

Impacto da Quantização das LSF sobre a Qualidade de Voz em Codecs a Baixas Taxas Operando em Redes IP

Fred Berkowicz, Rodrigo C. de Lamare e Abraham Alcaim

Resumo—Este artigo examina o impacto da quantização vetorial das LSFs sobre a qualidade de voz em codecs a baixas taxas operando em redes IP. São considerados diferentes esquemas de quantização vetorial (QV) multiestágio com busca em árvore envolvendo QV sem memória e QV preditiva chaveada com 2 e 4 classes. A distribuição de perda de quadros em redes IP foi modelada de acordo com o Modelo de Gilbert e a avaliação de desempenho foi realizada tanto em termos das distorções espectrais como da qualidade de voz resultante de codecs a baixas taxas. Esta última foi medida através da recomendação ITU-T P.862 de avaliação perceptiva de qualidade da voz (PESQ), além de testes de escuta informais.

Palavras-Chave—Quantizadores vetoriais multiestágio, redes IP, perda de quadros, avaliação perceptiva de qualidade da voz (PESQ).

Abstract—This paper investigates the impact of LSF vector quantisation over the voice quality in low rate codecs operating in IP networks. Tree-structured multistage vector quantisation (VQ) schemes involving memoryless VQ and switched-predictive VQ with 2 and 4 classes are considered. The packet loss frame distribution in IP networks was modelled according to the Gilbert Model and the performance was carried out both in terms of spectral distortions and the speech quality at the out put of low rate codecs. The latter one was measured by means of ITU-T recommendation P.862 of perceptual evaluation of speech quality (PESQ) and informal listening tests.

Keywords—Multistage vector quantisation, IP networks, frame loss, perceptual evaluation of speech quality (PESQ).

I. INTRODUÇÃO

Atualmente a maior parte dos algoritmos de codificação de voz a baixas taxas é baseada na técnica de análise de predição linear (ou análise LPC - "Linear Predictive Coding"), onde um sinal de excitação é aplicado a um filtro só de pólos, caracterizado pelos parâmetros LPC, que representa a informação da envoltória espectral do sinal de voz. Este artigo trata do problema da transmissão da informação da envoltória espectral da voz e seu impacto sobre a qualidade da voz em codecs a baixas taxas, usando esquemas de quantização vetorial (QV) em redes IP. Bolot [1], [2] estudou a distribuição de perda de pacotes na Internet e concluiu que ela poderia ser aproximada por um modelo de perda Markoviano, também conhecido como Modelo de Gilbert. Neste trabalho as frequências em linhas espectrais ou parâmetros LSF foram escolhidos para representar os coeficientes LPC, uma vez

que são mais adequadas para procedimentos de quantização e interpolação [3].

Os parâmetros LSF apresentam, em geral, uma correlação significativa entre quadros de voz sucessivos, especialmente para segmentos sonoros. Na quantização vetorial sem memória (QVSM) [4], [5], cada vetor de LSFs é quantizado de maneira independente de qualquer outro conjunto de LSFs. Entretanto, esta não é a melhor forma de codificar os parâmetros LSF, uma vez que ganhos podem ser conseguidos ao explorar a correlação entre conjuntos de LSFs adjacentes. Um esquema eficiente de codificação das LSFs, onde emprega-se quantização vetorial preditiva chaveada (QVPC) foi apresentado em [6], [7].

Neste artigo, o QVSM e esquemas de QVPC serão comparados em canais sujeitos a perdas de quadros, simulados de acordo com o Modelo de Gilbert. O desempenho dos quantizadores será avaliado utilizando a medida de distorção espectral, percentual de *outliers* e análise da qualidade de voz em codificadores a baixas taxas usando a medida PESQ da recomendação ITU-T P.862 [8] e testes de escuta informais. A Seção II deste artigo apresenta uma breve descrição do codec utilizado nas simulações. A Seção III descreve os diferentes tipos de quantizadores vetoriais examinados. A Seção IV trata dos problemas envolvidos na transmissão de LSFs em redes IP. A Seção V apresenta e discute os resultados de simulações e a Seção VI resume as principais conclusões deste artigo.

II. DESCRIÇÃO DO CODEC

O codec utilizado neste trabalho, proposto por de Lamare e Alcaim [9] e mostrado na Figura 1, é baseado em excitação mista multibandas e opera em uma taxa de bits média de 1,2Kb/s. O codificador realiza uma análise de predição linear a cada quadro de 20 ms e emprega um quantizador vetorial preditivo chaveado (esquema QVPC4 a ser descrito na Seção III), para codificar 10 parâmetros LSF com 21 bits por quadro. O ganho é quantizado uniformemente com 5 bits por quadro e a excitação é codificada com 3 bits por quadro. Os quadros de voz classificados como sonoros são divididos em 3 bandas de frequências, que são obtidas com bancos de filtros fixos. Em seguida, uma análise em sub-bandas é realizada a fim de classificar as sub-bandas em sonoras e surdas. A excitação mista consiste da soma de pulsos periódicos nas sub-bandas sonoras com ruído branco nas sub-bandas restantes. Para quadros surdos, é empregada a técnica de modelagem e síntese de fricativos e oclusivos proposta em [10]. Com o objetivo

Este trabalho foi financiado em parte pelo CNPq. Os autores estão com o CETUC/PUC-RIO, 22453-900, Rio de Janeiro - Brasil Telephone: +55-21-31141701. E-mails: fredberko@ig.com.br, delamare@infolink.com.br e alcaim@cetuc.puc-rio.br

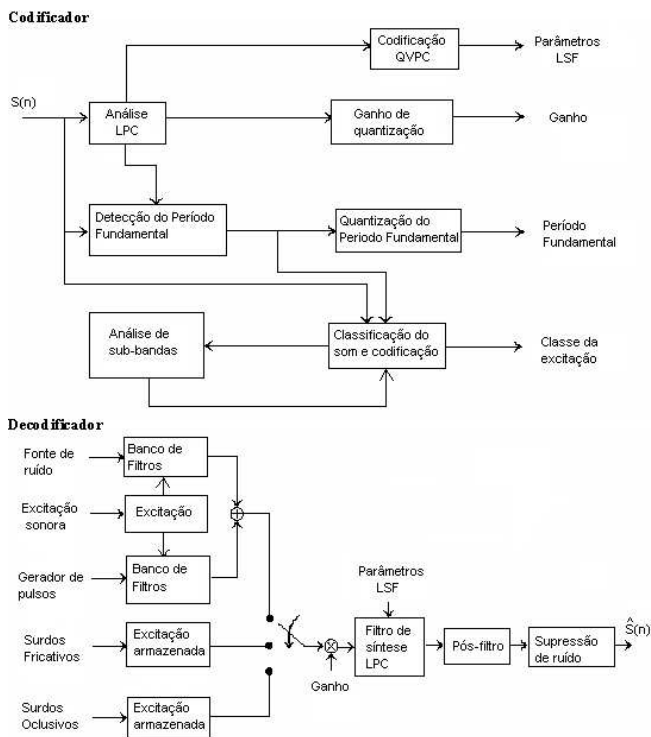


Fig. 1. Diagrama em blocos do codec.

de reduzir o ruído e melhorar a fala codificada são adotados um pós-filtro de restauração da envoltória espectral combinado com redução de ruído e um esquema de supressão de ruído. A alocação de bits do codificador é mostrada na Tabela I.

TABELA I
ALOCAÇÃO DE BITS

Parâmetros	Quadro sonoro	Quadro surdo
LSFs	21	0
Ganho	5	5
Excitação Fundamental	3	3
	6	0
Total bits/20 ms	35	8
Bit rate	1,75 kb/s	0,4 kb/s

III. DESCRIÇÃO DOS QUANTIZADORES VETORIAIS

O quantizador vetorial sem memória (QVSM), quantiza cada vetor de LSFs independente de qualquer outro conjunto de LSFs adjacentes [4]. Paliwal e Atal [5] demonstraram que um esquema de QV particionada sem memória é capaz de codificar de forma eficiente os parâmetros LSF com 24 bits por quadro. Um esquema de QV multiestágio sem memória com estrutura em árvore foi apresentado em [11] e seu desempenho superou o esquema de QV particionada sem memória. Uma técnica simples de melhorar o desempenho das estruturas em [11] e de obter vantagens na exploração da memória de uma fonte é usar quantização vetorial preditiva (QVP) [4]. Um preditor linear vetorial forma uma estimativa dos vetores de entrada como uma combinação linear de observações passadas, e o vetor residual de predição é quantizado vetorialmente e

pode ser expresso por:

$$\delta_{j+1} = \mathbf{f}_{j+1} - \hat{\mathbf{f}}_j \cdot \rho^t \quad (1)$$

onde ρ é o vetor com os coeficientes de correlação e $\hat{\mathbf{f}}_j$ é a versão quantizada de \mathbf{f}_j .

A correlação entre os vetores de LSF pode ser explorada pelo uso de QV com memória, como o QVP. Entretanto, existem situações de rápidas mudanças na envoltória espectral da voz e, portanto, baixas correlações entre os conjuntos de LSF adjacentes. Essa observação motivou a combinação de técnicas de QVSM e QVP para codificar quadros de baixa correlação separadamente dos quadros com alta correlação. Uma busca de ambos esquemas de QV é realizada para cada quadro e o melhor candidato, no que diz respeito a um critério de distorção, é codificado e transmitido [12].

Dentre os esquemas de quantização vetorial preditiva chaveada (QVPC), destacamos aquele denominado *safety-net*, proposto por Eriksson *et al.* [13]. Este sistema chaveia entre uma estrutura QV sem memória e uma estrutura QV preditiva, conseguindo uma quantização eficiente dos parâmetros LSF com apenas 20 bits por quadro. Outra técnica de QVPC, introduzida posteriormente por McCree e De Martin [14], usa 2 estruturas multiestágios de QVP com busca em árvore, onde cada preditor é projetado para um banco de dados de treinamento específico e opera com 21 bits por quadro.

Recentemente, foi desenvolvido um método de QVPC mais eficiente para quantização das LSF [6], [7]. Esse esquema, denominado QVPC4, combina QVSM e quantização vetorial preditiva (QVP) para codificar quadros com alta correlação temporal de forma separada de quadros com alta correlação temporal. Uma busca de diferentes QVs é realizada para cada quadro e o vetor de LSFs com melhor desempenho, no que se refere a um critério de distorção, é codificado e transmitido. Ele opera com 21 bits por quadro de 20 ms para quantização de $p = 10$ LSFs e usa 4 QVs multiestágios com busca em árvore, sendo 3 QVPs e 1 QVSM, onde cada QV possui 4 estágios. Para o QVPC, o procedimento de busca em árvore usa uma aproximação dos $M = 12$ melhores vetores-código. O erro quadrático-ponderado usado no QVPC, tanto para projeto dos dicionários como para realização da busca dentro de cada QV, emprega a função de ponderação definida por

$$\alpha_i(\mathbf{f}) = \frac{1}{f_i - f_{i-1}} + \frac{1}{f_{i+1} - f_i} \quad (3)$$

onde $i = 1, \dots, p$ e $f_0 = 0$ e $f_{p+1} = 0,5$.

A seleção do melhor QV, dentre os 4 usados neste sistema, é feita com base na distorção espectral (DE) definida por

$$DE = \left[\sum_{f=0}^{4000} \frac{1}{4000} \left(10 \log_{10} \left| \frac{S(f)}{\hat{S}(f)} \right| \right)^2 \right]^{1/2} \quad (dB) \quad (4)$$

onde $S(f)$ e $\hat{S}(f)$ representam a envoltória espectral original e quantizada, respectivamente. A base de dados para treinamento dos quantizadores foi dividida em 4 subconjuntos de acordo com a DE entre vetores LSF consecutivos. Os preditores usados nos esquemas QVP foram projetados com base nos coeficientes de correlação obtidos para cada subconjunto de vetores LSF de treinamento.

Os métodos de QVPC que serão considerados neste trabalho são apresentados na Tabela II. Seus desempenhos em diferentes taxas de bits, em termos da distorção espectral média, em canal livre de ruído são mostrados na Figura 2. Observa-se nessa figura que o QVPCP4 obteve melhor desempenho. Entretanto, testes realizados em presença de ruído [7] mostram que o QVPC4 tem melhor desempenho se comparado ao QVPCP4. Por outro lado, o QVSM apresentou o pior desempenho, mas em canais com altas taxas de erro de bit o QVSM mostrou-se mais eficiente [7]. Neste artigo escolhemos 3 quantizadores vetoriais para avaliarmos a influência da quantização das LSF sobre a qualidade de voz em codecs a baixas taxas operando em redes IP. O QVPC4 e o QVSM foram escolhidos, tendo em vista que testes em [7] mostraram que estes são mais robustos em presença de ruído. O terceiro quantizador vetorial que será utilizado na análise aqui apresentada é o QVPCP2 proposto por McCree e De Martin [14], de modo que se possa avaliar, também, o desempenho de um sistema composto só por QVPs em canais com perda de quadros. As simulações serão feitas utilizando 21 bits para codificar cada quadro de LSFs.

TABELA II
ESQUEMAS DE QVPC.

QVPC	Estrutura dos Quantizadores
QVPC	1 QVP and 1 QVSM .
QVPCP2	2 QVP.
QVPC4	3 QVP e 1 QVSM.
QVPCP4	4 QVP.

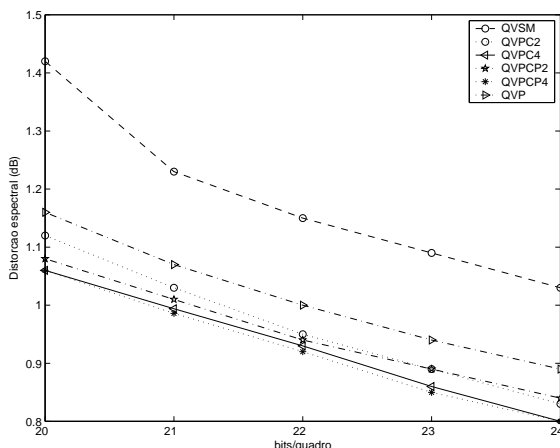


Fig. 2. Desempenho dos QVs em termos da DE média para diferentes taxas de bits.

IV. TRANSMISSÃO DAS LSFs EM REDES IP

O sistema considerado neste trabalho é uma rede IP em que no terminal de transmissão, um conjunto de parâmetros LSF é quantizado e codificado a cada quadro de voz em uma palavra-código de 21 bits. Cada seqüência de conjuntos de parâmetros LSF codificados está associada a uma seqüência de quadros, que por sua vez é caracterizada por uma seqüência de bits. Essa última é encapsulada em pacotes, de acordo com

o mecanismo de transmissão da rede, e enviada pelo canal. No destinatário, o enquadramento é desfeito pelo decodificador e os parâmetros LSF quantizados são recebidos.

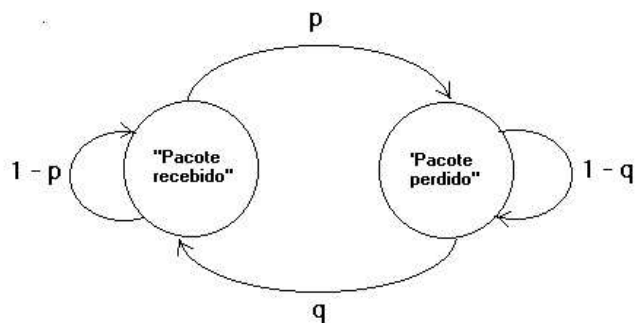


Fig. 3. Modelo de Gilbert.

Em uma rede IP, havendo congestionamento, poderá ocorrer a situação de *buffer overflow* nos switches ou roteadores levando ao descarte ou perda de pacotes, caracterizando um canal com perda ou apagamento de quadro (*Frame Erasure-FE*). As perdas de pacotes em redes IP normalmente ocorrem em rajadas. Supõe-se, sem perda de generalidade, que um quadro do quantizador é encapsulado em um pacote. Para avaliar o desempenho dos quantizadores, adotamos um modelo Markoviano de dois estados, também conhecido como Modelo de Gilbert, para representar o canal com perdas. Os dois estados se referem aos eventos “pacote recebido” e “pacote perdido”, respectivamente. Como mostrado na Figura 3, p é a probabilidade de transição do estado “pacote recebido” para o estado “pacote perdido”, e q a probabilidade de transição do estado “pacote perdido” para “pacote recebido”. A taxa de perda de quadro (TPQ), também conhecida como probabilidade de perda incondicional é dada por: $TPQp/(p+q)$. O comprimento médio da rajada (B) é dado por $B = 1/(1 - ppc)$, onde ppc é a probabilidade de perda condicional, que é, a probabilidade de transição do estado “pacote perdido” para “pacote perdido” ($ppc = 1 - q$).

Para combater os efeitos de perda de quadro nos codificadores existem algumas contramedidas. A primeira e mais simples delas é silenciar (*muting*) a saída de áudio enquanto não se estiver recebendo as LSFs perdidas por causa do descarte de quadros. Entretanto, pode-se obter um melhor desempenho se ao invés do *muting* for realizada a repetição (*repeating*) do último quadro de parâmetros LSF recebido. Uma solução um pouco mais atraente e adotada neste artigo é a interpolação dos parâmetros LSF. Devido às suas propriedades, as LSFs podem ser interpoladas linearmente para formar um conjunto válido de parâmetros LSF. Seja \mathbf{f}_i o conjunto de parâmetros LSF quantizados recebidos no quadro i . Considere o caso em que as L LSFs seguintes são perdidas e \mathbf{f}_{i+L+1} é recebida. Então, as LSFs interpoladas $\mathbf{f}_{i+1} \dots \mathbf{f}_{i+L}$ são dadas por:

$$\mathbf{f}_{i+x} = \frac{L+1-x}{L+1} \times \mathbf{f}_i + \frac{x}{L+1} \times \mathbf{f}_{i+L+1} \quad (5)$$

para $x = 1 \dots L$. Assim sendo, recebe-se, por exemplo, um conjunto de parâmetros LSF#1 e, por causa das imperfeições do canal, deixa-se de receber as LSF#2 e recebe-se as LSF#3.

A interpolação permite que se obtenha uma aproximação das LSF#2 às custas de um pequeno retardo adicional. Nota-se que é possível realizar a interpolação de mais de um conjunto de parâmetros LSFs às custas de um retardo cada vez maior [15], [16]. De acordo com [16], se n quadros consecutivos de duração t cada um, são perdidos, então o atraso devido à interpolação D_i é dado por

$$D_i = n \times t + RTT/2 \quad (6)$$

onde RTT (*Round Trip Time*) é o tempo que um pacote leva para ir e voltar de um destinatário (valores típicos para RTT são de 10-700ms). Ainda segundo [16], atrasos aceitáveis para aplicações de voz sobre IP (VoIP) devem ser menores que 800ms.

V. RESULTADOS E DISCUSSÃO

O codec utilizado nas simulações foi descrito na Seção II e os diferentes esquemas de quantização testados foram apresentados na Seção III deste artigo. A base de dados usada para treinamento dos QVs foi produzida por 20 locutores masculinos e 20 femininos, onde cada locutor pronunciou 2 conjuntos de dez frases obtidas de listas foneticamente balanceadas para o português falado no Rio de Janeiro [17], resultando em um total de 800 frases. Os desempenhos foram aferidos utilizando-se um conjunto de 60 frases distintas, também foneticamente balanceadas [17], produzidas por 2 locutores masculinos e 2 femininos, gerando uma coleção de 8992 quadros de LSFs, correspondendo a uma seqüência de $8992 \times 21 = 188832$ bits. Foram usados como critérios de avaliação objetiva a distorção espectral média explicitada por (4) e a distribuição dos *outliers*. São definidas duas classes de *outliers*: porcentagem de LSFs quantizadas com distorção espectral entre 2 e 4 dB e porcentagem de LSFs quantizadas com distorção espectral superior a 4 dB.

Para avaliar a qualidade da voz em Codecs operando sobre redes IP utilizamos a recomendação ITU-T P.862 de avaliação perceptiva de qualidade da voz (PESQ - *Perceptual evaluation of speech quality*), que consiste de uma técnica de medição objetiva para estimar a qualidade subjetiva que seria obtida em testes de escuta [8]. A saída do PESQ é a predição da qualidade percebida que seria dada a uma fala decodificada por um ouvinte em um teste de audição subjetivo como o MOS (*Mean Opinion Score*). A pontuação do PESQ é mapeada em uma escala tipo MOS, com um limite entre 1,0 e 4,5. Em situações de distorções extremamente altas, o resultado pode ficar abaixo de 1,0, mas isso é muito incomum. Atualmente o PESQ tornou-se um critério bem aceito para medição da qualidade de voz.

O modelo de perda de quadros, descrito na Seção IV, foi simulado para as condições de rede consideradas em um trabalho recente [1] e mostrada na Tabela III. As distribuições do comprimento da rajada de perdas de quadros para essas condições são apresentadas na Figura 4.

Os gráficos de distorção espectral *versus* TPQ e porcentagem de *outliers versus* TPQ são mostrados nas Figuras 5 e 6 respectivamente. Na Figura 5, observa-se que o desempenho do QVPC4 é superior a do QVSM para toda a faixa de taxa

TABELA III
PARÂMETROS DO MODELO DE GILBERT USADO PARA SIMULAR AS
CONDIÇÕES DA REDE COMO EM [1].

TPQ(%)	ppc	B	p	q
0	—	—	0	0
0,6	0,147	1,17	0,005	0,853
9	0,330	1,49	0,066	0,670
28,6	0,500	2,00	0,200	0,500
38,5	0,600	2,50	0,250	0,400

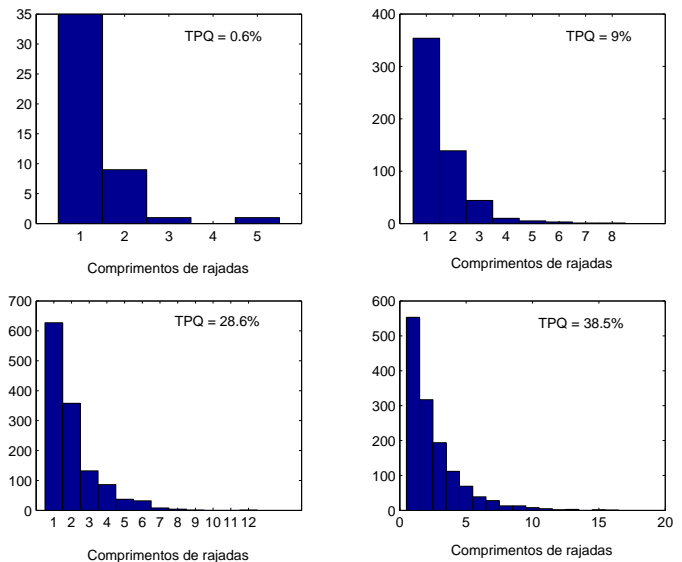


Fig. 4. Comprimento da rajada de perda de pacotes com os parâmetros de rede apresentados na Tabela II.

de perda de quadros considerada e em relação ao QVPCP2 o desempenho é um pouco superior. O melhor desempenho do QVPC4 é confirmado pela Figura 6, pois para toda a faixa de interesse, a quantidade de *outliers* com distorção espectral entre 2 e 4 dB do QVPC4 é menor em relação aos outros dois quantizadores utilizados nas simulações. Já a quantidade de *outliers* com distorção espectral maior do que 4 dB é comparável para os 3 quantizadores.

A Tabela IV e a Figura 7 apresentam os resultados em termos do PESQ versus TPQ, que mostram a influência da quantização na qualidade da voz em codecs a baixas taxas operando em redes IP. A avaliação feita com o PESQ mostrou que a qualidade da voz utilizando-se os quantizadores vetoriais QVPC4, QVPCP2 e QVSM para TPQ de até 28,6% é considerada boa, visto que uma pontuação acima de 2,0 para codecs a baixas taxas pode ser considerada satisfatória. Apenas com uma TPQ de 38,5% é que começamos a perceber uma maior degradação do sinal de voz. Além disso, o teste com o PESQ confirma que para toda a faixa de taxa de perda de quadro o QVPC4 é o melhor esquema de quantização para codecs a baixas taxas operando em redes IP. Entretanto, as diferenças são pequenas quando comparadas ao esquema QVPCP2. Testes de escuta informais validaram os resultados obtidos com o PESQ.

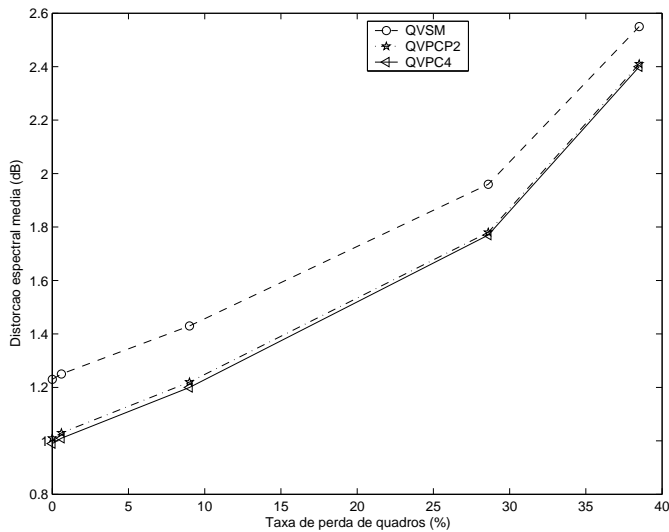


Fig. 5. Desempenho em termos de DE para TPQ.

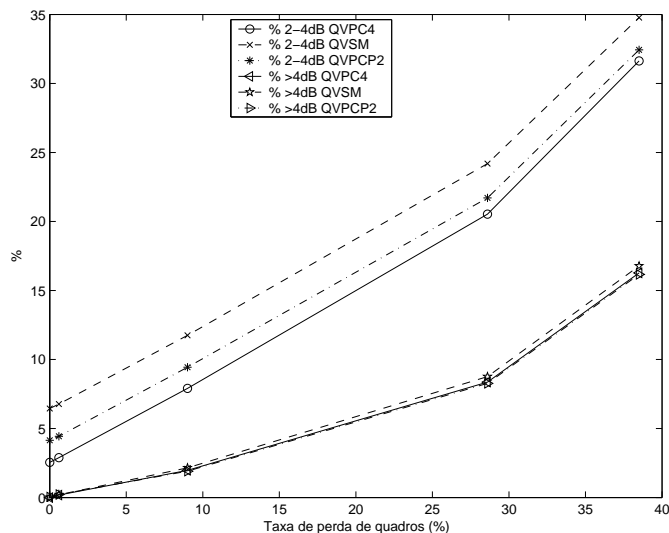


Fig. 6. Desempenho em termos de outliers para TPQ.

VI. CONCLUSÕES

Neste trabalho foi avaliada a influência da quantização das LSF sobre a qualidade de voz em codecs a baixas taxas operando em redes IP. Três esquemas de quantização vetorial foram escolhidos para as simulações. A perda de quadros em redes IP foi modelada de acordo com o Modelo de Gilbert e em condições empregadas em um trabalho recente [1]. Para combater os efeitos de perda de quadro de LSFs na qualidade da voz, a solução adotada neste artigo foi a interpolação dos parâmetros LSF. Através das simulações concluímos que o QVPC4 sempre forneceu melhor desempenho em todos os critérios de avaliação quando comparado ao QVSM e ao QVPCP2. Na avaliação da qualidade da voz utilizando o PESQ, o QVPC4 confirmou ser o melhor esquema de quantização vetorial a ser utilizado em codecs a baixas taxas operando em redes IP. Entretanto, todos os esquemas testados com o PESQ mostraram um bom desempenho considerando

TABELA IV

RESULTADO DO TESTE DE QUALIDADE DA VOZ UTILIZANDO O PESQ.

TPQ(%)	QVPC4	QVPCP2	QVSM
0	2.297	2.261	2.229
0.6	2.290	2.245	2.229
9	2.252	2.235	2.227
28.6	2.186	2.178	2.158
38.5	1.938	1.815	1.861

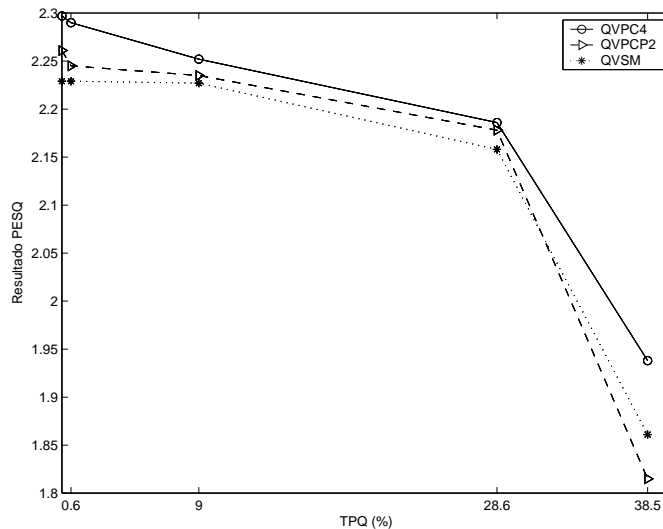


Fig. 7. Resultado do teste de qualidade da voz utilizando o PESQ versus TPQ (%).

taxas de perda de quadros de até 28,6%. Esses resultados foram corroborados por testes de escuta informais. Vale ressaltar que o pequeno ganho do QVPC4 sobre o QVPCP2 poderá desempenhar papel mais significativo caso sejam consideradas as perdas relativas aos outros parâmetros do codec.

REFERÊNCIAS

- [1] D. Quercia, L. Docio-Fernandez, C. Garcia-Mateo, L. Farinetti and J. C. De Martin, "Performance analysis of distributed speech recognition over IP networks on the AURORA database", *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 3820-3823, 2002.
- [2] J.-C. Bolot, "Characterizing end-to-end packet delay and loss in the Internet", *Proc. ACM SIGCOMM*, pp. 289-298, September 1993.
- [3] F. K. Soong e B.H. Juang, "Line spectrum pair (LSP) and speech data compression", *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 1984.
- [4] A. Gersho e R. M. Gray, *Vector quantization and signal compression*, Kluwer Academic Publishers, 1992.
- [5] K. K. Paliwal e B. S. Atal, "Efficient vector quantization of LPC parameters at 24 bits/frame", *IEEE Trans. Speech and Audio Processing*, vol. 1, no. 1, pp. 3-14, 1993.
- [6] R. C. de Lamare e A. Alcaim, "Analysis of LSF switched-predictive vector quantisers", *Proc. International Symposium on Signal Processing and its Applications*, Kuala-Lumpur, Malaysia, 2001.
- [7] R. C. de Lamare e A. Alcaim, "Noisy channel performance of LSF switched-predictive vector quantisers", *Proc. IEEE International Conference on Information, Communications and Signal Processing*, Singapore, 2001.
- [8] "Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs," Recommendation P.862, ITU-T, February 2001.
- [9] R. C. de Lamare e A. Alcaim, "Very low bit rate speech coding in tandem connections," *Electronics Letters*, vol. 39, no. 18, pp. 1356-1357, 2003.

- [10] R. C. de Lamare, L. M. da Silva e A. Alcaim, "Sound specific modelling and synthesis with a new postfiltering in low bit rate speech coding", *Proc. IEEE Int. Symp. on Circuits and Systems, Scottsdale, AZ, USA, 2002*.
- [11] W.P. LeBlanc, B. Battacharya, S.A. Mahmoud and V. Cupperman, "Efficient search and design procedures for robust multi-stage VQ of LPC parameters for 4 kb/s speech coding", *IEEE Trans. Speech and Audio Processing*, vol. 1, no. 4, pp. 373-385, 1993.
- [12] M. Yong, G. Davidsson and A. Gersho, "Encoding of LPC spectral parameters using switched-adaptive interframe vector prediction", *Proc. ICASSP-88*, vol. 1, pp. 402-405, New York, USA, 1988.
- [13] T. Eriksson, J. Lindén and J. Skoglund, "Exploiting Interframe Correlation in Spectral Quantization: A Study of Different Memory VQ Schemes", *Proc. ICASSP-96*, 1996.
- [14] A. McCree and J.C. De Martin, "A 1.7 KB/S Melp Coder with Improved Analysis and Quantization", *Proc. ICASSP-98*, 1998.
- [15] E. Daniel e K. Teague, "Federal standard 2.4 kbps MELP over IP", *Proc. 43rd IEEE Midwest Symp. on Circuits and Systems*, 2000.
- [16] J. Wang e J. Gibson, "Parameter interpolation to enhance the frame erasure robustness of CELP coders in packet networks", *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, 2001.
- [17] 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*, vol. 7, pp. 23-41, 1992.