

MINISTÉRIO DA EDUCAÇÃO E DO DESPORTO  
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
Escola de Engenharia  
Programa de Pós-Graduação em Engenharia de Minas, Metalúrgica e de Materiais-PPGEM

ANÁLISE E PROCESSAMENTO DE SINAIS DE VOZ DISFÔNICA ATRAVÉS DA  
TRANSFORMADA WAVELET DISCRETA

Eng. Adalberto Schuck Júnior, M.Sc.

Tese para obtenção do título de Doutor  
em Engenharia

Porto Alegre  
1998

ESCOLA DE ENGENHARIA  
BIBLIOTECA

MINISTÉRIO DA EDUCAÇÃO E DO DESPORTO  
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
Escola de Engenharia  
Programa de Pós-Graduação em Engenharia de Minas, Metalúrgica e de Materiais-PPGEM

ANÁLISE E PROCESSAMENTO DE SINAIS DE VOZ DISFÔNICA ATRAVÉS DA  
TRANSFORMADA WAVELET DISCRETA

Adalberto Schuck Júnior  
Engenheiro Eletricista, Mestre em Ciências em Engenharia Biomédica.

Trabalho realizado no Departamento de Engenharia Elétrica da Escola de Engenharia da UFRGS, dentro do Programa de Pós-Graduação em Engenharia de Minas, Metalúrgica e de Materiais - PPGEM, como parte dos requisitos para a obtenção do título de Doutor em Engenharia

Área de Concentração: Instrumentação Eletro-Eletrônica

Porto Alegre  
1998

Esta Tese foi julgada adequada para obtenção do título de Doutor em Engenharia, área de concentração em Instrumentação Eletro-Eletrônica, e aprovada em sua forma final, pelo Orientador e pela Banca Examinadora do Curso de Pós-Graduação.

Orientador: Prof. Dr. Marco Túlio M. B. de Vilhena

Banca Examinadora:

Prof. Dr. Rui Seara, LINSE, EEL, UFSC.

Prof. Dr. Milton Antônio Zaro, PROMEC, UFRGS.

Prof. Dr. Dario F. G. de Azevedo , Depto. Eng. Elétrica, PUC-RS.

Prof. Dr. Alwin Elbern, PPGEM, UFRGS.

Prof. Dr. Luís Frederico Pinheiro Dick  
Coordenador do PPGEM

*Dedico este trabalho aos meus pais, cujo amor, paciência, suporte, incentivo e educação foram imprescindíveis em toda minha senda.*

## AGRADECIMENTOS

Em uma tese de doutorado, o número de pessoas que colaboram de forma direta ou indireta costuma ser muito grande. Assim o foi neste trabalho. Portanto, de maneira mais geral, gostaria de agradecer a todos que de alguma forma auxiliaram ou acrescentaram algo a este trabalho. Mais especificamente, gostaria de agradecer nominalmente :

ao meu orientador, por ter conduzido este trabalho com maestria, tendo tido paciência com minhas deficiências e me apontando o caminho certo toda vez que surgia uma nebulosidade, se tornando para mim mais do que apenas um orientador, mas um mentor para a vida e um exemplo a ser seguido;

ao Prof. Dr. Milton Zaro, por ter sugerido este trabalho e por ter sido tão solícito toda vez que necessitei, colocando à disposição seu laboratório e pessoal e por sempre ter sido o amigo que é com seus conselhos e sabedoria;

ao Eng. Luciano Ricardo Maciel Silva, por ter realizado a impressionante façanha (embora coisa comum para ele) de programar, conferir, sugerir e auxiliar na interpretação dos diversos índices usados, mostrando-me ainda que a genialidade está associada com a simplicidade, e me fazendo sentir orgulho em ter achado um amigo assim. Estarei em eterna gratidão;

ao Dr. Nédio Steffen, por ter aberto seu consultório a mim nos 2 anos em que estive lá na etapa de coleta de dados, e por ter dado todo o suporte técnico relacionado com a parte de clínica otorinolaringológica. Foi uma viagem impressionante à um mundo totalmente novo;

a Dra. Adriana V. Feijó, por ter me ensinado os rudimentos da análise qualitativa e me apresentado o maravilhoso campo da fonoaudiologia, além de ter sido uma amiga sem igual;

ao Prof. Dr. David M. Simpson, por ter me iniciado no campo do processamento digital de sinais, em particular, os biomédicos, tornando simples de compreender algo que não é nada trivial. E ainda por ter dado suporte a um ex-aluno que nem orientado seu era, sem nenhum interesse outro;

a todos (realmente) colegas do IEE, mas em especial aos engenheiros Cristiano Leuckert, Marcos Zuccoloto, Julio C. V. dos Santos e Marcelo Negeiros, por terem tido a paciência de me "agüentar" e ainda sacrificarem o precioso tempo de suas próprias teses e projetos para me auxiliar na minha própria tese, sempre que foi preciso.

Muitas outras pessoas foram importantes no desenrolar desta tese, mas não haveria espaço suficiente para citar a todos. Assim, me perdoem os que sabem que me auxiliaram, mas não apareceram na lista acima e considerem-se como participantes deste tópico. Obrigado a todos, inclusive aos que, apesar dos esforços, não conseguiram barrar este trabalho.

## SUMÁRIO

LISTA DE FIGURAS.....	VII
LISTA DE TABELAS.....	XXII
RESUMO.....	XXIII
ABSTRACT.....	XXIV
1 INTRODUÇÃO.....	1
2 CONCEITOS BÁSICOS E REVISÃO DE LITERATURA.....	4
2.1-Fisiologia e formação da voz.....	4
2.2-Patologias da pregas vocais.....	12
2.3-Análise acústica da voz patológica e estado da arte.....	15
2.3.1- Análise acústica da voz patológica.....	15
2.3.2- Revisão de literatura e estado da arte.....	16
3 FUNDAMENTOS TEÓRICOS.....	33
3.1-A transformada Wavelet contínua (CWT).....	33
3.2-A transformada Wavelet discreta (DWT).....	42
3.3-A análise Multi-Resolução.....	44
4 MATERIAIS E MÉTODOS.....	52
4.1-Avaliação e gravação das vozes normais e patológicas.....	52
4.2-O equipamento utilizado.....	53
4.3-Programas computacionais utilizados.....	60
4.4-Métodos Utilizados.....	61
5 RESULTADOS OBTIDOS.....	66
6 DISCUSSÃO DOS RESULTADOS E CONCLUSÕES.....	96
REFERÊNCIAS BIBLIOGRÁFICAS.....	103
ANEXO 1: Sinais e suas DWTs dos pacientes do primeiro procedimento de aquisição.....	111
ANEXO 2: FFTs dos sinais dos pacientes, para o primeiro procedimento de aquisição.....	134
ANEXO 3: Sinais e suas DWTs dos pacientes do segundo procedimento de aquisição.....	166
ANEXO 4: Tabela com valores dos índices calculados.....	186
ANEXO 5: Tabela com valores dos índices calculados para os pacientes do segundo procedimento.....	187
ANEXO 6: O coeficiente de correlação linear.....	188

## LISTA DE FIGURAS

Figura 2.1. Diagrama mostrando os 3 sub-sistemas que compõe o sistema fonador. Em (a) vê-se o sistema supra-laríngeo em (b) a laringe e em (c) o sistema respiratório (extraído de Kent & Read, 1992).....	5
Figura 2.2. Diagrama esquemático que mostra as partes do sistema respiratório e a laringe (extraído de Boone & McFarlane, 1994).....	6
Figura 2.3. Diagrama que mostra as partes que compõe a laringe. Em (a) tem-se a vista lateral e em (b) tem-se a vista superior da laringe (extraído de Kent e Read, 1992).....	7
Figura 2.4. Esquema mostrando 1 ciclo de vibração das pregas vocais. Em (a) tem-se a representação da vista frontal e em (b) a representação da vista superior (extraído de Isshiki, 1989).....	8
Figura 2.5. Gráfico do (a) modelo do tipo de som glotal e (b) seu espectro de Fourier.....	9
Figura 2.6: Corte transversal de uma prega vocal, mostrando uma representação de suas diversas partes (extraído de Isshiki, 1989).....	9
Figura 2.7. Representação das partes do sistema articulatório , extraído de Kent e Read, 1992. ....	11
Figura 2.8. Esquema gráfico, no domínio freqüência, do modelo de formação de sons vocálicos por filtragem do som glotal.....	12
Figura 3.1: Gráfico de um sinal senoidal de freqüência variável e amortecido e seu espectrograma. A janela temporal é de 64ms (extraído de Bentley e McDonnell, 1994).....	35
Figura 3.2: Representação das células de resolução $\Delta t.\Delta f$ na STFT.....	36
Figura 3.3: Gráficos das funções de base: (a) de diferentes freqüências para a STFT, (b) diferentes dilatações para a CWT. Nota-se que em a) o aumento da freqüência não modifica a duração da função.....	40
Figura 3.4: Representação das células de resolução $\Delta t.\Delta(1/s)$ na Transformada Wavelet.....	41
Figura 3.5: Gráfico de (a) um sinal senoidal de freqüência variável e amortecido e, (b) seu escalograma (extraído de Bentley e McDonnell, 1994).....	41
Figura 3.6: Representação dos espaços de funções para "m" variando de 0 a 2.....	45
Figura 3.7: Representação dos espaços de funções $V_0$ e $V_1$ . A área hachurada representa o espaço cuja base é $\psi_{m,k}(t)$ .....	47
Figura 3.8: Representação do esquema de codificação em sub-bandas ("Sub-Band Coding").....	50

Figura 4.1: Diagrama de bloco do sistema usado nas gravações.....	54
Figura 4.2: Diagrama de blocos do sistema usado para o primeiro procedimento de aquisição. A placa DSP se encontra dentro do micro hospedeiro.....	55
Figura 4.3: Diagrama de blocos do sistema usado para o segundo procedimento de aquisição. A placa DSP se encontra dentro do micro hospedeiro, embora não seja usada.....	55
Figura 4.4: Sistema do Laboratórios de processamento de sinais e imagens (LaPSI-UFRGS) usado para a reprodução do sinal gravado.....	56
Figura 4.5: Diagrama esquemático do filtro passa-baixas analógico do tipo Butterworth de 10 <sup>a</sup> ordem. A frequência de corte calculada é de 11300Hz.....	57
Figura 4.6: Vista da placa DSP usada no pré-processamento dos sinais de voz.....	59
Figura 4.7: Sistema de conversão A/D dos sinais de voz.....	59
Figura 4.8: Representação da (a)Scaling Function e (b)Wavelet para a base de Haar.....	61
Figura 4.5: Representação do módulo das funções de transferência dos filtros digitais equivalentes associados aos coeficientes da DWT $c_0$ e $d_0$ .....	62
Figura 5.1: Gráfico da estimação do espectro de potência do ruído branco gravado e reproduzido sem Dolby, em escala linear.....	67
Figura 5.2: Gráfico da estimação do espectro de potência do ruído branco gravado e reproduzido com Dolby, em escala linear.....	67
Figura 5.3: Gráfico da estimação do espectro de potência do ruído branco gravado e reproduzido sem Dolby, em escala logarítmica.....	68
Figura 5.4: Gráfico da estimação do espectro de potência do ruído branco gravado e reproduzido com Dolby, em escala logarítmica.....	68
Figura 5.5: Gráfico da diferença entre as estimações do espectro de potência do ruído branco gravado e reproduzido sem Dolby e com Dolby, em escala de decibéis.....	69
Figura 5.6: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC10.....	69
Figura 5.7: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC15.....	70
Figura 5.8: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC26.....	70
Figura 5.9: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC29.....	71



Figura 5.10: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC35.....	71
Figura 5.11: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC45.....	72
Figura 5.12: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC48.....	72
Figura 5.13: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC49.....	73
Figura 5.14: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC55.....	73
Figura 5.15: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC56.....	74
Figura 5.16: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC59.....	74
Figura 5.17: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC62.....	75
Figura 5.18: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC63.....	75
Figura 5.19: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC67.....	76
Figura 5.20: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC68.....	76
Figura 5.21: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC74.....	77
Figura 5.22: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC10, com segmentação por janela retangular.....	77
Figura 5.23: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC15, com segmentação por janela retangular.....	78
Figura 5.24: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC26, com segmentação por janela retangular.....	78

- Figura 5.25: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC29, com segmentação por janela retangular.....79
- Figura 5.26: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC35, com segmentação por janela retangular.....79
- Figura 5.27: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC45, com segmentação por janela retangular.....80
- Figura 5.28: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC48, com segmentação por janela retangular.....80
- Figura 5.29: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC49, com segmentação por janela retangular.....81
- Figura 5.30: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC55, com segmentação por janela retangular.....81
- Figura 5.31: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC56, com segmentação por janela retangular.....82
- Figura 5.32: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC59, com segmentação por janela retangular.....82
- Figura 5.33: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC62, com segmentação por janela retangular.....83
- Figura 5.34: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC63, com segmentação por janela retangular.....83
- Figura 5.35: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC67, com segmentação por janela retangular.....84
- Figura 5.36: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC68, com segmentação por janela retangular.....84

Figura 5.37: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC74, com segmentação por janela retangular.....	85
Figura 5.38: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC10, com segmentação por janela Hamming.....	85
Figura 5.39: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC15, com segmentação por janela Hamming.....	86
Figura 5.40: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC26, com segmentação por janela Hamming.....	86
Figura 5.41: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC29, com segmentação por janela Hamming.....	87
Figura 5.42: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC35, com segmentação por janela Hamming.....	87
Figura 5.43: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC45, com segmentação por janela Hamming.....	88
Figura 5.44: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC48, com segmentação por janela Hamming.....	88
Figura 5.45: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC49, com segmentação por janela Hamming.....	89
Figura 5.46: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC55, com segmentação por janela Hamming.....	89
Figura 5.47: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC56, com segmentação por janela Hamming.....	90
Figura 5.48: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC59, com segmentação por janela Hamming.....	90

Figura 5.49: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC62, com segmentação por janela Hamming.....	91
Figura 5.50: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC63, com segmentação por janela Hamming.....	91
Figura 5.51: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC67, com segmentação por janela Hamming.....	92
Figura 5.51: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC68, com segmentação por janela Hamming.....	92
Figura 5.53: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC74, com segmentação por janela Hamming.....	93
Figura 5.54: Sinal aleatório gerado pelo MATHLAB e os coeficientes dos 6 níveis da DWT para a base de Haar.....	93
Figura 5.55: Senóide crescente de 0 a 8000Hz, gerada pelo MATHLAB e seus coeficientes dos 6 níveis da DWT, para a base Haar.....	94
Figura 5.56: Senóide crescente de 0 a 8000Hz, gerada pelo MATHLAB, segmentada pela janela Hamming e seus coeficientes dos 6 níveis da DWT, para a base Haar.....	94
Figura 5.57. Diagrama de espalhamento entre o Índice de Shoji e o índice calculado com o uso dos coeficientes $d_0$ e $c_0$ .....	95
Figura 5.58: Diagrama de espalhamento entre o índice Br. e o índice que usa a segunda derivada do coeficiente $d_0$ .....	96
Figura A1.1: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC01.....	111
Figura A1.2: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC02.....	111
Figura A1.3: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC04.....	112
Figura A1.4: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC05.....	112
Figura A1.5: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC06.....	113

Figura A1.6: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC08.....	113
Figura A1.7: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC09.....	114
Figura A1.8: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC11.....	114
Figura A1.9: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC13.....	115
Figura A1.10: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC14.....	115
Figura A1.11: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC16.....	116
Figura A1.12: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC17.....	116
Figura A1.13: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC18.....	117
Figura A1.14: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC19.....	117
Figura A1.15: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC21.....	118
Figura A1.16: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC22.....	118
Figura A1.17: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC23.....	119
Figura A1.18: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC24.....	119
Figura A1.19: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC27.....	120
Figura A1.20: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC28.....	120
Figura A1.21: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC30.....	121
Figura A1.22: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC31.....	121

Figura A1.23: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC32.....	122
Figura A1.24: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC34.....	122
Figura A1.25: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC36.....	123
Figura A1.26: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC41.....	123
Figura A1.27: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC42.....	124
Figura A1.28: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC43.....	124
Figura A1.29: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC44.....	125
Figura A1.30: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC46.....	125
Figura A1.31: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC47.....	126
Figura A1.32: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC50.....	126
Figura A1.33: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC51.....	127
Figura A1.34: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC53.....	127
Figura A1.35: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC54.....	128
Figura A1.36: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC57.....	128
Figura A1.37: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC58.....	129
Figura A1.38: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC60.....	129

Figura A1.39: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC61.....	130
Figura A1.40: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC64.....	130
Figura A1.41: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC66.....	131
Figura A1.42: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC69.....	131
Figura A1.43: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC70.....	132
Figura A1.44: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC71.....	132
Figura A1.45: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC72.....	133
Figura A1.46: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC73.....	133
Figura A2.1: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC01.....	134
Figura A2.2: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC02.....	134
Figura A2.3: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC04.....	135
Figura A2.4: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC05.....	135
Figura A2.5: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC06.....	136
Figura A2.6: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC08.....	136
Figura A2.7: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC09.....	137
Figura A2.8: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC10.....	137
Figura A2.9: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC11.....	138

Figura A2.10: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC13.....	138
Figura A2.11: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC14.....	139
Figura A2.12: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC15.....	139
Figura A2.13: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC16.....	140
Figura A2.14: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC17.....	140
Figura A2.15: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC18.....	141
Figura A2.16: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC19.....	141
Figura A2.17: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC21.....	142
Figura A2.18: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC22.....	142
Figura A2.19: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC23.....	143
Figura A2.20: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC24.....	143
Figura A2.21: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC26.....	144
Figura A2.22: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC27.....	144
Figura A2.23: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC28.....	145
Figura A2.24: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC29.....	145
Figura A2.25: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC30.....	146



Figura A2.26: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC31.....	146
Figura A2.27: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC32.....	147
Figura A2.28: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC34.....	147
Figura A2.29: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC35.....	148
Figura A2.30: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC36.....	148
Figura A2.31: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC37.....	149
Figura A2.32: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC41.....	149
Figura A2.33: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC42.....	150
Figura A2.34: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC43.....	150
Figura A2.35: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC44.....	151
Figura A2.36: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC45.....	151
Figura A2.37: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC46.....	152
Figura A2.38: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC47.....	152
Figura A2.39: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC48.....	153
Figura A2.40: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC49.....	153
Figura A2.41: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC50.....	154
Figura A2.42: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC51.....	154

Figura A2.43: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC53.....	155
Figura A2.44: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC54.....	155
Figura A2.45: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC55.....	156
Figura A2.46: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC56.....	156
Figura A2.47: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC57.....	157
Figura A2.48: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC58.....	157
Figura A2.49: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC59.....	158
Figura A2.50: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC60.....	158
Figura A2.51: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC61.....	159
Figura A2.52: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC62.....	159
Figura A2.53: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC63.....	160
Figura A2.54: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC64.....	160
Figura A2.55: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC66.....	161
Figura A2.56: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC67.....	161
Figura A2.57: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC68.....	162
Figura A2.58: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC69.....	162

Figura A2.59: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC70.....	163
Figura A2.60: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC71.....	163
Figura A2.61: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC72.....	164
Figura A2.62: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC73.....	164
Figura A2.63: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC74.....	165
Figura A3.1: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC01.....	166
Figura A3.2: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC02.....	166
Figura A3.3: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC04.....	167
Figura A3.4: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC05.....	167
Figura A3.5: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC06.....	168
Figura A3.6: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC08.....	168
Figura A3.7: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC09.....	169
Figura A3.8: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC11.....	169
Figura A3.9: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC13.....	170
Figura A3.10: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC14.....	170
Figura A3.11: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC17.....	171
Figura A3.12: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC18.....	171

Figura A3.13: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC19.....	172
Figura A3.14: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC21.....	172
Figura A3.15: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC23.....	173
Figura A3.16: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC24.....	173
Figura A3.17: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC27.....	174
Figura A3.18: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC28.....	174
Figura A3.19: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC30.....	175
Figura A3.20: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC31.....	175
Figura A3.21: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC32.....	176
Figura A3.22: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC33.....	176
Figura A3.23: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC36.....	177
Figura A3.24: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC41.....	177
Figura A3.25: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC42.....	178
Figura A3.26: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC43.....	178
Figura A3.27: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC44.....	179
Figura A3.28: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC46.....	179

- Figura A3.29: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC47.....180
- Figura A3.30: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC50.....180
- Figura A3.31: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC51.....181
- Figura A3.32: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC54.....181
- Figura A3.33: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC58.....182
- Figura A3.34: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC60.....182
- Figura A3.35: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC61.....183
- Figura A3.36: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC64.....183
- Figura A3.37: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC66.....184
- Figura A3.38: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC69.....184
- Figura A3.39: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC70.....185
- Figura A3.40: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC72.....185

## LISTA DE TABELAS

Tabela 2.1: Principais distúrbios da voz, classificados pela sua origem (extraído de Boone e McFarlane (1994)).....	14
Tabela 2.2: Resumo das técnicas de Análise Acústica da Voz Digitalizada (Kent e Read, 1992).....	15
Tabela 2.3: Resumo dos principais índices de "jitter".....	30
Tabela 2.4: Resumos dos principais índices de "shimmer".....	31
Tabela 2.5: Resumo dos principais índices de Relação Sinal-Ruído.....	31
Tabela 2.6: Outros índices propostos na literatura.....	32
Tabela 2.7: Índices relacionados com a soproidade da voz.....	32
Tabela 4.1: Descrição dos pacientes gravados.....	52
Tabela 4.2: Bandas de frequência dos coeficientes da DWT para o primeiro procedimento (fs=25000Hz).....	63
Tabela 4.3: Bandas de frequência dos coeficientes da DWT para o primeiro procedimento (fs=50000Hz).....	63
Tabela A4.1: Valores dos índices de soproidade clássicos e propostos para os pacientes...	186
Tabela A5.1: Valores dos índices de soproidade clássicos e propostos para os pacientes do segundo procedimento de aquisição.....	186

## RESUMO

O presente trabalho apresenta um resumo da fisiologia de produção da voz humana, das patologias mais comuns da laringe e seus principais efeitos sobre o som fonado, e apresenta diversos métodos quantitativos de avaliação do som de vozes patológicas. É então proposto um novo método de avaliação da soproidade da voz, baseado na Transformada Wavelet Discreta (DWT) através da análise multi-resolução, usando como base ortogonal de decomposição a base Haar. São feitas duas aquisições por dois diferentes procedimentos, dos sinais de voz de 64 pacientes. É mostrado que é possível se obter um índice acústico para a característica soproidade da voz por intermédio da DWT. Este índice é estatisticamente correlacionado com dois outros índices existentes para soproidade, para ambos os procedimentos de aquisição. O método serve tanto para auxílio ao diagnóstico como acompanhamento dos resultados obtido por um tratamento.

## ABSTRACT

This work shows a brief review of human voice production physiology, including the most common larynx pathologies and its effects in the voice quality, and the main methods of pathologic quantitative vocal fold assessment. It is proposed a novel method of breathiness of voice characteristic evaluation, based on the Discrete Wavelet Transform, using the orthonormal Haar basis as a reconstruction basis. Two procedures of data acquisition were used for the 64 subjects voice signals. Results are obtained and statistically compared with the ones obtained by classical methods, for both acquisition procedures. This method can be an auxiliary tool for the diagnosis as well as an assessment of a specific treatment.



## 1 INTRODUÇÃO:

Dentre os profissionais da área de saúde que lidam com a voz humana, tais como os otorrinolaringologistas e fonoaudiologistas, faz-se cada vez mais necessário o desenvolvimento de ferramentas que auxiliem no diagnóstico médico das patologias da laringe e das pregas vocais. O primeiro indicador de disfunção da laringe é a observação por parte do paciente de alteração na sua qualidade de voz, com a presença ou não de dor. Com a procura de auxílio ao profissional de saúde, é feita a identificação da patologia e prescrito o devido tratamento. Inicialmente o profissional de saúde tenta identificar estas patologias também pela qualidade sonora apresentada pelo paciente. Esta impressão sonora é classificada de modo qualitativo e o profissional treinado tem a capacidade de formar uma imagem prévia de alguns tipos de patologia e alterações apresentados pelo paciente, com base apenas na voz deste (Isshiki, 1989, Boone e McFarlane, 1994; Colton e Casper, 1996). A partir de então, são feitos exames clínicos, como por exemplo a videolaringoscopia direta e a videoestrobolaringoscopia (Kuhl, 1982, Hirano e Bless, 1993), que identificam e classificam rigorosamente estas patologias. O problema dos métodos qualitativos acima mencionados é que são subjetivos, podendo variar de médico para médico pela sua percepção auditiva e base de conhecimentos.

Vários métodos foram propostos por diversos autores para, através do sinal acústico da voz, classificar de maneira quantitativa a voz normal da voz patológica e, em alguns casos, tentar identificar e correlacionar previamente o tipo de patologia apresentada pelo paciente (Baken, 1987). A revisão bibliográfica destes métodos será apresentada na seção 2.3.2.

A transformada Wavelet, uma ferramenta matemática surgida em meados da década de 80 para análise de sinais estocásticos não-estacionários (Rioul e Vetterli, 1991; Meyer, 1993; Strang e Nguyen, 1996; Sheng, 1996), vem sendo aplicada em diversos campos da análise e processamento de sinais, substituindo a transformada de Gabor (Gabor, 1946) pela vantagem de ter resolução espectral e também a resolução temporal variáveis conforme a faixa de frequência a ser analisada (Grossmann e Morlet, 1984), como será apresentado no Capítulo 3. Hoje, são diversos os usos desta transformada na área de processamento de sinais e imagens biomédicos (Aldroubi e Unser, 1996). Entretanto, com base na pesquisa bibliográfica realizada até o momento, não se tem notícia de que esta ferramenta tenha sido utilizada para a análise da voz patológica, o que caracteriza a originalidade deste trabalho em âmbito internacional.

O objetivo desta tese é investigar a possibilidade de análise e identificação da voz patológica através da Transformada Wavelet Discreta, usando a base de Haar como base de decomposição, através da abordagem de análise multi-resolução, por ser esta transformada um avanço em relação ao Espectrograma (transformada de Gabor), que é a ferramenta que mais apresenta informações sobre a voz (Rontal, Rontal e Rolnick, 1975). É mostrado que, por intermédio desta transformada, é possível se obter um índice acústico para a característica soprosidade da voz, índice este que é estatisticamente correlacionado com dois índices existentes, calculados por outros métodos, para este mesmo fim.

Para atingir o objetivo proposto, o trabalho é organizado da maneira que segue:

No Capítulo 2 são apresentados os conceitos básicos da fisiologia de produção da voz humana, das patologias da voz e suas principais características e as formas de análise acústica

da voz. É também apresentada uma revisão bibliográfica dos diversos métodos e índices propostos para a classificação e identificação da voz patológica, sendo apresentado o estado-da-arte neste campo. No Capítulo 3 é apresentada uma introdução dos principais conceitos teóricos da transformada Wavelet contínua, da transformada Wavelet discreta e seu algoritmo de cálculo operado em microcomputadores pessoais. No Capítulo 4 são apresentados os equipamentos utilizados bem como a metodologia de gravação e processamento da voz dos pacientes normais e patológicos. O programa computacional e o algoritmo utilizado para o cálculo da análise multi-resolução também são descritos neste capítulo. É também descrito de forma sucinta o método de cálculo para os índices propostos que serão usados na comparação com os índices de soproidade sugeridos pela literatura. Finalmente, é feita a descrição do método estatístico empregado para a comparação. No Capítulo 5 são mostrados os resultados obtidos. No Capítulo 6 são comentados e discutidos os resultados obtidos e por intermédio destes, são evidenciadas as conclusões a que este trabalho chegou, bem como as sugestões de continuidade para futuras investigações. Por fim, são relacionadas as referências bibliográficas utilizadas neste trabalho.

## 2 CONCEITOS BÁSICOS E REVISÃO BIBLIOGRÁFICA:

Visando uma compreensão dos fenômenos envolvidos na formação da voz e dos diversos índices utilizados para a avaliação acústica da voz, são apresentados neste capítulo os conceitos básicos da fisiologia de produção da voz humana e os possíveis distúrbios desta. Ainda neste capítulo são apresentadas as técnicas gerais para a análise acústica da voz humana. Então é apresentada uma revisão bibliográfica dos diversos métodos e índices acústicos propostos para a classificação e a identificação de voz patológica, como também apresentado o estado-da-arte neste campo.

### 2.1 FISIOLOGIA E FORMAÇÃO DA VOZ:

Neste tópico é apresentado um resumo da fisiologia da produção da voz humana. O enfoque básico será o de apresentar o estudo fisiológico dos sistemas biológicos envolvidos na produção da fala, com objetivo de se obter modelos físicos e matemáticos dos mesmos. Discutir-se-á inicialmente a divisão do sistema fonador, sendo comentado cada um dos referidos sub-sistemas. A seguir é feita uma apresentação referente ao funcionamento das cordas (ou pregas) vocais e finalmente é mostrado como o som produzido é articulado e transformado em fala.

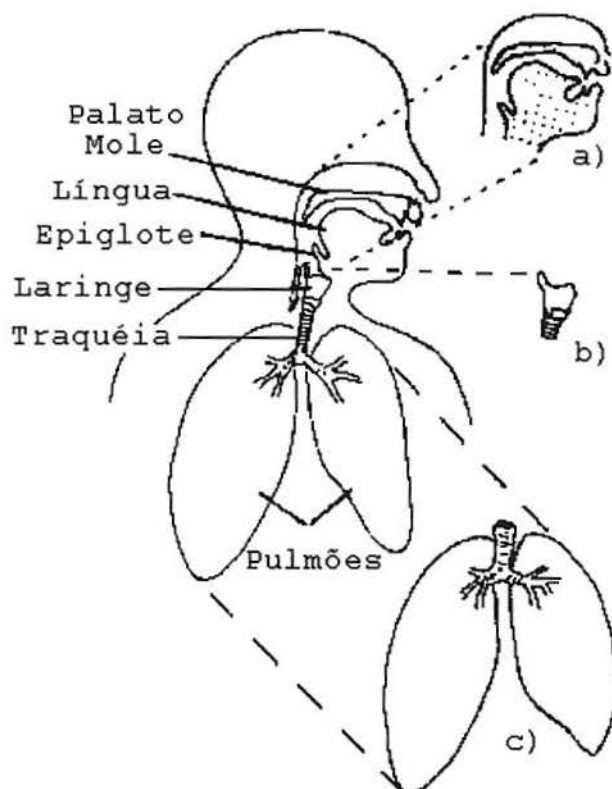
Para que a fala seja produzida, diversas estruturas orgânicas devem entrar em funcionamento coordenado. O movimento da língua, da mandíbula, o posicionamento dos lábios, a tensão das pregas vocais e a força de expiração são alguns dos diversos fatores que

entram na formação da fala. Ao conjunto de todas as estruturas envolvidas no processo de produção da fala, denomina-se sistema fonador.

O sistema fonador pode ser subdividido em 3 grandes sub-sistemas (Kent e Read, 1992; Boone e McFarlane, 1994), classificados conforme sua função fisiológica no processo de produção da voz humana:

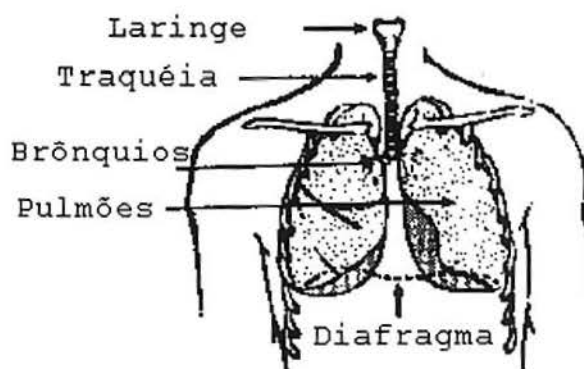
- a) Sub-sistema Respiratório;
- b) Sub-sistema Laringeal;
- c) Sub-sistema Supra-Laringeal.

A Figura 2.1 apresenta um diagrama incluindo os 3 sub-sistemas.



*Figura 2.1. Diagrama mostrando os 3 sub-sistemas que compõem o sistema fonador. Em (a) vê-se o sistema supra-laringeal; em (b) a laringe e; em (c) o sistema respiratório (extraído de Kent e Read, 1992).*

O sub-sistema Respiratório, que compreende a traquéia, os brônquios e os pulmões, é considerado a fonte de ar para a geração do som, uma vez que este é produzido na fase de expiração. Este sistema tem pequena contribuição na articulação da fala, mas seus efeitos afetam o timbre da voz e o tempo de fonação do falante (Kent e Read, 1992). Doenças tais como tuberculose e enfisema pulmonar, que reduzem a capacidade vital do indivíduo (fôlego curto), tornam o tipo de fala gerado por estes indivíduos de baixa intensidade e de curta duração. Contudo, para fins de modelagem, os efeitos de ressonância e modificação do timbre são incluídos como contribuições no som glotal, ou são incluídos na função de transferência que caracteriza o sistema supra-laríngeo (Kent e Read, 1992). A Figura 2.2 apresenta uma representação deste sub-sistema.



*Figura 2.2. Diagrama esquemático que mostra as partes do sistema respiratório e a laringe (extraído de Boone e McFarlane, 1994).*

O sub-sistema Laringeal, ou simplesmente laringe, é composto pelas cordas (ou pregas) vocais, pelo osso hióide, pela cartilagem cricóide, pelas cartilagens aritenóides, pela epiglote e pela cartilagem tireóide. Um diagrama simplificado mostrando as diversas partes que compõem o sub-sistema laringeal é apresentado na Figura 2.3 .

Este sub-sistema é o responsável pela produção dos sons de vogais (sons vocálicos) e de certas consoantes (sons consonantais vocálicos, /z/, /v/), através da vibração das pregas vocais. Durante a fase de expiração, quando do desejo da emissão de som por parte do falante,

um conjunto de músculos aduzem (aproximam) as pregas vocais até que elas se encostem uma à outra. Neste momento, o fluxo de ar ascendente proveniente dos pulmões e traquéia força as cordas vocais a vibrarem, permitindo o escape de ar e produzindo um som denominado de som glotal. Um ciclo de movimento completo das pregas vocais é mostrado na Figura 2.4 e um modelo do tipo de som gerado, bem como o módulo de seu espectro é mostrado na Figura 2.5.

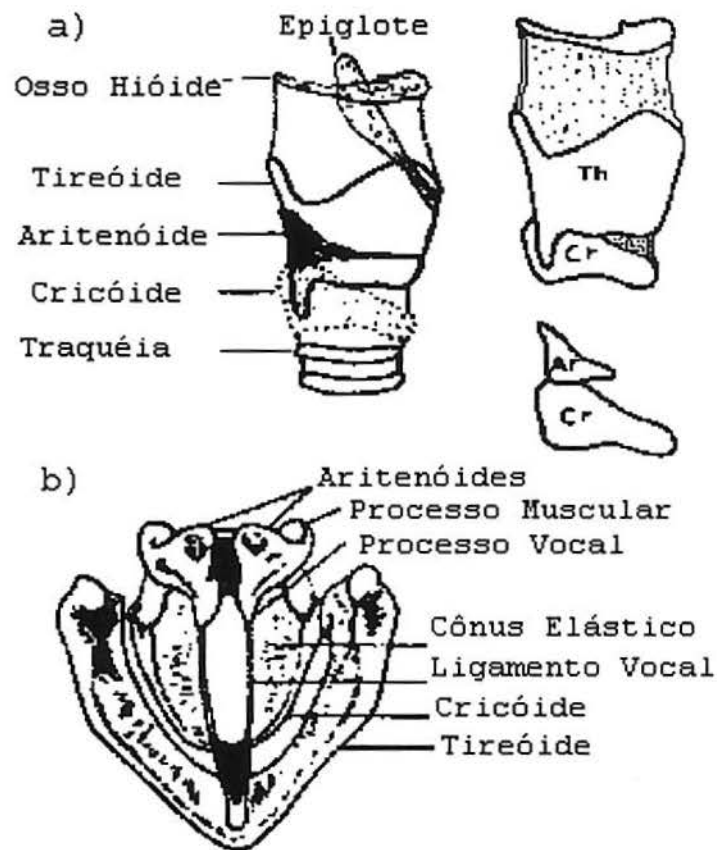


Figura 2.3. Diagrama que mostra as partes que compõem a laringe. Em (a) tem-se a vista lateral e em (b) tem-se a vista superior da laringe (extraído de Kent e Read, 1992).

É interessante observar que a força de fechamento das pregas vocais é variável conforme a vontade do falante, e pode ser tão intensa a ponto de interromper totalmente o fluxo de ar, tal como comumente acontece quando se levanta um peso excessivo, ou se faz um esforço que envolva força intensa. A força de adução das pregas é uma das principais variáveis de controle da intensidade de som emitido. Para altas intensidades, a força de

adução é aumentada, bem como a força de expiração, provocando uma maior amplitude de vibração das pregas vocais e conseqüentemente, do som glotal gerado.

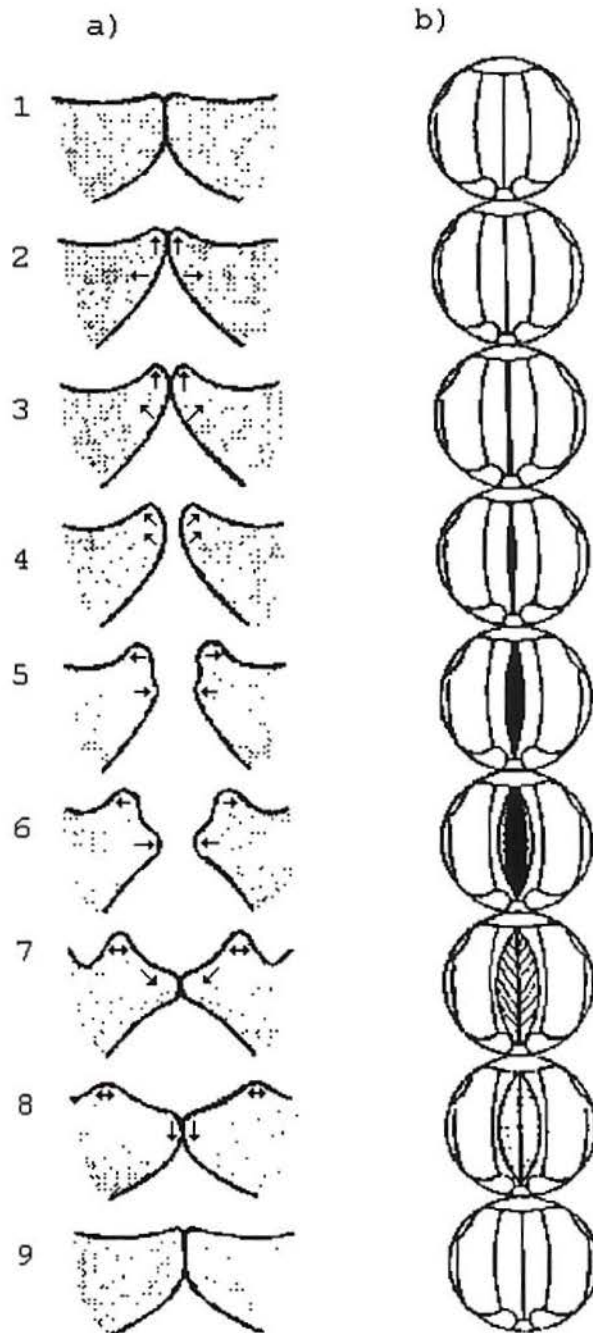


Figura 2.4. Esquema mostrando 1 ciclo de vibração das pregas vocais. Em a) tem-se a representação da vista frontal e em b), a representação da vista superior (extraído de Isshiki, 1989).



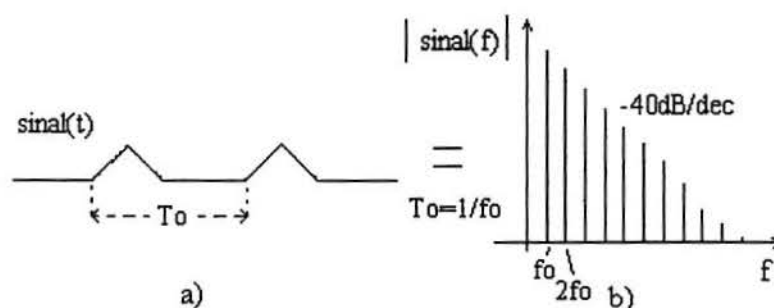


Figura 2.5: Gráfico do (a) modelo do tipo de som glotal e (b) seu respectivo módulo do espectro de frequência.

A Figura 2.6 apresenta a representação de um corte transversal de uma prega vocal, mostrando sua composição histológica básica.

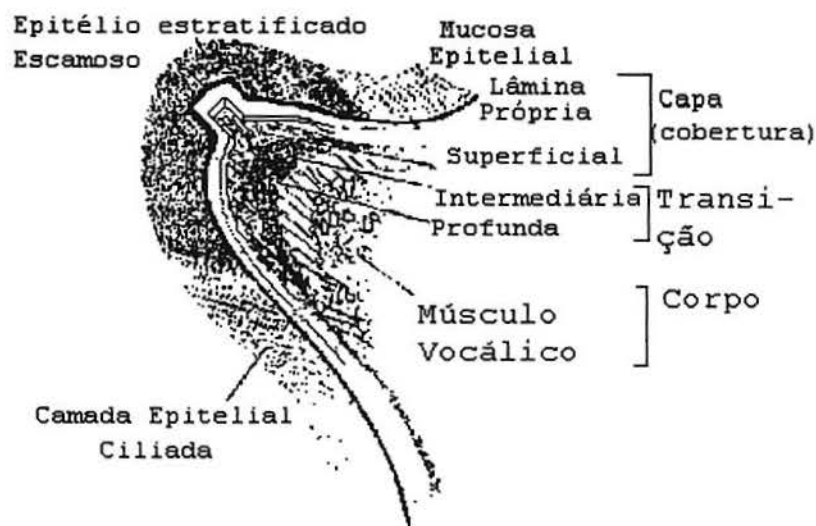


Figura 2.6: Corte transversal de uma prega vocal, mostrando uma representação de suas diversas partes (extraído de Isshiki, 1989).

O músculo vocálico é o principal responsável pela tensão longitudinal das pregas vocais. Esta tensão é a responsável pelas alterações da frequência fundamental da voz conforme a vontade, ou seja, pode-se passar de sons graves (baixa tensão) a sons agudos (alta tensão) e vice-versa. Mas esta tensão longitudinal não se relaciona com a força de aproximação das pregas vocais, a qual é executada por outros músculos.

A frequência fundamental média é de aproximadamente 125 Hz (ciclos por segundo) para o homem e de 220 Hz para mulheres. O "Pitch" é definido em Kent e Read (1992) como a percepção auditiva da frequência fundamental.

A composição das pregas vocais afeta profundamente a característica do som produzido. Para que um som possa ser considerado dentro de padrões normais, ambas as pregas devem ter massa aproximadamente igual, bem como tensão e rigidez ("stiffness") semelhantes. Assim, o padrão de vibração de ambas as cordas se assemelhará, criando um som quase periódico e harmônico, para sons vocálicos. Este tipo de som tem como característica um espectro de frequências composto por uma frequência fundamental e por seus harmônicos, que são múltiplos inteiros desta frequência fundamental (Kent e Read, 1992). A composição da prega contribui para a característica tímbrica da voz do falante, tal como a frequência fundamental, por exemplo. Pregas com mais massa geram sons de mais baixa frequência (graves), enquanto pregas com menos massa, sons de mais alta frequência (agudos).

Alterações nas pregas vocais que provoquem assimetria no padrão vibratório têm como consequência uma elevação dos níveis de harmônicos no sinal que não são múltiplos inteiros da frequência fundamental, o que caracteriza um sinal temporal mais aperiódico. Isto é percebido pelos ouvintes como um som mais desagradável, "rouco". O fumo, a ingestão de álcool e a própria idade causam alterações com o tempo da morfologia das pregas vocais, alterando o timbre de voz de uma pessoa.

O sub-sistema supra-laríngeo é também denominado de articulador, pois não produz sons vocálicos, mas modifica os sons glotais, atuando como uma cavidade ressonante variável

que altera o conteúdo harmônico dos sons glotais. Na Figura 2.7, tem-se uma representação das partes que compõem este sistema.

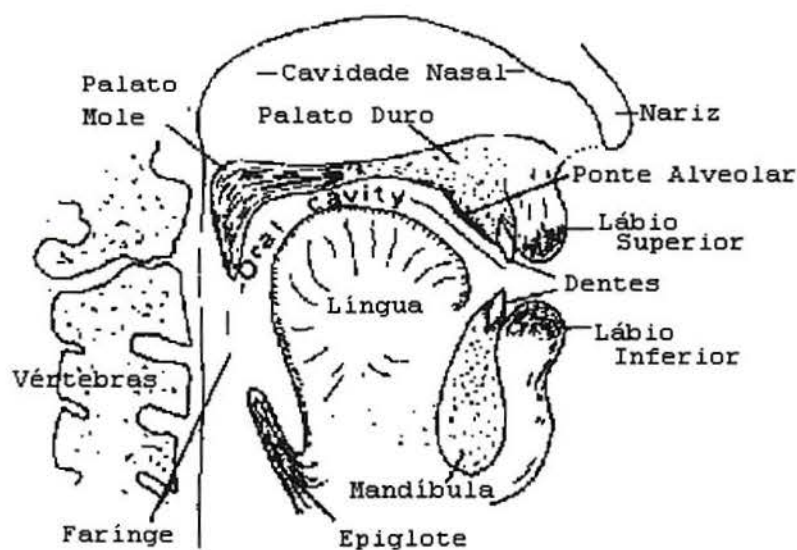


Figura 2.7. Representação das partes do sistema articulador (extraído de Kent e Read, 1992).

Note-se que, embora a cavidade ressonante não produza sons vocálicos, ela atenua certas faixas de frequências e reforça ou amplifica outras regiões de frequência do sinal glotal gerado pela fonte. Aos picos das regiões que reforçam o sinal glotal denominam-se formantes e são regiões de ressonância natural da cavidade ressonante deste sub-sistema. A alteração do timbre da voz está profundamente relacionada com estes formantes. Assim, por este motivo pode-se identificar uma vogal /a/ de uma vogal /u/, mesmo que ambas sejam fonadas com mesma frequência fundamental. Note-se que os formantes não têm relação com a frequência fundamental da voz, pois um é função das cavidades ressonantes do sistema articulador enquanto o outro é função da fonte do sinal (Kent e Read, 1992).

Em termos de modelagem, o sub-sistema supra-laríngeal pode ser representado como um filtro com uma determinada função de transferência que alterará um som simples gerado pela glote, para sons vocálicos e sons consonantais vocálicos. Conforme o teorema de fonte-

filtragem (Kent e Read, 1992), pode-se então modelar os sons vozeados da seguinte forma: tem-se uma fonte de sinal que gera um som semelhante ao som glotal, como mostrado na Figura 2.5 (a). Este sinal é então filtrado por um filtro cujas características copiam as frequências de ressonância do sistema articulador (incluídas as contribuições de ressonâncias do sistema respiratório e da saída da onda sonora pela boca). Um diagrama deste esquema pode ser visto na Figura 2.8.

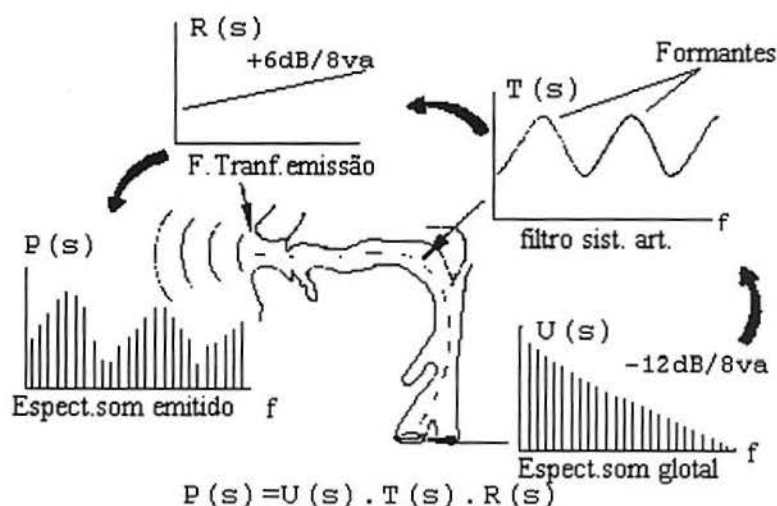


Figura 2.8: Esquema gráfico, no domínio frequência, do modelo de formação de sons vocálicos por filtragem do som glotal.

Revisões sobre a fisiologia e as teorias acústicas da produção de voz são encontradas em Kent e Read (1992), Boone e McFarlane (1994) e Isshiki (1989), dentre outros.

## 2.2-PATOLOGIAS DAS PREGAS VOCAIS:

Como foi visto na seção anterior, alterações das pregas vocais causam diretamente uma alteração da qualidade sonora da voz emitida. As maneiras de se classificar a qualidade da voz patológica são bastante variadas e não existe atualmente um consenso para esta

classificação. Tentando padronizar a avaliação perceptual da voz patológica, Isshiki (1989) sugere que a disfonia seja expressa em 5 diferentes índices ou impressões auditivas:

- a)Aspereza ("roughness");
- b)Soprosidade ("breathiness");
- c)Afonía ("asthenic") ;
- d)Intensidade ("grade") ;
- e)Hiperfuncionalidade ("strained").

Cada índice teria um sistema de graduação de 4 pontos (0=normal, 1=leve, 2=moderado, 3=intenso). Assim foi composto o sistema "GRABS", cujo uso é sugerido pelo "Japan Society of Logopedics and Phoniatics".

Outros termos são citados como indicadores subjetivos de alteração da qualidade vocal, tais como rudeza, estridência, afonia, etc... . Mas os mais usados são os anteriormente citados. Faz-se então necessário definir estes fatores. Conforme Boone e McFarlane (1994), voz áspera é aquela que apresenta tom desagradável, freqüentemente descrita como "rude, gutural, que raspa, que range". Voz soprosa é aquela que apresenta um escape de ar audível, freqüentemente acompanhada de perda da intensidade da voz. É quase sempre associada à falta de contato entre as pregas vocais, provocada por diversos fatores. Voz hiperfuncional é aquela causada pelo excesso de tensão dos mecanismos vocais tornando o som "estrangulado, apertado", sem eficiência. Conforme Isshiki (1989), os fatores mais marcantes da voz disfônica são os fatores aspereza e soprosidade .

Conforme Hirano e Bless (1993), não existe nenhuma relação consistente entre padrões vibratórios de pregas vocais normais e padrões vibratórios de pregas com uma determinada patologia, pois a vibração também depende do modo da fala e mecanismos compensatórios de fonação, bem como da extensão, grau e padrão histológico da doença. Entretanto, existe uma tendência vibratória geral oriunda do desvio nas propriedades mecânicas das pregas vocais causadas por uma dada doença.

Os distúrbios podem ser classificados como funcionais ou como orgânicos, dependendo de sua origem. Os distúrbios funcionais são aqueles que têm sua origem no mau uso ou abuso dos mecanismos vocais. Os distúrbios orgânicos têm sua origem em alterações orgânicas dos mecanismos vocais. Na Tabela 2.1 tem-se um resumo dos distúrbios conforme esta classificação.

Uma descrição mais detalhada de cada patologia, bem como as alterações causadas na qualidade da voz podem ser encontradas em Kuhl (1982), Boone e McFarlane (1994) e Colton e Casper (1996).

*Tabela 2.1: Principais distúrbios da voz, classificados pela sua origem (extraído de Boone e McFarlane, (1994)).*

Distúrbio Funcional:	Distúrbio Orgânico:
Afonia funcional	Granulomas
Diplofonia	Disartria
Disfonia espasmódica	Laringectomia
Disfonia ventricular	Papiloma
Espessamento da prega vocal	Sulco Vocálico
Falsete	Leucoplasia, Hiperkeratose
Laringite	Muda Vocal
Nódulos, Pólipos	Paralisia das Pregas vocais
Quebras de altura	Câncer
Espasmos abdutores	Laringite infecciosa
Úlceras de contato	Edema de Reike

## 2.3-ANÁLISE ACÚSTICA DA VOZ PATOLÓGICA E ESTADO-DA-ARTE:

Nesta seção mostrar-se-á uma revisão dos diversos parâmetros acústicos desenvolvidos para a detecção e avaliação de patologias da laringe.

### 2.3.1-ANÁLISE ACÚSTICA DA VOZ PATOLÓGICA:

De maneira geral, as técnicas de análise acústica digital, para o sinal de voz digitalizado ( Kent e Read, 1992) são resumidas na Tabela 3.1.

*Tabela 3.1: Resumo das técnicas de Análise Acústica da Voz Digitalizada (Kent e Read, 1992).*

Domínio Tempo:	Domínio da Frequência:
-Exibição da forma de onda do sinal;	-Espectro de frequências via "Fast Fourier Transform" (FFT) ou "Linear Predictive Coding" (LPC) de intervalos selecionados do sinal;
-Edição da forma de onda do sinal;	-Espectrogramas;
-Reprodução do sinal original ou editado;	-Rastreamento de "formantes";
-Análise da frequência fundamental (no tempo);	-Análise da frequência fundamental;
-Análise de perturbações da frequência fundamental e da amplitude da voz ("jitter" e/ou "shimmer").	-Cálculos da razão sinal-ruído;
	-Síntese da fala.

As opções mais frequentes de análise acústica do sinal de uma voz disfônica são: medidas de perturbação da frequência fundamental e amplitude da voz ("jitter" e "shimmer", respectivamente); análise espectrográfica da voz; diversas maneiras propostas para determinar a razão sinal-ruído e processamento do sinal com fins de extrair informações específicas. Tais técnicas são descritas na próxima seção.

Outras formas de análise são apresentadas usando diferentes tipos de sinais relacionados com a voz. Uma delas é a eletroglotografia (EGG) em que é rastreado o movimento das pregas vocais através da medição da impedância elétrica da região da laringe (Baken, 1987). Com este sinal pode-se observar o padrão vibratório de ambas as pregas de maneira independente, podendo-se verificar se existem assimetrias nos seus padrões vibratórios. Pode-se extrair também, em um ciclo de fundamental, quanto tempo as pregas permanecem fechadas e quanto tempo permanecem abertas. Da razão entre estes 2 valores pode-se calcular o índice OQ (Opening Quocient), que mostra o quão eficiente é o uso do sistema fonador. Outra forma de análise é através das medidas aerodinâmicas da fala, onde se faz uso de um pneumotacógrafo ou expirômetro para medida de fluxo de ar durante a fala. Este sinal é gravado conjuntamente com o próprio sinal de fala e de ambos se extraem parâmetros importantes como a pressão sub-glotal instantânea, a eficiência da fala (razão entre fluxo DC e o fluxo AC), entre outros (Isshiki, 1989).

Neste trabalho serão abordados apenas os métodos de processamento digital do sinal da voz.

### 2.3.2-REVISÃO DE LITERATURA E ESTADO-DA-ARTE:

É apresentado a seguir um resumo das técnicas de processamento de voz patológica com fins de detecção de patologias e auxílio no diagnóstico médico. No final desta seção, nas Tabelas 2.3 à 2.7, podem ser vistos resumos das fórmulas das principais técnicas apresentadas a seguir.



Com relação às medições de perturbações do valor do período da frequência fundamental (o que equivale a pequenas variações da própria frequência fundamental), denominada de "jitter", Lieberman (1961),(1963) definiu o parâmetro perturbação da frequência fundamental (PFF) como a diferença entre a duração de um período de frequência fundamental da voz e seu período subsequente (da onda temporal registrada de uma vogal sustentada). Baseado nesta definição, é calculado então o "Pitch Perturbation Factor" (PPF) como sendo a percentagem do número total de PFF que sejam maiores do que o valor limiar de 0,5 ms. A magnitude deste indicador foi usada com êxito para a diferenciação de vozes disfônicas de vozes normais (Iwata e von Lenden, 1970a e Hiki *et al.*, 1975).

Outra maneira de medir o "jitter" foi proposta por Hecker e Kreul (1971). Os autores propõem o "Directional Perturbation Factor" (DPF) que mede o número de vezes que a diferença entre a duração de um período de frequência fundamental da voz e seu período subsequente muda de sinal (em valor percentual), não se preocupando com a magnitude da mesma. Neste trabalho, algumas pessoas idosas, para uma sentença padrão lida, obtiveram valor médio de 33,3% com desvio padrão de 4,2%. Sorensen e Horii (1984) determinaram o DPF para 20 homens e 20 mulheres não-fumantes, para 3 tipos de vogais sustentadas (/a/, /u/, e /i/) obtendo médias e desvios-padrão desta medida para pessoas normais e comparando os resultados para as 3 vogais.

Horii (1979) demonstra em seu trabalho que a magnitude da perturbação da frequência fundamental está correlacionada com o valor médio da altura da frequência fundamental. É demonstrado que o "jitter" diminui com o aumento da frequência fundamental.

Com a finalidade de tentar normalizar a medida de "jitter" em relação à frequência fundamental, o autor sugere que se calcule o "Jitter Ratio" (JR), onde é feita uma média da diferença entre períodos sucessivos (perturbação média), dividido pelo período médio sucessivo, multiplicado por 1000. A efetividade desta compensação, entretanto, não foi avaliada.

O equivalente desta medida é sugerido por Hollien, Michel e Doherty (1973) e Murry e Doherty (1980), denominado de "Jitter Factor" (JF), onde é feita uma média da diferença entre períodos sucessivos (perturbação média), dividido pelo período médio sucessivo, multiplicado por 100, mas onde todos os valores usados são expressos em termos de frequência em lugar de tempo. Valores médios para pessoas normais são apresentados pelos autores.

Em outra forma de calcular o PPF, Honjo e Isshiki (1980) assumem que a percepção da frequência fundamental é melhor expressa em escala logarítmica da frequência fundamental (semitons musicais). Assim também se obtém uma medida de perturbação compensada em relação ao valor da frequência fundamental. É mostrado que a faixa para voz normal, para vogais sustentadas, está entre 0,08 a 0,23 semitons para homens e 0,14 a 0,25 semitons para mulheres. Valores de perturbação acima de 1 semitom indicam uma severa disfonia.

Tentando minimizar o efeito de variações lentas normais da frequência fundamental, Koike (1973) introduz o "Relative Average Perturbation" (RAP), também denominado "Frequency Perturbation Quocient" (FPQ) por Takahashi e Koike (1975), onde é calculada a

diferença entre a duração do período e a média da duração dos períodos anterior, atual e o posterior.

Os valores de voz normal médios desta medida são 0,0057 para homens e 0,0061 para mulheres. Davis (1979) tentou usar tal medida para discriminar voz normal da voz patológica.

Adotando outra abordagem derivada da estatística descritiva, Deal e Emanuel (1978) definem o "Period Variability Index" (PVI), que envolve o cálculo do coeficiente de variação (CV), que por sua vez é a média do quadrado da diferença entre períodos sucessivos, dividido pelo período médio quadrado.

O índice PVI é então calculado multiplicando-se o CV por 1000. É feito um estudo para diferentes vogais, obtendo-se valores médios, para adultos normais.

Koike (1969) e von Lenden e Koike (1970) investigam o parâmetro de perturbação na amplitude dos picos dominantes de cada período de fundamental ("shimmer"), para uma vogal sustentada. O sinal adquirido é o sub-glotal, ou seja, é posicionado um microfone de contato abaixo da laringe, no espaço ao final do osso esterno. O tipo de onda neste trecho, por não ter sofrido influências do trato buco-faríngeo, se aproxima mais de uma senóide pura do que o sinal da voz captado por um microfone em frente à boca.

Os autores marcam todos os picos dominantes de cada período de fundamental. A investigação foi feita para 30 períodos de fundamental consecutivos da parte mais estável da onda, para cada paciente. Em seguida, procura-se determinar se existe uma variação periódica nos picos de fundamental através do coeficiente de correlação serial (ou correlograma). O

correlograma é calculado para  $k$  variando de 1 até 15 períodos de fundamental. É mostrado que pacientes normais possuem modulação em amplitude com periodicidade dita de longo termo, cujo período da moduladora excede os 15 períodos de fundamental. Este fenômeno é atribuído a um "vibrato" natural da voz. Os pacientes com paralisia uni-lateral não apresentaram nenhum pico de alta correlação, exceto para  $k$  igual a 1, mostrando que a voz não apresenta nenhuma periodicidade significativa. Já os pacientes com neoplasias da laringe apresentaram uma série variada de periodicidades de curto termo, com período compreendido entre 3 e 15 períodos de fundamental. Entretanto 10% dos pacientes normais também apresentaram esta característica, evidenciando que este não pode ser usado para indicar neoplasia das pregas vocais, ainda que ele indique alguma disfunção patológica da laringe.

O "Directional Perturbation Factor" (DPF) definido por Hecker e Kreul (1971) também pode ser aplicado para as medidas de amplitude do pico dominante de cada período de fundamental como foi demonstrado por Sorensen e Horii (1984). Os valores obtidos para "shimmer" são comparados com os obtidos para a medida de "jitter", para as mesmas 3 vogais sustentadas, para pessoas normais, como descrito anteriormente. A idéia básica é tentar achar os valores normais para o DPF aplicado à amplitude dos picos de fundamental.

Koike *et al.* (1977) definem o "Perturbation Quocient" (PQ) que pode ser aplicado tanto às perturbações de amplitude quanto às perturbações de frequência. O PQ é apresentado como uma definição mais genérica do índice RAP, onde o número de pontos para o cálculo da média local do termo dentro do módulo, no numerador, é variável. Fazendo-se o número de pontos igual a 3, obtém-se o RAP. para medidas de período.

Deal e Emanuel (1978) definem o "Amplitude Variability Index", que é o equivalente do PVI, para amplitudes. Entretanto a principal diferença é que o sinal de voz é inicialmente processado por um filtro passa-faixas com largura de banda de 10Hz e frequência central no primeiro harmônico do sinal, com o intuito de se remover irregularidades dentro de cada período, obtendo-se assim um sinal mais próximo ao glotal. Em seguida, é calculado o "coeficiente de variação" (CV) dado pelo quadrado da diferença entre as amplitudes de períodos sucessivos, dividido pela amplitude média quadrada. Então o AVI é calculado por  $AVI = \log_{10}(CV.1000)$ . Este índice é então aplicado tanto para pacientes normais como para aqueles com patologias da laringe.

Horii (1980) sugere que o "shimmer" seja calculado pela média da divisão entre a amplitude de um período e o período posterior, em decibéis. Assim, o autor afirma obter um valor de "shimmer" independente do valor absoluto da amplitude do sinal.

Takahashi e Koike (1975) e Koike, Takahashi e Calcaterra (1977) sugerem o "Amplitude Perturbation Quocient" (APQ) como método para se eliminar as variações de longo termo da amplitude, que são normais. Este índice é equivalente ao RAP e ao PQ, definidos anteriormente, mas onde a média local usa 11 valores de amplitude. O sinal de voz, neste caso é o pré-traqueal, obtido por um microfone de contato colocado na pele logo abaixo da laringe, no espaço acima do fim do osso esterno. Valores médios são calculados para diversos pacientes.

Uma maneira geral de se obter informações sobre patologias da voz é através do espectrograma. Yanagihara (1967 a, b) realizou as avaliações de diversos pacientes com variados graus de rouquidão, para diversas vogais sustentadas, com espectrogramas de banda

estreita (45 Hz). Conforme a composição dos espectrogramas e a distribuição de ruído nestes, o autor conseguiu classificar a rouquidão em 4 tipos ou graus. Conforme Yanagihara (1967 b) os quatro tipos são:

a) Tipo I: os componentes harmônicos regulares estão misturados com o ruído principalmente nas regiões dos formantes das vogais;

b) Tipo II: os componentes de ruído no segundo formante das vogais da língua inglesa /ε/ e /i/ sobrepõem as componentes harmônicas, e um ruído adicional aparece na região acima de 3000 Hz;

c) Tipo III: o segundo formante das vogais /ε/ e /i/ são totalmente substituídos por ruído e, na faixa acima de 3000 Hz, o ruído adicional aumenta sua energia e sua extensão;

d) Tipo IV: o primeiro e segundo formante das vogais /a/, /ε/ e /i/ são substituídos por ruído e, na região do primeiro formante, todas as vogais têm suas componentes periódicas substituídas por ruído. Além disso, aparece mais um ruído adicional na faixa de alta frequência.

Iwata e von Lenden (1970b) sugerem em seu trabalho que diferentes disfunções da laringe podem levar a espectrogramas com diferentes características.

Rontal, Rontal e Rolnick (1975), através dos espectrogramas de banda larga, procuraram correlacionar os padrões gráficos obtidos para diversos tipos de patologia com os tipos qualitativos de voz disfônica, bem como da voz normal. É mostrado que os diversos

fenômenos fisiológicos da produção da voz são evidenciados nos espectrogramas. É ainda sugerido que os espectrogramas possam ser utilizados para se avaliar a reabilitação de pacientes em tratamento fonoterapêutico. Ao final do artigo são relatadas as limitações da análise através de espectrogramas.

Conforme resumido por Isshiki (1989), espectrogramas de banda estreita (45 Hz) para uma vogal sustentada de uma pessoa normal apresentam os harmônicos bem pronunciados, como linhas horizontais igualmente espaçadas. Para vozes patológicas, estas linhas aparecem tanto mais borradas quanto maior for a gravidade da disfonia ou rouquidão.

Entretanto, a característica qualitativa aspereza da voz, bem como as perturbações na frequência fundamental são melhor representadas pelos espectrogramas de banda larga (300Hz), onde os períodos de fundamental aparecem como barras verticais regularmente espaçadas e as perturbações no período de fundamental se mostram como barras verticais irregularmente espaçadas. A característica soprosidade aparece neste tipo de espectrograma como ruído aleatório espalhado por toda a banda de frequência. Baseado nessas observações dos espectrogramas, pode-se diferenciar as características aspereza e soprosidade de uma voz disfônica, como também dar um grau quantitativo de classificação para elas.

Assumindo-se que o sinal acústico de uma vogal sustentada é a soma de um componente periódico e de ruído randômico, são apresentadas diversas formas de se calcular a relação sinal-ruído, com o fim de se identificar patologias.

Com base no uso de espectrogramas, Emanuel e Sansone (1969) e Sansone e Emanuel (1970) propõem o índice "Spectral Noise Level" (SNL), calculado da seguinte forma: um

trecho de 3 segundos de uma vogal sustentada de intensidade com 75 dB SPL ("Sound Pressure Level") é analisada continuamente por um espectrógrafo cuja filtro de banda foi modificado de 45Hz para 3Hz (alta resolução em frequência). Os gráficos produzidos são então equivalentes aos produzidos pela banda de 45Hz, só que com maior resolução espectral. O ruído é então definido como o menor valor marcado (em dB SPL) de cada segmento de 100Hz, sendo isto feito para todo o espectro de 200 a 8000Hz. O índice SNL encontrado é correlacionado com os níveis da qualidade aspereza, sendo portanto útil para a identificação de voz patológica (Emanuel e Whitehead, 1979).

Kojima *et al.* (1980) propõem calcular a razão sinal-ruído (S/N) através da expansão por série de Fourier de 3 períodos de fundamental do sinal da vogal sustentada /a/, de duração total igual a  $T_1$ , criando-se um novo sinal periódico assumido como infinito no qual é feita a análise (a série de Fourier, apesar de considerar o sinal infinito, é definida para apenas um período do sinal). Como o período  $T_1$  do novo sinal contém 3 períodos de fundamental  $T_0$  do sinal original ( $T_0 = \frac{T_1}{3} \Leftrightarrow F_0 = 3.F_1$ ), os harmônicos relativos à parte periódica do sinal de voz aparecem nas frequências  $3F_1, 6F_1, 9F_1, \dots$ , para  $F_1 = 1/T_1$ . Os outros componentes entre estes são considerados o espectro do ruído aditivo. O sinal foi expandido até 5500 Hz. Sendo S o somatório do quadrado das intensidades das componentes do sinal e N o somatório do quadrado das intensidades das componentes do ruído, então a razão sinal-ruído para este trecho extraído do sinal original é dada por S/N. Este procedimento é repetido para os próximos 3 períodos do sinal e assim sucessivamente. Então a razão sinal-ruído média do sinal original ( $R_a$ ) é definida como 20 vezes o logaritmo da divisão entre o valor médio do quadrado do espectro do sinal e o valor médio do quadrado do espectro do ruído. A faixa de  $R_a$  obtida para vozes normais está entre 15,0 dB e 23,5 dB e a região estatística para a faixa



normal, a 5% de nível de significância, está entre 15,4 e 23,3 dB. Os autores mostram que este índice pode ser usado para diferenciar a voz normal da voz patológica e quantificar a disфонia e, quando usado em conjunto com outros indicadores, pode diferenciar a característica soprosidade da característica aspereza.

Kitajima (1981) sugere que para se calcular a razão de ruído ("Noise Ratio"), se faça a separação da componente periódica do sinal do ruído no domínio da frequência, usando o espectro obtido por FFT (Fast Fourier Transform) de 4096 pontos de um sinal, através de um complicado processo de estimação e filtragem.

São mostrados resultados e comparados com os obtidos via espectrogramas, achando-se uma correlação entre o ruído calculado por este método e o calculado com base nos espectrogramas. Entretanto, o processo de cálculo deste índice é extremamente complexo.

Visando simplificar o método de cálculo da razão entre a energia da componente harmônica e da componente de ruído aditivo, Yumoto, Gold e Baer (1992) sugerem que se separem todos períodos de fundamental de um trecho de vogal sustentada e calcule-se um sinal temporal médio de período de fundamental, que representará a parte harmônica do sinal.

A medida de energia acústica do componente harmônico (H) é definida por  $n$  vezes a integral do quadrado deste sinal médio, onde  $n$  é o número de períodos de fundamental da amostra.

O ruído em cada período de fundamental é definido como a diferença entre este período e o sinal médio. Então a energia total do ruído (N) é dada pela média desta diferença.

A razão entre H e N obtida em dBs resulta no índice "Harmonic-Noise Ratio" (H/N). Os valores médios normais encontrados pelos autores são 12,2dB para homens e 15,5dB para mulheres. A faixa do grupo normal total fica entre 7,0dB e 17,0dB com média de 11,9dB, ao nível de significância de 5%. Espera-se que 95% dos normais tenham H/N maior ou igual a 7,4dB (one-tail test). Já a faixa do grupo de pré-operatórios se apresenta entre -15,2dB e 9,6dB, com média de 1,6dB. Entretanto, 16,7% dos pacientes patológicos têm a razão H/N maior que 7,4dB. Finalmente os resultados são comparados aos obtidos por Kojima *et al.* (1980).

Assumindo que o sinal de uma vogal sustentada pode ser dividido em 3 componentes: O componente da frequência fundamental, os componentes dos harmônicos da frequência fundamental e o componente de ruído aditivo, Hiraoka *et al.* (1984) sugerem que se calcule os índices Intensidade de Harmônicos ("Harmonic-Intensity",  $H_T$ ) e Intensidade Relativa do Sinal ("Relative Signal Intensity",  $S_T$ ). É calculada a FFT de 4096 pontos da porção mais estável do sinal da vogal sustentada /a/ (aproximadamente 0,2 s), segmentado por janela Hamming e processada por um filtro passa-baixas com frequência de corte de 8kHz, sendo então determinada no espectro a frequência fundamental  $F_0$ . A seguir, os harmônicos da frequência fundamental são também determinados. O índice  $H_T$  é então calculado pelo somatório dos harmônicos da fundamental, não incluindo a fundamental, dividido pelo somatório de todos componentes espectrais, até 8 kHz, usando escala linear para todos casos.

O índice  $S_T$  é também calculado de forma semelhante, mas acrescentando-se o componente de frequência fundamental ao somatório do numerador.

Observa-se que  $S_r$  é equivalente aos índices  $R_a$  e  $H/N$  calculados em trabalhos anteriormente citados, só que agora é usada escala linear. Os valores de  $H_r$  e  $S_r$ , para pessoas normais foram de 80,3% e 87,1%, com limites inferiores de 67,2% e 76,5%, respectivamente, ao nível de significância de 5%. Entretanto, o índice  $H_r$  apresentou um número maior de identificação de patologias (90%) que o índice  $S_r$  (60%), mostrando-se um índice mais confiável.

Kasuya *et al.* (1986b) propõem o parâmetro Energia Normalizada de Ruído ("Normalized Noise Energy", NNE) como forma para se avaliar patologias da voz. O sinal da vogal /e/ sustentada é todo dividido em trechos e é calculada a DFT de cada trecho do sinal (segmentado por janela Hamming). O espectro obtido pode ser considerado como a soma de uma componente periódica  $S_m(k)$  e uma componente de ruído  $W_m(k)$ . Assim, o NNE é calculado em escala de dB (10 vezes o logaritmo) usando-se a média do quadrado do espectro da parte periódica do sinal e a média do quadrado do espectro do ruído aditivo.

Como já se tem o espectro do sinal, o problema é achar o estimador do espectro do ruído  $\hat{W}_m(k)$ . No espectro do sinal, quando segmentado por uma janela Hamming de largura relativamente grande, os picos representam a parte periódica do sinal e mais o ruído aditivo, e a região entre eles, somente de ruído aditivo. Na região de pico o ruído é estimado pelo valor médio entre o valor mínimo no início do pico (no vale entre um pico e outro) e o valor máximo do topo do pico. O índice é calculado na região do espectro que vai de uma frequência inferior  $f_L$  até uma frequência superior  $f_H$ .

Para se avaliar qual a melhor banda de frequência em que o índice se apresenta mais confiável, são experimentados diversos valores de  $f_L$  e  $f_H$ . Resultados são apresentados e os

autores afirmam que este índice tem melhor desempenho do que os índices anteriormente apresentados. Kasuya *et al.* (1986 a) fazem uma revisão e avaliação dos métodos propostos anteriormente, bem como comparam seu método proposto com outros.

Seguindo outra linha, Koike e Markel (1975) analisam a detecção de patologias do sinal de voz através do resíduo da filtragem inversa. Neste processo, é aplicado ao sinal um filtro cuja característica seja inversa à função de transferência do trato buco-faríngeo modelado por "Linear Predictive Code" (LPC ou também denominado modelagem auto-regressiva), obtendo-se assim um sinal residual. É observado que indivíduos normais apresentam como sinal residual um trem de impulsos periódicos, com ruído de baixa amplitude entre eles. Para pessoas com patologias na voz, os impulsos têm menos intensidade e o ruído entre os impulsos têm intensidade aumentada.

Deller Jr. e Anderson (1980) sugerem que se faça a classificação da voz utilizando-se as singularidades da função de transferência em  $Z$  do filtro digital inverso obtido através da modelagem LPC. É mostrado que o mapa das raízes no plano  $Z$  formam regiões diferentes ("clusters") para cada tipo de patologia, embora a dispersão de cada "cluster" possa acarretar em problemas de identificação da patologia.

Remacle e Trigaux (1991) fazem um estudo da frequência do sinal de voz de pacientes com nódulos nas cordas vocais, através do analisador de frequência de alta resolução Bruel e Kjaer modelo 2033. As análises vão até a frequência de 5 kHz e o sinal de voz é a vogal sustentada /a/. É concluído que esta análise sozinha não provê capacidade de diagnóstico.

Childers e Bae (1992) propõem um sistema automático de classificação de patologia da voz com base na modelagem LPC e em técnicas de identificação de palavras baseada em quantização vetorial. São utilizados sinais de voz bem como sinais eletroglotográficos. Tal sistema detecta patologias com taxas de acerto entre 69% a 75%, apresentando falsos positivos da ordem de 9,6%, para pacientes normais.

Tentando achar um indicador que melhor se adapte à característica soprosidade da voz, Fukazawa, El-Assuooty e Honjo (1988) propõem o índice  $B_T$ . O sinal de voz da vogal sustentada /a/ é processado por um filtro passa-altas com frequência de corte de 10000Hz, antes de ser digitalizado. O índice  $B_T$  é então calculado como a média do quadrado da diferença móvel de 3 pontos, dividido pelo quadrado do sinal. O índice é aplicado a pessoas normais sendo obtida a faixa de 8,3 a 75,7 com média 27,0 e desvio-padrão de 15,0. A média obtida dos pacientes patológicos foi de 213,3, embora não seja dada a faixa total obtida.

Também como medida do aspecto soprosidade da voz, Shoji *et al.* (1992) sugerem que se calcule a razão de potência de alta frequência (High-Frequency Power Ratio). Este índice é calculado com base na estimação do espectro de potência por FFT de 1024 pontos de vários trechos da vogal sustentada /a/, segmentada pela janela Blackman. A frequência de amostragem usada é de 44,1 kHz. A razão de potência é a divisão entre o somatório do espectro de potência acima de 6 kHz e o somatório do espectro total, dado em dBs. Outras frequências de corte foram experimentadas. Para valores do índice acima de -30 dB, a voz pode ser considerada soprosa. Finalmente, para frequências de corte acima de 4 kHz, os autores mostram que este índice pode separar vozes soprosas de vozes normais.

As Tabelas 2.3 a 2.7, apresentam um resumo das técnicas descritas anteriormente.

Tabela 2.3: Resumo dos principais índices de "jitter".

"Jitter"	Autor(es)	Cálculo	Descrição/ Observações
"Pitch Perturbation Factor" (PPF)	Lieberman (1961), (1963)	$pff =  T_i - T_{i+1} , i = 0, 1, \dots, N-1;$ $PPF(\%) = \frac{\sum n^{\circ} pff > 0,5 \text{ ms}}{N} \cdot 100$	$T_i$ é o valor em [s]. do i-ésimo período de fundamental; N é o número total de períodos de fundamental da amostra.
"Directional Perturbation Factor" (DPF)	Hecker e Kreul (1971)		É o número de vezes que a diferença entre a duração de um período de frequência fundamental da voz e seu período subsequente muda de sinal, em valor percentual
"Relative Average Perturbation" (RAP) ou "Frequency Perturbation Quocient" (FPQ)	-Koike (1973) -Takahashi e Koike (1975)	$RAP = \frac{1}{N-2} \sum_{i=1}^{N-1} \frac{ T_{i-1} + T_i + T_{i+1} - 3T_i }{\frac{1}{N} \sum_{i=1}^N T_i}$	$T_i$ é o valor do i-ésimo período de fundamental em [ms].
"Jitter Factor"	Hollien, Michel e Doherty (1973); Murry e Doherty (1980)	$JF = \frac{1}{N-1} \sum_{i=1}^{N-1} \frac{ F_i - F_{i+1} }{\frac{1}{N} \sum_{i=1}^N F_i} \cdot 100$	$F_i$ é o valor em [Hz] do i-ésimo período de frequência fundamental.
"Jitter Ratio" (JR)	Horii (1979)	$JR = \frac{1}{N-1} \sum_{i=1}^{N-1} \frac{ T_i - T_{i+1} }{\frac{1}{N} \sum_{i=1}^N T_i} \cdot 1000$	$T_i$ é o valor do i-ésimo período de fundamental em [ms].
Jitter em semi-tons	Honjo e Isshiki (1980)	$PPF_{\text{semi-ton}} = \Delta \bar{F} = \frac{\sum_{i=1}^N  F_i - F_{i+1} }{N}$ <p>onde</p> $F_{\text{semi-ton}} = 12 \cdot \log_2 \frac{f}{f_r} = 39,86 \cdot \text{LOG} \frac{f}{f_r}$	$F_i$ é o valor instantâneo em semi-tons do i-ésimo período de fundamental, f é a frequência a ser convertida e $f_r$ é a frequência de referência de valor 16,352 Hz
"Period Variability Index" (PVI)	Deal e Emanuel (1978)	PVI = CV.1000	$CV = \frac{1}{N} \frac{\sum (T_i - \bar{T})^2}{\bar{T}^2}; T_i \text{ é o } i\text{-ésimo período de fundamental em [ms] e } \bar{T} \text{ é o valor médio de período de fundamental para a amostra, em ms.}$

Tabela 2.4: Resumos dos principais índices de "shimmer".

"shimmer"	Autor(es)	Cálculo	Descrição
Correlograma	Koike (1969) e von Lenden e Koike (1970)	$r_k = \frac{1}{N-k} \sum_{i=1}^{N-k} \left( \frac{A_i - \bar{x}_1}{s_1} \right) \left( \frac{A_{i+k} - \bar{x}_2}{s_2} \right)$	$\bar{x}_1 = \sum_{i=1}^{N-k} \frac{A_i}{(N-k)}; \quad \bar{x}_2 = \sum_{i=k+1}^N \frac{A_i}{(N-k)}$ $s_1^2 = \sum_{i=1}^{N-k} \frac{(A_i - \bar{x}_1)^2}{(N-k)} \text{ e } s_2^2 = \sum_{i=k+1}^N \frac{(A_i - \bar{x}_2)^2}{(N-k)}$
"Perturbation Quocient" (PQ)	Koike et al. (1977)	$PQ = \frac{\frac{1}{N-k+1} \sum_{n=1}^{N-k+1} \left  \sum_{r=1}^k x(n+r-1) - x(n+n) \right }{\frac{1}{N} \sum_{n=1}^N x(n)}$	$m = \frac{1}{2}(k-1)$ ; $k$ é um número inteiro ímpar e $x(n)$ é uma seqüência de $1, \dots, N$ pontos que podem ser a amplitude do pico dominante de cada período de fundamental ("Amplitude Perturbation Quocient", APQ) ou pode ser o valor de cada período de fundamental ("Pitch Perturbation Quocient", PPQ).
"Amplitude Variability Index" (AVI)	Deal e Emanuel (1978)	$AVI = \log_{10}(CV.1000)$	$CV = \frac{\frac{1}{N} \sum_{i=1}^N (A_i - \bar{x})^2}{\bar{x}^2}$
"shimmer" em decibéis	Horii (1980)	$dB_{shimmer} = \frac{\sum_{i=1}^{N-1} \left  20 \log(A_i/A_{i+1}) \right }{N-1}$	$A_i$ indica o valor de pico de amplitude do $i$ -ésimo período de fundamental.
"Amplitude Perturbation Quocient" (APQ)	Takahashi e Koike (1975); Koike, Takahashi e Calcaterra (1977)	$APQ = \frac{\frac{1}{N-10} \sum_{i=6}^{N-5} \left  \frac{\sum_{j=-5}^{+5} A_{i+j}}{11} - A_i \right }{\frac{1}{N} \sum_{i=1}^N A_i}$	$A_i$ indica o valor de pico de amplitude do $i$ -ésimo período de fundamental

Tabela 2.5: Resumo dos principais índices da Razão Sinal-Ruído.

Índice	Autor(es)	Descrição
Razão sinal-ruído média ( $R_a$ )	Kojima et al. (1980)	Através de expansão em série de Fourier. São extraídos 3 períodos de fundamental do sinal da vogal sustentada /a/, de duração igual a $T_1$ . O novo período $T_1$ contém 3 períodos de fundamental, logo ( $T_{Fundamental} = T_1/3$ ); os harmônicos relativos à parte periódica do sinal de voz aparecem nas freqüências $3F_1, 6F_1, 9F_1, \dots$ . Os outros componentes entre estes são considerados espectro do ruído aditivo. O sinal foi expandido até 5.500 Hz. Sendo $S$ o somatório do quadrado das intensidades das componentes do sinal e $N$ o somatório do quadrado das intensidades das componentes do ruído, então a razão sinal-ruído para este trecho extraído do sinal original é dada por $S/N$ . Repete-se para os próximos 3 períodos do sinal e assim sucessivamente até que se complete 325 ms de sinal.
Razão de ruído ("Noise Ratio")	Kitajima (1981)	Separa-se a componente periódica do sinal do ruído, no domínio da freqüência, usando um processo que envolve o espectro obtido por DFT (Discrete Fourier Transform, calculado através de um algoritmo FFT).
"Harmonic-Noise Ratio" (H/N) calculado em [dBs]	Yumoto, Gold e Baer (1992)	Separa-se todos períodos de fundamental de um trecho de vogal sustentada e calcula-se um sinal média de período de fundamental, que representará a parte harmônica do sinal. O ruído é a diferença entre o sinal média e cada período de fundamental. O razão é a energia da parte harmônica pela média do ruído.
"-Harmonic-Intensity", $H_r$ "-Relative Signal Intensity", $S_r$ em porcentagem	Hiraoka et al. (1984)	É calculada a DFT de 4096 pontos da porção mais estável do sinal da vogal sustentada /a/ (aprox. 0,2 s) e é determinada no espectro a freqüência fundamental $F_0$ . A seguir, os harmônicos da freqüência fundamental são também determinados. A razão entre os harmônicos, não incluindo o componente fundamental e o ruído é $H_r$ e a razão entre os harmônicos com a fundamental e o ruído é $S_r$ .
"Normalized Noise Energy", NNE	Kasuya et al. (1986 a, b)	Estimação no domínio freqüência da parte harmônica e do ruído, através da FFT. Cálculo em dB.

Tabela 2.6: Outros índices propostos na literatura.

Índice	Autor(es)	Descrição
Resíduo da filtragem inversa	Koike e Markel (1975)	Aplica-se um filtro cuja característica seja inversa à função de transferência do trato buco-faríngeo (calculado pela modelagem Linear Predictive Coding, LPC), obtendo-se assim o sinal residual que se aproxima do sinal glotal.
Posição das singularidades da transformada Z do Filtro Inverso	Deller Jr. e Anderson (1980)	É feita a classificação da voz utilizando-se as raízes da transformada Z do filtro digital inverso obtido através da modelagem "Linear Predictive Code" (LPC). É mostrado que o mapa das raízes no plano Z formam diferentes "clusters" para diferentes patologias
Análise de Clusters	Childers e Bae (1992)	Sistema automático de classificação de patologia da voz com base na modelagem LPC e em técnicas de identificação de palavras baseada em quantização vetorial.

Tabela 2.7: Índices relacionados com a soproidade da voz.

Índice $B_r$	Fukazawa, El-Assuooty e Honjo (1988)	Índice para a "soproidade" da voz. O sinal de voz da vogal sustentada /a/ é filtrado por um filtro passa-altas com frequência de corte de 10.000 Hz, antes de ser digitalizado. $B_r = \frac{\sum [f(t_j) - 2.f(t_{j-1}) + f(t_{j-2})]^2}{\sum [f(t_j)]^2} \cdot 100$ , onde $f(t_j)$ é o j-ésimo ponto do sinal de voz $f(t)$ .
"High-Frequency Power Ratio"	Shoji et al.(1992)	Este índice é calculado com base na estimação do espectro de potência por FFT usando 1024 pontos de vários trechos da vogal sustentada /a/, segmentada com uma janela Blackman. A razão de potência é a divisão entre o somatório do espectro de potência acima de 6 kHz e o somatório do espectro total, em dBs.

Resumos destas técnicas podem ser encontrados em Kasuya *et al.* (1986 a), Baken (1987), Isshiki (1989) e em Esquenazi, Childers e Hicks (1990)



### 3 FUNDAMENTOS TEÓRICOS:

Conforme Daubechies (1992), a transformada Wavelet é uma transformação integral recente, pois é citado neste trabalho que as primeiras referências do nome Wavelet apareceram em Morlet *et al.* (1982), Morlet (1983) e Grossmann e Morlet (1984), muito embora tenham existido alguns trabalhos anteriores cuja filosofia era a mesma, tais como a decomposição em funções ortogonais de Haar (1910) e a resolução da identidade de Calderón (1964) em análise harmônica, dentre outros. Toda base teórica matemática para validação desta transformada foi feita por diversos pesquisadores da escola matemática francesa (Bentley e McDonnell, 1994). Sua aplicação no campo dos sinais digitais tornou-se possível graças aos trabalhos de Daubechies (1988, 1989) no desenvolvimento das bases de Wavelets ortonormais discretas, e Mallat (1989a,b) na sua análise multi-resolução. Nestes trabalhos são apresentadas as bases matemáticas da transformada Wavelet discreta e sua conexão com o processamento digital de sinais.

Neste Capítulo são mostrados os fundamentos matemáticos desta transformada, começando-se pela transformada Wavelet contínua, comparando-se esta com a transformada de Gabor. Em seguida é apresentada a transformada Wavelet discreta e finalmente, a análise multi-resolução.

#### 3.1- A TRANSFORMADA WAVELET CONTÍNUA (CWT):

A título de comparação, será inicialmente mostrada a transformada de Fourier de curta duração (Short-time Fourier Transform, STFT) ou transformada de Gabor, mostrando-se suas limitações e justificando-se desta feita a introdução da transformada Wavelet contínua .

Uma vez que a transformada de Fourier não é satisfatória para a análise de sinais cujo espectro varie no tempo ou sinais não estacionários, pois os limites de integração são de  $-\infty$  a  $+\infty$  (Bentley e McDonnel, 1994), foi proposta por Gabor (1946) a transformada de Fourier de curta duração (STFT) descrita matematicamente por:

$$\text{STFT}(\tau, \omega) = \int_{-\infty}^{\infty} f(t) \cdot g(t - \tau) \cdot e^{-j\omega \cdot t} dt \quad (\text{III.1})$$

onde  $g(t - \tau)$  é uma determinada janela temporal que multiplica a função  $f(t)$ . Note-se que esta janela se desloca de  $\tau$ .

Esta transformação descreve o sinal em ambos os domínios tempo e frequência ao mesmo tempo. Os gráficos obtidos por esta transformação costumam ter representada a variável tempo ( $\tau$ ) nas abscissas, a variável frequência ( $\omega$ ) nas ordenadas, e a amplitude do módulo, ou fase como uma escala de cinza. Um exemplo é mostrado na Figura 3.1.

A duração da janela temporal  $g(t)$  é dependente da maior duração em que o sinal pode ser considerado estacionário. Para sinais de voz, por exemplo este tempo é estimado em aproximadamente 40 [ms] (Deller, Jr., Proakis e Hansen, 1993). Entretanto observa-se que existe um compromisso entre a resolução temporal e a resolução espectral. Com janelas estreitas, obtém-se boa identificação temporal de eventos transientes, mas baixa resolução espectral. Em caso contrário, se obtém uma melhor identificação entre frequências adjacentes, mas baixa identificação temporal de eventos. Ambos os casos são considerados para mesma frequência de amostragem do sinal analisado.

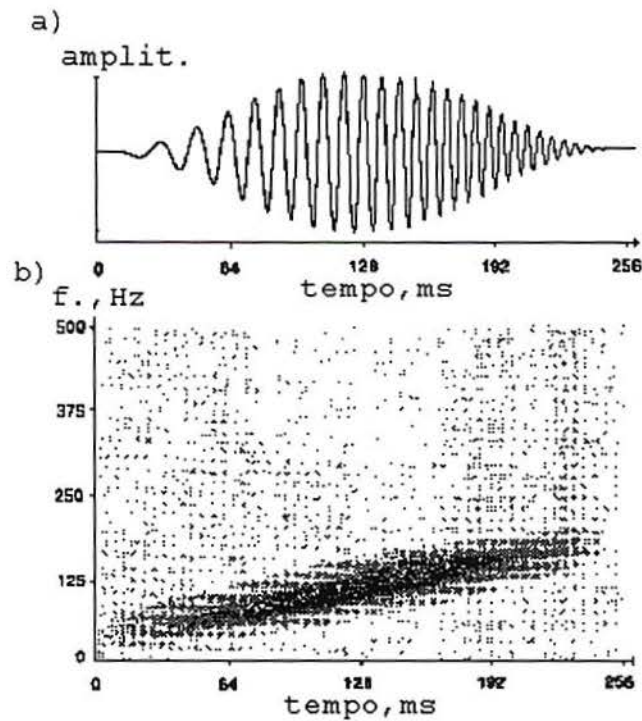


Figura 3.1: Gráfico de um sinal senoidal de frequência variável e amortecido e seu espectrograma. A duração temporal da janela é de 64ms (extraído de Bentley e McDonnell, 1994).

Este compromisso entre resolução em frequência e duração temporal da janela pode ser descrito pelo princípio da incerteza, ou inequação de Heisenberg aplicado à análise de sinais (Sheng, 1996; Rioul e Vetterli, 1991), dado por:

$$\Delta t \cdot \Delta f \geq \frac{1}{2} \quad (\text{III.2})$$

onde  $\Delta t$  é a duração da janela e  $\Delta f$  é a resolução em frequência.

Para uma dada frequência de amostragem, observa-se que, uma vez que a duração da janela  $\Delta t$  é escolhida, ela será fixa em todas faixas de frequência a serem analisadas, ou seja,  $\Delta t$  se mantém durante toda a análise. Por conseguinte, uma vez que a resolução em frequência  $\Delta f$  é proporcional ao inverso da duração da janela temporal, esta também será fixa. Assim, tanto  $\Delta t$  como  $\Delta f$  permanecem constantes ao longo da análise. Isto é melhor entendido quando

emprega-se a transformada discreta de Fourier para o cálculo da STFT. Tais células de resolução são mostradas na Figura 3.2.

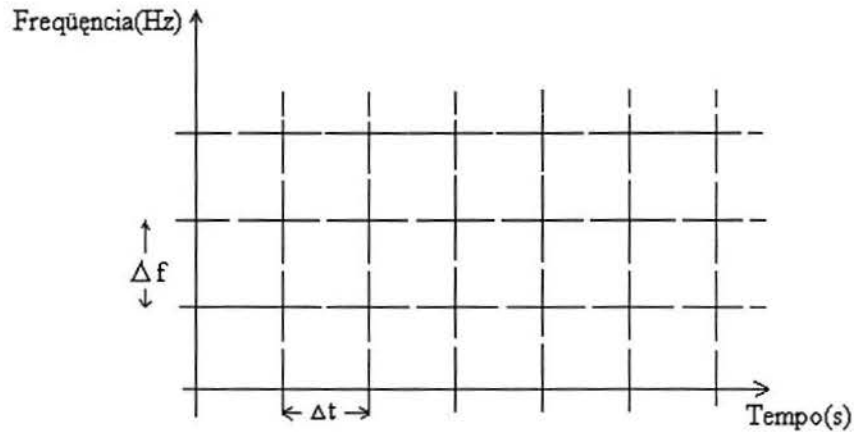


Figura 3.2: Representação das células de resolução  $\Delta t \cdot \Delta f$  na STFT, onde  $\Delta t$  é a duração da janela e  $\Delta f$  é a resolução em frequência.

Procurando superar este inconveniente de células de resolução fixas, Grossmann e Morlet (1984) desenvolveram a transformada Wavelet como uma maneira alternativa à STFT para a análise de sinais não-estacionários.

Seja o espaço  $L^2(\mathcal{R})$  de todas as funções  $f(t)$  que obedecem à seguinte condição  $\int |f(t)|^2 dt < \infty$ . Seja então uma função  $f(t)$  e uma função  $\Psi(t)$  ambas pertencentes ao espaço de funções  $L^2(\mathcal{R})$ . Então, a transformada Wavelet Contínua (CWT) é definida por:

$$\text{CWT}(a, b) = \frac{1}{\sqrt{|a|}} \langle f, \Psi_{a,b} \rangle = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(t) \cdot \Psi^* \left( \frac{t-b}{a} \right) dt \quad (\text{III.3})$$

onde  $\langle \rangle$  representa a operação matemática produto interno,  $a$  e  $b$  são números reais com  $a \neq 0$ ,  $\Psi_{a,b}(t)$  é o conjunto das funções chamadas de Wavelets e  $*$  representa o complexo conjugado da função Wavelet.

Note-se que a função Wavelet é transladada no tempo pelo parâmetro "b" e escalonada pelo parâmetro "a" (a função Wavelet é dilatada para  $a > 1$  e contraída para  $a < 1$ ). A variação do escalonamento e da translação cria um conjunto que será a base de representação da função a ser analisada. Assim, a base é composta por Wavelets que são cópias escalonadas e transladadas de uma Wavelet dita "original" ou "mãe".

O fator  $\frac{1}{\sqrt{|a|}}$  na expressão (III.3) é um fator de normalização da energia da função

Wavelet, para que:

$$\int_{-\infty}^{+\infty} |\Psi_{a,b}(t)|^2 dt = 1 \quad (\text{III.4})$$

Uma condição é geralmente associada às funções Wavelet (Sheng, 1996): é a chamada condição de admissibilidade, condição esta que garante a existência de uma transformada inversa (pois satisfaz a "Resolução de Identidade", Daubechies, 1992), e é dada por:

$$C_\psi = \int_{-\infty}^{+\infty} \frac{|H(w)|^2}{|w|^2} dw < \infty \quad (\text{III.5})$$

onde  $C_\psi$  é uma constante e  $H(w)$  é a transformada de Fourier da função Wavelet  $\Psi(t)$ .

O desenvolvimento da expressão (III.5) é mostrado em detalhes em Sheng (1996). Uma implicação desta condição é que a transformada de Fourier da função Wavelet seja nula na componente de frequência nula (DC), ou seja:

$$|H(w)|^2 \Big|_{w=0} = 0 \quad (\text{III.6})$$

Como conseqüência, a função no domínio tempo deverá ser oscilatória e ter valor médio igual a 0, ou seja:

$$\int_{-\infty}^{+\infty} \Psi(t) dt = 0 \quad (\text{III.7})$$

Assim, é mostrado que qualquer função que pertença ao espaço  $L^2(\mathfrak{R})$  e satisfaça a condição de admissibilidade pode ser uma função Wavelet. Desta forma um sinal decomposto pela transformação definida por (III.3) pode ser recuperado sem perda de informação, pela transformada inversa, que é definida por:

$$f(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle f, \Psi_{a,b} \rangle \Psi_{a,b}^* \frac{da db}{a^2} \quad (\text{III.8})$$

onde  $C_\psi = \int_{-\infty}^{+\infty} \frac{|H(w)|^2}{|w|^2} dw$ .

A transformada Wavelet pode ser vista como a decomposição de uma função  $f(t)$  em uma combinação linear de funções wavelets. Uma vez que a operação produto interno é semelhante à operação correlação, os coeficientes da transformada Wavelet indicam o quão semelhante a função  $f(t)$  é em relação a uma função de decomposição  $\Psi_{a,b}(t)$ . A transformada inversa mostra como a função original  $f(t)$  pode ser sintetizada pela soma das componentes (projeções) do sinal em wavelets. Isto funciona como uma transformação

ortogonal, embora a condição de admissibilidade não garanta ortogonalidade. Assim, dentro do conjunto de funções básicas Wavelets  $\Psi_{a,b}(t)$ , estas não são necessariamente ortogonais.

A condição de ortonormalidade que garante que a decomposição seja feita em componentes ortogonais é dada por:

$$\langle \Psi_{a_1,b_1}(t) \cdot \Psi_{a_2,b_2}(t) \rangle = \int \Psi_{a,b}(t) \cdot \Psi_{c,d}^*(t) dt = \begin{cases} 1, & \text{quando } (a_1 = a_2) \text{ e } (b_1 = b_2); \\ 0, & \text{outros casos.} \end{cases} \quad (\text{III.9})$$

onde  $a_1, a_2$  são as variáveis contínuas de escalonamento e  $b_1, b_2$  são as variáveis contínuas de deslocamento. Desta condição que se chega à expressão (III.4).

Observa-se que as funções Wavelets normalmente são funções limitadas no tempo, e cujos espectros de frequência são semelhantes aos de filtros do tipo passa-faixas. Também observa-se que, Wavelets com maior dilatação tem maior duração temporal e conseqüentemente, maior resolução de frequência do que Wavelets menos dilatadas. Como foi visto anteriormente, na STFT tanto a janela temporal como a largura da banda de frequência permanecem constantes.

Assim, a análise por Wavelets tem a característica de decompor a função a ser analisada por um conjunto de funções cuja duração da janela temporal é variável (e conseqüentemente a largura de banda do espectro), ao invés de valores fixos como na STFT. Esta comparação é mostrada na Figura 3.3.

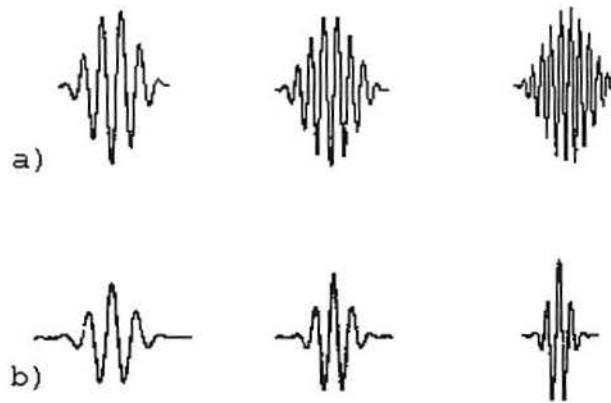


Figura 3.3: Gráficos das funções de base: (a) diferentes freqüências para a STFT, (b) diferentes dilatações para a CWT. Note-se que em (a) o aumento da freqüência não modifica a duração da função.

Uma vez que a duração da janela temporal varia conforme a escala escolhida (e conseqüentemente, a freqüência), também o tamanho das células de resolução é variável, conforme pode ser visto na Figura 3.4.

Os gráficos obtidos neste caso se assemelham aos da STFT, mas a variável freqüência é substituída pela variável de escalonamento ou dilatação "a" e o valor do módulo da CWT é apresentado em escala de cinza. Tais gráficos são também denominados Escalogramas.

Uma vez que muitas vezes os sinais a serem analisados são discretos, existem duas possibilidades de cálculo da CWT. Uma é a aproximação da própria equação (III.3), obtendo-se uma "falsa" Transformada Wavelet Discreta (Bentley e McDonnell, 1994), como é dado por:

$$\text{CWT}(i.T_s, a) = T_s \cdot \frac{1}{\sqrt{|a|}} \sum_n f(n.T_s) \cdot \Psi\left(\frac{[n-i]T_s}{a}\right); a \in Z \quad (\text{III.10})$$

Observa-se que o parâmetro "a" costuma ser dividido em oitavas (ou vozes), ou seja, cada valor de "a" é o dobro do anterior. Assim, variando-se "a" de forma discreta dentro



destes valores, obtém-se um escalograma comparável ao obtido na STFT. Assim, a escala das ordenadas costuma ser representada em escala logarítmica. Um exemplo de escalograma é mostrado na Figura 3.5, para o mesmo sinal da Figura 3.1.

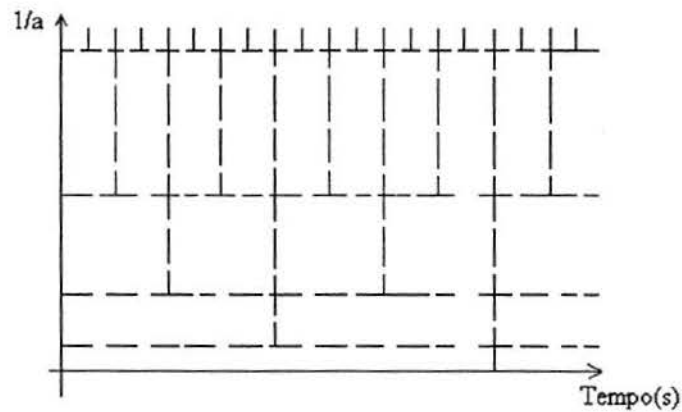


Figura 3.4: Representação das células de resolução  $\Delta t, \Delta(1/s)$  na Transformada Wavelet.

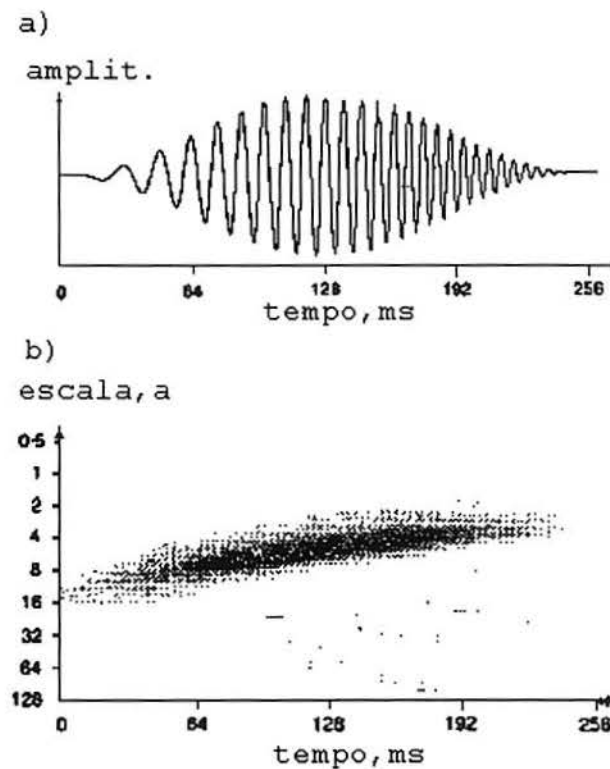


Figura 3.5: Gráfico de (a) um sinal senoidal de frequência variável e amortecido e (b) seu escalograma (extraído de Bentley e McDonnell, 1994).

Outra maneira de se calcular esta transformada por um computador digital é através da verdadeira transformada Wavelet discreta, como será mostrada na próxima seção.

### 3.2-A TRANSFORMADA WAVELET DISCRETA (DWT):

Sejam uma função  $f(t)$  e uma função Wavelet  $\Psi(t)$ , ambos pertencentes ao espaço de funções  $L^2(\mathcal{R})$ , e que a função Wavelet obedece a condição de admissibilidade dada por (III.5). Conforme Daubechies (1992) e Simpson (1993), a transformada Wavelet discreta desta função  $f(t)$  é dada por:

$$DWT(m, n) = d_{m,n} = \langle f, \Psi_{m,n} \rangle = \int_{-\infty}^{+\infty} f(t) \cdot \Psi_{m,n} dt \quad (\text{III.11})$$

onde  $\langle \rangle$  indica o produto interno e  $\Psi_{m,n}$  é a versão discreta da função Wavelet, dada por:

$$\Psi_{m,n} = \frac{1}{\sqrt{a_0^m}} \Psi\left(\frac{t}{a_0^m} - nb_0\right) = \frac{1}{\sqrt{a_0^m}} \Psi\left(\frac{t - n \cdot a_0^m \cdot b_0}{a_0^m}\right) \quad (\text{III.12})$$

onde  $a_0$  e  $b_0$  são constantes e as variáveis "m" e "n" pertencem ao conjunto dos Inteiros.

Note-se que o tempo "t" permanece como uma variável contínua. Agora os parâmetros que dão a dilatação e a translação são "m" e "n", respectivamente, e que variam em valores inteiros. As constantes  $a_0$  e  $b_0$  dão a variação da dilatação e do passo de translação, respectivamente.

Em Daubechies (1992) é mostrado que a condição necessária e suficiente para que exista uma reconstrução estável da função  $f(t)$  pela combinação linear dos coeficientes DWT(m,n) é que deva existir um  $A > 0$  e um  $B < \infty$ , "A" e "B" independentes de  $f(t)$  tais que:

$$A\|f\|^2 \leq \sum_{m,n} |\langle f, \Psi_{m,n} \rangle|^2 \leq B\|f\|^2 \quad (\text{III.13})$$

onde  $\|f\|^2$  é o quadrado da norma da função  $f(t)$  que corresponde à energia do sinal.

Isto quer dizer que o quadrado da norma da DWT(m,n) deve ter limites inferior e superior positivos. Quando (III.13) é satisfeita, diz-se que o conjunto de funções base  $\{\Psi_{m,n}\}$  é um "frame" e que "A" e "B" são os limites deste "frame". Quando  $A=B$ , a energia da DWT(m,n) é proporcional à energia do sinal e o "frame" é dito apertado ("tight"), e as wavelets se comportam exatamente como uma base ortonormal. Quando  $A=B=1$ , (III.13) se torna equivalente à relação de Parseval de conservação de Energia da transformada de Fourier (Sheng, 1996).

Para garantir que as componentes  $\Psi_{m,n}$  da base de representação sejam ortonormais entre si, é necessário que:

$$\langle \Psi_{m,n}, \Psi_{m,k} \rangle = \delta(n - k) = \begin{cases} 1, & n = k; \\ 0, & n \neq k. \end{cases} \quad (\text{III.14})$$

Uma escolha comumente feita para "a<sub>0</sub>" e "b<sub>0</sub>" é a de que  $a_0=2$  e  $b_0=1$ . Esta definição é demonstrada em Daubechies (1992). Também escolhe-se que os componentes  $\Psi_{m,n}$  do conjunto de base sejam funções ortonormais e reais. Então (III.12) fica:

$$\Psi_{m,n} = \frac{1}{\sqrt{2^m}} \Psi\left(\frac{x - 2^m \cdot n}{2^m}\right) ; m, n \in \text{inteiros.} \quad (\text{III.15})$$

Assim, tem-se uma escala de dilatação como uma potência de 2 ( $a_m = 2^m$ ), e passos de translação de 1 amostra da escala de dilatação ( $b_n = 2^m \cdot n$ ). A transformada Wavelet discreta feita com funções de base como em (III.16) é também denominada transformada Wavelet diádica ("Dyadic Wavelet Transform", Aldroubi e Unser, 1996). Esta escolha é a mais utilizada em várias análises, por simplificar os cálculos e relações (Daubechies, 1992).

### 3.3- A ANÁLISE MULTI-RESOLUÇÃO:

A discussão a seguir apresenta de forma semelhante à proposta em Simpson (1993) a análise multi-resolução desenvolvida por Mallat (1989).

Seja então um conjunto de funções  $f(t)$  pertencentes ao espaço de funções  $L^2(\mathfrak{R})$ . Para uma melhor compreensão considere uma função  $f(t)$  definida como a combinação linear de  $n$  funções ortogonais cujo espectro de frequência é a função de transferência de um filtro do tipo passa-baixas  $\phi_{m,n}(t)$ , definidas de forma semelhante à expressão (III.15). Então,  $f(t)$  é:

$$f(t) = \sum_n c_{m,n} \cdot \phi_{m,n}(t) \quad (\text{III.16})$$

Os coeficientes  $c_{m,n}$  são calculados conforme (III.6) como:

$$c_{m,n} = \langle f, \phi_{m,n} \rangle \quad (\text{III.17})$$

Seja  $V_0$  o espaço composto de todas as funções  $f(t)$  que são plenamente representadas por (III.16) à dilatação  $m=0$  (ou seja, funções não dilatadas). Logo:

$$f(t)|_{m=0} = \sum_n c_{0,n} \cdot \phi_{0,n}(t) \quad (\text{III.18})$$

e as funções  $\phi_{0,n}(t)$  formam a base deste espaço.

Seja também um outro espaço  $V_1$  composto de todas as funções representadas em um novo nível de dilatação  $m=1$  (cuja duração é o dobro das funções da base anterior). Assim, se  $f(t) \in V_0; f\left(\frac{t}{2}\right) \in V_1$ . Então  $V_1$  constitui-se em um sub-espaço de  $V_0$ . A base deste sub-espaço é formada por todas as funções  $\phi_{1,n}(t)$ . Este processo pode ser sucessivamente repetido, formando-se outros sub-espaços, para diferentes valores de  $m$ . A dilatação dos espaços formados sempre é  $2^m$  e  $V_0 \subset V_1 \subset V_2 \dots$ . Uma representação gráfica destes espaços é mostrada na Figura 3.6.

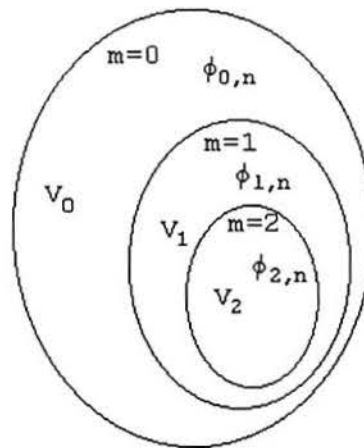


Figura 3.6: Representação dos espaços vetoriais para "m" variando de 0 a 2.

Se  $f(t) \in V_0$ , a reconstrução da função  $f(t)$  em um nível de resolução temporal "m" mais baixo ( $m > 0$ ) é uma aproximação mais pobre da função  $f(t)$ , notado por  $\hat{f}_m(t)$  e dado por:

$$\hat{f}_m(t) = \sum_n c_{m,n} \cdot \phi_{m,n}(t) \quad (\text{III.19})$$

Esta função reconstruída não será igual à função original, mas uma versão filtrada tendo a metade do número de pontos. Generalizando este conceito, a reconstrução de uma função que pertença a um espaço  $V_{m-1}$  em um nível m é dada por:

$$\hat{f}_{m-1}(t) = \sum_n c_{m,n} \cdot \phi_{m,n}(t) \quad (\text{III.20})$$

Para se reconstruir a função original, é necessário acrescentar os componentes que faltam (como  $\phi_{m,n}(t)$  são funções do tipo passa-baixas, aumentando-se a dilatação "m" diminui-se a banda de passagem). Assim, tem-se que:

$$\begin{aligned} f_{m-1}(t) &= \hat{f}_{m-1}(t) + \sum_k d_{m,k} \cdot \psi_{m,k}(t) \\ &= \sum_n c_{m,n} \cdot \phi_{m,n}(t) + \sum_k d_{m,k} \cdot \psi_{m,k}(t) \end{aligned} \quad (\text{III.21})$$

onde  $\psi_{m,k}(t)$  são as funções de base do espaço das funções representado no diagrama pelo anel entre  $V_{m-1}$  e  $V_m$ . Isto é mostrado na Figura 3.7.

As funções  $\psi_{m,n}(t)$  devem ser ortogonais entre si, como também devem ser ortogonais em relação às funções  $\phi_{m,k}(t)$  (Simpson, 1993). Assim:

$$\begin{aligned}
 \langle \phi_{m,n}, \phi_{m,k} \rangle &= \delta(n-k); \\
 \langle \psi_{m,n}, \psi_{m,k} \rangle &= \delta(n-k) \text{ e} \\
 \langle \phi_{m,n}, \psi_{m,k} \rangle &= 0
 \end{aligned}
 \tag{III.22}$$

onde  $\delta(n-k) = \begin{cases} 1, n=k; \\ 0, n \neq k \end{cases}$

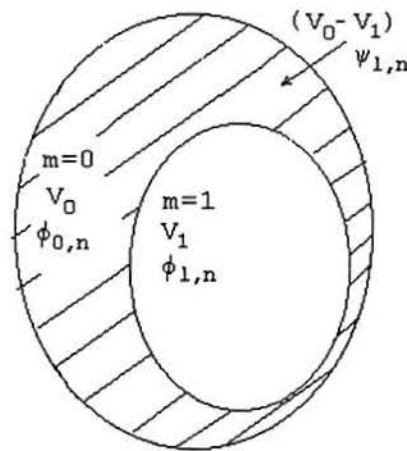


Figura 3.7: Representação dos espaços de funções  $V_0$  e  $V_1$ . A área hachurada representa o espaço vetorial cuja base é  $\psi_{m,k}(t)$ .

Em termos de processamento de sinais, se  $\phi_{m-1,n}(t)$  representa um filtro passa-baixas, então, devido às condições (III.22),  $\psi_{m,k}(t)$  representará um filtro passa-faixas onde estão as componentes espectrais que faltam para a completa reconstrução do sinal em um nível  $(m-1)$ .

As funções  $\phi(t)$  são denominadas funções de escalonamento ("Scaling Functions") e as funções  $\psi(t)$  são denominadas funções Wavelets primárias.

O cálculo dos coeficientes " $c_{m,n}$ " e " $d_{m,k}$ " saem também das condições de ortogonalidade (III.22) e são dados por:

$$\begin{aligned} c_{m,n} &= \langle f_m, \phi_{m,n} \rangle = \langle f_{m-1}, \phi_{m,n} \rangle \\ d_{m,n} &= \langle f_m, \psi_{m,n} \rangle = \langle f_{m-1}, \psi_{m,n} \rangle \end{aligned} \quad (\text{III.23})$$

Como os coeficientes também são funções contidas nos espaços de funções, pode-se representá-los da seguinte forma:

$$c_{m,k} = \sum_n c_{m-1,n} \cdot h(n-2k) \quad (\text{III.24})$$

$$\text{onde } h(n-2k) = \langle \phi_{m-1,n}, \phi_{m,k} \rangle = \frac{1}{\sqrt{2}} \int_{-\infty}^{+\infty} \phi(t-n) \phi\left(\frac{t}{2}-k\right) dt$$

De modo similar, tem-se que:

$$d_{m,k} = \sum_n d_{m-1,n} \cdot g(n-2k) \quad (\text{III.25})$$

$$\text{onde } g(n-2k) = \langle \phi_{m-1,n}, \psi_{m,k} \rangle = \frac{1}{\sqrt{2}} \int_{-\infty}^{+\infty} \phi(t-n) \psi\left(\frac{t}{2}-k\right) dt$$

Uma propriedade das funções  $h(n-2k)$  e  $g(n-2k)$ , demonstrada em Simpson (1993), é dada por:

$$h(n-2k) = \langle \phi_{m-1,n}, \phi_{m,k} \rangle = \langle \phi_{0,n}, \phi_{1,k} \rangle \quad (\text{III.26})$$

$$g(n-2k) = \langle \phi_{m-1,n}, \psi_{m,k} \rangle = \langle \phi_{0,n}, \psi_{1,k} \rangle \quad (\text{III.27})$$

As propriedades (III.26) e (III.27) mostram que as funções  $h(n-2k)$  e  $g(n-2k)$  são constantes em qualquer nível de dilatação, permitindo que se elabore um algoritmo rápido para o cálculo da transformada Wavelet discreta, pois não é necessário se recalcularem estas funções para cada novo nível de dilatação. Através de (III.24), (III.25), (III.26) e (III.27)



pode-se calcular os coeficientes de um nível de dilatação  $m$  a partir dos coeficientes do nível de dilatação imediatamente inferior  $m-1$ .

Para a base ortonormal mais simples, denominada de "base de Haar", a "scaling function" é definida por:

$$\phi(t) = \begin{cases} 1, & 0 \leq t < 1; \\ 0, & t \geq 1 \text{ e } t < 0 \end{cases} \quad (\text{III.28})$$

e sua função Wavelet é:

$$\psi(t) = \begin{cases} 1, & 0 \leq t < 1/2; \\ -1, & 1/2 \leq t < 1; \\ 0, & \text{qualquer outro } t. \end{cases} \quad (\text{III.29})$$

Os valores de  $h(n)$  e  $g(n)$  para esta base, calculados por (III.26) e (III.27) para 2 pontos ( $0 \leq n \leq 1$ ), são:

$$\begin{aligned} h(0) &= \frac{1}{\sqrt{2}}; & h(1) &= \frac{1}{\sqrt{2}} \\ g(0) &= \frac{1}{\sqrt{2}}; & g(1) &= -\frac{1}{\sqrt{2}} \end{aligned} \quad (\text{III.30})$$

É mostrado que os coeficientes  $h(n)$  e  $g(n)$  podem ser relacionados com os coeficientes de filtros digitais passa-baixas e passa-altas do tipo "Quadrature Mirror Filters" (QMF) (Mallat, 1989, Daubechies, 1992, Strang e Nguyen, 1996). O processo de filtragem e reconstrução de sinais com esta classe de filtros digitais foi inicialmente apresentado no artigo de Esteban e Galland (1977) com o nome de "Sub-band Coding", onde é mostrado que um sinal pode ser decomposto em 2 sub-bandas, uma passa-baixas e uma passa-altas, ambas com metade da banda original, e respectiva decimação por um fator de 2 ("down-sampling"). É

mostrado que para esta classe de filtros, é possível se reconstruir perfeitamente o sinal original, inserindo zeros nos sinais resultantes ("up-sampling") e passando os sinais resultantes pelos filtros novamente (processo de interpolação) e somando-se o resultado. Isto é mostrado na Figura 3.8.

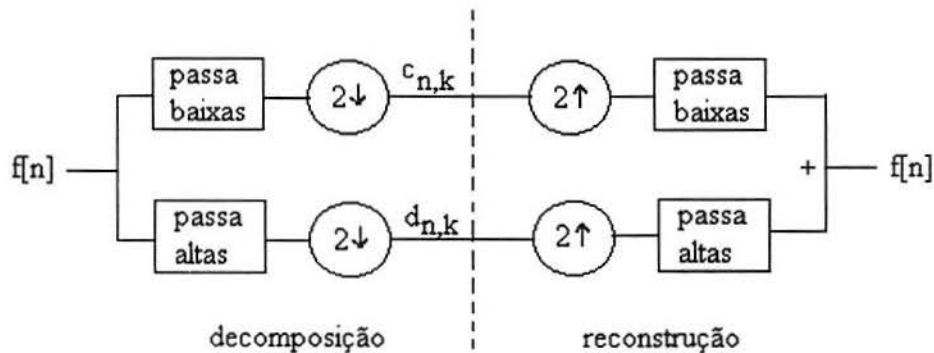


Figura 3.8: Representação do esquema de codificação em sub-bandas ("Sub-Band Coding").

O algoritmo usa uma estrutura do tipo pirâmide (ou árvore), onde, através da convolução discreta dos coeficientes "c" e "d" de um nível com os filtros "QMF", são calculados os coeficientes do nível seguinte, fazendo-se então a devida sub-amostragem (os coeficientes "c" e "d" de cada novo nível terão a metade dos pontos dos coeficientes do nível anterior).

Para a transformada inversa mostra-se que os coeficientes " $c_{m,n}$ " podem ser calculados com base nos coeficientes de um nível de dilatação superior " $m+1$ " por:

$$c_{m,n} = \sum_k c_{m+1,n} \cdot h(n - 2k) + \sum_k d_{m+1,k} \cdot g(n - 2k) \quad (\text{III.31})$$

Neste caso também é utilizado o algoritmo pirâmide, onde os coeficientes "c" e "d" são combinados para formar os coeficientes "c" do nível inferior (de resolução maior). Os

coeficientes "d" de todos níveis calculados e o coeficiente "c" do último nível calculado devem ser armazenados, de alguma forma, para que se possa obter as reconstruções em níveis de maior resolução.

Uma completa explanação sobre os algoritmos é mostrada em Cody (1992), Simpson (1993), Bentley e McDonnell (1994), Aldroubi e Unser (1996) e Poularikas (1996).

#### 4 MATERIAIS E MÉTODOS:

##### 4.1-AVALIAÇÃO E GRAVAÇÃO DAS VOZES NORMAIS E PATOLÓGICAS:

Inicialmente, os pacientes foram avaliados pelo Dr. Nédio Steffen, do Hospital da PUC-RS. A avaliação laringológica dos pacientes foi feita através de videolaringoscopia direta e de videoestrobolaringoscopia. Com esta avaliação, obteve-se o diagnóstico preciso da existência e tipo de patologia (ou de pregas vocais normais) do paciente (Kuhl, 1982, Hirano e Bless, 1993). A seguir, foram gravadas as amostras das vozes dos pacientes emitindo a vogal sustentada /a/, em tom natural e o mais confortável possível (a vogal cujas estruturas do trato vocal se mantêm mais estáveis; Yumoto e Gold, 1982). As gravações foram realizadas no consultório do referido especialista, em uma sala acusticamente adequada para tal. Para análise do sinal, digitalizou-se 3 segundos do trecho auditivamente mais estável da voz dos pacientes. Ao todo, foram avaliadas e gravadas as vozes de 64 pessoas, entre casos normais e patológicos. A descrição dos pacientes pode ser vista na Tabela 4.1.

*Tabela 4.1: Descrição dos pacientes gravados.*

No.	Pacientes	Descrição	Patologia
1	PAC01	Professora de canto, voz normal.	Micronódulo nas pregas vocais
2	PAC02	Cantor nativista, vibrato na emissão, voz normal.	Pregas normais
3	PAC04	Cantor da noite, voz muito rouca.	Laringite crônica, possível uso de drogas
4	PAC05	Mulher, voz soprosa.	Sulco vocálico
5	PAC06	Mulher, voz soprosa.	Edema pós-cirúrgico
6	PAC08	Rapaz pré-cirúrgico.	Voz supra-glótica (ventricular)
7	PAC09	Senhora, 25 dias sem falar	Pós-laringite por abuso vocal
8	PAC10	Senhor de 78 anos, voz quase normal.	Paralisia unilateral bem compensada
9	PAC11	Senhor pós-operatório de 6 anos, feita neocorda.	Laringectomia fronto-lateral
10	PAC13	Senhora, voz rouca e mal compensada.	Paralisia unilateral
11	PAC14	Senhora pós-operatória de cisto, em tratamento fonoterápico.	Sulco vocálico
12	PAC15	Senhora meia idade, fumante.	Leucoplasia+laringite crônica
13	PAC16	Moça c/ voz normal.	Rinite e Sinusite alérgica
14	PAC17	Senhor 66 anos, pós-operatório+radioterapia.	CA na laringe, lesão prega vocal direita
15	PAC18	Senhora fonastênica, cansaço na voz.	Pregas normais
16	PAC19	Estudante fonoaudiologia, voz normal.	Pregas normais
17	PAC21	Menina de 16 anos, voz rouca.	Cisto pequeno na prega vocal direita, secreção na traquéia
18	PAC22	Senhora , pós-operatório de 3 semanas.	Leucoplasia
19	PAC23	Senhora, fumante, voz rouca e grave.	Edema de Reike
20	PAC24	Mulher, voz muito soprosa e rouca.	Cisto, uma prega vocal inchada e sem ondas
21	PAC25	Senhor, pós operatório 45 laringectomia parcial.	Laringectomia fronto-lateral
22	PAC26	Mulher, voz limpa, tratamento fonoterápico.	Disfonia funcional+micro-nódulo

continuação da Tabela 4.1

No.	Pacientes	Descrição	Patologia
23	PAC27	Mulher, voz limpa mas levemente soprosa, tratamento fonoterápico.	Disfonia funcional
24	PAC28	Mulher, voz limpa, trat. fonoterápico.	Disfonia funcional
25	PAC29	Mulher, voz limpa, fina e levemente rouca, tratamento fonoterápico.	Disfonia funcional
26	PAC30	Mulher, voz grave e rouca.	Disfonia funcional+laringite
27	PAC31	Senhor, voz extremamente disfônica.	Paralisia prega vocal direita
28	PAC32	Senhora, voz grave e rouca, fumante.	Edema de Reike+laringite crônica
29	PAC34	Senhor, voz com sibilção.	Papilomas
30	PAC35	Senhora com voz grave e rouca.	Edema de Reike-pregas frouxas
31	PAC36	Senhora com voz grave.	Edema de Reike+formações polipóides
32	PAC37	Adolescente, voz levemente soprosa, pós-cirúrgica de papiloma recorrente.	Pregas vocais limpas
33	PAC41	Senhor meia idade, voz rouca.	neoplasia na laringe
34	PAC42	Senhor pós-cirúrgico de muito tempo, revisão. Voz muito rouca.	Laringectomia fronto-lateral
35	PAC43	Senhor 59 anos.	Paralisia unilateral prega esquerda
36	PAC44	Senhor, voz normal.	pregas vocais normais
37	PAC45	Mulher, voz normal.	pregas vocais normais
38	PAC46	Mulher, com voz muito soprosa, em fonoterapia	Paralisia prega vocal direita
39	PAC47	Senhor com voz normal, excelente gravação	pregas vocais normais
40	PAC48	Menina com voz normal.	pregas vocais normais
41	PAC49	Senhor com voz normal	pregas vocais normais
42	PAC50	Senhora, voz aparentemente normal.	micronódulo nas pregas vocais
43	PAC51	Estudante fonoaudiologia	Laringite + Uso Aerolyn
44	PAC53	Moça, voz rouca e soprosa.	Em teste
45	PAC54	Rapaz com voz normal (talvez alteração ressonância).	Linfonodo no pescoço
46	PAC55	Cantor nativista, Padrão Fry, voz normal.	pregas vocais normais
47	PAC56	Senhora, voz normal.	Prega vocal direita vascularizada
48	PAC57	Senhor, voz extremamente disfônica.	Laringectomia fronto-lateral
49	PAC58	Senhor, com quebra de fonação, voz rouca.	Granuloma abaixo prega vocal direita
50	PAC59	Cantor, voz rouca e de baixa intensidade, histórico de diversas patologias nas pregas vocais.	Leucoplasia+Irritação
51	PAC60	Senhor, voz rouca.	Laringectomia fronto-lateral
52	PAC61	Senhora voz normal.	pregas vocais normais
53	PAC62	Senhor fumante, pouco tempo de fonação	Leucoplasia+ Laringite crônica
54	PAC63	Senhora, voz extremamente disfônica.	Leucoplasia+ Laringite crônica
55	PAC64	Estudante fonoaudiologia, voz normal, nervosa	pregas vocais normais
56	PAC66	Senhor idoso	Presbifonia
57	PAC67	Estudante fonoaudiologia, ex-fumante	Secreção nas pregas
58	PAC68	Estudante fonoaudiologia, voz normal	pregas vocais normais
59	PAC69	Rapaz com diplofonia	Pólipo grande na prega vocal
60	PAC70	Senhor com voz normal, pós-cirúrgico de tireóide.	Pregas vocais limpas
61	PAC71	Criança.	Nódulo bilateral
62	PAC72	Senhora, pós-operatória de 3 dias.	pregas vocais normais
63	PAC73	Senhor, voz áspera.	Ceratose+Laringite crônica
64	PAC74	Adolescente, voz levemente soprosa.	Fenda Posterior aumentada

## 4.2-O EQUIPAMENTO UTILIZADO:

Para a gravação dos sinais de voz, foi utilizado um microfone dinâmico marca LeSon, modelo LM-58, com resposta em frequência linear de 50 Hz a 17 kHz, conforme o fabricante. O sinal do microfone foi amplificado para um sinal de linha ("line input") por um circuito baseado em 2 amplificadores operacionais "low noise" do tipo TL072, na configuração não-inversor, em cascata, com alimentação por bateria de 9V, e com resposta linear dentro da faixa do microfone. O sinal obtido é levado então às entradas de um gravador cassete profissional Technics model RS-B105, sendo que tanto para gravação como para reprodução foi usado sistema de redução de ruído da fita do tipo Dolby-NR, do próprio gravador. Para fins de teste, também foram feitas gravações sem o uso do sistema de redução de ruído. As fitas cassetes utilizadas foram fitas do tipo II (óxido de cromo) da Sony do tipo UX60 (fitas de 60 minutos), por terem melhor performance e serem mais resistentes que as fitas de 90 minutos. Esta etapa foi feita em campo. O diagrama de blocos desta etapa é mostrado na Figura 4.1.

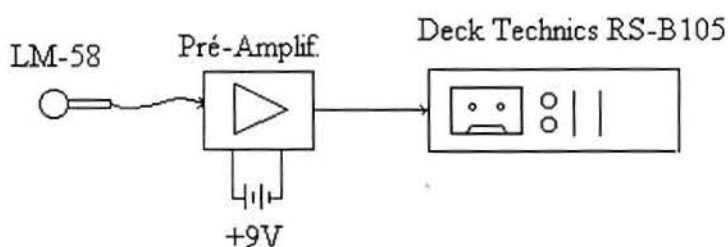


Figura 4.1: Diagrama de blocos do sistema usado nas gravações.

Para a reprodução, conversão analógica/digital e pré-processamento dos sinais de voz, foi usado o seguinte equipamento para a análise de sinais do laboratório de processamento de sinais e imagens (LaPSI-UFRGS): duplo deck Tascam 302, mesa de mixagem Mackie 1402 VLZ, um conversor analógico-digital (A/D) e digital-analógico (D/A), uma placa de processamento de sinais usando o processador DSP TMS320C25 da Texas Instruments, e um microcomputador hospedeiro do tipo PC-Pentium 133. Ambos a placa DSP e o conversor A/D-D/A foram desenvolvidos no Laboratório de Processamento de Sinais e Imagens (LaPSI)

do IEE-UFRGS. Foram usados 2 procedimentos para a aquisição dos sinais: o primeiro, através da técnica de sobre-amostragem ("oversampling") do sinal, filtragem passa-baixas digital e após, decimação ("downsampling") do sinal para a frequência de amostragem desejada; o segundo, por amostragem na frequência desejada com um filtro passa-baixas analógico "anti-aliasing" na entrada do conversor A/D e sem posterior filtragem digital. Os diagramas de bloco destes procedimentos são mostrados nas Figuras 4.2 e 4.3, respectivamente, e cada bloco será explicado com mais detalhes a seguir.

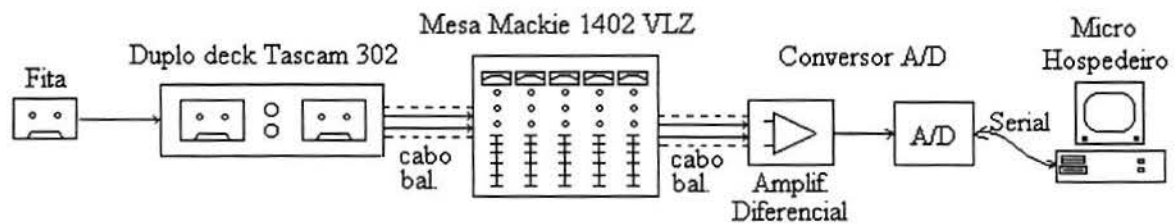


Figura 4.2: Diagrama de blocos do sistema usado para o primeiro procedimento de aquisição. A placa de DSP se encontra dentro do micro hospedeiro.

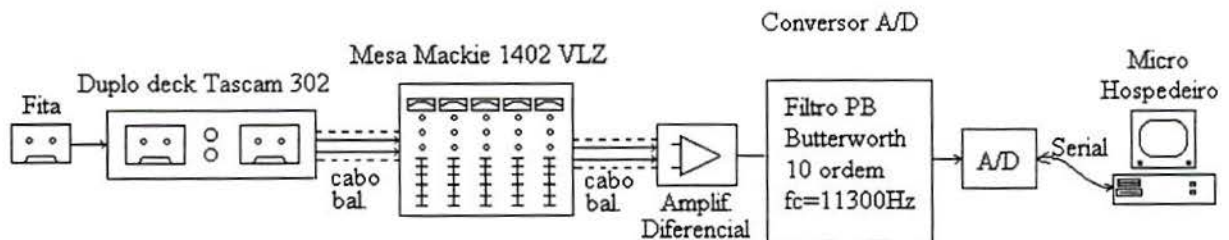


Figura 4.3: Diagrama de blocos do sistema usado para o segundo procedimento de aquisição. A placa de DSP se encontra dentro do micro hospedeiro, embora não seja usada.

Para a reprodução dos sinais gravados, foi utilizado o seguinte equipamento:

- 1 Double DECK TASCAM 302;
- 1 Mesa de Mixagem Mackie 1402 VLZ;

O sinal reproduzido da fita sai do deck para uma das entradas estéreo da mesa de mixagem Mackie e, de uma saída balanceada desta, o sinal é conectado ao conversor A/D, utilizando-se sua entrada balanceada. Estas precauções foram tomadas para minimizar a indução de ruído pelos cabos e conexões, antes da conversão A/D. Este sistema pode ser visto na Figura 4.4.



*Figura 4.4: Sistema do Laboratório de processamento de sinais e imagens (LaPSI-UFRGS) usado para a reprodução dos sinais gravados .*



Para a conversão analógica/digital dos sinais de voz, foi utilizado um sistema conversor analógico-digital (A/D) e digital-analógico (D/A) externo ao computador e alimentado por fonte independente desenvolvido no LAPSI-IEE-UFRGS, com as seguintes características:

- Conversor A/D de 14 bits, e capacidade de até 80.000 amostras por segundo (frequência de amostragem,  $f_s$ ), baseado no integrado AD7871 da Analog Devices.
- 2 saídas D/A de 8 bits, baseadas no integrado PM-7226, da Analog Devices.
- 1 porta de comunicação serial RS-232 com a placa de processamento de sinais DSP.
- 3 entradas de sinal: 2 normais (RCA) e 1 do tipo diferencial balanceada.

Utilizou-se a entrada diferencial balanceada, para reduzir-se a indução de ruído nos cabos entre a mesa e a entrada do conversor. Para o segundo método de aquisição, foi inserido um filtro passa-baixas analógico do tipo Butterworth de 10<sup>a</sup> ordem, com frequência de corte de 11300Hz, para limitar a banda do sinal a ser adquirido, e cujo diagrama esquemático é apresentado na Figura 4.5.

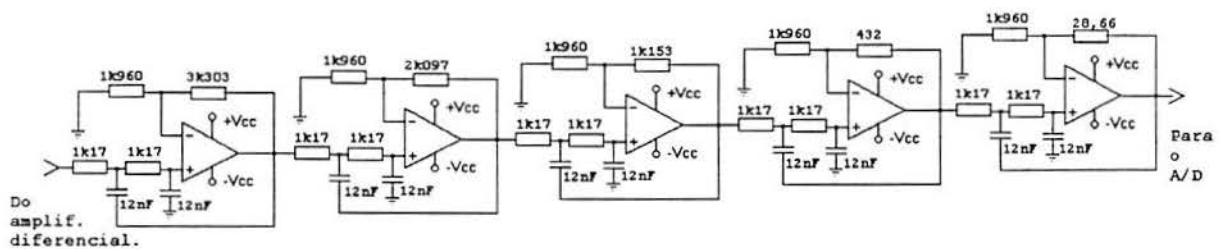


Figura 4.5: Diagrama esquemático do filtro passa-baixas analógico do tipo Butterworth de 10<sup>a</sup> ordem. A frequência de corte calculada é de 11300Hz.

Para o pré-processamento do sinal adquirido foi usada a placa de DSP e seu microcomputador hospedeiro, do tipo IBM-PC Pentium 133.

As principais características da placa DSP são:

- Processador TMS320C25, com "clock" de 40 MHz e ciclo de instruções de 100 ns.
- Memória de 32 kBytes, permitindo uso conjunto de RAM e EPROM.
- Barramento para expansão de memória.
- 2 interfaces de entrada e saída, uma com 8 e outra com 16 bits.
- 1 interface de comunicação serial, do tipo RS-232.
- Interface de comunicação com o microcomputador hospedeiro, no padrão PC-AT, com acesso em 16 bits.

A placa DSP pode ser vista na Figura 4.6 e o sistema de conversão AD (junto com a placa DSP) pode ser visto na Figura 4.7.

Para o primeiro procedimento de aquisição, foi feita a sobre-amostragem do sinal ("oversampling") em 3 vezes a frequência de amostragem ( $f_s$ ) desejada, filtragem do sinal com um filtro digital do tipo passa-baixas pela DSP e posterior decimação do sinal para a frequência de amostragem desejada (de 3 amostras 2 são desprezadas). Para um sinal cuja banda desejada é de 12kHz, a frequência de amostragem mínima pelo critério de Nyquist seria de 24 kHz. No sistema proposto, o sinal é amostrado em 75 kHz, filtrado digitalmente e então sub-amostrado para a taxa  $f_s=25\text{kHz}$ , obtendo-se uma banda de sinal efetiva de 12,25 kHz. O filtro passa-baixas digital utilizado é um filtro FIR ("Finite Impulse Response") do tipo janela Hamming de 63 pontos, com frequência de corte de  $0,16.f_s$  (12kHz).

Para o segundo procedimento de aquisição, a frequência de amostragem foi de 50kHz, sendo que a banda do sinal foi previamente limitada pelo filtro analógico passa-baixas, anteriormente descrito, e não foi feito nenhum processamento digital posterior.

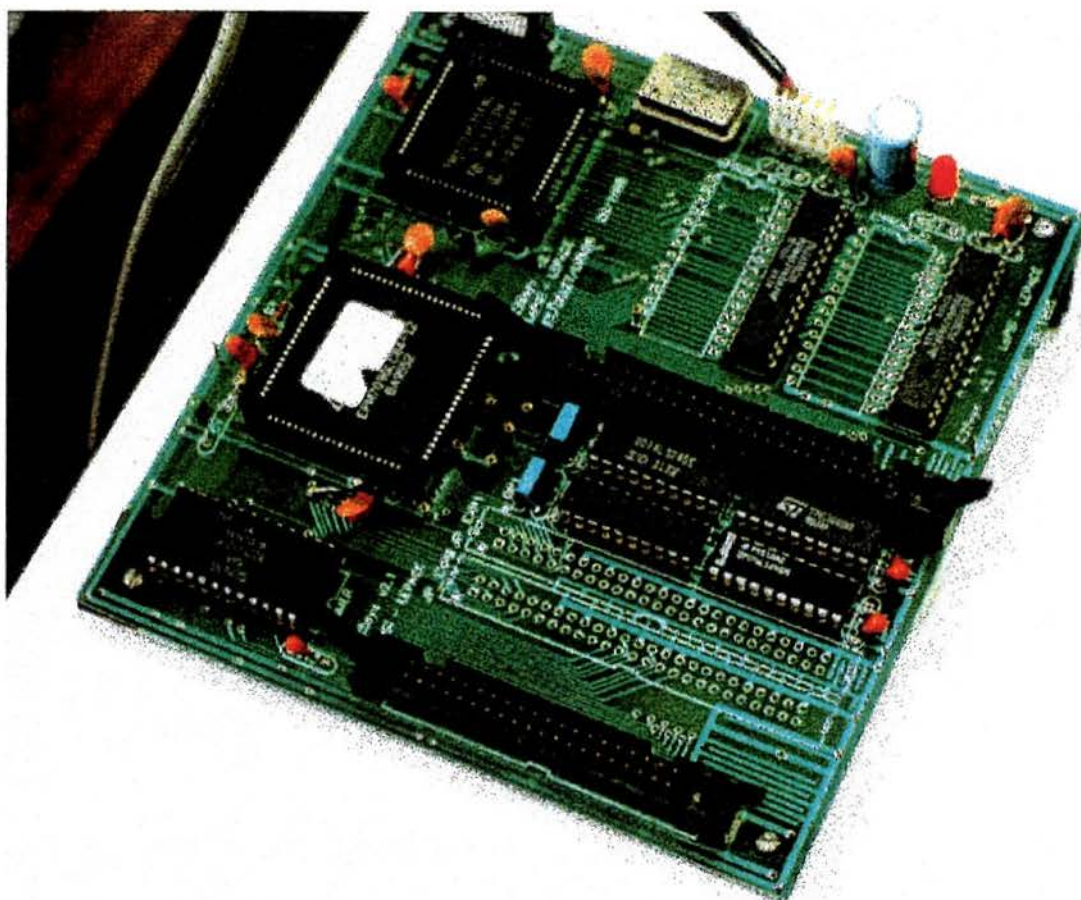


Figura 4.6: Vista da placa DSP usada no pré-processamento dos sinais de voz.

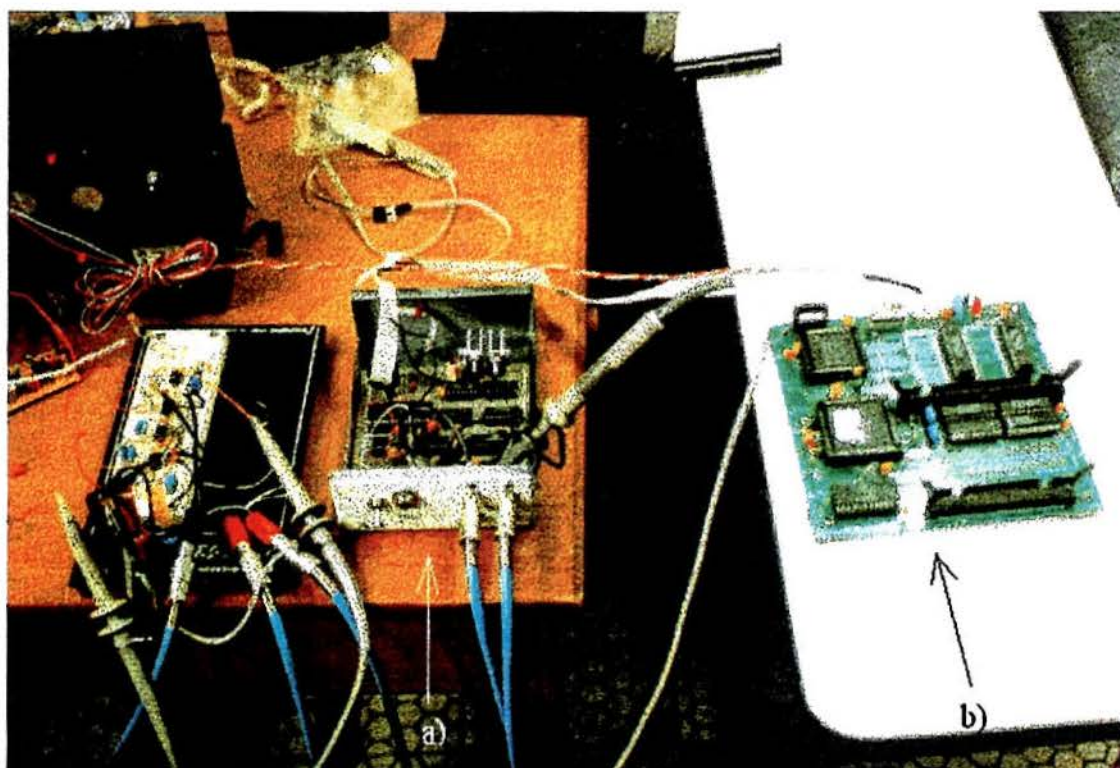


Figura 4.7: Sistema de conversão A/D dos sinais de voz. Pode-se ver o a) conversor A/D aqui já interligado com a b) placa DSP.

#### 4.3-PROGRAMAS COMPUTACIONAIS UTILIZADOS:

Para a análise dos sinais já digitalizados, foi desenvolvido no Laboratório de Medições Mecânicas do PROMEC-UFRGS um aplicativo DOS com a capacidade de exibição, edição e processamento do sinal. A transformada Wavelet discreta, calculada através do algoritmo sugerido em Cody (1992), foi incorporada a este aplicativo, e cujas equações para o cálculo dos coeficientes dos filtros digitais QMF  $h(n)$  e  $g(n)$  são dadas por:

$$h(0) = \frac{1}{4\sqrt{2}}[(1 + \cos \alpha + \operatorname{sen} \beta)(1 - \cos \alpha - \operatorname{sen} \beta) + 2 \operatorname{sen} \beta \cdot \cos \alpha] \quad (\text{IV.1})$$

$$h(1) = \frac{1}{4\sqrt{2}}[(1 - \cos \alpha + \operatorname{sen} \beta)(1 + \cos \alpha - \operatorname{sen} \beta) - 2 \operatorname{sen} \beta \cdot \cos \alpha] \quad (\text{IV.II})$$

$$h(2) = \frac{1}{2\sqrt{2}}[1 + \cos(\alpha - \beta) + \operatorname{sen}(\alpha - \beta)] \quad (\text{IV.III})$$

$$h(3) = \frac{1}{2\sqrt{2}}[1 + \cos(\alpha - \beta) - \operatorname{sen}(\alpha - \beta)] \quad (\text{IV.IV})$$

$$h(4) = \frac{1}{\sqrt{2}} - h(0) - h(2) \quad (\text{IV.V})$$

$$h(5) = \frac{1}{\sqrt{2}} - h(1) - h(3) \quad (\text{IV.VI})$$

e

$$g(n) = (-1)^n \cdot h(5 - n) \quad (\text{IV.VII})$$

onde variando-se  $\alpha$  e  $\beta$ , obtêm-se diferentes bases de Wavelets. Note-se que fazendo  $\alpha=\beta$ , obtêm-se os coeficientes para a base de Haar, anteriormente definidos em III.30.

Na análise dos sinais pela análise multi-resolução, foi escolhida a base de Haar, por ser a base ortonormal discreta real mais simples e de menor duração que se pode obter, e por esta

ser a base com menos influência de fase sobre os resultados (Antonini *et al.*,1992). Assim tem-se um ponto de partida para a análise dos resultados, verificando-se se a base mais simples pode ser utilizada para a análise de sinais de voz. Os gráficos desta base são mostrados na Figura 4.8.

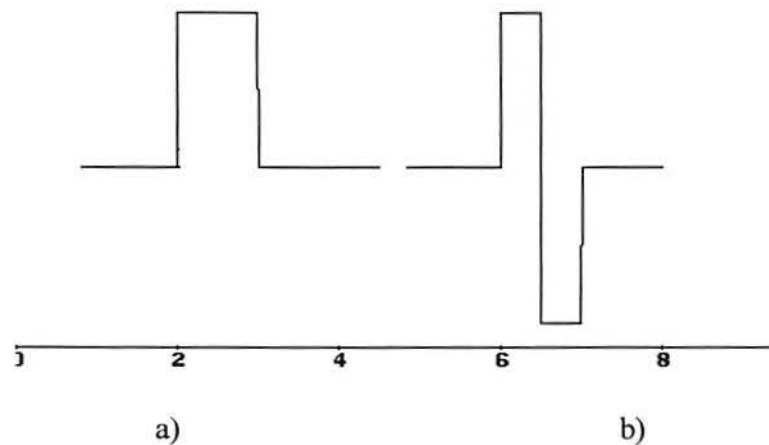


Figura 4.8-Representações da (a)Scaling Function e (b)Wavelet, da base de Haar.

Para a análise estatística dos resultados, foi utilizado o programa Excel 97, da Microsoft.

#### 4.4-MÉTODO UTILIZADO:

Inicialmente, para verificar se o sistema Dolby de redução de ruído de fitas alterava a fidelidade do sinal gravado, foi gerado ruído branco e este foi gravado ao mesmo tempo em 2 fitas cassete de mesmo tipo (Sony UX60, Type II), uma com a presença do sistema Dolby NR (ou Dolby B) e outra sem este sistema. Isto foi feito utilizando-se o duplo deck profissional TASCAM 302. A seguir, as gravações foram digitalizadas pelo sistema descrito na seção 4.2, mas com uma frequência de amostragem de 40 kHz (frequência de 80kHz com sobre-amostragem de 2 vezes), provendo uma banda efetiva de sinal de 20kHz. De ambos trechos

foram recortados 60.000 pontos e estimados os espectros de potência através do processo de Welch (Kay e Marple, 1981; Bendat E Piersol, 1986; Oppenheim e Schaffer, 1989), de média dos quadrados das FFTs do sinal segmentado por janela e com superposição ("overlapping") de 50% dos segmentos do sinal. Neste caso, as FFTs são de 4096 pontos, com segmentação dos trechos de sinal pela janela Hanning e deslocamentos da janela de segmentação do sinal de 2048 pontos. Os espectros de potência resultantes em escala linear e logarítmica ( $10.\log$ ) e também a subtração de um espectro do outro, em escala logarítmica são mostrados no próximo capítulo.

Todos sinais de voz foram digitalizados como explanado na seção 4.2. A seguir é recortado um trecho de 16384 pontos do trecho digitalizado, para o primeiro procedimento de aquisição e 32768 pontos, para o segundo procedimento. Para os sinais adquiridos pelo primeiro procedimento, é calculada a DWT para 5 níveis diferentes de dilatação ( $0 \leq m \leq 4$ ). Para fins de comparação é também calculada a transformada discreta de Fourier, via FFT. Os resultados da DWT para os pacientes PAC10, PAC15, PAC26, PAC29, PAC35, PAC45, PAC48, PAC49, PAC55, PAC56, PAC59, PAC62, PAC63, PAC67, PAC68 e PAC74, são apresentados no próximo capítulo, enquanto os gráficos da DWT dos outros pacientes são apresentados no Anexo 1. Os gráficos dos módulos das FFT são apresentados no Anexo 2.

Para os sinais obtidos pelo segundo procedimento, também é calculada a transformada discreta de Wavelet para 6 níveis diferentes de dilatação ( $0 \leq m \leq 5$ ). Para fins de investigação, os sinais dos paciente PAC10, PAC15, PAC26, PAC29, PAC35, PAC45, PAC48, PAC49, PAC55, PAC56, PAC59, PAC62, PAC63, PAC67, PAC68 e PAC74, são também segmentados por uma janela Hamming e após, calculada a DWT para 6 níveis de dilatação. Os gráficos obtidos da DWT destes pacientes acima descritos, para as segmentações com

janelas retangular e Hamming são apresentados no próximo capítulo, enquanto o resto dos gráficos da DWT, para segmentação com janela retangular são apresentados no Anexo 3.

As frequências das bandas de passagem dos coeficientes da DWT, para os 2 procedimentos de aquisição, são mostrados nas Tabelas 4.2 e 4.3:

*Tabela 4.2: Bandas de frequência dos coeficientes da DWT para o primeiro procedimento ( $f_s=25000$ ).*

Coeficiente:	Banda Frequência [Hz]:	Coeficiente:	Banda Frequência [Hz]:
C0	(0-6250)	D0	(6250-12500)
C1	(0-3125)	D1	(3125-6250)
C2	(0-1562,5)	D2	(1562,5-3125)
C3	(0-781,25)	D3	(781,25-1562,5)
C4	(0-390,62)	D4	(390,62-1562,5)

*Tabela 4.3: Bandas de frequência dos coeficientes da DWT para o segundo procedimento ( $f_s=50000$ ).*

Coeficiente:	Banda Frequência [Hz]:	Coeficiente:	Banda Frequência [Hz]:
C0	(0-12500)	D0	(12500-25000)
C1	(0-6250)	D1	(6250-12500)
C2	(0-3125)	D2	(3125-6250)
C3	(0-1562,5)	D3	(1562,5-3125)
C4	(0-781,25)	D4	(781,25-1562,5)
C5	(0-390,62)	D5	(390,62-1562,5)

Procurando investigar uma possível influência dos sistemas de amostragem no cálculo da DWT, foram gerados pelo MATLAB (MATHWORK Co.) 2 sinais: primeiro, um sinal aleatório e segundo, uma senóide de frequência crescente no tempo (0 a 8kHz). Ambos foram gerados para uma frequência de amostragem de 50000Hz, e então, calculadas as suas DWTs, para 6 níveis de dilatação, com a mesma base usada para os outros sinais. Os resultados são mostrados no capítulo seguinte.

São também calculados os índices "High Frequency Power Ratio" (Shoji, 1992a) e o índice Br (Fukazawa *et al.*, 1988), ambos índices de soproidade.

Para a frequência de amostragem escolhida do primeiro procedimento de aquisição, o cálculo do primeiro nível de dilatação da transformada Wavelet discreta (coeficientes " $c_0$ " e " $d_0$ ") divide o espectro do sinal bem ao meio, na frequência de 6,25kHz, como é mostrado na Figura 4.4. Esta frequência vem de acordo com a frequência de corte ótima usada no cálculo do índice "high frequency power ratio" (Shoji, 1992a), explanando o por quê da escolha desta frequência de amostragem. Para o segundo procedimento de aquisição, os coeficientes correspondentes para a comparação acima são " $c_1$ " e " $d_1$ ".

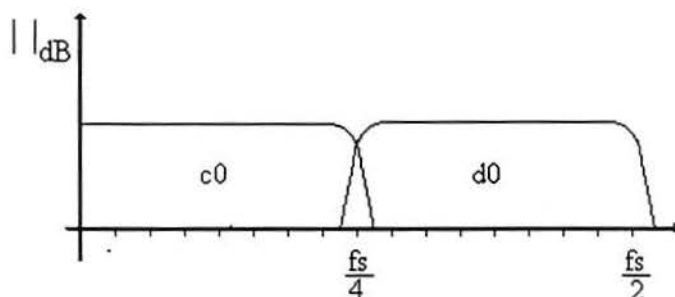


Figura 4.4- Representação do módulo das funções de transferência dos filtros digitais equivalentes associados aos coeficientes da DWT " $c_0$ " e " $d_0$ " para o primeiro procedimento de aquisição. " $f_s$ " é a frequência de amostragem.

Para o primeiro procedimento de aquisição é calculado o quociente entre o valor rms do coeficiente " $d_0$ " e o valor rms do coeficiente " $c_0$ ", para se comparar com o índice "high frequency power ratio". É também calculado o quociente entre o valor rms da segunda derivada do coeficiente " $d_0$ " e o valor rms do sinal, para ser comparado com o índice Br.

Para fins de comparação dos procedimentos de aquisição, para o segundo procedimento de aquisição são calculados: o quociente entre o valor rms do coeficiente " $d_1$ " e



o valor rms do coeficiente " $c_1$ ", para se comparar com o índice "high frequency power ratio"; e o quociente entre o valor rms da segunda derivada do coeficiente " $d_1$ " e o valor rms do sinal, para ser comparado com o índice Br.

Os valores obtidos destes 2 quocientes para todos pacientes são arranjados em uma planilha do aplicativo Excel 97 e então são comparados estatisticamente com os valores dos 2 índices de soproidade, também inseridos nesta planilha, através de diagramas de espalhamento e cálculo do coeficiente de correlação, calculados por este mesmo programa. A planilha com os resultados para o primeiro procedimento de aquisição é apresentada no Anexo 4, e para o segundo procedimento de aquisição, no Anexo 5.

No Anexo 6, é apresentada a teoria sobre o coeficiente de correlação e seus limites de significância.

## 5 RESULTADOS OBTIDOS:

Neste capítulo são mostrados todos resultados obtidos. Nas Figuras 5.1 e 5.2 são apresentados os espectros de potência em escala linear do sinal de ruído branco, gravado e reproduzido sem Dolby, e gravado e reproduzido com Dolby, respectivamente. Nas Figuras 5.3 e 5.4 são mostrados os espectros de potência, em [dBs], do sinal de ruído branco gravado e reproduzido sem Dolby, e gravado e reproduzido com Dolby, respectivamente. Na Figura 5.5 é apresentado o gráfico da diferença entre as as curvas das Figuras 5.3 e 5.4, em [dBs]. Nas Figuras 5.6 a 5.21, são mostrados os gráficos dos trechos de sinal de voz e os 5 níveis de decomposição da DWT, para a base Haar, dos sinais adquiridos pelo primeiro procedimento, dos pacientes PAC10, PAC15, PAC26, PAC29, PAC35, PAC45, PAC48, PAC49, PAC55, PAC56, PAC59, PAC62, PAC63, PAC67, PAC68 e PAC75. Nas Figuras 5.22 a 5.37, são mostrados os gráficos dos trechos de sinal de voz e os 6 níveis de decomposição da DWT, para a base Haar, dos sinais adquiridos pelo segundo procedimento, dos mesmos pacientes acima, com segmentação por janela retangular, e nas Figuras 5.38 a 5.53, são mostrados os gráficos dos trechos de sinal de voz e os 6 níveis de decomposição da DWT, para a base Haar, dos sinais adquiridos pelo segundo procedimento, dos mesmos pacientes acima, mas para segmentação por janela Hamming. Na Figura 5.54 é mostrado o sinal aleatório gerado pelo MATHLAB, e seus 6 níveis de decomposição da DWT, com base de Haar. Na Figura 5.55 é apresentada a senóide crescente gerada pelo MATHLAB, e seus 6 níveis de decomposição da DWT, com base de Haar. Na Figura 5.56 é apresentada a mesma senoide crescente gerada pelo MATHLAB, mas segmentada por uma janela Hamming, e seus 6 níveis de decomposição da DWT, com base de Haar. Finalmente, nas Figuras 5.57 e 5.58 são plotados os gráficos dos diagramas de dispersão ("Scatter Plots") entre o índice de Shoji e o índice calculado com o uso dos coeficientes " $d_0$ " e " $c_0$ " e entre o índice Br e o índice que usa a segunda derivada do coeficiente " $d_0$ ", para o primeiro procedimento de aquisição.

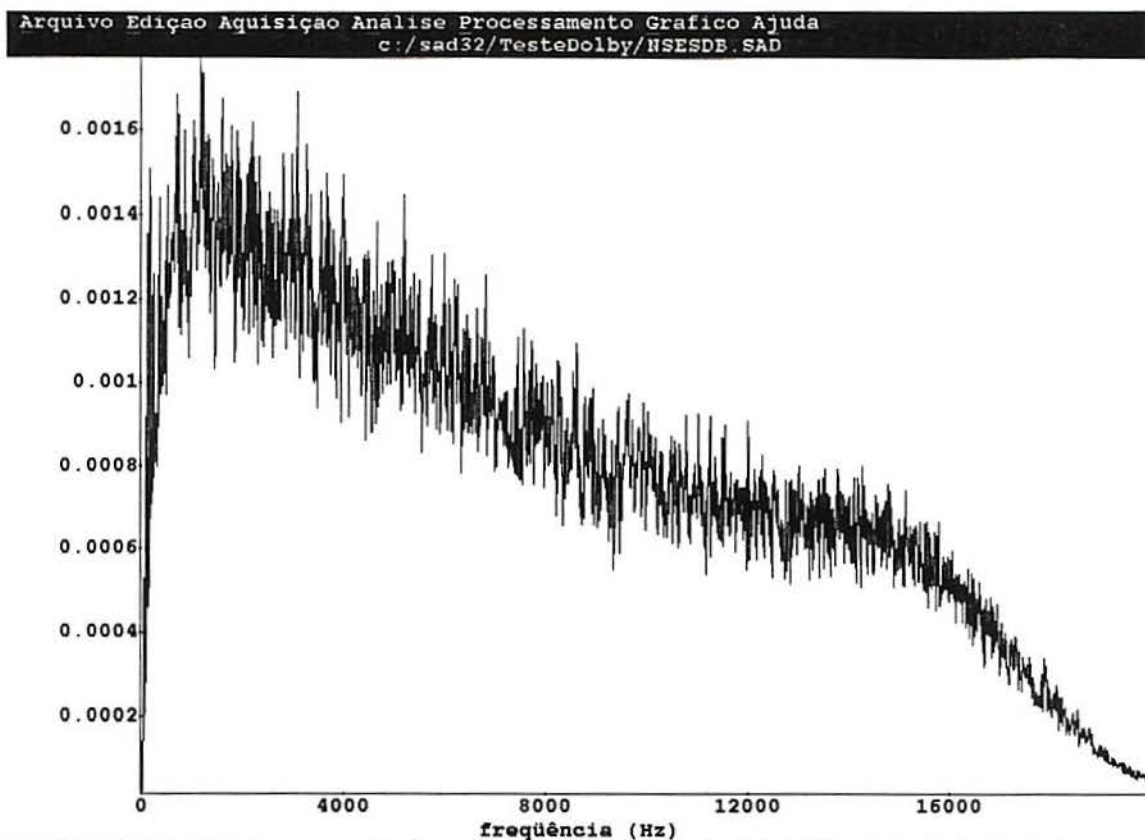


Figura 5.1: Gráfico da estimativa do espectro de potência do ruído branco gravado e reproduzido sem Dolby, em escala linear.

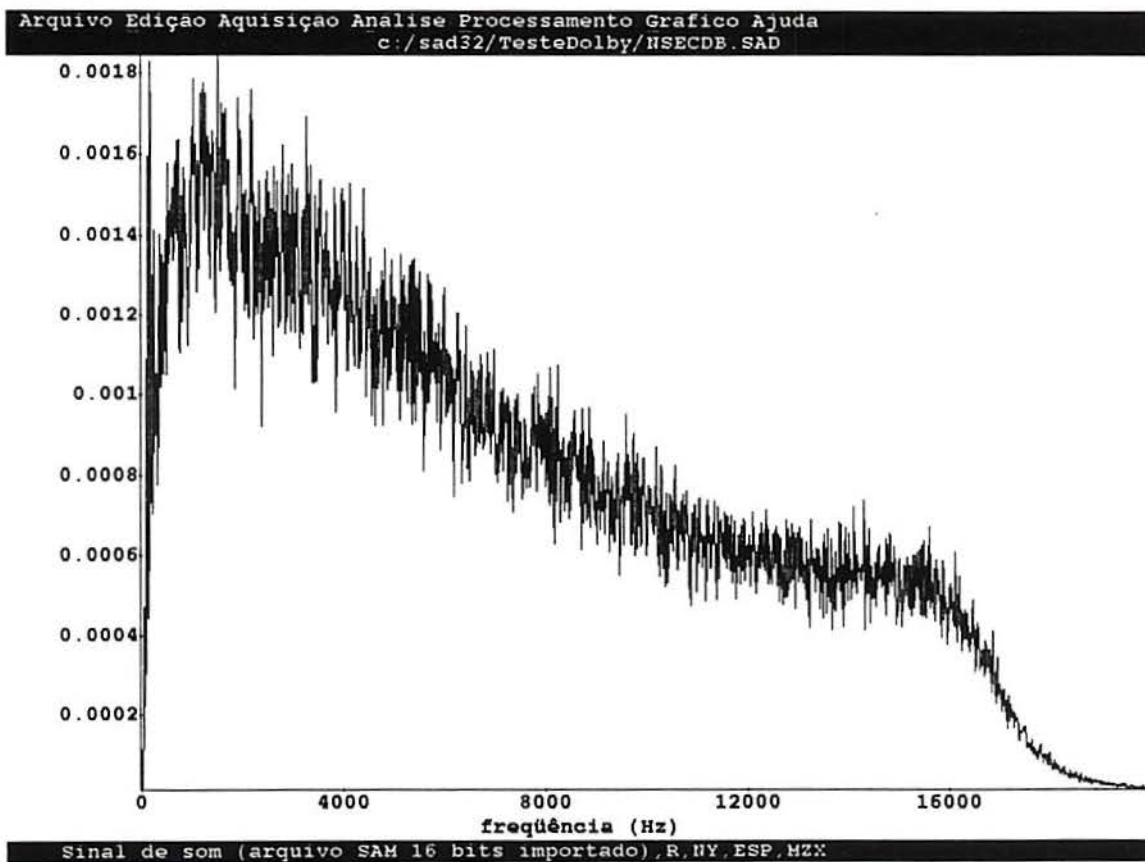


Figura 5.2: Gráfico da estimativa do espectro de potência do ruído branco gravado e reproduzido com Dolby, em escala linear.

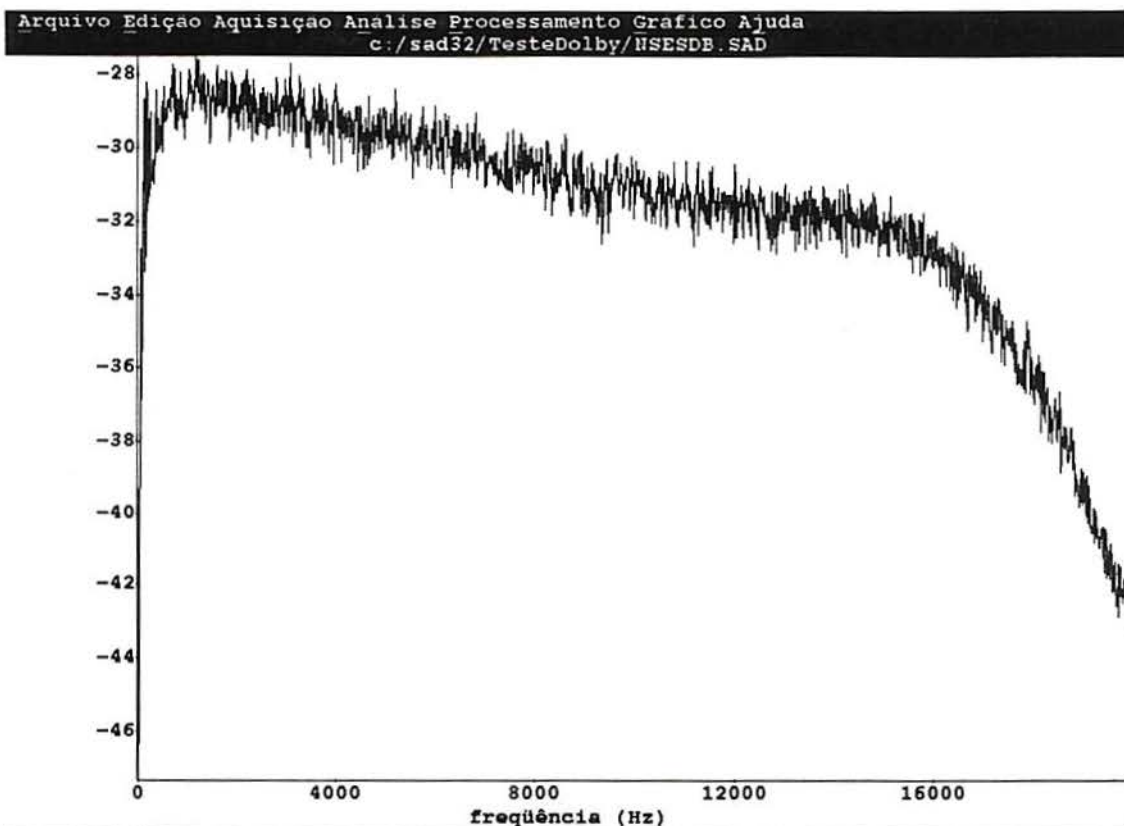


Figura 5.3: Gráfico da estimativa do espectro de potência do ruído branco gravado e reproduzido sem Dolby, em escala logarítmica.

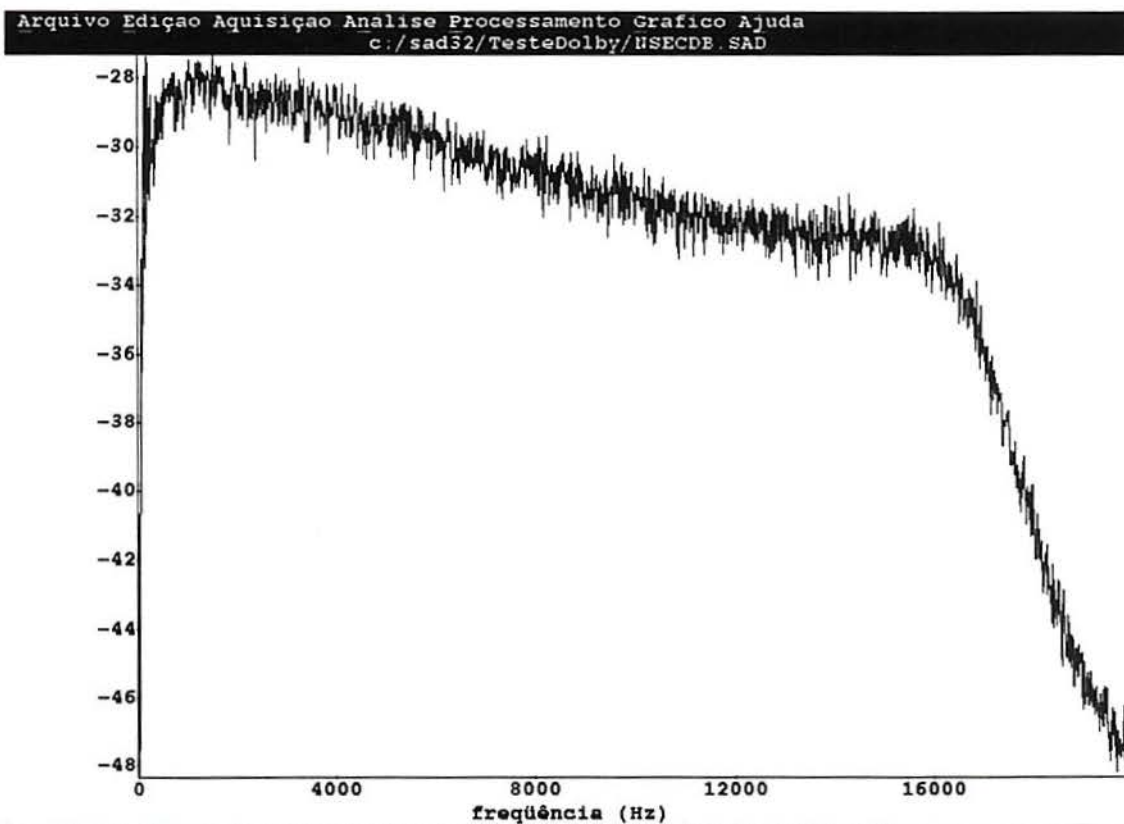


Figura 5.4: Gráfico da estimativa do espectro de potência do ruído branco gravado e reproduzido com Dolby, em escala logarítmica.

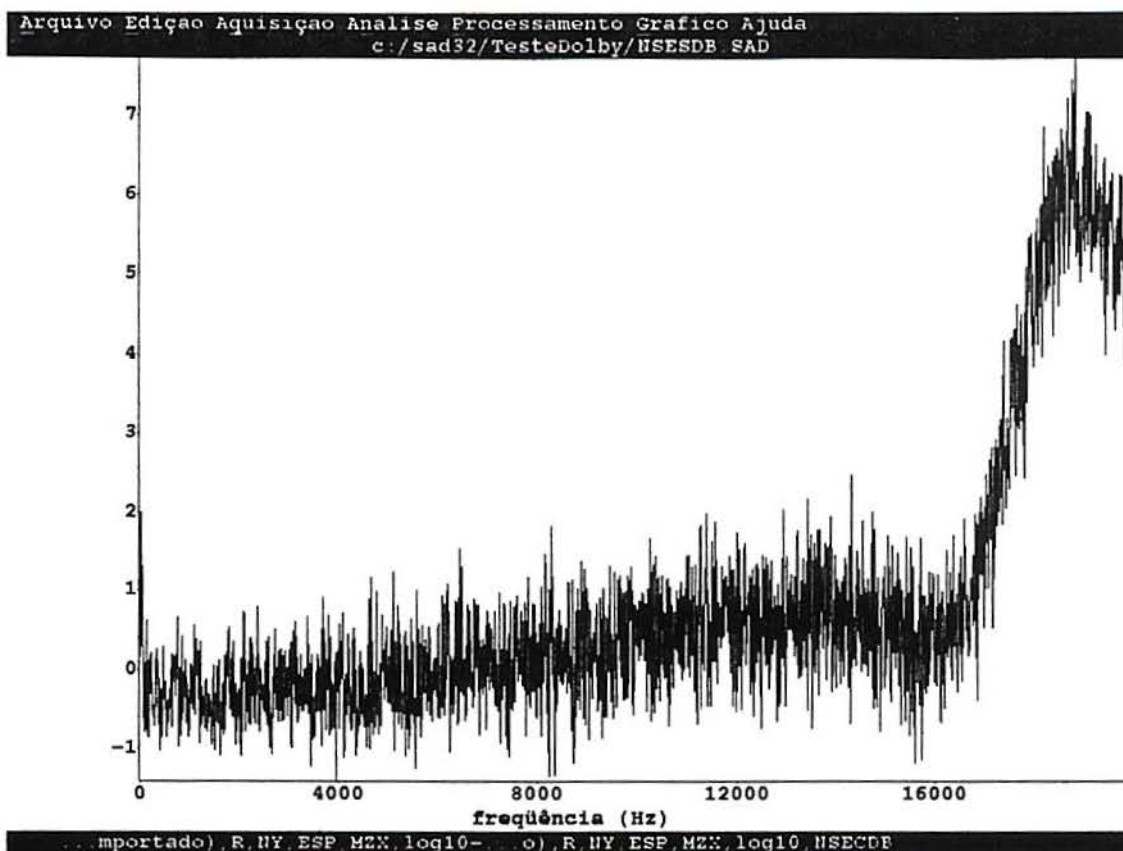


Figura 5.5: Gráfico da diferença entre as estimções do espectro de potência do ruído branco gravado e reproduzido sem Dolby e com Dolby, em escala de decibéis.

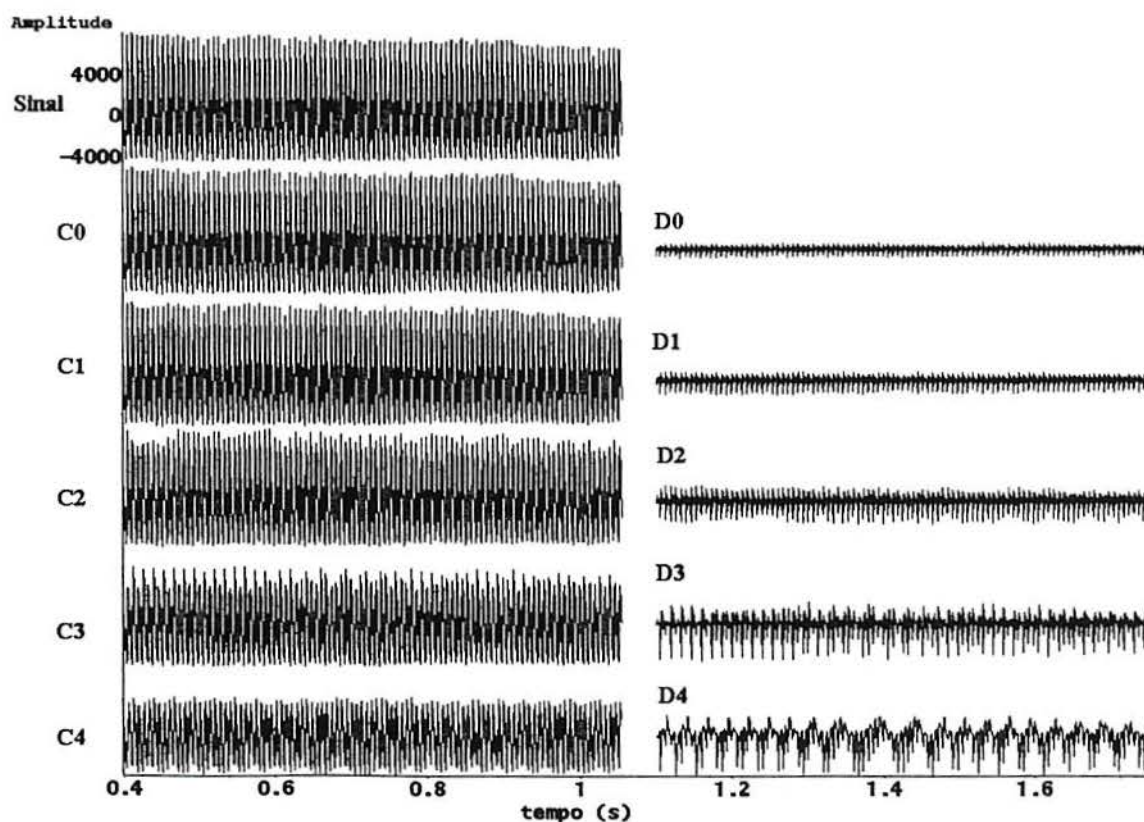


Figura 5.6: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC10.

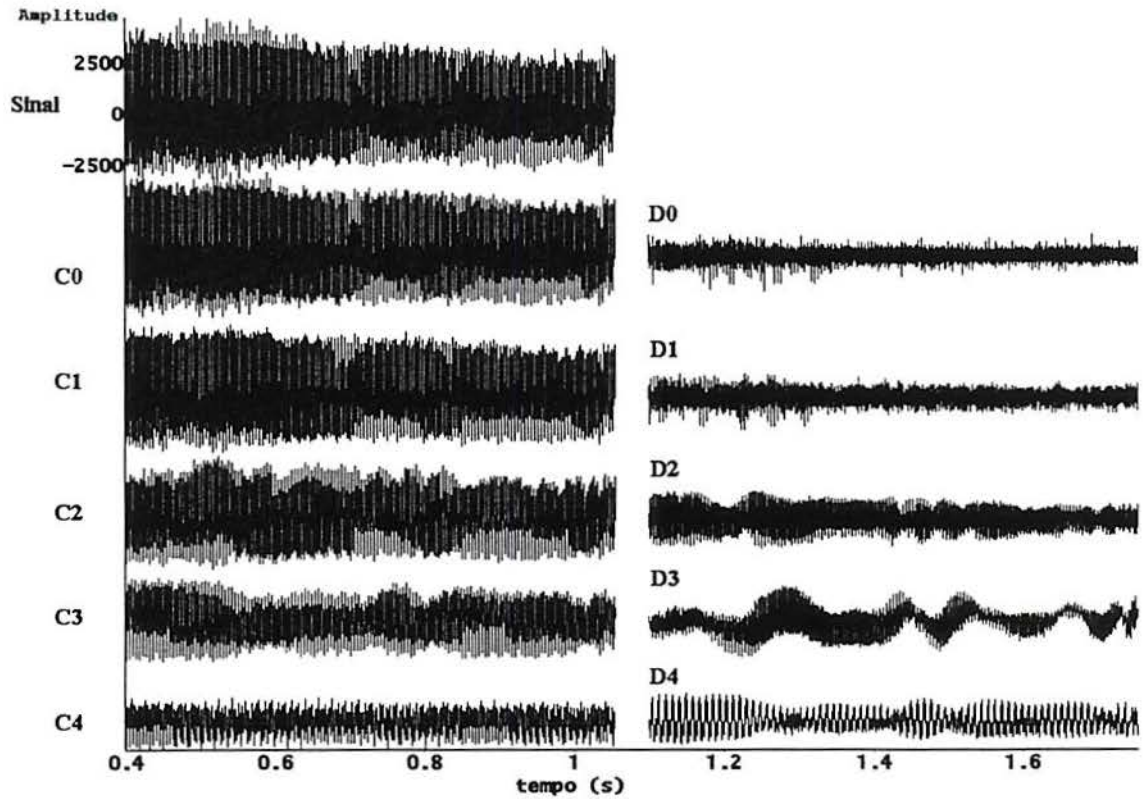


Figura 5.7: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC15.

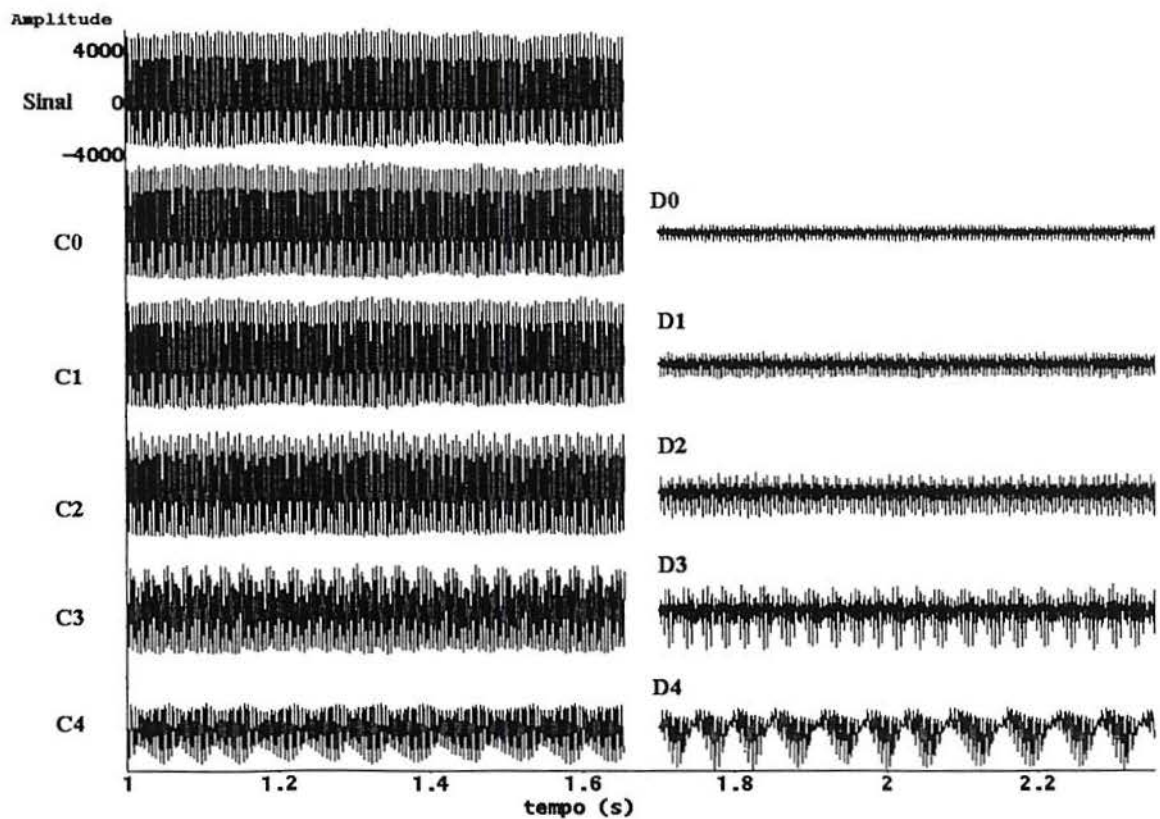


Figura 5.8: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC26.

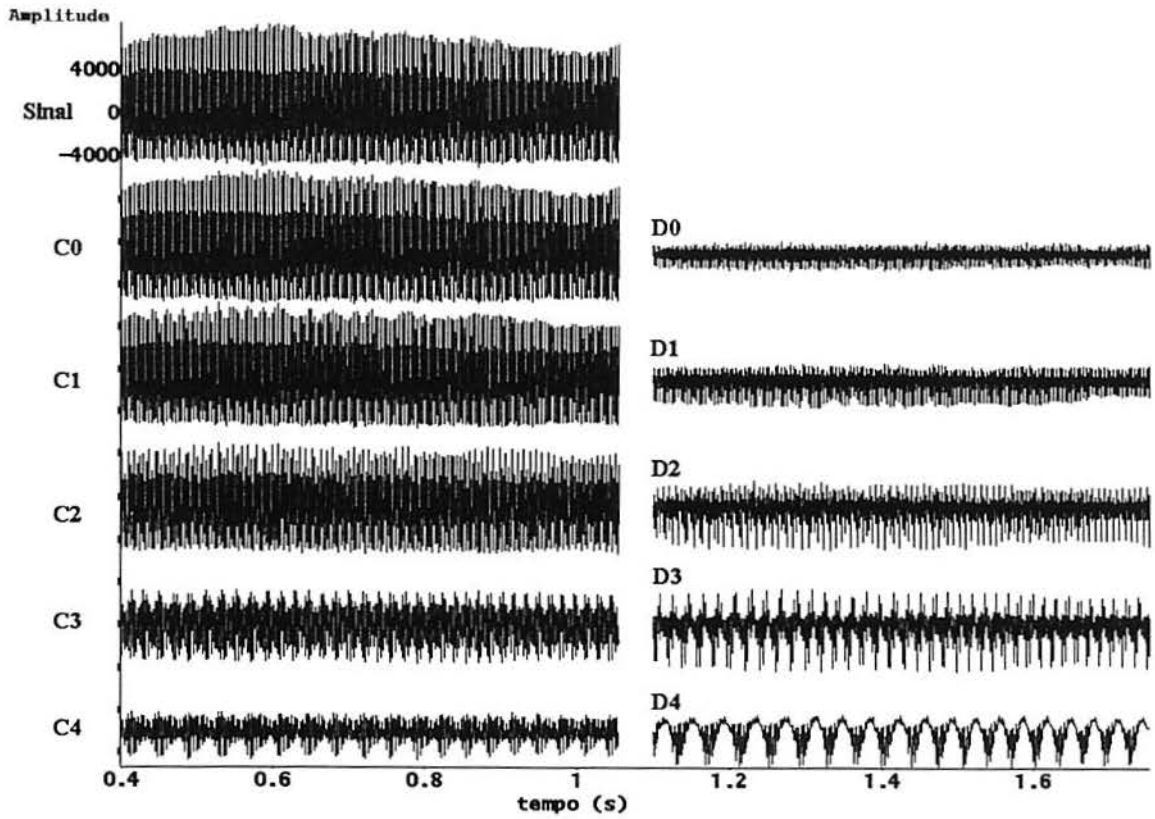


Figura 5.9: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC29.

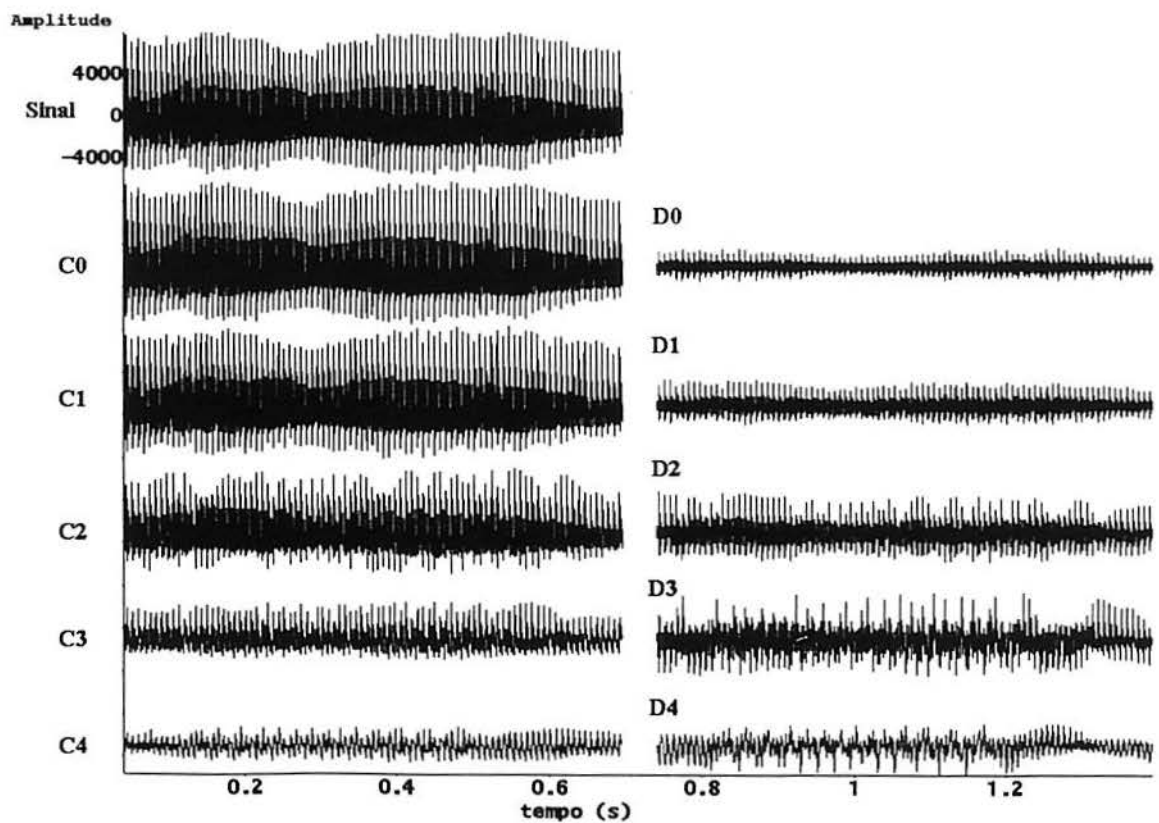


Figura 5.10: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC35.

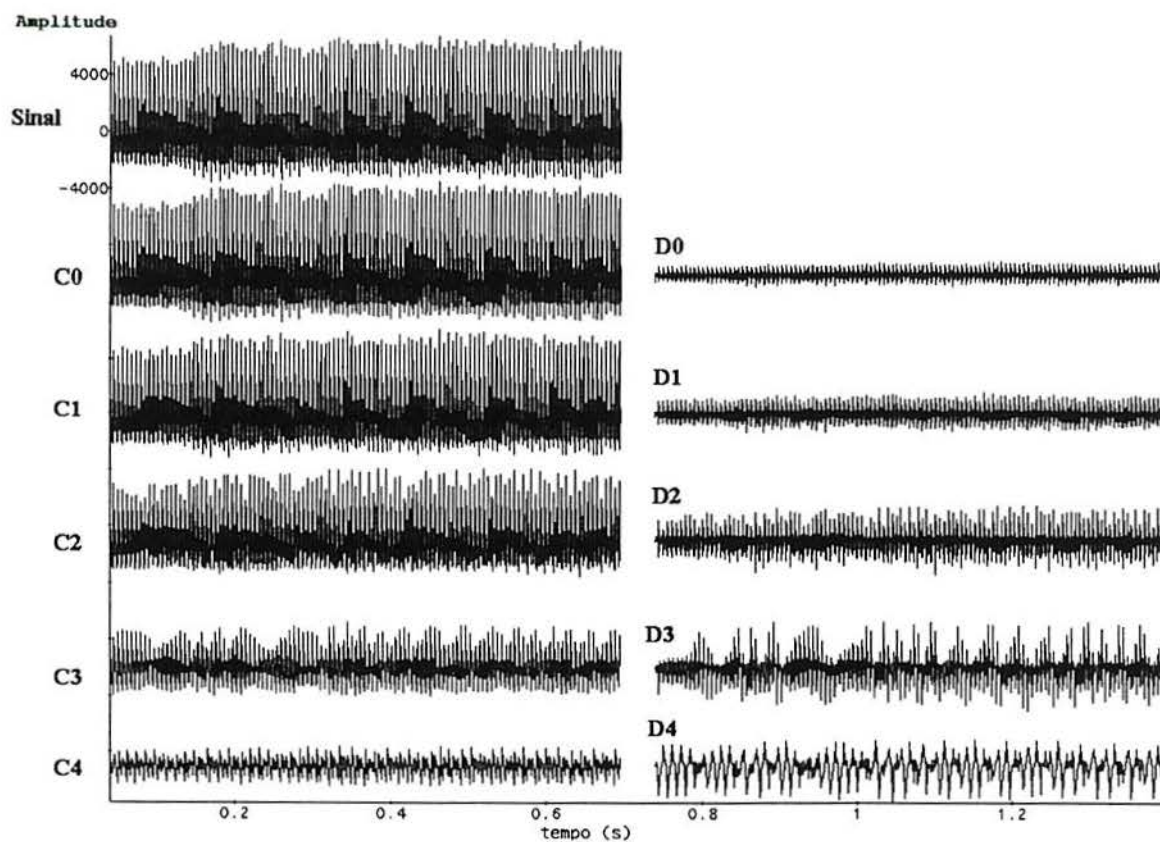


Figura 5.11: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC45.

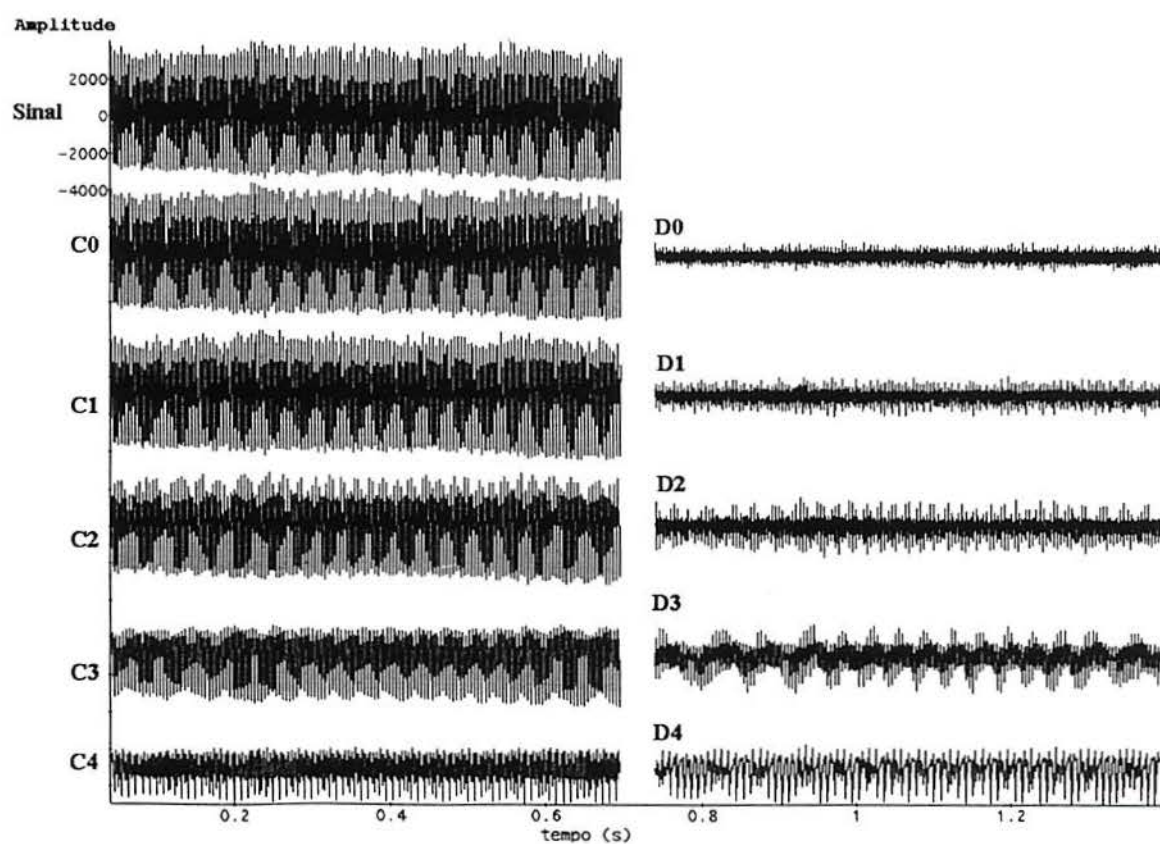


Figura 5.12: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC48.



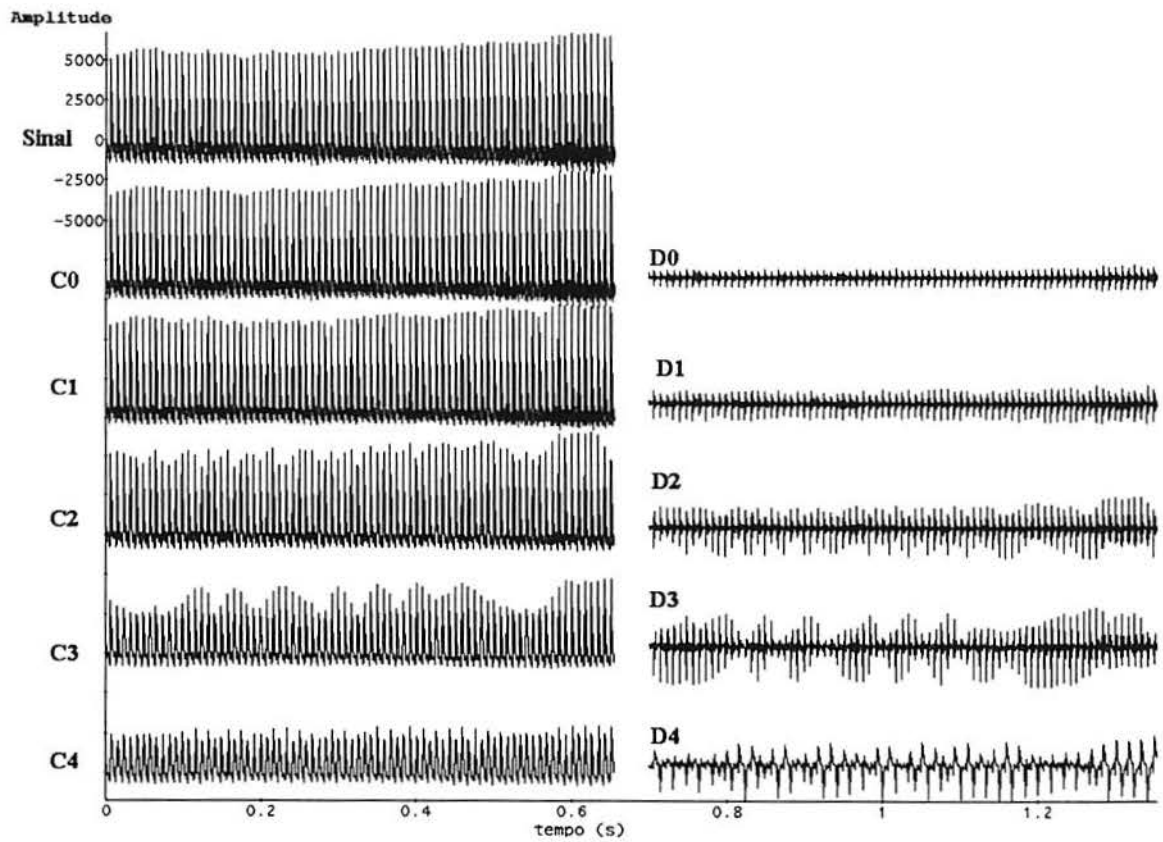


Figura 5.13: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC49.

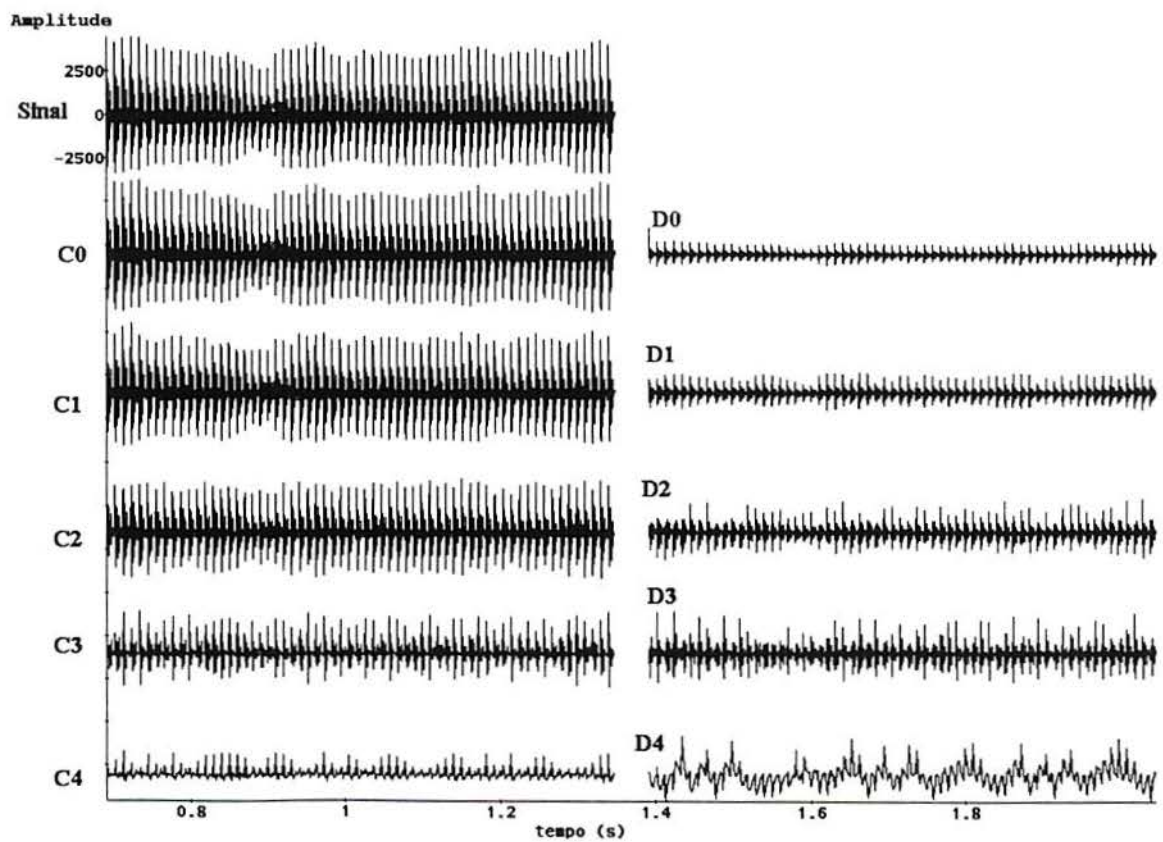


Figura 5.14: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC55.

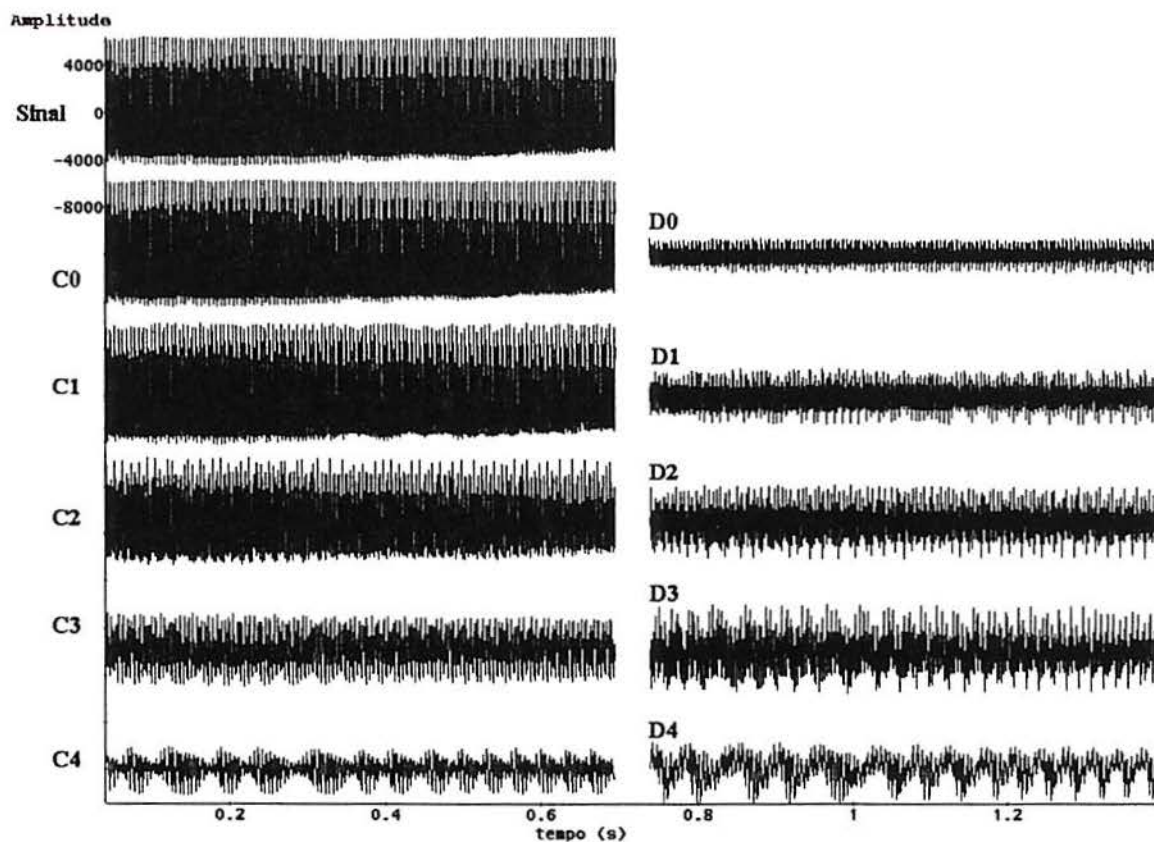


Figura 5.15: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC56.

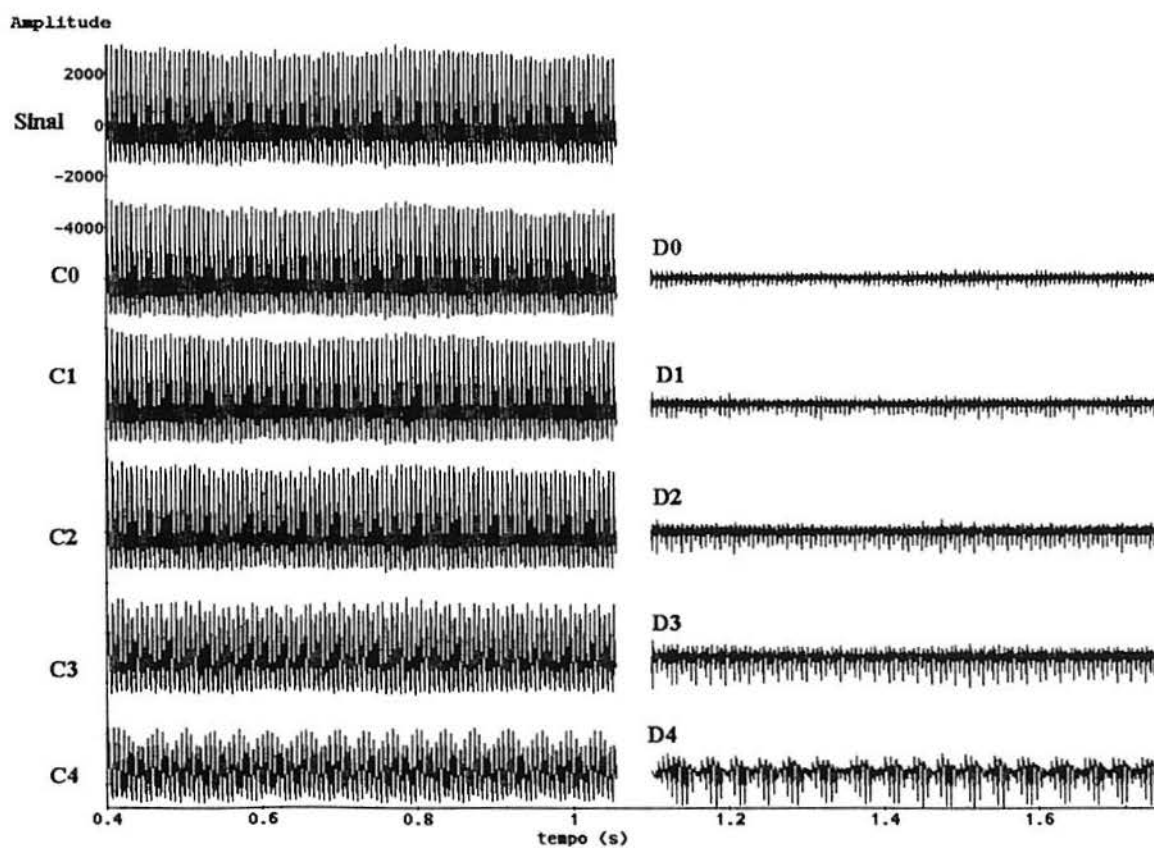


Figura 5.16: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC59.

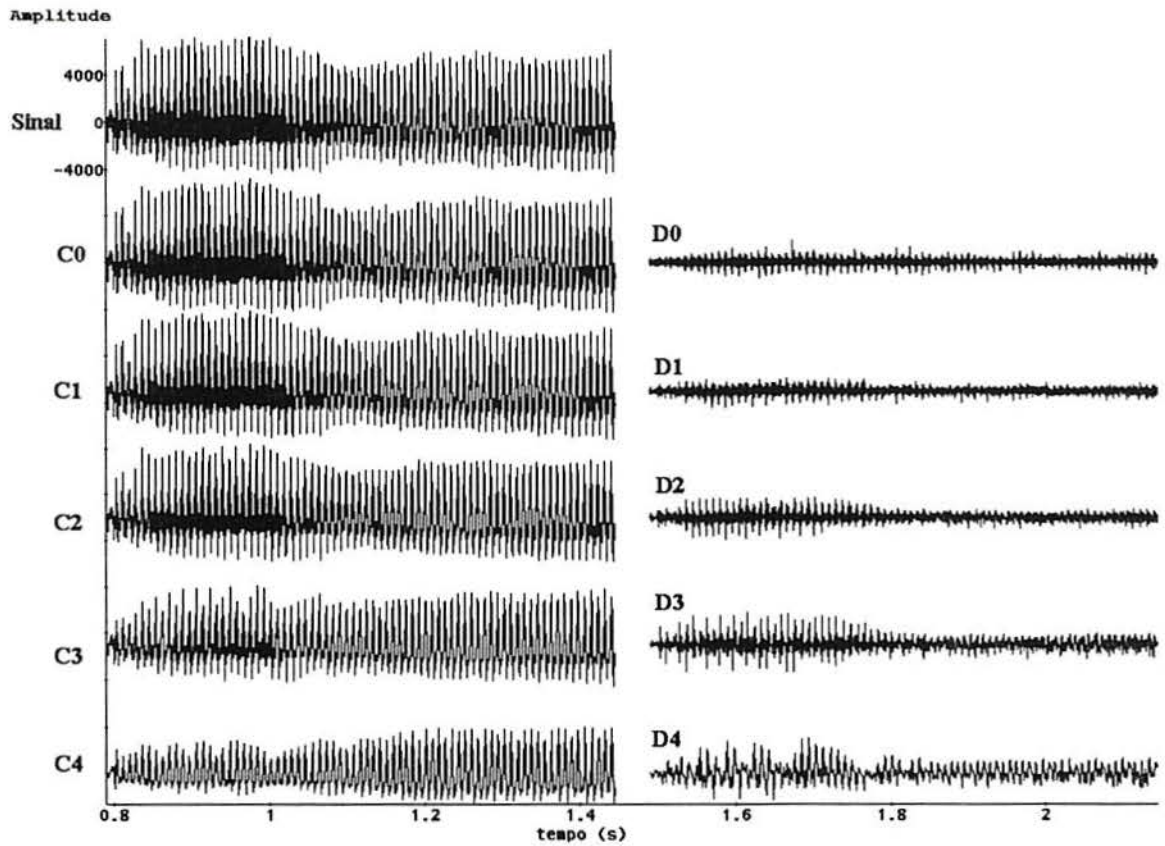


Figura 5.17: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC62.

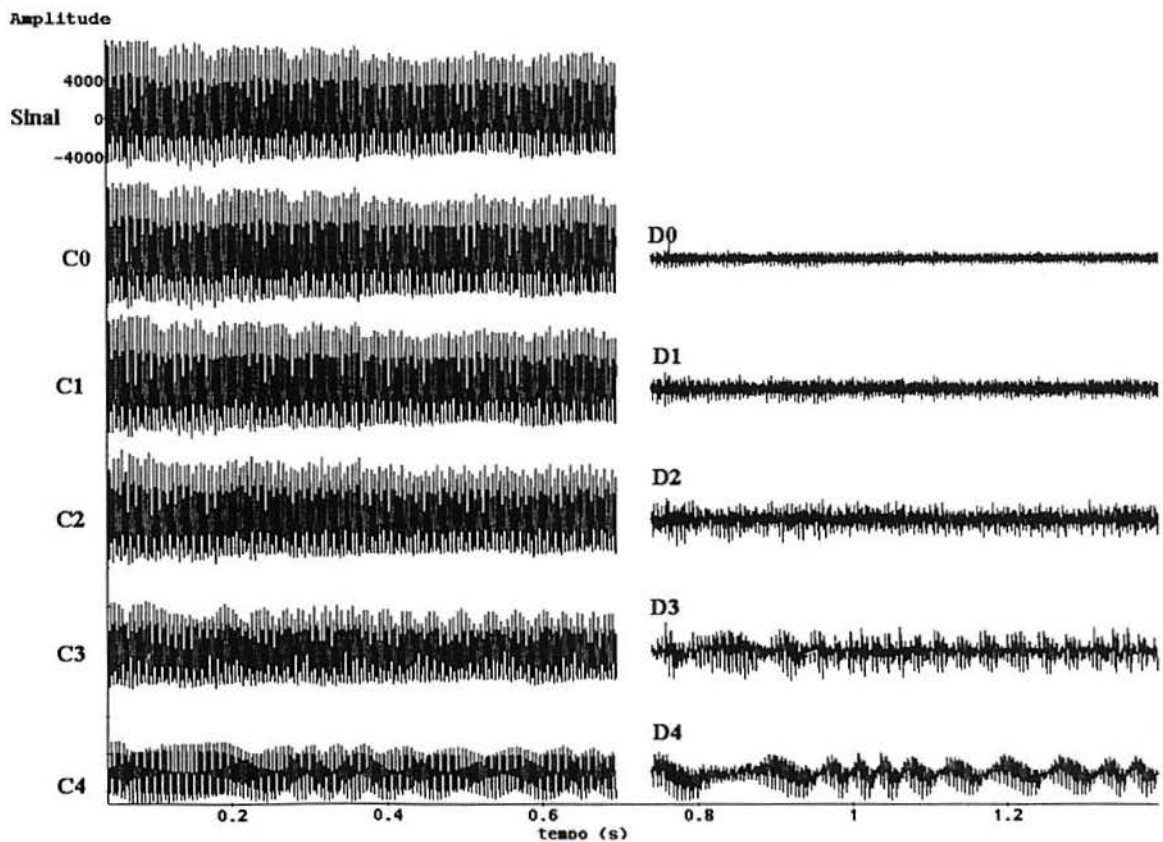


Figura 5.18: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC63.

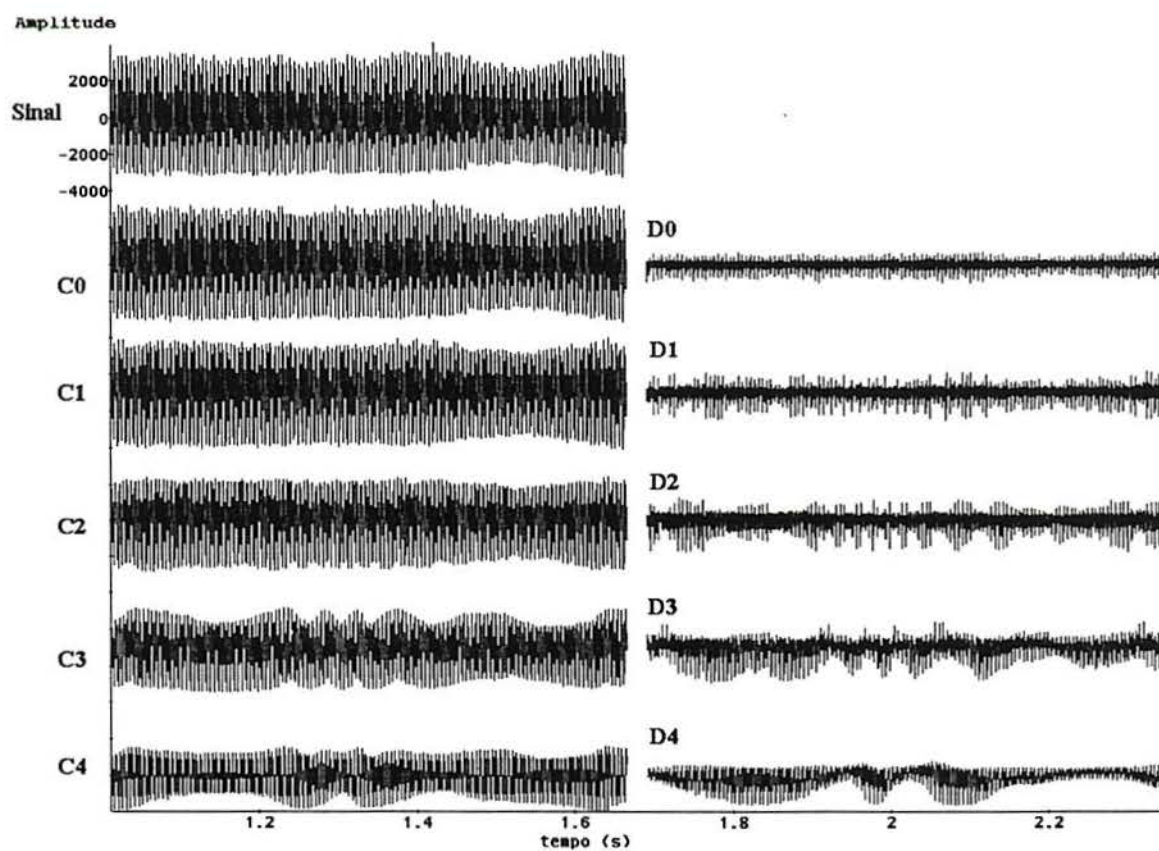


Figura 5.19: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC67.

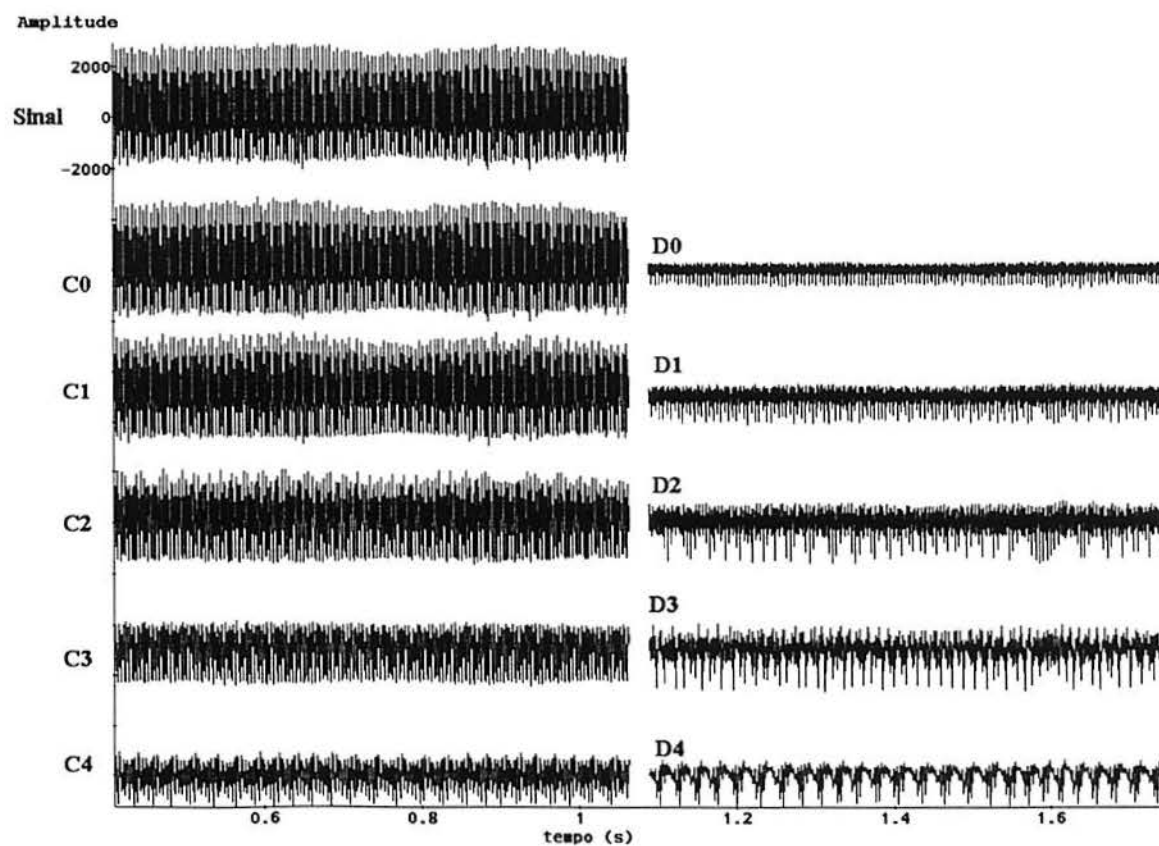


Figura 5.20: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC68.

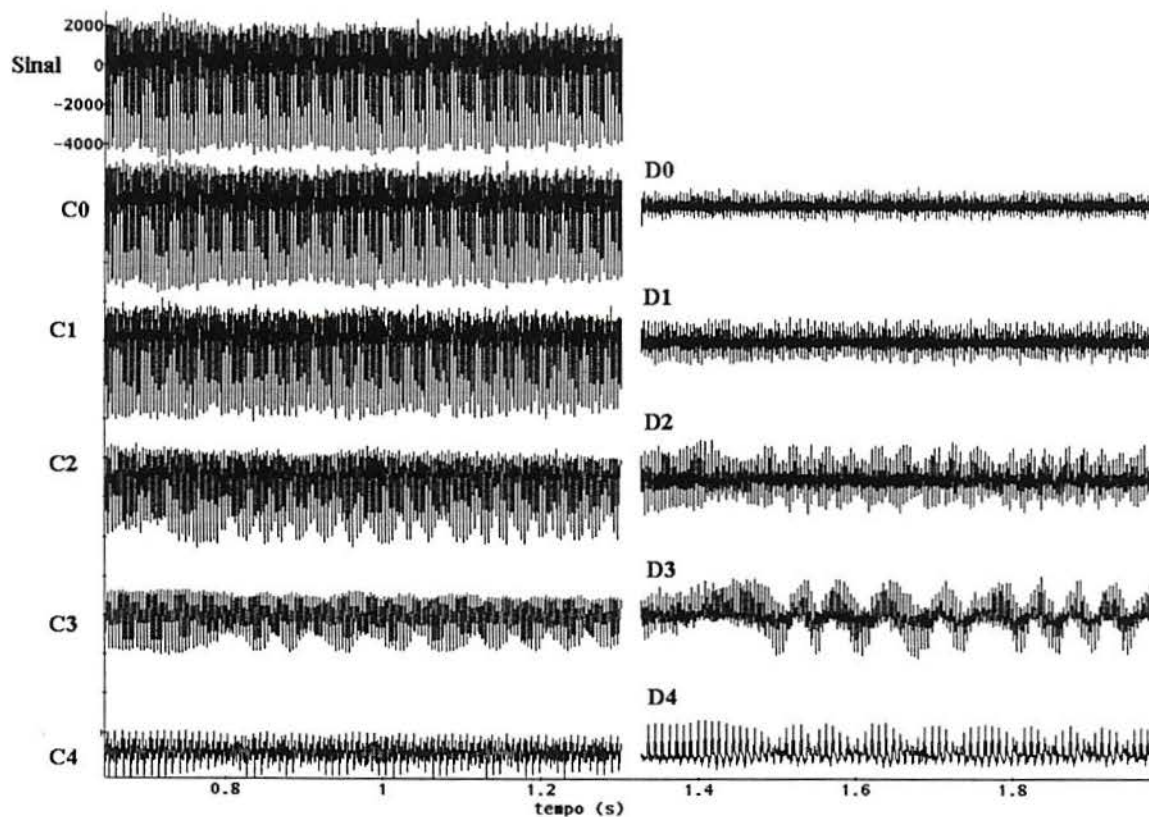


Figura 5.21: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC74.

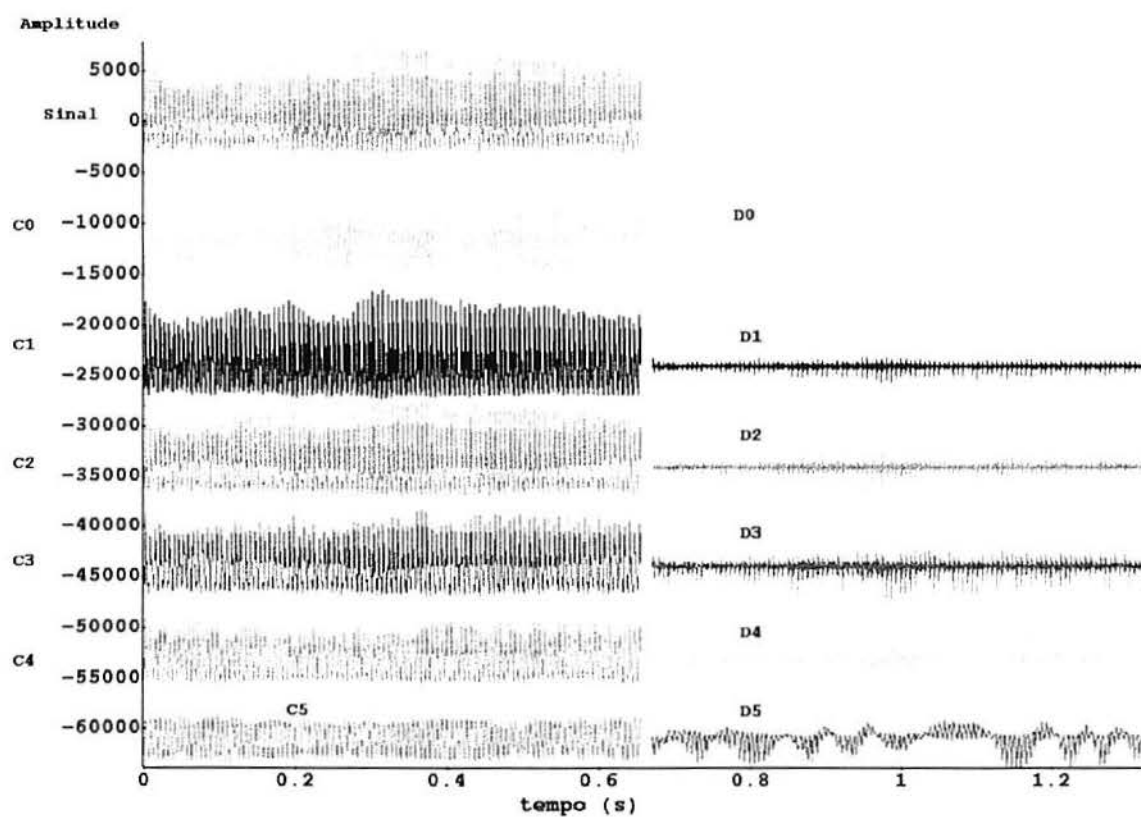


Figura 5.22: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC10, com segmentação retangular.

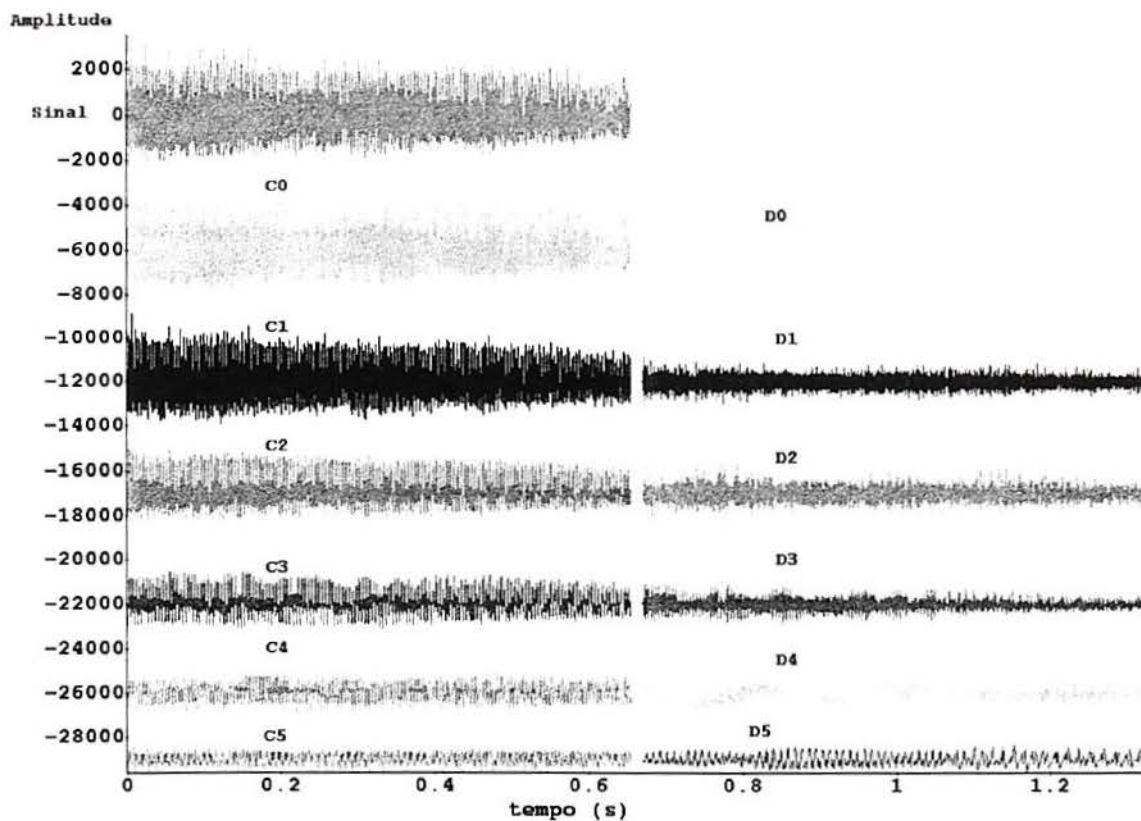


Figura 5.23: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC15, com segmentação retangular.

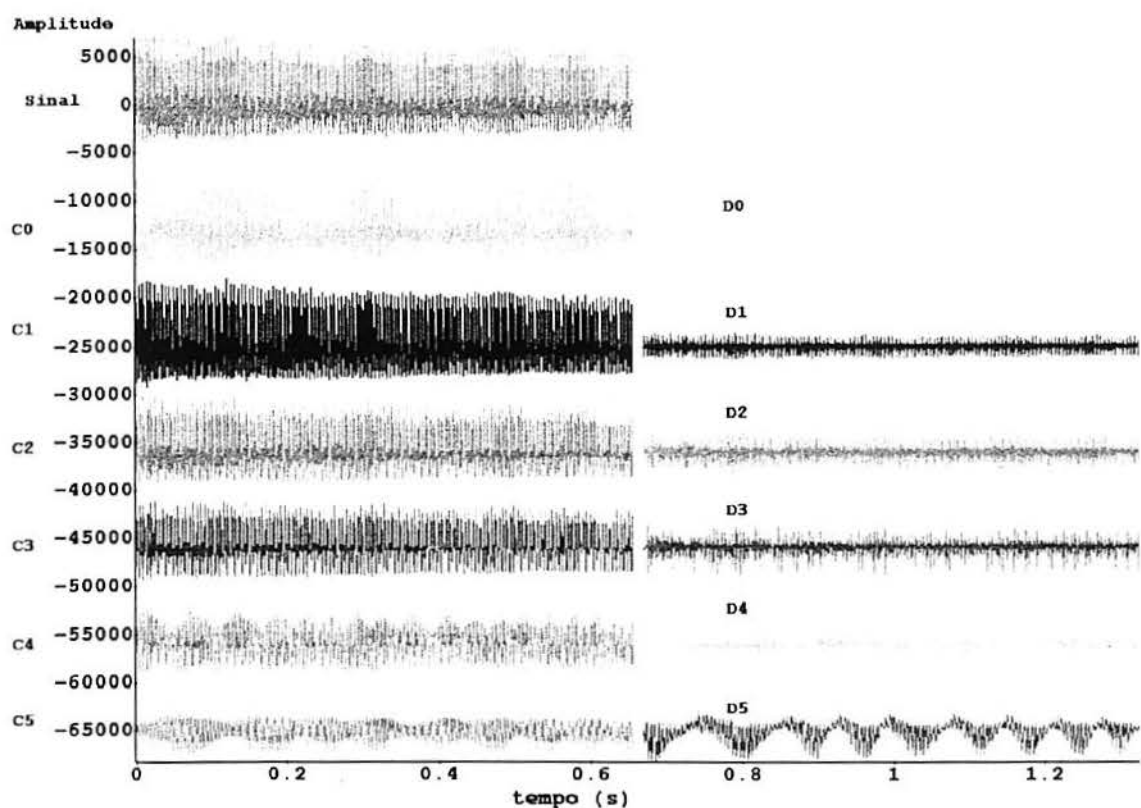


Figura 5.24: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC26.

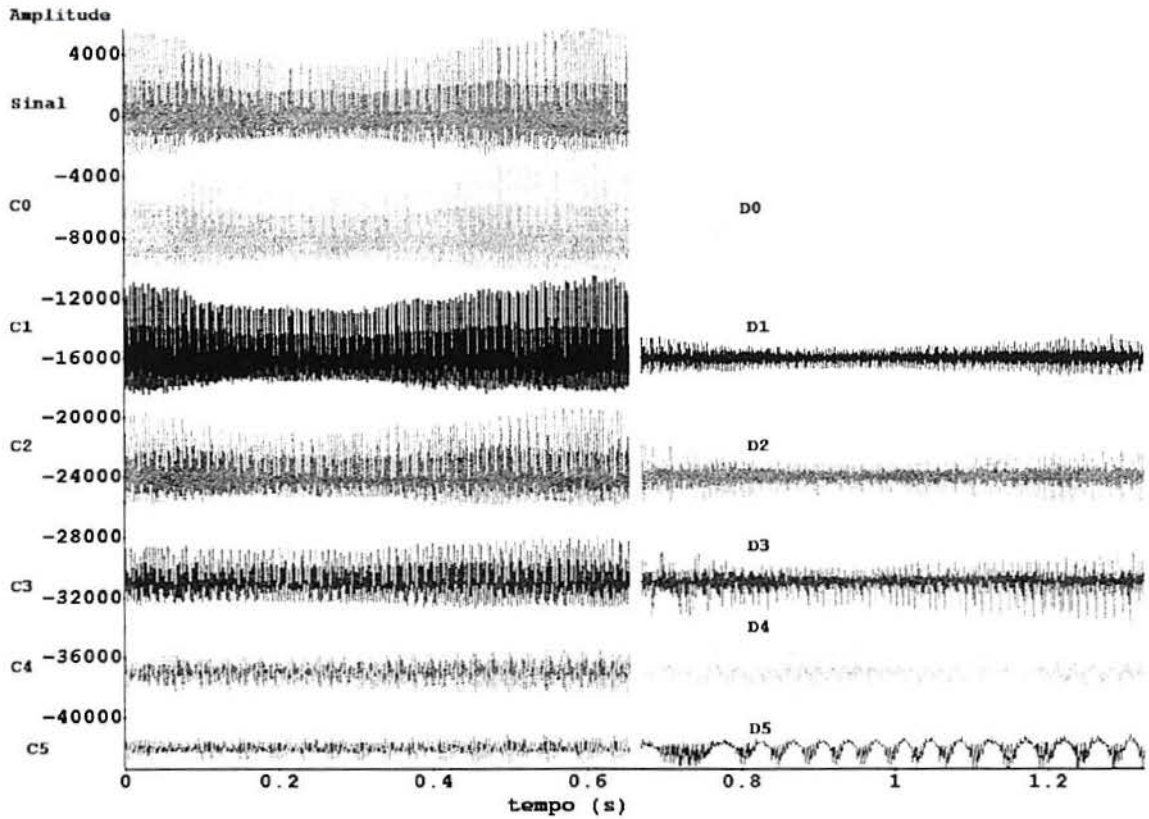


Figura 5.25: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC29.

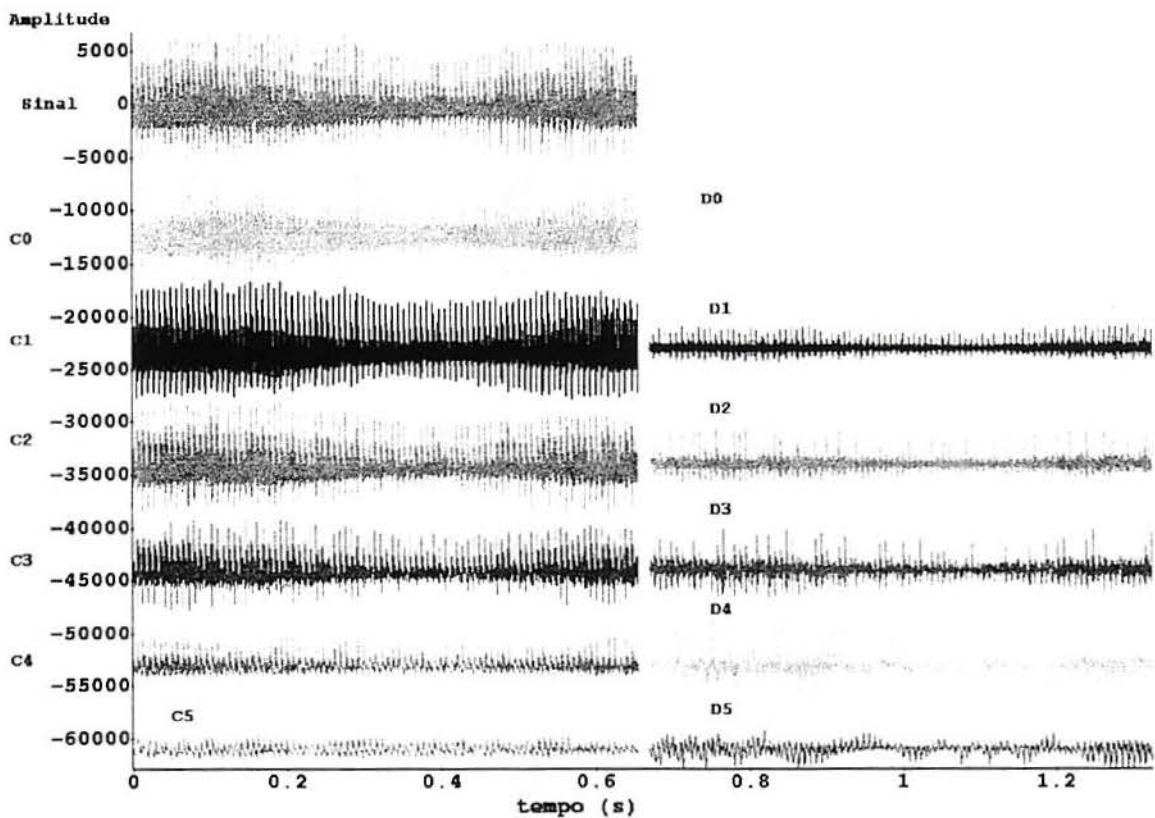


Figura 5.26: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC35.

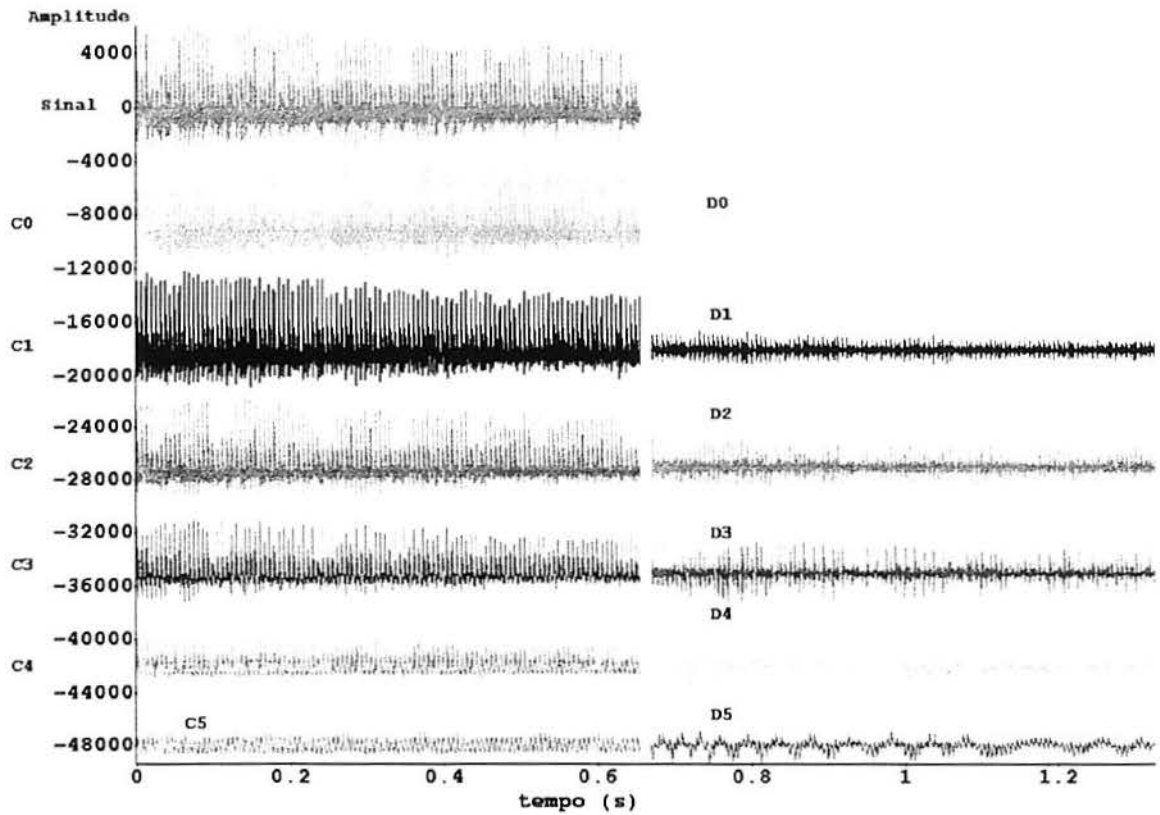


Figura 5.27: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC45.

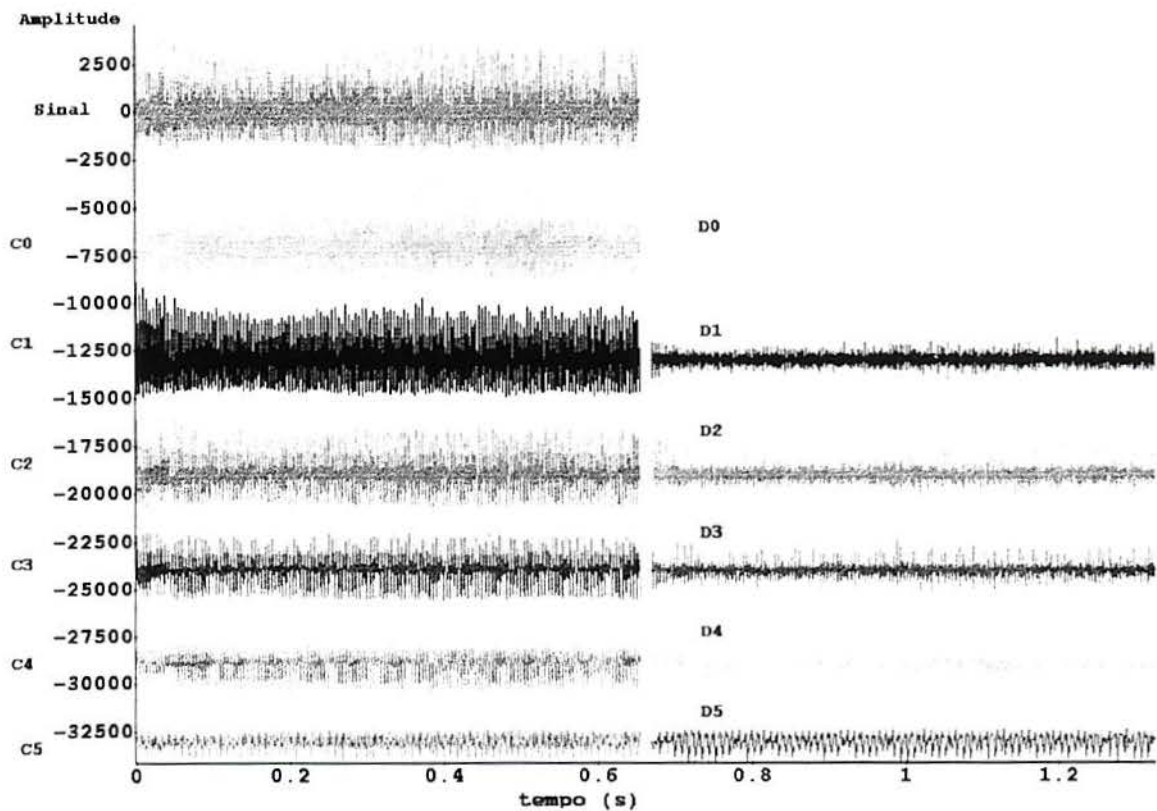


Figura 5.28: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC48.



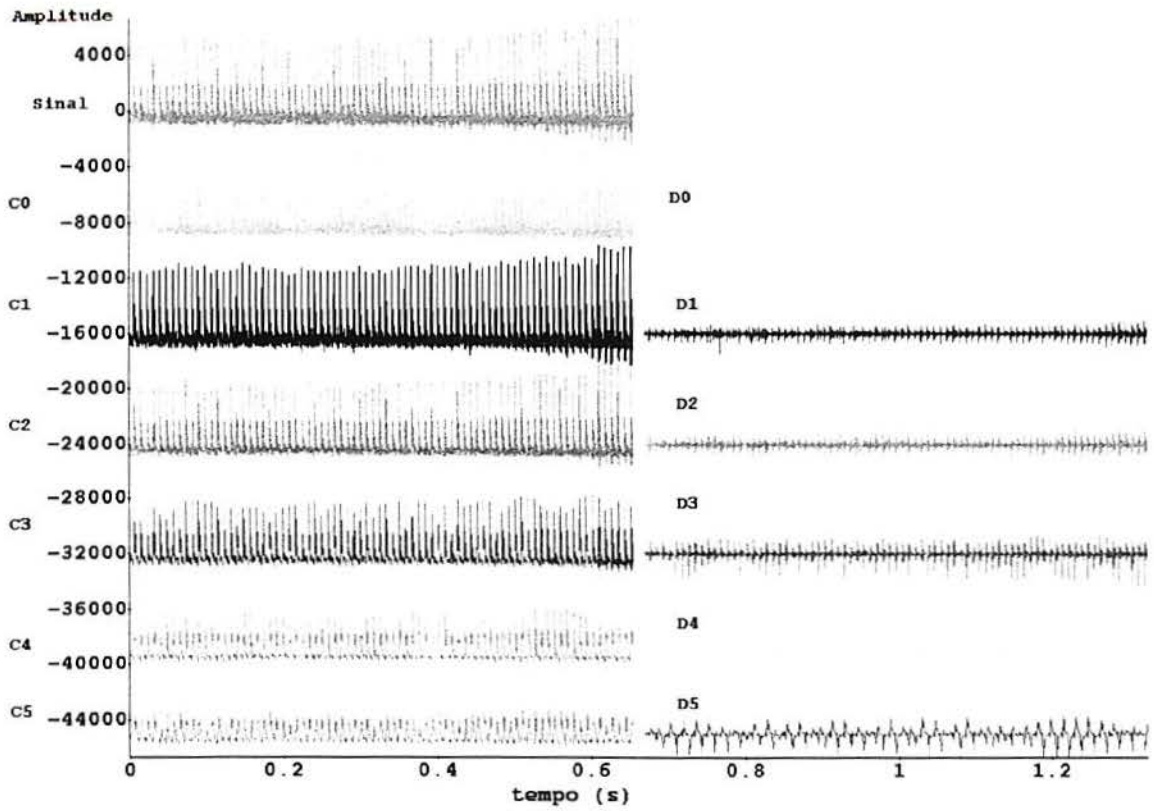


Figura 5.29: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC49.

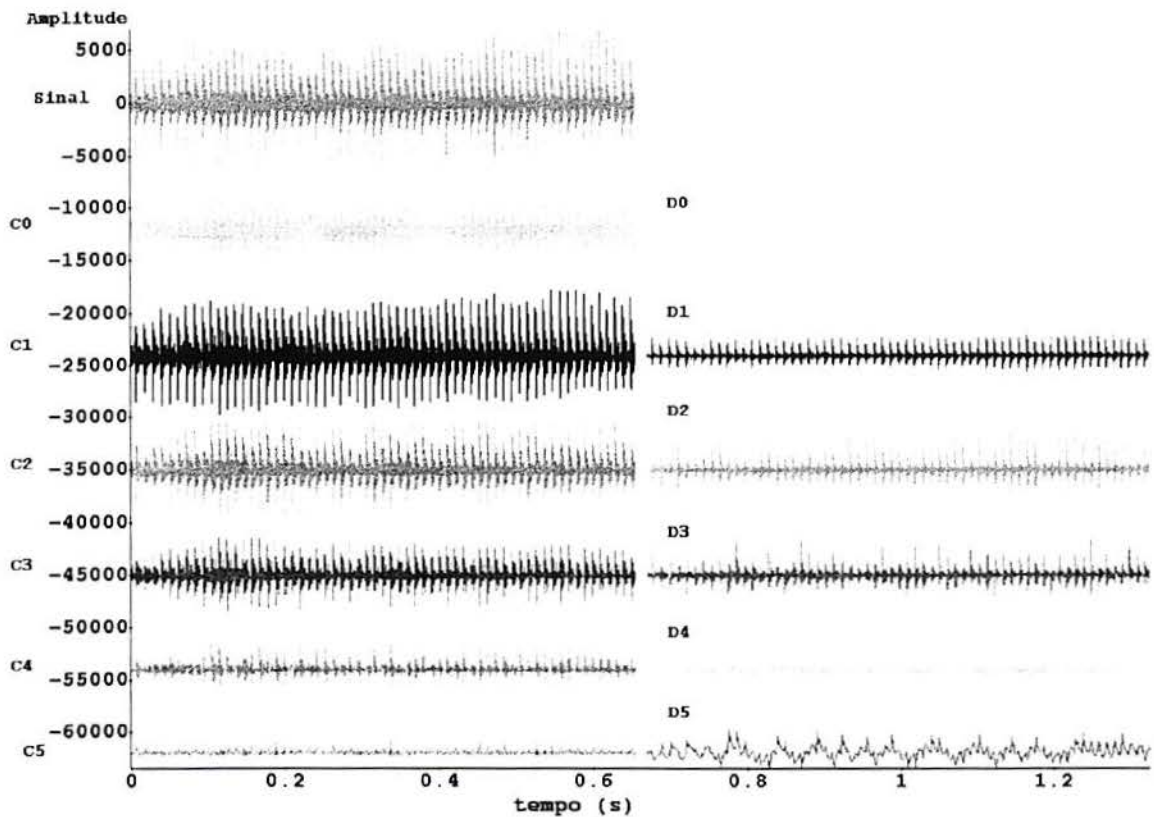


Figura 5.30: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC55.

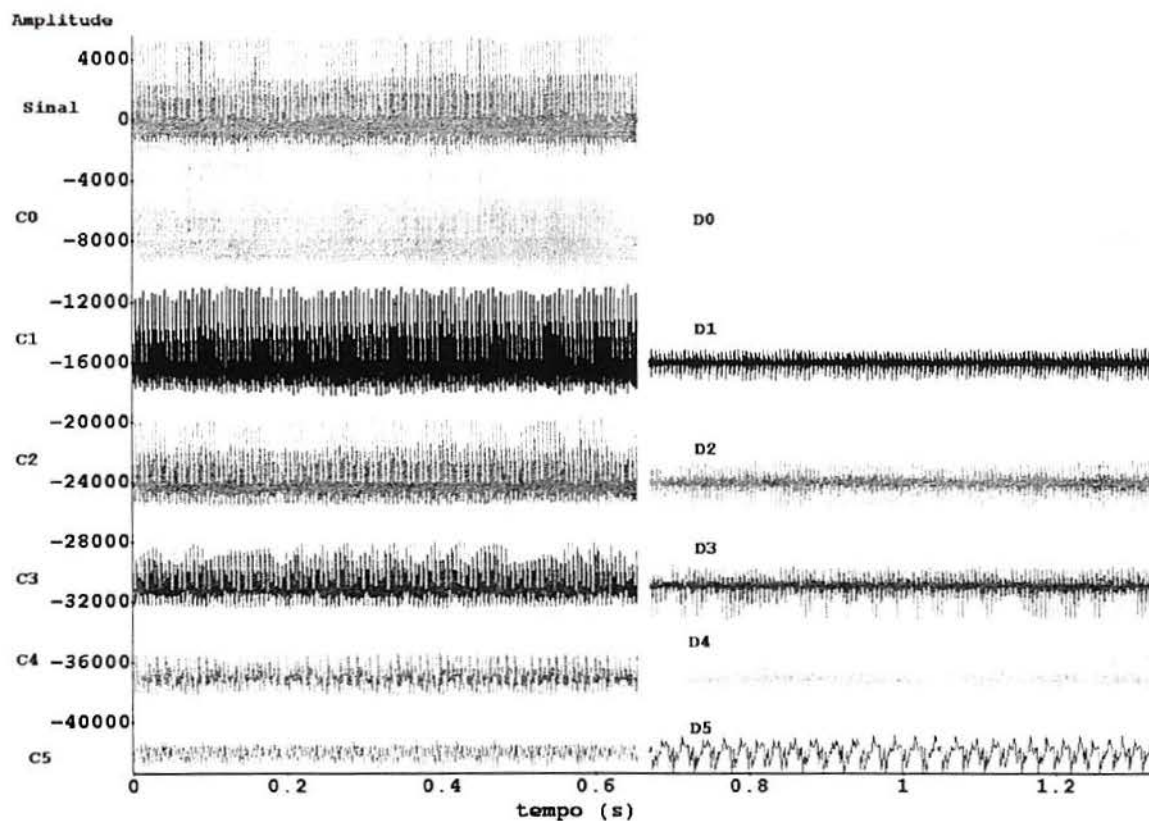


Figura 5.31: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC56.

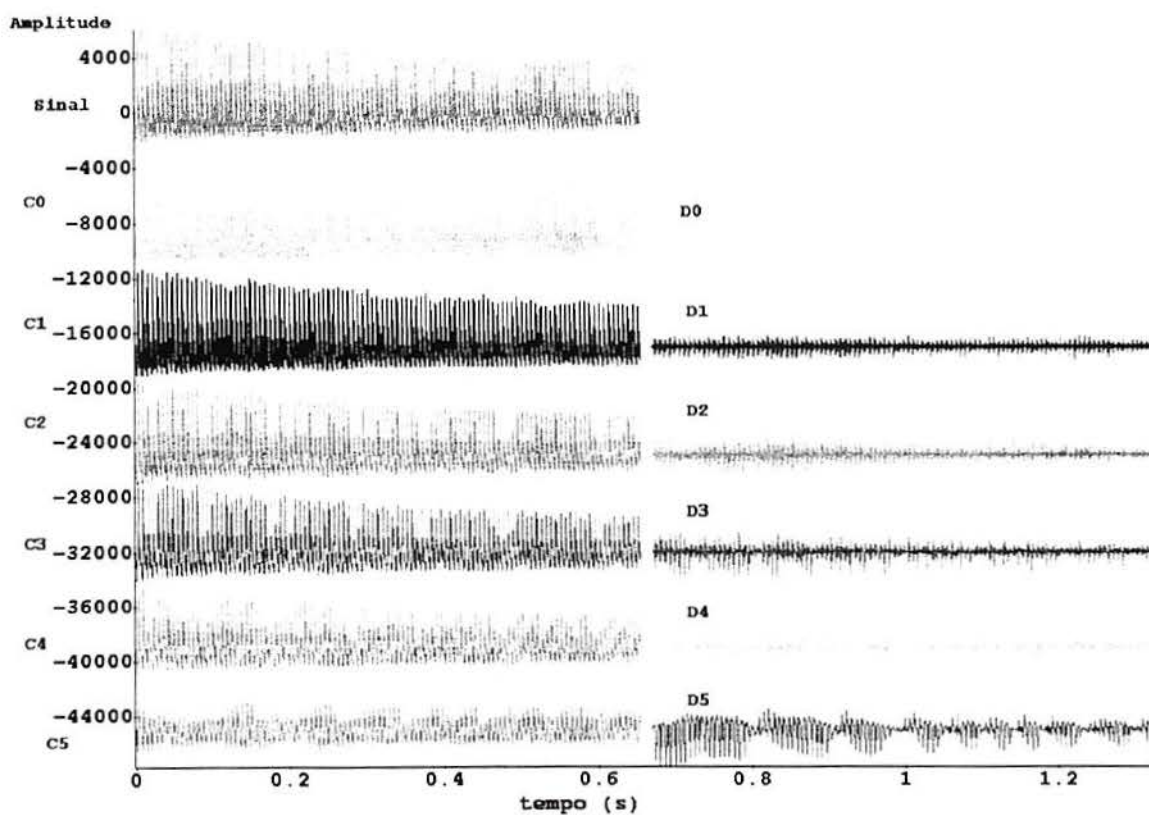


Figura 5.32: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC59.

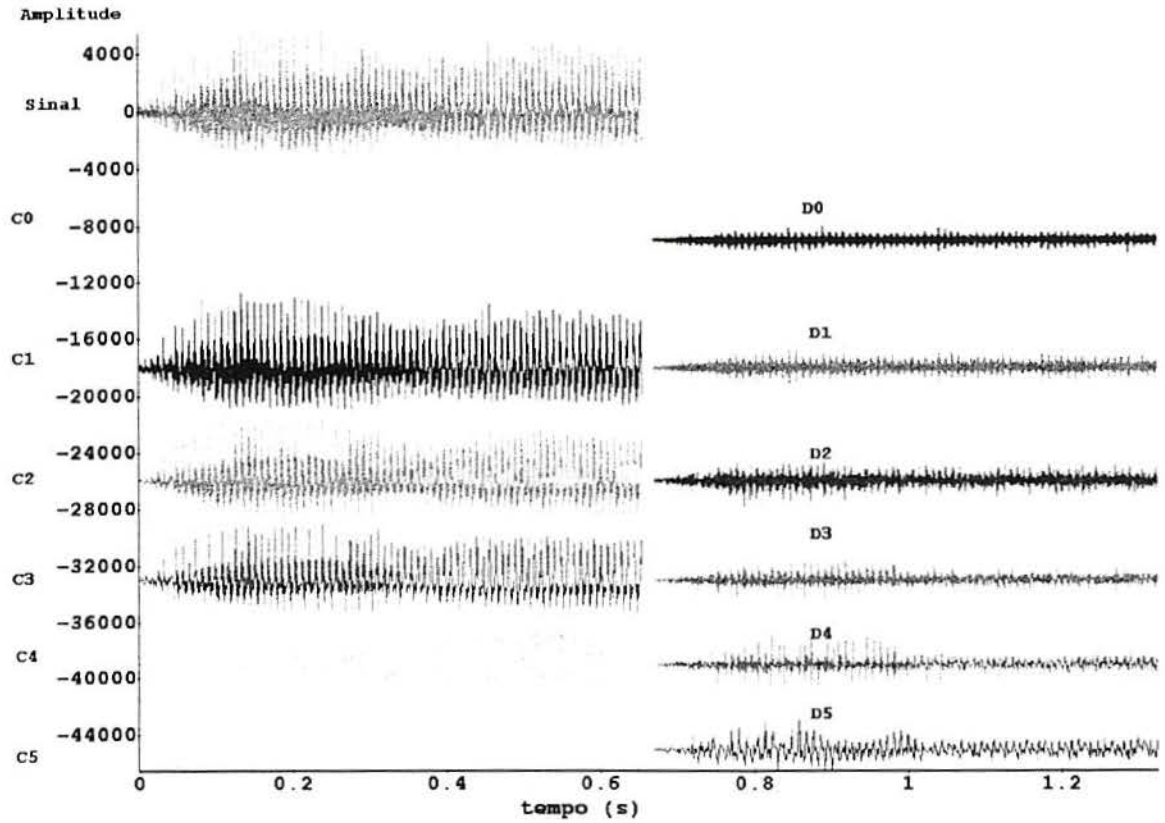


Figura 5.33: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC62.

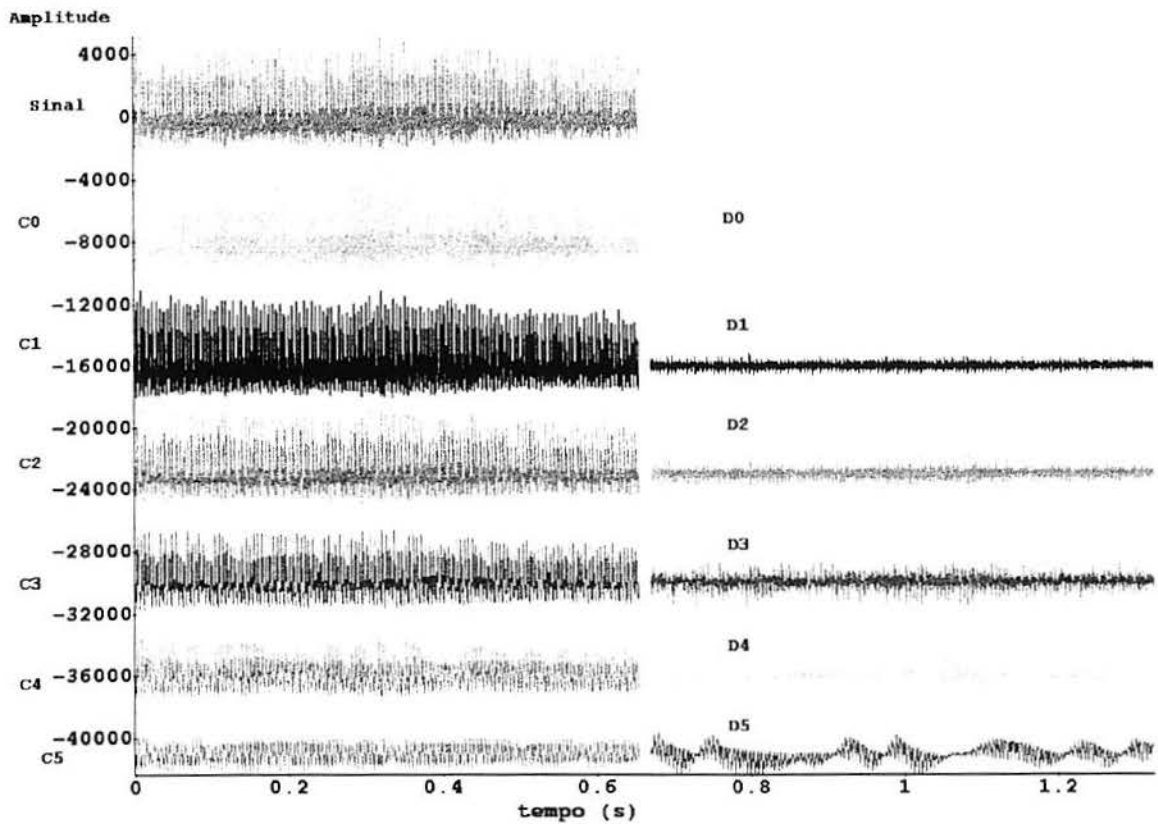


Figura 5.34: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC63.

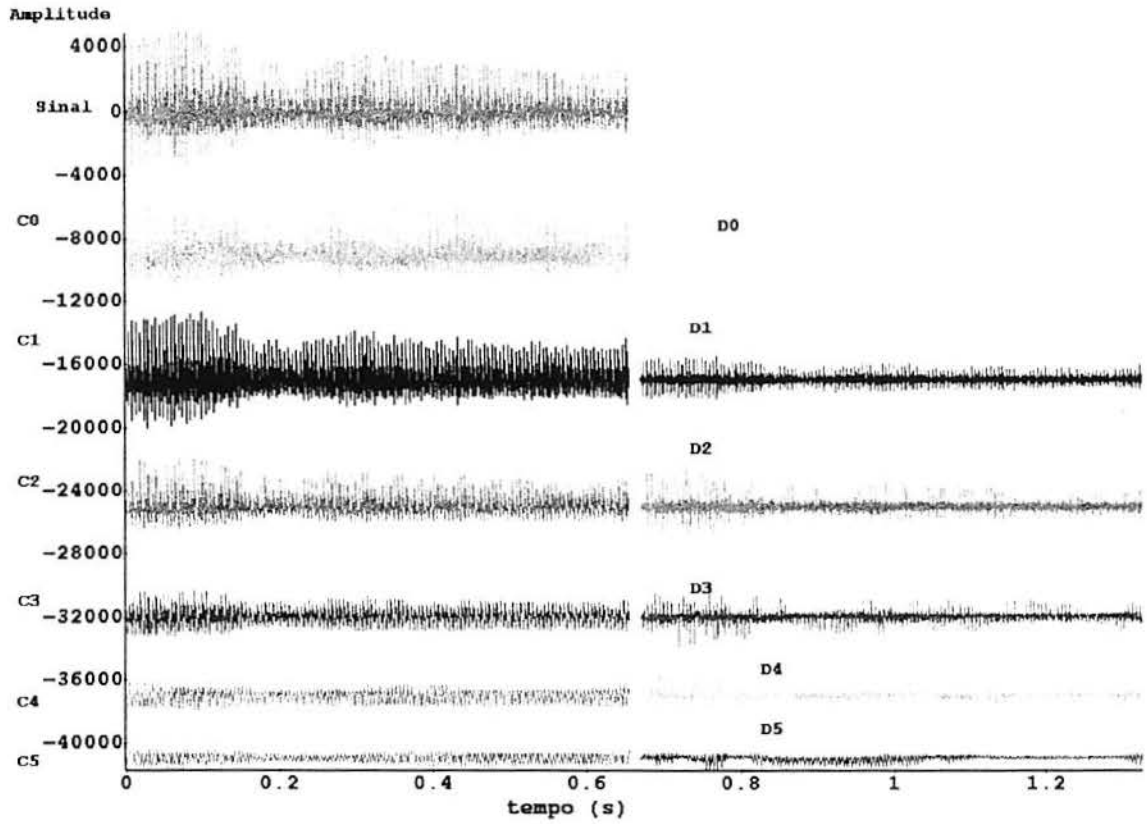


Figura 5.35: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC67.

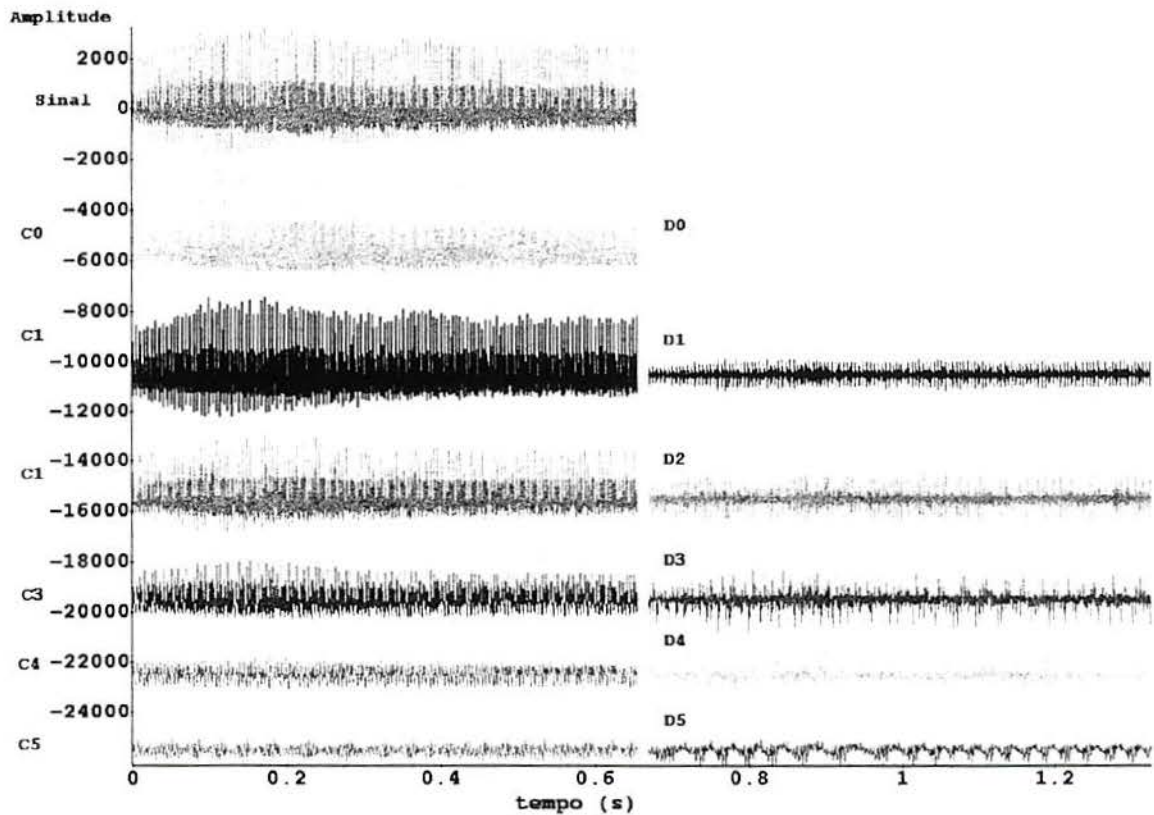


Figura 5.36: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC68.

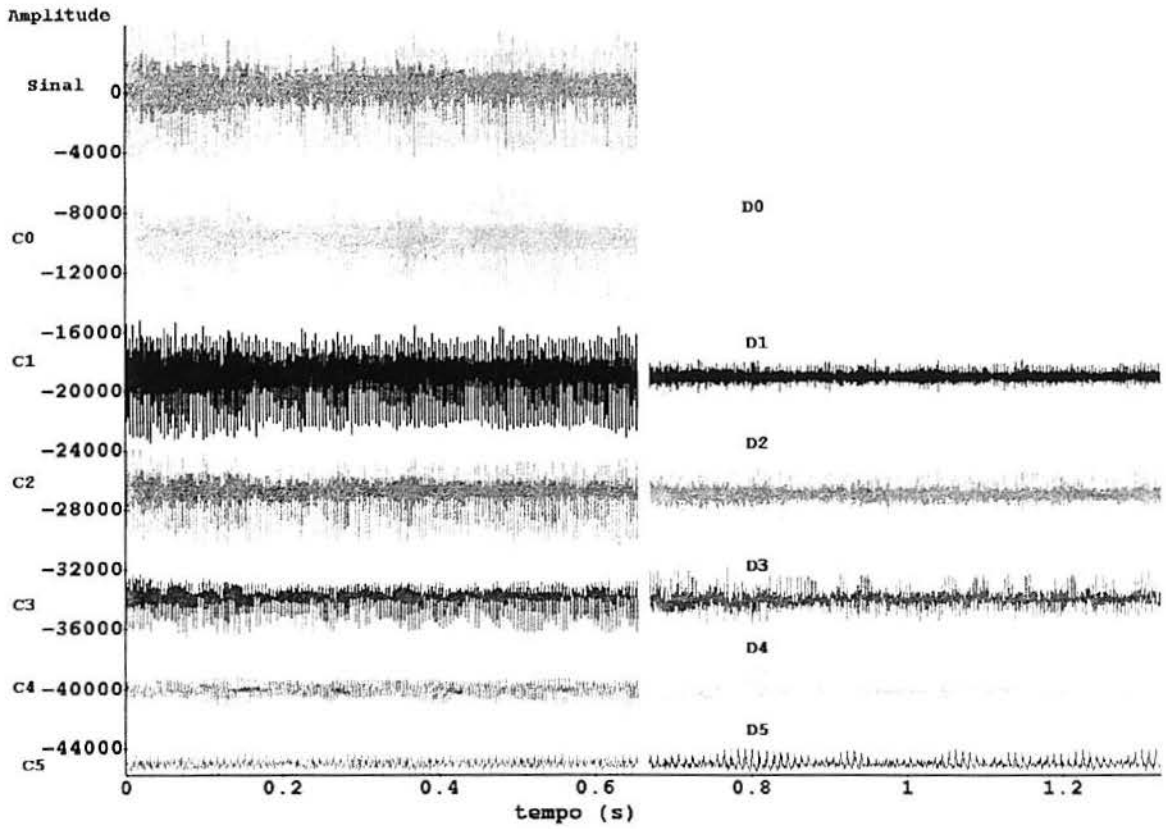


Figura 5.37: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC74.

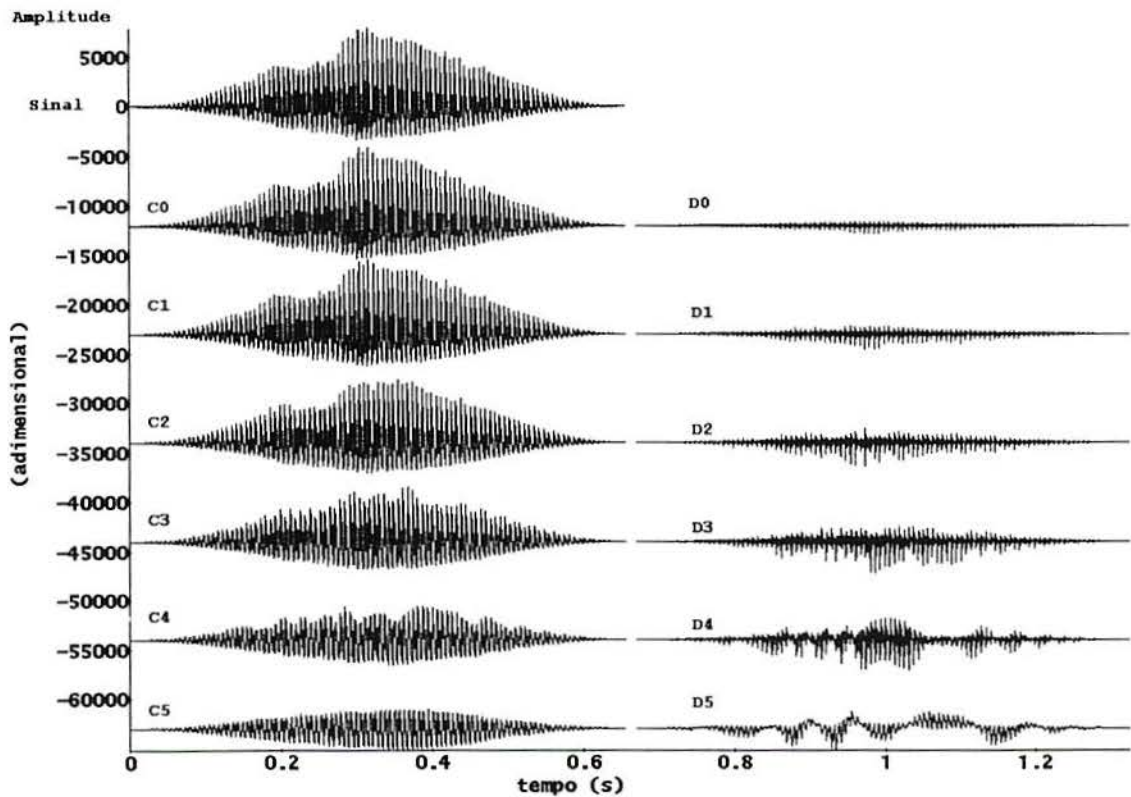


Figura 5.38: Gráficos do trecho de 32762 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC10, com segmentação por janela Hamming.

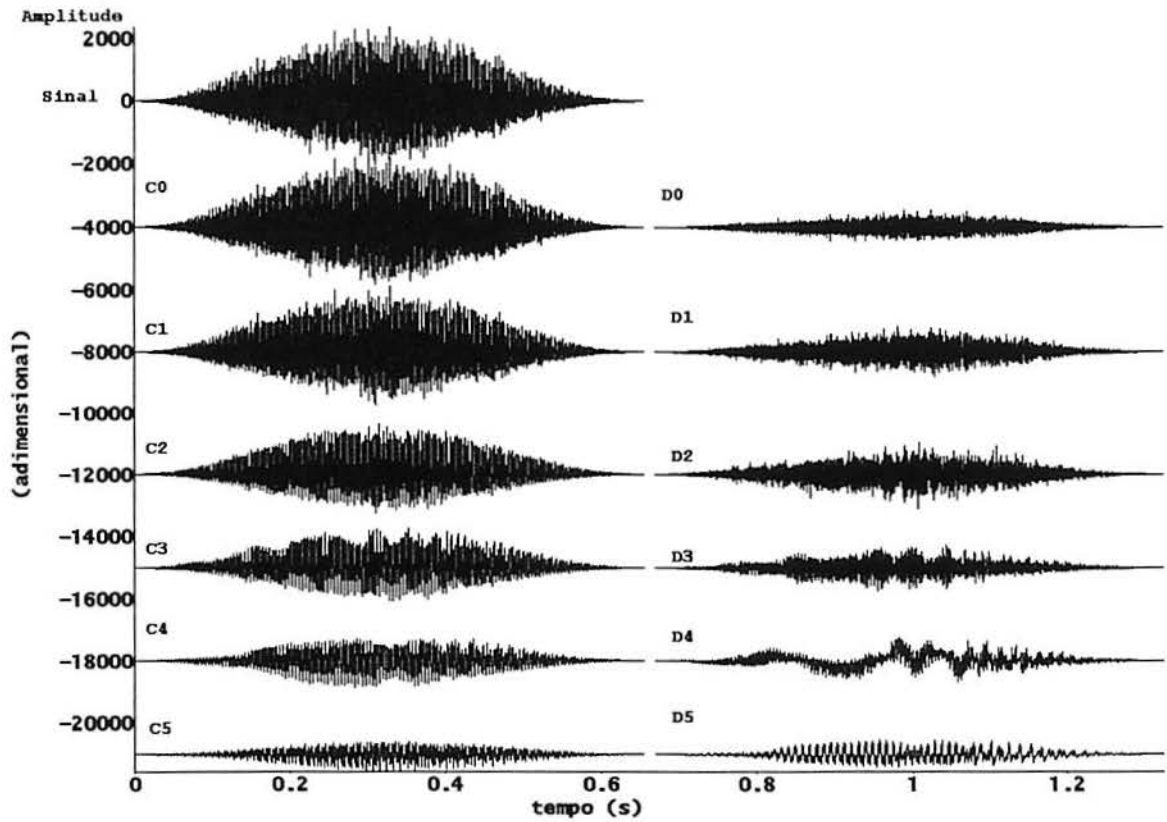


Figura 5.39: Gráficos do trecho de 32762 pontos do sinal e os coeficientes dos 6 níveis da DWT, para a base de Haar do paciente PAC15, com segmentação por janela Hamming.

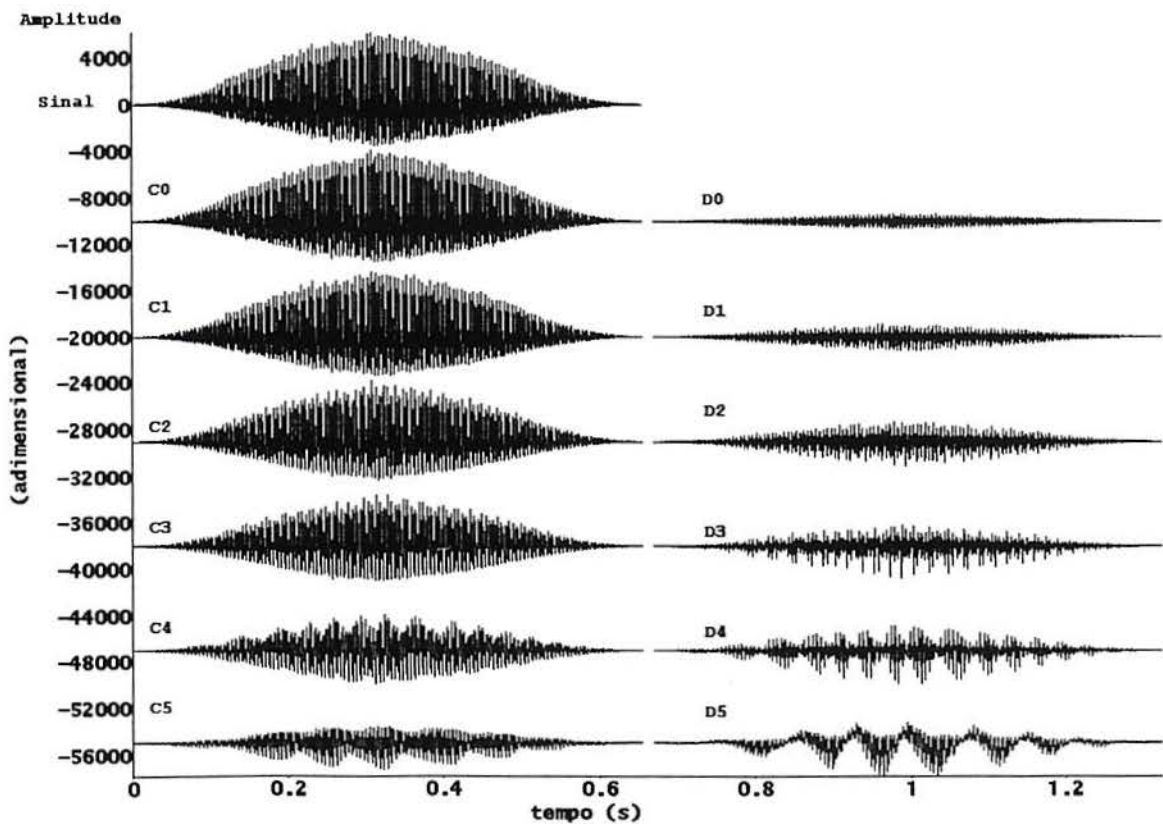


Figura 5.40: Gráficos do trecho de 32762 pontos do sinal e os coeficientes dos 6 níveis da DWT, para a base de Haar do paciente PAC26, com segmentação por janela Hamming.

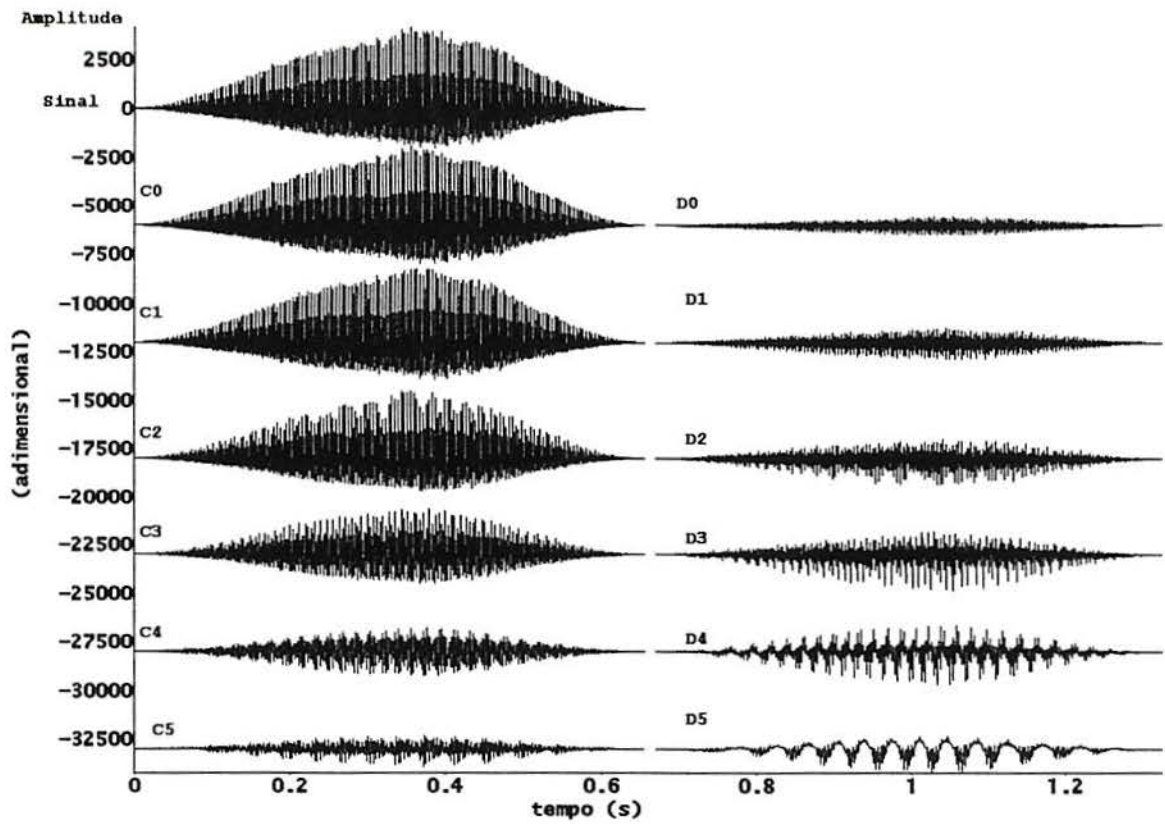


Figura 5.41: Gráficos do trecho de 32762 pontos do sinal e os coeficientes dos 6 níveis da DWT, para a base de Haar do paciente PAC29, com segmentação por janela Hamming.

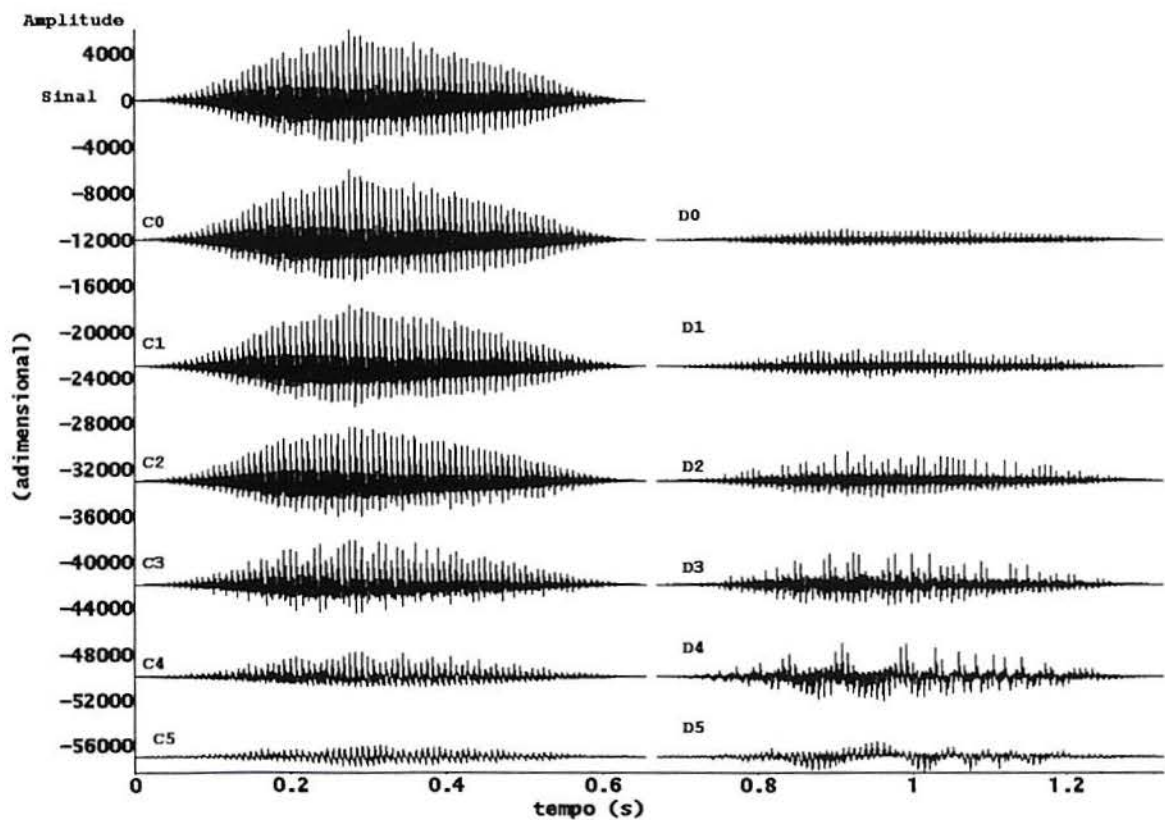


Figura 5.42: Gráficos do trecho de 32762 pontos do sinal e os coeficientes dos 6 níveis da DWT, para a base de Haar do paciente PAC35, com segmentação por janela Hamming.

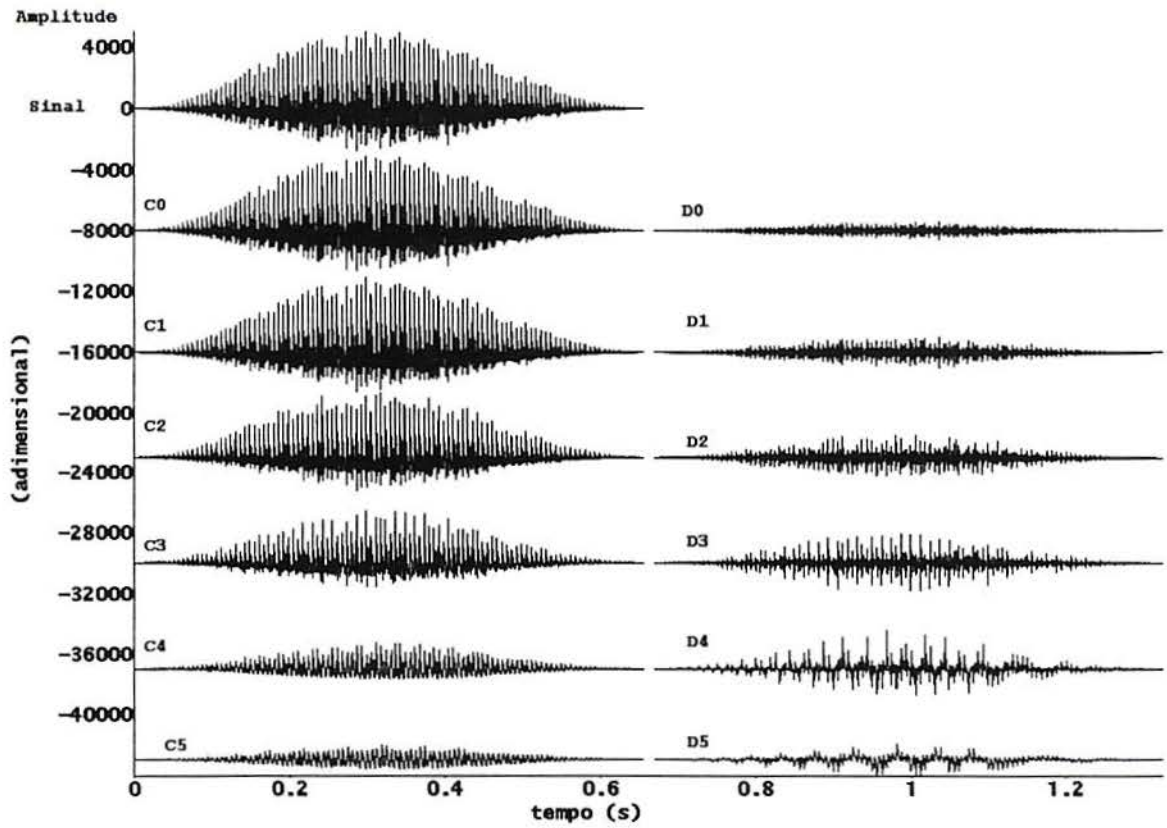


Figura 5.43: Gráficos do trecho de 32762 pontos do sinal e os coeficientes dos 6 níveis da DWT, para a base de Haar do paciente PAC45, com segmentação por janela Hamming.

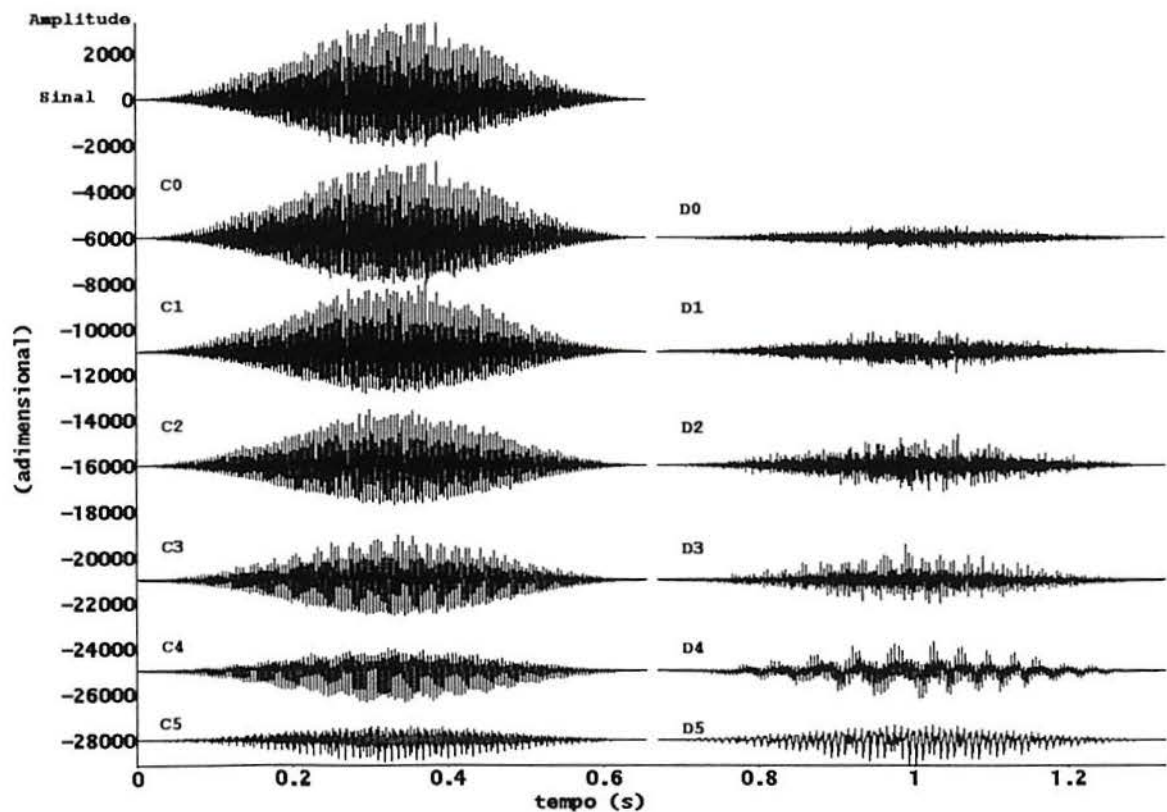


Figura 5.44: Gráficos do trecho de 32762 pontos do sinal e os coeficientes dos 6 níveis da DWT, para a base de Haar do paciente PAC48, com segmentação por janela Hamming.



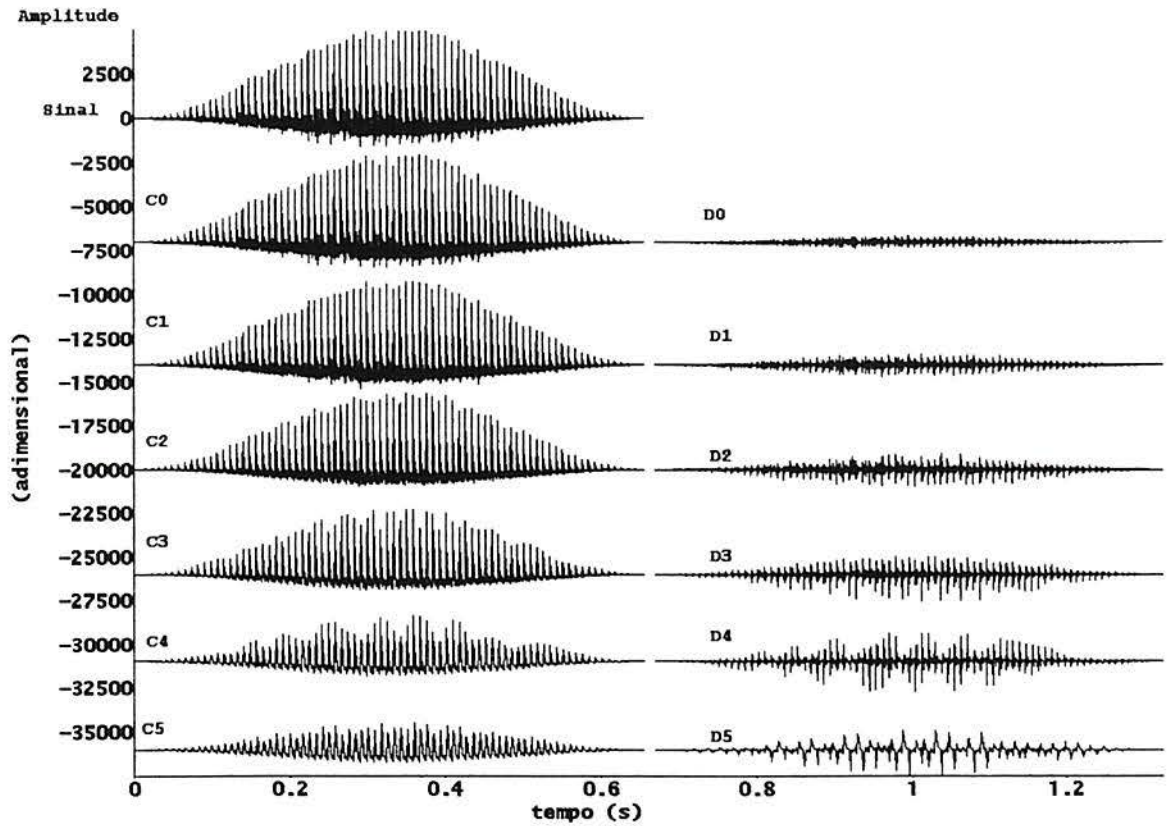


Figura 5.45: Gráficos do trecho de 32762 pontos do sinal e os coeficientes dos 6 níveis da DWT, para a base de Haar do paciente PAC49, com segmentação por janela Hamming.

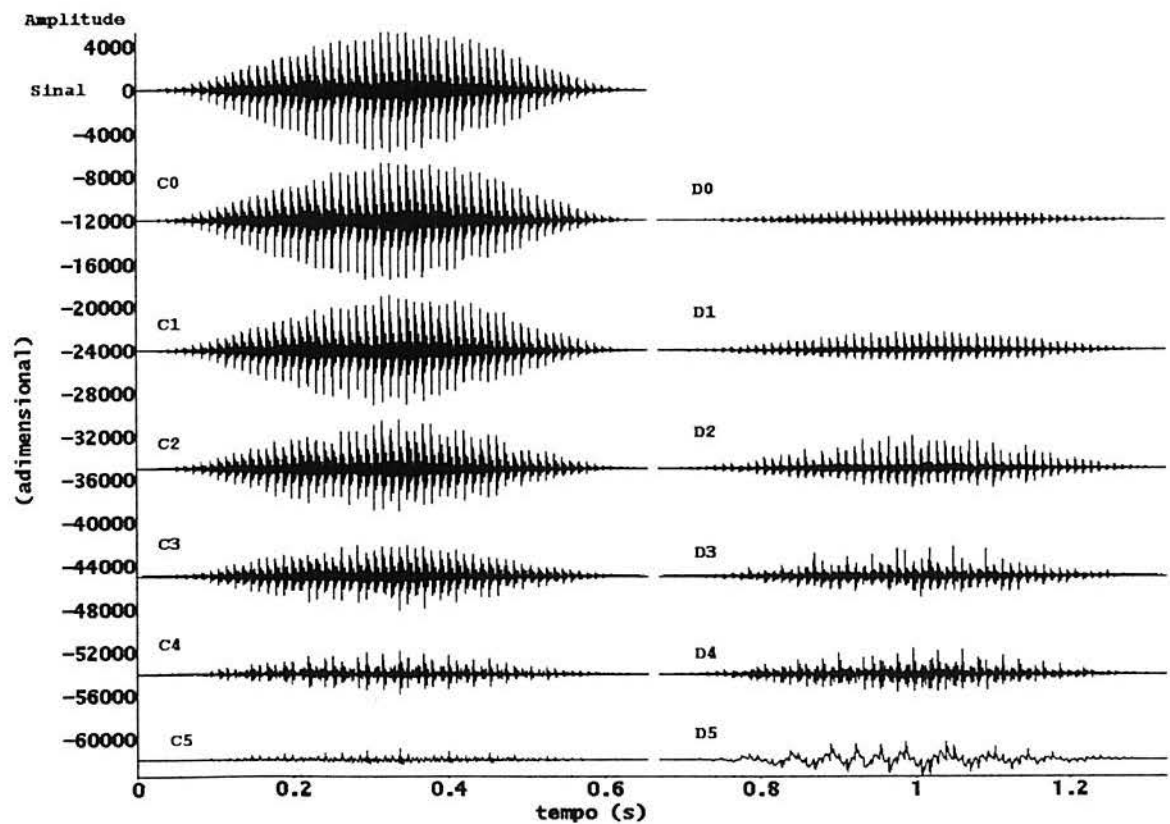


Figura 5.46: Gráficos do trecho de 32762 pontos do sinal e os coeficientes dos 6 níveis da DWT, para a base de Haar do paciente PAC55, com segmentação por janela Hamming.

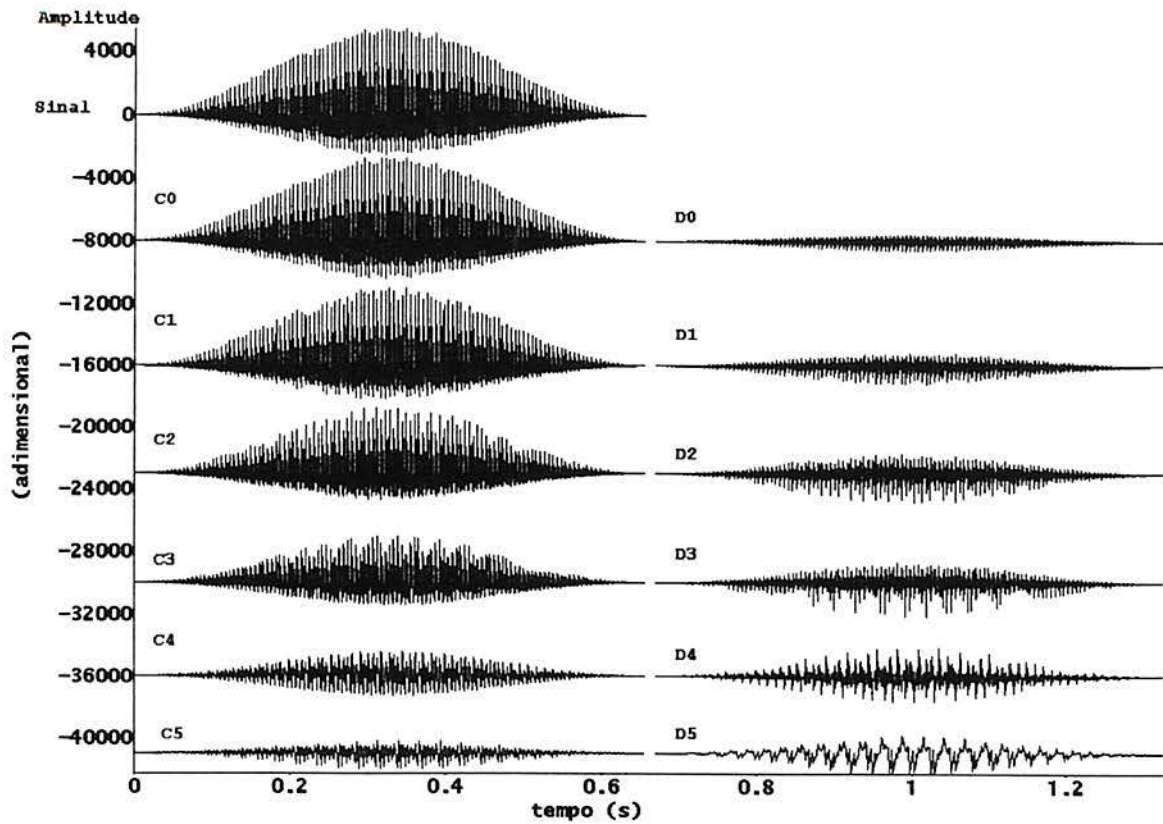


Figura 5.47: Gráficos do trecho de 32762 pontos do sinal e os coeficientes dos 6 níveis da DWT, para a base de Haar do paciente PAC56, com segmentação por janela Hamming.

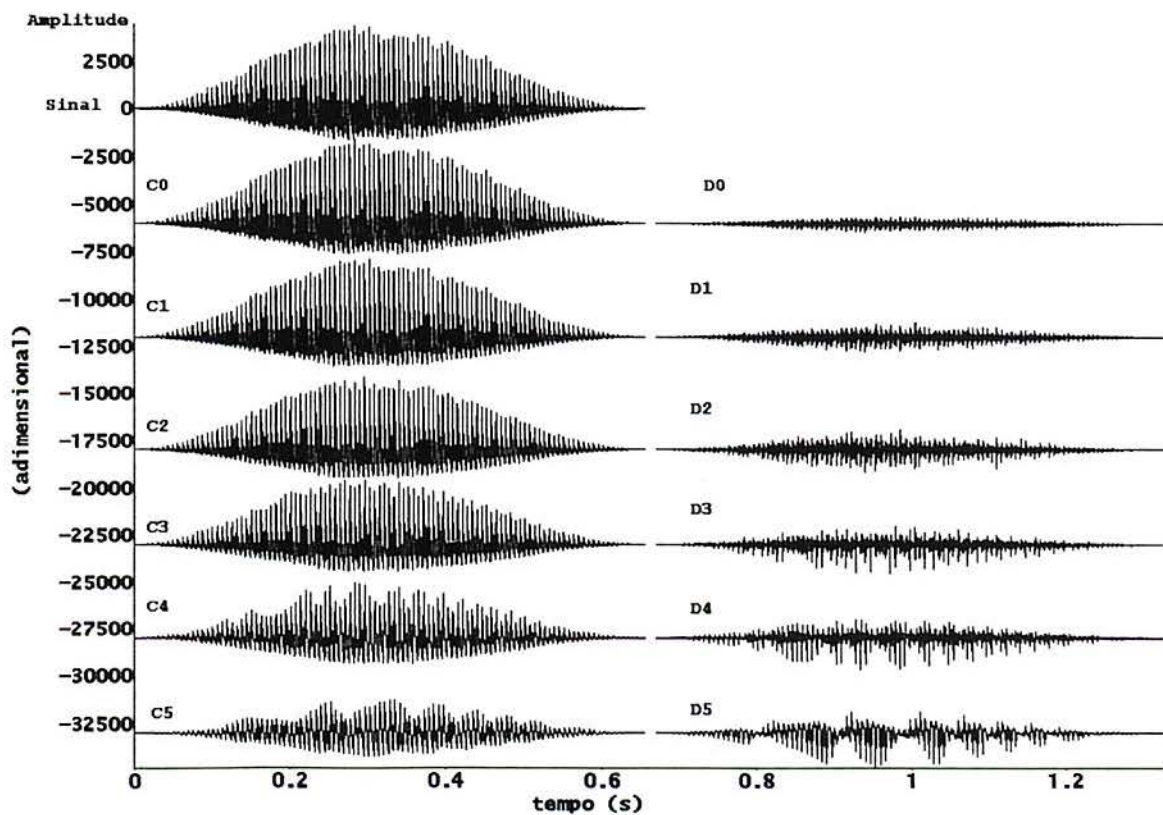


Figura 5.48: Gráficos do trecho de 32762 pontos do sinal e os coeficientes dos 6 níveis da DWT, para a base de Haar do paciente PAC59, com segmentação por janela Hamming.

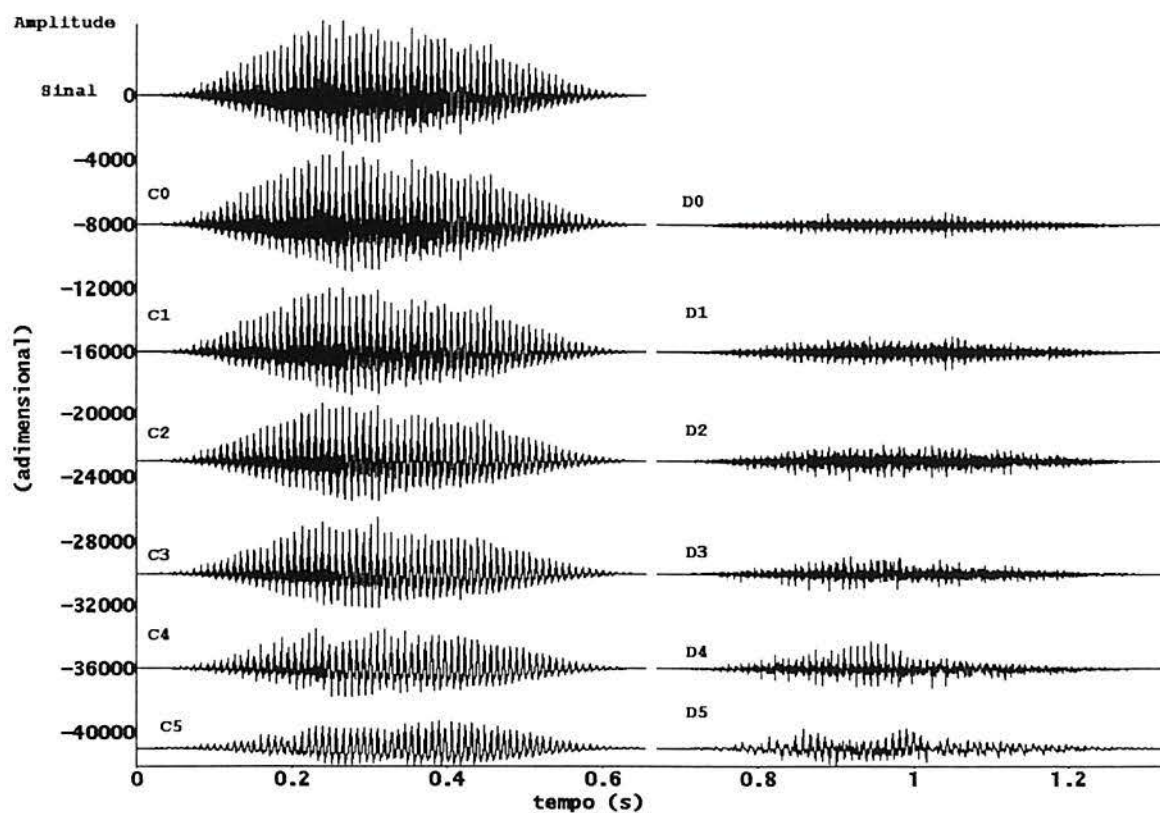


Figura 5.49: Gráficos do trecho de 32762 pontos do sinal e os coeficientes dos 6 níveis da DWT, para a base de Haar do paciente PAC62, com segmentação por janela Hamming.

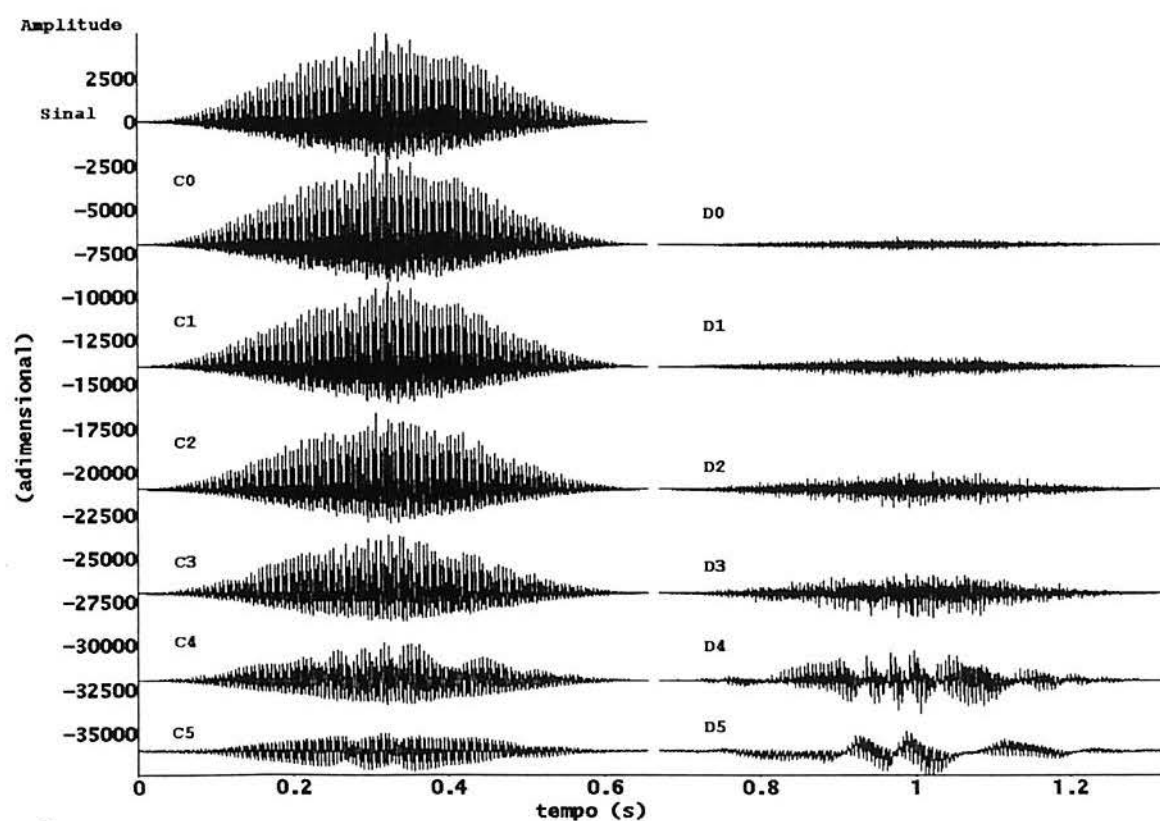


Figura 5.50: Gráficos do trecho de 32762 pontos do sinal e os coeficientes dos 6 níveis da DWT, para a base de Haar do paciente PAC63, com segmentação por janela Hamming.

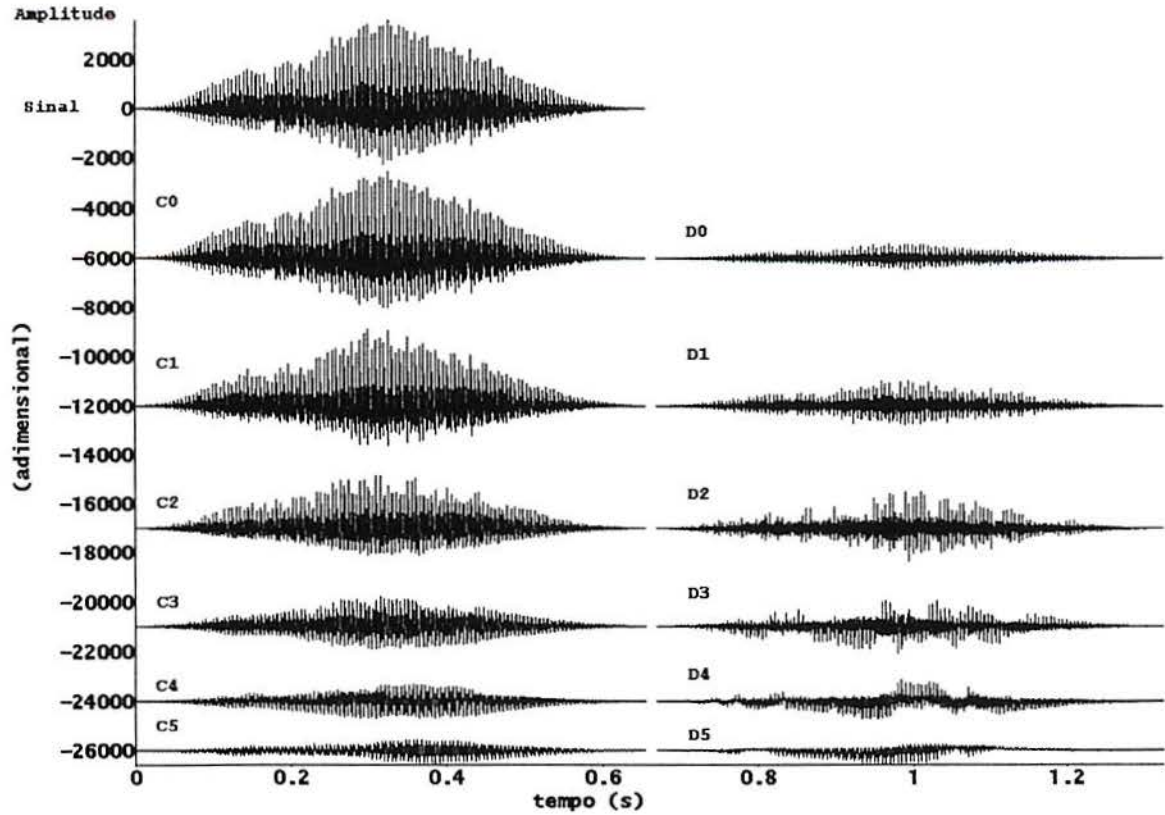


Figura 5.51: Gráficos do trecho de 32762 pontos do sinal e os coeficientes dos 6 níveis da DWT, para a base de Haar do paciente PAC67, com segmentação por janela Hamming.

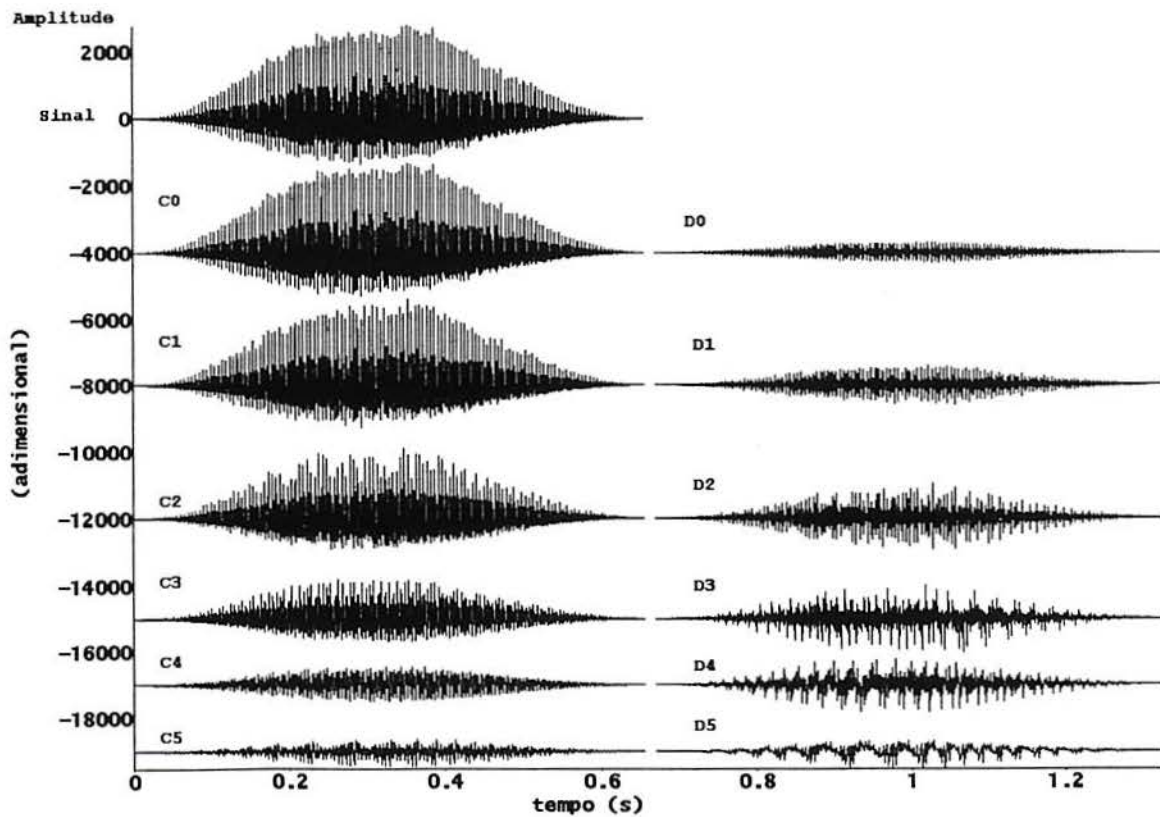


Figura 5.52: Gráficos do trecho de 32762 pontos do sinal e os coeficientes dos 6 níveis da DWT, para a base de Haar do paciente PAC68, com segmentação por janela Hamming.

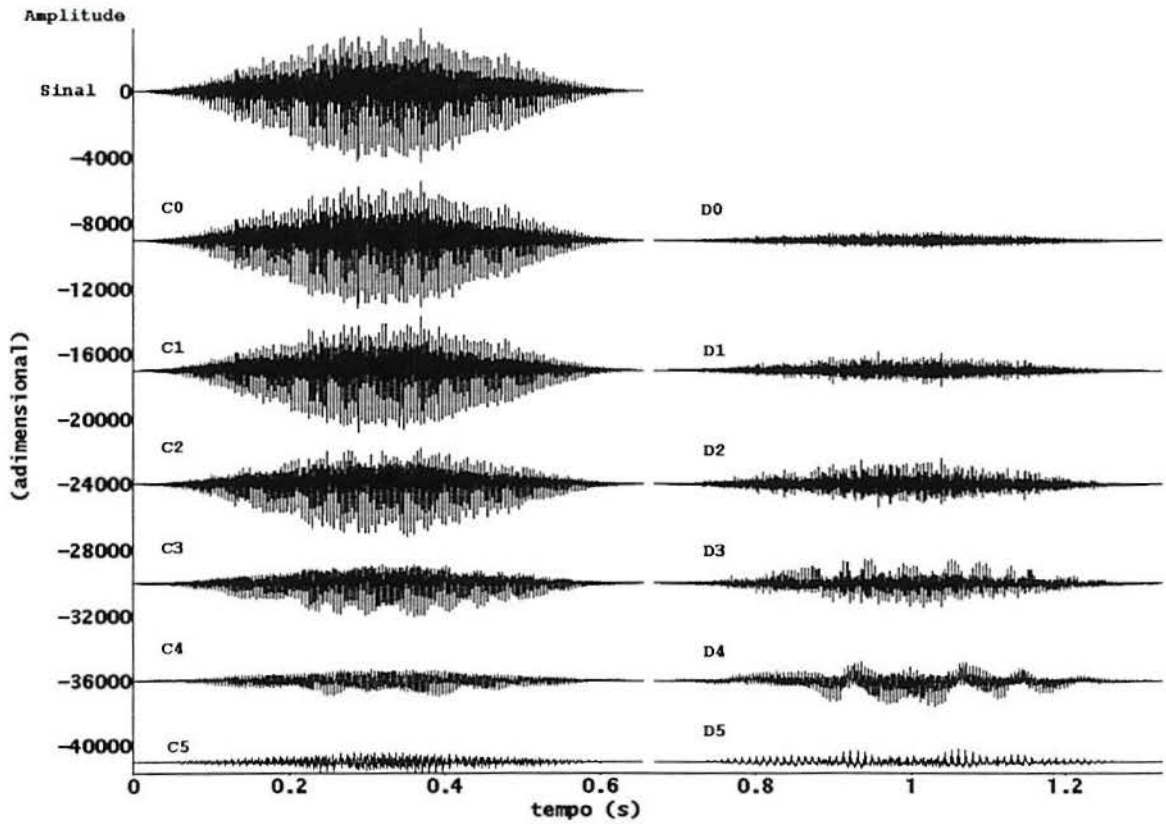


Figura 5.53: Gráficos do trecho de 32762 pontos do sinal e os coeficientes dos 6 níveis da DWT, para a base de Haar do paciente PAC74, com segmentação por janela Hamming.

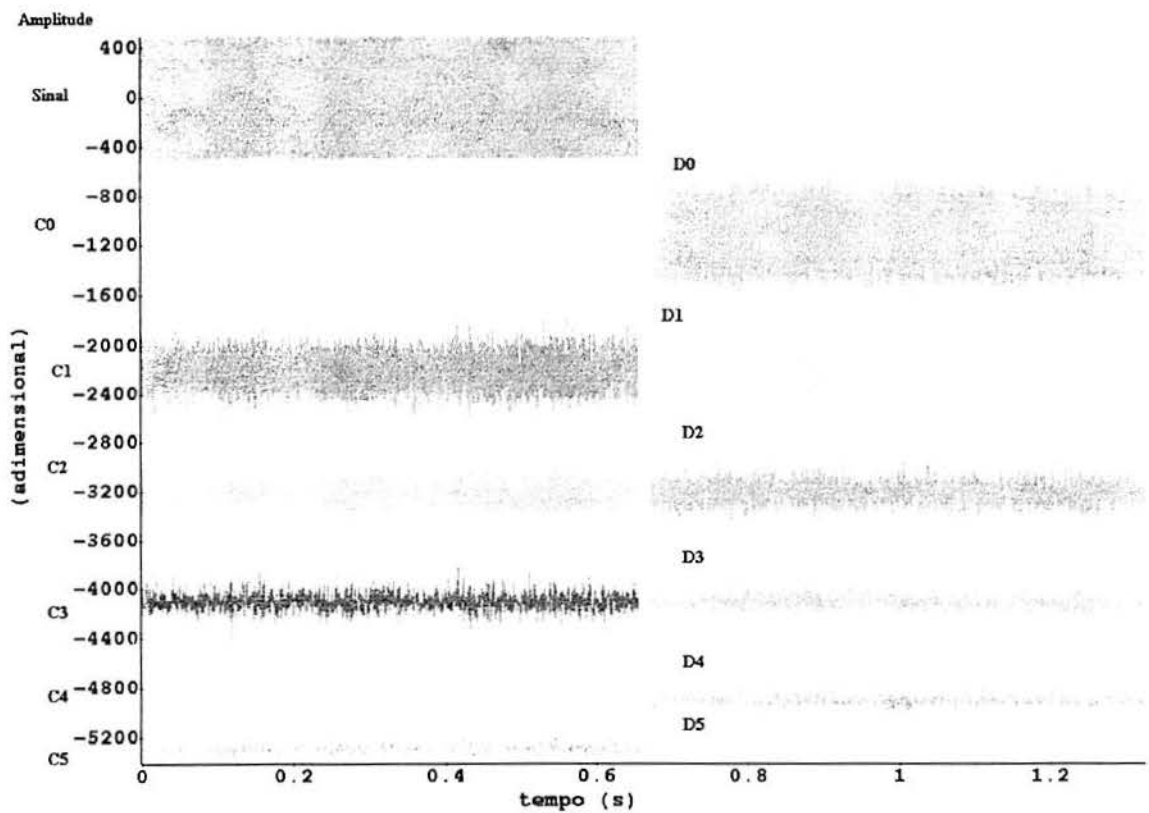


Figura 5.54: Sinal aleatório gerado pelo MATHLAB e os coeficientes dos 6 níveis da DWT, para a base de Haar.

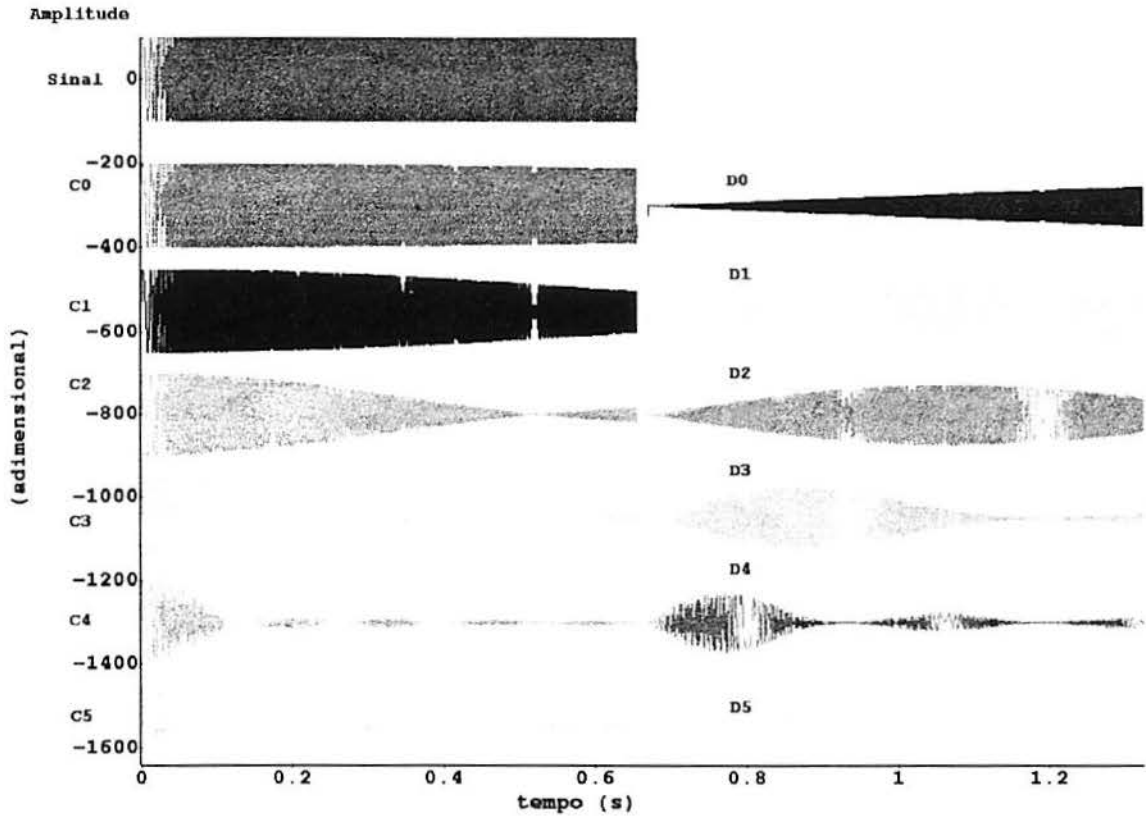


Figura 5.55: Senóide crescente de 0 a 8000 Hz, gerada pelo MATHLAB e seus coeficientes dos 6 níveis da DWT, para a base de Haar.

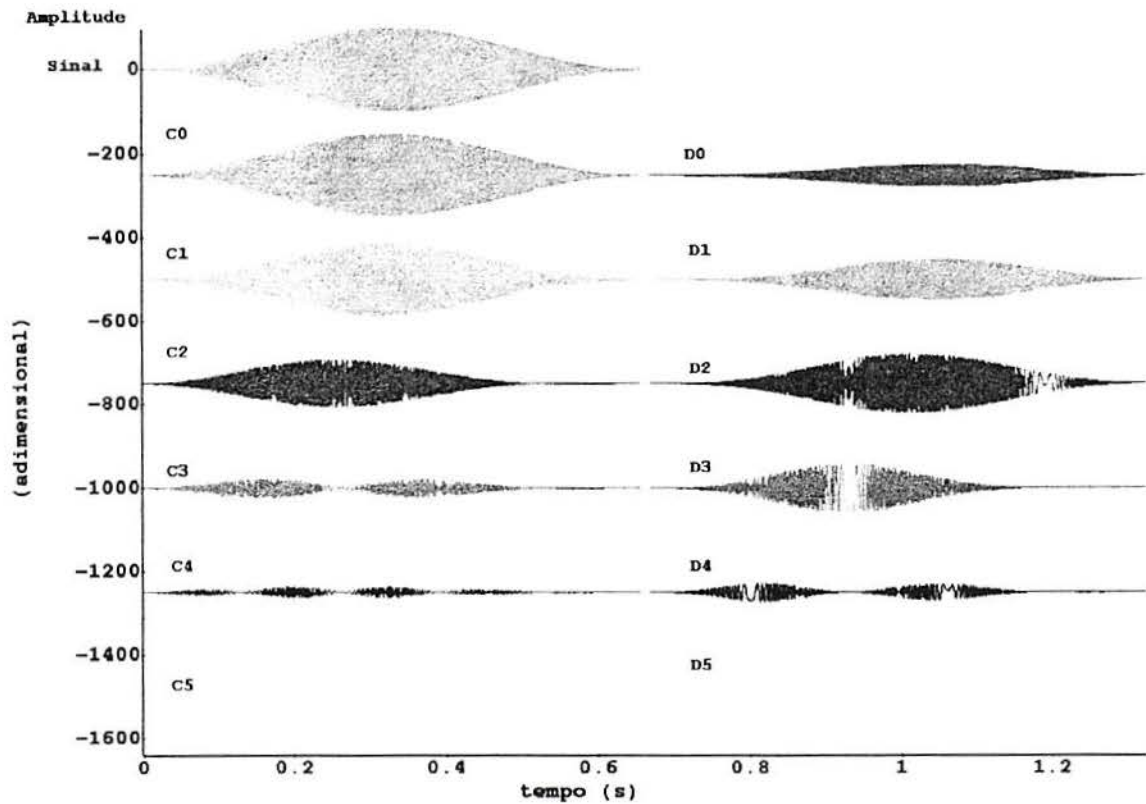


Figura 5.56: Senóide crescente de 0 a 8000 Hz, gerada pelo MATHLAB, segmentada pela janela Hamming, e seus coeficientes dos 6 níveis da DWT, para a base de Haar.

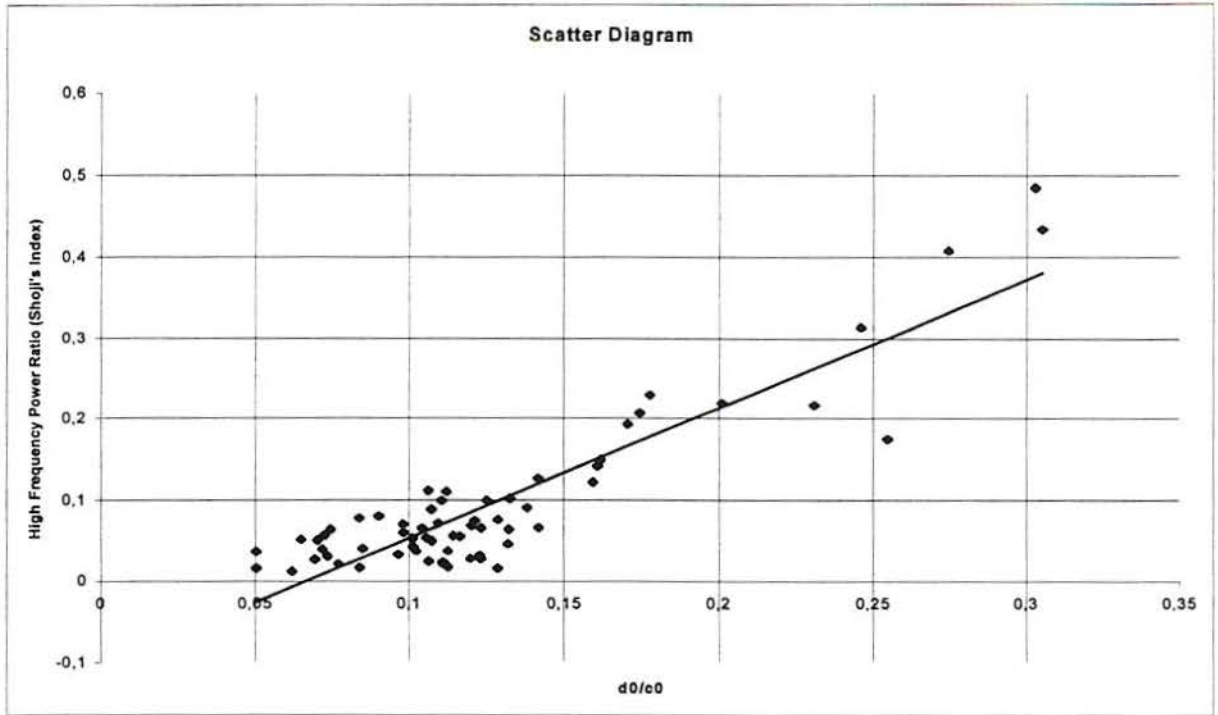


Figura 5.57: Diagrama de espalhamento entre o Índice de Shoji e o índice calculado com o uso dos coeficientes " $d_0$ " e " $c_0$ ".

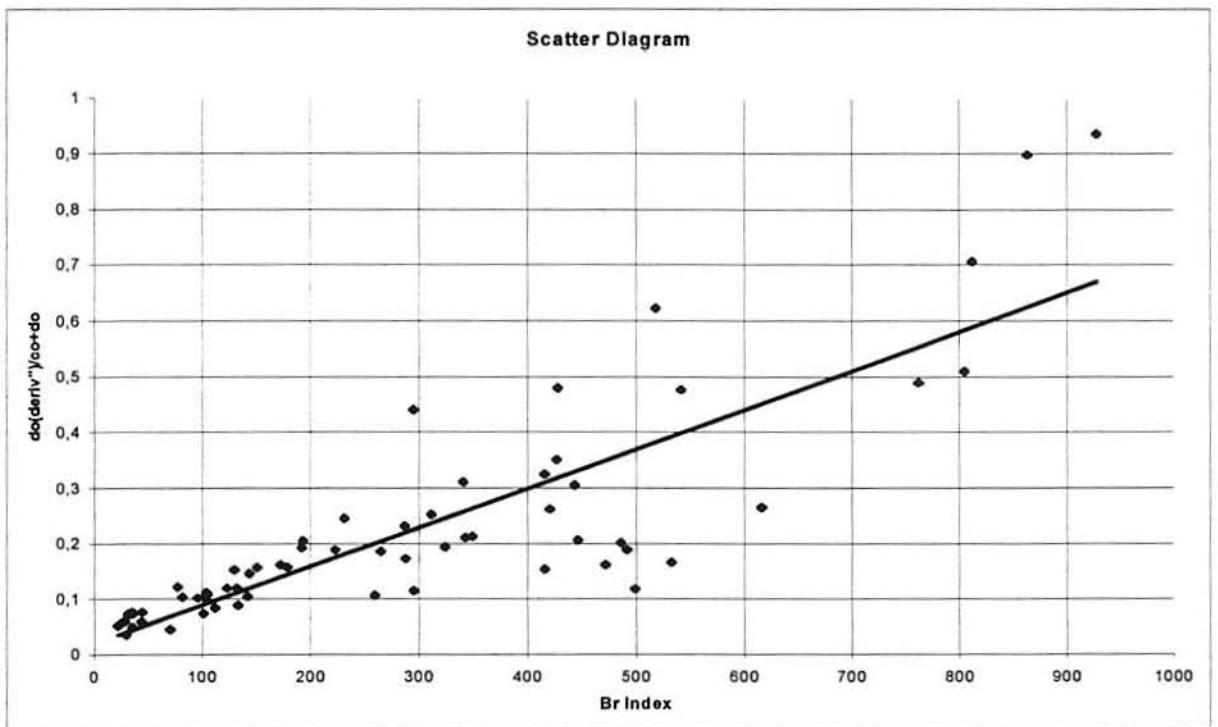


Figura 5.58: Diagrama de espalhamento entre o índice Br. e o índice que usa a segunda derivada do coeficiente " $d_0$ ".

## 6 DISCUSSÃO DOS RESULTADOS E CONCLUSÕES:

Inicialmente, observando-se os estimadores dos espectros de potência do ruído branco gravado e reproduzido sem Dolby e gravado e reproduzido com Dolby (Figuras 5.1 e 5.2, respectivamente), pode-se notar uma diferença significativa apenas na parte superior dos espectros, claramente acima dos 16000 Hz. Esta diferença fica mais evidente nos gráficos em [dBs] (Figuras 5.3 e 5.4, respectivamente), mostrando a declividade mais acentuada causada pelo sistema Dolby. Para se verificar o quanto um estimador difere do outro, foi feita a diferença entre os espectros de potência, em [dBs] (Figura 5.5), mostrando-se que esta permanece dentro da faixa compreendida entre -1dB e +1dB até os 12000 Hz. A diferença apenas supera a faixa dos 3dB, para frequências acima dos 16000Hz. Uma vez que a banda efetiva do sinal utilizada neste trabalho é de 12500Hz (em decorrência das taxas de amostragem escolhidas), conclui-se que a utilização de sistema Dolby não afeta a fidelidade do sinal analisado, na banda de frequências utilizada.

Observando os gráficos dos coeficientes das decomposições pela DWT dos sinais adquiridos pelo primeiro procedimento, nota-se a presença de um "efeito de modulação" ou variações irregulares de longo termo na amplitude dos coeficientes de mais alta dilatação ("d<sub>3</sub>", "d<sub>4</sub>" e "c<sub>4</sub>") não existentes no sinal original. Este fenômeno fica bem explicitado principalmente nas Figuras 5.6 a 5.21. Conforme a literatura (citada a seguir), isto se deve a 2 efeitos relacionados com o processo de decomposição por transformadas Wavelet discretas diádicas com bases reais:

- 1) o efeito de "Phase Sensitivity", ou sensibilidade à fase em bases reais, significando que os valores dos coeficientes variam com o alinhamento das oscilações das funções de



decomposição com mesma frequência que as oscilações do sinal (Simpson, 1993). Este efeito pode ser compensado utilizando-se bases ortonormais complexas, que exploram a diferença de fase de 90 graus entre o componente real e o componente imaginário da base de decomposição (Newland, 1995);

2) o efeito de "aliasing" do processo de filtragem e decimação usado no cálculo dos coeficientes da DWT. Isto se dá pelo fato de que os filtros QMF digitais utilizados (coeficientes da base de Haar) não serem ideais, não tendo uma transição brusca entre a banda de passagem e a banda de rejeição, permitindo que parte do sinal na banda de rejeição vaze, causando o "aliasing" quando da decimação (Strang e Nguyen, 1996, pg 92). Pelo fato de se usar o coeficiente de um nível anterior para se calcular os próximos níveis, este efeito vai se acumulando ao longo das decomposições, se tornando mais evidente nos coeficientes de dilatações mais altas (partes de baixa frequência do espectro do sinal). Este é um problema intrínseco à análise multi-resolução com bases reais. A solução sugerida pela literatura é a utilização de uma base não ortogonal (com redundância, Sheng, 1996).

Visando testar se o efeito acima não foi causado por problemas do sistema de aquisição, foi feita uma nova aquisição, usando-se um segundo procedimento de aquisição (descrito no Capítulo 4). Em seguida foi aplicada a DWT, agora para 6 níveis de dilatação. Os gráficos para os mesmos pacientes são mostrados nas Figuras 5.22 a 5.37. Pode-se observar que o fenômeno permanece, embora as irregularidades mudem um pouco.

A seguir, para se testar se o fenômeno não estaria sendo causado por vazamento de frequências causado pela segmentação do sinal por janela retangular ("leakage"), segmentou-se os mesmos trechos dos sinais do segundo procedimento por uma janela Hamming e fez-se

novamente a DWT para 6 níveis de dilatação. Os resultados são apresentados nas Figuras 5.38 a 5.53. Observa-se que o fenômeno permanece o mesmo, apesar da segmentação por janela Hamming.

Finalmente, gerou-se sinteticamente 2 sinais, através do aplicativo MATHLAB 5.0 (MathWork, Co): primeiro, um sinal randômico e segundo uma senóide de frequência crescente no tempo, de 0 a 8000Hz, ambos com frequência de amostragem de 50000Hz (2 sinais teóricos, sem influência de sistemas de aquisição). Foi aplicada a DWT para os 2 sinais, ambos com 6 níveis de dilatação (Figuras 5.54 e 5.55, respectivamente). Observa-se que para o sinal randômico, não aparece o efeito. Isto acontece porque os coeficientes da DWT, calculados por produto interno (de forma semelhante à correlação cruzada, Sheng, 1996), mostram a correlação entre o sinal e as funções de base, no caso, ruído aleatório e as funções de Haar, respectivamente. Neste caso, não pode haver correlação entre eles e os coeficientes da DWT serão também aleatórios, o que é confirmado pela Figura 5.54.

Já para a senóide crescente, o fenômeno de "modulação" aparece de novo (Figura 5.55), agora em função da frequência crescente, mostrando a não-idealidade dos filtros digitais. Após, a senóide crescente foi segmentada pela janela Hamming e foi feita nova DWT, para 6 níveis de dilatação (apresentado na Figura 5.56). Observa-se que o efeito permaneceu o mesmo, apesar da nova segmentação. Isto reforça a hipótese de que os 2 efeitos relacionados ao processo de DWT anteriormente citados são os responsáveis pela geração das irregularidades de amplitude de longo termo (o fenômeno de "modulação").

Um outro problema da análise multi-resolução com bases ortonormais citado na literatura (Sheng, 1996), mas não investigado neste trabalho, é a falta de invariância com

deslocamento ("shift invariance"), significando que se o sinal de entrada é deslocado de uma unidade de tempo, os coeficientes da transformada não são simplesmente versões deslocadas dos coeficientes anteriormente calculados, mas também variam de valor. Isto explicaria porque as irregularidades de amplitude de longo termo dos sinais do primeiro procedimento de aquisição são diferentes das irregularidades dos sinais do segundo procedimento, pois foram adquiridos trechos diferentes do sinal de cada paciente (da passagem da fita cassete para o computador, é impossível se conseguir iniciar exatamente no mesmo ponto 2 vezes). Para este problema, é citado que Wavelets não ortonormais do tipo "quadratic spline" tem sido usadas para se evitar este problema em análises de sinais em tempo real e reconhecimento de padrões (Sheng, 1996).

Devido ao de irregularidades de longo termo ("efeito de modulação") nos coeficientes de maior dilatação ( $c_3$  e  $d_3$ ), fica evidente que não é interessante se utilizar a técnica da análise multi-resolução com bases discretas reais ortonormais para tentar criar índices que possam ser correlacionados com índices relativos à variações na frequência fundamental da voz como "Jitter" e "Shimmer". Para este fim, seria melhor uma base complexa, talvez com redundância.

Entretanto, este trabalho não é afetado pelo problema falta de invariância com deslocamento ("shift invariance"), pois para o cálculo dos indicadores de soproidade sugeridos, utilizou-se o valor proporcional ao rms dos coeficientes da DWT, tendo-se então uma média temporal quadrática de 8192 pontos,.

Observa-se , principalmente pelos pacientes PAC05, PAC08, PAC13, PAC16, PAC21, PAC34 e PAC46 (Figuras A1.4, A1.6, A1.9, A1.11, A1.15, A1.22, A1.24 e A1.30) com altos

índices de ruído de banda larga associado ao sinal de voz, característica do escape turbulento de ar causado pela má oclusão das pregas vocais e associado à característica soprosidade da voz (Shoji, 1991 ) têm a relação entre a amplitude do coeficiente de alta frequência " $d_0$ " e a amplitude do coeficiente de baixa frequência " $c_0$ " aumentada em relação aos outros pacientes. Ainda, pacientes com pouco ruído aditivo e morfologia do período de fundamental clara , principalmente PAC01, PAC17, PAC26, PAC27, PAC44, PAC45, PAC47, PAC48, PAC49, PAC53 e PAC54 (Figuras A1.1, A1.12, 5.24, A1.19, A1.29, 5.27, A1.31, 5.28, 5.29, A1.34 e A1.35), apresentam maiores amplitudes nos níveis de alta dilatação (componentes de baixa frequência " $d_3$ " e " $d_4$ ") do que nos níveis de baixa dilatação e alta frequência (" $d_0$ " e " $d_1$ "). Estas observações sugerem que um índice de soprosidade possa ser extraído da relação entre os coeficientes " $d_0$ " e " $c_0$ ", para o primeiro procedimento de aquisição, uma vez que neste nível de resolução, tem-se o espectro do sinal dividido em 2, (a porção de alta e baixa frequência do sinal, respectivamente).

São então propostos 2 maneiras de se calcular um índice de soprosidade e ambas as maneiras são comparadas aos índices clássicos de soprosidade propostos na literatura.

Olhando-se para o diagrama de espalhamento da Figura 5.57, observa-se que existe uma correlação entre o índice proposto por Shoji *et al.* (1992a) e o índice calculado pelo quociente entre os valores rms do coeficiente " $d_0$ " e " $c_0$ ", para o primeiro procedimento de aquisição. O valor do coeficiente de correlação linear entre os índices, para os 63 pacientes é 0.905, o que significa estatisticamente uma excelente correlação (Anexo 6).

Também pode-se observar pelo diagrama de espalhamento da Figura 5.58 que existe uma correlação entre o índice Br e o índice que usa o valor rms da segunda derivada do

coeficiente " $d_0$ ", para o primeiro procedimento de aquisição. Neste caso, o coeficiente de correlação é 0,848, o qual expressa também uma boa correlação entre os índices. Para o segundo procedimento de aquisição, a correlação linear entre o índice proposto por Shoji *et al.* (1992a) e o índice calculado pelo quociente entre o valor rms do coeficiente " $d_1$ " e o de " $c_1$ " foi de 0,63, e a correlação entre o índice Br e o índice que usa o valor rms da segunda derivada do coeficiente " $d_1$ " foi de 0,5, ambas também satisfatórias para o número de pacientes usados. Assim, pode-se utilizar a transformada Wavelet Discreta, mais especificamente a análise multi-resolução, para se encontrar um índice relacionado com a soproside da voz.

Pelos resultados obtidos e pela discussão apresentada, pode-se concluir que:

1) a análise multi-resolução com bases ortonormais discretas e reais apresenta problemas para a análise temporal em tempo real de sinais e reconhecimento de padrões (Sheng, 1996) e que devido aos fenômenos de "aliasing" e sensibilidade de fase, não é interessante tentar-se extrair índices que sejam correlacionados com fenômenos da frequência fundamental como "jitter" e "shimmer";

2) embora com os problemas supracitados, pode-se obter índices relacionados com a soproside da voz com bom coeficiente de correlação com os índices propostos e testados na literatura, com a vantagem de que o cálculo daqueles através da transformada discreta ser, neste caso, mais simples do que os 2 índices propostos pela literatura.

Tem-se como sugestões futuras e prosseguimento natural deste trabalho:

1) utilizar a transformada Wavelet contínua, com uma base complexa não ortonormal, para a análise dos sinais de voz, tentando-se índices que se correlacionem com o "jitter" e "shimmer".

2) utilizar-se o conceito de Wavelet adaptativa, tal como em Mallet *et al.* (1997), para se tentar classificar patologias através do sinal de voz.

3) utilizar a lógica "fuzzy" para se implementar um classificador de tipo de patologia que utilize conjuntamente diversos índices de análise acústica, calculados através de diferentes técnicas, fazendo-se uma verdadeira análise multi-variáveis.

## REFERÊNCIAS BIBLIOGRÁFICAS:

- AIRAINERaumentada em relação a, R. AND KLINGHOLZ, F. (1993):"Quantitative Evaluation of Phonetograms in the case of Functional Dysphonia". J. of Voice, vol.7, n.2, pp.136-141.
- ALDROUBI, A. AND UNSER, M. (1996): "Wavelets in Medicine and Biology,". CRC Press, Inc., Boca raton, FL.
- ATAL, B. S. AND HANAUER, S. L. (1971):"Speech Analysis and Synthesis by Linear Prediction of the Speech Wave". J. Acoust. Soc. Am. , vol.50, n.2 (part 2), pp.637-655.
- BAKEN, R. J. (1987):"Clinical Measurement of Speech and Voice,".Allyn and Bacon, a division of Simon and Schuster, Inc., Needham Heights, MA.
- BECKETT, R. L. (1969):"Pitch Perturbation as a Function of Subjective Vocal Constriction". Folia Phoniatr., vol.21, pp.416-425.
- BENDAT, J. S. AND PIERSOL, A. G. (1986): "Random Data: Analysis and Measurement Procedures, 2ed.,". A Wiley-IntersciencePublication, John Wiley & Sons, Inc, Ny.
- BENTLEY, P. M. AND MCDONNELL, J. T. E. (1994):"Wavelet transforms: An Introduction". Electronics & Communication Engineering Journal, pp.175-186, August.
- BERRY, D. A. ; HERZEL, H. ; TITZE, I. R. AND KRISCHER, K. (1994):"Interpretation of Biomechanical Simulations of Normal and Chaotic Vocal Fold Oscillations with Empirical Eigenfunctions". J. Acoust. Soc. Am. , vol.95, n.6, pp.3595-3604, June.
- BOONE, D. R. E MCFARLANE, S. C. (1994):"A Voz e a Terapia Vocal-5ª ed.,".Editora Artes Médicas, Porto Alegre, RS.
- CASEY, S. D. AND WALNUT, D. F. (1994):"Systems of Convolution Equations, Deconvolution, Shannon Sampling, and the Wavelet and Gabor Transforms". SIAM Review, vol.36(4), pp.537-577, December.
- CHILDERS, D. G. AND BAE, K. S. (1992):"Detection of Laryngeal Function using Speech and Electroglottographic Data". IEEE Trans. Biom. Eng., vol.39, n.1, pp.19-25, January.
- CODY, M. A. (1992):"The Fast Wavelet Transform". Dr. Dobb's Journal, pp. 16-28 and 100-101, April.
- CODY, M. A. (1993):"A Wavelet Analyser". Dr. Dobb's Journal, pp.44-54,84-84; April.
- CODY, M. A. (1994):"The Wavelet Packet Transform". Dr. Dobb's Journal, pp. 44-54, 100; April.
- COLEMAN, R. F. (1971):"Effect of Waveform Changes upon Roughness perception". Folia Phoniatr., vol.23, pp.314-322.

COLTON, R. H. e KASPER, J. K. (1996): "Compreendendo os Problemas da Voz: Uma Perspectiva Pisiológica ao Diagnóstico e ao Tratamento,".Editora Artes Médicas, Porto Alegre, RS.

CROWE, J. A.; GIBSON, N. M.; WOOLFSON, M. S. AND SOMEKH, M. G. (1992):"Wavelet Transform as a potential tool for ECG Analysis and Compression", J. Biom. Eng., vol.14, pp.268-272, May.

DAUBECHIES, I. (1988):"Orthonormal Bases of Compactly Supported Wavelets", Communications on Pure and Applied Mathematics, vol. XLI, 909-996.

DAUBECHIES, I. (1989):"Orthonormal Bases of Wavelets with Finite Support-Connection with Discrete Filters", In Wavelets: Time-Frequency Methods and Phase Space, pp.38-67, Springer-Verlag, Berlin.

DAUBECHIES, I. (1990):"The Wavelet Transform, Time-Frequency Localization and Signal Analysis", IEEE Trans. on Inform. Theory, vol.36, pp. 961-1005.

DAUBECHIES, I. (1992): "Ten Lectures on Wavelets,". Society for Industrial and Applied Mathematics (SIAM) book n.61, Philadelphia, Pennsylvania.

DAVIS, R. E. (1979):"Acoustic Characteristics of Normal and Pathologic Voices", Speech and Language: Advances in Basic Research and Practice, vol.1, pp.271-335, Academic Press, N.J. .

DEAL, R. E. AND EMANUEL, F. W. (1978):"Some Waveform and Spectral Features of Vowel Roughness", J. of Speech and Hear. Res., vol.21, pp.250-264.

DELLER JR., J. R. AND ANDERSON, D. J. (1980):"Automatic Classification of Laryngeal Dysfunction using the Roots of the Digital Inverse Filter", IEEE Trans. on Biom. Eng. , Vol.BME 27, pp.714-721.

DELLER JR., J. R. ; PROAKIS, J. G. AND HANSEN, J. H. L. (1993):"Discrete-Time Processing of Speech Signals,". MacMillan Publishing Company, New York, NY.

EMANUEL, F. W. AND AUSTIN, D. (1981):"Identification of Normal and Abnormal Rough Vowel by Spectral Noise Level Measurements", J. of Communication Disorders, vol.14, pp.75-85.

EMANUEL, F. W. ; LIVELY, M. A. AND MCCOY, J. F. (1973):"Spectral Noise Levels and Roughness Severity Ratings for Vowels produced by males and females", Folia Phoniatr., vol.25, pp.110-120.

EMMANUEL, F. W. AND WHITEHEAD, R. L. (1979):"Harmonic Levels and Vowel Roughness", J. of Speech Hear. Research, vol.22, pp.829-840.

ESKENAZI, L. ; CHILDERS, D. G. AND HICKS, D. M. (1990):"Acoustic Correlates of Vocal Quality", J. Speech Hear. Research, vol.33, pp.298-306.



- FUKAZAWA, T. ; EL-ASSUOOTY, A. AND HONJO, I. (1988):"A New Index for Evaluation of the Turbulent Noise in Pathological Voice". J. Acoust. Soc. Am., vol.83, n.3, pp.1189-1193, March.
- GABOR, D. (1946):"Theory of communication". Journal IEE, vol.93, pp. 429-457.
- GOLD, B. AND RABINER, L. (1969):"Parallel processing Techniques for Estimating Pitch Periods of Speech in the Time Domain". J. of the Acoust. Soc. Am., vol.46, pp.442-448.
- GROSSMANN, A. AND MORLET, J. (1984):"Decomposition of Hardy Functions into square-integrable Wavelets of constant Shape". SIAM J. of Math., vol 15, pp.723-736.
- GROSSMANN, A. ; KRONLAND-MARTINET, R. AND MORLET, J. (1989):"Reading and Understanding Continuous Wavelet Transform". In Wavelets: Time-Frequency Methods and Phase Space, pp.2-20, Springer-Verlag, Berlin.
- HAMMARBERG, B. ; FRITZELL, G. ; GAUFFIN, J. ; SUNDBERG, J. AND WEDIN, L. (1980):"Perceptual and Acoustic Correlates of Abnormal Voice Qualities". Acta Otolaryngol. ,vol.90, pp.441-451.
- HECKER, M. H. L. AND KREUL, E. J. (1971):"Descriptions of the Speech of Patients with Cancer of the Vocal Folds. Part I: measures of the fundamental frequency". J. Acoust. Soc Am., vol.49, pp.1275-1282.
- HERMANSKY, H. (1990):"Perceptual Linear Predictive (PLP) Analysis of Speech". J. Acoust. Soc. Am. , vol.87, n.4, pp.1738-1752, April.
- HIKI, S.; MATSUOKA, K. ; IMAIZUMI, S.; HIRANO, M. AND MATSUSHITA, H. (1975):"A Study on Acoustical Analysis of Hoarseness". J Acoust. Soc. Jpn, vol.31, pp.504-506.
- HIRANO, M. AND BLESS, D. M. (1993):"Videostroboscopic Examination of the Larynx". Singular Publishing Group, Inc; San Diego, CA.
- HIRAOKA, N. ; KITAZOE, Y. ; UETA, H. ; TANAKA, S. AND TANABE, M. (1984):"Harmonic-Intensity Analysis of Normal and Hoarse Voices". J. Acoust. Soc. Am. , vol.76, n.4, pp.1648-1651, December.
- HOLLIEN, H. ; MICHEL, J. AND DOHERTY, E. T. (1973):"A Method for Analysing Vocal Jitter in Sustained Phonation". J. of Phonetics, vol.1, pp.85-91.
- HONJO, I. AND ISSHIKI, N. (1980):"Laryngoscopic and Voice Characteristics of Aged Persons". Archives of Otolaryngology, vol.106, pp.149-150.
- HORII, Y. (1979):"Fundamental Frequency Perturbation observed in sustained phonation". J. of Speech and Hearing Research, vol.22, pp.5-19.
- HORII, Y. (1980):"Vocal Shimmer in sustained phonation". J. of Speech and Hearing Research, vol.23, pp.202-209.

- HORII, Y. (1982): "Jitter and Shimmer Differences among sustained vowel phonations", J. of Speech and Hearing Research, vol.25, pp.12-14.
- HORII, Y. (1983): "Automatic Analysis of Voice Fundamental Frequency and Intensity using a VisiPitch", J. of Speech and Hearing Research, vol.26, pp.467-471.
- HORII, Y. (1985): "Jitter and Shimmer in Sustained Vocal Fry Phonation", Folia Phoniatr., vol. 37, pp.81-86.
- ISHIZAKA, K AND ISSHIKI, N. (1976): "Computer Simulation of Pathological Vocal-Cord Vibration", J. Acoust. Soc. Am., vol.60, n.5, pp.1193-1198, November.
- ISSHIKI, N. (1989): "Phonosurgery-Theory and Practice", Springer-Verlag, Tokyo, Japan.
- ISSHIKI, N. ; YANAGIHARA, N. AND MORIMOTO, M. (1966): "Approach to the Objective Diagnosis of Hoarseness", Folia Phoniatr., vol.18, pp.393-400.
- IWATA, S. (1972): "Periodicities of Pitch Perturbations in Normal and Pathological Larynges", Laryngoscope, vol.82, pp.87-96.
- IWATA, S. AND VON LENDEN, H. (1970 a): "Pitch Perturbations in Normal and Pathologic Voices", Folia Phoniatr., vol.22, pp.413-424.
- IWATA, S. AND VON LENDEN, H. (1970 b): "Voice Prints in Laryngeal Diseases", Arch. Otolaryngol., vol.91, pp.346-351.
- JAFARI, M. ; TILL, J. A. ; TRUESDELL, L. F. AND LAW-TILL, C. B. (1993): "Time-Shift, Trial, and Gender Effects on Vocal Perturbation Measures", J. of Voice, vol.7, n.4, pp.326-336.
- KAISER, G. (1994): "A Friendly Guide to Wavelets", Birkhäuser Boston, Cambridge, MA.
- KASUYA, H. ; OGAWA, S. ; KIKUCHI, Y. AND EBIHARA, S. (1986a): "An Acoustic Analysis of Pathological Voice and Its Application to the Evaluation of Laryngeal Pathology", Speech Communication, vol.5, n.2, pp.171-181, June.
- KASUYA, H. ; OGAWA, S. ; MASHIMA, K. AND EBIHARA, S. (1986b): "Normalized Noise Energy as an Acoustic Measure to Evaluate Pathologic Voice", J. Acoust. Soc Am., vol.80, n.5, pp.1329-1334, November.
- KAY, S. M. AND MARPLE JR., S. T. (1981): "Spectrum Analysis- A Modern Perspective", Proceedings of the IEEE, vol.69, n.11, pp.1380-1419, November.
- KENT, R. D. AND READ, C. (1992): "The Acoustics Analysis of Speech", Singular Publishing Group, Inc., San Diego, CA.
- KITAJIMA, K. (1981): "Quantitative Evaluation of the Noise Level in the Pathologic Voice", Folia Phoniatr., vol.33, pp.115-124.

- KITAJIMA, K. ; TANABE, M. AND ISSHIKI, N. (1975): "Pitch Perturbation in Normal and Pathological Voice". *Studia Phonologica*, vol.9, pp.25-32.
- KLATT, D. H. AND KLATT, L. C. (1990): "Analysis, Synthesis, and Perception of Voice Quality Variations among Female and Male Talkers". *J. Acoust. Soc. Am.*, vol.87, n.2, pp.820-857, February.
- KLINGHOLZ, F. AND MARTIN F. (1983): "Speech Wave Aperiodicities at Sustained Phonation in Functional Dysphonia". *Folia Phoniatr.*, vol.35, pp.322-327.
- KLINGHOLZ, F. AND MARTIN F. (1985): "Quantitative Evaluation of Shimmer and Jitter". *J. of Speech and Hearing Research*, vol.28, pp.169-174.
- KLINGHOLZ, F. ; PENNING, R. AND LIEBHARDT, E. (1988): "Recognition of Low-Level Alcohol Intoxication form Speech Signal". *J. Acoust. Soc. Am.*, vol.84, n.3, pp.929-935, September.
- KOIKE, Y. (1969): "Vowel Amplitude Modulations in Patients with Laryngeal Diseases". *J. Acoust. Soc. Am.*, vol.45, n.4, pp.839-844.
- KOIKE, Y. (1973): "Applications of some measures for the evaluation of laryngeal dysfunction". *Studia Phonol.*, vol.7, pp.17-23.
- KOIKE, Y. AND MARKEL, J. D. (1975): "Application of Inverse Filtering for Detecting Laryngeal Pathology". *Ann. Otol. Rhinol. Laryngol.*, vol.84, pp.117-124.
- KOIKE, Y. ; TAKAHASHI, H. AND CALCATERA, T. C. (1977): "Acoustic Measures for Detecting Laryngeal Pathology". *Acta Otolaryngologica*, vol.84, pp.105-117.
- KOJIMA, H; GOULD, W. J.; LAMBIASE, A. AND ISSHIKI, N. (1980): "Computer Analysis of Hoarseness". *Acta Otolryngol.*, vol.89, pp.547-554.
- KUHL, I. (1982): "Manual prático de laringologia", Editora da Universidade, livro-texto 11, Porto Alegre.
- LIEBERMAN, P. (1961): "Perturbations in Vocal Pitch". *J. Acoust. Soc. Am.*, vol.33, pp.597-603.
- LIEBERMAN, P. (1965): "Some Acoustic Measures of the Fundamental Periodicity of normal and Pathologic larynges". *J. Acoust. Soc. Am.*, vol.35, pp.344-353.
- LINDBLOM, B. E. F. AND SUNDBERG, J. E. F. (1971): "Acoustical Consequences of Lip, Tongue, Jaw, and Larynx Movement". *J. Acoust. Soc. Am.*, vol.50, n.4 (part 2), pp.1166-1179.
- LUDLOW, C. ; COULTER, D. AND GENTGES, F. (1983): "The Differential Sensitivity of Measure of Fundamental Frequency Perturbation to Laryngeal Neoplasms and neuropathologies". *Vocal Fold Physiology: Contemporary Research and Clinical Issues.*, Cap.33, pp.381-392. College-Hill, San Diego, CA.

- LUFT, J. A. (1994): "Reconhecimento automático de voz para palavras isoladas e independente do locutor,". Tese de mestrado aprovada, apresentada no PPGEMM-UFRGS em 1994.
- MA, C. AND O'SHAUGHNESSY, D. (1994): "A Perceptual Study of Source Coding of Fourier Phase and Amplitude of the Linear Predictive Coding Residual of Vowel Sounds,". J. Acoust. Soc. Am., vol.95, n.4, pp.2231-2239, April.
- MAKHOUL, J. (1975): "Linear Prediction: A Tutorial Review,". Proceedings of the IEEE, vol.63, n.4, pp.561-580, April.
- MALLAT, S. G. (1989A): "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation,". IEEE Trans. on Pattern Analysis and Machine Int., vol 11(7), pp.674-693, July.
- MALLAT, S. G. (1989b): "Multifrequency channel decompositions of images and wavelet models,". IEEE Trans. Acoust., Speech and Signal Process., vol. 37, pp. 20919-2110.
- MALLET, Y.; COOMANS, D.; KAUTSKY, J. and DE VEL, O. (1997): "Classification using Adaptive Wavelets for Feature Extraction,". IEEE Trans. on Pattern Anal. Mach. Intell., vol.19 (10), pp.1058-1066, october.
- MEYER, Y. (1993): "Wavelets: Algorithms and Applications,". Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.
- MORLET, J. (1983): "Sampling Theory and Wave Propagation,". NATO ASI Séries, vol.1, Issues in Acoustical Signal/Image Processing and Recognition, C. H. Chen editor-in-chief, Springer-Verlag, Berlin, pp.233-261.
- MORLET, J.; ARENS, G.; FOURGEAU, I. and GIARD, D. (1982): "Wave propagation and Sampling Theory,". Geophysics, 47, pp.203-236.
- MURRY, T. (1979): "Speaking Fundamental Frequency Characteristics associated with Voice Pathologies,". J. Speech and Hearing Res. , vol.43, pp.374-379.
- MURRY, T AND DOHERTY, E.T. (1980): "Selected Acoustic Characteristics of Pathologic and Normal Speakers,". J. of Speech and Hearing Research, vol.23, pp.361-369.
- NEWLAND, D. E. (1995): "Progress in the Application of Wavelet Theory to Vibration Analysis,". DE-vol.84 (1), Design Engineering Technical Conferences, volume 3, part A, pp.1313-1322, ASME 1995.
- NOLL, A. M. (1964): "Short-term Spectrum and "Cepstrum" Techniques for Vocal Pitch Detection,". J. Acoust. Soc. Am. , vol.36, pp.296-302.
- NOLL, A. M. (1967): "Cepstrum Pitch Determination,". J. Acoust. Soc. Am. , vol.41, n.2, pp.293-309.
- OPPENHEIM, A. V. AND SCHAFFER, R. W. (1989): "Discrete-Time Signal Processing,". Prentice-Hall, Inc, a Simon & Schuster division, Englewood Cliff, NJ.

- REMACLE, M. AND TRIGAUX, I. (1991): "Characteristics of Nodules Through the High-Resolution Frequency Analyzer", Folia Phoniatr., vol.43, pp.53-59.
- RIOUL, O. AND VETTERLI, M. (1991): "Wavelets and Signal Processing", IEEE Signal Processing Magazine, vol.8 (4), pp.14-37, October.
- RONTAL, E. ; RONTAL, M. AND ROLNICK, M. I. (1975): "Objective Evaluation of Vocal Pathology using Voice Spectrography", Ann. Otol. Rhinol. Laryngol., Vol.84, pp.662-671.
- SCHAFFER, R. AND RABINER, L. R. (1970): "System for Automatic Formant Analysis of Voiced Speech", J. Acoust. Soc. Am. , vol.47, n.2 (part 2), pp.634-648.
- SCHAFFER, R. W. AND RABINER, L. R. (1975): "Digital Representations of Speech Signals", Proceedings of the IEEE, vol.63, n.4, pp.662-677, April.
- SHENG, Y (1996): "Wavelet Transform". Em The Transforms and Applications Handbook, POULARIKAS, A. D.(editor-in-chief),CRC Press in cooperation with IEEE Press, Boca Raton, FL, pp747-827.
- SHOJI, K. ; REGENBOGEN, E. ; YU, J. D. AND BLAUGRUND, S. M. (1991): "High-Frequency Components of Normal Voice", Journal of Voice, vol. 5, n.1, pp. 29-35.
- SHOJI, K. ; REGENBOGEN, E. ; YU, J. D. AND BLAUGRUND, S. M. (1992a): "High-Frequency Power Ratio of Breathly Voice", Laryngoscope 102, pp.267-271, March.
- SHOJI, K. ; REGENBOGEN, E. ; YU, J. D. AND BLAUGRUND, S. M. (1992b): "H-Index: A New Measure of Glottal Efficiency for the Pathologic Voice", Laryngoscope 102, pp.1113-1117, October.
- SIMPSON, D. M. (1993): "An Introduction to the Discrete Orthogonal Wavelet Transform", Revista Brasileira de Engenharia, Caderno de engenharia Biomédica, vol. RBE 9 (1), pp.57-81.
- SORENSEN, D. AND HORII, Y. (1983): "Frequency and Amplitude Perturbations in the Voices of Female Speakers", J. of Communication Disorders, vol.16, pp.57-61.
- SORENSEN, D. AND HORII, Y. (1984): "Directional Perturbation Factors for Jitter and for Shimmer", J. of Communication Disorders, vol.17, pp.143-151.
- STRANG, G. (1989): "Wavelets and Dilatation Equations: A Brief Introduction", SIAM Review, vol.31(4), pp.614-627.
- STRANG, G. AND NGUYEN, T. (1996): "Wavelets and Filter Banks". Wellesley-Cambridge Press, Wellesley, MA.
- TAKAHASHI, H. AND KOIKE, Y. (1975): "Some Perceptual Dimensions and Acoustic Correlates of Pathologic Voices", Acta Otolaryngol. , Suppl.338, pp.1-24.

TITZE, I. (1976): "On the Mechanism of Vocal-Fold Vibration", J. Acoust. Soc. Am., vol.60, n.6, pp.1366-1380, December.

TITZE, I. R. (1989): "Physiologic and Acoustic Differences between Male and Female Voices", J. Acoust. Soc. Am., vol.85, n.4, pp.1699-1707, April.

TITZE, I. R. (1993): "Vocal Fold Physiology: Frontiers in Basic Science", Singular Publishing Group, San Diego, CA.

TITZE, I. R. (1994): "Towards Standards in Acoustic Analysis of Voice", J. of Voice, vol.8, n.1, pp.1-7.

TITZE, I. R. AND STRONG, W. J. (1975): "Normal Modes in vocal cord Tissues", J. Acoust. Soc. Am., vol.57, n.3, pp.736-744, March.

TITZE, I. R. AND TALKIN, D. T. (1979): "A Theoretical Study of the Effects of Various Laryngeal Configurations on the Acoustics of Phonation", J. Acoust. Soc. Am., vol.66, n.1, pp.60-74, July.

von LENDEN, H. AND KOIKE, Y. (1970): "Detection of Laryngeal Disease by Computer Technique", Arch. Otolaryngol., vol.91, pp.3-10.

YANAGIHARA, N. (1967): "Significance of Harmonic changes and Noise Components in Hoarseness", J. of Speech Hearig. Res., vol. 10, pp.531-541.

YOUNG, R. K. (1993): "Wavelet Theory and its Applications". Kluwer Academic Publishers, Norwell, MA, USA.

YUMOTO, E. AND GOULD, W. J. (1982): "Harmonic-to-Noise Ratio as an Index of the Degree of Hoarseness", J. Acoust. Soc. Am., vol.71, n.6, pp.1544-1549, June.

ANEXO 1: Sinais e suas DWTs dos pacientes do primeiro procedimento de aquisição:

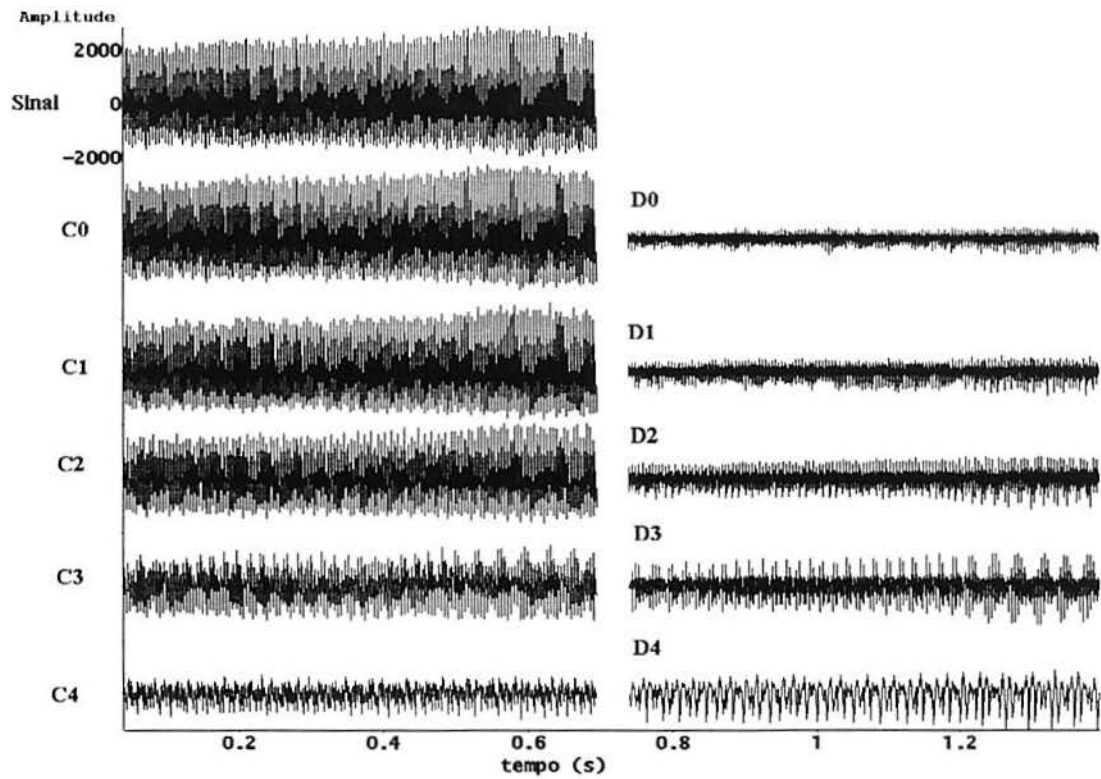


Figura A1.1: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC01.

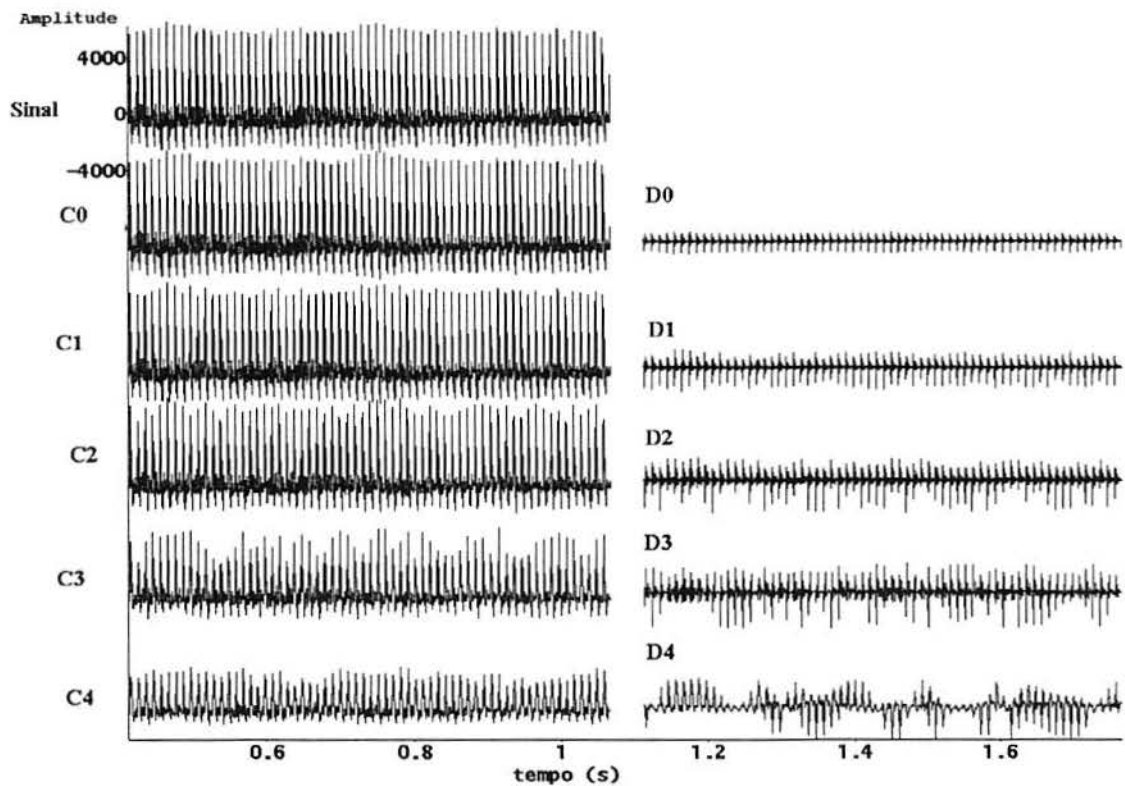


Figura A1.2: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC02.

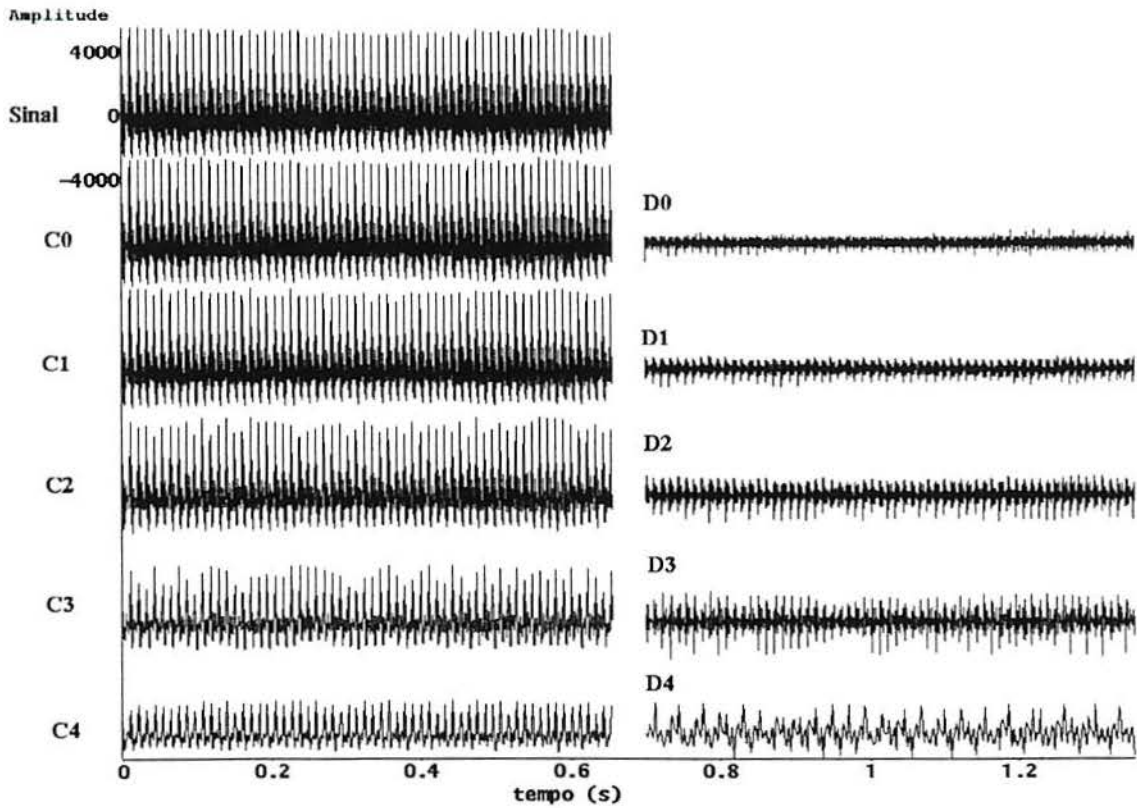


Figura A1.3: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC04.

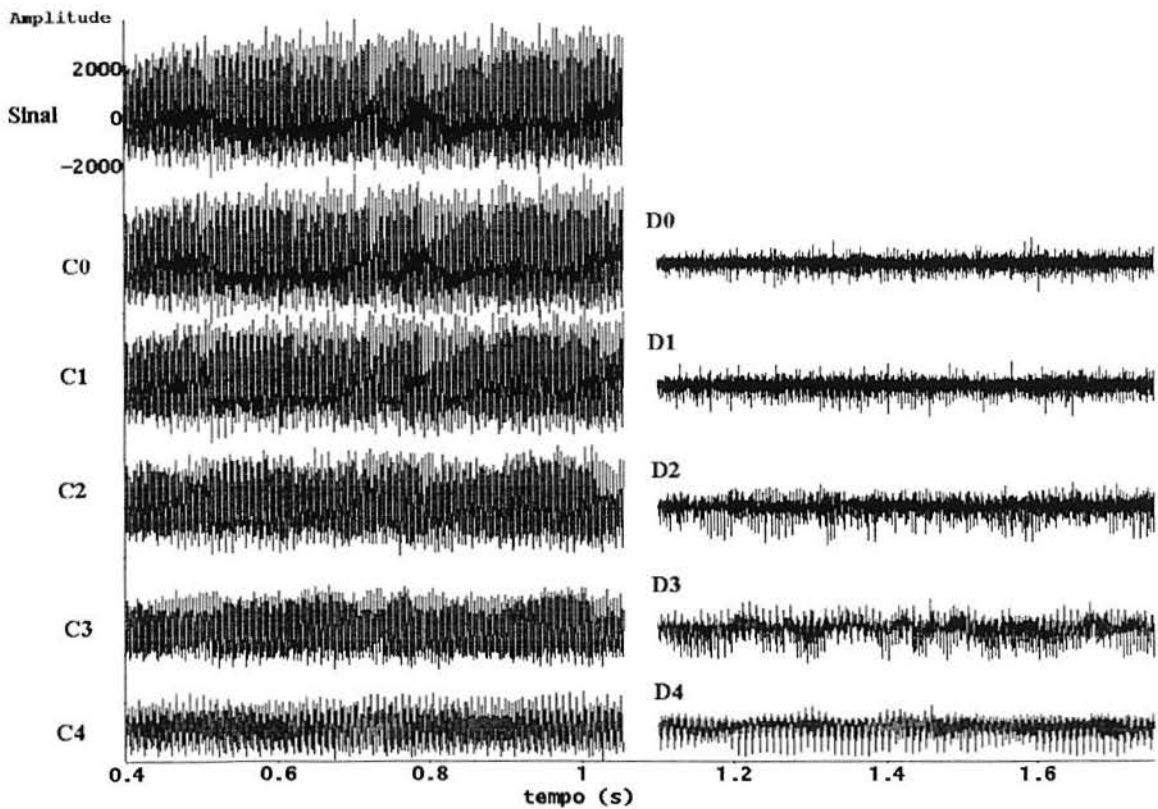


Figura A1.4: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC05.



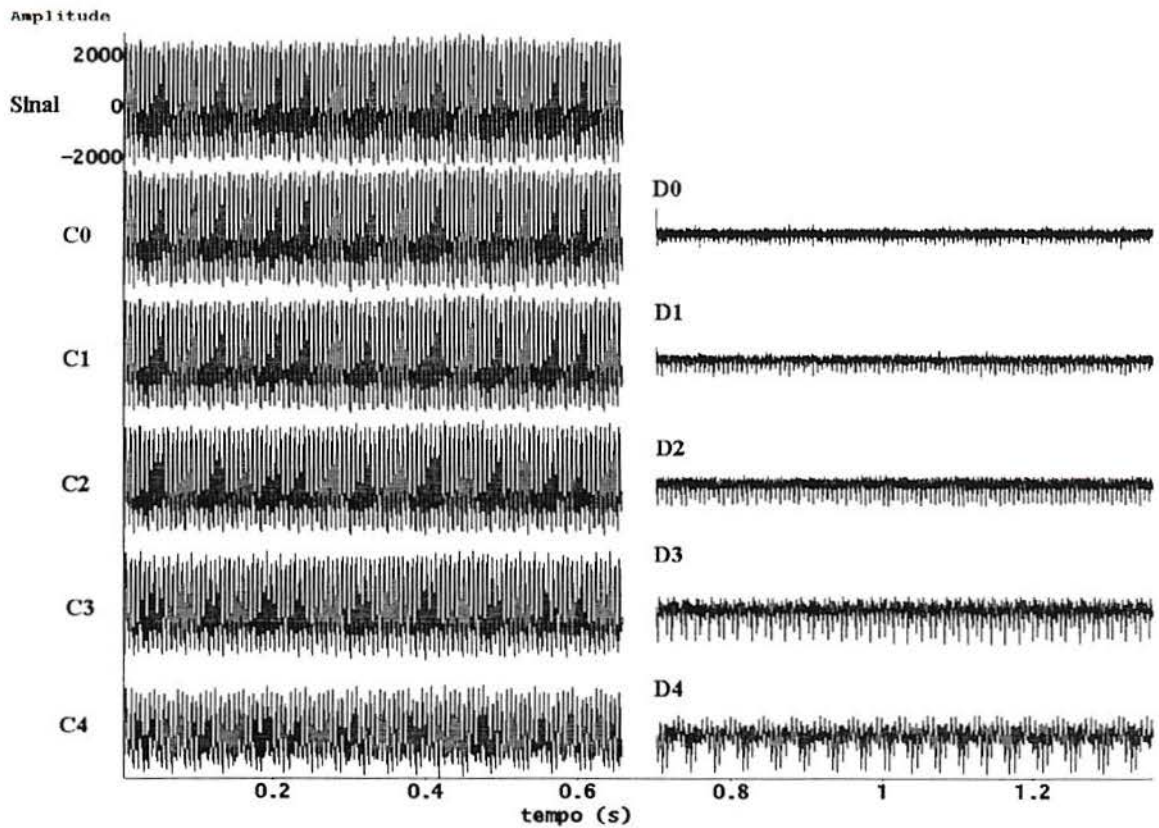


Figura A1.5: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC06.

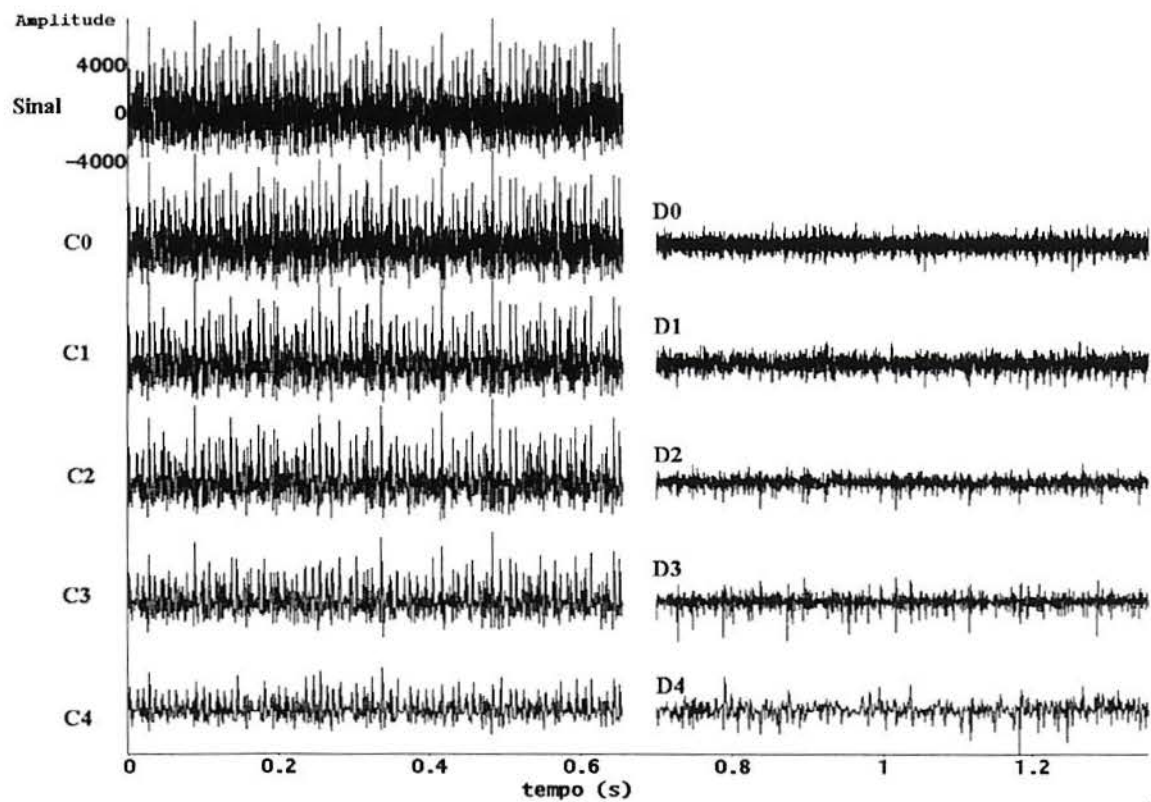


Figura A1.6: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC08.

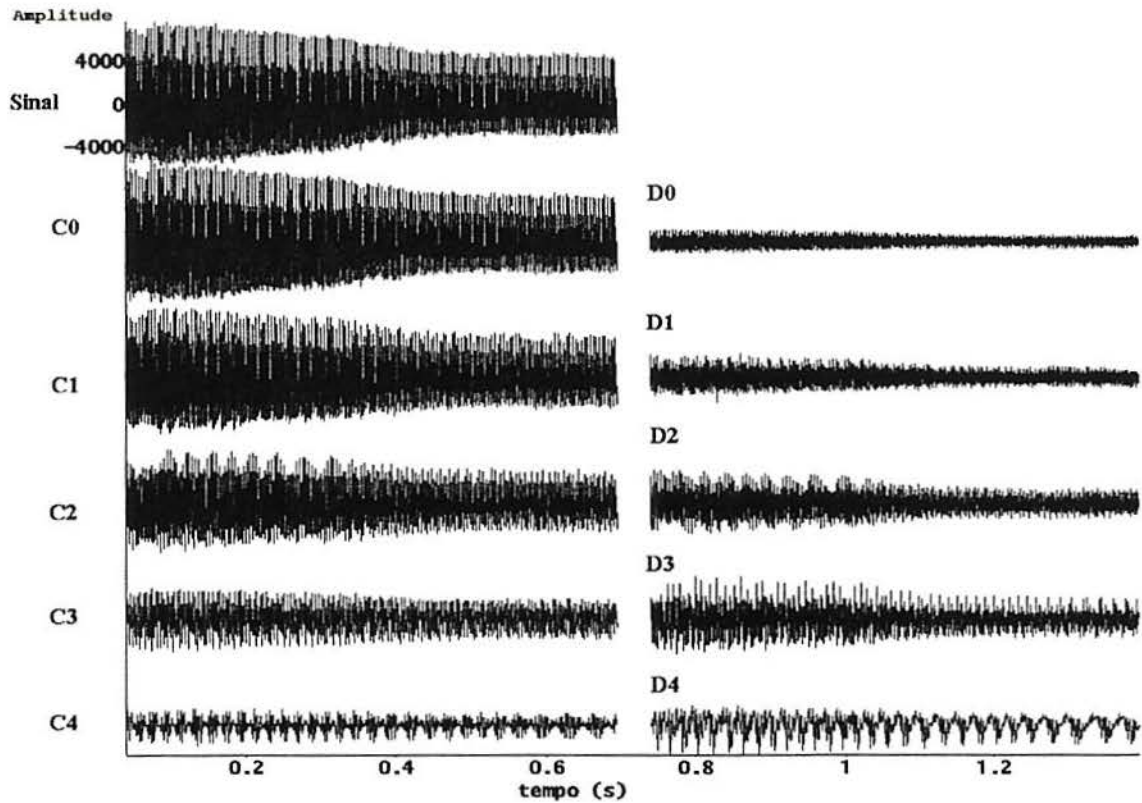


Figura A1.7: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC09.

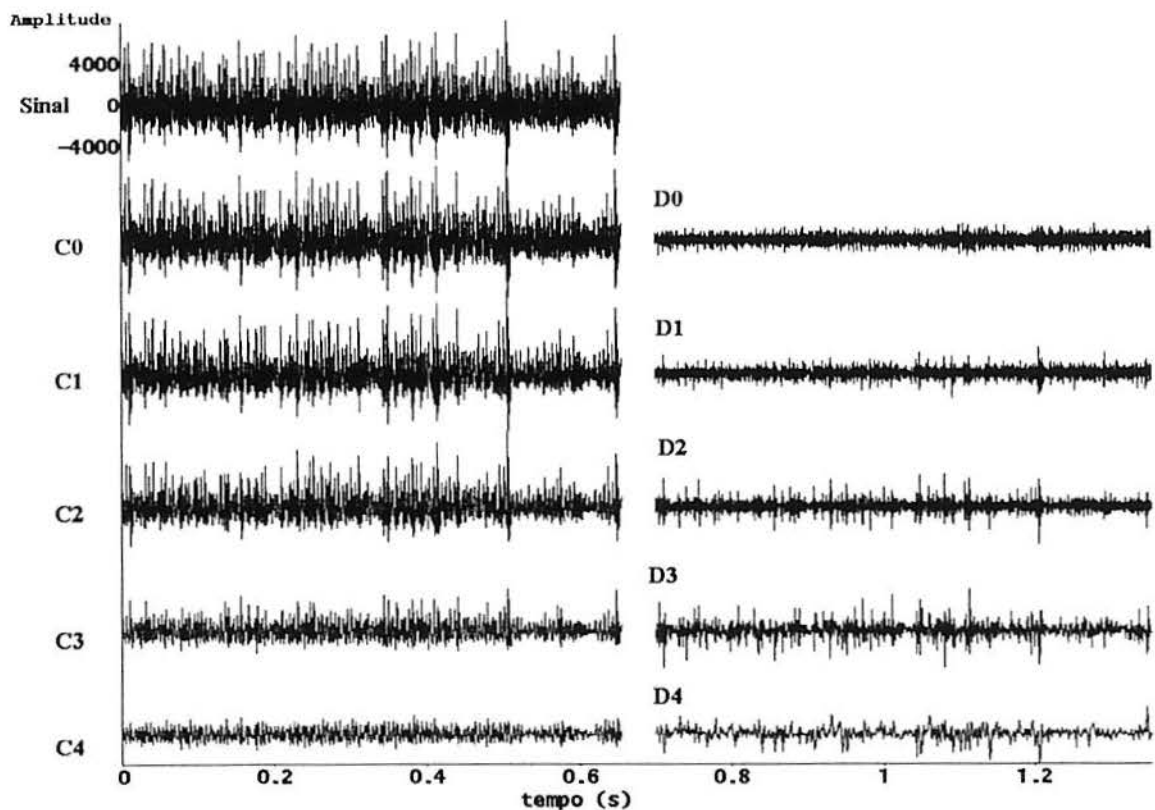


Figura A1.8: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC11.

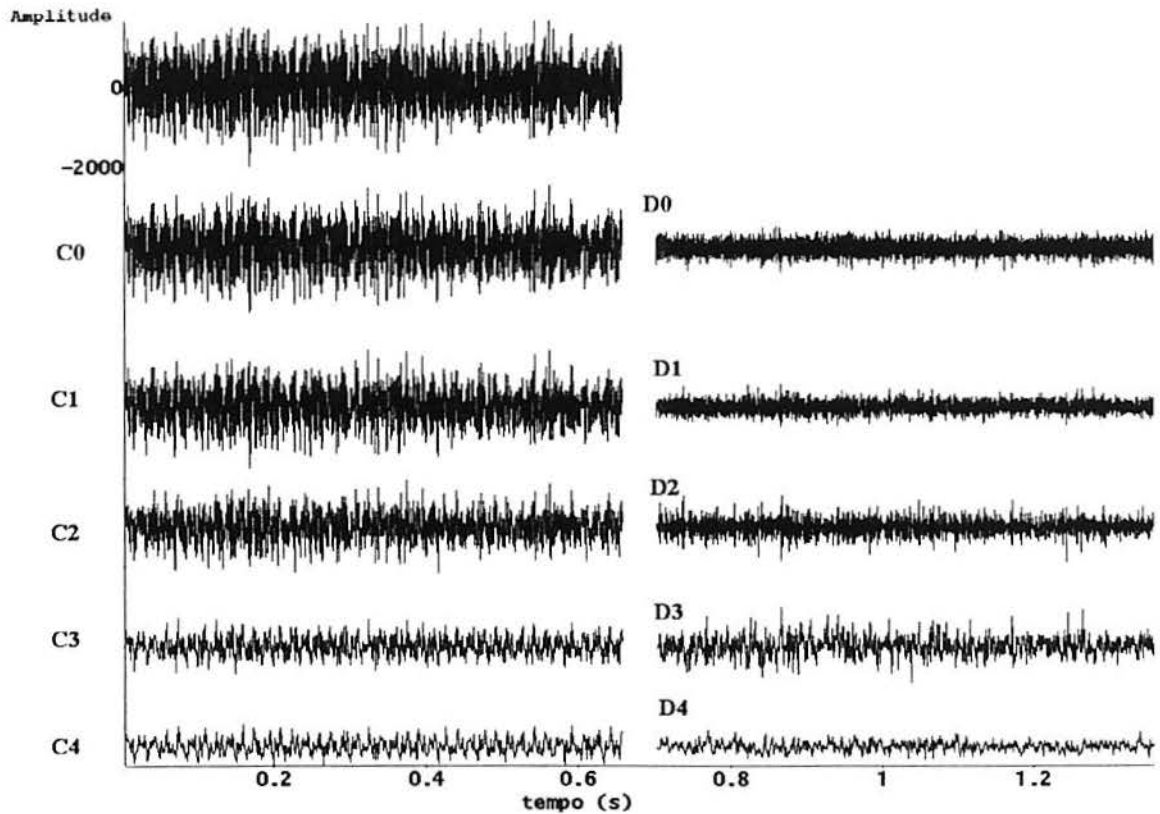


Figura A1.9: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC13.

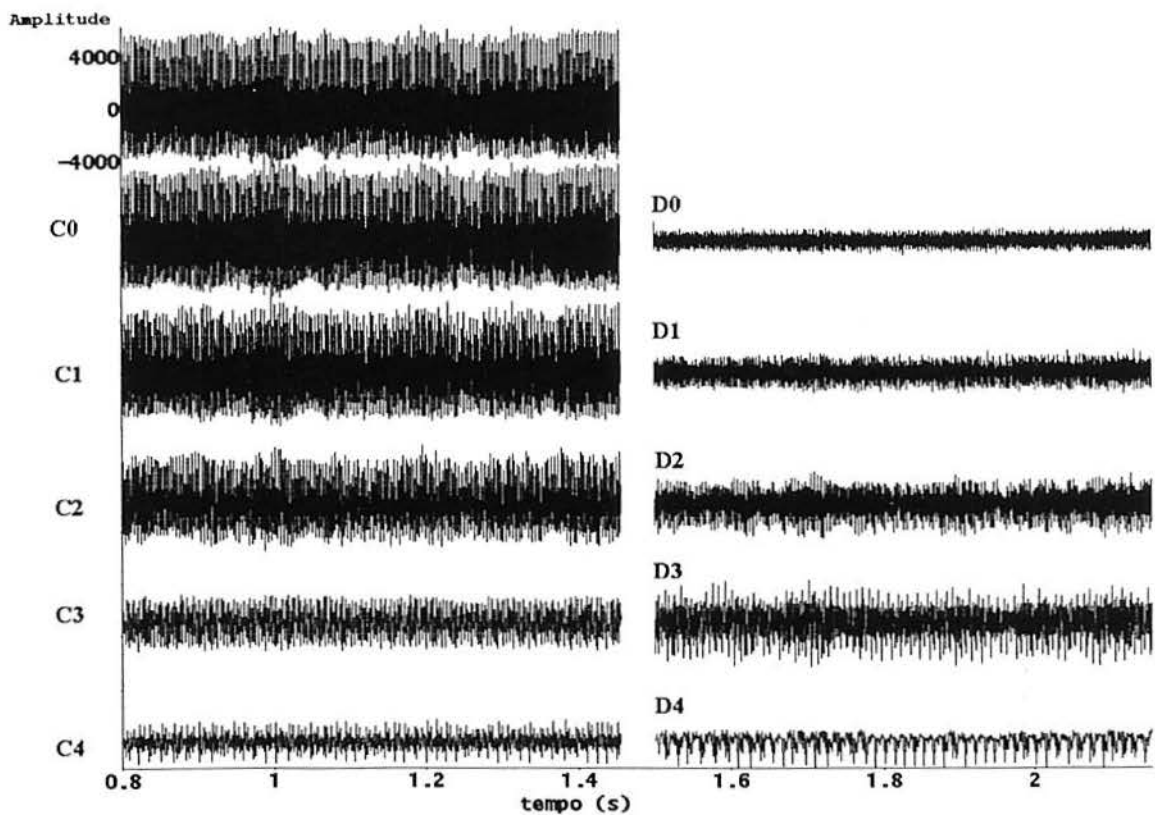


Figura A1.10: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC14.

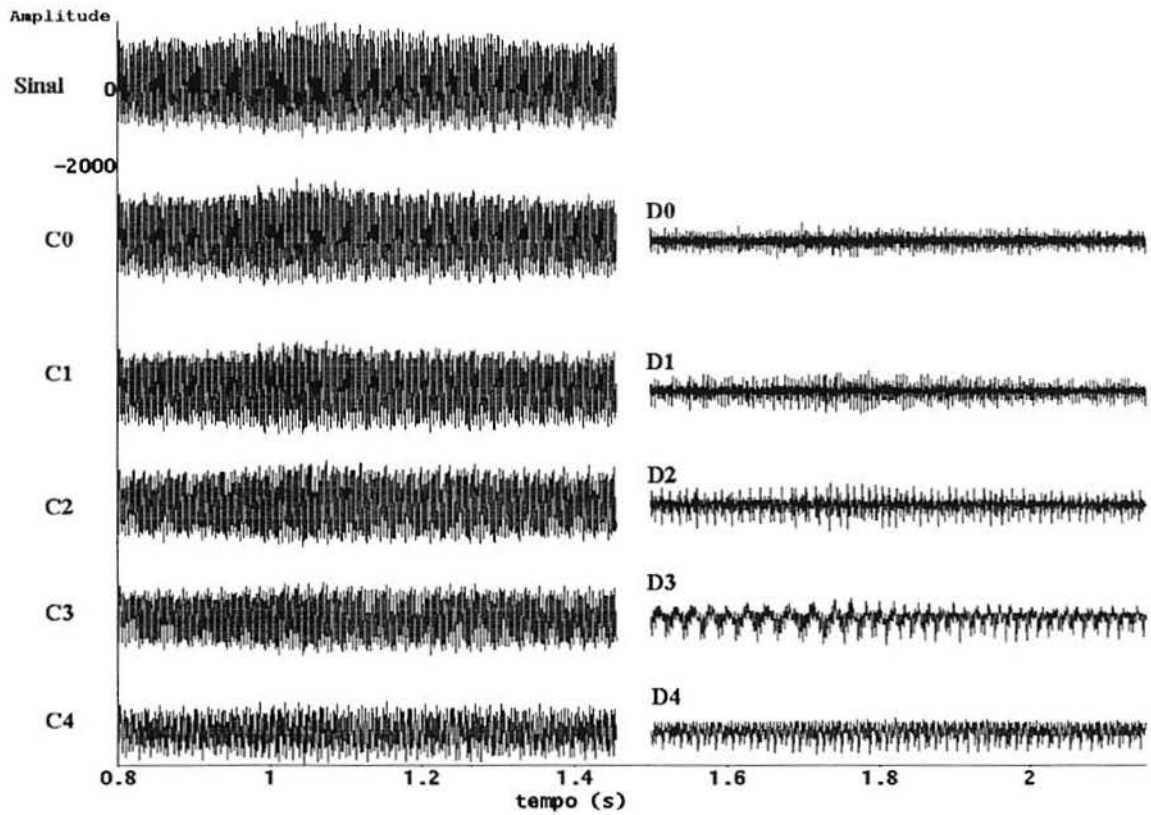


Figura A1.11: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC16.

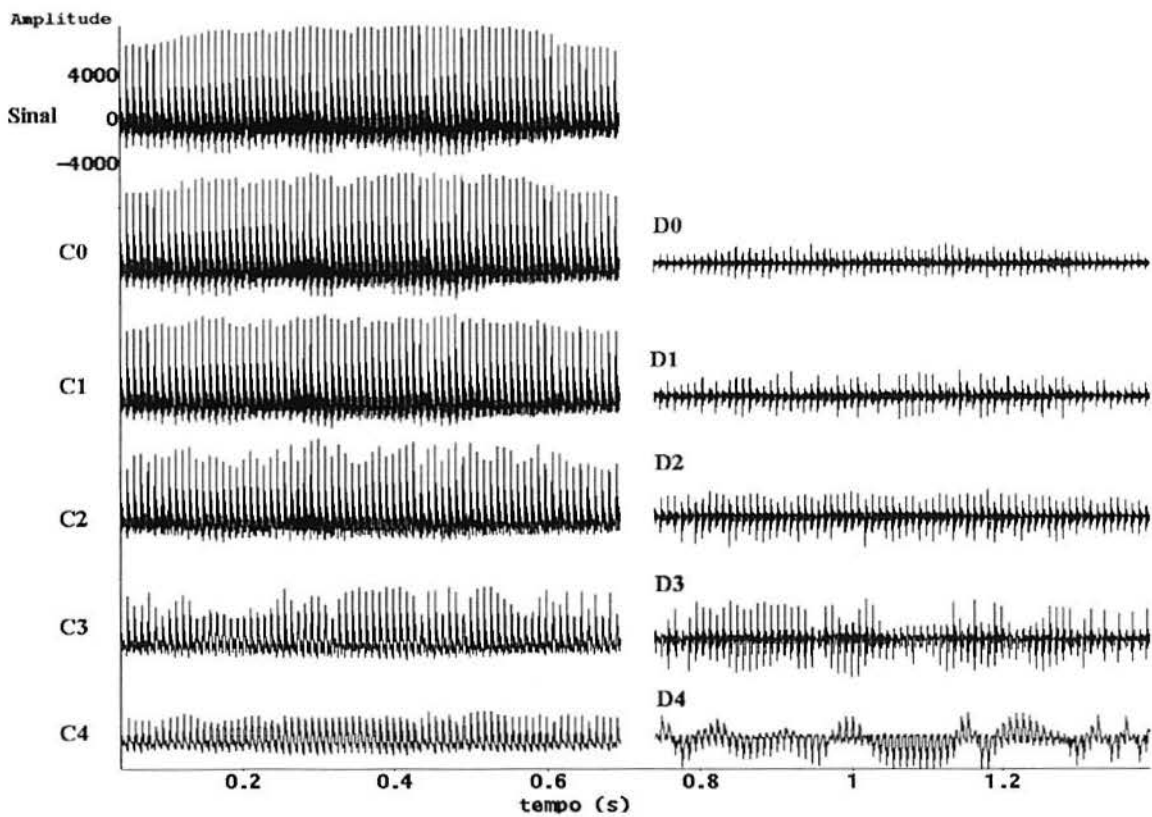


Figura A1.12: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC17.

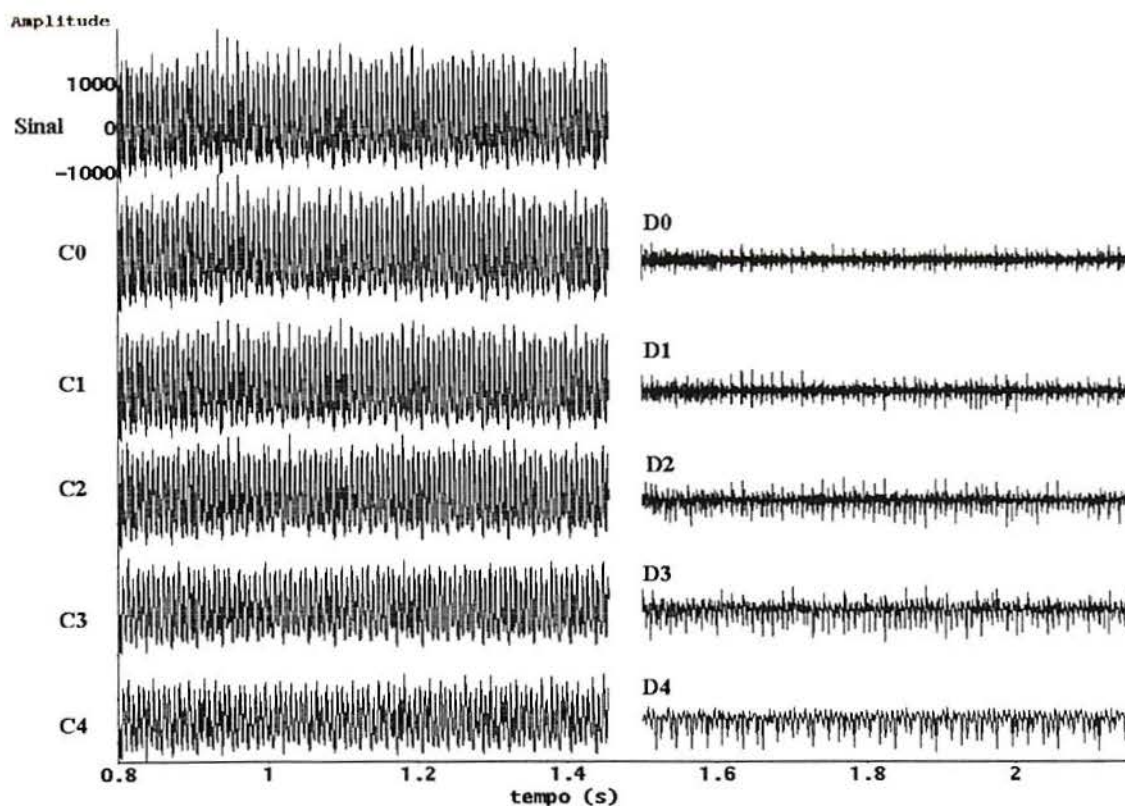


Figura A1.13: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC18.

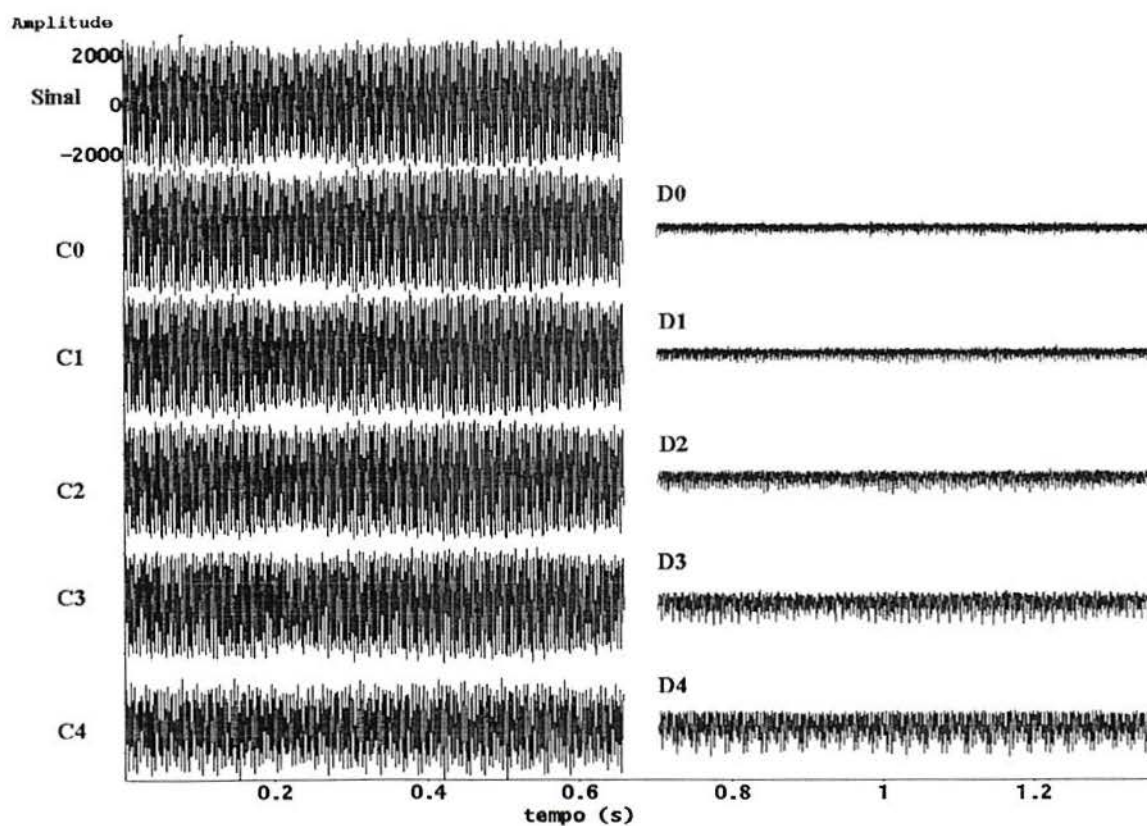


Figura A1.14: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC19.

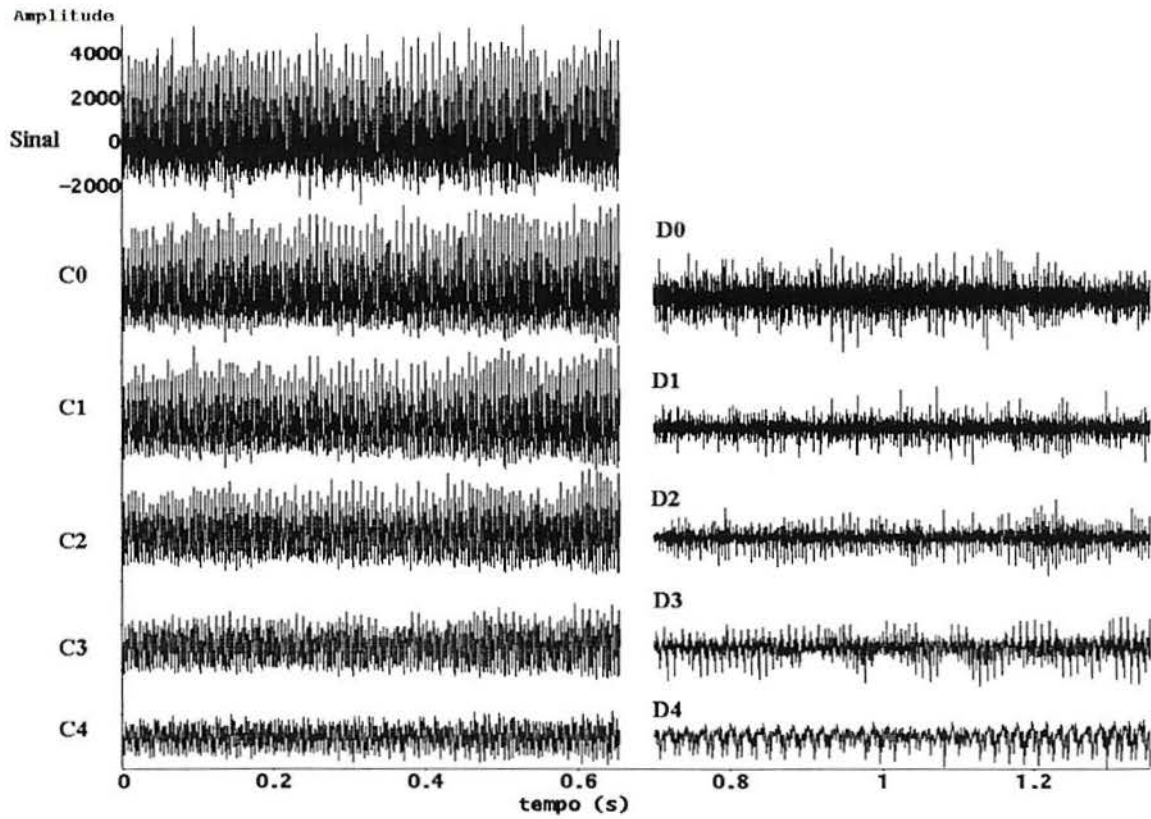


Figura A1.15: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC21.

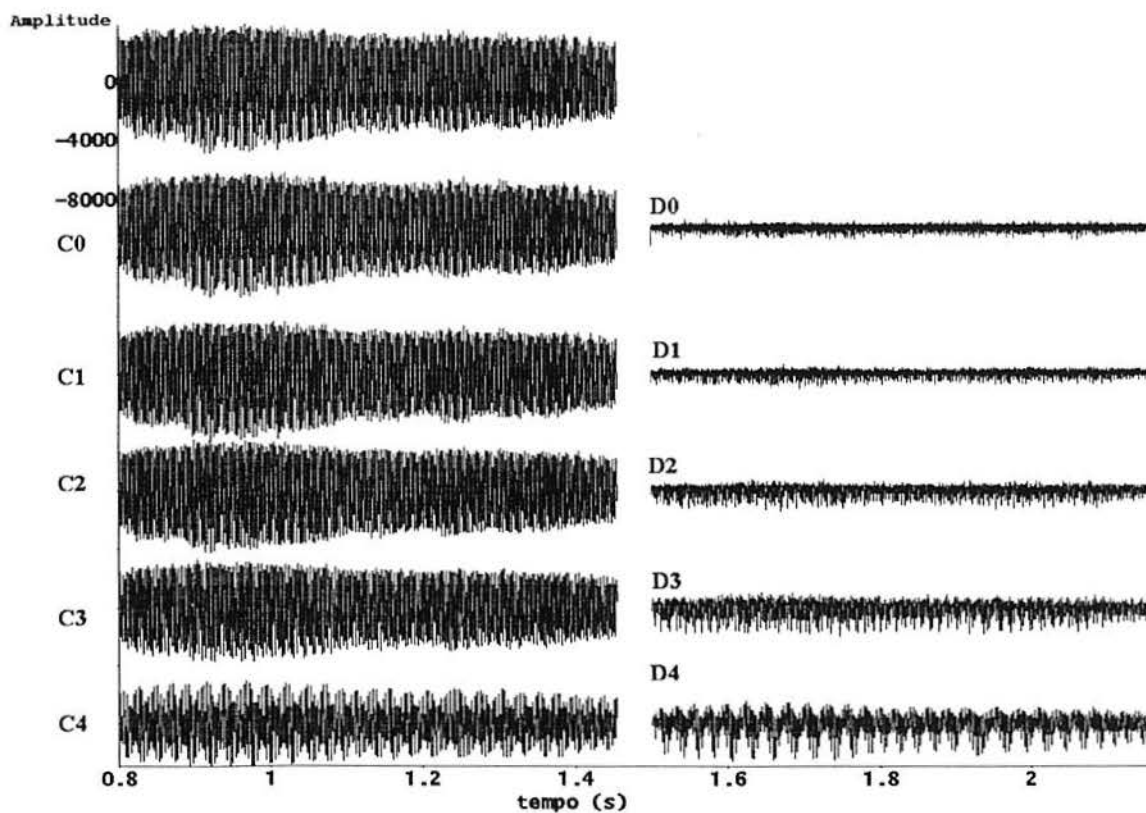


Figura A1.16: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC22.

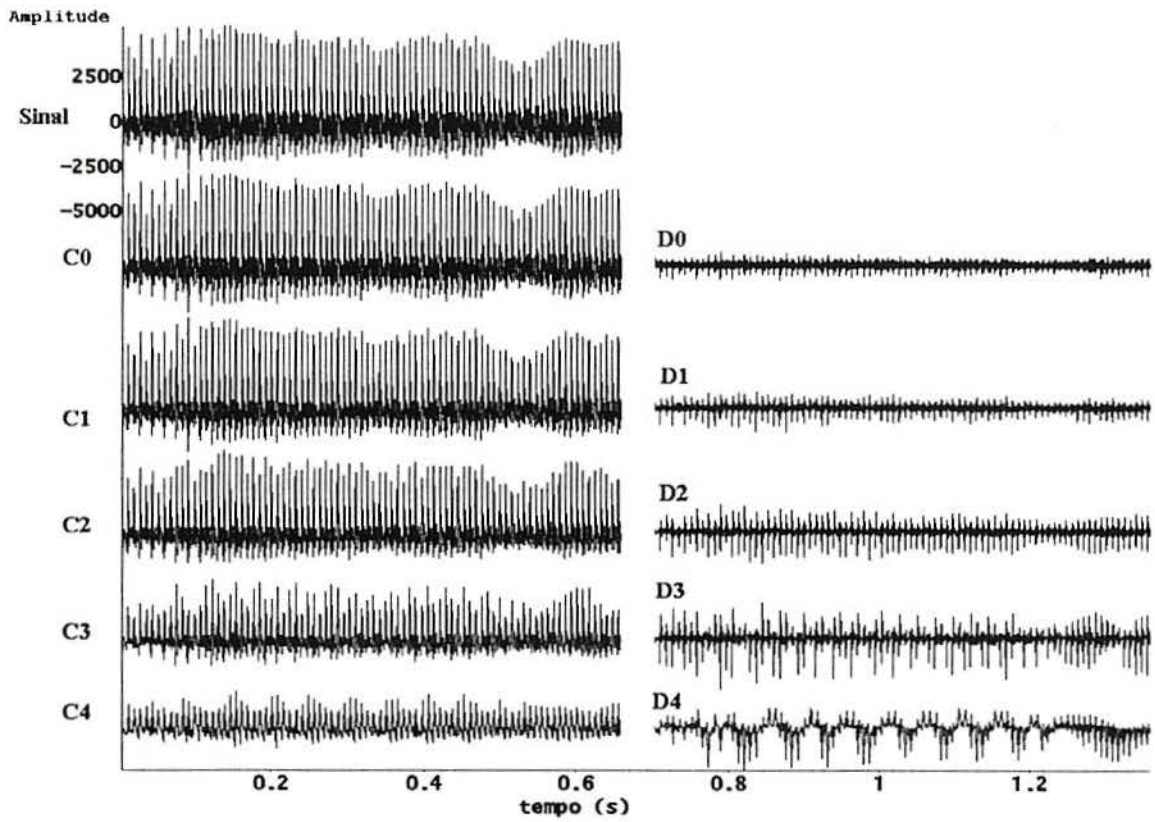


Figura A1.17: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC23.

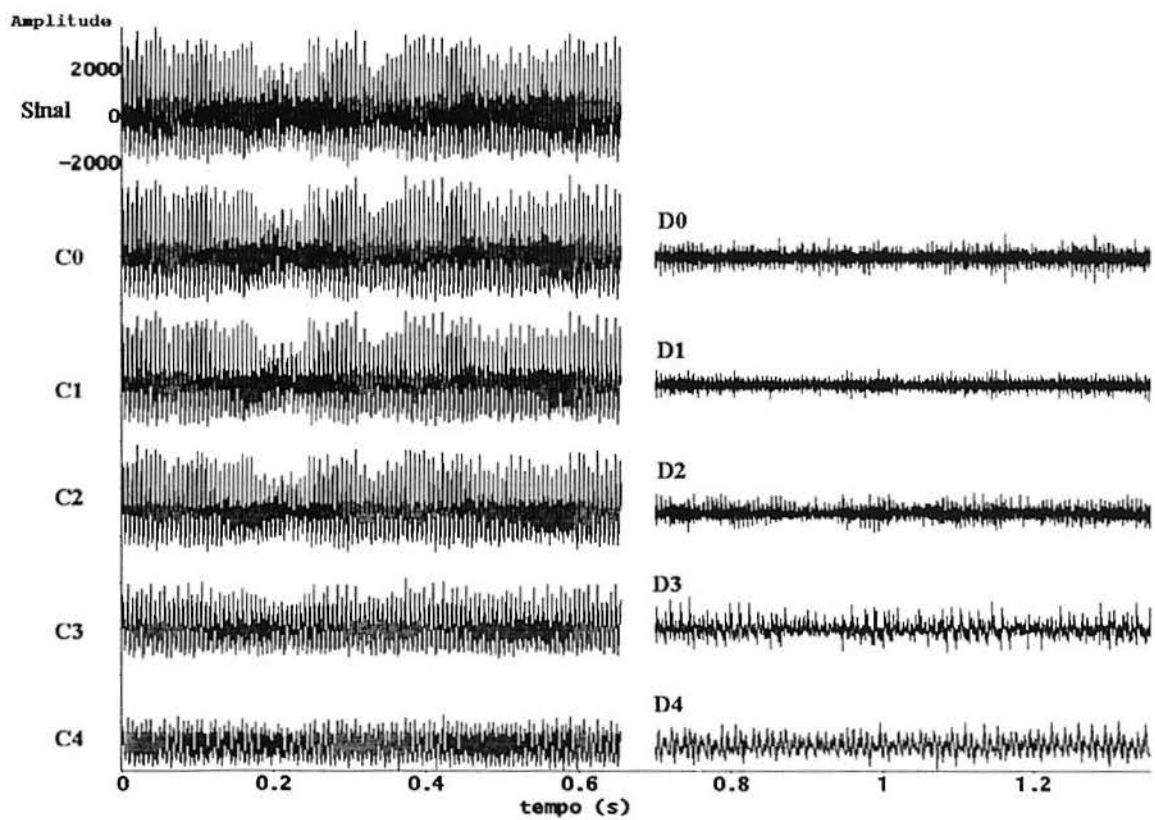


Figura A1.18: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC24.

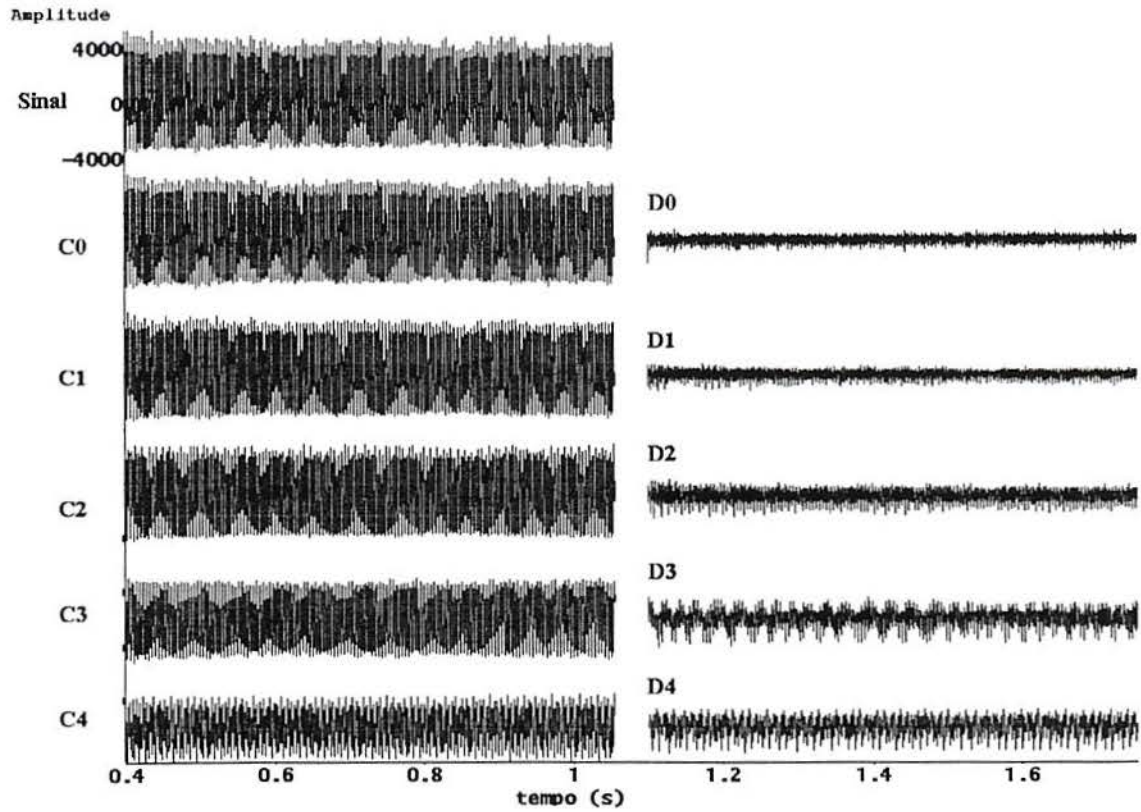


Figura A1.19: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC27.

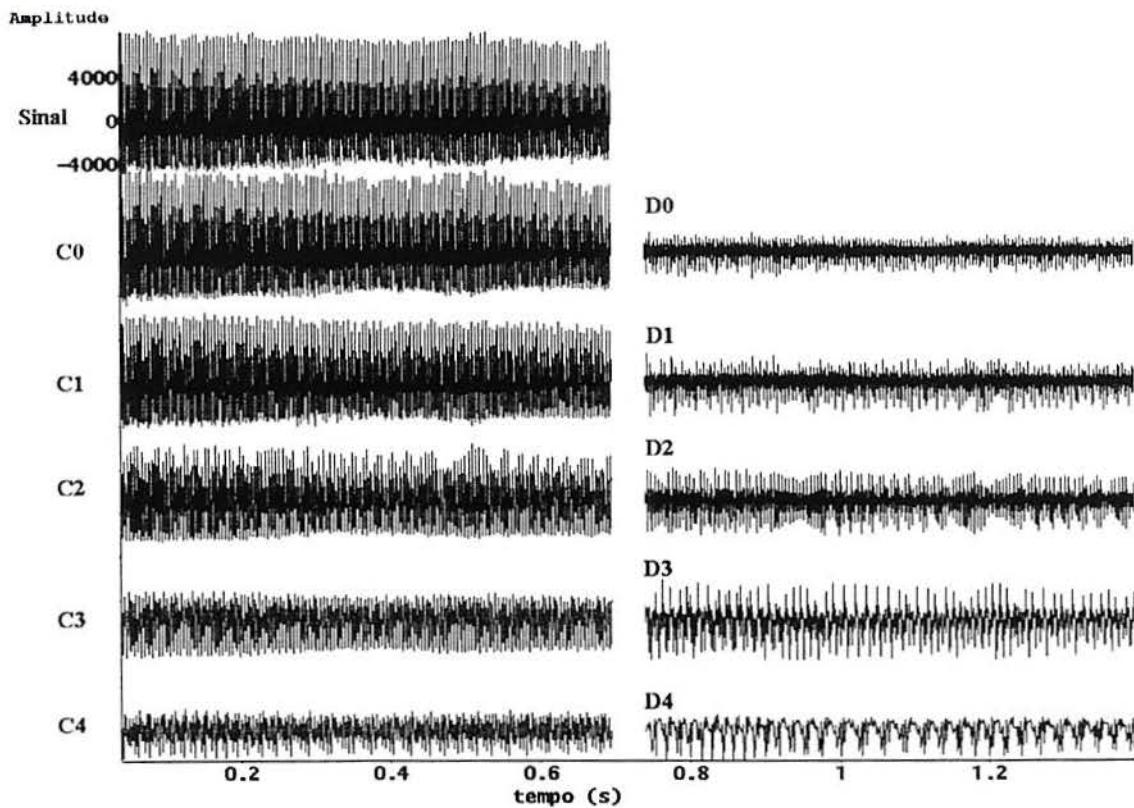


Figura A1.20: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC28.



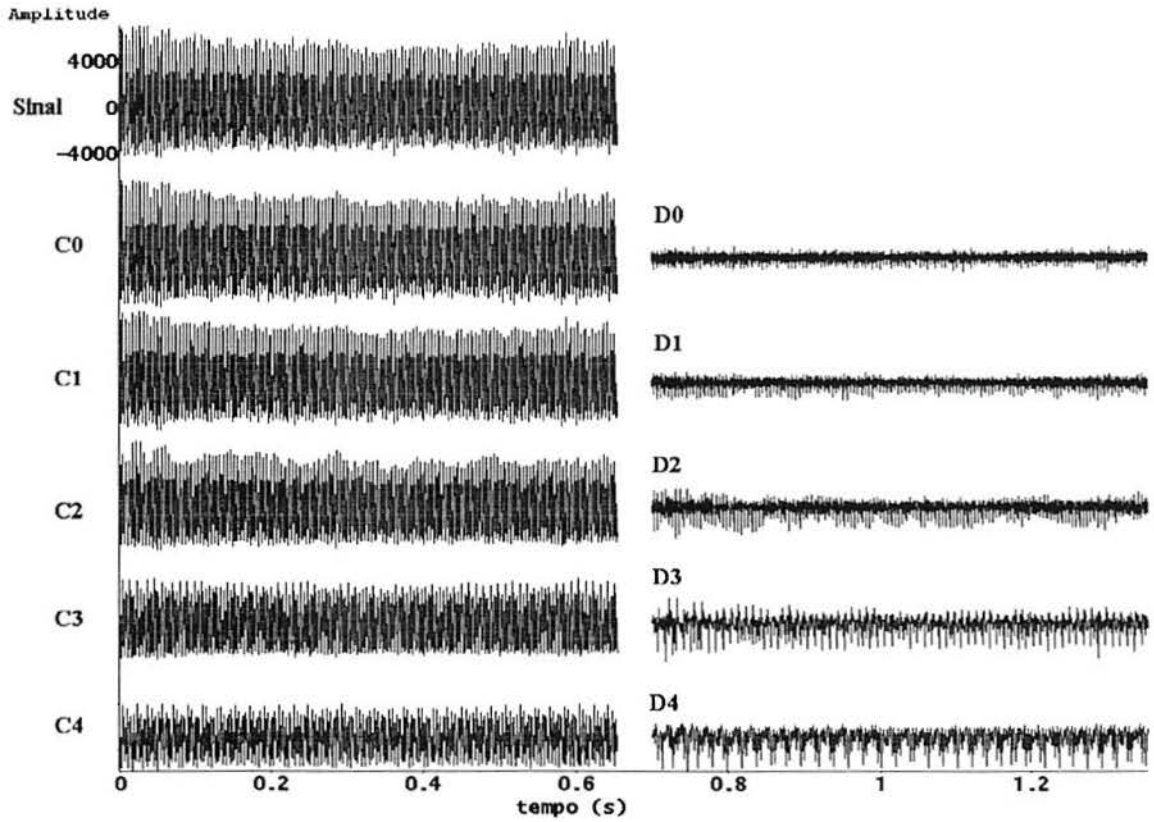


Figura A1.21: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC30.

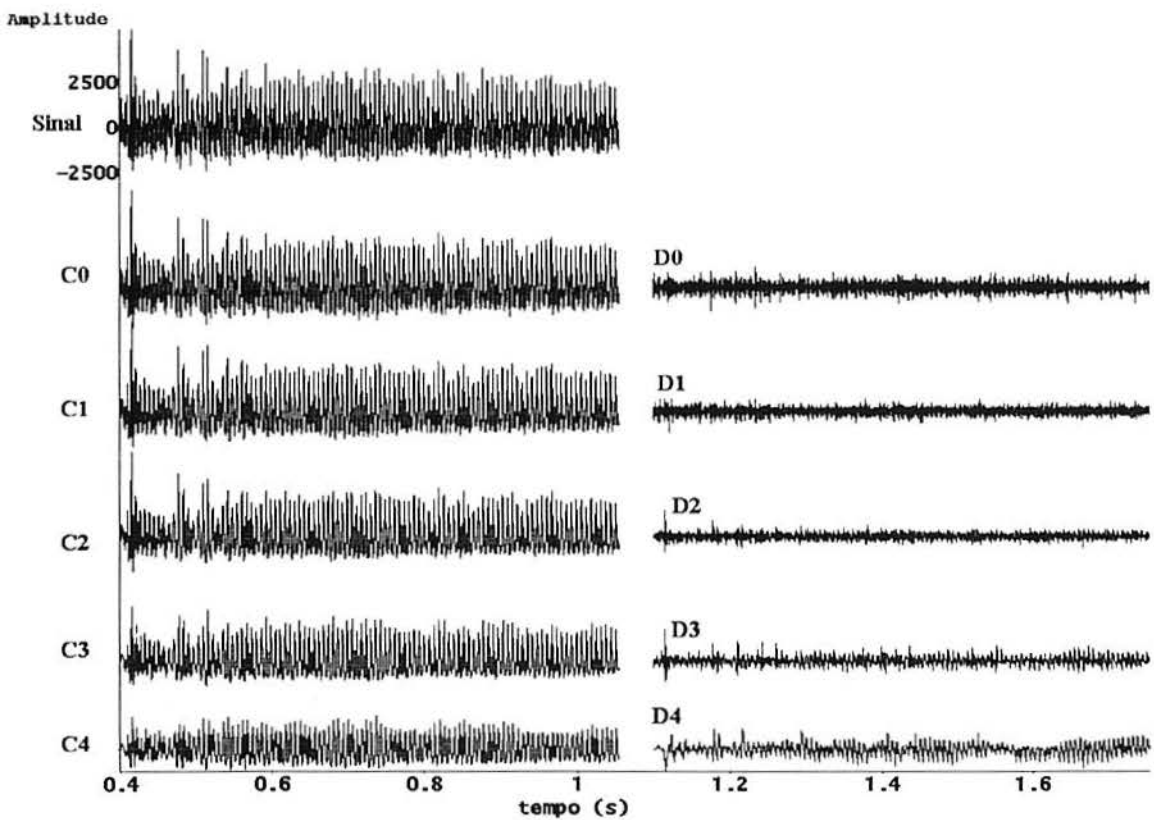


Figura A1.22: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC31.

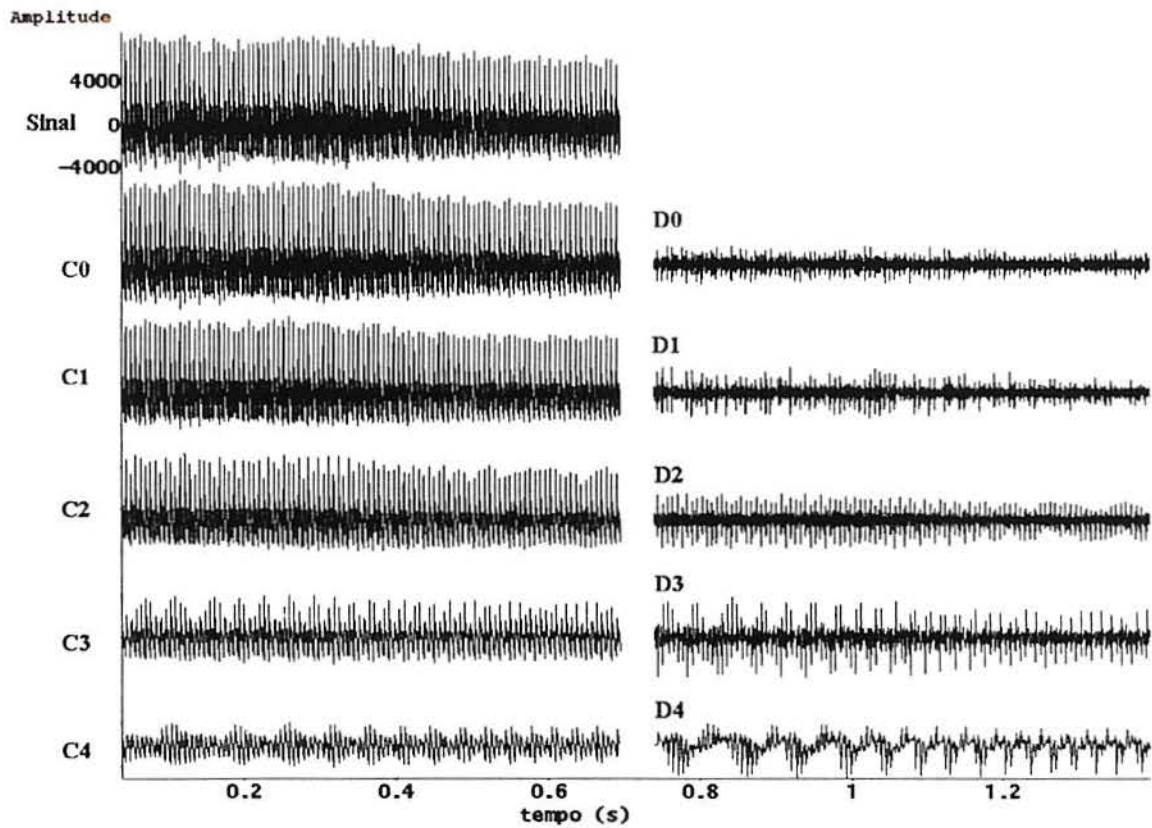


Figura A1.23: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC32.

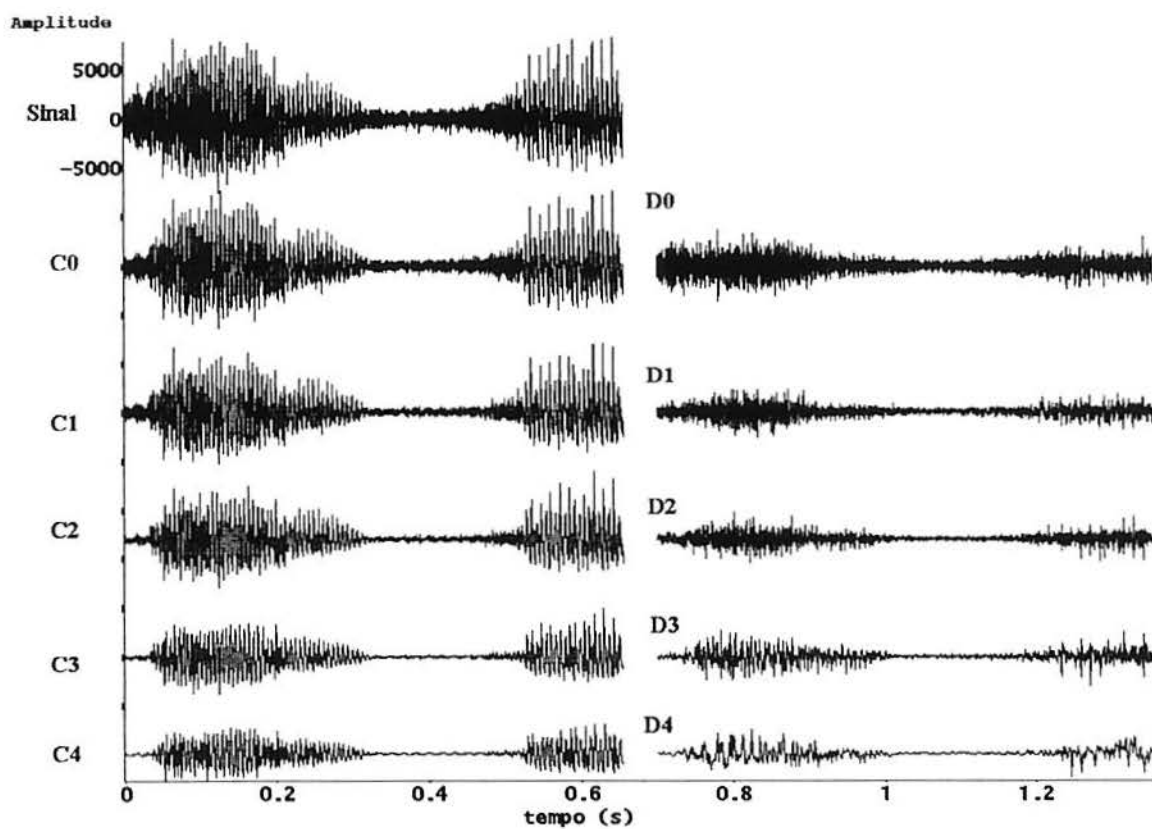


Figura A1.24: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC34.

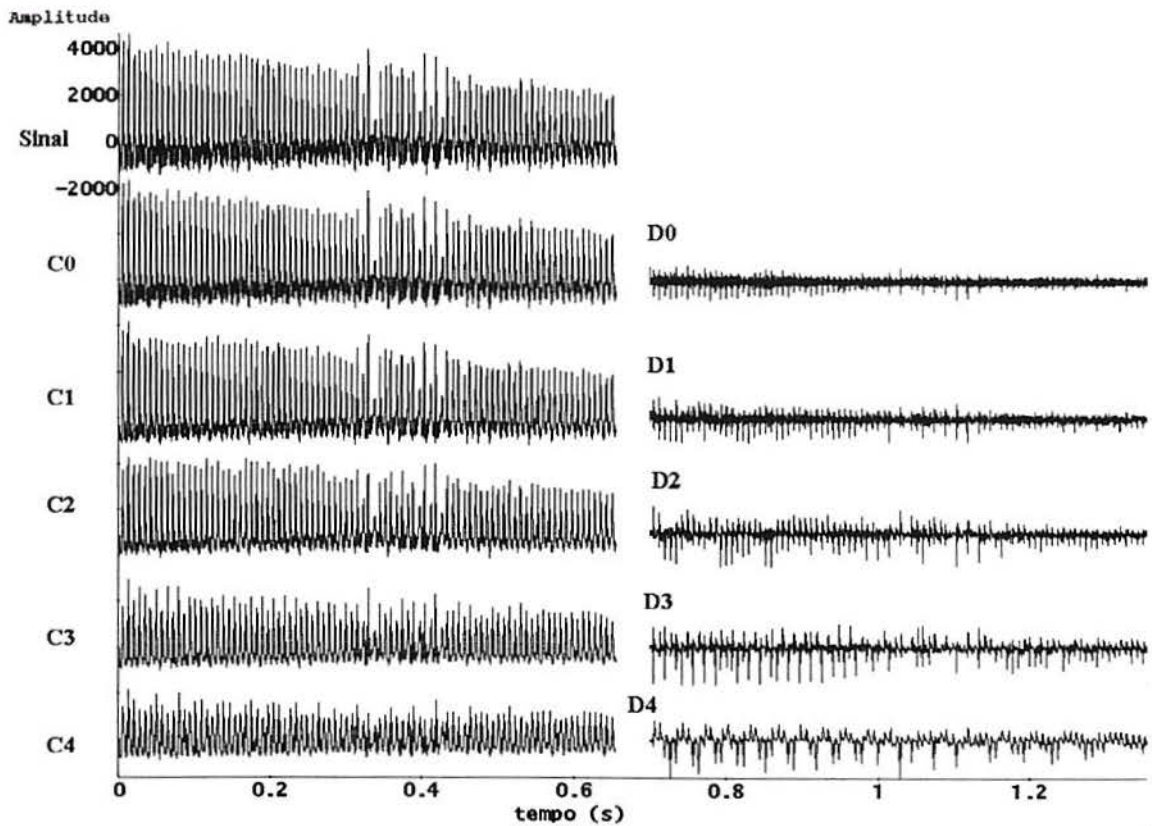


Figura A1.25: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC36.

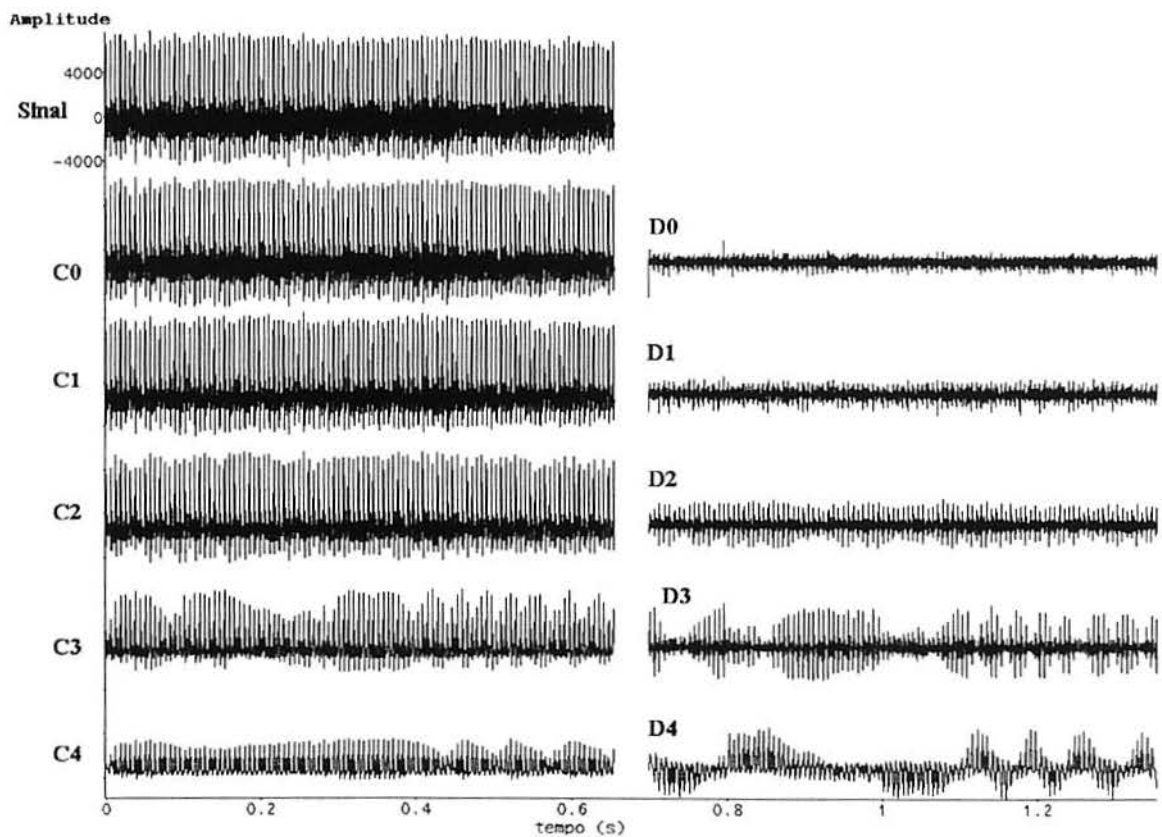


Figura A1.26: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC41.

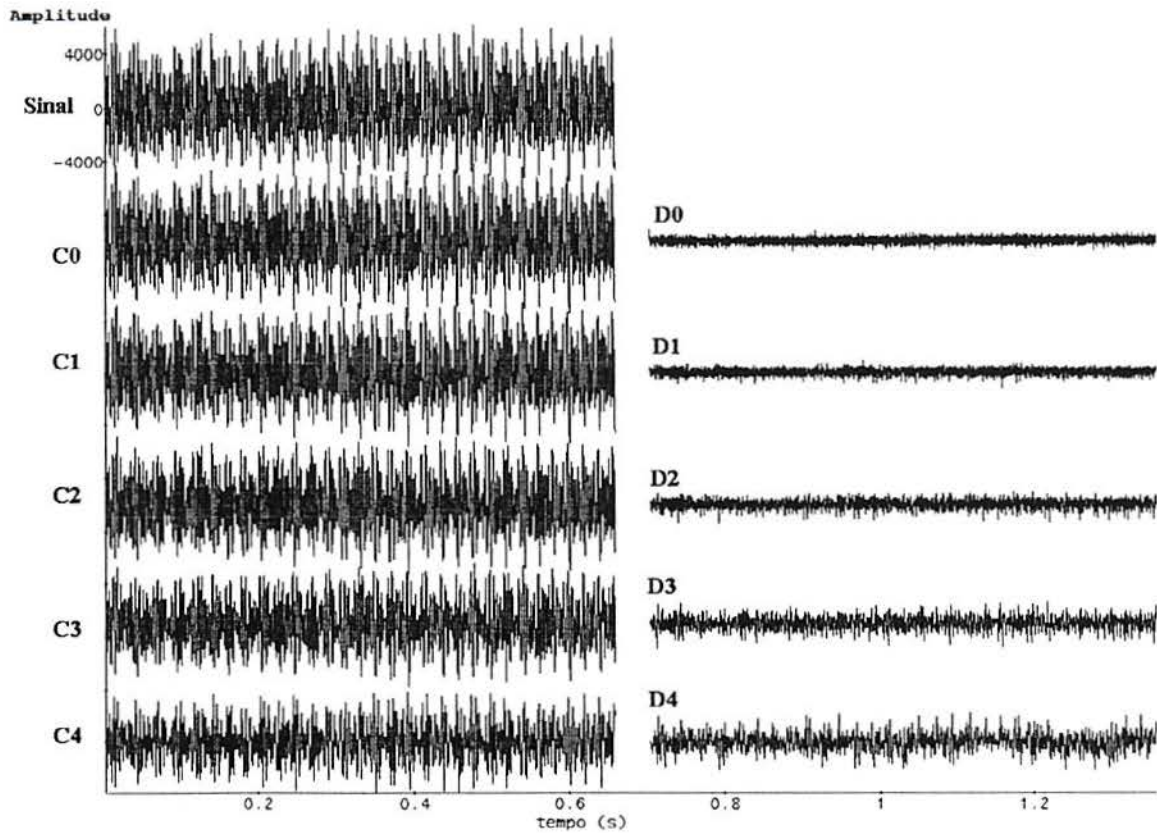


Figura A1.27: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC42.

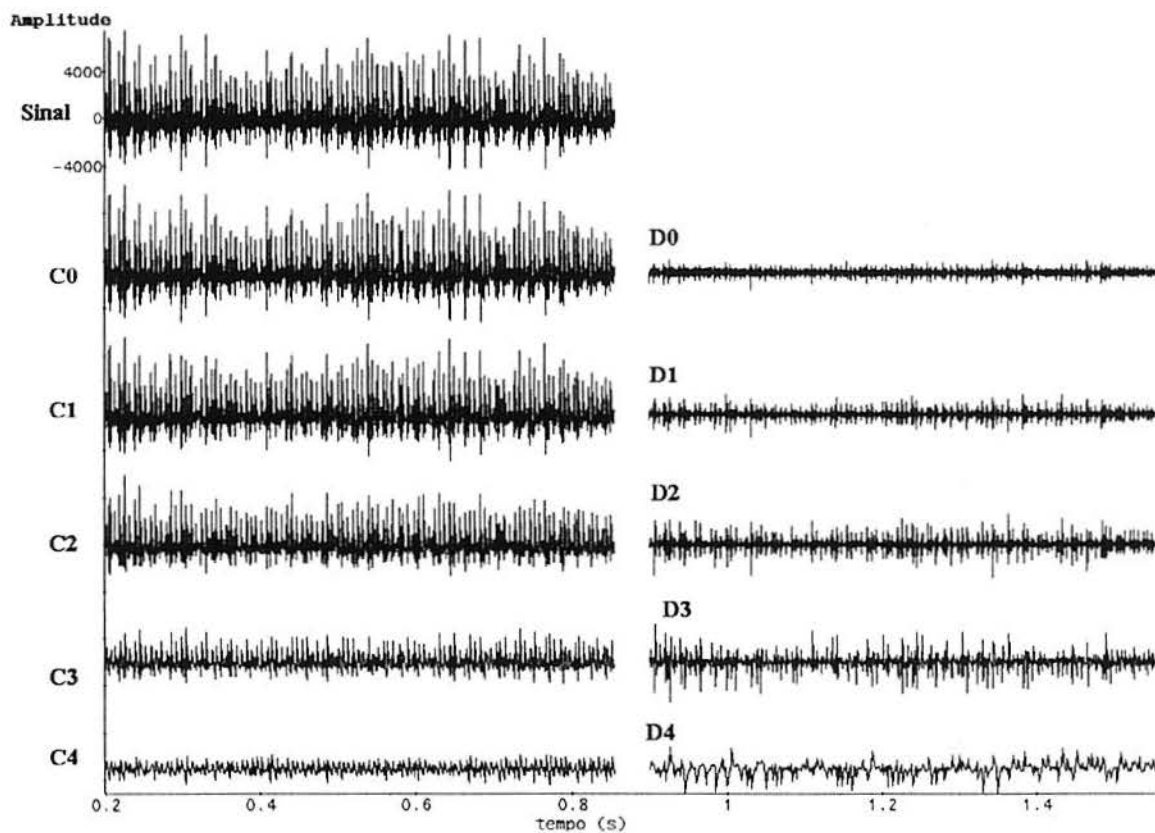


Figura A1.28: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC43.

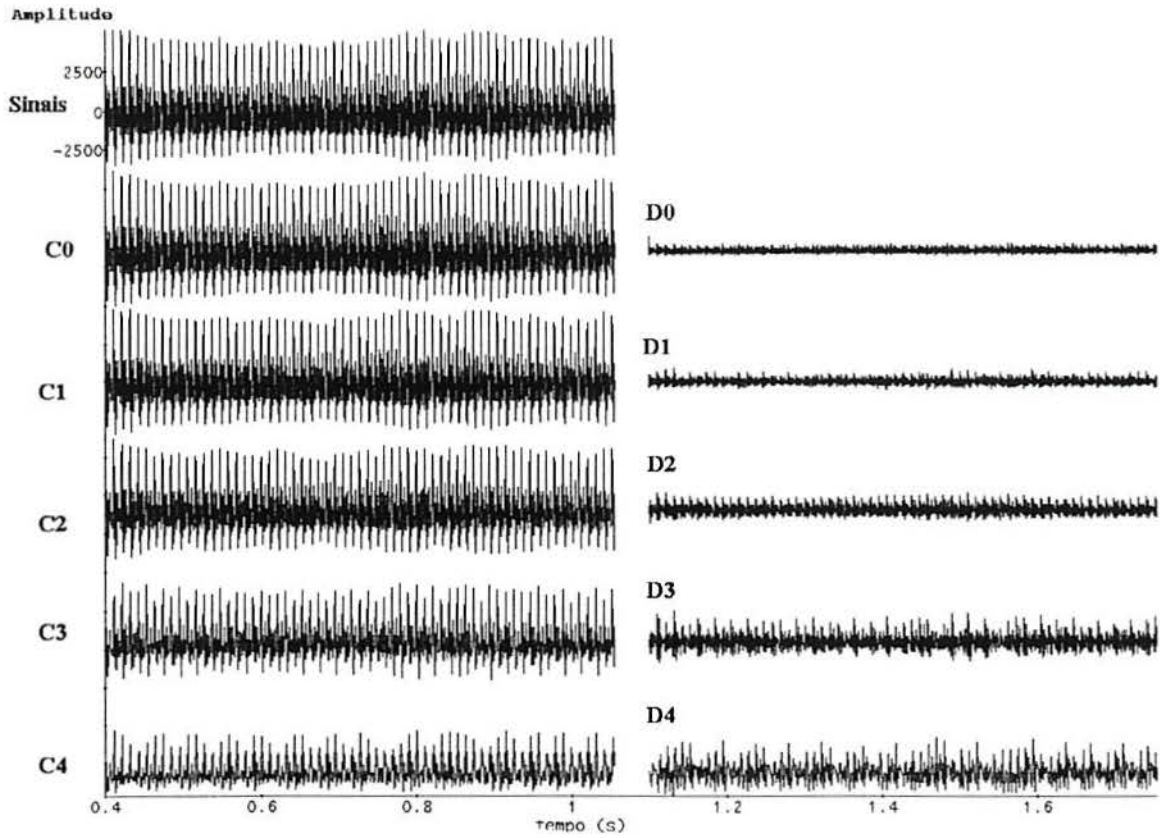


Figura A1.29: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC44.

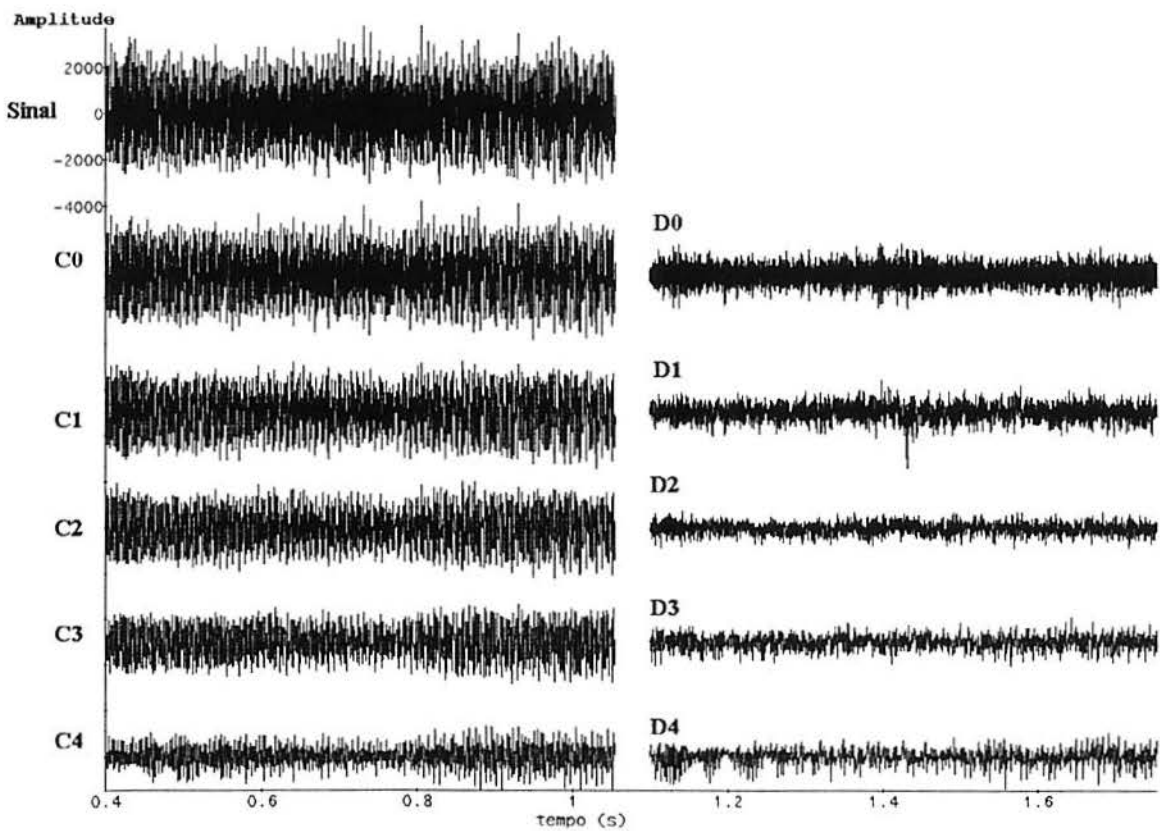


Figura A1.30: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC46.

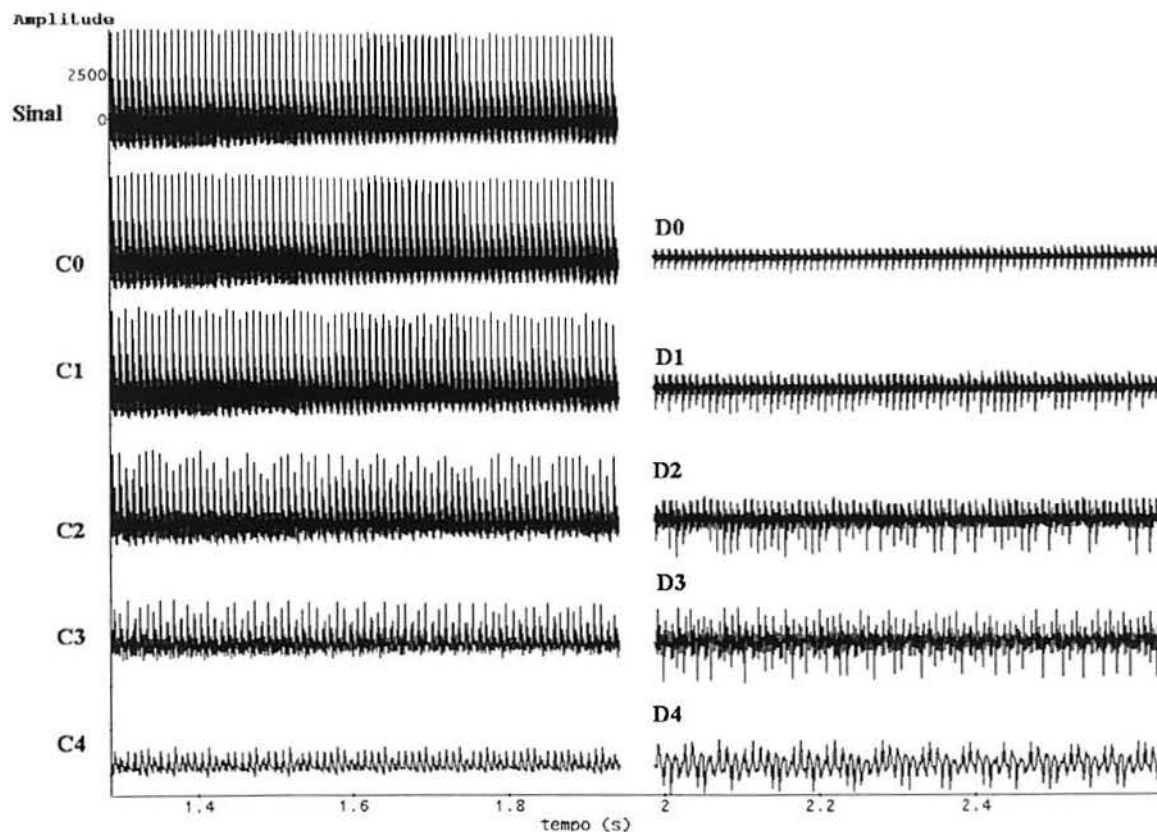


Figura A1.31: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC47.

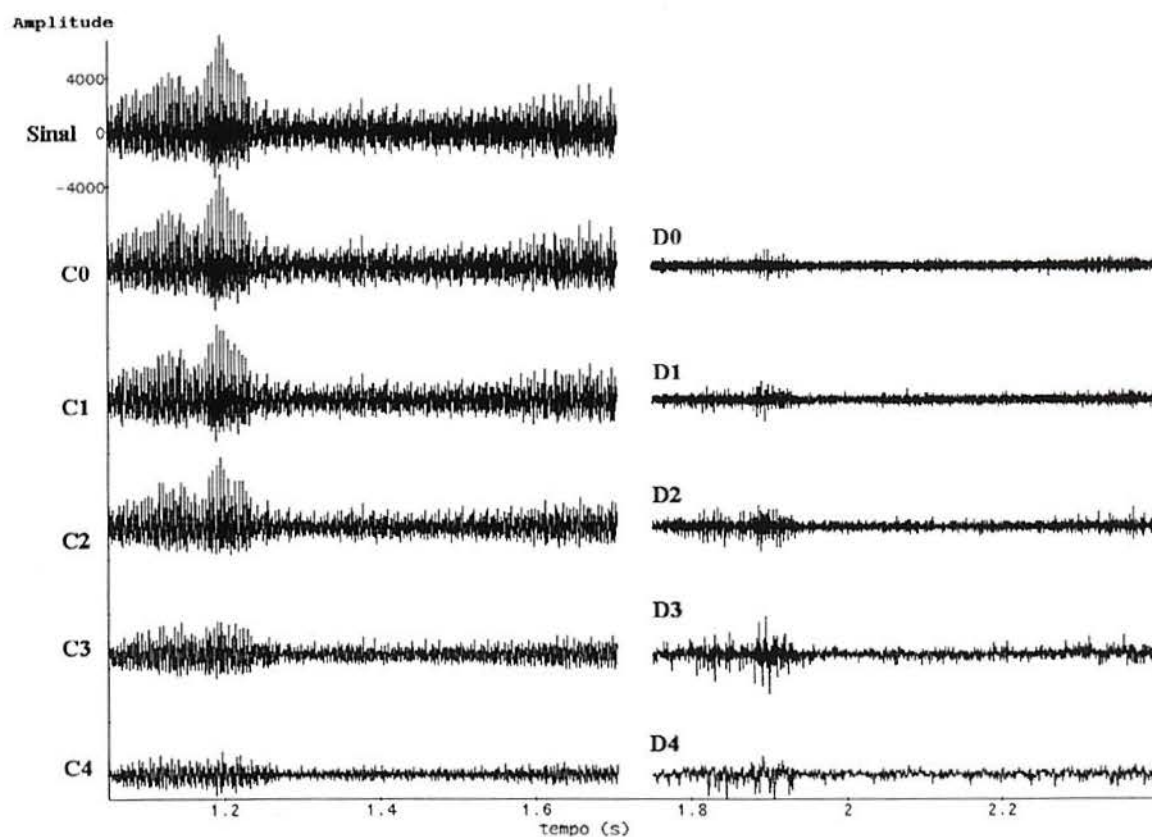


Figura A1.32: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC50.

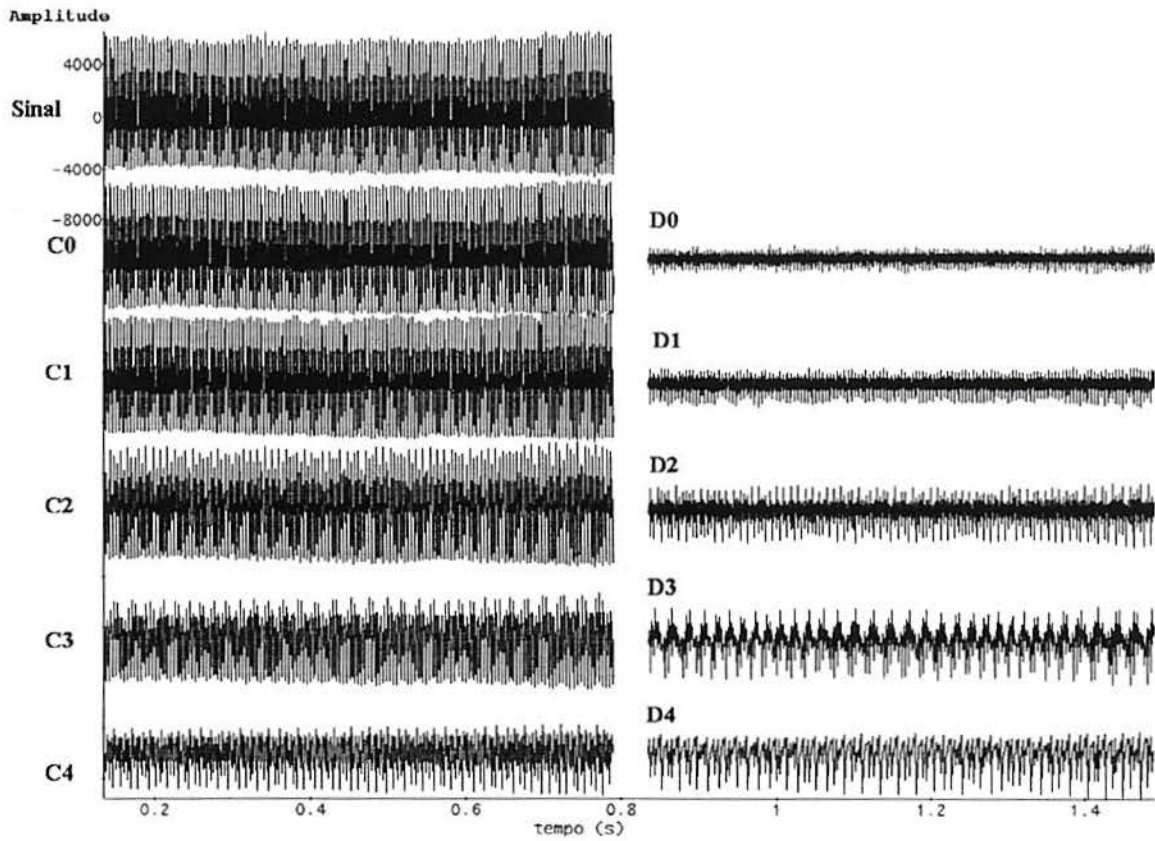


Figura A1.33: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC51.

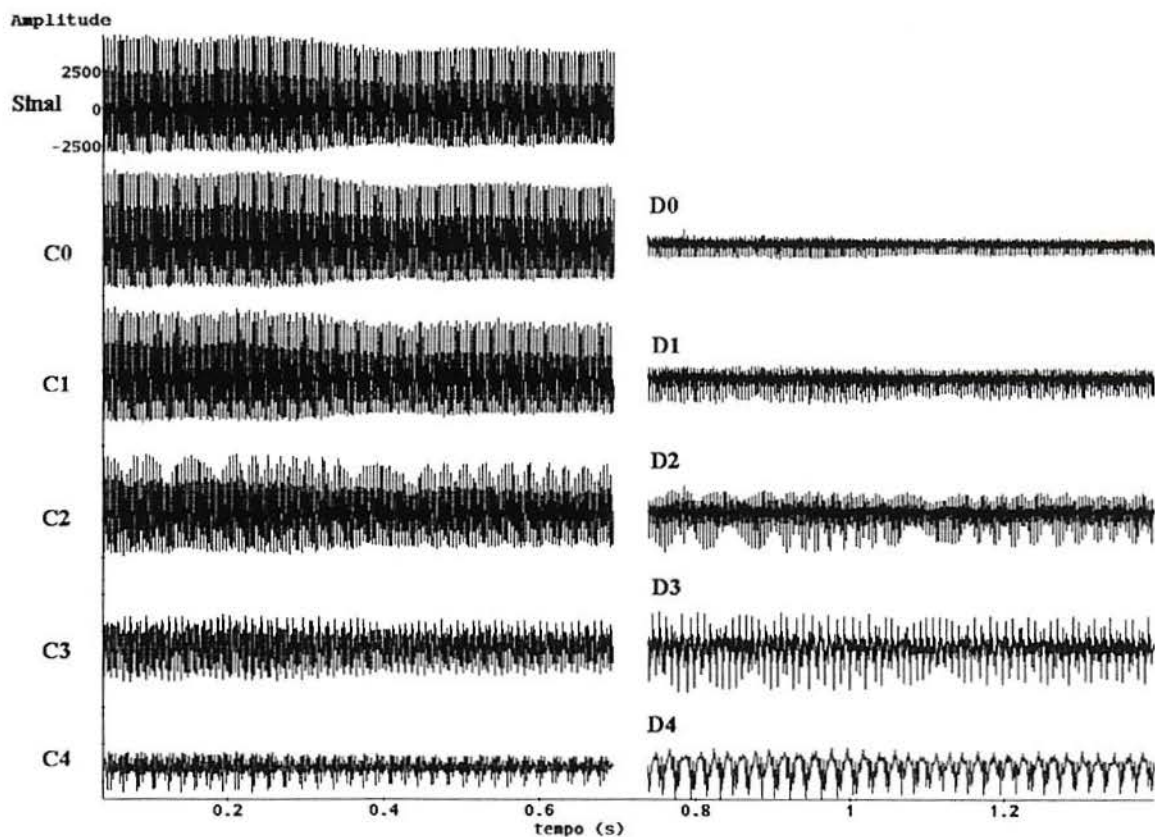


Figura A1.34: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC53.

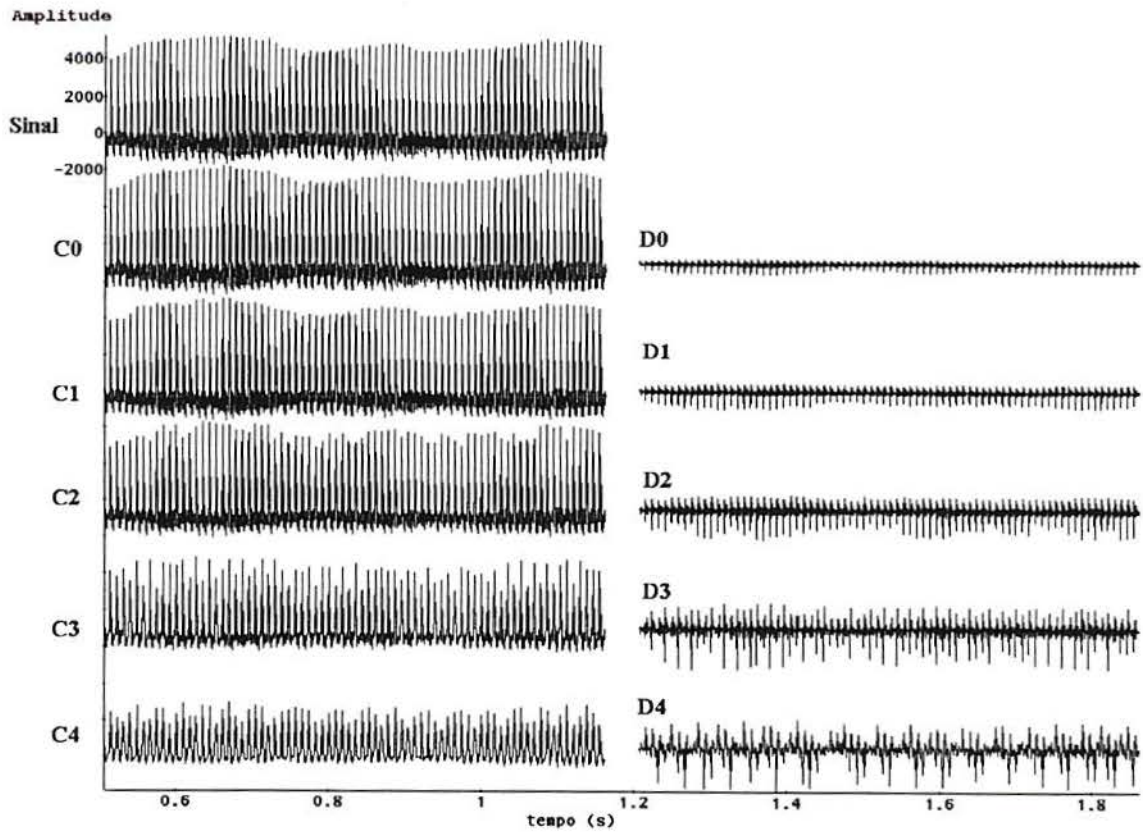


Figura A1.35: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC54.

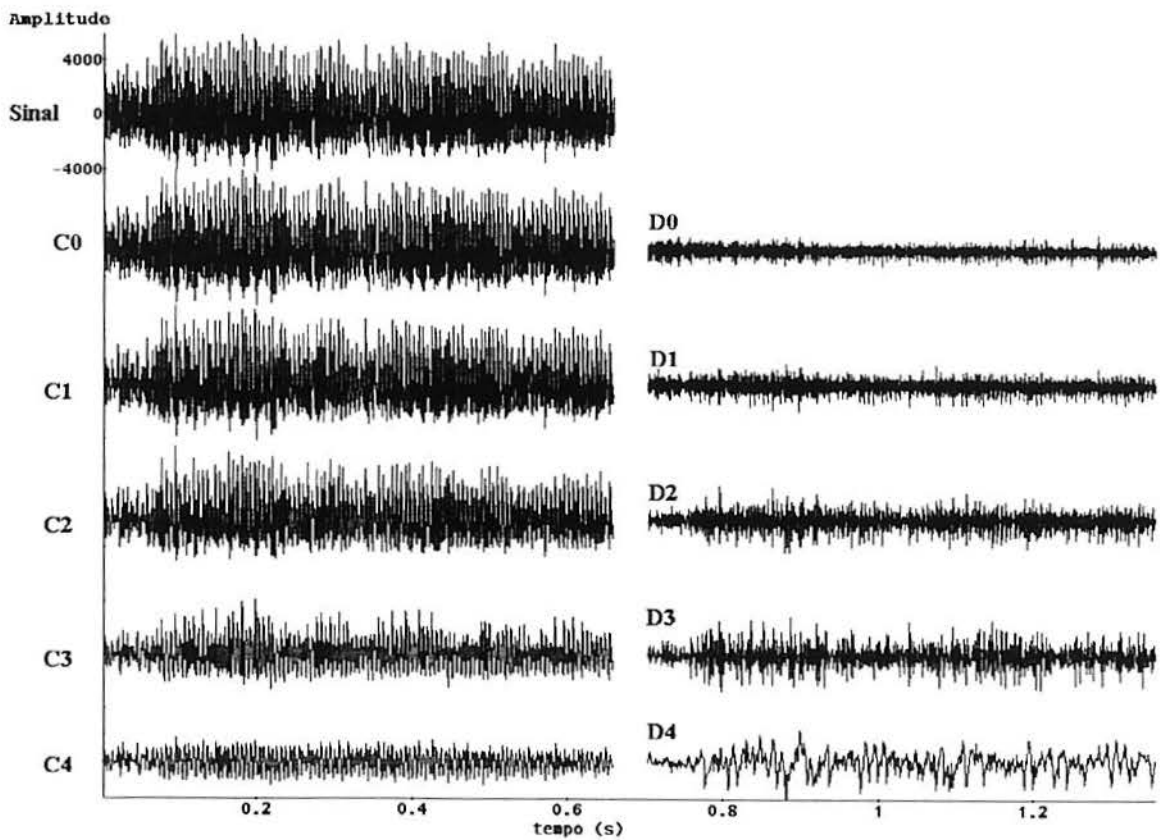


Figura A1.36: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC57.



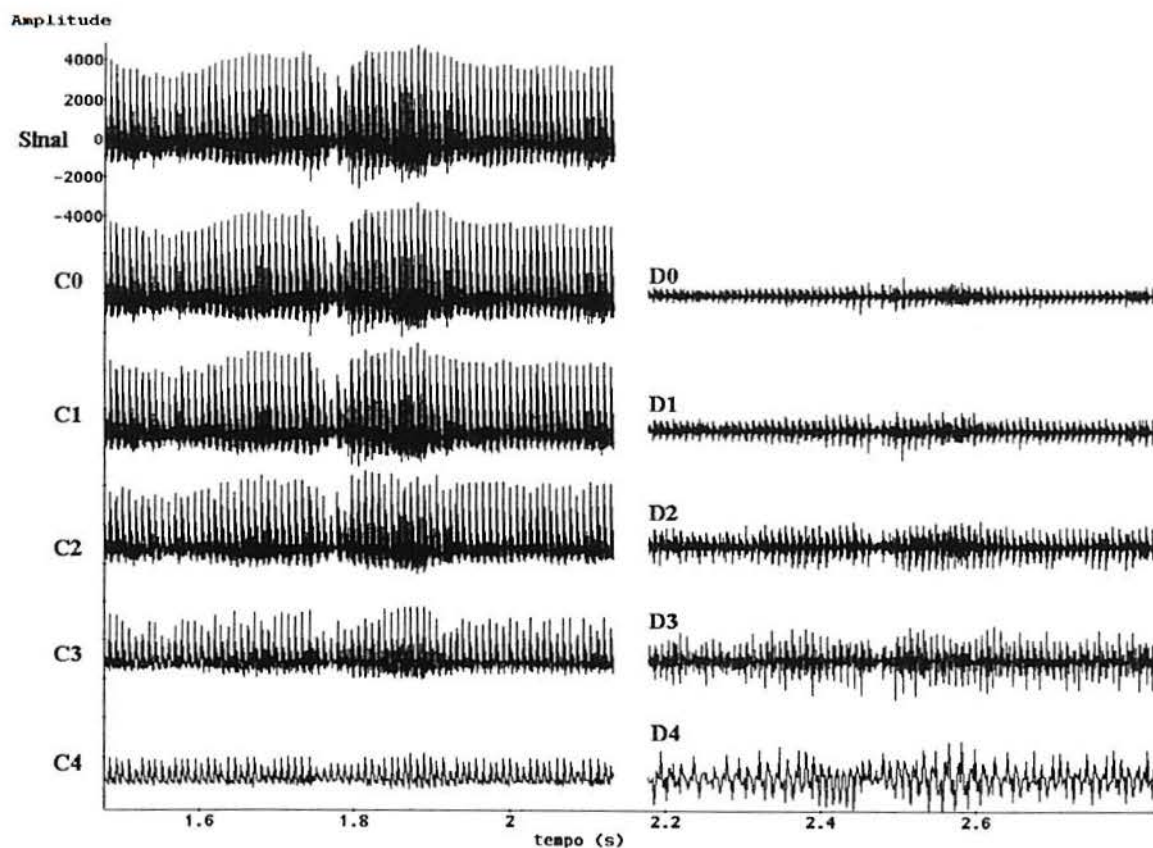


Figura A1.37: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC58.

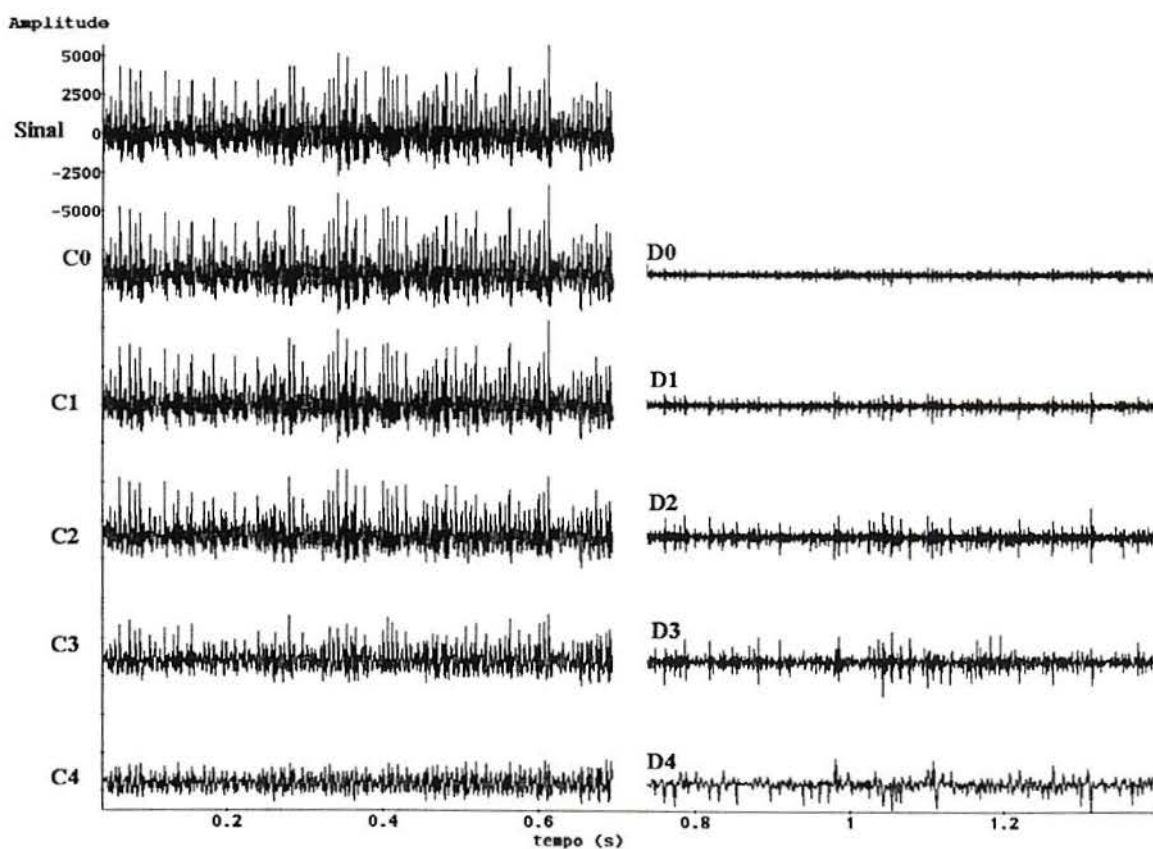


Figura A1.38: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC60.

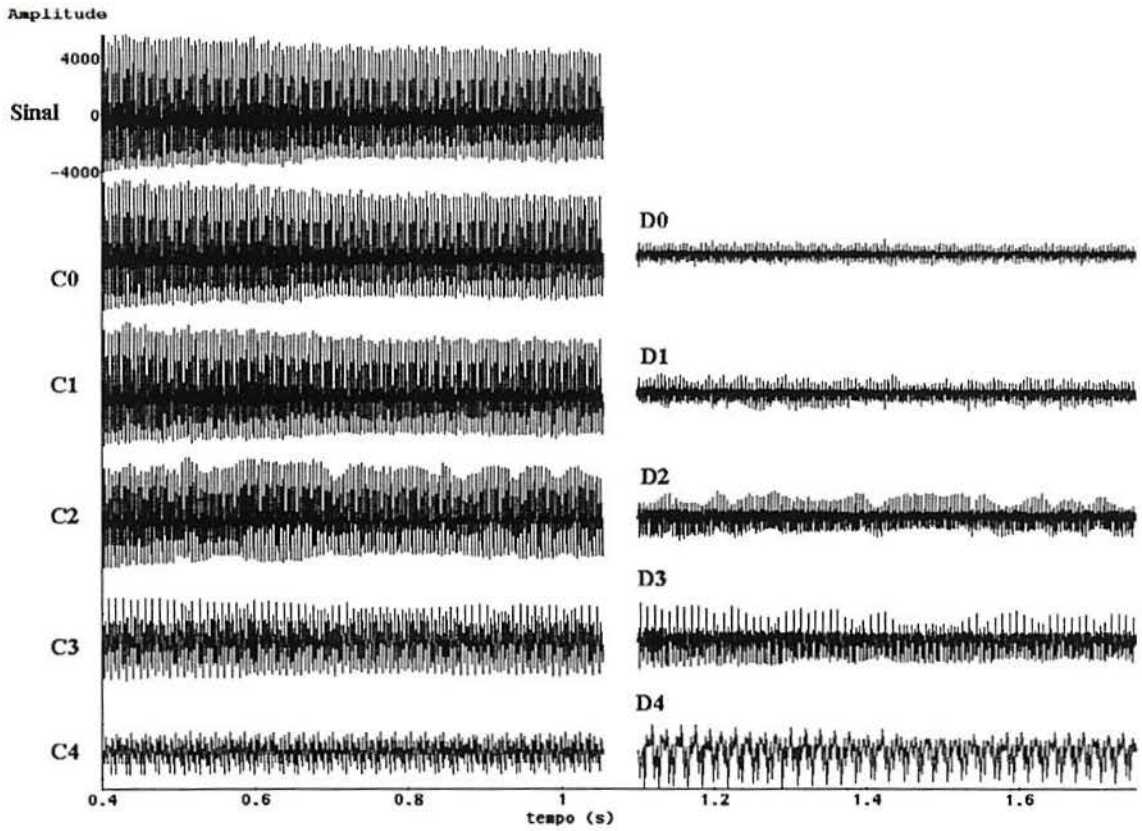


Figura A1.39: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC61.

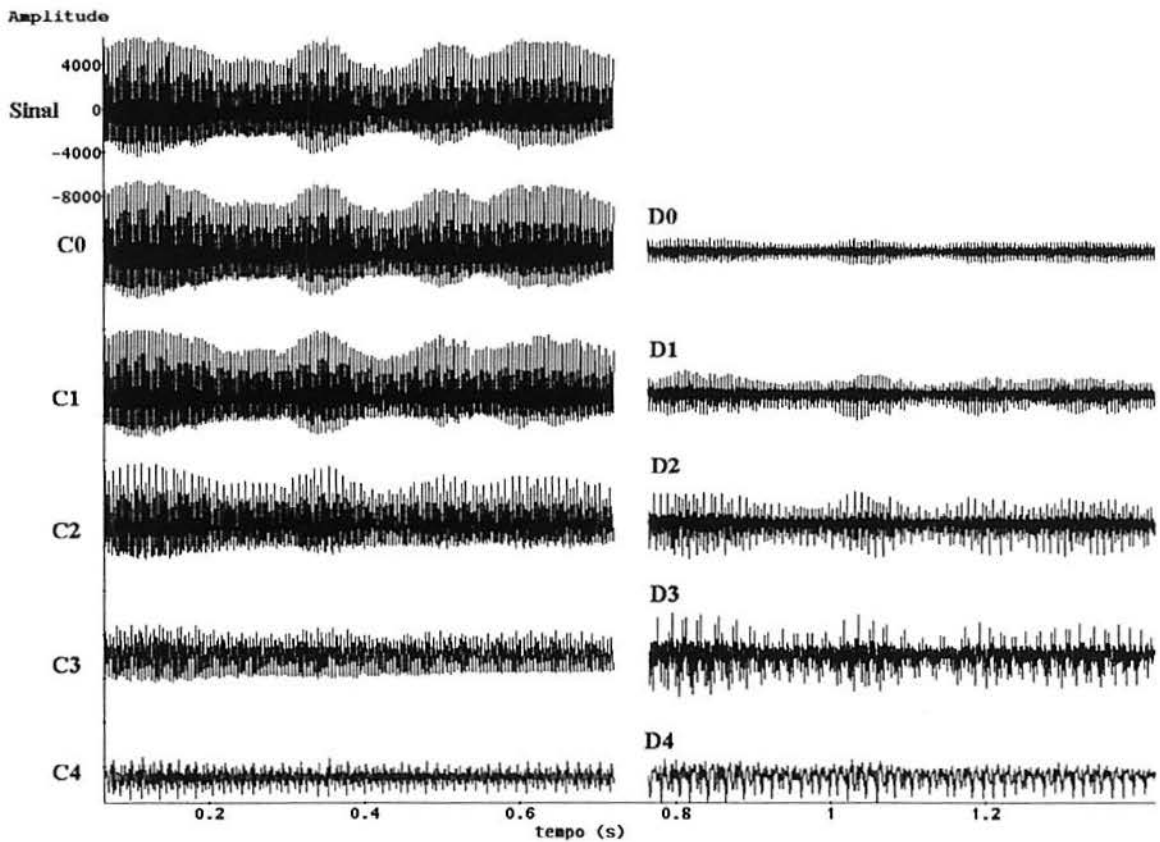


Figura A1.40: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC64.

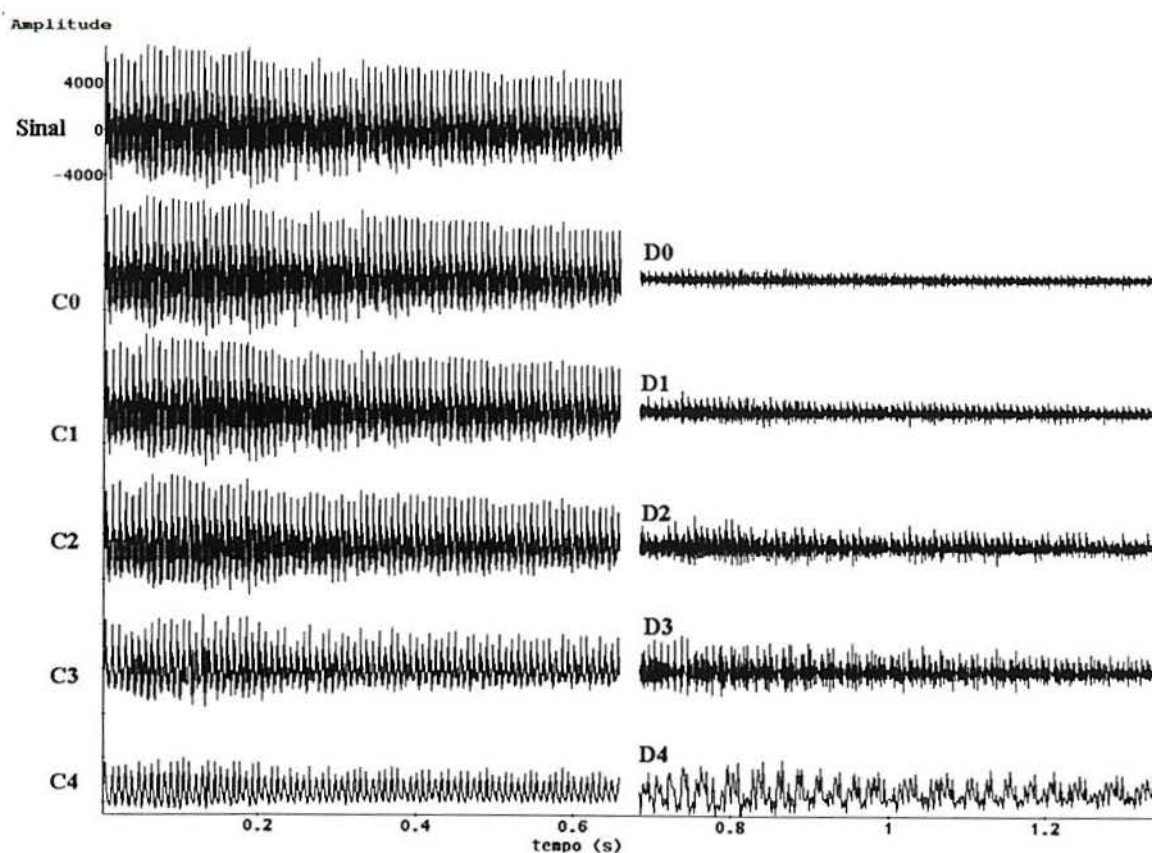


Figura A1.41: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC66.

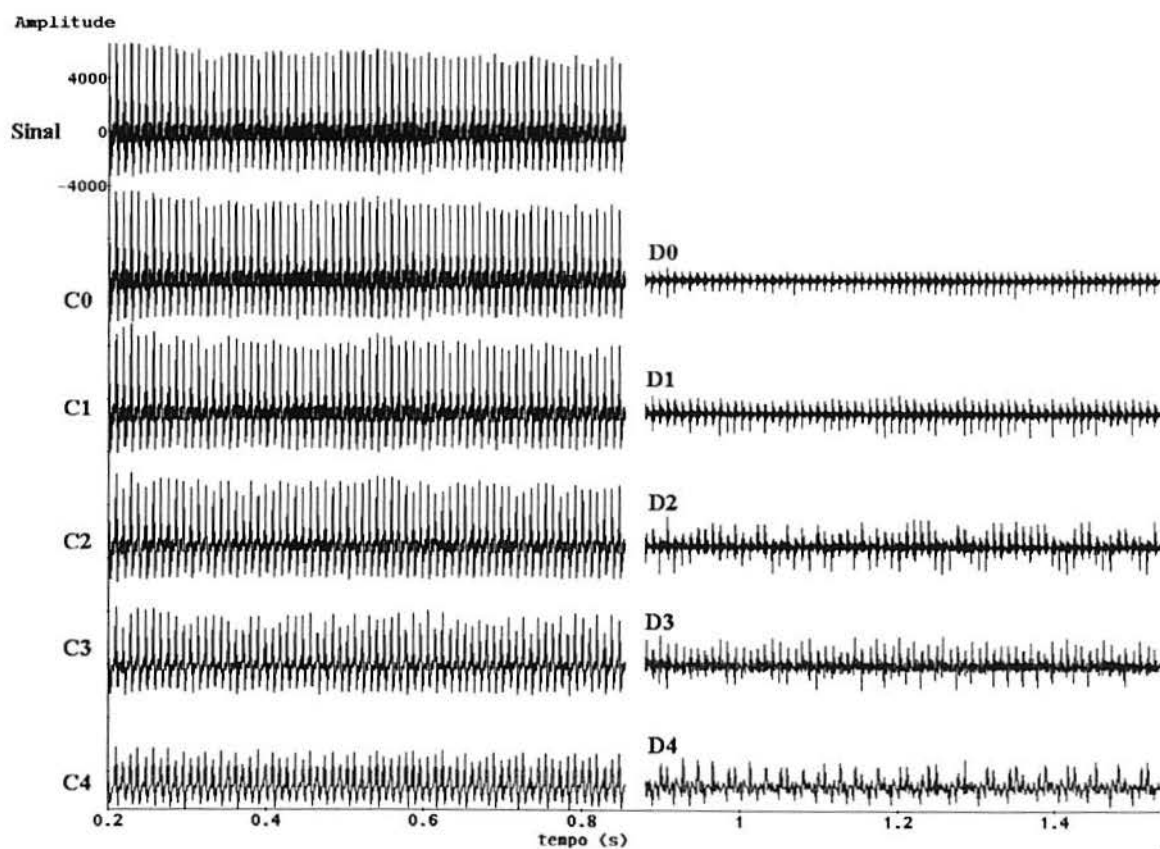


Figura A1.42: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC69.

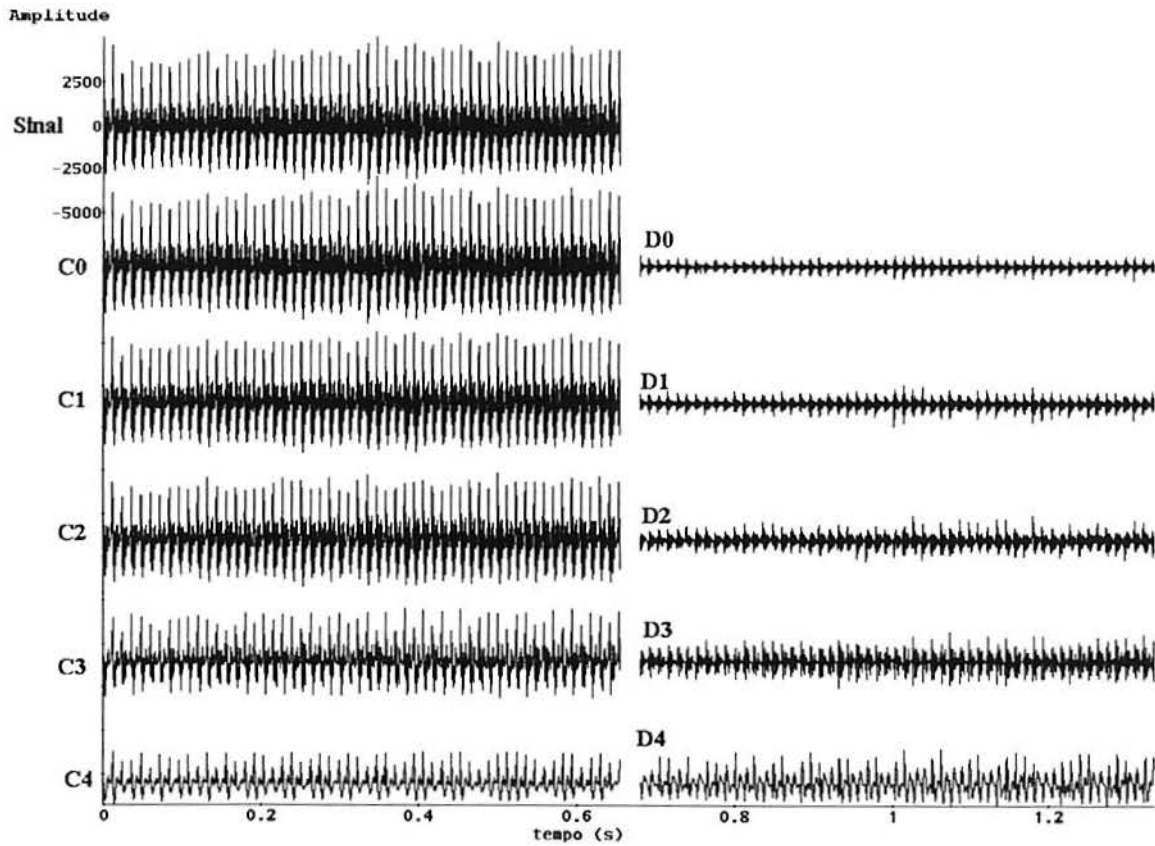


Figura A1.43: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC70.

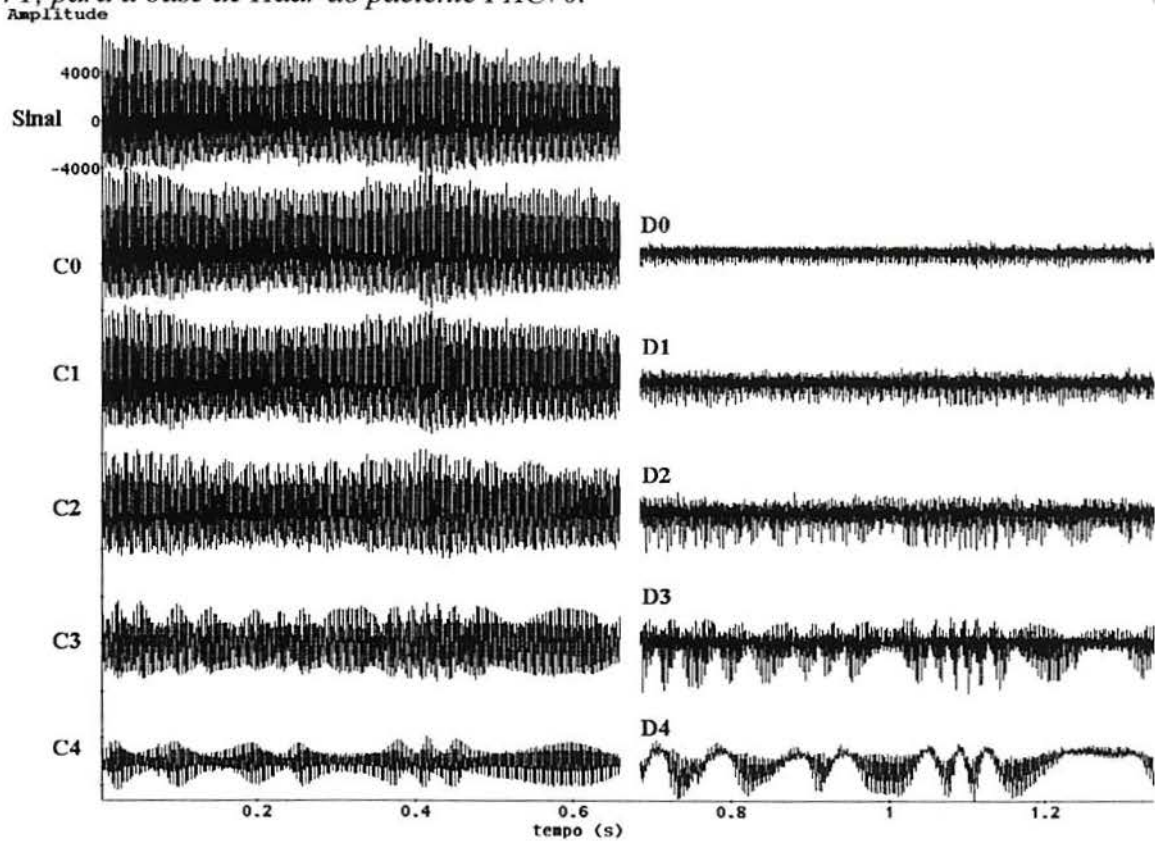


Figura A1.44: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC71.

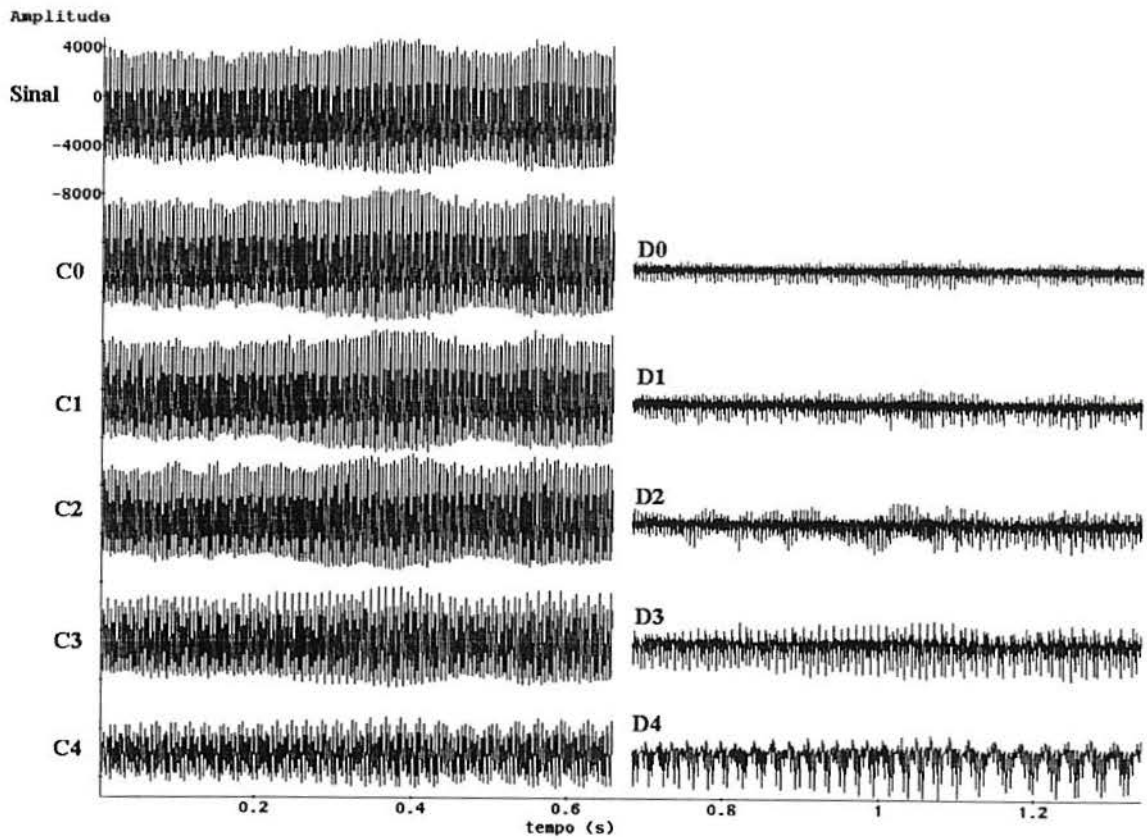


Figura A1.45: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC72.

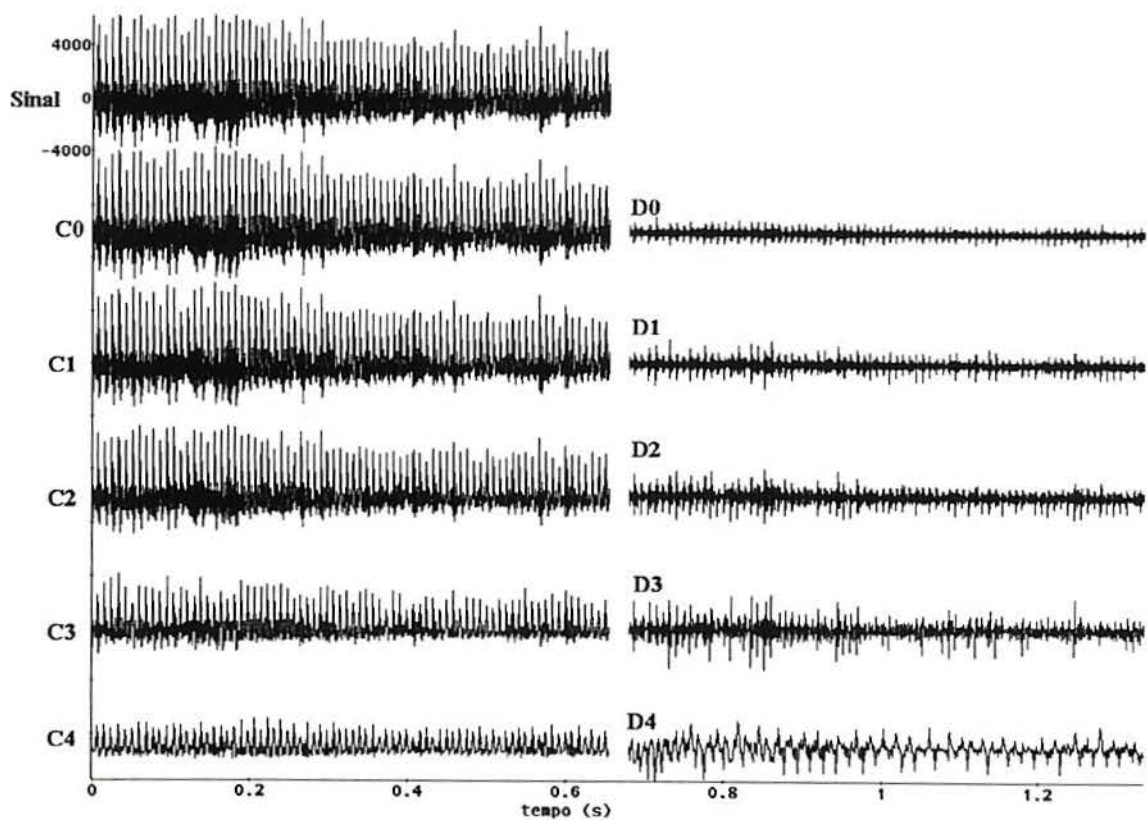


Figura A1.46: Gráficos do trecho de 16384 pontos do sinal e dos 5 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC73.

## ANEXO 2: FFTs dos sinais dos pacientes, para o primeiro procedimentos de aquisição

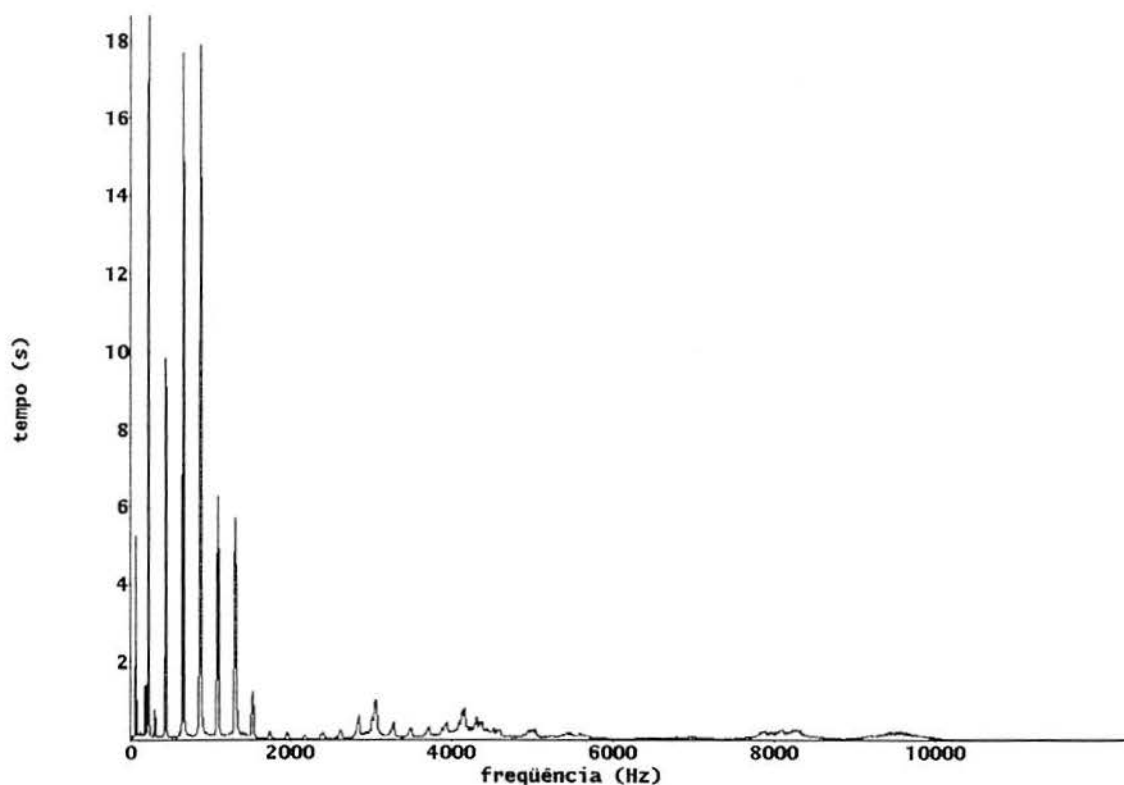


Figura A2.1: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC01.

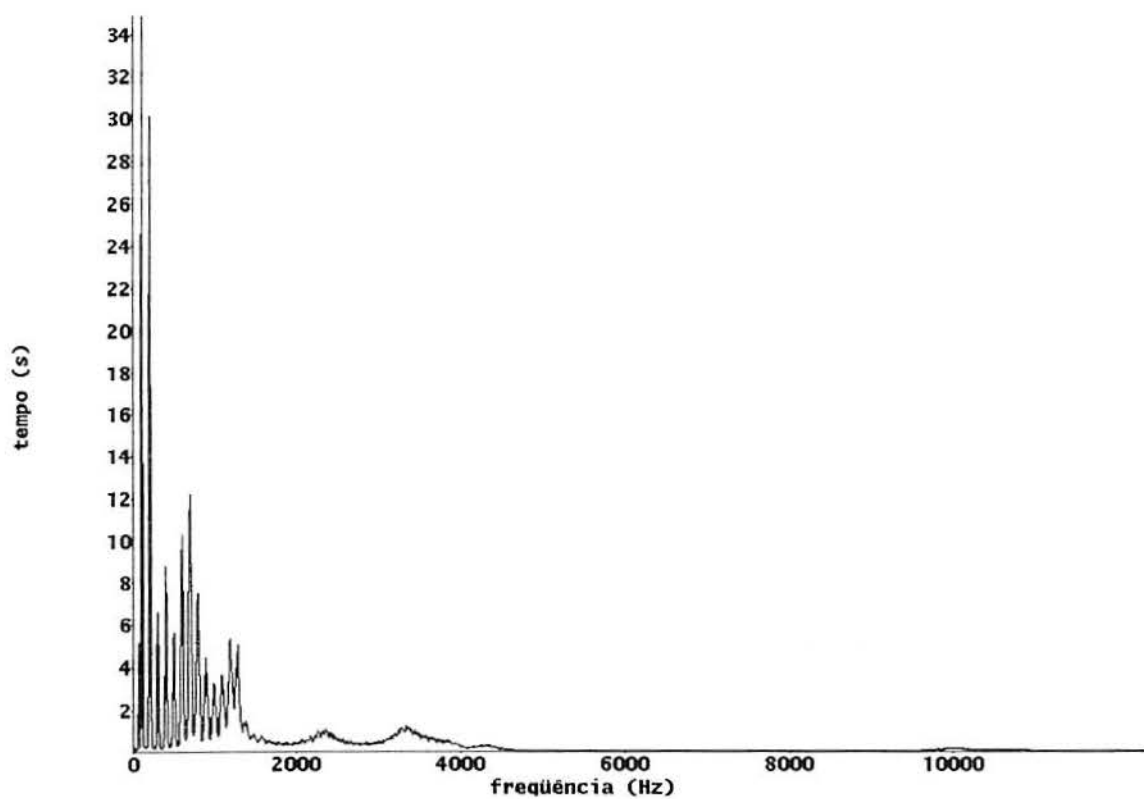


Figura A2.2: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC02.

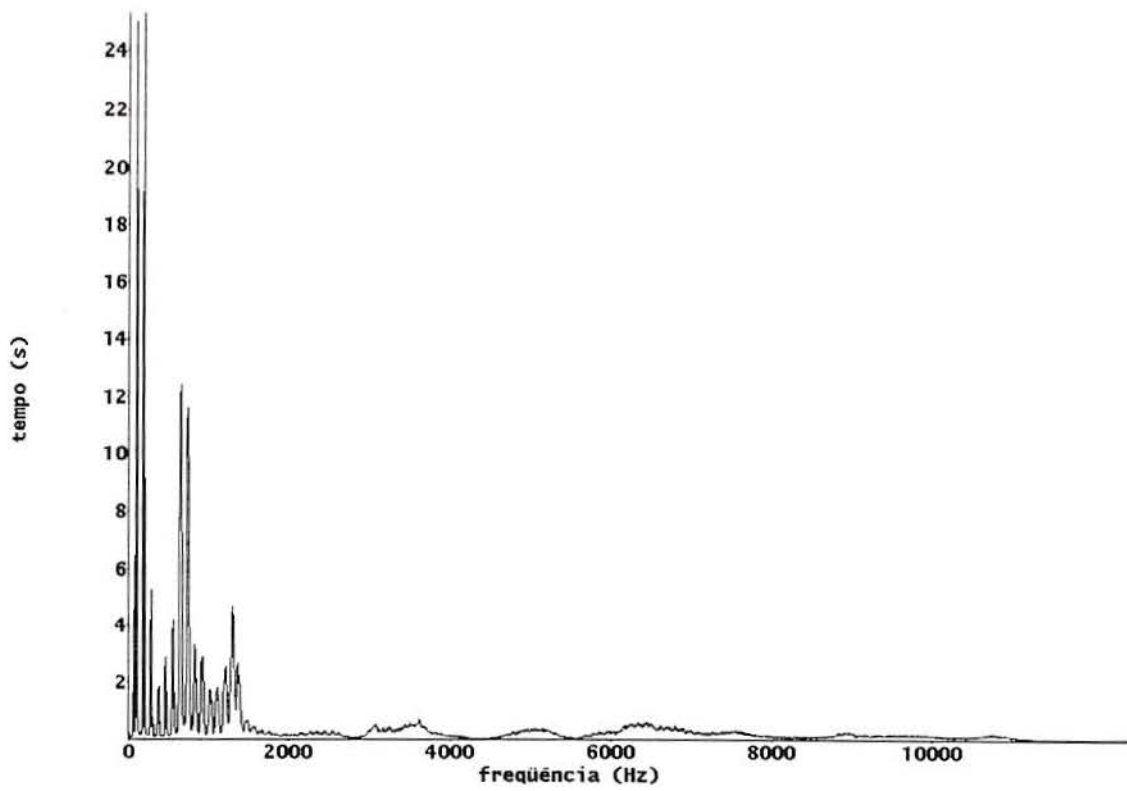


Figura A2.3: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC04.

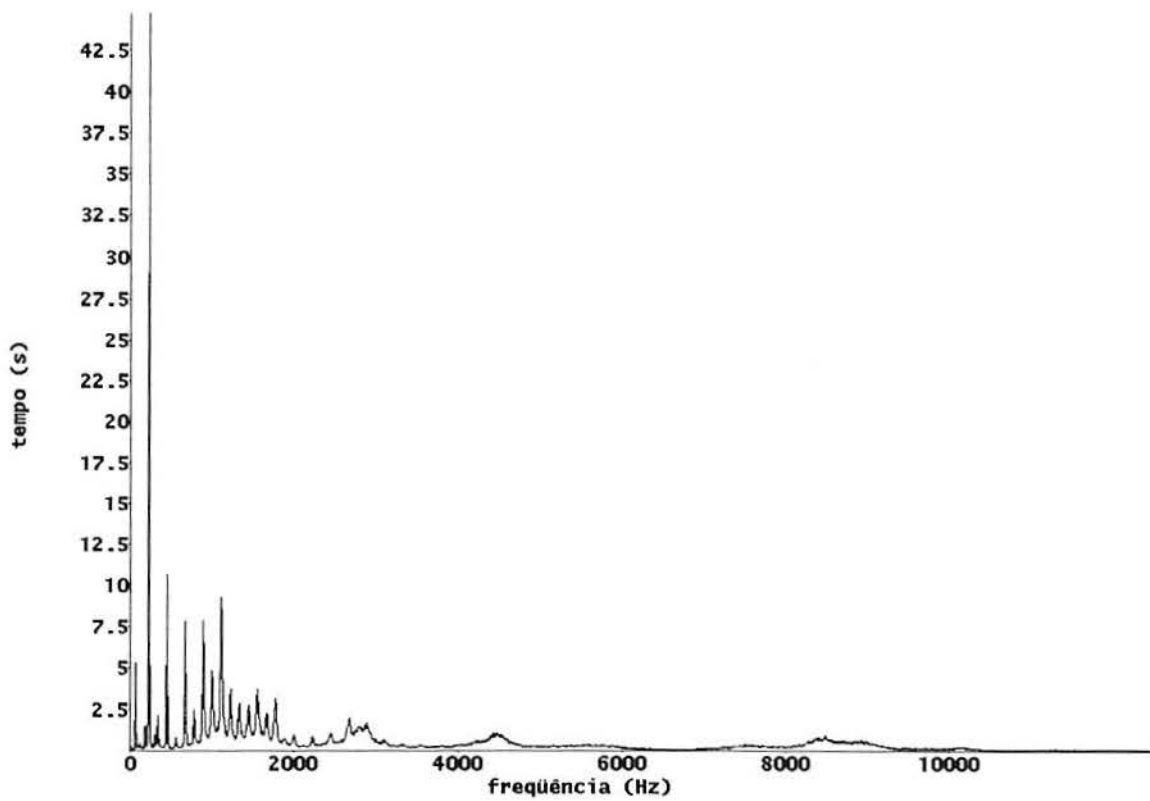


Figura A2.4: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC05.

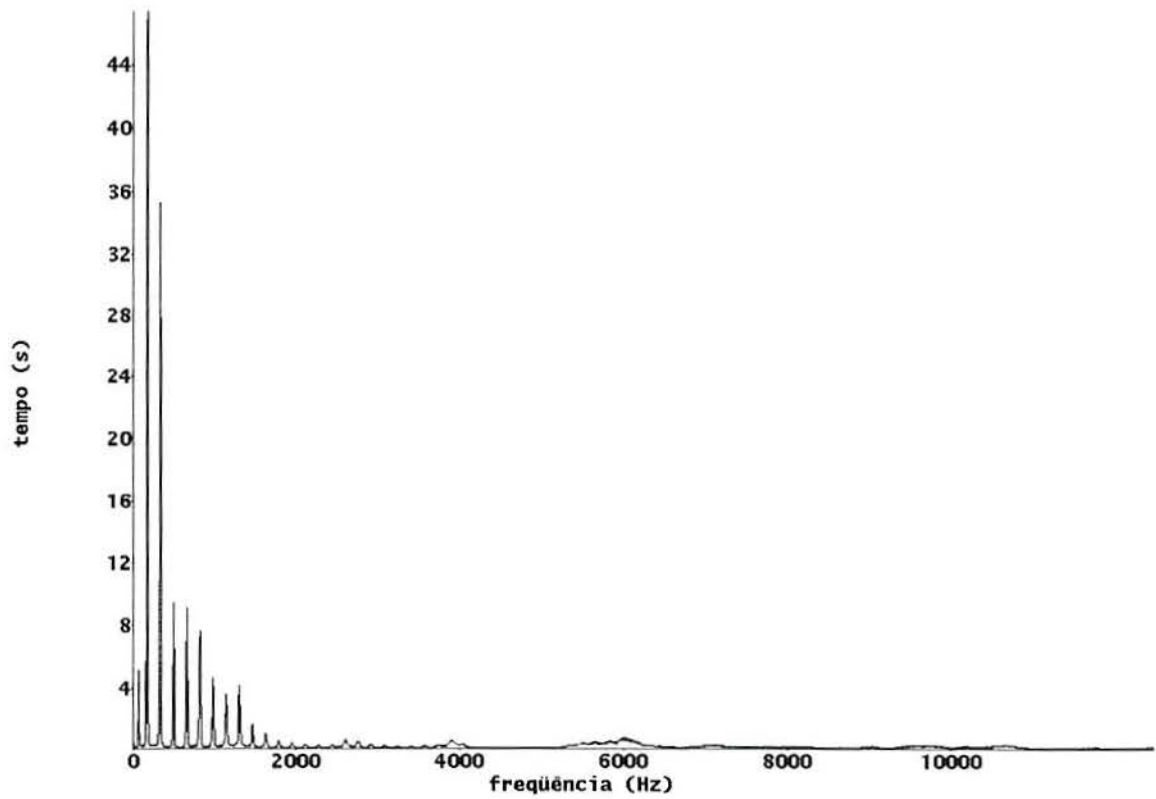


Figura A2.5: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC06.

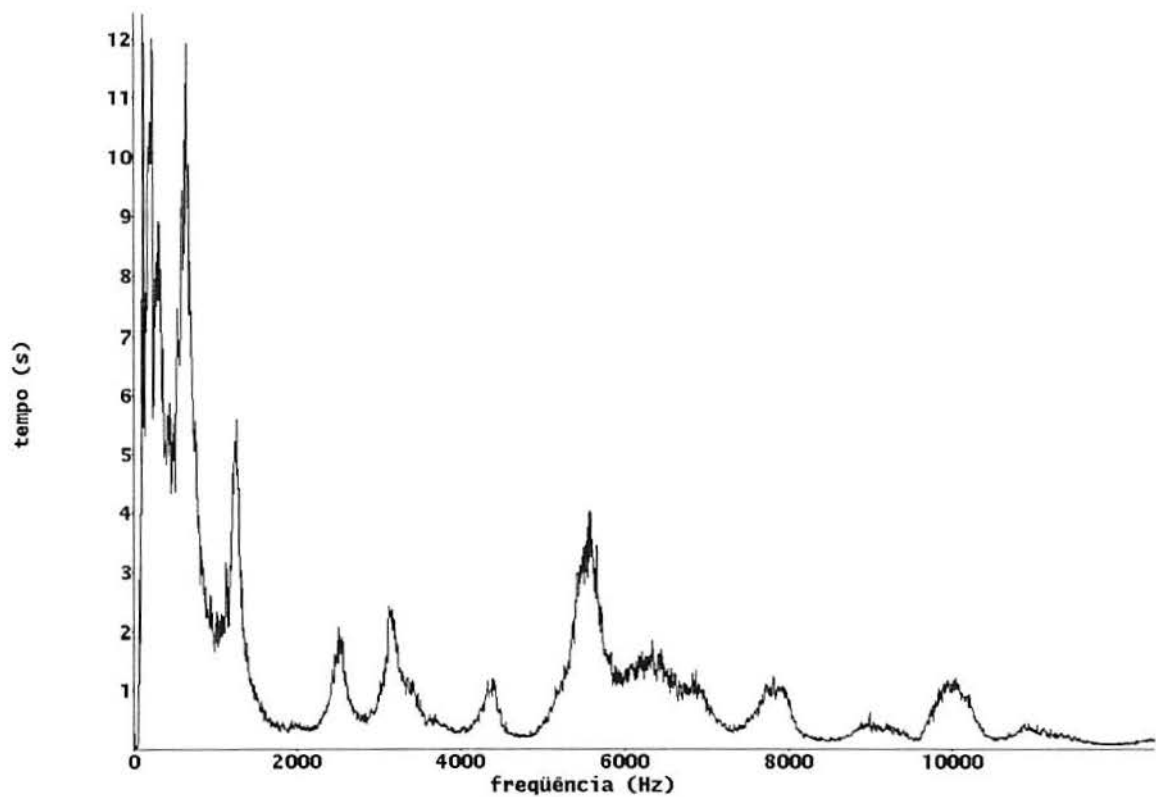


Figura A2.6: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC08



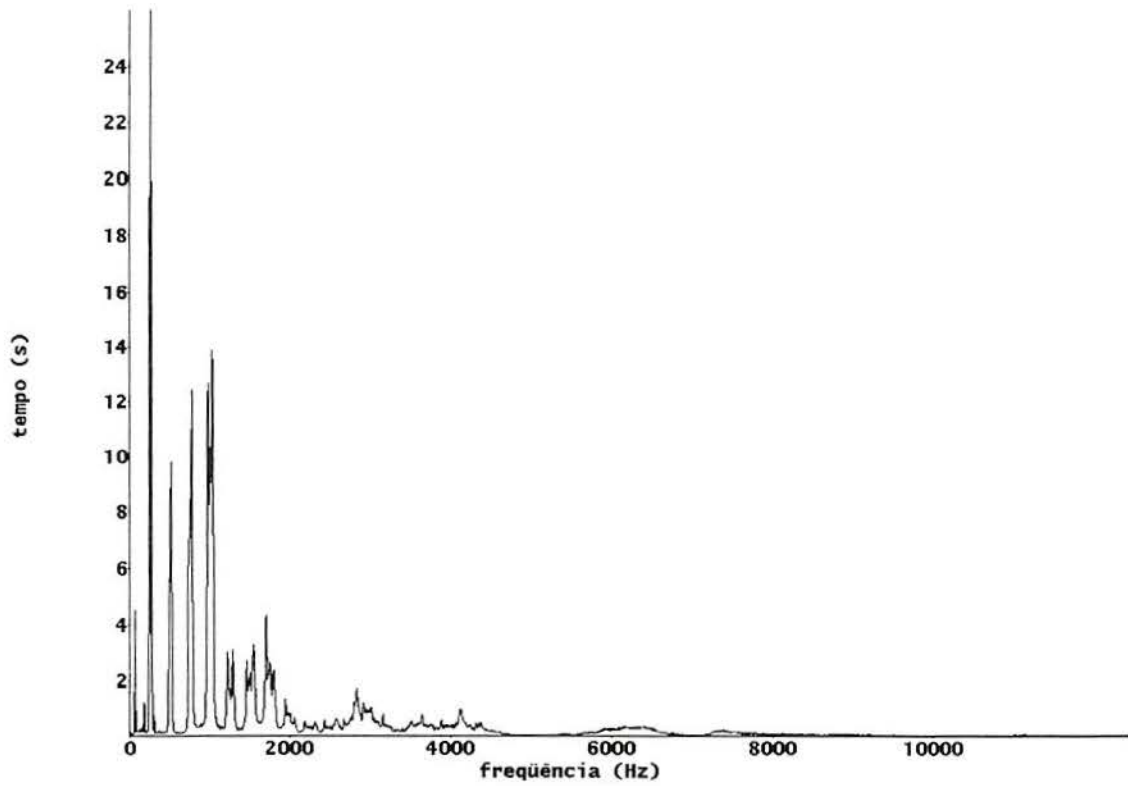


Figura A2.7: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC09.

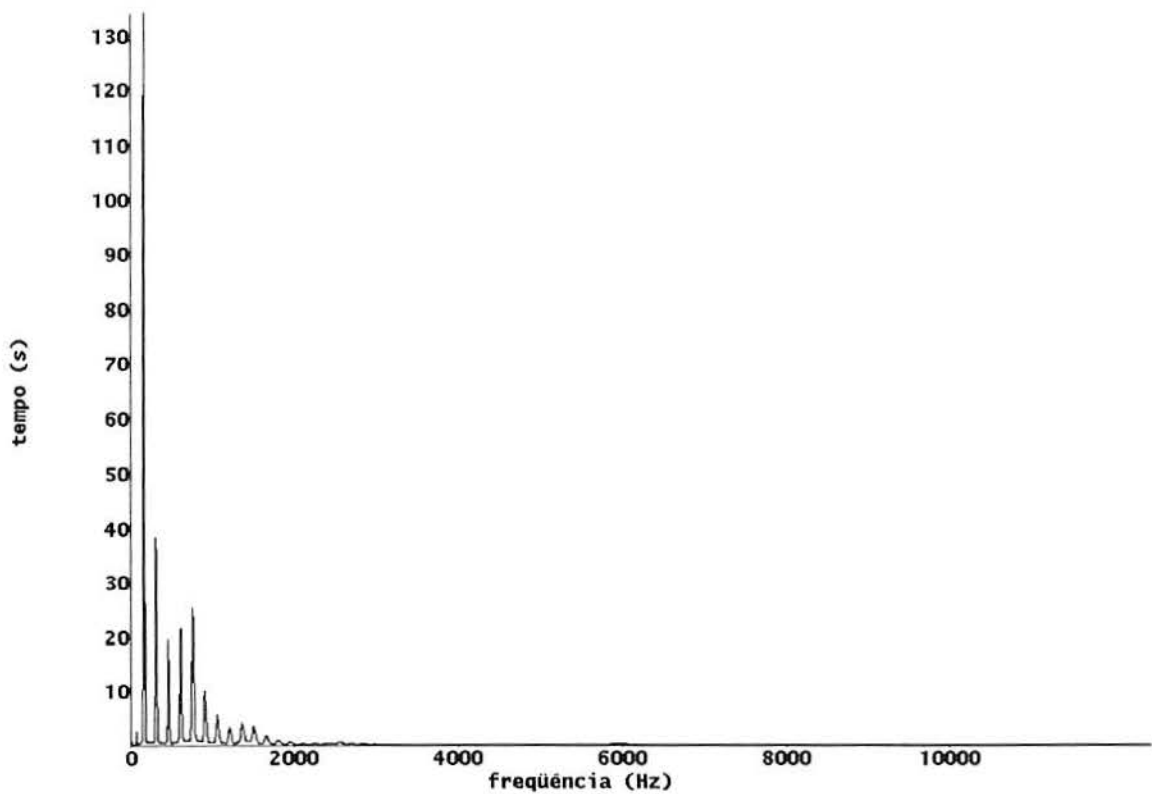


Figura A2.8: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC10.

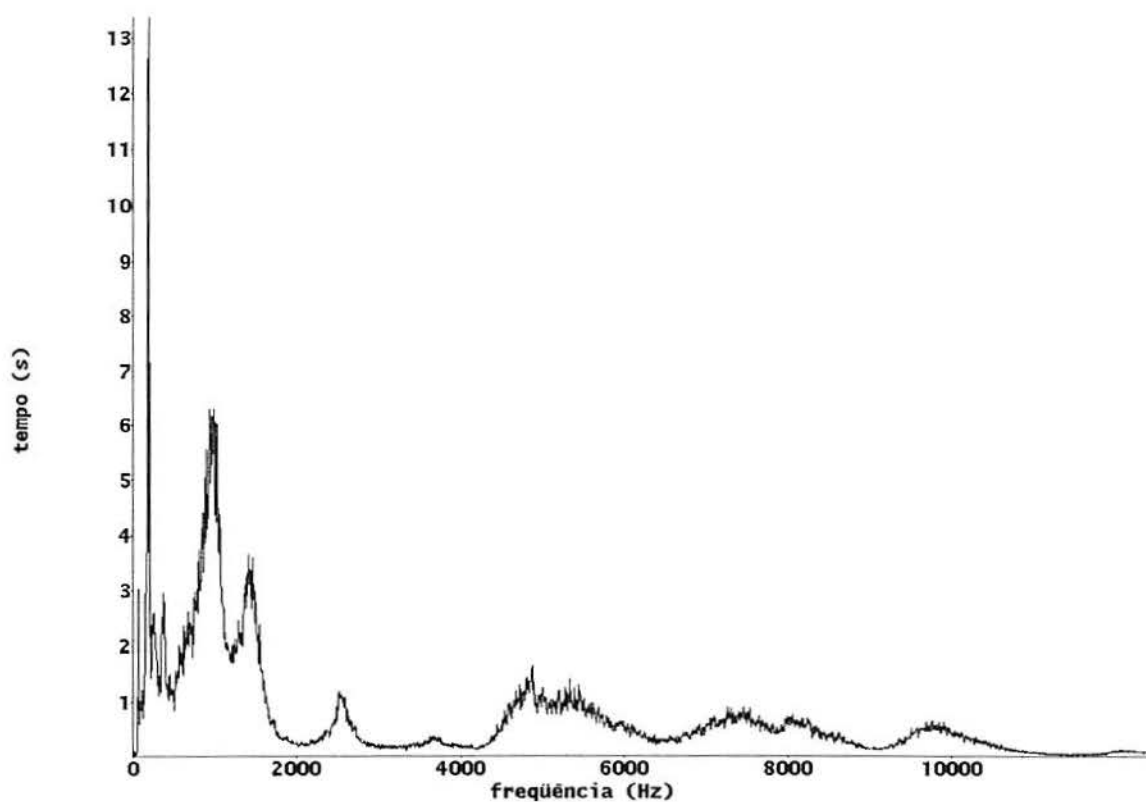


Figura A2.9: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC11.

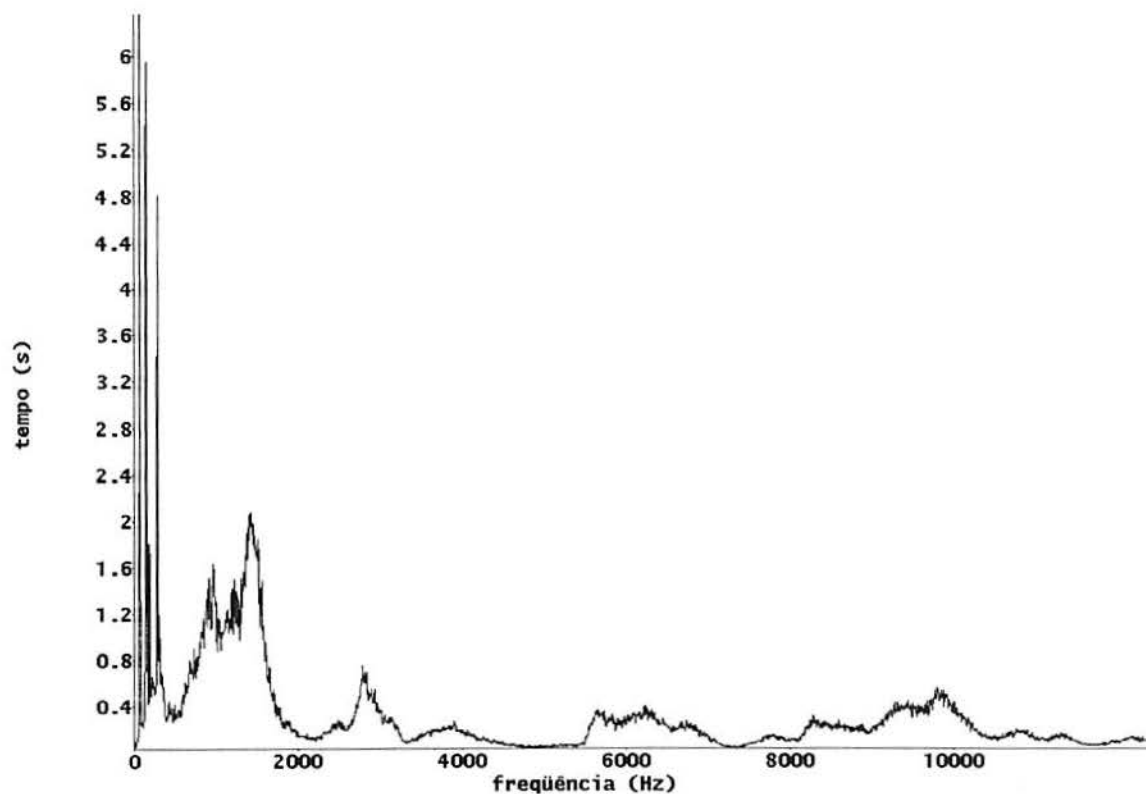


Figura A2.10: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC13.

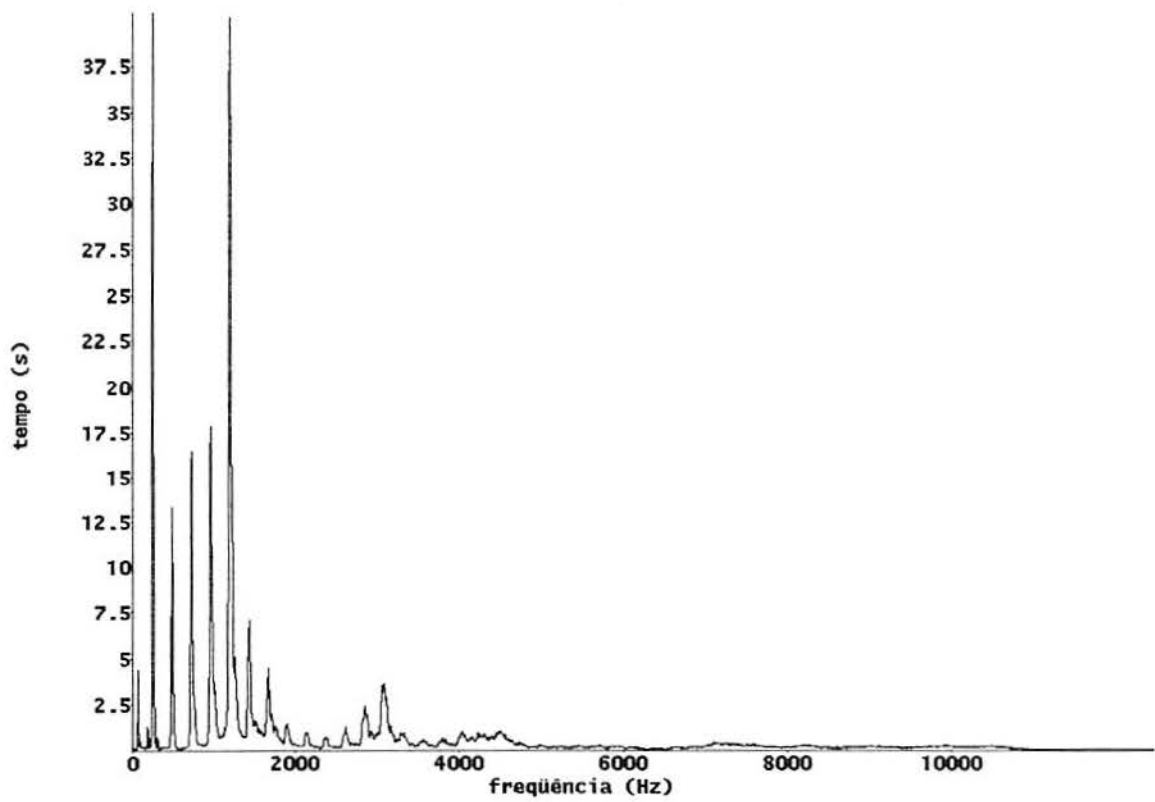


Figura A2.11: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC14.

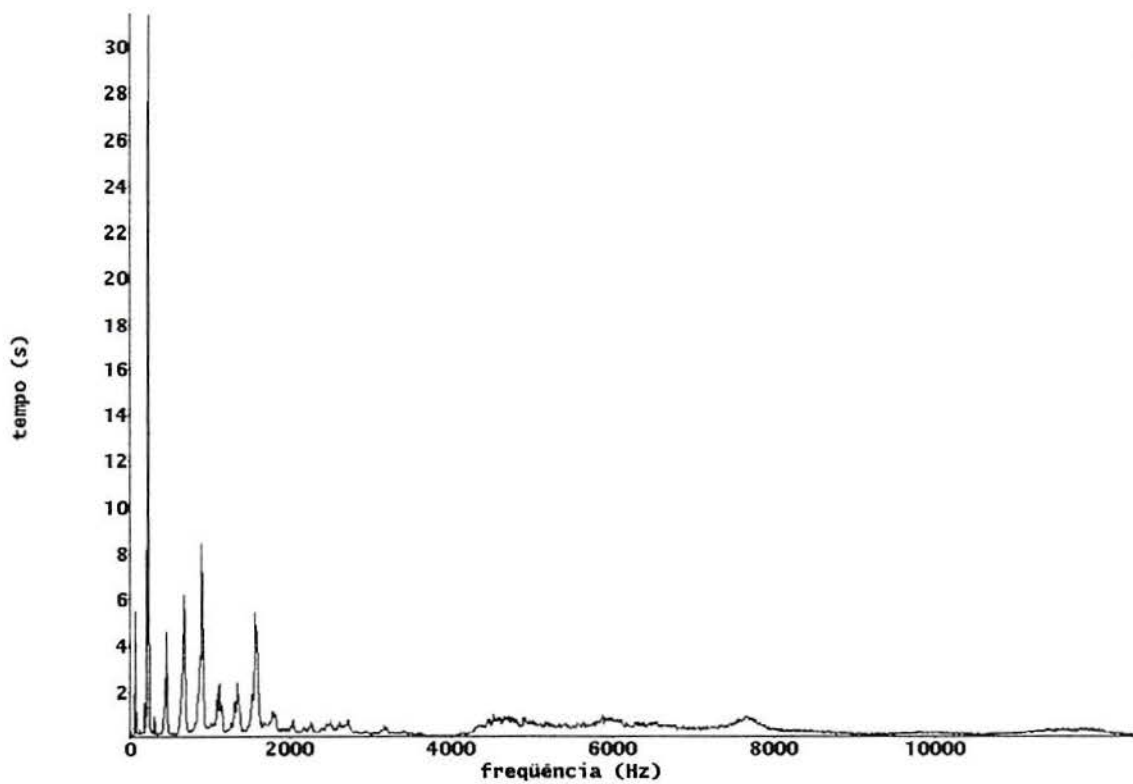


Figura A2.12: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC15.

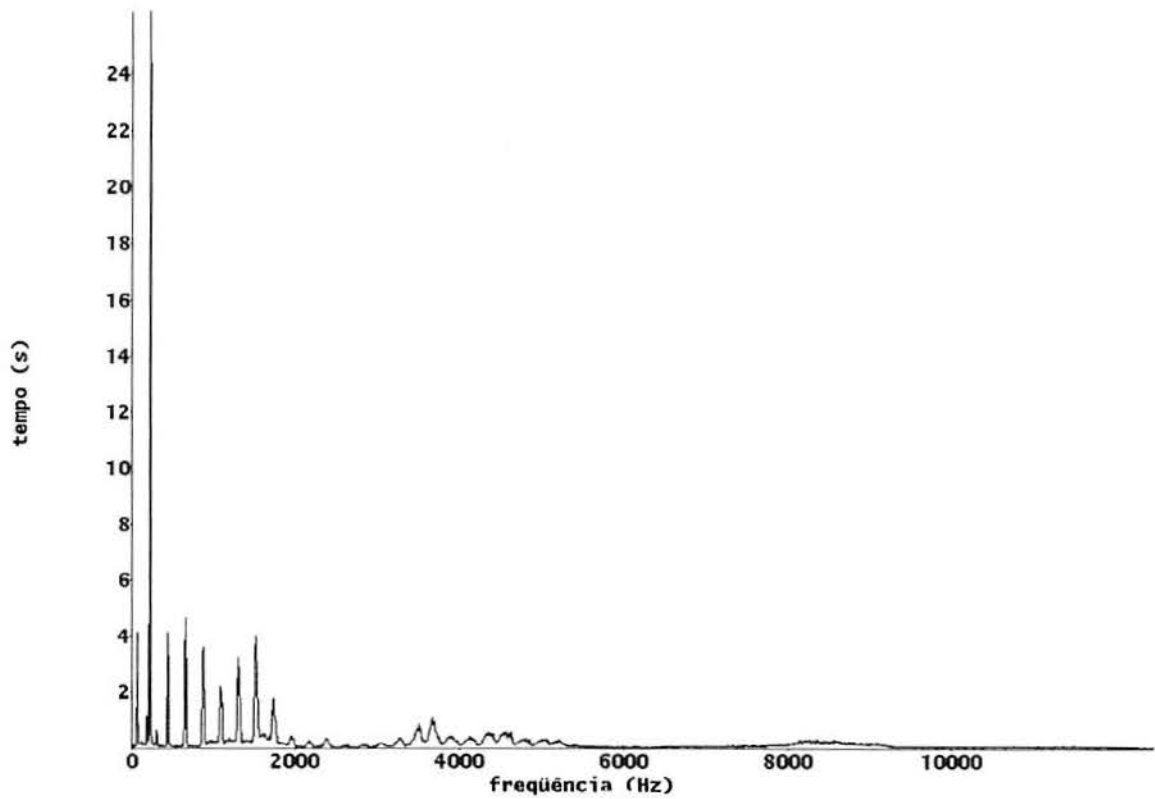


Figura A2.13: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC16.

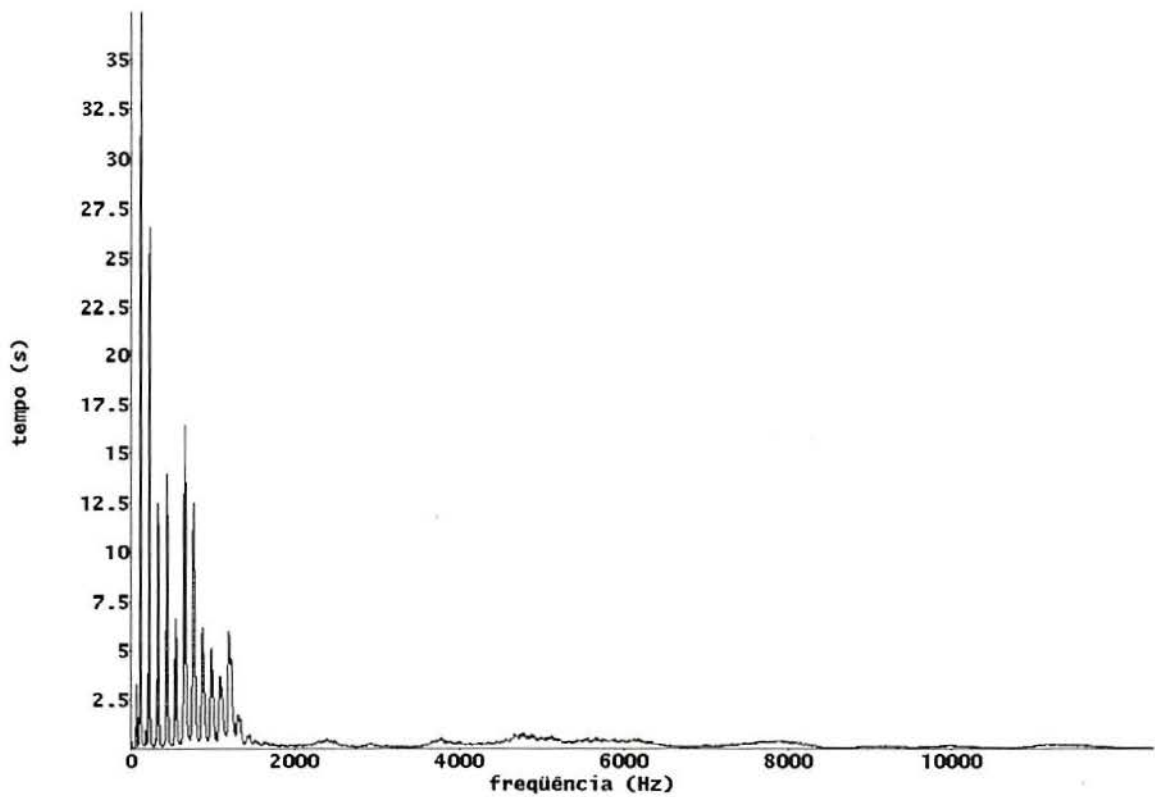


Figura A2.14: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC17.

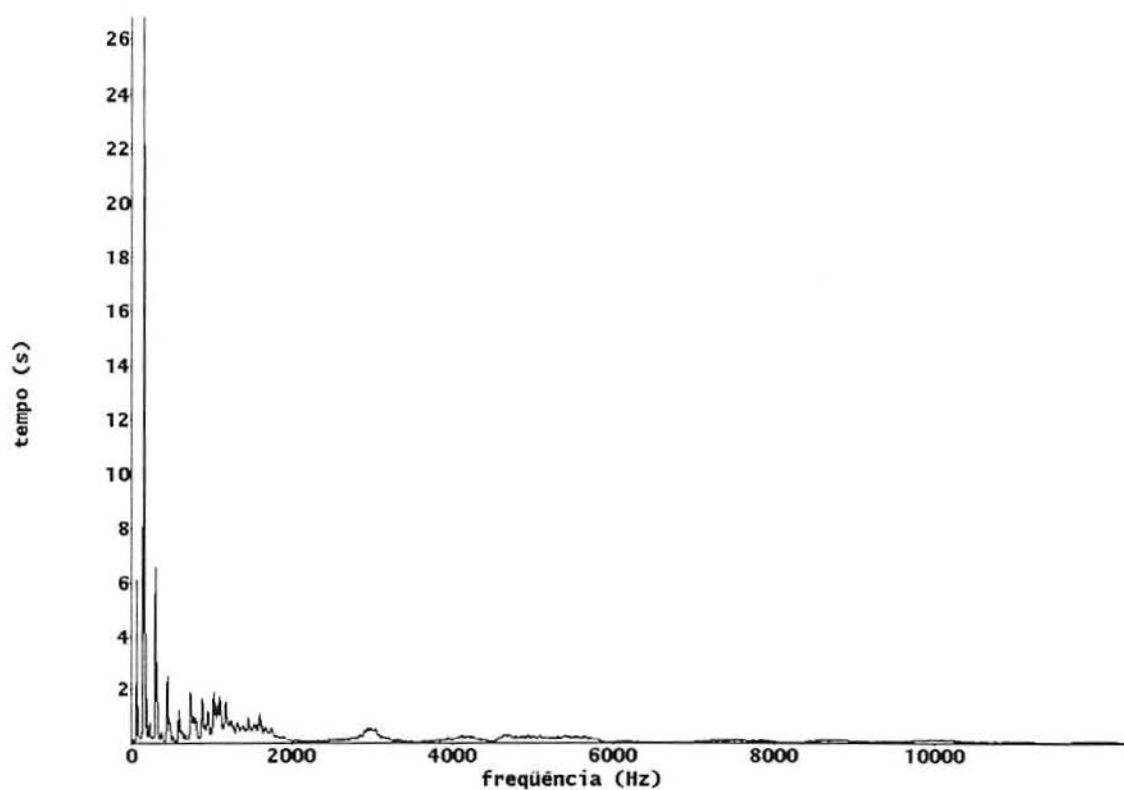


Figura A2.15: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC18.

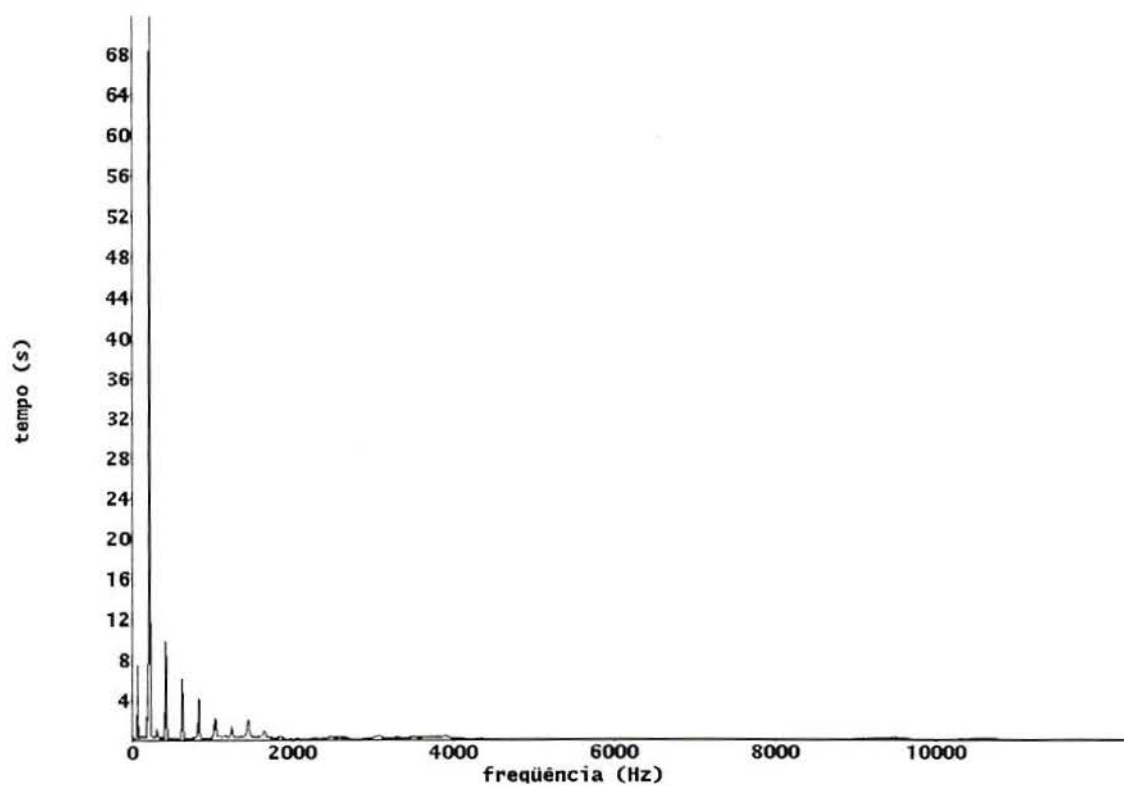


Figura A2.16: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC19.

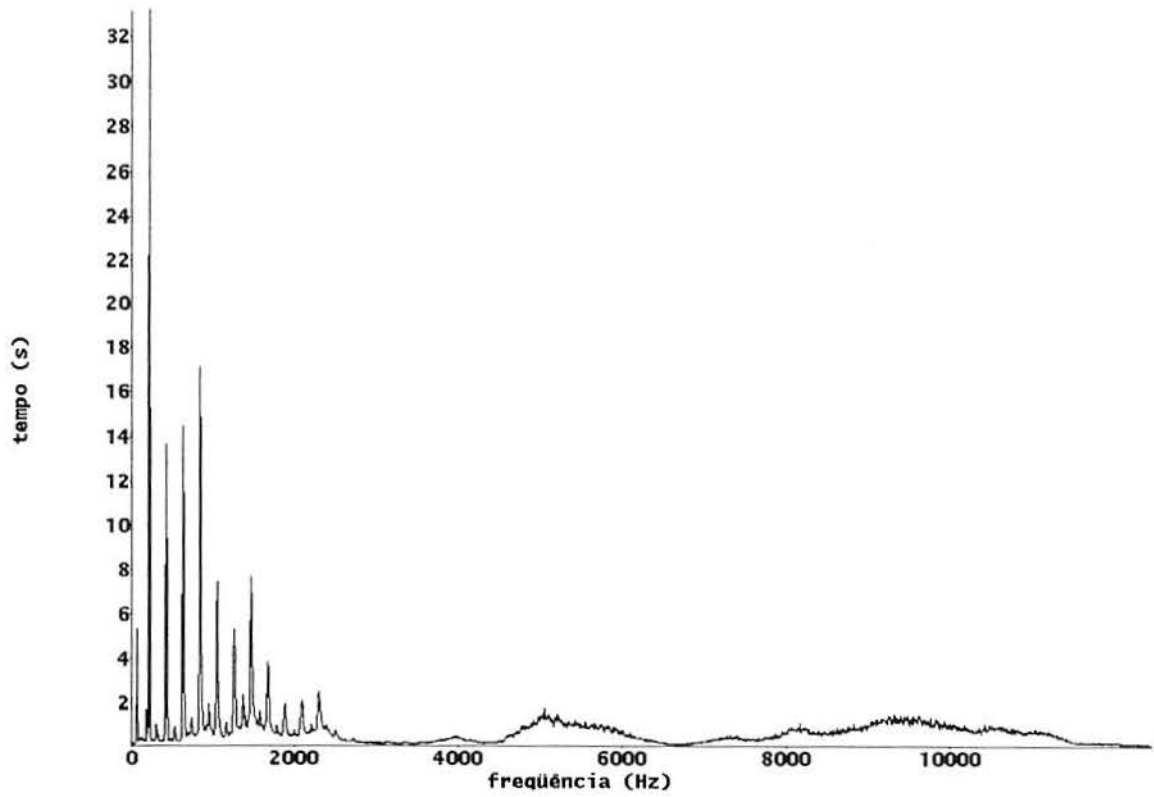


Figura A2.17: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC21.

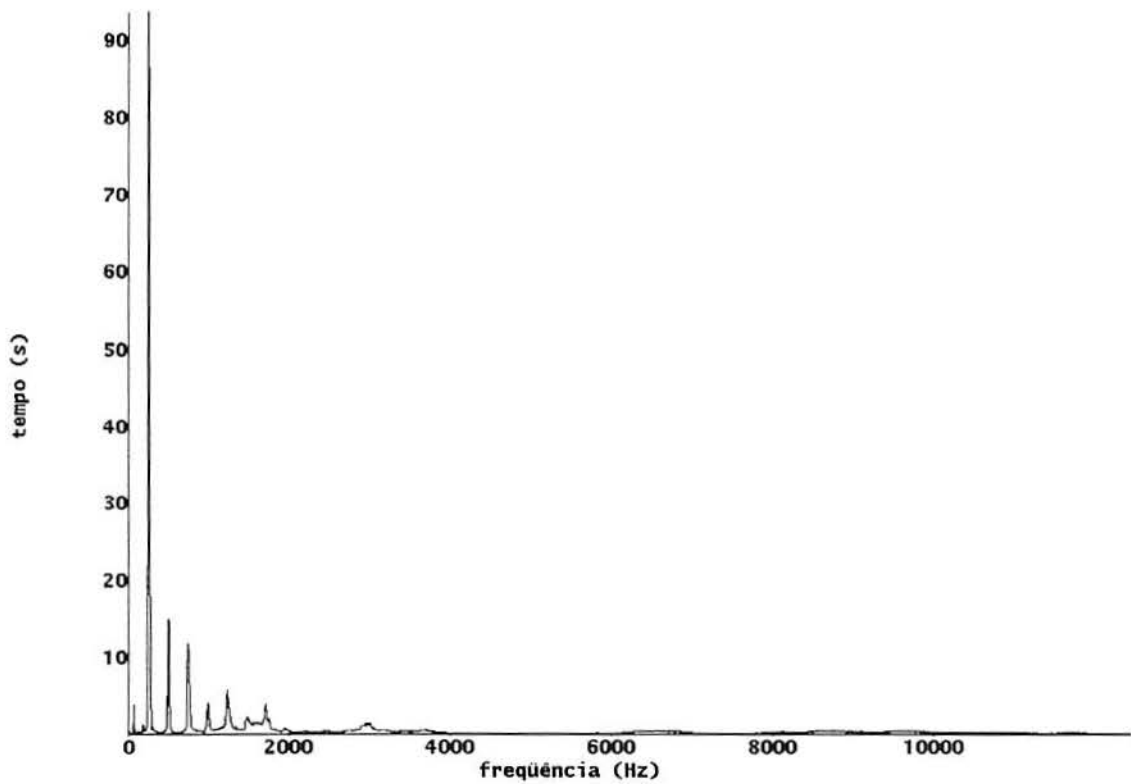


Figura A2.18: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC22.

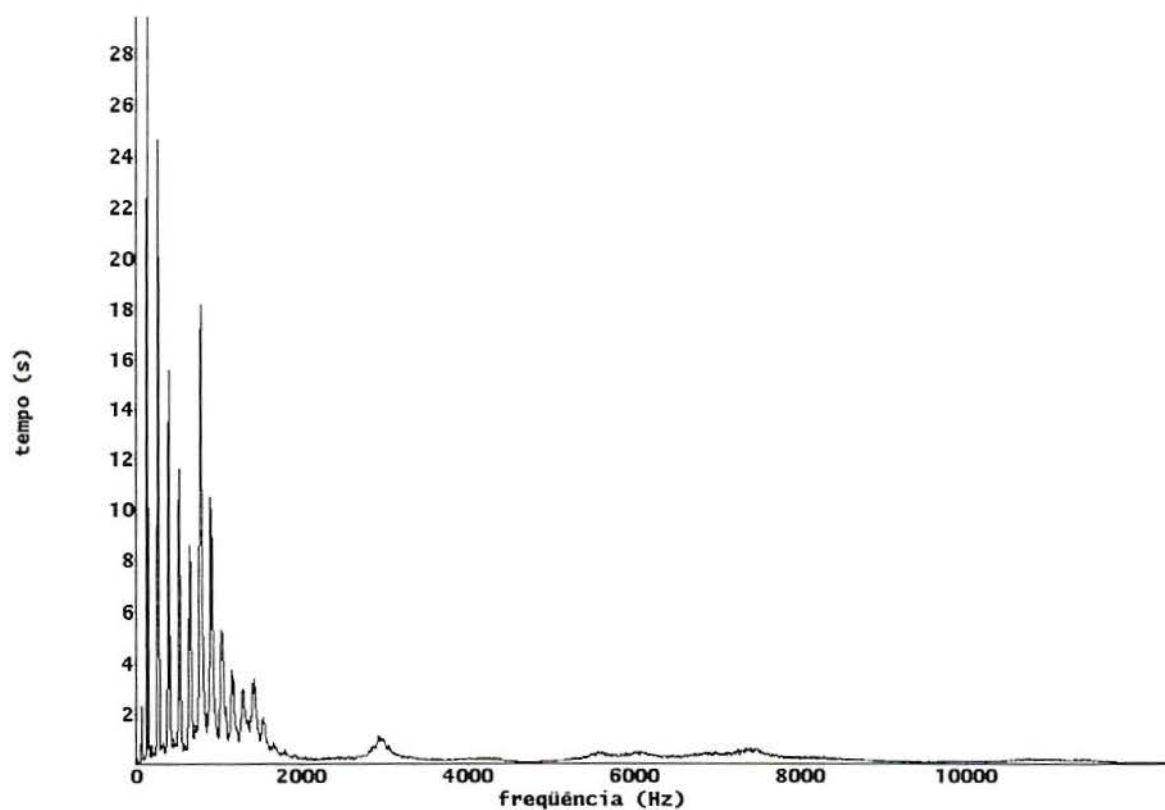


Figura A2.19: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC23.

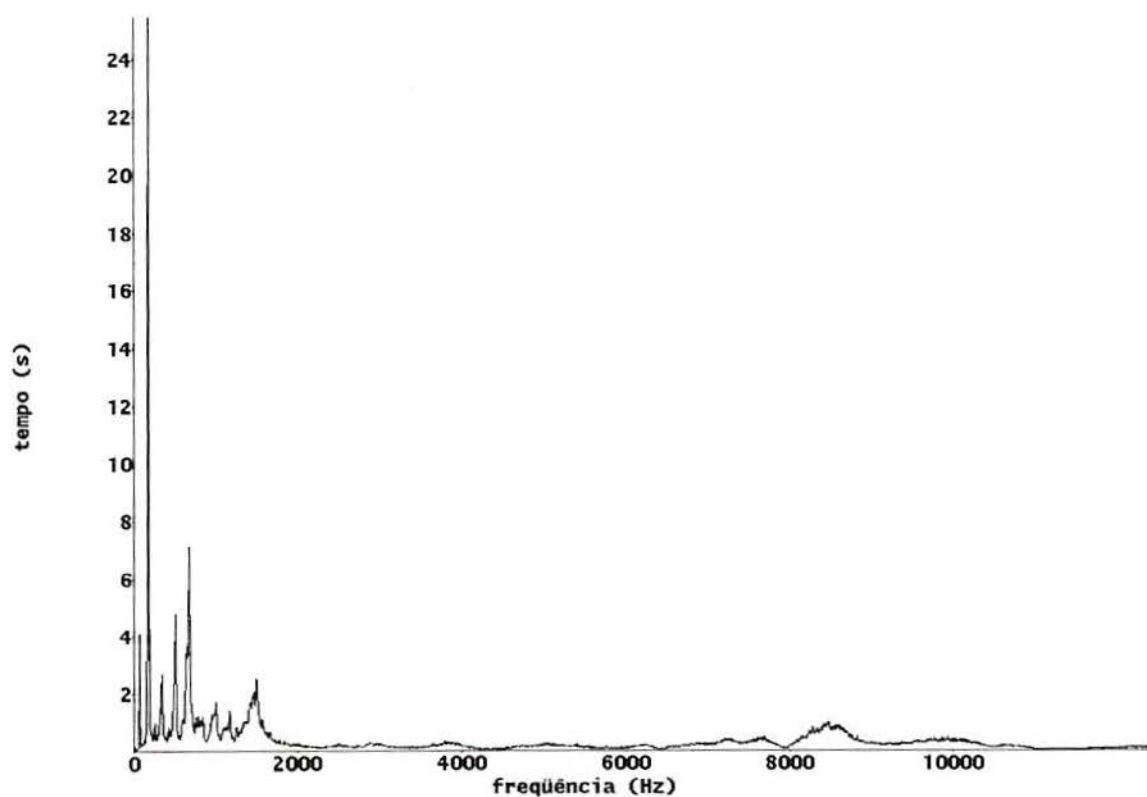


Figura A2.20: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC24.

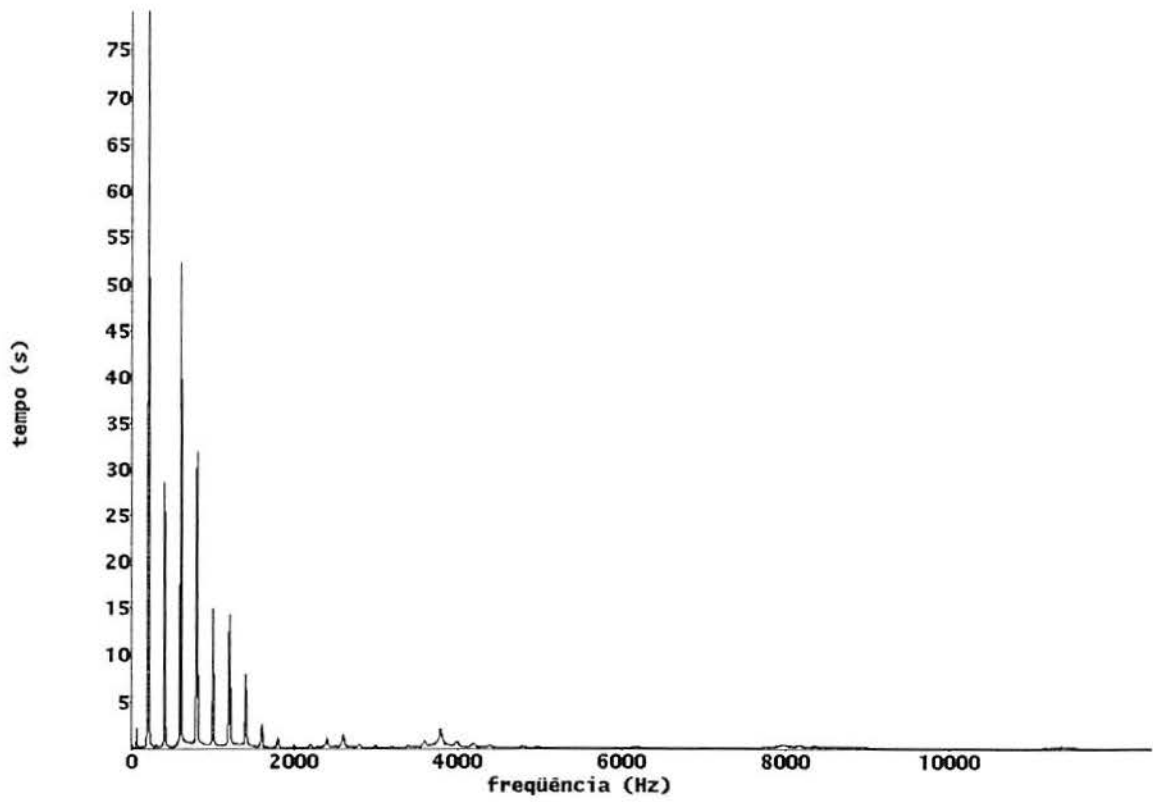


Figura A2.21: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC26.

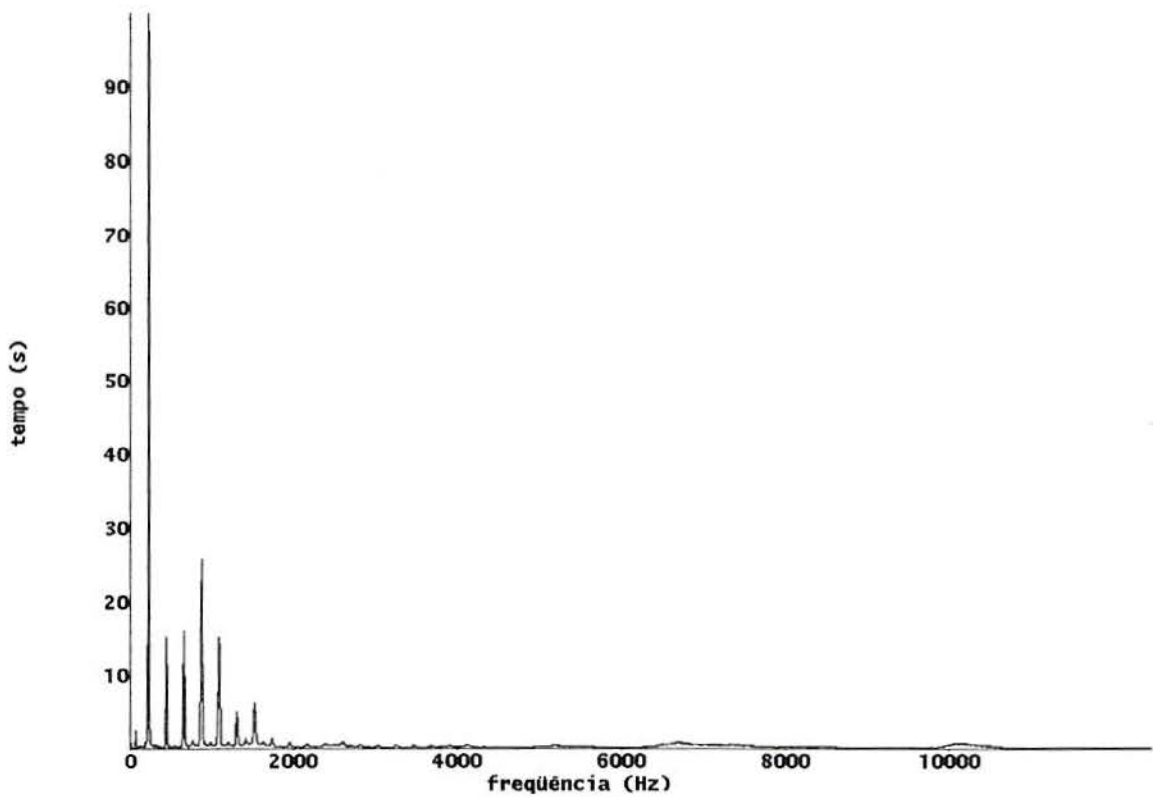


Figura A2.22: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC27.



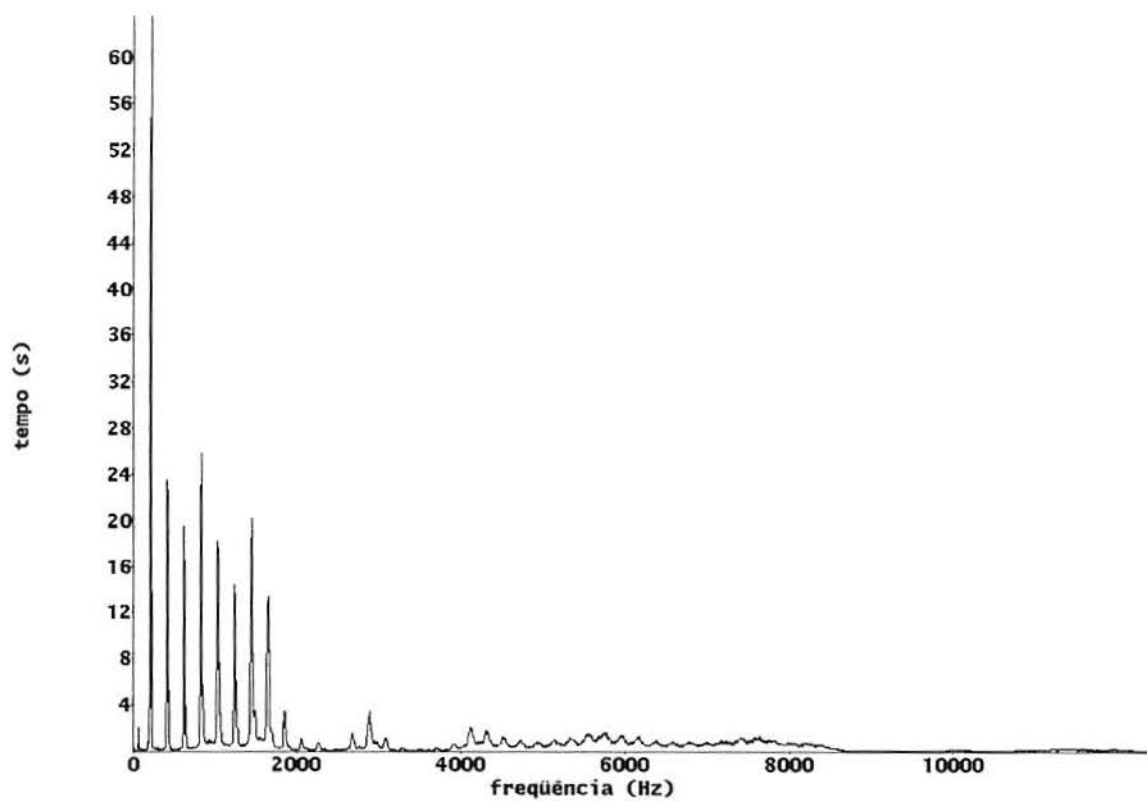


Figura A2.23: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC28.

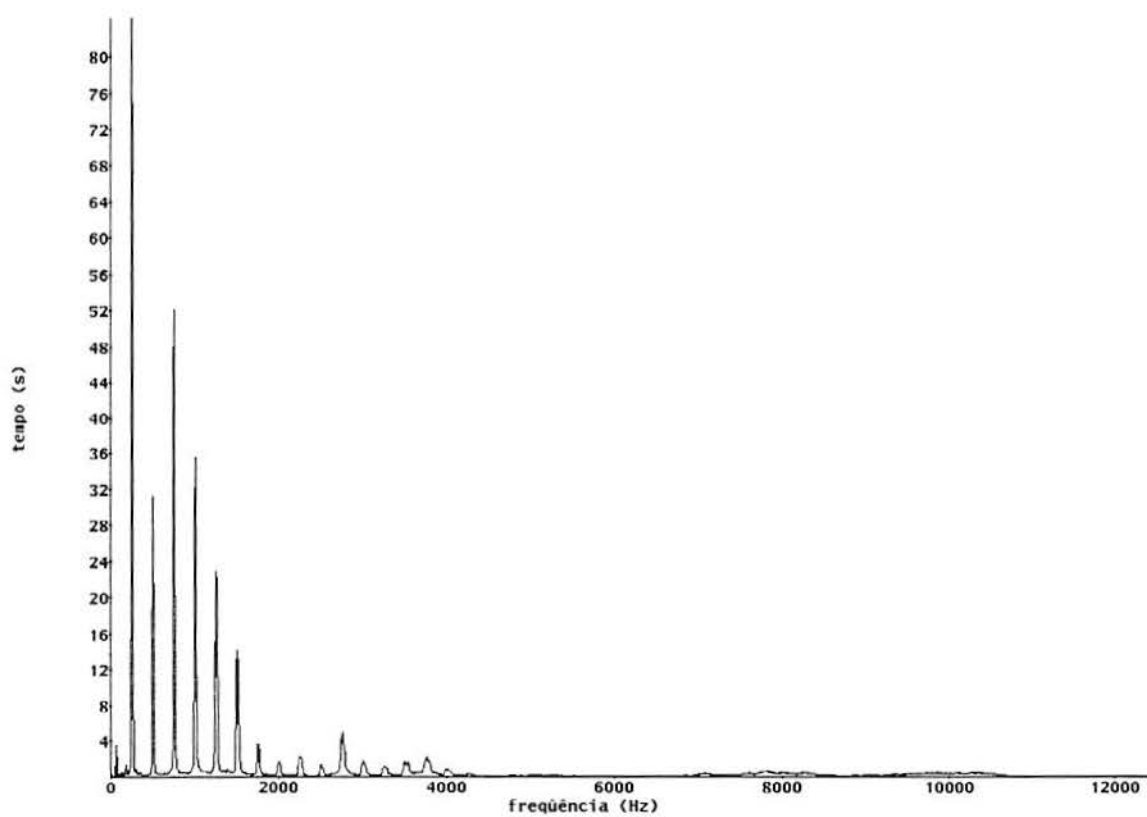


Figura A2.24: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC29.

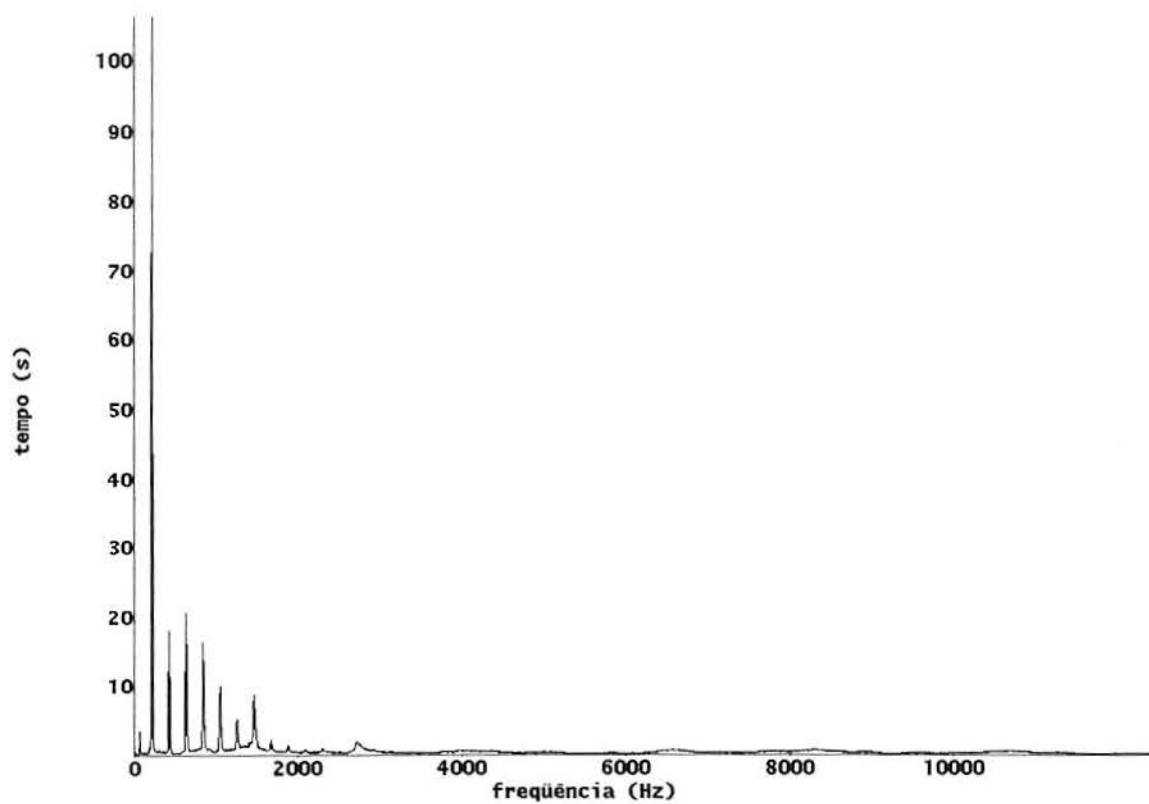


Figura A2.25: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC30.

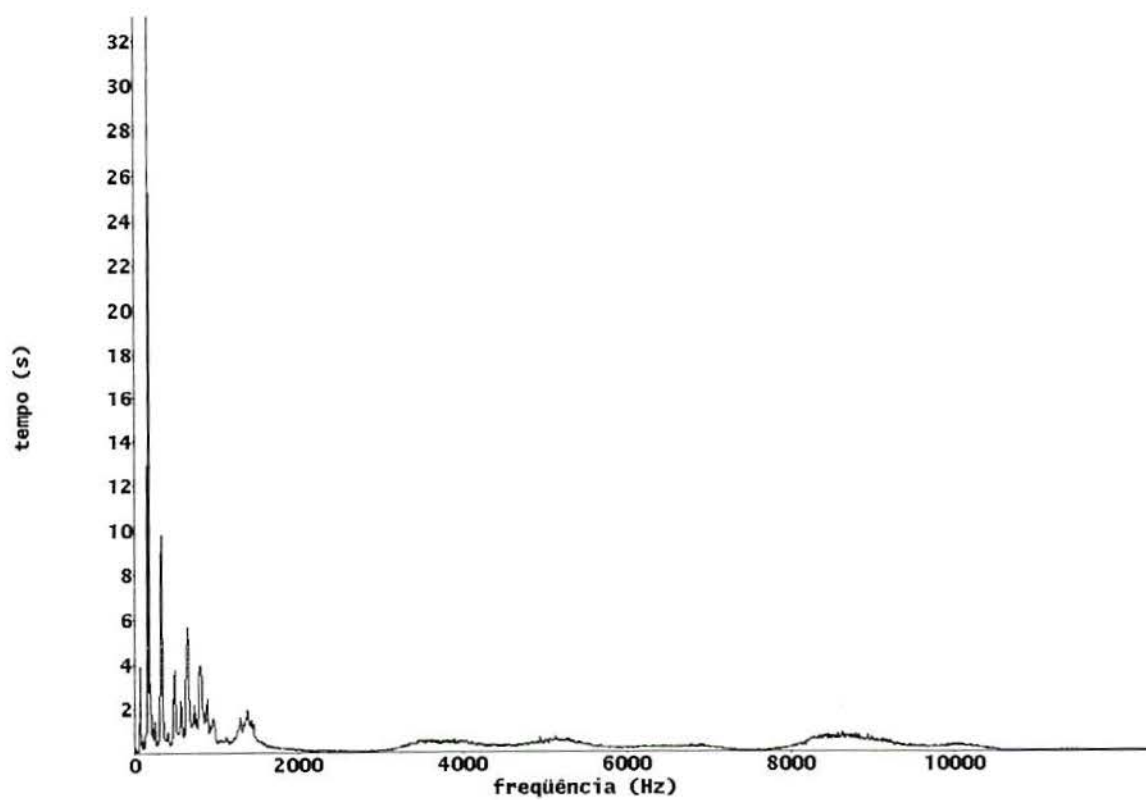


Figura A2.26: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC31.

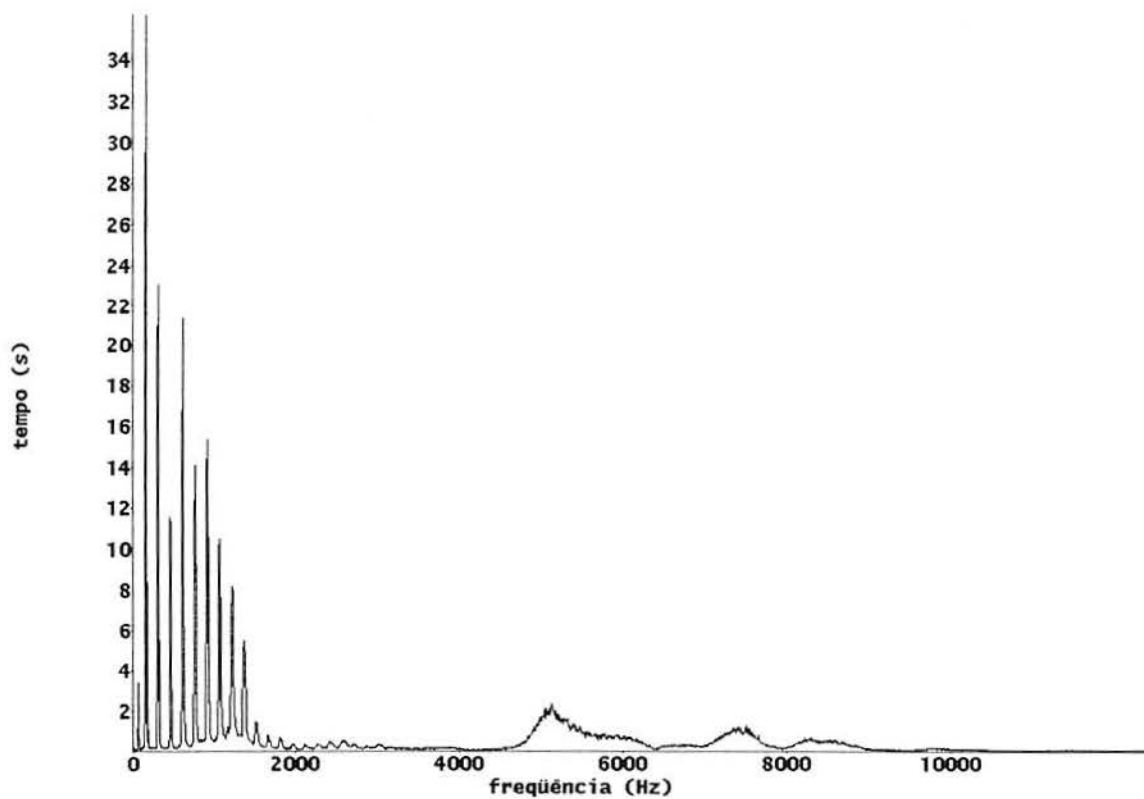


Figura A2.27: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC32.

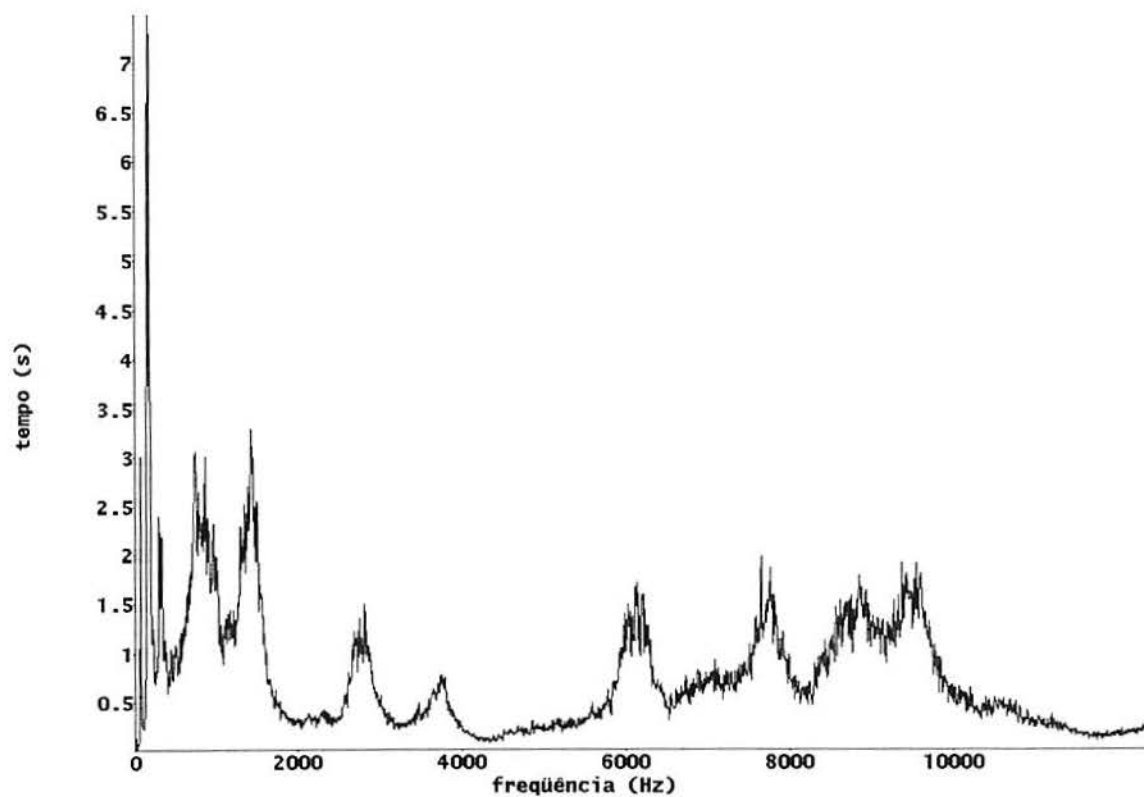


Figura A2.28: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC34.

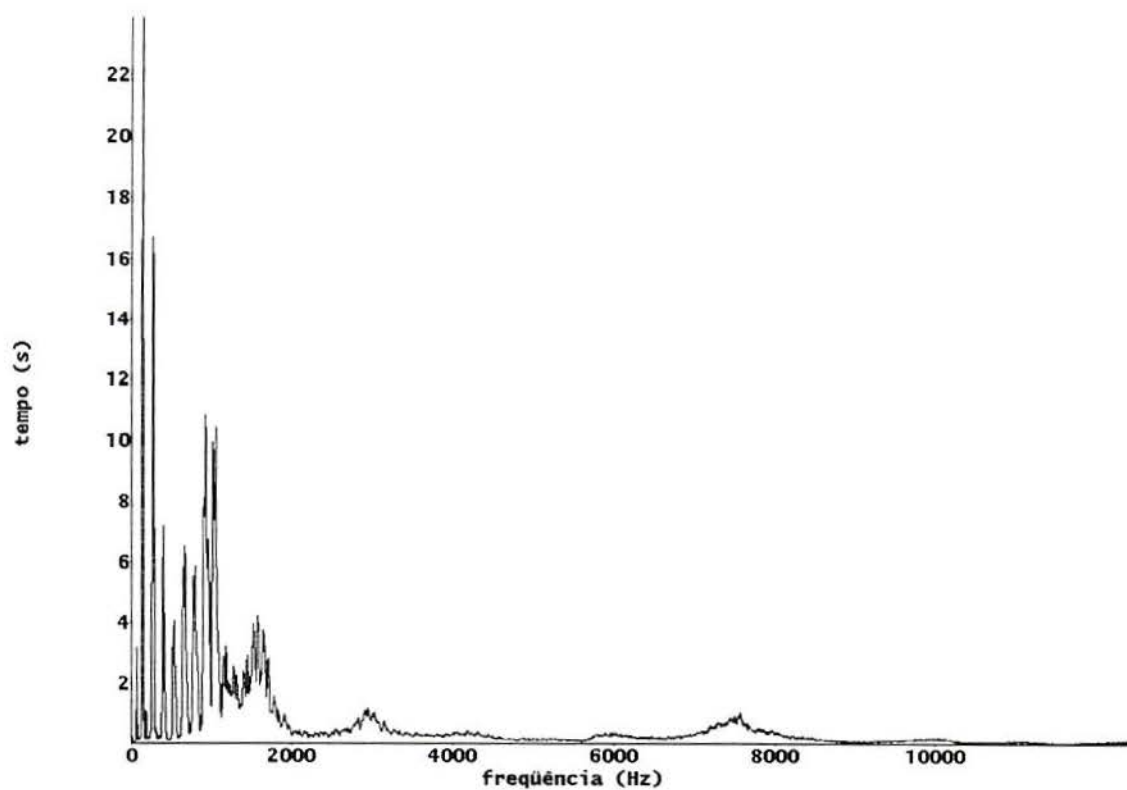


Figura A2.29: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC35.

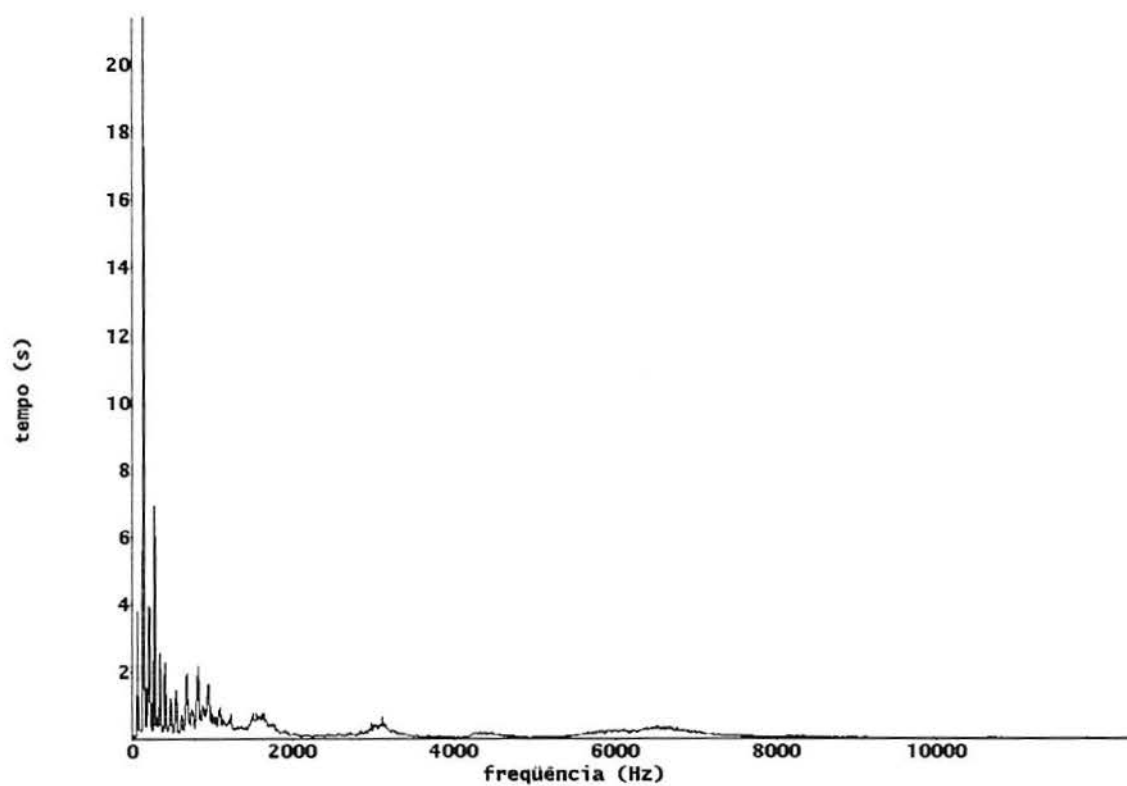


Figura A2.30: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC36.

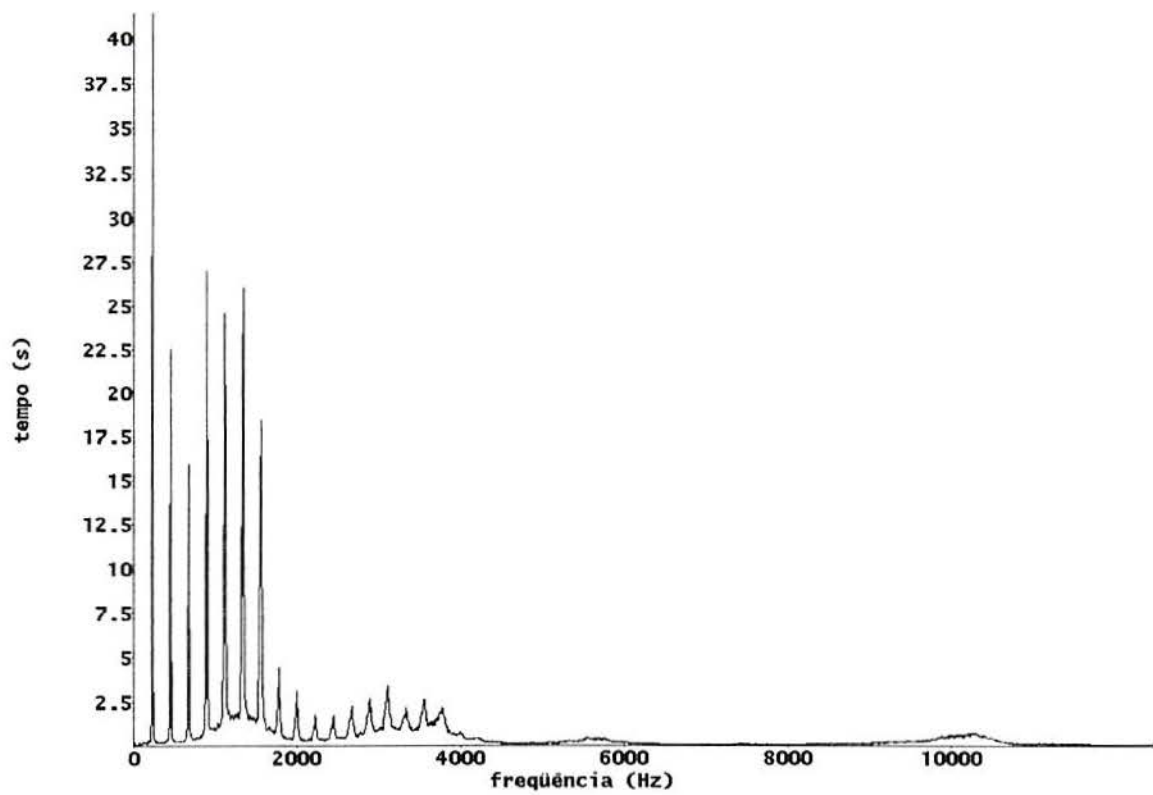


Figura A2.31: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC37.

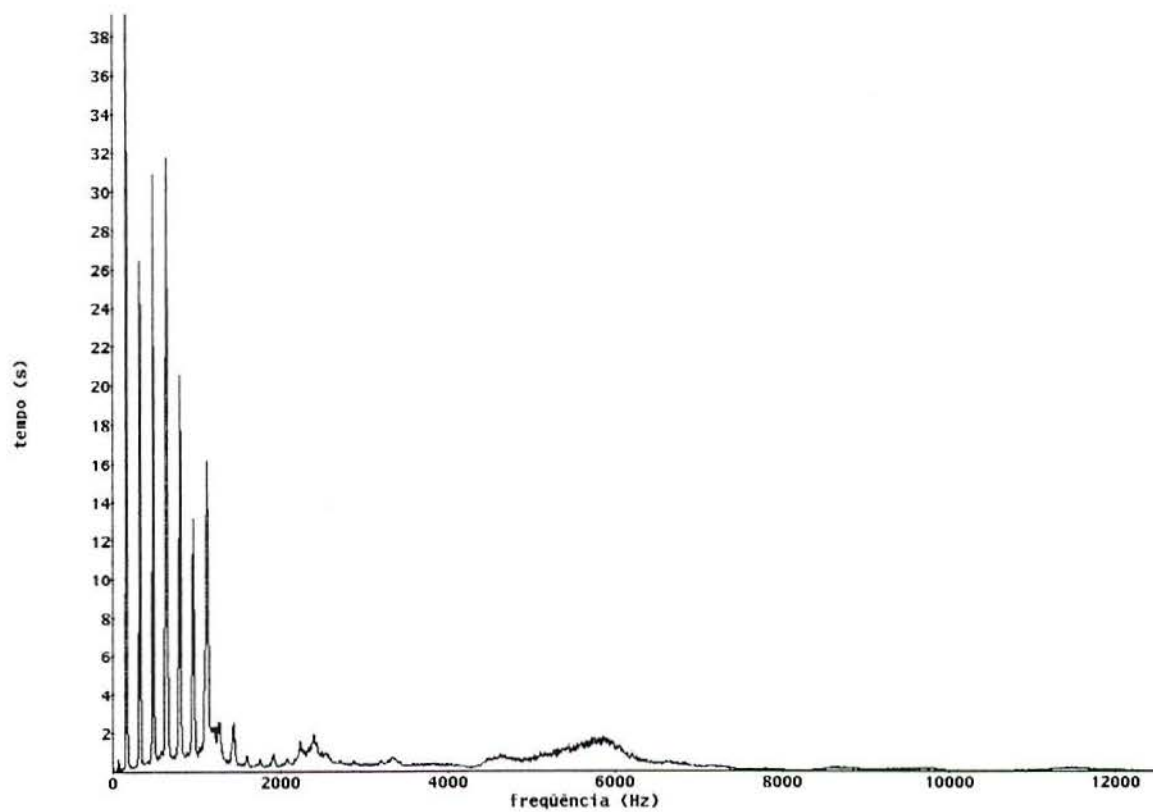


Figura A2.32: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC41.

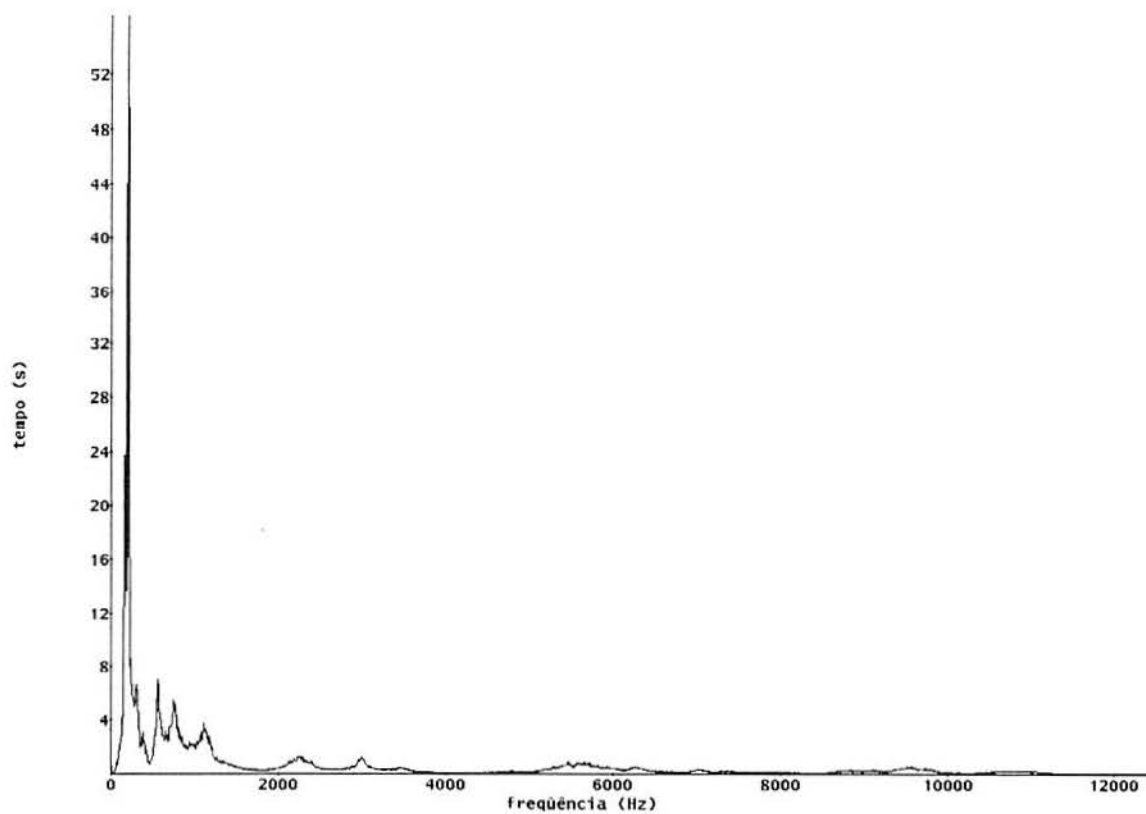


Figura A2.33: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC42.

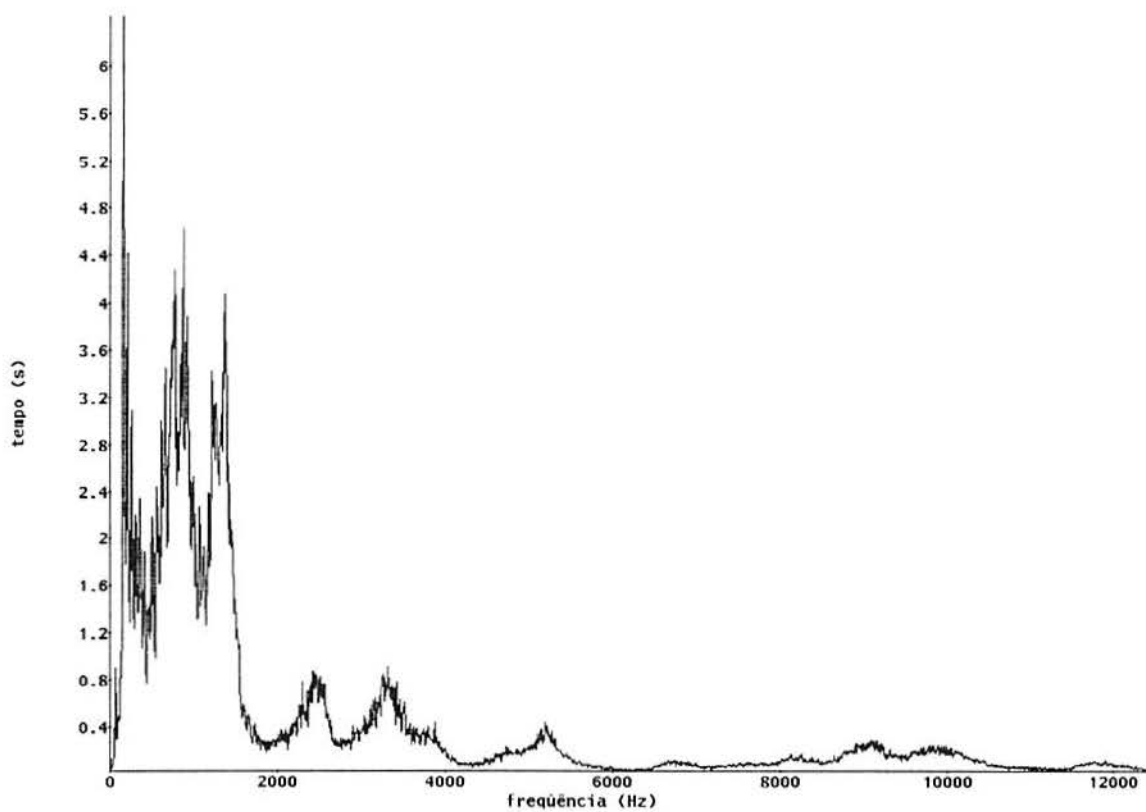


Figura A2.34: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC43.

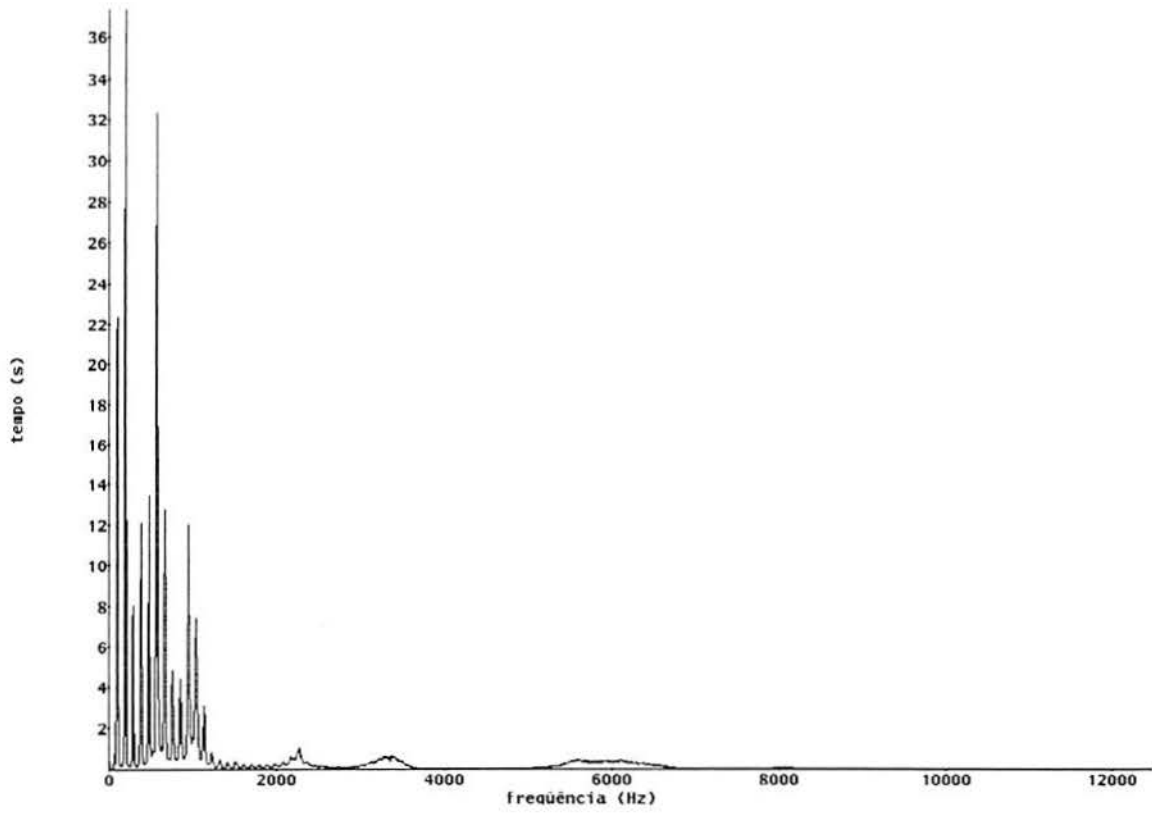


Figura A2.35: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC44.

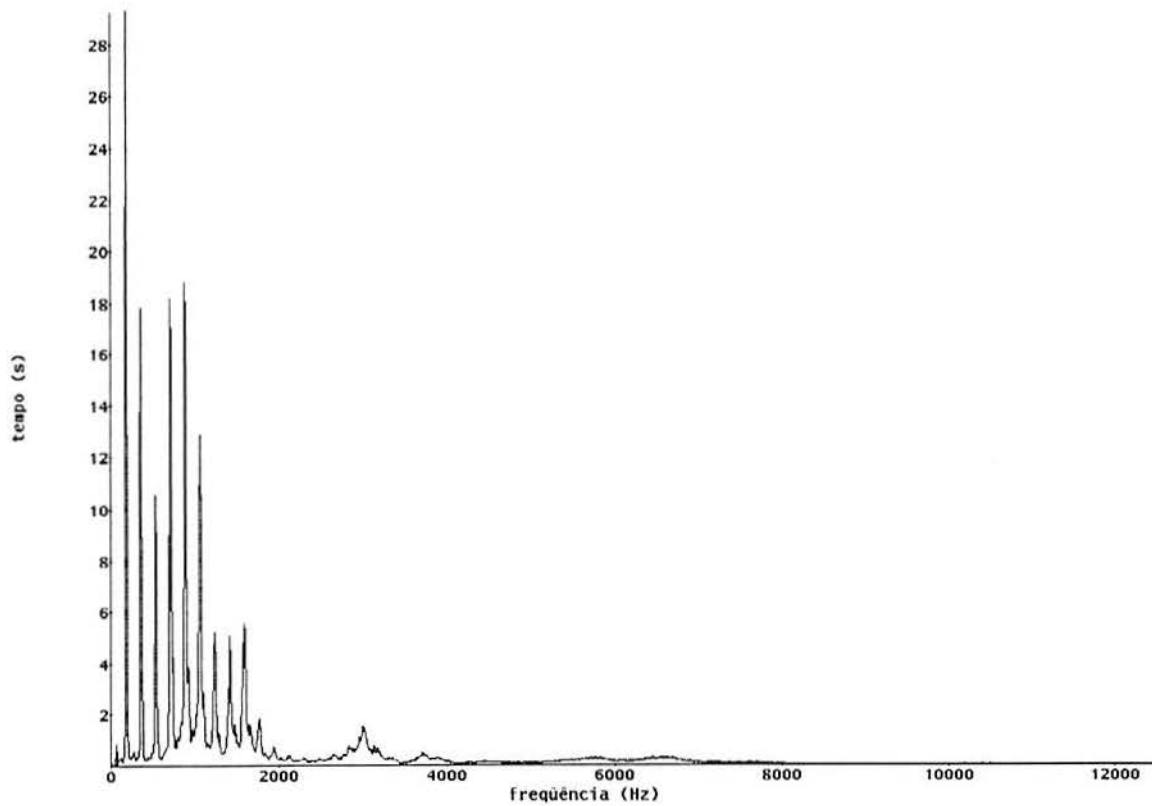


Figura A2.36: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC45.

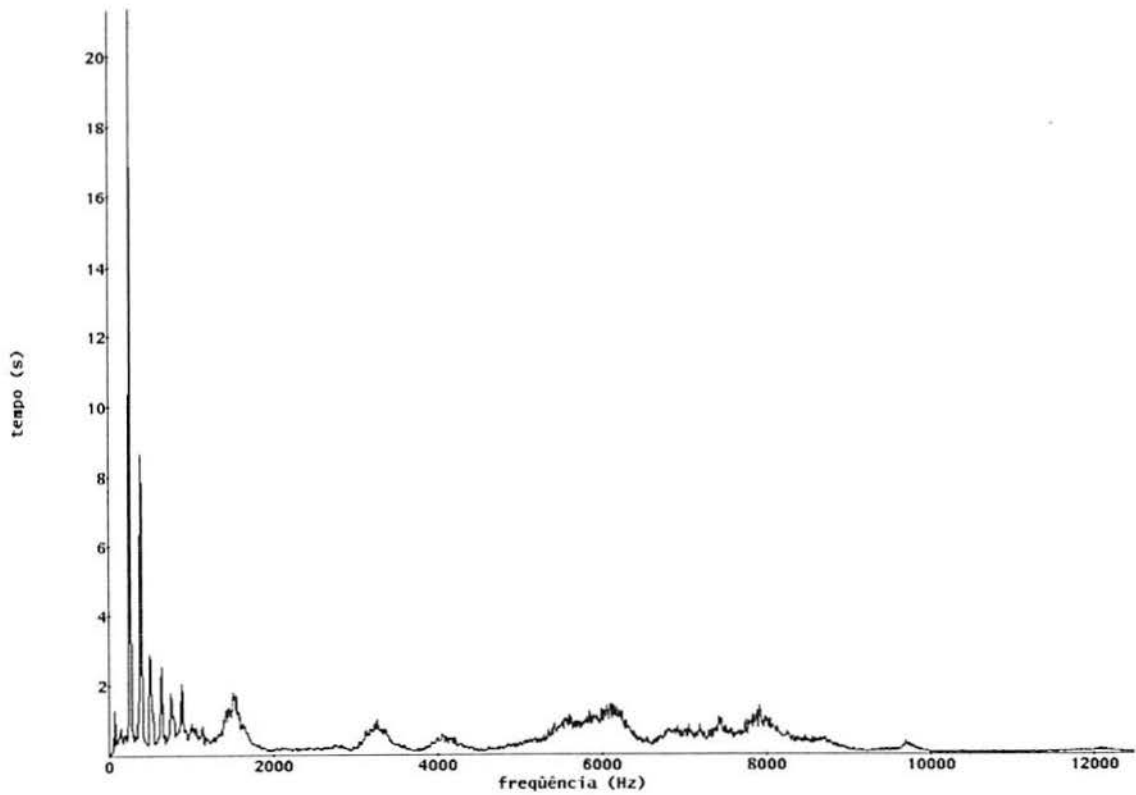


Figura A2.37: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC46.

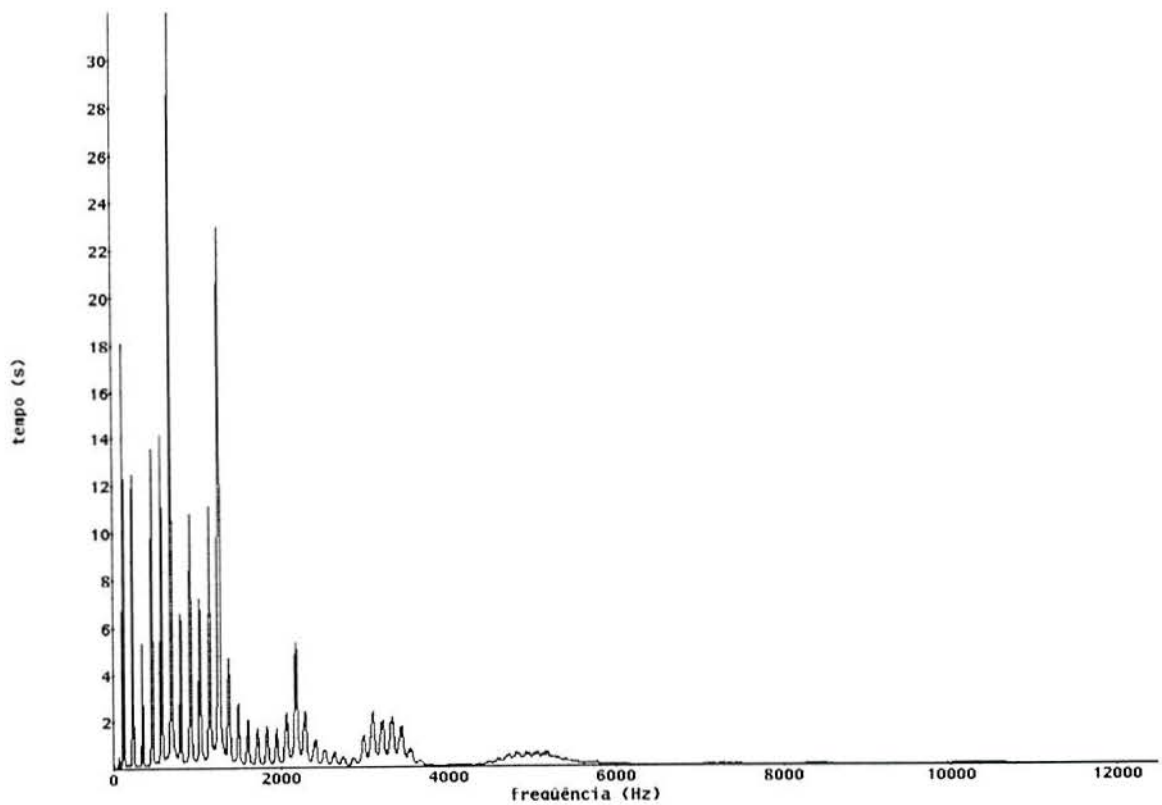


Figura A2.38: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC47.



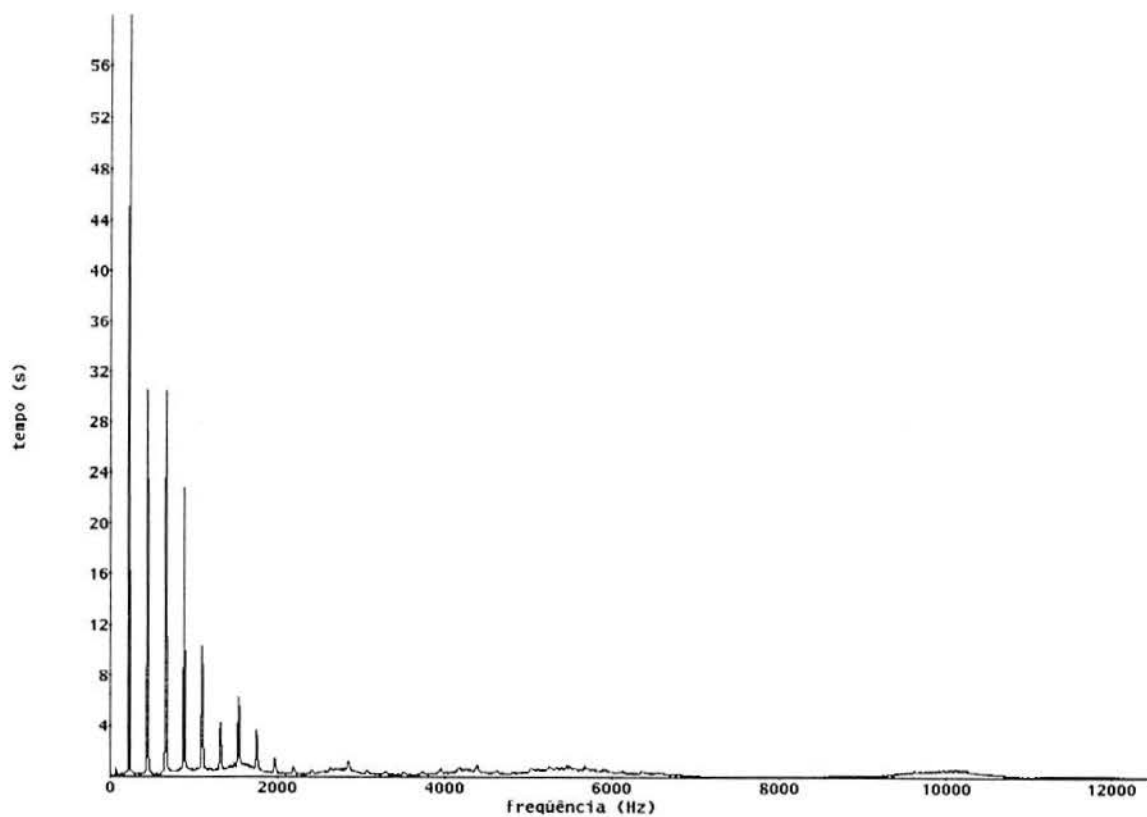


Figura A2.39: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC48.

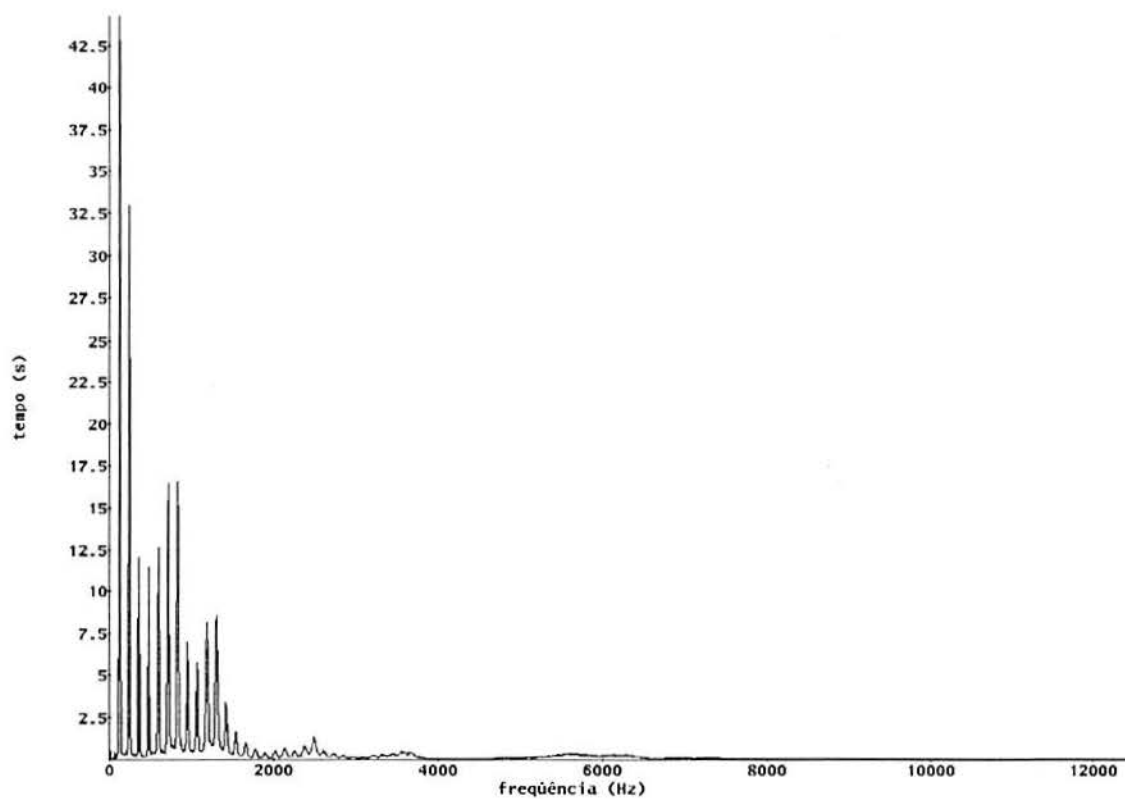


Figura A2.40: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC49.

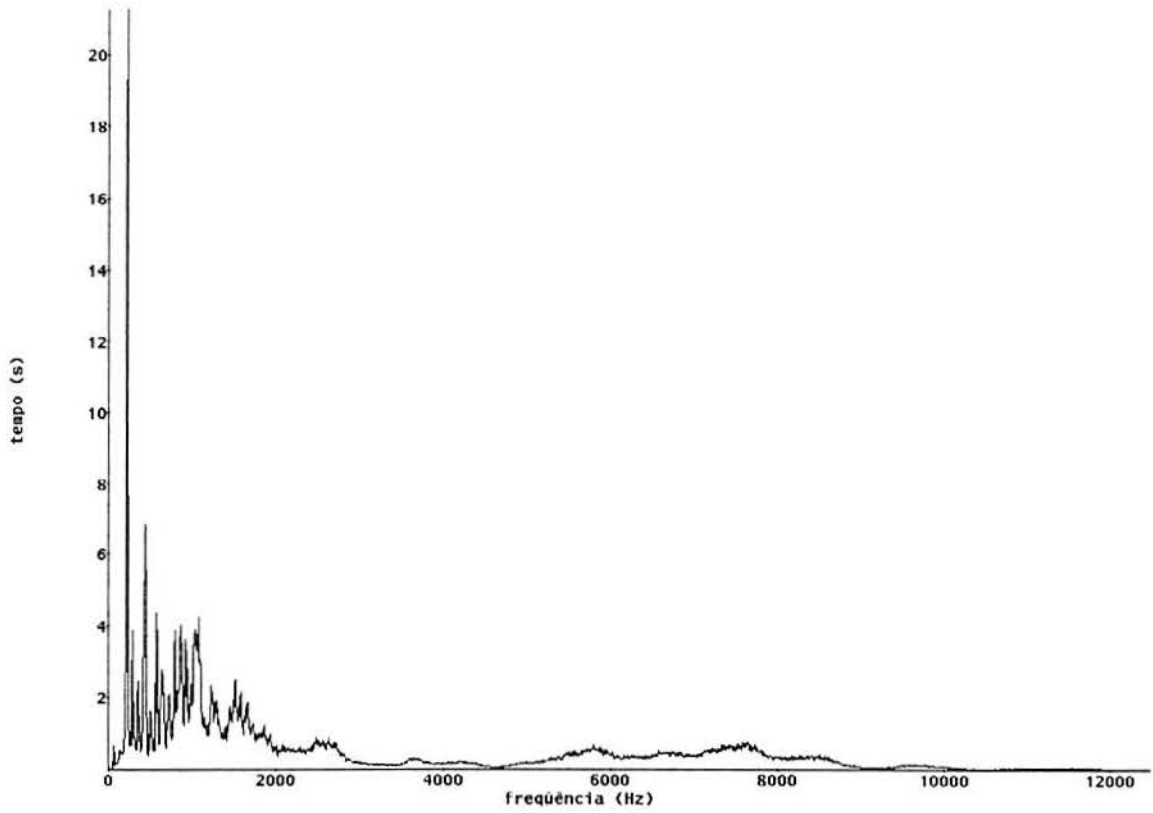


Figura A2.41: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC50.

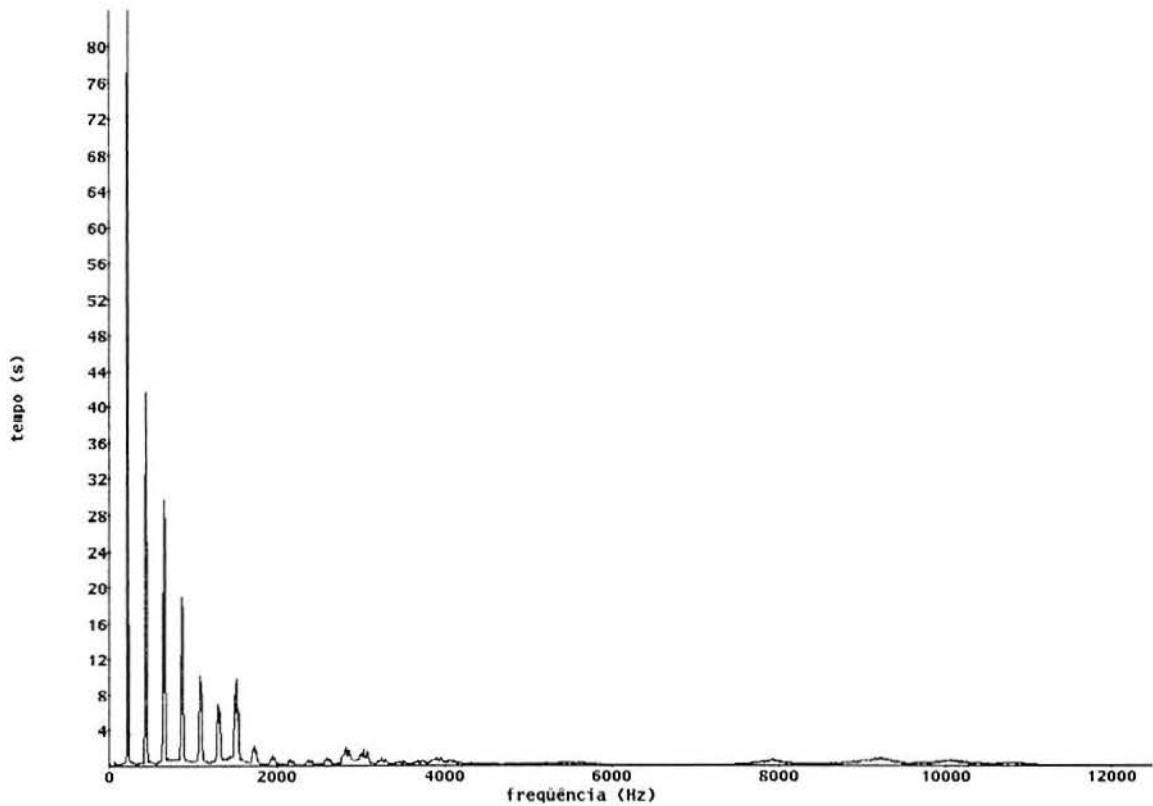


Figura A2.42: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC51.

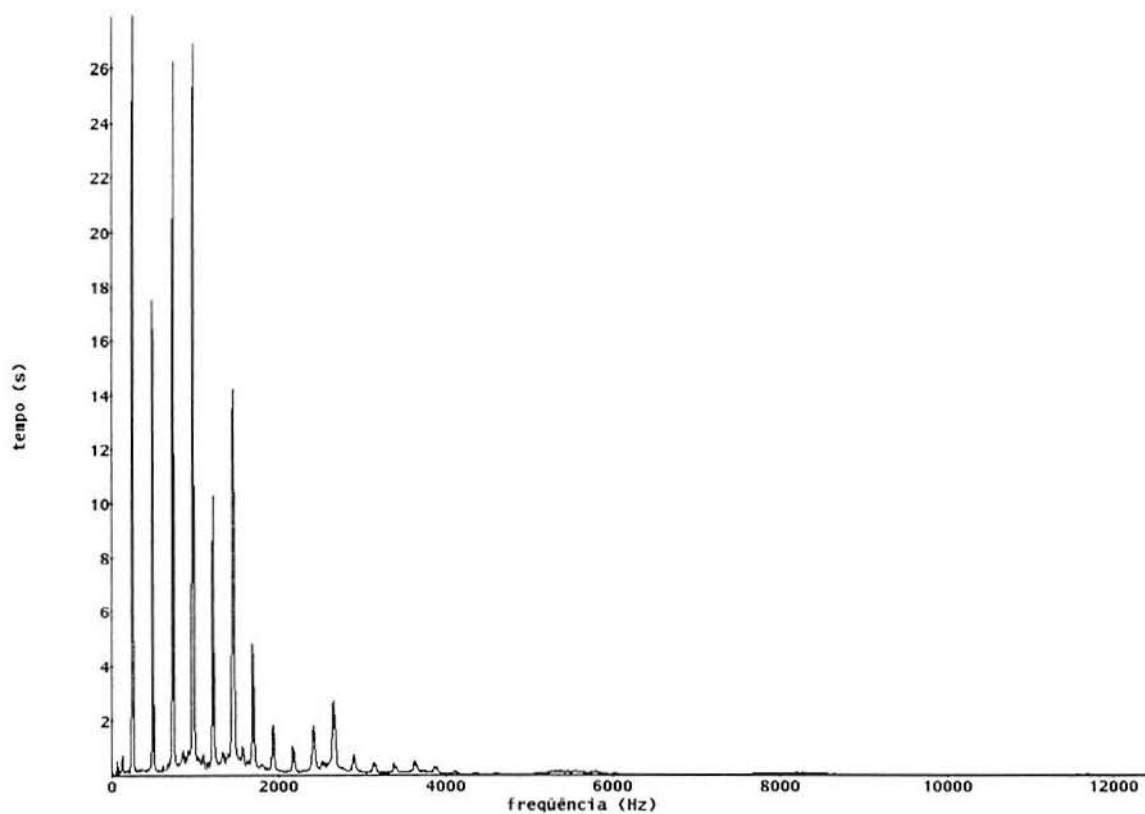


Figura A2.43: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC53.

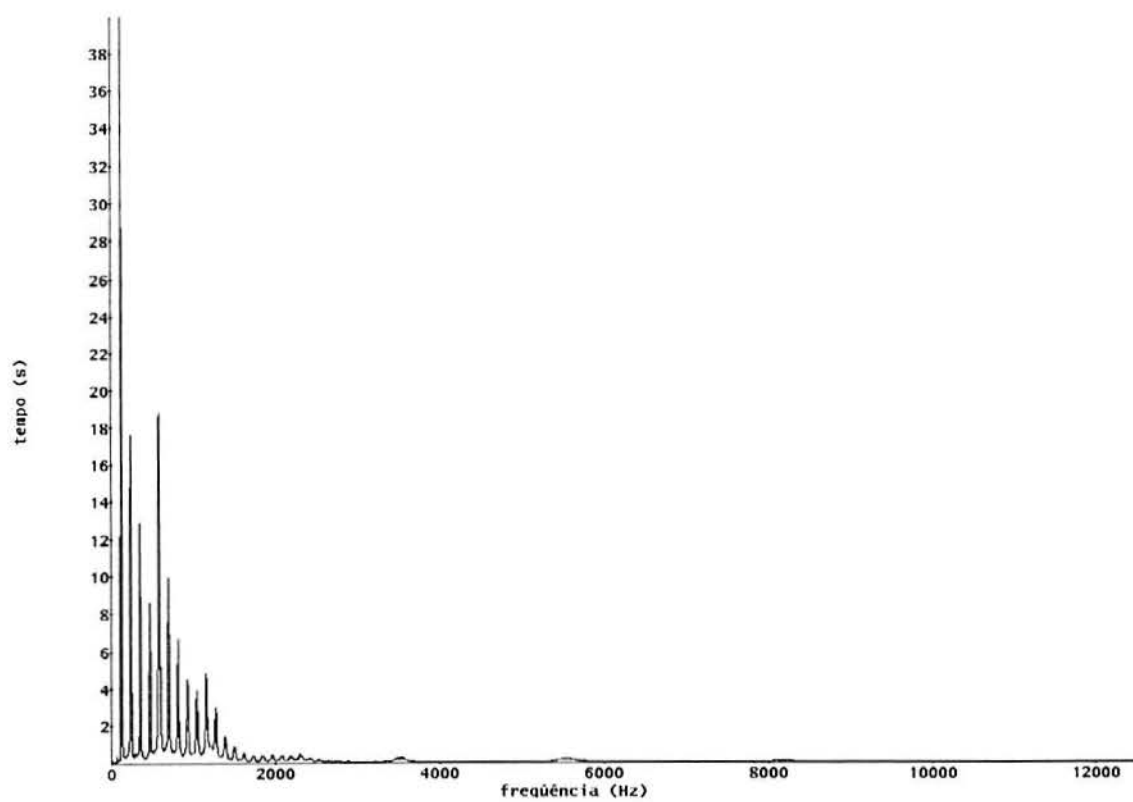


Figura A2.44: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC54.

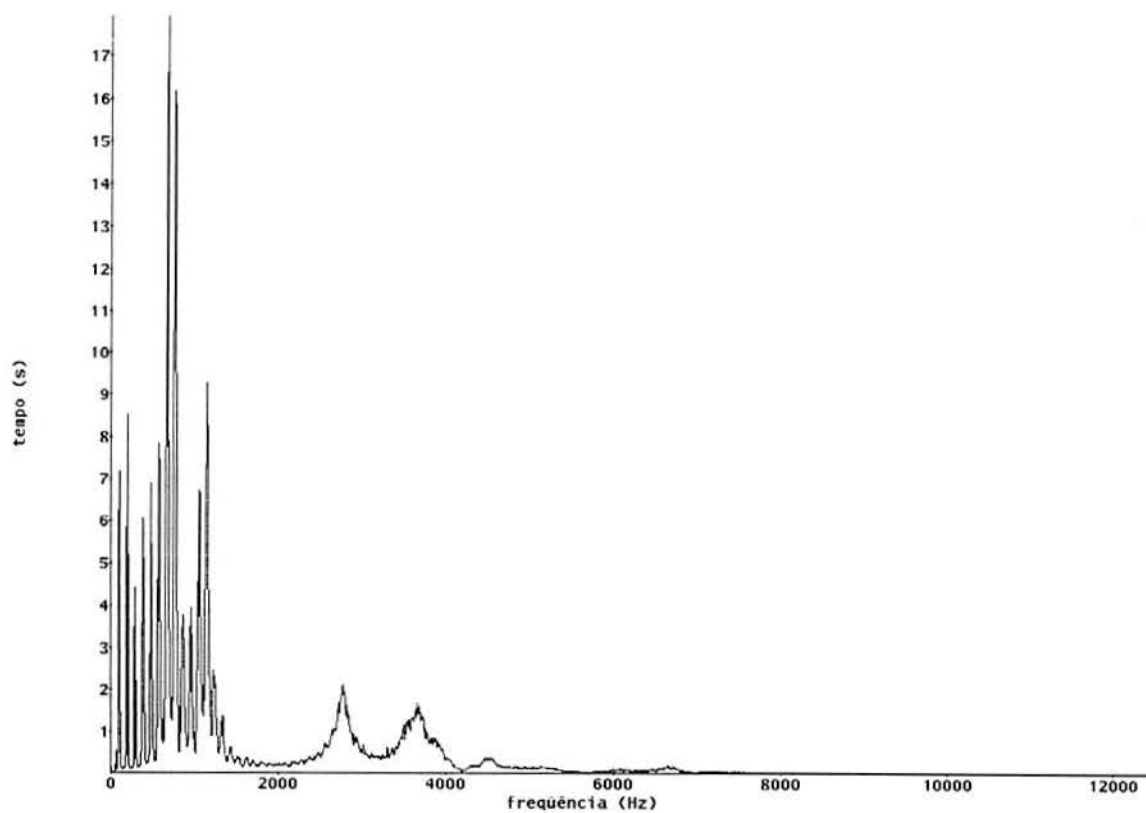


Figura A2.45: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC55.

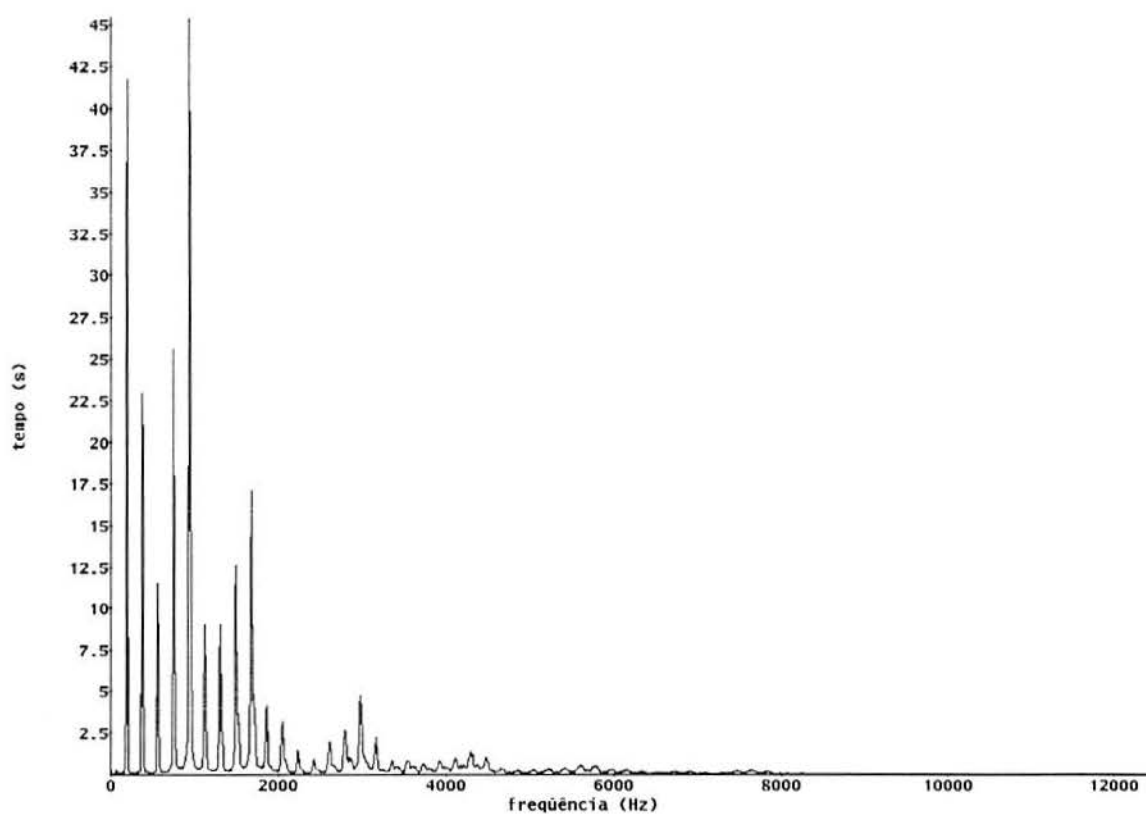


Figura A2.46: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC56.

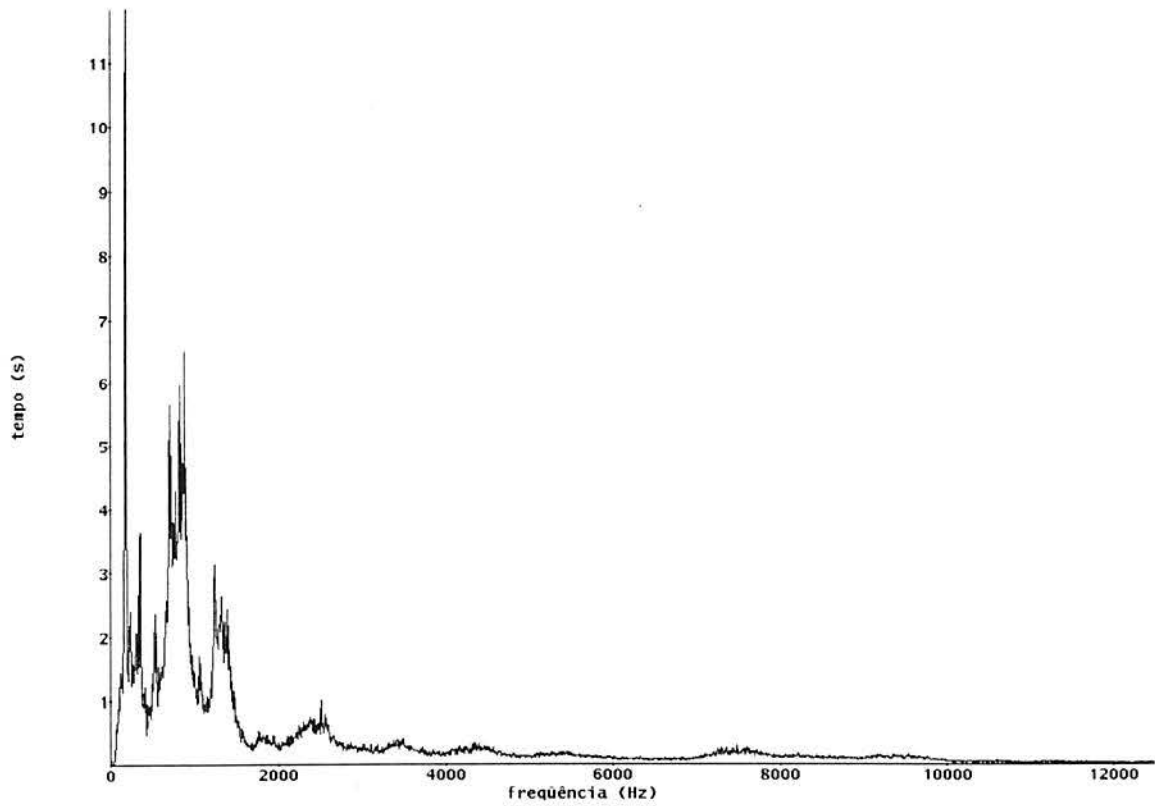


Figura A2.47: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC57.

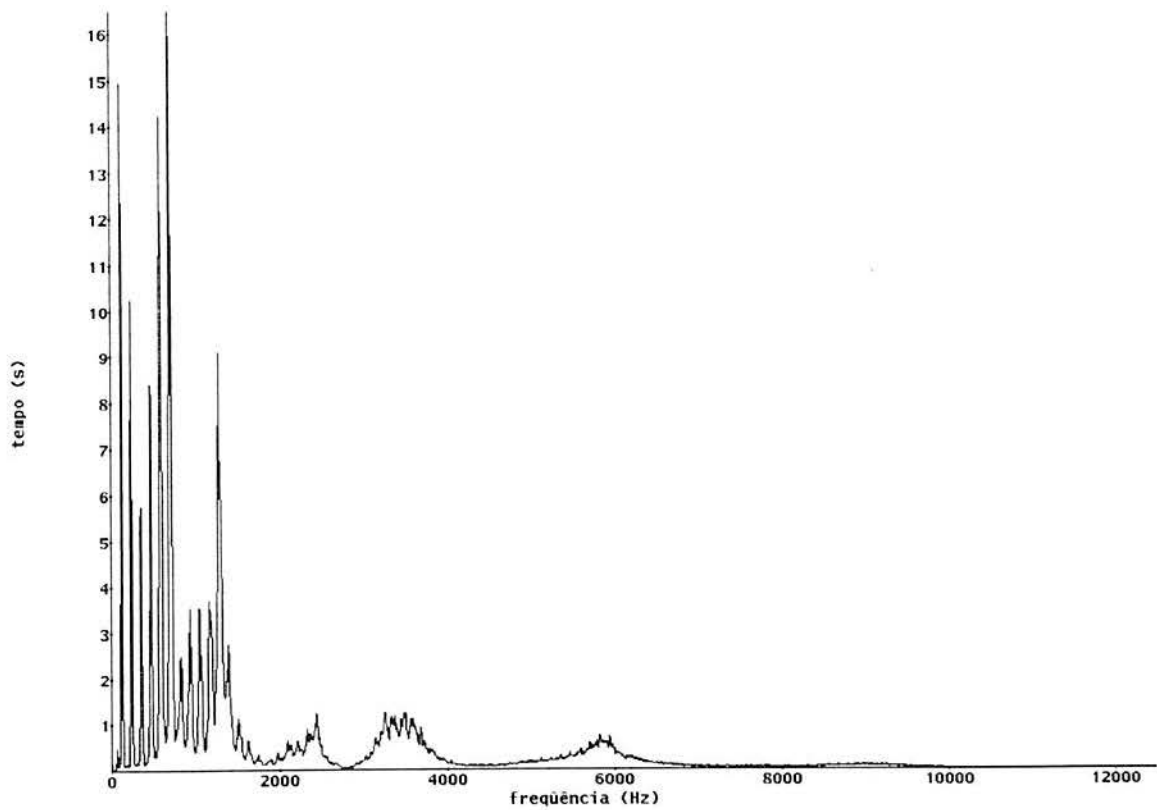


Figura A2.48: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC58.

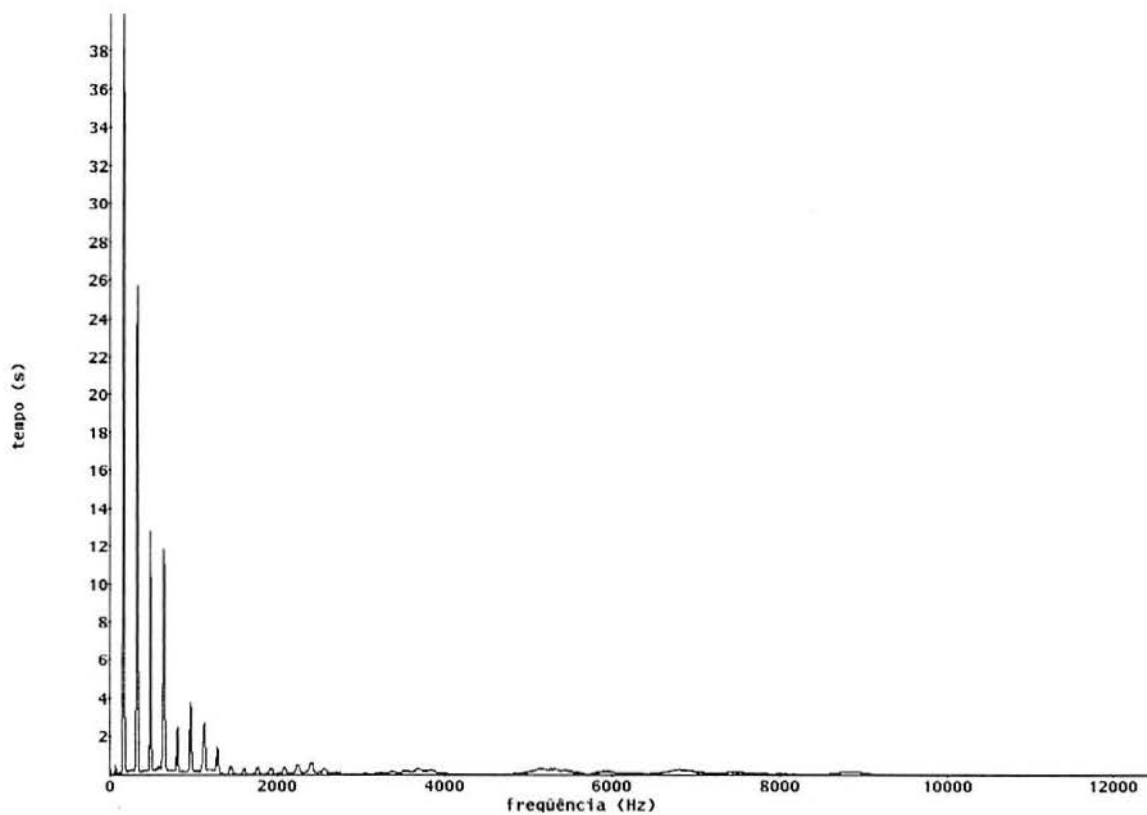


Figura A2.49: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC59.

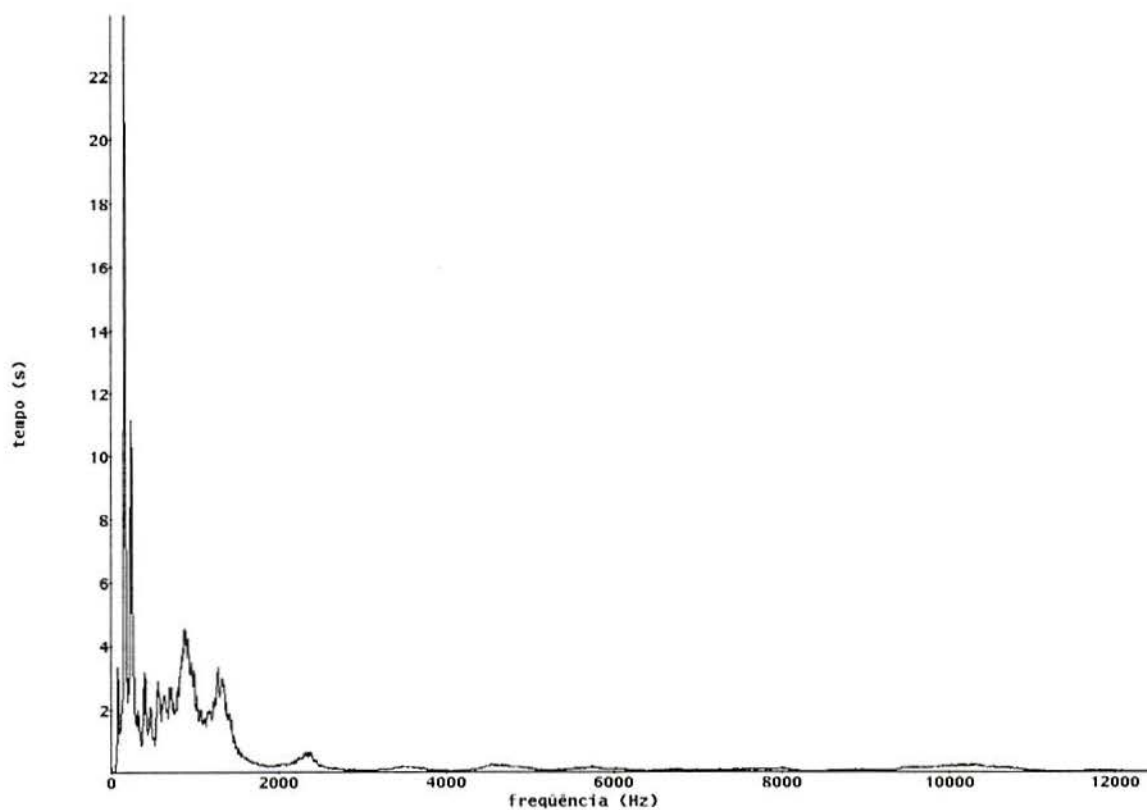


Figura A2.50: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC60.

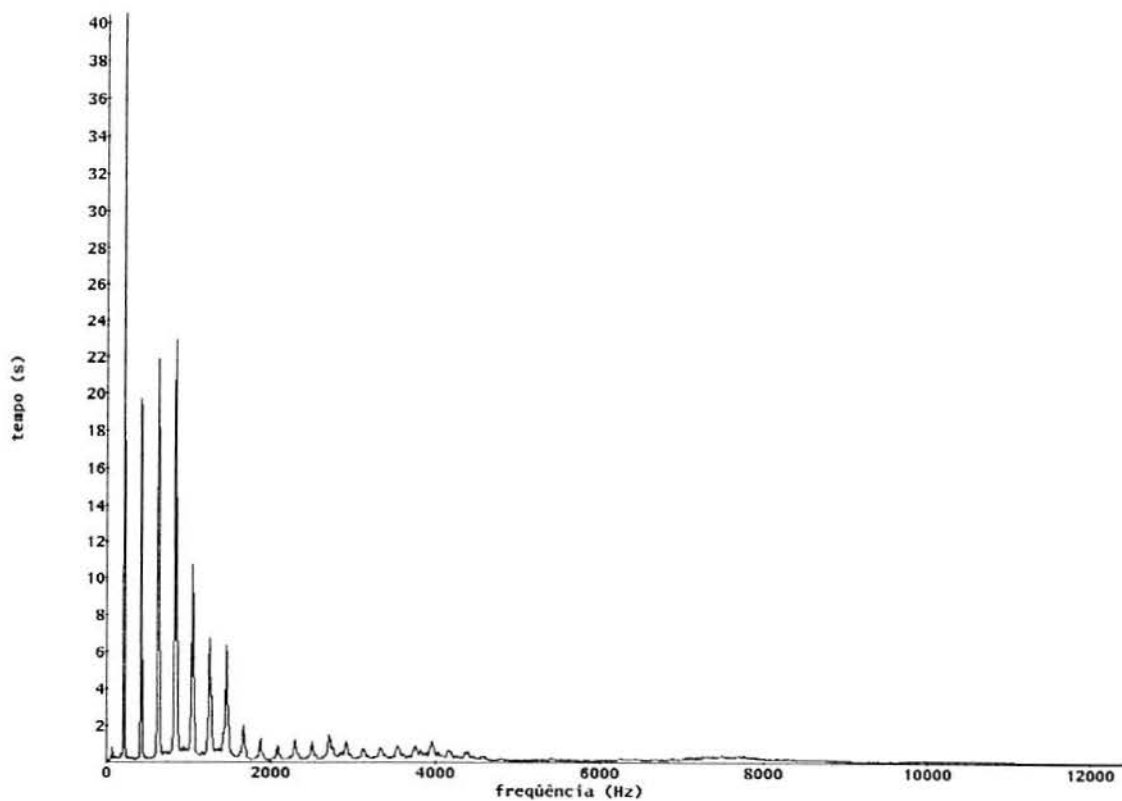


Figura A2.51: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC61.

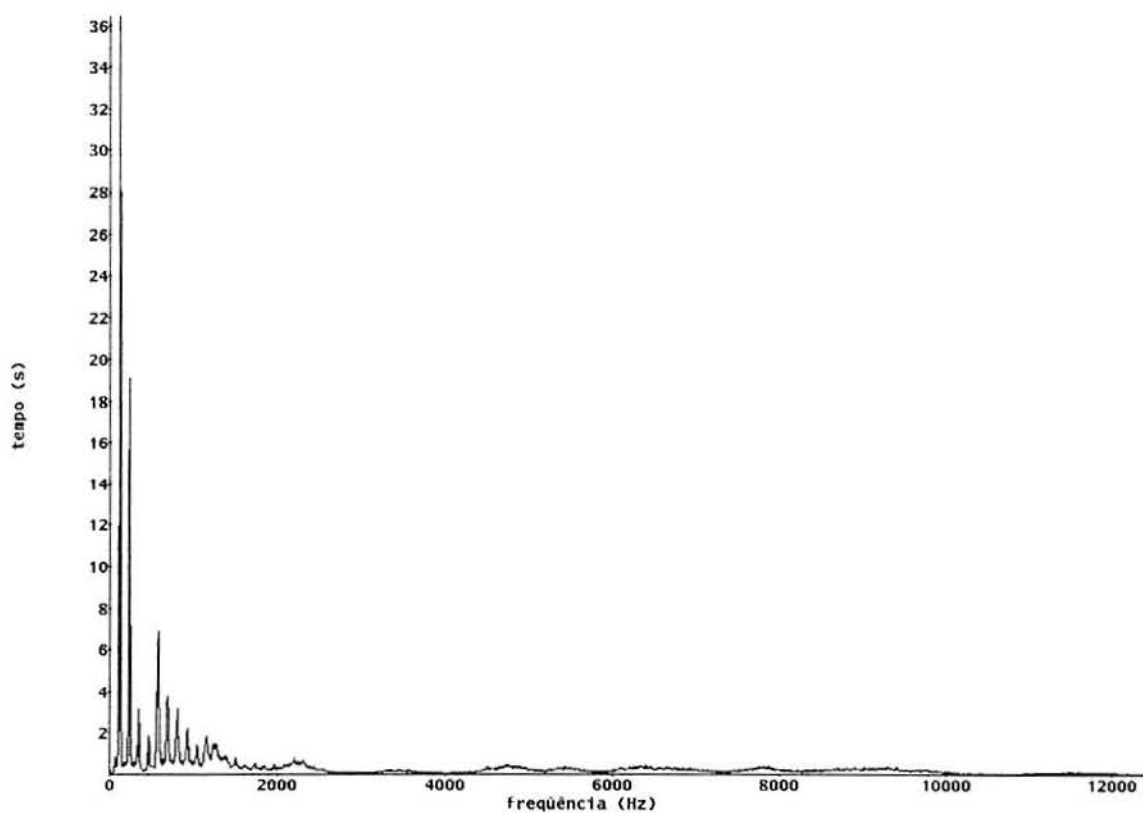


Figura A2.52: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC62.

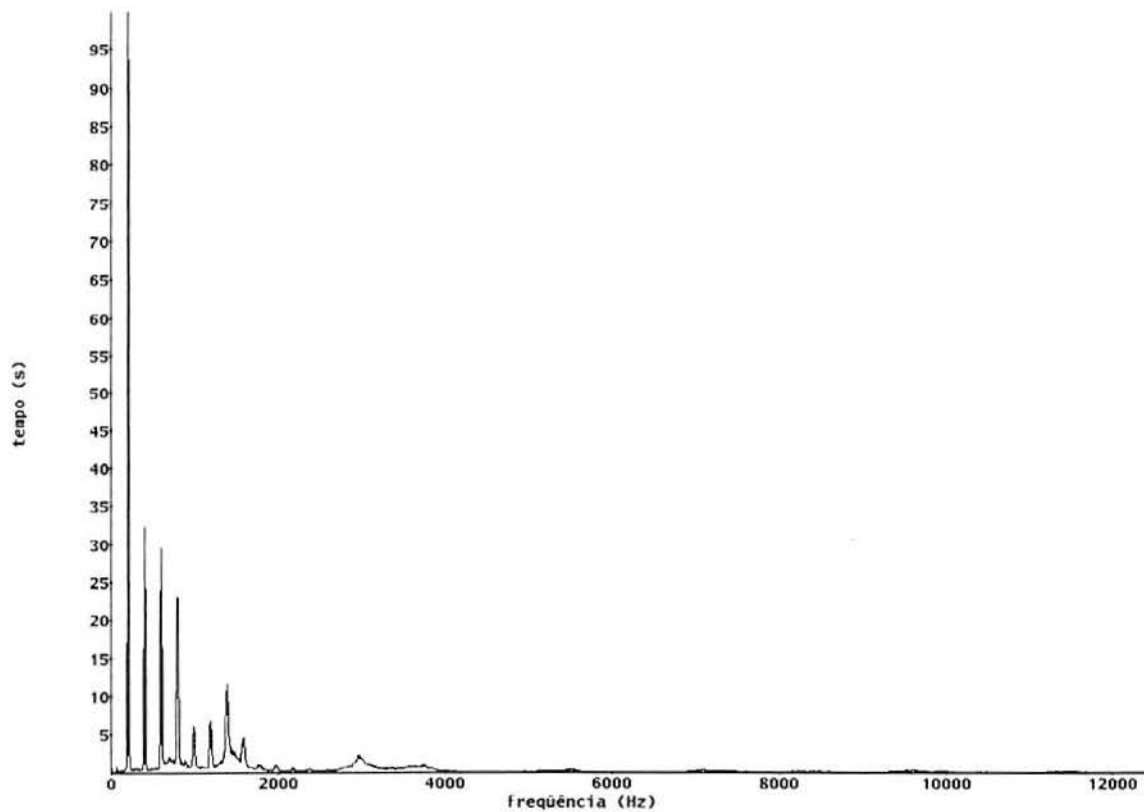


Figura A2.53: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC63.

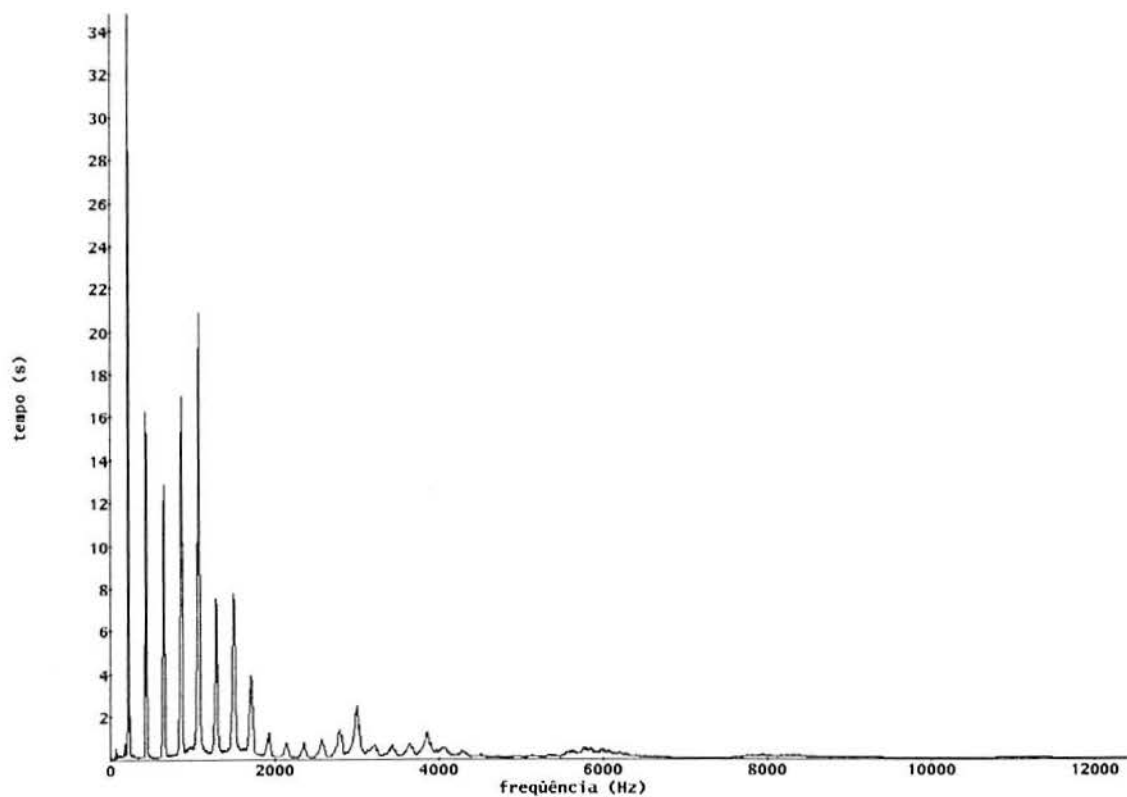


Figura A2.54: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC64.



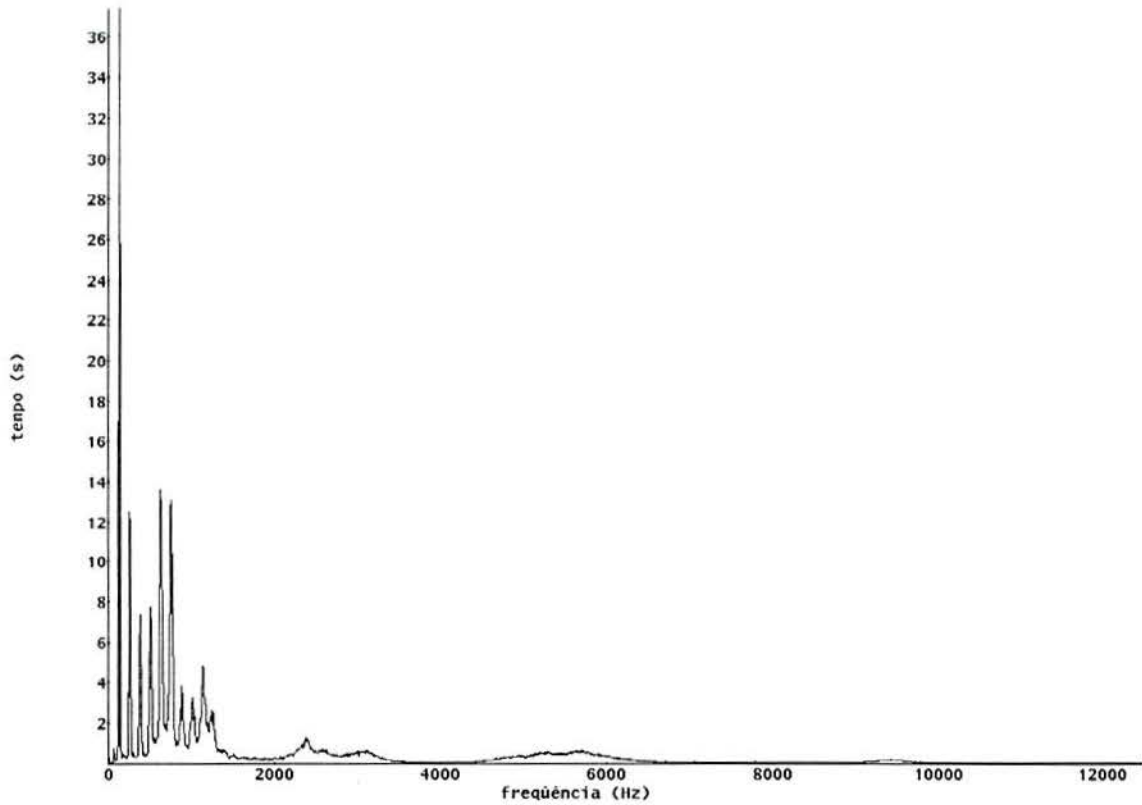


Figura A2.55: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC66.

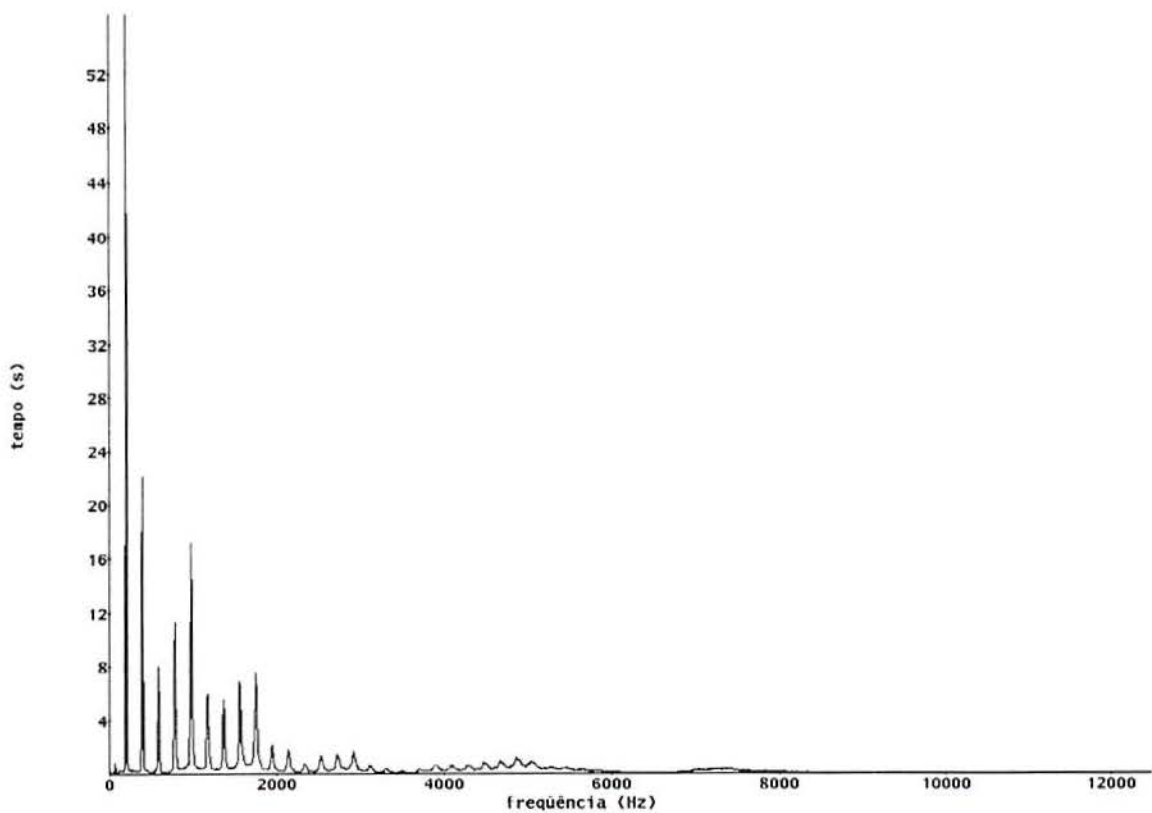


Figura A2.56: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC67.

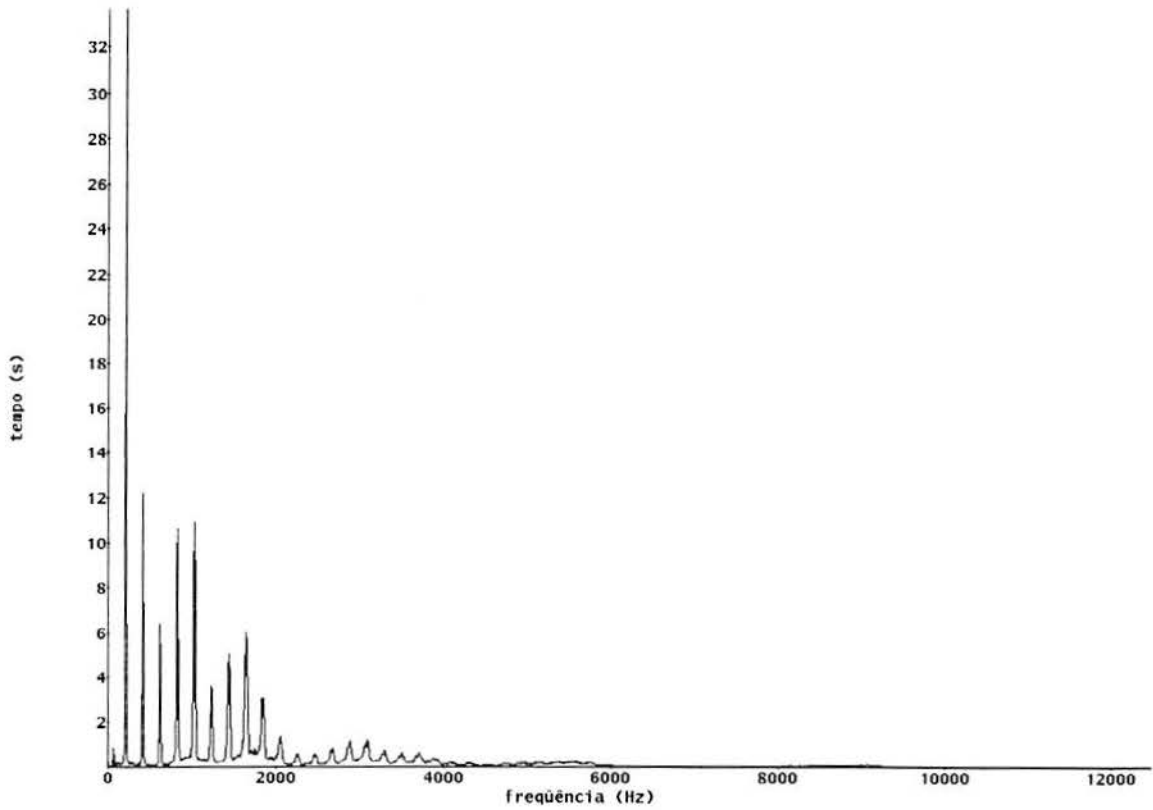


Figura A2.57: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC68.

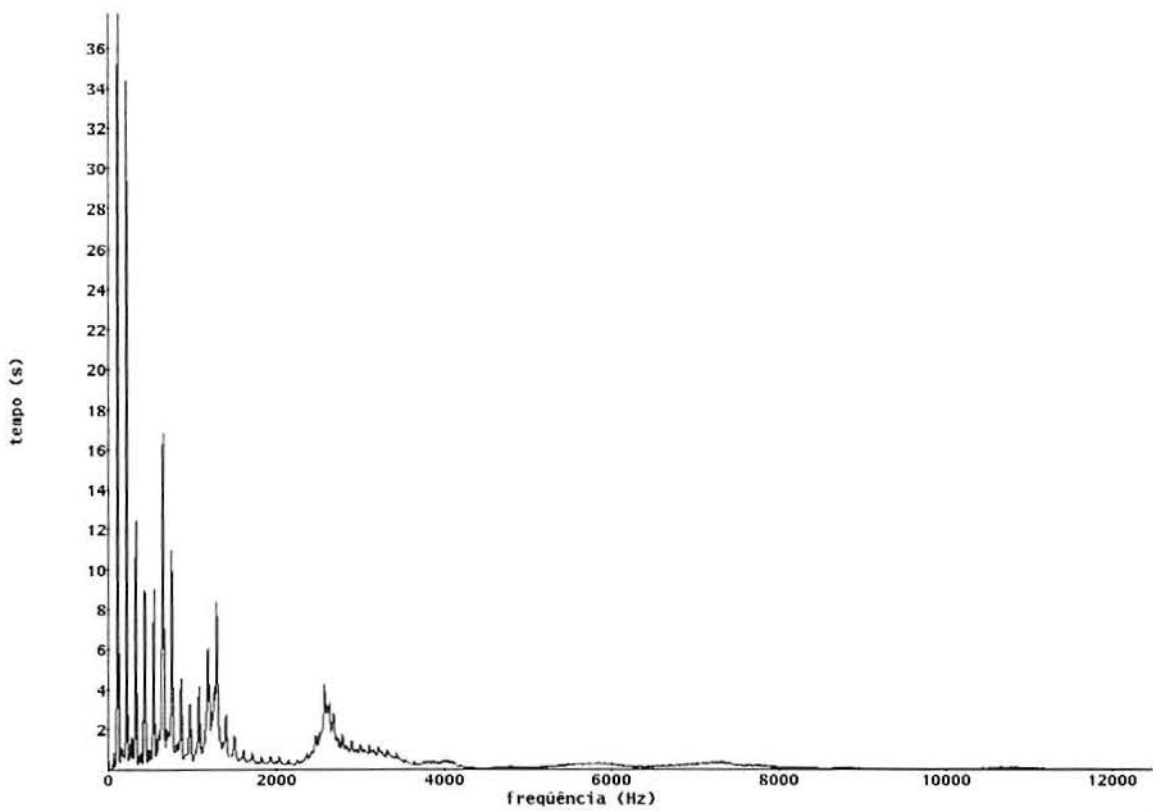


Figura A2.58: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC69.

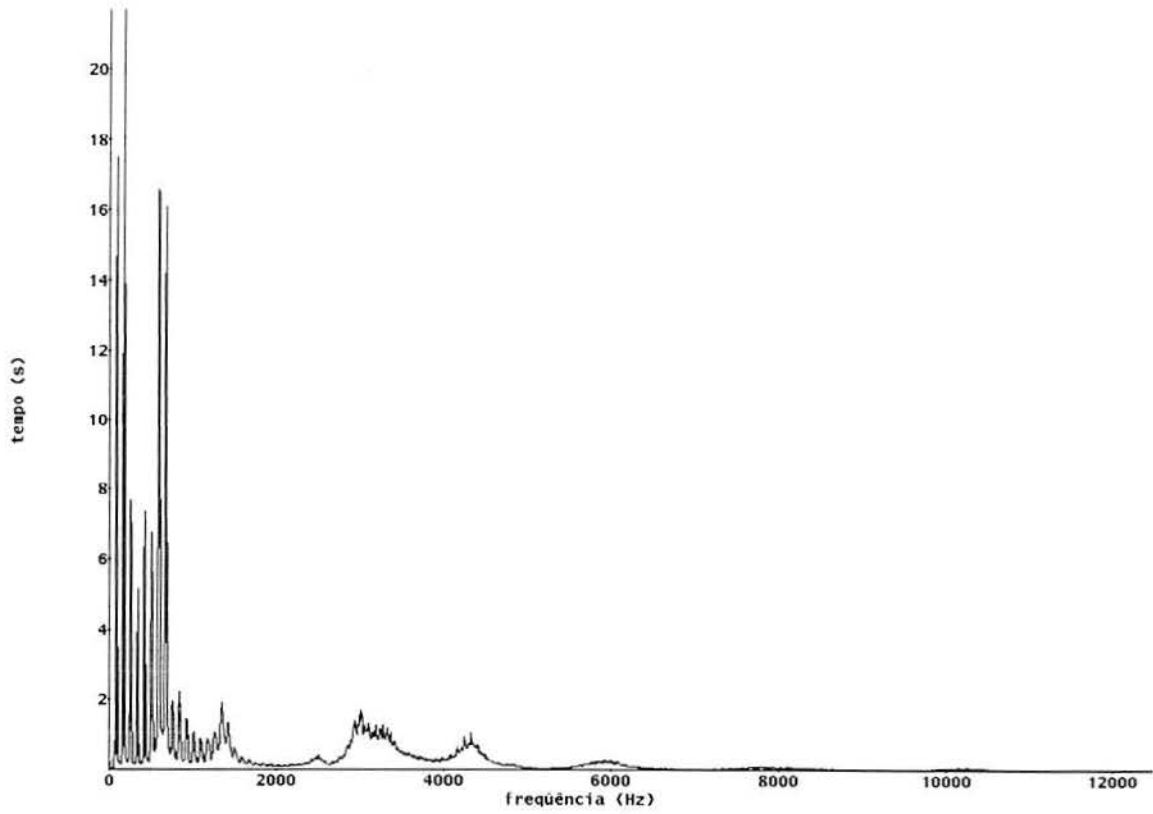


Figura A2.59: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC70.

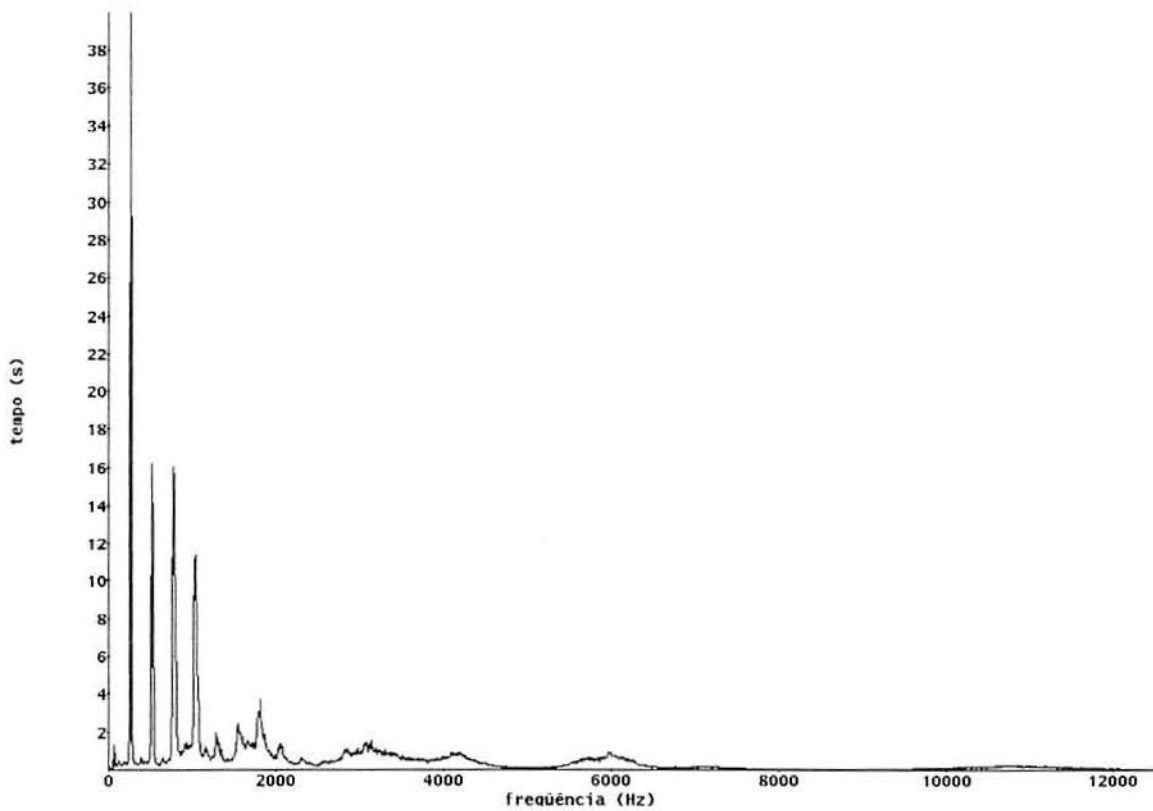


Figura A2.60: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC71.

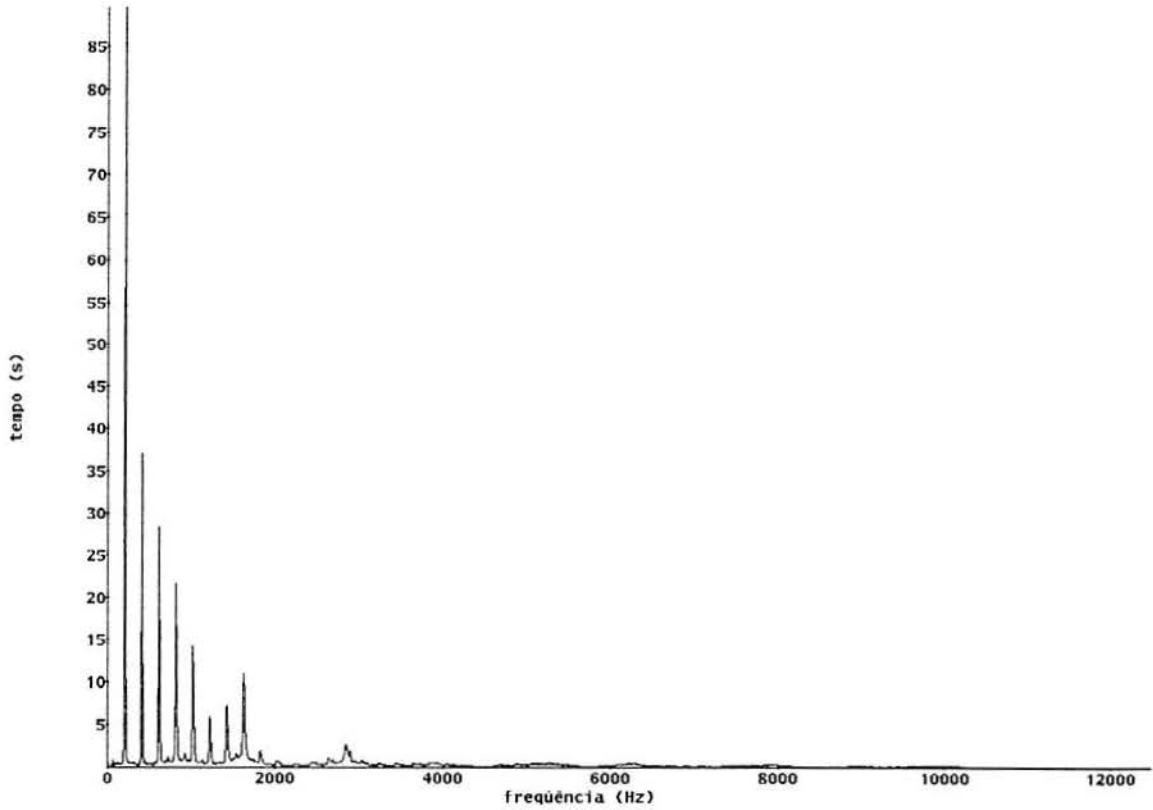


Figura A2.61: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC72.

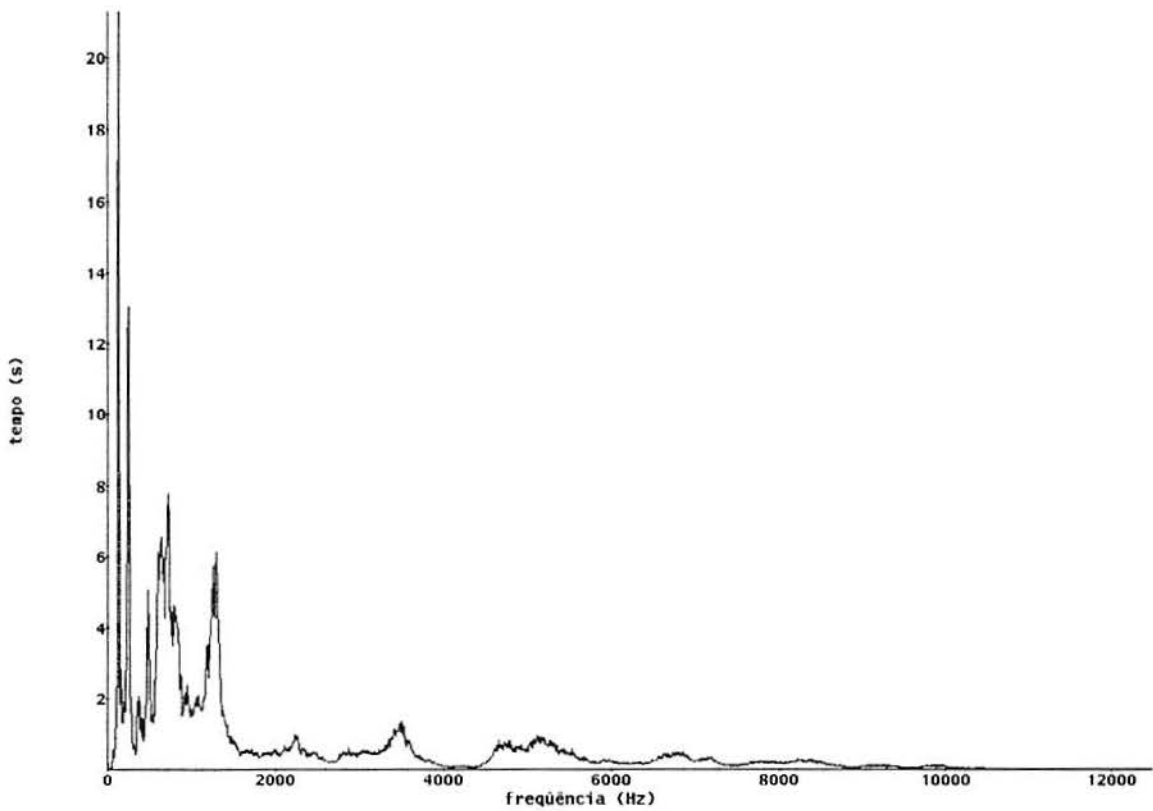
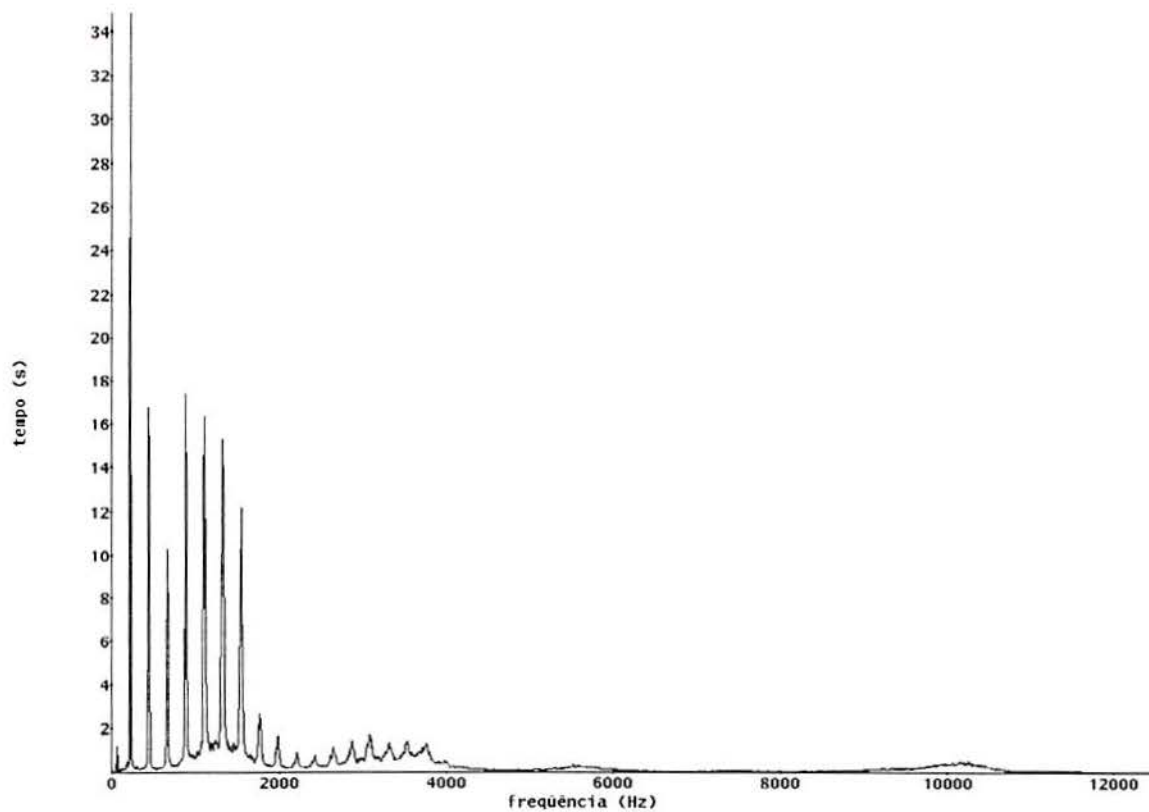


Figura A2.62: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC73.



*Figura A2.63: Gráficos do módulo da DFT do trecho de 16384 pontos do sinal do paciente PAC74.*

ANEXO 3: Sinais e suas DWTs dos pacientes do segundo procedimento de aquisição:

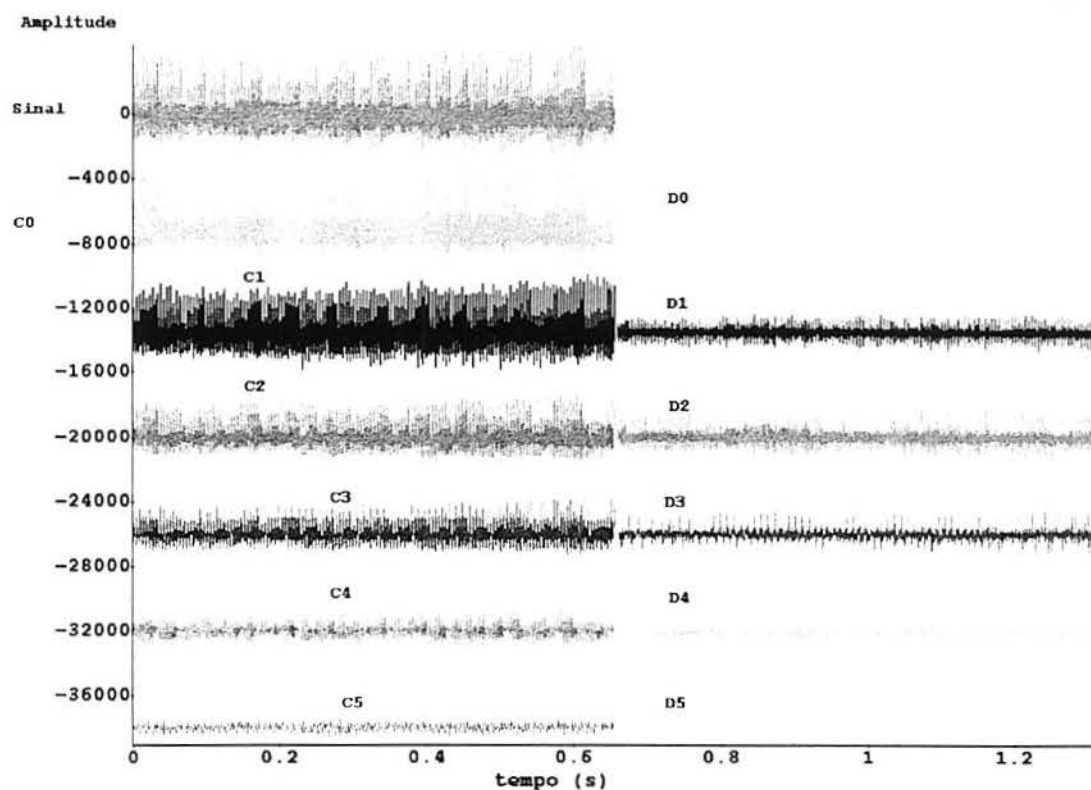


Figura A3.1: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC01.

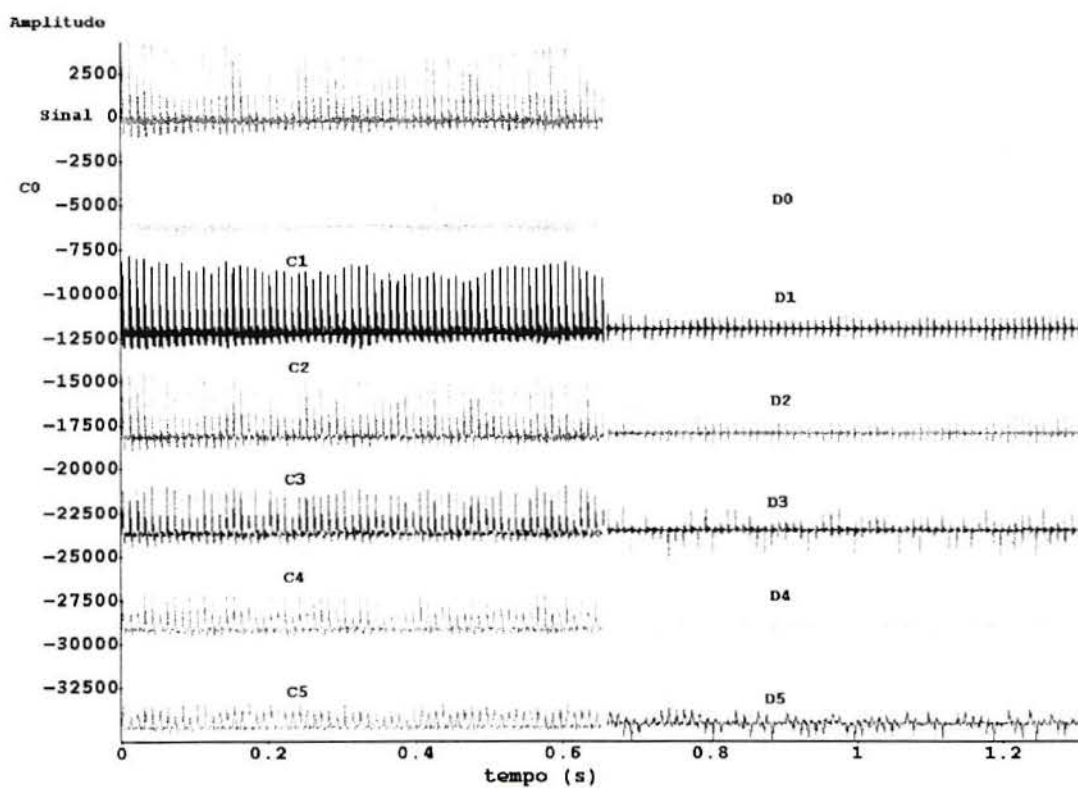


Figura A3.2: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC02.

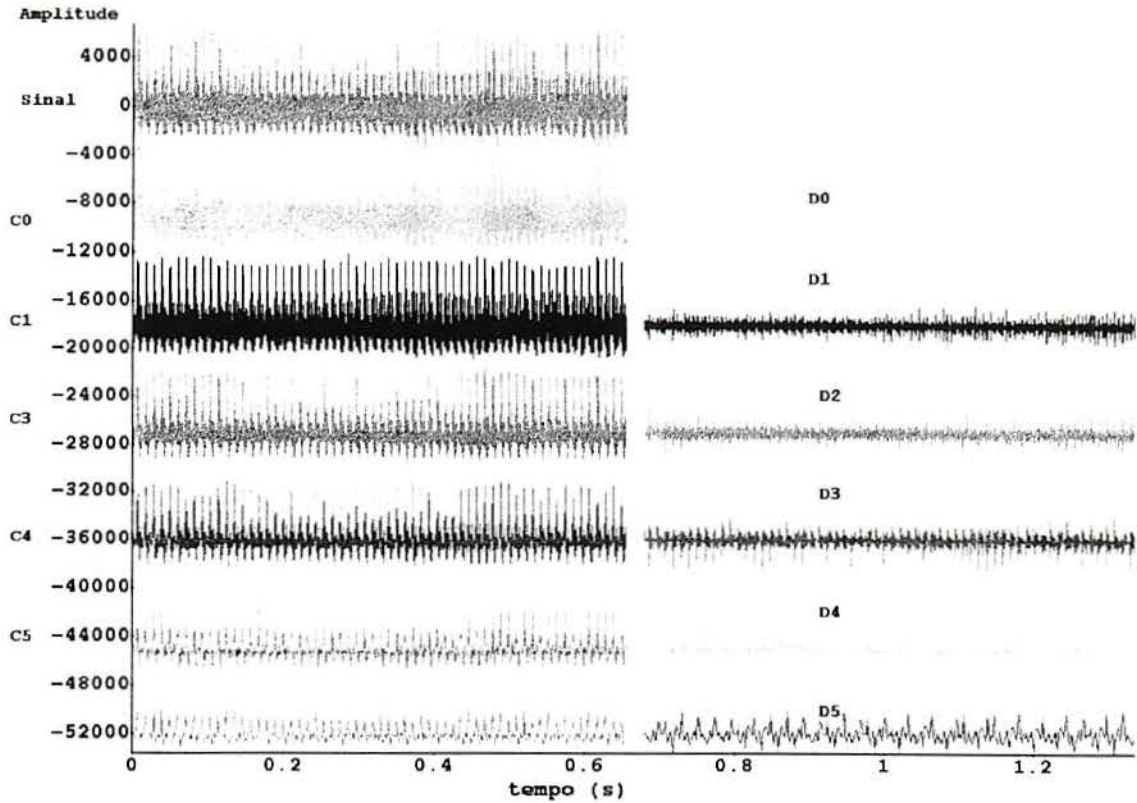


Figura A3.3: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC04.

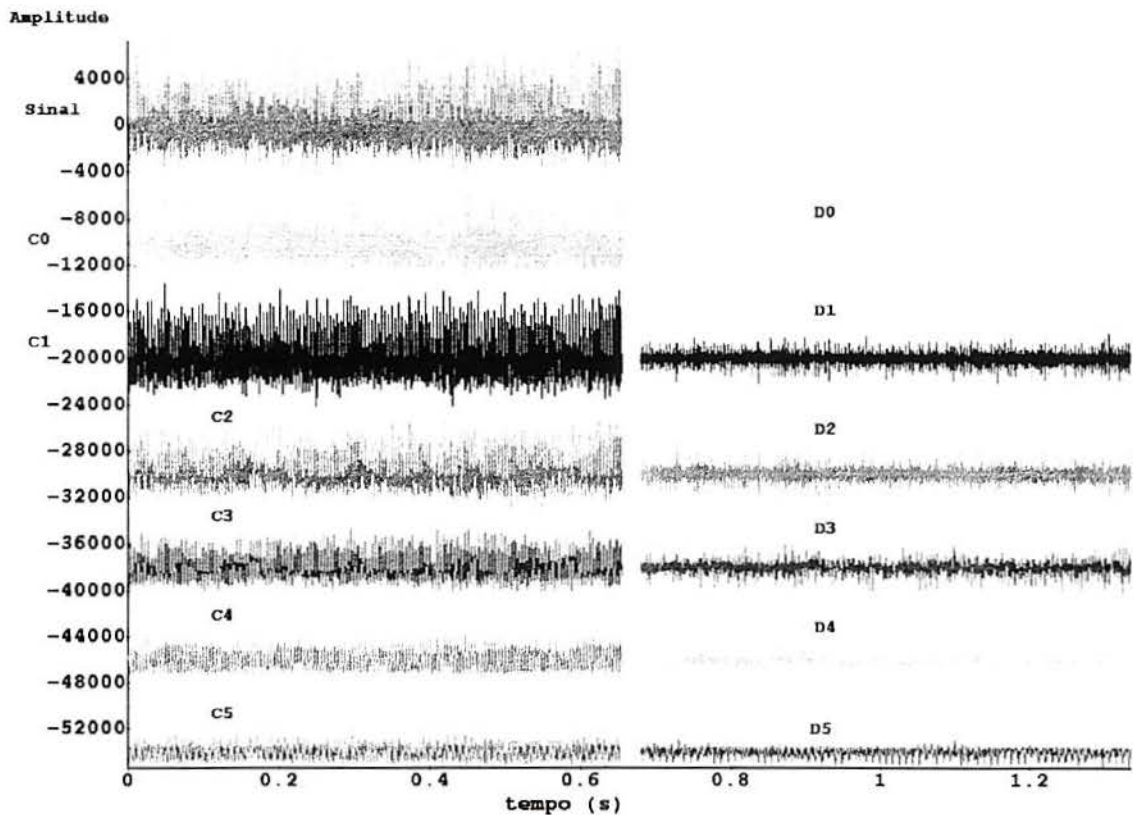


Figura A3.4: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC05.

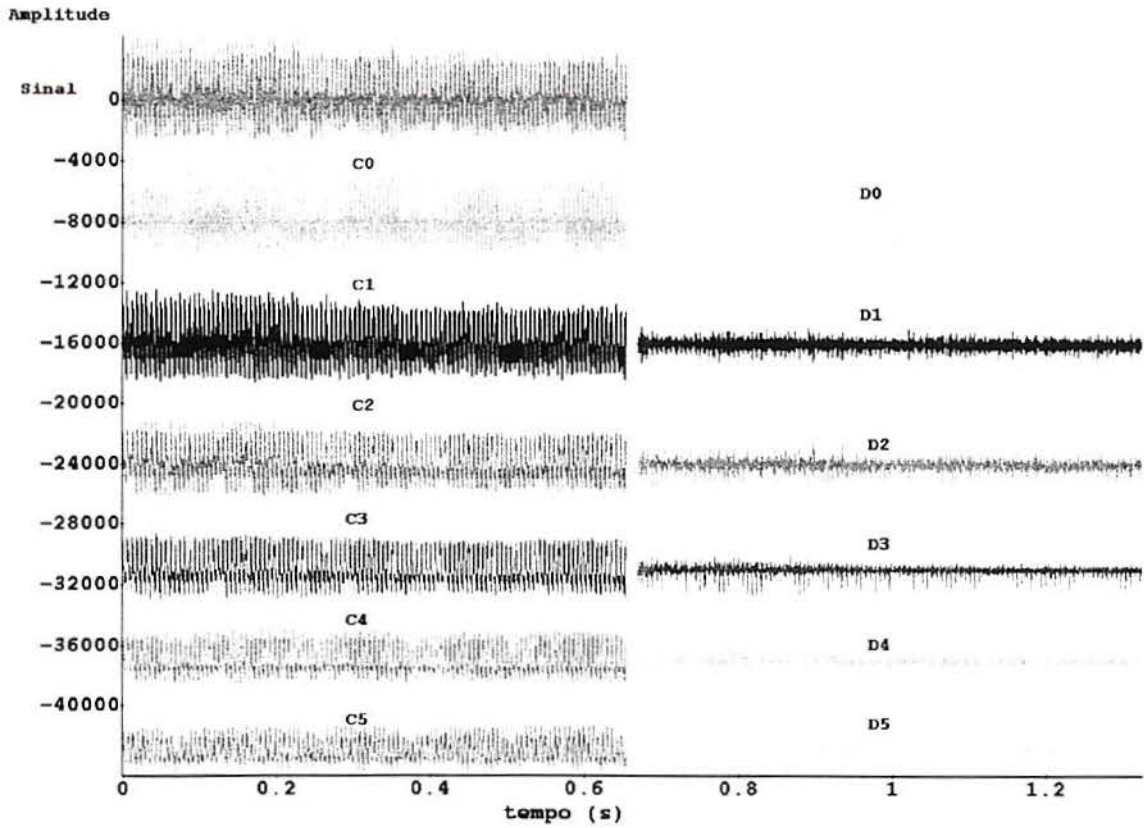


Figura A3.5: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC06.

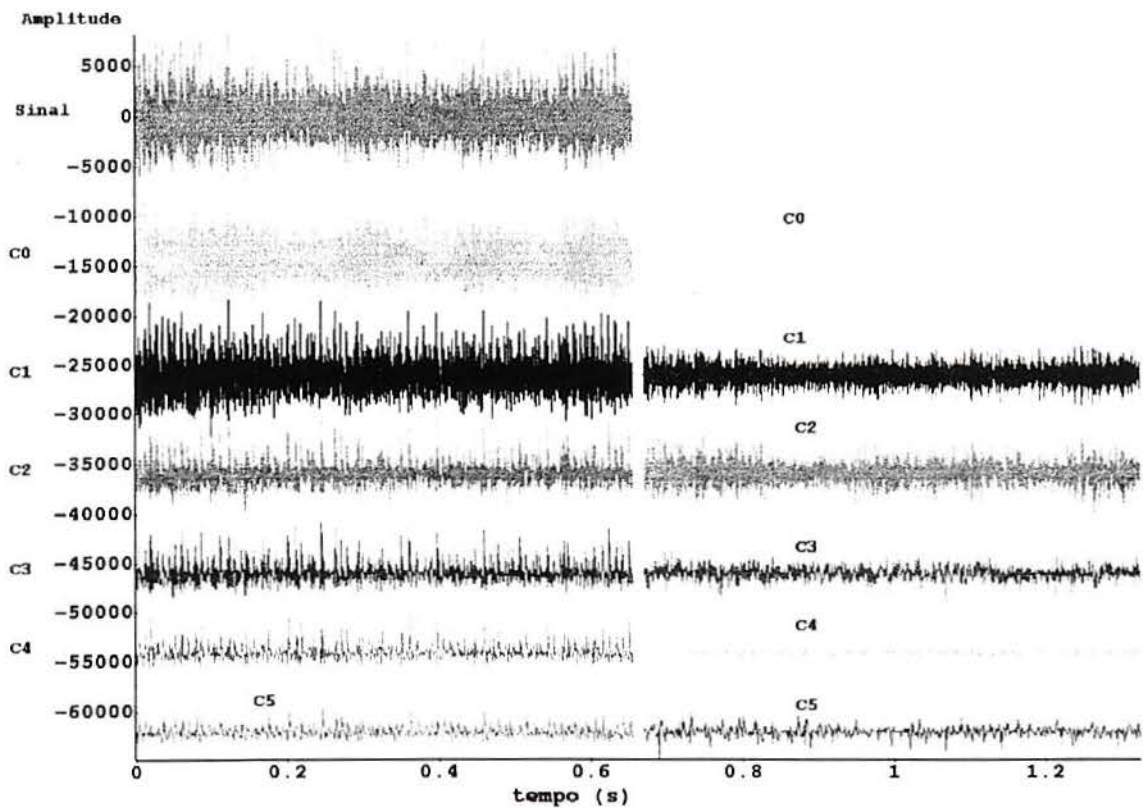


Figura A3.6: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC08.



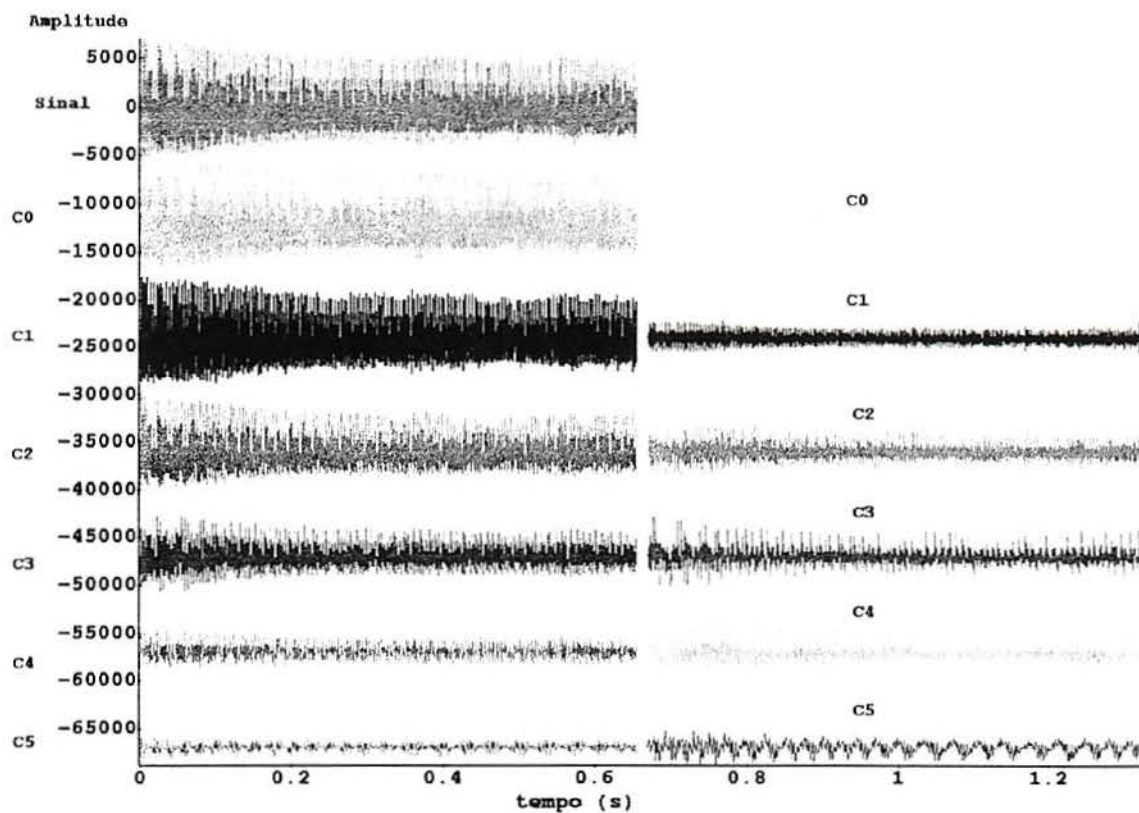


Figura A3.7: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC09.

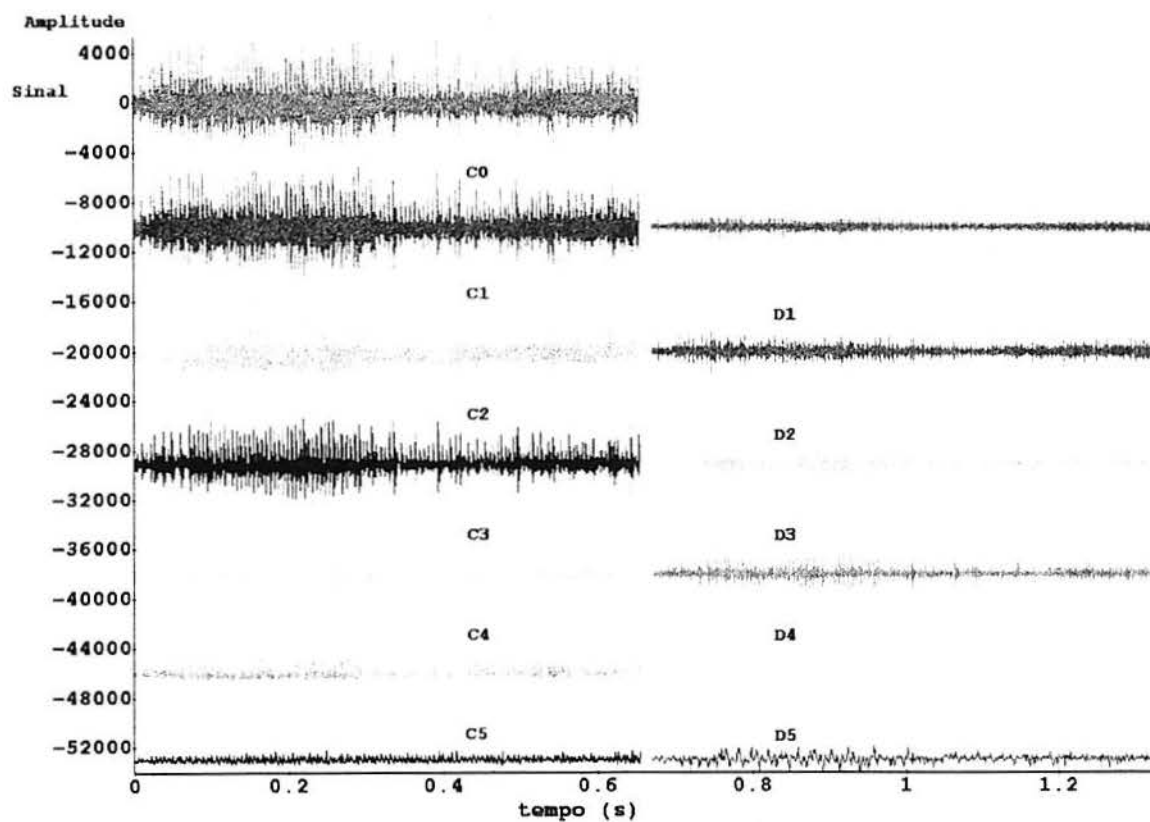


Figura A3.8: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC11.

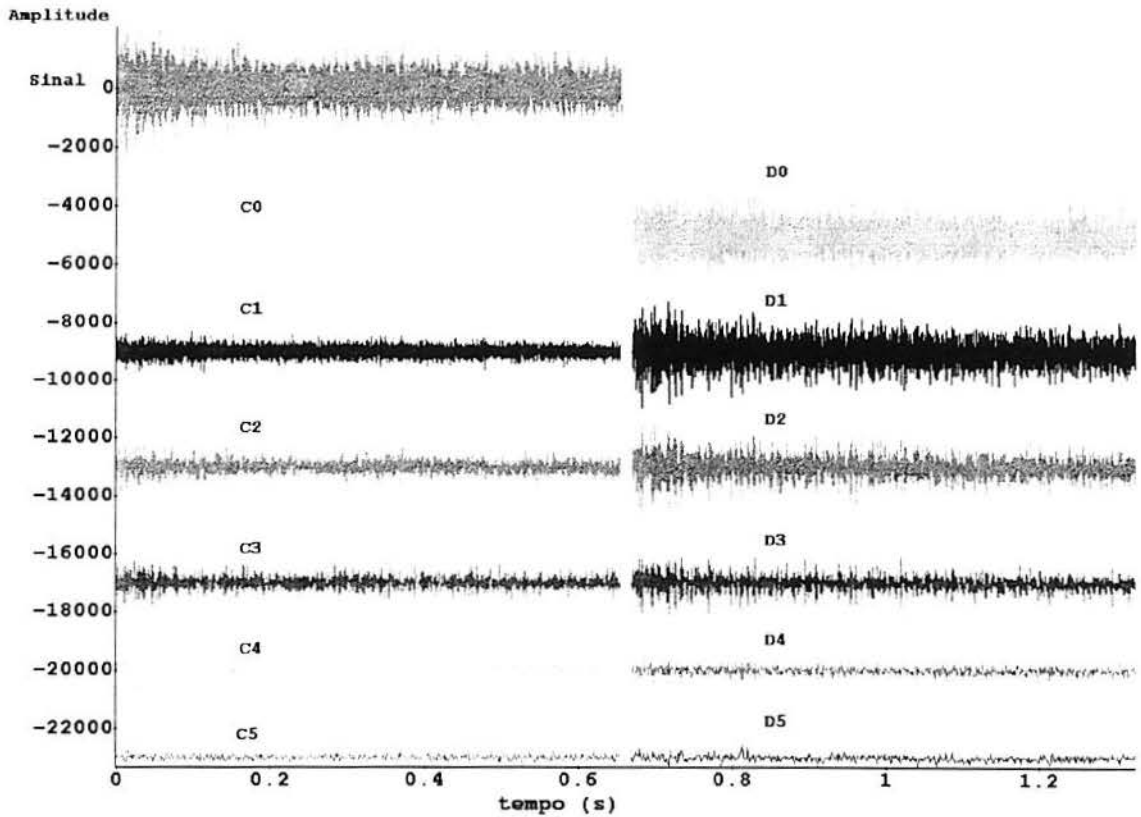


Figura A3.9: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC13.

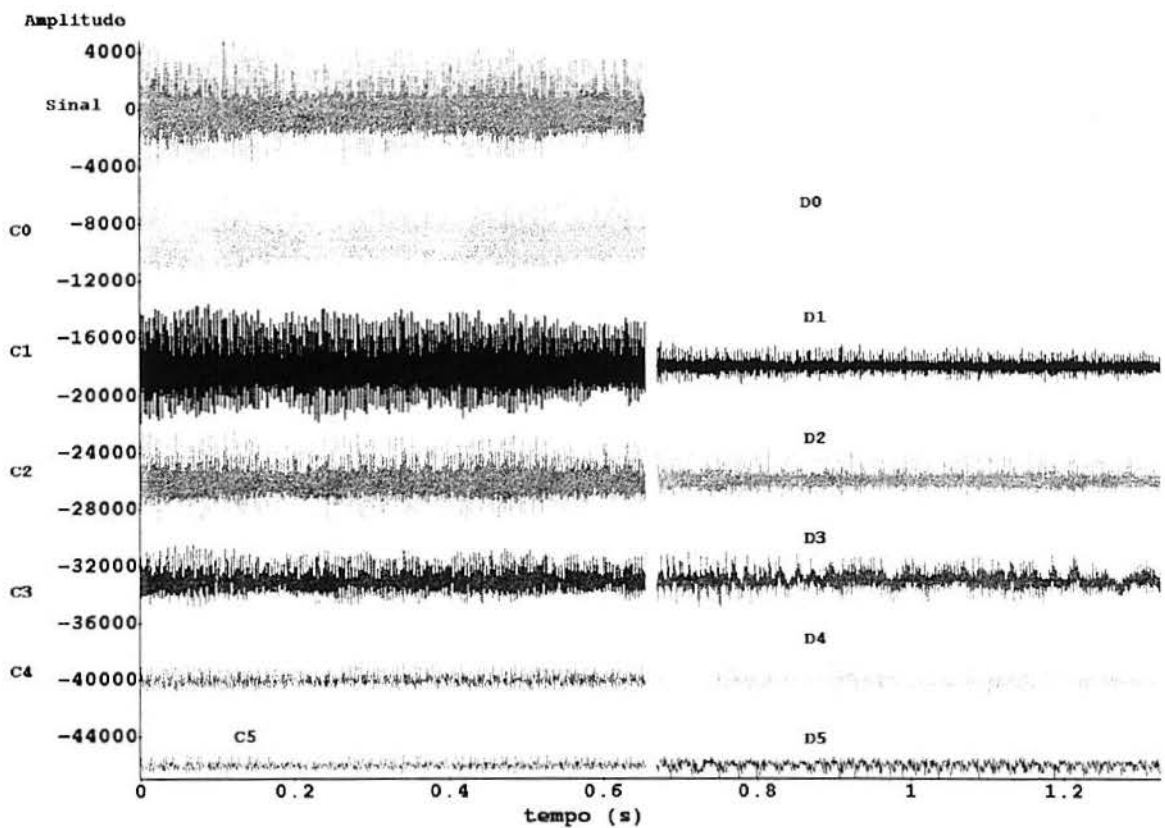


Figura A3.10: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC14.

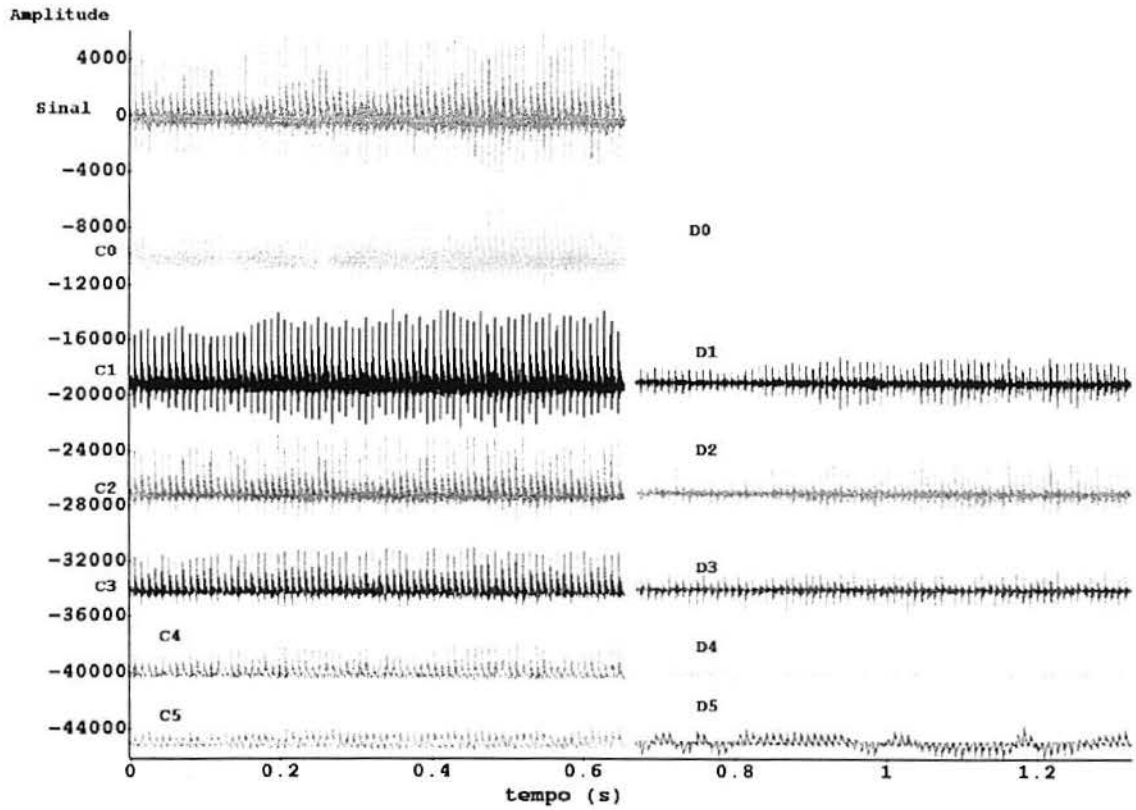


Figura A3.11: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC17.

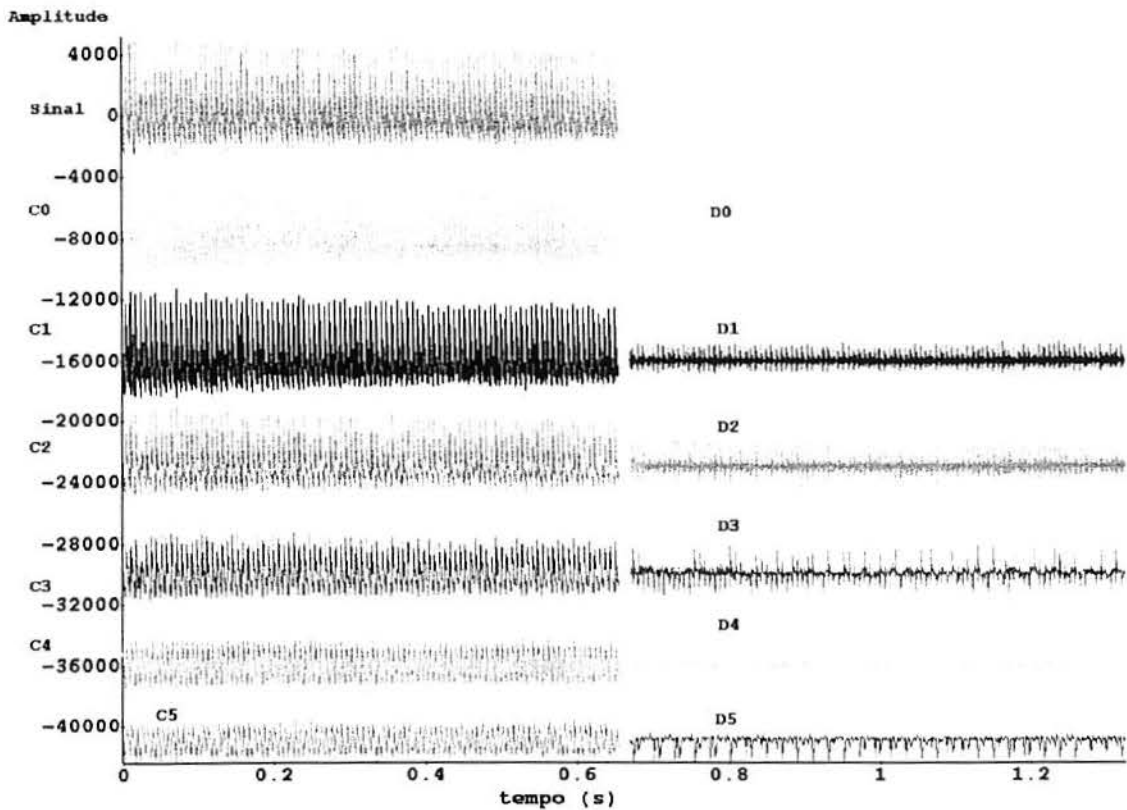


Figura A3.12: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC18.

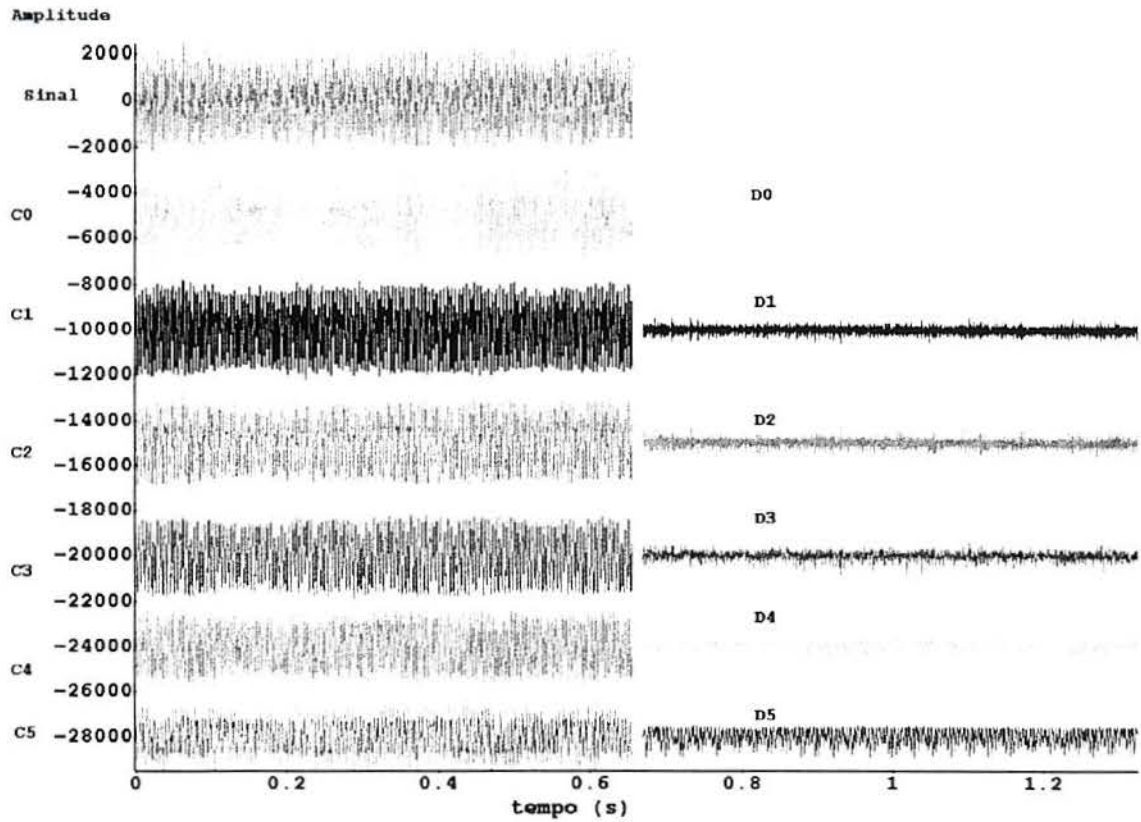


Figura A3.13: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC19.

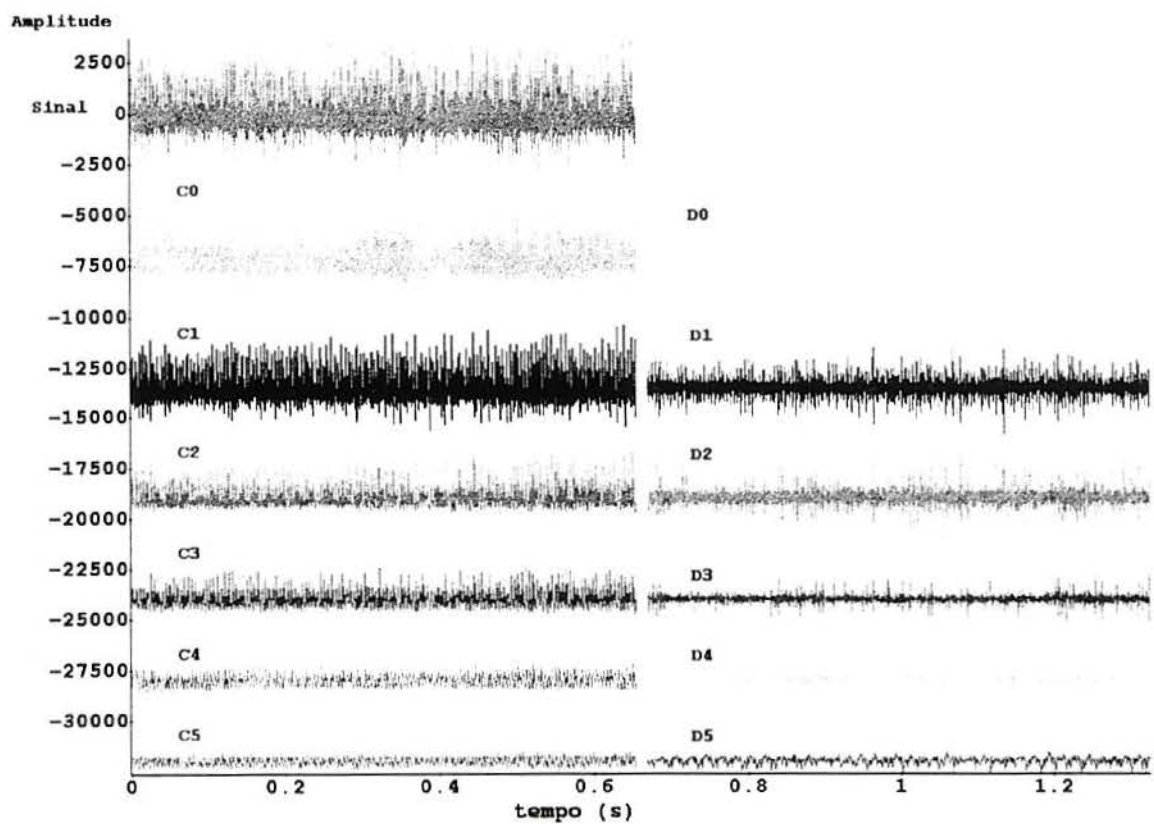


Figura A3.14: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC21.

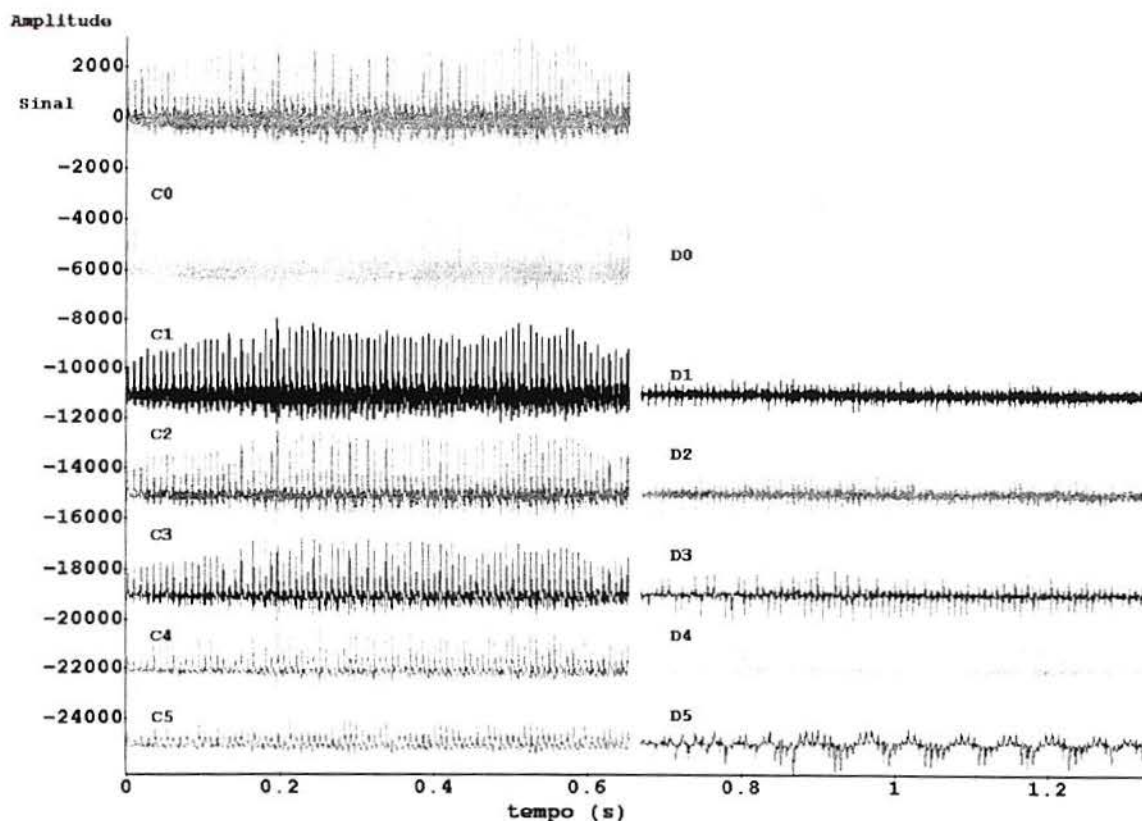


Figura A3.15: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC23.

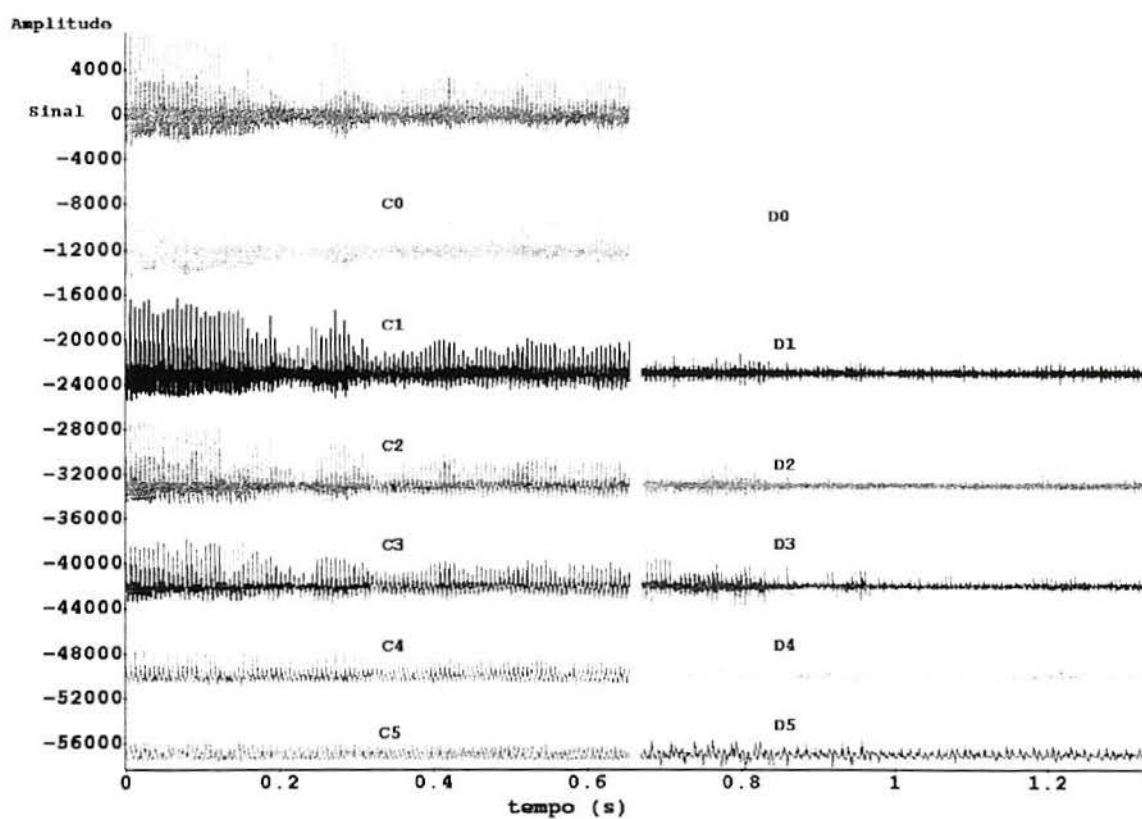


Figura A3.16: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC24.

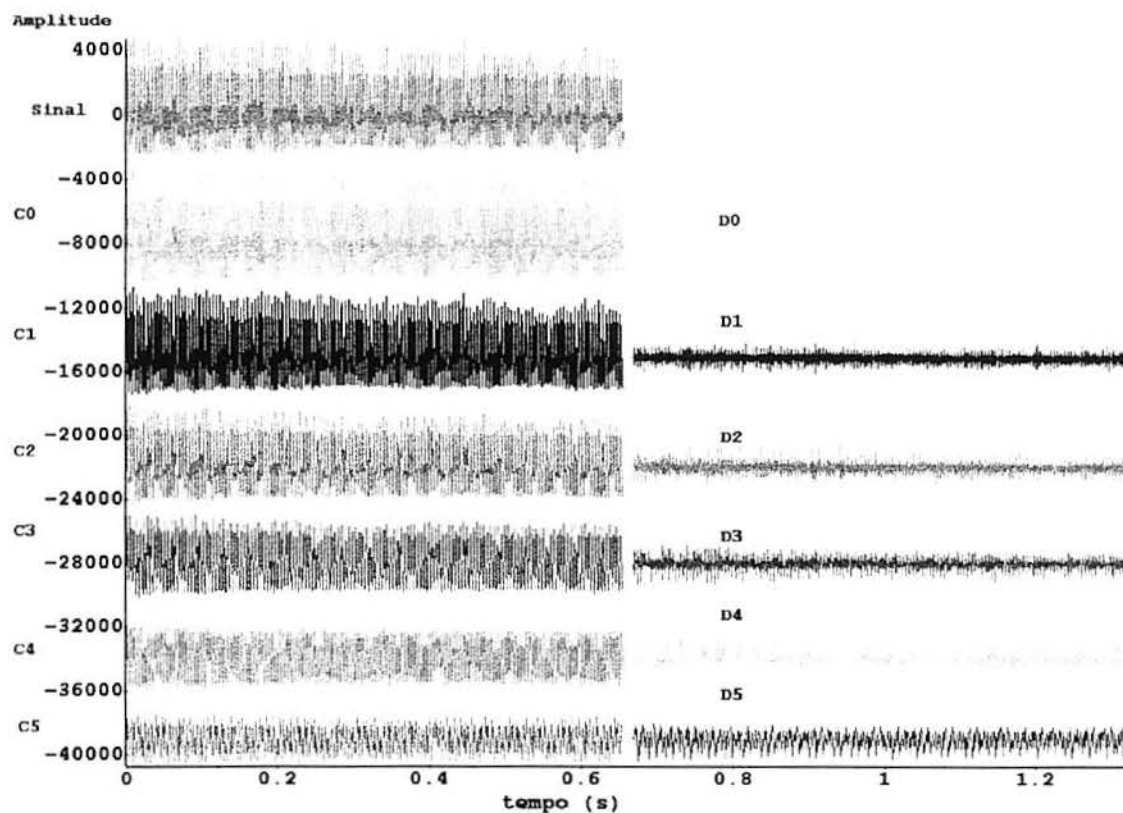


Figura A3.17: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC27.

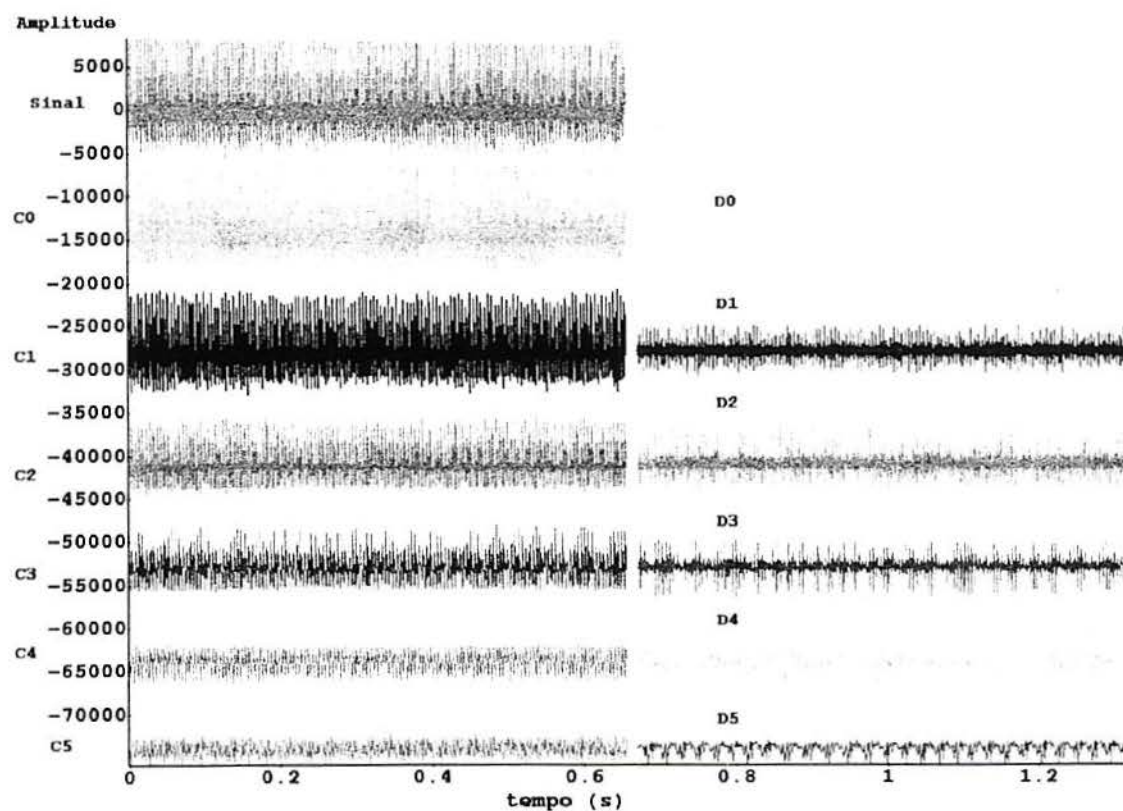


Figura A3.18: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC28.

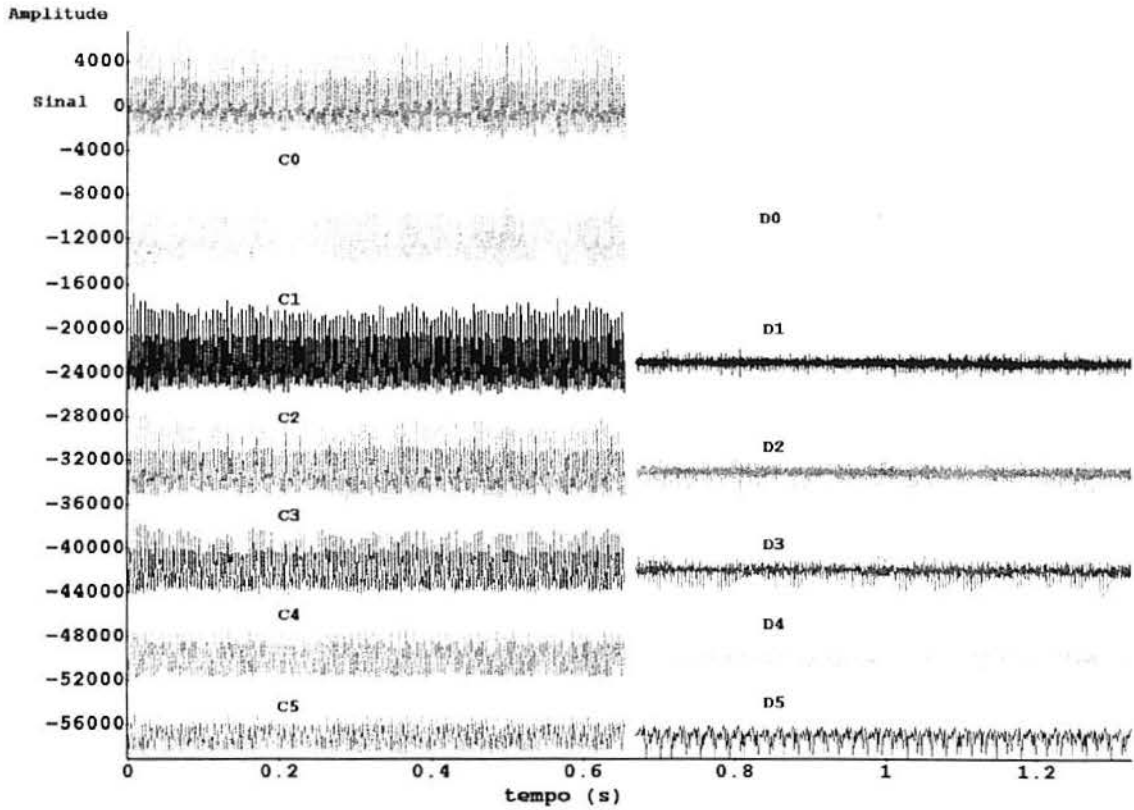


Figura A3.19: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC30.

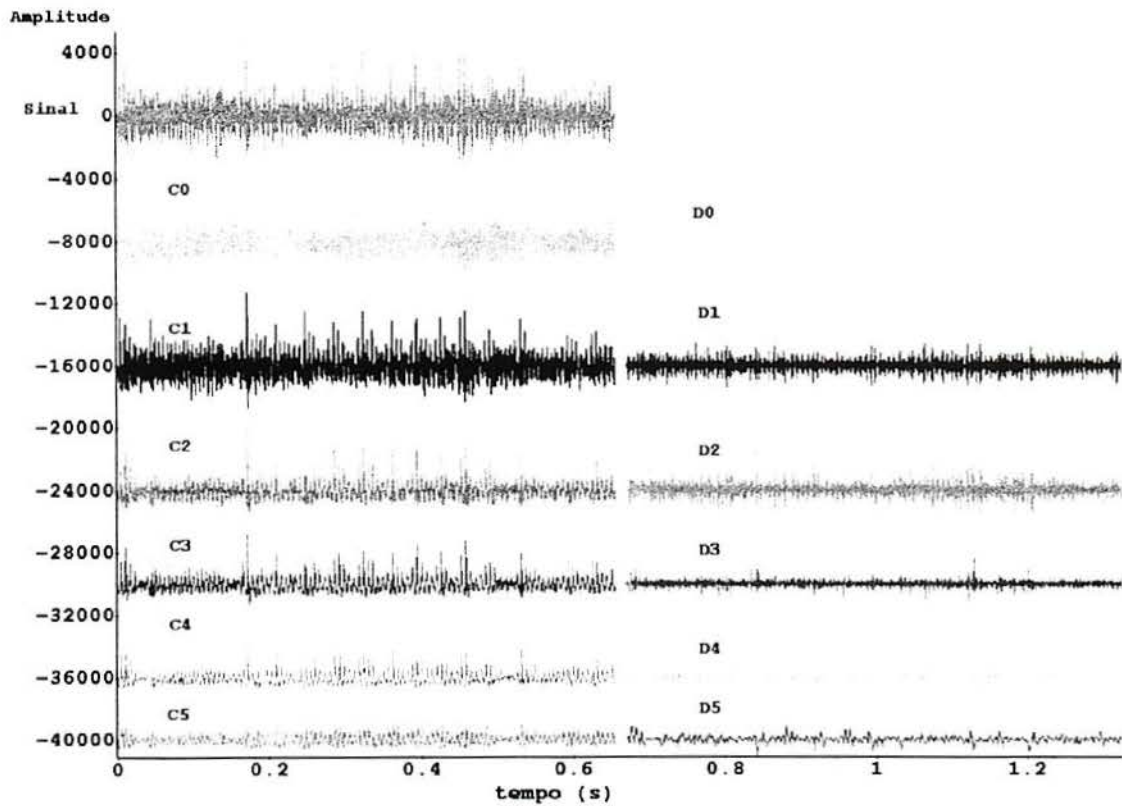


Figura A3.20: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC31.

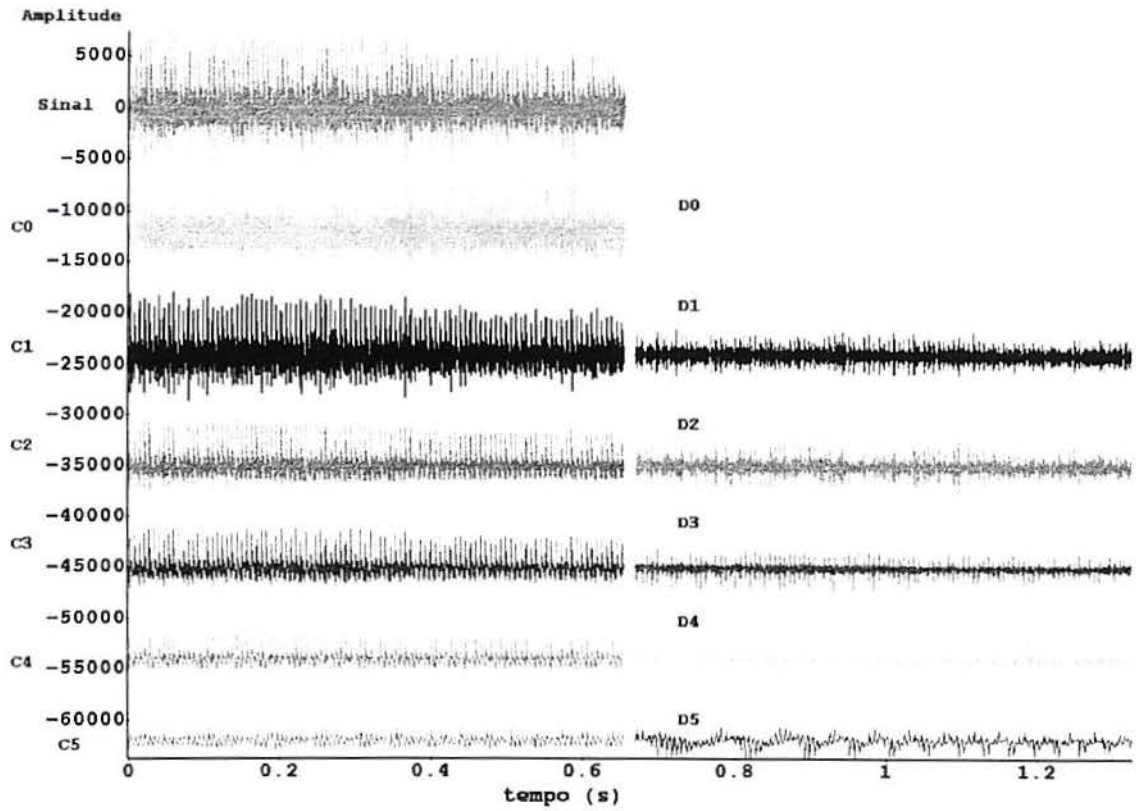


Figura A3.21: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC32.

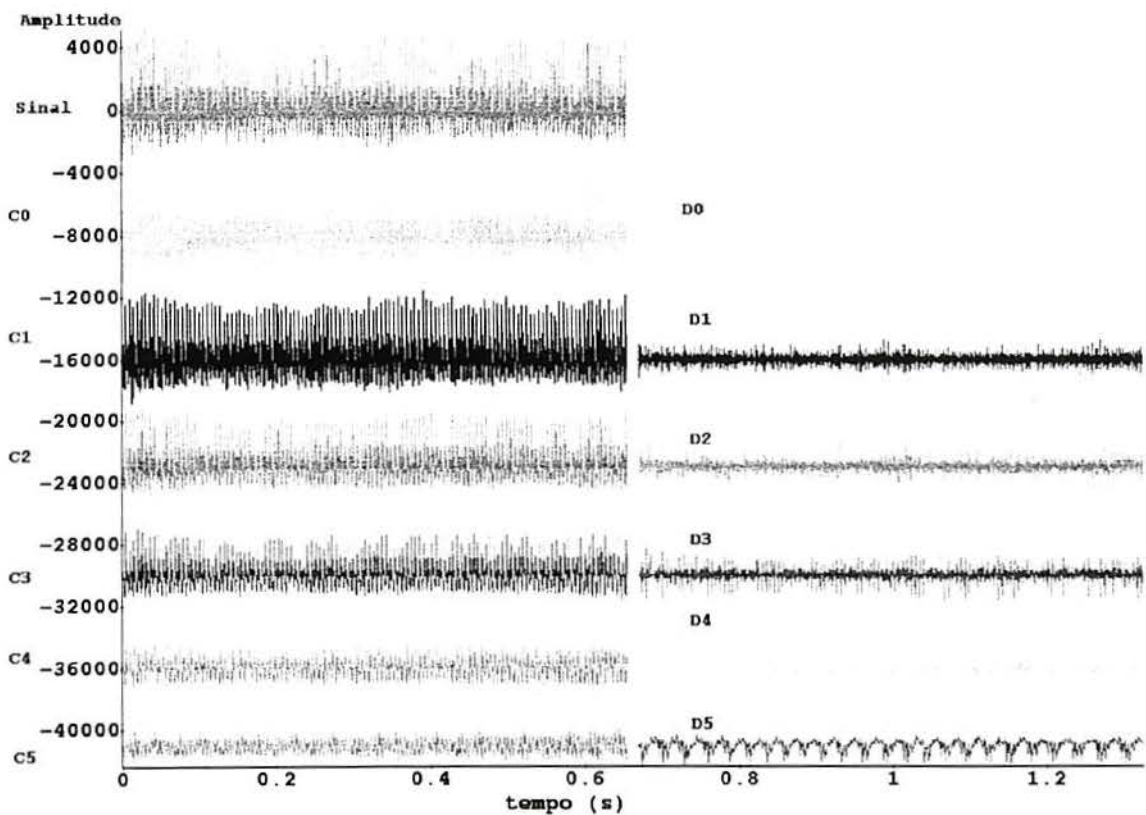


Figura A3.22: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC33.



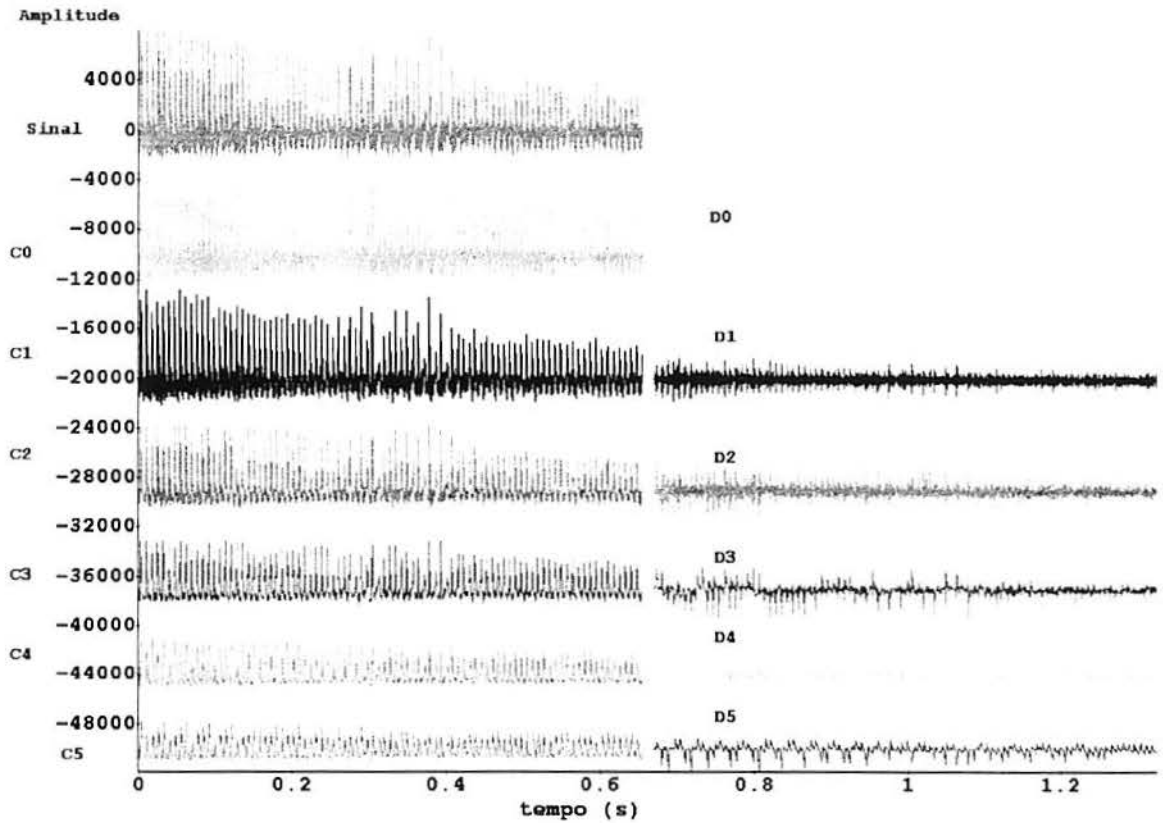


Figura A3.23: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC36.

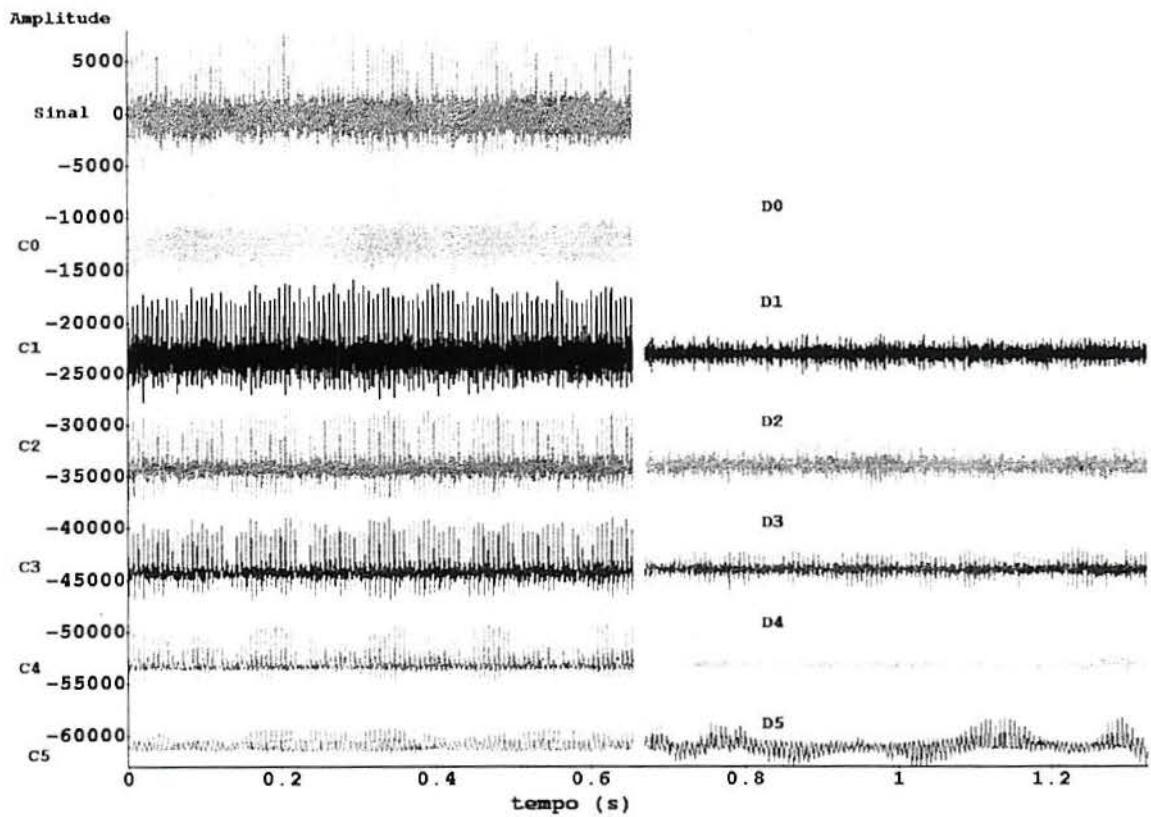


Figura A3.24: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC41.

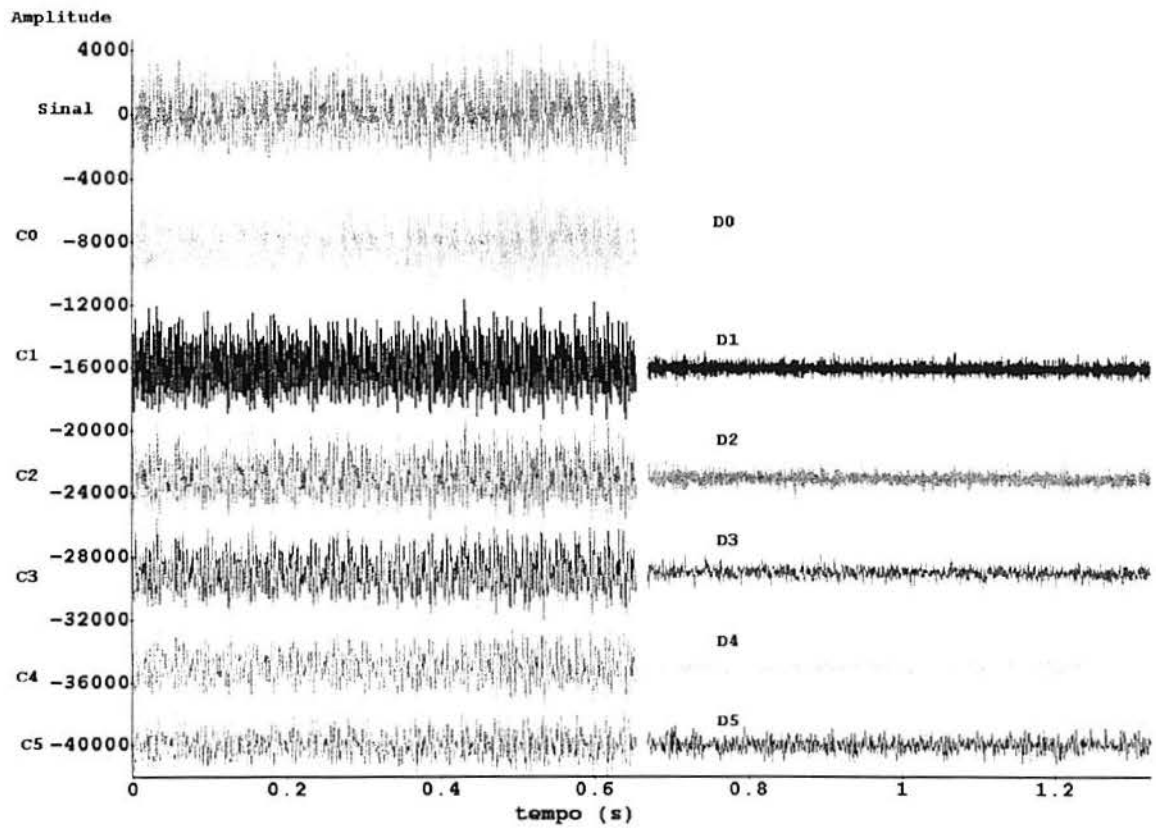


Figura A3.25: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC42.

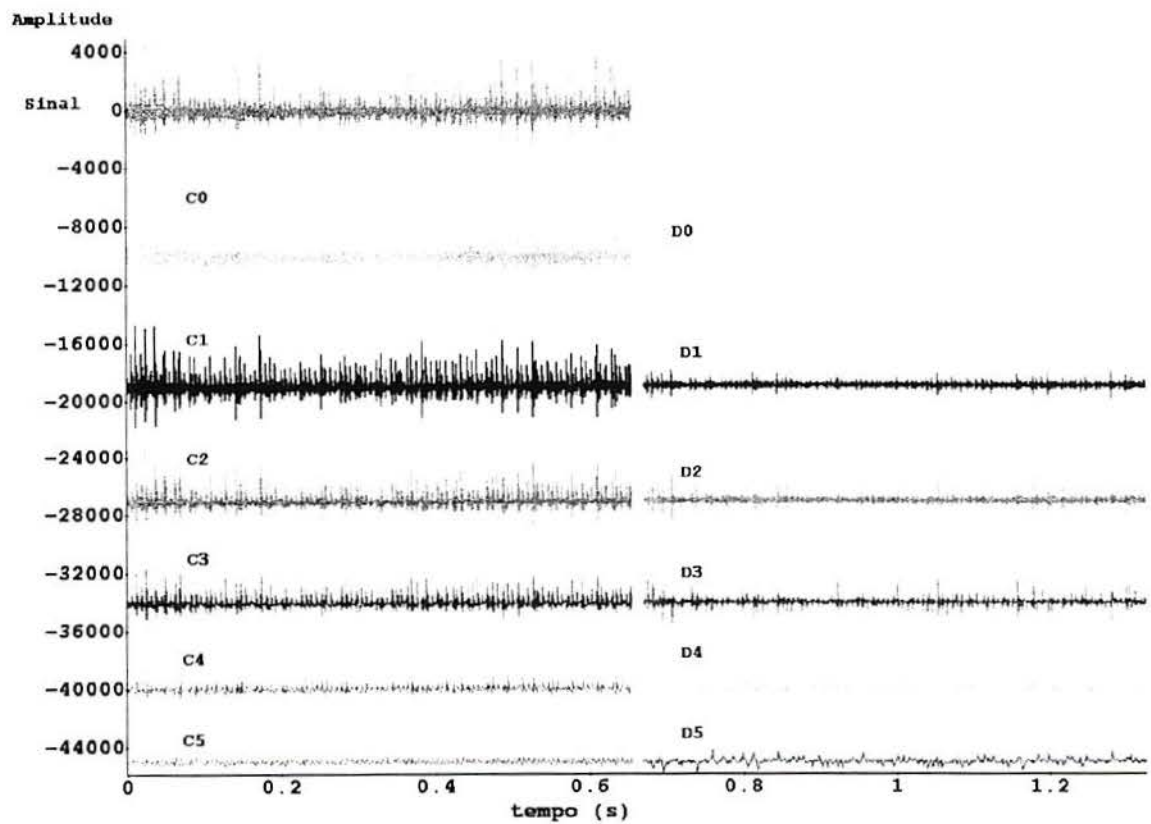


Figura A3.26: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC43.

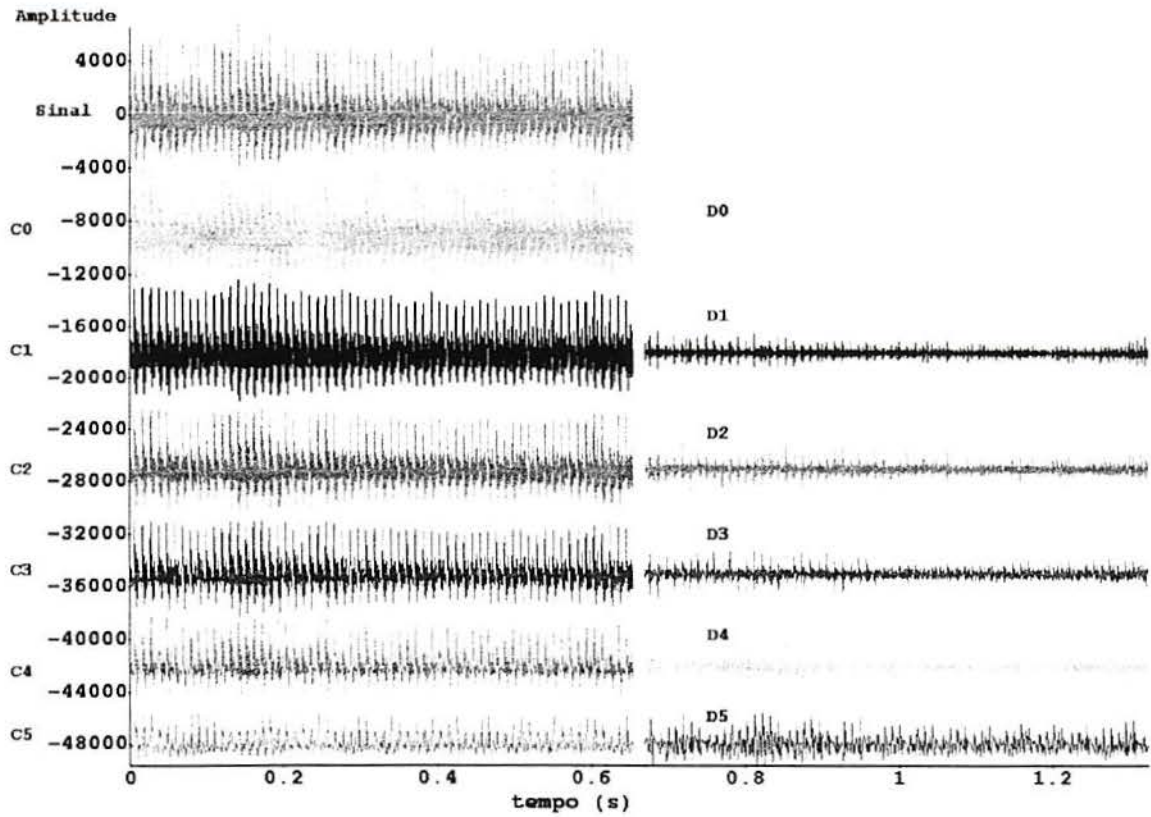


Figura A3.27: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC44.

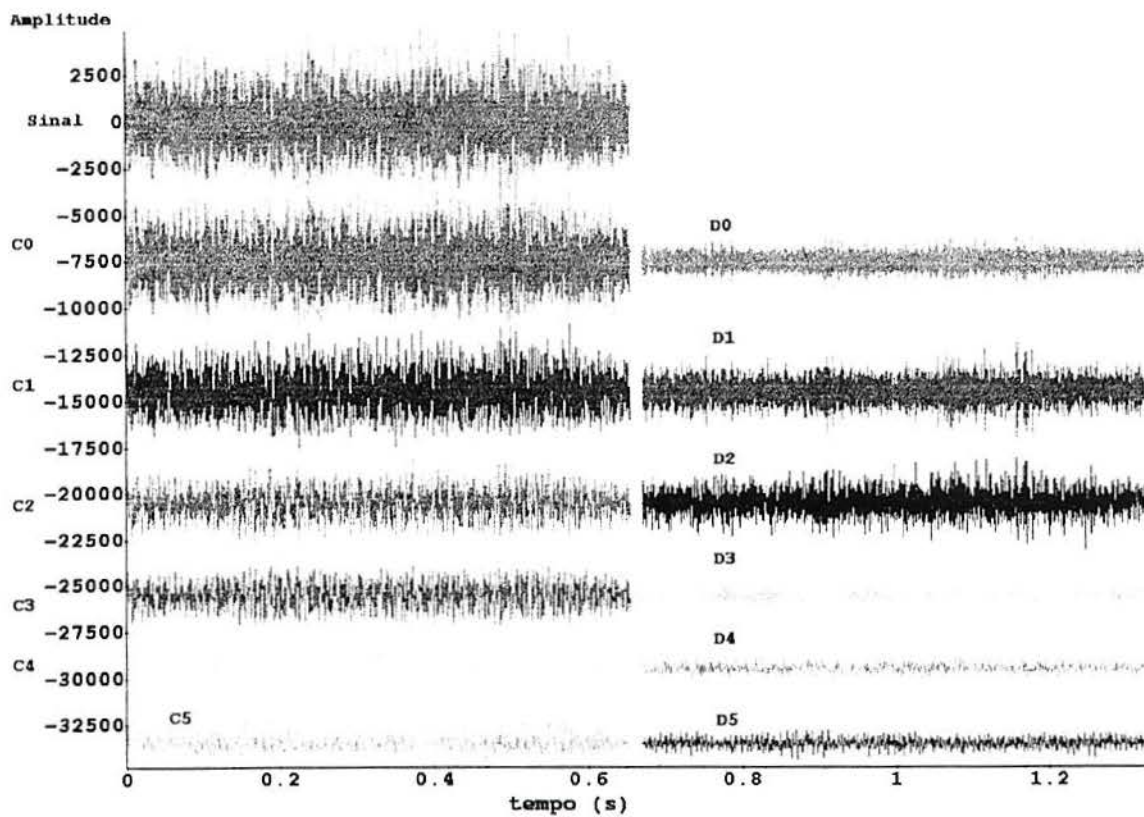


Figura A3.28: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC46.

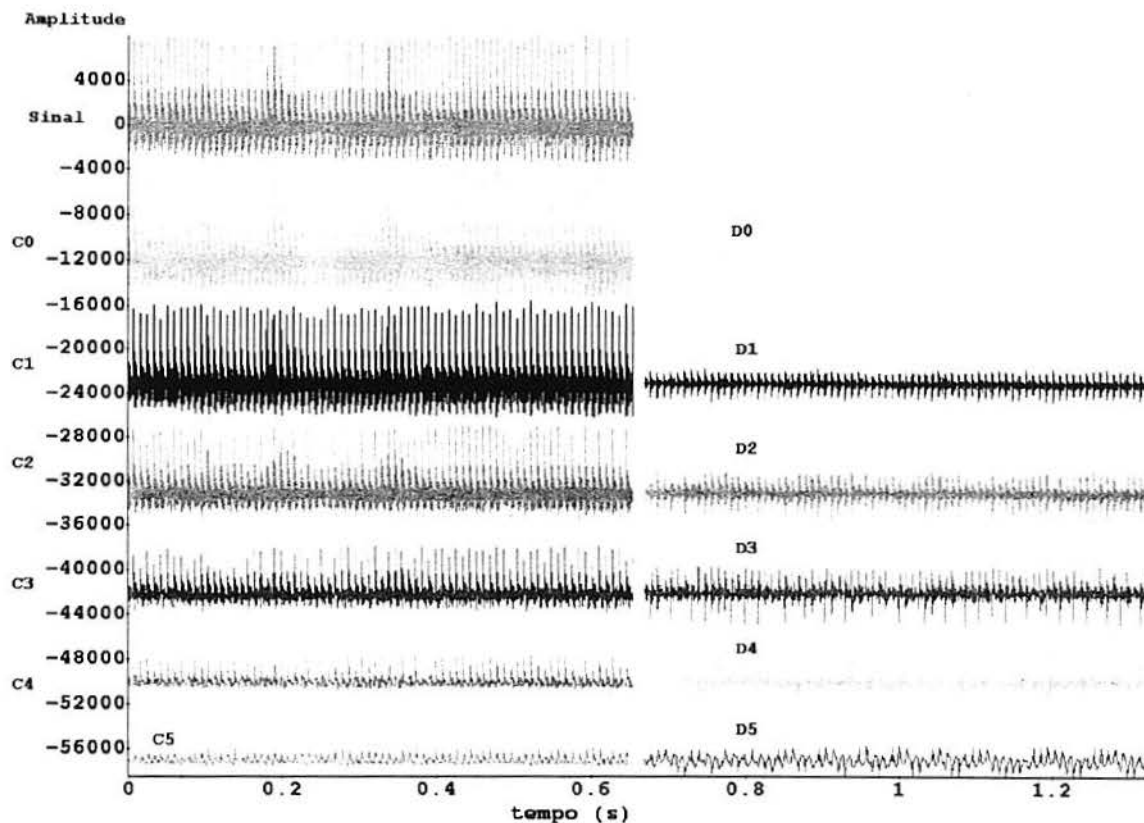


Figura A3.29: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC47.

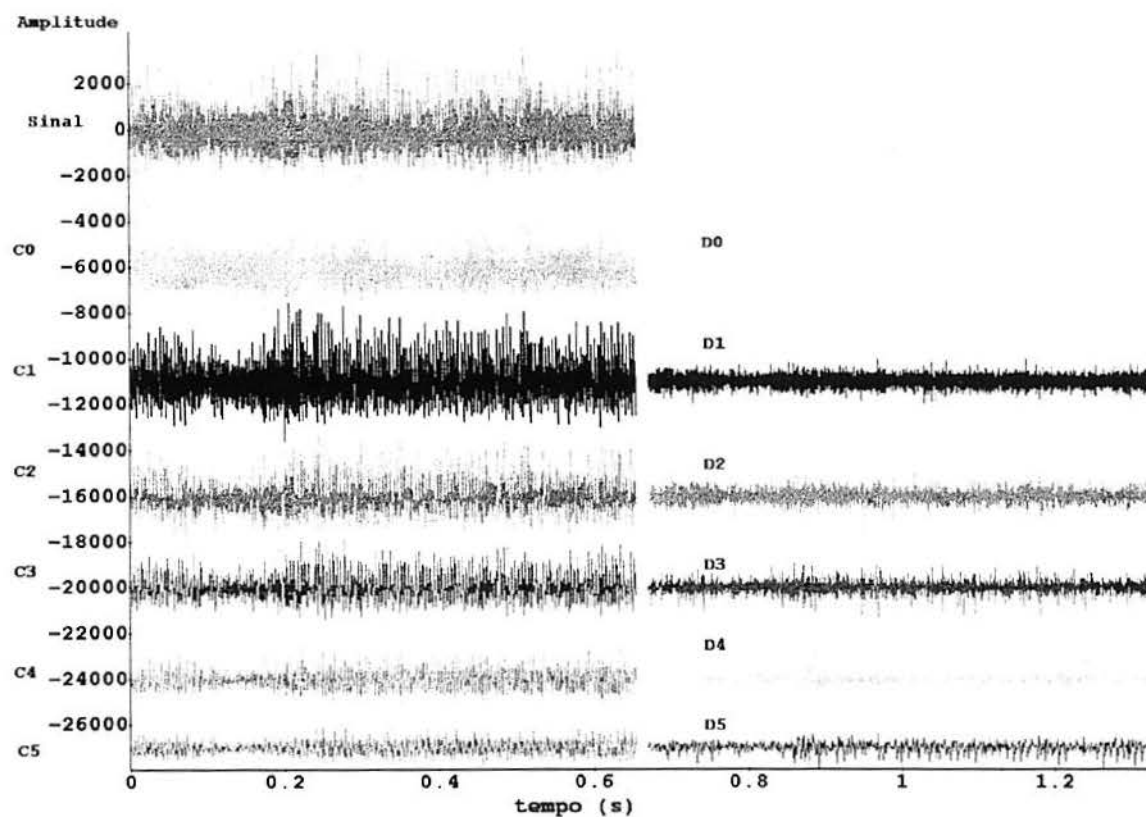


Figura A3.30: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC50.

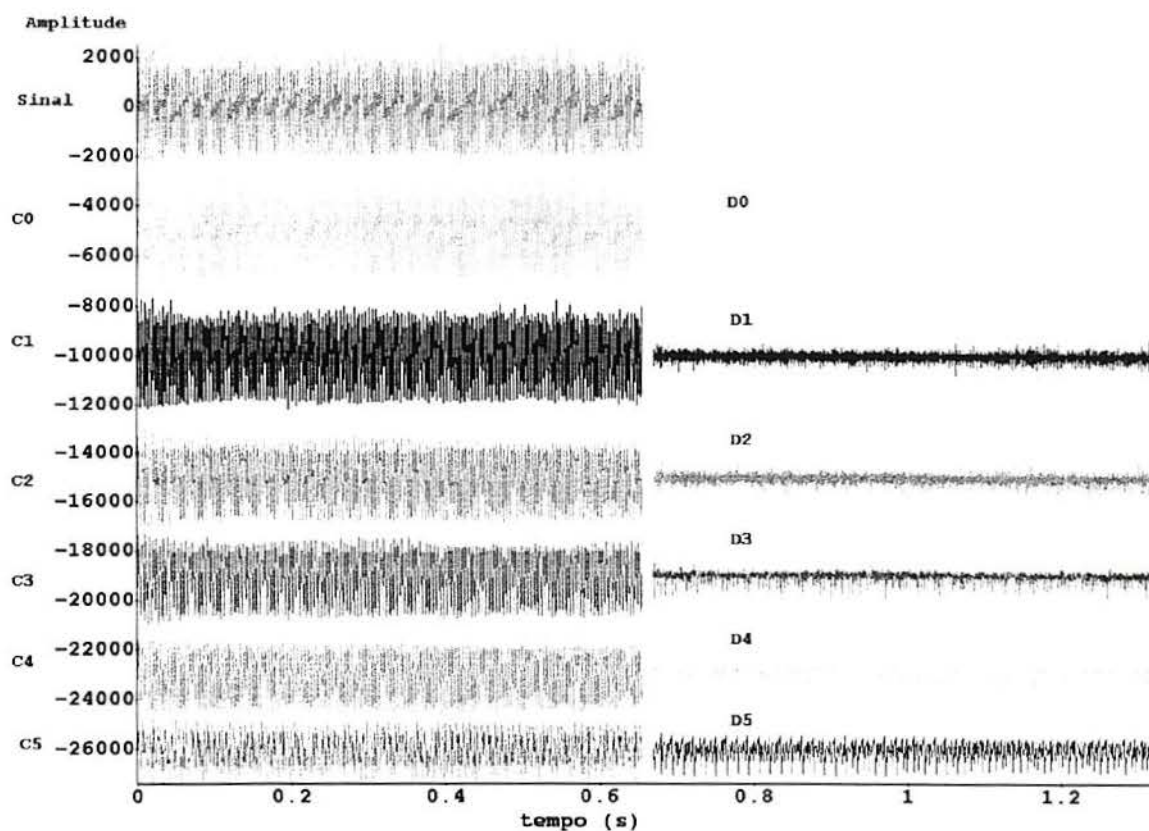


Figura A3.31: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC51.

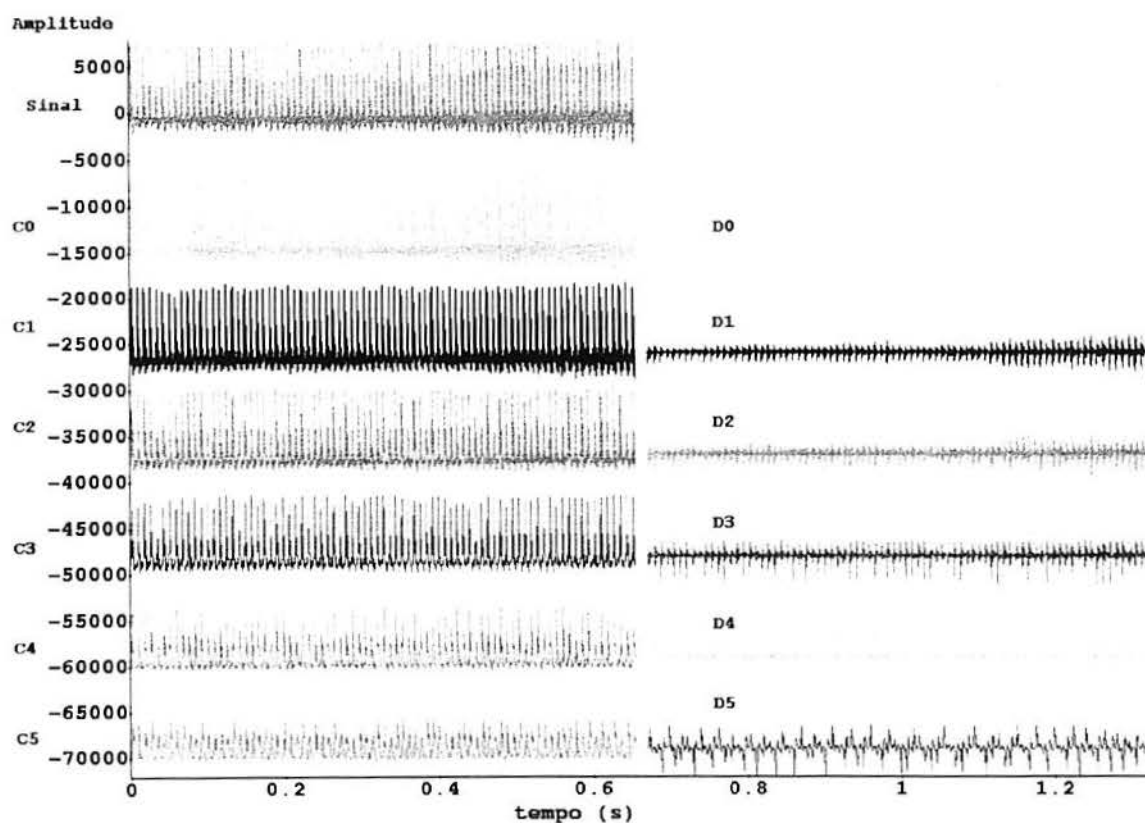


Figura A3.32: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC54.

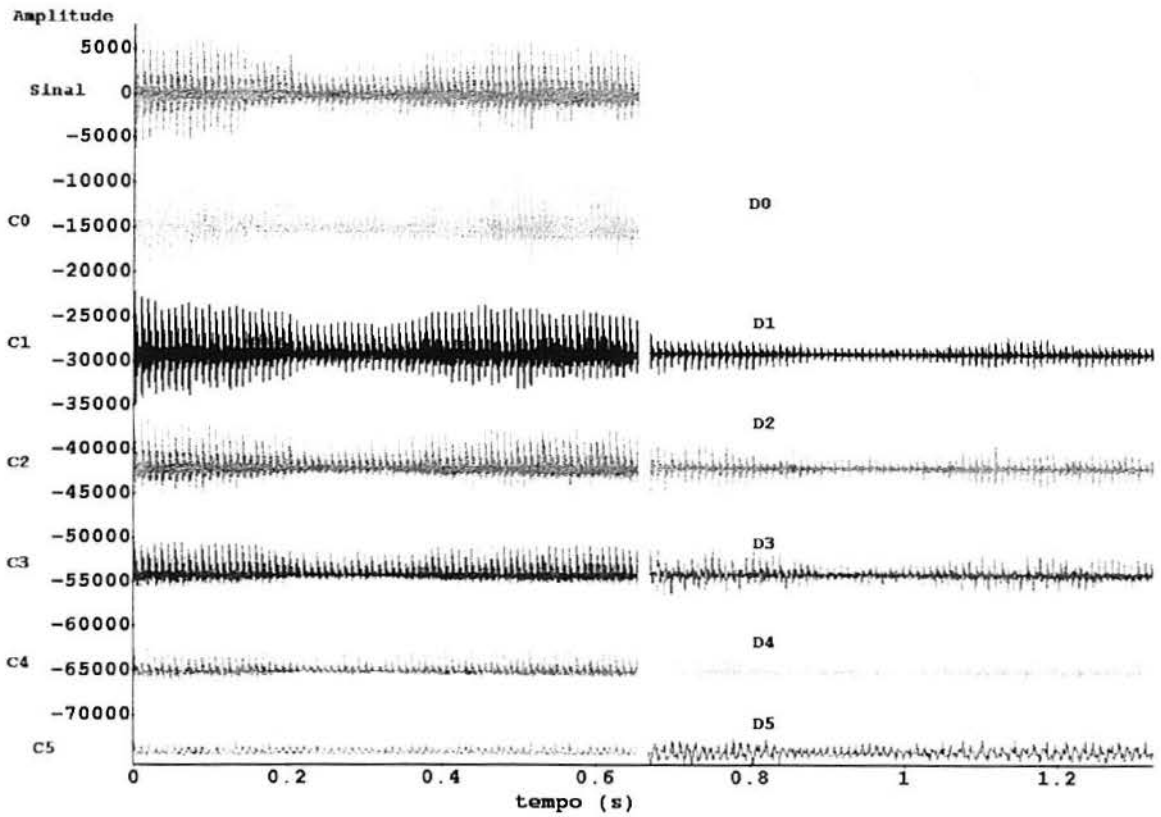


Figura A3.33: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC58.

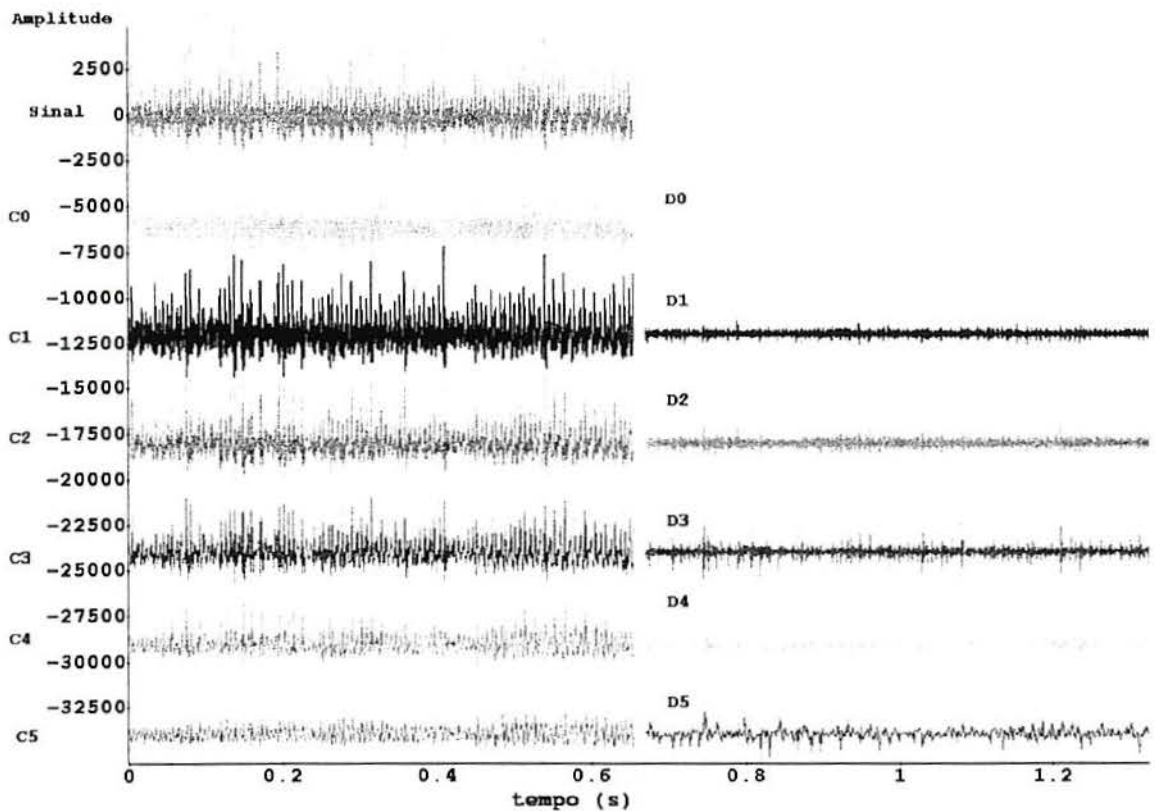


Figura A3.34: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC60.

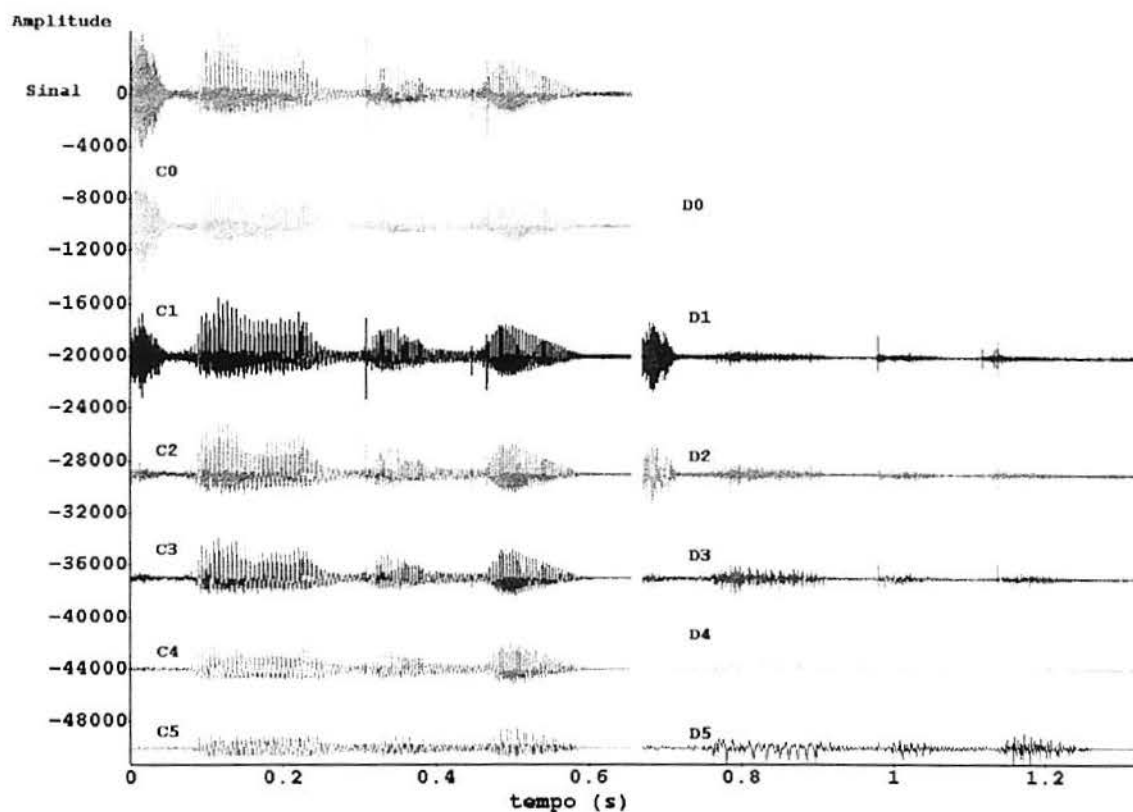


Figura A3.35: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC61.

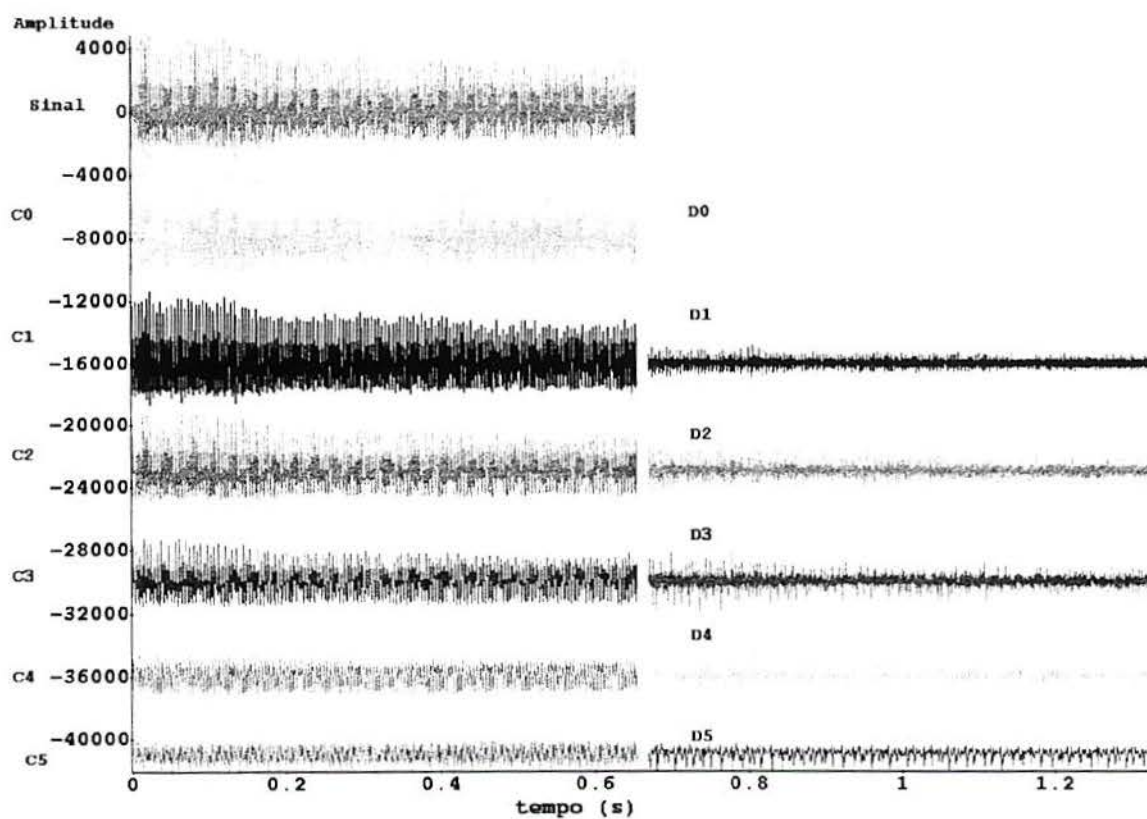


Figura A3.36: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC64.

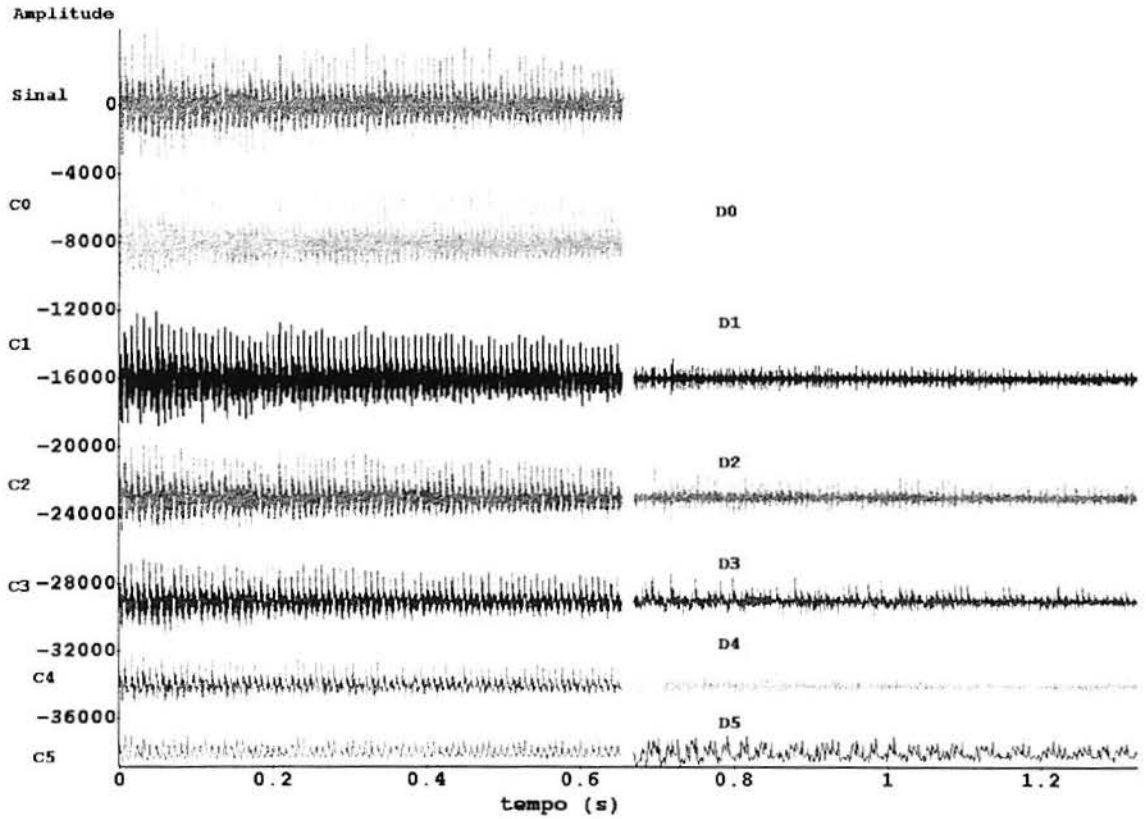


Figura A3.37: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC66.

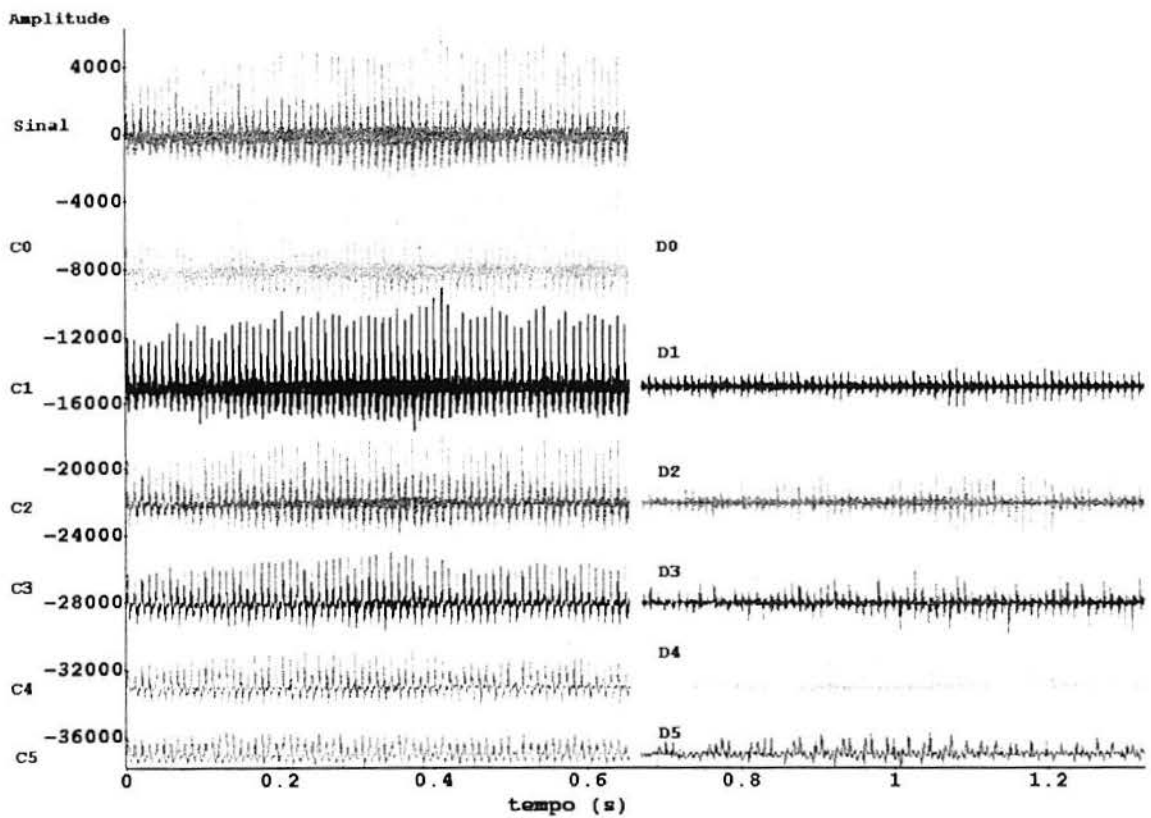


Figura A3.38: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC69.



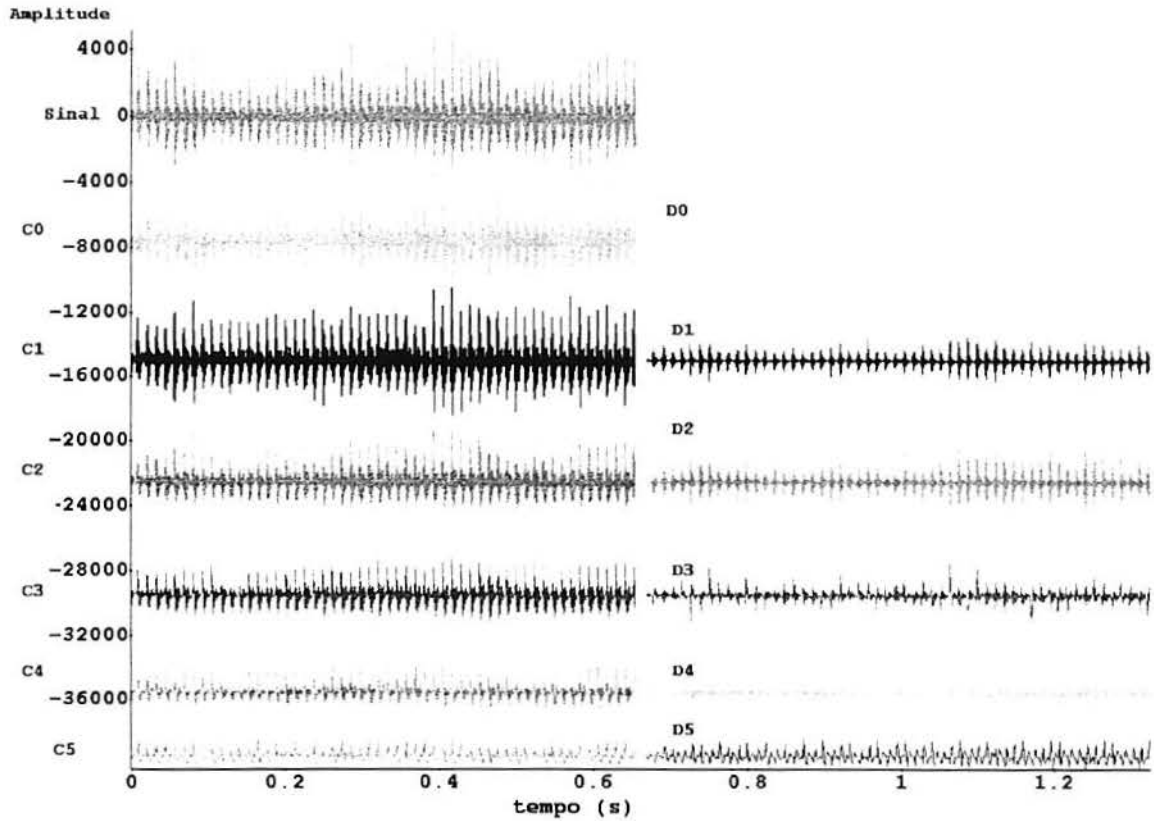


Figura A3.39: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC70.

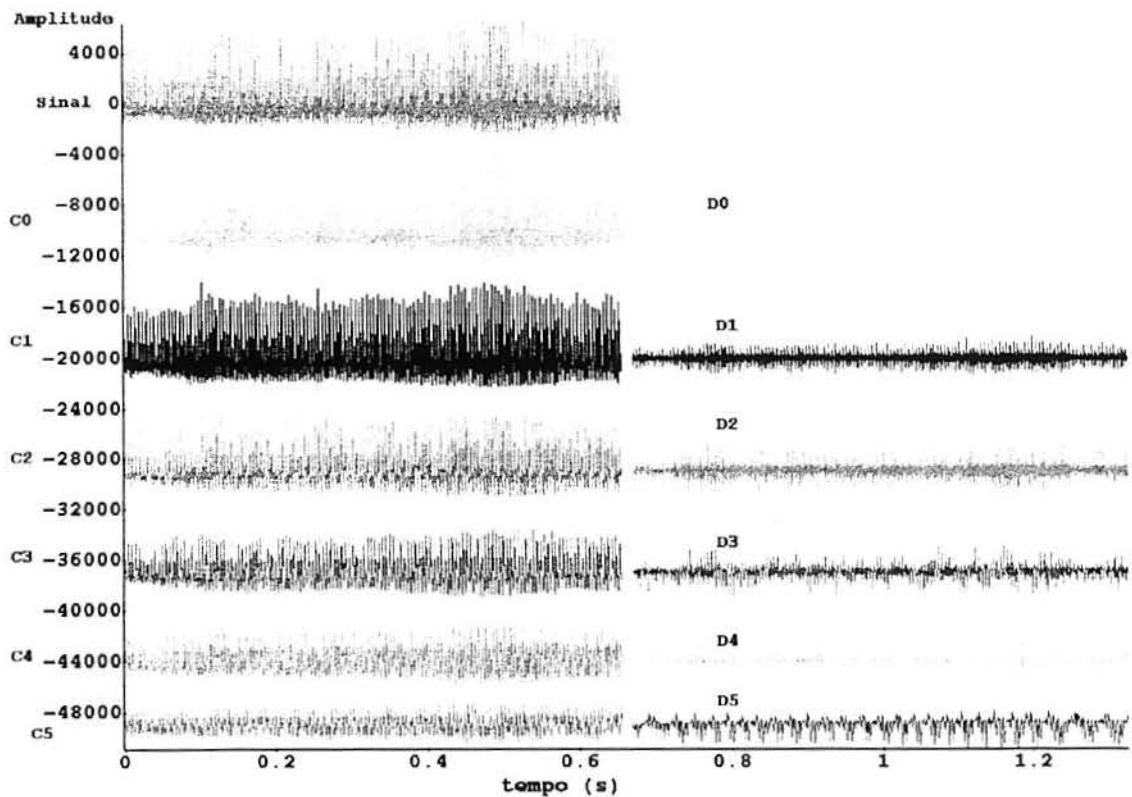


Figura A3.40: Gráficos do trecho de 32768 pontos do sinal e dos 6 níveis dos coeficientes da DWT, para a base de Haar do paciente PAC72.

ANEXO 4:

**Tabela A4.1: Valores dos índices de soprosideade clássicos e propostos para os pacientes:**

No.	Pacientes	RMS Sinal	RMS C0	RMS D0	d0/c0	rms pb0-6kHz	rms pa6-12.5kHz	pa/pb	rms do-deriv2	do"/rms sinal	Índice Br
1	PAC01	836,372	829,491	106,882	0,114	835,033	47,284	0,057	152,363	0,182	172,200
2	PAC02	1449,662	1443,490	131,949	0,084	1449,440	25,376	0,018	91,972	0,063	43,470
3	PAC04	1324,385	1314,839	157,545	0,107	1319,186	117,229	0,089	157,543	0,119	259,200
4	PAC05	1133,170	1117,977	184,337	0,142	1124,204	142,272	0,127	398,464	0,352	443,300
5	PAC06	1285,387	1281,208	103,034	0,074	1282,724	82,690	0,064	213,810	0,166	415,700
6	PAC08	1578,188	1511,599	453,510	0,231	1542,204	335,098	0,217	944,286	0,598	427,800
7	PAC09	2181,113	2160,127	301,783	0,123	2179,956	71,017	0,033	186,274	0,085	34,870
8	PAC10	3069,295	3064,967	162,349	0,050	3068,860	51,086	0,017	144,291	0,047	69,890
9	PAC11	1464,740	1420,557	357,012	0,201	1430,663	314,112	0,220	847,542	0,579	541,200
10	PAC13	445,419	423,453	138,090	0,246	424,907	133,610	0,314	396,473	0,890	811,800
11	PAC14	2031,707	2008,672	304,501	0,132	2029,475	95,139	0,047	282,455	0,139	76,600
12	PAC15	1352,824	1328,652	254,578	0,161	1339,406	190,082	0,142	515,108	0,381	415,300
13	PAC16	601,842	595,777	84,978	0,125	598,844	59,986	0,100	178,778	0,297	420,400
14	PAC17	1764,205	1747,720	240,604	0,121	1759,245	132,230	0,075	376,623	0,213	222,800
15	PAC18	608,447	604,892	65,648	0,098	606,931	42,907	0,071	138,858	0,228	446,300
16	PAC19	1370,356	1368,380	72,468	0,050	1369,406	51,001	0,037	169,842	0,124	499,000
17	PAC21	1069,567	979,350	429,928	0,305	980,853	426,502	0,435	1320,110	1,234	927,700
18	PAC22	2103,943	2097,925	158,485	0,070	2099,464	107,511	0,051	376,137	0,179	532,500
19	PAC23	1084,320	1076,084	133,384	0,110	1078,974	107,511	0,100	258,071	0,238	349,100
20	PAC24	861,204	841,796	181,793	0,178	839,331	192,865	0,230	522,443	0,607	804,400
21	PAC25	2292,651	2169,318	741,313	0,255	2258,457	394,466	0,175	1282,297	0,559	294,800
22	PAC26	2163,767	2157,741	160,801	0,069	2162,937	59,743	0,028	170,343	0,079	100,900
23	PAC27	2078,480	2069,844	189,037	0,084	2072,075	163,066	0,079	427,378	0,206	491,300
24	PAC28	2330,296	2289,620	433,475	0,159	2313,164	282,049	0,122	671,270	0,288	231,100
25	PAC29	2531,054	2509,445	329,861	0,116	2527,132	140,862	0,056	414,816	0,164	143,000
26	PAC30	2310,712	2299,434	227,459	0,090	2303,175	186,392	0,081	510,869	0,221	485,800
27	PAC31	981,688	960,494	202,853	0,174	961,303	199,021	0,207	570,264	0,581	761,500
28	PAC32	1863,222	1829,408	353,103	0,162	1842,749	275,435	0,149	680,921	0,365	340,500
29	PAC34	1619,858	1485,712	645,448	0,303	1457,243	707,372	0,485	1915,931	1,183	863,200
30	PAC35	2134,788	2111,974	311,228	0,128	2128,502	163,652	0,077	370,548	0,174	129,300
31	PAC36	886,513	879,519	110,875	0,112	881,146	97,392	0,111	184,062	0,208	264,600
32	PAC37	1904,913	1879,465	310,338	0,142	1900,710	126,493	0,067	451,220	0,237	192,600
33	PAC41	1978,387	1960,263	267,116	0,120	1973,703	136,051	0,069	350,984	0,177	178,200
34	PAC42	2171,613	2164,773	169,213	0,072	2168,033	124,690	0,058	379,746	0,175	471,700
35	PAC43	1200,081	1185,020	189,542	0,138	1195,130	108,920	0,091	348,476	0,290	311,100
36	PAC44	1361,783	1357,727	104,942	0,072	1360,691	54,490	0,040	121,849	0,089	111,700
37	PAC45	1694,998	1683,133	199,721	0,106	1694,464	42,545	0,025	96,244	0,057	21,780
38	PAC46	1022,582	956,405	361,888	0,275	946,671	386,638	0,408	821,167	0,803	518,300
39	PAC47	1073,477	1062,021	156,422	0,128	1073,331	17,786	0,017	90,239	0,084	31,520
40	PAC48	1419,836	1409,051	172,885	0,109	1416,152	102,200	0,072	334,744	0,236	342,700
41	PAC49	1528,059	1522,780	126,830	0,077	1527,689	33,731	0,022	79,208	0,052	35,080
42	PAC50	857,600	840,058	172,584	0,170	842,028	162,709	0,193	356,231	0,415	426,800
43	PAC51	2367,867	2353,982	255,890	0,098	2363,500	143,778	0,061	452,383	0,191	287,300
44	PAC53	1561,461	1549,257	192,628	0,111	1561,055	35,660	0,023	103,973	0,067	27,410
45	PAC54	1249,551	1246,841	82,248	0,062	1249,458	15,748	0,013	46,748	0,037	29,370
46	PAC55	967,955	960,295	121,535	0,112	967,791	17,771	0,018	61,195	0,063	24,960
47	PAC56	2749,137	2722,400	381,173	0,123	2748,037	77,806	0,028	218,656	0,080	30,880
48	PAC57	1644,504	1625,710	247,685	0,132	1635,942	167,597	0,102	435,924	0,265	286,800
49	PAC58	990,854	983,021	124,353	0,112	990,130	37,959	0,038	114,382	0,115	81,170
50	PAC59	1062,029	1059,468	73,505	0,065	1060,592	55,184	0,052	130,820	0,123	295,100
51	PAC60	894,334	888,361	103,171	0,104	892,400	58,804	0,066	192,469	0,215	323,700
52	PAC61	1703,848	1693,196	190,165	0,101	1701,408	91,106	0,054	224,534	0,132	131,700
53	PAC62	1962,201	1948,504	231,148	0,106	1949,976	218,692	0,112	580,263	0,296	615,900
54	PAC63	2693,665	2685,196	212,846	0,073	2692,359	83,939	0,031	256,738	0,095	132,800
55	PAC64	1882,260	1865,011	253,252	0,120	1881,485	53,831	0,029	154,805	0,082	34,750
56	PAC66	1729,796	1719,998	183,273	0,096	1728,818	58,206	0,034	193,836	0,112	101,900
57	PAC67	1420,740	1411,622	160,362	0,102	1419,701	54,336	0,038	160,323	0,113	95,430
58	PAC68	1008,958	1001,218	124,653	0,111	1008,666	24,176	0,024	85,718	0,085	43,890
59	PAC69	1422,275	1412,491	166,527	0,105	1420,197	76,890	0,054	189,757	0,133	122,300
60	PAC70	1111,903	1104,918	124,271	0,101	1110,857	48,280	0,043	132,554	0,119	104,500
61	PAC71	2132,088	2116,921	253,886	0,107	2129,399	107,026	0,050	265,463	0,125	103,400
62	PAC72	2277,010	2267,280	210,199	0,085	2275,084	93,575	0,041	259,542	0,114	141,500
63	PAC73	1290,295	1277,793	179,119	0,123	1287,481	85,126	0,066	229,548	0,178	150,500
64	PAC74	1256,956	1242,426	188,989	0,132	1254,336	81,099	0,065	276,895	0,220	191,400

## ANEXO 5:

Tabela A5.1: Valores dos índices de soproclassicos e propostos para os pacientes do segundo procedimento de aquisição:

Pacientes	RMS Sinal	RMS sinal janelado	RMS C1	RMS D1	rms D1-deriv2	PB:rms 0-6kHz	PA:rms 6-12.5kHz	D1/C1	pa/pb	Índice Br	D1"/rms sinal
PAC01	884.77	541.72	505.84	170.72	503.59	539.77	45.93	0,337	0,085	13.23	0.930
PAC02	643.44	377.7	370.26	66.07	139.8	377.54	10.93	0,178	0,029	6.58	0.370
PAC04	1397.92	859.61	832.2	187.8	587.78	840.95	178.13	0,226	0,212	27.54	0.684
PAC05	1475.23	904.69	848.99	272.17	768.15	887.6	174.99	0,321	0,197	32.38	0.849
PAC06	1199.26	730.32	708.46	154.41	547.72	719.85	123.22	0,218	0,171	29.52	0.750
PAC08	1839.48	1054.26	896.4	482.07	1708.65	1008.58	306.95	0,538	0,304	27.63	1.621
PAC09	1680.66	930.33	897.55	216.88	492.64	928.39	59.98	0,242	0,065	5.34	0.530
PAC10	2055.3	1302.24	1294.33	127.24	232.32	1301.16	52.95	0,098	0,041	11.42	0.178
PAC11	910.35	570.63	534.5	172.88	558.61	559.55	109.42	0,323	0,196	31.56	0.979
PAC13	438.9	260.12	232.25	99.56	291.31	235.22	110.94	0,429	0,472	74.49	1.120
PAC14	1226.09	749.51	719.59	185.14	465.41	747.36	56.75	0,257	0,076	9.64	0.621
PAC15	774.81	473.67	437.41	158.11	527.69	461.57	106.20	0,361	0,230	26.1	1.114
PAC17	905.61	584.96	538.63	200.11	680.87	578.41	87.2	0,372	0,151	18.04	1.164
PAC18	1114.22	668.4	650.82	139.19	353.03	666.29	53.05	0,214	0,080	11.89	0.528
PAC19	1050.84	655.35	650.02	72.75	179.82	654.19	39.09	0,112	0,080	30.58	0.274
PAC21	744.6	463.07	370.77	236.47	775.1	420.88	193.09	0,638	0,060	62.45	1.674
PAC23	534.28	358.48	347.29	76.84	247.02	351.47	70.49	0,221	0,459	35.23	0.689
PAC24	890.89	456.53	434.45	118.81	335.87	438.27	127.76	0,273	0,201	59.02	0.736
PAC25	2402.93	1632.68	1381.31	759.87	2679.06	1594.79	349.42	0,550	0,219	20.25	1.641
PAC26	2059.49	1245.9	1224.02	205.69	492.96	1245.42	34.57	0,168	0,028	6	0.396
PAC27	1335.01	792.21	779.36	123.13	370.96	784.01	113.64	0,158	0,145	41.63	0.468
PAC28	2094.18	1255.64	1174.1	389.28	1265.23	1229.13	256.55	0,332	0,209	22.1	1.008
PAC29	1205.3	672.08	655.32	131.65	264.86	670.12	51.26	0,201	0,076	21.15	0.394
PAC30	1594.85	972.34	957.04	150.03	391.51	965.82	112.4	0,157	0,116	32.09	0.403
PAC31	717.55	450.26	401.05	176.48	568.59	430.96	130.4	0,440	0,303	41.94	1.263
PAC32	1359.8	845.34	760.13	321.67	1147.39	823.35	191.53	0,423	0,233	26.75	1.357
PAC33	965.67	578.85	559.36	130.3	408.27	570.93	95.36	0,233	0,167	24.81	0.705
PAC34	1288.53	537.28	440.44	260.04	844.66	432.4	318.85	0,590	0,737	72.71	1.572
PAC35	1601.25	849.02	826.97	168.9	388.09	841.51	112.63	0,204	0,134	22.07	0.457
PAC36	1124.37	652.34	623.84	166.4	522.94	633.67	154.94	0,267	0,245	24.11	0.802
PAC41	1619.03	1012.55	951.05	303.51	1065.41	1002.87	139.66	0,319	0,139	18.56	1.052
PAC42	1255.66	755.23	738.74	136.29	438.63	750.54	84.05	0,184	0,112	30.45	0.581
PAC43	500.78	281.47	270.46	68.29	168.46	279.65	31.93	0,252	0,114	21.97	0.599
PAC44	1274.51	756.85	743.12	125.16	403.43	751.98	85.7	0,168	0,114	18.55	0.533
PAC45	1119.11	676.32	658.65	135.53	352.18	674.41	50.82	0,206	0,075	9.96	0.521
PAC46	1159.82	739.72	607.51	361.85	1317.79	644.43	363.13	0,596	0,563	42.97	1.781
PAC47	1342.47	821.5	789.31	201.57	488.84	821.19	22.59	0,255	0,028	6.03	0.595
PAC48	903.59	556.31	536.39	128.97	374.39	553.11	59.51	0,240	0,108	24.06	0.673
PAC49	1107.65	653.1	645.6	87.07	220.15	651.96	38.38	0,135	0,059	9.29	0.337
PAC50	725.58	468.14	440.31	138.21	457.26	452.22	121.02	0,314	0,268	34.58	0.977
PAC51	979.77	597.08	590.37	70.71	246.79	595.93	36.9	0,120	0,062	29.17	0.413
PAC54	1884.06	1138.38	1119.98	179.12	534.16	1137.36	48.17	0,160	0,042	16.34	0.469
PAC55	1276.82	791.21	758.66	198.96	456.52	790.5	33.55	0,262	0,042	4.28	0.577
PAC56	1277.84	792.84	770.43	165.62	415.84	792.23	31.07	0,215	0,039	6.32	0.524
PAC58	1213.97	612.94	583.49	165.81	435.31	611.53	41.48	0,284	0,068	8.45	0.710
PAC59	1325.95	783.43	772.97	112.51	298.38	782.02	46.92	0,146	0,060	11.53	0.381
PAC60	738.77	447.91	439.24	77.21	204.12	446.73	32.31	0,176	0,072	18.62	0.456
PAC61	738.27	381.54	374.54	63.22	185.3	379.67	37.67	0,169	0,099	39.78	0.486
PAC62	1183.23	728.04	706.09	153.94	507.24	717.16	125.36	0,218	0,175	31.55	0.697
PAC63	1152.5	729.31	720.43	100.77	181.44	729	21.22	0,140	0,029	5.13	0.249
PAC64	997.62	586.9	572.1	115.29	312.66	584.82	49.26	0,202	0,084	9.73	0.533
PAC66	696.72	411.42	396.34	96.81	302.35	410.82	22.19	0,244	0,054	13.95	0.735
PAC67	715.95	405.79	381.91	120.95	362.02	405.16	22.51	0,317	0,056	10.18	0.892
PAC68	600.1	371.02	359.19	82.38	172.45	370.88	9.97	0,229	0,027	4.45	0.465
PAC69	789.64	519.46	503.29	113.4	276.62	516.81	52.36	0,225	0,101	10.52	0.533
PAC70	750.66	466.92	441.1	135.09	366.83	465.08	41.36	0,306	0,089	7.93	0.786
PAC72	1330.2	827.99	804.86	171.96	425.31	826.3	52.91	0,214	0,064	7.66	0.514
PAC73	1116.55	628.48	580.39	211.57	683.23	623.95	75.28	0,365	0,121	12.84	1.087
PAC74	1081.55	632.87	607.35	157.28	348.4	631.51	41.44	0,259	0,066	13.18	0.551

## ANEXO 6: O Coeficiente de Correlação Linear:

Uma das formas que se tem de verificar se 2 variáveis que foram medidas simultaneamente estão inter-relacionadas e, medir o grau de associação entre estas 2 variáveis é o "Coeficiente de Correlação Linear".

Suponha que  $n$  pares de medidas  $(x_n, y_n)$  são feitas em 2 variáveis aleatórias  $X$  e  $Y$ , simultaneamente. O Coeficiente de Correlação é calculado por:

$$r = \frac{\sum (x_n - \bar{x})(y_n - \bar{y})}{\sqrt{\left\{ \sum (x_n - \bar{x})^2 \right\} \left\{ \sum (y_n - \bar{y})^2 \right\}}}$$

O coeficiente de correlação assume valores  $-1 \leq r \leq 1$ , 1 indicando que a relação entre as 2 variáveis é perfeitamente linear, ou seja, todos os pontos de um diagrama de espalhamento caem numa reta com declividade positiva,  $-1$  indicando que todos pontos caem em uma reta com declividade negativa e  $,0$ , não existe correlação entre as variáveis.

O "verdadeiro coeficiente de correlação" é dado pela expressão:

$$r = \frac{S_{X,Y}}{S_X \cdot S_Y}$$

onde  $S_{X,Y}$  é a Covariância de  $X$  e  $Y$  e definido por  $S_{XY} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$ ;  $S_X$  e  $S_Y$  são

os Desvios Padrões das variáveis  $X$  e  $Y$ , dados por  $S_X = \sqrt{\left[ \frac{\sum (x_i - \bar{x})^2}{n-1} \right]}$ ,

$S_Y = \sqrt{\left[ \frac{\sum (y_i - \bar{y})^2}{n-1} \right]}$ . Assim, o coeficiente de correlação fica  $r = \frac{1}{n-1} \sum_{i=1}^n \left( \frac{x_i - \bar{x}}{S_X} \right) \left( \frac{y_i - \bar{y}}{S_Y} \right)$ ,

que é idêntico à expressão inicial.

É freqüentemente necessário se fazer um teste de hipótese para se verificar se o coeficiente de correlação linear obtido é significativamente diferente de 0. É mostrado que se o verdadeiro coeficiente de correlação é 0, então a estatística  $\frac{r\sqrt{(n-2)}}{\sqrt{(1-r^2)}}$  tem uma distribuição t de student com n-2 graus de liberdade.

Assim, o coeficiente de correlação será  $\neq 0$  ao nível de significância  $\alpha$ , se

$$\left| \frac{r\sqrt{(n-2)}}{\sqrt{(1-r^2)}} \right| \geq t_{\alpha/2; (n-2)}$$

A tabela abaixo dá os valores críticos absolutos de coeficiente de correlação linear para diferentes tamanhos amostrais:

*Tabela A6.1: valores críticos absolutos em função do tamanho da amostra*

Tamanho amostral n:	Valor crítico absoluto de r:
5	0,88
10	0,63
15	0,51
20	0,44
25	0,39
30	0,36
50	0,28
100	0,20