

Construção de Dicionários Voltados para a Redução da Complexidade Computacional da Etapa de Codificação da Quantização Vetorial

Francisco Madeiro¹, Waslon T. A. Lopes², Benedito G. Aguiar Neto², Marcelo S. Alencar²

¹Universidade Católica de Pernambuco – Recife, PE, Brasil

²Universidade Federal de Campina Grande – Campina Grande, PB, Brasil*
madeiro@dei.unicap.br, {waslon,bganeto,malencar}@dee.ufcg.edu.br

Abstract

This work presents a methodology for designing codebooks for the purpose of reducing the computational complexity of the vector quantization (VQ) encoding phase. The methodology consists on efficiently exploring a symmetry observed in some signals. The proposed methodology is evaluated by simulations involving voice waveform coding based on VQ.

1. Introdução

A quantização vetorial [1, 2], que pode ser vista como uma extensão da quantização escalar em um espaço multidimensional, encontra-se fundamentada na Teoria da Distorção Versus Taxa [3], 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 outras palavras, essa teoria ressalta a superioridade da quantização vetorial (QV) sobre a quantização escalar [4]. Matematicamente, a quantização vetorial pode ser definida como um mapeamento Q de um vetor de entrada \mathbf{x} pertencente ao espaço euclidiano K -dimensional, \mathbb{R}^K , em um vetor pertencente a um subconjunto finito W de \mathbb{R}^K , ou seja,

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

O dicionário $W = \{\mathbf{w}_i; i = 1, 2, \dots, N\}$ 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 tamanho do dicionário, isto é, o número de vetores-código (ou número de níveis, em analogia com a quantização escalar).

Em um sistema de compressão de sinais baseado em quantização vetorial, apresentado na Figura 1, um quantizador vetorial pode ser visto como a combinação de duas funções: um codificador de fonte e um decodificador de fonte. Dado um vetor $\mathbf{x} \in \mathbb{R}^K$, da fonte a ser codificada, o codificador calcula a distorção $d(\mathbf{x}, \mathbf{w}_i)$ entre o vetor de entrada (vetor a ser quantizado) e cada vetor-código \mathbf{w}_i , $i = 1, 2, \dots, N$ do dicionário W . A regra ótima para codificação é a regra do vizinho mais próximo [5],

na qual uma representação binária do índice I , denotada por \mathbf{b}_I , é transmitida ao decodificador de fonte se o vetor-código \mathbf{w}_I corresponder à menor distorção [6], isto é, se \mathbf{w}_I for o vetor-código que apresentar a maior similaridade com \mathbf{x} dentre todos os vetores-código do dicionário. Em outras palavras, o codificador usa a regra de codificação $\mathcal{C}(\mathbf{x}) = \mathbf{b}_I$ se $d(\mathbf{x}, \mathbf{w}_I) < d(\mathbf{x}, \mathbf{w}_i), \forall i \neq I$. Ao receber a representação binária $\mathbf{b}_I = (b_1(I), b_2(I), \dots, b_m(I))$ do índice I , o decodificador de fonte, que dispõe de uma cópia do dicionário W , simplesmente procura pelo I -ésimo vetor-código e produz o vetor \mathbf{w}_I como a reprodução (versão quantizada) de \mathbf{x} . Em outras palavras, é utilizada a seguinte regra de decodificação: $\mathcal{D}(\mathbf{b}_I) = \mathbf{w}_I$. A representação do vetor de entrada \mathbf{x} pelo vetor-código \mathbf{w}_I é consubstanciada na expressão $\mathbf{w}_I = Q(\mathbf{x})$.

A quantização vetorial, portanto, constitui uma técnica de compressão com perdas, visto que o sinal reconstruído é uma versão degradada do sinal original. A distorção introduzida na representação dos vetores da fonte pelos correspondentes vetores-código depende da qualidade do dicionário W . A taxa de codificação do quantizador vetorial, que mede o número de bits por componente do vetor, é dada por $R = \frac{m}{K} = \frac{1}{K} \log_2 N$. Observe que estão sendo consideradas palavras-binárias \mathbf{b}_i de comprimento fixo m bits. Em codificação de forma de onda de voz (e. g. [7, 8]), R é expressa em bit/amostra. Em se tratando de codificação de imagens (e. g. [6, 9]), R é expressa em bits por *pixel* (bpp).

No presente trabalho é apresentada uma nova metodologia para a construção de dicionários, visando a redução da complexidade computacional envolvida na etapa de codificação da QV. O método proposto introduz uma organização estruturada nos dicionários projetados, com o objetivo de reduzir o número de multiplicações, adições, subtrações e comparações realizadas na etapa de codificação por distorção mínima (etapa de determinação do vizinho mais próximo). É apresentado um método eficiente de codificação, que explora a organização estruturada dos dicionários projetados. A metodologia proposta consiste em explorar eficientemente uma simetria observada em alguns sinais.

O restante do artigo encontra-se organizado de acordo com as seções a seguir. Na Seção 2 é apresentada uma breve abordagem da complexidade computacional da QV. A simetria apresentada nos sinais de voz, explora-

* Os autores gostariam de expressar os agradecimentos à CAPES e ao CNPq pelo apoio financeiro ao trabalho.

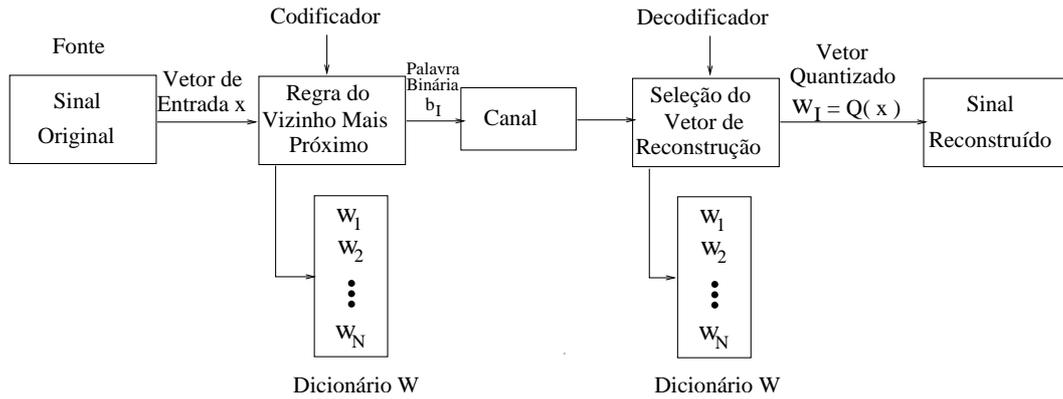


Figura 1: Codificação/decodificação em um sistema de codificação baseado em quantização vetorial.

da no presente trabalho, é discutida na Seção 3. A Seção 4 descreve a metodologia proposta para o projeto de dicionários. A Seção 5 descreve o método para redução do número de operações realizadas na etapa de codificação da QV. Resultados e comentários finais são apresentados, respectivamente, nas Seções 6 e 7.

2. Complexidade Computacional da QV

O problema da determinação do vizinho mais próximo desempenha papel importante em diversas aplicações, como por exemplo reconhecimento de padrões e compressão de sinais envolvendo quantização vetorial. Em virtude da necessidade de comparação direta de um padrão de teste com cada padrão de referência armazenado, a regra do vizinho mais próximo implica elevada complexidade computacional, o que tem despertado o interesse de diversas pesquisas relacionadas à análise e à concepção de métodos de procura eficientes.

Em quantização vetorial, em particular, a complexidade computacional existente na etapa de codificação dos vetores a serem quantizados constitui um problema crítico.

Para codificar um dado vetor de entrada, é necessário encontrar sua distância para cada um dos N vetores-código do dicionário e depois comparar as distâncias de modo a encontrar o vetor-código mais semelhante, ou seja, o vizinho mais próximo. No método convencional de busca total (BT), a codificação de cada vetor de entrada requer, portanto, N cálculos de distância (distorção) e $N - 1$ comparações. Ao ser utilizada como medida de distorção o erro quadrático, isto é,

$$d(\mathbf{x}, \mathbf{w}_i) = \sum_{j=1}^K (x_j - w_{ij})^2, \quad (2)$$

em que w_{ij} é a j -ésima componente do vetor-código \mathbf{w}_i e x_j é a j -ésima componente do vetor de entrada \mathbf{x} , cada cálculo de distância requer K multiplicações, K subtrações e $K - 1$ adições. Portanto, para codificar cada vetor de entrada, KN multiplicações, KN subtrações, $(K - 1)N$ adições e $N - 1$ comparações devem ser computadas. A complexidade de um quantizador

vetorial pode ser alternativamente expressa em termos de N multiplicações, N subtrações, $(1 - \frac{1}{K})N$ adições e $(N - 1)/K$ comparações por amostra.

Em resumo, a complexidade computacional de um quantizador vetorial de dimensão K e taxa R requer um número de operações por amostra da ordem de $N = 2^{KR}$ para cada vetor de entrada se uma busca completa (busca total, busca exaustiva) é realizada ao longo do dicionário.

A seção a seguir aborda a simetria dos sinais de voz. Ressalta-se, entretanto, que essa forma de simetria também pode ser observada em outros sinais.

3. A Simetria dos Sinais de Voz

A aquisição (resolução 8,0 bit/amostra e taxa de amostragem 8 kHz) dos sinais de voz utilizados neste trabalho foi realizada usando uma estação de trabalho Sun®, instalada com utilitários de processamento de áudio.

A Figura 2 apresenta o histograma de um sinal de voz correspondente a 10 frases foneticamente balanceadas, extraídas de [10], e pronunciadas por 10 locutores diferentes, sendo 5 masculinos e 5 femininos. A Figura 3 apresenta a representação do sinal no espaço de padrões euclidiano bidimensional. Conforme se pode observar, um sinal de voz apresenta as seguintes características: predominância de vetores com componentes de baixa amplitude; correlação entre amostras consecutivas (conforme se observa na direção correspondente à componente principal do sinal). Os sinais de voz apresentam uma simetria interessante: existe uma espécie de correspondência entre os vetores de voz, no sentido de que a um determinado vetor \mathbf{x} de voz em geral está associado um vetor aproximadamente igual ao correspondente simétrico $-\mathbf{x}$. Conforme mostra a Tabela 1, nos sinais de voz a simetria também se apresenta no fato de aproximadamente metade dos vetores terem média¹ positiva e metade terem média negativa.

¹Ao longo deste trabalho, a média de um vetor deve ser entendida como a média aritmética de suas componentes.

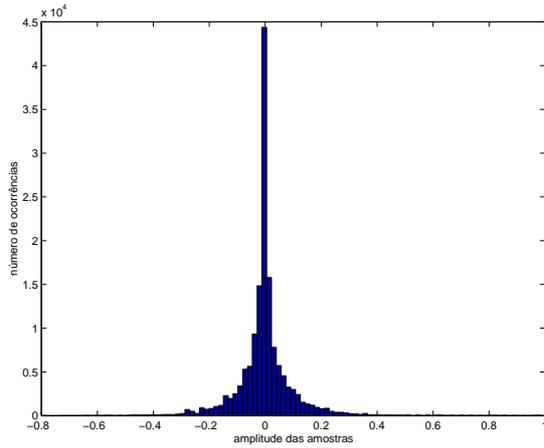


Figura 2: Histograma de um sinal de voz.

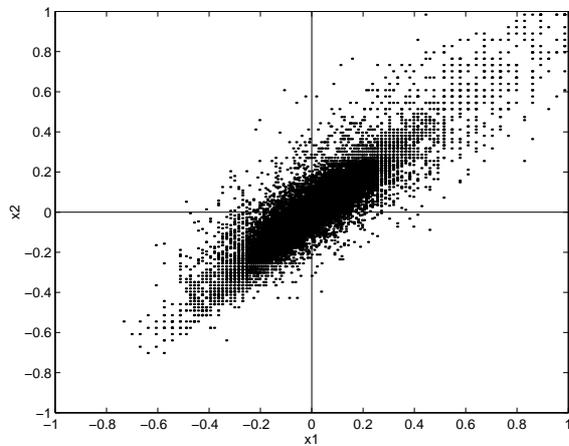


Figura 3: Sinal de voz consistindo de 10 frases foneticamente balanceadas (18,76s, 75040 vetores). As coordenadas x_1 e x_2 representam a primeira e a segunda componentes dos vetores de treino $\mathbf{x} \in \mathbb{R}^2$, respectivamente.

4. O Projeto dos Dicionários

A metodologia proposta neste trabalho para o projeto de dicionários procura fazer com que os dicionários projetados *reflitam* a simetria existente nos sinais de voz. O conjunto S dos vetores de treino K -dimensionais é dividido em dois subconjuntos, S_{POS} e S_{NEG} , em que S_{POS} é constituído pelos vetores de treino que apresentam média positiva e S_{NEG} é constituído pelos vetores de treino que apresentam média negativa.

O subconjunto S_{POS} é utilizado para obter os primeiros $N/2$ vetores-código, por meio de algum algoritmo de treinamento de dicionários, como por exemplo, algoritmo LBG (Linde-Buzo-Gray) [5], algoritmos de aprendizagem não-supervisionada [11–13] e algoritmos *fuzzy* [14].

Os primeiros $N/2$ vetores-código do dicionário, isto é, os vetores \mathbf{w}_i , $1 \leq i \leq N/2$, portanto, têm componentes cujo valor médio é positivo. Esses vetores serão denotados a partir de agora por $\mathbf{w}_{i,\text{POS}}$, com $1 \leq i \leq N/2$, em que o subscrito POS é uma alusão ao valor médio positivo.

Tabela 1: Percentagem dos vetores K -dimensionais com média positiva e com média negativa.

K	Média positiva	Média negativa
2	47,32%	52,68%
3	47,25%	52,65%
4	47,19%	52,81%
5	47,23%	52,77%
6	47,27%	52,73%
7	47,43%	52,57%
8	47,88%	52,12%

Tabela 2: Dicionário com 8 vetores-código bidimensionais $\mathbf{w}_i = [w_{i1} \ w_{i2}]^T$, com $1 \leq i \leq 8$. A palavra-binária do i -ésimo vetor-código, \mathbf{w}_i , é denotada por \mathbf{b}_i , enquanto que w_{ij} representa a j -ésima componente do vetor \mathbf{w}_i .

i	w_{i1}	w_{i2}	\mathbf{b}_i
1	0,1271	0,1288	000
2	0,0628	0,0621	001
3	0,2560	0,2536	010
4	0,0103	0,0105	011
5	-0,0103	-0,0105	100
6	-0,2560	-0,2536	101
7	-0,0628	-0,0621	110
8	-0,1271	-0,1288	111

Os demais $N/2$ vetores-código do dicionários são obtidos como descrito a seguir. Para cada vetor-código $\mathbf{w}_{i,\text{POS}}$ é obtido um correspondente vetor-código $\mathbf{w}_{N+1-i,\text{NEG}}$ tal que

$$\mathbf{w}_{i,\text{POS}} = -\mathbf{w}_{N+1-i,\text{NEG}}, 1 \leq i \leq N/2. \quad (3)$$

Os últimos $N/2$ vetores-código do dicionário, portanto, têm componentes cujo valor médio é negativo. Esses vetores serão denotados a partir de agora por $\mathbf{w}_{i,\text{NEG}}$, com $N/2 + 1 \leq i \leq N$, em que o subscrito NEG é uma alusão ao valor médio negativo.

Portanto, em decorrência da Equação (3), existe uma simetria marcante nos dicionários obtidos com a metodologia descrita neste trabalho: associado a um vetor-código \mathbf{w}_i , $1 \leq i \leq N/2$, existe um vetor-código \mathbf{w}_{N+1-i} tal que $\mathbf{w}_{N+1-i} = -\mathbf{w}_i$.

A Tabela 2 apresenta um conjunto de 8 vetores-código obtidos com a metodologia descrita neste trabalho. Os primeiros $N/2$ vetores-código foram obtidos por meio de aprendizagem competitiva, usando como conjunto de treino os vetores de treino pertencentes a S_{POS} . Observe que os vetores-código obtidos *refletem* a simetria dos sinais de voz: metade deles tem média positiva e metade tem média negativa; cada vetor-código tem seu correspondente simétrico.

Em virtude da simetria introduzida nos dicionários por meio da metodologia de projeto de dicionário propos-

ta neste trabalho, faz-se necessário armazenar apenas metade dos vetores-código. De fato, armazenando-se apenas os vetores-código $w_{i,POS}$, os vetores-código $w_{i,NEG}$ são facilmente determinados. Com isso, reduz-se pela metade os requisitos de memória de armazenamento para o dicionário. A simetria dos dicionários também contribui para a utilização de uma estratégia para redução da complexidade computacional da etapa de codificação (determinação do vizinho mais próximo) da QV, conforme mostra a seção a seguir. No presente trabalho, apenas a metade do dicionário correspondente aos vetores $w_{i,POS}$ é efetivamente armazenada na memória de referência do codificador. O decodificador, por outro lado, dispõe dos N vetores-código do dicionário.

5. Método para Redução da Complexidade Computacional da QV

O método de codificação considerado neste trabalho para reduzir o número de operações realizadas na etapa de codificação de um determinado vetor de entrada (vetor da fonte a ser quantizada) x é descrito a seguir: dado x , se $\text{média}(x) \geq 0$ então² o codificador realiza uma busca pelo vizinho mais próximo de x apenas nos vetores-código $w_{i,POS}$, isto é, nos vetores-código efetivamente armazenados na memória de referência do codificador. Neste caso, é transmitida para o decodificador uma palavra-binária iniciada por 0, indicando para o decodificador que o vetor-código a ser produzido como representante de x trata-se de um vetor do tipo $w_{i,POS}$, seguido de $\log_2(N/2)$ bits necessários para representar o índice i do vetor $w_{i,POS}$ selecionado da memória de referência. Por outro lado, se $\text{média}(x) < 0$, a busca pelo vizinho mais próximo deveria contemplar uma procura dentre os vetores-código $w_{i,NEG}$. Como estes vetores não se encontram armazenados na memória de referência do codificador, a codificação consiste em comparar o vetor $-x$ (vetor simétrico de x) com os vetores-código $w_{i,POS}$. Uma vez determinado o melhor representante para $-x$, é transmitida uma palavra-binária iniciada por 1, indicando que o vetor-código a ser produzido pelo decodificador como representante de x trata-se de um vetor do tipo $w_{i,NEG}$, seguido de uma seqüência de $\log_2(N/2)$ bits: cada bit dessa seqüência é o complemento do bit correspondente na seqüência de $\log_2(N/2)$ bits necessários para representar o índice i do vetor $w_{i,POS}$ selecionado como o mais próximo de $-x$ segundo um critério de distorção mínima. É importante observar que, devido à simetria introduzida no dicionário, o fato de x ter como vizinho mais próximo um determinado vetor-código $w_{N+1-i,NEG}$ implica dizer que o simétrico $-x$ tem como vizinho mais próximo $w_{i,POS} = -w_{N+1-i,NEG}$.

Para ilustrar o método de codificação proposto neste trabalho, considere o dicionário da Tabela 2 (disponível para o decodificador). Já o codificador dispõe da Tabela 3, correspondente aos primeiros $N/2$ vetores-

código, do tipo $w_{i,POS}$, da Tabela 2. Suponha que o sistema de comunicação receba como entrada o vetor $x = [0,0657 \ 0,0692]^T$. Após a avaliação da média³ das componentes desse vetor, o codificador decide que o vetor-código a ser selecionado como o representante (como a versão quantizada) de x trata-se de um vetor do tipo $w_{i,POS}$. Determina-se, assim, que o primeiro bit da palavra-binária a ser transmitida para o decodificador é 0. A busca do vizinho mais próximo é então realizada no dicionário efetivamente armazenado (Tabela 3) na memória de referência do codificador. Seguindo o critério de distorção mínima o vetor $[0,0628 \ 0,0621]^T$, com representação binária 01, é selecionado. O codificador, portanto, transmite para o decodificador a palavra-binária 001, em que o primeiro bit informa que a versão quantizada de x é do tipo $w_{i,POS}$ (isto é, um dos primeiros $N/2$ vetores do dicionário de N vetores do decodificador) e os dois últimos bits são uma representação binária do índice do vetor selecionado (da tabela do codificador) como vizinho mais próximo de x . No outro lado do sistema de comunicação baseado em QV, o decodificador, que dispõe do dicionário da Tabela 2, ao receber a palavra-binária 001, produz como saída o vetor $[0,0628 \ 0,0621]^T$.

Suponha, agora, que o sistema de comunicação tenha como entrada o vetor $x = [-0,1125 \ -0,1257]^T$. Após avaliar a média desse vetor, o codificador decide que a versão quantizada de x trata-se de um vetor do tipo $w_{i,NEG}$. Determina-se, assim, que o primeiro bit da palavra-binária a ser transmitida para o decodificador é 1. É realizada uma busca pelo vizinho mais próximo de $-x$ dentre os $N/2$ vetores-código disponíveis (Tabela 3) para o codificador: é selecionado o vetor-código $[0,1271 \ 0,1288]^T$ por ser o mais próximo (semelhante) a $[0,1125 \ 0,1257]^T = -x$. O codificador, portanto, transmite para o decodificador a palavra-binária 111, em que o primeiro bit informa que o vetor-código selecionado como a versão quantizada do vetor de entrada (vetor da fonte) x é do tipo $w_{i,NEG}$ e os dois últimos bits correspondem ao complemento da palavra-binária 00 que identifica o vetor $[0,1271 \ 0,1288]^T$ no dicionário efetivamente disponível ao codificador. No outro lado do sistema de comunicação, o codificador, ao receber a palavra-binária 111, produz como saída o vetor $[-0,1271 \ -0,1288]^T$.

5.1. Número de Operações do Método Proposto

O método de codificação considerado no presente trabalho será denotado a partir de agora por 1/2DIC, em virtude de o codificador utilizar apenas metade do dicionário previamente projetado.

Para codificar um determinado vetor de entrada, o método 1/2DIC testa, inicialmente, a média desse vetor.

²Para um vetor $x = [x_1 \ x_2 \ \dots \ x_K]^T$, em que T denota a operação de transposição, $\text{média}(x) = \frac{1}{K} \sum_{j=1}^K x_j$.

³Tendo em vista que a soma positiva de todas as componentes de um vetor implica média positiva das componentes, em lugar de determinar a média das componentes do vetor o método de codificação proposto calcula apenas a soma das componentes. Economiza-se, assim, uma operação de divisão para cada vetor de entrada.

Tabela 3: Vetores-código bidimensionais efetivamente armazenados na memória de referência do codificador.

i	w_{i1}	w_{i2}	Representação binária
1	0,1271	0,1288	00
2	0,0628	0,0621	01
3	0,2560	0,2536	10
4	0,0103	0,0105	11

Para tanto, é necessário apenas avaliar a soma das componentes desse vetor:

$$\begin{cases} \text{soma}(\mathbf{x}) \geq 0 \Rightarrow \text{média}(\mathbf{x}) \geq 0 \\ \text{soma}(\mathbf{x}) < 0 \Rightarrow \text{média}(\mathbf{x}) < 0 \end{cases} \quad (4)$$

O teste da média requer, portanto, $K - 1$ adições (uma vez que a soma das componentes do vetor de entrada é dada por $\text{soma}(\mathbf{x}) = \sum_{j=1}^K x_j$) e 1 comparação (com 0).

Considerando a medida de distorção erro quadrático da Equação (2) e tendo em vista que o codificador realiza a busca pelo vizinho mais próximo em apenas $N/2$ vetores-código, observa-se que para codificar cada vetor de entrada são necessárias ainda $K(N/2)$ multiplicações, $K(N/2)$ subtrações, $(K - 1)(N/2)$ adições e $(N/2) - 1$ comparações.

O número total de adições do método de codificação 1/2DIC é, portanto, dado por $(K - 1) + (K - 1)(N/2)$, em que o primeiro termo da soma diz respeito ao “teste da média” e o segundo se refere ao processo de procura pelo vizinho mais próximo no dicionário com $N/2$ vetores-código. O número total de comparações, por sua vez, é igual a $1 + (N/2 - 1)$, em que o primeiro termo da soma se deve ao “teste da média” e o segundo termo é concernente às comparações de distância (distorção quadrática) do processo de busca pelo vizinho mais próximo.

A Tabela 4 apresenta um resumo do número total de operações envolvido no método de busca total (busca exaustiva pelo vizinho mais próximo, realizada de forma convencional em um dicionário com N vetores-código), denotado por BT, e no método 1/2DIC (em que apenas metade do dicionário, correspondente a $N/2$ vetores-código, é efetivamente utilizada pelo codificador).

Tabela 4: Número de operações requerido para codificar um vetor, ao serem utilizados os métodos BT e 1/2DIC, em função de K e N .

	N ^o de operações	
	BT	1/2DIC
×	KN	$KN/2$
-	KN	$KN/2$
+	$(K - 1)N$	$(K - 1)(1 + N/2)$
Comp.	$N - 1$	$N/2$

6. Resultados

A Tabela 4 mostra que o método 1/2DIC apresenta uma economia de 50% em termos do número de multiplicações (bem como subtrações) com relação ao método convencional BT.

Conforme se pode observar na Tabela 4, fixado K , à medida que N aumenta o método 1/2DIC tende a levar a uma economia de 50% em termos do número de adições em relação ao método BT. Para $K = 4$ e $N = 128$, por exemplo o método 1/2DIC requer 195 adições, o que representa uma economia de 49,21% em relação às 384 adições requeridas pelo método BT.

No que diz respeito ao número de comparações, por meio da Tabela 4 é possível observar que para $N = 32$ e $N = 64$ são obtidas, respectivamente, economias de 48,38% e 49,20% com a utilização do método 1/2DIC em lugar do método BT. A Tabela 4 mostra que a economia do número de comparações do método 1/2DIC com relação ao método BT tende a 50% à medida que N aumenta.

Foram realizadas simulações que tiveram como objetivo avaliar a qualidade, por meio da relação sinal-ruído segmental (SNRseg) [15], do sinal de voz reconstruído ao se utilizar o método 1/2DIC. Constatou-se que o método 1/2DIC leva a uma queda, em geral pequena, de SNRseg do sinal reconstruído quando comparado com o método BT.

7. Comentários Finais

Neste trabalho foi apresentada uma nova metodologia para o projeto de dicionários para quantização vetorial, concebidos de tal maneira que os vetores-código reflitam a simetria apresentada por alguns sinais, dentre os quais os sinais de voz. A metodologia consiste em utilizar aproximadamente metade dos vetores do conjunto de treino (aqueles cujas componentes têm valor médio positivo) para obter, por meio de um algoritmo iterativo de treinamento de dicionário, metade do dicionário, correspondente a $N/2$ vetores-código. A outra metade dos vetores-código do dicionário é obtida de uma forma simples: esses demais $N/2$ vetores-código são determinados (calculados) de tal maneira que sejam os correspondentes simétricos dos primeiros $N/2$ vetores-código, obtidos pelo algoritmo de treinamento.

A simetria dos dicionários projetados foi explorada de forma a reduzir os requisitos de memória exigida para armazenar os vetores-código bem como para reduzir significativamente a complexidade computacional (número de operações) envolvida na etapa de codificação da quantização vetorial.

Com o método de codificação considerado no presente trabalho, o codificador do sistema de comunicação baseado em QV realiza uma busca pelo vizinho mais próximo em apenas metade dos vetores-código do dicionário. Ao se utilizar como medida de distorção a distância quadrática, o método proposto leva a uma economia de 50% do número de multiplicações e subtrações

quando comparado com o método convencional de busca total (BT). Além disso, ao ser considerado um dicionário de dimensão K e tamanho N , o método proposto requer $(K - 1)(1 + N/2)$ adições e $N/2$ comparações para codificar um vetor de entrada, enquanto o método BT requer $(K - 1)N$ adições e $N - 1$ comparações. Observou-se, no entanto, que o método de codificação considerado neste trabalho pode introduzir uma queda (em geral pequena) de relação sinal-ruído segmental (SNRseg) dos sinais de voz reconstruídos, quando comparado com o valor de SNRseg obtido com a utilização do método BT.

Referências

- [1] A. Gersho and R. M. Gray. *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, Boston, MA, 1992.
- [2] R. M. Gray. Vector Quantization. *IEEE ASSP Magazine*, pages 4–29, April 1984.
- [3] T. Berger. *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Prentice-Hall, Englewood Cliffs, NJ, 1971.
- [4] R. M. Gray and D. L. Neuhoff. Quantization. *IEEE Transactions on Information Theory*, 44(6):2325–2383, October 1998.
- [5] Y. Linde, A. Buzo, and R. M. Gray. An algorithm for vector quantizer design. *IEEE Transactions on Communications*, 28(1):84–95, January 1980.
- [6] B. Ramamurthi and A. Gersho. Classified vector quantization of images. *IEEE Transactions on Communications*, 34(11):1105–1115, November 1986.
- [7] H. Abut, R. M. Gray, and G. Rebolledo. Vector quantization of speech and speech-like waveforms. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 30(3):423–435, June 1982.
- [8] A. Gersho and V. Cuperman. Vector quantization: A pattern-matching technique for speech coding. *IEEE Communications Magazine*, pages 15–20, December 1983.
- [9] N. M. Nasrabadi and R. A. King. Image coding using vector quantization: A review. *IEEE Transactions on Communications*, 36(8):957–971, August 1988.
- [10] A. Alcaim, J. A. Solewicz, and J. A. Moraes. Frequência de ocorrência dos fones e listas de frases foneticamente balanceadas no português falado no Rio de Janeiro. *Revista da Sociedade Brasileira de Telecomunicações*, 7(1):23–41, Dezembro 1992.
- [11] T. Kohonen. The self-organizing map. *Proceedings of the IEEE*, 78(9):1464–1480, September 1990.
- [12] A. K. Krishnamurthy, S. C. Ahalt, D. E. Melton, and P. Chen. Neural networks for vector quantization of speech and images. *IEEE Journal on Selected Areas in Communications*, 8(8):1449–1457, October 1990.
- [13] E. Yair, K. Zeger, and A. Gersho. Competitive learning and soft competition for vector quantizer design. *IEEE Transactions on Signal Processing*, 40(2):294–309, February 1992.
- [14] N. B. Karayiannis and P.-I. Pai. Fuzzy vector quantization algorithms and their applications in image compression. *IEEE Transactions on Image Processing*, 4(9):1193–1201, September 1995.
- [15] N. S. Jayant and P. Noll. *Digital Coding of Waveforms*. Prentice-Hall, Englewood Cliffs, NJ, 1984.