

# AVALIAÇÃO DOS DICIONÁRIOS PCA PARA USO EM TRANSMISSÃO DE VOZ POR CANAL RUIDOSO

F. Madeiro Waslon T. A. Lopes Marcelo S. Alencar Benedito G. Aguiar Neto

Laboratório de Comunicações, DEE, COPELE – Universidade Federal da Paraíba  
58.109-970 - Campina Grande, PB, Brasil  
{madeiro,waslon,malencar,bganeto}@dee.ufpb.br

## RESUMO

O projeto de dicionários desempenha um papel importante para o bom desempenho de sistemas de comunicações baseados em quantização vetorial. Recentemente [1], mostrou-se que um algoritmo para projeto de dicionários denominado PCA (*principal component analysis*) apresenta diversas vantagens com relação ao algoritmo LBG (Linde-Buzo-Gray). No presente trabalho, mostra-se que, além das vantagens relatadas em [1], o algoritmo PCA produz dicionários que apresentam, como característica inerente, uma menor sensibilidade aos erros de canal quando comparados aos dicionários projetados com o algoritmo LBG. Por meio de simulações envolvendo transmissão de voz por um canal binário simétrico, este trabalho mostra que a queda de qualidade do sinal reconstruído, em relação a uma transmissão sem erro, é menor ao serem utilizados dicionários PCA.

## 1. INTRODUÇÃO

Dentre os algoritmos utilizados para projeto de dicionários aplicados à codificação de sinais baseada em quantização vetorial (QV), o algoritmo LBG (Linde-Buzo-Gray) [2] destaca-se por sua ampla utilização. Em [1], mostrou-se que o algoritmo para projeto de dicionários denominado PCA (*principal component analysis*) apresenta-se como uma alternativa adequada de projeto – diversas avaliações apontaram uma superioridade do algoritmo PCA sobre o algoritmo LBG.

No presente trabalho é realizada uma avaliação da *qualidade inerente* dos dicionários PCA sob o ponto de vista da robustez aos erros de transmissão de voz por um canal ruidoso. Nesse contexto, cumpre salientar que, em se tratando de compressão de voz baseada em QV, os algoritmos PCA e LBG visam gerar dicionários de qualidade, no sentido de que levem a sinais reconstruídos que apresentem a menor distorção para uma determinada taxa de codificação. A dinâmica dos algoritmos PCA e LBG não leva em consideração a questão da sensibilidade dos dicionários projetados aos erros de transmissão. Daí utilizar-se, com efeito, no presente trabalho, o termo *robustez inerente* dos dicionários aos erros de canal.

A avaliação dos dicionários PCA quanto à robustez aos erros de canal é levada a efeito utilizando-se o modelo de canal binário simétrico. Para efeito de comparação são também utilizados dicionários LBG com o mesmo modelo de canal. Resultados de simulação apresentados neste trabalho mostram que os erros de transmissão introduzidos pelo canal levam a uma menor perda de qualidade do sinal de voz reconstruído ao serem utilizados dicionários PCA. Fica apontada, portanto, uma superioridade dos dicionários PCA sobre os dicionários LBG no que diz respeito à *robustez inerente* aos erros de canal.

Este artigo encontra-se organizado da seguinte forma. A Seção 2 apresenta uma breve descrição de QV. Na Seção 3 é abordado o problema de QV em canais ruidosos, sendo apresentada uma figura de mérito utilizada para avaliar a robustez inerente dos dicionários aos erros de transmissão. A Seção 4 descreve o algoritmo PCA. Resultados e conclusão são apresentados, respectivamente, nas Seções 5 e 6.

## 2. QUANTIZAÇÃO VETORIAL

A quantização vetorial [3,4] pode ser definida como um mapeamento  $Q$  de um vetor de entrada  $x$  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  $W = \{w_i; i = 1, 2, \dots, N\}$  é o conjunto de vetores de reprodução (também denominados vetores-códigos ou vetores de reconstrução),  $K$  é a dimensão do quantizador vetorial e  $N$  é o tamanho do dicionário, isto é, o número de vetores-códigos.

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

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

em que cada célula  $S_i$  é definida da seguinte maneira:

$$S_i = \{x : Q(x) = w_i\}. \quad (3)$$

O vetor-código  $w_i$  constitui o vetor representativo de todos os vetores de entrada pertencentes à célula  $S_i$ .

Em um sistema de compressão de sinais baseado em QV, apresentado na Figura 1, um quantizador vetorial pode ser visto como a combinação de duas funções: um codificador e um decodificador. Dado um vetor  $x \in R^K$ , da fonte a ser codificada, o codificador calcula a distorção  $d(x, w_i)$  entre esse vetor de entrada (vetor a ser quantizado) e cada vetor-código do dicionário  $W$ . A regra ótima para codificação é a regra do vizinho mais próximo, na qual uma representação binária do índice  $I$ , denotada por  $b_I$ , é transmitida ao decodificador se o vetor-código  $w_I$  corresponder à menor distorção, isto é, se  $w_I$  for o vetor-código que apresenta a maior similaridade com  $x$  dentre todos os vetores-códigos do dicionário. Em outras palavras, o codificador usa a regra de codificação  $\mathcal{C}(x) = b_I$  se  $d(x, w_I) < d(x, w_i), \forall i \neq I$ . Ao receber a representação binária  $b_I = (b_1(I), b_2(I), \dots, b_m(I))$  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  $w_I$  como a reprodução (versão quantizada) de  $x$ . Em outras palavras, é utilizada a seguinte regra de decodificação:  $D(b_I) = w_I$ .

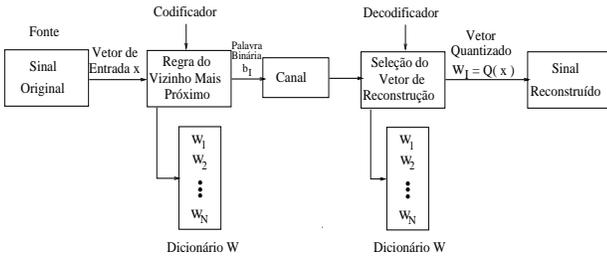


Figura 1: Codificação/decodificação em um sistema de codificação baseado em quantização vetorial simples.

A taxa de codificação de um quantizador vetorial, que mede o número de bits por componente de 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.

Uma das questões principais no projeto de dicionários para QV é o compromisso entre taxa e distorção. O alvo a ser perseguido é a obtenção de um dicionário que minimize a distorção média (para uma determinada taxa) introduzida pela aproximação de cada vetor de entrada por um dos vetores-códigos. Dentre os métodos para projeto de dicionários, o algoritmo LBG constitui a técnica mais amplamente utilizada.

### 3. QV EM CANAIS RUIDOSOS

Quando o sistema de comunicação envolve a utilização de um canal ruidoso, o desempenho do dicionário pode ser seriamente prejudicado.

Considere o sistema apresentado na Figura 2. O objetivo do sistema de comunicação é transmitir uma seqüência de vetores  $X = \{x\} \subset R^K$  através de um canal ruidoso, utilizando, para tanto, QV dos vetores de entrada (vetores da fonte)  $x$ . Seja  $b_i \in \{0, 1\}^m$  a palavra-binária de  $m$  bits enviada pelo codificador. Devido a erros de canal, considere que o decodificador recebe a palavra-binária  $b_j$ , produzindo como saída o vetor reconstruído  $\hat{x} = w_j$ .

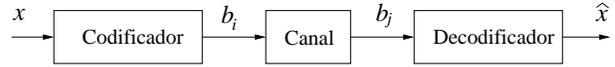


Figura 2: Sistema de comunicação.

Seja  $p_x$  a função densidade de probabilidades  $K$ -dimensional da fonte  $X$ . Seja  $p_i$  a probabilidade *a priori* de o vetor  $w_i$  ser selecionado como a reconstrução de  $x$  e  $p_{j|i}$  a probabilidade de o decodificador receber  $b_j$  dado que  $b_i$  foi enviado pelo codificador. A distorção total introduzida ao se transmitir a fonte  $X$  através do canal ruidoso pode ser expressa como [5, 6]

$$D = D_Q + D_C = \sum_{i=1}^N \int_{S_i} p_x \|x - w_i\|^2 dx + \sum_{i=1}^N \sum_{j=1}^N p_i p_{j|i} \|w_i - w_j\|^2, \quad (4)$$

em que  $D_Q$  e  $D_C$  representam, respectivamente, a distorção de quantização e a distorção devida aos erros de canal, enquanto  $\|w_i - w_j\|^2$  denota a distância euclidiana entre os vetores-códigos  $w_i$  e  $w_j$ .

Uma figura de mérito, denominada índice de desordem [7], pode ser utilizada para avaliar o grau de organização do dicionário tendo em vista o problema de transmissão por canal ruidoso. O índice de desordem é definido como

$$I_{\text{dis}} = \sum_{i=1}^N \sum_{j \in H^1(i)} \|w_i - w_j\|^2, \quad (5)$$

em que  $j \in H^1(i)$  denota o conjunto de palavras  $b_j$  tais que a distância de Hamming para  $b_i$  é 1.

O índice de desordem acumula as distâncias euclidianas entre vetores-códigos cujas palavras-binárias apresentam distância de Hamming igual a 1 bit<sup>1</sup>. A formulação matemática de  $I_{\text{dis}}$  encontra seu respaldo teórico na seguinte afirmação [8]: as técnicas de atribuição de índices (QV robusta) [5] devem ser levadas a efeito no sentido de assegurar que palavras-binárias com pequenas (grandes) distâncias

<sup>1</sup>Para baixas probabilidades de erro de bit de um canal ruidoso, a probabilidade de ocorrência de múltiplos erros nas palavras-binárias é pequena quando comparada com a probabilidade de ocorrência de erro em um bit.

de Hamming correspondam a vetores-códigos com pequenas (grandes) distâncias euclidianas [9, 10]. Assim, quanto menor  $I_{dis}$  de um dicionário, maior a robustez inerente do dicionário à transmissão ruidosa, isto é, menor a sensibilidade do dicionário aos erros que o canal introduz nas palavras-binárias transmitidas pelo codificador do sistema de comunicação baseado em QV.

#### 4. PROJETO DE DICIONÁRIOS USANDO ANÁLISE DE COMPONENTES PRINCIPAIS

O algoritmo PCA utiliza os autovalores e os autovetores (componentes principais) da matriz de covariância de um sinal de voz típico para calcular os vetores-códigos do dicionário. Consiste da seguinte seqüência de passos:

1. Defina a dimensão  $K$  e o tamanho do dicionário  $N$ ;
2. A partir da matriz de covariância  $C_{xx}$  de ordem  $K$  de um sinal de voz típico,  $\mathbf{X}$ , determine os autovalores  $\lambda_i$  e os autovetores  $\mathbf{z}_i$ ;  $i = 1, 2, \dots, K$ , obtidos por

$$C_{xx}\mathbf{z}_i = \lambda_i\mathbf{z}_i; \quad (6)$$

3. Defina  $L \leq K$  vetores  $\mathbf{z}_i$  (componentes principais) ao longo de cujas direções os vetores-códigos devem ser alocados. O número  $L$  é escolhido de acordo com o valor percentual relativo  $\lambda_{i,p}$  de cada autovalor  $\lambda_i$ :

$$\lambda_{i,p} = \frac{\lambda_i}{\sum_{i=1}^K \lambda_i}. \quad (7)$$

Apenas os  $L$  autovetores mais significativos, os quais definem as  $L$  direções principais, são escolhidos, de acordo com os valores mais significativos de  $\lambda_{i,p}$ ;

4. Para cada  $\mathbf{z}_i$  escolhido, determine um vetor  $\tilde{\mathbf{z}}_i = r_i\mathbf{z}_i$  ( $i = 1, 2, \dots, L$ ), em que o escalar  $r_i$  é o recíproco do valor absoluto da componente de maior valor absoluto. O vetor  $\tilde{\mathbf{z}}_i$ , portanto, tem no mínimo uma componente com valor absoluto 1 e as demais componentes apresentam valores absolutos situados na faixa de 0 a 1;

5. Seja  $N_i$  o número de vetores-códigos a serem alocados ao longo da direção definida pelo vetor  $\tilde{\mathbf{z}}_i$ . Escolha cada  $N_i$  como uma fração de  $N$ , proporcionalmente a  $\lambda_{i,p}$ , de modo que

$$N = \sum_{i=1}^L N_i; \quad (8)$$

6. Finalmente, aloque  $N_i$  vetores-códigos  $\mathbf{w}_{i,n_i}$  em cada  $i$ -ésima direção principal de acordo com

$$\mathbf{w}_{i,n_i} = f(n_i, \lambda_i)\tilde{\mathbf{z}}_i, \quad (9)$$

em que  $i = 1, 2, \dots, L$ ,  $n_i = 1, 2, \dots, N_i$  e os escalares  $f(n_i, \lambda_i)$  são determinados assumindo-se que cada  $i$ -ésima

direção principal está relacionada a um conjunto de vetores de entrada cujas componentes têm distribuição gaussiana com média igual à média do sinal de voz e variância  $\sigma_i^2 = \lambda_i$ . Os escalares  $f(n_i, \lambda_i)$ ,  $n_i = 1, 2, \dots, N_i$ , são determinados de tal forma que a área sob a função densidade de probabilidades gaussiana seja igualmente dividida em  $N_i$  intervalos.

Dentre as principais características do algoritmo PCA, podem ser destacadas [1, 11]: simplicidade e rapidez, geração de dicionários estruturados, inexistência de parâmetros a serem ajustados e independência de dicionário inicial. De fato, o algoritmo PCA é bastante simples e intuitivo: os vetores-códigos são “calculados”, de acordo com a estatística do sinal (de acordo com os autovalores e autovetores da matriz de covariância de um sinal de voz), diferindo consideravelmente de abordagens tradicionais, tais como os algoritmos LBG e de Kohonen [12], que atualizam iterativamente os vetores-códigos utilizando um conjunto de treino. O algoritmo PCA também não apresenta o incômodo problema de ajuste de parâmetros e não requer a definição de um dicionário inicial, como ocorre nos algoritmos LBG e de Kohonen. Além disso, testes subjetivos informais relatados em [11] mostraram que, em se tratando de QV de forma de onda com dimensão  $K = 2$ , os sinais de voz reconstruídos utilizando-se dicionários PCA apresentam qualidade superior à apresentada pelos sinais reconstruídos com uso de dicionários LBG.

#### 4.1. Robustez Inerente aos Erros de Canal

Em virtude de apresentarem uma organização estruturada, os dicionários PCA, conforme mostram os resultados da Seção 5, apresentam uma maior robustez inerente aos erros de transmissão quando comparados aos dicionários LBG. A Figura 3 mostra que o algoritmo PCA gera dicionários organizados (estruturados), ao contrário do algoritmo LBG, que gera dicionários sem estrutura alguma, conforme ilustra a Figura 4. Observa-se nas Tabelas 1(a) e 1(b) que a estrutura dos dicionários PCA contribui para que eles apresentem uma menor sensibilidade aos erros de canal quando comparados aos dicionários LBG. De fato, as Tabelas 1(a) e 1(b) mostram que nos dicionários PCA vetores-código com pequenas/grandes distâncias euclidianas têm uma melhor correspondência (relativamente ao algoritmo LBG) com palavras-binárias com pequenas/grandes distâncias de Hamming. A Tabela 1(a) mostra que os dicionários PCA têm uma característica interessante: vetores-códigos simétricos têm palavras-binárias com máxima distância de Hamming, isto é, com distância de Hamming igual a  $\log_2 N$  bits, o que também justifica a qualidade dos dicionários PCA no que diz respeito à robustez inerente aos erros de canal.

Tabela 1: Dicionários PCA e LBG com 16 vetores-códigos bidimensionais  $w_i = [w_{i1} \ w_{i2}]^T$ , com  $1 \leq i \leq 16$ . A palavra-binária do  $i$ -ésimo vetor-código,  $w_i$ , é denotada por  $b_i$ , enquanto que  $w_{ij}$  representa a  $j$ -ésima componente do vetor  $w_i$ .

$i$	$w_{i1}$	$w_{i2}$	$b_i$
1	0,5784	0,5784	0000
2	0,1365	0,1365	0001
3	0,1020	0,1020	0010
4	0,0765	0,0765	0011
5	0,0550	0,0550	0100
6	0,0357	0,0357	0101
7	0,0176	0,0176	0110
8	0,0010	0,0010	0111
9	-0,0010	-0,0010	1000
10	-0,0176	-0,0176	1001
11	-0,0357	-0,0357	1010
12	-0,0550	-0,0550	1011
13	-0,0765	-0,0765	1100
14	-0,1020	-0,1020	1101
15	-0,1365	-0,1365	1110
16	-0,5784	-0,5784	1111

(a) PCA.

$i$	$w_{i1}$	$w_{i2}$	$b_i$
1	0,0006	-0,0003	0000
2	-0,1391	-0,1414	0001
3	-0,0740	-0,0901	0010
4	0,0468	0,0437	0011
5	0,0576	-0,2353	0100
6	0,1532	-0,0554	0101
7	-0,3344	-0,3209	0110
8	-0,2202	-0,2224	0111
9	-0,2029	-0,0461	1000
10	-0,0446	-0,0288	1001
11	0,2635	0,3676	1010
12	0,5461	0,5536	1011
13	0,1871	0,1941	1100
14	0,1223	0,1106	1101
15	-0,0178	0,1747	1110
16	0,3902	0,2229	1111

(b) LBG.

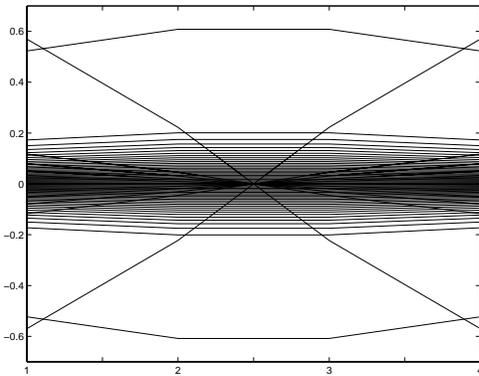


Figura 3: Dicionário PCA com  $K = 4$  e  $N = 64$ . Cada curva do conjunto de 64 curvas é obtida conectando-se os pontos correspondentes às componentes dos vetores-códigos.

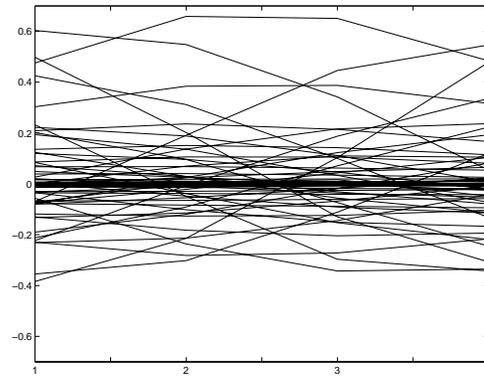


Figura 4: Dicionário LBG com  $K = 4$  e  $N = 64$ . Cada curva do conjunto de 64 curvas é obtida conectando-se os pontos correspondentes às componentes dos vetores-códigos.

## 5. RESULTADOS

Nesta seção são apresentados resultados de simulação envolvendo transmissão de voz por um canal binário simétrico (BSC, *binary symmetric channel*). Para avaliação da robustez inerente dos dicionários PCA e LBG aos erros de canal, utilizou-se o sinal de voz (com média  $4,6129 \cdot 10^{-5}$ , desvio-padrão 0,1142, valor máximo 0,9844 e valor mínimo -0,7334) constituído de 29120 amostras (3,64 s), correspondente ao conjunto de sentenças “*O sol*

*ilumina a fachada de tarde. Trabalhou mais do que podia*”.

A aquisição (resolução 8,0 bit/amostra e taxa de amostragem 8 kHz) dos sinais de voz utilizados nas simulações<sup>2</sup> foi realizada usando uma estação de trabalho Sun<sup>®</sup>, instalada com utilitários de processamento de áudio. A qualidade dos sinais de voz reconstruídos foi avaliada utilizando-se a relação sinal-ruído segmental (SNRseg) [13, 14].

O primeiro conjunto de avaliações teve como objetivo

<sup>2</sup>Foram utilizados sinais distintos para projeto e para avaliação dos dicionários.

realizar uma análise comparativa dos dicionários PCA e LBG sob o ponto de vista do índice de desordem. A Tabela 2 mostra claramente que os dicionários PCA apresentam índice de desordem inferior ao apresentado pelos dicionários LBG, o que traduz (conforme mostram os resultados apresentados ao final desta seção) o fato de os dicionários PCA serem mais adequados que os dicionários LBG para o caso de o sistema de comunicação baseado em QV utilizar um canal ruidoso. Em virtude de apresentarem um menor índice de desordem, os dicionários PCA são menos sensíveis aos erros de canal que os dicionários LBG. Em outras palavras, os erros de canal (que introduzem degradação nas palavras  $b_i$  enviadas pelo codificador) fazem com que os vetores-códigos PCA decodificados incorretamente (correspondentes a palavras-binárias erroneamente recebidas como  $b_j$ ) se aproximem, em média, dos vetores-códigos PCA concernentes a uma transmissão sem erro.

Considerando, por exemplo, que o codificador tenha transmitido a palavra-binária 1001, devido a um erro introduzido pelo canal ruidoso, admita-se que o decodificador tenha recebido a palavra-binária 1011, que dista de 1 bit da palavra-binária transmitida. Assim, utilizando o dicionário PCA da Tabela 1(a), o decodificador produz como saída o vetor  $[-0,0550 \ -0,0550]^T$  quando de fato deveria produzir como saída o vetor  $[-0,0176 \ -0,0176]^T$ , caso a transmissão tivesse sido realizada sem erro. Por outro lado, usando o dicionário LBG da Tabela 1(b), o decodificador produz como saída o vetor  $[0,5461 \ 0,5536]^T$  quando de fato deveria produzir como saída o vetor  $[-0,0446 \ -0,0288]^T$ , caso a transmissão tivesse sido realizada sem erro. Observa-se, neste exemplo, que o erro de transmissão é mais crítico no caso LBG. De fato, ao ser utilizado o dicionário LBG, a transmissão sem erro leva a um vetor com duas componentes negativas de baixa amplitude – o erro de canal considerado leva a um vetor com componentes positivas de alta amplitude. No caso de ser utilizado dicionário PCA, tanto a transmissão sem erro quanto a transmissão com erro levam a um vetor com amostras negativas de baixa amplitude.

Tabela 2: Índice de desordem para os dicionários PCA e LBG, para diversos valores de  $K$  e  $N$ .

$K$	$N$	$I_{dis}$	
		LBG	PCA
2	16	15,08	11,60
2	32	37,43	18,96
2	64	63,02	42,51
2	128	105,62	72,26
4	16	16,87	14,59
4	32	38,29	22,92
4	64	64,31	37,97
4	128	92,56	66,30

A Tabela 3 apresenta valores de SNRseg do sinal reconstruído (mais precisamente, o valor médio de SNRseg resultante de 50 transmissões do sinal para cada uma das diversas probabilidades de erro de bit,  $\epsilon$ , consideradas) para dicionários PCA e LBG com  $N = 16, 32, 64$  e 128 vetores de reconstrução de dimensão  $K = 2$ .

A queda de qualidade provocada pelo canal ruidoso no sinal de voz reconstruído pode ser avaliada por meio de

$$RP_{\text{SNRseg}}(\epsilon) = \frac{\text{SNRseg}(\text{tsr}) - \text{SNRseg}(\epsilon)}{\text{SNRseg}(\text{tsr})} \times 100\% \quad (10)$$

em que  $RP_{\text{SNRseg}}(\epsilon)$  denota a redução percentual (em relação a uma transmissão sem erro, denotada por tsr) da SNRseg do sinal reconstruído após transmissão por um canal BSC com probabilidade de erro de bit igual a  $\epsilon$ , enquanto  $\text{SNRseg}(\text{tsr}) = \text{SNRseg}(\epsilon = 0,0)$  denota a SNRseg do sinal reconstruído após transmissão sem erro e  $\text{SNRseg}(\epsilon)$  representa a SNRseg do sinal reconstruído após transmissão por um canal BSC com probabilidade de erro de bit igual a  $\epsilon$ . Portanto, fixados  $N, K$  e  $\epsilon$ , diz-se que o dicionário PCA apresenta uma maior robustez inerente aos erros de canal que a apresentada pelo dicionário LBG se o valor de  $RP_{\text{SNRseg}}(\epsilon)$  obtido com o dicionário PCA for inferior ao valor de  $RP_{\text{SNRseg}}(\epsilon)$  obtido com o dicionário LBG. Nesse sentido, a Tabela 4 mostra que os dicionários PCA são menos sensíveis aos erros de canal que os dicionários LBG, em virtude de apresentarem, para todos os valores de  $N, K$  e  $\epsilon$  considerados, menores valores de  $RP_{\text{SNRseg}}(\epsilon)$ . Para  $K = 2, N = 32$  e  $\epsilon = 0,0005$ , por exemplo, o dicionário PCA apresenta

$$RP_{\text{SNRseg}}(0,0005) = \frac{11,01 - 10,37}{11,01} \times 100\% = 5,81\%$$

enquanto que o algoritmo LBG apresenta

$$RP_{\text{SNRseg}}(0,0005) = \frac{10,33 - 9,37}{10,33} \times 100\% = 9,29\%.$$

## 6. CONCLUSÃO

No presente trabalho mostrou-se que a organização estruturada dos dicionários PCA os qualifica como mais adequados que os dicionários LBG para QV de voz através de canais ruidosos. Simulações realizadas com um canal BSC mostraram que os erros de transmissão levam a uma menor queda de qualidade do sinal de voz reconstruído (com relação à transmissão sem erro) ao serem utilizados dicionários PCA em substituição a dicionários LBG.

## 7. AGRADECIMENTOS

Os autores expressam os agradecimentos à CAPES e ao CNPq pelo apoio financeiro ao trabalho.

Tabela 3: SNRseg (dB) em função da probabilidade de erro de bit,  $\epsilon$ , para os dicionários LBG e PCA com dimensão  $K = 2$  e tamanho  $N$ .

$N$	Dicionário	SNRseg							
		$\epsilon = 0,0$	$\epsilon = 0,0001$	$\epsilon = 0,0005$	$\epsilon = 0,001$	$\epsilon = 0,005$	$\epsilon = 0,01$	$\epsilon = 0,05$	$\epsilon = 0,1$
16	LBG	7,52	7,40	6,87	6,21	2,58	-0,17	-7,32	-10,19
	PCA	8,86	8,73	8,27	7,74	4,32	1,39	-7,04	-10,15
32	LBG	10,33	10,15	9,37	8,49	3,46	-0,03	-7,98	-10,91
	PCA	11,01	10,87	10,37	9,78	6,08	3,00	-4,89	-8,07
64	LBG	13,55	13,26	12,34	11,18	4,95	0,98	-7,39	-10,18
	PCA	12,35	12,17	11,42	10,56	5,65	1,99	-6,15	-9,05
128	LBG	15,81	15,54	14,54	13,51	7,09	2,67	-5,82	-8,57
	PCA	13,41	13,19	12,43	11,54	6,67	3,41	-4,44	-7,43

Tabela 4: Redução percentual de SNRseg em relação à transmissão sem erro, para alguns valores de probabilidade de erro de bit,  $\epsilon$ , para dicionários LBG e PCA com dimensão  $K = 2$  e tamanho  $N$ .

$N$	Dicionário	RP <sub>SNRseg</sub> ( $\epsilon$ )			
		$\epsilon = 0,0001$	$\epsilon = 0,0005$	$\epsilon = 0,001$	$\epsilon = 0,005$
16	LBG	1,59%	8,64%	17,42%	65,69%
	PCA	1,46%	6,65%	12,64%	51,24%
32	LBG	1,74%	9,29%	17,81%	66,50%
	PCA	1,27%	5,81%	11,17%	44,77%
64	LBG	2,14%	8,92%	17,49%	63,46%
	PCA	1,45%	7,53%	14,49%	54,25%
128	LBG	1,70%	8,03%	14,54%	55,15%
	PCA	1,64%	7,30%	13,94%	50,26%

## 8. REFERÊNCIAS

- [1] Madeiro, F., Vilar, R. M., Aguiar Neto, B. G., and Alencar, M. S. An Algorithm for Voice Waveform VQ Codebook Design Based on PCA. *Revista da Sociedade Brasileira de Telecomunicações*, Vol. 15, No. 2, pp. 59–69, December 2000.
- [2] Linde, Y., Buzo, A., and Gray, R. M. An Algorithm for Vector Quantizer Design. *IEEE Transactions on Communications*, Vol. COM - 28, No. 1, pp. 84–95, January 1980.
- [3] Gersho, A. and Gray, R. M. *Vector Quantization and Signal Compression*. Kluwer Academic Publishers, Boston, MA, 1992.
- [4] Gray, R. M. Vector Quantization. *IEEE ASSP Magazine*, pp. 4–29, April 1984.
- [5] Farvardin, N. A Study of Vector Quantization for Noisy Channels. *IEEE Transactions on Information Theory*, Vol. 36, No. 4, pp. 799–809, July 1990.
- [6] Knagenhjelm, P. and Agrell, E. The Hadamard Transform – A Tool for Index Assignment. *IEEE Transactions on Information Theory*, Vol. 42, No. 4, pp. 1139–1151, July 1996.
- [7] Lopes, W. T. A., Madeiro, F., Alencar, M. S., and Aguiar Neto, B. G. Simulated Annealing for Robust VQ: Improving Image Transmission through a Fading Channel. *Proceedings of the Vth Brazilian Symposium on Neural Networks (IEEE SBRN'2000)*, Rio de Janeiro – RJ, Brazil, pp. 243–248, November 2000.
- [8] Gray, R. M. and Neuhoff, D. L. Quantization. *IEEE Transactions on Information Theory*, Vol. 44, No. 6, pp. 2325–2383, October 1998.
- [9] Zeger, K. and Gersho, A. Pseudo-Gray Coding. *IEEE Transactions on Communications*, Vol. 38, No. 12, pp. 2147–2158, December 1990.
- [10] Lopes, W. T. A., Madeiro, F., Alencar, M. S., and Aguiar Neto, B. G. An index assignment algorithm for improving the transmission of vector-quantized images over a Rayleigh fading channel. *CD-ROM of the XVIII Brazilian Telecommunications Symposium (SBrT'2000)*, Gramado – RS, Brazil, September 2000.
- [11] Madeiro, F. Projeto e Avaliação de Dicionários para Quantização Vetorial de Voz e Imagem. *Proposta de Tese, Universidade Federal da Paraíba, Departamento de Engenharia Elétrica*, 2001.
- [12] Kohonen, T. *Self-Organization and Associative Memory (3rd ed)*. Springer-Verlag, Berlin, 1989.
- [13] Jayant, N. S. and Noll, P. *Digital Coding of Waveforms*. Prentice-Hall, Englewood Cliffs, NJ, 1984.
- [14] Deller Jr., J. R., Proakis, J. G., and Hansen, J. H. L. *Discrete-time Processing of Speech Signals*. Macmillan Publishing Co., 1993.