

CODIFICAÇÃO DE VOZ BASEADA EM WAVELETS E QUANTIZAÇÃO VETORIAL

F. MADEIRO M. S. VAJAPAYAM B. G. AGUIAR NETO M. S. ALENCAR

Laboratório de Comunicações – Departamento de Engenharia Elétrica – Universidade Federal da Paraíba
 58.109-970 - Campina Grande, PB, Brasil Phone: +55 83 3101410 Fax: +55 83 3101418
 {madeiro,madhavan,bganeto,malencar}@dee.ufpb.br

ABSTRACT

Embora a transformada *wavelet* discreta (DWT) seja amplamente utilizada em diversos sistemas de compressão de imagens, levando a excelentes resultados de codificação, pouco estudo tem sido dedicado à avaliação das potencialidades da DWT em compressão de voz. No presente trabalho, é avaliado o desempenho de um sistema de codificação de voz baseado em DWT e quantização vetorial. Testes de preferência realizados com 20 avaliadores mostram que o uso da DWT permite obter sinais reconstruídos com boa qualidade a baixas taxas de codificação. Além disso, neste trabalho, é apresentada uma investigação da importância da componente de aproximação e das componentes de detalhe, resultantes de uma decomposição *wavelet* multiresolucional, na qualidade do sinal reconstruído.

1. INTRODUÇÃO

O principal objetivo da compressão de sinais de voz é reduzir o número de bits necessários para se representar adequadamente o sinal, visando minimizar os requisitos de largura de faixa e de capacidade de armazenamento. Apesar de alguns sistemas não apresentarem grandes limitações de largura de faixa, como é o caso das redes de comunicações por fibra ótica, e embora a evolução tecnológica esteja continuamente contribuindo para o surgimento de memórias com grandes capacidades de armazenamento, a compressão de sinais de voz desempenha um papel importante, devido a uma série de fatores, como por exemplo [1], [2]:

- nas redes digitais de serviços integrados (ISDN, *integrated services digital networks*), as técnicas de compressão permitem uma integração eficiente de sinais e dados;
- a enorme utilização dos sistemas multimídia tem levado ao aumento da demanda no tocante ao armazenamento de voz, música, imagens, vídeo e dados em forma comprimida;
- em telefonia móvel celular, a largura de faixa é severamente limitada, o que tem motivado muitos estudos em compressão de voz;
- nos sistemas de reconhecimento de fala e de locutor e nos sistemas de resposta vocal, vocabulários maiores podem ser armazenados por meio da redução dos requisitos de memória necessários para cada padrão de voz.

Motivado pelos excelentes resultados de compressão apresentados pelos sistemas de codificação de imagens que utili-

zam a combinação da transformada *wavelet* discreta (DWT, *discrete wavelet transform*) e quantização vetorial (QV), este trabalho tem como objetivo avaliar o desempenho de um sistema de codificação de forma de onda de voz baseado em DWT e QV. Neste contexto, é utilizado um conjunto de dicionários multiresolucionais destinados à codificação dos coeficientes *wavelet* resultantes da aplicação de uma decomposição *wavelet* multiresolucional no sinal de voz. Vale salientar que, muito embora alguns trabalhos tenham se dedicado à utilização de *wavelets* no contexto de *codificação paramétrica de voz* (coeficientes LPC), não se tem conhecimento de trabalhos referentes à aplicação de *wavelets* no contexto de *codificação de forma de onda de voz* baseada em DWT e QV.

No presente trabalho, além da avaliação da qualidade do sinal de voz em diversas taxas de codificação, é apresentada uma investigação da importância (contribuição) das componentes de aproximação e de detalhe, resultantes de uma decomposição *wavelet* multiresolucional, na qualidade do sinal reconstruído.

Este trabalho encontra-se organizado de acordo com as Seções a seguir. Uma visão geral da quantização vetorial (QV) é apresentada na Seção 2. A transformada *wavelet* discreta (DWT) é brevemente abordada na Seção 3. Na Seção 4 é descrito o sistema de codificação de voz que utiliza DWT e QV. Os resultados e a conclusão do trabalho são apresentados, respectivamente, nas Seções 5 e 6.

2. QUANTIZAÇÃO VETORIAL

A quantização vetorial [2], [3], que pode ser vista como uma extensão da quantização escalar em um espaço multidimensional, encontra-se fundamentada na Teoria da Distorção Versus Taxa, formulada por Shannon, segundo a qual um melhor desempenho é obtido codificando-se blocos de amostras (isto é, vetores) ao invés de amostras individuais (isto é, escalares). Em diversos sistemas de codificação de voz e imagem, a quantização vetorial tem sido utilizada com sucesso, permitindo elevadas taxas de compressão.

Matematicamente, a quantização vetorial consiste de um mapeamento Q de um vetor de entrada \vec{x} pertencente ao espaço Euclidiano K -dimensional, R^K , em um vetor pertencente a um subconjunto finito W de R^K , ou seja,

$$Q : R^K \rightarrow W. \quad (1)$$

O dicionário $W = \{\vec{w}_i; i = 1, 2, \dots, N\}$ é o conjunto de vetores de reprodução (também denominados vetores-código ou vetores de reconstrução), K é a dimensão do quantizador vetorial e N é o número de vetores-código (ou número

de níveis, em analogia com a quantização escalar). A taxa de codificação do quantizador vetorial, que representa o número de bits por componente do vetor, é dada por $R = \frac{1}{K} \log_2 N$. Em codificação de forma de onda de voz, R é expressa em *bit/amostra*. Em se tratando de codificação de imagens, R é expressa *bits por pixel (bpp)*.

O mapeamento Q introduz um particionamento de R^K em N células (denominadas regiões de Voronoi) S_i , $i = 1, 2, \dots, N$, tais que:

$$\bigcup_{i=1}^N S_i = R^K \text{ e } S_i \cap S_j = \emptyset \text{ para } i \neq j. \quad (2)$$

Cada célula S_i é definida da seguinte maneira:

$$S_i = \{\vec{x} : Q(\vec{x}) = \vec{w}_i\}. \quad (3)$$

Deste modo, o vetor-código \vec{w}_i constitui o vetor representativo de todos os vetores de entrada pertencentes à célula S_i . Como a quantização vetorial realiza um mapeamento de padrões de entrada (vetores de entrada \vec{x}) semelhantes em padrões de saída (vetores-código \vec{w}_i) semelhantes, ela pode ser vista como uma forma de reconhecimento de padrões, em que um padrão de entrada é “aproximado” por um padrão de referência, pertencente a um conjunto predefinido (dicionário) de padrões (vetores de reconstrução) de referência [2], [4].

O objetivo das técnicas de projeto de quantizadores vetoriais é obter um dicionário ótimo, que minimize a distorção média introduzida pela aproximação do conjunto de vetores de entrada (vetores a serem quantizados) pelos correspondentes vetores de reprodução (vetores-código).

3. TRANSFORMADA WAVELET DISCRETA

A transformada *wavelet* discreta (DWT) pode ser descrita a partir do banco de filtros ilustrados na Fig. 1. Os filtros $h_0(n)$ e $h_1(n)$ correspondem aos filtros *wavelet* de análise, enquanto que $g_0(n)$ e $g_1(n)$ correspondem aos filtros *wavelet* de síntese. Um sinal de entrada $x(n)$ é convoluído com o filtro passa-baixa $h_0(n)$ e com o filtro passa-alta $h_1(n)$. O sinal resultante de cada convolução é então submetido a uma decimação (sub-amostragem) de ordem 2. Geram-se, assim, no processo de decomposição (ou análise), o sinal de aproximação $a_1(n)$ e o sinal de detalhe $d_1(n)$. Em outras palavras, $a_1(n)$ contém os coeficientes *wavelet* correspondentes à componente de aproximação do sinal, ao passo que $d_1(n)$ contém os coeficientes *wavelet* referentes à componente de detalhe do sinal. No processo de reconstrução (ou síntese), procede-se uma interpolação (super-amostragem) de ordem 2, seguida de convolução com os filtros passa-baixa $g_0(n)$ e passa-alta $g_1(n)$. Satisfeitas as condições de reconstrução perfeita, o sinal $x(n)$ é reconstruído de forma exata, como a soma das saídas dos filtros *wavelet* de síntese, ou seja, $y(n) = x(n)$.

Em se tratando de uma decomposição *wavelet* multiresolucional (multinível) [5], [6], [7], o próximo nível de decomposição (resolução mais grosseira) é obtido mediante a convolução do sinal $a_1(n)$ com um par idêntico de filtros $h_0(n)$ e $h_1(n)$, seguida da decimação de ordem 2. Geram-se, assim, os sinais $a_2(n)$ e $d_2(n)$, em que o índice 2 diz

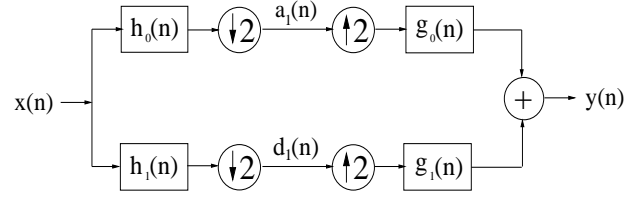


Fig. 1 - Estrutura básica do banco de filtros e das operações de decimação e interpolação envolvidos na DWT.

respeito ao segundo nível de decomposição. Esse processo pode ser repetido iterativamente, e o número de pares de filtros *wavelet* de análise define o número de níveis de decomposição (ou número de níveis de resolução), denotado por L , da transformada *wavelet* discreta. Deste modo, são obtidos, ao final de L níveis de decomposição, um sinal de aproximação $a_L(n)$, com resolução reduzida por um fator de 2^L em relação ao sinal de entrada, e os sinais de detalhe $d_L(n)$, $d_{L-1}(n)$, \dots , $d_1(n)$. Em outras palavras, $d_i(n)$, com $1 \leq i \leq L$, contém os coeficientes *wavelet* correspondentes às componentes de detalhe no i -ésimo nível de decomposição, enquanto que $a_L(n)$ contém os coeficientes *wavelet* referentes à componente de aproximação no L -ésimo nível de decomposição. A Fig. 2(a) ilustra a estrutura piramidal de uma decomposição *wavelet* de 3 níveis.

Para se realizar a etapa de síntese, ilustrada na Fig. 2(b), os sinais $d_i(n)$ e $a_i(n)$, após interpolação de ordem 2, são submetidos aos filtros *wavelet* de síntese $g_0(n)$ e $g_1(n)$ para permitirem a reconstrução de $a_{i-1}(n)$, que constitui a componente de aproximação no próximo nível de resolução mais alta. Uma vez asseguradas as condições de reconstrução perfeita, $y(n) = x(n)$. O conjunto de operações (convolução e interpolação) que permitem a reconstrução do sinal a partir dos coeficientes *wavelet* $a_L(n)$ e $d_i(n)$, com $1 \leq i \leq L$, constitui a transformada *wavelet* discreta inversa (IDWT, *inverse discrete wavelet transform*).

4. SISTEMA DE CODIFICAÇÃO DE VOZ BASEADO EM WAVELETS E QUANTIZAÇÃO VETORIAL

Uma visão geral do sistema de codificação de voz baseado em *wavelets* e quantização vetorial é apresentado na Fig. 3. Após aplicação da transformada *wavelet* discreta (DWT), ou seja, de uma decomposição *wavelet* multiresolucional (L níveis de decomposição), as componentes de detalhe, $d_i(n)$, com $1 \leq i \leq L$, e a componente de aproximação, $a_L(n)$, são quantizadas mediante utilização de um banco de dicionários multiresolucionais. Em outras palavras, em cada nível de resolução, os coeficientes *wavelet* são quantizados por meio de um dicionário específico. O sinal reconstruído é obtido aplicando-se a transformada *wavelet* discreta inversa (IDWT) nas componentes de aproximação e de detalhe reconstruídas (quantizadas).

Considere uma decomposição *wavelet* multiresolucional de um sinal $x(n)$, constituído de M amostras. O número de coeficientes *wavelet* em $d_i(n)$, com $1 \leq i \leq L$, é igual a $\frac{M}{2^i}$, e o número de coeficientes *wavelet* em $a_L(n)$ é igual a

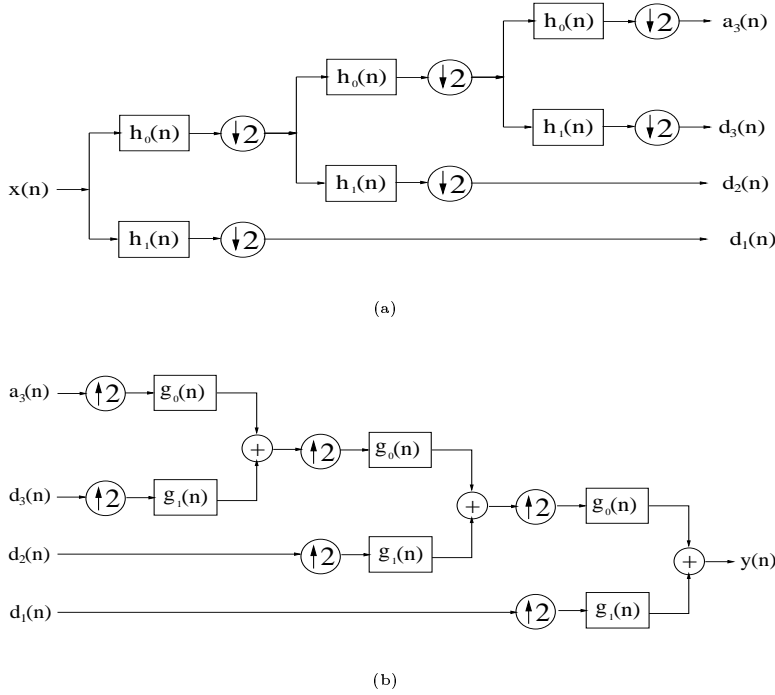


Fig. 2 - Estrutura piramidal de uma decomposição *wavelet* multiresolucional em 3 níveis: (a) operação de análise ou decomposição; (b) operação de síntese ou reconstrução.

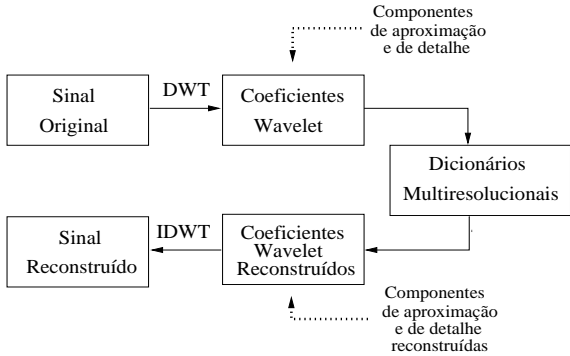


Fig. 3 - Diagrama de blocos do sistema de codificação de voz baseado em *wavelets* e quantização vetorial.

$\frac{M}{2^L}$. Seja R_{d_i} a taxa (em *bit/amostra* ou, mais precisamente, em *bit/coeficiente*) do quantizador vetorial alocada para o sinal $d_i(n)$ e seja R_{a_L} a taxa (em *bit/amostra*) do quantizador vetorial alocada para o sinal $a_L(n)$. A taxa de codificação total R_T (em *bit/amostra*) do sistema de codificação de voz utilizando DWT, com L níveis de decomposição, e QV é dada por

$$R_T = \frac{1}{M} \left(\frac{M}{2^L} \cdot R_{a_L} + \sum_{i=1}^L \frac{M}{2^i} \cdot R_{d_i} \right), \quad (4)$$

isto é,

$$R_T = \frac{R_{a_L}}{2^L} + \sum_{i=1}^L \frac{R_{d_i}}{2^i}. \quad (5)$$

A Tabela I exemplifica um esquema de alocação de taxa (em *bit/amostra*) para as componentes $a_3(n)$, $d_3(n)$, $d_2(n)$ e $d_1(n)$, em um sistema de codificação de voz usando DWT (com 3 níveis de decomposição) e QV. Observe-se que os valores de número de níveis N e dimensão K do quantizador vetorial definem a taxa para cada componente. De acordo com a Equação (5), a taxa de codificação total para a estratégia de alocação em questão é

$$R_T = \frac{R_{a_3}}{8} + \frac{R_{d_1}}{2} + \frac{R_{d_2}}{4} + \frac{R_{d_3}}{8}, \quad (6)$$

ou seja,

$$R_T = \frac{2}{8} + \frac{0,75}{2} + \frac{1}{4} + \frac{1,5}{8} = 1,0625 \text{ bit/amostra}. \quad (7)$$

5. RESULTADOS

Os sinais de voz, originalmente codificados a 8,0 *bit/amostra*, utilizados nas simulações correspondem a sentenças extraídas de um conjunto de frases foneticamente balanceadas da língua portuguesa [8]. Os dicionários multiresolucionais utilizados neste trabalho foram projetados com o uso do algoritmo LBG [9]. Utilizou-se, na DWT, a família Daubechies 6 [6]. Os resultados apresentados nesta

TABELA I

UM EXEMPLO DE ALOCAÇÃO DE TAXA PARA AS COMPONENTES $a_3(n)$, $d_3(n)$, $d_2(n)$ E $d_1(n)$. OS VALORES DE N E K DO QUANTIZADOR VETORIAL DEFINEM A TAXA PARA CADA COMPONENTE.

Componente	N	K	taxa (<i>bit/amostra</i>)
$a_3(n)$	256	4	$R_{a_3} = 2,0$
$d_3(n)$	64	4	$R_{d_3} = 1,5$
$d_2(n)$	256	8	$R_{d_2} = 1,0$
$d_1(n)$	64	8	$R_{d_1} = 0,75$

seção foram obtidos com o uso de uma decomposição *wavelet* multiresolucional de 3 níveis. A qualidade dos sinais reconstruídos foi avaliada utilizando-se a relação sinal-ruído segmental (SNR_{seg}) [1], [10].

O primeiro conjunto de simulações consistiu em avaliar a importância das componentes de aproximação e de detalhe, isto é, dos sinais $a_3(n)$, $d_3(n)$, $d_2(n)$ e $d_1(n)$, na qualidade do sinal reconstruído. Para tanto, é avaliada a qualidade do sinal obtido mediante aplicação da IDWT com a exclusão de uma ou mais componentes. A Tabela II mostra que $a_3(n)$ corresponde à componente mais importante para a qualidade do sinal reconstruído. De fato, o menor valor de SNR_{seg} da Tabela II corresponde ao sinal reconstruído após a exclusão de $a_3(n)$. Por outro lado, $d_1(n)$ corresponde ao conjunto de coeficientes *wavelet* menos importante para a qualidade do sinal reconstruído – observa-se, na Tabela II, que o maior valor de SNR_{seg} diz respeito à exclusão de $d_1(n)$. É possível observar que, dentre as componentes de detalhe, $d_3(n)$ constitui a mais importante para a qualidade do sinal reconstruído.

Vale salientar que os testes subjetivos realizados com diversos sinais reconstruídos indicaram que a exclusão de $d_1(n)$ implica queda praticamente imperceptível de qualidade do sinal reconstruído. É importante ressaltar que $d_1(n)$ contém 50% dos coeficientes *wavelet* obtidos com a DWT, de modo que a taxa de codificação de $d_1(n)$ desempenha um papel importante na definição da taxa de codificação total do sistema de compressão de voz em consideração, conforme mostra a Equação (5).

Observou-se, durante sessões de avaliação subjetiva, que a exclusão da componente $d_2(n)$, bem como $d_3(n)$, apesar de não comprometer a inteligibilidade do sinal reconstruído, confere um certo caráter metálico à voz. No que diz respeito à exclusão de $a_3(n)$, o caráter metálico presente no sinal reconstruído torna-se mais perceptível; a inteligibilidade, por sua vez, continua alta.

O segundo conjunto de experimentos consistiu em avaliar diferentes esquemas de alocação de bits para os sinais (componentes de aproximação e de detalhe) resultantes da decomposição *wavelet* multiresolucional. Para tanto, foram projetados diversos dicionários multiresolucionais, variando-se a dimensão K e o número de níveis N dos quantizadores vetoriais. A Tabela III apresenta algumas estratégias de alocação avaliadas e os respectivos valores de SNR_{seg} obtidos para os sinais reconstruídos. Conforme mencionado anteriormente, a exclusão de $d_1(n)$ implica queda pratica-

TABELA II

IMPORTÂNCIA DAS COMPONENTES DE APROXIMAÇÃO E DE DETALHE (DE UMA DECOMPOSIÇÃO WAVELET MULTIRESOLUCIONAL DE 3 NÍVEIS) NA QUALIDADE DO SINAL RECONSTRUÍDO.

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

TABELA III

ESTRATÉGIAS DE ALOCAÇÃO DE BITS OBTIDAS COM A VARIAÇÃO DA DIMENSÃO K E DO NÚMERO DE NÍVEIS N DOS DICIONÁRIOS MULTIRESOLUCIONAIS. A TAXA R CORRESPONDENTE É EXPRESSA EM *bit/amostra*. A COMPONENTE $d_1(n)$ FOI CODIFICADA A 0,0 *bit/amostra*. SÃO APRESENTADOS OS VALORES DE SNR_{seg} DO SINAL RECONSTRUÍDO PARA DIVERSAS TAXAS DE CODIFICAÇÃO TOTAL, R_T (EM *bit/amostra*).

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

mente imperceptível na qualidade subjetiva do sinal reconstruído. Deste modo, foi adotada uma taxa de codificação de 0,0 *bit/amostra* para a componente $d_1(n)$, ou seja, esta componente foi excluída quando da realização da IDWT.

A Tabela IV apresenta valores de SNR_{seg} para algumas taxas de um sistema de codificação utilizando quantização vetorial simples, ou seja, no domínio original (própria forma de onda, sem uso de DWT) do sinal de voz. Comparando-se as Tabelas III e IV, observa-se que, apenas para a taxa de 1,0 *bit/amostra*, a codificação de voz utilizando QV simples levou a um resultado de SNR_{seg} superior ao obtido com o uso de DWT + QV (transformada *wavelet* discreta seguida de quantização vetorial). A realização de testes de escuta, entretanto, indicou que, para esta taxa, o sinal

reconstruído mediante uso de DWT + QV apresenta qualidade subjetiva superior à apresentada pelo sinal obtido com o uso de QV simples. As sessões de avaliação de qualidade subjetiva também apontaram uma superioridade da codificação DWT + QV sobre QV simples para as taxas de 2,0 *bit/amostra* e 0,875 *bit/amostra*.

TABELA IV
DESEMPENHO DA QV SIMPLES (SEM USO DE DWT). VALORES DE SNR_{seg} DO SINAL RECONSTRUÍDO PARA ALGUMAS TAXAS DE CODIFICAÇÃO, $R = \frac{1}{K} \log_2 N$.

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

No que diz respeito à avaliação de desempenho fora da seqüência de treino, a Tabela V apresenta valores de SNR_{seg} para os sistemas de codificação usando DWT + QV e QV simples. Observa-se que, para todas as taxas de codificação consideradas, a QV simples leva a maiores valores de SNR_{seg} . Entretanto, os testes subjetivos indicaram que, para todas estas taxas, os sinais reconstruídos com uso de DWT + QV apresentam qualidade superior à apresentada pelos sinais reconstruídos mediante utilização de QV simples. De fato, a Tabela VI mostra o resultado de testes de preferência [11] realizados com 20 avaliadores: para a taxa de 2,0 *bit/amostra*, 95% dos avaliadores consideraram a qualidade do sinal obtido com DWT + QV superior à qualidade do sinal obtido com QV simples; para a taxa de 1,0 *bit/amostra*, 85% dos avaliadores manifestaram sua preferência pelo sinal obtido com uso de DWT + QV; para a taxa de 0,875 *bit/amostra*, 65% dos avaliadores apontaram a melhor qualidade do sinal obtido com uso de DWT + QV, 10% registraram preferência pelo sinal obtido com QV simples e 25% registraram inexistência de distinção de qualidade.

Neste contexto, portanto, convém mencionar que, apesar de SNR_{seg} ser muito utilizada para avaliar o desempenho de codificadores de forma de onda de voz, esta medida não se apresenta como uma alternativa adequada para análise comparativa de desempenho entre a codificação que utiliza DWT + QV (em que a quantização é levada a efeito no domínio da transformada *wavelet* discreta, ou seja, quantização dos coeficientes *wavelet*) e a codificação que usa QV simples (em que a quantização é realizada diretamente na forma de onda, ou seja, nas próprias amostras do sinal).

6. CONCLUSÃO

Motivado pelos excelentes resultados de compressão obtidos com a aplicação da transformada *wavelet* discreta (DWT) em codificação de imagens, e em virtude de não se ter conhecimento de estudos referentes à utilização da DWT em codificação de forma de onda de voz, este trabalho se propôs a avaliar o desempenho de um sistema de codificação de voz baseado em DWT e quantização vetorial (QV). Para

TABELA V
DESEMPENHO FORA DA SEQÜÊNCIA DE TREINO. VALORES DE SNR_{seg} (EM DECIBÉIS) DO SINAL RECONSTRUÍDO PARA ALGUMAS TAXAS DE CODIFICAÇÃO TOTAL, R_T , PARA OS SISTEMAS UTILIZANDO DWT + QV E QV SIMPLES.

R_T (<i>bit/amostra</i>)	DWT + QV	QV simples
2,0	13,20	13,47
1,0	6,72	9,42
0,875	6,47	8,54

TABELA VI
DESEMPENHO FORA DA SEQÜÊNCIA DE TREINO. RESULTADOS, EM TERMOS DE PERCENTAGENS, DO TESTE DE PREFERÊNCIA REALIZADO COM 20 AVALIADORES. ESCALA DE PONTUAÇÃO: A – A QUALIDADE DO SINAL OBTIDO COM DWT + QV É SUPERIOR À QUALIDADE DO SINAL OBTIDO COM QV SIMPLES; B – A QUALIDADE DO SINAL OBTIDO COM QV SIMPLES É SUPERIOR À OBTIDA COM DWT + QV; C – NÃO SE DISTINGUE A QUALIDADE DOS SINAIS OBTIDOS COM DWT + QV E QV SIMPLES.

R (<i>bit/amostra</i>)	A	B	C
2,0	95%	5%	0%
1,0	85%	15%	0%
0,875	65%	10%	25%

tanto, investigou-se, inicialmente, a importância das componentes de aproximação e de detalhe, resultantes da decomposição *wavelet* multiresolucional, na qualidade do sinal reconstruído. Em seguida, foram avaliados diferentes esquemas de alocação de bits para as componentes resultantes da DWT. Testes de preferência realizados com 20 avaliadores indicaram que a qualidade do sinal reconstruído mediante utilização de DWT + QV (quantização vetorial no domínio da DWT, ou seja, quantização dos coeficientes *wavelet*) é superior à qualidade do sinal obtido com uso de QV simples (quantização vetorial na própria forma de onda do sinal, ou seja, quantização das amostras do sinal).

7. REFERÊNCIAS

- [1] Jayant, N. S. and Noll, P. *Digital Coding of Waveforms*. Prentice-Hall, Englewood Cliffs, NJ, 1984.
- [2] Gersho, A. and Gray, R. M. *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, Boston, MA, 1992.
- [3] Gray, R. M. Vector Quantization. *IEEE ASSP Magazine*, pages 4–29, April 1984.
- [4] Kosko, B. *Neural Networks and Fuzzy Systems*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1992.
- [5] Mallat, S. G. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, pages 674–696, July 1989.
- [6] Burrus, C. S., Gopinath, R. A. and Guo, H. *Introduction to Wavelets and Wavelet Transforms*. Prentice-Hall, New Jersey, 1998.

- [7] Strang, G. and Nguyen, T. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, 1996.
- [8] Alcaim, A., Solewicz, J. A. and Moraes, J. A. Frequência de Ocorrência dos Fones e Listas de Frases Foneticamente Balanceadas no Português Falado no Rio de Janeiro. *Revista da Sociedade Brasileira de Telecomunicações*, Vol. 7, No. 1, pages 23–41, December 1992.
- [9] Linde, Y., Buzo, A. and Gray, R. M. An Algorithm for Vector Quantizer Design. *IEEE Transactions on Communications*, Vol. COM-28, No. 1, pages 84–95, January 1980.
- [10] Deller Jr., J. R., Proakis, J. G. and Hansen, J. H. L. *Discrete-time Processing of Speech Signals*. Macmillan Publishing Co., 1993.
- [11] Aguiar Neto, B. G. *Processamento e Transmissão Digital de Voz (Apostila)*. Curso de Pós Graduação em Engenharia Elétrica, Campina Grande - PB, 1995.