

**Fred Berkowicz Borges**

**Codificadores de Voz a  
Baixas Taxas Operando em  
Ambientes Ruidosos e Redes  
IP**

**DISSERTAÇÃO DE MESTRADO**

**DEPARTAMENTO DE ENGENHARIA ELÉTRICA  
Programa de Pós-graduação em  
Engenharia Elétrica**

Rio de Janeiro  
Março de 2005

PONTIFÍCIA UNIVERSIDADE CATÓLICA  
DO RIO DE JANEIRO



**Fred Berkowicz Borges**

**Codificadores de Voz a Baixas Taxas  
Operando em Ambientes Ruidosos e Redes  
IP**

**Dissertação de Mestrado**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica da PUC-Rio

Orientador: Prof. Abraham Alcaim

Rio de Janeiro  
Março de 2005



**Fred Berkowicz Borges**

**Codificadores de Voz a Baixas  
Taxas Operando em Ambientes  
Ruidosos e Redes IP**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

**Prof. Abraham Alcaim**

Orientador

Centro de Estudos em Telecomunicações - PUC-Rio

**Prof. Fernando Gil Vianna Resende Jr.**

UFRJ

**Prof. Rodrigo Caiado de Lamare**

Centro de Estudos em Telecomunicações - PUC-Rio

**Prof. José Eugenio Leal**

Coordenador Setorial do Centro

Técnico Científico - PUC-Rio

Rio de Janeiro, 18 de março de 2005

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

### **Fred Berkowicz Borges**

Graduou-se em Engenharia Elétrica com ênfase em Telecomunicações pela Universidade do Estado do Rio de Janeiro em 2002.

#### Ficha Catalográfica

Borges, Fred Berkowicz

Codificadores de Voz a Baixas Taxas Operando em Ambientes Ruidosos e Redes IP/ Fred Berkowicz Borges; orientador: Abraham Alcaim. — Rio de Janeiro : PUC–Rio, Departamento de Engenharia Elétrica, 2005.

87 f. ; 29,7 cm

1. Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica.

Inclui referências bibliográficas.

1. Engenharia Elétrica – Teses. 2. Quantizadores vetoriais multiestágio. 3. LSF. 4. Redes IP. 5. Perda de quadros. 6. Avaliação perceptiva de qualidade da voz (PESQ). 7. Wavelet Denoising I. Alcaim, Abraham. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. III. Título.

CDD: 621.3

## Agradecimentos

Agradeço a Deus.

Ao Professor Abraham Alcaim pela orientação no desenvolvimento deste trabalho.

Aos professores José Mauro, Weiler, João Célio, Boisson e Grivet pelo aprendizado proporcionado.

Aos funcionários do Departamento de Engenharia Elétrica, em especial à Alcina, pelo apoio.

Ao Cnpq e à PUC-Rio, pelos auxílios concedidos.

Aos meus pais Edi e Kátia e a minha irmã Alinne, pelo incentivo e por estar ao meu lado sempre, incentivando e torcendo por minhas vitórias.

A Erika e a sua família Ricardo, Moema e Flávia, pela compreensão e apoio dado nos momentos difíceis.

Aos amigos da PUC-Rio: Aureo, José Antônio, Juan, Kiuza, Marcelle, Marcio, Leonardo Brega, Renato e Vladimir pela ajuda e companherismo.

E um agradecimento especial aos amigos Danilo, Tiago, Fabian e Rodrigo pela grande contribuição dada a esta dissertação.

## Resumo

Borges, Fred Berkowicz; Alcaim, Abraham. **Codificadores de Voz a Baixas Taxas Operando em Ambientes Ruidosos e Redes IP**. Rio de Janeiro, 2005. 87p. Dissertação de Mestrado — Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Este trabalho examina o impacto da quantização vetorial das LSFs sobre a qualidade de voz em codecs a baixas taxas operando em redes IP e em diversos ambientes ruidosos. 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. Ainda neste trabalho, foi avaliada a qualidade da voz codificada após a utilização de uma técnica de supressão de ruído baseada em transformadas wavelets (Wavelet Denoising).

## Palavras-chave

Quantizadores vetoriais multiestágio; LSF; redes IP; perda de quadros; avaliação perceptiva de qualidade da voz (PESQ); Wavelet Denoising.

## Abstract

Borges, Fred Berkowicz; Alcaim, Abraham. **Low Rate Codecs Operating in Noisy Environment and IP Networks.** Rio de Janeiro, 2005. 87p. MSc. Dissertation — Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

This work 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. In this work, we also evaluated the quality of the coded speech after employing Wavelet Denoising.

## Keywords

Multistage vector quantisation; LSF; IP networks; frame loss; perceptual evaluation of speech quality (PESQ); Wavelet Denoising.

## Conteúdo

1	Introdução	<b>12</b>
1.1	Motivação e Objetivos	12
1.2	Organização da Dissertação	13
2	Modelo do Sistema	<b>15</b>
2.1	Análise LPC e Diferentes Esquemas de Quantização das LSFs	16
2.2	Descrição do Codec	23
2.3	Transmissão e Recuperação dos Parâmetros LSF em Redes com Perdas de Pacotes	27
3	Análise de Quantizadores Vetoriais das LSF em Codecs de Voz em Ambientes Ruidosos e Redes IP	<b>31</b>
3.1	Condições das Simulações	32
3.2	Análise dos Esquemas de QV das LSFs em Ambientes Não-Ruidosos	34
3.3	Análise dos Esquemas de QV das LSF em Ambientes Ruidosos	37
3.4	Conclusões	59
4	Supressão de Ruído Baseada em Wavelets	<b>60</b>
4.1	Uma Introdução à Teoria de Wavelets	61
4.2	Supressão de Ruído Usando Wavelet	63
4.3	Qualidade da Voz Utilizando Wavelet Denoising Antes da Codificação	69
4.4	Conclusões	80
5	Conclusões e Sugestões para Trabalhos Futuros	<b>81</b>
5.1	Resumo da Dissertação	81
5.2	Sugestões para Trabalhos Futuros	82
	Referências Bibliográficas	<b>83</b>

## Lista de Figuras

2.1	Diagrama ilustrativo do sistema estudado.	16
2.2	Diagrama ilustrando a interpolação dos parâmetros LSF.	18
2.3	Desempenho dos QVs em termos da DE média para diferentes taxas de bits.	22
2.4	Diagrama em blocos do codec.	26
2.5	Sinal de voz digitalizado.	27
2.6	Diagrama em blocos do sistema de transmissão e recepção de pacotes IP.	28
2.7	Modelo de Gilbert.	29
3.1	Modelo básico do PESQ.	33
3.2	Comprimento da rajada de perda de pacotes para os parâmetros de rede apresentados na Tabela 3.1.	34
3.3	Desempenho em termos de DE para TPQ em ambientes não ruidosos.	35
3.4	Desempenho em termos de <i>outliers</i> para TPQ em ambientes não ruidosos.	36
3.5	Resultado do teste de qualidade da voz utilizando o PESQ <i>versus</i> TPQ (%) em ambientes não ruidosos.	37
3.6	Desempenho em termos de DE em diferentes RSR (Ruído de fábrica) e TPQ = 0%.	38
3.7	Desempenho em termos de DE em diferentes RSR (Ruído de falatório) e TPQ = 0%.	39
3.8	Desempenho em termos de DE em diferentes RSR (Ruído branco) e TPQ = 0%.	39
3.9	Desempenho em termos de <i>outliers</i> em diferentes RSR (Ruído de fábrica) e TPQ = 0%.	40
3.10	Desempenho em termos de <i>outliers</i> em diferentes RSR (Ruído de falatório) e TPQ = 0%.	41
3.11	Desempenho em termos de <i>outliers</i> em diferentes RSR (Ruído de branco) e TPQ = 0%.	41
3.12	Desempenho em termos de DE em diferentes RSR (Ruído de fábrica) e TPQ = 9%.	42
3.13	Desempenho em termos de DE em diferentes RSR (Ruído de falatório) e TPQ = 9%.	43
3.14	Desempenho em termos de DE em diferentes RSR (Ruído de branco) e TPQ = 9%.	43
3.15	Desempenho em termos de <i>outliers</i> em diferentes RSR (Ruído de fábrica) e TPQ = 9%.	44
3.16	Desempenho em termos de <i>outliers</i> em diferentes RSR (Ruído de falatório) e TPQ = 9%.	45
3.17	Desempenho em termos de <i>outliers</i> em diferentes RSR (Ruído de branco) e TPQ = 9%.	45

3.18	Desempenho em termos de DE em diferentes RSR (Ruído de fábrica) e TPQ = 28,6%.	46
3.19	Desempenho em termos de DE em diferentes RSR (Ruído de falatório) e TPQ = 28,6%.	47
3.20	Desempenho em termos de DE em diferentes RSR (Ruído de branco) e TPQ = 28,6%.	47
3.21	Desempenho em termos de <i>outliers</i> em diferentes RSR (Ruído de fábrica) e TPQ = 28,6%.	48
3.22	Desempenho em termos de <i>outliers</i> em diferentes RSR (Ruído de falatório) e TPQ = 28,6%.	49
3.23	Desempenho em termos de <i>outliers</i> em diferentes RSR (Ruído de branco) e TPQ = 28,6%.	49
3.24	Desempenho em termos de DE em diferentes RSR (Ruído de fábrica) e TPQ = 38,5%.	50
3.25	Desempenho em termos de DE em diferentes RSR (Ruído de falatório) e TPQ = 38,5%.	51
3.26	Desempenho em termos de DE em diferentes RSR (Ruído de branco) e TPQ = 38,6%.	51
3.27	Desempenho em termos de <i>outliers</i> em diferentes RSR (Ruído de fábrica) e TPQ = 38,5%.	52
3.28	Desempenho em termos de <i>outliers</i> em diferentes RSR (Ruído de falatório) e TPQ = 38,5%.	53
3.29	Desempenho em termos de <i>outliers</i> em diferentes RSR (Ruído de branco) e TPQ = 38,5%.	53
3.30	Resultado do PESQ em vozes não codificadas em ambientes ruidosos e sem perda de quadros.	55
3.31	Resultado do PESQ com vozes codificadas em ambientes ruidosos utilizando o QVPC4 e sem perda de quadros.	56
3.32	Resultado do PESQ para vozes codificadas em presença de ruído de fábrica utilizando QVPC4 e com perda de quadros.	59
4.1	Comparação entre uma wavelet (db10) e uma senóide	61
4.2	Exemplo de árvore de decomposição <i>wavelet</i> (DWT de 3 níveis de um sinal).	63
4.3	Árvore de decomposição por pacotes <i>wavelet</i> .	63
4.4	Funções de Limiar: (a) <i>Hard-Thresholding</i> e (b) <i>Soft-Thresholding</i> .	65
4.5	Transformada wavelet de 5 níveis para (a) Ruído Branco e (b) Ruído Colorido.	67
4.6	Diagrama ilustrativo do sistema com <i>wavelet denoising</i> .	69
4.7	Comparação da voz com ruído de falatório utilizando <i>wavelet denoising</i> e não utilizando.	72
4.8	Comparação da voz com ruído de fábrica utilizando <i>wavelet denoising</i> e não utilizando.	72
4.9	Comparação da voz com ruído branco utilizando <i>wavelet denoising</i> e não utilizando.	73
4.10	Resultado PESQ para voz codificada com ruído de falatório utilizando <i>wavelet denoising</i> e não utilizando.	75

4.11	Resultado PESQ para voz codificada com ruído de fábrica utilizando <i>wavelet denoising</i> e não utilizando.	75
4.12	Resultado PESQ para voz codificada com ruído branco utilizando <i>wavelet denoising</i> e não utilizando.	76

## Lista de Tabelas

2.1	Esquemas de QVPC.	21
2.2	Alocação de bits/estágio dos esquemas de QV com 4 estágios.	22
2.3	Alocação de Bits	26
3.1	Parâmetros do Modelo de Gilbert usado para simular as condições da Rede com perda de quadros.	34
3.2	Resultado do teste de qualidade da voz utilizando o PESQ.	36
3.3	Resultado do teste de qualidade da voz em presença de ruído e com TPQ = 0% utilizando o PESQ.	56
3.4	Resultado do teste de qualidade da voz em presença de ruído com TPQ = 9% utilizando o PESQ.	57
3.5	Resultado do teste de qualidade da voz em presença de ruído com TPQ = 28,6% utilizando o PESQ.	58
3.6	Resultado do teste de qualidade da voz em presença de ruído com TPQ = 38,5% utilizando o PESQ.	58
4.1	Tabela comparativa do resultado PESQ, para vozes não-codificadas, em ambientes ruidosos, com e sem a utilização de wavelet denoising.	71
4.2	Tabela comparativa do resultado PESQ, para vozes codificadas com TPQ = 0%, em ambientes ruidosos, com e sem a utilização de wavelet denoising.	74
4.3	Tabela comparativa do resultado PESQ, para vozes codificadas com TPQ = 9%, em ambientes ruidosos, com e sem a utilização de wavelet denoising.	77
4.4	Tabela comparativa do resultado PESQ, para vozes codificadas com TPQ = 28,6%, em ambientes ruidosos, com e sem a utilização de wavelet denoising.	77
4.5	Tabela comparativa do resultado PESQ, para vozes codificadas com TPQ = 38,5%, em ambientes ruidosos, com e sem a utilização de wavelet denoising.	78
4.6	Comparação A/B para vozes com ruído de falatório.	79
4.7	Comparação A/B para vozes com ruído branco.	79