

Leonardo Vidal Batista

Compressão de Sinais Eletrocardiográficos
Baseada na Transformada Cosseno Discreta

Tese submetida à Coordenação de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Campina Grande como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências no domínio da Engenharia Elétrica.

Área de Concentração: Processamento da Informação

Orientadores:

Elmar Uwe Kurt Melcher

Luis Carlos Carvalho

Campina Grande - PB, Brasil

28 de agosto de 2002



B333c Batista, Leonardo Vidal.
Compressão de sinais eletrocardiográficos baseada na transformada cosseno discreta / Leonardo Vidal Batista. - Campina Grande, 2002.
115 f.

Tese (Doutorado em Engenharia Elétrica) - Universidade Federal de Campina Grande, Centro de Ciências e Tecnologia, 2002.

Referências.
"Orientação : Prof. Dr. Elmar Uwe Kurt Melcher, Prof. Dr. Luis Carlos Carvalho".

1. Processamento de Sinais. 2. Eletrocardiograma. 3. Compressão de Dados. 4. Tese - Engenharia Elétrica. I. Melcher, Elmar Uwe Kurt. II. Carvalho, Luis Carlos. III. Universidade Federal de Campina Grande - Campina Grande (PB). IV. Título

CDU 621.391:616.12-073.97(043)

COMPRESSÃO DE SINAIS ELETROCARDIOGRÁFICOS BASEADA NA
TRANSFORMADA COSSENO DISCRETA

LEONARDO VIDAL BATISTA

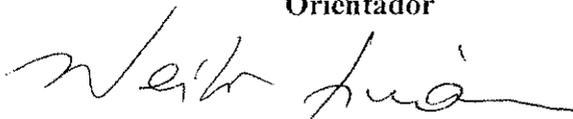
Tese Aprovada em 23.08.2002



PROF. ELMAR UWE KURT MELCHER, Dr., UFCG
Orientador



PROF. LUIS CARLOS CARVALHO, Dr., UFPB
Orientador



PROF. WEILER ALVES FINAMORE, Dr., CETUC-PUC-RIO
COMPONENTE DA BANCA



PROF. JURANDIR NADAL, Dr., UF RJ
COMPONENTE DA BANCA



PROF. JOÃO MARQUES DE CARVALHO, Ph.D., UFCG
COMPONENTE DA BANCA



PROF. GURDIP SINGH DEEP, Ph.D., UFCG
COMPONENTE DA BANCA

CAMPINA GRANDE – PB
Agosto - 2002

DEDICATÓRIA

A minha esposa, Jaqueline, e aos meus
filhos, Tatiana e Paulo Vitor.

AGRADECIMENTOS

Gostaria de externar minha gratidão a todos que ajudaram a compor este trabalho:

Aos Professores Elmar Uwe Kurt Melcher e Luis Carlos Carvalho, modelos de profissionalismo e competência, além de excelentes orientadores. Nossas incontáveis reuniões foram invariavelmente iluminadoras e estimulantes. Como se não bastasse, ainda me honraram com sua amizade.

Aos membros da Banca Examinadora da presente tese, pela inestimável contribuição na forma de sugestões e críticas sempre pertinentes, e pelas palavras estimulantes. De fato, ultrapassaram os officios próprios de membros de Banca Examinadora e ajudaram a moldar esta tese. Alguns exemplos: o Professor João Marques de Carvalho acompanhou o trabalho desde o início e participou de todos os meus seminários de Projeto e Pesquisa; o Professor Gurdip Singh Deep me emprestou um material importante e atual sobre compressão de dados; o Professor Weiler Alves Finamore, meu orientador no mestrado e professor em várias disciplinas, suscitou meu interesse em compressão de dados; o Professor Jurandir Nadal nos enviou o MIT-BIH *Arrhythmia Database* e um artigo sobre compressão de sinais de ECG, de sua autoria.

Ao Professor Mário Toscano de Brito Filho, pela minuciosa análise dos efeitos da compressão, pelo método aqui proposto, nos sinais de ECG.

Ao amigo JanKees van der Poel, pelos muitos artigos sobre compressão de sinais de ECG que me emprestou e pelas discussões sobre este fascinante assunto, que lançaram as sementes do presente trabalho.

Ao Professor Antonio Marcus Nogueira, coordenador da COPELE durante a maior parte do tempo em que por lá estive, sempre eficiente e prestimoso na resolução das raras pendências relacionadas ao curso e ao meu afastamento.

Ao Professor Benemar Alencar de Souza, pela eficiência na coordenação da COPELE.

Ao Professor Mirabeau Dias. Por ocasião do processo de afastamento das minhas atividades docentes para cursar o doutorado, Mirabeau, então chefe do Departamento de Informática, muito me ajudou e estimulou. Aos Professores Hamilton Soares da Silva e José Antonio Gomes de Lima, pelo estímulo e pelas ocasiões em que me ajudaram espontânea e desinteressadamente. Aos Professores Antonio Carlos Cavalcanti, Augusto Antonio Pinheiro Neto e Lucídio dos Anjos F. Cabral pelos auxílios prestados antes e durante o doutorado, pela confiabilidade e pela amizade. A todos os demais colegas do Departamento de Informática, pelo apoio.

Ao Professor Yaniv Zigel, que me enviou sua dissertação de mestrado sobre compressão de sinais de ECG.

Aos funcionários da COPELE, pela dedicação e eficiência e, em especial, a Ângela de Lourdes Ribeiro Matias, pedra angular da instituição. Aos funcionários do Departamento de Informática, do NETEB e da PRPG, sem os quais o bom desenvolvimento da vida profissional de um “afastado” seria impraticável.

Aos colegas da Pós-Graduação em Engenharia Elétrica da UFCG e, em especial, ao amigo Marcos Moraes, pela cooperação e estímulo.

À CAPES, pelo apoio financeiro.

Aos amigos de longa data Silas Batista Meneses e Sandra Valéria Alves Sales de Meneses, pelo apoio e compreensão.

Aos meus pais, Malaquias Batista Filho e Germana Vidal Batista, e a minha segunda mãe, Carmelita de Siqueira Lau, por tudo.

A minha mulher Jaqueline Brito Vidal Batista, pelo amor e pelo suporte que me permitiu produzir em meio a um turbilhão de três anos de idade chamado Tatiana. A minha filha Tatiana Brito Vidal Batista e ao meu filho Paulo Vítor Brito Vidal Batista, que tantas vezes me arrancaram das agruras da pesquisa para as diligências da paternidade. Devo a eles estas recargas periódicas na bateria mental.

RESUMO

A utilização extensiva de eletrocardiogramas digitais produz grande quantidade de dados, conferindo importância à aplicação de técnicas eficientes de compressão para reduzir os requisitos de armazenamento ou transmissão dos dados gerados. Eventuais distorções introduzidas pelo processamento devem ser cuidadosamente controladas, para que não conduzam a diagnósticos errôneos. O presente trabalho propõe soluções para o problema da compressão de eletrocardiogramas, apresentando um novo compressor baseado na quantização ótima com zona morta dos coeficientes da transformada cosseno discreta. No método proposto, o sinal a ser comprimido é particionado em blocos de tamanho fixo aos quais se aplica a transformada. Os blocos são quantizados usando-se um vetor de quantização e um vetor de limiares que são especificamente determinados para cada sinal. Estes vetores são definidos, via multiplicadores de Lagrange, de forma que a entropia estimada é minimizada para uma dada distorção no sinal reconstruído. A estratégia de otimização proposta é uma adaptação para eletrocardiogramas de uma técnica aplicada anteriormente à compressão de imagens. O algoritmo finaliza com uma etapa de codificação de entropia, para a qual são propostas cinco variantes, três das quais utilizando codificação aritmética e as demais utilizando codificação de Golomb-Rice. A PRD (*Percent Root-Mean-Square Difference*) foi adotada para mensurar a distorção nos sinais reconstruídos. Para avaliar o desempenho do compressor proposto, seções de dois minutos de todos os 96 registros do MIT-BIH *Arrhythmia Database* foram comprimidos em diferentes níveis de PRD e as razões de compressão (RC) correspondentes foram computadas. Traçados de sinais originais e processados são também exibidos para inspeção visual da qualidade das reconstruções. Foram efetuados testes para avaliar o efeito do tamanho dos blocos no compromisso entre RC e PRD. Com a variante de melhor desempenho do codificador de entropia, uma RC média de 11,3:1 foi atingida para PRD fixa em 2,5%. Experimentos com registros do MIT-BIH *Arrhythmia Database* usados na literatura permitiram comparações diretas com vários outros compressores. As comparações revelaram que o compromisso RC-PRD do método proposto é equivalente ou superior ao dos melhores compressores de ECG conhecidos.

ABSTRACT

The extensive use of digital electrocardiograms produces large amounts of data and thus efficient techniques to compress electrocardiographic signals are important in order to reduce the amount of data to be recorded or transmitted. Eventual distortions introduced by the compression process must be carefully controlled, so that they do not lead to erroneous diagnostics. The present work proposes some solutions to the problem of electrocardiographic signal compression, presenting a new electrocardiogram compressor based on optimum dead-zone quantization of discrete cosine transform coefficients. In the proposed method, the signal to be compressed is partitioned in blocks of fixed size, which are then transformed. Each DCT block is quantized using a quantization vector and a threshold vector specifically determined for each signal. These vectors are defined, via Lagrange multipliers, so that the estimated entropy is minimized for a given distortion in the reconstructed signal. The proposed optimization strategy is an adaptation for electrocardiograms of a technique previously applied for image compression. The algorithm ends up with an entropy coding stage, for which five variants are described. Three of these variants use arithmetic coding and the others use Golomb-Rice coding. The *Percent Root-Mean-Square Difference* (PRD) was adopted to measure the distortion in the reconstructed signals. To assess the performance of the proposed compressor, 2-minute sections of all 96 records of the MIT-BIH Arrhythmia Database were compressed at different PRD values, and the corresponding compression ratios (CR) were computed. Traces of original and processed signals are also shown, allowing visual inspection of the quality of the reconstruction. Furthermore, the effect of block size on the compromise between CR and PRD was empirically tested. An average CR of 11.3:1 was achieved for PRD equal to 2.5%, with the best-performance entropy coding variant. Experiments with records from the MIT-BIH Arrhythmia Database used in the literature allowed direct comparisons with various compressors. The comparisons revealed that the CR-PRD compromise of the proposed method is equivalent or superior to that of the best electrocardiogram compressors known.

SUMÁRIO

Dedicatória.....	ii
Agradecimentos	iii
Resumo.....	v
Abstract.....	vi
Sumário.....	vii
Índice de Figuras.....	ix
Índice de Tabelas	xi
Índice de Equações.....	xii
Lista de Abreviaturas	xiv
Capítulo 1 Introdução.....	1
1.1 Motivação.....	1
1.2 Objetivos	7
Capítulo 2 O Eletrocardiograma	8
2.1 Origem do eletrocardiograma	8
2.2 O ECG de 12 derivações.....	9
2.3 Morfologia do ECG	11
2.4 O ECG digital.....	13
2.5 O banco de sinais de ECG do MIT-BIH.....	14
Capítulo 3 Conceitos Básicos de Compressão de Dados.....	17
3.1 Medidas de distorção	20
3.2 Informação e entropia	22
3.3 Função taxa-distorção	23
3.4 Taxa-distorção operacional.....	23
Capítulo 4 Técnicas de Compressão de Dados	31
4.1 Modelagem.....	32
4.2 Codificação de entropia	36
4.3 Codificação por comprimento de seqüência.....	42
4.4 Quantização.....	43
4.5 Transformadas.....	48

4.6	O Paradigma Transformada + Quantização + Codificação	53
Capítulo 5	Desenvolvimentos Recentes em Compressão Baseada em DWT e DCT	56
5.1	Compressão Baseada em DWT.....	57
5.2	Compressão Baseada em DCT.....	61
5.3	Compressão com DCT e DWT combinadas	66
5.4	Conclusão.....	67
Capítulo 6	Metodologia	70
6.1	O compressor proposto	70
6.2	O descompressor.....	84
6.3	Metodologia da Avaliação Experimental.....	84
Capítulo 7	Resultados	87
Capítulo 8	Discussão e Conclusões	97
Referências	107

ÍNDICE DE FIGURAS

Figura 2.1 Derivações precordiais	11
Figura 2.2 Um ciclo típico de ECG.....	12
Figura 3.1 Um compressor/descompressor genérico	18
Figura 3.2 Plano $R-D$ operacional.....	24
Figura 3.3 Polígono convexo gerado por pontos operacionais.....	26
Figura 3.4 Região de busca da solução ótima.....	27
Figura 3.5 O ponto P_j , localizado fora da BCI, é ótimo dentro do intervalo de tolerância.....	28
Figura 3.6 Interpretação gráfica da minimização lagrangiana.....	29
Figura 4.1 Diagrama em blocos de um compressor/descompressor baseado em transformada + quantização + codificação.....	53
Figura 6.1 Diagrama em blocos do compressor/descompressor proposto.....	71
Figura 6.2 Histograma normalizado dos coeficientes 2 do registro 100/MLII, após quantização e mapeamento, e distribuição geométrica unilateral para comparação.....	80
Figura 6.3. Histograma dos coeficientes 1 do registro 100/MLII, após quantização. (a)Antes do mapeamento. (b) Após o mapeamento.....	81
Figura 6.4. Histograma dos coeficientes 5 do registro 100/MLII, após quantização. (a)Antes do mapeamento. (b) Após o mapeamento.....	81
Figura 7.1 RC média de A3 em função do tamanho do bloco L_b , para PRD de 1,5%, 2,0%, 2,5%, 3,0%	88
Figura 7.2. RC média de GR1 em função do tamanho do bloco L_b , para PRD de 1,5%, 2,0%, 2,5%, 3,0%.....	88
Figura 7.3. RC média atingida por GR2 em função do tamanho do bloco L_b , para PRD de 1,5%, 2,0%, 2,5%, 3,0%.....	89
Figura 7.4. RC média atingida por GR2 em função da PRD.....	91
Figura 7.5. (a) Detalhe do registro 100/MLII; (b) reconstrução com PRD = 2,5%, RC = 12,1:1	91
Figura 7.6. (a) Seção do registro 100/MLII; reconstruções com (b) PRD = 1,5%, RC = 5,9:1; (c) PRD = 2,0%, RC = 8,1:1; (d) PRD = 2,5%, RC = 10,7:1; (e) PRD = 3,0%, RC = 12,6:1; (f) PRD = 4,0%, RC = 18,1:1; (g) PRD = 5%, RC = 21,8:1.....	92

Figura 7.7. (a) Seção do registro 107/V1; (b) reconstrução com PRD = 2,5%, RC = 9,9:1 93

Figura 7.8. (a) Seção do registro 108/V1; (b) reconstrução com PRD = 2,5%, RC = 12,1:1 93

Figura 7.9. (a) Seção do registro 109/MLII; (b) reconstrução com PRD = 2,5%, RC = 11,6:1 ... 94

ÍNDICE DE TABELAS

Tabela 4.1. Exemplos de códigos de Golomb e de Golomb-Rice.	40
Tabela 7.1. RCs médias, máximas e mínimas, e desvios padrões das RCs obtidas com A2, A3, GR1 e GR2 para PRDs de 1,5%, 2,0% , 2,5% e 3,0%.	89
Tabela 7.2. Ganhos de compressão das variantes propostas para PRD = 2,5%.	90
Tabela 7.3. Comparação entre as PRDs produzidas por GR2 e por vários compressores da literatura.	95
Tabela 7.4. PRDs (%) geradas pelos métodos CAB/2D-DCT, 1D-DCT, e A3.	96

ÍNDICE DE EQUAÇÕES

Equação	
(2.1)	10
(2.2)	10
(3.1)	18
(3.2)	20
(3.3)	20
(3.4)	20
(3.5)	20
(3.6)	21
(3.7)	21
(3.8)	21
(3.9)	22
(3.10)	22
(3.11)	22
(3.12)	22
(3.13)	23
(3.14)	25
(3.15)	26
(3.16)	29
(3.17)	29
(3.18)	30
(4.1)	34
(4.2)	34
(4.3)	34
(4.4)	34
(4.5)	34
(4.6)	34
(4.7)	35
(4.8)	35
(4.9)	38

(4.10).....	38
(4.11).....	38
(4.12).....	39
(4.13).....	39
(4.14).....	39
(4.15).....	40
(4.16).....	44
(4.17).....	44
(4.18).....	45
(4.19).....	46
(4.20).....	46
(4.21).....	47
(4.22).....	49
(4.23).....	50
(4.24).....	50
(4.25).....	51
(4.26).....	51
(4.27).....	51
(4.28).....	51
(6.1).....	72
(6.2).....	72
(6.3).....	72
(6.4).....	74
(6.5).....	74
(6.6).....	74
(6.7).....	75
(6.8).....	75
(6.9).....	75
(6.10).....	75
(6.11).....	75
(6.12).....	76
(6.13).....	76

LISTA DE ABREVIATURAS

ADPCM:	adaptive differential pulse code modulation
ASEC:	analysis by synthesis ECG compressor
AZTEC:	amplitude zone time epoch coding
BCI:	borda convexa inferior
CAB:	cut and align beats
D:	distorção
DCI:	desfibriladores cardíacos implantáveis
DCT:	<i>discrete cosine transform</i>
DPCM:	differential pulse code modulation
DWT:	<i>discrete wavelet transform</i>
ECG:	eletrocardiograma
EEG:	eletroencefalograma
EGIC:	eletrograma intracardíaco
EZDCT:	<i>embedded zerotree DCT</i>
EZW:	<i>embedded zerotree wavelet</i>
GWAVQ:	<i>gold washing adaptive vector quantization</i>
IID:	independente e identicamente distribuído
JPEG:	<i>joint photographic experts group</i>
KLT:	<i>Karhunen-Loeve transform</i>
LTP:	<i>long-term prediction</i>
MLII:	<i>modified limb lead II</i>
MPEG:	<i>moving pictures experts group</i>
PRD:	<i>percent root-mean-square difference</i>
QUMC:	quantizador uniforme, <i>midtread</i> , reconstrução central
QUMO:	quantizador uniforme, <i>midtread</i> , reconstrução ótima
QURO:	quantizador uniforme, <i>midriser</i> , reconstrução ótima
QVA:	quantização vetorial adaptativa
QZMC:	quantizador com zona morta, reconstrução central
R:	taxa de bits

RC: razão de compressão
RLE: *run length encoding*
SAPA: *scan-along polygonal approximation*
SPIHT: *set partitioning in hierarchical trees*
TCW: terminal central de Wilson
WT: *wavelet transform*

Capítulo 1

INTRODUÇÃO

1.1 MOTIVAÇÃO

A coleta de sinais biológicos através de polígrafos permite a realização de uma série de exames e experimentos fisiológicos e farmacológicos. Os polígrafos são analógicos por natureza e, tipicamente, registram em papel os dados coletados. Estes registros são posteriormente analisados por um profissional, que se encarrega de extrair parâmetros importantes para o experimento em questão, o que frequentemente constitui uma tarefa árdua, repetitiva e propensa a erros. Para sinais de eletroencefalogramas, por exemplo, longos registros em papel, atingindo até 60 metros de comprimento, são comuns (Pradhan e Dutt, 1994).

A conversão dos dados para o formato digital permite a utilização de computadores para a extração automática ou semi-automática dos parâmetros relevantes; o emprego de uma ampla variedade de técnicas de processamento digital de sinais, como a filtragem para redução de ruídos de origens diversas; o armazenamento em meios óticos ou magnéticos; e a busca e recuperação eficiente das informações de interesse. Além disso, tornou-se rotineira a transmissão de sinais digitais, muitas vezes através de enlaces de baixa velocidade (Furht e Perez, 1988). Como destacam Husoy e Gjerde (1996), após a coleta do sinal, pode ser desejável armazenar os registros do paciente para posterior comparação e avaliação; armazenar os registros em bancos de dados para pesquisa e ensino; transmitir o registro por enlaces de dados para avaliação remota; ou armazenar o registro em dispositivos portáteis de baixa capacidade de armazenamento.

Segundo Womble *et al.* (1977), já em 1977 a Biblioteca Eletrocardiográfica Central da USAFSAM (*United States Air Force School of Aerospace Medicine*) possuía mais de 800.000 eletrocardiogramas (ECGs) arquivados, com mais de 100 novos registros em microfilmes e papel sendo adicionados diariamente à biblioteca. Womble *et al.* (1977) relatam a necessidade de conversão dos dados para formato digital, com vistas a automatizar a obtenção de parte das medidas extraídas dos traçados pelos cardiologistas, tornar os dados mais acessíveis, e reduzir os problemas causados por ruídos na transmissão rotineira de ECGs através de linhas telefônicas, na forma de sinais de FM.

Os sistemas de registro computadorizado de sinais biológicos costumam gerar elevadas quantidades de dados (Lai, 1993), impondo dificuldades para o armazenamento ou transmissão dos resultados. Um eletrocardiógrafo digital de 12 derivações, operando com valores de 12 bits e taxa de amostragem de 500 Hz, produz mais de 30 Mbytes por hora. O eletroencefalógrafo do Departamento de Neurologia do Hospital Santa Clara, em Trento, Itália, é capaz de operar com 32 canais simultâneos, a uma taxa de amostragem máxima de 1 KHz, com 8 bits por amostra, podendo assim produzir até 110 Mbytes de dados por hora (Antoniol e Tonella, 1997). A configuração típica para exames de rotina emprega 20 canais, 128 Hz de taxa de amostragem e 8 bits por amostra, o que resulta em um fluxo de aproximadamente 9 Mbytes por hora.

Obviamente, quanto maior a quantidade de dados gerada, maiores os custos com armazenamento e transmissão. Especialmente críticos são os registros de longa duração. Os dispositivos portáteis denominados *Holters* são capazes de adquirir e armazenar eletrocardiogramas ou eletroencefalo-

gramas (EEG) por períodos que se estendem tipicamente por 24 ou 48 horas. Tais exames são necessários quando os sintomas são esporádicos, tornando difícil a detecção dos indícios clínicos em registros breves. Os aparelhos de Holter devem apresentar dimensões reduzidas e baixo consumo de energia elétrica, para que possam ser transportados pelos pacientes (Antoniol e Tonella, 1997). Os sinais de EEG podem ser usados, por exemplo, para localizar a origem das descargas espasmódicas em pacientes epilépticos candidatos a cirurgia para remoção dos focos de espasmos. Como os sintomas de epilepsia são, muitas vezes, largamente espaçados, pode ser necessário estender o registro do EEG por vários dias (Sirne *et al.*, 1999). Mais recentemente, surgiram também dispositivos portáteis não-invasivos para registro de longa duração de sinais de pressão sanguínea contínua, como o Portapres (Graaf *et al.*, 1997).

Alguns dispositivos, como os desfibriladores cardíacos implantáveis (DCI) e os *smart cards*, possuem capacidade de armazenamento reduzida, em comparação com a quantidade de dados de interesse para as aplicações a que se destinam. Os DCIs destinam-se a detectar e tratar arritmias potencialmente perigosas (Coggins e Jabri, 1999). Para tanto, estes dispositivos monitoram a atividade elétrica cardíaca através de sensores posicionados diretamente sobre a parede interna do coração, aplicando estímulos elétricos quando condições anormais são detectadas. O sinal colhido para monitorização é conhecido como eletrograma intracardíaco (EGIC). O EGIC usa tipicamente apenas uma derivação ventricular e, em alguns casos, uma derivação atrial adicional. Dada a limitada capacidade de armazenamento dos DCIs, apenas alguns poucos batimentos podem ser registrados durante os episódios arrítmicos, o que limita substancialmente a informação disponível ao cardiologista para diagnóstico e ajuste da operação do dispositivo. Em geral, um registro de aproximadamente 5 a 30 s é efetuado para cada evento isolado que ocasiona a ativação do DCI, e a capacidade total de armazenamento situa-se na faixa de 5 a 10 minutos (Crawford, 1999).

A inclusão de alguns sinais, como o ECG e o EEG, em *smart cards*, juntamente com outras informações a respeito do paciente, tem recebido atenção nos últimos anos (Madhukar e Murthy, 1993). Entretanto, o elevado custo de armazenamento em *smart cards* permite apenas a inclusão de registros muito breves e com qualidade reduzida, o que limita a utilidade e, em consequência, a aplicação rotineira desta prática.

Todos estes fatos mostram a importância de utilizar algum método de compressão de dados que mantenha as características essenciais do registro original. Antoniol e Tonella (1997) relatam que determinadas técnicas de compressão permitiram o desenvolvimento de um sistema capaz de transmitir sinais de EEG em tempo real, da residência do paciente para o hospital, através de linhas telefônicas. O registro é coletado por uma enfermeira e os médicos podem se comunicar com o paciente por telefone durante a coleta.

Um esquema de compressão *sem perdas* permite a reconstrução perfeita do sinal original a partir do sinal comprimido, devendo ser aplicado quando qualquer alteração nos dados originais é inaceitável, como na compressão de textos ou arquivos executáveis. Na compressão sem perdas, portanto, a informação é integralmente preservada, procurando-se unicamente reduzir a redundância na representação dos dados. Por outro lado, na compressão de imagens estáticas e vídeos, por exemplo, pode-se aplicar compressão *com perdas*, onde parte da informação original é eliminada com o objetivo de obter uma maior compressão. Com a perda de informação, os dados originais não podem ser perfeitamente restaurados, ocorrendo distorções na reconstrução. Para sinais biológicos utilizados para diagnósticos, aceita-se algum grau de distorção em determinadas situações, mas é essencial que as características importantes para o diagnóstico sejam preservadas. Nestes casos, a reconstrução perfeita é uma característica altamente desejável mas não obrigatória (Lai, 1993; Antoniol e Tonella, 1997; Koski, 1997).

O ECG é um dos sinais fisiológicos mais utilizados para diagnóstico e monitorização. A enorme quantidade de eletrocardiogramas coletados diariamente em todo o mundo confere especial interesse à compressão deste tipo de sinal. Qualquer processo de redução de dados de ECG deve preservar as características relevantes do registro original, uma vez que alterações na forma ou amplitude das ondas e complexos ou na duração dos intervalos e segmentos que o compõem podem conduzir a diagnósticos errôneos.

As técnicas baseadas em transformada, quantização e codificação de entropia têm sido extensamente estudadas e aplicadas, principalmente para compressão com perdas de sinais biológicos, áudio, imagens e vídeo. A popularidade de tais métodos procede do bom compromisso entre o grau de compressão e a distorção resultante. Contudo, as dificuldades inerentes à definição de uma estratégia de quantização adequada levam com frequência a abordagens demasiadamente simplificadas, com conseqüente degradação na relação entre compressão e distorção. A etapa de

codificação consiste geralmente em algum tipo de codificação de comprimento de seqüência (*Run Length Encoding* - RLE) associado a um codificador aritmético ou de Huffman. Algumas estratégias de quantização e codificação serão apresentadas no Capítulo 4.

As transformadas cosseno discreta (*Discrete Cosine Transform* – DCT) e *wavelet* discreta (*Discrete Wavelet Transform* - DWT) são as mais comumente empregadas em compressores modernos. Por exemplo, o JPEG (*Joint Photographic Experts Group*) (Wallace, 1991), padrão ISO para compressão de imagens estáticas de tom contínuo, emprega a DCT, enquanto que o JPEG-2000 (Marcellin *et al.*, 2000; Usevitch, 2001) utiliza a DWT. O MPEG (*Moving Pictures Experts Group*) (Hans, 1998), padrão internacional de vídeo (imagens em movimento e áudio) utiliza uma versão modificada da DCT (*Modified Discrete Cosine Transform*, MDCT). Dentre os compressores de ECG, pode-se citar o SPIHT-ECG (*Set Partitioning in Hierarchical Trees Algorithm*) (Lu *et al.*, 2000), baseado na DWT, e o CAB (*Cut and Align Beats*) com DCT bidimensional (Lee e Buckley, 1999).

Durante a década de 80 e a primeira metade da década de 90, a DCT foi a transformada mais aplicada em compressão de sinais. A existência de algoritmos eficientes para calculá-la e sua capacidade de concentrar a energia do sinal em poucos coeficientes são alguns dos motivos para a popularidade desta transformada. A DCT tende a apresentar desempenho ótimo de compressão para sinais estacionários Markov-1 quando o coeficiente de correlação entre amostras adjacentes tende a 1 (Rao e Yip, 1990). Na prática, observa-se bom desempenho mesmo em situações em que o sinal desvia-se consideravelmente deste modelo. A DCT tem demonstrado bom desempenho em um amplo espectro de aplicações, incluindo sinais de áudio, vídeo, videofone, imagens estáticas, HDTV (*High Definition Television*), imagens SAR (*Synthetic Aperture Radar*), sinais biológicos, etc.

Em 1993, Shapiro observou uma estruturação inerente nos coeficientes da DWT, explorando-a com grande sucesso no compressor EZW (*Embedded Zerotree Wavelet*) (Shapiro, 1993). Os excelentes resultados obtidos pelo EZW impulsionaram o interesse pela DWT na área de compressão de dados. De acordo com Usevitch (2001), o EZW marca o início da moderna compressão baseada em *wavelet*, visto que os desenvolvimentos subseqüentes incorporam conceitos fundamentais do EZW. O algoritmo SPIHT (Said e Pearlman, 1996), por exemplo, estendeu as idéias introduzidas no EZW, obtendo o melhor compromisso taxa-distorção dentre todos os métodos

de compressão de imagens propostos até o presente momento. Os compressores EZW e SPIHT exploram a estruturação dos coeficientes *wavelet* através de uma abordagem adequada para as etapas de quantização e codificação.

Há atualmente um justificado interesse em DWT para compressão de dados. Entretanto, como ressaltam Laurence e Monro (1997), o sucesso das abordagens baseadas em *wavelets* deriva do emprego de técnicas de quantização e codificação que exploram a organização observada nos coeficientes da transformada, e não necessariamente de uma superioridade inerente à DWT.

Zou e Gallagher (1994) através de inspeção visual das reconstruções efetuadas por dois métodos básicos de compressão de ECG, sendo um deles baseado em DCT e o outro em DWT, concluem que, para uma mesma razão de compressão, a qualidade visual dos sinais reconstruídos é minimamente superior na versão baseada em DWT. Mais especificamente, estes autores afirmam que a DWT é mais eficiente na decomposição dos complexos QRS, enquanto que a DCT é superior na representação das ondas P e segmentos ST dos sinais, que variam mais suavemente no tempo. Com base nesta constatação, Zou e Gallagher propõem um esquema que combina DCT e DWT, e concluem que uma melhor qualidade visual é obtida com o esquema proposto do que com a aplicação isolada das transformadas.

Xiong *et al.* (1996), apresentaram uma adaptação para DCT da idéia básica de Shapiro, concluindo que o esquema resultante, conhecido como EZDCT (*Embedded Zerotree DCT*), é superior ao EZW. Davis e Chawla (1997) descrevem um procedimento para otimizar, independentemente da transformada empregada, o processo de partição em árvores de zero do EZW. O procedimento foi utilizado em um compressor de imagens de baixa complexidade baseado em DCT. Sem utilizar codificação de entropia, os resultados do método de Davis e Chawla foram superiores aos do JPEG (Wallace, 1991) e aos da versão sem codificação de entropia do EZDCT de Xiong *et al.* (1996).

Uma outra adaptação do EZW para DCT, apresentado por Laurence e Monro (1997), atingiu resultados muito próximos àqueles obtidos com o algoritmo SPIHT. Também significativos são os resultados de uma outra modificação do EZW para DCT, também denominado EZDCT (Monro e Dickson, 1997). O desempenho do EZDCT de Monro e Dickson é semelhante ao do EZW, desde compressão sem perdas até níveis de distorção elevados.

Ratnakar (1997) partiu do esquema básico do JPEG e desenvolveu uma estratégia de quantização mais elaborada para os coeficientes DCT. Mesmo sem adaptar para a DCT as idéias inovadoras do EZW, Ratnakar reporta resultados semelhantes aos dos melhores compressores baseados em *wavelets*.

Uma das maiores autoridades mundiais em *wavelets*, Wim Sweldens, do AT&T *Bell Laboratories*, discorreu sobre algumas ameaças ao desenvolvimento adequado dos campos de pesquisa relacionados a *wavelets* (Sweldens, 1996). Dentre eles, cita as promessas excessivamente otimistas do potencial das *wavelets*, ocasionadas pelo modismo exagerado. Ainda segundo Sweldens, as afirmativas feitas por proponentes da transformada *wavelet*, inclusive na área de compressão, nem sempre se concretizam, de forma que deveria haver uma maior precaução com tais assertivas.

Os parágrafos anteriores sugerem que as etapas de quantização e codificação podem ser fatores mais importantes para o grau de compressão do que a opção entre DCT e DWT. O bom desempenho da DCT em diferentes problemas de compressão e o comportamento muito próximo ao ótimo verificado por diversos estudos teóricos (Rao e Yip, 1990), reduzem as possibilidades de se obter ganhos substanciais de compressão com outras transformadas.

1.2 OBJETIVOS

O presente trabalho tem como principal objetivo desenvolver um novo método de compressão de ECG baseado em DCT, priorizando-se a preservação da qualidade dos sinais reconstruídos. Serão investigadas estratégias cuidadosamente elaboradas para a quantização dos coeficientes da transformada e para a codificação da seqüência de coeficientes quantizados, procurando obter-se uma relação satisfatória entre o grau de compressão e a distorção entre os sinais originais e reconstruídos. Pretende-se mostrar empiricamente que esta abordagem produz resultados semelhantes ou superiores aos dos melhores métodos de compressão de ECG descritos na literatura.

Capítulo 2

O ELETROCARDIOGRAMA

Muitas anomalias cardíacas podem ser diagnosticadas através do eletrocardiograma. Alguns aspectos básicos da gênese e registro dos sinais eletrocardiográficos serão discutidos a seguir.

2.1 ORIGEM DO ELETROCARDIOGRAMA

Uma fibra muscular em repouso encontra-se em um estado polarizado, com carga negativa em seu interior e positiva no exterior. Quando estimulada por meios mecânicos, químicos ou elétricos, ocorre na região excitada um fenômeno denominado despolarização, que consiste em trocas iônicas súbitas entre os meios interno e externo, com uma conseqüente inversão de polaridade que se propaga ao longo de toda a fibra. Ao final da despolarização, ocorre a repolarização, com

novas trocas iônicas entre os meios interno e externo da fibra, restaurando-se o estado de repouso.

A propagação da despolarização pode ser tratada como um dipolo deslocando-se no sentido da propagação, com o pólo positivo à frente. Analogamente, a propagação da onda de repolarização corresponde a um dipolo deslocando-se no sentido da propagação, com o pólo negativo à frente (Guyton, 1984; Carvalho, 1998).

No coração, a excitação sincronizada de um grande número de fibras musculares gera um campo elétrico cujos efeitos podem ser medidos na superfície do corpo. O eletrocardiograma, ou ECG, é o registro da evolução temporal das diferenças de potencial decorrentes do campo elétrico gerado durante os ciclos de despolarização e repolarização das fibras cardíacas, captadas por eletrodos colocados na superfície do corpo. Os valores medidos dependem da posição dos eletrodos em relação ao sentido do fluxo dos dipolos, da distância dos eletrodos até estes, e da intensidade das correntes elétricas associadas ao fluxo dos dipolos.

2.2 O ECG DE 12 DERIVAÇÕES

Os sistemas padronizados de distribuição dos eletrodos sobre a superfície do corpo para o registro do ECG são conhecidos como *derivações eletrocardiográficas*. O eletrocardiograma padrão possui 12 derivações, das quais oito são linearmente independentes, distribuídas em três sistemas descritos nas seções seguintes.

2.2.1 Derivações clássicas do plano frontal

Nas derivações clássicas do plano frontal, também denominadas de derivações de Einthoven ou derivações *standard* do plano frontal, as diferenças de potencial são medidas como se segue:

- Derivação I (D1): o terminal negativo do eletrocardiógrafo é ligado ao braço direito e o positivo ao esquerdo.

- Derivação II (D2): o terminal negativo do eletrocardiógrafo é ligado ao braço direito e o positivo à perna esquerda.
- Derivação III (D3): o terminal negativo do eletrocardiógrafo é ligado ao braço esquerdo e o positivo à perna esquerda.

Da lei das tensões de Kirchhoff, decorre que, em qualquer instante, a tensão elétrica medida pela derivação D2 é igual à soma das tensões medidas pelas derivações D1 e D3:

$$D2 = D1 + D3 \quad (2.1)$$

A Equação (2.1) mostra que uma das derivações é redundante; pode-se empregar duas delas e calcular a terceira. Muitos eletrocardiógrafos medem D1 e D2 e computam D3 (Tompkins, 1993).

2.2.2 Derivações unipolares aumentadas do plano frontal

No sistema de derivações unipolares aumentadas do plano frontal, dois membros são conectados por intermédio de um resistor ao terminal negativo do eletrocardiógrafo, sendo o terminal positivo ligado ao terceiro membro. As três derivações obtidas são:

- Derivação aVR: terminal positivo no braço direito.
- Derivação aVL: terminal positivo no braço esquerdo.
- Derivação aVF: terminal positivo na perna esquerda.

Pode-se mostrar que as três derivações unipolares aumentadas do plano frontal podem ser calculadas a partir das derivações D1, D2 e D3 (Tompkins, 1993). Tem-se, por exemplo, a seguinte relação entre aVL, D1 e D3:

$$aVL = \frac{D1 - D3}{2} \quad (2.2)$$

2.2.3 Derivações precordiais

As seis derivações precordiais exploram o plano horizontal ou transverso. Neste sistema, o braço direito, o braço esquerdo e a perna esquerda são conectados por intermédio de resistências, a um único ponto, denominado *terminal central de Wilson* (TCW). O terminal negativo do eletrocardiógrafo é ligado ao TCW e os terminais positivos são posicionados em seis pontos bem definidos no precórdio, originando as derivações V1, V2, V3, V4, V5 e V6. A Figura 2.1 mostra a distribuição das derivações precordiais.

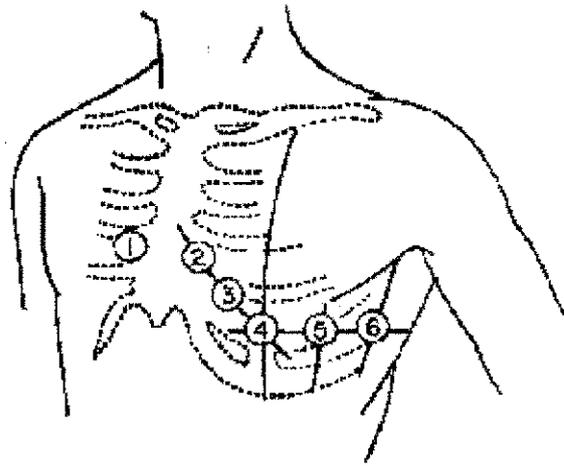


Figura 2.1 Derivações precordiais

2.3 MORFOLOGIA DO ECG

O ECG é normalmente composto por uma onda P; um complexo QRS, formado pelas ondas Q, R e S; e uma onda T. Pode haver ainda, após a onda T, uma onda de baixa amplitude, denominada onda U, cuja origem não é perfeitamente clara. A Figura 2.2 apresenta um ciclo típico de um traçado de ECG, juntamente com suas partes constituintes.

As relações temporais entre os segmentos, intervalos e ondas que constituem o ECG, além da morfologia das ondas, são dados importantes para um diagnóstico correto. Assim, as técnicas de compressão de sinais de ECG devem preservar as características fundamentais do sinal original. Uma descrição sumária das ondas, segmentos e intervalos mais importantes do ECG, suas dura-

ções e amplitudes típicas, para as derivações clássicas do plano frontal, são fornecidas a seguir (Guyton, 1984).

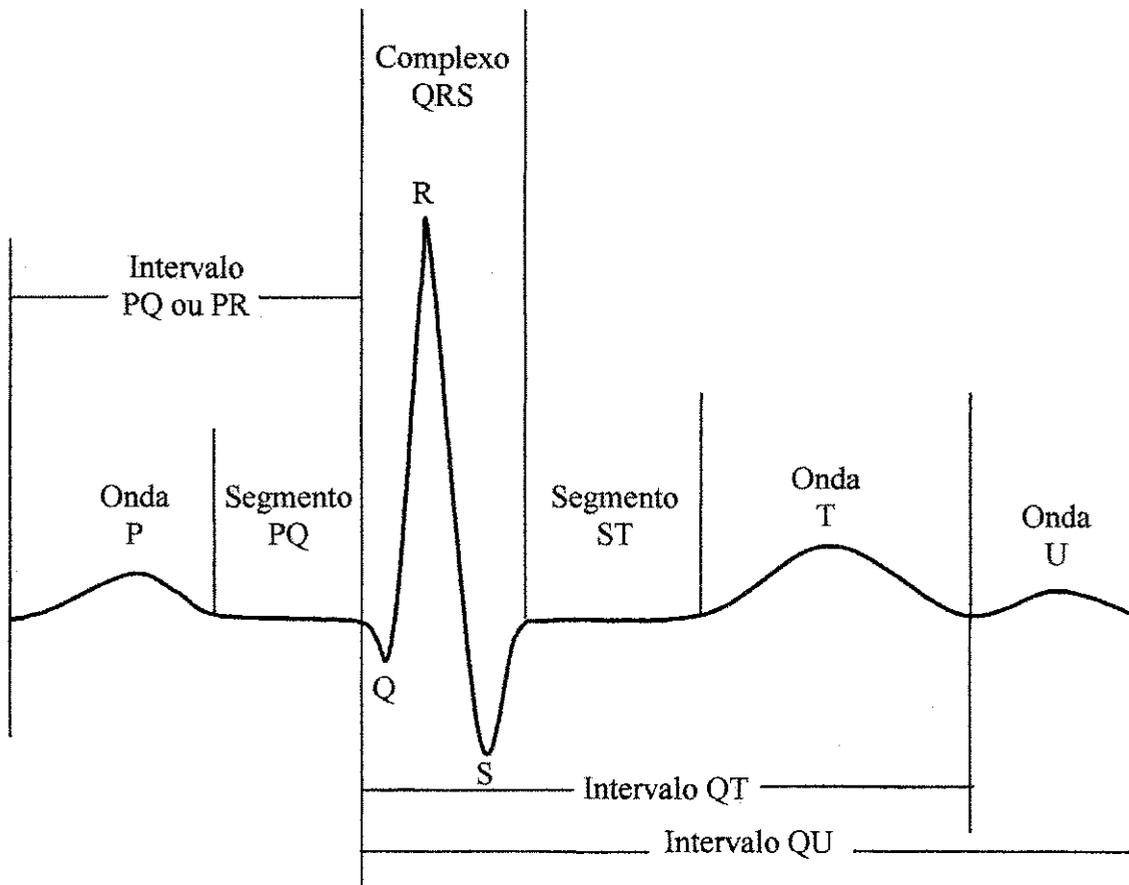


Figura 2.2 Um ciclo típico de ECG

- **Onda P.** De aspecto arredondado, representa a despolarização atrial, durando de 0,05 a 0,10 s, com amplitude entre 0,1 e 0,3 mV.
- **Intervalo PQ ou PR.** Vai do início da onda P ao início do complexo QRS, e representa o tempo entre o início da contração atrial e o início da contração ventricular. Dura de 0,12 a 0,20 s. Nos casos em que a onda Q não está presente, o que ocorre com muita frequência, o começo do complexo QRS coincide com o começo da onda R. Por este motivo, o intervalo PQ é denominado também de intervalo PR.

- **Complexo QRS.** Representa a despolarização ventricular, durando entre 0,10 e 0,12 s, com amplitude em torno de 1mV.
- **Onda Q.** Quando presente, ocorre no início do complexo QRS e, no ECG normal, deve durar menos que 0,04 s.
- **Onda S.** Deflexão negativa que ocorre no final do complexo QRS, devendo durar, no ECG normal, menos que 0,08 s.
- **Onda T.** Representa a repolarização ventricular, durando cerca de 0,15 s, com amplitude entre 0,2 e 0,3 mV.
- **Intervalo QT.** Estende-se do início do complexo QRS ao final da onda T. Sua duração média é de 0,3 s.
- **Intervalo R-R.** Estende-se do pico da onda R de um batimento até o pico da onda R do batimento seguinte. A frequência cardíaca, expressa em batimentos por minuto, é o inverso do valor do intervalo R-R, em minutos. A duração típica do intervalo R-R é de 0,83 s, correspondendo a uma frequência cardíaca de $60/0,83$, ou aproximadamente 72 batimentos por minuto.

2.4 O ECG DIGITAL

Para evitar o efeito de *aliasing*, a faixa de frequências do sinal deve ser limitada por filtragem analógica antes do processo de digitalização. A largura de faixa do filtro e, em consequência, a taxa de amostragem empregada, dependem fundamentalmente da aplicação que se deseja dar ao sinal. Para análise de arritmias a faixa é de 0,5 a 50 Hz. Na prática clínica, com o ECG padrão de doze derivações, a largura de faixa típica é de 0,05 a 100 Hz. Para medir potenciais tardios, eventos de frequência elevada e amplitude reduzida que ocorrem após o complexo QRS, a faixa estende-se até 500 Hz (Tompkins, 1993). Pelo teorema de Nyquist, neste caso, a taxa mínima de amostragem deve ser de 1000 amostras/s.

A *American Heart Association* sugere que, em geral, o ECG seja digitalizado com uma taxa mínima de 500 amostras/s, com no mínimo 12 bits por amostra (Bailey *et al.*, 1990). Com esta qualidade de digitalização, as oito derivações linearmente independentes gerariam uma taxa de $8 \times 500 \times 12 = 48000$ bits/s, ou 21.600.000 bytes/hora. Assim, um Holter de 24 horas com oito derivações geraria 518,4 milhões de bytes. Em tais situações, restrições de armazenamento obrigam a uma redução no número de derivações, na taxa de amostragem e no número de bits por amostra. Esquemas atuais de compressão de ECG, capazes de atingir razões de compressão da ordem de 10:1 com baixa distorção, proporcionam uma alternativa interessante à simples redução da qualidade da digitalização.

2.5 O BANCO DE SINAIS DE ECG DO MIT-BIH

O desempenho de um compressor de ECG é fortemente influenciado por diversas características do sinal, como frequência de amostragem, número de bits por amostra, nível de ruído, e morfologia. Assim, uma comparação empírica precisa do desempenho de técnicas diferentes exige que se utilize o mesmo conjunto de sinais de teste. Entretanto, muitos resultados divulgados na literatura foram obtidos com registros colhidos em laboratórios locais, que não podem ser facilmente acessados por outros pesquisadores, o que dificulta uma comparação justa entre os métodos. Para minimizar o problema, Jallaledine (1990) sugere a utilização de bancos padrões de ECG, disponíveis aos pesquisadores de todo o mundo, como o banco da *American Heart Association* (Hermes *et al.*, 1980) e o MIT-BIH *Arrhythmia Database* (Moody e Mark, 1988).

Uma excelente ilustração de como os níveis de ruído e morfologia afetam o compromisso entre o grau de compressão e a distorção é proporcionada pelos resultados do compressor baseado em DCT desenvolvido por van der Poel (1999), e testado com um conjunto de 34 sinais provenientes do MIT-BIH *Arrhythmia Database*. Todos os sinais deste banco foram digitalizados a 360 amostras/s, com 11 bits/amostra. Para um mesmo nível de distorção, um dos sinais foi comprimido a uma razão de 2,8:1, e um outro a 28,9:1. Vê-se com isto que a informação sobre a razão de compressão e o nível de distorção têm relevância reduzida na aferição da qualidade do compressor, quando dissociadas da especificação exata dos sinais de teste.

O banco de eletrocardiogramas digitais MIT-BIH *Arrhythmia Database*, um projeto conjunto do *Massachusetts Institute of Technology* (MIT) e do *Beth Israel Hospital* (BIH) tem sido extensamente empregado para testes empíricos de vários algoritmos. Segundo Moody e Mark (1988), no decurso de oito anos após sua disponibilização inicial, este banco foi comprovadamente utilizado por uma centena de grupos acadêmicos e industriais em todo o mundo. Para facilitar a comparação direta com outros compressores, a avaliação experimental das técnicas desenvolvidas no presente trabalho empregará o MIT-BIH *Arrhythmia Database*, pois este tem sido o banco padrão mais utilizado na obtenção dos resultados descritos na literatura.

O MIT-BIH *Arrhythmia Database* contém um total de 48 registros. Os 23 primeiros, numerados de 100 a 124 (não há os números 110 e 120), foram selecionados aleatoriamente de um conjunto de mais de 4000 Holters de longa duração obtidos pelo BIH *Arrhythmia Laboratory* entre 1975 e 1979. Os demais 25 registros, numerados de 200 a 234 (10 números não estão presentes), foram selecionados do mesmo conjunto de forma a incluir arritmias e anormalidades de condução raras mas importantes. Cada registro contém dois canais, coletados quase simultaneamente (a diferença de tempo entre os canais é da ordem de poucos microssegundos) e dura aproximadamente 30,1 minutos. Na maioria dos casos, o canal superior é uma derivação II modificada (*modified limb lead II*, ou MLII, na terminologia do MIT-BIH *Arrhythmia Database*), obtida colocando-se os eletrodos sobre o tórax, e não nos membros. O canal inferior é normalmente uma derivação V1 modificada (ocasionalmente V2 ou V5, e em apenas um caso, V4), com os eletrodos também colocados no tórax.

Os registros foram originalmente armazenados em fitas magnéticas, em formato analógico. Para o processo de digitalização, alguns registros foram reproduzidos em tempo real; outros, a duas vezes a velocidade real. O sinal de saída da unidade de reprodução foi filtrado por um filtro passa-faixa de 0,1 a 100 Hz, referente à velocidade real. Vários artefatos no domínio da frequência, devidos a componentes específicos do equipamento de gravação e de reprodução, foram identificados (Moody e Mark, 1988). Os sinais foram amostrados a 360 Hz por canal (relativamente ao tempo real), para facilitar a implementação de filtros *notch* para redução do ruído de 60 Hz, introduzido principalmente durante a reprodução para a digitalização. Nos registros reproduzidos a duas vezes a velocidade real, este ruído surge em 30 Hz, com relação ao tempo real. O conversor analógico/digital utilizou 11 bits por amostra, com faixa de tensões de -5 a +5 mV. Assim, os

valores das amostras digitalizadas vão de 0, para tensões de entrada iguais ou menores que -5 mV, a 2047, para tensões iguais ou maiores que 5 mV, com 1024 representando 0 mV.

Capítulo 3

CONCEITOS BÁSICOS DE COMPRESSÃO DE DADOS

As técnicas de compressão partem do princípio de que as fontes usuais de informação geram dados redundantes e codificados de maneira pouco eficiente. Com base nesta premissa, procura-se explorar as características da fonte de forma a reduzir a redundância e obter uma codificação mais apropriada para os dados. O objetivo de um compressor é reduzir a representação da informação gerada pela fonte, procurando preservar a informação em si.

Se a fonte gera uma mensagem x codificada com n bits, um compressor constrói uma nova representação z , utilizando m bits, com m idealmente menor que n . O descompressor, por sua vez, constrói uma aproximação \tilde{x} para x , a partir de z . A Figura 3.1 apresenta esquematicamente um compressor/descompressor genérico. Seguindo a convenção usual, em diversos pontos deste trabalho o termo *compressor* designa o par compressor/descompressor. Na Figura 3.1, o canal é o meio para o qual a mensagem comprimida é enviada. O canal pode ser, por exemplo, um disco rígido ou uma linha telefônica.

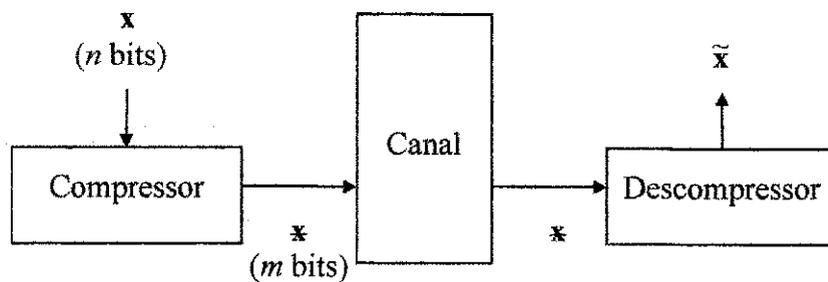


Figura 3.1 Um compressor/descompressor genérico

Um indicador muito aplicado para avaliar o grau de compressão obtido é a *razão de compressão*, RC, dada por:

$$RC = \frac{n}{m} \quad (3.1)$$

Alternativamente, pode-se empregar a *taxa de bits*, R , para mensurar a compressão. A taxa de bits, ou simplesmente *taxa*, de uma mensagem é o número médio de bits utilizados para representar cada elemento da mensagem. Um texto composto por caracteres codificados em ASCII estendido, por exemplo, apresenta uma taxa de 8 bits/elemento. Se este texto é comprimido com uma RC de 2:1, tem-se $R = 8/2 = 4$ bits/elemento.

Na compressão sem perdas a informação é completamente preservada e, assim, a mensagem original pode ser perfeitamente reconstruída a partir de x , ou seja, $\tilde{x} = x$. Na compressão com perdas os dados são irreversivelmente alterados, ocorrendo perda de informação potencialmente importante (Sayood, 1996; Nelson e Gailly, 1996). Neste caso, pode-se apenas construir uma aproximação para x a partir de x . Na compressão com perdas, portanto, $\tilde{x} \neq x$, e a qualidade da reconstrução depende da relevância da informação descartada. A distorção introduzida pelo processamento é aferida por alguma *métrica de distorção*, D . Algumas métricas usuais serão apresentadas posteriormente. Em geral, maiores distorções permitem obter maiores razões de compressão.

A compressão sem perdas é aplicada quando a alteração de um único bit é considerada inaceitável, como na compressão de textos ou arquivos executáveis. Por outro lado, na compressão de

imagens estáticas e em movimento, por exemplo, distorções consideráveis podem ser introduzidas de maneira imperceptível ou quase imperceptível para o sistema visual humano, aumentando assim as oportunidades de compactação (Wallace, 1991).

No caso de sinais biológicos utilizados para diagnósticos, reconstrução perfeita é uma característica altamente desejável mas não obrigatória (Lai, 1993; Antoniol e Tonella, 1997; Koski, 1997). Se a necessidade de uma maior RC exigir a introdução de distorção, estas devem ser cuidadosamente controladas. Diversas medidas matemáticas de distorção têm sido propostas, mas nenhuma delas é capaz de indicar com precisão o grau de manutenção das características clínicas importantes do sinal original (Lai, 1993). Assim, tais medidas não devem substituir a apresentação dos sinais a especialistas, após o processamento, para aferição visual da qualidade clínica resultante.

Para uma dada mensagem, um compressor não parametrizado produz invariavelmente a mesma taxa e distorção. Os compressores parametrizados, por outro lado, oferecem um maior grau de controle sobre os resultados, uma vez que a distorção e a taxa podem ser ajustados pelos parâmetros. Denotaremos a taxa e a distorção geradas por uma compressão parametrizada de uma mensagem x por $R(\Theta)$ e $D(\Theta)$, onde Θ representa a lista de parâmetros do compressor. Para simplificar a notação, sempre que o contexto deixar claro que estamos nos referindo a compressores parametrizados, representaremos a taxa por R e a distorção por D .

Dadas uma mensagem x e uma métrica de distorção, o problema da compressão de dados consiste em desenvolver métodos que procurem atingir a menor taxa possível sem ultrapassar uma dada distorção ou, alternativamente, que procurem atingir a menor distorção possível sem ultrapassar uma dada taxa. Assim expressos, ambos os casos podem ser tratados como um problema mais genérico de minimização com restrição. As abordagens para solucioná-los são, portanto, conceitualmente idênticas e, por este motivo, será investigado neste trabalho unicamente o problema de minimização da taxa para uma distorção especificada. Como a qualidade da reconstrução será priorizada no presente estudo, esta opção permitirá um maior controle sobre a distorção resultante.

Para um compressor parametrizado específico, o problema da compressão ótima de uma mensagem pode ser formulado como um problema de minimização com restrição:

Formulação 3.1: Dada uma distorção D_T , encontrar uma instância de parâmetros do compressor que minimize R , com $D = D_T$.

3.1 MEDIDAS DE DISTORÇÃO

Sejam x_n e \tilde{x}_n , $n = 0, 1, \dots, N-1$, os valores das amostras, em mV , dos sinais original, x e reconstruído, \tilde{x} , respectivamente. Há diversas medidas que permitem aferir a distorção entre x_n e \tilde{x}_n . Uma das mais populares é o *erro médio quadrático* (*mean square error* - MSE):

$$\text{MSE} = \frac{\sum_{n=0}^{N-1} (x_n - \tilde{x}_n)^2}{N} \quad (3.2)$$

Uma outra medida de distorção diretamente relacionada a esta é a *raiz do erro médio quadrático* (*root mean square error* - RMSE):

$$\text{RMSE} = \sqrt{\frac{\sum_{n=0}^{N-1} (x_n - \tilde{x}_n)^2}{N}} \quad (3.3)$$

A distorção na reconstrução dos sinais de ECG é frequentemente medida pelo *percentual da raiz média quadrática das diferenças* (*Percent Root-Mean-Square Difference* - PRD) (Abenstein e Tompkins, 1982):

$$\text{PRD} = \frac{\text{RMSE}}{\text{RMSV}} \times 100\% \quad (3.4)$$

Onde RMSV é o valor RMS (*Root Mean Square*) do sinal

$$\text{RMSV} = \sqrt{\frac{\sum_{n=0}^{N-1} x_n^2}{N}} \quad (3.5)$$

Como esta medida é muito sensível ao nível DC do sinal original, uma segunda definição da PRD, que será aqui denominada PRD2, substitui o valor RMS pelo desvio padrão, σ , do sinal:

$$\sigma = \sqrt{\frac{\sum_{n=0}^{N-1} (x_n - \bar{x})^2}{N}} \quad (3.6)$$

onde

$$\bar{x} = \frac{1}{N} \sum_{n=0}^{N-1} x_n \quad (3.7)$$

A PRD2 é então dada por:

$$\text{PRD2} = \frac{\text{RMSE}}{\sigma} \times 100\% \quad (3.8)$$

Apesar de largamente utilizadas como medida de distorção, a PRD e a PRD2 não indicam com precisão a qualidade da reconstrução (Zigel e Cohen, 1999). Em outras palavras, baixos valores de PRD ou PRD2 não garantem preservação total das características essenciais do sinal original e, assim, o sinal descomprimido deve ser também avaliado por inspeção visual. Recentemente, uma nova medida de distorção, denominada *distorção diagnóstica ponderada* (*Weighted Diagnostic Distortion - WDD*), foi proposta (Zigel *et al.*, 1997). Esta medida é baseada nas características PQRST importantes para o diagnóstico e seu cálculo exige a segmentação do sinal. Apesar da boa correlação da WDD com a percepção de cardiologistas, a elevada complexidade envolvida em seu cálculo limita sua aplicabilidade (Zigel e Cohen, 1999).

A especificação da qualidade dos compressores de sinais de ECG deve levar em consideração o nível de distorção introduzido no sinal pelo processamento, a taxa de compressão média obtida, e a complexidade computacional.

3.2 INFORMAÇÃO E ENTROPIA

Considere-se uma fonte de informação S produzindo elementos cujos valores são selecionados de um conjunto $A = \{a_0, a_1, \dots, a_{M-1}\}$. Este conjunto é o *alfabeto* de S e os seus elementos são conhecidos como *letras* ou *símbolos*. Um elemento x gerado por S pode ser considerado uma variável aleatória que assume o valor a_i com probabilidade $P(x = a_i)$, $i = 0, 1, \dots, M-1$. Define-se a *auto-informação*, $I(a_i)$, associada à ocorrência do símbolo a_i como

$$I(a_i) = \log_2 \frac{1}{P(x = a_i)} \text{ bits} \quad (3.9)$$

Se x_1, x_2, \dots, x_{N-1} são os elementos gerados por S , a *entropia* da fonte é definida como

$$H(S) = \lim_{N \rightarrow \infty} \frac{1}{N} G_N \text{ bits/símbolo} \quad (3.10)$$

onde

$$G_N = - \sum_{i_0=0}^{M-1} \sum_{i_1=0}^{M-1} \dots \sum_{i_{N-1}=0}^{M-1} P(x_0 = a_{i_0}, x_1 = a_{i_1}, \dots, x_{N-1} = a_{i_{N-1}}) \log_2 P(x_0 = a_{i_0}, x_1 = a_{i_1}, \dots, x_{N-1} = a_{i_{N-1}}) \quad (3.11)$$

Se as variáveis aleatórias que constituem a mensagem são independentes e identicamente distribuídas (IID), a Equação (3.11) pode ser simplificada para expressar a *entropia de ordem 1* de S (Sayood, 1996):

$$H(S) = - \sum_{i=0}^{M-1} P(x = a_i) \log_2 P(x = a_i) \text{ bits/símbolo} \quad (3.12)$$

Shannon mostrou que é possível codificar sem perdas a saída de uma fonte qualquer de informação com um número médio de bits por símbolo arbitrariamente próximo à entropia, mas não inferior a ela (Shannon, 1948). Assim, pode-se usar a entropia para avaliar a eficiência da codificação efetuada, sendo o código ótimo aquele cujo comprimento médio é igual à entropia. O problema de como construir tal código foi resolvido com a introdução da codificação aritmética (Bell *et al.*, 1990), que será apresentada no Capítulo 4.

3.3 FUNÇÃO TAXA-DISTORÇÃO

A teoria taxa-distorção define uma função $R(D)$, denominada *função taxa-distorção*, que representa a taxa mínima que pode ser atingida para uma dada fonte de informação (estatisticamente caracterizada) em função da distorção. A taxa mínima para compressão sem perdas, $R(0)$, é a entropia da fonte. Pode-se mostrar que $R(D)$ é uma função convexa não crescente de D . Para fontes com determinadas funções densidade de probabilidade, foram definidas formas analíticas para a função $R(D)$. Por exemplo, para uma fonte gaussiana com média zero e variância σ^2 , usando-se o erro médio quadrático (EMQ) como métrica de distorção, tem-se (Hoang, 97):

$$R(D) = \begin{cases} \frac{1}{2} \log_2 \frac{\sigma^2}{D}, & 0 \leq D \leq \sigma^2 \\ 0, & D > \sigma^2 \end{cases} \quad (3.13)$$

Destaque-se que a teoria taxa-distorção define um limite inferior para a taxa em função da distorção para uma dada fonte. Para tanto, é preciso caracterizar estatisticamente a fonte de informação de maneira analiticamente tratável. Maiores detalhes podem ser encontrados na literatura (Jayant e Noll, 1984).

3.4 TAXA-DISTORÇÃO OPERACIONAL

Em situações práticas, os resultados da teoria taxa-distorção nem sempre produzem resultados suficientemente exatos, pois as fontes de informação reais podem desviar-se significativamente das hipóteses simplificadoras (elementos estatisticamente independentes, função densidade de probabilidade analiticamente tratável, estacionaridade, etc.) necessárias para obtenção da função taxa-distorção. Como alternativa a esta abordagem, pode-se medir as taxas e distorções reais atingidas por um compressor parametrizado para um número finito, ainda que potencialmente grande, de instâncias dos parâmetros Θ , caracterizando assim o *comportamento R-D operacional* do compressor.

3.4.1 O Plano $R-D$ operacional

Cada instância de Θ produz um único ponto, denominado *ponto operacional*, no plano $R-D$. Considere-se que instância de parâmetros Θ_i gera o ponto P_i de coordenadas (D_i, R_i) . A Figura 3.2 apresenta alguns pontos hipotéticos em um plano $R-D$ operacional.

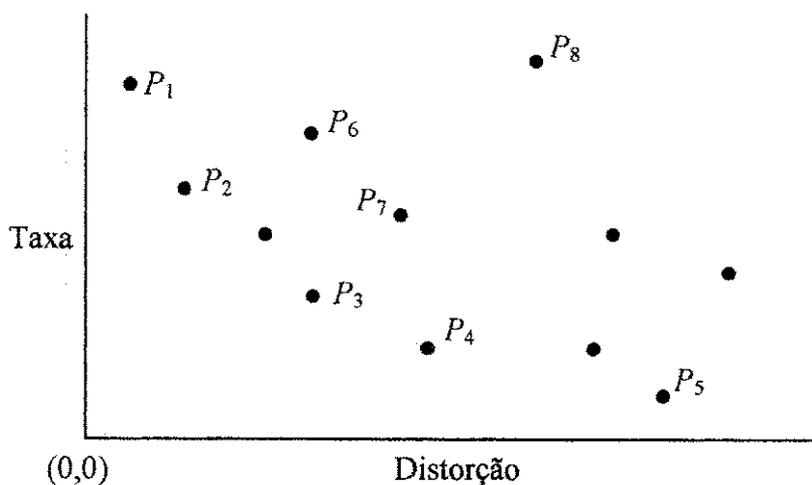


Figura 3.2 Plano $R-D$ operacional

A distribuição dos pontos neste plano indica que a escolha dos parâmetros pode afetar fortemente os resultados. Por exemplo, os pontos P_3 e P_6 apresentam a mesma distorção, mas a taxa correspondente a P_3 é aproximadamente a metade daquela associada a P_6 . Logo, P_3 é claramente superior a P_6 em termos de compromisso $R-D$. A diferença é ainda mais drástica quando comparamos P_3 com P_8 , pois em P_3 temos simultaneamente distorção e taxa significativamente menores que em P_8 . Uma comparação de qualidade entre P_2 e P_3 não é óbvia, pois o primeiro apresenta uma distorção menor mas, ao mesmo tempo, uma taxa maior que o segundo. Se a qualidade da reconstrução deve ser priorizada, P_2 poderia representar uma melhor escolha; se a ênfase recai sobre uma maior compressão, P_3 seria talvez a opção preferida.

Devido à natureza discreta do plano $R-D$ operacional, a Formulação 3.1 só têm solução quando há pelo menos um ponto operacional com distorção igual à distorção alvo D_T . A Formulação 3.2, dada a seguir, é mais apropriada ao caso discreto:

Formulação 3.2: Dada uma distorção D_T e uma tolerância τ , encontrar uma instância de parâmetros do compressor que minimize R , com $D_T - \tau \leq D \leq D_T + \tau$.

3.4.2 Conjunto convexo gerado pelos pontos operacionais

Seja A um conjunto de pontos $\{P_i\}$ em \mathbb{R}^2 . Uma vizinhança ϵ em torno de P_i é o conjunto de pontos pertencentes ao interior do círculo de centro P_i e raio ϵ . Um ponto P_i é um *ponto limite* do conjunto A se qualquer vizinhança ϵ em torno de P_i , com $\epsilon > 0$ arbitrariamente pequeno, contenha pontos que pertencem e pontos que não pertencem a A . Um conjunto é *fechado* quando contém todos os seus pontos limites.

Uma *combinação convexa* de dois pontos P_i, P_j é um ponto $P = \alpha P_i + (1 - \alpha)P_j$, $0 \leq \alpha \leq 1$. O conjunto A é *convexo* se a combinação convexa de quaisquer dois pontos de A também pertence a A . Isto implica que o segmento de reta ligando dois pontos quaisquer de um conjunto convexo também pertence ao conjunto. Intuitivamente, um conjunto convexo é “sólido” (sem “buracos”) e não reentrante (Hadley, 1982).

Um ponto P é um *ponto extremo* de um conjunto convexo A se e somente se não houver outros pontos $P_i, P_j \in A$, $P_i \neq P_j$, tais que $P = \alpha P_i + (1 - \alpha)P_j$, $0 < \alpha < 1$. Note-se que desigualdades estritas são impostas para α . Assim, um ponto extremo não pode situar-se em nenhum segmento de reta unindo outros dois pontos do conjunto. Todo ponto extremo é um ponto limite, mas nem todo ponto limite é um ponto extremo. Por exemplo, no conjunto convexo formado por um triângulo e seu interior, todos os pontos do contorno são limites, mas apenas os vértices do triângulo são extremos.

Uma combinação convexa de m pontos P_0, \dots, P_{m-1} é um ponto P dado por

$$P = \sum_{i=0}^{m-1} \alpha_i P_i \quad (3.14)$$

onde $\alpha_i \geq 0$, $i = 0, \dots, m - 1$, e

$$\sum_{i=0}^{m-1} \alpha_i = 1 \quad (3.15)$$

Um *polígono convexo* gerado por m pontos P_0, \dots, P_{m-1} em \mathcal{R}^2 é o conjunto produzido por todas as combinações convexas dos pontos geradores. Os pontos geradores que não podem ser produzidos por combinação convexa de outros pontos formam os extremos (vértices) do polígono.

A Figura 3.3 mostra o polígono convexo gerado pelos pontos operacionais de um plano R - D hipotético. Os pontos com a menor distorção e a menor taxa possíveis são representadas por $P_a = (D_a, R_a)$ e $P_b = (D_b, R_b)$, respectivamente. Nenhum ponto $P = (D, R)$ com $R > R_a$ pode ser uma solução para as Formulações 1 e 2, pois o ponto P_a representa um melhor compromisso R - D , com $R_a < R$ e $D_a < D$. Analogamente, nenhum ponto $P = (D, R)$ com $D > D_b$ pode ser uma solução para as Formulações 1 e 2, uma vez que o ponto P_b representa um melhor compromisso R - D , com $R_b < R$ e $D_b < D$. Conclui-se então que a solução ótima deverá ser procurada no conjunto convexo mostrado na Figura 3.4.

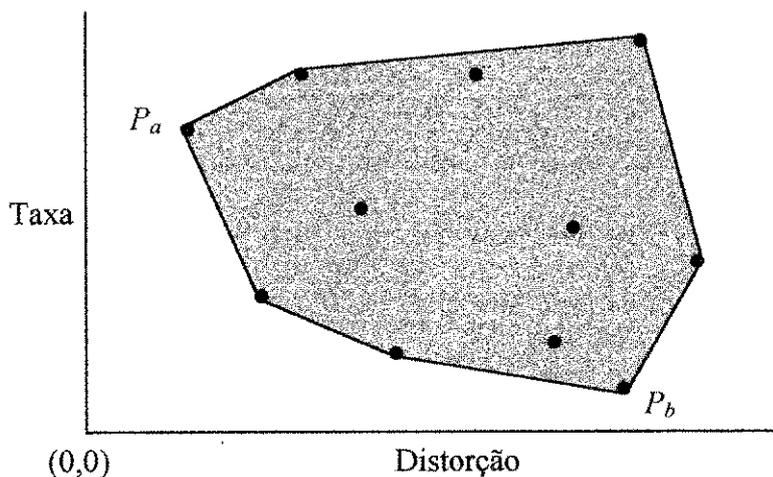


Figura 3.3 Polígono convexo gerado por pontos operacionais

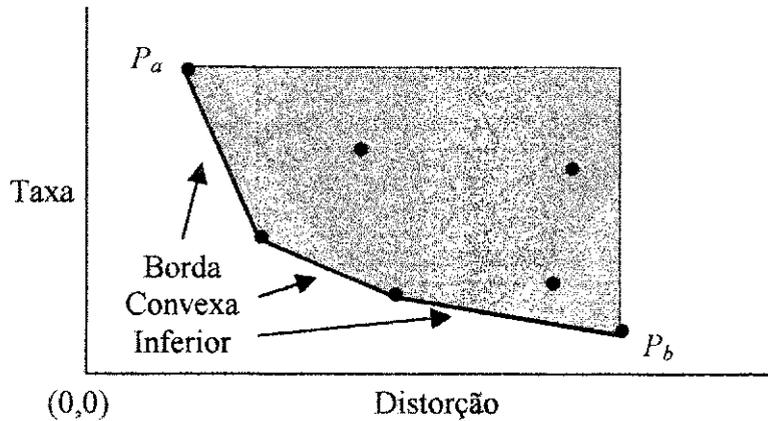


Figura 3.4 Região de busca da solução ótima

A borda convexa inferior (BCI), realçada por uma linha mais espessa na Figura 3.4, possui especial interesse para o problema da compressão ótima. Qualquer ponto operacional $P = (D, R)$ localizado sobre a BCI apresenta taxa mínima para a distorção D , sendo portanto a solução para a Formulação 3.1 com $D_T = D$. Por outro lado, como já discutido anteriormente, dada uma distorção alvo arbitrária D_T , nem sempre existe um ponto operacional cuja distorção seja igual a D_T . A Formulação 3.2 deve então ser utilizada, com uma tolerância adequada, muitas vezes empiricamente definida. Neste caso, um ponto operacional $P = (D, R)$ sobre a BCI, tal que $D_T - \tau \leq D \leq D_T + \tau$, será ótimo para a distorção D , mas não necessariamente ótimo sobre todo o intervalo $[D_T - \tau, D_T + \tau]$. Na Figura 3.5, que ilustra esta situação, $P_i = (D_i, R_i)$, localizado sobre a BCI, é a solução ótima para a distorção D_i , mas P_j apresenta a menor distorção sobre o intervalo de tolerância, sendo conseqüentemente a solução ótima à luz da Formulação 3.2.

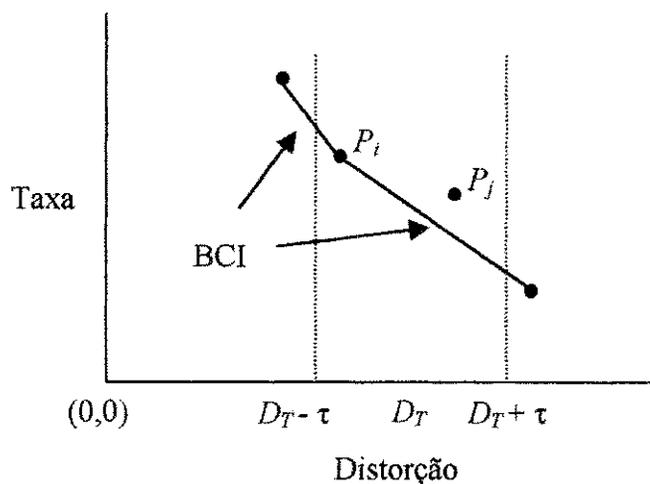


Figura 3.5 O ponto P_j , localizado fora da BCI, é ótimo dentro do intervalo de tolerância.

Sabendo-se que qualquer ponto operacional $P = (D, R)$ localizado na BCI é ótimo para a distorção D , e que buscas convexas podem ser eficientemente efetuadas (Ratnakar, 1997), muitos métodos práticos restringem a procura da solução ótima à BCI. Quando o conjunto de pontos operacionais é suficientemente denso, condição que freqüentemente se verifica, a solução ótima de borda convexa é normalmente muito próxima à solução ótima para a Formulação 3.2 (Ortega, 1994; Ratnakar, 1997).

3.4.3 Minimização lagrangiana

Para buscas sobre a BCI, o problema de otimização da taxa com distorção restrita pode ser transformado em um problema de otimização sem restrição usando-se *minimização lagrangiana*:

Formulação 3.3: Dado um parâmetro real $\lambda \geq 0$, encontrar uma instância de parâmetros do compressor que minimize o lagrangiano $J(\Theta) = R(\Theta) + \lambda D(\Theta)$.

O parâmetro λ é denominado multiplicador de Lagrange. Para λ fixo, $J = R + \lambda D$ representa uma reta no plano R - D com declividade $-\lambda$, interceptando o eixo R em J e o eixo D em J/λ . O método de minimização lagrangiana pode ser graficamente interpretado como a procura pelo último ponto operacional interceptado pela varredura, em direção à origem do plano R - D , realizada por esta reta, conforme mostra a Figura 3.6. Nesta figura, o ponto P_2 é o último interceptado pela reta

durante a varredura sendo, por conseguinte, o ponto ótimo para o valor de λ utilizado. A interpretação gráfica deixa claro que tal ponto localiza-se sempre sobre a BCI (Hoang, 1997).

O valor de λ define o compromisso entre taxa e distorção. Para $\lambda = 0$, a reta é horizontal, e a varredura é feita de cima para baixo. Conseqüentemente, o último ponto interceptado será o correspondente à menor taxa, independente da distorção. Para $\lambda \rightarrow \infty$, a reta é vertical, e a varredura ocorre da direita para a esquerda. Neste caso, o último ponto interceptado será o correspondente à menor distorção, independente da taxa. Em outras palavras, a otimização resulta na menor taxa possível quando $\lambda = 0$, e na menor distorção possível quando $\lambda \rightarrow \infty$. Nas situações intermediárias, à medida que o valor de λ aumenta, o valor ótimo de R aumenta e o de D diminui.

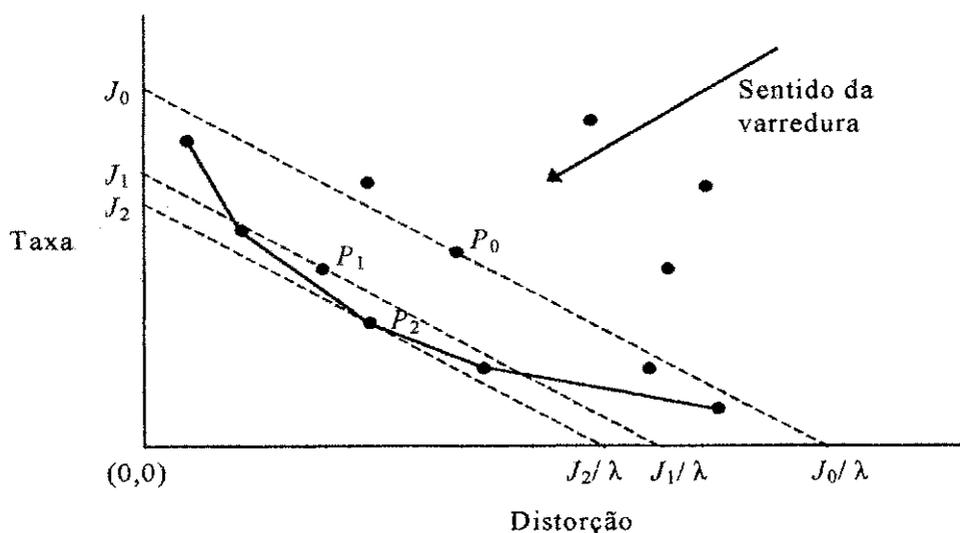


Figura 3.6 Interpretação gráfica da minimização lagrangiana

Seja $\Theta^*(\lambda)$ uma instância de parâmetros que minimiza $J(\Theta)$:

$$R(\Theta^*(\lambda)) + \lambda D(\Theta^*(\lambda)) \leq R(\Theta) + \lambda D(\Theta) \quad (3.16)$$

ou, equivalentemente,

$$R(\Theta) - R(\Theta^*(\lambda)) \geq \lambda(D(\Theta^*(\lambda)) - D(\Theta)) \quad (3.17)$$

Logo, para $D(\Theta) \leq D(\Theta^*(\lambda))$, tem-se

$$R(\Theta^*(\lambda)) \leq R(\Theta) \tag{3.18}$$

Vê-se com isto que a instância de parâmetro $\Theta^*(\lambda)$ que minimiza o lagrangiano minimiza também a taxa para $D(\Theta) \leq D(\Theta^*(\lambda))$. Assim, qualquer solução para o problema proposto na Formulação 3.3 é também uma solução para o problema proposto na Formulação 3.1, com $D_T = D(\Theta^*(\lambda))$. Deve-se notar, contudo, que nem sempre é possível encontrar um valor de λ tal que a distorção $D(\Theta^*(\lambda))$ coincida exatamente com a distorção alvo especificada. Na prática, pode-se procurar um valor para λ tal que $D_T - \tau \leq D(\Theta^*(\lambda)) \leq D_T + \tau$, onde τ é uma tolerância especificada. Se o conjunto de pontos operacionais é suficientemente denso, a tolerância pode ser muito pequena. O *método da bisseção* (Ortega, 1994) é um procedimento iterativo que permite determinar um valor adequado para λ .

Capítulo 4

TÉCNICAS DE COMPRESSÃO DE DADOS

Este capítulo apresenta um conjunto de ferramentas fundamentais para a compressão de dados. Serão tratados os métodos de modelagem, codificação, quantização e transformada. Em última instância, a investigação de modelos estatísticos e técnicas de predição e interpolação precisas, bem como de transformadas com características propícias à compressão e de abordagens de quantização bem planejadas, objetivam passar ao codificador uma mensagem com a menor entropia possível. Quando o estudo não é puramente teórico, havendo também interesse na implementação em computador, tenta-se reduzir ao máximo a entropia respeitando-se as limitações dos recursos computacionais disponíveis. O capítulo finaliza apresentando o importante paradigma de compressão baseado em transformada, quantização e codificação.

4.1 MODELAGEM

A separação do processo de compressão em modelagem e codificação foi um dos grandes avanços na teoria e prática da compressão de dados nas últimas décadas (Bell *et al.*, 1990), permitindo tratar uma enorme variedade de compressores como a associação entre alguns poucos modelos e codificadores diversos. As próximas seções discutem alguns modelos estatísticos e as técnicas de modelagem denominadas predição e interpolação.

4.1.1 Modelagem Estatística

O conceito de entropia e, conseqüentemente, de código ótimo, está diretamente relacionado com a definição das probabilidades associadas aos símbolos gerados pela fonte, bem como com as considerações a respeito de dependência estatística. Em outras palavras, a entropia é calculada a partir da definição de um *modelo estatístico* para a fonte de informação. Na verdade, qualquer processo de codificação está associado a um modelo subjacente, que pode ou não ser explicitamente definido. Por exemplo, a atribuição de um número fixo de bits a cada um dos elementos de uma seqüência, como no código ASCII de 8 bits, é ótimo para fontes que seguem um modelo IID com a mesma probabilidade de ocorrência para todos os símbolos do alfabeto.

Diz-se que um modelo que captura mais precisamente as características reais da fonte “reduz sua entropia”, aumentando as oportunidades de compressão. Por exemplo, se os elementos de uma mensagem são estatisticamente dependentes, a entropia associada a um modelo que expresse este comportamento será menor do que aquela associada a um modelo IID.

Enquanto a codificação ótima é um problema já solucionado, a modelagem permanece um ativo campo de pesquisa. Dada a natureza intrinsecamente estatística das comunicações e dos processos físicos de interesse prático, a definição de um modelo que leve à menor entropia para uma fonte genérica é um problema sem solução definida (Bell *et al.*, 1990), e ganhos de compressão podem sempre ser obtidos com a construção de modelos mais precisos.

Considere-se uma fonte com alfabeto $A = \{a_0, a_1, \dots, a_{M-1}\}$. Um modelo estatístico simples consiste em assumir que os elementos gerados pela fonte são independentes e assumem o valor a_i com

probabilidade $P(x = a_i)$, $i = 0, 1, \dots, M-1$. Quando a suposição de independência não é satisfatória, os *modelos de Markov* estão entre os mais comumente usados para representar a relação entre os símbolos. Uma mensagem segue um modelo de Markov de k -ésima ordem se a distribuição de probabilidades de um elemento depende unicamente dos k símbolos que o antecedem. É necessário, portanto, estimar as probabilidades $P(x_n = a_i | x_{n-1}, x_{n-2}, \dots, x_{n-k})$, $i = 0, 1, \dots, M-1$. Os k símbolos precedentes constituem o *contexto* a partir do qual a probabilidade de ocorrência do próximo elemento é estimada, e por este motivo os modelos de Markov de k -ésima ordem são também conhecidos como *modelos de contexto finito*. Note-se que o IID é um modelo de Markov de ordem zero.

Se uma fonte gera símbolos que dependem estatisticamente dos valores presentes em um contexto C de tamanho L , um modelo de Markov de ordem k , com $k \leq L$, é capaz de representar com maior precisão as características da fonte do que um modelo de ordem $k-1$ e, conseqüentemente, conduzem a uma maior redução na entropia. Seria natural empregar modelos de ordem elevada para obter boas compressões, mas alguns problemas práticos limitam a aplicabilidade da idéia. Com um modelo de ordem k e um alfabeto de tamanho M , tem-se M^k possíveis contextos, ou seja, o número de contextos diferentes cresce exponencialmente com a ordem do modelo. Se $M = 256$ e $k = 5$, haveria aproximadamente um trilhão de contextos diversos. Um número elevado de contextos leva a uma série de problemas práticos, relacionados aos requisitos de memória e ao cálculo das estimativas e possível transmissão para o decodificador das probabilidades condicionais (Bell *et al.*, 1990). Por este motivo, os modelos de Markov de ordem zero e de primeira ordem são os mais utilizados em situações reais.

4.1.2 Predição e Interpolação

Como alternativa ao uso de modelos estatísticos complexos para fontes com elementos dependentes, pode-se procurar aproximar o comportamento da fonte através de um modelo determinístico, cujo objetivo é proporcionar uma aproximação precisa para o valor de um elemento através da aplicação de uma função determinística $f(C)$ a um contexto C formado pelos valores reais de outros elementos da mensagem. A função f é um mapa entre contextos e aproximações. Como o mapa é fixo e conhecido pelo codificador e pelo decodificador, apenas os erros entre os valores reais dos elementos e as aproximações precisam ser codificados. Se o esquema operar satisfato-

riamente, a entropia dos erros será menor do que a da mensagem original. Pode-se obter um maior grau de compressão, com evidente perda de informação, desprezando-se os erros menores que um determinado limiar e codificando-se apenas os elementos que se desviam significativamente do comportamento previsto pela função $f(C)$, ou ainda quantizando os erros (Ruttimann e Pipberger, 1979; Jalaleddine *et al.*, 1990).

Se o contexto empregado para construir a aproximação para o valor de um dado elemento contém apenas símbolos que o antecedem na mensagem, o processo é chamado de *predição*. Por outro lado, se há símbolos antecedentes e subseqüentes no contexto, tem-se uma *interpolação*. Apresenta-se a seguir alguns exemplos de predição e interpolação reportados na literatura.

Os *preditores polinomiais* efetuam uma extrapolação polinomial, um elemento por vez. Dada uma mensagem $\{x_n\}$, a regra de extrapolação para um preditor polinomial de ordem k é (Jalaleddine *et al.*, 1990):

$$\tilde{x}_i = x_{i-1} + d_{i-1} + d_{i-1}^2 + \dots + d_{i-1}^k \quad (4.1)$$

onde \tilde{x}_i é a aproximação para x_i , e

$$d_{i-1} = x_{i-1} - x_{i-2} \quad (4.2)$$

$$d_{i-1}^k = d_{i-1}^{k-1} - d_{i-2}^{k-1} \quad (4.3)$$

As predições para $k = 0, 1$ e 2 são dadas, respectivamente, pelas Equações (4.4), (4.5) e (4.6)

$$\tilde{x}_i = x_{i-1} \quad (4.4)$$

$$\tilde{x}_i = 2x_{i-1} + x_{i-2} \quad (4.5)$$

$$\tilde{x}_i = 3x_{i-1} - 3x_{i-2} + x_{i-3} \quad (4.6)$$

Um *preditor linear* de ordem k calcula as aproximações a partir de uma combinação linear dos símbolos em um contexto composto pelos k símbolos prévios:

$$\tilde{x}_i = \sum_{j=1}^k a_j x_{i-j} \quad (4.7)$$

onde $a_j, j = 1, 2, \dots, k$, são pesos calculados de maneira a minimizar o erro médio quadrático entre as predições e os valores reais. Considerando que os erros têm média zero, o erro médio quadrático é igual à variância dos erros. Se os erros possuem uma distribuição gaussiana, minimizar a variância equivale a minimizar a entropia (Ruttimann e Pipberger, 1979). Para outras distribuições unimodais, pode-se mostrar que a entropia decresce monotonicamente com a variância. Testes efetuados com sinais de ECG (Ruttimann e Pipberger, 1979) mostraram que não se obtém redução significativa na variância dos erros com preditores lineares de ordem maior que 2.

Um *interpolador linear* pode ser descrito de forma semelhante a um preditor linear, usando um contexto que inclui símbolos antecedentes e subseqüentes:

$$\tilde{x}_i = \sum_{j=1}^k (a_j x_{i-j} + b_j x_{i+j}) \quad (4.8)$$

Novamente, a_j e $b_j, j = 1, 2, \dots, k$, são pesos calculados para que o erro médio quadrático entre os valores interpolados e os reais seja mínimo. Testes efetuados com sinais de ECG (Ruttimann e Pipberger, 1979) indicam que um interpolador com $k = 1$ é significativamente mais eficiente na redução da entropia que um preditor de ordem 2.

Além da redução na variância, a predição e a interpolação geram seqüências de erros mais descorrelacionados que os elementos da seqüência original, permitindo o uso de modelos estatísticos simples (Carpentieri *et al.*, 2000). Neste sentido, a predição e a interpolação têm um papel análogo ao das transformadas, descritas na Seção 4.5.

4.2 CODIFICAÇÃO DE ENTROPIA

No presente trabalho, *codificação* é a representação de dados através de identificadores binários. Os identificadores são denominados *palavras-códigos*, e o conjunto de todas as palavras-códigos para uma determinada fonte constitui um *código*. O *comprimento* de uma palavra-código é o seu número de bits. Em um *código de comprimento fixo*, todas as palavras-códigos têm o mesmo comprimento, enquanto que em um código de comprimento variável, as palavras-códigos podem ter comprimentos diferentes. Atribuindo-se menos bits à representação dos valores com maior probabilidade de ocorrência, é possível reduzir o comprimento médio dos identificadores, reduzindo assim a taxa das mensagens geradas pela fonte. Este é princípio básico dos codificadores de entropia.

Dada uma mensagem codificada, a mensagem original poderá ser recuperada sem ambigüidade se o código utilizado for *unicamente decodificável*, ou seja, se a mensagem codificada pode ser decodificada de uma, e apenas uma, maneira. Esta decodificação pode ser um processo complexo. Códigos unicamente decodificáveis são facilmente decodificáveis quando são também *códigos de prefixo*, ou seja, quando nenhuma de suas palavras-códigos é prefixo de outra (Sayood, 1996).

4.2.1 Codificação de Huffman

Em 1952, Huffman apresentou um procedimento sistemático para a construção de códigos de prefixo ótimos, também conhecidos como *códigos de mínima redundância*, dadas as probabilidades previamente definidas na etapa de modelagem (Huffman, 1952). A redundância do código de Huffman é definida como a diferença entre o comprimento médio das palavras-código, l , e a entropia da fonte, H , calculada a partir das probabilidades fornecidas pelo modelo. Pode-se mostrar que esta redundância é superiormente limitada por $p + 0,086$ (Bell *et al.*, 1990), onde p é a probabilidade do símbolo mais provável e, em conseqüência, l situa-se no intervalo $[H, H + p + 0,086)$. Portanto, quanto mais p se aproxima de 1, maior a redundância do código. O comprimento médio de um código de Huffman é igual à entropia apenas quando todas as probabilidades são potências de 0,5.

O algoritmo de codificação de Huffman associa uma árvore ponderada a cada letra do alfabeto da fonte de informação. Inicialmente, cada árvore possui um único nó, com peso igual à probabilidade de ocorrência do símbolo a ela associado. A cada iteração do algoritmo, as duas árvores de menor peso são substituídas por uma nova árvore cujo peso é a soma dos pesos das primeiras. A árvore de menor peso se torna a subárvore esquerda e a outra se torna a subárvore direita da nova árvore. Na ordenação dos pesos, empates são resolvidos por qualquer regra sistemática. O procedimento pára quando resta apenas uma única árvore. A palavra-código para qualquer letra é obtida percorrendo-se esta árvore desde a raiz até a folha correspondente à letra em questão, registrando 0 para cada ramo esquerdo e 1 para cada ramo direito.

A redundância residual do código de Huffman decorre da atribuição de um número inteiro de bits a cada letra do alfabeto. Pode-se tornar o processo mais eficiente agrupando-se em blocos os elementos gerados pela fonte e tratando estes blocos como os símbolos a serem codificados. Quanto maior o bloco, menor a redundância (Bell *et al.*, 1990). Este procedimento é limitado a poucos símbolos em cada bloco, por uma série de motivos práticos, como o crescimento exponencial da tabela de probabilidades.

4.2.2 Codificação de Golomb e de Golomb-Rice

Em 1966, Golomb propôs um esquema simples de codificação de valores inteiros não negativos (Weinberger *et al.*, 1996). Dado um parâmetro inteiro positivo m , o código de Golomb de um inteiro não negativo n é a concatenação da *codificação unária* de $\lfloor n / m \rfloor$ com a *codificação binária ajustada* de $n \bmod m$.

A codificação unária de um inteiro não negativo v é uma seqüência de v bits com valor '1' finalizada por um bit com valor '0'. Para descrever a codificação binária ajustada (Howard, 1994), considere-se um alfabeto de tamanho $m = 2^k + b$, composto pela seqüência de inteiros $0, 1, 2, \dots, 2^k + b - 1$, onde k é o maior inteiro tal que $2^k \leq m$, e b é um inteiro não negativo. Na representação binária convencional, cada símbolo deste alfabeto seria codificado com $\lceil \log_2 m \rceil$ bits. Assim, as palavras-código teriam k bits se $b = 0$, e $k+1$ bits se $b \neq 0$. Para $b = 0$, a codificação binária ajustada é idêntica à codificação binária convencional mas, para $b \neq 0$, parte dos símbolos é codificada com k bits, e o restante com $k+1$ bits. Mais especificamente, a codificação binária ajustada

de um símbolo no intervalo $[0, 1, \dots, 2^k - b - 1]$ é simplesmente sua representação binária convencional com k bits, enquanto a codificação binária ajustada de um símbolo de valor w fora deste intervalo é a representação binária, com $k+1$ bits, de $w + m - 2b$. Para $m = 5$, por exemplo, tem-se $k = 2$ e $b = 1$. Neste caso, os símbolos 0, 1 e 2, na codificação binária ajustada, tornam-se 00, 01 e 10, respectivamente, enquanto que os símbolos 3 e 4 são codificados como 110 e 111.

O código de Golomb é ótimo (Gallager e Voorhis, 1975) para distribuições de probabilidade geométricas, também conhecidas como *geométricas unilaterais*, dadas por

$$P(x = n) = (1 - \rho)\rho^n \quad (4.9)$$

onde $P(x = n)$ é a probabilidade de x assumir o valor inteiro $n \geq 0$, e $0 < \rho < 1$. Pode-se mostrar (Mood *et al.*, 1974) que a variância desta distribuição é dada por

$$\sigma^2 = \frac{\rho}{(1 - \rho)^2} \quad (4.10)$$

Para distribuições geométricas unilaterais, definindo-se o parâmetro m como

$$m = \left\lceil \frac{\log(1 + \rho)}{\log \rho^{-1}} \right\rceil \quad (4.11)$$

o código de Golomb produz o menor comprimento médio dentre todos os códigos unicamente decodificáveis (Weinberger *et al.*, 1996).

A distribuição contínua de Laplace, também conhecida como exponencial bilateral ou dupla, centrada em zero, e a distribuição discreta geométrica bilateral centrada em zero têm sido frequentemente consideradas boas aproximações para a distribuição de erros de predição e de coeficientes DCT e *wavelet* (Weinberger *et al.*, 1996; Seroussi e Weinberger, 1997; Ratnakar, 1997; Wu, 1997; Hans, 1998; Motta *et al.*, 2000; Usevitch, 2001). Com base em experimentos envolvendo uma grande variedade de imagens, Reininger e Gibson (1983) concluíram que os coeficientes AC da DCT são mais precisamente modelados por distribuições laplacianas do que por qualquer outra distribuição, enquanto que os coeficientes DC são mais apropriadamente modela-

dos por distribuições gaussianas. Os erros de predição no esquema clássico de compressão denominado DPCM (*Delta Pulse Code Modulation*), bem como os coeficientes DCT dos erros de predição no esquema híbrido DPCM-DCT são invariavelmente modelados pela distribuição de Laplace (Rao e Yip, 1990).

A distribuição de Laplace de média zero é caracterizada pela densidade de probabilidade (Mood *et al.*, 1974):

$$f(x) = \frac{1}{2\beta} e^{-\frac{|x|}{\beta}} \quad (4.12)$$

com $\beta > 0$. A distribuição geométrica bilateral é dada por

$$P(x = n) = \frac{(1-\rho)\rho^{|n|}}{2} \quad (4.13)$$

com $0 < \rho < 1$.

O seguinte mapeamento de valores (Weinberger *et al.*, 1996) permite converter distribuições geométricas bilaterais em distribuições aproximadamente geométricas unilaterais, viabilizando o emprego do código de Golomb:

$$M(n) = \begin{cases} 2n & n \geq 0 \\ 2|n| - 1 & n < 0 \end{cases} \quad (4.14)$$

Rice (Howard, 1994) deteve-se no estudo do subconjunto dos códigos de Golomb com parâmetro $m = 2^k$. Este caso especial, que se tornou conhecido como codificação de Golomb-Rice, permite uma simplificação ainda maior, com o código para n sendo obtido pela concatenação da representação unária do valor n deslocado k bits para a direita, com os k bits menos significativos de n . Para distribuições geométricas unilaterais, a razão de compressão média obtida com um código de Golomb-Rice com parâmetro k ótimo é próximo à entropia (Seroussi e Weinberger, 1997).

Dada uma seqüência cujos elementos seguem uma distribuição geométrica bilateral, uma estimativa do valor ótimo do parâmetro k a ser empregado na codificação da seqüência obtida após o mapeamento expresso pela Equação (4.14) é (Weinberger *et al.*, 1996):

$$k = \lceil \log_2 a \rceil \quad (4.15)$$

onde a é a média dos módulos dos valores da seqüência antes do mapeamento.

A codificação de Golomb e de Golomb-Rice tem sido muito aplicada à compressão de imagens (Weinberger *et al.*, 1996; Motta *et al.*, 2000), áudio (Hans, 1998), ECG Koski (1997) e outros sinais. A Tabela 4.1 exemplifica a codificação de Golomb e de Golomb-Rice para diversos valores de n e m .

Tabela 4.1. Exemplos de códigos de Golomb e de Golomb-Rice.

n	$m=1$ $k=0$	$m=2$ $k=1$	$m=3$	$m=4$ $k=2$	$m=5$	$m=6$	$m=7$
0	0	00	00	000	000	000	000
1	10	01	010	001	001	001	0010
2	110	100	011	010	010	0100	0011
3	1110	101	100	011	0110	0101	0100
4	11110	1100	1010	1000	0111	0110	0101
5	111110	1101	1011	1001	1000	0111	0110
6	1111110	11100	1100	1010	1001	1000	0111
7	11111110	11101	11010	1011	1010	1001	1000
8	111111110	111100	11011	11000	10110	10100	10010
...

4.2.3 Codificação aritmética

Shannon mostrou que a representação binária mais eficiente para um símbolo que ocorre com probabilidade P requer um número de bits igual à auto-informação associada ao símbolo, ou seja, $-\log_2 P$ bits (ver Seção 3.2). Nos códigos de Huffman, Golomb e Golomb-Rice, entretanto, as palavras-códigos têm sempre um número inteiro de bits, o que impede que o limite teórico de compressão, dado pela entropia, seja atingido, exceto em casos específicos.

A ineficiência decorrente desta limitação torna-se substancial quando a probabilidade de ocorrência de um dos símbolos é superior à 0,5. Se uma fonte gera, por exemplo, três símbolos a , b e c , com probabilidades $P(a) = 98/100$, $P(b) = 1/100$ e $P(c) = 1/100$, a codificação ideal atribuiria 0,029 bits ao símbolo a e 6,645 bits a cada um dos outros símbolos, resultando em um comprimento médio, igual à entropia da fonte, de 0,161 bits/símbolo. O código de Huffman, por sua vez, atribuiria 1 bit ao símbolo a e 2 bits a cada um dos outros símbolos, para um comprimento médio de 1,02 bits/símbolo, mais de seis vezes acima da entropia.

O codificador aritmético elimina a associação entre símbolos individuais e palavras-códigos de comprimento inteiro e, com isto, é capaz de praticamente igualar a entropia da fonte em todos os casos. A codificação aritmética pode ser descrita como se segue:

1. Cria-se um *intervalo corrente* iniciado com $[0, 1)$
2. Para cada elemento da mensagem,
 - 2.1 Particiona-se o intervalo corrente em subintervalos, um para cada letra do alfabeto. O tamanho do subintervalo associado a uma dada letra é proporcional à probabilidade de que esta letra seja o próximo elemento da mensagem, de acordo com o modelo assumido.
 - 2.2 O subintervalo correspondente à letra que é realmente o próximo elemento é selecionado como novo intervalo corrente.
3. Codifica-se a mensagem com o menor número de bits necessário para distinguir o intervalo corrente final de todos os outros possíveis intervalos correntes finais.

Decorre imediatamente do algoritmo que o comprimento do intervalo corrente final é igual ao produto das probabilidades das letras que constituem a mensagem, que é também a probabilidade de ocorrência desta mensagem específica. Representando por p esta probabilidade, o número de bits necessários para codificar a mensagem, no passo 3 do algoritmo, é $\lceil -\log_2 p \rceil$ (Howard, 1994).

Há dois problemas práticos principais para a implementação de um codificador aritmético seguindo a descrição básica apresentada: o cálculo dos limites do intervalo corrente para mensagens genéricas (potencialmente grandes) exige aritmética de altíssima precisão; e nenhuma codificação é produzida antes que toda a mensagem tenha sido processada. Vários mecanismos para transmissão incremental e aritmética de precisão fixa foram desenvolvidos ao longo dos anos (Bell *et al.*, 1990). Howard (1994) provou analiticamente que, com os mecanismos propostos por Witten, Neal e Cleary (1987), a redução na eficiência de compressão é insignificante. Pode-se, desta forma, construir codificadores aritméticos executáveis em computadores atuais e capazes de atingir o limite teórico de compressão dado pela entropia, para todos os efeitos práticos.

Uma descrição detalhada dos aspectos práticos da codificação aritmética, incluindo implementações completas na linguagem de programação "C", é fornecida por Bell *et al.* (1990).

4.3 CODIFICAÇÃO POR COMPRIMENTO DE SEQÜÊNCIA

Quando há uma elevada probabilidade de ocorrência de símbolos iguais consecutivos em uma mensagem, uma técnica eficiente de compressão é a codificação por comprimento de seqüência (*Run-Length Encoding* - RLE). A idéia consiste em substituir as seqüências de símbolos repetidos por contadores contendo o número de símbolos na seqüência. Pode-se empregar um símbolo especial para indicar que o próximo elemento na seqüência é um contador, e não um símbolo da fonte. É necessário, também, indicar qual é o símbolo que se repete. Assim, cada seqüência será representada por três elementos: um indicador, um contador e um símbolo. O RLE é a base de alguns formatos gráficos, como o PCX (Binder, 1994).

Há casos em que apenas um determinado símbolo apresenta alta probabilidade de ocorrer em longas seqüências. Por exemplo, a quantização dos coeficientes DCT gera tipicamente seqüências de zeros, podendo-se empregar RLE exclusivamente para estas seqüências, eliminando a necessidade de informar qual o símbolo que se repete. Tais seqüências podem ser representadas por apenas dois elementos: um indicador e um contador.

Quando a probabilidade de ocorrências de seqüências de zero é especialmente elevada, pode-se reduzir ainda mais a representação RLE considerando-se que cada seqüência de zero é interrompida por um elemento não nulo, e a cada elemento não nulo segue-se uma seqüência de zeros. Sempre que dois ou mais elementos não nulos ocorrerem em sucessão, considera-se para efeito de codificação que eles são separados por uma seqüência de zeros de comprimento zero. Esta alternância obrigatória torna o indicador desnecessário, e permite que as seqüências sejam representadas unicamente pelo contador (Wallace, 1991).

4.4 QUANTIZAÇÃO

Muitos sinais de interesse são, por natureza, de amplitude e tempo contínuos. Os conversores A/D, através das operações de amostragem e quantização, permitem a representação discreta destas formas de onda. Neste contexto, a quantização é o processo de transformar amplitudes analógicas em discretas (Jayant e Noll, 1984), sendo um fator determinante para a qualidade do sinal digital e para a taxa de bits necessária à sua representação.

O processo de quantização não se aplica unicamente à conversão de sinais analógicos em digitais. Se o sinal é digital por natureza, ou foi previamente convertido a esta forma, pode-se empregar quantização com o objetivo de reduzir a sua representação binária. Neste caso, a quantização atua essencialmente por redução de alfabeto: uma amplitude (ou um grupo de amplitudes), selecionada de um conjunto com M elementos, é transformada pela quantização em uma amplitude (ou um grupo de amplitudes) selecionada de um conjunto com K elementos, $K < M$. Quando os elementos dos conjuntos representam amplitudes individuais, tem-se uma *quantização escalar*; quando representam grupos de amplitudes, tem-se uma *quantização vetorial*.

Teoricamente, a quantização vetorial é mais eficiente, em termos de taxa-distorção, principalmente quando as amostras são estatisticamente dependentes ou correlacionadas. Por outro lado, a quantização escalar apresenta menor complexidade computacional. Em conseqüência, é comum o emprego de quantização escalar, com a inclusão de uma etapa prévia de decorrelação. Quando bem projetada, esta associação muitas vezes apresenta um compromisso $R-D$ superior ao obtido por muitos esquemas de quantização vetorial, cuja implementação prática pode exigir simplificações intensas.

4.4.1 Quantização Escalar

Um *quantizador escalar de K níveis* é uma função $Q(x)$ que mapeia dos reais para um conjunto de K números $\{r_0, r_1, \dots, r_{K-1}\}$:

$$Q(x) = r_k \text{ se } d_k \leq x < d_{k+1}, k = 0, \dots, K-1 \quad (4.16)$$

com $d_0 = -\infty$, $d_{K-1} = \infty$, e $d_0 < r_0 < d_1 < r_1 < \dots < r_{K-1} < d_K$. Os números d_k , $k = 0, \dots, K$, são denominados *níveis de decisão*; os intervalos $[d_k, d_{k+1})$, $k = 0, \dots, K-1$, são os *intervalos de decisão*; e os números r_k , $k = 0, \dots, K-1$, são os *níveis de reconstrução*. No processo de quantização, portanto, um elemento com valor x , gerado pela fonte, será substituído por r_k se $d_k \leq x < d_{k+1}$.

Quando os níveis de decisão são igualmente espaçados, diz-se que a quantização é *uniforme*, e o espaçamento é denominado *tamanho do passo de quantização*. Os níveis de reconstrução não são necessariamente uniformemente espaçados, podendo ser definidos de forma a minimizar a distorção. Uma simplificação usual é posicionar os níveis de reconstrução no centro dos intervalos de decisão:

$$r_k = \frac{d_k + d_{k+1}}{2} \quad (4.17)$$

Como a quantização é um mapeamento de muitos-para-um, o processo é irreversível. Se a quantização produz um nível de reconstrução r_k em um determinado momento, tudo o que se pode concluir é que o valor originalmente gerado pela fonte localizava-se no intervalo $[d_k, d_{k+1})$, mas não há como saber qual era o valor exato. Em outras palavras, há perda de informação no proces-

so, ou seja, a quantização gera distorção, e qualquer compressor que a utilize é, conseqüentemente, um compressor com perdas.

O projeto de um quantizador envolve a definição da quantidade e dos valores dos níveis de decisão e de reconstrução. Estes fatores, por sua vez, afetam a taxa de bits e a distorção. O problema da quantização ótima pode ser expresso de maneira análoga à Formulação 3.1: dada uma distorção D_T , definir os níveis de decisão e de reconstrução de forma a minimizar R , com $D = D_T$. Para algumas fontes com f.d.p. conhecidas, quantizadores ótimos podem ser analiticamente projetados (Sayood, 1996).

Em um compressor/descompressor, a quantização é dividida em duas etapas. A primeira ocorre durante a compressão, e consiste em substituir os valores gerados pela fonte por identificadores (valores inteiros codificados em binário, por exemplo) para os intervalos de decisão em que se localizam. Na segunda etapa, que ocorre durante a descompressão, estes identificadores são substituídos pelos níveis de reconstrução correspondentes aos intervalos. Tornou-se comum empregar o termo *quantização* para designar a primeira etapa do processo, e o termo *dequantização* para a segunda etapa (Wallace, 1991). Por este motivo, os identificadores dos intervalos, gerados durante o processo de quantização, são muitas vezes chamados de *valores quantizados*, enquanto os níveis de reconstrução são chamados de *valores dequantizados*. Esta será a terminologia utilizada neste trabalho.

Quando um dos níveis de decisão possui valor zero, o quantizador é denominado de *midriser*; caso contrário, tem-se um quantizador *midtread*. Por sua simplicidade, o quantizador uniforme *midtread* com níveis de reconstrução no centro dos intervalos de decisão, que será identificado por QUMC (Quantizador Uniforme, *Midtread*, reconstrução Central), é provavelmente o quantizador mais utilizado para compressão de imagens (Ratnakar, 1997). O QUMC pode ser implementado através de divisão seguida por arredondamento para o inteiro mais próximo. Se x representa um valor gerado pela fonte de informação, o valor quantizado \hat{x} é obtido através da seguinte operação:

$$\hat{x} = x // q \quad (4.18)$$

onde // denota divisão seguida por arredondamento para o inteiro mais próximo e q é um valor positivo. Pela própria definição da operação, $\hat{x} \in \mathbb{Z}$. O valor dequantizado, ou nível de reconstrução, \tilde{x} , é dado por:

$$\tilde{x} = q \hat{x} \quad (4.19)$$

Um número real qualquer, v , é arredondado para o inteiro i se e somente se $i-0,5 \leq v < i+0,5$. Logo, na divisão seguida por arredondamento, um valor x será quantizado para i , e conseqüentemente dequantizado para o nível de reconstrução qi , quando $i-0,5 \leq x/q < i+0,5$, ou seja, $(i-0,5)q \leq x < (i+0,5)q$. Vê-se, portanto, que o nível de reconstrução qi encontra-se posicionado no centro do intervalo de decisão $[(i-0,5)q, (i+0,5)q)$, para $i = \dots-2, -1, 0, 1, 2, \dots$. Como o espaçamento entre os níveis de decisão é q , a quantização por divisão seguida por arredondamento para o inteiro mais próximo é uniforme, sendo q o tamanho do passo de quantização.

Pode-se elaborar um pouco mais o processo empregando *quantização com zona morta*, dada por

$$\hat{x} = \begin{cases} 0 & \text{se } |x| < t \\ x // q & \text{caso contrário} \end{cases} \quad (4.20)$$

onde t é um valor positivo real que define o limiar abaixo do qual todos os elementos são quantizados para zero. Dado o valor quantizado \hat{x} , o nível de reconstrução é dado pela Equação (4.19).

Ratnakar (1997) comparou analiticamente a relação $R-D$ de quatro esquemas de quantização para densidades laplacianas: o QUMC; o quantizador com zona morta e níveis de reconstrução no centro dos intervalos de decisão (Quantizador com Zona Morta, reconstrução Central - QZMC); o quantizador uniforme *midtread* com níveis de reconstrução definidos de forma a minimizar a distorção (Quantizador Uniforme, *Midtread*, reconstrução Ótima - QUMO); e o quantizador uniforme *midriser* com níveis de reconstrução definidos de maneira a minimizar a distorção (Quantizador Uniforme, *midRiser*, reconstrução Ótima - QURO). O QUMO e o QZMC apresentam desempenho destacadamente superior, enquanto que o QURO apresenta o pior desempenho, notadamente para distorções elevadas. As curvas $R-D$ do QUMO e do QZMC praticamente coincidem em todos os níveis de distorção analisados, sendo o QUMO marginalmente superior. De fato, Papat (1990) mostrou que, para fontes gaussianas, laplacianas e gama, o QUMO apresenta

performance próxima à ótima para uma larga faixa de taxa de bits. Por outro lado, o QZMC é de implementação muito mais simples, e o descompressor necessita conhecer apenas o valor de q para efetuar corretamente a dequantização.

4.4.2 Quantização Vetorial

Shannon (1948) mostrou que codificar elementos agrupados em vetores é teoricamente mais eficiente que codificá-los isoladamente. Este resultado permanece válido mesmo quando os elementos gerados são independentes. Na quantização vetorial, a mensagem gerada pela fonte é particionada em blocos ou vetores com L elementos consecutivos. Um número K de vetores, denominados *vetores-código*, é selecionado, dentre todos aqueles que podem ser gerados pela fonte, e armazenados em uma tabela conhecida como *codebook*, disponível para o compressor e para o descompressor. Sejam \mathbf{c}_k , $k = 0, 1, \dots, K-1$, o k -ésimo vetor-código do *codebook*; c_{kj} , $j = 0, 1, \dots, L-1$, o j -ésimo elemento de \mathbf{c}_k ; \mathbf{v}_i , $i = 0, 1, \dots$, o i -ésimo vetor na mensagem produzida pela fonte; e v_{ij} , $j = 0, 1, \dots, L-1$, o j -ésimo elemento de \mathbf{v}_i . Para codificar \mathbf{v}_i , o quantizador calcula a distorção $D(\mathbf{v}_i, \mathbf{c}_k)$, $k = 0, 1, \dots, K-1$, entre este vetor e cada um dos vetores-código. A distorção é freqüentemente dada pelo erro médio quadrático:

$$D(\mathbf{v}_i, \mathbf{c}_k) = \frac{1}{L} \sum_{j=0}^{L-1} (v_{ij} - c_{kj})^2 \quad (4.21)$$

Se $D(\mathbf{v}_i, \mathbf{c}_k)$ é mínima para $k = k'$, o índice k' é anexado à mensagem comprimida. Durante a descompressão, k' é substituído pelo vetor-código correspondente, $\mathbf{c}_{k'}$.

Note-se que, enquanto o compressor realiza uma série de comparações para encontrar o vetor-código que mais se aproxima do vetor de entrada, o descompressor precisa apenas retirar de uma tabela o vetor indexado. Isto faz da quantização vetorial uma opção atrativa para aplicações em que os recursos computacionais disponíveis durante a descompressão são consideravelmente menores que os disponíveis durante a compressão.

Para códigos de comprimento fixo, os índices são codificados com $\log_2(K)$ bits. Como cada índice representa, na mensagem codificada, um vetor com L elementos, a taxa resultante é de $\log_2(K)/L$ bits/elemento.

Teoricamente, quanto maior L , mais eficiente a quantização. Na prática, a dimensão dos vetores é limitada por considerações de eficiência de processamento. O desempenho depende ainda do tamanho do *codebook*; da escolha dos vetores-código; e da medida de distorção utilizada. O *codebook* é gerado por um procedimento de treinamento, durante o qual uma quantidade normalmente elevada de mensagens é examinada. Os vetores considerados mais representativos (de acordo com o método de treinamento empregado) do conjunto de mensagens são selecionados como vetores-código. Desde sua introdução em 1980, o algoritmo LBG tem sido um método de treinamento muito popular (Sayood, 1996), mas uma grande variedade de técnicas têm sido apresentadas. No compressor de ECG desenvolvido por McAuliffe (1993), por exemplo, o *codebook* é criado com o auxílio de redes neurais.

Se as mensagens de treinamento representam adequadamente as características encontradas nas mensagens reais, e se o *codebook* é bem definido, a distorção causada pela quantização será baixa. Em muitas situações reais, no entanto, o elevado grau de variabilidade de aspectos relevantes das mensagens geradas pela fonte exigiria um conjunto de treinamento e um *codebook* de grandes dimensões, implicando tempo de treinamento e de compressão elevado. Tal é o caso dos sinais de ECG (Miaou e Yen, 2000). Para atenuar o problema, pode-se empregar quantização vetorial adaptativa (QVA). Na QVA, quando a distorção entre um vetor de entrada e o vetor-código mais próximo a ele (no sentido de menor erro médio quadrático, por exemplo) é superior a um limiar *dth*, o vetor de entrada é inserido no *codebook*, em substituição a um dos seus vetores-código, de acordo com alguma regra de atualização (Miaou e Yen, 2000).

4.5 TRANSFORMADAS

Na área de compressão de dados, as transformadas são principalmente aplicadas a sinais e imagens. A utilidade da transformada para a compressão depende de sua capacidade de gerar coeficientes descorrelacionados e de concentrar a maior parte da energia do sinal em um pequeno

número de coeficientes (Rao e Yip, 1990). A transformada pode ser vista como um pré-processamento que reduz a correlação entre as amostras, permitindo com isto o emprego eficiente de quantização escalar e uma codificação baseada em modelos simples (Goyal, 2001). O par transformada/quantização escalar pode apresentar desempenho semelhante ao da quantização vetorial, com complexidade de compressão consideravelmente menor (Ratnakar, 1997). De fato, a despeito da profícua e bem consolidada teoria que fundamenta o projeto dos quantizadores, a maior parte das técnicas modernas combina a simplicidade da quantização escalar e uniforme com predição ou transformada (Goyal, 2001).

Quando o sinal se comporta como um processo estocástico do tipo Markov-1, pode-se mostrar que a transformada Karhunen-Loeve (*Karhunen-Loeve Transform* - KLT) é ótima. No entanto, a implementação da KLT é elaborada, exigindo a estimação da matriz de autocovariância da seqüência a ser comprimida e sua diagonalização, bem como o cálculo da transformada propriamente dita, de forma que transformadas subótimas são utilizadas na prática. A DCT e a DWT, duas das transformadas mais empregadas em esquemas práticos de compressão de dados, são apresentadas a seguir.

4.5.1 A DCT

Há quatro versões da DCT, identificadas aqui como DCT-I, DCT-II, DCT-III e DCT-IV (Rao e Yip, 1990). Por sua maior capacidade de decorrelação e concentração de energia em poucos coeficientes, a DCT-II é a mais utilizada para compressão de dados.

Se x_n , $n = 0, 1, \dots, N-1$, são os elementos de uma seqüência \mathbf{x} , a DCT-II unidimensional de \mathbf{x} gera um vetor \mathbf{X} cujos coeficientes X_k , $k = 0, 1, \dots, N-1$, são dados por:

$$X_k = \left(\frac{2}{N}\right)^{1/2} c_k \sum_{n=0}^{N-1} x_n \cos\left[\frac{(2n+1)k\pi}{2N}\right], \quad k = 0, 1, \dots, N-1 \quad (4.22)$$

onde

$$c_k = \begin{cases} (1/2)^{1/2} & \text{para } k = 0 \\ 1 & \text{para } k = 1, 2, \dots, N-1 \end{cases} \quad (4.23)$$

Esta operação leva vetores de um determinado domínio (comumente, do domínio do tempo ou do espaço) para o domínio da frequência (Wallace, 1991). O coeficiente X_0 , que representa o componente de frequência zero e está relacionado diretamente com o valor médio de x , é conhecido como *coeficiente DC* (de *direct current*); os demais são os *coeficientes AC* (de *alternating current*).

A DCT apresenta desempenho próximo ao ótimo para processos do tipo Markov-1 com valores positivos elevados de correlação entre amostras adjacentes (Rao e Yip, 1990). Este fato, aliado à existência de algoritmos rápidos para o cálculo da DCT, tornou esta transformada extremamente popular na área de compressão de dados. Quando aplicada a sinais e imagens, a DCT normalmente concentra a maior parte da energia em poucos coeficientes. Para sinais de ECG, os resultados de Ahmed *et al.* (1975) mostram que a distribuição de variâncias dos coeficientes DCT e KLT é muito semelhante, o que confirma a boa capacidade de concentração de energia da DCT também para este tipo específico de sinal.

A seqüência x pode ser recuperada aplicando-se a X a DCT-II inversa:

$$x_n = \left(\frac{2}{N}\right)^{1/2} \sum_{k=0}^{N-1} c_k X_k \cos\left[\frac{(2n+1)k\pi}{2N}\right], \quad n = 0, 1, \dots, N-1 \quad (4.24)$$

Mesmo havendo algoritmos eficientes para o cálculo da DCT, o custo computacional quando a seqüência x contém muitos elementos pode ser elevado. Quando aplicada para compressão de dados, é prática comum particionar a seqüência em blocos consecutivos com tamanho fixo L_b , e efetuar o processamento bloco a bloco. Se L_b é muito pequeno tem-se uma degradação na RC; blocos demasiadamente grandes, por outro lado, exigem um longo tempo de processamento. Diversos experimentos indicam que aumentar o tamanho dos blocos acima de um certo ponto resulta em melhorias marginais ou mesmo em degradações no compromisso taxa-distorção, enquanto que o tempo de processamento aumenta significativamente (Nelson e Gailly, 1996; Wallace, 1991; Lee e Buckley, 1999). O valor ótimo para L_b depende da seqüência a ser comprimida. O padrão JPEG para compressão de imagens, por exemplo, emprega blocos com 8×8 *pixels*, apesar

de existirem indícios de que é possível melhorar consideravelmente a relação entre compressão e distorção utilizando-se blocos com tamanho 16x16 (Nelson e Gailly, 1996).

4.5.2 A DWT

As *wavelets*, aqui representadas por $\psi_{a,b}(t)$, são funções geradas por dilatações e translações de uma função ψ denominada *wavelet-base*:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (4.25)$$

com $a \in \mathbf{R}^+$ e $b \in \mathbf{R}$.

A transformada *wavelet* de uma função $f(t)$ gera uma seqüência de coeficientes W_{ab} dados por:

$$W_{ab} = \int_{-\infty}^{\infty} \psi_{a,b}(t) f(t) dt \quad (4.26)$$

com $a, b \in \mathbf{Z}$.

A função $f(t)$ pode ser recuperada através da transformada inversa:

$$f(t) = \sum_{j,k} W_{jk} \psi_{j,k}(t) \quad (4.27)$$

com $j, k \in \mathbf{Z}$.

Para $a = 2^\omega$ e $b = n2^\omega$, com $\omega, n \in \mathbf{Z}$ tem-se:

$$\psi_{\omega,n}(t) = \frac{1}{\sqrt{2^\omega}} \psi\left(\frac{t-2^\omega n}{2^\omega}\right) \quad (4.28)$$

É possível escolher ψ de tal forma que $\{\psi_{o,n}\}$ constitui uma base ortonormal (Wei, 1995). Quando a escolha de ψ é apropriada, a transformada *wavelet* concentra a maior parte da energia do sinal em poucos coeficientes e, com isto, muitos coeficientes podem ser eliminados com pequena distorção na reconstrução (Lu *et al.*, 2000).

As transformadas *wavelet* são comumente implementadas como bancos de filtros de reconstrução perfeita estruturados em árvore. Como no caso da DCT, o sinal de entrada é dividido em blocos que são então independentemente transformados. Para tanto, cada bloco é filtrado por um par de filtros de análise, sendo um passa-baixas e um passa-altas, e reamostrado em uma frequência igual à metade da original. As operações de filtragem e reamostragem são aplicadas recursivamente ao sinal passa-baixas reamostrado, gerando coeficientes *wavelet* organizados em camadas (Lu *et al.*, 2000).

A definição dos filtros, que equivale à definição das *wavelets*, é uma etapa crítica para o resultado da compressão, devendo ser efetuada de acordo com as características específicas do sinal em questão (Strang e Nguyen, 1996; Hilton, 1997). Não há até o presente momento resultados teóricos que indiquem que *wavelet* é mais apropriada para compressão de um dado sinal; a escolha é normalmente feita a partir da comparação experimental do desempenho de várias *wavelets* (Hilton, 1997). Deve-se destacar que a escolha dos filtros envolve um compromisso entre compressão, geralmente maior com filtros longos, e eficiência computacional, obviamente maior com filtros mais curtos. Além disso, o tamanho dos blocos e o número de camadas da transformada também afetam a compressão e o tempo de processamento, e devem ser selecionados apropriadamente. Vê-se, portanto, que a aplicação eficiente da DWT exige a definição de mais fatores do que a DCT, onde a única consideração essencial, sob o ponto de vista de taxa de bits e complexidade computacional, é o tamanho dos blocos. A necessidade da definição cuidadosa destes parâmetros, essencial para obter-se um bom compromisso taxa-distorção, torna a utilização eficiente da DWT mais complexa que a da DCT.

4.6 O PARADIGMA TRANSFORMADA + QUANTIZAÇÃO + CODIFICAÇÃO

A Figura 4.1 apresenta o diagrama em blocos de um compressor/descompressor baseado no paradigma transformada + quantização + codificação. O sinal de entrada é primeiramente particionado em blocos. O sinal comprimido é composto pelos coeficientes das transformadas dos blocos, quantizados e codificados. Durante a descompressão, a transformada inversa é aplicada aos coeficientes decodificados e dequantizados, gerando uma aproximação \tilde{x} para x . Devido à quantização, em geral $\tilde{x} \neq x$.

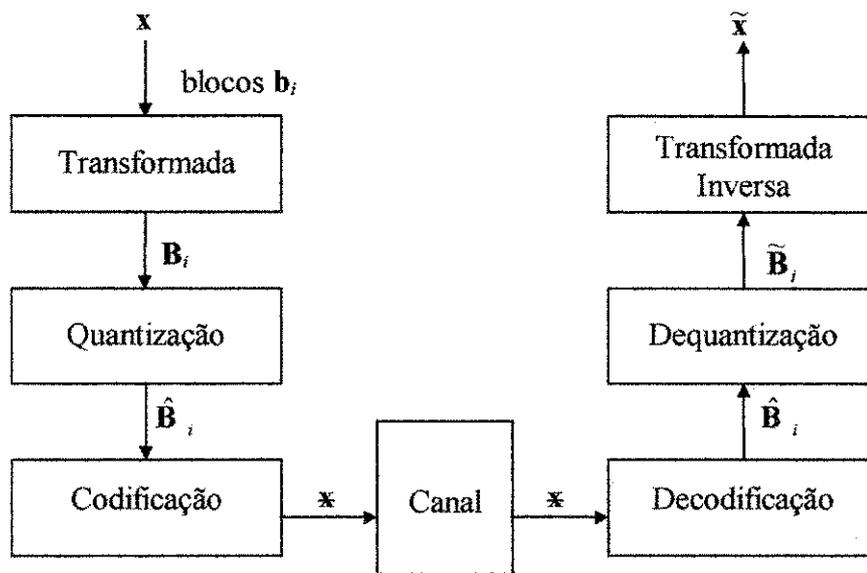


Figura 4.1 Diagrama em blocos de um compressor/descompressor baseado em transformada + quantização + codificação

A eficiência do esquema, em termos de taxa e distorção, depende da transformada, da estratégia de quantização e da codificação. Uma transformada apropriada gera coeficientes pouco correlacionados e concentra a maior parte da energia do bloco em um pequeno número de coeficientes. A quantização pode então ser efetuada de forma diferenciada: os coeficientes que concentram uma parcela menor da energia de x são menos importantes para a sua reconstrução, podendo ser quantizados mais grosseiramente, ou seja, com maior distorção e menor taxa. Apesar da descorrelação proporcionada pela transformada não implicar independência (a não ser para o caso gaussiano), observa-se na prática que o aumento de compressão da quantização vetorial com

relação à escalar, para elementos descorrelacionados, normalmente não compensa o acréscimo na complexidade (Ratnakar, 1997).

Após a quantização, um modelo estatístico dos coeficientes quantizados é construído e passado ao codificador. Tipicamente observa-se que, nos esquemas de compressão baseados em transformada e quantização, os blocos quantizados apresentam grandes seqüências de coeficientes nulos, de maneira que a etapa final é muitas vezes uma combinação de RLE e codificação de entropia (Wallace, 1991).

O projeto de um compressor deste tipo envolve, entre outros fatores, a especificação dos níveis de decisão e de reconstrução. Quando há liberdade para a determinação independente do número e dos valores dos níveis para cada coeficiente da transformada, o projeto de um quantizador ótimo é um problema de elevada complexidade (Ratnakar, 1997). Na prática, algumas limitações são impostas para a abordagem de quantização, com o intuito de reduzir a liberdade na escolhas dos parâmetros e, conseqüentemente, a complexidade da tarefa. Descreve-se a seguir, em linhas gerais, algumas soluções adotadas em compressores práticos para o projeto de quantizadores.

Uma abordagem simples consiste em definir um quantizador uniforme ótimo com a restrição de utilização de um único tamanho de passo de quantização para todo o conjunto de coeficientes da transformada. Sabendo-se que o comportamento estatístico dos diversos coeficientes apresenta tipicamente amplas variações, tal restrição pode ter um forte impacto negativo na relação $R-D$.

Um esquema um pouco mais elaborado consiste em restringir a distorção em cada coeficiente a um valor único especificado. Nesta situação, coeficientes diferentes serão em geral quantizados com tamanhos de passos de quantização diferentes e, para quantização com zona morta, os limiares podem também variar de um coeficiente para outro. O problema consiste agora em definir um *vetor de quantização* e um *vetor de limiares*, cujos elementos são, respectivamente, os tamanhos de passos de quantização e os limiares empregados.

Os resultados desta abordagem nem sempre são satisfatórios. Sejam $R^i(D)$ e $R^j(D)$, respectivamente, funções contínuas com derivada contínuas representando aproximações para as bordas convexas inferiores dos planos $R-D$ operacionais (ver Seções 3.4.1 e 3.4.2) de dois coeficientes i e j quaisquer. Tal aproximação é aceitável quando o plano $R-D$ é suficientemente denso. Se a

distorção alvo é D_T , pela abordagem em questão ambos os coeficientes deverão ser quantizados com um nível de distorção D_T , com taxa média resultante $(R^i(D_T) + R^j(D_T))/2$. Considere-se agora um acréscimo infinitesimal na distorção do coeficiente j , compensado por um decréscimo idêntico na distorção do coeficiente i , de forma que a distorção média não se altera. Se

$$\frac{dR^i(D)}{dD} < \frac{dR^j(D)}{dD}$$

a redução na taxa de j devida ao aumento na sua distorção será maior do que o aumento na taxa de i devido à redução em sua distorção, obtendo-se com isto redução na taxa média sem alteração na distorção média. Quando as distorções nos coeficientes são tais que

$$\frac{dR^i(D)}{dD} = \frac{dR^j(D)}{dD}$$

obtem-se a taxa média mínima para a distorção média D_T . O resultado pode ser estendido para um número qualquer de coeficientes (Goyal, 2001). Logo, o projeto de quantizadores com distorção única para todos os coeficientes pode também levar a uma degradação considerável no compromisso $R-D$, principalmente no caso frequente em que as derivadas das BCIs, calculadas no ponto correspondente à distorção alvo, diferem amplamente entre os coeficientes.

O projeto de um quantizador uniforme com zona morta, com tamanhos de passo de quantização e limiares diferentes para cada coeficiente, definidos de maneira que a distorção média não ultrapasse um dado valor, envolve um maior nível de complexidade. Para buscas restritas à borda convexa, a minimização da taxa média para uma dada distorção média alvo pode ser obtida por minimização lagrangiana (Ratnakar, 1997) independentemente aplicada a cada um dos planos operacionais dos coeficientes, com o mesmo valor para o multiplicador λ . Para cada coeficiente, a solução ótima consiste no último ponto interceptado pela varredura, em direção à origem, efetuada por uma reta com declividade $-\lambda$. O ponto ótimo para cada um dos coeficientes localiza-se, portanto, na região da BCI com derivada aproximadamente igual a $-\lambda$ (a igualdade não é exata devido à natureza discreta dos planos operacionais), o que está de acordo com os resultados apresentados nos parágrafos anteriores.

Capítulo 5

DESENVOLVIMENTOS RECENTES EM COMPRESSÃO BASEADA EM DWT E DCT

Neste capítulo, serão apresentados alguns compressores baseados em DWT e em DCT descritos na literatura especializada nos últimos dez anos, conferindo-se especial atenção às etapas responsáveis pelo controle das distorções. Além de proporcionar uma visão geral das soluções adotadas na prática para os problemas inerentes a este paradigma de compressão, este capítulo fornece um indicativo de que muitos métodos descritos na literatura adotam simplificações, visando reduzir as exigências computacionais, que podem conduzir a degradações substanciais na relação taxa-distorção. Em consequência, técnicas que empregam uma quantização cuidadosamente elaborada, como o SPIHT e o RD-OPT, descritos nas Seções 5.1 e 5.2, respectivamente, sobressaem-se quando comparadas com aquelas que incorrem em uma simplificação excessiva do processo de controle de perdas.

5.1 COMPRESSÃO BASEADA EM DWT

(Crowe *et al.* 1992)

No compressor de ECG proposto por Crowe *et al.* (1992), os coeficientes *wavelet* com valor inferior a um limiar t predefinido são descartados. O sinal comprimido contém os coeficientes preservados, juntamente com os seus índices. Para os testes, foram usados sinais de ECG amostrados a 500 Hz, contendo 1026 amostras. O grau de compressão foi aferido pela razão entre o número de amostras do sinal original e o número de coeficientes *wavelet* preservados. Observe-se que esta medida é imprecisa, visto que os índices dos coeficientes preservados também são incluídos no sinal comprimido, e os próprios coeficientes devem ser armazenados em ponto flutuante.

Três valores de t foram utilizados, levando a razões de compressão de 4,1:1, 9,9:1, e 27,0:1. Nas palavras dos autores, as reconstruções preservam adequadamente os complexos QRS, mas distorcem consideravelmente as ondas P e T. Na razão 27:1, a forma da *wavelet*-base é claramente visível no sinal reconstruído.

(Shapiro, 1993)

No EZW (Shapiro, 1993), os coeficientes *wavelet* de um bloco de imagem são estruturados em uma árvore onde cada coeficiente c tem quatro filhos na escala seguinte de resolução mais elevada, em posições correspondentes à mesma localização espacial que c .

A idéia fundamental do EZW deriva diretamente da seguinte constatação empírica: quando um coeficiente é menor que um limiar t , há uma probabilidade elevada de que todos os seus descendentes sejam também menores que t . Partindo desta observação, Shapiro estabeleceu para a etapa de quantização do EZW a simples anulação dos coeficientes menores que t . Para cada coeficiente zerado após a comparação com t , há uma alta probabilidade de que os seus descendentes sejam também zerados, originando uma *árvore de zeros (zerotree)*, que é codificada com um símbolo especial. A taxa desejada é atingida experimentando-se uma sucessão de limiares crescentes. Quando divulgado, o algoritmo proposto por Shapiro, denominado EZW (*embedded zerotree wavelet*), foi considerado comparável ou superior aos melhores compressores de imagens até então conhecidos.

(Said e Pearlman, 1996)

O compressor SPIHT (*set partitioning in hierarchical trees*) (Said e Pearlman, 1996) é uma extensão das idéias do EZW. Os valores dos limiares são potências de dois, correspondendo a uma ordenação por planos de bits. O ganho de compressão do SPIHT com relação ao EZW provém do elaborado esquema de codificação empregado. A codificação é tão eficiente que a relação taxa-distorção é apenas marginalmente melhorada com o emprego adicional de codificação de entropia. O algoritmo SPIHT é considerado um dos mais eficientes compressores de imagens da atualidade. Recentemente, Lu *et al.* (2000) adaptaram com sucesso o método para ECG, obtendo um compromisso taxa-distorção semelhante aos melhores resultados da literatura.

(Strang e Nguyen, 1996)

No compressor de ECG de Strang e Nguyen (1996), o sinal é dividido em blocos adjacentes aos quais é aplicada uma transformada *wavelet* com quatro níveis e fase linear. Os coeficientes são uniformemente quantizados de acordo com um esquema de alocação de bits para as sub-bandas, derivado de um estimador espectral, e codificados com um esquema de codificação de entropia semelhante ao adotado no padrão JPEG, associando RLE e codificação de Huffman (Wallace, 1990). Com blocos de tamanho 2000, foi reportada uma RC de 7,9:1 com PRD de 3,9%.

(Hilton, 1997)

Hilton (1997) propôs uma adaptação do EZW para o caso unidimensional de compressão de ECG. Oito tipos diferentes de *wavelet* foram testados, a maioria comprovadamente eficiente para compressão de imagens. Seis delas apresentaram desempenho semelhante, mas a transformada *wavelet* usada no compressor de imagens de impressões digitais do FBI apresentou compromisso RC-PRD levemente superior para todos os valores de RC testados (de 2,0:1 a 30,0:1). Para RC = 10,0:1, esta transformada atingiu uma PRD de aproximadamente 5%, enquanto que a *wavelet* de Haar gerou a maior distorção, com PRD aproximadamente igual a 7,5%. O conjunto de testes consistiu em 168 trechos, com 20,48 s, de registros provenientes do MIT-BIH *Compression Test Database*. Para avaliar os efeitos da compressão sobre o diagnóstico, 75 trechos de 30 s provenientes de três diferentes bancos de ECGs (*MIT-BIH Supraventricular Arrhythmia Database*, *MIT-BIH ST Change Database*, e *Creighton University Ventricular Tachyarrhythmia Database*) foram comprimidos a razões de 8:1, 16:1 e 24:1, e os trechos originais e descomprimidos foram aleatoriamente apresentados a um cardiologista para diagnóstico. Diagnósticos idênticos

efetuados a partir de um sinal original e de um sinal reconstruído foram computados como “diagnósticos corretos” para o sinal reconstruído. A porcentagem de sinais “corretamente diagnosticados” foi de 92% para RC = 8,0:1; 80% para RC = 16,0:1; e apenas 44% para RC = 24,0:1. As PRDs correspondentes a estas RCs não foram informadas.

(Ramakrishnan e Supratim Saha, 1997)

Ramakrishnan e Supratim Saha (1997) propuseram um esquema de compressão de ECG baseado em predição linear de coeficientes *wavelet*. Nesse método, os batimentos são primeiramente delimitados, seus períodos e amplitudes são normalizados e uma transformada *wavelet* é então aplicada a cada batimento. Os n coeficientes de amplitudes mais elevadas são integralmente preservados e os demais são eliminados, de maneira semelhante à utilizada pelo compressor de Crowe *et al.* (1992), descrito no início desta seção.

Devido à normalização, coeficientes *wavelet* na mesma localização dentro de batimentos diferentes apresentam alta correlação, o que é explorado por uma etapa de predição linear interbatimentos dos coeficientes preservados. Assim, o compressor explora tanto a correlação intrabatimento, através da própria transformada *wavelet*, quanto a correlação interbatimentos, através da predição linear. Nos experimentos com sinais amostrados a 250 Hz com 12 bits por amostra, o compressor atingiu uma RC média de 17,4:1, com uma PRD média de 11,3%.

(Chen e Itoh, 1998)

O compressor de ECG proposto por Chen e Itoh (1998) é baseado na quantização otimizada de coeficientes *wavelet*. O sinal é primeiramente particionado em blocos de tamanho 1024. Um tamanho de passo de quantização inteiro é definido para cada coeficiente, através de um procedimento iterativo, de forma que a distorção, única para todos os coeficientes, é igual a um erro médio quadrático especificado, dentro de uma certa tolerância. Como discutido na Seção 4.6, muitas vezes os resultados produzidos por esta abordagem são insatisfatórios. O algoritmo LZW (Bell *et al.*, 1990) compõe a etapa de codificação de entropia. Para os primeiros 10 minutos do canal superior do registro 200 do MIT-BIH *Arrhythmia Database*, foi reportada uma RC de 12,5:1 com PRD de 6,8%.

(Chagas *et al.*, 2000)

O algoritmo EZW foi também adaptado para compressão de ECG por Chagas *et al.* (2000). O método resultante emprega transformada *wavelet* biortogonal 9-7, de fase linear, com extensão simétrica do sinal, o que reduz a ocorrência de artefatos causados por descontinuidades nas fronteiras dos blocos, e codificação aritmética adaptativa. Os registros completos do MIT-BIH *Arrhythmia Database* foram utilizados para os testes empíricos. Com a RC fixa em 12,0:1, obteve-se uma PRD média de 3,6%. A partir da inspeção dos traçados originais e reconstruídos, os autores concluíram que a qualidade do sinal é bem preservada até esse nível de compressão.

(Miaou e Yen, 2000)

Um compressor de ECG baseado em quantização vetorial *gold washing* adaptativa (*gold washing adaptive vector quantization*, GWAVQ) foi apresentado por Miaou e Yen (2000). Em uma das versões implementadas pelos autores, os sinais foram diretamente quantizados no domínio do tempo (GWAVQ direto); na outra, a quantização foi efetuada sobre os coeficientes da transformada *wavelet* Daubechies-8 dos sinais, com cinco níveis (WT + GWAVQ). Miaou e Yen apontam os dois principais problemas da quantização vetorial: a elevada complexidade computacional e a possível inadequação do conjunto de treinamento utilizado para a construção do *codebook* aos sinais a serem comprimidos. No trabalho em questão, a redução no tempo de processamento é deixada a cargo de inovações tecnológicas nos computadores, com a esperança de que, nas palavras dos autores, “eventualmente, a quantização vetorial possa realmente tornar-se prática na compressão de dados de ECG”.

Sinais de ECG provenientes de derivações ou patologias diferentes podem apresentar amplas variações morfológicas. Conseqüentemente, *codebooks* pré-treinados e fixos podem ser pouco eficientes para a compressão de ECG. No GWAVQ, o *codebook* é periodicamente atualizado de acordo com um mecanismo denominado de *gold washing*. O GW procura adaptar o *codebook* às características do sinal, inserindo novos vetores-códigos e descartando aqueles pouco utilizados. A eficiência do esquema depende da escolha de um limiar de distorção, *dth*, que determina como o *codebook* será atualizado. O GWAVQ inclui um algoritmo para ajustar o *dth* automaticamente, de acordo com a distorção permitida pelo usuário.

Os 15 primeiros minutos do canal superior dos registros 101, 111, 208, e 228 do MIT-BIH *Arrhythmia Database* foram usados como sinais de teste. Apenas os resultados do WT + GWAVQ serão aqui apresentados, dado que estes foram claramente melhores que os do GWAVQ direto. O WT + GWAVQ obteve RC = 4,6:1, com PRD2 = 3,3%; RC = 9,4:1, com PRD2 = 6,0%; e RC = 12,4:1, com PRD2 = 8,2%.

(Ahmeda e Abo-Zahhad, 2001)

A aplicação de predição linear de coeficientes *wavelet* para compressão de sinais de ECG foi também investigada por Ahmeda e Abo-Zahhad (2001). No método proposto por estes autores não há delimitação de batimentos, e a predição do valor de um dado coeficiente é efetuada levando em consideração os dois coeficientes que o antecedem. Erros de predição menores que um dado limiar são anulados, e a seqüência resultante é codificada por um RLE modificado. A medida de distorção utilizada é denominada pelos autores de PRD, mas a fórmula difere da usualmente empregada na literatura pela ausência da raiz quadrada. Os testes envolvendo pequenos trechos (em torno de 2000 amostras) dos registros 103 e 223 do MIT-BIH *Arrhythmia Database* resultaram em uma RC média de 20,0:1 com distorção de 3,8%. A RC foi calculada com base na suposição errônea de que os sinais de teste têm 12 bits por amostra, e não 11 bits. A distorção de 3,8% equivale a uma PRD de 19,5%, em sua formulação usual. Os traçados apresentados indicam, de fato, uma distorção substancial no sinal reconstruído.

5.2 COMPRESSÃO BASEADA EM DCT

(Davis e Chawla, 1997)

Davis e Chawla (1997) analisaram os motivos do excelente desempenho do EZW e do SPIHT, mostrando que, para uma importante classe de processos aleatórios que descrevem um modelo contextual simples para o comportamento dos coeficientes *wavelet*, a codificação por árvores de zero é ótima. A análise indica que estes esquemas associam implicitamente um tipo de quantização vetorial restrita, que explora eficientemente as interdependências existentes entre os coeficientes *wavelet* e seus descendentes, com uma codificação de entropia de baixa complexidade.

Assim, apesar de simples, a estratégia de quantização e codificação adotada nestes compressores é adequada à estruturação observada nos coeficientes *wavelet*.

Com base neste estudo, Davis e Chawla descreveram um procedimento para otimizar o processo de partição em árvores de zero do EZW para um conjunto de imagens de treinamento. O método, que independe da transformada, foi empregado em um compressor de imagens de baixa complexidade, baseado em DCT e sem codificação de entropia. Os resultados do compressor de Davis e Chawla foram substancialmente superiores aos do JPEG (Wallace, 1991) e marginalmente superiores aos da versão sem codificação de entropia do EZDCT de Xiong *et al.* (1996). Ressalte-se que este último, de acordo com os testes realizados por seus proponentes, apresenta resultados melhores que os do EZW.

(Ratnakar, 1997)

Muitos procedimentos de otimização de vetores ou matrizes de quantização são descritos na literatura especializada. Na abordagem mais comum, constrói-se a matriz de quantização multiplicando-se uma matriz padrão por um escalar. Para atingir uma taxa ou distorção específica, o valor do multiplicador é repetidamente alterado, e os dados são comprimidos para calcular a taxa e descomprimidos para calcular a distorção, até que o alvo desejado é atingido. Este processo é ineficiente e pode levar a um compromisso taxa-distorção pobre (Ratnakar e Livny 1994). Em diversas abordagens mais sofisticadas, que efetuam a busca pela solução ótima utilizando algum tipo de heurística, a avaliação de cada matriz de quantização e limiares também exige que a imagem passe por todo o ciclo de compressão-descompressão (Ratnakar, 1997).

Ratnakar (1997) caracterizou analiticamente o desempenho de quatro estratégias de quantização para uma fonte laplaciana, concluindo que a quantização com zona morta, com níveis de reconstrução no centro dos intervalos de decisão, alia baixa complexidade a uma boa relação taxa-distorção. Com base nos resultados deste estudo, o autor desenvolveu um compressor de imagens, denominado RD-OPT, que emprega quantização com zona morta ótima dos coeficientes DCT de blocos de dimensões 8x8. Diversas estratégias destinadas a efetuar eficientemente a procura pelos vetores de quantização e de limiares ótimos são empregadas, resultando em um reduzido tempo de processamento. A busca pela solução ótima é restrita à borda convexa inferior do

conjunto de pontos operacionais. Após a definição dos vetores ótimos, o RD-OPT segue os passos definidos no padrão JPEG (Wallace, 1991).

Usando a métrica psicovisual de distorção conhecida como PQS (*Picture Quality Scale*) (Miyahara *et al.*, 1996), para taxas de bits na faixa “moderadamente baixas” até “moderadamente altas”, o RD-OPT atinge os mesmos resultados que os melhores compressores baseados em *wavelet* (Algazi e Estes, 1997). Sabendo-se que o padrão JPEG-2000, baseado em *wavelets*, supera largamente o desempenho do JPEG (Marcellin *et al.*, 2000), os resultados do RD-OPT, que difere do JPEG exclusivamente na etapa de quantização, mostram a eficiência da abordagem empregada e a importância dessa etapa no processo de compressão.

(Lee e Buckley, 1999)

Além da dependência estatística entre amostras consecutivas, o sinal de ECG apresenta também dependência interbatimentos. O CAB (*Cut and Align Beats*) com DCT bidimensional (Lee e Buckley, 1999) obtém resultados expressivos explorando este fato. Nesta abordagem, denominada CAB/2D-DCT, após a detecção dos complexos QRS, o sinal de ECG é particionado em batimentos, que são alinhados em uma matriz bidimensional. Cada batimento constitui uma linha da matriz, com um número apropriado de zeros adicionados ao final dos batimentos para uniformizar os tamanhos.

A seguir, a matriz é dividida em blocos de tamanho $L \times L$, e a DCT bidimensional é aplicada aos blocos. Se X_{ij} representa um coeficiente DCT de um bloco qualquer, o coeficiente quantizado \hat{X}_{ij} é dado simplesmente por $X_{ij} // q$, onde $//$ representa divisão seguida por arredondamento, e q é o tamanho do passo de quantização, único para todos os coeficientes, e calculado a partir das variâncias estimadas dos coeficientes. Mais uma vez, como exposto na Seção 4.6, diferenças no comportamento estatístico dos coeficientes, freqüentemente verificadas na prática, fazem com que a utilização de um único passo de quantização degrade sensivelmente o compromisso taxa-distorção resultante.

A etapa final do CAB/2D-DCT consiste na codificação RLE dos coeficientes quantizados. Além dos coeficientes quantizados e codificados, o número de zeros adicionados a cada batimento e o

tamanho do passo de quantização devem ser anexados ao sinal comprimido para permitir a descompressão.

O conjunto de testes consistiu nos 10 primeiros minutos dos registros 100/II, 100/V5, 119/II e 119/VI do MIT-BIH *Arrhythmia Database*, com a resolução e a taxa de amostragem previamente alteradas para 12 bits/amostra e 250 Hz, respectivamente. Com estes sinais, o CAB/DCT-2D atingiu RC = 4,0:1 com PRD = 2,3%; RC = 6,0:1 com PRD = 3,5%; e RC = 12,0:1 com PRD = 6,1%.

(Batista, Melcher e Carvalho, 2000)

A importância dos coeficientes DCT de sinais de ECG tende a decrescer com o aumento da frequência, o que permite atenuar e mesmo anular os componentes de frequências mais elevadas sem distorções consideráveis no sinal. Quanto maior a atenuação/eliminação de componentes de alta frequência, maior a RC e a distorção, principalmente nas transições mais bruscas do sinal.

O algoritmo *Fan* (Lai, 1993), um dos mais populares métodos de compressão de ECG, apresenta efeitos complementares aos acima descritos, com distorções menos perceptíveis em regiões com variações abruptas de amplitude, como o complexo QRS. O *Fan* substitui as amostras do sinal original por segmentos de retas, de tal forma que a diferença entre as amostras eliminadas e os pontos correspondentes nos segmentos é sempre menor que um determinado limiar ε . Em geral, quanto maior ε , maiores são a compressão e a distorção resultantes. O arquivo comprimido consiste nas amplitudes das amostras iniciais e finais de cada segmento, e nos comprimentos dos segmentos, dados pelo número de pontos que os constituem. Ao contrário do que ocorre quando as componentes de frequência mais elevadas da DCT são atenuadas, com algoritmo *Fan* normalmente há uma maior preservação das variações bruscas no sinal do que das variações suaves.

Esta complementaridade na preservação das componentes de frequência é explorada pelo compressor *Fan*+DCT (Batista, Melcher e Carvalho, 2000). Neste método, uma aproximação poligonal do sinal é efetuada por um compressor *Fan* modificado, e a DCT do erro entre a aproximação e o sinal original é calculada e processada por um filtro passa-baixas. Os coeficientes DCT resultantes são então codificados por uma associação de RLE com codificação de entropia. As modificações efetuadas no algoritmo *Fan* visam preservar, com maior eficiência, os máximos e míni-

mos locais do sinal. Os parâmetros do compressor consistem no erro máximo, ε , permitido entre a aproximação poligonal e o sinal original e na frequência de corte f_c do filtro. Estes parâmetros permitem controlar a RC e a PRD resultantes.

Quando aplicadas isoladamente, as técnicas de aproximação poligonal podem introduzir alterações importantes na morfologia do sinal, produzindo reconstruções de aparência pouco natural. O *Fan+*DCT reduz o problema incluindo, na sequência comprimida, componentes de baixa frequência do erro entre a aproximação poligonal e o sinal original. O descompressor pode então adicionar estes componentes de baixa frequência à aproximação poligonal, gerando um sinal com menor distorção e com um aspecto morfológico mais próximo do traçado eletrocardiográfico usual.

Para os testes, foram incluídos os primeiros 60 s dos registros 100/MLII, 101/MLII, 102/V5, 103/MLII e 107/V1 do MIT-BIH *Arrhythmia Database*. Para uma PRD de 3%, obteve-se uma RC média de 5,3:1, com boa qualidade visual das reconstruções.

(van der Poel *et al.*, 2001)

No compressor de ECG proposto por van der Poel *et al.* (2001), o sinal é particionado em blocos de 4096 amostras, sobre os quais se aplica a DCT. Os coeficientes DCT são quantizados por intermédio de uma função de quantização composta por 4096 valores em ponto flutuante que crescem linearmente começando em 1 e terminando em um valor máximo variável v_{max} . O valor na posição i da função de quantização é o tamanho do passo de quantização do coeficiente i de todos os blocos. Assim, o efeito final é o de quantizar com passos maiores os coeficientes de frequências mais elevadas. O valor v_{max} controla a distorção e a RC. Os coeficientes quantizados são então limitados ao intervalo $[-126, 126]$ de uma forma reversível e os blocos são então codificados usando RLE associado à codificação de Huffman ou ao algoritmo LZW (Bell *et al.*, 1990).

O conjunto de testes consistiu de 34 ECGs, sendo 28 do MIT-BIH *Arrhythmia Database* e seis colhidos no Laboratório de Processamento de Sinais Biológicos e Instrumentação Biomédica do Núcleo de Estudos e Tecnologia em Engenharia Biomédica (NETEB) da UFPB, a 500 amostras/s e 12 bits/amostra. Com a codificação de Huffman, obteve-se uma RC média de 10,2:1, com PRD próxima a 5%.

5.3 COMPRESSÃO COM DCT E DWT COMBINADAS

(Zou e Gallagher, 1994)

Partindo do pressuposto de que a DWT é mais eficiente na decomposição do complexo QRS, enquanto que a representação por DCT é superior no restante do sinal, Zou e Gallagher (1994) propõem o seguinte esquema de compressão de ECG, que combina DWT e DCT, procurando explorar a complementaridade das duas transformadas:

1. Calcular a DWT do sinal e salvar e indexar o número predefinido de coeficientes de magnitudes mais elevadas.
2. Calcular a DWT inversa dos coeficientes salvos.
3. Subtrair o resultado do passo 2 do sinal original.
4. Calcular a DCT do resultado do passo 3 e salvar e indexar o número predefinido de coeficientes de magnitudes mais elevadas.

Como se pode notar, o mecanismo de controle da distorção, extremamente simples, consiste na eliminação de um número predefinido de coeficientes de menor amplitude. A inspeção visual dos traçados apresentados indica que, preservando-se 91 de um total de 512 coeficientes, sendo 35 da DCT e 56 da DWT (*wavelet* Daubechies D4), o sinal reconstruído é de boa qualidade. A RC resultante não é informada, mas a necessidade de representar com boa precisão os coeficientes preservados, bem como a inclusão dos seus índices na seqüência comprimida, contribuem para que a RC seja consideravelmente menor que a razão entre o número original de coeficientes e o número de coeficientes preservados ($512/91 = 5,6$). A conclusão mais importante do trabalho é que a combinação das duas transformadas eleva a qualidade visual da reconstrução para um mesmo grau de compressão, o que sugere que a premissa básica, que estipula que DCT e DWT possuem características complementares que podem ser conjuntamente exploradas, é correta.

5.4 CONCLUSÃO

O paradigma transformada + quantização + codificação tem sido extensamente aplicado à compressão com perdas de sinais biológicos, áudio, imagens e vídeo. Sua popularidade deve-se essencialmente ao bom compromisso entre taxa, distorção e demanda de recursos computacionais. Diversos compressores atuais de ECG seguem este esquema, muitas vezes com excelentes resultados e exigências de recursos computacionais modestas ou, pelo menos, aceitáveis.

Para a transformada, há um grande interesse atual na DWT, principalmente devido aos excelentes resultados do EZW (Shapiro, 1993) e do SPIHT (Said e Pearlman, 1996). Entretanto, a superioridade destes compressores com relação àqueles baseados em DCT parece decorrer fundamentalmente da eficiente combinação quantização-codificação do EZW e do SPIHT, que exploram a organização observada nos coeficientes DWT, e não de uma superioridade inerente a esta transformada. Diversos resultados indicam que estratégias de quantização-codificação semelhantes, quando aplicadas à DCT, produzem resultados semelhantes aos obtidos com a DWT (Zou e Gallagher, 1994; Xiong *et al.*, 1996; Davis e Chawla, 1997; Laurence e Monro, 1997, Monro e Dickson, 1997). Além disso, mesmo com uma abordagem de quantização e de codificação completamente diversa, o RD-OPT, baseado em DCT, atingiu desempenho semelhante ao dos melhores compressores baseados em DWT (Ratnakar, 1997).

Considerando o exposto nos parágrafos anteriores, além do bom desempenho da DCT, reafirmado ao longo de sua história por uma série de estudos práticos e teóricos, e sabendo-se ainda que sua aplicação é intrinsecamente mais simples que a da DWT, propõe-se no presente trabalho um método de compressão baseado em DCT. Com a transformada definida, os esforços se concentram na busca de estratégias de quantização e codificação adequadas. O método é detalhadamente descrito no próximo capítulo.

As dificuldades associadas com a definição da quantização têm muitas vezes conduzido a simplificações excessivas, com conseqüente degradação na relação taxa-distorção. Como visto neste capítulo, alguns compressores utilizam um único limiar de decisão para todos os coeficientes, anulando os valores inferiores ao limiar e preservando os demais (Crowe *et al.*, 1992; Zou e Gallagher, 1994; Ramakrishnan e Supratim Saha, 1997). Em outros métodos, como o proposto por Chen e Itoh (1998), define-se um tamanho de passo de quantização inteiro para cada coeficiente,

de forma que a distorção, única para todos os coeficientes, seja igual a um dado erro médio quadrático, dentro de uma certa tolerância. Há também compressores, como o CAB/2D-DCT (Lee e Buckley, 1999), em que o tamanho do passo de quantização é único para todos os coeficientes.

O RD-OPT (Ratnakar, 1997), por outro lado, emprega quantização com zona morta, sendo os parâmetros definidos por um procedimento eficiente de otimização, com liberdade para estabelecer tamanhos de passos de quantização e limiares ótimos diferentes para os diversos coeficientes. Esta é uma abordagem de peculiar generalidade dentre os compressores práticos descritos na literatura especializada. De fato, as abordagens citadas no parágrafo antecedente são casos particulares do processo de otimização do RD-OPT. Em outras palavras, nas situações reais específicas em que tais esquemas são ótimos, a estratégia de controle de perdas do RD-OPT será equivalente, por exemplo, ao critério de preservação/descarte de coeficientes; ao uso de um tamanho de passo de quantização único, para todos os coeficientes, sem zona morta; ou ao emprego de distorção única para todos os coeficientes. Com isto, o controle de distorção do RD-OPT terá, em qualquer caso, desempenho igual ou superior aos citados no parágrafo anterior.

Para fontes laplacianas, a curva $R-D$ associada à quantização com zona morta com níveis de reconstrução no centro dos intervalos de decisão, utilizada no RD-OPT, praticamente coincide com a da quantização uniforme com níveis de reconstrução posicionados de forma a minimizar a distorção (Ratnakar, 1997); sabe-se ainda que a performance desta última abordagem é próxima à ótima para fontes gaussianas, laplacianas e gama (Popat, 1990). Os resultados experimentais do RD-OPT, que utiliza o mesmo esquema de codificação do JPEG, são semelhantes aos dos melhores compressores de imagens conhecidos.

O processo de otimização da quantização efetuado pelo RD-OPT, além de permitir uma notável liberdade na definição dos parâmetros, possui implementação eficiente para o nível tecnológico atual, conferindo interesse à investigação de sua aplicabilidade à compressão de ECG. A etapa de quantização do compressor descrito neste capítulo é uma adaptação do RD-OPT para este caso.

A etapa final, de codificação de entropia, é também importante para o desempenho do sistema como um todo, permitindo aumentar a RC sem qualquer acréscimo na distorção introduzida pela quantização. A caracterização estatística dos coeficientes DCT, fundamental para a construção

de códigos eficientes, depende da fonte de informação original. Para imagens, por exemplo, é usual supor que os coeficientes AC seguem distribuições laplacianas (Rao e Yip, 1990; Jayant e Noll, 1984), enquanto que o coeficiente DC segue uma distribuição gaussiana ou Rayleigh (Rao e Yip, 1990). A ocorrência freqüente de seqüências de zeros nos blocos quantizados sugere a inclusão de algum mecanismo de eliminação dos elementos nulos. Normalmente, estas seqüências concentram-se sobretudo nas regiões de freqüência mais elevadas dos blocos, e muitos mecanismos habitualmente aplicados, como o RLE do JPEG (Wallace, 1991), são ineficientes no início do bloco, onde pode haver até mesmo uma expansão na representação binária dos coeficientes. O compressor proposto, descrito no capítulo seguinte, reordena os coeficientes quantizados de forma a aumentar a probabilidade de ocorrência de valores nulos no final dos blocos, e elimina então a seqüência de zeros no final de cada bloco.

Capítulo 6

METODOLOGIA

Neste capítulo, propõe-se um método de compressão de ECG baseado na quantização ótima com zona morta dos coeficientes DCT. Diversos esquemas diferentes de codificação de entropia, empregando codificação aritmética ou de Golomb-Rice em associação com um esquema simples de eliminação de coeficientes quantizados nulos, são propostos e testados.

6.1 O COMPRESSOR PROPOSTO

A Figura 6.1 mostra o diagrama em blocos do compressor proposto. A compressão é efetuada em quatro etapas principais:

1. Partição do ECG em N_b blocos consecutivos \mathbf{b}_i , $i = 0, 1, \dots, N_b - 1$, cada um com L_b amostras

2. Cálculo da DCT de cada bloco
3. Quantização ótima com zona morta dos coeficientes DCT
4. Codificação de entropia dos coeficientes DCT quantizados

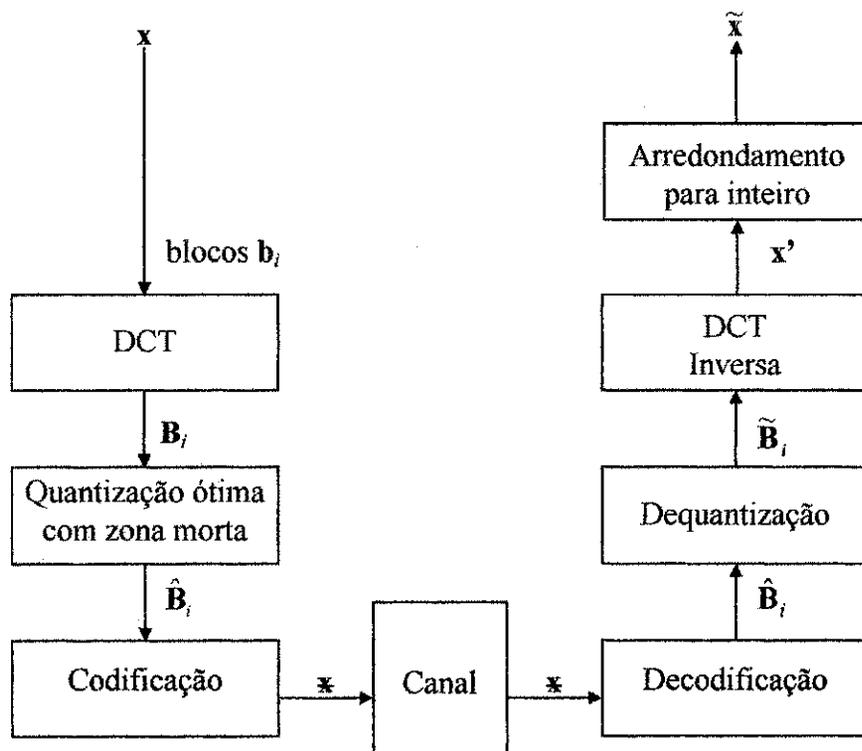


Figura 6.1 Diagrama em blocos do compressor/descompressor proposto

6.1.1 Partição em blocos e DCT

Nos compressores baseados na DCT, os sinais a serem comprimidos são particionados em blocos de tamanho tipicamente inferior a 4096 amostras, e a DCT é aplicada separadamente a cada uma destes blocos. Isto é feito não apenas para reduzir o tempo de processamento, mas também para permitir uma melhor adaptação às características locais dos sinais. Alguns estudos indicam que aumentos no tamanho dos blocos acima de um certo ponto levam a um ganho de compressão modesto, quando comparado com o aumento no tempo de processamento, ou até mesmo a uma redução na compressão (Lee e Buckley, 1999; Nelson e Gailly, 1996). Para um cálculo mais efi-

ciente da DCT, o tamanho do bloco deve ser uma potência de dois. Foram efetuados testes, descritos no Capítulo 7, para determinação do efeito do tamanho dos blocos na relação RC-PRD. Na descrição que se segue, o tamanho dos blocos é representado por L_b .

A DCT do tipo II (DCT-II) foi selecionada devido a sua maior capacidade de concentrar a energia do sinal em poucos coeficientes. Se b_{in} , $n = 0, 1, \dots, L_b-1$, são as L_b amostras no bloco \mathbf{b}_i , a DCT-II unidimensional deste bloco gera um bloco transformado \mathbf{B}_i constituído por L_b coeficientes B_{im} , $m = 0, 1, \dots, L_b - 1$, dados por:

$$B_{im} = \left(\frac{2}{L_b}\right)^{1/2} c_m \sum_{n=0}^{L_b-1} b_{in} \cos\left[\frac{(2n+1)m\pi}{2L_b}\right], \quad m = 0, 1, \dots, L_b - 1 \quad (6.1)$$

onde $c_m = 1$ para $1 \leq m \leq L_b - 1$ e $c_0 = (1/2)^{1/2}$.

Dado \mathbf{B}_i , \mathbf{b}_i pode ser recuperado aplicando-se a DCT-II inversa:

$$b_{in} = \left(\frac{2}{L_b}\right)^{1/2} \sum_{m=0}^{L_b-1} c_m B_{im} \cos\left[\frac{(2n+1)m\pi}{2L_b}\right], \quad n = 0, 1, \dots, L_b - 1 \quad (6.2)$$

6.1.2 Quantização ótima com zona morta

Para a quantização com zona morta, usa-se um *vetor de quantização*, \mathbf{q} , com elementos q_m , $m = 0, 1, \dots, L_b - 1$, e um vetor de limiares, \mathbf{t} , com elementos t_m , $m = 0, 1, \dots, L_b - 1$. O elemento \hat{B}_{im} do bloco DCT quantizado $\hat{\mathbf{B}}_i$ é obtido pela operação:

$$\hat{B}_{im} = \begin{cases} 0, & \text{se } |B_{im}| < t_m \\ B_{im} // q_m, & \text{caso contrário} \end{cases} \quad (6.3)$$

$m = 0, 1, \dots, L_b - 1$
 $i = 0, 1, \dots, N_b - 1$

onde // representa divisão seguida por arredondamento para o inteiro mais próximo. A dequantização consiste simplesmente na multiplicação de cada coeficiente quantizado pelo componente

correspondente de q . A definição de q e t influencia fortemente a relação taxa-distorção. Um quantizador mal projetado pode levar a razões de compressão reduzidas associadas a distorções elevadas. Descreve-se a seguir um procedimento que permite definir q e t de maneira a minimizar a entropia estimada dos coeficientes quantizados para uma dada distorção alvo.

Para um dado sinal, seja $H(\mathbf{q}, \mathbf{t})$ a entropia de ordem zero dos coeficientes DCT quantizados usando os vetores \mathbf{q} e \mathbf{t} , e $D(\mathbf{q}, \mathbf{t})$ uma medida da distorção introduzida no sinal pela quantização. A Formulação 6.1, que é análoga à Formulação 3.2, expressa o problema em questão:

Formulação 6.1: Dada uma distorção D_T e uma tolerância τ , encontrar uma instância de (\mathbf{q}, \mathbf{t}) que minimize $H(\mathbf{q}, \mathbf{t})$, com $D_T - \tau \leq D(\mathbf{q}, \mathbf{t}) \leq D_T + \tau$.

Os valores máximos permitidos para o tamanho de passo de quantização e para os limiares de zona morta serão representados por Q_{max} e T_{max} , respectivamente. Seguindo uma prática comum, os valores de q_m são inteiros no intervalo $[1, Q_{max}]$ (Wallace, 1991). Pela Equação (6.3), percebe-se que todo coeficiente B_m com valor inferior a $q_m/2$ será anulado pela divisão por q_m seguida por arredondamento, o que torna desnecessário testar limiares t_m inferiores ou iguais a $q_m/2$ no procedimento de otimização. No procedimento aqui adotado, os limiares podem assumir valores no intervalo $(q_m/2, T_{max}]$, com incrementos de 0,5. Este valor de 0,5 permite uma implementação muito eficiente para a otimização (Ratnakar, 1997).

Para $L_b = 64$ e $Q_{max} = 64$, uma solução por busca exaustiva exigiria, apenas para a definição do vetor \mathbf{q} ótimo, o teste de $64^{64} \approx 10^{115}$ instâncias diferentes de \mathbf{q} . Os próximos parágrafos descrevem um procedimento que permite reduzir a complexidade do problema a níveis práticos. No processo de otimização aqui descrito, todo o registro de ECG a ser comprimido é usado na determinação dos vetores \mathbf{q} e \mathbf{t} ótimos.

Como discutido no Capítulo 3, se o conjunto de pontos operacionais no plano $H-D$ gerados pela quantização com todos as instâncias possíveis de (\mathbf{q}, \mathbf{t}) é suficientemente denso, a busca pela solução ótima pode restringir-se à borda convexa inferior do conjunto convexo gerado pelos pontos operacionais. Neste caso, a otimização pode ser efetuada através da minimização do lagrangiano

$$J = H(\mathbf{q}, \mathbf{t}) + \lambda D(\mathbf{q}, \mathbf{t}) \quad (6.4)$$

para um dado valor do multiplicador de Lagrange λ . O valor de λ que leva ao $D(\mathbf{q}, \mathbf{t})$ desejado, pode ser eficientemente encontrado pelo método da bissecção (Ortega, 1994).

Para a otimização, a distorção $D(\mathbf{q}, \mathbf{t})$ é medida pelo erro médio quadrático, MSE, dado pela Equação (3.2). Como a DCT é ortonormal, o MSE entre \mathbf{x} e \mathbf{x}' é igual ao MSE entre \mathbf{B}_i e $\tilde{\mathbf{B}}_i$ (ver Figura 6.1). Em outras palavras, o MSE entre \mathbf{x} e \mathbf{x}' pode ser calculado no domínio da DCT, o que elimina a necessidade de efetuar a DCT inversa dos coeficientes dequantizados para medir a distorção no domínio do tempo. Entretanto, a operação de arredondamento efetuada sobre \mathbf{x}' para gerar o sinal reconstruído $\tilde{\mathbf{x}}$, indicada na Figura 6.1, introduz distorções extras que fazem com que o MSE entre \mathbf{B}_i e $\tilde{\mathbf{B}}_i$ não coincida exatamente com o MSE entre \mathbf{x} e $\tilde{\mathbf{x}}$. O efeito do arredondamento é normalmente pequeno, e pode ser desconsiderado em muitas situações práticas, de modo que uma boa aproximação para a distorção entre o sinal original e o recuperado pode ser efetivamente calculada no domínio da DCT. A opção adotada no presente trabalho será descrita nos parágrafos finais desta seção.

Há uma relação direta entre a PRD e o MSE, que pode ser extraída das Equações (3.2), (3.3) e (3.4):

$$\text{PRD} = \frac{\sqrt{\text{MSE}}}{\text{RMSV}} \times 100\% \quad (6.5)$$

onde RMSV é o valor RMS do sinal original, dado pela Equação (3.5).

Se N_b é o número de blocos em que o sinal foi particionado, a distorção $D(\mathbf{q}, \mathbf{t})$ no domínio da DCT, medida pelo MSE, é dada por:

$$D(\mathbf{q}, \mathbf{t}) = \frac{1}{L_b N_b} \sum_{i=0}^{N_b-1} \sum_{m=0}^{L_b-1} (B_{im} - q_m \hat{B}_{im})^2 \quad (6.6)$$

Notando que o erro médio quadrático devido à quantização do k -ésimo coeficiente de todos os blocos, $D_k(q_k, t_k)$, é dado por

$$D_k(q_k, t_k) = \frac{1}{N_b} \sum_{i=0}^{N_b-1} (B_{ik} - q_k \hat{B}_{ik})^2 \quad (6.7)$$

pode-se reescrever a Equação (6.6) como

$$D(\mathbf{q}, \mathbf{t}) = \frac{1}{L_b} \sum_{m=0}^{L_b-1} D_m(q_m, t_m) \quad (6.8)$$

Se o k -ésimo coeficiente quantizado assume valor v em $n_k(v)$ dos N_b blocos, sua entropia, $H_k(q_k, t_k)$, é dada por

$$H_k(q_k, t_k) = - \sum_v p_k(v) \log_2 p_k(v) \quad (6.9)$$

onde $p_k(v) = n_k(v) / N_b$.

A entropia dos coeficientes quantizados será estimada pelo seguinte modelo simplificado:

$$H(\mathbf{q}, \mathbf{t}) = \frac{1}{L_b} \sum_{m=0}^{L_b-1} H_m(q_m, t_m) \quad (6.10)$$

Nos resultados experimentais apresentados por Ratnakar (1997), o erro na entropia calculada pela Equação (6.10) foi normalmente inferior a 0,02 bits/símbolo, o que indica a precisão do modelo.

Substituindo as Equações (6.8) e (6.10) na Equação (6.4), tem-se

$$J = \frac{1}{L_b} \sum_{m=0}^{L_b-1} H_m(q_m, t_m) + \lambda \frac{1}{L_b} \sum_{m=0}^{L_b-1} D_m(q_m, t_m) \quad (6.11)$$

ou

$$J = \frac{1}{L_b} \sum_{m=0}^{L_b-1} [H_m(q_m, t_m) + \lambda D_m(q_m, t_m)] \quad (6.12)$$

e a minimização de J reduz-se à minimização de

$$J_m = H_m(q_m, t_m) + \lambda D_m(q_m, t_m), m = 0, 1, \dots, L_b - 1 \quad (6.13)$$

Com isto, a minimização pode ser efetuada independentemente para cada coeficiente. O processo de minimização descrito atua, portanto, com um conjunto de pontos operacionais específico para cada coeficiente, e a busca pelos valores ótimos dos tamanhos de passo de quantização e limiares de zona morta restringe-se à BCI associada a cada um destes conjuntos.

É importante realçar o grau de redução de complexidade decorrente da abordagem descrita nos parágrafos anteriores, com a conversão do problema de minimização com restrição proposto na Formulação 6.1 para um problema de minimização sem restrição independentemente efetuada para cada coeficiente. Mantendo-se fixo o valor de t_m , considere-se o problema da definição do tamanho do passo de quantização q_m ótimo, a ser utilizado para quantizar os elementos do conjunto formado pelos m -ésimos coeficientes quantizados de todos os blocos, $\{\hat{B}_{0,m}, \hat{B}_{1,m}, \dots, \hat{B}_{N_b-1,m}\}$. Se q_m é um inteiro entre 1 e Q_{max} , é necessário apenas calcular o lagrangiano J_m , dado pela Equação (6.13), para os Q_{max} valores possíveis de q_m , e selecionar aquele que minimiza J_m . O processo é repetido para $m = 0, 1, \dots, L_b-1$. Se $L_b = Q_{max} = 64$, um total de $64 \times 64 (= 4096)$ testes com os elementos de \mathbf{q} serão requeridos para a otimização. Este número deve ser comparado com as $64^{64} (\approx 10^{115})$ possíveis instâncias de \mathbf{q} que seriam analisadas na minimização por busca exaustiva. Esta redução de complexidade, combinada com o uso de histogramas, cálculos incrementais de distorção e entropia e outras técnicas (Ratnakar, 1997), permitem realizar uma busca eficiente pelos vetores \mathbf{q} e \mathbf{t} ótimos.

Cabe aqui salientar um detalhe de importância prática. Conforme já afirmado, o MSE entre o sinal original e o sinal reconstruído é aproximadamente igual ao MSE entre os coeficientes DCT do sinal original e os coeficientes DCT dequantizados. Se a aproximação é considerada satisfatória, a PRD no domínio do tempo pode ser calculada por intermédio das Equações (6.5) e (6.8), não havendo necessidade de efetuar a DCT inversa. Em geral, a diferença na PRD antes e após o

arredondamento é pequena. Em testes com sinais do MIT-BIH *Arrhythmia Database*, esta diferença foi menor que 0,2 pontos percentuais, para PRDs em torno de 2,5%. Contudo, se a tolerância para a PRD alvo é muito reduzida, estas distorções adicionais podem levar a PRD efetivamente atingida após a reconstrução do sinal para fora da faixa desejada.

Apesar do erro entre a PRD calculada no domínio da DCT e a PRD real ser pequeno, adotou-se neste trabalho o procedimento descrito a seguir para um controle mais preciso da distorção resultante.

Na determinação do conjunto de pontos operacionais e da BCI associada a cada coeficiente, o MSE é calculado no domínio da frequência. Os vetores \mathbf{q} e \mathbf{t} que minimizam os lagrangeanos J_m , $m = 0, 1, \dots, L_b - 1$, dados pela Equação (6.13), são definidos para um determinado valor de λ . Conforme já discutido, a busca por estes vetores está restrita às BCIs. Os coeficientes DCT são quantizados e dequantizados empregando-se as instâncias de \mathbf{q} e \mathbf{t} encontradas, a DCT inversa é computada, os valores resultantes são arredondados para os inteiros mais próximos e a PRD *real* é calculada e comparada com a PRD alvo. Caso a diferença ultrapasse a tolerância especificada, o método da bissecção é aplicado, gerando um novo valor de λ . Uma nova busca sobre as BCIs é realizada, e o processo inteiro se repete até que a diferença entre a PRD real e a PRD alvo seja inferior ou igual à tolerância. Esta opção permite um controle mais preciso da PRD real, mas causa também um aumento no tempo de processamento, devido principalmente à necessidade de calcular a DCT inversa dos blocos dequantizados em cada iteração do método da bissecção.

6.1.3 Codificação de Entropia - Codificação Aritmética

Três esquemas diferentes de codificação de entropia envolvendo o emprego de codificação aritmética foram investigados, gerando três variações do método proposto, aqui identificados por A1, A2 e A3. No esquema A1 (Batista, Carvalho e Melcher, 2000), os coeficientes quantizados, pré-processados conforme descrito adiante, são codificados por um codificador aritmético semi-adaptativo de ordem zero (Bell *et al.*, 1990). Para possibilitar a decodificação aritmética, uma tabela contendo as frequências de ocorrência dos valores que podem ser assumidos pelos coeficientes quantizados precisa ser enviada ao arquivo comprimido. Se a tabela ocupar uma parcela considerável do arquivo comprimido, a razão de compressão sofrerá uma redução substancial. Na prática, observa-se que a ampla maioria dos coeficientes é quantizada para valores que po-

dem ser representados com 8 bits ou menos, o que sugere o pré-processamento reversível descrito a seguir para limitar o tamanho da tabela a 256 linhas.

Cada coeficiente quantizado para um valor na faixa $[-127, 126]$ é representado por um único byte. Um valor $v \leq -128$ é substituído por uma seqüência de a bytes, cada um com valor -128 , onde a é o maior inteiro tal que $128a \leq |v|$, seguida por um byte com valor $v+128a$. Similarmente, um valor $v \geq 127$ é substituído por uma seqüência de b bytes, cada um com valor 127 , onde b o maior inteiro tal que $127b \leq v$, seguida por um byte com valor $v-127b$. Por exemplo, o valor -300 seria substituído pela seqüência de três bytes $(-128, -128, -44)$ e o valor 127 por $(127, 0)$. Esta operação, que é facilmente reversível e não introduz perdas, gera uma seqüência de valores limitados à faixa $[-128, 127]$, que é codificada pelo codificador aritmético. Se o número de ocorrências de cada valor possível nessa seqüência é representado em dois bytes, a tabela com os contadores para cada um dos 256 valores possíveis ocupará $256 \cdot 2 \cdot 8 = 1572$ bits no arquivo comprimido. Na eventualidade do número de ocorrência de algum valor ultrapassar o maior valor que pode ser representado com 2 bytes, $2^{16} - 1$, todos os contadores de ocorrências são divididos por um mesmo número e arredondados, calculado de forma que, após a divisão e arredondamento, o maior contador tenha valor igual a $2^{16} - 1$, um procedimento semelhante ao descrito por Bell *et al.* (1990).

Para a dequantização, necessária durante a descompressão, o vetor q ótimo também precisa ser incluído no arquivo comprimido. Com $L_b = 64$ e q_m limitado a valores inteiros na faixa de 1 a 64, por exemplo, o vetor q ótimo ocupa $64 \cdot \log_2 64 = 384$ bits. Com isto, a informação extra que deve ser incluída no arquivo comprimido com A1 totaliza $1572 + 384 = 1956$ bits

No esquema A2 (Batista, Melcher e Carvalho, 2001), explora-se o fato de que os blocos quantizados normalmente terminam em uma grande seqüência de coeficientes nulos. Se c_i é o número do último coeficiente não nulo no bloco quantizado \hat{B}_i , a seqüência final de coeficientes nulos é eliminada e a diferença entre n_i e n_{i-1} é enviada ao arquivo comprimido para permitir identificar as fronteiras entre blocos consecutivos durante a descompressão. Para o primeiro bloco, n_0 é enviado sem alteração. Como no JPEG (Wallace, 1991), para explorar a correlação entre coeficientes DC de blocos adjacentes, os coeficientes DC quantizados são codificados diferencialmente, ou seja, o coeficiente DC do bloco i , \hat{B}_{i0} , é substituído pela diferença entre seu próprio valor e o do coeficiente DC do bloco anterior, $\hat{B}_{i0} - \hat{B}_{i-1,0}$. Novamente, \hat{B}_{00} não é alterado. Os elementos da seqüência resultante são então limitados a valores na faixa $[-128, 127]$, exatamente como descrito anteriormente, e codificados por um codificador aritmético semi-adaptativo de ordem zero. Da mesma forma que com A1, para $L_b = 64$, a informação extra que deve ser incluída no arquivo

comprimido com A2, para permitir a decodificação aritmética e dequantização, totaliza 1956 bits.

O esquema A3 (Batista, Carvalho e Melcher, 2001) redistribui os coeficientes com o intuito de aumentar a probabilidade de ocorrência dos coeficientes nulos no final dos blocos e, conseqüentemente, aumentar o comprimento médio das seqüências finais de zeros. Seja z_c o número de valores nulos no conjunto formado pelos c -ésimos coeficientes quantizados de todos os blocos, $\{\hat{B}_{0c}, \hat{B}_{1c}, \dots, \hat{B}_{N_b-1,c}\}$, $c = 0, 1, \dots, L_b$. Na redistribuição, os coeficientes quantizados são crescentemente ordenados de acordo com o valor de z_i , o que tende a deixar no final dos blocos aqueles coeficientes com probabilidade mais elevada de assumir valor zero. Alguns testes indicaram que, no contexto específico do compressor aqui proposto, a reordenação combinada com a eliminação da seqüência final de zeros é mais eficiente que os esquemas de compressão RLE mais populares. Após a reordenação, segue-se o mesmo esquema de codificação de entropia utilizado em A2.

Na decodificação, a ordem original deve ser restabelecida, para permitir a aplicação correta da DCT inversa. Para tanto, deve-se incluir no arquivo comprimido a informação necessária. No compressor aqui proposto esta informação é dada por uma lista com os números originais dos coeficientes. Uma lista com os valores 1, 0, 2, 10, 3, ..., por exemplo, indica que após a reordenação, a posição 0 está sendo ocupada pelo coeficiente 1; a posição 1 pelo coeficiente 0; a posição 2 pelo próprio coeficiente 2; a posição 3 pelo coeficiente 10, e assim por diante. A lista contendo a informação sobre a ordenação dos coeficientes ocupa $L_b \times \log_2 L_b$ bits, ou 384 bits com $L_b = 64$. Neste caso, com A3, $1956 + 384 = 2340$ bits extras devem ser adicionados ao arquivo comprimido para permitir a decodificação aritmética, a restauração da ordem original dos coeficientes e a dequantização.

6.1.4 Codificação de Entropia – Codificação de Golomb-Rice

Em A1, A2 e A3 as probabilidades associadas aos valores dos coeficientes quantizados são globalmente determinadas. No entanto, coeficientes diferentes apresentam em geral variâncias muito diversas e, assim, a freqüência de ocorrência de um determinado valor quantizado pode diferir amplamente de coeficiente para coeficiente. Idealmente, as probabilidades e, conseqüentemente, a codificação deveriam ajustar-se a tais variações.

Para fontes gaussianas, é possível projetar a transformada de forma que os coeficientes quantizados tenham aproximadamente a mesma distribuição e sejam praticamente independentes (Goyal, 2001), o que permite a aplicação da mesma codificação de entropia para cada coeficiente. Muitos

dos esquemas práticos baseados em transformada, incluindo os compressores A1, A2 e A3 aqui propostos, não são otimizados para um único codificador de entropia. A solução óbvia seria projetar um codificador específico para cada coeficiente quantizado. No caso da codificação aritmética, isto significaria incluir no arquivo comprimido L_b tabelas de contadores de ocorrências, uma para cada um dos L_b coeficientes diferentes. Com $L_b = 64$ e tabelas de contadores com as mesmas dimensões das de A1, $64 \cdot 1572 = 100.608$ bits seriam ocupados por estas tabelas no arquivo comprimido, o que degradaria sensivelmente a RC, a não ser para sinais muito longos.

A codificação de Golomb e de Golomb-Rice, por outro lado, é definida apenas por um único parâmetro m , que viabiliza todo o processo de codificação e decodificação, dispensando-se assim a necessidade da inclusão de tabelas de contadores no arquivo comprimido. Pode-se com isto definir, para o conjunto formado pelos c -ésimos coeficientes quantizados de todos os blocos, $\{\hat{B}_{0c}, \hat{B}_{1c}, \dots, \hat{B}_{N_b-1,c}\}$, um valor de m específico, m_c , $c = 0, 1, \dots, L_b$, e efetuar a codificação do conjunto mapeado segundo a Equação (4.14). A eficiência da codificação dependerá do grau em que a distribuição de valores no conjunto mapeado se aproxima de uma distribuição geométrica unilateral.

A Figura 6.2 apresenta o histograma normalizado dos coeficientes 2 quantizados, $\{\hat{B}_{02}, \hat{B}_{12}, \dots, \hat{B}_{(N_b-1),2}\}$ do registro 100/MLII do MIT-BIH *Arrhythmia Database*, após quantização e mapeamento, superposto a uma distribuição geométrica unilateral. A comparação das duas curvas indica que o histograma apresenta comportamento aproximadamente geométrico. Para obtenção de todos os histogramas desta seção, o registro foi quantizado com PRD de 2,5% e $L_b = 64$.

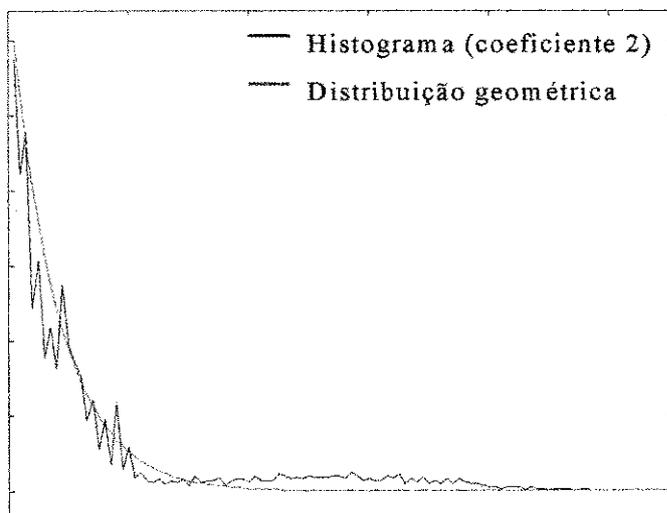


Figura 6.2 Histograma normalizado dos coeficientes 2 do registro 100/MLII, após quantização e mapeamento, e distribuição geométrica unilateral para comparação.

A semelhança entre os histogramas dos coeficientes quantizados não-mapeados e distribuições geométricas bilaterais centradas em zero ocorre para a maior parte dos coeficientes, mas em alguns casos os histogramas ou apresentam uma clara bimodalidade ou não estão centrados em zero. Na literatura, tanto a bimodalidade quanto o desvio da média zero foram também observados em histogramas de erros de predição condicionados a contextos (Seroussi e Weinberger, 1997). Nestas situações, indicadas na Figura 6.3 (bimodalidade) e na Figura 6.4 (desvio do zero), a codificação de Golomb sofre uma redução de eficiência.

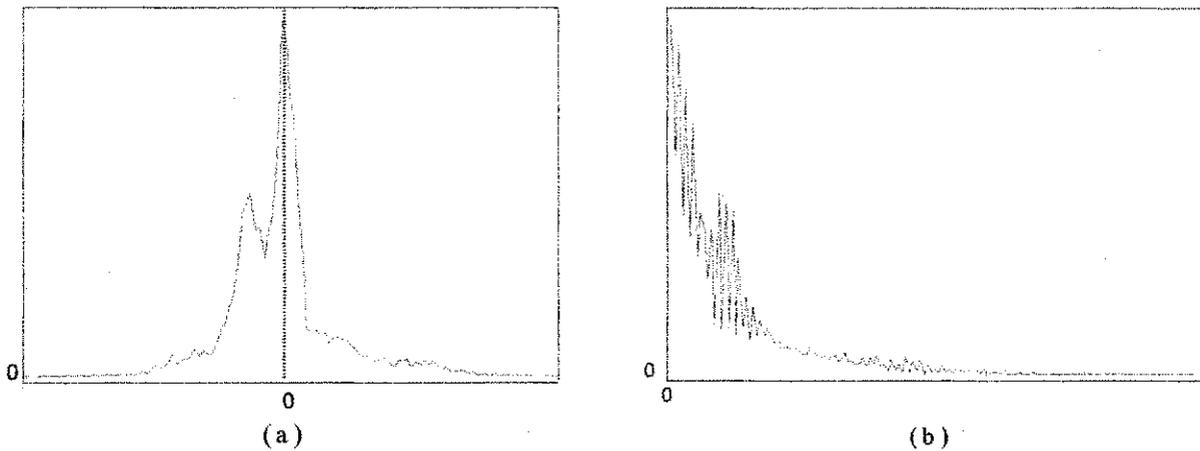


Figura 6.3. Histograma dos coeficientes 1 do registro 100/MLII, após quantização. (a) Antes do mapeamento. (b) Após o mapeamento.

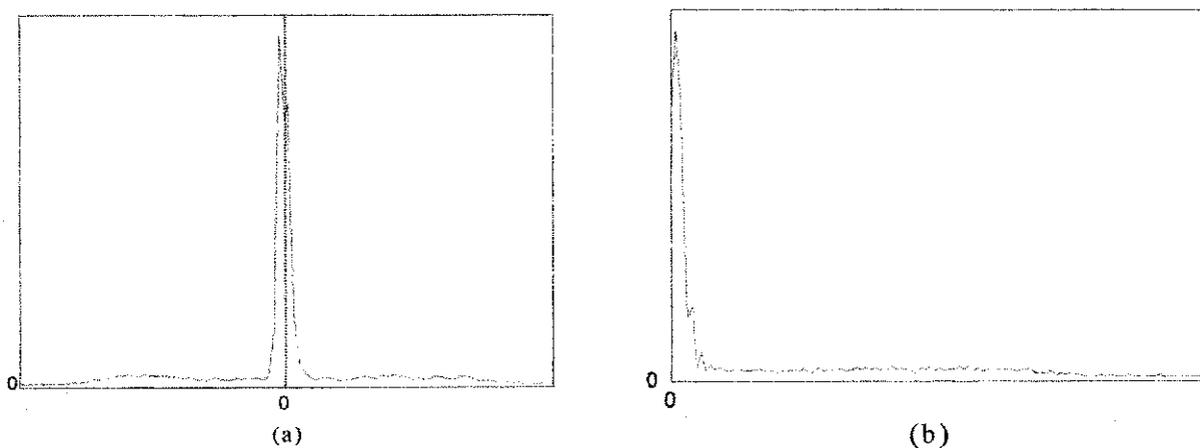


Figura 6.4. Histograma dos coeficientes 5 do registro 100/MLII, após quantização. (a) Antes do mapeamento. (b) Após o mapeamento.

Duas versões do método proposto empregando codificação de Golomb-Rice foram investigadas. A primeira, denominada GR1, consiste na substituição do codificador aritmético de A3 por L_b codificadores de Golomb-Rice, com um parâmetro k_c especificamente definido para cada conjunto de coeficientes quantizados $\{\hat{B}_{0c}, \hat{B}_{1c}, \dots, \hat{B}_{N_b-1,c}\}$, $c = 0, 1, \dots, L_b-1$, e um codificador de Golomb-Rice adicional, com parâmetro k' , para o conjunto de indicadores de último coeficiente não-nulo dos blocos, codificados diferencialmente como em A3. Os valores de k_c , $c = 0, 1, \dots, L_b-1$, e de k' são calculados pela aplicação da Equação (4.15) a cada conjunto citado.

O vetor \mathbf{q} , os parâmetros k_c e k' e a informação sobre a ordenação dos coeficientes são necessários para a decodificação e a dequantização, devendo ser adicionados ao arquivo comprimido. Para reduzir a representação destes dados extras, os elementos do vetor \mathbf{q} , os parâmetros k_c e os elementos da lista de índices ordenados são codificados diferencialmente, mapeados conforme a Equação (4.14) e a seqüência resultante é codificada pelo método de Golomb-Rice com parâmetro $k = 0$.

Quando o tamanho dos blocos é inferior ao tamanho médio do intervalo R-R, a distribuição de energia dos componentes de frequência varia substancialmente de um bloco para outro. As diferenças serão particularmente marcantes entre blocos com e sem complexos QRS. Alguns estudos mostram que a razão sinal-ruído do complexo QRS com relação aos demais componentes do sinal de ECG, incluindo ruídos e artefatos de movimento, é máxima em torno de 17 Hz (Afonso, 1993).

Desta afirmativa decorre que os coeficientes correspondentes a uma faixa em torno de 17 Hz possuem em geral maior energia ou variância em blocos contendo complexos QRS. Este fato permite discriminar entre a presença ou ausência destes complexos, sendo utilizado pelos algoritmos de detecção de QRS, que normalmente contêm um filtro passa-faixa em torno de 17 Hz. Assim, pode-se supor que o histograma do conjunto de coeficientes quantizados $\{\hat{B}_{0c}, \hat{B}_{1c}, \dots, \hat{B}_{(N_b-1),c}\}$ apresenta uma mistura de pelo menos duas distribuições geométricas de variâncias muito diferentes. A argumentação pode ser estendida considerando-se as diferenças no conteúdo de frequências de blocos contendo as demais ondas ou segmentos do ECG, mas as diferenças nestes casos são menos marcantes.

Como a codificação de Golomb é ótima para distribuições geométricas unilaterais, e não para misturas de distribuições deste tipo, a segunda versão do compressor empregando codificação de Golomb-Rice, denominada GR2, utiliza um método simples para separar os blocos em duas classes. Um bloco é atribuído à classe 0 quando a soma dos valores absolutos dos n coeficientes

contíguos localizados a partir da posição i do bloco é inferior à média desta soma em todos os blocos; caso contrário, o bloco é atribuído à classe 1.

Foram realizados testes para definir o valor do número de coeficientes n e da posição i , para blocos de tamanho 64. Com uma PRD fixa, a RC foi maximizada com o emprego de três coeficientes, de número 5, 6 e 7, correspondendo à faixa de frequências [14,3 Hz – 20 Hz]. O centro deste intervalo situa-se em 17,2 Hz, que coincide quase exatamente com a frequência central do filtro passa-faixa usado por muitos detectores de QRS, conforme citado anteriormente (Afonso, 1993). Vê-se, assim, que o critério aqui aplicado corresponde em última análise a uma classificação em termos da presença ou ausência de complexos QRS nos blocos. Nos testes empíricos do compressor, para todos os tamanhos de blocos, foi utilizado um conjunto de três coeficientes discriminadores, com o coeficiente central correspondendo sempre a aproximadamente 17 Hz.

Após a classificação dos blocos, efetua-se uma codificação independente para cada classe. Isto significa que a reordenação dos coeficientes e a codificação diferencial dos níveis DC e dos indicadores de último coeficiente não-nulo são efetuadas de forma independente em cada classe. Agora, para cada classe, tem-se um codificador para cada coeficiente, e mais um para os indicadores de último coeficiente não-nulo, com um total de $2L_b + 2$ codificadores de Golomb-Rice independentes. A informação necessária à decodificação é incluída de forma análoga à empregada em GR1, mas agora é necessário enviar ao decodificador mais um bit extra por bloco para identificar a classe.

De acordo com a descrição acima, pode-se considerar que as variantes GR1 e GR2 incluem uma espécie de modelagem condicionada a contextos. Em outros termos, em GR1 a modelagem dos coeficientes quantizados depende das posições por eles ocupadas no bloco; em GR2, a modelagem depende da posição do coeficiente e da classe a que pertence o bloco. A posição e a classe são, portanto, os contextos de modelagem dos coeficientes. Há ainda um contexto adicional para a modelagem dos indicadores de último coeficiente não-nulo, em GR1, e dois contextos adicionais para a modelagem destes indicadores em GR2.

A substituição do codificador aritmético por codificadores de Golomb-Rice representa uma redução substancial na complexidade computacional da etapa de codificação. O desempenho em termos de compromisso $R-D$, conforme já comentado, depende fundamentalmente do grau em que os histogramas nos diversos contextos apresentam comportamento semelhante ao de distribuições geométricas. Se as discrepâncias entre os histogramas e as distribuições geométricas não forem demasiadamente acentuadas, é possível que as vantagens decorrentes da aplicação de codificadores especificamente projetados para cada contexto compensem estes eventuais descasa-

mentos. Se esta hipótese revelar-se verdadeira, pode-se esperar que o compromisso $R-D$ de GR1 e GR2 seja semelhante ou superior ao de A3.

6.2 O DESCOMPRESSOR

A descompressão envolve os seguintes passos:

1. Decodificação de entropia.
2. Dequantização de $\hat{\mathbf{B}}_i$, $i = 0, 1, \dots, N_b - 1$, usando o \mathbf{q} ótimo.
3. Cálculo da DCT inversa dos blocos dequantizados, seguido por arredondamento dos valores para os inteiros mais próximos.
4. Concatenação dos blocos no domínio do tempo produzidos no passo 3.

Para o passo 1, há um processo de decodificação diferente para cada variante proposta. Este processo realiza as operações inversas do codificador, e pode ser inferido da própria descrição deste último.

6.3 METODOLOGIA DA AVALIAÇÃO EXPERIMENTAL

O compressor foi avaliado experimentalmente por intermédio de uma série de testes com um conjunto de sinais constituído por seções de registros do MIT-BIH *Arrhythmia Database*. Os códigos executáveis foram gerados no ambiente de programação C++ Builder 3.0, da Borland, e o processamento foi efetuado em um AMD K6-III de 400 MHz, com o sistema operacional Windows 98, da Microsoft.

Adotou-se a PRD, dada pela Equação (3.4), como métrica de distorção, e a RC, dada pela Equação (3.1), como indicador do grau de compressão. A tolerância para PRD alvo e para RC alvo

foram sempre de 0,049. Com PRD alvo de até 3%, alguns testes revelaram que os vetores q e t ótimos não apresentam elementos com valores superiores a 64, o que permitiu definir os valores máximos para os elementos destes vetores, ou seja, $Q_{max} = T_{max} = 64$. Nos experimentos com PRDs superiores a 3%, definiu-se $Q_{max} = T_{max} = 256$.

O efeito do tamanho dos blocos, L_b , no compromisso RC-PRD do método proposto foi aferido. O corpo de testes foi constituído pelos dois primeiros minutos de todos os registros do MIT-BIH *Arrhythmia Database*, totalizando 96 sinais, com 196 minutos de duração e 4.147.200 amostras. Para um cálculo computacionalmente mais eficiente da DCT, os tamanhos testados limitaram-se a potências de dois. Os sinais foram comprimidos pelas variantes A3, GR1 e GR2, com PRD alvo de 1,5%, 2,0%, 2,5% e 3,0%, e as RCs médias foram calculadas, para blocos de tamanho 16, 32, 64, 128, 256 e 512.

A PRD foi limitada ao valor máximo de apenas 3% porque o presente trabalho prioriza a qualidade dos sinais reconstruídos. Com uma tolerância de 0,04, foi possível atingir a PRD alvo em todos os casos. As curvas RC x L_b resultantes são apresentadas, para os diversos valores de PRD. Para as variantes menos elaboradas e de menor desempenho, A1 e A2, foram efetuados testes similares envolvendo os dois primeiros minutos dos registros 100/MLII, 100/V5 e 119/MLII. As curvas RC x L_b de A1 e A2 apresentam comportamento muito semelhante às de A3 e, por este motivo, não são exibidas.

Em todos os casos, a maior RC foi atingida com $L_b = 64$. Por este motivo, este tamanho foi selecionado para tabelar as RCs das variantes A2, A3, GR1 e GR2, para PRDs de 1,5%, 2,0%, 2,5% e 3,0%. A curva RC-PRD da variante de melhor desempenho, GR2, foi traçada para PRD variando de 1,5% a 10%, a fim de permitir uma avaliação mais completa do seu comportamento *R-D*. A variante A1, a primeira desenvolvida no transcurso do presente trabalho, possui uma etapa de codificação de entropia muito simples e, assim, não foram feitos testes completos para definição de sua relação RC-PRD. Para esta variante, os dois primeiros minutos dos registros 100/MLII, 100/V5 e 119/MLII foram comprimidos com RC alvo de 6,0:1, e a PRD alvo foi calculada.

Para uma melhor avaliação dos ganhos provenientes das sucessivas modificações efetuadas na etapa de codificação, o ganho de compressão de cada variante com relação às demais foi tabela-

do. O ganho de compressão de um método em relação a outro é definido como a razão entre a RC obtida pelo primeiro método e a RC obtida pelo segundo método, para uma dada PRD.

São exibidos diversos traçados de sinais originais, normais e patológicos, e suas reconstruções por GR2 em vários níveis de distorção, para avaliação visual dos efeitos da compressão. Os sinais selecionados contêm uma série de diferentes características, incluindo ruídos severos, e foram submetidos a dois especialistas para inspeção visual da qualidade.

Para comparação direta da variante de melhor desempenho, GR2, com a de outros compressores de ECG, foram selecionados diversos resultados da literatura que reportam a RC e a PRD (ou PRD2) obtidas com seções precisamente definidas de registros do MIT-BIH *Arrhythmia Database*. Os mesmos sinais foram comprimidos por GR2 para a mesma RC e as PRDs foram comparadas. Os resultados reportados na literatura foram em alguns casos obtidos após alteração na taxa de amostragem ou no número de *bits* por amostra do sinal. Sabendo que tais alterações podem afetar substancialmente os resultados, os sinais foram igualmente alterados antes da compressão por GR2. Quando a taxa de amostragem foi reduzida para 250 Hz, o tamanho do bloco ótimo, utilizado por GR2 nas comparações, se reduziu para 32.

Uma comparação entre a variante A3 e dois outros métodos de compressão de ECG baseados em DCT, propostos por Lee e Buckley (1999), foi incluída. A vantagem fundamental de A3 em relação aos compressores de Lee e Buckley consiste na utilização de uma etapa de quantização mais elaborada, de forma que estas comparações fornecem um bom indicativo dos ganhos provenientes do esquema de quantização empregado em todas as variantes do método proposto no presente trabalho.

Capítulo 7

RESULTADOS

O presente capítulo reporta os resultados experimentais obtidos com o compressor proposto em todas as suas variantes.

As Figuras 7.1, 7.2 e 7.3 exibem as curvas de RC em função do tamanho do bloco, L_b , para as variantes A3, GR1 e GR2, respectivamente. A escala do eixo L_b é logarítmica para uma melhor visualização. As curvas foram obtidas por interpolação linear entre os pontos efetivamente calculados.

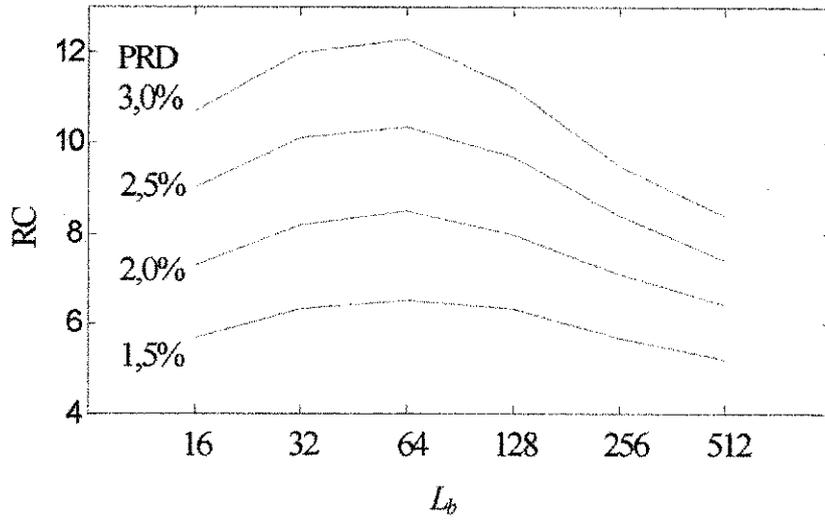


Figura 7.1 RC média de A3 em função do tamanho do bloco L_b , para PRD de 1,5%, 2,0%, 2,5%, 3,0%

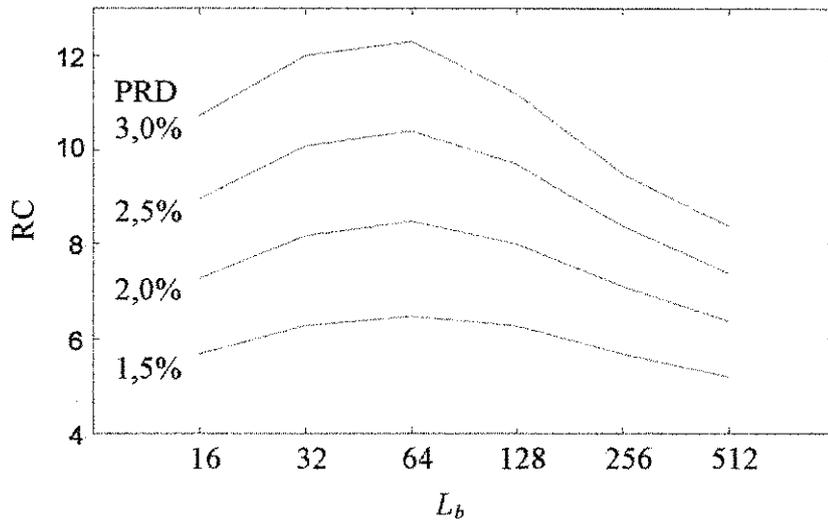


Figura 7.2. RC média de GR1 em função do tamanho do bloco L_b , para PRD de 1,5%, 2,0%, 2,5%, 3,0%

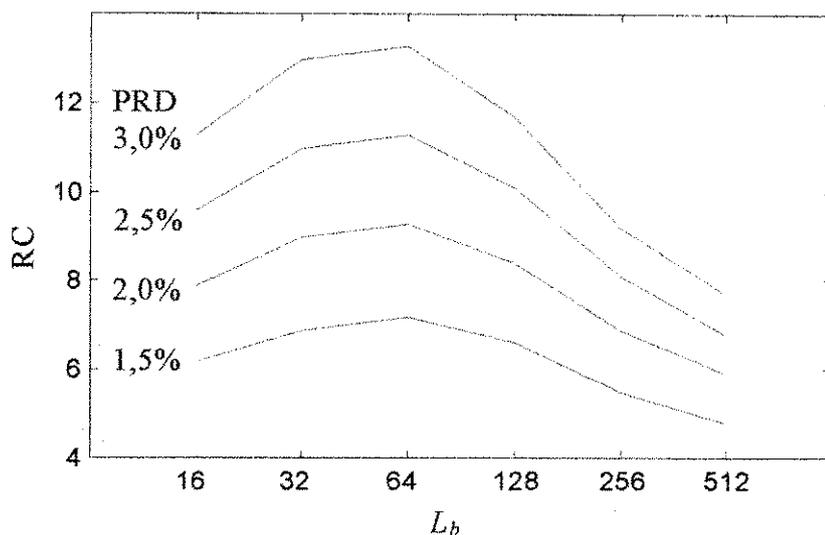


Figura 7.3. RC média atingida por GR2 em função do tamanho do bloco L_b , para PRD de 1,5%, 2,0%, 2,5%, 3,0%

A Tabela 7.1 apresenta, para cada valor de PRD, as RCs médias, máximas e mínimas, bem como os desvios padrões das RCs, obtidas por A2, A3, GR1 e GR2. No caso da variante A1, foram efetuados testes apenas com os dez primeiros minutos dos registros 100/MLII, 100/V5 e 119/MLII, com a RC fixa em 6,0:1, de modo que esta variante não foi incluída na Tabela 7.1, construída a partir de testes muito mais completos. Para RC = 6,0:1, A1 produziu uma PRD média de 1,8% com os sinais citados.

Tabela 7.1. RCs médias, máximas e mínimas, e desvios padrões das RCs obtidas com A2, A3, GR1 e GR2 para PRDs de 1,5%, 2,0%, 2,5% e 3,0%.

PRD (%)	RC média				Desvio Padrão				RC Máxima				RC Mínima			
	A2	A3	GR1	GR2	A2	A3	GR1	GR2	A2	A3	GR1	GR2	A2	A3	GR1	GR2
1,5	6,2	6,4	6,5	7,2	3,2	3,7	3,6	4,0	17,3	21,0	21,8	23,1	2,5	2,5	2,8	2,8
2,0	7,9	8,4	8,5	9,3	4,2	5,0	4,8	5,2	26,2	28,4	27,8	30,1	2,6	2,6	2,9	3,0
2,5	9,3	10,3	10,4	11,3	4,9	6,2	5,9	6,2	29,4	36,1	34,7	36,6	2,8	2,8	3,2	3,2
3,0	10,9	12,3	12,3	13,3	5,8	7,3	6,8	7,1	35,3	43,7	40,3	43,5	3,1	3,0	3,4	3,4

A Tabela 7.2 apresenta os ganhos de compressão de cada variante listada na primeira linha (A2, A3, GR1 e GR2) com relação às variantes listadas na primeira coluna. A última coluna, por exemplo, fornece os ganhos de compressão de GR2 com relação a A2, A3, GR1 e GR2. Os ganhos foram calculados a partir das RCs fornecidas na Tabela 7.1. Utilizando os dez primeiros minutos dos registros 100/MLII, 100/V5 e 119/MLII para os testes, o ganho de compressão de GR2 em relação a A1 foi de 1,57, para uma PRD média de 1,8%.

Tabela 7.2. Ganhos de compressão das variantes propostas para PRD = 2,5%

Referência	A2	A3	GR1	GR2
A2	1,00	1,10	1,12	1,22
A3	0,90	1,00	1,01	1,10
GR1	0,89	0,99	1,00	1,09
GR2	0,82	0,91	0,92	1,00

A Figura 7.4 mostra a curva RC x PRD da variante GR2, para valores de PRD variando de 1,5% a 10%. Novamente, a curva foi obtida por interpolação linear entre os pontos efetivamente calculados.

Para PRD alvo de 2,5% a implementação de GR2 levou 12,7 minutos para processar todo o conjunto de testes, que contém 192 minutos de sinal, correspondentes a 4.147.200 amostras. Este tempo inclui a leitura dos sinais a partir do disco rígido, a otimização para a PRD alvo, a compressão, a descompressão e o cálculo das RCs e PRDs.

As Figuras 7.5 a 7.9 exibem reconstruções de registros comprimidos com a variante GR2. A Figura 7.5 mostra um único ciclo do registro 100/MLII, e o sinal reconstruído com PRD de 2,5% e RC de 12,1:1, permitindo uma observação detalhada da qualidade da reconstrução em regiões importantes do sinal. A Figura 7.6 apresenta os primeiros 6 s do registro 100/MLII e as reconstruções com PRD de 1,5%, 2,0%, 2,5%, 3,0%, 4,0% e 5,0%. As Figuras 7.7, 7.8 e 7.9 mostram seções de 6 s dos registros 107/V1, 108/V1 e 109/MLII, respectivamente, e as reconstruções com PRD igual a 2,5%. A Figura 7.8 ilustra a performance de GR2 na presença de ruído severo.

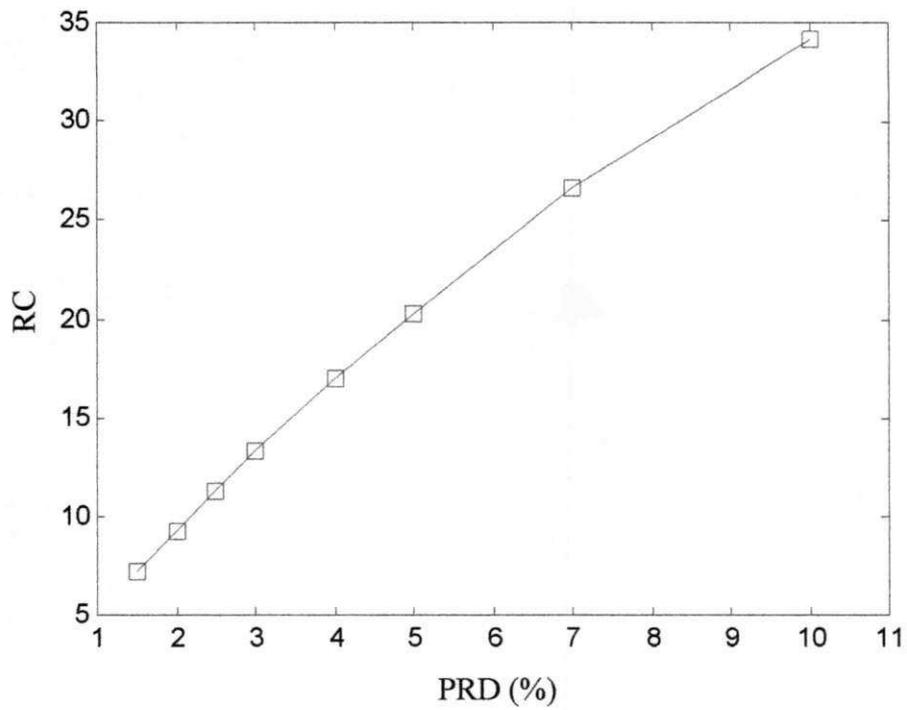


Figura 7.4. RC média atingida por GR2 em função da PRD.

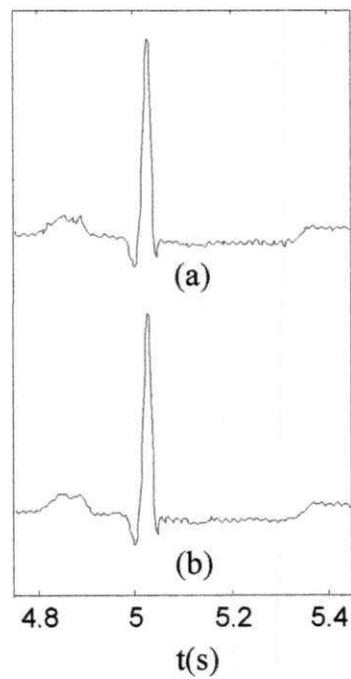


Figura 7.5. (a) Detalhe do registro 100/MLII;
(b) reconstrução com PRD = 2,5%, RC = 12,1:1

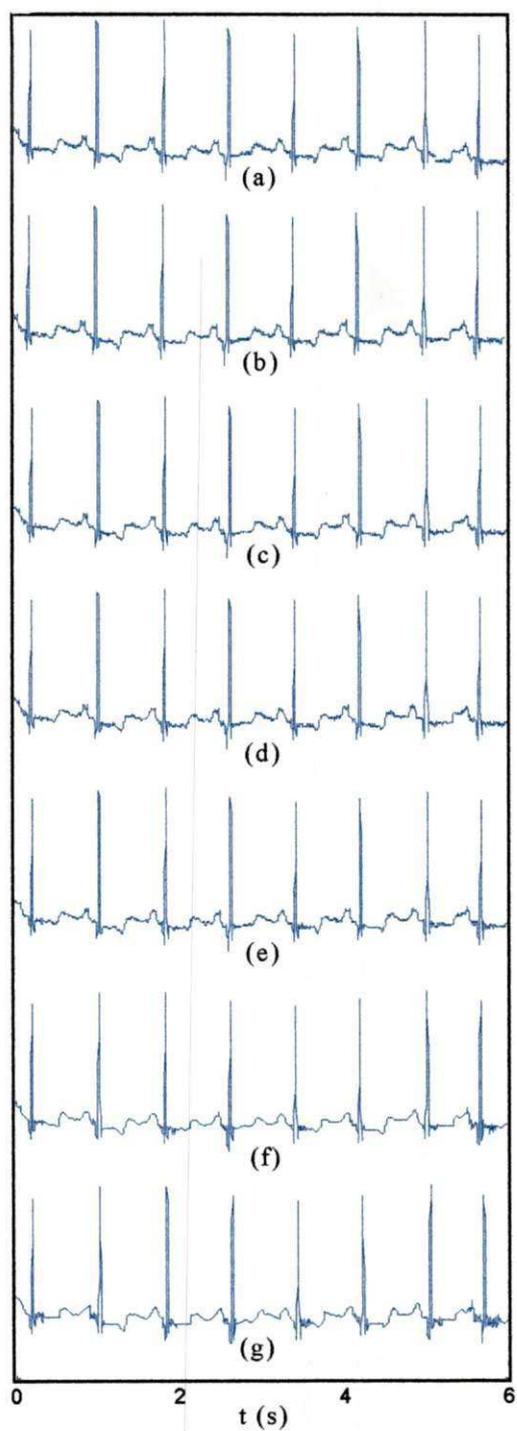


Figura 7.6. (a) Seção do registro 100/MLII; reconstruções com
 (b) PRD = 1,5%, RC = 5,9:1; (c) PRD = 2,0%, RC = 8,1:1; (d) PRD = 2,5%, RC = 10,7:1;
 (e) PRD = 3,0%, RC = 12,6:1; (f) PRD = 4,0%, RC = 18,1:1; (g) PRD = 5%, RC = 21,8:1.

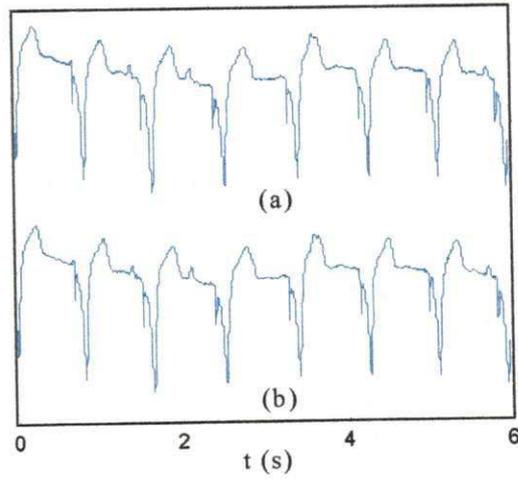


Figura 7.7. (a) Seção do registro 107/V1; (b) reconstrução com $PRD = 2,5\%$, $RC = 9,9:1$

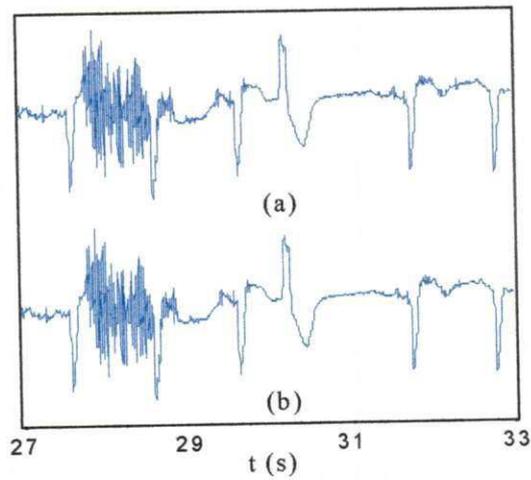


Figura 7.8. (a) Seção do registro 108/V1; (b) reconstrução com $PRD = 2,5\%$, $RC = 12,1:1$

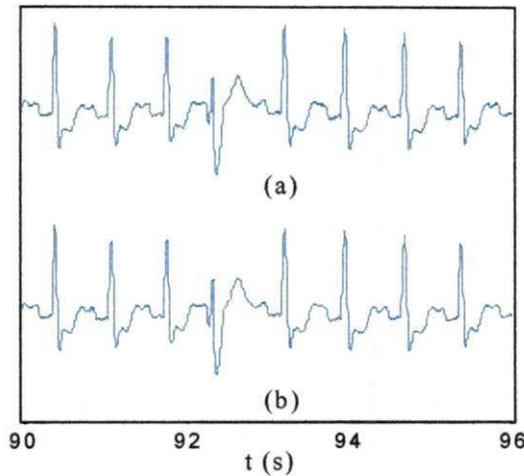


Figura 7.9. (a) Seção do registro 109/MLII; (b) reconstrução com PRD = 2,5%, RC = 11,6:1

A Tabela 7.3 permite comparar as PRDs (ou PRD2) obtidas por GR2 e por diversos compressores descritos na literatura. Em cada comparação, as RCs e os sinais de teste são idênticos. Números marcados por um asterisco são valores de PRD2. A última coluna da tabela traz uma descrição breve do tamanho e variedade do corpo de teste usado. Nas situações em que, considerando-se o especificado na literatura, foram efetuadas alterações na taxa de amostragem ou no número de *bits* por amostra dos sinais, os valores modificados são indicados entre parênteses na última coluna da tabela. Os resultados referentes aos três primeiros compressores apresentados nesta tabela, AZTEC, SAPA e LTP, foram obtidos por Zigel *et al.*(1997).

A Tabela 7.4 expõe, para três valores de RC, as PRDs obtidas com a variante A3 e com dois métodos apresentados por Lee e Buckley (1999): o CAB/2D-DCT, descrito no Capítulo 5, e uma técnica baseada na quantização uniforme ótima com passo único para todos os coeficientes da DCT unidimensional, aqui identificada como 1D-DCT.

Tabela 7.3. Comparação entre as PRDs produzidas por GR2 e por vários compressores da literatura

Da Literatura	GR2	RC	Conjunto de teste
Compressor	PRD	PRD	
AZTEC (Cox <i>et al.</i> , 1968)	15,5*	3,7 *	6,9 36 sinais de 1min. (250Hz)
SAPA (Ishijima <i>et al.</i> , 1983)	9,6*	3,7 *	6,9 36 sinais de 1min. (250Hz)
LTP (Nave e Cohen, 1993)	3,4*	2,3*	4,6 36 sinais de 1min. (250Hz)
	7,3*	3,7 *	6,9
	13,1*	8,7*	13,8
Sub-bandas/F16B (Husoy e Gjerde, 1996)	2,8*	2,5*	7,3 6 sinais de 14,22s.
Quantização adaptativa ótima / <i>wavelet</i> (Chen e Itoh, 1998)	6,8	5,3	12,5 1 sinal de 10min.
	8,9	7,0	16,0
	13,3	10,5	22,9
Quantização vetorial média-forma sem codificação de entropia (Cárdenas-Barreras e Lorenzo- Ginori, 1999)	4,1	2,0	9,6 96 sinais de 1min. (500 Hz)
	5,6	3,1	14,3
	8,3	4,2	19,1
	8,6	5,1	22,0
Quantização vetorial média-forma incluindo efeitos teóricos da codi- ficação de entropia (Cárdenas- Barreras e Lorenzo-Ginori, 1999)	4,1	2,8	13,1 96 sinais de 1min. (500 Hz)
	5,6	4,3	19,4
	8,3	7,0	28,7
	8,6	8,9	34,5
ADPCM (Lee e Buckley, 1999)	2,6	0,8	4,0 4 sinais de 10 min. (250 Hz,
	6,9	1,4	6,0 12 bits/amostra)
1D-DCT (Lee e Buckley, 1999)	4,1	0,8	4,0 4 sinais de 10 min. (250 Hz,
	7,5	1,4	6,0 12 bits/amostra)
	15,1	2,9	12,0

(*) PRD2

Tabela 7.3. (Cont.)

Compressor	Da Literatura	PRD	RC	Conjunto de teste
	PRD	GR2		
CAB/2D-DCT (Lee e Buckley, 1999)	2,5	0,8	4,0	4 sinais de 10 min. (250 Hz, 12 bits/amostra)
	3,5	1,4	6,0	
	6,1	2,9	12,0	
	10,6	6,8	24,0	
SPIHT-ECG (Lu <i>et al.</i> , 2000)	1,2	0,9	4,0	22 sinais de 10 min.
	3,0	2,2	10,0	
	6,5	4,9	20,0	
Quantização vetorial GW / <i>wavelet</i> (Miaou e Yen, 2000)	3,3*	1,5*	4,6	4 sinais de 15 min.
	6,3*	2,9*	9,4	
	8,2*	4,0*	12,4	
EZW com codificação aritmética adaptativa (Chagas <i>et al.</i> , 2000)	3,6	3,0	12,0	Banco completo.
Aproximação poligonal / DCT (Ba- tista, Melcher e Carvalho, 2000)	3,0	1,6	5,3	6 sinais de 1 min.
Fan (Nygaard <i>et al.</i> , 2001)	6,1	3,3	11,0	6 sinais de 10 min.
Min-Ave-ORD/VCL ótimo (Nygaard <i>et al.</i> , 2001)	3,5	3,3	11,0	6 sinais de 10 min.

(*) PRD2

Tabela 7.4. PRDs (%) geradas pelos métodos CAB/2D-DCT, 1D-DCT, e A3

RC	PRD		
	CAB/2D-DCT	1D-DCT	A3
4,0	2,5	4,1	1,1
6,0	3,5	7,5	1,7
12,0	6,1	15,1	4,0

Capítulo 8

DISCUSSÃO E CONCLUSÕES

Devido à grande quantidade de dados gerada pela crescente popularização do ECG digital, a compressão de sinais eletrocardiográficos tem se constituído em uma ativa área de pesquisa nos últimos 30 anos.

O desempenho de um compressor com perdas pode ser avaliado observando-se a relação entre três aspectos normalmente conflitantes: o grau de compressão, a qualidade dos sinais reconstruídos e o tempo de processamento. Em geral, para um dado compressor, mantendo-se o tempo de processamento fixo, a qualidade da reconstrução tende a diminuir à medida que aumenta a razão de compressão. Analogamente, aumentar o grau de compressão sem degradação da qualidade exige, em geral, o emprego de um esquema de compressão mais complexo, o que se reflete em um maior tempo de processamento.

Dada a importância diagnóstica do eletrocardiograma, a qualidade da reconstrução é um fator crítico. Diversos trabalhos divulgados na literatura especializada enfocam prioritariamente a obtenção de razões de compressão elevadas, em detrimento da qualidade da reconstrução. PRDs superiores a 10%, e até a 20%, têm sido reportadas. Com estes níveis de distorção, a aplicabilidade dos compressores restringe-se a situações em que uma forte degradação na qualidade é aceitável, como na análise de arritmias.

No presente trabalho, ao contrário, enfatiza-se a qualidade da reconstrução. Com uma PRD de apenas 3%, atingiu-se uma RC média de 13,3:1. Com esta mesma PRD, alguns sinais foram comprimidos com RCs superiores a 40,0:1 pelo método aqui proposto, que combina transformada DCT, quantização uniforme com zona morta ótima e codificação de entropia. Foram propostas cinco versões para a etapa de codificação de entropia, originando cinco variantes do método, A1, A2, A3, GR1 e GR2. As quatro últimas foram avaliadas por intermédio de experimentos com os dois minutos iniciais de todos os registros do MIT-BIH *Arrhythmia Database*. O tempo para compressão e descompressão de cada sinal limitou-se em média a aproximadamente 1/15 da duração do sinal, o que indica a aplicabilidade do método em situações práticas.

Os resultados mostram que o método permite atingir boas razões de compressão com distorção muito baixa. A variante de melhor desempenho, GR2, atingiu uma RC média de 11,3:1, com PRD de 2,5%. Para esta PRD, constatou-se uma elevada qualidade dos traçados reconstruídos, como pode ser confirmado pela inspeção dos trechos expostos nas Figuras 7.5 a 7.9. Observando-se a Figura 7.5, percebe-se que a degradação na qualidade começa a tornar-se claramente perceptível com PRDs superiores a 3%. Como o presente trabalho preconiza a manutenção da qualidade da reconstrução, nas demais figuras a PRD alvo ficou restrita a 2,5%. Neste nível de distorção, todos os traçados apresentados indicam uma excelente preservação da morfologia dos complexos QRS e das características clinicamente importantes dos sinais, conforme atestado por dois especialistas.

O efeito do tamanho dos blocos no compromisso RC-PRD foi experimentalmente aferido e os resultados mostraram que, para todas as PRDs testadas, a RC foi maximizada para $L_b = 64$, como se pode ver nas Figuras 7.1, 7.2 e 7.3. Considerando-se ainda que o tempo de processamento da DCT aumenta com o tamanho dos blocos, conclui-se que a utilização de blocos com mais de 64 amostras não é, a princípio, uma boa opção. Por outro lado, blocos maiores reduzem o efeito de

blocagem, particularmente importante com PRDs elevadas. No presente trabalho, onde a ênfase recai sobre a qualidade da reconstrução, as PRDs são mantidas em níveis reduzidos, e tais artefatos são muito pouco visíveis. Conseqüentemente, não há na presente situação vantagens em aumentar os blocos para 128 ou mais amostras.

As curvas RC x PRD de A3 e GR1 coincidem quase exatamente para todas as PRDs e tamanhos de bloco testados. Uma vez que a variante GR1 é computacionalmente menos complexa, sua utilização é mais vantajosa. GR2, por outro lado, mostra uma RC maior que a das demais variantes para $L_b = 64$ mas, acima deste valor, sua superioridade reduz-se à medida que L_b aumenta. Este é um comportamento esperado, pois o aumento de L_b tende a reduzir a variação, de um bloco para outro, da distribuição de energia dos componentes de frequência, que foi a própria justificativa para a separação dos blocos em duas classes.

Não há, na literatura, muitos estudos a respeito dos efeitos do tamanho do bloco em compressores de ECG baseados em transformadas. Em geral, supõe-se que aumentar o tamanho dos blocos aumenta a RC para uma dada distorção, mas também aumenta o tempo de processamento e os requisitos de memória. Como conseqüência deste pressuposto, que não foi confirmado pelos testes aqui realizados e expostos na Figura 7.1, utiliza-se comumente o maior bloco possível que não ultrapasse limitações computacionais predefinidas.

Lu *et al.* (2000), por exemplo, sugerem que os blocos devem ser suficientemente longos para conter vários batimentos e, ao mesmo tempo, suficientemente pequenos para que o tempo de processamento e a utilização de memória sejam aceitáveis. Com base nesta argumentação optaram, no compressor SPIHT-ECG, por blocos com 1024 amostras. O SPIHT-ECG, a propósito, é uma adaptação, para sinais de ECG, do SPIHT (Said e Pearlman, 1996), considerado um dos melhores algoritmos de compressão com perdas de imagens da atualidade (Lu *et al.*, 2000). Outros autores não informam os critérios adotados para a definição do tamanho dos blocos, como Chen e Itoh (1998) e Miaou e Yen (2000), que também utilizam 1024 amostras por bloco.

No compressor de ECG baseado em *wavelet* de Strang e Nguyen (1996), experimentos com diversos sinais eletrocardiográficos mostraram que um aumento excessivo no tamanho dos blocos degrada o compromisso RC-PRD. Um bom compromisso foi atingido com 2000 amostras por bloco.

Lee e Buckley (1999) investigaram, para o compressor CAB/2D-DCT, o efeito do tamanho dos blocos sobre a PRD, com diversos valores de RC mas, devido à abordagem bidimensional empregada, os resultados não se aplicam diretamente ao caso unidimensional típico dos compressores de ECG. Os testes envolveram os registros 100/MLII, 100/V5, 119MLII e 119/V1, e foram avaliados blocos de tamanho 4x4, 8x8, 16x16, 32x32 e 64x64, para RCs de 4:1, 6:1, 12:1 e 24:1. Os autores não se detiveram na análise dos resultados mas fica evidente que o compromisso R-D não necessariamente melhora a medida em que o tamanho dos blocos aumenta, o que concorda com os resultados expostos na Figura 7.1. Com o CAB/2D-DCT, com nenhum valor de RC o maior tamanho de bloco testado, 64x64, resultou na menor PRD. Em geral, as menores distorções foram obtidas com blocos de tamanho 16x16 e 32x32. Estes e os resultados do compressor aqui proposto sugerem que blocos relativamente pequenos em comparação com, por exemplo, os blocos de 2000 amostras do compressor de Strang e Nguyen (1996), são mais adequados para compressão de ECG baseada em DCT.

A Tabela 7.2 permite aferir os ganhos provenientes das alterações efetuadas na etapa de codificação de entropia. Para $PRD = 2,5\%$ vê-se que o ganho de A3 com relação a A2 é de 1,1, ou 10%. Este ganho deve-se à ordenação dos coeficientes, pois esta é a única diferença entre A3 e A2. Analogamente, o ganho de GR1 em relação a A3 provém da substituição da codificação aritmética pelo conjunto de codificadores de Golomb-Rice. Neste caso, o ganho é de apenas 1,01, ou 1%, mas a redução na complexidade da codificação é uma vantagem adicional associada a GR1. O efeito da divisão dos blocos em duas classes reflete-se no ganho de 1,09, ou 9%, de GR2 em relação a GR1. Finalmente, o ganho de 1,22, ou 22%, de GR2 em relação a A2, representa o efeito conjunto da ordenação de coeficientes, divisão dos blocos em duas classes e codificação de Golomb-Rice. Estes resultados mostram que é possível obter bons ganhos na razão de compressão, sem qualquer alteração na distorção, com elaborações simples na etapa de codificação de entropia.

Em geral, nas comparações com outros compressores de ECG clássicos e modernos, GR2 produziu distorções, medidas pela PRD ou PRD2 consideravelmente menores. Exceto em um caso, que será analisado adiante, a PRD ou PRD2 gerada por GR2 foi sempre inferior àquelas geradas pelos demais compressores. Alguns poucos métodos apresentaram desempenho similar ao de GR2: o compressor por sub-bandas com filtro F16B, o Min-Ave-ORD/VCL ótimo e o esquema baseado em quantização vetorial média-forma (*mean-shape vector quantizer* - MSVQ) incluindo

os efeitos teóricos da etapa de codificação de entropia. É interessante notar que estes dois últimos são métodos no domínio do tempo.

O Min-Ave-ORD/VLC ótimo (Nygaard *et al.*, 2001) constrói uma aproximação poligonal ótima, no sentido taxa-distorção operacional (*operational rate-distortion* – ORD). A aproximação poligonal construída minimiza o erro médio entre esta e o sinal original, de forma que, pela própria natureza do método, nenhuma outra técnica baseada em interpolação linear produzirá um erro menor para a mesma taxa de bits. As amplitudes e posições dos segmentos de retas são diferencialmente codificadas e os valores resultantes são codificados por um codificador de comprimento variável otimizado (*variable length coder* – VLC).

O MSVQ (Cárdenas-Barreras e Lorenzo-Ginori, 1999) particiona o sinal em pequenos segmentos e subtrai de cada um deles a sua média. Os segmentos de média zero assim obtidos são vetorialmente quantizados, enquanto as médias são quantizadas como escalares. O método não inclui codificação de entropia.

Em adição aos resultados efetivamente obtidos pelo MSVQ, Cárdenas-Barreras e Lorenzo-Ginori informam as RCs que seriam *teoricamente* atingidas se fosse utilizada codificação de entropia. Estas projeções decorrem provavelmente de estimativas baseadas na entropia da sequência gerada pelo MSVQ, e não levam em consideração as informações extras que devem ser adicionadas ao arquivo comprimido para permitir a decodificação. Se esta informação ocupa um espaço fixo no arquivo comprimido, sua importância na redução da RC real cresce com o grau de compressão, pois a proporção entre o espaço ocupado pela informação e o tamanho do arquivo comprimido cresce com a diminuição do tamanho deste último.

Isto poderia explicar a maior degradação, para RCs elevadas, do compromisso RC-PRD de GR2, que inclui todas as informações necessárias à decodificação. De fato, apenas com a maior RC testada, ou seja, 34,5:1, a relação RC-PRD de GR2 foi inferior a do MSVQ com o efeito teórico da codificação de entropia incluído. Neste caso, a PRD produzida pelo MSVQ foi de 8,6%, e a PRD produzida por GR2 foi de 8,9%. Ressalte-se que esta foi a única situação desfavorável a GR2 em todas as comparações.

Apesar do já mencionado efeito da não inclusão da informação necessária à decodificação pelo MSVQ com codificação de entropia teórica, sua superioridade na RC de 34,5:1 parece decorrer mais do seu comportamento pouco usual na transição de RC de 28,7:1 para 34,5:1 do que de uma eventual degradação na relação RC-PRD de GR2 para RCs elevadas. Com o MSVQ, a PRD aumentou 2,7 pontos percentuais quando a RC passou de 19,4:1 para 28,7:1, e apenas 0,3 pontos percentuais quando a RC passou de 28,7:1 para 34,5:1. Com GR2, o aumento na PRD foi também de 2,7 pontos percentuais na transição da RC de 19,4:1 para 28,7:1, e de 1,9 ponto percentual na transição da RC de 28,7:1 para 34,5:1. Como pode ser observado na Tabela 7.3, o MSVQ sem codificação de entropia apresenta relação RC-PRD marcadamente inferior à de GR2, para todos os valores de RC testados.

Dentre os compressores baseados em *wavelet*, a combinação EZW-codificação aritmética adaptativa, proposta por Chagas *et al.* (2000), com PRD = 3,6% para RC = 12,0:1, foi a que mais se aproximou do desempenho da variante GR2, que produziu PRD = 3,0% para esta mesma RC.

Atente-se particularmente para a comparação entre A3 e o compressor 1D-DCT de Lee e Buckley (1999), apresentada na Tabela 7.4. Para uma RC de 12,0:1, por exemplo, a PRD resultante da compressão por A3 é 3,8 vezes menor que aquela produzida pelo 1D-DCT. A diferença fundamental de A3 em relação a este método reside na otimização da quantização que, no 1D-DCT, é restrita a um passo de quantização único para todos os coeficientes. A diferença marcante nos resultados mostra a importância do procedimento de otimização da quantização com zona morta, que permite definir limiares e tamanhos de passo de quantização diferentes para os diversos coeficientes. Esta abordagem fez com que A3 superasse a estratégia bidimensional do CAB/2D-DCT, como pode ser visto na Tabela 7.4.

Considerando-se os resultados obtidos, conclui-se que o objetivo do presente trabalho foi atingido. Particularmente, as comparações do método proposto com os compressores baseados em *wavelet*, em termos de compromisso RC-PRD, revelaram-se notadamente favoráveis ao primeiro.

Além da proposta de um novo método de compressão de ECG, uma outra contribuição do presente trabalho reside na demonstração experimental de que é possível projetar compressores baseados em DCT com desempenho semelhante aos melhores compressores divulgados na literatu-

ra. Especificamente, a respeito da discussão DCT x DWT, apesar desta última ser comprovadamente uma excelente ferramenta para a compressão de dados, as comparações indicaram que métodos não baseados em *wavelets* foram os que mais se aproximaram do desempenho do compressor aqui proposto, em termos de relação RC-PRD.

Como já sugerido em outros estudos, na aplicação à compressão de dados a DCT e a DWT comportam-se de maneira semelhante, com boa capacidade para descorrelacionar e concentrar a energia dos sinais em poucos coeficientes. As eventuais diferenças de desempenho dependem fundamentalmente de um projeto adequado para as etapas de quantização e de codificação de entropia. A simplificação da quantização e da codificação é a principal motivação para o emprego de transformadas na compressão de dados; se for possível definir um modelo prático para alimentar o codificador de entropia, capaz de reproduzir de forma suficientemente precisa as dependências potencialmente complexas existentes entre os elementos da seqüência original, a etapa de transformada pode tornar-se dispensável (Goyal, 2001).

Em algumas das comparações com compressores da literatura foi necessário alterar o número de bits por amostra e a frequência de amostragem dos sinais. Observou-se então que uma redução de 360Hz para 250Hz causou uma redução no tamanho ótimo dos blocos de 64 para 32. O efeito do número de bits por amostra e da frequência de amostragem dos sinais sobre as curvas RC x L_b deverá ser proximamente aprofundado.

O compromisso RC-PRD de A3 é muito superior ao dos compressores 1D-DCT e CAB/2D-DCT de Lee e Buckley (1999). Como a diferença fundamental entre estes compressores e A3 é a otimização da quantização, estes resultados mostram a importância do procedimento de otimização do método proposto. Além disso, posto que a superioridade do CAB/2D-DCT sobre o 1D-DCT decorre fundamentalmente da conversão para uma organização bidimensional, e dado o melhor desempenho de GR2 com relação às outras variantes do método proposto, será importante investigar os efeitos da combinação da estratégia de otimização da quantização aqui empregada com a codificação de GR2 e a abordagem bidimensional do CAB/2D-DCT.

A estratégia adotada define uma única instância de q e de t para cada sinal. Contudo, há diferenças marcantes entre o conteúdo de frequência dos blocos contendo complexos QRS, por exemplo, e o dos demais blocos. Por um raciocínio semelhante ao que conduziu à variante GR2, com

a separação dos blocos em duas classes e a codificação independente de cada uma delas, a otimização da quantização poderia ser feita de maneira independente para cada classe. Ao final da otimização, seriam definidos vetores q e t ótimos específicos para cada classe. Além de um eventual ganho de compressão, pode-se esperar também um aumento substancial no tempo de processamento, dada a necessidade de efetuar um procedimento de otimização para cada classe.

A codificação de Golomb-Rice é um caso particular da codificação de Golomb que, por sua vez, é ótima para distribuições geométricas. Para estas distribuições, o ganho de compressão médio da codificação de Golomb com parâmetro m ótimo, em relação à codificação de Golomb-Rice com parâmetro k ótimo, é reduzido, não compensando, em algumas situações o acréscimo na complexidade da construção do código (Seroussi e Weinberger, 1997). Esta afirmativa, entretanto, aplica-se principalmente quando há impreterível necessidade de manter a demanda por recursos computacionais em níveis muito limitados. No compressor aqui proposto a complexidade global é dominada pela etapa de otimização da quantização, de modo que o leve acréscimo de complexidade proveniente da substituição da codificação de Golomb-Rice pela de Golomb seria desprezível.

Melhorias na relação RC-PRD de GR1 e GR2 podem decorrer ainda de ações corretivas para aqueles histogramas que diverjam substancialmente da hipótese de distribuição geométrica bilateral centrada em zero, como a correção do desvio com relação ao zero, além de outras medidas citadas por Seroussi e Weinberger (1997).

Como já afirmado, a complexidade do método proposto é dominada pela etapa de otimização da quantização. Há diversas formas de atuar sobre esta etapa para reduzir o tempo de processamento, provavelmente por um fator superior a dez. Uma maior tolerância para a PRD alvo, por exemplo, reduziria o número de iterações do método da bissecção. Além disso, conforme discutido no fim da Seção 6.1.2, para uma tolerância suficientemente elevada, a PRD poderia ser calculada no domínio da DCT, eliminando a necessidade de calcular a DCT inversa em cada iteração.

Na construção das BCIs, durante a etapa de definição dos vetores q e t ótimos, é necessário calcular distorções, dadas pelo MSE, e entropias para cada um dos coeficientes. Um acréscimo de t para $t + 0,5$ no limiar de zona morta altera apenas os valores situados no intervalo $[t, t + 0,5)$, o que permite um cálculo incremental muito eficiente das distorções e entropias (Ratnakar, 1997).

Alterações no tamanho do passo de quantização, por outro lado, podem afetar todos os valores, o que exige que as distorções e entropias sejam integralmente recalculadas. Com base nestes fatos, Ratnakar propôs uma versão de menor complexidade para seu compressor de imagens, eliminando alguns tamanhos de passo de quantização na busca pelos vetores ótimos. Em vez de construir as BCIs para os 64 tamanhos de passo inteiros entre 1 e 64, pode-se construí-las apenas para tamanhos de passo múltiplos de 4, por exemplo.

Os experimentos com imagens demonstraram que a relação entre o tempo de processamento e o número de tamanhos possíveis para os passos de quantização é aproximadamente linear, e que a diferença no compromisso *R-D* do compressor RD-OPT (Ratnakar, 1997) com 255 tamanhos de passo inteiros entre 1 e 255 e com apenas 14 tamanhos de passo (2, 4, 6, 8, 10, 14, 18, 22, 30, 38, 46, 96, 128 e 255) é imperceptível. Com isto, o tempo de processamento é reduzido por um fator de aproximadamente dez sem degradação perceptível de qualidade. Uma alternativa não investigada por Ratnakar seria limitar os tamanhos de passo a potências de dois, o que traria o benefício extra de permitir a substituição das divisões necessárias à quantização por simples deslocamentos binários, com redução adicional no tempo de processamento.

Seria também interessante verificar o grau de degradação no comportamento *R-D* do compressor proposto com a eliminação da zona morta, bem como o efeito de determinar os vetores q e t ótimos usando apenas parte do registro. Para um sinal de 30 minutos, por exemplo, os dois primeiros minutos poderiam ser empregados para a otimização, e os vetores definidos seriam aplicados na quantização do sinal inteiro.

Por intermédio da transformada DCT, o compressor proposto explora a correlação entre as amostras adjacentes do sinal, denominada de *correlação intracanal*. Quanto maior esta correlação, maior o poder da DCT em concentrar a energia do sinal em um número reduzido de coeficientes de baixa frequência. No caso de sinais eletrocardiográficos com múltiplas derivações coletadas simultaneamente, ou quase simultaneamente, existe também correlação entre os canais (Mammen e Ramamurthi, 1990; Cohen e Zigel, 1998), ou *correlação intercanais*. A inclusão, no compressor proposto, de algum mecanismo capaz de explorar adequadamente a correlação intercanais pode resultar em consideráveis ganhos de compressão.

Seria relevante ainda analisar a possibilidade de adaptar o método proposto para utilização com outras transformadas práticas, incluindo DWT e LOT (*lapped orthogonal transform*) (Rao e Yip, 1990). Na LOT, que guarda forte relação com a DCT, as funções de base se superpõem em blocos adjacentes e decaem suavemente para zero nas fronteiras dos blocos tornando a transformada virtualmente livre dos efeitos de blocagem observáveis na compressão baseada em DCT com RCs muito elevadas. A LOT exige o cálculo da DCT de cada bloco e os cálculos adicionais da DCT das amostras ímpares dos blocos e da DST (*discrete sine transform*) dos coeficientes gerados por esta última DCT, além de algumas operações extras. Ao todo, a LOT perfaz aproximadamente o dobro das operações da DCT, produzindo o mesmo número de coeficientes que esta. Alguns experimentos com imagens indicam que a LOT aumenta a qualidade subjetiva das reconstruções com baixas taxas de bit, em comparação com a DCT (Rao e Yip, 1990).

Finalmente, observa-se que o método proposto não explora as características específicas dos sinais eletrocardiográficos. Por exemplo, não há necessidade de detecção dos complexos QRS. Esta generalidade permite sua aplicação a outros sinais unidimensionais, tais como sinais eletroencefalográficos e eletromiográficos. A eficiência do método poderia ser testada, por exemplo, nos sinais tipicamente presentes nos registros multicanais de apnéia do sono (Goldberger *et al.*, 2000). Estes registros podem envolver a coleta simultânea de sinais de ECG, EEG, eletromiograma, pressão sanguínea, respiração, saturação arterial de oxigênio e outros.

REFERÊNCIAS

ABENSTEIN, J. P.; AND TOMPKINS, W. J. "A New Data-reduction Algorithm for Real-time ECG Analysis." *IEEE Transactions on Biomedical Engineering*, v. BME-29, n. 1, pp. 43-48, 1982.

AFONSO, V. X. "ECG QRS Detection." In: *Biomedical Digital Signal Processing*, Ed. W. Tompkins, New Jersey: Prentice Hall, p. 236-264, 1993.

AHMED, N.; MILNE, P. J.; AND HARRIS, S. G. "Electrocardiographic Data Compression Via Orthogonal Transforms." *IEEE Transactions on Biomedical Engineering*, v. BME-22, n. 6, pp. 484-487, 1975.

ALGAZI, V. R.; AND ESTES, R. R. "Comparative Performance of Wavelet and JPEG Coders at High Quality." *Very High Resolution and Quality Imaging, Proceedings of the SPIE*, v. 3025, 1997.

ANTONIOL, G.; AND TONELLA, P. "EEG Data Compression Techniques." *IEEE Transactions on Biomedical Engineering*, v. 44, n. 2, p. 105-114, 1997.

BAILEY, J. J.; BERSON, A. S.; GARSON, A.; HORAN, L. G., MACFARLANE, P. W., MORTARA, D. W. AND ZYWIEZ, C. "Recommendations for standardization and specifications in automated electrocardiography: bandwidth and digital signal processing." *American Heart Association Special Report, Circulation*, v. 81, p. 730-739, 1990.

BATISTA, L. V.; CARVALHO, L. C.; AND MELCHER, E. U. K. "An ECG Signal Compressor Based on Optimized Quantization of Discrete Cosine Transform Coefficients." *Anais do II Congresso Latino-americano de Engenharia Biomédica*, Havana, 2001.

BATISTA, L. V.; CARVALHO, L. C.; E MELCHER, E. U. K. "Um Método para Compressão de Eletrocardiogramas Baseado na Quantização Otimizada dos Coeficientes da Transformada Cosseno Discreta". *Anais do XVII Congresso Brasileiro de Engenharia Biomédica*, Florianópolis, p. 1311-1315, 2000.

BATISTA, L. V.; MELCHER, E. U. K.; AND CARVALHO, L. C. "Compression of ECG Signals by Optimized Quantization of Discrete Cosine Transform Coefficients." *Medical Engineering & Physics*, v. 23, n. 2, p. 127-134, 2001.

BATISTA, L. V.; MELCHER, E. U. K.; E CARVALHO, L. C. "Um esquema de compressão de ECG utilizando seleção de picos e transformada cosseno discreta". *Revista Brasileira de Engenharia Biomédica*, v. 16, n. 1, p. 39-48, 2000.

BELL, T. C.; CLEARY, J. G.; AND WITTEN, I. H. *Text Compression*, Englewood Cliffs:Prentice Hall, 1990.

BINDER, F. V. *Multimídia: Animação Gráfica e Sons Utilizando Linguagem C*. São Paulo: Érica, 1994.

CÁRDENAS-BARRERAS, J. L.; AND LORENZO-GINORI, J. V. "Mean-Shape Vector Quantizer for ECG Signal Compression." *IEEE Transactions on Biomedical Engineering*, v. 46, n. 1, p. 62-70, 1999.

CARPENTIERI, B.; WEINBERGER, M. J.; AND SEROUSSI, G. "Lossless Compression of Continuous-Tone Images." *Proceedings of the IEEE*, v. 88, n. 11, p. 1797-1809, 2000.

CARVALHO, L. C. Apostila da disciplina "Instrumentação Médico-Hospitalar". NETEB/MEB. Universidade Federal da Paraíba, João Pessoa, 1998.

CHAGAS, A. V.; DA SILVA, E. A. B.; AND NADAL, J. "ECG Data Compression Using Wavelets." *Computers in Cardiology*, n. 27, p. 423:426.

CHEN, J.; AND ITOH, S. "A Wavelet Transform-Based ECG Compression Method Guaranteeing Desired Signal Quality." *IEEE Transactions on Biomedical Engineering*, v. 45, n. 12, p. 1414-1419, 1998.

COGGINS, R. J.; AND JABRI, M. A. "A Low-Complexity Intracardiac Electrogram Compression Algorithm." *IEEE Transactions on Biomedical Engineering*, v. 46, n. 1, p. 82-91, 1999.

COHEN, A.; AND ZIGEL, Y. "COMPRESSION OF MULTICHANNEL ECG THROUGH MULTICHANNEL LONG-TERM PREDICTION." *IEEE Engineering in Medicine and Biology*, Janeiro/Fevereiro, 1998.

COX, J. R.; NOLLE, F. M.; FOZZARD, H. A.; AND OLIVER, JR., G. C. "AZTEC: A Preprocessing Program for Real-Time ECG Rhythm Analysis." *IEEE Transactions on Biomedical Engineering*, BME-15, p. 128-129, 1968.

CRAWFORD, M. H.; BERNSTEIN, S. J.; DEEDWANIA, P. C.; DIMARCO, J. P.; FERRICK, K. J.; GARSON, JR, A. *et al.* "ACC/AHA guidelines for ambulatory electrocardiography: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee on Ambulatory Electrocardiography)." *Journal of the American College of Cardiology*, v. 34, n. 3, p. 912-948, 1999.

CROWE, J. A.; GIBSON, N. M.; WOOLFSON, M. S.; AND SOMEKH, M. G. "Wavelet Transform as a Potential Tool for ECG Analysis and Compression." *Journal of Biomedical Engineering*, v. 11, p. 268-272, 1992.

DAVIS, G. M.; AND CHAWLA, S. "Image Coding Using Optimized Significance Tree Quantization." *Proceedings of the Data Compression Conference*, Snowbird, p. 387-396, March, 1997.

FURHT, B.; AND PEREZ, A. "An Adaptive Real-Time ECG Compression Algorithm with Variable Threshold." *IEEE Transactions on Biomedical Engineering*, v. 35, n. 6, p. 489-492, 1988.

GALLAGER, R. AND VOORHIS, D. V. "Optimal source codes for geometrically distributed integer alphabets." *IEEE Transactions on Information Theory*, v. IT-21, p. 228-230, March, 1975.

GOLDBERGER, A. L.; AMARAL, L. A. N.; GLASS L.; HAUSDORFF, J. M.; IVANOV, P. C. H.; MARK, R. G.; *et al.* "PhysioBank, PhysioToolkit, and Physionet: Components of a New Research Re-

source for Complex Physiologic Signals." *Circulation* 101:215e-220e [Circulation Electronic Pages, <http://circ.ahajournals.org/cgi/content/full/101/23/e215>], June 13, 2000.

GOYAL, V. K. "Theoretical Foundations of Transform Coding." *IEEE Signal Processing Magazine*, v. 18, n. 5, p. 9-21, 2001

GRAAF, P. M. A.; GOUDOEVER, J. V.; AND WESSELING, K. H. "Compressed Storage of Arterial Pressure Waveforms by Selection of Significant Points." *Medical & Biological Engineering & Computing*, v. 35, n. 5, 1997.

GUYTON, A. C. *Tratado de Fisiologia Médica*. Sexta edição, Rio de Janeiro: Interamericana, 1984.

HADLEY, G. *Programação Linear*. Rio de Janeiro: Guanabara Dois, 1982.

HANS, M. C. *Optimization of Digital Audio for Internet Transmission*. Ph.D. thesis, Georgia Institute of Technology, 1998.

HERMES, R.; GESELOWITZ, D. B.; AND OLIVER, G. "Development, Distribution, and Use of the American Heart Association Database for Ventricular Arrhythmia Detectors." *Computers in Cardiology*, p. 263-266, 1980.

HILTON, M. L. "Wavelet and Wavelet Packet Compression of Electrocardiograms." *IEEE Transactions on Biomedical Engineering*, v. 44, n. 5, p. 394-402, 1997.

HOANG, D. T. *Fast and Efficient Algorithms for Text and Video Compression*. Ph.D. thesis, Department of Computer Science, Brown University, Providence, March 1997.

HOWARD, P. G. *The Design and Analysis of Efficient Lossless Data Compression Systems*, Ph.D. thesis, Brown University, 1994.

HUFFMAN, D. A. "A Method for the Construction of Minimum Redundancy Codes." *Proceedings of the Institute of Radio Engineers*, n. 40, p. 1098-1101, 1952.

HUSOY, J. H.; AND GJERDE, T. "Computationally Efficient Sub-band Coding of ECG Signals." *Medical Engineering & Physics*, v. 18, n. 2, p. 132-142, 1996.

ISHIJIMA, M.; SHIN, S.; HOSTETTER, G. H.; AND SKLANSKY, J. "Scan-Along Polygonal Approximation for Data Compression of Electrocardiograms." *IEEE Transactions on Biomedical Engineering*, v. 30, n. 11, p. 723-729, 1983.

JALALEDDINE, S. M. S.; HUTCHENS, C. G.; STRATTAN, R. D.; AND COBERLY, W. A. "ECG Data Compression Techniques - A Unified Approach." *IEEE Transactions on Biomedical Engineering*, v. 37, n. 4, p. 329-343, 1990.

JAYANT, N. S.; AND NOLL, P. *Digital Coding of Waveforms*. Englewood Cliffs: Prentice Hall, 1984.

KOSKI, A. "Lossless ECG Encoding." *Computer Methods and Programs in Biomedicine*, v. 52, n. 1, p. 23-33, 1997.

LAI, K. "Data Reduction Techniques." In: *Biomedical Digital Signal Processing*, Ed. W. Tompkins, New Jersey: Prentice Hall, p. 193-215, 1993.

LAURANCE, N. K.; AND MONRO, D. M. "Embedded DCT Coding with Significance Masking." *Proceedings of the ICASSP*, Munique, v. 4, p. 2717:2720, April 1997.

LEE, H.; AND BUCKLEY, K. M. "ECG Data Compression Using Cut and Align Beats Approach and 2-D Transforms." *IEEE Transactions on Biomedical Engineering*, v. 46, n. 5, p. 556-564, March 1999.

LU, Z.; KIM, D. Y.; AND PEARLMAN, W. A. "Wavelet Compression of ECG Signals by Set Partitioning in Hierarchical Trees Algorithm." *IEEE Transactions on Biomedical Engineering*, v. 47, n. 7, p. 849-856, 2000.

MADHUKAR, B.; AND MURTHY, I. S. N. "ECG Data Compression by Modeling." *Computers and Biomedical Research*, n. 26, p. 310-317, 1993.

MAMMEN, C. P.; AND RAMAMURTHI, B. "Vector Quantization for Compression of Multichannel ECG." *IEEE Transactions on Biomedical Engineering*, v. 37, n. 9, p. 821-825, 1990.

MARCELLIN, M. W.; GORMISH, M. J.; BILGIN, A.; AND BOLIEK, M. P. "An Overview of JPEG-2000." *Proceedings of the Data Compression Conference*, Snowbird, p. 523-541, March 2000.

MCAULIFFE, J. D. "Data Compression of the Exercise ECG using a Kohonen Neural Network." *Journal of Electrocardiology*, v. 26 – Supplement, 1993.

MIAOU, S.; AND YEN, H. "Quality Driven Gold Washing Adaptive Vector Quantization and its Application to ECG Data Compression." *IEEE Transactions on Biomedical Engineering*, v. 47, n. 2, p. 209-218, 2000.

MIYAHARA, M.; KOTANI, K.; E ALGAZI, V. R. *Objective Picture Quality Scale (PQS) for Image Coding*. Technical Report, Center for Image Processing and Integrated Computing, University of California, Davis, 1996.

MONRO, D. M.; AND DICKSON, G. J. "Zerotree Coding of DCT Coefficients." *Proceedings of the IEEE International Conference on Image Processing*, v. 2, p. 625-628, 1997.

MOOD, A. M.; GRAYBILL, F. A.; AND BOES, D. C. *Introduction to the Theory of Statistics*. Third edition, Tokyo: McGraw-Hill, 1975

MOODY, G. B.; AND MARK, R. G. *MIT-BIH Arrhythmia Database Directory*. Second edition, BMEC TR010 (revised), Massachusetts Institute of Technology, Biomedical Engineering Center, August 1988.

MOTTA, G.; STORER, J. A.; AND CARPENTIERI, B. "Lossless Image Coding via Adaptive Linear prediction and Classification." *Proceedings of the IEEE*, v. 88, n. 11, p. 1790-1796, 2000.

NAVE, G.; AND COHEN, A. "ECG Compression using Long Term Prediction." *IEEE Transactions on Biomedical Engineering*, v. 40, n. 9, p. 877-885, 1993.

NELSON, M.; AND GAILLY, J. L. *The Data Compression Book*. Second edition, New York: M&T Books, 1996.

NYGAARD, R.; MELNIKOV, G.; AND KATSAGGELOS, A. K. "A Rate Distortion Optimal ECG Coding Algorithm." *IEEE Transactions on Biomedical Engineering*, v. 48, n. 1, p. 28-39, 2001.

ORTEGA, A. *Optimization Techniques for Adaptive Quantization of Image and Video Under Delay Constraints*. Ph.D. thesis, Columbia University, 1994.

POPAT, A. C. *Scalar Quantization with Arithmetic Coding*. Master thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, June 1990.

PRADHAN, N.; AND DUTT, D. N. "Data Compression by Linear Prediction for Storage and Transmission of EEG Signals." *International Journal of Biomedical Computing*, n. 35, p. 207-217, 1994.

RAMAKRISHNAN, A. G.; AND SUPRATIM SAHA. "ECG Coding by Wavelet-Based Linear Prediction." *IEEE Transactions on Biomedical Engineering*, v. 44, n. 12, p.1253-1261, 1997.

RAO, K. R.; AND YIP, P. *Discrete Cosine Transform – Algorithms, Advantages, Applications*. San Diego: Academic Press, 1990.

RATNAKAR, V. *Quality-Controlled Lossy Image Compression*. Ph.D. thesis, University of Wisconsin – Madison, 1997.

RATNAKAR, V.; AND LIVNY, M. "Performance of Customized DCT Quantization Tables on Scientific Data." *Proceedings of the Science Information Management and Data Compression Workshop*, NASA Conference Publication 3277, pages 1-8, September 1994.

REININGER, R. C.; AND GIBSON, J. D. "Distribution of the Two Dimensional DCT Coefficients for Images." *IEEE Transaction on Communications*, v. COM-31, p. 835-839, 1983

RUTTIMANN, U. E.; AND PIPBERGER, H. V. "Compression of the ECG by Prediction or Interpolation and Entropy Encoding." *IEEE Transactions on Biomedical Engineering*, v. BME-26, n. 11, p. 613-623, 1979.

SAID, A.; AND PEARLMAN, W. A. "A New, Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees." *IEEE Transactions on Circuits and Systems II*, v. 6, p. 243-250, 1996.

SAYOOD, K. *Introduction to Data Compression*. San Francisco: Morgan Kaufmann, 1996.

SEROUSSI, G.; AND WEINBERGER, M. J. "On Adaptive Strategies for an Extended Family of Golomb-type Codes." *Proceedings of the Data Compression Conference*, Snowbird, p. 131-140, March, 1997.

SHANNON, C. E. "A Mathematical Theory of Communication." *Bell Syst. Tech. J.*, v. 27, p. 379-423, 1948.

SHAPIRO, J. M. "Embedded Image Coding Using Zerotrees of Wavelet Coefficients." *IEEE Transactions on Signal Processing*, v. 41, n. 12, p. 3445-3462, 1993.

SIRNE, R. O.; ISAACSON, S. I.; AND D'ATELLIS, C. E. "A Data-Reduction Process for Long-Term EEGs." *IEEE Engineering in Medicine and Biology*, January/February 1999.

STRANG, G.; AND NGUYEN, T. *Wavelets and Filter Banks*. Wellesley: Wellesley-Cambridge, 1996.

TOMPKINS, W. J. "Electrocardiography." In: *Biomedical Digital Signal Processing*, Ed. W. Tompkins, New Jersey: Prentice Hall, p. 24-54, 1993.

USEVITCH, B. E. "A Tutorial on Modern Lossy Wavelet Image Compression: Foundations of JPEG 2000." *IEEE Signal Processing Magazine*, v. 18, n. 5, p. 22-35, 2001

VAN DER POEL, J. *Compressão de Sinais de Eletrocardiograma*, Dissertação de Mestrado, Mestrado em Engenharia Biomédica, NETEB/UFPB, João Pessoa, maio, 1999.

VAN DER POEL, J.; BATISTA, L. V.; E CARVALHO, L. C. "Compressão de Sinais Eletrocardiográficos baseada em Transformada Discreta do Cosseno e Codificação de Entropia". *Revista Brasileira de Engenharia Biomédica*, v. 17, n. 1, p. 29-39, 2001.

WALLACE, G. K. "The JPEG Still Picture Compression Standard," *Communications of the ACM*, v. 34, n. 4, p. 30-44, 1991.

WEI, D. *Image Data Compression Using Wavelet Decomposition*. Master thesis, Department of Electrical and Computing Engineering, Rice University, 1995.

WEINBERGER, M. J.; SEROUSSI, G.; AND SAPIRO, G. LOCO-I: "A Low Complexity, Context-Based, Lossless Image Compression Algorithm." *Proceedings of the IEEE Data Compression Conference*, Snowbird, p. 140-149, March, 1996.

WITTEN, I. H.; NEAL, R. M.; AND CLEARY, J. G. "Arithmetic Coding for Data Compression." *Communications of the ACM*, n. 30, p. 520-540, 1987.

WOMBLE, M. E.; HALLIDAY, J. S.; MITTER, S. K.; LANCASTER, M. C.; AND TRIEBWASSER, J. H. "Data Compression for Storing and Transmitting ECG's/VCG's." *Proceedings of the IEEE*, v. 65, n. 5, p. 702-706, 1977.

WU, X. "Lossless Compression of Continuous-Tone Images via Context Selection, Quantization, and Modeling." *IEEE Transactions on Image Processing*, v. 6, n. 5, p. 656-664, 1997.

XIONG, Z.; GULERYUZ, O.; AND ORCHARD, M. T. "A DCT-based Embedded Image Coder." *IEEE Signal Processing Letters*, v. 3, n. 11, p. 289-290, 1996.

ZIGEL, Y.; AND COHEN, A. "On the Optimal Distortion Measure for ECG Compression." *Proceedings of the EMBEC*, Vienna, p. 1618-1619, November 1999.

ZIGEL, Y.; COHEN, A.; ABU-FUL, A.; WAGSHAL, A.; AND KATZ, A. "Analysis by Synthesis ECG Signal Compression." *Computers in Cardiology*, v. 24, p. 279-292, 1997.

ZOU, F.; AND GALLAGHER, R. R. "ECG Data Compression with Wavelet and Discrete Cosine Transforms", *ISA*, p. 57-62, 1994.