow W, \quad (5.2)$$

em que $W = \{ \vec{w}_i \mid i = 1, 2, \dots, N \}$ é o conjunto de vetores de reprodução e N é o número de vetores em W ; cada \vec{w}_i é chamado vetor-código e W é chamado dicionário do quantizador vetorial.

O mapeamento é caracterizado pela partição $C = \{ C_i : \cup C_i = R^k, i = 1, 2, \dots, N \}$ de R^k ; de modo que um vetor de entrada k -dimensional real $\vec{x} \in C_i$ é associado ao vetor representante \vec{w}_i , ou seja,

$$C_i = \{ \vec{x} : Q(\vec{x}) = \vec{w}_i \}. \quad (5.3)$$

O vetor código \vec{w}_i pode ser visto como o rótulo de uma classe de padrões dos vetores de entrada que pertencem à célula de Voronoi C_i . Em outras palavras, cada \vec{w}_i deve ser escolhido de modo que seja o melhor vetor representativo para todos os vetores pertencentes à célula C_i , segundo algum critério de distância ou distorção mínima. A Figura 5.1 ilustra um exemplo de partição realizada no espaço R^2 .

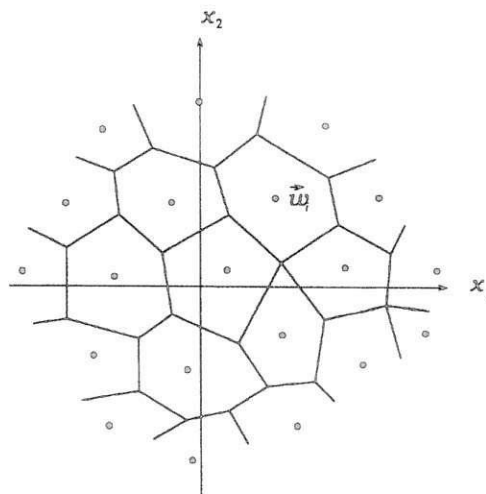


Figura 5.1: Partição em R^2 realizada pela quantização vetorial.

O processo de codificação de um sinal por quantização vetorial pode ser dividido em duas etapas. Na primeira etapa, para todo vetor de entrada \vec{x}_i , um vetor-código $\vec{w}_i \in W$ é selecionado como a representação de \vec{x}_i . Esse processo é denominado fase de quantização (ou fase de procura no dicionário), denotada por $Q(\vec{x}_i) = \vec{w}_i$. Em seguida, o vetor-código \vec{w}_i escolhido é representado por alguns símbolos (em geral, o endereço do vetor-código no dicionário) e transmitido ao longo do canal. Esta segunda etapa é a fase de codificação.

A decodificação, que ocorre no outro lado do canal, consiste em receber os símbolos transmitidos após a codificação e usá-los para selecionar os vetores-código do dicionário de reprodução, que é uma cópia de W .

O número médio de bits necessário para representar os símbolos na fase de codificação é chamado taxa do quantizador. Denomina-se distorção do quantizador o erro médio de quantização introduzido ao se representar o sinal de entrada X por sua versão quantizada $Q(X)$.

Para um sinal de voz, para a codificação utilizando-se quantização vetorial com um dicionário de N vetores k -dimensionais, a taxa do quantizador R é expressa por [12]:

$$R = \frac{1}{k} \log_2 N \text{ bit / amostra (bpa)}. \quad (5.4)$$

Para imagens, a taxa R é expressa em *bits/pixel* (bpp).

Pode-se conseguir uma redução da distorção do quantizador através de um aumento do número de vetores-código no dicionário. Isso, no entanto, segundo (5.4), leva a um aumento da taxa R (se a dimensão dos vetores for mantida fixa). Como há um compromisso entre taxa e distorção do quantizador, é muito importante projetarem-se dicionários ótimos, que minimizem a distorção média introduzida pela aproximação de cada vetor de entrada por um dos vetores-código. Para isso duas condições devem ser satisfeitas: a condição de vizinho mais próximo para a partição e a condição de centróide para os vetores-código [10, 7].

5.3 Projeto de Dicionários para Quantização Vetorial

Vários parâmetros devem ser considerados para o projeto de um quantizador vetorial, como [12]:

- taxa;
- qualidade;
- complexidade;
- custo.

A taxa está diretamente relacionada com a qualidade, como já foi visto. Ao se reduzir a taxa R , há uma queda de qualidade. Para uma mesma taxa, porém, o dicionário que apresenta maior dimensão k é o que possui melhor qualidade. Mas nesse caso, a complexidade (em termos do número de operações necessárias para a codificação) do dicionário também aumenta, já que ele é da ordem de 2^{kR} [12]. Com o aumento da complexidade, há uma elevação do custo.

Entre as várias técnicas existentes para o projeto de dicionários para quantização vetorial, podem ser citadas:

- algoritmo LBG, Linde-Buzo-Gray [14];
- rede neural de Kohonen [12,15, 16, 17];
- algoritmos competitivos de redes neurais [18];
- *stochastic relaxation* [13].

O algoritmo LBG é a técnica mais amplamente utilizada para o projeto de quantizadores, e será descrito a seguir.

5.3.1 Algoritmo LBG

O tradicional algoritmo LBG [10, 14] utiliza uma seqüência típica de dados para treinamento do dicionário, e consiste da seguinte seqüência de passos [12]:

- *Passo 1*- condição inicial: inicialize com qualquer configuração desejada. Vá para o *Passo 2* se a inicialização ocorreu com um conjunto de vetores-código (dicionário inicial); vá para o *Passo 3* se a inicialização ocorreu com a partição dos dados;
- *Passo 2*- particionamento: aloque cada dado (ou vetor de entrada) na respectiva classe segundo o critério do vetor código mais próximo;
- *Passo 3*- atualização do dicionário: compute os novos vetores-código como os centróides das classes de dados;
- *Passo 4*- teste de convergência: alterna-se *Passo 2* e *Passo 3* até a convergência do processo.

O algoritmo LBG, portanto, caracteriza-se por uma atualização iterativa do dicionário visando satisfazer as condições de centróide e vizinho mais próximo. Com isso a função distorção decresce monotonicamente, mas geralmente ela pode conter múltiplos mínimos locais [19]. Com isso o mínimo global normalmente não é atingido e os dicionários produzidos não são ótimos. Além disso, os dicionários levam geralmente a células de Voronoi que são subutilizadas (ou até mesmo vazias), e o desempenho do dicionário final é bastante dependente do dicionário/partição inicial [12].

Apesar das desvantagens mencionadas, ainda assim a técnica LBG é a mais utilizada, suas principais vantagens sendo a simplicidade de implementação, razoável velocidade de convergência, e o fato de não requerer informações do usuário além da condição de teste para a convergência.

A seguir, será descrito um outro algoritmo para projetos de dicionários de QV, denominado de SSC.

5.3.2 Algoritmo SSC

Algoritmos não-supervisionados de redes neurais vêm sendo utilizados com sucesso para o projeto de dicionários de QV. Dentre eles, um dos mais simples é o SSC [25] (*Synaptic Space Competitive*), cuja seqüência de passos é:

- *Passo 1-* Condição inicial: inicialize aleatoriamente um dicionário W com vetores k -dimensionais $\vec{\omega}_i$ ($1 \leq i \leq N$);
- *Passo 2-* Apresente o vetor de treino \vec{x} ;
- *Passo 3-* Encontre o vetor-código vencedor $\vec{\omega}_{i^*}$ (de acordo com o critério de distância mínima $d(\vec{x}, \vec{\omega}_{i^*}) < d(\vec{x}, \vec{\omega}_i), \forall i \neq i^*$);
- *Passo 4-* Atualiza-se o vetor-código $\vec{\omega}_{i^*}$ na direção \vec{x} , de acordo com a regra de treinamento:

$$\Delta\omega_{i^*j} = \eta(n) \cdot (x_j - \omega_{i^*j}), 1 \leq j \leq k . \quad (5.5)$$

Em (5.5), $\eta(n)$ é a taxa de aprendizado, n é a iteração atual, x_j é a j -ésima componente do vetor de treino \vec{x} , ω_{i^*j} é a j -ésima componente de $\vec{\omega}_{i^*}$, e $\Delta\omega_{i^*j}$ é a modificação introduzida na j -ésima componente do neurônio $\vec{\omega}_{i^*}$. Os passos de 2 a 4 são repetidos até que todos os vetores da seqüência de treinamento sejam utilizados. Após isso, realiza-se uma nova passagem da seqüência de treino, caso o número de iterações (especificado pelo usuário) não tenha sido ainda alcançado. A taxa de aprendizado $\eta(n)$ é reduzida em função da iteração n de alguma forma pré-estabelecida. Dessa forma ocorrem variações maiores dos vetores-código do dicionário nas primeiras iterações do que nas últimas.

A seguir serão discutidas algumas formas de avaliação de sinais de voz e imagem comprimidos.

5.4 Avaliação de Desempenho

Existem, basicamente, duas categorias de medidas de qualidade de desempenho: medidas objetivas e medidas subjetivas. As medidas objetivas são baseadas em comparações matemáticas feitas entre o sinal original e o processado [20]. Já as medidas

subjetivas baseiam-se em comparações entre os dois sinais, realizadas por grupos de pessoas.

5.4.1 Medidas Objetivas

As medidas objetivas são importantes na avaliação do desempenho em algoritmos de compressão por permitirem uma avaliação rápida, serem mais facilmente implementadas e com menor dispêndio de tempo (em relação às medidas subjetivas), além de poderem ser repetidas várias vezes na fase de desenvolvimento de um algoritmo [12]. Caracterizam-se por realizarem uma comparação direta entre os sinais original e processado. Para isso, devem apresentar duas características fundamentais. Primeiramente, devem ter significado subjetivo, de modo que pequenas e grandes distorções objetivas correspondam a avaliações subjetivas boas e ruins respectivamente. Segundo, elas devem ser amenas do ponto vista da análise matemática.

Duas medidas de distorção muito utilizadas em sistemas de compressão de voz são a relação sinal-ruído (*SNR*) e a relação sinal-ruído segmental (*SNR*). Elas serão descritas a seguir.

Relação Sinal-Ruído

A relação sinal-ruído (*SNR*, *Signal to Noise Ratio*) e suas variações são apropriadas apenas em sistemas em que se deseja que a forma de onda do sinal após o processamento seja a mais próxima possível da original [12].

Seja $x(n)$ o sinal original, $y(n)$ o sinal processado e $e(n) = x(n) - y(n)$ o sinal erro no instante de tempo n . A energia do sinal original é dada por:

$$E_x = \sum_n x^2(n). \quad (5.6)$$

A energia do sinal erro é:

$$E_e = \sum_n e^2(n) = \sum_n [x(n) - y(n)]^2. \quad (5.7)$$

A relação sinal-ruído, em *dB*, vale, portanto:

$$SNR = 10 \log_{10} \frac{E_x}{E_e} = 10 \log_{10} \frac{\sum_n x^2(n)}{\sum_n [x(n) - y(n)]^2}. \quad (5.8)$$

Relação Sinal-Ruído Segmental

A medida SNR , apesar da sua simplicidade, apresenta um problema: a igual ponderação de todos os erros no domínio do tempo [12]. Por este motivo, o efeito do ruído, que é maior nos segmentos de baixa amplitude (sons fricativos surdos), fica mascarado pelas altas energias nos segmentos de alta amplitude (segmentos vocais, sonoros), resultando num alto valor de SNR , que não reflete a realidade.

É necessário, portanto, uma medida de qualidade mais refinada, que é obtida tomando-se a média da SNR calculada em curtos intervalos (janelas) de tempo. Assim, é definida a relação sinal-ruído segmental:

$$SNR_{seg} (dB) = E[SNR(j)(dB)], \quad (5.9)$$

em que $SNR(j)$ é a SNR convencional para o segmento ou janela de tempo j . A SNR_{seg} é calculada por:

$$SNR_{seg} = \frac{1}{J} \sum_{j=0}^{J-1} 10 \log_{10} \left[\sum_{n=m_j-N-1}^{m_j} \frac{x^2(n)}{[x(n) - y(n)]^2} \right], \quad (5.10)$$

em que m_0, m_1, \dots, m_{J-1} são os instantes finais para as J janelas de tempo, de comprimento típico de 15 a 25 ms.

Relação Sinal Ruído de Pico ($PSNR$)

Embora sejam reportadas, freqüentemente, críticas à correlação da relação sinal-ruído de pico ($PSNR$) com critérios de avaliação subjetiva, a grande maioria das avaliações de desempenho de algoritmos de codificação de imagem utiliza a $PSNR$ como medida de qualidade objetiva. Ela é definida como 10 vezes o logaritmo na base 10 da razão entre o quadrado do valor de pico de amplitude de entrada (v_p^2) e o erro médio quadrático (MSE , *Mean Square Error*):

$$PSNR = 10 \log_{10} \left[\frac{(v_p^2)}{MSE} \right]. \quad (5.11)$$

Para o caso de imagens codificadas a 8 bpp,

$$PSNR = 10 \log_{10} \left[\frac{255^2}{MSE} \right]. \quad (5.12)$$

5.4.2 Medidas Subjetivas

Apesar das vantagens apresentadas pelas medidas objetivas de avaliação, as medidas subjetivas são imprescindíveis para uma avaliação definitiva de um sistema de processamento de voz, áudio ou imagem. Como utilizam um número relativamente grande de pessoas, no entanto, os métodos subjetivos são em geral bastante dispendiosos de tempo e de difícil implementação.

Um teste subjetivo de escuta ou visualização é baseado em opiniões individuais de cada pessoa participante no teste. Os avaliadores devem ser escolhidos entre pessoas de diferentes formações profissionais, e também possíveis usuários do sistema a ser avaliado. É recomendado que pelo menos 15 avaliadores não sejam especialistas na área do projeto. Cada sessão de avaliação não deve durar mais do que 30 minutos e deve começar mostrando ao avaliador alguns exemplos típicos de sinais com diferentes escalas de degradação, para aclimatá-lo [21].

Duas das técnicas de avaliação subjetivas mais usuais em sistemas de voz e imagem são os *testes de preferência* e os *testes de qualidade absoluta*, a serem descritos a seguir.

Testes de Preferência

Os testes de preferência consistem em comparações feitas entre pares. O resultado da avaliação pode ser dada em forma de um conceito para cada uma das possibilidades de comparação, ou seja:

- A- A qualidade do primeiro é melhor do que a do segundo;
- B- A qualidade do segundo é melhor do que a do primeiro;
- C- A qualidade de ambos não se distingue.

Teste de Qualidade Absoluta: Escore de Opinião Média (MOS)

Em sistemas de compressão de voz e imagem, uma medida de desempenho bastante utilizada denomina-se MOS (*Mean Opinion Score*). No teste MOS, cada avaliador atribui um escore de qualidade segundo a escala graduada ilustrada na Tabela 5.1 [20]. Após a realização de todos os testes, calcula-se a média aritmética dos escores obtidos e o valor MOS final é determinado.

Escore	Qualidade	Nível de Distorção
5	Excelente	Imperceptível
4	Boa	Apenas perceptível mas não incômodo
3	Razoável	Perceptível e pouco incômodo
2	Pobre	Incômodo
1	Ruim	Muito incômodo

Tabela 5.1: Escala para o teste MOS.

6. Simulações Realizadas e Resultados

6.1 Simulações Envolvendo Sinais de Voz

De modo a aplicar os conhecimentos adquiridos, foi realizada uma etapa prática durante o estágio, em que foi implementado e avaliado um sistema de compressão de forma de onda de voz baseado em quantização vetorial e DWT. Para isso, utilizou-se um conjunto de dicionários multiresolucionais para a codificação através da QV dos coeficientes wavelet obtidos a partir da aplicação de uma DWT em um sinal de voz.

Embora a combinação de DWT e QV já venha sendo utilizada há algum tempo no contexto da compressão de imagens, produzindo excelentes resultados, não se tem conhecimento de trabalhos em que essa técnica tenha sido utilizada para codificação de forma de onda de voz. Esse fato foi uma das principais motivações para esta fase do estágio.

Os resultados apresentados englobam a avaliação da qualidade do sinal reconstruído sob várias taxas de codificação, e também uma investigação da importância (contribuição) de cada nível de decomposição wavelet para a qualidade deste sinal.

O sistema de codificação utilizado está ilustrado na Figura 6.1. Numa primeira etapa do processamento é feita a DWT multiresolucional do sinal de voz original em L níveis de decomposição, gerando $L+1$ componentes: uma de aproximação, $a_L(n)$, e L de detalhe, $d_j(n)$, com $1 \leq j \leq L$. Em seguida, é realizada a quantização vetorial sobre os coeficientes da DWT, através de um banco de dicionários multiresolucionais, de modo que cada componente da decomposição é quantizada por um dicionário específico. Finalmente, é realizada uma IDWT (*Inverse DWT*) sobre as componentes de detalhe e de aproximação reconstruídas, obtendo-se o sinal reconstruído.

Como o número total de coeficientes em uma decomposição multiresolucional é o mesmo que a quantidade de amostras do sinal original, consequentemente a taxa de codificação do sinal também será igual à taxa de codificação dos coeficientes. Portanto, se o sinal original apresenta S amostras e são realizadas L níveis de decomposição, o número de coeficientes da componente de aproximação $a_L(n)$ será igual a $S/2^L$, enquanto que para cada componente de detalhe $d_j(n)$ será igual a $S/2^j$, $1 \leq j \leq L$.

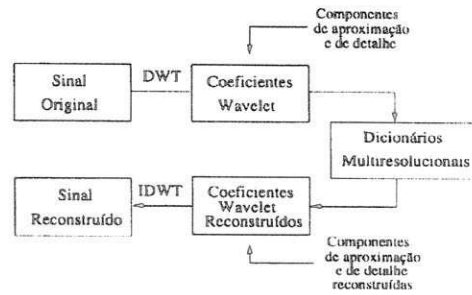


Figura 6.1: Diagrama de blocos do sistema de codificação baseado em DWT+QV

Se a taxa (em bits/coeficiente) do quantizador vetorial para $a_L(n)$ for R_{a_L} e para cada $d_j(n)$ for R_{d_j} , então o número total de bits (N_b) utilizado na codificação é

$$N_b = \frac{S}{2^L} R_{a_L} + \sum_{j=1}^L \frac{S}{2^j} R_{d_j}. \quad (6.1)$$

Portanto, a taxa de codificação total, R_T , em bits por amostra (que é igual à taxa em bits por coeficiente), é dada por

$$R_T = \frac{N_b}{S} = \frac{R_{a_L}}{2^L} + \sum_{j=1}^L \frac{R_{d_j}}{2^j}. \quad (6.2)$$

Um exemplo de esquema de alocação de taxa para cada componente de uma decomposição em 3 níveis está ilustrado na Tabela 6.1. Nesta tabela, os valores de N e K definem o número de vetores-código no dicionário e a dimensão dos vetores para o quantizador vetorial, respectivamente. A taxa para cada componente é calculada usando (5.4). A partir de (6.2) obtém-se que a taxa de codificação total para o exemplo é dada por:

$$R_T = \frac{R_{a_3}}{8} + \frac{R_{d_1}}{2} + \frac{R_{d_2}}{4} + \frac{R_{d_3}}{8}, \quad (6.3)$$

ou seja, após a substituição dos valores da Tabela (6.1) :

$$R_T = \frac{4}{8} + \frac{0,875}{2} + \frac{1,5}{4} + \frac{2}{8} = 1,5625 \text{ bits/amostra}. \quad (6.4)$$

Componente	N	K	Taxa (<i>bit/amostra</i>)
$a_3(n)$	256	2	$R_{a_3} = 4,0$
$d_3(n)$	256	4	$R_{d_3} = 2,0$
$d_2(n)$	64	4	$R_{d_2} = 1,5$
$d_1(n)$	128	8	$R_{d_1} = 0,875$

Tabela 5.1: Exemplo de alocação de taxa para as componentes da decomposição em 3 níveis.

6.1.1 Resultados

Para as simulações, foram utilizados sinais de voz codificados a 8,0 *bits/amostra* correspondentes a sentenças extraídas de um conjunto de frases foneticamente balanceadas da língua portuguesa [22]. Para o projeto dos dicionários multiresolucionais utilizou-se o algoritmo LBG, e para a DWT foram utilizadas as famílias Daubechies'6, Daubechies'10, Coif'3, Biortogonal4.4 e de Haar. Foram utilizados 3 níveis de decomposição wavelet multiresolucional em todos os casos. As estruturas piramidais de decomposição e síntese wavelet estão apresentadas na Figura 6.2.

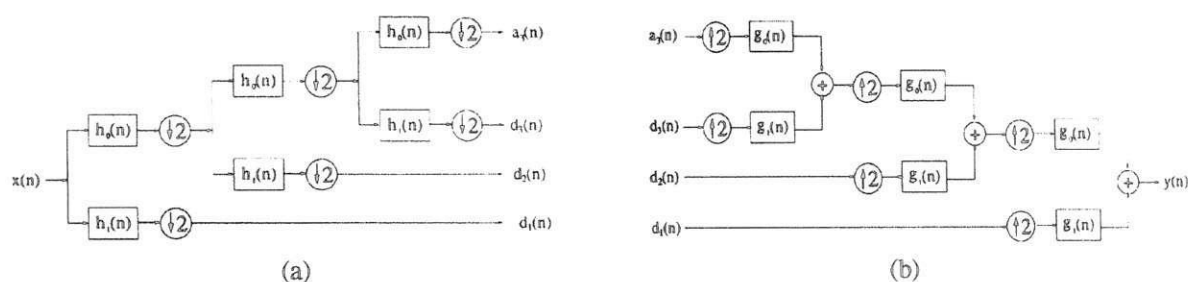


Figura 6.2: Estruturas de análise (a) e síntese (b) para uma decomposição wavelet.

Avaliou-se a qualidade dos sinais reconstruídos através da relação sinal-ruído segmental (SNR_{seg}) e de testes subjetivos de preferência.

Num primeiro conjunto de simulações, foi avaliada a importância da componente de aproximação e das de detalhe na qualidade do sinal reconstruído. Isso foi feito aplicando-se a IDWT sobre os coeficientes obtidos pela DWT em 3 níveis de decomposição com a exclusão de uma ou mais componentes. Os resultados de relação sinal-ruído segmental obtidos estão apresentados na Tabela 6.2.

De acordo com a Tabela 6.2, observa-se que $a_3(n)$ corresponde à componente mais importante para qualidade do sinal reconstruído, já que o menor valor de SNR_{seg} é obtido quando esse nível é excluído. Inversamente, $d_1(n)$ corresponde à componente de menor importância, pois sua exclusão resulta no maior valor de SNR_{seg} . Além disso, dentre as componentes de detalhe, verifica-se que a ordem decrescente de importância para a reconstrução é $d_3(n) > d_2(n) > d_1(n)$.

Tipo de reconstrução	SNR_{seg} (dB)
Excluindo-se $a_3(n)$	1,95
Excluindo-se $d_3(n)$	10,21
Excluindo-se $d_2(n)$	16,54
Excluindo-se $d_1(n)$	20,67
Excluindo-se $d_1(n)$ e $d_2(n)$	14,20
Excluindo-se $d_1(n)$ e $d_3(n)$	9,20
Excluindo-se $d_2(n)$ e $d_3(n)$	8,85
Excluindo-se $d_1(n)$, $d_2(n)$ e $d_3(n)$	8,13

Tabela 6.2: Valores de SNR_{seg} obtidos após a exclusão de uma ou mais componentes na reconstrução do sinal.

Os testes subjetivos realizados para a avaliação da importância das componentes de decomposição confirmaram as indicações dos testes objetivos. De fato, observou-se que a exclusão de $d_1(n)$ implica queda pouco perceptível na qualidade do sinal reconstruído em relação ao original. Essa é uma constatação interessante, pois $d_1(n)$ contém 50% dos coeficientes da DWT e, em consequência disso, sua taxa de codificação apresenta um peso elevado no cálculo da taxa total, conforme mostrado em (6.2).

Os testes subjetivos realizados com a exclusão das componentes $d_2(n)$ e $d_3(n)$ indicaram que, embora a inteligibilidade do sinal não fosse comprometida, havia um caráter metálico da voz no sinal reconstruído, mais acentuado na exclusão de $d_3(n)$ do que $d_2(n)$. Com a exclusão apenas de $a_3(n)$, o caráter metálico da voz torna-se ainda mais perceptível, embora a inteligibilidade seja ainda mantida.

Numa segunda etapa de simulações, foram avaliados diferentes esquemas de alocação de bits para os sinais resultantes da decomposição wavelet multiresolucional. Para isso foram projetados diversos dicionários multiresolucionais para QV, variando-se a dimensão K e o número de vetores-código N . As estratégias de alocação avaliadas estão apresentadas na Tabela 6.3, juntamente com os valores de SNR_{seg} obtidos para os sinais reconstruídos. A taxa utilizada para a codificação de $d_1(n)$ foi de 0,0 *bit/amostra* (ou seja, $d_1(n)$ foi excluída), o que é justificável pois, conforme mencionado anteriormente, a exclusão dessa componente causa queda muito pouco perceptível na qualidade do sinal reconstruído.

	$d_2(n)$	$d_3(n)$	$a_3(n)$	R_T (bpa)	SNR_{seg} (dB)
N	256	256	256		
K	2	2	2	2,0	14,84
R (bpa)	4,0	4,0	4,0		
N	256	256	256		
K	4	4	4	1,0	10,81
R (bpa)	2,0	2,0	2,0		
N	64	256	256		
K	4	4	4	0,875	10,39
R (bpa)	1,5	2,0	2,0		
N	64	64	256		
K	4	4	4	0,8125	9,39
R (bpa)	1,5	1,5	2,0		
N	256	64	256		
K	8	4	4	0,6865	9,54
R (bpa)	1,0	1,5	2,0		
N	256	128	128		
K	8	4	4	0,6875	8,69
R (bpa)	1,0	1,75	1,75		

Tabela 6.3: Resultados das estratégias de alocação de bits para as componentes $a_3(n)$, $d_2(n)$ e $d_3(n)$. A componente $d_1(n)$ foi excluída.

Na Tabela 6.4 são apresentados valores de SNR_{seg} para algumas taxas em sistemas de codificação de voz baseados em quantização vetorial simples no domínio original, sem uso de DWT. Observa-se, comparando-se as Tabelas 6.3 e 6.4, que os resultados de SNR_{seg} são bastante próximos nos dois sistemas. No entanto, a realização de testes de escuta indicou que, para as taxas investigadas, o sinal reconstruído usando-se DWT+QV apresenta qualidade subjetiva superior a obtida pelo uso de QV simples. Os testes de preferência

foram realizados com 20 pessoas. Em cada teste foram apresentados 2 sinais reconstruídos, um utilizando DWT+QV e outro apenas QV. O conceito atribuído foi “A” quando o sinal obtido por DWT+QV era melhor que o obtido por QV simples, e “B” caso contrário. Se os dois sinais apresentassem qualidade equivalente atribuía-se o conceito “C”.

R (bit/amostra)	K	N	SNR_{seg} (dB)
2,0	4	256	14,63
1,0	8	256	11,06
0,875	8	128	9,61

Tabela 6.4: Desempenho do sistema baseado em QV sem DWT para algumas taxas.

O melhor desempenho do esquema de codificação utilizando DWT+QV fica evidenciado nas 3 primeiras linhas da Tabela 6.5, em que os dois sinais comparados estão codificados a uma mesma taxa. Para as duas últimas linhas, a taxa no esquema DWT+QV é menor, e, mesmo assim, o seu desempenho é superior.

Taxa do sistema DWT + QV	Taxa do sistema QV simples	Conceito A	Conceito B	Conceito C
2,0	2,0	100%	0%	0%
1,0	1,0	95%	0%	5%
0,875	0,875	55%	10%	35%
0,875	1,0	55%	30%	15%
0,6875	1,0	50%	30%	20%

Tabela 6.5: Resultados dos testes de preferência entre os dois esquemas de codificação utilizados, utilizando-se o sinal da seqüência de treino. As taxas de codificação estão em bits por amostra.

Com relação ao desempenho fora da seqüência de treino, ou seja, quando o sinal a ser reconstruído é diferente daquele usado para projeto dos dicionários, a Tabela 6.6 mostra os resultados obtidos para a SNR_{seg} no sistema utilizando DWT+QV, enquanto que na Tabela 6.7 estão apresentados os resultados para o sistema utilizando QV simples.

R_T (bit / amostra)	SNR_{seg}
2,0	13,20
1,0	6,72
0,875	6,47
0,8125	5,95
0,6875	5,91

Tabela 6.6: Desempenho fora da seqüência de treino para o sistema DWT+QV.

R_T (bit / amostra)	SNR_{seg}
2,0	13,47
1,0	9,42
0,875	8,54

Tabela 6.7: Desempenho fora da seqüência de treino para o sistema QV simples.

Comparando-se os resultados das Tabelas 6.6 e 6.7, observa-se que o sistema baseado em QV simples apresenta desempenho superior em termos de relação sinal-ruído segmental. No entanto, conforme pode ser observado pela Tabela 6.8, os testes subjetivos indicam que a qualidade do sinal reconstruído utilizando DWT+QV é, na verdade, superior. Esses experimentos constataam, portanto, a inadequação da SNR_{seg} como medida objetiva de comparação de qualidade entre os dois esquemas de codificação investigados. É essencial e obrigatório, portanto, a realização de testes de escuta para uma completa avaliação dos codificadores.

Taxa do sistema DWT + QV	Taxa do sistema QV simples	Conceito A	Conceito B	Conceito C
2,0	2,0	95%	5%	0%
1,0	1,0	85%	15%	0%
0,875	0,875	65%	10%	25%
0,875	1,0	60%	20%	20%
0,6875	1,0	45%	40%	15%

Tabela 6.7: Resultados dos testes de preferência entre os dois esquemas de codificação utilizados, utilizando-se um sinal fora da seqüência de treino. As taxas de codificação estão em bits por amostra.

Os resultados apresentados até agora foram obtidos utilizando a família Daubechies'6 (Db6), mas foram também conduzidas algumas investigações envolvendo outras famílias: Daubechies'10 (Db10), Coiflet'3 (Coif3), Biortogonal4.4 (Bior4.4) e de Haar. A Tabela 6.8

apresenta os resultados de SNR_{seg} obtidos para essas famílias sob 3 taxas de codificação. Os resultados indicam um desempenho muito próximo das várias famílias, excetuado-se a família de Haar, que apresenta o pior desempenho. De fato, os testes de escuta realizados constataram diferenças praticamente imperceptíveis em termos de qualidade subjetiva entre as famílias Db6, Db10, Coif3 e Bior4.4. A família de Haar, por outro lado, apresentou uma queda de qualidade do sinal de voz bastante acentuada.

R_T (bpa)	Db6	Db10	Coif3	Bior4.4	Haar
2,0	14,84 dB	15,09 dB	15,15 dB	15,26 dB	12,16 dB
1,0	10,81 dB	10,83 dB	11,02 dB	11,11 dB	9,30 dB
0,875	10,39 dB	10,43 dB	10,59 dB	10,67 dB	9,40 dB

Tabela 6.8: Valores de SNR_{seg} obtidos utilizando famílias diferentes para DWT.

6.2 Simulações Envolvendo Imagens

O procedimento mais usual para se obter uma representação wavelet hierárquica de uma imagem foi introduzido por Mallat [4]. A DWT é implementada passando a imagem por um filtro passa-baixas L e um passa-altas H em ambas as direções, vertical e horizontal, e fazendo-se uma decimação por um fator 2 na saída. Como resultado, obtém-se as sub-bandas LH_1 , HL_1 , HH_1 , e LL_1 , no primeiro nível de resolução, correspondente à escala mais fina.. Para se obter as próximas sub-bandas, de menor resolução, o processo é repetido utilizando-se a sub-banda passa-baixas LL_1 , que gera as sub-bandas LH_2 , HL_2 , HH_2 , e LL_2 . O próximo nível de resolução é obtido a partir de LL_2 , e assim por diante. A Figura 6.3 mostra a decomposição wavelet de uma imagem em três níveis.

6.2.1 Resultados

O desempenho da codificação foi avaliado em imagens de 256x256 pixels, monocromáticas, 8 bpp (*bits/pixel*). O algoritmo utilizado para a quantização vetorial dos coeficientes foi o SSC. As imagens avaliadas foram: Lena, Airplane (F-16), Mandrill e Barbara. A medida de qualidade objetiva utilizada foi a relação sinal-ruído de pico (PSNR) [10].

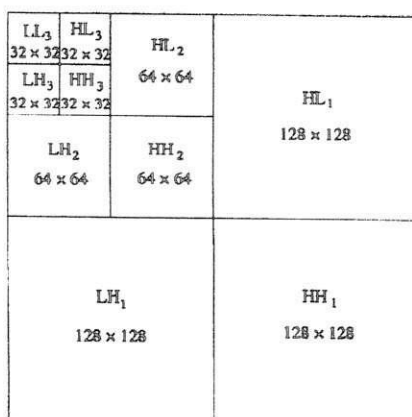


Figura 6.3: Decomposição wavelet de uma imagem 256x256 em três níveis.

Uma decomposição wavelet em 3 níveis foi realizada utilizando a família Daubechies'6. A taxa de codificação utilizada foi de 0,3125 bpp, e ela foi obtida codificando-se a sub-banda LL_3 a 8 bpp com quantização escalar, e para as demais sub-bandas no nível de resolução 3 (HL_3 , LH_3 e HH_3) foram utilizados dicionários SSC multiresolucionais de tamanho $N = 256$ e dimensão $K = 4$ (blocos 2×2); o nível de resolução 2 (HL_2 , LH_2 e HH_2) foi codificado utilizando dicionários SSC multiresolucionais de tamanho $N = 256$ e dimensão $K = 16$ (blocos 4×4). O nível de resolução 1 foi descartado no processo de reconstrução, por apresentar contribuição muito pequena para a qualidade das imagens [23].

A Tabela 6.9 mostra os valores de PSNR obtidos para as imagens utilizando dicionários multiresolucionais SSC, o algoritmo JPEG [24] e QV simples utilizando a técnica SSC. Os resultados indicam a superioridade da técnica baseada em wavelets/QV, o que foi confirmado através de testes subjetivos. A Figura 6.4 mostra que, de fato, a distorção introduzida pela técnica wavelet/QV é menos desagradável que a da JPEG ou da QV convencional.

Imagem	JPEG	QV Convencional	Wavelet/QV
Lena	27,72	25,60	29,02
Airplane	25,12	23,50	26,32
Mandrill	20,40	20,62	21,67
Barbara	22,29	21,77	22,32

Tabela 6.9: Resultados de codificação (PSNR) para uma taxa total de 0,3125 bpp.

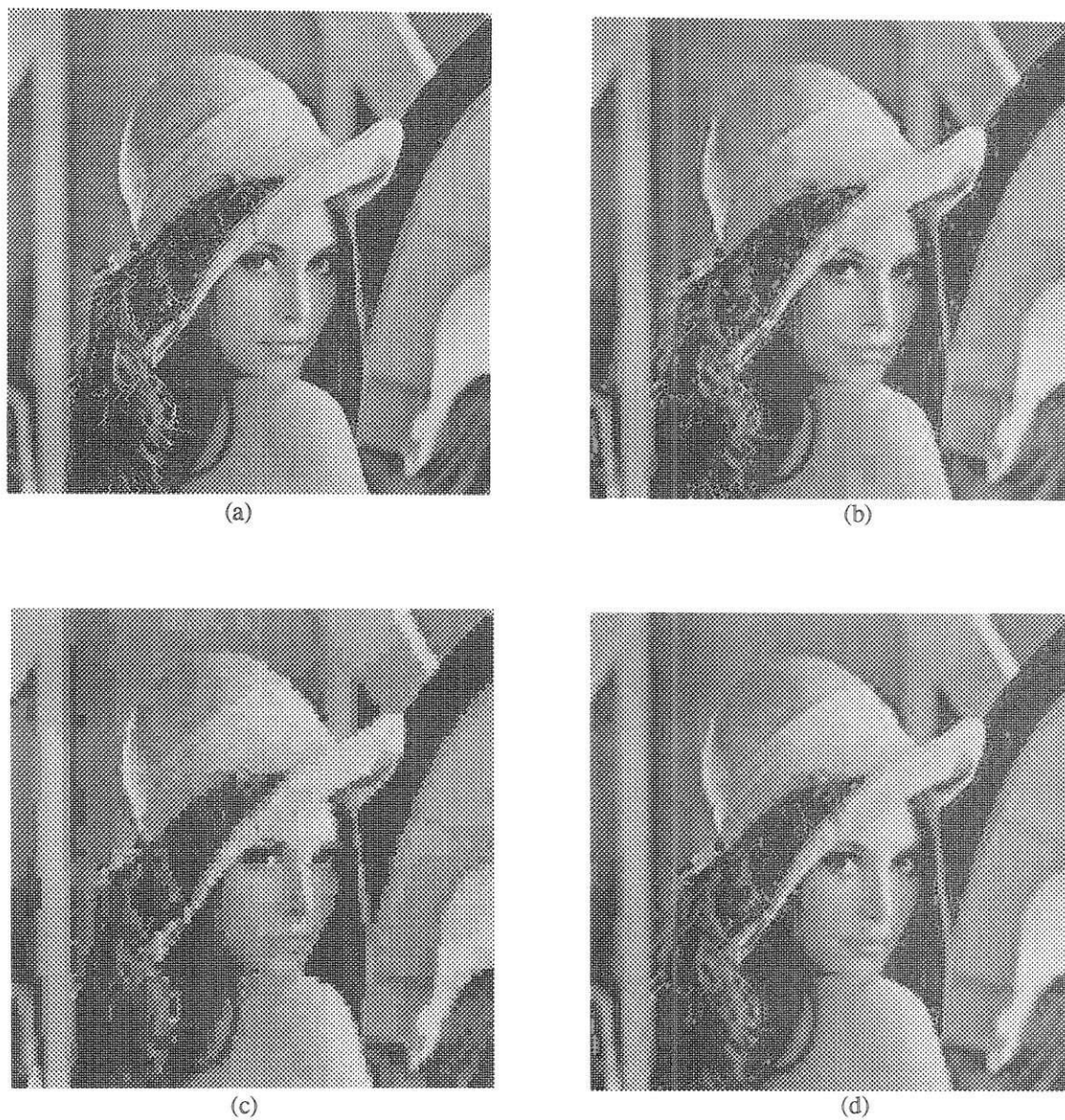


Figura 6.4: Imagens Lena: original (a) e codificadas a 0.3125 bpp utilizando as técnicas JPEG (b), QV simples (c), e Wavelet/QV (d).

7. Conclusões

O estágio foi plenamente satisfatório, pois contribuiu de maneira decisiva para a ampliação dos conhecimentos do aluno, além de permitir-lhe oferecer sua contribuição através das seções tutoriais deste relatório como também da seção de resultados obtidos. Dessa forma, as atividades realizadas estabelecem uma boa base para a continuação dos estudos através de uma pós-graduação, que será iniciada a partir de agora.

Todos objetivos traçados para todas as etapas do plano de estágios foram alcançados. Durante a realização do estágio foram, inclusive, incorporadas atividades que não estavam previstas de início, mas que se fizeram necessárias, o que sem dúvida enriqueceu o trabalho e ressalta a importância de se ser flexível e atentar para todas as possibilidades durante a resolução de um problema, características indispensáveis a um bom engenheiro. Neste contexto, merece menção especial a etapa do estágio referente à realização dos testes de preferência, que embora trabalhosa e dispendiosa de tempo, revelou-se extremamente gratificante tanto no aspecto profissional, por confirmar o melhor desempenho do sistema proposto, como também no aspecto pessoal, pela colaboração, participação e interesse demonstrados por todos os usuários do LAPS e colegas do PET.

Portanto, mais uma vez enfatiza-se que a realização do estágio é uma etapa de fundamental importância para a formação do engenheiro, seja qual for a sua especialidade. Isso porque apenas uma atividade como o estágio permite que o aluno tenha o primeiro contato com o mundo profissional, sendo desafiado a buscar sempre a melhor solução para os problemas que lhe são apresentados no dia-a-dia da sua profissão, de modo a melhor servir à sua comunidade, além de, sem dúvida, contribuir para a sua formação moral e intelectual.

Referências Bibliográficas

- [1] Burrus, C. S., Gopinath, R. A. and Guo, H. *Introduction to Wavelets and Wavelet Transforms*. Prentice-Hall, New Jersey, 1998.
- [2] Sweldens, Wim. Wavelets: what next? *Proceedings of the IEEE*, 84(4):680-685, April 1996.
- [3] Daubechies, I. Ten Lectures on Wavelets. SIAM, Philadelphia, PA 1992. Notes from the 1990 *CBMS-NSF Conference on Wavelets and Applications* at Lowell, MA.
- [4] Mallat, S. G. A theory of multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 11(7):674-693, July 1989.
- [5] Daubechies, I. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 41:909-996, November 1988.
- [6] Daubechies, I. The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*, 36(5):961-1005, September 1990. Also a Bell Labs Technical Report.
- [7] Mallat, S. G. Multiresolution approximation and wavelet orthonormal bases of L^2 . *Transactions of the American mathematical Society*, 315:69-87, 1989.
- [8] Gopinath, R. A., Odegard, J. E., and Burrus, C. S. On the correlation structure of multiplicity M scaling functions and wavelets. In *Proceedings of the IEEE International Symposium on Circuits and Systems*, pages 959-962, ISCAS-92, San Diego, CA, May 1992.
- [9] Jayant, N. S. and Noll, P. *Digital Coding of Waveforms*. Prentice-Hall, Englewood Cliffs, NJ, 1984.
- [10] Gersho, A. and Gray, R. M. *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, Boston, MA, 1992.
- [11] Gray, R. M. Vector Quantization. *IEEE ASSP Magazine*, pages 4-29, April 1984.
- [12] Junior, F. M. B. *Quantização Vetorial Aplicada a Sinais de Voz e Imagem*, Dissertação de Mestrado, DEE/UFPB, Março, 1998.
- [13] Zeger, K., Vaisey, J. and Gersho, A. Globally Optimal Vector Quantizer Design by Stochastic Relaxation. *IEEE Transactions on Signal Processing*, Vol. 40, No.2, pages 310-322, February 1992.

- [14] Linde, Y. and Buzo, A. and Gray, R. M. An Algorithm for Vector Quantizer Design. *IEEE Transactions on Communications*, Vol. COM – 28, No.1, pages 84-95, January, 1980.
- [15] Kohonen, T. Self-Organization and Associative Memory (3rd ed). Springer-Verlag, Berlin, 1989.
- [16] Kohonen, T. The Self-Organizing Map. *Proceedings of the IEEE*, Vol. 78, No.9, pages 1464-1480, September 1990.
- [17] Haykin, S. Neural Networks – A Comprehensive Foundation, IEEE Press, Englewood Cliffs – NJ, 1994.
- [18] Hertz, J and Krogh, A. and Palmer, R. G. Introduction to the Theory of Neural Computation. Addison-Wesley, Redwood City, CA, 1992.
- [19] Gray, R. M. and Karnin, E. D. Multiple Local Optima in Vector Quantizers. *IEEE Transactions on Information Theory*, Vol. IT-28, No.2, pages 256-261, Mar 1982.
- [20] Deller Jr., J. R., Proakis, J. G. and Hansen, J. H. L. *Discrete-time Processing of Speech Signals*. MacMillan Publishing Co., 1993.
- [21] Aguiar Neto, B. G. *Processamento e Transmissão Digital de Voz (Apostila)*. Curso de Pós-Graduação em Engenharia Elétrica, Campina Grande – PB, 1995.
- [22] Alcain, A., Solewicz, J. A. and Moraes, J. A. Frequência de Ocorrência dos Fones e Listas de Frases Foneticamente Balanceadas no Português Falado no Rio de Janeiro. *Revista da Sociedade Brasileira de Telecomunicações*, 1990.
- [23] Averbuch, A. Lazar, D. and Israeli, M. Image Compression Using Wavelet Transform and Multiresolution Decomposition. *IEEE Transactions on Image Processing*, Vol. 5, No.1, pages 4-15, January, 1996.
- [24] Wallace, G. K. The JPEG Still Picture Compression Standard. *Communications of the ACM*, Vol. 34, pages 30-44, April, 1991.
- [25] França, R. M. Aguiar Neto, B. G. Voice Waveform Vector Quantization Using a Competitive Algorithm.