

TESE

1153

ESCOLA FEDERAL DE ENGENHARIA DE ITAJUBÁ

CLASSIFICADOR DE SINAIS DE VOZ
USANDO CONJUNTOS APROXIMADOS

Sandra Bianca Henriques

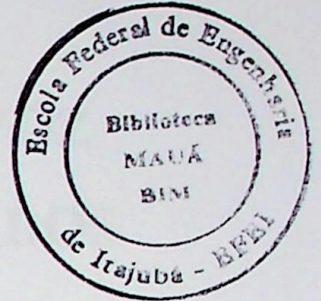
2001

CLASS.	621.391(043.2)
CUTTER.	H518c
TOMBO.	1153



ESCOLA FEDERAL DE ENGENHARIA DE ITAJUBÁ

Programa de Pós-Graduação em Engenharia Elétrica



CLASSIFICADOR DE SINAIS DE VOZ USANDO CONJUNTOS APROXIMADOS

Dissertação apresentada à Escola
Federal de Engenharia de Itajubá para
obtenção do título de Mestre em
Engenharia Elétrica.

Sandra Bianca Henriques

Itajubá/MG

2001

AGRADECIMENTOS

Os pais Manoel e Elizabeth, meus irmãos Anderson e Leônidas e minha família pelo apoio e incentivo na busca de novos ideais e objetivos.

Os professores, Prof. Dr. Luiz Eduardo Borges de Melo e Prof. Dr. Germano Landini, pelo incentivo, dedicação e orientação, dando suporte, paciência e apoio.

E a Vagner De Lencastre, pelo apoio, incentivo e ajuda nas atividades de pesquisa.

Meus amigos e familiares e por terem possibilitado a realização de meus estudos e pesquisas.

Aqueles que contribuíram de alguma forma para a realização desta obra.

Aos meus pais.

CLASS.	621.391(043.2)
CUTTER	H518.c
TOMBO	1153



ESCOLA FEDERAL DE ENGENHARIA DE ITAJUBÁ
Programa de Pós-Graduação em Engenharia Elétrica



USANDO CONJUNTOS APROXIMADOS CLASSIFICADOR DE SINAIS DE VOZ

Dissertação apresentada à Escola
Federal de Engenharia de Itajubá para
obtenção do título de Mestre em
Engenharia Elétrica.

Sandra Bianca Henriques

Itajubá/MG

2001

Sandra Bianca Henriques

**CLASSIFICADOR DE SINAIS DE VOZ
USANDO CONJUNTOS APROXIMADOS**

Dissertação apresentada à Escola Federal
de Engenharia de Itajubá para obtenção
do título de Mestre em Engenharia Elétrica.

Área de concentração:

Automação e Sistemas Elétricos Industriais

Orientador:

Luiz Eduardo Borges da Silva

Co-orientador:

Germano Lambert Torres

Itajubá/MG

2001

AGRADECIMENTOS

Aos meus pais Manoel e Elizabeth; meus irmãos Anderson e Leonardo e minha cunhada Maria pelo apoio e incentivo na busca de meus ideais e objetivos.

Aos meus orientadores, Prof. Dr. Luiz Eduardo Borges de Sousa e Prof. Dr. Germano Lorenzini Torres, exemplos de dedicação e virtude, pelo incentivo, participação e apoio.

A Eliana e Vinícius De Lencastre pela amizade, atenção e apoio nos momentos difíceis.

Aos professores, pelo ensino e por terem possibilitado a ampliação de meus conhecimentos.

A todos aqueles que contribuíram direta ou indiretamente.

Aos meus pais.

AGRADECIMENTOS

Aos meus pais Manoel e Elizabeth, meus irmãos Anderson e Leandro e minha cunhada Marília pelo apoio e incentivo, na busca de meus ideais e objetivos.

Aos meus orientadores, Prof. Dr. Luiz Eduardo Borges da Silva e Prof. Dr. Germano Lambert Torres, exemplo de dedicação e virtude, pelo incentivo, participação e apoio.

À Eliane e Vitorio De Lorenci, pela amizade, atenção e apoio nos momentos difíceis.

Aos professores, pelo auxílio e por terem possibilitado a ampliação de meus conhecimentos.

A todos aqueles que contribuíram direta ou indiretamente.

Este trabalho apresenta o desenvolvimento de um classificador baseado na técnica de Conjuntos Aproximados ("Rough Sets"), aplicado a sinais de voz que se propõe a auxiliar no reconhecimento de um locutor quando o mesmo pronuncia uma dada vogal.

O trabalho está dividido basicamente em duas etapas: a primeira é responsável pela extração de características dos sinais de voz e a segunda etapa está relacionada ao desenvolvimento do classificador, a qual incorpora o tratamento das informações obtidas na primeira etapa, sendo ambas relevantes para a classificação de padrões de informações.

Na primeira etapa foram implementados alguns métodos, relacionados à área de Processamento de Sinais, a fim de extrair parâmetros característicos dos sinais de voz, sendo um deles as frequências fundamentais. Os parâmetros obtidos estão relacionados às características físicas dos locutores envolvidos no conjunto de treinamento e foram usados para estruturar o Sistema de Informação ou a Base de Conhecimento. O Sistema de Informação foi estruturado em formato de banco de dados, sendo este capaz de armazenar um conjunto de padrões ou informações. Estes padrões, também referenciados como um conjunto de exemplos, foram obtidos diretamente da aquisição da voz dos diversos locutores.

A segunda etapa foi responsável pelo desenvolvimento do classificador e se baseou na teoria de Conjuntos Aproximados, que permitiu o tratamento de aspectos como redundância, irrelevância e incerteza, encontradas nas informações extraídas dos sinais de voz.

A teoria de Conjuntos Aproximados foi aplicada ao Sistema de Informação para extrair conhecimento relevante das características obtidas dos sinais de voz e estabelecer regras capazes de classificar uma nova informação ou padrão de dados, permitindo conseqüentemente o reconhecimento de locutor. Assim, a idéia fundamental incorporada por essa etapa é a transformação de um conjunto de exemplos, obtidos previamente, em um conjunto de regras utilizadas para classificação.

Para isso, foram realizados tratamentos no banco de dados tais como: redefinição de valores em faixas de valores e redução de informação redundante.

Os resultados foram obtidos com a utilização de diferentes grupos de locutores para estruturar os Sistemas de Informação, os quais permitiram obter um conjunto de regras que foram usadas para a classificação de novos padrões relacionados aos locutores.

This work presents a classifier based on Rough Sets technique which is applied on voice signals to help in the future speaker recognition.

This work is divided in two parts: the first one is responsible for the extraction of characteristics of the voice signals, and the second is related with the classifier's development, which incorporates the treatment of the information obtained in the first part.

In the first part some methods were implemented, related to the Digital Signal Processing, in order to extract characteristic parameters of the voice signals, just as the fundamental frequency. The obtained parameters are related with the physics characteristics of the speakers involved in training groups, and the parameters were used to structure Information System or Knowledge Base. The Information System was structured in database format capable to store a set of patterns or information.

The second part was responsible for the development of the classifier based on Rough Sets technique. This technique allow the treatment of aspects as redundancy, irrelevance, and uncertainty, found in the extracted information of the voice signals.

The theory of Rough Sets was applied in databases to extract knowledge of the obtained characteristics of the voice signals, and to establish rules capable to classify a new information or pattern of data, and consequently allowing the speaker recognition. Thus, the fundamental idea incorporated by this part is the transformation of a set of examples, obtained previously, in a set of rules used for classification. For sake of explanation, some reductions and treatments were made in the databases.

The results had been obtained for different groups of speakers. The information System derived from the exemplos have been used to obtain a set of rules that was used for the classification of new patterns related to speakers.

SUMÁRIO

Agradecimentos.....	iv
Resumo.....	v
Abstract.....	vi
Lista de Figuras.....	x
Lista de Tabelas.....	xii
Capítulo 1 – Introdução.....	1
1.1 – Considerações Iniciais.....	1
1.2 – O problema.....	2
1.3 – Classificação da Tecnologia de Reconhecimento de Locutor.....	3
1.4 – Objetivo.....	4
1.5 – Estrutura Básica do Sistema Proposto.....	4
1.6 – Organização da Dissertação.....	5
Capítulo 2 – Aparelho Vocal Humano.....	6
2.1 – Introdução.....	6
2.2 – Constituição do Aparelho Vocal Humano.....	6
2.3 – Funcionamento do Aparelho Fonador.....	8
2.4 – Produção da Voz.....	8
2.5 – O Sinal de Voz.....	10
2.6 – Parâmetros de Representação da Voz.....	12
Capítulo 3 - Transformada Rápida de Fourier.....	14
3.1 – Introdução.....	14
3.2 – Considerações sobre a Análise de Fourier.....	14
3.3 – Transformada de Fourier.....	15
3.4 – Transformada Discreta de Fourier.....	17
3.5 – Discretização da Transformada de Fourier.....	19
3.6 – Transformada Rápida de Fourier (FFT).....	20
3.6.1 – Desenvolvimento do Algoritmo da Transformada Rápida de Fourier....	21
3.6.2 – Propriedades dos gráficos de fluxo de sinais da FFT.....	27
3.6.3 – Determinação do Fator W^P	31
3.6.4 – Ordenação da FFT.....	31
Capítulo 4 – Métodos de Análise do Sinal de Voz.....	33

4.1 – Introdução.....	33
4.2 – Determinação da Periodicidade.....	33
4.3 – Análise de Picos.....	36
4.4 – Análise da Energia.....	37
4.5 – Taxa de Cruzamento de Zeros.....	38
4.6 – Autocorrelação.....	39
4.7 – Considerações sobre o “Lag”.....	42
4.8 – Análise por Predição Linear.....	42
4.9 – Análise Espectral.....	43
4.10 – Processamento Homomórfico do Sinal de Voz.....	44
4.11 – Análise Cepstral.....	44
4.12 – Coeficientes Mel Cepstrais (“Mel Frequency Cepstrum”).....	48
4.12.1 – Considerações sobre a Escala Mel.....	49
Capítulo 5 – Aquisição das Características do Sinal de Voz.....	51
5.1 – Introdução.....	51
5.2 – Representação Digital dos Sinais de Voz.....	51
5.2.1 – Preparação dos Dados.....	53
5.2.2 – Tratamento dos Sinais de Voz.....	54
5.3 – Extração das Características.....	56
5.4 – Armazenamento de Padrões.....	58
5.5 – Tabelas de Dados Contendo Características dos Locutores.....	59
Capítulo 6 – Teoria dos Conjuntos Aproximados (“Rough Sets”).....	70
6.1 – Introdução.....	70
6.2 – Teoria de Conjuntos Aproximados.....	71
6.2.1 – Conjuntos e Relações.....	72
6.2.2 – Conjuntos Aproximados.....	72
6.2.3 – Classificação Aproximada de Um Conjunto.....	75
6.2.4 – Regiões em um Conjunto Aproximado.....	77
6.2.5 – Propriedades das Aproximações em um Espaço Aproximado.....	80
6.2.6 – Igualdade Aproximada de Conjuntos.....	80
6.2.7 – Precisão de uma Aproximação.....	81
6.2.8 – Independência de Atributos.....	82
6.3 – Sistemas de Representação de Conhecimento (SRC).....	83

6.4 – Tabelas de Decisão.....	85
6.5 – Dependência entre atributos abordada como relação entre condição e decisão....	86
Capítulo 7 – Algoritmos Baseados na Teoria dos Conjuntos Aproximados.....	95
7.1 – Introdução.....	95
7.2– Classificação dos Dados em Faixas.....	96
7.3 – Simplificação da Tabela de Decisão.....	97
7.4 – Grau de Dependência.....	98
7.5 – Obtenção de Redutos.....	101
7.6 – Composição do Conjunto de Regras.....	103
Capítulo 8 – Aplicação e Resultados.....	105
8.1 – Introdução.....	105
8.2 – Aplicação.....	105
8.3 – Treinamento de Conjuntos de Dados.....	106
8.4 – Testes e Resultados.....	119
Capítulo 9 – Conclusão.....	126
9.1 – Principais Considerações.....	126
9.2 – Sugestões para Trabalhos Futuros.....	128
Apêndice A – Faixas usadas para Redefinir os Atributos Observados.....	130
Apêndice B – Sistemas de Representação de Conhecimento com Atributos Redefinidos.	137
Apêndice C – Regras obtidas pelos Conjuntos de Treinamento.....	146
Apêndice D – Regras Mínimas dos Conjuntos de Treinamento.....	160
Apêndice E – O Sistema.....	165
Referências Bibliográficas.....	170

Lista de Figuras

Figura 2.1 - Aparelho Vocal Humano.....	7
Figura 2.2 - Modelo de Produção de Voz em Tempo Discreto.....	9
Figura 2.3 - Forma de onda de um sinal de voz.....	11
Figura 3.1 - Interpretação da Transformada de Fourier.....	16
Figura 3.2 - Comparação entre FFT e cálculo direto, quanto ao número de multiplicações.....	25
Figura 3.3 - Fluxos de computação dos vetores para $N=4$	26
Figura 3.4 - Exemplo de Nós Duais.....	28
Figura 3.5 - Exemplo de quais nós se deve pular devido computação desnecessária.....	30
Figura 3.6 - Exemplo de “ <i>Bit-Reversing</i> ” para $N=16$	32
Figura 4.1 - Segmento Vozeado obtido a partir da pronúncia do fonema /a/.....	34
Figura 4.2 - Segmento Não Vozeado obtido a partir da pronúncia do fonema /s/.....	34
Figura 4.3 - Faixas de frequência da Frequência Fundamental Masculina e Feminina.....	35
Figura 4.4 - Exemplo de janelamento de um sinal de voz.....	38
Figura 4.5 - Autocorrelação de um Segmento Vozeado.....	40
Figura 4.6 - Autocorrelação de um Segmento Vozeado, contendo ressonâncias.....	40
Figura 4.7 - Exemplo de aplicação do “center clipping”.....	41
Figura 4.8 - Modelo de análise através de predição linear.....	43
Figura 4.9 - Cepstro obtido a partir de um segmento de 25 ms ilustrado na figura 4.1.....	47
Figura 4.10 - Cálculo dos Coeficientes Mel Cepstrais (MFCCs).....	48
Figura 4.11 - Exemplo de magnitude do espectro do banco de filtros.....	50
Figura 5.1 - Reconhecimento de Locutor em Diagrama de Blocos.....	51
Figura 5.2 - Exemplo de um sinal de voz	55
Figura 5.3 - Aplicação da janela de Hamming para um sinal de voz.....	56
Figura 5.4 - Densidade espectral de um segmento vozeado.....	58
Figura 5.5 - Frequências de maior relevância.....	59
Figura 6.1 - Conjunto Procurado, Aproximação Superior e Inferior.....	71
Figura 6.2 - Espaço Aproximado $A=(U,R)$	73
Figura 6.3 - Aproximação Inferior de X em $A=(U,R)$	76
Figura 6.4 - Aproximação Superior de X em $A=(U,R)$	76
Figura 6.5 – Região Positiva.....	77
Figura 6.6 – Região Negativa.....	78

Figura 6.7 - Região Duvidosa.....	78
Figura 6.8 - Espaço aproximado e $\text{pos}(X)$, $\text{neg}(X)$ e $\text{dub}(X)$	79
Figura 6.9 - Conjuntos X_1 , X_2 , e X_3 , com a mesma aproximação inferior e superior.....	81
Figura 6.10 - Região positiva, ou conjuntos induzidos por atributos de condição e decisão.....	87
Figura 6.11 - Espaço aproximado $A=(U, \tilde{C})$ a partir do SRC fornecido pela tabela 6.7.....	91
Figura 6.12 - Espaço aproximado considerando atributos de condição.....	93
Figura 6.13 - Espaço aproximado considerando atributos de condição e decisão.....	94
Figura 8.1 - Resultado da tabela 8.12.....	121
Figura 8.2 - Resultado da tabela 8.14.....	122
Figura 8.3 - Resultado da tabela 8.16.....	123
Figura 8.4 - Resultado da tabela 8.18.....	125
Figura E.1 - Tela Principal do Sistema.....	165
Figura E.2 - Lista de arquivos do tipo WAV.....	166
Figura E.3 - Exibição do sinal de voz.....	166
Figura E.4 - Sinal segmentado.....	167
Figura E.5 - Características contida no sinal.....	167
Figura E.6 – Exemplo de uma tabela de dados.....	168
Figura E.7 – Exemplo de valores redefinidos de uma tabela.....	168
Figura E.8 – Busca de redutos.....	169

Lista de Tabelas

Tabela 2.1 – Partes do Aparelho Vocal Humano.....	7
Tabela 5.1 - Lista de Locutores.....	53
Tabela 5.2 - Tabela de dados com características do locutor Jusefina.....	60
Tabela 5.3 - Tabela de dados com características do locutor Sandra.....	61
Tabela 5.4 - Tabela de dados com características do locutor Marília.....	62
Tabela 5.5 - Tabela de dados com características do locutor Anderson.....	63
Tabela 5.6 - Tabela de dados com características do locutor Marcelo.....	64
Tabela 5.7 - Tabela de dados com características do locutor Mauro.....	65
Tabela 5.8 - Tabela de dados com características do locutor Maria José.....	66
Tabela 5.9 - Tabela de dados para teste com características de diversos locutores.....	67
Tabela 6.1 - Conjunto de Dados.....	74
Tabela 6.2 - Agrupamento de Objetos Idênticos, considerando a_1 , a_2 , e a_3	74
Tabela 6.3 - Agrupamento de Objetos Idênticos, considerando a_1 , e a_2	75
Tabela 6.4 - Verificação da Independência de Atributos.....	82
Tabela 6.5 - Exemplo de Sistema de Representação de Conhecimento.....	84
Tabela 6.6 - Exemplo de Tabela de Decisão, onde $C=\{a, b, c, d\}$ representa a classe de condições e $D=\{e\}$ representa a classe de decisão.....	86
Tabela 6.7 - SRC de avaliação de carros.....	88
Tabela 6.8 - Análise de dependência dos atributos de condição em relação a decisão.....	90
Tabela 6.9 - Conjuntos elementares de D'	90
Tabela 6.10 - SRC para atributos de condição Preço e Velocidade Máxima.....	91
Tabela 6.11 - Análise de dependência dos atributos de condição em relação à decisão avaliação.....	92
Tabela 6.12 - SRC para atributos de condição Quilometragem e Velocidade Máxima.....	92
Tabela 7.1 - Exemplo de Base de Dados.....	96
Tabela 7.2 - SRC com redefinição dos valores dos atributos.....	97
Tabela 7.3 - SRC após eliminar exemplos idênticos.....	98
Tabela 7.4 - SRC contendo apenas atributos indispensáveis.....	100
Tabela 7.5 – Regras obtidas a partir do SRC representado pela tabela 7.1.....	104
Tabela 8.1 - Sistema de Representação de Conhecimento do conjunto de treinamento1.....	106
Tabela 8.2 - Exemplo de identificação de exemplos semelhantes.....	110

Tabela 8.3 - SRC após passos 1, 2, 3 e 4.....	110
Tabela 8.4 - Redutos iniciais do conjunto de treinamento 1.....	112
Tabela 8.5 - Conjuntos mínimos do conjunto de treinamento 1.....	112
Tabela 8.6 - SRC agrupado por F1, F2, Saída (sem padrões repetidos).....	113
Tabela 8.7 - SRC do conjunto de treinamento 2 após passos 1, 2, 3 e 4.....	115
Tabela 8.8 - Conjuntos mínimos do conjunto de treinamento 2.....	116
Tabela 8.9 - SRC do conjunto de treinamento 3 após passos 1, 2, 3 e 4.....	117
Tabela 8.10 - Conjuntos mínimos do conjunto de treinamento 3.....	119
Tabela 8.11 - Teste de classificação para locutores do conjunto de treinamento 1.....	120
Tabela 8.12 - Resumo da tabela 8.11.....	121
Tabela 8.13 - Novo teste de classificação para locutores do conjunto de treinamento 1.....	121
Tabela 8.14 - Resumo da tabela 8.13.....	122
Tabela 8.15 - Teste de classificação para locutores do conjunto de treinamento 2.....	122
Tabela 8.16 - Resumo da tabela 8.15.....	123
Tabela 8.17 - Teste de classificação para locutores do conjunto de treinamento 3.....	124
Tabela 8.18 - Resumo da tabela 8.17.....	125
Tabela A.1 - Faixa de valores para o atributo F1.....	130
Tabela A.2 - Faixa de valores para o atributo F2.....	131
Tabela A.3 - Faixa de valores para o atributo F3.....	131
Tabela A.4 - Faixa de valores para o atributo F4.....	132
Tabela A.5 - Faixa de valores para o atributo F5.....	133
Tabela A.6 - Faixa de valores para o atributo F6.....	134
Tabela A.7 - Faixa de valores para o atributo F7.....	135
Tabela A.8 - Faixa de valores para o atributo Ffund.....	136
Tabela B.1 - SRC do conjunto de treinamento 1 com valores redefinidos.....	137
Tabela B.2 - SRC do conjunto de treinamento 2 com valores redefinidos.....	139
Tabela B.3 - SRC do conjunto de treinamento 3 com valores redefinidos.....	141
Tabela C.1 - Total de regras do conjunto de treinamento 1.....	146
Tabela C.2 - Total de regras do conjunto de treinamento 2.....	152
Tabela C.3 - Total de regras do conjunto de treinamento 3.....	155
Tabela D.1 - Regras mínimas do conjunto de treinamento 1.....	160
Tabela D.2 - Regras mínimas do conjunto de treinamento 2.....	161
Tabela D.3 - Regras mínimas do conjunto de treinamento 3.....	162

INTRODUÇÃO

Este trabalho apresenta uma contribuição na tecnologia do reconhecimento de locutor, a qual envolve o manuseio de sinais de voz e a obtenção das características para a estruturação de um sistema capaz de classificar tais sinais.

O sistema é composto de duas importantes etapas: a extração das características dos sinais de voz e o desenvolvimento do classificador. A extração das características envolve a utilização de técnicas de Processamento de Sinais, enquanto o desenvolvimento do classificador envolve a utilização de uma técnica da Inteligência Artificial que usa a ferramenta matemática denominada Teoria dos Conjuntos Aproximados (“Rough Sets”).

1.1 – Considerações Iniciais

A área de processamento de sinais, auxiliada pelo grande avanço tecnológico, tem apresentado um aumento significativo em pesquisas, principalmente na tecnologia de reconhecimento envolvendo a voz humana, sendo esta uma área que possui grande potencialidade em diversas aplicações.

O uso da voz humana em pesquisas científicas tem recebido destaque em duas situações: no reconhecimento de voz e no reconhecimento de locutor. O reconhecimento de voz consiste em reconhecer palavras isoladas, frases ou textos, enquanto que o reconhecimento de locutor consiste em reconhecer uma pessoa a partir de sua voz, onde esse último tipo de reconhecimento está associado a inúmeras aplicações, tais como: serviços de segurança, transações bancárias via telefone, “voice mail”, controle de acesso a informações confidenciais, acesso remoto a computadores, entre outros [1].

Esse trabalho propõe uma contribuição para a segunda situação, o reconhecimento de locutor.

De acordo com o desenvolvimento das pesquisas e o avanço da tecnologia, pode-se identificar três fases para o reconhecimento de locutor, são elas: reconhecimento de locutor por ouvinte, reconhecimento de locutor por inspeção visual, e o reconhecimento automático de locutor realizado por computador. Cabe ressaltar que, para todas as situações são usados padrões de informação compostos por atributos ou variáveis que caracterizam um locutor,

onde tais padrões são utilizados em duas etapas, a primeira de treinamento e a segunda de reconhecimento.

O reconhecimento de locutor por ouvinte surgiu há muitas décadas atrás quando não se dispunha de tecnologia para tarefas de reconhecimento, sendo este método responsável pelas primeiras pesquisas envolvendo a voz humana [2]. Cabe ressaltar que, esse tipo de reconhecimento apresenta grande robustez quanto a ruídos. Entretanto sua subjetividade e os avanços tecnológicos foram os responsáveis pelo desenvolvimento de outros métodos para o reconhecimento. Já o reconhecimento de locutor por inspeção visual é realizado por meio de espectrogramas, conhecidos também como “voiceprint”, os quais permitem representar graficamente o espectro de frequência do sinal de voz e suas mudanças ao longo do tempo [3]. Porém, é importante lembrar que, o bom desempenho para esse tipo de reconhecimento irá depender da experiência do pesquisador responsável pela inspeção visual [4].

O reconhecimento automático de locutor realizado por computador, ou RAL, surgiu devido aos avanços na área computacional e na área de processamento de sinais, aspecto esse que permitiu o uso de diversas técnicas tanto para representação, quanto para estruturação e auxílio aos sistemas de reconhecimento. A tecnologia de RAL pode ainda ser dividida em: identificação de locutor e verificação de locutor, que são classificações quanto ao modo de reconhecimento [5].

1.2 – O problema

Reconhecer uma pessoa através de sua voz pode parecer uma tarefa simples, quando se considera a identificação de um locutor a partir da audição humana, isso devido o ser humano ser capaz de reconhecer, desde o nascimento, as pessoas à sua volta através de suas vozes.

Porém, tal tarefa pode representar um problema, quando se considera o uso de computadores para a classificação de sinais de voz, e conseqüentemente o reconhecimento de locutor.

O problema em questão está relacionado à parametrização de aspectos como a fisiologia de cada pessoa e a maneira de falar, as quais devem ser consideradas quando na utilização de sinais de voz, pois representam aspectos que caracterizam um locutor.

A parametrização da fisiologia e o modo de falar de uma pessoa foram responsáveis pelo desenvolvimento de inúmeros métodos, as quais tentavam e tentam definir quais e como são encontrados os melhores parâmetros de representação. Entretanto, embora o número de

pesquisas seja grande até hoje, não há uma definição de quais características ou parâmetros devem ser utilizados, isso ocorre devido os sinais de voz apresentarem grandes variações, incorporando informações que nem sempre são precisas [4].

Além das questões referentes à voz, devem-se considerar aspectos relacionados à forma de obtenção dos sinais, devido à introdução de ruídos que podem vir a interferir no conteúdo das informações.

Assim, pode-se dizer que a tarefa de reconhecimento de locutor não é um caso simples a ser resolvido. Por isso, existe a necessidade de pesquisar por novos algoritmos ou por algoritmos modificados com o objetivo de tratar a redundância ou irrelevância na representação dos sinais de voz, bem como o conteúdo de informações incertas e imprecisas [6]. Esse trabalho visa a aplicação da teoria de Conjuntos Aproximados em um Sistema de Informação que contém informações sobre sinais de voz.

1.3 – Classificação da Tecnologia de Reconhecimento de Locutor

Como mencionado anteriormente, a tecnologia de reconhecimento de locutor realizada por computador pode ser dividida em: identificação de locutor e verificação de locutor, que são classificações quanto ao modo de reconhecimento.

Na identificação é possível classificar um sinal de voz, ou padrão de informação, a partir de um conjunto de informações. Nesse caso se considera um conjunto de N locutores para a aquisição dos sinais, os quais serão usados como referência para a classificação. Assim, ao adquirir um novo sinal, ou padrão de entrada, este é comparado com sinais ou padrões referentes aos locutores do conjunto mencionado. Após a comparação o sistema deve indicar um valor referente a qual padrão representa melhor a informação de entrada, identificando qual dos locutores de referência pronunciou o sinal de voz.

Na verificação o padrão referente ao sinal de entrada possui um atributo que representa o código de identificação do locutor, além das características que representam o locutor. Nesse caso, o novo sinal de entrada, ou de teste, será comparado com o padrão que possui o mesmo código de locutor. Assim, haverá uma comparação entre o sinal de entrada e os sinais de referência, de forma que o sistema deverá indicar se os dois sinais ou padrões são semelhantes, decidindo se determinado sinal pertence ou não ao locutor referenciado pelo código de identificação.

1.4 – Objetivos

O objetivo principal desse trabalho é o desenvolvimento de um sistema classificador de sinais de voz baseado na Teoria dos Conjuntos Aproximados (“Rough Sets”), bem como a aplicação de técnicas de Processamento de Sinais para a extração das características dos sinais de voz.

O trabalho também se propõe a tratar aspectos como redundância, irrelevância e incerteza, contidas nas informações extraídas dos sinais de voz, as quais são utilizadas para estruturar um Sistema de Informação.

Considerando os aspectos mencionados sobre a verificação e identificação de locutor, o trabalho se propõe a contribuir quando na identificação de um locutor, após a classificação de novos padrões de informação.

1.5 – Estrutura Básica do Sistema Proposto

O sistema proposto se baseia em características do Reconhecimento Automático de Locutor realizado por computador, e possui como estrutura básica duas etapas importantes: a extração das características dos sinais de voz e o desenvolvimento do classificador baseado na Teoria de Conjuntos Aproximados.

A primeira etapa realiza a obtenção dos sinais de voz, que são armazenados em arquivos do tipo WAV, e realiza a extração das características dos sinais obtidos, as quais irão gerar padrões de informação que caracterizam o locutor. Os padrões de informação obtidos são usados para estruturar Sistemas de Representação de Conhecimento, ou SRC, os quais permitem incorporar aspectos como redundância, irrelevância, incerteza e imprecisão na representação dos sinais de voz.

Ainda na primeira etapa são implementados métodos da área de Processamento de Sinais para a extração de parâmetros ou informações referentes aos sinais de voz. Os parâmetros estão relacionados às características físicas dos locutores e são usados para estruturar o Sistema de Informação ou Sistemas de Representação de Conhecimento – SRC.

A segunda etapa realiza o desenvolvimento do classificador e faz uso dos SRCs obtidos na primeira etapa, e se baseia da Teoria dos Conjuntos Aproximados, que é uma ferramenta matemática capaz de executar tarefas como: tratamento de redundância, irrelevância, incerteza e imprecisão de informação incorporada pelos SRCs, redução da base

de conhecimento com a eliminação de informações supérfluas sem perder o poder discriminatório da base de conhecimento original e transformação das informações da base de conhecimento em um conjunto de regras. Cabe ressaltar que, a execução de tais tarefas é realizada por algoritmos, os quais são responsáveis pela comparação de padrões e geração de regras, que permitirão classificar e reconhecer novos padrões de informações, ou novos sinais de voz, referentes a um locutor. Assim, a classificação no sistema proposto irá se basear nos Sistemas de Representação de Conhecimento, já que as informações de um grupo de locutores serão transformadas em um conjunto de regras.

Após realizar tais etapas espera-se que o classificador proposto seja capaz de reduzir a base de conhecimento, bem como classificar novos padrões de sinais de voz. Para isso o locutor responsável pela pronuncia do novo padrão deve estar envolvido em um grupo de locutores.

1.6 – Organização da dissertação

O trabalho está organizado da forma a seguir.

O *Capítulo 2* apresenta as características do Aparelho Fonador Humano que é responsável pela produção da voz.

O *Capítulo 3* descreve a Transformada Rápida de Fourier (FFT), usada para determinar o espectro de frequência de um sinal e que permite extrair algumas características importantes.

O *Capítulo 4* apresenta uma visão geral dos métodos mais usados para análise de sinais de voz, com a finalidade de obter características capazes de representar um locutor.

O objetivo *Capítulo 5* é a descrição da etapa de aquisição e extração das características dos sinais de voz, os quais são usados para estruturar Sistemas de Representação de Conhecimento, ou SRC, para o sistema proposto.

O *Capítulo 6* apresenta conceitos da Teoria de Conjuntos Aproximados, uma ferramenta matemática usada para tratar a incerteza contida em SRCs.

O *Capítulo 7* descreve os algoritmos baseados na Teoria dos Conjuntos Aproximados, a fim de estruturar o classificador de sinais de voz.

No *Capítulo 8* são apresentados os resultados obtidos a partir da aplicação do classificador em sinais de voz.

Finalmente, no *Capítulo 9* conclui-se o trabalho ressaltando os aspectos que contribuíram para o desenvolvimento de um classificador a partir de uma técnica pouco utilizada para sinais de voz, e algumas sugestões para futuros trabalhos.

APARELHO VOCAL HUMANO

2.1 – Introdução

Este capítulo apresenta os principais conceitos envolvendo o processo de elaboração da voz humana, que consiste em um “som ou um conjunto de sons emitidos pelo aparelho fonador” [7]. Além disso, apresenta o Aparelho Fonador Humano que é responsável pela produção da voz.

2.2 – Constituição do Aparelho Vocal Humano

O Aparelho Fonador (Vocal) Humano, mostrado na figura 2.1 [8], é o nome dado ao conjunto dos órgãos responsáveis pela produção de sons, ou formação dos fonemas, incluindo a voz cantada e a voz falada.

As partes que constituem o aparelho fonador são:

- Os **pulmões**, os **brônquios** e a **traquéia**, que são os órgãos respiratórios que fornecem a corrente de ar, matéria-prima da fonação, onde a traquéia funciona como suporte para vibração das cordas vocais;
- A **laringe**: que é o órgão mais importante da fonação, pode ser considerada como um vibrador. É na laringe que se localizam, a glote, e as **cordas vocais** (pregas musculares que produzem o som utilizado na fala);
- As cavidades (**faringe**, **boca** e **fossas nasais**): funcionam como caixa de ressonância, uma vez que a cavidade bucal pode variar de forma e de volume, graças aos movimentos dos órgãos ativos, sobretudo da língua que participa de todos os sons produzidos e é muito importante na fonação.

A figura 2.1 mostra ainda:

- Os **lábios**: que permite a articulação de sons bilabiais (/b/, /p/, /m/) e labiodentais (/f/, /v/);
- O **palato duro** (céu da boca): responsável pela projeção da voz;
- A **cavidade nasal**: responsável pela vibração e amortização do som;
- A **úvula**: responsável pelo acoplamento acústico entre o trato vocal e o trato nasal.

O trato vocal, também chamado de tubo ressonante, é formado pela boca, laringe, e faringe, situando-se entre as cordas vocais e os lábios. O comprimento do trato vocal varia de pessoa para pessoa, no entanto seu comprimento médio em uma pessoa adulta do sexo masculino é de aproximadamente 17 cm, enquanto que para o sexo feminino tem um tamanho de 14 cm.

O trato nasal é considerado como um tubo auxiliar e pode ser conectado ou desconectado ao trato vocal a partir do movimento da úvula e seu comprimento médio é de 12 cm.

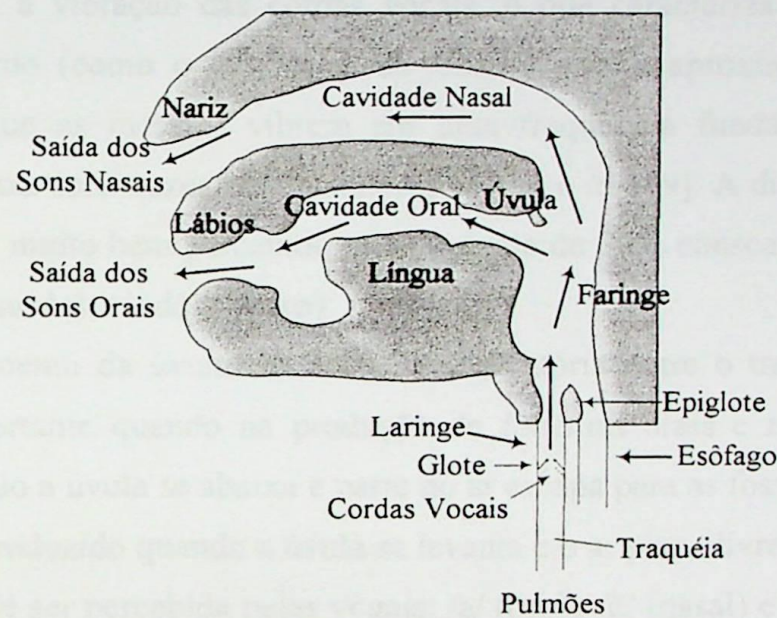


Figura 2.1 - Aparelho Vocal Humano [8]

De forma simplificada é possível dividir o aparelho vocal humano em 4 partes, como mostra a tabela 2.1.

Tabela 2.1 - Partes do Aparelho Vocal Humano

Parte	Componentes	Função
Produtora	Pulmões, músculos (abdominais, intercostais, extensores da coluna), diafragma	Produção de ar que pressiona a laringe, produzindo som nas cordas vocais
Vibradora	Laringe	Produzir som fundamental
Ressonadora	Cavidade nasal, faringe, boca	Ampliar o som
Articuladora	Lábios, língua, palato mole, palato duro, mandíbula	Articulam e dão sentido ao som, transformando sons em orais e nasais

2.3 – Funcionamento do Aparelho Fonador

No aparelho fonador, o ar é expelido dos pulmões por via dos brônquios, penetra na traquéia e chega à laringe. Ao chegar na glote, que fica na altura do pomo de Adão ou gogó, encontra o primeiro obstáculo à sua passagem. Esse obstáculo criado pela glote é capaz de proporcionar a abertura entre as duas pregas musculares das paredes superiores da laringe, conhecidas como cordas vocais, que quando vibram emitem os sons.

Assim, a glote pode se *abrir*, afastando as cordas vocais e permitindo que o ar passe livremente sem a vibração das cordas vocais, o que caracteriza a produção de som não vozeado ou surdo (como o /s/); ou pode *fechar*, com a aproximação das cordas vocais, fazendo com que as mesmas vibrem em uma frequência fundamental, conhecida como “pitch”, e produza sons vozeados ou sonoros (como o /d/) [9]. A distinção entre som surdo e sonoro pode ser muito bem percebido na pronúncia de duas consoantes: /p/ pê (surdo), /b/ bê (sonoro), /t/ tê (surdo), /d/ dê (sonoro).

O movimento da úvula, na corrente expiratória entre o trato vocal e o trato nasal, também é importante quando na produção de fonemas orais e nasais. O fonema nasal é produzido quando a úvula se abaixa e parte do ar escapa para as fossas nasais, enquanto que o fonema oral é produzido quando a úvula se levanta e o ar passa livremente em direção à boca. A diferença pode ser percebida pelas vogais: /a/ (oral), /ã/ (nasal) e conforme as palavras: /a/ lá (oral), /ã/ lâ (nasal), e /a/ mato (oral), /ã/ manto (nasal).

A faringe situa-se entre a parte superior do esôfago e a boca, e conduz o ar para as fossas nasais e a boca, permitindo a emissão e ampliação dos sons.

É preciso ressaltar que os órgãos envolvidos na produção da voz, bem como o sistema de audição humana, possui limitações, onde a banda de frequência típica do sinal de voz é da ordem de 7 à 8 KHz.

2.4 – Produção da Voz

A boa produção vocal depende da integridade do aparelho fonador, além da ação coordenada e precisa da musculatura laríngea, e pode ser criada de diferentes formas [8]:

- *Pela excitação do trato vocal ou nasal*, que dão origem aos sons vozeados a partir de pulsos de ar produzidos pelas cordas vocais, onde os pulsos são quase-periódicos;

- *Pela passagem forçada de ar*, que dão origem aos sons não vozeados (fricativo, como o /s/) a partir da compressão produzida pelos articuladores; seria uma turbulência. A excitação, nesse tipo de criação de som, pode ser comparada ao ruído;
- *Pela expansão inesperada de ar*, que dão origem aos sons explosivos ou oclusivos (podendo ser vozeados, como o /d/, /b/, ou não vozeados, como o /p/, /q/), causada pelo aumento de pressão após o fechamento total do trato vocal seguido de uma repentina abertura.

Como mencionado anteriormente, os sons são produzidos a partir de vibrações das cordas vocais. Essas vibrações, também conhecidas como pulsos glotais ou excitação glotal, são provocadas pelo fluxo de ar proveniente dos pulmões. Assim, as vibrações das cordas vocais sofrem modificações ao passar através dos tratos.

Um modelo para a geração dos sinais de voz pode ser criado considerando um filtro digital variável no tempo controlado por coeficientes que representam parâmetros do trato vocal, como mostra a figura 2.2.

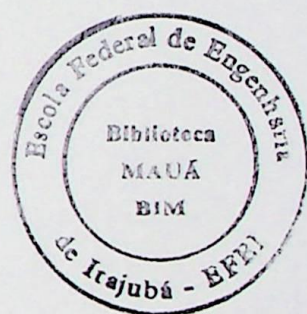
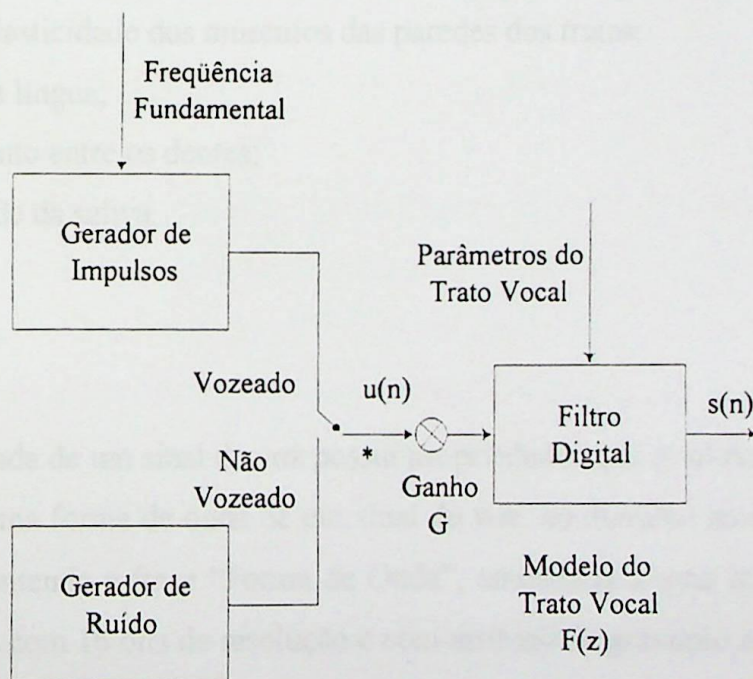


Figura 2.2 - Modelo de Produção de Voz em Tempo Discreto [4]

O filtro do modelo mostrado pela figura 2.2 apresenta uma aproximação da resposta em frequência do trato vocal e pode ser representado por uma função de transferência. O filtro é responsável pela forma do espectro do sinal e por produzir diferentes sons. Assim, as

modificações das vibrações das cordas vocais podem ser representadas através de uma função de transferência para o filtro digital que relacione a excitação das cordas vocais (entrada), com a produção final da voz (saída). Essa função pode ser representada pela equação 2.1 [10].

$$F(z) = \frac{G}{1 - \sum_{k=1}^N \alpha_k \cdot z^{-k}} \quad (2.1)$$

Onde α_k são parâmetros que dependem das características e dimensões do trato vocal e G representa o ganho do sistema.

No caso da produção de voz, o filtro digital $F(z)$ possui características constantes, já que o trato vocal também apresenta características um tanto quanto constantes, para pequenos intervalos de tempo.

O modelamento do aparelho vocal tende a representar características individuais, as quais dependem de uma série de fatores, dentre os quais se destacam [10]:

- Variações nas formas físicas dos tratos vocal e nasal;
- Ressonâncias das cordas vocais;
- Massa e elasticidade dos músculos das paredes dos tratos;
- Posição da língua;
- Espaçamento entre os dentes;
- Viscosidade da saliva.

2.5 – O Sinal de Voz

A forma de onda de um sinal de voz possui propriedades que se alteram com o tempo. A figura 2.3 ilustra uma forma de onda de um sinal de voz, no domínio tempo, com duração de 2,16 segundos, contendo a frase “Forma de Onda”, amostrada a uma taxa de 8KHz com amplitude quantizada com 16 bits de resolução e com atributo de gravação em modo de canal mono.

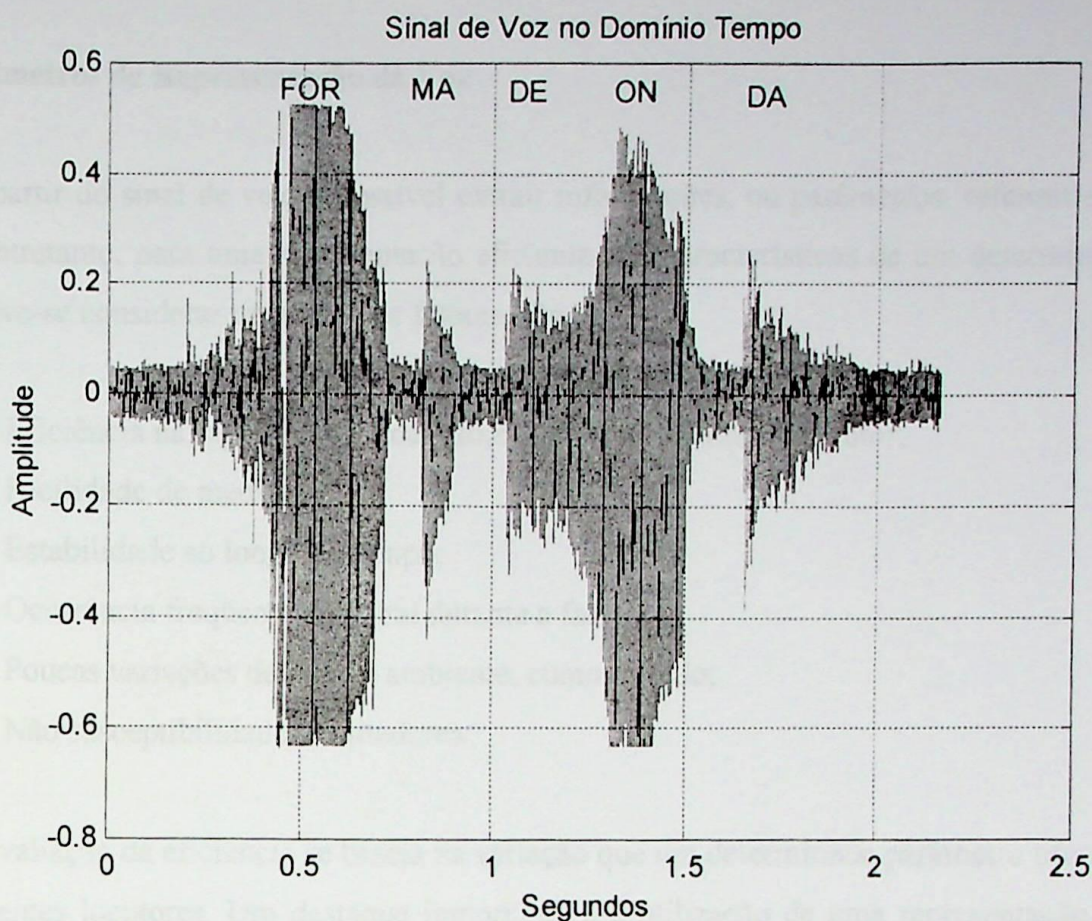


Figura 2.3 - Forma de onda de um sinal de voz

Os sinais de voz são formados pelas frequências de ressonância dos tratos vocal e nasal. Essas frequências são chamadas de frequências formantes ou simplesmente formadoras, as quais variam segundo modificações do formato do trato. Essas modificações ocorrem ao longo do tempo durante a pronúncia das palavras (sons) e aproximadamente em períodos de milisegundos (ms). Em aplicações práticas é comum desconsiderar variações para períodos inferiores a 10 ms [9]. Nos estudos da produção dos diferentes tipos de sons (ditongos, fricativas), um aspecto a ser ressaltado é a importância dos sons das vogais e dos sons nasalados (produzidos a partir de ressonâncias ao longo do trato nasal) na aplicação do Reconhecimento Automático de Locutor. Durante a produção desses sons, praticamente não há variação no formato do trato vocal. Desta forma, as características físicas exclusivas de cada locutor podem ser melhores extraídas para tarefas de reconhecimento. Entretanto, para o caso dos sons nasalados, alterações fisiológicas temporárias, como a gripe, sinusites, entre outros, podem alterar algumas dessas características individuais [5].

2.6 – Parâmetros de Representação da Voz

A partir do sinal de voz, é possível extrair informações, ou parâmetros, referentes ao locutor. Entretanto, para uma representação eficiente das características de um determinado locutor, deve-se considerar uma série de fatores, são eles:

- Eficiência na representação da informação dependente do locutor;
- Facilidade de medida;
- Estabilidade ao longo do tempo;
- Ocorrência freqüente e natural durante a fala;
- Poucas variações devido ao ambiente, como o ruído;
- Não susceptibilidade a imitadores.

A avaliação da eficiência se baseia na variação que um determinado parâmetro oferece entre diferentes locutores. Um destaque importante é a utilização de uma representação de eficiência de parâmetros isolados, a F-ratio. Tal representação (medida) é bastante para a escolha de parâmetros que melhor representam o locutor e pode ser representada pela equação 2.2 [10].

$$F_{\text{ratio}} = \frac{\text{var}(\text{global})}{\text{var}(\text{loc})} \quad (2.2)$$

Onde $\text{var}(\text{global})$ é a variância média do sinal de voz inter-locutor e $\text{var}(\text{loc})$ é a variância média do sinal intra-locutor.

A representação F-ratio é normalmente utilizada para descartar parâmetros que contém pouca informação que caracterize o locutor. Entretanto, devido essa medida ser realizada em parâmetros isolados, as relações entre parâmetros não são consideradas no cálculo da F-ratio.

Os parâmetros acústicos de uma determinada locução podem ser obtidos através de diferentes tipos de processamento, ou métodos. Entre eles pode-se citar:

- Análise de Picos;
- Cálculo da energia do sinal de voz;

- Determinação do “pitch”;
- Análise espectral com banco de filtros;
- Análise espectral de curta duração;
- Análise Cepstral;
- A transformada Wavelet;
- Predição Linear, entre outros.

A Transformada de Fourier é uma ferramenta matemática muito utilizada em diversas áreas, tais como medicina, engenharia, meteorologia, produção, entre outras. Porém, sua aplicação para a Transformada Rápida de Fourier, ou FFT (Fast Fourier Transform), é feita por a mesma maneira de grande quantidade de tempo para a sua implementação e sua taxa é muito baixa para muitas aplicações.

Outra grande ferramenta matemática é a Transformada Rápida de Fourier, ou FFT (Fast Fourier Transform). A FFT é uma versão mais rápida da DFT, sendo a FFT mais rápida e eficiente que a DFT. A FFT é uma ferramenta matemática muito utilizada em diversas áreas, tais como medicina, engenharia, meteorologia, produção, entre outras. Porém, sua aplicação para a Transformada Rápida de Fourier, ou FFT (Fast Fourier Transform), é feita por a mesma maneira de grande quantidade de tempo para a sua implementação e sua taxa é muito baixa para muitas aplicações.

2.1 – Considerações sobre a Análise de Fourier

A análise de Fourier é uma ferramenta matemática muito utilizada em diversas áreas, tais como medicina, engenharia, meteorologia, produção, entre outras. Porém, sua aplicação para a Transformada Rápida de Fourier, ou FFT (Fast Fourier Transform), é feita por a mesma maneira de grande quantidade de tempo para a sua implementação e sua taxa é muito baixa para muitas aplicações.

Especificamente a análise de Fourier é utilizada para analisar a potência de um sinal em função da frequência.

Logo, a análise de Fourier é utilizada para analisar a potência de um sinal em função da frequência.

TRANSFORMADA RÁPIDA DE FOURIER

3.1 – Introdução

Este capítulo tem como objetivo apresentar a Transformada de Fourier e suas propriedades básicas, a fim de proporcionar um melhor entendimento da Transformada Rápida de Fourier, que será o método usado para determinar o espectro de frequência de um sinal [11].

A Transformada de Fourier é uma ferramenta matemática muito utilizada em diversos campos, tais como: medicina, sismologia, manutenção preditiva, entre outros. Porém, essa afirmação não é válida para a Transformada Discreta de Fourier, ou DFT (“Discrete Fourier Transform”), já que a mesma necessita de grande quantidade de tempo para a sua computação, o que torna o seu uso inviável para muitas aplicações.

O algoritmo da Transformada Rápida de Fourier, ou FFT (“Fast Fourier Transform”), foi desenvolvido a fim de minimizar o tempo de computação da DFT. Assim, a FFT representa um algoritmo computacional altamente eficiente, capaz de executar os cálculos inclusos na DFT de forma muito mais rápida [12]. Entretanto, para que se entenda a FFT é preciso considerar princípios observados na Análise de Fourier, os quais são apresentados a seguir.

3.2 – Considerações sobre a Análise de Fourier

A análise de Fourier é uma ferramenta matemática que permite decompor sinais em ondas de senos e co-senos, ou componentes independentes do tempo, aos quais são chamadas de componentes senoidais, que somadas permitem obter o sinal original. Assim, ao utilizar essa ferramenta matemática é possível investigar quais frequências compõem um sinal, já que cada componente senoidal, da soma de senos e co-senos, representa apenas uma frequência. Dessa forma, aplica-se a análise de Fourier em sinais no domínio do tempo, para que se possa investigar suas características no domínio da frequência.

Em relação à análise de Fourier é importante lembrar que a mesma pode ser dividida em duas técnicas, quando se considera a análise de *sinais contínuos*, são elas:

- *Série de Fourier*: utilizada para analisar sinais periódicos e;

- *Transformada de Fourier*: que tem como aplicação básica a análise de sinais que possuem energia finita, tais como um pulso retangular ou uma exponencial decrescente. Entretanto, a Transformada de Fourier também pode ser utilizada para analisar sinais periódicos, ao invés de recorrer à série de Fourier. Contudo, os sinais periódicos possuem energia infinita, e a integral utilizada na transformada de Fourier não teria valor finito, caso não se recorresse ao uso de impulsos, os quais simplificam muitas derivações. Nesse caso, o uso do impulso permite que a transformada aumente seu campo de aplicação, representando também sinais periódicos. Com isso, a transformada de Fourier, apresentada a seguir, acaba superando a série de Fourier em termos de potencial e aplicações [12].

3.3 – Transformada de Fourier

A essência da Transformada de Fourier, assim como a Análise de Fourier, é a decomposição de um sinal em componentes senoidais.

É certo que existem inúmeras técnicas para se decompor um sinal. Entretanto o uso de ondas de seno e co-seno, se justifica pois as mesmas são consideradas mais simples, além de possuírem uma propriedade denominada *fidelidade senoidal*, na qual a frequência e a forma de onda permanecem a mesma, mudando somente a amplitude e a fase do sinal [3].

Cabe lembrar que para a Transformada de Fourier, pode-se definir mais duas técnicas ao analisar *sinais discretos*, as quais diferem das duas técnicas citadas anteriormente (*Série de Fourier* e *Transformada de Fourier*), são elas [3]:

- *Transformada de Fourier de Tempo Discreto*: utilizada para analisar sinais amostrados no tempo ou discretos e que não possuem periodicidade e;
- *Transformada Discreta de Fourier* (DFT): usados em casos que tratam sinais discretos que possuem periodicidade.

Um exemplo da Transformada de Fourier é ilustrado pela figura 3.1, na qual um sinal é decomposto em duas componentes senoidais individuais, cada uma com sua frequência e amplitude. A figura permite visualizar as duas componentes, que somadas originam o sinal inicial. Além disso, é possível visualizar a amplitude e frequência de cada senóide, as quais são obtidas após a transformação de Fourier; por convenção se exhibe também a amplitude e frequência negativas. Assim, a Transformada de Fourier permite representar uma função ou sinal no domínio frequência.

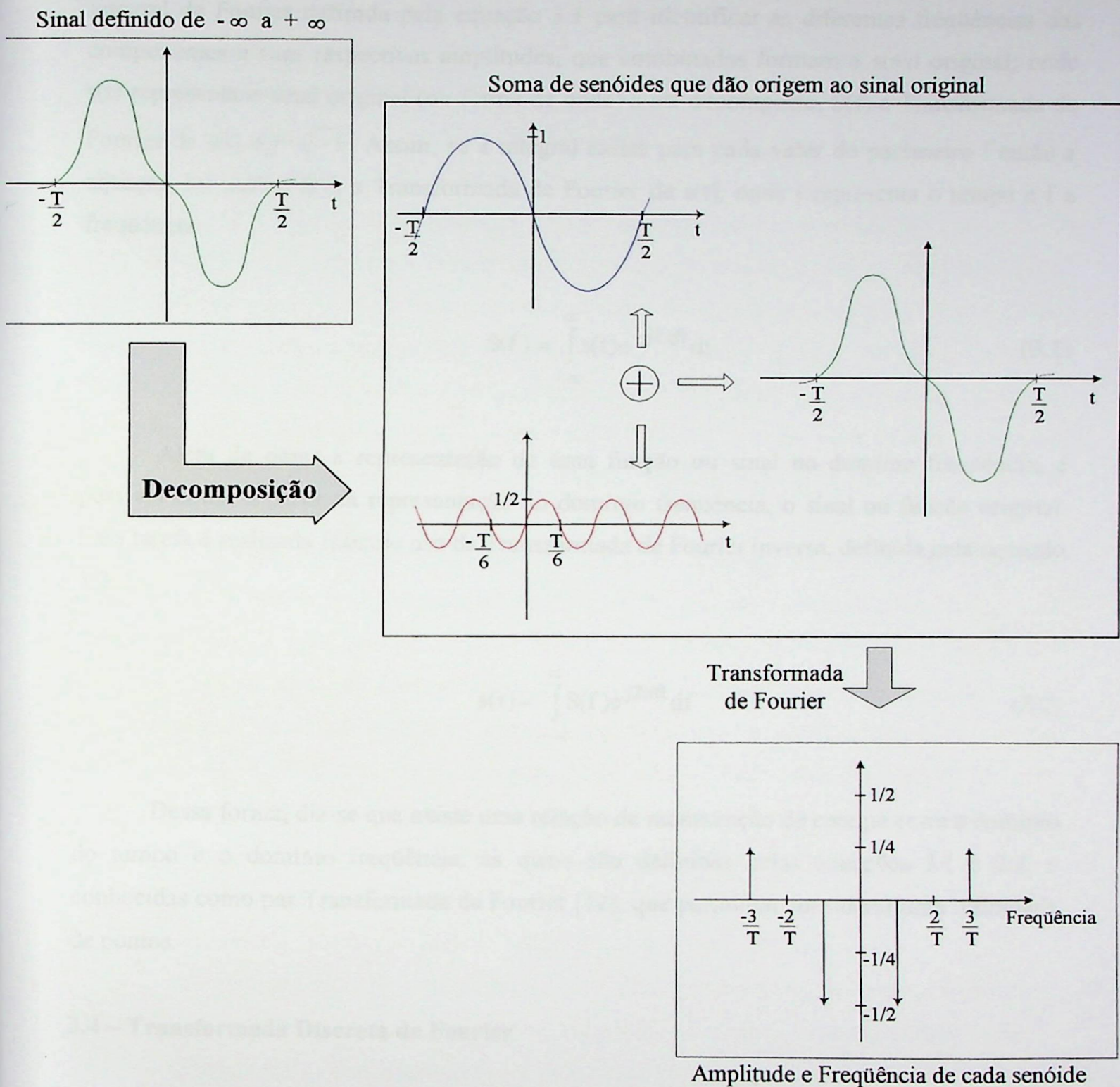


Figura 3.1 - Interpretação da Transformada de Fourier

Matematicamente, a Transformada de Fourier ilustrada pela figura 3.1, se baseia na integral de Fourier definida pela equação 3.1 para identificar as diferentes frequências das componentes e suas respectivas amplitudes, que combinadas formam o sinal original; onde $s(t)$ representa o sinal original (ou forma de onda) a ser decomposto, $S(f)$ a Transformada de Fourier de $s(t)$ e $j = \sqrt{-1}$. Assim, se a integral existe para cada valor do parâmetro f então a equação 3.1 define $S(f)$ a Transformada de Fourier de $s(t)$, onde t representa o tempo e f a frequência.

$$S(f) = \int_{-\infty}^{\infty} s(t)e^{-j2\pi ft} dt \quad (3.1)$$

Além de obter a representação de uma função ou sinal no domínio frequência, é possível obter, a partir da representação no domínio frequência, o sinal ou função original. Essa tarefa é realizada fazendo uso da Transformada de Fourier inversa, definida pela equação 3.2.

$$s(t) = \int_{-\infty}^{\infty} S(f)e^{j2\pi ft} df \quad (3.2)$$

Dessa forma, diz-se que existe uma relação de manutenção de energia entre o domínio do tempo e o domínio frequência, as quais são definidas pelas equações 3.1 e 3.2, e conhecidas como par Transformada de Fourier [12], que permitem considerar uma infinidade de pontos.

3.4 – Transformada Discreta de Fourier

Com o crescimento do uso de computadores, tornou-se cada vez mais importante a análise de Fourier de um sinal já digitalizado. A Transformada Discreta de Fourier (DFT) é a ferramenta fundamental utilizada para esses casos.

A Transformada Discreta de Fourier é usada para sinais amostrados no tempo, de forma que o computador ao calcular a DFT deve tratar apenas um *conjunto finito de pontos*. Porém, esta característica da DFT faz com que passem a existir diversas peculiaridades em

seu resultado, tornando importante interpretar adequadamente a DFT, pois a mesma consiste em uma aproximação do resultado que seria obtido ao se usar a análise de Fourier.

A primeira etapa para o uso da DFT consiste na discretização de $s(t)$, ou sinal no domínio do tempo, onde o sinal é amostrado e truncado em N amostras, a fim de permitir o cálculo da sua transformada com uso do computador digital. Nessa etapa as integrais das equações 3.1 e 3.2 podem ser aproximadas, utilizando métodos clássicos de integração numérica, como a regra dos trapézios, a fim de permitir uma aproximação do resultado da DFT à Transformada de Fourier. Porém, os cálculos necessários exigiriam um enorme esforço computacional caso fossem utilizados um número significativo de pontos para a representação do sinal, o que constitui uma grande desvantagem no uso da DFT, visto que quanto maior o número de pontos usados, melhor a aproximação obtida pelo cálculo numérico. Dessa forma, não seria uma solução conveniente simplesmente diminuir o número de pontos usados na discretização das funções ou sinais, a melhor solução seria a busca de formas eficientes para o cálculo numérico das equações.

Assim, a fim de reduzir o esforço computacional, algumas alternativas são utilizadas para o cálculo da DFT. Uma dessas alternativas é a utilização do valor de N como sendo uma potência de dois, a qual permite alguma simplificação no cálculo da DFT, tornando-a mais eficiente. Cabe ressaltar que a necessidade do uso de alternativas para o cálculo da DFT foi a responsável pelo desenvolvimento de algoritmos para calcular eficientemente a DFT. Esses algoritmos são conhecidos como algoritmos de FFT (“Fast Fourier Transform”), ou algoritmos da Transformada Rápida de Fourier. Porém dois aspectos devem ser considerados:

- Existem vários algoritmos de FFT, alguns concentrando esforços para diminuir o número de multiplicações, outros para diminuir a quantidade de memória necessária, outros que permitem que N não seja potência de dois, etc.
- O resultado numérico obtido pela FFT é igual (a menos que existam erros devido à precisão finita da máquina) ao obtido pela DFT. Assim, a Transformada Rápida de Fourier (FFT) consiste apenas em uma forma rápida para calcular a DFT.

3.5 – Discretização da Transformada de Fourier

Como mencionado anteriormente, a etapa de discretização faz com que o sinal $s(t)$ seja amostrado e truncado em N amostras, onde o valor de N irá representar o número de pontos da DFT. Assim, as integrais das equações 3.1 e 3.2 irão tratar um conjunto finito de pontos.

O primeiro passo para a discretização consiste na amostragem do sinal $s(t)$. O sinal amostrado pode então ser escrito como $s(t)\Delta_0(t)$, que equivale a convolução do sinal no domínio tempo com $\Delta_0(t)$, uma função de amostragem também no domínio do tempo.

A equação 3.3 representa $\Delta_0(t)$, onde T é o intervalo de amostragem. Cabe ressaltar que, dependendo da escolha do valor de T pode ocorrer o efeito “aliasing”, porém para evitar esse efeito basta obedecer ao Teorema de Nyquist [11], [13], [14]. Já a equação 3.4 representa a convolução do sinal com a função de amostragem.

$$\Delta_0(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT) \quad (3.3)$$

$$\begin{aligned} s(t)\Delta_0(t) &= s(t) \sum_{k=-\infty}^{\infty} \delta(t - kT) \\ &= \sum_{k=-\infty}^{\infty} s(kT) \delta(t - kT) \end{aligned} \quad (3.4)$$

Em seguida, o sinal amostrado é truncado a fim de se obter um conjunto finito de pontos para sua representação. Esse passo, apresentado pela equação 3.5, pode ser realizado a partir da multiplicação do sinal por uma função retangular $r(t)$, tal como mostra a equação 3.6.

$$\begin{aligned} s(t)\Delta_0(t)r(t) &= \left[\sum_{k=-\infty}^{\infty} s(kT) \delta(t - kT) \right] r(t) \\ &= \sum_{k=0}^{N-1} s(kT) \delta(t - kT) \end{aligned} \quad (3.5)$$

$$r(t) = \begin{cases} 1, & -\frac{T}{2} < t < T_0 - \frac{T}{2} \\ 0, & \text{outros casos} \end{cases} \quad (3.6)$$

O passo final consiste em amostrar a Transformada de Fourier do sinal obtido a partir da equação 3.5, ou seja, deve-se amostrar $|S(f) \cdot \Delta_0(f) \cdot R(f)|$. No domínio do tempo essa tarefa equivale a convolução do sinal truncado da equação 3.5 e uma função

$$\Delta_1(t) = T_0 \sum_{r=-\infty}^{\infty} \delta(t - rT_0), \text{ que pode ser verificada pelo relacionamento } [s(t)\Delta_0(t)r(t)] * \Delta_1(t).$$

Após o total tratamento matemático e considerando a seqüência com duração limitada (para cálculo em computador) se obtém a equação 3.7, onde N é o número de amostras do sinal $x(k)$ [13]. Assim a equação 3.7 representa a Transformada Discreta de Fourier de um sinal $x(k)$ no domínio do tempo, enquanto a equação 3.8 representa a sua forma inversa. Dessa forma, as equações 3.7 e 3.8, representam o par Transformada Discreta de Fourier.

$$X(n) = \sum_{k=0}^{N-1} x(k) e^{-j2\pi kn/N} \quad (3.7)$$

$$x(k) = \frac{1}{N} \sum_{n=0}^{N-1} X(n) e^{j2\pi kn/N} \quad (3.8)$$

A total representação matemática dos passos utilizados para a obtenção da Transformada Discreta de Fourier, bem como uma representação gráfica e suas propriedades, podem ser verificadas em BRIGHAM [12].

3.6 – Transformada Rápida de Fourier (FFT) [12]

A Transformada Rápida de Fourier (FFT) é um algoritmo eficiente capaz de computar a Transformada Discreta de Fourier (DFT) de forma muito mais rápida. Essa característica é obtida a partir da redução de cálculos repetidos realizados pela DFT [14]. Essas repetições existem devido à natureza periódica do peso, $e^{-j2\pi kn/N}$, da equação 3.7.

Assim, substituindo $e^{-j2\pi kn/N}$ por W^{nk} é possível reescrever a equação 3.7, que irá gerar a equação 3.9.

$$X(n) = \sum_{k=0}^{N-1} x_0(k) W^{nk}, \quad n=0,1,\dots,N-1 \quad (3.9)$$

As repetições mencionadas podem ser observadas a partir de um exemplo simples de fatoração de matrizes, o qual é usado para justificar a importância do algoritmo FFT. A fatoração de matrizes pode ser representada por diagramas de fluxo de sinal, que são usados para construir a lógica de um programa que implementa a FFT, a qual será apresentada a seguir.

A partir da equação 3.9, percebe-se que existe a computação de N equações. Tomando como exemplo N=4 é possível obter as seguintes equações:

$$\begin{aligned} X(0) &= x_0(0)W^0 + x_0(1)W^0 + x_0(2)W^0 + x_0(3)W^0 \\ X(1) &= x_0(0)W^0 + x_0(1)W^1 + x_0(2)W^2 + x_0(3)W^3 \\ X(2) &= x_0(0)W^0 + x_0(1)W^2 + x_0(2)W^4 + x_0(3)W^6 \\ X(3) &= x_0(0)W^0 + x_0(1)W^3 + x_0(2)W^6 + x_0(3)W^9 \end{aligned} \quad (3.10)$$

As equações de 3.10 podem ser escritas pela seguinte forma (3.11), a qual obedece ao seguinte equação: $X(n) = W^{nk} x_0(k)$.

$$\begin{bmatrix} X(0) \\ X(1) \\ X(2) \\ X(3) \end{bmatrix} = \begin{bmatrix} W^0 & W^0 & W^0 & W^0 \\ W^0 & W^1 & W^2 & W^3 \\ W^0 & W^2 & W^4 & W^6 \\ W^0 & W^3 & W^6 & W^9 \end{bmatrix} \begin{bmatrix} x_0(0) \\ x_0(1) \\ x_0(2) \\ x_0(3) \end{bmatrix} \quad (3.11)$$

Nesse caso W, e possivelmente $x_0(k)$ são complexos, então se tem N^2 multiplicações complexas e $(N)(N-1)$ adições complexas. Porém, o algoritmo da FFT permite reduzir o número de multiplicações e adições, a partir da eliminação dos cálculos repetidos.

3.6.1 – Desenvolvimento do Algoritmo da Transformada Rápida de Fourier

A fim de facilitar o entendimento do algoritmo da FFT, considera-se um número de pontos para $x_0(k)$ de acordo com a relação $N=2^Y$, onde Y é um número inteiro. O primeiro passo do algoritmo da FFT está relacionado à equação 3.11. Assim, tendo como número de pontos $N=4$ e $Y=2$ se obtém a equação 3.12.

$$\begin{bmatrix} X(0) \\ X(1) \\ X(2) \\ X(3) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & W^1 & W^2 & W^3 \\ 1 & W^2 & W^0 & W^2 \\ 1 & W^3 & W^2 & W^1 \end{bmatrix} \begin{bmatrix} x_0(0) \\ x_0(1) \\ x_0(2) \\ x_0(3) \end{bmatrix} \quad (3.12)$$

Onde W^{nk} é obtido a partir da relação $W^{nk} = W^{nk \bmod N}$. Assim, se $N=4$, com $n=2$ e $k=3$, então $W^6 = W^2$, relação esta que pode ser verificada pela equação 3.13.

$$\begin{aligned} W^{nk} &= W^6 = \exp\left[\left(\frac{-j2\pi}{4}\right)(6)\right] = \exp[-j3\pi] = \\ &= \exp[-j\pi] = \exp\left[\left(\frac{-j2\pi}{4}\right)(2)\right] = W^2 = W^{nk \bmod N} \end{aligned} \quad (3.13)$$

O segundo passo consiste na fatoração da matriz quadrada da equação 3.12, que resultará na equação 3.14. Cabe lembrar que o resultado da multiplicação da equação 3.14 é o mesmo da equação 3.12, com exceção das linhas 1 e 2 que foram trocadas (as linhas são numeradas por 0, 1, 2 e 3). A troca de linhas da equação 3.14 ocorre devido à reestruturação do vetor coluna $X(n)$ a fim de possibilitar a fatoração, onde essa troca é denotada por 3.15.

$$\begin{bmatrix} X(0) \\ X(2) \\ X(1) \\ X(3) \end{bmatrix} = \begin{bmatrix} 1 & W^0 & 0 & 0 \\ 1 & W^2 & 0 & 0 \\ 0 & 0 & 1 & W^1 \\ 0 & 0 & 1 & W^3 \end{bmatrix} \begin{bmatrix} 1 & 0 & W^0 & 0 \\ 0 & 1 & 0 & W^0 \\ 1 & 0 & W^2 & 0 \\ 0 & 1 & 0 & W^2 \end{bmatrix} \begin{bmatrix} x_0(0) \\ x_0(1) \\ x_0(2) \\ x_0(3) \end{bmatrix} \quad (3.14)$$

$$\overline{X(n)} = \begin{bmatrix} X(0) \\ X(2) \\ X(1) \\ X(3) \end{bmatrix} \quad (3.15)$$

Dessa forma, diz-se que a fatoração é a chave da eficiência do algoritmo da FFT, devido desmembrar os cálculos existentes. Assim, a partir da equação 3.14 nota-se que é possível diminuir o número de cálculos, embora o resultado esteja trocado, o que não representa um problema.

A partir da equação 3.14 pode-se ainda examinar a quantidade de computações existentes na mesma. Tomando como exemplo o vetor coluna $x_1(k)$, mostrado pela equação 3.16, o qual é semelhante ao produto das duas matrizes à direita da equação 3.14, pode-se verificar que o elemento $x_1(0)$ é computado por uma multiplicação e uma adição complexas

(W^0 não é substituído pela unidade para que se obtenha um resultado generalizado), como mostra a equação 3.17.

$$\begin{bmatrix} x_1(0) \\ x_1(1) \\ x_1(2) \\ x_1(3) \end{bmatrix} = \begin{bmatrix} 1 & 0 & W^0 & 0 \\ 0 & 1 & 0 & W^0 \\ 1 & 0 & W^2 & 0 \\ 0 & 1 & 0 & W^2 \end{bmatrix} \begin{bmatrix} x_0(0) \\ x_0(1) \\ x_0(2) \\ x_0(3) \end{bmatrix} \quad (3.16)$$

$$x_1(0) = x_0(0) + W^0 x_0(2) \quad (3.17)$$

O elemento $x_1(1)$ também é obtido por uma multiplicação e uma adição complexas. Já o elemento $x_1(2)$ é obtido por apenas uma adição complexa, após considerar que $W^0 = -W^2$; como mostra a equação 3.18.

$$\begin{aligned} x_1(2) &= x_0(0) + W^2 x_0(2) \\ &= x_0(0) - W^0 x_0(2) \end{aligned} \quad (3.18)$$

Para esse caso, a multiplicação complexa $W^0 x_0(2)$ já foi computada na determinação de $x_1(0)$ na equação 3.17. Seguindo o mesmo raciocínio, $x_1(3)$ é obtido por uma adição complexa e nenhuma multiplicação complexa. Assim, o vetor $x_1(k)$ é obtido por quatro adições complexas e duas multiplicações complexas.

Continuando a computação de 3.14 obtém-se nesse momento a equação 3.19.

$$\begin{bmatrix} X(0) \\ X(2) \\ X(1) \\ X(3) \end{bmatrix} = \begin{bmatrix} x_2(0) \\ x_2(1) \\ x_2(2) \\ x_2(3) \end{bmatrix} = \begin{bmatrix} 1 & W^0 & 0 & 0 \\ 1 & W^2 & 0 & 0 \\ 0 & 0 & 1 & W^1 \\ 0 & 0 & 1 & W^3 \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_1(1) \\ x_1(2) \\ x_1(3) \end{bmatrix} \quad (3.19)$$

Onde o termo $x_2(0)$ é determinado por uma multiplicação e uma adição complexa, como mostra a equação 3.20.

$$x_2(0) = x_1(0) + W^0 x_1(1) \quad (3.20)$$

Já o elemento $x_2(1)$ é obtido por uma adição devido $W^0 = -W^2$. Pela mesma razão, o elemento $x_2(2)$ é determinado por uma multiplicação e uma adição complexas, e $x_2(3)$ apenas por uma adição complexa.

Assim, a computação de $\overline{X(n)}$ obtida através da equação 3.14 requer um total de quatro (4) multiplicações complexas e oito (8) adições complexas. A mesma computação de $\overline{X(n)}$ a partir da equação 3.11 requer dezesseis (16) multiplicações complexas e doze (12) adições complexas. Porém, é possível perceber que o processo de fatoração de matrizes, que introduz zeros nas matrizes fatoradas, reduz o número necessário de multiplicações para a obtenção do resultado.

Dessa forma, devido utilizar o fator $N=2^r$, o algoritmo de FFT é simplesmente um procedimento de fatoração de matrizes $N \times N$, em r matrizes (com $N \times N$), de modo que cada matriz fatorada tem uma propriedade especial de minimização do número de multiplicações e adições complexas. Nesse caso, o processo de fatoração reduz o número de multiplicações pelo fator de 2, reduzindo também o tempo de computação necessário, o que justifica a eficiência do algoritmo da FFT.

Ao considerar os resultados dos exemplos anteriores, conclui-se então que a FFT gera $Nr/2 = 4$ multiplicações complexas e $Nr = 8$ adições complexas, enquanto que pelo método direto seriam necessárias N^2 multiplicações complexas e $N(N-1)$ adições complexas.

Em relação ao tempo de computação conclui-se que o mesmo é proporcional ao número de multiplicações. Assim, o tempo de computação da FFT em relação ao método direto, é dado pela equação 3.21.

$$\frac{N^2}{Nr/2} = \frac{2N}{r} \quad (3.21)$$

Assim, para $N=1024=2^{10}$ existe uma redução computacional de mais de 200 para 1. A figura 3.2 ilustra a relação entre o número de multiplicações necessárias usando o algoritmo da FFT e o número de multiplicações usando o método direto.

Porém, é importante lembrar que o procedimento de fatoração de matrizes traz uma divergência, pois a partir da computação da equação 3.14 obtém-se a equação 3.15 ao invés de $X(n)$, como mostra a equação 3.22.

$$\overline{X(n)} = \begin{bmatrix} X(0) \\ X(2) \\ X(1) \\ X(3) \end{bmatrix} \text{ ao invés de } X(n) = \begin{bmatrix} X(0) \\ X(1) \\ X(2) \\ X(3) \end{bmatrix} \quad (3.22)$$

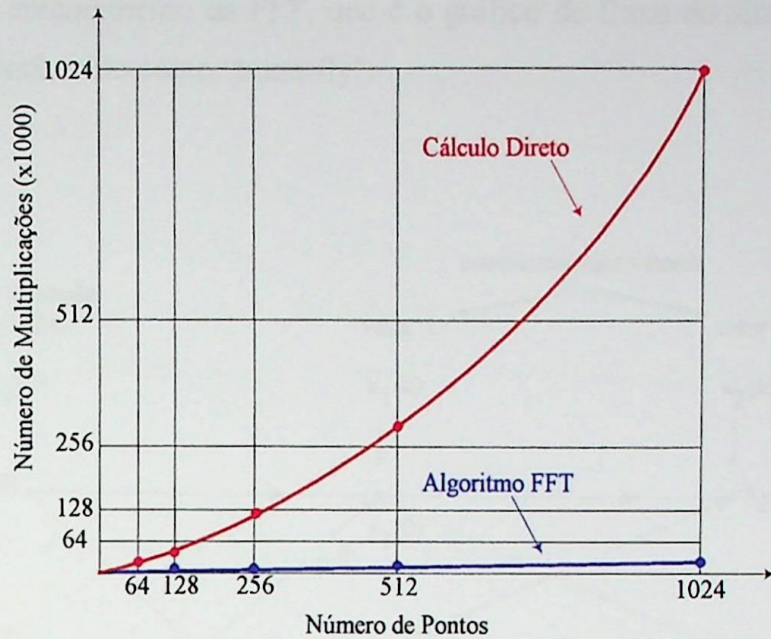


Figura 3.2 - Comparação entre FFT e cálculo direto, quanto ao número de multiplicações.

Assim, é necessário elaborar uma reestruturação, a qual passa a ser inerente ao processo de fatoração de matrizes. Cabe lembrar que essa reestruturação nada mais é do que a ordenação de $\overline{X(n)}$ para obter $X(n)$.

A ordenação é realizada da seguinte maneira: para reescrever $\overline{X(n)}$ será necessário substituir o argumento n por seu equivalente em valor binário, como mostra a equação 3.23.

$$\begin{bmatrix} X(0) \\ X(2) \\ X(1) \\ X(3) \end{bmatrix} \text{ será dado por } \begin{bmatrix} X(00) \\ X(10) \\ X(01) \\ X(11) \end{bmatrix} \quad (3.23)$$

Nesse caso, se os argumentos binários forem invertidos, executando o chamado “*bit reverso*” (por exemplo: 10 irá gerar 01, 01 irá gerar 10 e assim sucessivamente), tem-se a equação 3.24. Assim, o resultado será obtido pela ordenação da FFT.

$$\overline{\overline{X(n)}} = \begin{bmatrix} X(00) \\ X(10) \\ X(01) \\ X(11) \end{bmatrix} \text{ inverterá para } = \begin{bmatrix} X(00) \\ X(01) \\ X(10) \\ X(11) \end{bmatrix} = X(n) \quad (3.24)$$

O exemplo citado anteriormente pode ainda ser visualizado a partir de uma ferramenta muito usada para o entendimento da FFT, que é o gráfico de fluxo do sinal, ilustrado pela figura 3.3, também conhecido como “butterfly”.

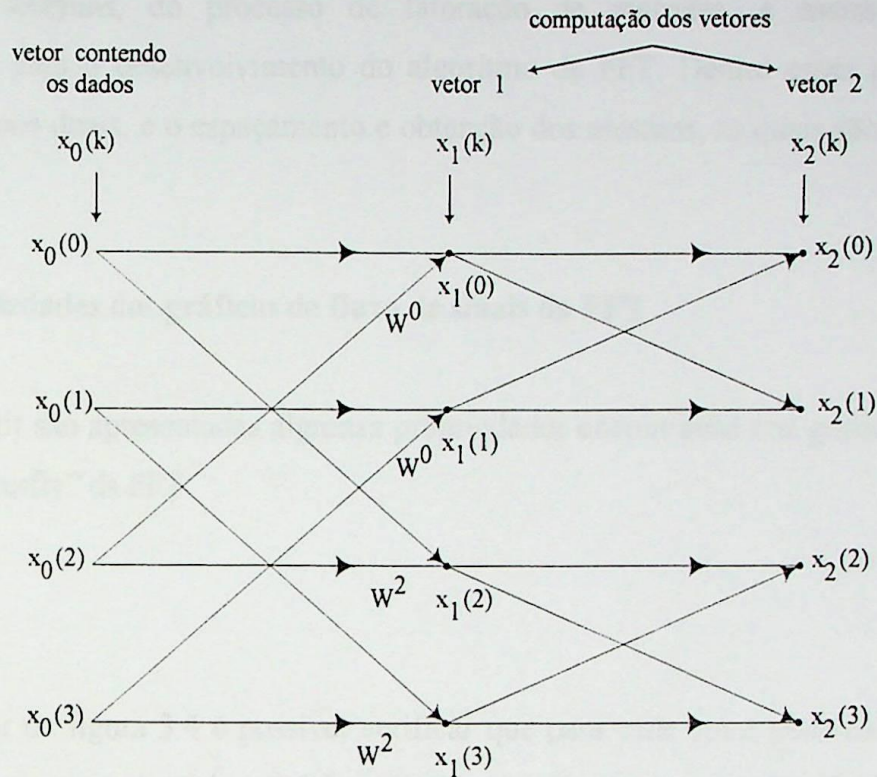


Figura 3.3 - Fluxos de computação dos vetores para N=4

A figura 3.3. apresenta a computação da FFT para N=4, onde $x_0(k)$ representa o vetor de dados, $x_1(k)$ obtido a partir da equação 3.16 e $x_2(k)$ que corresponde ao vetor $\overline{X(n)}$ (obtido pela equação 3.19). Assim, irão existir Y vetores, os quais foram computados ao longo do algoritmo da FFT, onde cada elemento do vetor é chamado de nó.

Em cada nó percebe-se que há a entrada de dois fluxos (linhas sólidas) representando os *caminhos de transmissão* de nós prévios. Cada nó representa um valor de um nó multiplicado por um fator W^p , onde seu resultado será usado para o nó do próximo vetor. Cada fator W^p aparece ao final da seta, a qual representa o fluxo do caminho de transmissão. A ausência desse fator indica que $W^p=1$.

Considerando o nó $x_1(2)$ da figura 3.3, e de acordo com os passos mencionados anteriormente, é possível obter a equação 3.25 que possui o mesmo valor da equação 3.18. Nesse caso, a mesma regra é mantida para todos os nós.

$$x_1(2)= x_0(0)+ W^2 x_0(2)$$

(3.25)

Dessa forma, o gráfico de fluxo do sinal ou “butterfly”, é um ótimo método para a representação das computações necessárias no algoritmo de fatoração de matrizes da FFT (3.14), e sua utilização permite descrever mais facilmente o processo de fatoração de matrizes quando o valor de N é maior.

A figura 3.4 ilustra um gráfico para $N=16$, permitindo verificar as propriedades, consideradas comuns, do processo de fatoração de matrizes, e assim facilitando o entendimento para o desenvolvimento do algoritmo da FFT. Dentre essas propriedades se destacam: os nós duais, e o espaçamento e obtenção dos mesmos, as quais são apresentadas a seguir.

3.6.2 – Propriedades dos gráficos de fluxo de sinais da FFT

A seguir são apresentadas algumas propriedades encontradas nos gráficos de fluxo de sinais ou “butterfly” da FFT.

a) Nós Duais

A partir da figura 3.4 é possível verificar que para cada vetor pode-se encontrar dois nós que possuem como entrada caminhos de transmissão com a mesma origem, tais como os nós $x_1(0)$ e $x_1(8)$ que são computados usando-se os nós $x_0(0)$ e $x_0(8)$, os quais não são utilizados em nenhuma outra computação. Esses nós ($x_1(0)$ e $x_1(8)$) são definidos como *par de nós duais*.

b) Espaçamento e obtenção de Nós Duais

O espaçamento entre os pares de nós duais é medido verticalmente. Ao observar a figura 3.4, e tomando como exemplo o vetor calculado quando $L=1$ (L representa o laço da computação dos vetores), pode-se perceber que o par de nós duais ($x_1(0)$ e $x_1(8)$) é separado por k , obtido a partir da equação 3.26, assim $k=8=N/2^L$.

$$k = \frac{N}{2^L} \quad (3.26)$$

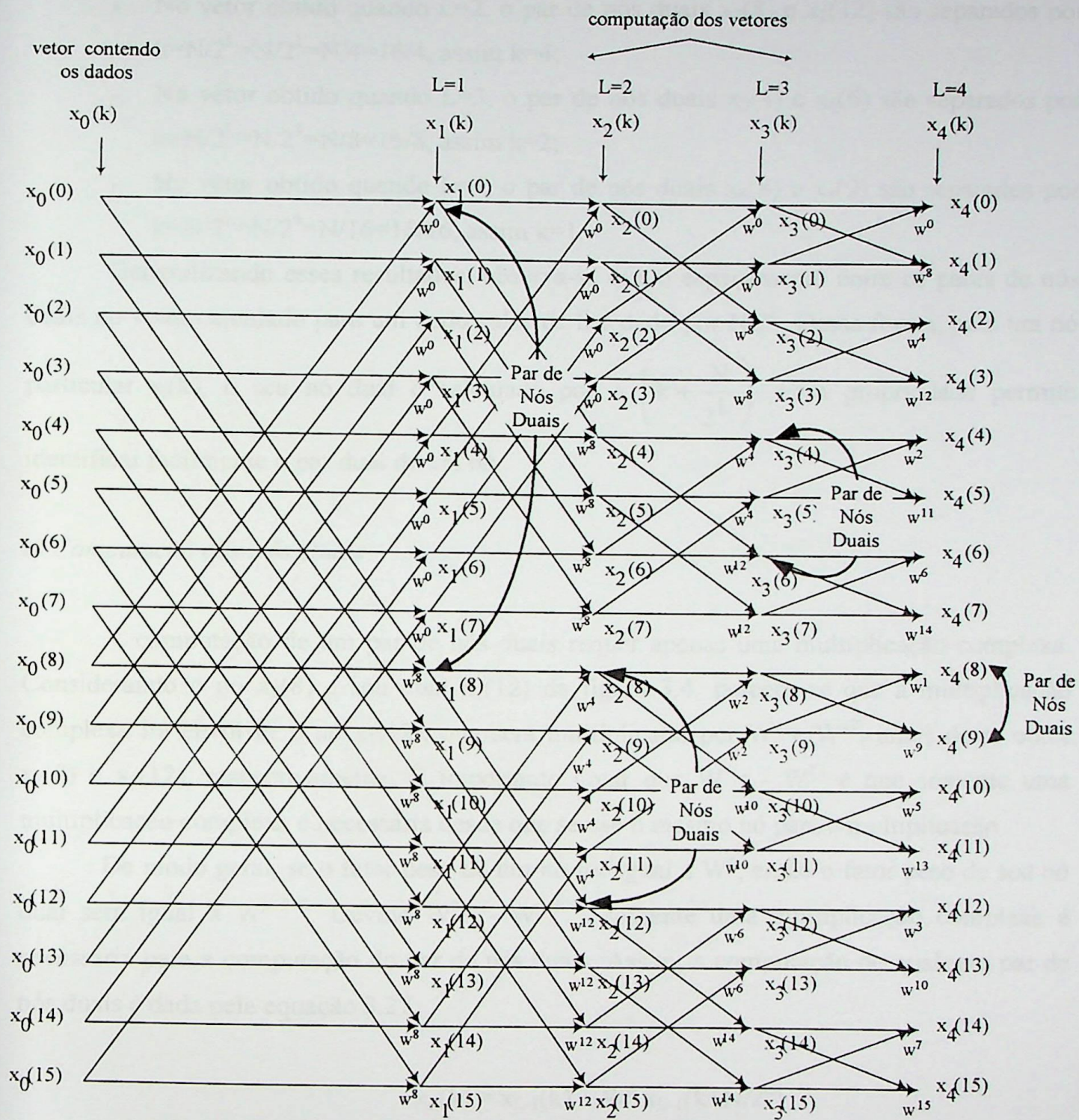


Figura 3.4 - Exemplo de Nós Duais

Dessa forma, os demais nós duais destacados na figura 3.4, são obtidos da seguinte maneira:

- No vetor obtido quando $L=2$, o par de nós duais $x_2(8)$ e $x_2(12)$ são separados por $k=N/2^L=N/2^2=N/4=16/4$, assim $k=4$;
- No vetor obtido quando $L=3$, o par de nós duais $x_3(4)$ e $x_3(6)$ são separados por $k=N/2^L=N/2^3=N/8=16/8$, assim $k=2$;
- No vetor obtido quando $L=4$, o par de nós duais $x_4(8)$ e $x_4(9)$ são separados por $k=N/2^L=N/2^4=N/16=16/16$, assim $k=1$.

Generalizando esses resultados, observa-se que o espaçamento entre os pares de nós duais no vetor calculado para um dado valor de L é dado por $N/2^L$. Dessa forma, para um nó particular $x_l(k)$, o seu nó dual é calculado por $x_l\left(k + \frac{N}{2^L}\right)$. Essa propriedade permite identificar facilmente o par dual de um nó.

c) Computação dos Nós Duais

A computação de um par de nós duais requer apenas uma multiplicação complexa. Considerando o nó $x_2(8)$ e seu dual $x_2(12)$ da figura 3.4, percebe-se que a multiplicação complexa irá envolver o nó $x_1(12)$, que será multiplicado por W^4 e W^{12} , antes de se obter $x_2(8)$ e $x_2(12)$, respectivamente. É importante notar que $W^4 = -W^{12}$ e que somente uma multiplicação complexa é necessária desde que se use o mesmo nó para a multiplicação.

De modo geral, se o fator peso de um nó for igual à W^p , então o fator peso de seu nó dual será igual à $W^{p+N/2}$. Devido $W^p = -W^{p+N/2}$, somente uma multiplicação complexa é necessária para a computação do par de nós duais. Assim, a computação de qualquer par de nós duais é dada pela equação 3.27.

$$\begin{aligned} x_L(k) &= x_{L-1}(k) + W^p x_{L-1}(k+N/2^L) \\ x_L(k+N/2^L) &= x_{L-1}(k) - W^p x_{L-1}(k+N/2^L) \end{aligned} \quad (3.27)$$

É importante lembrar que a existência de pares de nós duais elimina cálculos de vários outros nós. Assim, se o espaçamento entre os nós e seus respectivos duais for de $N/2^L$, então se deve pular (“skip”) os próximos $N/2^L$ nós, como mostra a figura 3.5, pois seus cálculos já foram efetuados.

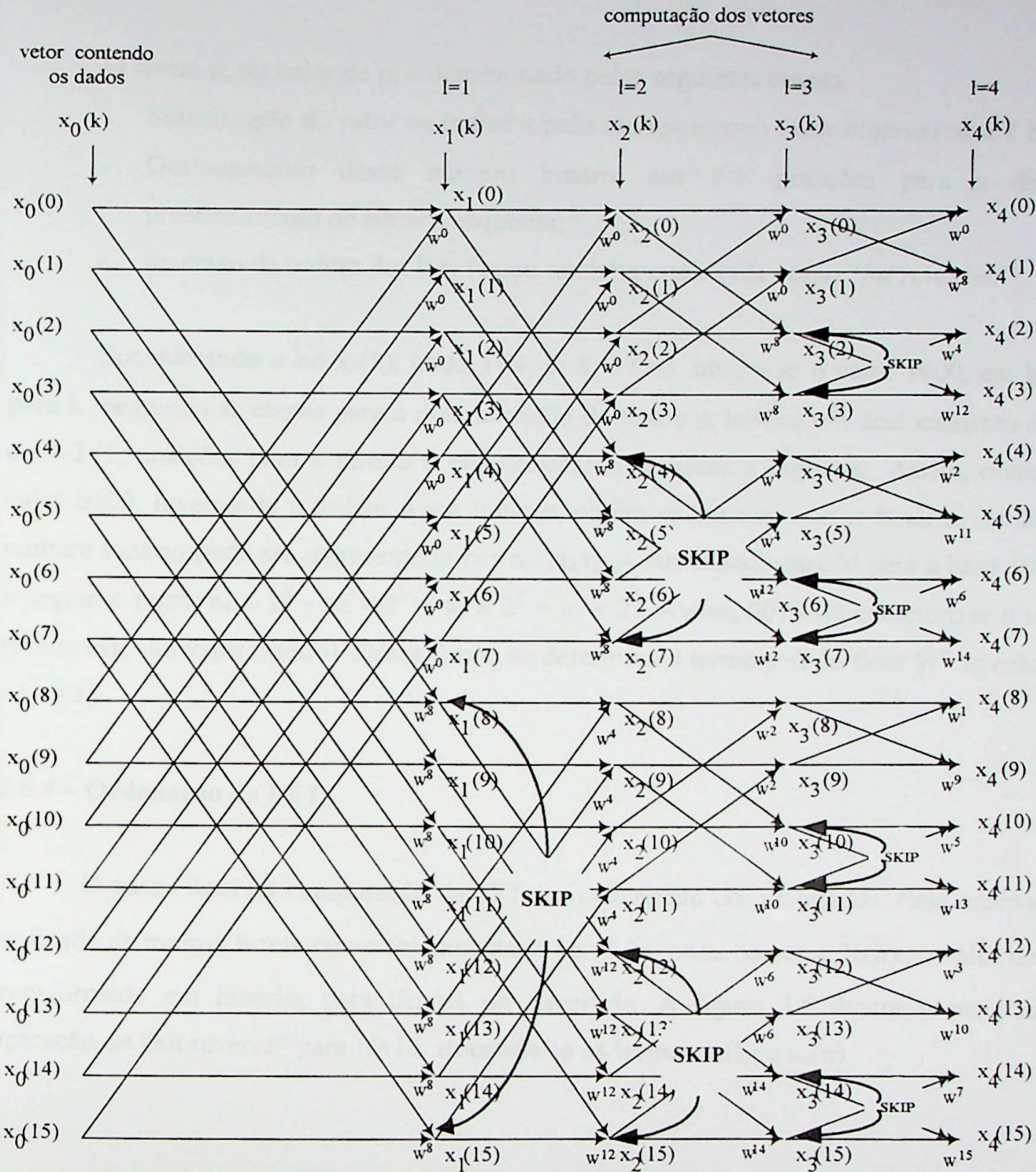


Figura 3.5 - Exemplo de quais nós se deve pular devido computação desnecessária.

3.6.3 – Determinação do Fator W^p

O termo p , ou valor de p , é determinado pelas seguintes etapas:

- Substituição do valor ou índice k pelo seu respectivo valor binário (com Y bits);
- Deslocamento desse número binário em $Y-L$ posições para a direita e preenchimento de zeros à esquerda;
- Inversão da ordem dos bits (etapa também conhecida como “*bit reversed*”).

Considerando o nó $x_3(8)$, onde $Y=4$, $k=8$, e $L=3$, obtém-se o valor 1000, em binário, para k . Seguindo as etapas para a determinação do termo p , haverá um deslocamento de $Y-L$, ou $4-3=1$, posições para a direita e preenchimento de zeros à esquerda. Assim, obtém-se o valor 0100. Invertendo a ordem desse binário, obtém-se um novo valor binário, ou 0010. O número binário pode ser representado por $a_4a_3a_2a_1$. A sua transformação para a base 10 segue a seguinte regra: $a_4 \times 2^3 + a_3 \times 2^2 + a_2 \times 2^1 + a_1 \times 2^0$. Assim, 0010 irá caracterizar o valor 2 (sendo este um valor inteiro). Dessa forma se determina o termo $p=2$ do fator W^p ao calcular o nó $x_3(8)$.

3.6.4 – Ordenação da FFT

O passo final na computação da FFT é a *ordenação* dos resultados. Essa ordenação é realizada da mesma forma como foi executado por 3.24, onde no vetor $\overline{X(n)}$, o valor de n foi transformado em binário, para depois ser invertido. A figura 3.6 ilustra o resultado da operação de “bit reverse” para $N=16$, trocando-se os termos $x_4(k)$ e $x_4(n)$.

Figura 3.6 – Exemplo de bit reverse para $N=16$

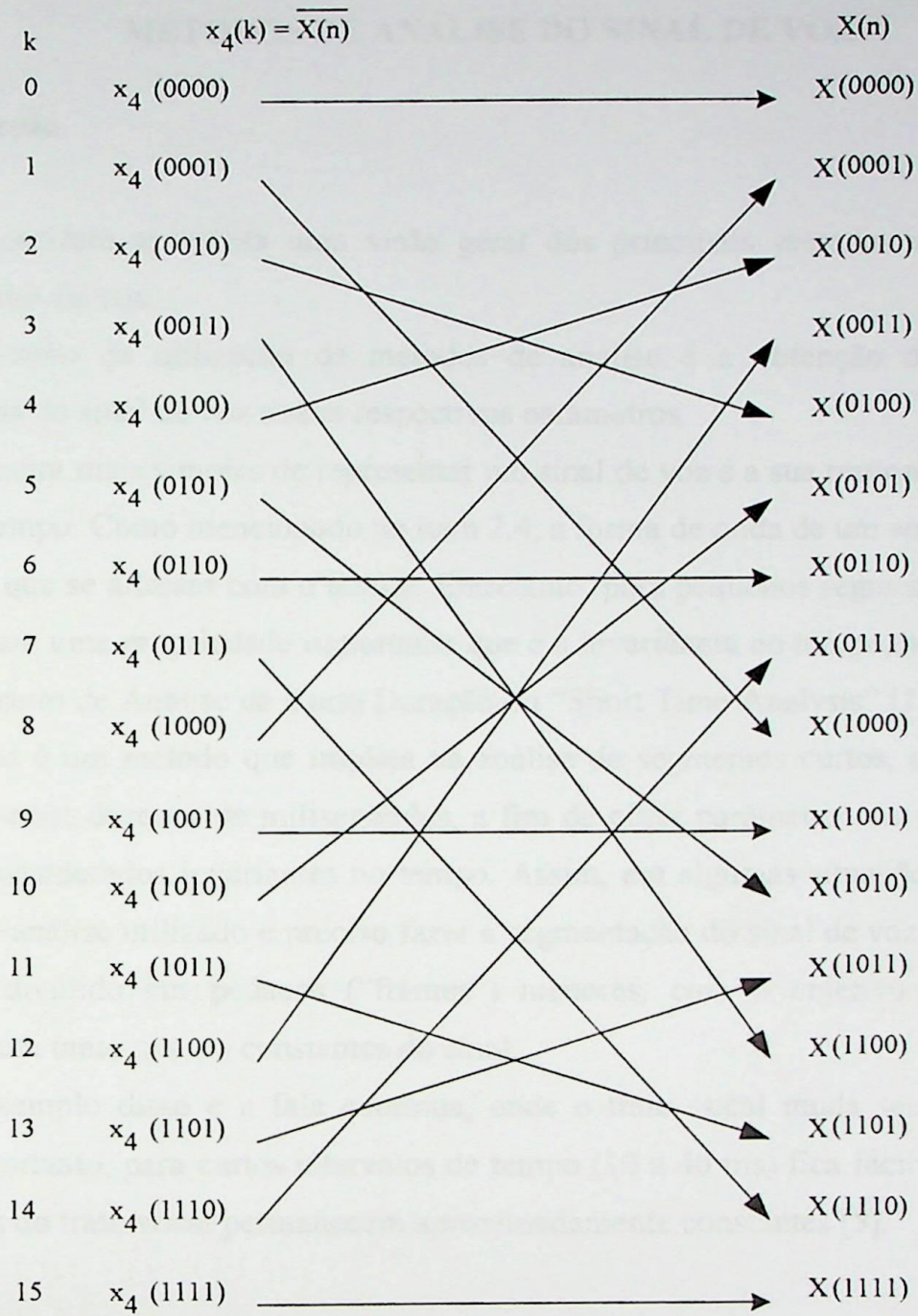


Figura 3.6 - Exemplo de “Bit-Reversing” para $N=16$

MÉTODOS DE ANÁLISE DO SINAL DE VOZ

4.1 – Introdução

Esse capítulo apresenta uma visão geral dos principais métodos utilizados para a análise de sinais de voz.

O objetivo da utilização de métodos de análise é a obtenção de características representativas do sinal de voz e seus respectivos parâmetros.

A maneira mais simples de representar um sinal de voz é a sua representação direta no domínio do tempo. Como mencionado no item 2.4, a forma de onda de um sinal de voz possui propriedades que se alteram com o tempo. Entretanto, para pequenos segmentos, os sinais de voz, apresentam uma propriedade importante que é a invariância no tempo, nos quais pode-se aplicar o conceito de Análise de Curta Duração ou “Short Time Analysis” [15]. A Análise de Curta Duração é um método que implica na análise de segmentos curtos, ou segmentos na ordem de algumas dezenas de milissegundos, a fim de obter parâmetros do sinal de voz que possam ser considerados invariantes no tempo. Assim, em algumas situações e dependendo do método de análise utilizado é preciso fazer a segmentação do sinal de voz, de forma que o mesmo seja dividido em pedaços (“frames”) menores, com o objetivo de observar as propriedades um tanto quanto constantes do sinal.

Um exemplo disso é a fala contínua, onde o trato vocal muda seu formato muito lentamente. Portanto, para curtos intervalos de tempo (10 a 40 ms) fica fácil observar que as características do trato vocal permanecem aproximadamente constantes [5].

4.2 – Determinação da Periodicidade

A periodicidade de um segmento de voz é o parâmetro mais utilizado para se identificar entre um segmento vozeado (“voiced”) e um segmento não vozeado (“unvoiced”). É importante ressaltar que o segmento vozeado possui periodicidade, enquanto que o segmento não vozeado não apresenta esse parâmetro. Assim, se o segmento for vozeado é possível identificar o período fundamental [16].

A partir das figuras 4.1 e 4.2 é possível observar a diferença entre um segmento vozeado e um segmento não vozeado, onde apenas o vozeado apresenta periodicidade. A periodicidade encontrada na figura 4.1 é de 4,5 ms, o que caracteriza uma frequência de

222Hz. Assim, a frequência do “pitch”, muitas vezes designada como F_0 , pode ser determinada, no domínio do tempo, pela medida direta do período da forma de onda. O inverso da frequência do “pitch” ($\tau = 1/F_0$), identifica o período do “pitch”.

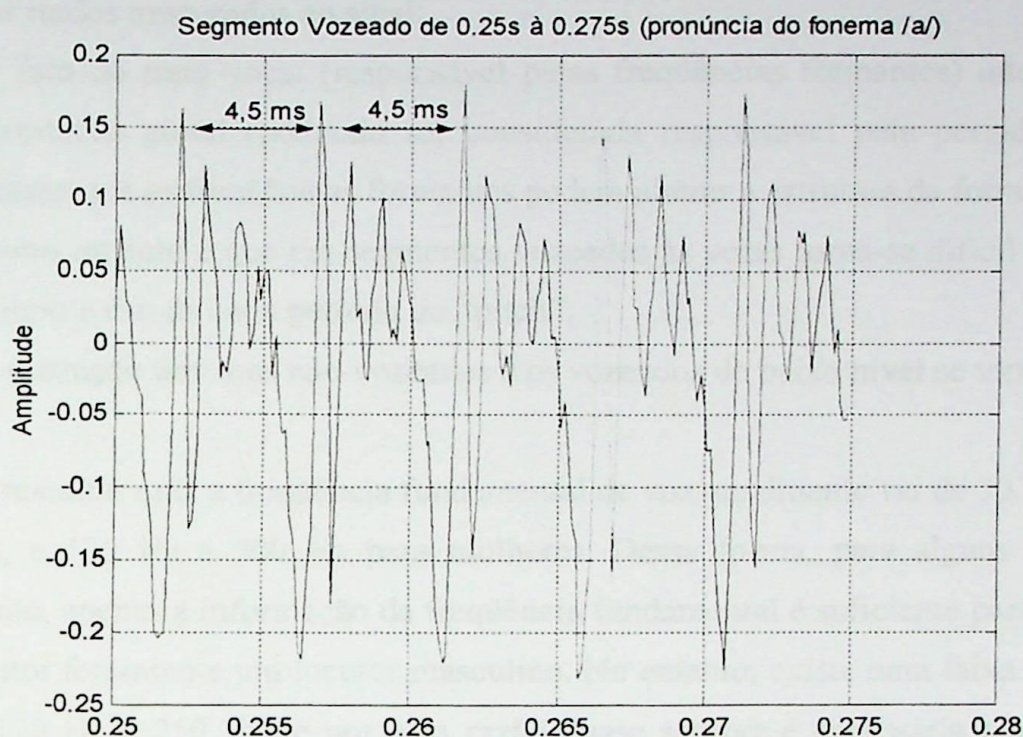


Figura 4.1 - Segmento Vozeado obtido a partir da pronúncia do fonema /a/

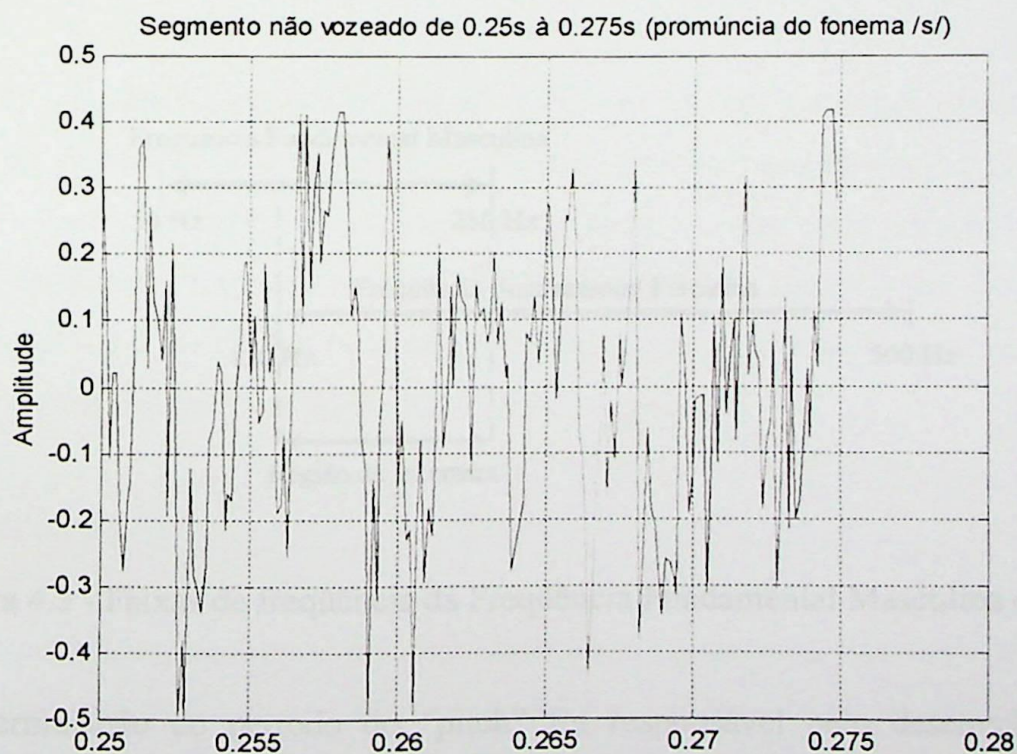


Figura 4.2 - Segmento Não Vozeado obtido a partir da pronúncia do fonema /s/



Entretanto, existem alguns aspectos que podem dificultar a medição da frequência fundamental ou “pitch”, sendo esta a frequência de vibrações das cordas vocais, como mencionado no capítulo 2. Dentre esses aspectos estão:

- os ruídos misturados ao sinal;
- o fato do trato vocal (responsável pelas frequências formantes) interagir com a excitação glotal (podendo ser considerada responsável pela periodicidade), de forma que as frequências formantes podem alterar a estrutura da forma de onda;
- outro aspecto é que em segmentos vozeados às vezes torna-se difícil identificar o início e fim de cada período do “pitch”;
- a distinção dos sons não-vozeados e os vozeados de baixo nível se torna difícil.

Cabe ressaltar que, a frequência fundamental de voz geralmente vai de 50 Hz a 250 Hz para homens, e 120 Hz a 500 Hz para mulheres. Dessa forma, para alguns sistemas de reconhecimento, apenas a informação da frequência fundamental é suficiente para diferenciar entre um locutor feminino e um locutor masculino. No entanto, existe uma faixa de incerteza (que vai de 120 Hz a 250 Hz) e por essa razão quase sempre é necessária a utilização de outras características do sinal de voz, além da frequência fundamental.

A figura 4.3 ilustra as faixas de frequência da Frequência Fundamental Masculina e Feminina [5].

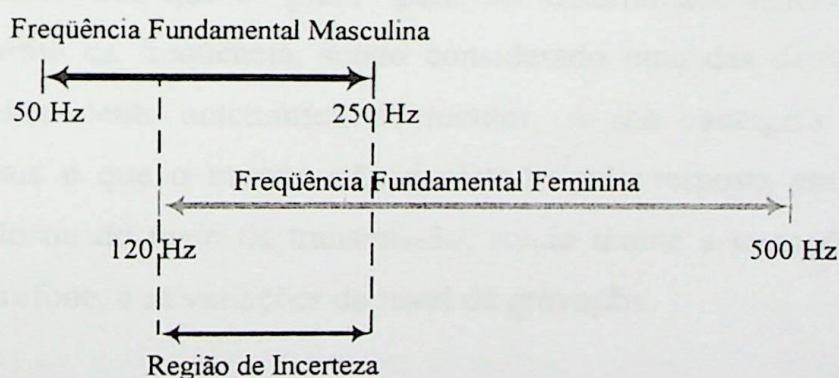


Figura 4.3 - Faixas de frequência da Frequência Fundamental Masculina e Feminina

A determinação do período do “pitch” foi responsável pelo desenvolvimento de inúmeros algoritmos, ou métodos, para sua estimação e detecção; isso se deve ao fato da dificuldade de obtenção de um valor confiável e preciso. A falta de periodicidade dos

segmentos não vozeados e a periodicidade dos segmentos vozeados (observada de forma clara pelas figuras 4.1 e 4.2), pode ser identificada através desses algoritmos.

No entanto, os algoritmos encontrados na literatura não possuem uma performance tão robusta como a percepção humana, e a escolha do algoritmo irá depender de aspectos como:

- complexidade e facilidade de implementação;
- precisão na estimativa do período do “pitch”;
- velocidade da operação;
- e precisão na identificação entre segmentos vozeados e não vozeados.

Dentre os algoritmos mais utilizados para a identificação do “pitch” estão [16]:

- algoritmo “Gold-Rabiner Pitch Tracker”: que usa a informação no domínio tempo para extrair o período do “pitch”;
- algoritmo “Autocorrelation Pitch Tracker”, ou simplesmente autocorrelação: que confia na propriedade da autocorrelação para extrair a informação do “pitch”;
- algoritmo “Homomorphic Pitch Tracker”: que usa o cepstrum como meio para separar o trato vocal da excitação;
- além dos algoritmos “Average Magnitude Difference Function (AMDF) Pitch Tracker” (ou Método da Função Diferença da Magnitude Média), o “Sift algorithm for Pitch Extraction”, e “Post-Processing Techniques for Pitch Detection”.

Dessa forma observa-se que o “pitch” pode ser determinado tanto no domínio do tempo, como no domínio da frequência, sendo considerado uma das características mais importantes no reconhecimento automático de locutor. A sua vantagem em relação às características espectrais é que o mesmo não é afetado pela resposta em frequência do dispositivo de gravação ou do meio de transmissão, sendo imune a variações da distância entre o locutor e o microfone, e as variações do nível de gravação.

4.3 – Análise de Picos

A análise de picos está relacionada à determinação da periodicidade e é efetuada a partir da representação do sinal de voz no domínio tempo. Esse método implica no uso de pequenos segmentos, os quais permite a verificação dos picos de maior relevância do sinal. Além dos picos existem os vales que são os pontos de menor concentração de energia, onde a perturbação da envoltória do espectro é que permite diferenciar os sons vozeados.

É importante lembrar que os picos estão associados à frequência fundamental e às frequências formantes do trato [15]. As frequências formantes, ou frequências de ressonância do tubo acústico, apresentam maior concentração de energia, de forma que a sua presença deforma a envoltória do espectro da fonte de excitação. As frequências formantes são determinadas pela forma do trato vocal para os sons orais, e dos tratos vocal e nasal para os sons nasalados.

Assim, a Análise de Picos é um método usado para se entender a propriedade de periodicidade do sinal, e identificar a frequência fundamental (“pitch period”).

4.4 – Análise da Energia

A análise da energia é outro método muito utilizado e permite distinguir os trechos de voz dos trechos de silêncio, em uma determinada locução. Isso é realizado a partir da variação de energia.

Na utilização desse método o sinal deve ser segmentado, considerando que não há grandes variações na voz [9].

Os intervalos ou segmentos são então ponderados por uma função de janelamento. A ponderação efetuada pelo janelamento destaca o trecho central de um segmento. Dentre as janelas encontradas na literatura estão: janela Retangular, Hamming, Hanning, Kaiser, Blackman e Bartlett (triangular), onde a mais utilizada é a janela de Hamming, a qual é dada pela equação 4.1, onde N é a largura da janela em número de amostras. Geralmente se escolhe N para segmentos de 20 a 40 ms, assim se considerarmos um sinal amostrado a 8 KHz teremos um valor de $N=240$ amostras para segmentos de 30ms [16]. A figura 4.4 mostra um exemplo simples de janelamento, na qual existe uma certa porcentagem de superposição entre as janelas. Entretanto, o janelamento pode ser realizado com ou sem superposição, onde a superposição aumentará a correlação entre janelas sucessivas evitando variações bruscas entre as “características” extraídas de janelas adjacentes. Porém a superposição exige maior tempo de processamento.

$$w(n) = \begin{cases} 0.54 - 0.46 \cos \frac{2\pi n}{N}, & 0 \leq n \leq N-1 \\ 0, & \text{outros casos} \end{cases} \quad (4.1)$$

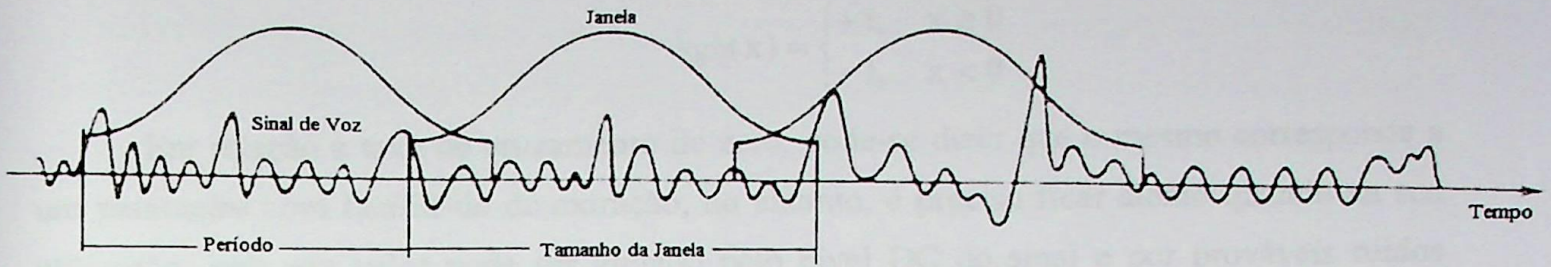


Figura 4.4 - Exemplo de janelamento de um sinal de voz

A análise de energia também permite separar segmentos vozeados, no entanto é muito sensível a grandes variações no nível do sinal por ser calculada pelo somatório do quadrado do valor de cada amostra, como mostra a equação 4.2, onde $x(n)$ representa as amostras do sinal de voz e $w(n)$ a janela escolhida.

$$E(n) = \sum_{m=-\infty}^{\infty} x^2(n)w(n - m) \quad (4.2)$$

Cabe lembrar que a utilização desse método requer cuidados que estão relacionados com a escolha da janela e o seu comprimento (ou tamanho do intervalo a ser analisado). Por exemplo, se o comprimento da janela for menor que o intervalo da frequência fundamental, é possível que se perda características do sinal.

4.5 – Taxa de Cruzamento de Zeros

A taxa de Cruzamento de Zero (“zero crossing rate”) é um método de análise no domínio do tempo e corresponde a mudança da amplitude do sinal de voz de positivo para negativo, ou vice-versa; sendo considerado um parâmetro de fácil análise e medida. Assim, diz-se que ocorre Cruzamento de Zero quando amostras consecutivas apresentam sinais algébricos diferentes [9].

A equação 4.3 define a taxa de Cruzamento de Zero (“ Z_s ”) para um sinal de voz (“ S ”), com uma janela de largura (“ LJ ”).

$$Z_s = \frac{1}{LJ} \sum_{n=2}^{LJ} \frac{|\text{sgn}(S(n)) - \text{sgn}(S(n-1))|}{2} \quad (4.3)$$

onde

$$\text{sgn}(x) = \begin{cases} +1, & x \geq 0 \\ -1, & x < 0 \end{cases}$$

Em relação à taxa de cruzamento de zero, pode-se dizer que o mesmo corresponde a um parâmetro com facilidade de extração, no entanto, é preciso ficar atento quando na sua utilização, pois seu valor pode ser afetado pelo nível DC do sinal e por prováveis ruídos existentes no processo de digitalização.

Nesse método um segmento não vozeado do sinal de voz é identificado quando se tem uma taxa de cruzamento de zero alta. Já o segmento vozeado é identificado quando se tem uma taxa de cruzamento de zero baixa.

4.6 – Autocorrelação

A autocorrelação pode ser usada para determinar a estrutura de qualquer forma de onda, incluindo os sinais de voz. A periodicidade da função de autocorrelação indica periodicidade no sinal analisado, e o contorno da função de autocorrelação pode indicar a aleatoriedade do sinal.

O algoritmo de autocorrelação, que usa a função de correlação, pode ser utilizado tanto para a extração do “pitch” como em conjunto com outros métodos para a identificação da periodicidade. No entanto, esse método pode gerar um resultado insatisfatório, pois o sinal pode apresentar ressonâncias do trato vocal, de forma que o “pitch” possa ser confundido com as oscilações de tais ressonâncias.

A função de autocorrelação de um segmento do sinal de voz janelado $s(n)$, começando com $n=0$ e possuindo um tamanho de N amostras, é definida pela equação 4.4:

$$R(k) = \sum_{n=0}^{N-1} s(n)s(n+k) \quad k=0,1,2,\dots \quad (4.4)$$

Onde a variável k é o índice do tempo, a qual também é chamada de “lag”. Em sons vozeados, com muita periodicidade, existirão picos em “lags” correspondentes ao período do “pitch”, enquanto que nos sons não-vozeados não irão aparecer esses picos; assim a partir dessas observações, o uso da função de autocorrelação permite a classificação do sinal, ou segmento.

Um aspecto importante da função de autocorrelação é que a mesma contém todas as informações associadas ao trato vocal, inclusive as informações necessárias para identificar o “pitch”. No entanto, qualquer informação adicional pode levar a decisões confusas e erradas. A figura 4.5 ilustra um exemplo onde a informação do “pitch” é obtida de forma clara, sendo esta indicada pelas setas. Já a figura 4.6 ilustra uma situação onde a obtenção da informação do “pitch” está confusa, isso devido às ressonâncias do trato vocal.

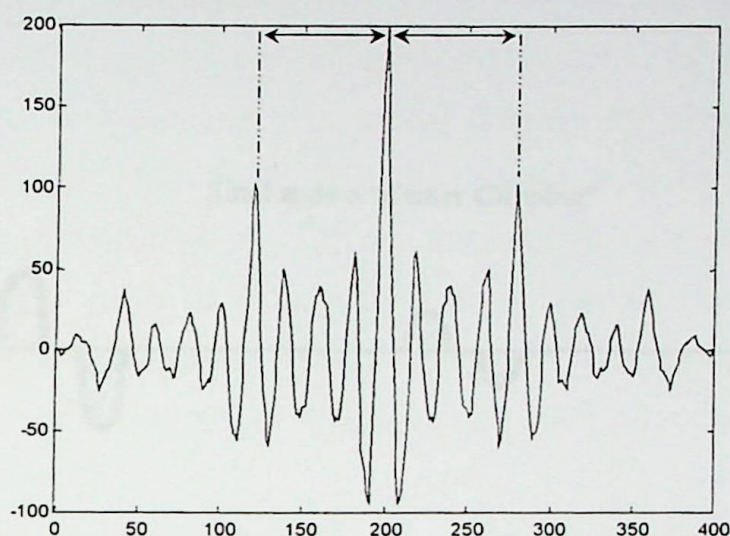


Figura 4.5 - Autocorrelação de um Segmento Vozeado

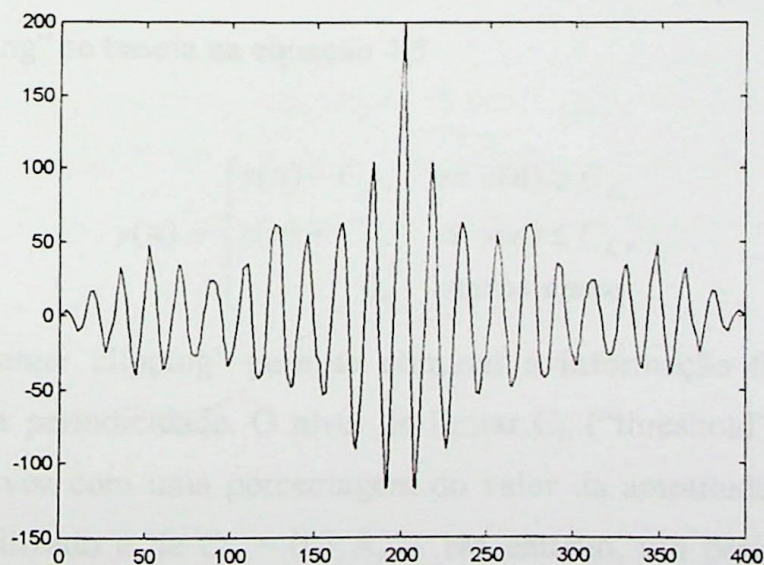


Figura 4.6 - Autocorrelação de um Segmento Vozeado, contendo ressonâncias

Para evitar esse problema e obter apenas a representação da informação do “pitch”, é preciso que o sinal passe por um corte (“center clipping”). O “center clipping” omite ou elimina o sinal entre certos níveis e aqueles que não são eliminados começam do nível zero. A figura 4.7 ilustra em exemplo de aplicação do “Center Clipping”.

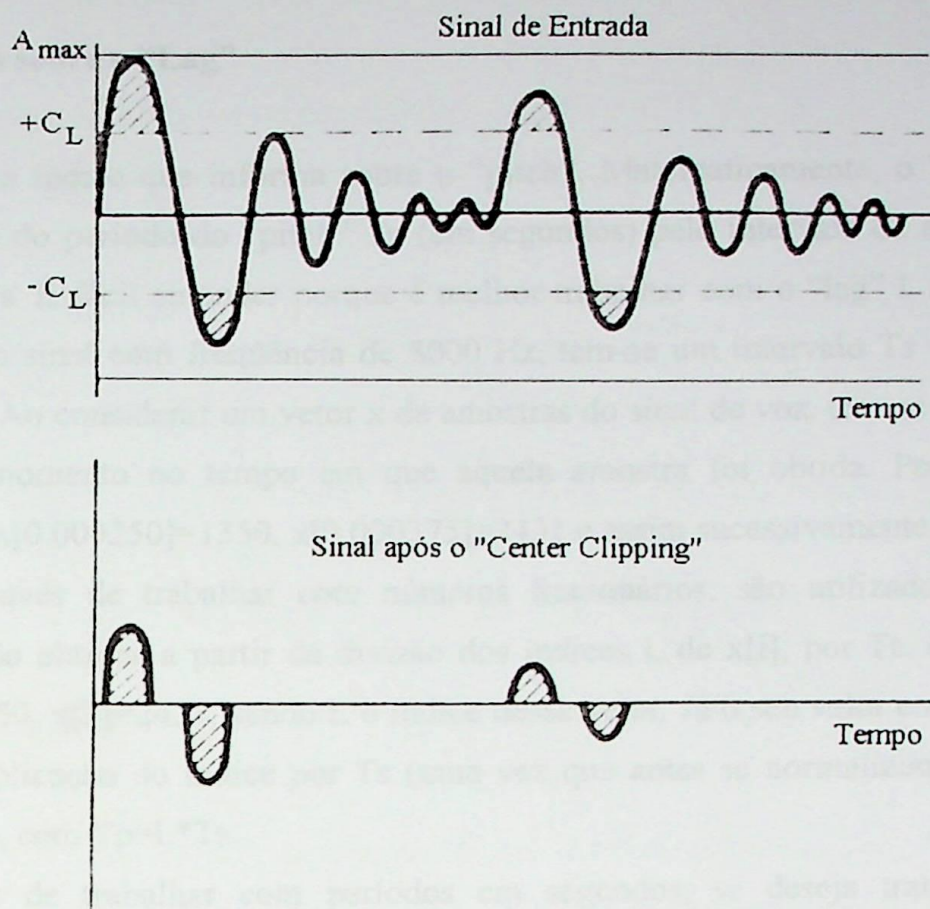


Figura 4.7 - Exemplo de aplicação do “center clipping”

O “center clipping” se baseia na equação 4.5.

$$y(n) = \begin{cases} s(n) - C_L, & \text{se } s(n) \geq C_L, \\ s(n) + C_L, & \text{se } s(n) \leq -C_L, \\ 0, & \text{outros casos.} \end{cases} \quad (4.5)$$

O efeito do “center clipping” permite eliminar a informação formante mantendo apenas a informação da periodicidade. O nível do limiar C_L (“threshold”) é calculado para todos os segmentos de voz com uma porcentagem do valor da amplitude máxima do sinal, $A_{\max}$. O valor típico utilizado é de $C_L = 0.3 A_{\max}$. No entanto, um perigo surge quando a amplitude do sinal varia consideravelmente durante o segmento de voz, onde um limiar alto poderia causar algum tipo de perda.

4.7 – Considerações sobre o “Lag”

O “lag” é um índice que informa sobre o “pitch”. Matematicamente, o “lag” L é o resultado da divisão do período do “pitch” T_p (em segundos) pelo intervalo de amostragem T_s , ou seja, $L=T_p/T_s$. É fácil entender porque é melhor trabalhar com o “lag” L ao invés de T_p . Ao amostrar um sinal com frequência de 8000 Hz, tem-se um intervalo $T_s = 0.000125$ entre cada amostra. Ao considerar um vetor x de amostras do sinal de voz, tem-se o índice de x que indicará o momento no tempo em que aquela amostra foi obtida. Por exemplo: $x[0.000125]=1254$, $x[0.000250]=1350$, $x[0.000375]=2431$ e assim sucessivamente.

Assim, ao invés de trabalhar com números fracionários, são utilizados números inteiros, os quais são obtidos a partir da divisão dos índices i , de $x[i]$, por T_s , obtendo-se: $x[1]=1254$, $x[2]=1350$, $x[3]=2431$, sendo L o índice desse vetor. Já o seu valor em segundos, é obtido pela multiplicação do índice por T_s (uma vez que antes se normalizou os índices dividindo-se por T_s), com $T_p=L*T_s$.

Se ao invés de trabalhar com períodos em segundos, se deseja trabalhar com frequências em Hz, a frequência de amostragem seria $F_s=1/T_s$ e a frequência do “pitch” seria $F_p=1/T_p$. Deste modo, a fórmula para o “lag” seria: $L=F_s/F_p$, tendo Hertz (Hz) como unidade para F_s e F_p .

Considerando como exemplo uma frequência de amostragem de 16KHz, para o “lag” de 24 obtém-se um valor de 666.67 Hz, e para o “lag” igual a 160 obtém-se o valor de 100 Hz. Portanto para o intervalo de 24 a 160 para o “lag”, tem-se um intervalo de 100 a 666 Hz para F_p .

4.8 – Análise por Predição Linear

A Análise por Predição Linear, muitas vezes referenciada como coeficientes LPC, (“Linear Prediction Coefficients”) foi muito utilizada no início dos anos 80 em sistemas de reconhecimento de voz; coincidindo com o início da utilização dos Modelos de Markov para o modelamento acústico. Assim, esse método pode ser considerado como um dos mais conhecidos para a codificação dos sinais de voz e se baseia no princípio de predição linear.

A idéia básica na codificação através da predição linear – LPC (“Linear Predictive Coding”) – é a de que uma dada amostra do sinal de voz pode ser aproximada através de uma

combinação linear de amostras anteriores. Através da minimização do quadrado da diferença entre a amostra atual e as amostras obtidas pela predição linear, pode-se determinar os coeficientes da combinação linear ou os coeficientes do preditor.

A figura 4.8 ilustra um modelo de análise através de predição linear.

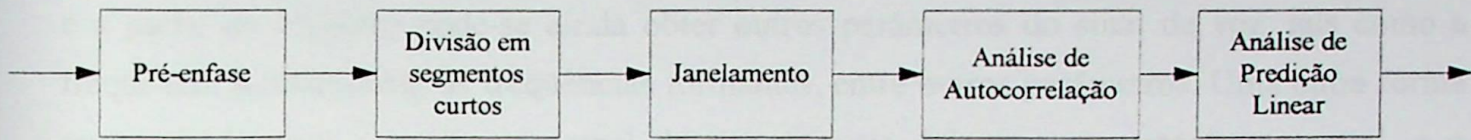


Figura 4.8 - Modelo de análise através de predição linear

A pré-ênfase é utilizada para tratar, ou enfatizar, altas frequências [16].

4.9 – Análise Espectral

A análise espectral é o método mais importante na área de processamento de sinais e é utilizada para representar um sinal no domínio da frequência, tendo como principais métodos o banco de filtros (“Filter Bank”) e a Transformada de Fourier de Curta Duração (“Short-Time Fourier Transform” – STFT), como também a Transformada Rápida de Fourier.

O banco de filtros, para a representação do sinal de voz no domínio frequência, usa uma série de filtros passa-banda (cerca de 15 a 20 filtros) igualmente espaçados ao longo da faixa de frequências (banda espectral) do sinal de voz (entre zero e a frequência de Nyquist), onde a saída dos filtros pode ser considerada como uma integração de vários componentes de Fourier do sinal de entrada. O banco de filtros também pode ser usado como ferramenta na obtenção de outros parâmetros, como é o caso dos coeficientes Mel Cepstrais, que usa um banco de filtros distribuídos de acordo com a escala Mel [17].

Entretanto, o algoritmo Transformada Rápida de Fourier se mostra superior ao modelo de banco de filtros, pois pode calcular eficientemente a Transformada Discreta de Fourier e de forma muito mais rápida (como mencionado no capítulo 3).

Portanto, ao optar pela análise espectral, geralmente se usa como principal ferramenta a Transformada de Fourier, onde a análise do espectro em frequência para a análise de curta duração, usa a Transformada de Fourier de segmentos curtos, e consecutivos do sinal de voz. Assim, esse método provê uma boa representação para o espectro em frequência do sinal

analisado, podendo ainda usar janelas, com o mesmo comprimento, de forma que pequenas partes do sinal são analisadas separadamente.

O STFT permite ainda a representação tridimensional (magnitude, tempo, frequência) do sinal de voz.

Assim, um sinal de voz pode ser representado diretamente pela representação espectral e a partir do espectro pode-se ainda obter outros parâmetros do sinal de voz, tais como a frequência fundamental, as frequências formantes, entre outros parâmetros. Uma outra forma muito usada para a análise do sinal de voz a partir de sua representação espectral, é o espectrograma, que nada mais é do que uma outra maneira de visualizar o sinal de voz, consistindo na representação da variação das frequências de um trecho do sinal de voz ao longo do tempo.

4.10 – Processamento Homomórfico do Sinal de Voz

O processamento homomórfico [18] é outro método de análise do sinal de voz e é muito utilizado quando se deseja separar dois sinais pela operação de convolução (ou também a multiplicação). Esse método usa como ferramenta fundamental o cepstrum ou cepstro.

Como mencionado anteriormente, o modelo ilustrado pela figura 2.2 é usado para a produção do sinal de voz. A análise desse sinal de voz pelo processamento homomórfico, em segmentos curtos o suficiente no tempo, é efetuada a partir da separação das componentes de uma convolução da função excitação (pulsos quase periódicos ou ruído branco) com a função de resposta impulsiva do trato vocal, ou seja, seria uma deconvolução proveniente do método apresentado pela figura 2.2. Dessa forma, a deconvolução homomórfica seria aplicada para separar estes sinais.

O cepstrum é talvez o processamento homomórfico [14] mais popular devido ser muito útil para a deconvolução.

4.11 – Análise Cepstral

A análise cepstral de sinais de voz foi motivada originalmente [19] pelo desejo de separar (dos sons vozeados) o conteúdo da excitação glotal (rica em conteúdo harmônico) do conteúdo do trato vocal, permitindo assim a estimação das frequências fundamentais e as frequências ressonantes (formantes). Esse método também é muito usado para extrair sinais de ecos de sons [20].

O cepstrum $C_s(n)$ de uma seqüência $s(n)$ é definido pela equação 4.6.

$$C_s(n) = \mathfrak{F}^{-1} \{ \log | \mathfrak{F} \{ s(n) \} | \} \quad (4.6)$$

Onde $\mathfrak{F}\{\bullet\}$ é a Transformada de Fourier e $\mathfrak{F}^{-1}\{\bullet\}$ a sua inversa. Assim, o cepstrum é a inversa do logaritmo do espectro de uma função ou sinal, de forma que os coeficientes extraídos desse processamento são chamados de coeficientes cepstrais.

Como o cepstrum se baseia na Transformada de Fourier, adotou-se uma terminologia para evitar possíveis confusões entre os componentes de frequência usuais de uma função tempo, e a onda de “frequência” do logaritmo do espectro. Essa terminologia inclui a paráfrase “quefrequency” para descrever a “frequência” da onda espectral, onde as “qüefrências” possuem como unidade ciclos por hertz, ou simplesmente, segundos. Dessa forma, os coeficientes cepstrais se encontram no domínio da “qüefrência”, por analogia a frequência. Lembrando que o próprio nome Cepstrum deriva do nome Espectrum (Espectro) [21].

Tomando como base o modelo de produção de voz da figura 2.2, é possível perceber que o sinal de voz é produzido pela convolução (denotada por $*$) da excitação glotal (trem de impulsos ou ruído branco, dependendo se o som é vozeado ou não) com a resposta ao impulso discreto do trato vocal. A equação 4.7 mostra a representação dessa convolução, onde $e[n]$ representa a excitação (Gerador de Ruído) e $\theta[n]$ a resposta ao impulso do trato vocal (Gerador de Impulsos). Onde o uso da análise cepstral permite obter subsídios para que se possa analisar separadamente a excitação e o trato vocal.

$$s[n] = e[n] * \theta[n] \quad (4.7)$$

No domínio da frequência, a convolução representada pela equação 4.7, se torna uma simples multiplicação, como mostra a equação 4.8, o que dificulta a separação da excitação do trato vocal (resposta do filtro), pois não é simples separar (filtrar) sinais num processo de multiplicação. No entanto, é mais simples projetar filtros para separar sinais considerando apenas parcelas de uma soma.

$$S(\omega) = E(\omega) \cdot \theta(\omega) \quad (4.8)$$

Sabendo que $\log AB = \log A + \log B$, é possível transformar a relação de multiplicação da equação 4.8 em uma relação de adição a partir do uso do logaritmo, esse passo é mostrado pela equação 4.9.

$$\text{Log } S(\omega) = \log E(\omega) + \log \theta(\omega) \quad (4.9)$$

Dessa forma, obtém-se os valores de $S(\omega)$, $E(\omega)$, e $\theta(\omega)$ como sendo valores complexos. Para evitar o uso de logaritmos complexos basta utilizar o módulo (magnitude) antes do uso do logaritmo, como mostra a equação 4.10.

$$\text{Log } |S(\omega)| = \log |E(\omega)| + \log |\theta(\omega)| \quad (4.10)$$

Da equação 4.10 observa-se que o logaritmo da excitação e o logaritmo do trato vocal estão combinados linearmente, assim é possível utilizar técnicas lineares para analisá-los separadamente. Obtém-se então a equação 4.11, que permite retornar ao domínio do tempo a partir da $\mathfrak{T}^{-1}\{\bullet\}$.

$$C_s(n) = C_e(n) + C_\theta(n) \quad (4.11)$$

Dessa forma, o processamento do trato vocal ($C_\theta(n)$), que apresenta “variações mais lentas”, encontra-se nas “qüefrências” mais baixas, enquanto que excitação ($C_e(n)$) por apresentar “variações mais rápidas” se encontra nas “qüefrências” mais altas, assim ambos podem ser facilmente separados a partir da análise cepstral.

A figura 4.9 ilustra o cepstrum do segmento vozeado obtido a partir da pronúncia do fonema /a/ e ilustrado pela figura 4.1.

No domínio da “qüefrência”, os componentes $C_e(n)$ e $C_\theta(n)$ estarão bem separados, permitindo separar um componente do outro, nesse caso é possível, por exemplo, separar e remover o componente $C_e(n)$ de $C_s(n)$, a fim de analisar os coeficientes $C_\theta(n)$, ou resposta ao impulso discreto do trato vocal. Esse processo é chamado de “liftering” (em analogia a “filtering”) e nesse caso usa-se um “low-time lifter”.

Para eliminar um dos dois componentes (trato vocal e excitação) do sinal de voz, é necessário identificar o limiar entre os dois componentes. Esse limiar é identificado por n_0 e deve ser menor que o período da frequência fundamental do sinal a ser analisado [9]. Assim, n_0 é calculado por $n_0 = F_s / \text{picth}$.

Considerando o segmento de 25 ms ilustrado na figura 4.1, cujo “pitch” encontrado foi de 222Hz, tem-se um valor $n_0=36$ já que o sinal de voz em questão possui como frequência de amostragem $F_s=8000\text{Hz}$. Assim, n_0 deve ser menor que 36.

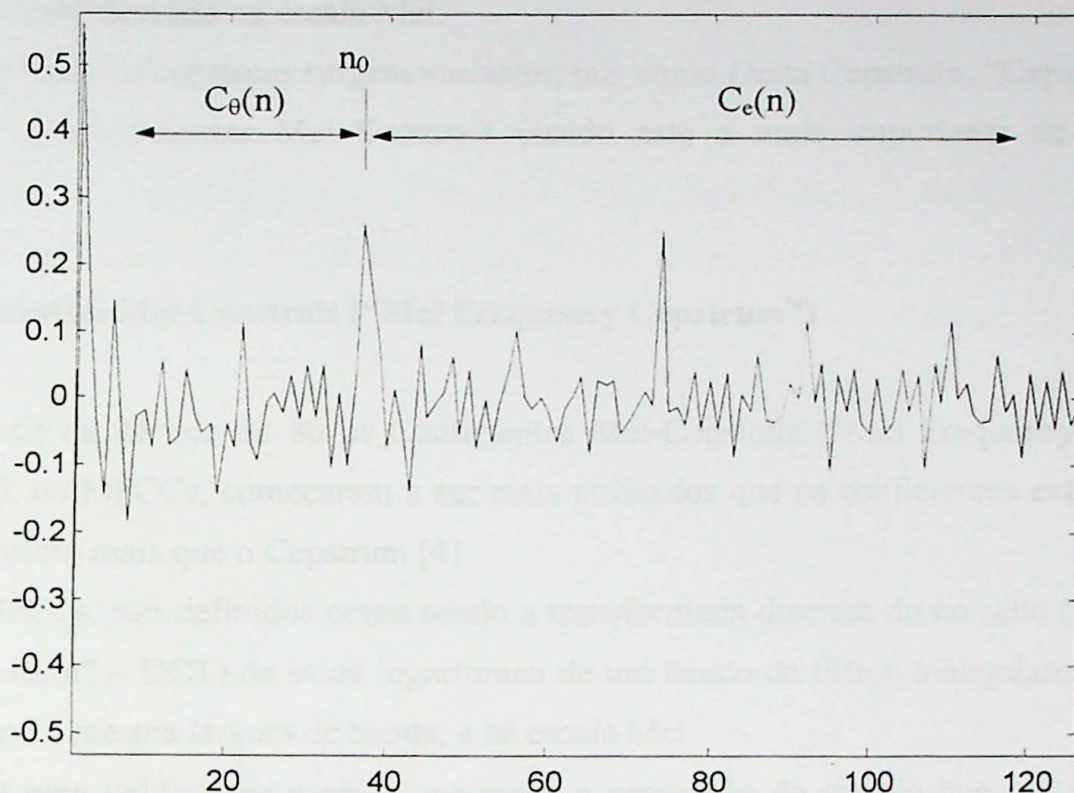


Figura 4.9 - Cepstrum obtido a partir de um segmento de 25 ms ilustrado na figura 4.1.

Como mencionado anteriormente, os sinais de voz mudam no tempo, precisando realizar a segmentação dos mesmos, dessa forma para cada segmento é calculado o cepstrum correspondente, a fim de obter um cepstrum médio para o sinal em questão. Esses componentes de cepstrum são denominados de cepstrum de curto tempo ou duração (“short-time cepstrum”).

O cepstrum, para segmentos vozeados, possui picos com maior relevância e que correspondem ao período do “pitch”, assim, para análise da voz, o cepstrum atende de forma excelente como um novo método para a determinação do “pitch”. É preciso lembrar ainda que existem dois tipos básicos de cepstrum: o cepstrum complexo e o cepstrum real. Entretanto, apenas o cepstrum real é usado no processamento de voz pois o mesmo descarta a informação da fase. Já o cepstrum complexo, considerado mais preciso, carrega a informação da fase, no entanto, a sua aplicação é mais onerosa em termos computacionais, assim o cepstrum real tem sido mais utilizado.

Os coeficientes cepstrais podem ainda ser calculados a partir dos coeficientes de predição linear (LPC). Assim, existem duas formas de calcular o cepstrum real, são elas: o cepstrum LPC e o cepstrum FFT. No cepstrum LPC, os coeficientes cepstrais são obtidos dos coeficientes LPC, e no cepstrum FFT de uma FFT; onde no reconhecimento de voz o cepstrum FFT está baseado na escala Mel.

Em relação ao cepstrum surgem variantes, tais como Delta Cepstrum, “Cepstral Mean Subtraction” e Coeficientes Mel Cepstrais (sendo este a mais importante de todas as variantes).

4.12 – Coeficientes Mel Cepstrais (“Mel Frequency Cepstrum”)

A partir da década de 80 os Coeficientes Mel-Cepstrais (“Mel Frequency Cepstral Coefficients”), ou MFCCs, começaram a ser mais utilizados que os coeficientes extraídos do LPC, e até mesmo mais que o Cepstrum [4].

Os MFCCs, são definidos como sendo a transformada discreta do co-seno (“Discrete Cosine Transform” – DCT) da saída logarítmica de um banco de filtros triangulares onde os filtros possuem a mesma largura de banda, e na escala Mel.

Mel é uma unidade de medida que mede a percepção do ouvido humano para uma determinada frequência [22], e foi obtida de forma experimental. Através de experimentos [17], [23] citados no item 4.12.1, notou-se que até 1000Hz, a escala Mel é aproximadamente linear e acima deste valor é logarítmica, isso em relação à escala de frequência em Hz. Assim, 1000 Mels equivale a 1000 Hz. A equação 4.12 define a relação da escala Mel para a frequência em Hz.

$$\text{Mel} = 2595 \log_{10} \left(1 + \frac{f}{700} \right)$$

(4.12)

Assim, os MFCCs são calculados como mostra a figura 4.10 [17]:

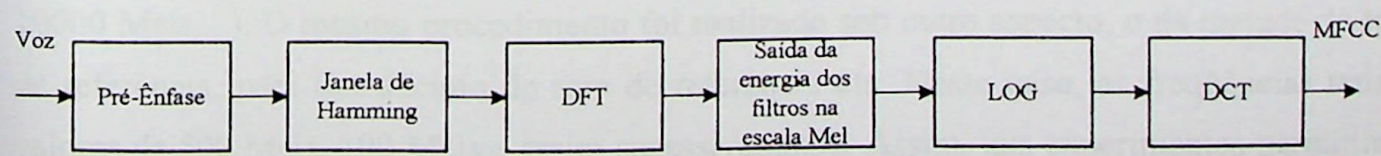


Figura 4.10 - Cálculo dos Coeficientes Mel Cepstrais (MFCCs)

Existem duas maneiras simples de analisar os MFCCs [17]: (a) como o processamento de um banco de filtros adaptados para as particularidades da voz; e (b) como uma modificação do cepstrum convencional, ou a técnica de deconvolução baseada no processamento homomórfico. Esses dois aspectos podem ser considerados complementares, facilitando assim o entendimento dos MFCCs.

Assim, em relação aos MFCCs pode-se dizer que os mesmos estão sendo mais utilizados por apresentarem melhores resultados, além de possuírem uma propriedade da DCT, que ordena os coeficientes obtidos por ordem de variância, colocando aqueles que apresentam maior variância em primeiro lugar. Dessa forma, os coeficientes obtidos via DCT provavelmente carregam mais informações do que os correspondentes obtidos através da FFT inversa [15].

4.12.1 – Considerações sobre a Escala Mel

O termo Mel está relacionado ao campo da psicoacústica, ciência que estuda a percepção auditiva humana, onde estudos comprovam que a percepção humana das frequências de tons puros ou de sinais de voz, não segue uma escala linear [17], [23]. Baseados nessa informação foram definidos frequências individuais de tons puros, onde para cada tom (altura do som) com frequência f medida em Hz definiu-se um tom individual medido em uma escala denominada Mel, que é a unidade de medida da frequência percebida de um tom. Assim, foram realizados experimentos para a identificação das medidas referentes às frequências.

Para a frequência de 1kHz, com 40 dB acima do limiar de audição do ouvido, foi definido o valor de 1000 Mels [24]. Para outras frequências foram definidos valores através de experimentos. Nesses experimentos, os ouvintes ajustavam a frequência física de um tom até que a frequência percebida fosse igual a duas vezes a frequência de referência. Posteriormente, o mesmo experimento foi realizado, só que para 10 vezes a frequência de referência e assim sucessivamente (as frequências obtidas agora teriam valores de 2000 Mels, 10000 Mels,...). O mesmo procedimento foi realizado sob outro aspecto, o da metade do tom de referência, para um décimo do tom de referência etc. Nesse caso, as frequências teriam valores de 500 Mels, 100 Mels e assim sucessivamente. Assim, tais experimentos permitiram obter um mapeamento entre a escala de frequência usual (em Hz) e a escala de frequências percebida (em Mel), observando que esse mapeamento é aproximadamente linear (abaixo de 1000Hz) e logarítmica (acima de 1000Hz), como mencionado anteriormente.

Baseados no mapeamento das escalas, foram iniciadas pesquisas relacionadas aos benefícios da utilização da escala de frequências percebida pelo ouvido humano no reconhecimento da voz, onde inicialmente foi realizada uma super amostragem no eixo das frequências, a partir de uma FFT com muitos pontos.

Após essa tarefa, foi realizada uma experiência onde foram selecionadas as frequências correspondentes à escala Mel, zerando todas as outras frequências, o que iniciou o estudo de bandas críticas juntamente com a escala em questão. Nesse caso, a banda crítica também está relacionada à percepção humana.

Porém, é preciso lembrar que a percepção humana não permite identificar algumas frequências de sons complexos, pois as mesmas não podem ser identificadas individualmente, dentro de certas bandas. Esse aspecto ocorre quando uma frequência particular (por exemplo f_0) é influenciada pela energia da banda crítica das frequências em torno de f_0 ; já quando uma componente cai fora da banda crítica, é possível identificá-la facilmente.

Cabe ressaltar que, o valor da banda varia de 10 a 20% da frequência central do som, começando em torno de 100 Hz para frequências abaixo de 1000 Hz e aumentando em escala logarítmica, acima de 1000 Hz [23], [25].

Dessa forma, considerando aspectos relacionados à escala Mel e a banda crítica, foram realizadas algumas modificações na representação das medidas de distâncias espectrais, onde inicialmente foi realizada uma ponderação da escala de frequências para a escala Mel, adicionando a idéia de banda crítica na definição de distorção espectral. Assim, ao invés de usar apenas o logaritmo da magnitude das frequências, passou-se a usar o logaritmo da energia total das bandas críticas em torno das frequências Mel. Para a aproximação desse cálculo foi utilizado um banco de filtros espaçados uniformemente em uma escala não linear (escala Mel). A figura 4.11 ilustra um exemplo de banco de filtros.

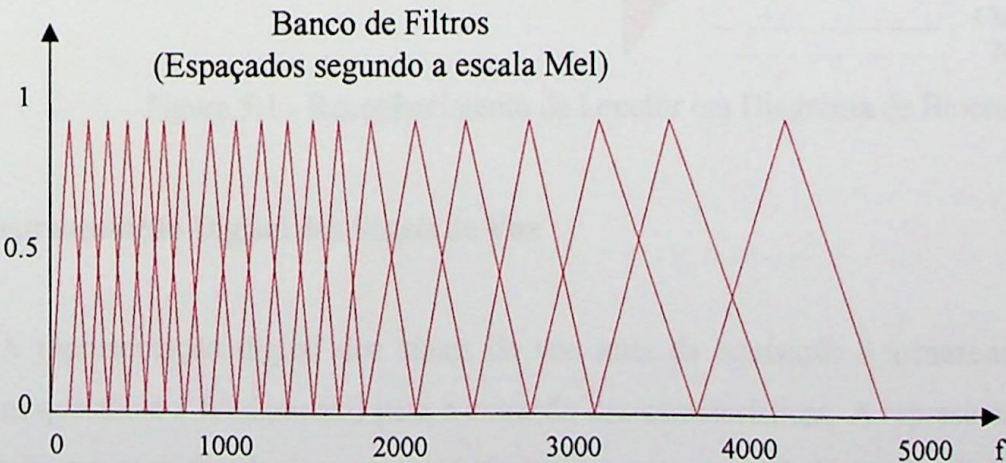


Figura 4.11 - Exemplo de magnitude do espectro do banco de filtros

AQUISIÇÃO DAS CARACTERÍSTICAS DO SINAL DE VOZ

5.1 – Introdução

Um Sistema de Reconhecimento de Locutor realizado por computador é dividido em duas etapas importantes. A primeira etapa está relacionada à aquisição e extração das características (“features”) relevantes dos sinais de voz (pronunciados por diferentes locutores); e a segunda etapa que está relacionada à classificação de locuções, e consiste nas fases de treinamento (comparação de padrões e geração de regras) e teste. Onde os algoritmos utilizados por essas etapas foram implementados no programa MATLAB (software da The MathWoks, Inc”).

O objetivo desse capítulo é apresentar a etapa de aquisição e extração das características dos sinais de voz, anterior ao Sistema de Reconhecimento de Locutor. O sistema é ilustrado pelo diagrama de blocos da figura 5.1.

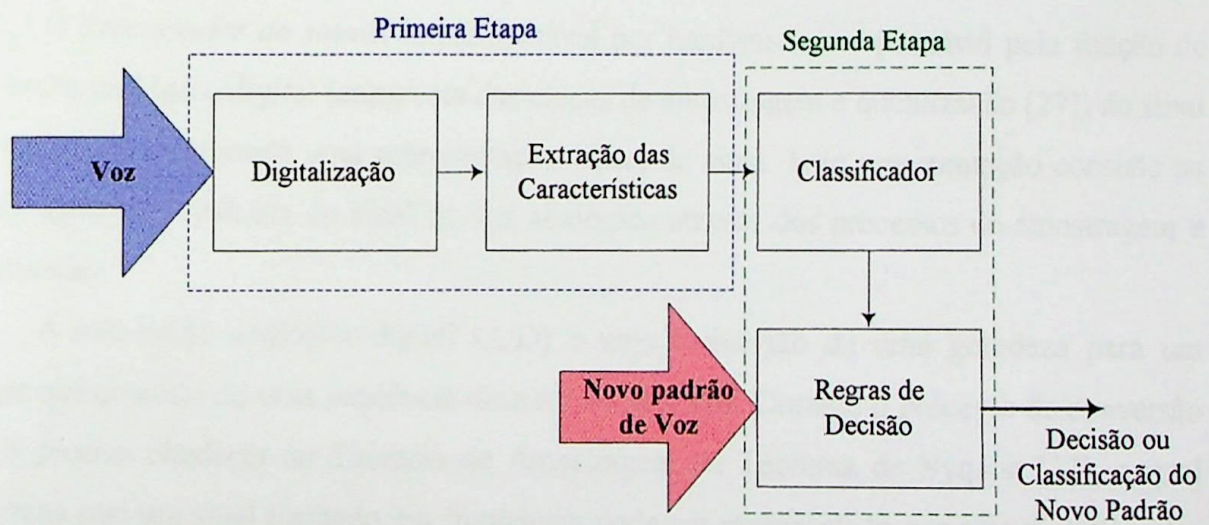


Figura 5.1 - Reconhecimento de Locutor em Diagrama de Blocos

5.2 – Representação Digital dos Sinais de Voz

A representação digital dos sinais de voz trata da aquisição e armazenamento dos sinais em questão e é fundamental para a extração das características. A representação digital pode ainda ser considerada como um pré-processamento no sinal de voz (realizada antes da

extração das características) responsável por tarefas como: conversão analógico/digital, amplificação, melhoria da relação sinal/ruído, entre outros.

O processamento do sinal de voz é realizado através de um computador a fim de capturar o sinal sonoro transformando-o em informação digital. Através da placa de som acoplada ao computador, é possível converter um sinal analógico em um sinal digital, tornando-o assim, um sinal discreto, e portanto, passível de ser manuseado por um sistema computadorizado.

Dessa forma, considera-se que a etapa de processamento de sinal é dividida basicamente em duas fases, são elas: *a aquisição dos sinais de voz*, e *o processador de sinais* propriamente ditos.

A *aquisição dos sinais de voz* é realizada a partir de um microfone, que é responsável pela conversão do sinal acústico em um sinal elétrico analógico, sendo responsável também pela “qualidade” do sinal de voz a ser processado. A qualidade está relacionada à fidelidade do conteúdo de frequências e à razão Sinal/Ruído do sinal. Como os sinais obtidos através do microfone são sinais analógicos, é necessário realizar a conversão A/D que transforma uma certa tensão elétrica em uma informação correspondente digital, sendo considerada uma discretização de um processo contínuo [26].

O *processador de sinais*, implementável por hardware, é responsável pela função de conversão analógico/digital (composta das etapas de amostragem e quantização [27]) do sinal de voz, a fim de permitir uma representação visual do sinal. Essa representação consiste na preservação da envoltória do sinal de voz analógico através dos processos de amostragem e quantização.

A conversão analógico/digital (A/D) é uma conversão de uma grandeza para um código que consiste de uma sequência discreta de números. Durante o processo de conversão A/D é preciso obedecer ao Teorema de Amostragem ou Teorema de Nyquist [27], o qual determina que um sinal limitado em frequência pode ser representado por amostras tomadas periodicamente no tempo, desde que essas amostras sejam tomadas numa taxa de pelo menos duas vezes maior do que a maior frequência contida no sinal amostrado. Assim, a frequência de amostragem deve ser superior a duas vezes a frequência máxima do sinal, de modo que o sinal contínuo possa ser recuperado a partir das amostras e sem perda de informação. Se a condição do Teorema de Nyquist não for mantida, sendo a amostragem mal feita, haverá o surgimento de “aliasing”, podendo haver perda de informação através de distorções. Com a distorção, falsas componentes harmônicas são criadas dentro do espectro, gerando assim sons confusos e distorcidos. Cabe lembrar que, a etapa de amostragem pode ser feita sem perda de

informação, no entanto na etapa de quantização podem surgir erros de quantização, sendo que esse erro pode ser baixo, se o número de níveis de quantização utilizado for elevado. A forma mais simples de conversão A/D é aquela que utiliza um quantizador uniforme, onde todos os intervalos ou níveis de quantização são iguais; conhecido também como PCM (“Pulse Code Modulation”) com quantização uniforme [28].

5.2.1 – Preparação dos Dados

No sistema proposto, a gravação dos sinais de voz realiza-se a partir da conversão analógico digital com uma taxa de amostragem de 8KHz e amostras em 16 bits, em canal no modo mono. Os sinais adquiridos são armazenados (gravados) em arquivos no formato .WAV da Microsoft WAVE. Para o sistema em questão utilizou-se um grupo de 7 locutores, representados por 4 mulheres e 3 homens, todos entre 25 e 40 anos, como mostra a tabela 5.1.

Tabela 5.1 - Lista de Locutores

Código de Identificação	Nome	Idade
01	Jusefina Oliveira Braga	30
02	Sandra Bianca Henriques	26
03	Marília Sanjad Henriques	25
04	Anderson Marcos Henriques	27
05	Adans Marcelo dos Santos	26
06	Mauro Rodrigues Leite Neto	34
07	Maria José Noronha Silvério	40

A fim de obter um certo grau de estabilidade nos sinais de voz, foi determinado que o conteúdo da forma de onda caracterizasse a vogal “A”, ou seja, os locutores deveriam pronunciar a vogal “A” de maneira clara, natural e contínua. Isso permite que o sinal possua poucas mudanças, de forma que a onda acústica seja considerada estacionária. Assim, inicialmente não foi necessário o uso de janelamento com superposição.

A escolha da pronúncia da vogal deve-se ao fato das vogais caracterizar melhor os locutores, e por apresentar alto grau de periodicidade.

Após a gravação e digitalização das amostras de voz, nesse momento com comprimento variável (em segundos), definiu-se que os sinais teriam a duração de 1 (um)

segundo, de forma que os mesmos passam por um corte (definindo-se o início e o fim da locução), e posteriormente por uma segmentação para que seja possível gerar padrões que caracterizam o locutor. A determinação de um tamanho fixo para os sinais de voz evita a variabilidade do tempo de pronúncia, fato que poderia necessitar de um ajustamento temporal.

Assim, para cada arquivo de som no formato .WAV (com duração de 1 segundo) são gerados 4 segmentos de 0.25 segundos (25ms) e cada segmento irá gerar um padrão referente ao locutor. Para cada locutor foram gravados 15 repetições da locução, das quais 10 repetições seriam utilizadas para o treinamento e 5 repetições seriam usadas para teste. O comprimento escolhido para os segmentos permite verificar a periodicidade do sinal e assim satisfazer as condições para a utilização da Transformada de Fourier.

5.2.2 – Tratamento dos Sinais de Voz

Após obter os arquivos em formato WAV e antes da extração das características propriamente dita, são executadas tarefas como leitura do arquivo WAV (realizado através do comando `wavread`), normalização do sinal, determinação do tamanho dos segmentos (mencionado no item 5.2.1), aplicação da janela de Hamming, definição do tamanho da Transformada de Fourier, entre outras que estão relacionadas ao método usado para análise.

O comando `wavread` do Matlab permite ler arquivos com extensão do tipo WAV, considerando os valores para a amplitude na faixa de $[-1,+1]$. Assim, o comando `[sinal,freq,nbits]=wavread(nome_do_arquivo)` irá retornar na variável `sinal` os dados amostrados, em `freq` a frequência de amostragem, e em `nbits` o número de bits usados para codificar os dados.

A normalização do sinal é necessária devido à possibilidade de gravação com níveis diferentes de volume, e consiste em equalizar as amplitudes dos sinais de voz. A tarefa de normalização coloca os valores das amostras num determinado intervalo [13]. O intervalo aplicado foi de $[-0.5,+0.5]$, ou seja, de -500 mV a $+500\text{mV}$ (1 V, pico a pico), conforme [5]. A normalização é obtida então da seguinte forma:

```
AmpMx = max(sinal);  
AmpMn = min(sinal);  
ValorMax=max(AmpMx,abs(AmpMn));  
Sinal=sinal*0.5/ ValorMax;
```


Após a leitura do arquivo e normalização do sinal de voz, é necessário determinar o tamanho, em milisegundos, de uma janela ou segmento que será usado para a obtenção das características, sendo que o tamanho por “default” é de 25 ms. Porém, para cada segmento é necessária a aplicação de uma janela.

A janela escolhida foi a janela de Hamming, sendo esta necessária devido existir descontinuidade nos limites ou bordas. A descontinuidade resultará em componentes de alta frequência, de forma que essas componentes podem vir a ser maiores que o limite do teorema de amostragem [13]. A técnica usada para tratar essa questão é chamada de janelamento (“windowing”) e consiste em atenuar o efeito do fenômeno de “Gibbs” (“ripple” ou ondulação em amplitude na resposta em frequência da janela) no truncamento da série de Fourier da janela considerada.

A figura 5.2 ilustra um exemplo de sinal de voz e a figura 5.3 a aplicação da janela de Hamming o sinal da figura 5.2.

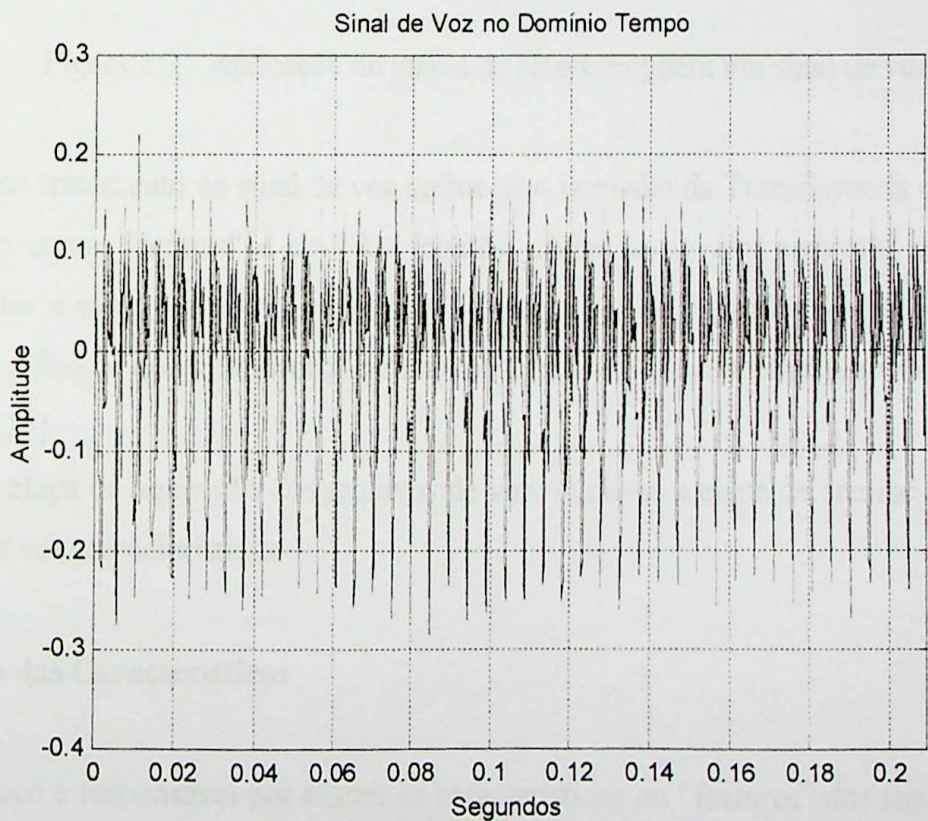


Figura 5.2 - Exemplo de um sinal de voz

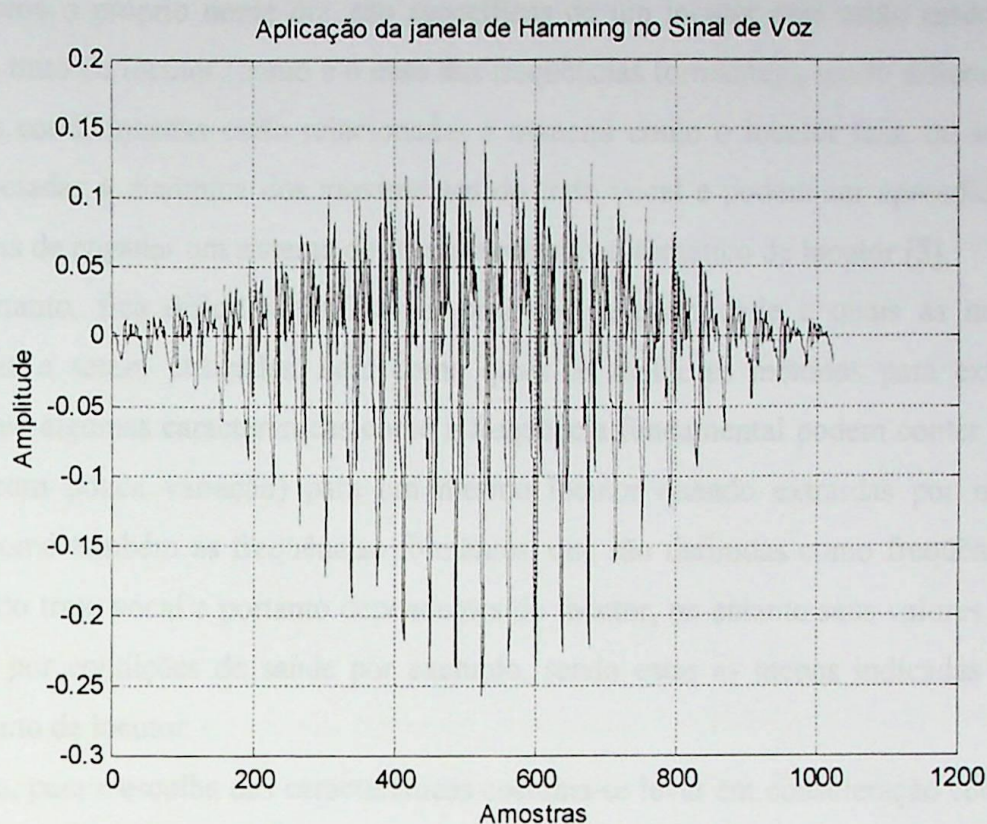


Figura 5.3 - Aplicação da janela de Hamming para um sinal de voz

Ainda no tratamento do sinal de voz define-se o tamanho da Transformada de Fourier. O valor usado como “default” é de 1024 pontos, assim se o valor indicado possuir um tamanho inferior a esse, o sistema irá aplicar a técnica de “zero padding”, a qual adiciona zeros no sinal, a fim de que a informação contenha como total de amostras um valor que seja potência de 2 [13].

Após a etapa de aquisição dos arquivos de voz, inicia-se a etapa de análise dos sinais, a fim de extrair suas características.

5.3 – Extração das Características

Este bloco é responsável por extrair as características ou “features” dos sinais de voz, a fim de reduzir a redundância de informações contidas no sinal. Tais características devem conter informações referentes aos locutores com o objetivo de reconhecê-los. Para isso podem ser utilizados vários métodos de análise, os quais foram abordados no capítulo 4. É preciso lembrar ainda que, as diferenças físicas e comportamentais entre os locutores, diferenciam os tipos de características.

Basicamente, existem dois tipos de características: as inerentes e as condicionadas. As inerentes, como o próprio nome diz, são específicas de um locutor pois estão associadas à anatomia do trato do locutor (como é o caso das frequências formantes), sendo difíceis de ser imitadas. As condicionadas estão relacionadas à maneira como o locutor fala, ou seja, são aquelas associadas à dinâmica dos movimentos do trato vocal e podem ser aprendidas, por isso são fáceis de enganar um sistema de reconhecimento automático de locutor [5].

Entretanto, fica difícil identificar, a princípio, a quantidade e quais as melhores características a serem utilizadas, bem como quais os melhores métodos para extraí-las, mesmo porque algumas características como a frequência fundamental podem conter valores diferentes (com pouca variação) para um mesmo locutor quando extraídas por métodos diferentes; como também as frequências formantes, que são definidas como frequências de ressonância do trato vocal e portanto dependentes do locutor, no entanto seus valores podem ser afetados por condições de saúde por exemplo, sendo estas as menos indicadas para o reconhecimento de locutor.

Assim, para a escolha das características costuma-se levar em consideração condições como: facilidade de medida, grande variação entre diferentes locutores sendo consistente para um mesmo locutor, estabilidade e não ser alterada por condições de saúde do locutor, não ser afetada pelo dispositivo de gravação e nem por ruídos, e não ser vulnerável a imitações.

Estudos indicam que o ideal é utilizar poucas características independentes, com pouca variação para um mesmo locutor e muita variação para locutores diferentes.

Após testes com algumas características verificou-se que a frequência fundamental atende às condições mencionadas anteriormente, além disso possui pouca variação para um mesmo locutor e não é afetada pelo dispositivo de gravação, estando também livre de variações do nível de gravação do sinal e à distância locutor/microfone, como mencionado anteriormente.

Como características para os locutores utilizou-se a frequência do “pitch” e as frequências de maior relevância. O “pitch” é uma característica importante para um sistema de reconhecimento de locutor e sua obtenção foi realizada a partir do método cepstrum. Foi implementado também o método que trabalha no domínio do tempo, com o uso de limiares inferior e superior baseados nas amplitudes das amostras para determinação dos picos do sinal de voz no domínio do tempo, permitindo assim a identificação da periodicidade, e uma comparação entre o “pitch” e a periodicidade em questão.

Além do “pitch” são consideradas as frequências de maior relevância, como mencionado anteriormente, sendo obtidas após o cálculo da densidade espectral de potência,

devido ao fato de tais frequências apresentarem um certo padrão para um mesmo locutor e por apresentar diferenças para locutores diferentes. A Transformada Rápida de Fourier é usada como ferramenta fundamental para extração das frequências de maior relevância de cada segmento. Dessa forma, a identificação de um locutor também se baseia nas informações, ou características, obtidas a partir do espectro do sinal de voz. As figuras 5.4 e 5.5 ilustram exemplos que permitem observar as frequências de maior relevância, a partir da densidade espectral. Os gráficos permitem visualizar 7 (sete) frequências consideradas em testes iniciais, identificadas como F1, F2, F3, F4, F5, F6 e F7. É importante ressaltar que o número de características pode ser variável.

As frequências obtidas são então usadas para estruturar um arquivo que irá conter as características do sinal de voz de um determinado locutor.

5.4 – Armazenamento de Padrões

A fase preliminar à segunda etapa do Sistema de Reconhecimento de Locutor é denominada armazenamento de padrões, a qual é responsável pela preparação da base de dados, ou base de conhecimento.

Entretanto, é preciso lembrar que o armazenamento depende do tipo de método usado na classificação. Para o sistema proposto torna-se necessário armazenar as informações em tabelas de dados, as quais são utilizadas na fase de treinamento com o objetivo de comparar padrões e gerar regras.

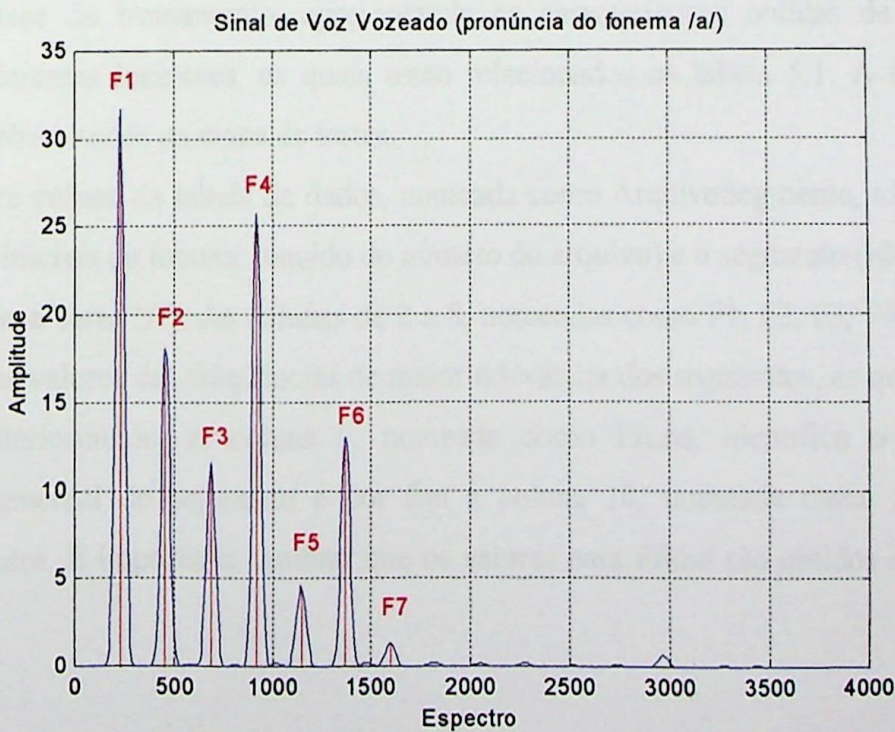


Figura 5.4 - Densidade espectral de um segmento vozeado

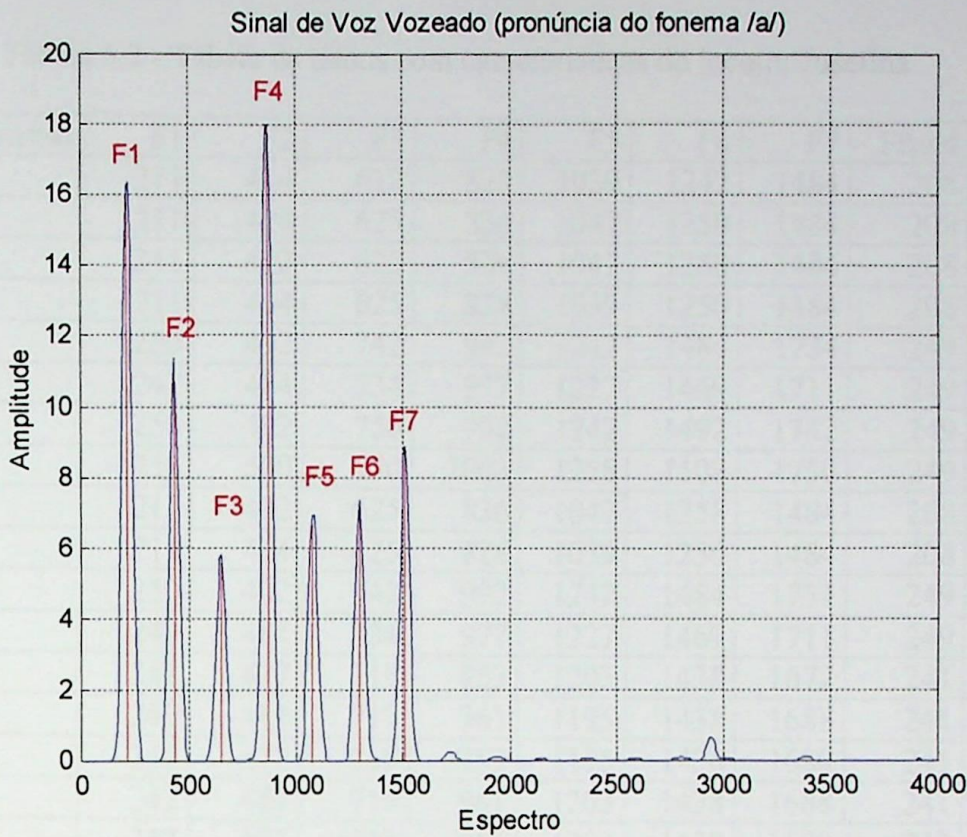


Figura 5.5 - Frequências de maior relevância

5.5 – Tabelas de Dados Contendo Características dos Locutores

As tabelas 5.2, 5.3, 5.4, 5.5, 5.6, 5.7 e 5.8, representam as tabelas de dados que são usadas para a fase de treinamento, apresentando as características obtidas de arquivos gravados por diferentes locutores, os quais estão relacionados na tabela 5.1. A tabela 5.9 apresenta os padrões usados na etapa de testes.

A primeira coluna da tabela de dados, nomeada como `ArquivoSegmento`, identifica o arquivo (com as iniciais do locutor seguido do número do arquivo) e o segmento (identificado pelo número após a barra “/”). As colunas de 2 a 8, nomeadas como `F1`, `F2`, `F3`, `F4`, `F5`, `F6` e `F7`, identificam os valores das frequências de maior relevância dos segmentos, as quais foram mencionadas anteriormente. A coluna 9, nomeada como `Ffund`, identifica o valor da frequência fundamental do segmento e por fim a coluna 10, nomeada como `Saida`, irá identificar o locutor. É importante lembrar que os valores para `Ffund` são obtidos a partir da análise cepstral.

Tabela 5.2 - Tabela de dados com características do locutor Jusefina

ArquivoSegmento	F1	F2	F3	F4	F5	F6	F7	Ffund	Saida
J1/1	211	414	617	828	1039	1242	1484	208	1
J1/2	211	414	625	836	1047	1250	1484	208	1
J1/3	211	422	625	836	1047	1258	1484	208	1
J1/4	211	414	625	828	1039	1250	1484	208	1
J2/1	250	492	742	992	1242	1484	1734	249	1
J2/2	242	484	734	977	1227	1469	1711	249	1
J2/3	250	492	750	992	1242	1492	1742	249	1
J2/4	250	500	750	1008	1258	1508	1750	249	1
J3/1	211	422	625	836	1047	1258	1484	208	1
J3/2	211	414	625	828	1039	1250	1484	208	1
J3/3	250	492	742	992	1242	1484	1734	249	1
J3/4	242	484	734	977	1227	1469	1711	249	1
J4/1	242	477	719	953	1203	1438	1672	241	1
J4/2	242	484	719	961	1195	1438	1688	241	1
J4/3	242	477	719	953	1195	1438	1680	241	1
J4/4	242	484	719	961	1203	1438	1688	241	1
J5/1	242	477	719	953	1203	1438	1680	242	1
J5/2	242	484	727	969	1211	1453	1695	242	1
J5/3	242	484	727	977	1211	1453	1695	242	1
J5/4	242	484	727	969	1203	1453	1695	242	1
J6/1	234	469	695	930	1175	1391	1625	234	1
J6/2	234	469	703	938	1118	1406	1641	234	1
J6/3	234	469	711	945	1180	1414	1656	234	1
J6/4	234	469	703	930	1172	1398	1633	234	1
J7/1	242	477	719	953	1203	1438	1680	245	1
J7/2	242	492	734	984	1227	1469	1711	245	1
J7/3	242	492	734	984	1234	1469	1719	245	1
J7/4	250	500	750	1000	1242	1492	1742	245	1
J8/1	250	500	758	1008	1258	1508	1758	252	1
J8/2	250	508	758	1008	1266	1516	1766	252	1
J8/3	250	500	758	1008	1258	1516	1758	252	1
J8/4	258	508	766	1016	1273	1523	1781	252	1
J9/1	250	516	773	1023	1281	1539	1797	257	1
J9/2	258	516	766	1023	1281	1539	1797	257	1
J9/3	258	508	766	1016	1273	1531	1781	257	1
J9/4	250	516	773	1031	1281	1547	1797	257	1
J10/1	258	516	773	1023	1281	1539	1797	256	1
J10/2	258	508	766	1016	1273	1531	1789	256	1
J10/3	258	508	766	1023	1281	1531	1789	256	1
J10/4	258	508	766	1023	1273	1531	1789	256	1

Tabela 5.3 - Tabela de dados com características do locutor Sandra

ArquivoSegmento	F1	F2	F3	F4	F5	F6	F7	Ffund	Saida
Sb1/1	219	438	656	875	1094	1313	1531	215	2
Sb1/2	219	430	648	867	1078	1289	1508	215	2
Sb1/3	211	430	641	859	1063	1281	1492	215	2
Sb1/4	219	430	648	859	1078	1297	1508	215	2
Sb2/1	219	438	656	875	1094	1313	1531	219	2
Sb2/2	219	438	656	875	1102	1320	1539	219	2
Sb2/3	219	438	656	883	1094	1313	1531	219	2
Sb2/4	219	430	648	867	1078	1297	1516	219	2
Sb3/1	219	438	656	867	1094	1313	1531	218	2
Sb3/2	219	438	656	875	1086	1313	1523	218	2
Sb3/3	219	438	656	867	1086	1305	1523	218	2
Sb3/4	219	438	656	875	1094	1313	1531	218	2
Sb4/1	219	445	664	883	1109	1328	1555	220	2
Sb4/2	227	445	672	891	1125	1344	1570	223	2
Sb4/3	219	445	664	891	1109	1328	1555	220	2
Sb4/4	227	445	672	891	1117	1344	1563	223	2
Sb5/1	227	453	680	906	1133	1359	1586	226	2
Sb5/2	227	445	672	898	1125	1344	1578	226	2
Sb5/3	227	453	672	898	1117	1344	1570	226	2
Sb5/4	227	453	680	906	1133	1359	1586	226	2
Sb6/1	227	445	672	891	1125	1344	1570	225	2
Sb6/2	227	445	672	891	1117	1344	1563	225	2
Sb6/3	227	445	672	898	1125	1352	1570	225	2
Sb6/4	227	453	680	906	1133	1359	1594	225	2
Sb7/1	227	453	680	906	1133	1359	1586	228	2
Sb7/2	227	453	672	898	1125	1352	1570	228	2
Sb7/3	227	461	688	914	1141	1367	1602	228	2
Sb7/4	227	453	688	914	1141	1375	1594	228	2
Sb8/1	227	453	688	914	1141	1367	1594	228	2
Sb8/2	227	453	680	906	1133	1359	1586	228	2
Sb8/3	227	453	680	906	1133	1359	1594	228	2
Sb8/4	227	453	680	906	1133	1367	1594	228	2
Sb9/1	234	461	695	922	1156	1383	1617	228	2
Sb9/2	227	453	688	906	1141	1367	1602	228	2
Sb9/3	227	453	688	914	1141	1367	1594	228	2
Sb9/4	227	453	680	898	1125	1352	1578	228	2
Sb10/1	219	438	656	875	1094	1313	1531	215	2
Sb10/2	219	430	648	867	1078	1289	1508	215	2
Sb10/3	211	430	641	859	1063	1281	1492	215	2
Sb10/4	219	430	648	859	1078	1297	1508	215	2

Tabela 5.4 - Tabela de dados com características do locutor Marília

ArquivoSegmento	F1	F2	F3	F4	F5	F6	F7	Ffund	Saida
Marilia1/1	211	422	633	844	1055	1258	1477	211	3
Marilia1/2	211	422	633	844	1055	1266	1477	211	3
Marilia1/3	219	422	633	844	1055	1266	1477	211	3
Marilia1/4	211	422	633	844	1047	1266	1477	211	3
Marilia2/1	211	430	648	867	1078	1297	1516	213	3
Marilia2/2	211	430	641	859	1078	1289	1500	213	3
Marilia2/3	203	414	625	836	1039	1250	1461	213	3
Marilia2/4	211	422	625	836	1047	1258	1469	213	3
Marilia3/1	211	422	633	844	1055	1273	1477	213	3
Marilia3/2	211	422	633	852	1063	1273	1484	213	3
Marilia3/3	211	430	641	859	1070	1281	1500	213	3
Marilia3/4	219	430	648	859	1070	1289	1500	213	3
Marilia4/1	219	438	656	883	1102	1320	1539	218	3
Marilia4/2	219	438	664	883	1102	1320	1547	218	3
Marilia4/3	219	438	648	875	1086	1305	1523	218	3
Marilia4/4	219	438	656	867	1086	1305	1516	218	3
Marilia5/1	219	430	641	859	1070	1289	1500	214	3
Marilia5/2	211	430	641	859	1070	1289	1500	214	3
Marilia5/3	219	438	656	867	1086	1297	1516	214	3
Marilia5/4	211	430	633	852	1063	1273	1492	214	3
Marilia6/1	219	445	664	891	1109	1336	1555	220	3
Marilia6/2	227	453	680	906	1133	1352	1586	227	3
Marilia6/3	234	461	695	922	1156	1391	1617	227	3
Marilia6/4	219	453	672	898	1125	1344	1570	220	3
Marilia7/1	227	453	680	906	1133	1359	1586	223	3
Marilia7/2	211	438	656	867	1094	1305	1523	220	3
Marilia7/3	227	445	672	898	1125	1352	1570	223	3
Marilia7/4	234	461	695	922	1156	1383	1617	223	3
Marilia8/1	227	445	672	898	1125	1352	1570	235	3
Marilia8/2	242	469	703	938	1172	1405	1641	241	3
Marilia8/3	234	469	695	930	1164	1398	1633	235	3
Marilia8/4	242	477	719	961	1203	1438	1680	241	3
Marilia9/1	219	438	656	875	1094	1313	1539	220	3
Marilia9/2	227	453	688	914	1141	1367	1594	227	3
Marilia9/3	234	461	695	922	1156	1383	1609	227	3
Marilia9/4	219	445	664	891	1109	1336	1555	227	3
Marilia10/1	219	430	648	859	1078	1289	1508	219	3
Marilia10/2	227	445	664	883	1109	1328	1555	221	3
Marilia10/3	227	445	664	883	1102	1328	1547	221	3
Marilia10/4	211	438	656	867	1086	1305	1523	219	3

Tabela 5.5 - Tabela de dados com características do locutor Anderson

ArquivoSegmento	F1	F2	F3	F4	F5	F6	F7	Ffund	Saida
Andso1/1	141	273	406	547	688	820	961	143	4
Andso1/2	133	266	398	530	664	797	938	133	4
Andso1/3	133	266	398	530	664	797	930	133	4
Andso1/4	133	266	391	523	655	781	914	133	4
Andso2/1	133	266	391	523	656	789	922	133	4
Andso2/2	133	258	391	516	648	781	906	133	4
Andso2/3	133	273	406	539	672	805	938	133	4
Andso2/4	141	273	406	547	680	820	954	143	4
Andso3/1	141	281	430	563	703	844	984	141	4
Andso3/2	141	281	422	563	703	844	984	141	4
Andso3/3	141	281	422	563	703	844	992	141	4
Andso3/4	141	273	414	555	688	828	961	141	4
Andso4/1	141	281	414	547	695	836	969	143	4
Andso4/2	141	273	414	547	680	820	961	143	4
Andso4/3	133	273	406	531	672	813	945	133	4
Andso4/4	133	266	398	531	664	797	930	133	4
Andso5/1	156	305	460	609	766	914	1060	154	4
Andso5/2	156	313	461	617	768	921	1078	154	4
Andso5/3	150	305	453	602	758	914	1070	154	4
Andso5/4	148	305	453	609	758	906	1050	154	4
Andso6/1	133	266	398	530	664	797	930	133	4
Andso6/2	133	266	391	523	655	781	914	133	4
Andso6/3	133	273	406	531	672	813	945	133	4
Andso6/4	133	266	398	531	664	797	930	133	4
Andso7/1	156	305	460	609	766	914	1060	154	4
Andso7/2	156	313	461	617	768	921	1078	154	4
Andso7/3	133	266	398	530	664	797	930	133	4
Andso7/4	133	266	391	523	655	781	914	133	4
Andso8/1	133	266	391	523	655	781	914	133	4
Andso8/2	133	266	391	523	656	789	922	133	4
Andso8/3	133	258	391	516	648	781	906	133	4
Andso8/4	133	273	406	539	672	805	938	133	4
Andso9/1	141	273	406	547	680	820	954	143	4
Andso9/2	141	281	430	563	703	844	984	141	4
Andso9/3	141	281	422	563	703	844	984	141	4
Andso9/4	141	281	422	563	703	844	992	141	4
Andso10/1	133	266	391	523	655	781	914	133	4
Andso10/2	133	266	391	523	656	789	922	133	4
Andso10/3	141	273	406	547	680	820	954	143	4
Andso10/4	141	281	430	563	703	844	984	141	4

Tabela 5.6 - Tabela de dados com características do locutor Marcelo

ArquivoSegmento	F1	F2	F3	F4	F5	F6	F7	Ffund	Saida
Marc1/1	117	234	352	469	586	711	820	119	5
Marc1/2	117	234	359	484	602	719	836	119	5
Marc1/3	117	242	367	477	602	727	844	119	5
Marc1/4	125	242	356	477	602	719	844	121	5
Marc2/1	117	234	352	484	586	711	828	117	5
Marc2/2	117	234	352	453	594	703	828	117	5
Marc2/3	117	234	352	461	586	703	933	117	5
Marc2/4	117	234	359	477	602	719	836	117	5
Marc3/1	141	234	367	602	719	844	960	145	5
Marc3/2	125	242	367	500	602	727	844	121	5
Marc3/3	117	242	360	484	609	727	852	119	5
Marc3/4	117	242	352	484	594	719	836	119	5
Marc4/1	125	250	374	492	617	734	859	125	5
Marc4/2	125	242	363	477	602	727	852	125	5
Marc4/3	125	250	356	484	617	742	875	125	5
Marc4/4	125	250	421	508	633	750	883	125	5
Marc5/1	125	242	367	492	617	742	859	125	5
Marc5/2	125	250	375	508	633	755	883	125	5
Marc5/3	125	242	383	492	609	734	852	125	5
Marc5/4	125	250	360	500	625	742	883	125	5
Marc6/1	117	242	367	477	602	727	844	119	5
Marc6/2	125	242	356	477	602	719	844	121	5
Marc6/3	117	234	352	484	586	711	828	117	5
Marc6/4	117	234	352	453	594	703	828	117	5
Marc7/1	117	242	352	484	594	719	836	119	5
Marc7/2	125	242	367	492	617	742	859	125	5
Marc7/3	117	234	359	477	602	719	836	117	5
Marc7/4	117	242	367	477	602	727	844	119	5
Marc8/1	117	242	360	484	609	727	852	119	5
Marc8/2	117	242	352	484	594	719	836	119	5
Marc8/3	125	250	374	492	617	734	859	125	5
Marc8/4	125	242	363	477	602	727	852	125	5
Marc9/1	117	234	352	469	586	711	820	119	5
Marc9/2	117	234	359	484	602	719	836	119	5
Marc9/3	117	242	367	477	602	727	844	119	5
Marc9/4	125	242	356	477	602	719	844	121	5
Marc10/1	117	234	359	477	602	719	836	117	5
Marc10/2	117	242	367	477	602	727	844	119	5
Marc10/3	117	242	352	484	594	719	836	119	5
Marc10/4	125	242	356	477	602	719	844	121	5

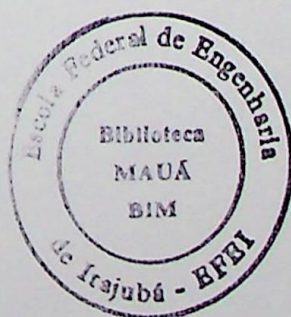


Tabela 5.7 - Tabela de dados com características do locutor Mauro

ArquivoSegmento	F1	F2	F3	F4	F5	F6	F7	Ffund	Saida
Mauro1/1	133	273	383	539	672	805	938	138	6
Mauro1/2	133	266	383	531	664	797	930	138	6
Mauro1/3	133	273	406	539	672	813	945	138	6
Mauro1/4	133	266	375	530	664	789	922	138	6
Mauro2/1	125	258	375	515	625	760	890	129	6
Mauro2/2	125	250	383	505	633	758	891	129	6
Mauro2/3	133	258	391	520	648	773	910	129	6
Mauro2/4	128	258	383	515	648	766	891	129	6
Mauro3/1	141	273	406	555	688	828	961	145	6
Mauro3/2	141	273	414	555	688	828	961	145	6
Mauro3/3	133	266	398	539	680	813	945	137	6
Mauro3/4	133	273	406	547	680	820	953	137	6
Mauro4/1	164	313	407	492	656	813	977	161	6
Mauro4/2	164	324	490	648	813	973	1136	161	6
Mauro4/3	164	320	484	640	800	961	1120	161	6
Mauro4/4	160	320	480	640	800	960	1120	160	6
Mauro5/1	164	328	492	656	820	992	1156	164	6
Mauro5/2	164	328	492	656	820	992	1148	164	6
Mauro5/3	164	328	484	648	813	977	1141	164	6
Mauro5/4	164	320	484	641	805	961	1125	164	6
Mauro6/1	188	367	555	742	930	1117	1300	184	6
Mauro6/2	188	367	555	742	922	1109	1297	184	6
Mauro6/3	180	367	547	734	914	1094	1280	180	6
Mauro6/4	180	367	547	734	914	1100	1280	180	6
Mauro7/1	164	336	500	664	836	1000	1172	168	6
Mauro7/2	172	336	500	672	844	1008	1180	168	6
Mauro7/3	172	336	508	680	844	1016	1180	168	6
Mauro7/4	164	328	492	664	828	992	1156	168	6
Mauro8/1	133	273	383	539	672	805	938	138	6
Mauro8/2	133	266	383	531	664	797	930	138	6
Mauro8/3	133	273	406	539	672	813	945	138	6
Mauro8/4	133	266	375	528	660	790	922	138	6
Mauro9/1	164	328	490	660	824	1000	1155	167	6
Mauro9/2	164	336	492	664	828	1000	1164	167	6
Mauro9/3	164	328	508	664	828	984	1164	167	6
Mauro9/4	164	328	492	664	836	1000	1164	167	6
Mauro10/1	180	352	531	703	883	1066	1234	176	6
Mauro10/2	172	352	523	703	883	1234	1410	176	6
Mauro10/3	172	352	516	703	875	1227	1405	176	6
Mauro10/4	172	352	516	700	875	1219	1390	176	6

Tabela 5.8 - Tabela de dados com características do locutor Maria José

ArquivoSegmento	F1	F2	F3	F4	F5	F6	F7	Ffund	Saida
ZZ1/1	203	398	602	797	984	1203	1398	201	7
ZZ1/2	195	398	602	797	984	1195	1391	199	7
ZZ1/3	203	398	594	797	1000	1195	1391	203	7
ZZ1/4	195	398	594	797	969	1195	1398	199	7
ZZ2/1	195	398	594	797	984	1188	1391	199	7
ZZ2/2	203	398	594	797	1023	1195	1398	201	7
ZZ2/3	203	398	602	805	992	1203	1406	201	7
ZZ2/4	203	398	594	797	984	1195	1391	203	7
ZZ3/1	203	406	602	805	992	1211	1406	201	7
ZZ3/2	203	406	609	813	1008	1211	1414	205	7
ZZ3/3	203	398	594	797	992	1195	1398	203	7
ZZ3/4	195	398	594	797	1000	1188	1391	199	7
ZZ4/1	203	406	609	813	1000	1211	1414	201	7
ZZ4/2	203	406	602	805	1000	1211	1406	201	7
ZZ4/3	203	406	609	805	1016	1211	1414	201	7
ZZ4/4	203	406	602	805	992	1211	1406	201	7
ZZ5/1	211	422	633	836	1000	1258	1469	209	7
ZZ5/2	211	422	633	836	1000	1258	1469	209	7
ZZ5/3	211	422	625	836	1000	1258	1469	209	7
ZZ5/4	211	422	633	844	1001	1258	1469	209	7
ZZ6/1	203	406	609	813	1008	1227	1430	203	7
ZZ6/2	203	406	609	813	1016	1219	1422	203	7
ZZ6/3	203	406	609	820	1031	1227	1430	203	7
ZZ6/4	203	406	609	820	1016	1227	1430	203	7
ZZ7/1	203	398	594	797	992	1195	1398	201	7
ZZ7/2	195	398	602	797	1008	1195	1391	199	7
ZZ7/3	195	398	594	789	1000	1188	1383	199	7
ZZ7/4	195	398	594	797	1008	1195	1391	199	7
ZZ8/1	211	422	633	844	1060	1273	1485	213	7
ZZ8/2	211	422	633	844	1050	1266	1470	213	7
ZZ8/3	211	422	641	852	1063	1274	1483	213	7
ZZ8/4	211	422	641	852	1063	1274	1480	213	7
ZZ9/1	195	383	578	773	969	1156	1352	192	7
ZZ9/2	188	383	570	758	953	1141	1328	192	7
ZZ9/3	195	383	578	766	969	1156	1344	192	7
ZZ9/4	195	383	578	773	953	1156	1352	192	7
ZZ10/1	195	398	594	797	992	1195	1391	198	7
ZZ10/2	195	391	594	789	977	1180	1375	198	7
ZZ10/3	195	398	594	789	984	1188	1383	198	7
ZZ10/4	195	398	594	797	992	1188	1391	198	7

Tabela 5.9 - Tabela de dados para teste com características de diversos locutores

ArquivoSegmento	F1	F2	F3	F4	F5	F6	F7	Ffund	Saida
testemari11-1	211	422	633	836	1047	1258	1469	210	3
testemari11-2	211	430	648	859	1078	1289	1508	213	3
testemari11-3	211	422	633	844	1055	1273	1484	211	3
testemari11-4	211	422	633	844	1055	1266	1477	211	3
testemari12-1	211	414	625	828	1039	1250	1453	211	3
testemari12-2	211	422	625	836	1047	1258	1469	211	3
testemari12-3	211	422	633	844	1055	1266	1477	214	3
testemari12-4	211	422	625	836	1047	1258	1469	211	3
testemari13-1	211	422	633	844	1047	1258	1469	211	3
testemari13-2	211	422	633	836	1047	1258	1469	211	3
testemari13-3	203	414	617	820	1023	1234	1438	202	3
testemari13-4	203	406	617	820	1023	1230	1438	203	3
testemari14-1	219	430	641	859	1070	1289	1500	219	3
testemari14-2	227	445	672	891	1117	1344	1563	227	3
testemari14-3	219	430	648	859	1078	1297	1516	219	3
testemari14-4	219	438	656	883	1102	1320	1539	219	3
testemari15-1	211	422	625	836	1047	1258	1469	211	3
testemari15-2	211	422	633	844	1047	1258	1469	211	3
testemari15-3	211	422	633	836	1047	1258	1469	211	3
testemari15-4	203	414	617	820	1023	1234	1438	203	3
Testeandso1-1	141	273	406	547	688	820	960	141	4
Testeandso1-2	141	273	406	547	680	820	953	141	4
Testeandso1-3	133	266	398	539	672	813	945	133	4
Testeandso1-4	133	266	391	523	664	789	922	133	4
Testeandso2-1	133	258	391	516	648	781	914	133	4
Testeandso2-2	133	266	406	539	672	805	938	133	4
Testeandso2-3	133	273	406	547	680	820	953	133	4
Testeandso2-4	133	273	406	547	680	820	953	133	4
Testeandso3-1	141	273	422	563	703	844	984	141	4
Testeandso3-2	141	281	422	555	703	844	984	141	4
Testeandso3-3	141	273	414	555	688	820	961	141	4
Testeandso3-4	133	273	406	539	672	813	922	133	4
Testeandso4-1	133	273	406	547	680	813	953	133	4
Testeandso4-2	133	273	406	539	680	813	945	133	4
Testeandso4-3	133	273	406	539	680	820	953	133	4
Testeandso4-4	133	266	406	539	672	813	945	133	4
Testeandso5-1	141	273	406	547	688	820	961	141	4
Testeandso5-2	141	273	406	547	680	820	953	141	4
Testeandso5-3	133	266	398	539	672	813	945	133	4
Testeandso5-4	133	266	391	523	664	789	922	133	4
Testeandso6-1	133	266	398	539	664	797	930	133	4
Testeandso6-2	133	266	398	539	664	797	930	133	4
Testeandso6-3	133	273	406	539	672	805	945	133	4
Testeandso6-4	133	266	391	523	656	789	925	133	4
Testemarc1-1	125	242	367	492	617	734	859	125	5
Testemarc1-2	125	250	367	492	617	742	867	125	5

Tabela 5.9: Continuação

Testemarc1-3	117	250	375	492	625	742	867	117	5
Testemarc1-4	125	250	375	492	617	750	867	125	5
Testemarc2-1	125	250	375	492	625	742	867	125	5
Testemarc2-2	125	242	367	492	609	734	852	125	5
Testemarc2-3	125	250	375	500	617	742	867	125	5
Testemarc2-4	125	250	367	500	617	742	867	125	5
Testemarc3-1	125	258	391	508	648	789	910	125	5
Testemarc3-2	133	258	391	508	648	789	907	133	5
Testemarc3-3	133	258	383	508	641	789	900	133	5
Testemarc3-4	133	258	398	508	656	789	912	133	5
Testemarc4-1	125	250	383	508	647	789	897	125	5
Testemarc4-2	125	258	391	508	648	789	910	125	5
Testemarc4-3	125	258	383	508	641	789	900	125	5
Testemarc4-4	125	258	383	508	641	766	891	125	5
Testemarc5-1	125	250	375	492	625	742	870	125	5
Testemarc5-2	125	242	367	492	609	734	860	125	5
Testemarc5-3	125	250	372	492	617	742	860	125	5
Testemarc5-4	125	250	367	500	625	750	870	125	5
Tesjj1-1	234	461	695	930	1156	1391	1625	234	1
Tesjj1-2	219	445	672	891	1117	1344	1563	219	1
Tesjj1-3	234	453	680	906	1125	1352	1586	234	1
Tesjj2-1	242	469	719	953	1196	1432	1672	242	1
Tesjj2-2	250	492	734	977	1220	1464	1711	250	1
Tesjj4-1	258	508	766	1023	1275	1534	1790	258	1
Tesjj3-2	242	492	734	977	1227	1470	1712	242	1
Tesjj5-1	258	516	781	1039	1297	1555	1720	258	1
Tessb2-1	211	422	625	836	1039	1250	1484	211	2
Tessb1-1	227	453	680	906	1133	1352	1578	227	2
Ts1-1	219	438	650	867	1080	1297	1523	218	2
Ts1-2	211	425	640	852	1060	1273	1500	219	2
Ts1-3	219	440	660	880	1100	1320	1540	218	2
Ts1-4	219	430	648	867	1080	1305	1500	217	2
Ts2-1	219	438	660	883	1103	1325	1545	219	2
Ts2-2	219	438	650	865	1080	1300	1520	219	2
Ta1-1	125	256	360	490	630	738	860	128	4
Ta1-2	125	255	356	490	620	733	850	126	4
Ta2-1	125	260	370	500	630	750	908	125	4
Ta2-2	120	250	366	500	620	730	850	120	4
Ta3-1	130	255	390	510	630	765	890	129	4
Ta3-2	130	260	400	520	650	780	905	130	4
Maur1-1	128	258	383	515	648	766	891	127	6
Maur1-2	141	273	406	555	688	828	961	142	6
Maur1-3	141	273	414	555	688	828	961	142	6
Maur1-4	133	266	398	539	680	813	945	133	6
Maur2-1	133	273	406	547	680	820	953	133	6
Maur2-2	164	313	407	492	656	813	977	165	6
Maur2-3	164	324	490	648	813	973	1136	164	6

Tabela 5.9: Continuação

Maur2-4	164	320	484	640	800	961	1120	164	6
Maur3-1	160	320	480	640	800	960	1120	160	6
Maur3-2	164	328	492	656	820	992	1156	164	6

TEORIA DOS CONJUNTOS APROXIMADOS (“ROUGH SETS”)

6.1 – Introdução

Os Sistemas de Informação, também conhecidos como Sistemas de Representação de Conhecimento (SRC), Tabelas de Dados, e Tabelas de Condição/Ação, são apresentados em forma de tabelas para representar o conhecimento, e são essenciais para a Inteligência Artificial; onde as colunas dos SRCs são chamadas de atributos e as linhas representam um objeto e contém informações ou dados que caracterizam um objeto.

O objetivo deste capítulo é apresentar a Teoria de Conjuntos Aproximados (TCA), que como o próprio nome diz se baseia na Teoria dos Conjuntos. Assim, são apresentados os principais conceitos da TCA, e sua aplicação no tratamento de incerteza em Sistemas de Representação de Conhecimento.

A principal noção dos conjuntos aproximados é que a imprecisão na representação dos dados faz com que se descubram padrões, entre os dados, considerados obscuros por não possuir muitos detalhes. A base da teoria dos conjuntos aproximados é a análise dos limites de discernibilidade, ou distinção, de um subconjunto X de objetos em um universo de discurso U caracterizado por um conjunto de atributos A . O menor grupo discernível de objetos é chamado de conjunto elementar, onde cada conjunto elementar inclui objetos com as mesmas características. Para cada conjunto X e seus subconjuntos, está associado: uma aproximação inferior de X , que contém a união dos conjuntos elementares que estão completamente incluídos em X e a aproximação superior de X , que contém a união dos conjuntos elementares que possuem uma interseção não vazia com X . O conjunto que contém objetos de uma aproximação superior que não são membros da aproximação inferior (ou a diferença das aproximações) é conhecido como região limite (“boundary”), ou região duvidosa. Um conjunto X é chamado de aproximado (“rough”) se a aproximação superior de X não for igual à aproximação inferior de X , ou se a região limite não é vazia; esses conceitos estão relacionados ao conjunto procurado [29].

Cabe lembrar que, as regras induzidas por uma aproximação inferior de um conjunto são consideradas regras certas, enquanto as regras induzidas por uma aproximação superior de um conjunto são consideradas regras possíveis [30].

A figura 6.1 ilustra um exemplo de conjunto procurado, sua aproximação inferior e aproximação superior.

Outro importante conceito em conjuntos aproximados é que o mesmo busca reduzir o conjunto de atributos inicial, onde a partir da redução, ou do conjunto mínimo de atributos, é possível classificar objetos com a mesma precisão que o conjunto original de atributos. A eliminação de atributos redundantes pode ajudar na identificação de “poderosas” regras, bem como regras de classificação não redundantes.

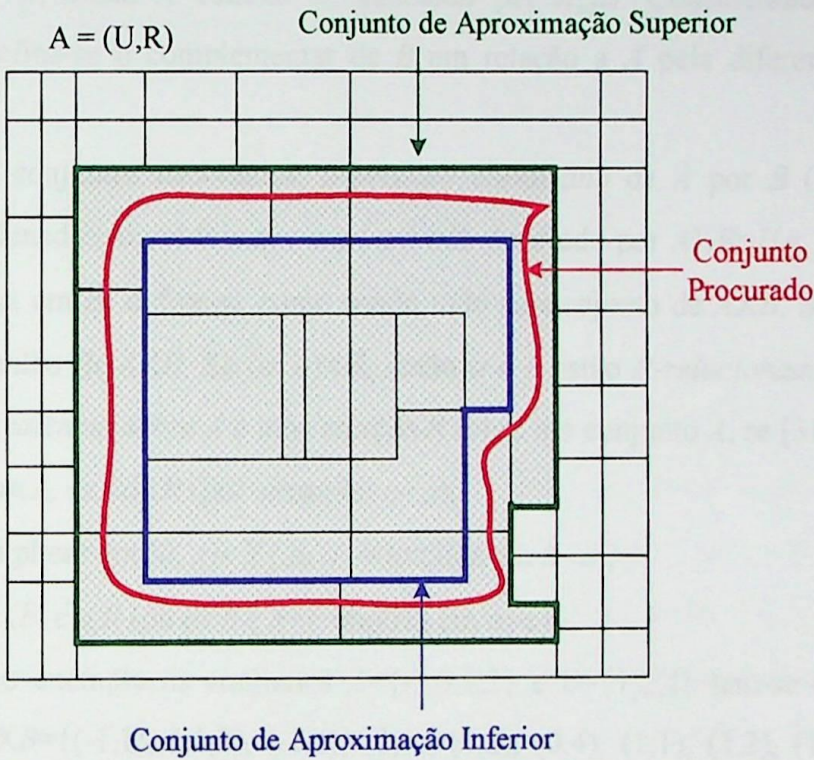


Figura 6.1 - Conjunto Procurado, Aproximação Superior e Inferior

6.2 – Teoria dos Conjuntos Aproximados

O Conjunto Aproximado foi proposto por Pawlak [29] como um modelo matemático para manipular as informações incertas, vagas e imprecisas encontradas em aplicações de Inteligência Artificial.

A fim de obter uma melhor compreensão dos conceitos relacionados à Teoria de Conjuntos Aproximados, serão apresentadas algumas definições referentes à Teoria dos Conjuntos, já que os conjuntos aproximados podem ser considerados como uma extensão da teoria dos conjuntos.

6.2.1 – Conjuntos e Relações

Intuitivamente pode-se visualizar conjunto como sinônimo de coleção. Um elemento é uma unidade indivisível que compõe a estrutura de um conjunto. Assim, um *conjunto* é uma coleção de objetos, denominados de elementos.

Dados dois conjuntos A e B , o conjunto B é *subconjunto* de A se qualquer elemento de B é também elemento de A ($\forall x \in B \rightarrow x \in A$), denota-se “ B é subconjunto de A ” por $B \subseteq A$. Se B é subconjunto de A , então A contém B , denotado por $A \supseteq B$. Considerando B como subconjunto de A , define-se o complementar de B em relação a A pela diferença $A - B$, denotado por $-B$.

Dados A e B conjuntos não-vazios, o *produto cartesiano* de A por B (AXB), é o conjunto de pares ordenados (a, b) tais que $a \in A$ e $b \in B$, denotado por $AXB = \{(a, b) \mid a \in A \text{ e } b \in B\}$. A *relação* de A em B , define-se como sendo todo subconjunto de AXB . Se R é uma relação de A em B , então $R \subseteq AXB$. Se $(a, b) \in R$, então a e b estão *R-relacionados* ou aRb . Uma *Relação de Equivalência sobre A* é uma relação R sobre um conjunto A , se [31]:

- para cada $a \in A$, $(a, a) \in R$ (por exemplo, $a=a$),
- $(a, b) \in R$ implicar em $(b, a) \in R$ (ou $a=b$ implica em $b=a$),
- $(a, b) \in R$, e $(b, c) \in R$ (ou $a=b$ e $b=c$ implica em $a=c$).

Tomando como exemplo os conjuntos $A = \{-1, 0, 1, 2\}$ e $B = \{1, 2, 4\}$ tem-se o seguinte produto cartesiano $AXB = \{(-1, 1), (-1, 2), (-1, 4), (0, 1), (0, 2), (0, 4), (1, 1), (1, 2), (1, 4), (2, 1), (2, 2), (2, 4)\}$. A partir do produto cartesiano AXB pode-se obter várias relações, tais como $R_1 = \{(-1, 2), (0, 1), (2, 2)\}$, $R_2 = \{(a, b) \in AXB \mid a=b\}$ assim $R_2 = \{(1, 1), (2, 2)\}$, etc.

6.2.2 – Conjuntos Aproximados

Um espaço aproximado é um par ordenado $A = (U, R)$, onde U é um conjunto não vazio, chamado de conjunto universo e R é uma relação de equivalência sobre U , denominada relação de indiscernibilidade, onde não se consegue distinguir seus elementos. A figura 6.2 ilustra um espaço aproximado.

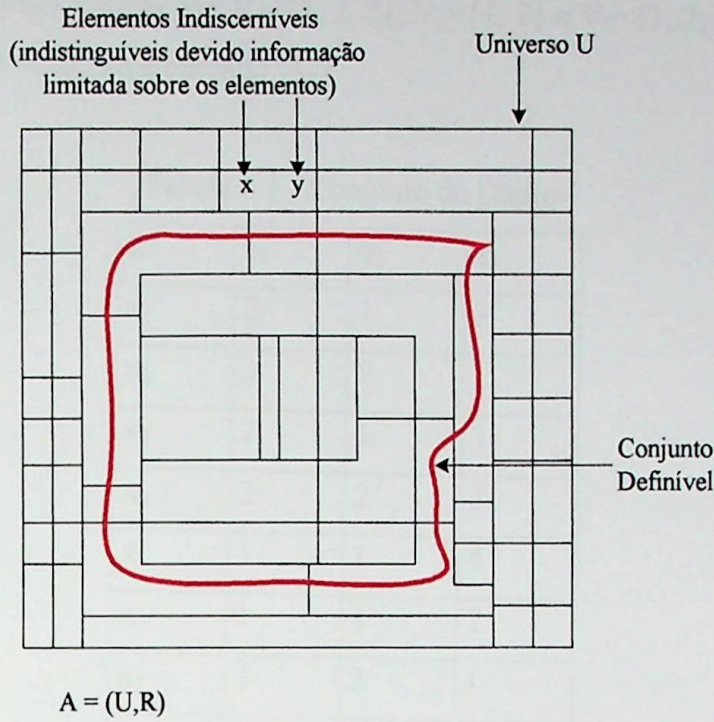


Figura 6.2 - Espaço Aproximado $A=(U,R)$

As classes de equivalência induzidas pela relação R em U , denotada por U/R , são chamadas de conjuntos elementares. Um conjunto vazio $\emptyset$ sempre será um conjunto elementar para todo espaço aproximado A . Sendo E um conjunto elementar em A , e $x, y \in E$, então x e y são indiscerníveis em E , ou seja, não se consegue distinguir x de y . Nesse caso, a classe de equivalência definida por x é a mesma que a classe de equivalência definida por y , de forma que $[x]_R = [y]_R$.

Assim, para todo conjunto de atributos $B \subset A$ (onde B está contido em A), uma relação de indiscernibilidade $Ind(B)$ é definida da seguinte maneira: dois objetos, x_i e x_j , são indiscerníveis pelo conjunto de atributos B em A , se $b(x_i) = b(x_j)$ para todo $b \in B$. A classe de equivalência de $Ind(B)$, ou conjunto elementar em B , representa o menor grupo discernível de objetos.

Para melhor entender o conceito de relação de indiscernibilidade, considere o exemplo 6.1.

Exemplo 6.1: O conjunto de dados da tabela 6.1 representa um sistema de informação. A partir da terminologia da teoria dos conjuntos aproximados tem-se: o conjunto universo $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$ como sendo o conjunto de objetos, e os atributos $A = \{a_1, a_2, a_3\}$ como sendo o conjunto de variáveis. O domínio dos atributos, ou conjunto de valores

possíveis para o atributo, é dado por: $V_1=\{1, 2, 3\}$, $V_2=\{1, 2\}$ e $V_3=\{1, 2, 3, 4\}$. Cada linha da tabela representa um conjunto elementar [32].

Tabela 6.1 - Conjunto de Dados

U	a ₁	a ₂	a ₃
x ₁	2	1	3
x ₂	3	2	1
x ₃	2	1	3
x ₄	2	2	3
x ₅	1	1	4
x ₆	1	1	2
x ₇	3	2	1
x ₈	1	1	4
x ₉	2	1	3
x ₁₀	3	2	1

No exemplo 6.1 é fácil notar que existem dois objetos idênticos no conjunto de dados, são eles x_1 e x_3 , de forma que não é possível distinguir os mesmos se baseando apenas na informação disponível. A partir da tabela 6.1 é possível agrupar os objetos idênticos, considerando todas as variáveis; assim obtém-se a tabela 6.2.

Tabela 6.2 - Agrupamento de Objetos Idênticos, considerando a_1 , a_2 , e a_3

U/R	a ₁	a ₂	a ₃
{x ₁ , x ₃ , x ₉ }	2	1	3
{x ₂ , x ₇ , x ₁₀ }	3	2	1
{x ₄ }	2	2	3
{x ₅ , x ₈ }	1	1	4
{x ₆ }	1	1	2

A tabela 6.3 mostra um outro agrupamento, ou conjuntos elementares resultantes, na consideração de apenas dois atributos (a_1 , a_2), de forma que a relação de indiscernibilidade seria limitada para o subconjunto $C = \{a_1, a_2\}$.

Tabela 6.3 - Agrupamento de Objetos Idênticos, considerando a_1 , e a_2

U/C	a_1	a_2
$\{x_1, x_3, x_9\}$	2	1
$\{x_2, x_7, x_{10}\}$	3	2
$\{x_4\}$	2	2
$\{x_5, x_6, x_8\}$	1	1

A união finita de conjuntos elementares define um conjunto *definível* em A , mostrado na figura 6.2. Onde a construção de conjuntos elementares é o primeiro passo na classificação com conjuntos aproximados.

Exemplo 6.2: Seja $U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$ e $U/R = \{\{x_1, x_3\}, \{x_2, x_5, x_6\}, \{x_4\}\}$. $A = (U, R)$ é um espaço aproximado, devido R ser uma relação de equivalência sobre U . Dessa forma obtém-se os seguintes conjuntos elementares em A :

$$D_1 = \{x_1, x_3\}, D_2 = \{x_2, x_5, x_6\}, D_3 = \{x_4\}, D_4 = \emptyset$$

Os conjuntos definíveis em A , além dos próprios conjuntos elementares são:

$$D_5 = \{x_2, x_4, x_5, x_6\}, D_6 = \{x_1, x_3, x_4\}, D_7 = \{x_1, x_2, x_3, x_5, x_6\}, D_8 = \{x_1, x_2, x_3, x_4, x_5, x_6\} = U$$

6.2.3 – Classificação Aproximada de Um Conjunto

Seja $A = (U, R)$ um espaço aproximado e dado um subconjunto arbitrário de objetos de U , ou $X \subseteq U$, a fim de classificar o conjunto X com base nas descrições dos conjuntos elementares de A , e medir se o conjunto de descrições da classe de equivalência pode refletir as funções de pertinência de objetos em X , obtém-se os seguintes conceitos:

- *Aproximação Inferior* de X em A , ou $A_{inf}(X)$, é a união dos conjuntos elementares que estão contidos em X , ou seja, é o maior conjunto definível em A , contido em X , como mostra a figura 6.3. Denota-se a aproximação inferior por:

$$A_{inf}(X) = \{x \in U \mid [x]_R \subseteq X\}$$

Figura 6.3 - Aproximação Inferior de X em $A = (U, R)$

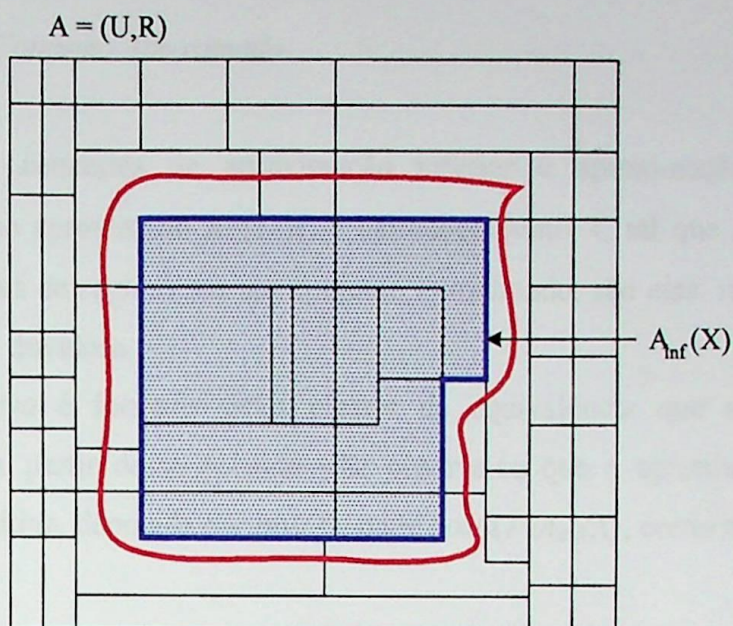


Figura 6.3 - Aproximação Inferior de X em $A=(U,R)$

- *Aproximação Superior* de X em A , ou $A_{sup}(X)$, é a união dos conjuntos elementares que possuem interseção não vazia em X , ou seja, é o menor conjunto definível em A contendo X , como mostra a figura 6.4. Denota-se a aproximação superior por:

$$A_{sup}(X) = \{x \in U \mid [x]_R \cap X \neq \emptyset\}$$

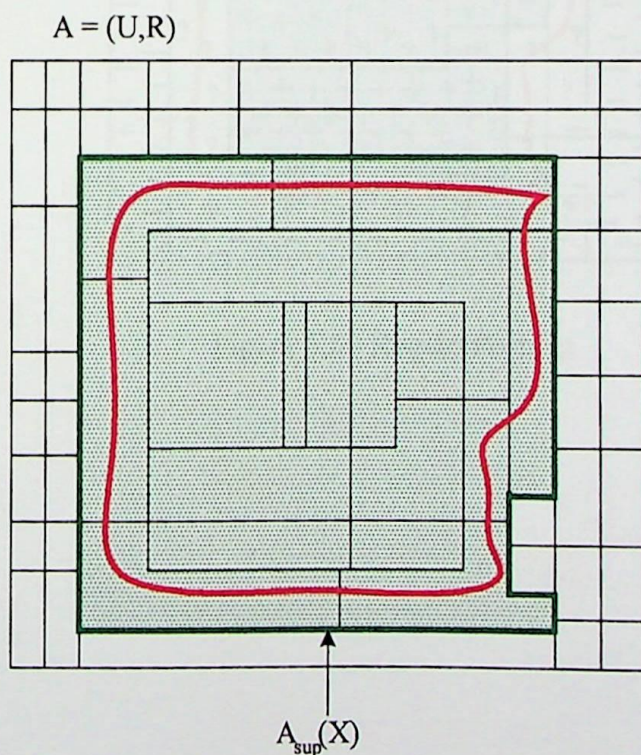


Figura 6.4 - Aproximação Superior de X em $A=(U,R)$

6.2.4 – Regiões em um Conjunto Aproximado

Com base nos conceitos de aproximação inferior e aproximação superior, e considerando um espaço aproximado $A=(U,R)$ e um subconjunto X , tal que $X \subseteq U$, pode-se identificar três categorias de regiões em um conjunto aproximado, são elas: região positiva, região negativa e região duvidosa.

A *região positiva* é formada pelas classes de equivalência que estão contidas integralmente em X . A partir dessa consideração observa-se que a aproximação inferior caracteriza a região positiva, denotada por $pos(X)$, onde $pos(X)=A_{inf}(X)$, como mostra a figura 6.5.

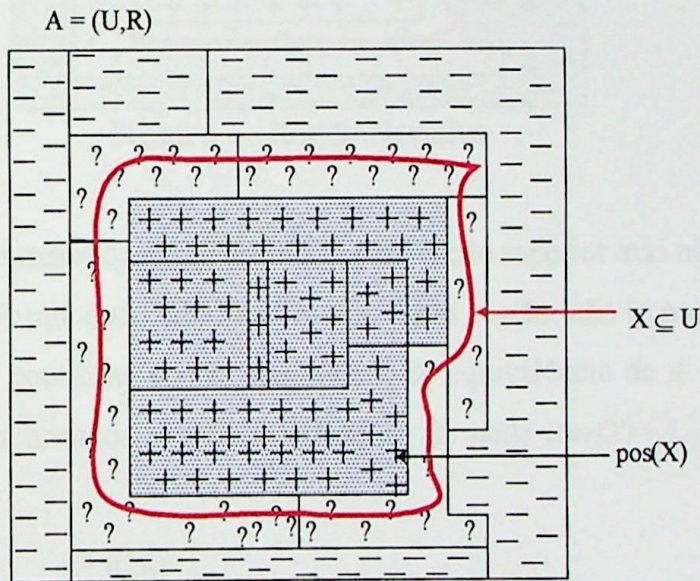


Figura 6.5 - Região Positiva

Dado um conjunto elementar E de A , se E não estiver contido na aproximação superior de X , então se conclui que nenhum de seus elementos pertence a X . Essa região de X é denotada por $neg(X)$, isto é, *região negativa* de X , onde $neg(X) = U - A_{sup}(X)$, ilustrada pela figura 6.6.

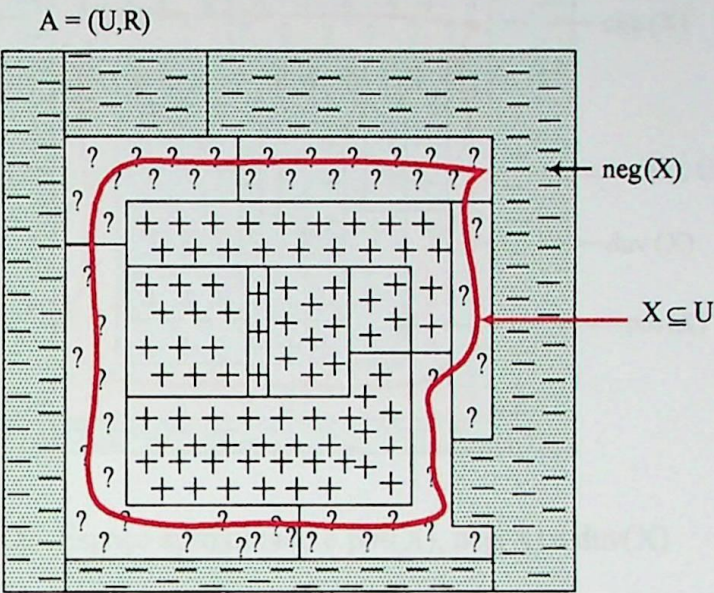


Figura 6.6 - Região Negativa

Existem ainda os elementos que pertencem à aproximação superior mas não pertencem à aproximação inferior, de forma que para um elemento dessa região, não se pode afirmar se ele pertence ao conjunto X com base apenas nas classes de equivalência de A . Essa região, denotada por $div(X)$, é chamada de *região duvidosa* de X , onde $div(X) = A_{sup}(X) - A_{inf}(X)$, ilustrada pela figura 6.7.

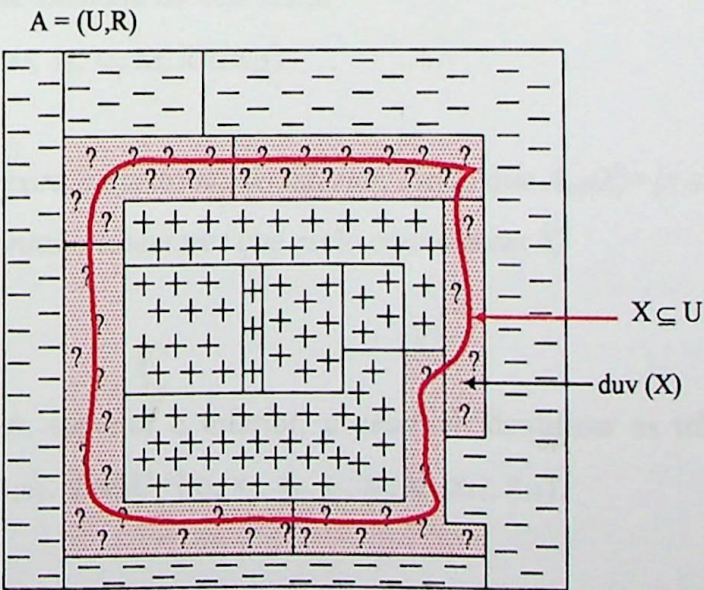


Figura 6.7 - Região Duvidosa

A figura 6.8 ilustra um espaço aproximado e suas regiões: positiva, negativa e duvidosa.

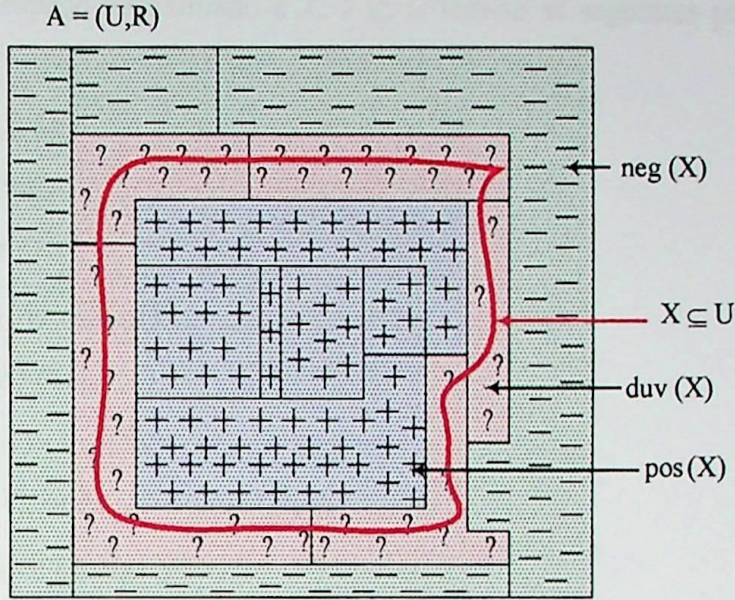


Figura 6.8 - Espaço aproximado e $pos(X)$, $neg(X)$ e $duv(X)$

Exemplo 6.3: Seja o universo $U = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}\}$, os conjuntos elementares $U/R = \{\{x_1, x_2, x_3\}, \{x_4\}, \{x_5, x_6, x_7\}, \{x_8, x_9, x_{10}\}, \{x_{11}, x_{12}\}, \{x_{13}, x_{14}, x_{15}\}\}$ e $X = \{x_2, x_3, x_4, x_5, x_{12}\}$. Então tem-se:

- A aproximação superior, onde $A_{sup}(X) = \{x \in U \mid [x]_R \cap X \neq \emptyset\}$, será composta pela união de todos os conjuntos elementares que possuem alguma interseção não vazia em X , como mencionado no item 6.2.3.

$$A_{sup}(X) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_{11}, x_{12}\}$$

- O mesmo acontece para a aproximação inferior, sendo que $A_{inf}(X) = \{x \in U \mid [x]_R \subseteq X\}$ (união dos conjuntos elementares que estão contidos em X)

$$A_{inf}(X) = \{x_4\}$$

Tendo as aproximações, superior e inferior, é possível identificar as três regiões:
 $pos(X) = \{x_4\}$; $neg(X) = \{x_8, x_9, x_{10}\}$; $duv(X) = \{x_1, x_2, x_3, x_5, x_6, x_7, x_{11}, x_{12}\}$.

6.2.5 – Propriedades das Aproximações em um Espaço Aproximado

Seja $A=(U, R)$ um espaço aproximado e $X, Y \subseteq U$ tem-se as seguintes propriedades para A .

- $A_{inf}(X) \subseteq X \subseteq A_{sup}(X)$
- $A_{inf}(\emptyset)=A_{sup}(\emptyset)=\emptyset$
- $A_{inf}(U)=A_{sup}(U)=U$
- $A_{inf}(X \cup Y) \supseteq A_{inf}(X) \cup A_{inf}(Y)$
- $A_{inf}(X \cap Y)=A_{inf}(X) \cap A_{inf}(Y)$
- $A_{inf}(X-Y) \subseteq A_{inf}(X) - A_{inf}(Y)$
- $A_{inf}(X)= - A_{sup}(-X)$, onde $-X$ é o complementar de X
- Se $X \subseteq Y$ então $A_{inf}(X) \subseteq A_{inf}(Y)$
- $A_{sup}(X \cup Y)=A_{sup}(X) \cup A_{sup}(Y)$
- $A_{sup}(X \cap Y) \subseteq A_{sup}(X) \cap A_{sup}(Y)$
- $A_{sup}(X-Y) \supseteq A_{sup}(X) - A_{sup}(Y)$
- $A_{sup}(X)= - A_{inf}(-X)$
- Se $X \subseteq Y$ então $A_{sup}(X) \subseteq A_{sup}(Y)$

Como $X \cup (-X)=U$ e $X \cap (-X)=\emptyset$, então tem-se:

- $A_{sup}(X) \cup A_{inf}(-X)=A_{sup}(-X) \cup A_{inf}(X)=A_{sup}(X) \cup A_{sup}(-X)=U$
- $A_{inf}(X) \cap A_{inf}(-X)=A_{sup}(X) \cap A_{inf}(-X)=A_{inf}(X) \cap A_{sup}(-X)=\emptyset$
- $A_{sup}(X) \cap A_{sup}(-X)=div(X)=div(-X)$

6.2.6 – Igualdade Aproximada de Conjuntos

Considere o espaço aproximado $A=(U, R)$ e $X \subseteq U$. É possível definir claramente as fronteiras do conjunto X , isso é realizado a partir das descrições dos conjuntos elementares de A , o que leva ao conceito de *Conjuntos Aproximados* ou “*Rough Sets*”.

Assim, um conjunto aproximado em A , ou conjunto aproximado se A é conhecido, é a família de todos os subconjuntos de U , os quais possuem as mesmas aproximação inferior e superior em A . A figura 6.9 ilustra um exemplo de conjuntos com as mesmas aproximações.

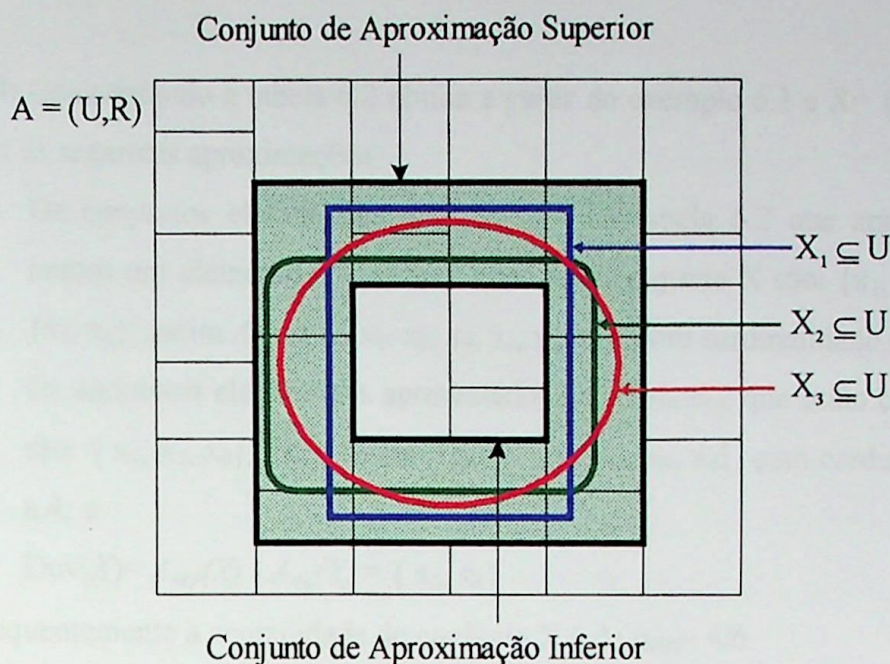


Figura 6.9 - Conjuntos X_1 , X_2 , e X_3 , com a mesma aproximação inferior e superior

Assim o conceito de igualdade aproximada abrange as seguintes considerações:

- X e Y são aproximadamente inf-iguais se $A_{inf}(X) = A_{inf}(Y)$;
- X e Y são aproximadamente sup-iguais se $A_{sup}(X) = A_{sup}(Y)$; e
- X e Y são aproximadamente iguais se $A_{inf}(X) = A_{inf}(Y)$ e $A_{sup}(X) = A_{sup}(Y)$.

A partir da figura 6.9 pode-se perceber que todos os conjuntos (X_1 , X_2 e X_3) são aproximadamente iguais, assim no espaço ilustrado pela figura qualquer um dos conjuntos representam o conjunto original, sendo chamados de conjuntos aproximados.

6.2.7 – Precisão de uma Aproximação

A medida de precisão, ou acuracidade, de uma aproximação é dada pela equação 6.1, que representa a relação entre a cardinalidade da aproximação inferior de X pela cardinalidade da aproximação superior de X . Assim, a cardinalidade de um conjunto é o número de objetos contidos na aproximação.

$$\mu_B = \frac{\text{card}(A_{inf}(X))}{\text{card}(A_{sup}(X))} = \frac{|A_{inf}(X)|}{|A_{sup}(X)|} \quad (6.1)$$

Onde $0 \leq \mu_B(X) \leq 1$, de forma que se $\mu_B(X) = 1$ então X é definível em U , e se $\mu_B(X) < 1$ então X é indefinível em U .

Exemplo 6.4: Considerando a tabela 6.2 obtida a partir do exemplo 6.1 e $X = \{x_1, x_3, x_4, x_5, x_9\}$, se obtém as seguintes aproximações:

- Os conjuntos elementares apresentados na tabela 6.2 que apresentam pelo menos um elemento em comum com o subconjunto X são: $\{x_1, x_3, x_9\}$, $\{x_4\}$, $\{x_5, x_8\}$, assim $A_{sup}(X) = \{x_1, x_3, x_9, x_4, x_5, x_8\}$, com cardinalidade igual a 6;
- Os conjuntos elementares apresentados na tabela 6.2 que estão contidos em X são: $\{x_1, x_3, x_9\}$, $\{x_4\}$, assim $A_{inf}(X) = \{x_1, x_3, x_9, x_4\}$, com cardinalidade igual a 4; e
- $Duv(X) = A_{sup}(X) - A_{inf}(X) = \{x_5, x_8\}$

Consequentemente a acuracidade do conjunto X é de $\mu_{(B)} = 4/6$.

6.2.8 – Independência de Atributos

A independência de um atributo é verificada a partir de sua eliminação da tabela (que representa o conhecimento) e se essa remoção altera o número de conjuntos elementares do sistema de representação de conhecimento.

Exemplo 6.5: Considerando e a tabela 6.1 verifica-se que o agrupamento dos atributos a_1, a_2 , e a_3 , gera a tabela 6.2. A partir da tabela 6.2 é possível verificar a independência dos atributos, para isso será necessário observar a mudança da quantidade de conjuntos elementares quando na eliminação de cada atributo um a um. Essa observação é dada pela tabela 6.4 [32].

Tabela 6.4 - Verificação da Independência de Atributos

	Atributo Eliminado			
	Nenhum	a_1	a_2	a_3
Número de Conjuntos Elementares	5	5	4	4

A partir da tabela 6.4 percebe-se que quando se elimina, ou remove, o atributo a_1 da tabela 6.2 o número de conjuntos elementares permanece o mesmo que quando não se elimina nenhum atributo. Entretanto, tanto na eliminação do atributo a_2 , como na eliminação do atributo a_3 , existe uma alteração da quantidade de conjuntos elementares. Assim, pode-se afirmar que o atributo a_1 é *supérfluo*, enquanto os atributos a_2 e a_3 são *indispensáveis*.

6.3 – Sistemas de Representação de Conhecimento (SRC)

Um Sistema de Representação de Conhecimento (SRC) ou Sistema de Informação é definido como uma 4-upla $S=(U, Q, V, \rho)$, onde [29]:

- U : universo finito de S , composto por elementos de U , denominados *objetos*. Os objetos são caracterizados por um conjunto de *atributos* e seus respectivos *valores*;
- Q : conjunto finito de *atributos* que caracterizam os elementos de U ;
- V : é o conjunto de valores de atributos, ou domínio (como mencionado no exemplo 6.1), dado por $V=\bigcup_{q \in Q} V_q$ sendo V_q o conjunto de valores do atributo q .
- ρ : dado por $\rho: U \times Q$, é uma função de descrição tal que $\rho(x, q) \in V_q$ para $x \in U$ e $q \in Q$. É a descrição do domínio.

Assim, para um Sistema de Representação de Conhecimento pode-se identificar as seguintes observações:

1. O número de descrições possíveis em S está relacionado ao número de descrições diferentes em V_q para cada atributo, assim quanto maior o número de descrições diferentes em V_q maior o número de descrições possíveis em S ;
2. Os objetos, ou exemplos, são caracterizados por vários atributos, dessa forma a função ρ_x será definida como uma seqüência de valores de atributos;
3. Tendo a relação de equivalência, ou relação de indiscernibilidade, no universo U definida por $P \subseteq Q$ (denotada por $\tilde{P}$), dois objetos $x, y \in U$ são indiscerníveis com relação a $P \subseteq Q$ se e somente se todo atributo $p \in P$. Se $P=Q$, então x e y são indiscerníveis em S , onde $\tilde{P}$ é uma relação de equivalência em S , denotado por $\tilde{S}$, para qualquer subconjunto de atributos $P \subseteq Q$. Assim, $S=(U, Q, V, \rho)$ caracteriza um único espaço aproximado $A=(U, \tilde{S})$.

Um Sistema de Representação de Conhecimento pode ser representado por uma tabela de dados. As colunas da tabela representarão os atributos, enquanto que as linhas da tabela irão representar a descrição dos objetos, ou exemplos. É preciso ressaltar que a classe dos objetos sempre será representada pela última coluna.

A tabela de dados pode ser obtida como resultado de medidas, observações ou representação de conhecimento de um agente ou um grupo de agentes [29].

É fácil entender como cada atributo pode estar associado a uma relação de equivalência. Considerando o atributo “cor”, é possível classificar todos os objetos de um universo em categorias para os que possuem a mesma cor.

Exemplo 6.6: A tabela 6.5 mostra um exemplo de SRC [33, 34] que contém um Sistema de Representação de Conhecimento sobre carros disponíveis em uma agência, onde a representa o preço do carro, b representa a Quilometragem, c representa o tamanho do carro, e d a velocidade máxima.

Tabela 6.5 - Exemplo de Sistema de Representação de Conhecimento

Carro	a	b	c	d
x_1	Alto	Baixa	Grande	Baixa
x_2	Baixo	Baixa	Grande	Baixa
x_3	Alto	Alta	Compacto	Baixa
x_4	Alto	Baixa	Grande	Baixa
x_5	Alto	Baixa	Grande	Alta
x_6	Baixo	Alta	Grande	Alta

A partir do SRC da tabela 6.5 obtém-se:

- $U = \{ x_1, x_2, x_3, x_4, x_5, x_6 \};$
- $A = \{ a, b, c, d \};$

E suponha que :

- $V_a = \{ \text{alto, baixo} \};$
- $V_b = \{ \text{alta, baixa} \};$
- $V_c = \{ \text{grande, compacto} \};$
- $V_d = \{ \text{alta, média, baixa} \};$

Entre as descrições fornecidas pelo sistema tem-se:

- $\rho(x_6, a) = \text{baixo};$
- $\rho(x_1, c) = \text{grande};$
- $\rho(x_4, a) = \text{alto};$
- $\rho(x_2, a) = \text{baixo};$

- $\rho(x_5, a) = \text{alto}$.

É possível observar que existem elementos distintos com a mesma descrição, como é o caso de x_4 e x_5 , assim o sistema é dito *não-seletivo*, pois não se consegue distinguir todos os elementos do sistema com base em suas descrições. Se fosse possível distinguir todos os elementos o sistema seria *seletivo*. A partir da tabela 6.5 é possível observar ainda que não existe nenhum objeto $x \in U$ tal que $\rho_x(d) = \text{média}$, dessa forma o sistema é chamado de *não-maximal*, pois existem descrições previstas pelo sistema que não são usadas pelos objetos. O sistema seria chamado de *maximal* se todas as descrições previstas fossem utilizadas.

6.4 – Tabelas de Decisão

Um Sistema de Representação de Conhecimento formado por atributos de condição (ou classes de condições) e atributos de decisão (ou classes de decisão) é chamado de *tabela de decisão*, e pode ser considerado um tipo especial de SRC.

Assim os atributos Q , onde $q \in Q$, mencionados anteriormente são divididos em *condições* e *decisões*. Dessa forma $Q = C \cup D$, onde C representa o subconjunto de Q que são condições e D o subconjunto de Q que são decisões.

As tabelas de decisão têm sido usadas para gerar regras a fim de facilitar a tomada de decisão, com isso é possível representar decisões em um dado sistema para o qual ainda não há regras explícitas na tomada de decisão [35].

Exemplo 6.7: Considere $S = (U, Q, V, \rho)$ como o SRC do exemplo 6.6. Para o sistema em questão é adicionado um novo atributo, obtido por intermédio de um avaliador da agência de carros. Assim, o atributo e irá representar uma avaliação referente às condições dos carros, de forma que S passa a ser uma tabela de decisão, com o atributo e representando uma decisão $D = \{e\}$, onde $V_e = \{\text{ruim, bom, excelente}\}$, como mostra a tabela 6.6.

Tabela 6.6 - Exemplo de Tabela de Decisão, onde $C=\{a, b, c, d\}$ representa a classe de condições e $D=\{e\}$ representa a classe de decisão

Carro	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>
x_1	Alto	Baixa	Grande	Baixa	Bom
x_2	Baixo	Baixa	Grande	Baixa	Bom
x_3	Alto	Alta	Compacto	Baixa	Ruim
x_4	Alto	Baixa	Grande	Baixa	Bom
x_5	Alto	Baixa	Grande	Alta	Excelente
x_6	Baixo	Alta	Grande	Alta	Bom

6.5 – Dependência entre atributos abordada como relação entre condição e decisão

Como mencionado no item 6.2.8, é possível identificar atributos indispensáveis e atributos supérfluos.

Assim a Teoria de Conjuntos Aproximados é utilizada para a analisar a dependência entre atributos de um sistema de representação de conhecimento S , a fim de identificar e eliminar os atributos redundantes. Isso permite a minimização do conjunto de atributos, onde é possível verificar se a partir de um conjunto de descrições de objetos, pode-se classificar esses objetos com um mínimo de atributos, preservando o poder discriminatório [36].

Análise de dependência é usada para determinar se um grupo de atributos (condições) pode caracterizar os valores de um outro atributo (decisão). A dependência entre as condições e decisão indica o quanto a decisão pode ser baseada nos valores dos atributos das condições (...) Em dados onde o atributo de decisão pode ser unicamente caracterizado em termos de condições, o grau de dependência é igual a um e a dependência é considerada totalmente funcional. O grau de dependência próximo de um, por exemplo, que caracteriza uma situação onde a maioria das decisões pode ser determinada pelas condições, de maneira não ambígua, estabelece as bases para hipóteses que evidenciam relações significativas entre condições e decisão. A dependência próxima de zero sugere uma relação fraca ou mesmo não existente. ¹

¹ Transcrição da tradução incluída em [35].

A dependência entre atributos, relacionados à condição e decisão, é tratada considerando um conjunto de atributos de condições e um conjunto de atributos de decisão [37], onde se identifica:

- conjuntos elementares $C'=\{C'_1, C'_2, C'_3,..., C'_n\}$ da relação $\tilde{C}$
- conjuntos elementares $D'=\{D'_1, D'_2, D'_3,..., D'_n\}$ da relação $\tilde{D}$

Nesse caso será necessário identificar a região positiva de D' , no espaço aproximado. Essa região é obtida pela união das aproximações inferiores de todos os conjuntos elementares (C') induzidos pela relação D' , de forma que se considere a região de U como única da classe D . Isso é definido pela equação 6.2, e visualizado pela figura 6.10.

$$pos(C, D) = \bigcup_{i=1}^n \{A_{C_{inf}}(D'_i) | D'_i \in D'\} \quad (6.2)$$

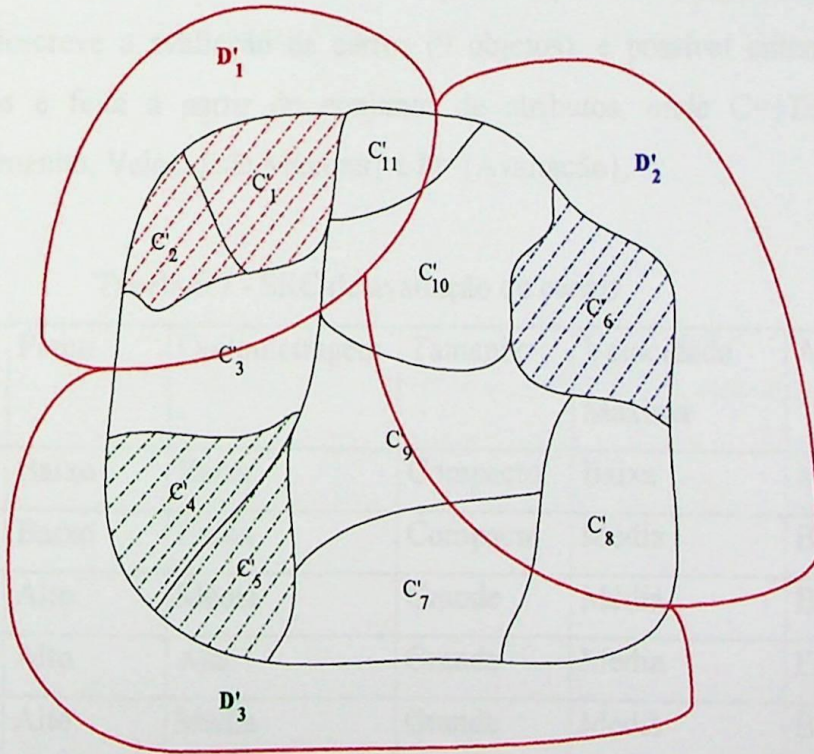


Figura 6.10 - Região positiva, ou conjuntos induzidos por atributos de condição e decisão.

A partir da figura 6.10 é possível perceber que no espaço aproximado, onde $D'=\{D'_1, D'_2, D'_3\}$ e $C'=\{C'_1, C'_2, C'_3, C'_4, C'_5, C'_6, C'_7, C'_8, C'_9, C'_{10}, C'_{11}\}$, existem conjuntos

elementares que são somente (únicos) de determinadas classes, são eles: $C'_1, C'_2, C'_4, C'_5, C'_6$. Assim a região positiva é determinada por:

$$\begin{aligned} \text{pos}(C,D) &= A_{C_{\text{inf}}}(D'_1) \cup A_{C_{\text{inf}}}(D'_2) \cup A_{C_{\text{inf}}}(D'_3) = (C'_1, C'_2) \cup (C'_4) \cup (C'_5, C'_6) = \\ &= C'_1 \cup C'_2 \cup C'_4 \cup C'_5 \cup C'_6 \end{aligned}$$

Baseado na informação da região positiva é possível obter o *grau de dependência* de um conjunto de atributos de decisão com relação ao conjunto de atributos de condição. O cálculo do grau de dependência (gdep) é dado pela equação 6.3, onde seus valores se enquadram no intervalo $[0, 1]$. Assim, se $\text{gdep}(C,D)=1$ então a dependência é funcional ou completa, se $\text{gdep}(C,D)=0$ então os conjuntos de atributos C e D são independentes e por fim se $0<\text{gdep}(C,D)<1$ então a dependência é parcial [37].

$$\text{gdep}(C,D) = \frac{|\text{pos}(C,D)|}{|U|} \tag{6.3}$$

Exemplo 6.8: Considerando o novo Sistema de Representação de Conhecimento da tabela 6.7, que também descreve a avaliação de carros (9 objetos), é possível entender como a avaliação de carros é feita a partir do conjunto de atributos, onde $C=\{\text{Turbo, Preço, Quilometragem, Tamanho, Velocidade Máxima}\}$ e $D=\{\text{Avaliação}\}$.

Tabela 6.7 - SRC de avaliação de carros

	U	Turbo	Preço	Quilometragem	Tamanho	Velocidade Máxima	Avaliação
C'_1	e_1	Não	Baixo	Baixa	Compacto	Baixa	Média
C'_2	e_2	Não	Baixo	Baixa	Compacto	Média	Bom
C'_3	e_3	Sim	Alto	Média	Grande	Média	Bom
C'_4	e_4	Sim	Alto	Alta	Grande	Média	Excelente
C'_5	e_5	Sim	Alto	Média	Grande	Média	Bom
C'_6	e_6	Não	Baixo	Baixa	Compacto	Média	Bom
C'_7	e_7	Sim	Alto	Média	Grande	Alta	Excelente
C'_8	e_8	Sim	Alto	Alta	Grande	Alta	Excelente
C'_9	e_9	Não	Baixo	Baixa	Compacto	Alta	Bom

No espaço aproximado é possível identificar os seguintes conjuntos:

$$C' = U / \tilde{C} = \{C'_1, C'_2, C'_3, C'_4, C'_5, C'_6, C'_7, C'_8, C'_9\} = \{\{e_1\}, \{e_2\}, \{e_3\}, \{e_4\}, \{e_5\}, \{e_6\}, \{e_7\}, \{e_8\}, \{e_9\}\};$$

$D' = U / \tilde{D} = \{D'_1, D'_2, D'_3\} = \{\{e_1\}, \{e_2, e_3, e_5, e_6, e_9\}, \{e_4, e_7, e_8\}\}$, onde os conjuntos de elementos estão associados a avaliação média, bom e excelente, respectivamente; e

$$\begin{aligned} \text{pos}(C, D) &= A_{\text{Cinf}}(D'_1) \cup A_{\text{Cinf}}(D'_2) \cup A_{\text{Cinf}}(D'_3) \\ &= (C'_1) \cup (C'_2 \cup C'_3 \cup C'_5 \cup C'_6 \cup C'_9) \cup (C'_4 \cup C'_7 \cup C'_8) = \\ &= C'_1 \cup C'_2 \cup C'_3 \cup C'_5 \cup C'_6 \cup C'_9 \cup C'_4 \cup C'_7 \cup C'_8 = U \end{aligned}$$

A partir da obtenção dos conjuntos das condições, decisões e região positiva, é possível calcular o grau de dependência. Para o exemplo 6.8 tem-se o seguinte grau de dependência:

$$\text{gdep}(C, D) = \frac{9}{9} = 1$$

O grau de dependência igual a 1 caracteriza a dependência completa, assim pode-se concluir que o atributo avaliação é afetado pela interação de todos, ou de alguns dos atributos contidos em C. Como apenas alguns atributos também podem afetar a avaliação, então é necessário verificar quais combinações de atributos afetam a avaliação e quais não afetam, ou são irrelevantes.

Exemplo 6.9: A partir da tabela 6.7 é possível analisar apenas alguns atributos de condição em relação ao atributo de decisão (avaliação). Considerando $C = \{\text{Preço}, \text{Velocidade Máxima}\}$ e $D = \{\text{Avaliação}\}$ do espaço aproximado, obtém-se a tabela 6.8

Tabela 6.8 - Análise de dependência dos atributos de condição em relação a decisão

U	Preço	Velocidade Máxima	Avaliação
e ₁	Baixo	Baixa	Média
e ₂	Baixo	Média	Bom
e ₃	Alto	Média	Bom
e ₄	Alto	Média	Excelente
e ₅	Alto	Média	Bom
e ₆	Baixo	Média	Bom
e ₇	Alto	Alta	Excelente
e ₈	Alto	Alta	Excelente
e ₉	Baixo	Alta	Bom

Para a análise em questão é preciso primeiramente isolar o atributo $D=\{\text{avaliação}\}$ para a verificação de seus conjuntos elementares. A tabela 6.9 apresenta os conjuntos elementares mencionados.

Tabela 6.9 - Conjuntos elementares de D'

	U	Avaliação
D' ₁	e ₁	Média
D' ₂	e ₂ , e ₃ , e ₅ , e ₆ , e ₉	Bom
D' ₃	e ₄ , e ₇ , e ₈	Excelente

Assim os conjuntos elementares são: $D'_1=\{e_1\}$, $D'_2=\{e_2, e_3, e_5, e_6, e_9\}$, $D'_3=\{e_4, e_7, e_8\}$. Esse resultado é semelhante ao resultado obtido pela análise de todos os atributos de condição com relação ao atributo de decisão, isso se dá devido não existir objetos repetidos quando se consideram todos os atributos. A figura 6.11 ilustra o espaço aproximado $A=(U, \tilde{C})$ formado a partir do SRC fornecido pela tabela 6.7, considerando-se $C=\{\text{Turbo, Preço, Quilometragem, Tamanho, Velocidade Máxima}\}$

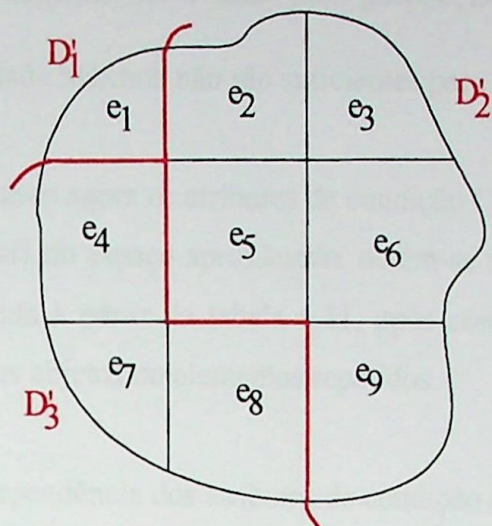


Figura 6.11 - Espaço aproximado $A=(U, \tilde{C})$ a partir do SRC fornecido pela tabela 6.7

A análise para $C=\{\text{Preço, Velocidade Máxima}\}$ e $D=\{\text{Avaliação}\}$ é realizada a partir da tabela 6.8. Considerando apenas os atributos $C=\{\text{Preço, Velocidade Máxima}\}$ é possível obter a tabela 6.10 após agrupar os objetos ou elementos repetidos.

Tabela 6.10 - SRC para atributos de condição Preço e Velocidade Máxima

	U	Preço	Velocidade Máxima
C'_1	e_1	Baixo	Baixa
C'_2	e_2, e_6	Baixo	Média
C'_3	e_3, e_4, e_5	Alto	Média
C'_4	e_7, e_8	Alto	Alta
C'_5	e_9	Baixo	Alta

$$C'=U/\tilde{C}=\{C'_1, C'_2, C'_3, C'_4, C'_5\}=\{\{e_1\}, \{e_2, e_6\}, \{e_3, e_4, e_5\}, \{e_7, e_8\}, \{e_9\}\};$$

$$D'=U/\tilde{D}=\{D'_1, D'_2, D'_3\}=\{\{e_1\}, \{e_2, e_3, e_5, e_6, e_9\}, \{e_4, e_7, e_8\}\}$$

$$\begin{aligned} \text{pos}(C,D) &= A_{\text{Cinf}}(D'_1) \cup A_{\text{Cinf}}(D'_2) \cup A_{\text{Cinf}}(D'_3) \\ &= (C'_1) \cup (C'_2 \cup C'_3) \cup (C'_4 \cup C'_5) = \\ &= \{e_1\} \cup \{e_2, e_6\} \cup \emptyset \cup \{e_7, e_8\} \cup \{e_9\} = \\ &= \{e_1, e_2, e_6, e_7, e_8, e_9\} \end{aligned}$$

Onde o grau de dependência é dado por: $gdep(C,D)=\frac{6}{9}=0,67$, de forma que os atributos Preço e Velocidade Máxima não são suficientes para caracterizar a Avaliação.

Exemplo 6.10: Considerando agora os atributos de condição $C=\{\text{Quilometragem, Velocidade Máxima}\}$ e $D=\{\text{Avaliação}\}$ do espaço aproximado, obtém-se a tabela 6.11, a partir da tabela 6.7. A tabela 6.12 é obtida a partir da tabela 6.11, após considerar apenas os atributos de condição (C) agrupando os objetos ou elementos repetidos.

Tabela 6.11-Análise de dependência dos atributos de condição em relação à decisão avaliação

U	Quilometragem	Velocidade Máxima	Avaliação
e ₁	Baixa	Baixa	Média
e ₂	Baixa	Média	Bom
e ₃	Média	Média	Bom
e ₄	Alta	Média	Excelente
e ₅	Média	Média	Bom
e ₆	Baixa	Média	Bom
e ₇	Média	Alta	Excelente
e ₈	Alta	Alta	Excelente
e ₉	Baixa	Alta	Bom

Tabela 6.12 - SRC para atributos de condição Quilometragem e Velocidade Máxima

	U	Quilometragem	Velocidade Máxima
C' ₁	e ₁	Baixa	Baixa
C' ₂	e ₂ , e ₆	Baixa	Média
C' ₃	e ₃ , e ₅	Média	Média
C' ₄	e ₄	Alta	Média
C' ₅	e ₇	Média	Alta
C' ₆	e ₈	Alta	Alta
C' ₇	e ₉	Baixa	Alta



$$C' = U / \tilde{C} = \{C'_1, C'_2, C'_3, C'_4, C'_5, C'_6, C'_7\} = \\ = \{\{e_1\}, \{e_2, e_6\}, \{e_3, e_5\}, \{e_4\}, \{e_7\}, \{e_8\}, \{e_9\}\};$$

$$D' = U / \tilde{D} = \{D'_1, D'_2, D'_3\} = \{\{e_1\}, \{e_2, e_3, e_5, e_6, e_9\}, \{e_4, e_7, e_8\}\}$$

$$\text{pos}(C, D) = A_{\text{Cinf}}(D'_1) \cup A_{\text{Cinf}}(D'_2) \cup A_{\text{Cinf}}(D'_3) \\ = (C'_1) \cup (C'_2 \cup C'_3 \cup C'_4) \cup (C'_5 \cup C'_6 \cup C'_7) = \\ = \{\{e_1\} \cup \{e_2, e_6\} \cup \{e_3, e_5\} \cup \{e_4\} \cup \{e_7\}, \{e_8\} \cup \{e_9\}\} \\ = \{e_1, e_2, e_6, e_3, e_5, e_4, e_7, e_8, e_9\}$$

Onde o grau de dependência é dado por: $\text{gdep}(C, D) = \frac{9}{9} = 1$, assim Quilometragem e

Velocidade Máxima definem a Avaliação. Dessa forma pode-se dizer que esse conjunto de atributos, ou conjunto mínimo de atributos, possui o mesmo poder discriminatório que o conjunto original. Nesse caso, o conjunto mínimo é também conhecido como reduto, sendo que é possível encontrar vários redutos para um mesmo sistema.

A figura 6.12 representa o espaço aproximado obtido pelo SRC da tabela 6.12, que considera os atributos de condição $C = \{\text{Quilometragem, Velocidade Máxima}\}$, enquanto a figura 6.13 representa o SRC considerando os atributos de condição $C = \{\text{Quilometragem, Velocidade Máxima}\}$ e o atributo de decisão $D = \{\text{Avaliação}\}$.

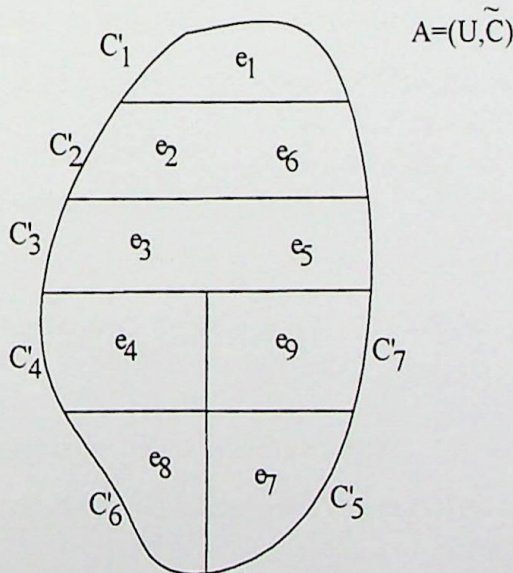


Figura 6.12 - Espaço aproximado considerando atributos de condição

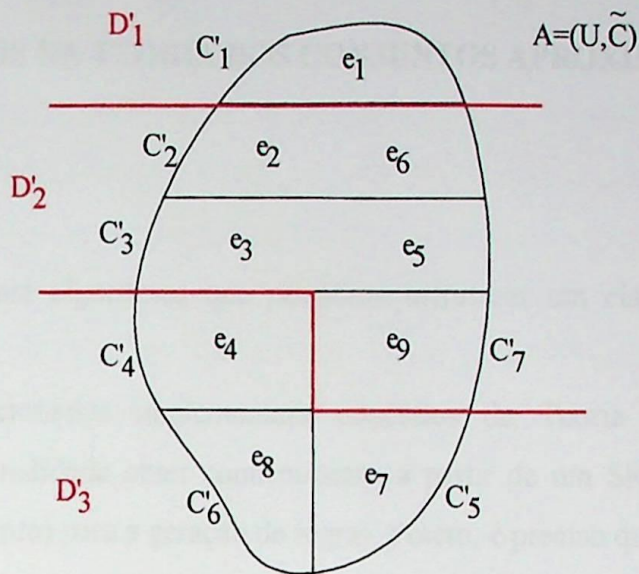


Figura 6.13 - Espaço aproximado considerando atributos de condição e decisão

ALGORITMOS BASEADOS NA TEORIA DOS CONJUNTOS APROXIMADOS

7.1 – Introdução

Esse capítulo apresenta algoritmos que permitem estruturar um classificador para sinais de voz.

Os algoritmos mencionados implementam conceitos da Teoria de Conjuntos Aproximados e têm como finalidade obter conhecimento a partir de um SRC (Sistema de Representação de Conhecimento) para a geração de regras. Porém, é preciso que o SRC esteja no formato de tabelas de decisão para que seja possível reduzir a base de conhecimento.

A redução, também considerada uma simplificação da tabela de decisão ou do conjunto de exemplos, permite agilizar a performance de alguns sistemas quando na tomada de decisão.

Os algoritmos apresentados aqui são capazes de fornecer a redução mencionada e podem ser representados pelos seguintes passos [38]:

Passos Prévio:

1. Transformação ou redefinição dos valores dos atributos em faixas;
2. Eliminação de atributos idênticos;
3. Eliminação de padrões ou exemplos idênticos (é importante lembrar que a repetição de padrões pode ou não surgir após a redefinição de valores);

Conjunto de algoritmos:

1. Algoritmo para o cálculo do grau de dependência;
2. Algoritmo para obtenção de redutos (capaz de buscar conjuntos de Redutos (redução), ou conjuntos com o mesmo poder discriminatório que o conjunto inicial);
3. Algoritmo para a composição do conjunto de regras.

Onde os algoritmos para o cálculo do grau de dependência e para obtenção de redutos permitem eliminar atributos dispensáveis.

7.2 – Classificação dos Dados em Faixas

A classificação de dados em faixas é responsável pela transformação ou redefinição dos valores dos atributos em faixas de valores, as quais são criadas a partir de uma certa métrica. Nesse momento inicia-se o processo de uso da Teoria dos Conjuntos Aproximados com a definição de faixas para os diversos atributos da tabela contendo o SRC.

A tabela 7.1 representa uma base de dados obtida de [38], que será considerada um SRC qualquer e usada para exemplificar os passos prévios, bem como a aplicação dos algoritmos.

O SRC em questão possui 9 atributos de condição (identificados por A, B, C, D, E, F, G, H e I) e um atributo de decisão (identificado por S). A coluna identificada por Nr. representa o conjunto universo U (que também pode ser vista como uma coluna de índices para os exemplos ou padrões).

Tabela 7.1 - Exemplo de Base de Dados ou SRC

Nr.	Atributos de Condição									S
	A	B	C	D	E	F	G	H	I	
1	67	27	62	51	71	75	17	51	32	1
2	17	57	68	71	67	67	28	17	28	3
3	32	63	61	52	17	18	71	21	27	3
4	22	27	72	72	63	61	29	50	42	2
5	55	13	63	52	18	27	70	21	17	3
6	47	27	72	51	35	51	27	24	27	3
7	72	22	64	67	63	63	20	51	52	2
8	58	22	73	57	18	18	70	17	25	3
9	28	52	67	71	67	63	21	47	52	1
10	28	45	65	67	63	67	25	18	25	3
11	22	17	63	63	63	71	71	35	51	2
12	21	22	68	61	65	73	17	42	51	2
13	42	23	65	57	55	58	18	72	62	2
14	25	51	71	63	62	62	21	18	18	3
15	68	17	72	51	63	62	27	50	57	1
16	72	27	61	72	67	64	28	45	57	2
17	70	51	72	52	22	28	25	62	67	1
18	18	18	78	72	71	63	17	32	52	2
19	31	61	75	50	28	27	71	28	17	3
20	42	58	77	51	18	25	67	27	27	3
21	27	53	67	63	71	72	17	32	57	1
22	63	13	73	51	67	67	25	57	52	1

Para o SRC usado como exemplo, e apresentado pela tabela 7.1, foram definidos três faixas de valores, as quais foram aplicadas para todos os atributos de condição. A métrica usada para a definição de tais faixas é mostrada abaixo:

- Quando o valor do atributo for inferior a 30, associa-se a ele o valor 1 (um);
- Quando o valor do atributo estiver entre a faixa de valores 30 e 60, associa-se a ele o valor 2 (dois);
- E quando o valor do atributo for superior a 60, associa-se a ele o valor 3 (três).

A tabela 7.2 representa a tabela 7.1 após a redefinição dos valores dos atributos.

Tabela 7.2 - SRC com redefinição dos valores dos atributos.

Nr.	Atributos de Condição									S
	A	B	C	D	E	F	G	H	I	
1	3	1	3	2	3	3	1	2	2	1
2	1	2	3	3	3	3	1	1	1	3
3	2	3	3	2	1	1	3	1	1	3
4	1	1	3	3	3	3	1	2	2	2
5	2	1	3	2	1	1	3	1	1	3
6	2	1	3	2	2	2	1	1	1	3
7	3	1	3	3	3	3	1	2	2	2
8	2	1	3	2	1	1	3	1	1	3
9	1	2	3	3	3	3	1	2	2	1
10	1	2	3	3	3	3	1	1	1	3
11	1	1	3	3	3	3	3	2	2	2
12	1	1	3	3	3	3	1	2	2	2
13	2	1	3	2	2	2	1	3	3	2
14	1	2	3	3	3	3	1	1	1	3
15	3	1	3	2	3	3	1	2	2	1
16	3	1	3	3	3	3	1	2	2	2
17	3	2	3	2	1	1	1	3	3	1
18	1	1	3	3	3	3	1	2	2	2
19	2	3	3	2	1	1	3	1	1	3
20	2	2	3	2	1	1	3	1	1	3
21	1	2	3	3	3	3	1	2	2	1
22	3	1	3	2	3	3	1	2	2	1

7.3 – Simplificação da Tabela de Decisão

A simplificação das tabelas de decisão é de grande importância pois a partir desta é possível, por exemplo, reduzir o número de atributos de condição, ou ainda eliminar exemplos duplicados, como também eliminar atributos supérfluos. A redução de atributos obtida pela

simplificação permite que uma decisão se baseie em um menor número de condições, de forma que se elimine a verificação de atributos desnecessários.

Considerando a eliminação de atributos de condição, é possível observar que a tabela 7.2 apresenta dois atributos que podem ser eliminados por serem repetidos, são eles: o atributo F que é semelhante ao atributo E, e o atributo I, sendo este semelhante ao atributo H.

Além disso, é possível eliminar vários exemplos devido à existência de redundância, como é o caso dos exemplos 5 e 8, os quais são idênticos; assim é possível eliminar um deles. A tabela 7.3 mostra o SRC após eliminar os atributos semelhantes e todos os exemplos repetidos, de forma que é possível perceber que houve uma redução do conjunto inicial de exemplos.

Tabela 7.3 - SRC após eliminar exemplos idênticos

Nr.	A	B	C	D	E	G	H	S
1	3	1	3	2	3	1	2	1
2	1	2	3	3	3	1	1	3
3	2	3	3	2	1	3	1	3
4	1	1	3	3	3	1	2	2
5	2	1	3	2	1	3	1	3
6	2	1	3	2	2	1	1	3
7	3	1	3	3	3	1	2	2
8	1	2	3	3	3	1	2	1
9	1	1	3	3	3	3	2	2
10	2	1	3	2	2	1	3	2
11	3	2	3	2	1	1	3	1
12	2	2	3	2	1	3	1	3

7.4 – Grau de Dependência

O grau de dependência é o valor usado para analisar a dependência de um atributo de decisão em relação a um conjunto de atributos de condição, a fim de determinar se os conjuntos de valores dos atributos de condição podem determinar o valor do atributo de decisão [39].

Nesse caso, é possível verificar se a tabela de decisão (ou SRC) possui somente atributos indispensáveis. Para isso é preciso eliminar os atributos um a um, a fim de observar se a eliminação de um desses atributos irá influenciar na classificação dos padrões, o que representa uma análise de consistência. Por exemplo, a partir da eliminação de um atributo é possível observar se a tabela continua com a mesma classificação gerada pela tabela inicial.

A partir da tabela 7.3, e eliminando os atributos um a um, tem-se as seguintes observações:

- Ao eliminar o atributo A verifica-se que os padrões indicados pelos índices 4 e 7 irão apresentar os mesmos valores para os atributos B, C, D, E, G e H, caracterizando padrões iguais e com os mesmos valores para o atributo de decisão S. Assim, o que se pode concluir é que o atributo A é dispensável, devido à classificação permanecer a mesma quando na sua eliminação;
- Ao eliminar o atributo B verifica-se que existem padrões com os mesmos valores para A, C, D, E, G e H (como é o caso dos exemplos 3, 5 e 12), e com a mesma saída para o atributo S. Entretanto, os padrões 4 e 8, também com os mesmos valores para os atributos de condição, possuirão valores diferentes para o atributo de decisão S, assim conclui-se que o atributo B é indispensável;
- Ao eliminar o atributo C verifica-se que não existirão padrões iguais, considerando o agrupamento dos atributos A, B, D, E, G e H, assim o atributo C é dispensável;
- Ao eliminar o atributo D verifica-se que os padrões indicados pelos índices 1 e 7 irão apresentar os mesmos valores para os atributos A, B, C, E, G e H, porém apresentam valores diferentes para o atributo de decisão S. Nesse caso o atributo D é indispensável;
- Ao eliminar o atributo E verifica-se que não existirão padrões iguais, considerando o agrupamento dos atributos A, B, C, D, G e H, assim o atributo E é dispensável;
- Ao eliminar o atributo G verifica-se que os padrões indicados pelos índices 4 e 9 irão apresentar os mesmos valores para os atributos A, B, C, D, E e H, caracterizando padrões iguais e com os mesmos valores para o atributo de decisão S. Assim, conclui-se que o atributo G também é dispensável;
- Ao eliminar o atributo H verifica-se que existem padrões com os mesmos valores para A, B, C, D, E e G, como é o caso dos exemplos 2 e 8, e 6 e 10, porém esses padrões possuem valores diferentes para o atributo de decisão S. Assim o atributo H é indispensável.

A tabela 7.4 apresenta o SRC contendo apenas os atributos indispensáveis, e sem a repetição de padrões.

Tabela 7.4 - SRC contendo apenas atributos indispensáveis

Nr.	B	D	H	S
1	1	2	2	1
2	2	3	1	3
3	3	2	1	3
4	1	3	2	2
5	1	2	1	3
6	2	3	2	1
7	1	2	3	2
8	2	2	3	1
9	2	2	1	3

O algoritmo para o cálculo do grau de dependência inclui os seguintes passos [40]:

1. Inicia uma variável (nomeada por exemplo como Grau_dep), com valor zero atribuído a esta variável, com o objetivo de armazenar o valor do Grau de Dependência;
2. Cria uma tabela A1, baseada no SRC, contendo os atributos de condição desejados (ou conjuntos de atributos, os quais serão analisados a fim de verificar se os mesmos podem determinar o valor do atributo de decisão). Esses atributos de condição devem ser agrupados, sem considerar o atributo decisão, a fim de analisar a quantidade de exemplos repetidos. Um aspecto importante é que para se determinar o grau de dependência será necessário incluir uma coluna na tabela A1 (contendo um contador, o qual pode ser nomeado de Cnt) para armazenar a informação referente ao número de repetições (ou quantidade) de um determinado exemplo;
3. Cria uma tabela B1 contendo os mesmos atributos de condição da tabela A1, mais o atributo de decisão. Para a tabela B1 também será necessário incluir uma coluna para armazenar a informação referente ao número de repetições de um determinado exemplo (Cnt). É importante ressaltar que nesse caso é necessário considerar a saída, ou o atributo decisão, que permite perceber que os valores de Cnt estão relacionados aos atributos de condição mais o atributo de decisão;
4. Apontar para a primeira linha da tabela A1, e fazer uma “varredura” na tabela B1, a fim de comparar as informações armazenadas. Nesse momento é possível

verificar os valores dos atributos agrupados (ou exemplos iguais das duas tabelas) bem como os valores das colunas Cnt, a fim de verificar se a quantidade de exemplos semelhantes é a mesma. Se os valores forem iguais então um determinado exemplo, tanto da tabela A1 quanto da tabela B1, possuem a mesma saída no atributo de decisão, caracterizando uma interseção entre as tabelas A1 e B1. Após verificar o primeiro exemplo da tabela A1, passa-se para o segundo exemplo, e assim sucessivamente até o ponteiro encontrar o fim da tabela A1. É importante lembrar que, deve-se criar uma variável (que pode ser nomeada de Cnt_interseção) que irá armazenar a somatória dos valores da coluna cnt referente aos exemplos que possuem interseção entre as tabelas A1 e B1;

5. Calcular o grau de dependência do conjunto de atributos. O cálculo é realizado a partir da equação 7.1, ou seja, pela divisão da variável Cnt_interseção (quantidade, ou somatória, de exemplos da interseção das tabelas A1 e B1) pela quantidade inicial de exemplos da tabela original (ou quantidade de dados do conjunto universo U).

$$\text{Grau_dep} = \frac{\text{Cnt_interseção}}{U} \quad (7.1)$$

7.5 – Obtenção de Redutos

Um reduto de R, sendo R uma família de relações equivalentes, é definido como um conjunto reduzido de atributos que conservam a mesma classificação indutiva da família R, ou conjunto R [41]. Em outras palavras, um reduto pode ser considerado como um conjunto reduzido de atributos de condição, o qual possui o mesmo poder discriminatório que o conjunto inicial.

Dessa forma, um reduto é a estrutura mínima de representação de conhecimento existente em uma base de dados. Entretanto, quando se identifica um primeiro reduto é importante lembrar que este pode não ser único, já que é possível encontrar vários redutos em uma base de conhecimento, isso dependendo dos atributos utilizados para a sua representação.

O algoritmo para a obtenção de redutos se baseia principalmente no cálculo do grau de dependência, e também na possibilidade de eliminação de atributos passo a passo, a fim de identificar atributos supérfluos. Assim, para um determinado conjunto de atributos será

aplicado o algoritmo para a obtenção de redutos de forma recursiva para que possam ser identificados subconjuntos mínimos do conjunto de atributos em questão.

O algoritmo para a obtenção de redutos é composto dos seguintes passos [40]:

1. Após a execução do algoritmo para o cálculo do grau de dependência, no SRC original, verifica-se a existência de um conjunto de atributos cujo Grau_dep é igual a 1 (um);
2. Identificado o conjunto de atributos de condição com Grau_dep=1, monta-se uma nova tabela, contendo as colunas referentes aos conjuntos de atributos de condição, mais o atributo de decisão de referência. A criação da nova tabela é realizada a partir do SRC original, o que irá gerar um novo SRC. É preciso ressaltar que os passos do algoritmo para a obtenção de redutos são realizados para todos os conjuntos de atributos com grau de dependência igual a 1, e que foram obtidos a partir do SRC original;
3. Após a obtenção da nova tabela, é necessário eliminar as linhas ou exemplos repetidos. Nesse momento, é executado o algoritmo para o cálculo do grau de dependência, devido tal algoritmo verificar o grau de dependência eliminando-se os atributos de condição um a um;
4. Esse procedimento é executado de forma recursiva até que sejam encontrados subconjuntos com o mesmo grau de dependência (igual a 1) do conjunto original. Caso o grau de dependência de um subconjunto seja igual a 1, este passa a ser o conjunto original, sendo então executado novamente o algoritmo para o cálculo do grau de dependência, e assim sucessivamente até que sejam verificados todos os subconjuntos possíveis;
5. O processo termina quando se encontra um conjunto mínimo que tenha grau de dependência igual a 1, ou seja, com o mesmo poder discriminatório que o conjunto original.

7.6 – Composição do Conjunto de Regras

A partir de uma tabela de decisão, ou um Sistema de Representação de Conhecimento, é possível extrair um conjunto de regras para a tomada de decisão. As regras podem ser exibidas na forma abaixo:

Se conjunto_de_antecedente **Então** consequente,

Podendo ser representado também da seguinte forma:

Se $(AC = VC)$ **Então** $(AD = VD)$, onde: $AC = \text{atributo_de_condição}$

$VC = \text{valor_atributo_de_condição}$

$AD = \text{atributo_de_decisão}$

$VD = \text{valor_atributo_de_decisão}$

Ou ainda possuir mais termos como antecedente:

Se $(AC_i = VC_{Aci}) \& \dots \& (AC_n = VC_{Acn})$ **Então** $AD = VD$, onde o símbolo $\&$ representa o “e” lógico.

O algoritmo para a composição do conjunto de regras contém os passos abaixo [40]:

1. Deve-se selecionar um reduto, ou conjunto mínimo de atributos de condição com o mesmo poder discriminatório que o conjunto original;
2. Montar uma nova tabela de decisão, contendo as colunas referentes aos conjuntos de atributos de condição que compõem o reduto, mais o atributo de decisão. A criação dessa tabela também é realizada a partir do SRC original. Nesse momento deve-se eliminar as linhas repetidas;
3. Para cada linha da nova tabela é necessário verificar a menor regra consistente, a partir da eliminação de cada atributo de condição um a um. Uma regra é considerada consistente quando não existir outra regra na tabela com o mesmo antecedente e consequente diferente, ou seja, duas condições com os mesmos valores (ou condições iguais) levando a decisões diferentes, o que caracteriza regras inconsistentes;
4. Após obter as regras consistentes, verificar se não existem regras repetidas. Caso existam regras duplicadas será necessário desconsiderar as repetições, considerando apenas uma das regras.

Assim, quando na existência de regras consistentes e regras inconsistentes, considera-se apenas as regras consistentes, eliminando-se as possíveis repetições de regras consistentes.

Dessa forma, as regras obtidas a partir do SRC representado pela tabela 7.1, são indicadas pela tabela 7.5. Lembrando que o SRC em questão foi utilizado para exemplificar e descrever os passos prévios e o conjunto de algoritmos.

Tabela 7.5 - Regras obtidas a partir do SRC representado pela tabela 7.1.

No	Atributos			S
	B	D	H	
1	-	2	2	1
2	2	-	3	1
3	2	-	2	1
4	1	3	-	2
5	1	-	3	2
6A	3	-	-	3
6B	-	-	1	3
7	-	-	1	3
8	-	-	1	3
9	-	-	1	3
10	-	-	1	3

APLICAÇÃO E RESULTADOS

8.1 – Introdução

Este capítulo apresenta um exemplo ilustrativo do classificador proposto, cuja finalidade é a classificação de sinais de voz, que pode ser utilizado na segunda etapa (mencionada no capítulo 5) de um Sistema de Reconhecimento de Locutor.

Para exemplificar a aplicação do classificador foram utilizados dados reais, obtidos na etapa de aquisição das características dos sinais de voz, onde tais dados foram usados para estruturar os Sistemas de Representação de Conhecimento, com o objetivo de viabilizar a inicialização da fase de treinamento.

A obtenção dos resultados é realizada após a fase de treinamento, mais especificamente na fase de testes, quando são aplicados novos padrões de dados em um conjunto de regras obtidas na fase de treinamento, e após realizar uma comparação e verificação da saída gerada pelo sistema com o valor referente ao atributo saída, o qual identifica o locutor do novo padrão.

8.2 – Aplicação

A aplicação do classificador nos sinais de voz, ou dados referentes a tais sinais, permite obter resultados para avaliação do mesmo. Os resultados estão relacionados à classificação de uma locução pronunciada por um locutor e conseqüentemente a sua identificação, considerando que esse locutor pertence a um conjunto de N saídas possíveis, ou N locutores de referência.

Como descrito no capítulo 5, cada locutor está associado a um código de identificação, que pode ser visualizado a partir da tabela 5.1. Assim, a diferenciação dos padrões que compõem os Sistemas de Representação de Conhecimento é realizada através do código mencionado.

Para que o sistema proposto seja capaz de identificar um locutor em um grupo de locutores, foi necessário obter padrões referentes aos locutores do grupo. A forma como se obteve padrões é descrita no capítulo 5. Assim, usou-se um determinado número de locuções para criar as tabelas que irão representar os SRCs das simulações, ou conjuntos de

treinamento. É importante lembrar que o número inicial de padrões, ou conjunto de dados, de um SRC está relacionado à quantidade de locutores incluídos no grupo.

8.3 – Treinamento de Conjuntos de Dados

A aplicação inicia com a fase de treinamento de conjunto de dados, que se baseia nos algoritmos apresentados no capítulo 7. O objetivo dessa fase é a obtenção de um conjunto de regras, as quais expressam o conhecimento existente. Cabe ressaltar que as regras foram geradas a partir de simulações, ou treinamentos, os quais realizam a comparação entre padrões.

As tabelas 5.2, 5.3, 5.4, 5.5, 5.6, 5.7 e 5.8, obtidas na etapa de aquisição ou extração das características (“features”) dos sinais de voz, foram utilizadas para estruturar os Sistemas de Representação de Conhecimento. A partir de tais tabelas pode-se perceber que no sistema proposto foram usados 8 atributos (F1, F2, F3, F4, F5, F6, F7 e Ffund) para caracterizar os padrões referentes aos sinais de voz, entretanto é importante lembrar que o número de características pode ser variável.

a) Conjunto de Treinamento 1

Inicialmente se realizou um conjunto de treinamento que incluiu sinais de voz de três locutores, com o objetivo de distingui-los quando na aquisição de um novo padrão de informação. Dentre esses locutores estão: um locutor do sexo feminino (identificado pelo código 3), e dois locutores do sexo masculino (identificados pelo código de saída 4 e 5).

Como mencionado no capítulo 5, para cada locutor foram realizadas gravações de 10 arquivos, a fim de gerar 40 padrões referentes à cada locutor. Assim, para o primeiro conjunto de treinamento foi utilizado um total de 120 padrões, como mostra a tabela 8.1, que na realidade nada mais é do que a união das tabelas 5.3, 5.4 e 5.5, considerando os 8 atributos F1, F2, F3, F4, F5, F6, F7 e Ffund, e que agora passa a ser referenciada como um Sistema de Representação de Conhecimento.

Tabela 8.1 - Sistema de Representação de Conhecimento do conjunto de treinamento 1

Índice	F1	F2	F3	F4	F5	F6	F7	Ffund	Saída
1	211	422	633	844	1055	1258	1477	211	3
2	211	422	633	844	1055	1266	1477	211	3
3	219	422	633	844	1055	1266	1477	211	3

Tabela 8.1 - Continuação

4	211	422	633	844	1047	1266	1477	211	3
5	211	430	648	867	1078	1297	1516	213	3
6	211	430	641	859	1078	1289	1500	213	3
7	203	414	625	836	1039	1250	1461	213	3
8	211	422	625	836	1047	1258	1469	213	3
9	211	422	633	844	1055	1273	1477	213	3
10	211	422	633	852	1063	1273	1484	213	3
11	211	430	641	859	1070	1281	1500	213	3
12	219	430	648	859	1070	1289	1500	213	3
13	219	438	656	883	1102	1320	1539	218	3
14	219	438	664	883	1102	1320	1547	218	3
15	219	438	648	875	1086	1305	1523	218	3
16	219	438	656	867	1086	1305	1516	218	3
17	219	430	641	859	1070	1289	1500	214	3
18	211	430	641	859	1070	1289	1500	214	3
19	219	438	656	867	1086	1297	1516	214	3
20	211	430	633	852	1063	1273	1492	214	3
21	219	445	664	891	1109	1336	1555	220	3
22	227	453	680	906	1133	1352	1586	227	3
23	234	461	695	922	1156	1391	1617	227	3
24	219	453	672	898	1125	1344	1570	220	3
25	227	453	680	906	1133	1359	1586	223	3
26	211	438	656	867	1094	1305	1523	220	3
27	227	445	672	898	1125	1352	1570	223	3
28	234	461	695	922	1156	1383	1617	223	3
29	227	445	672	898	1125	1352	1570	235	3
30	242	469	703	938	1172	1405	1641	241	3
31	234	469	695	930	1164	1398	1633	235	3
32	242	477	719	961	1203	1438	1680	241	3
33	219	438	656	875	1094	1313	1539	220	3
34	227	453	688	914	1141	1367	1594	227	3
35	234	461	695	922	1156	1383	1609	227	3
36	219	445	664	891	1109	1336	1555	227	3
37	219	430	648	859	1078	1289	1508	219	3
38	227	445	664	883	1109	1328	1555	221	3
39	227	445	664	883	1102	1328	1547	221	3
40	211	438	656	867	1086	1305	1523	219	3
41	141	273	406	547	688	820	961	143	4
42	133	266	398	530	664	797	938	133	4
43	133	266	398	530	664	797	930	133	4
44	133	266	391	523	655	781	914	133	4
45	133	266	391	523	656	789	922	133	4
46	133	258	391	516	648	781	906	133	4
47	133	273	406	539	672	805	938	133	4
48	141	273	406	547	680	820	954	143	4
49	141	281	430	563	703	844	984	141	4
50	141	281	422	563	703	844	984	141	4
51	141	281	422	563	703	844	992	141	4
52	141	273	414	555	688	828	961	141	4
53	141	281	414	547	695	836	969	143	4
54	141	273	414	547	680	820	961	143	4
55	133	273	406	531	672	813	945	133	4
56	133	266	398	531	664	797	930	133	4
57	156	305	460	609	766	914	1060	154	4
58	156	313	461	617	768	921	1078	154	4
59	150	305	453	602	758	914	1070	154	4
60	148	305	453	609	758	906	1050	154	4

Tabela 8.1 - Continuação

61	133	266	398	530	664	797	930	133	4
62	133	266	391	523	655	781	914	133	4
63	133	273	406	531	672	813	945	133	4
64	133	266	398	531	664	797	930	133	4
65	156	305	460	609	766	914	1060	154	4
66	156	313	461	617	768	921	1078	154	4
67	133	266	398	530	664	797	930	133	4
68	133	266	391	523	655	781	914	133	4
69	133	266	391	523	655	781	914	133	4
70	133	266	391	523	656	789	922	133	4
71	133	258	391	516	648	781	906	133	4
72	133	273	406	539	672	805	938	133	4
73	141	273	406	547	680	820	954	143	4
74	141	281	430	563	703	844	984	141	4
75	141	281	422	563	703	844	984	141	4
76	141	281	422	563	703	844	992	141	4
77	133	266	391	523	655	781	914	133	4
78	133	266	391	523	656	789	922	133	4
79	141	273	406	547	680	820	954	143	4
80	141	281	430	563	703	844	984	141	4
81	117	234	352	469	586	711	820	119	5
82	117	234	359	484	602	719	836	119	5
83	117	242	367	477	602	727	844	119	5
84	125	242	356	477	602	719	844	121	5
85	117	234	352	484	586	711	828	117	5
86	117	234	352	453	594	703	828	117	5
87	117	234	352	461	586	703	933	117	5
88	117	234	359	477	602	719	836	117	5
89	141	234	367	602	719	844	960	145	5
90	125	242	367	500	602	727	844	121	5
91	117	242	360	484	609	727	852	119	5
92	117	242	352	484	594	719	836	119	5
93	125	250	374	492	617	734	859	125	5
94	125	242	363	477	602	727	852	125	5
95	125	250	356	484	617	742	875	125	5
96	125	250	421	508	633	750	883	125	5
97	125	242	367	492	617	742	859	125	5
98	125	250	375	508	633	755	883	125	5
99	125	242	383	492	609	734	852	125	5
100	125	250	360	500	625	742	883	125	5
101	117	242	367	477	602	727	844	119	5
102	125	242	356	477	602	719	844	121	5
103	117	234	352	484	586	711	828	117	5
104	117	234	352	453	594	703	828	117	5
105	117	242	352	484	594	719	836	119	5
106	125	242	367	492	617	742	859	125	5
107	117	234	359	477	602	719	836	117	5
108	117	242	367	477	602	727	844	119	5
109	117	242	360	484	609	727	852	119	5
110	117	242	352	484	594	719	836	119	5
111	125	250	374	492	617	734	859	125	5
112	125	242	363	477	602	727	852	125	5
113	117	234	352	469	586	711	820	119	5
114	117	234	359	484	602	719	836	119	5
115	117	242	367	477	602	727	844	119	5
116	125	242	356	477	602	719	844	121	5
117	117	234	359	477	602	719	836	117	5

Tabela 8.1 - Continuação

118	117	242	367	477	602	727	844	119	5
119	117	242	352	484	594	719	836	119	5
120	125	242	356	477	602	719	844	121	5

Classificação dos Dados em Faixas

Como descrito no capítulo 7, todo processo se inicia definindo-se faixas para os diversos atributos observados. Assim, o primeiro passo do algoritmo dos Conjuntos Aproximados permite redefinir os valores (de acordo com uma certa métrica) de cada coluna (ou atributo) da tabela 8.1 que contém o SRC. Neste caso, os valores de cada frequência serão substituídos por valores inteiros que irão identificar uma certa faixa de valores na qual a frequência se enquadra, gerando assim novas tabelas.

Para as tabelas usadas como Sistemas de Representação de Conhecimento, foram definidas várias faixas de valores para cada atributo, como mostram as tabelas A.1, A.2, A.3, A.4, A.5, A.6, A.7, A.8, incluídas no apêndice A.

A tabela B.1 do apêndice B representa a tabela 8.1, porém com seus valores redefinidos.

Simulação 1

Ao obter o SRC do conjunto de treinamento foram realizados os seguintes passos responsáveis pela simulação, são eles:

1. Transformação dos valores dos atributos em faixas, a qual já foi mencionada;
2. Eliminação de atributos idênticos;
3. Eliminação de padrões ou exemplos idênticos (é importante lembrar que a repetição de padrões pode ou não surgir após a redefinição de valores);
4. Eliminação de Atributos Dispensáveis;
5. Busca do conjunto de Redutos (ou conjunto com o mesmo poder discriminatório que o conjunto inicial);
6. Composição do Conjunto de Regras.

Assim, na simulação 1 foram realizados os passos abaixo:

Passo 1: Cada atributo foi redefinido em faixas de valores;

Passo 2: Com a redefinição das faixas observou-se que o atributo Ffund possui valores semelhantes ao atributo F1, o que caracteriza um atributo idêntico. Dessa forma o atributo Ffund foi eliminado da tabela;

Passo 3: A existência de padrões semelhantes permitiu uma redução de 57,5% da base de dados do SRC apresentado pela tabela 8.1, restando apenas os padrões diferentes, os quais correspondem a um montante de 51 padrões (ou 42,5% da base de dados inicial). A Tabela 8.2, composta pelos 5 primeiros padrões da tabela 8.1, é usada para demonstrar esse passo, sendo possível perceber como se realizou redução da base de dados após a redefinição das faixas e a eliminação do atributo Ffund. Nota-se que após tais passos a tabela apresenta padrões idênticos, como os padrões 1, 2 e 3, de forma que os padrões 2 e 3 podem ser eliminados da tabela;

Tabela 8.2 - Exemplo de identificação de exemplos semelhantes

Índice	F1	F2	F3	F4	F5	F6	F7	Saída
1	7	8	13	17	20	23	28	3
2	7	8	13	17	20	23	28	3
3	7	8	13	17	20	23	28	3
4	7	8	13	17	19	23	28	3
5	7	9	13	18	21	24	29	3

A tabela 8.3 apresenta os 51 padrões mencionados.

Tabela 8.3 - SRC após passos 1, 2, 3 e 4

U	F1	F2	F3	F4	F5	F6	F7	Saída
1	7	8	13	17	20	23	28	3
2	7	8	13	17	19	23	28	3
3	7	9	13	18	21	24	29	3
4	7	9	13	18	21	24	28	3
5	7	8	12	17	19	22	27	3
6	7	8	12	17	19	23	27	3
7	7	8	13	18	20	23	28	3
8	7	9	13	18	20	24	28	3
9	7	9	14	19	22	25	30	3
10	7	9	13	18	21	25	29	3
11	7	9	14	18	21	25	29	3
12	7	9	14	18	21	24	29	3
13	7	9	13	18	20	23	28	3
14	7	9	14	19	22	26	31	3

Tabela 8.3 - Continuação

15	8	10	15	20	23	27	32	3
16	8	10	15	20	24	28	33	3
17	7	10	14	19	22	26	31	3
18	8	9	14	19	22	27	31	3
19	9	10	16	21	24	29	34	3
20	8	10	15	21	24	28	34	3
21	9	11	16	22	26	30	36	3
22	7	9	14	18	21	25	30	3
23	8	9	14	19	22	26	31	3
24	8	9	14	19	22	26	30	3
25	4	2	4	5	5	5	7	4
26	3	2	3	5	4	4	6	4
27	3	2	3	4	4	4	5	4
28	3	2	3	4	3	4	5	4
29	3	2	4	5	4	5	6	4
30	4	3	5	6	6	6	8	4
31	4	3	4	6	6	6	8	4
32	4	2	4	6	5	6	7	4
33	4	3	4	5	5	6	7	4
34	4	4	6	8	8	9	11	4
35	4	4	6	8	8	9	12	4
36	4	4	6	8	8	9	10	4
37	2	1	2	2	1	1	1	5
38	2	1	2	3	2	1	2	5
39	2	1	2	3	2	2	2	5
40	3	1	2	3	2	1	2	5
41	2	1	2	3	1	1	2	5
42	2	1	2	2	1	1	2	5
43	2	1	2	2	1	1	6	5
44	4	1	2	8	6	6	7	5
45	3	1	2	3	2	2	2	5
46	2	1	2	3	2	2	3	5
47	3	1	2	3	2	2	3	5
48	3	1	4	4	3	2	4	5
49	3	1	2	4	3	3	4	5
50	3	1	3	3	2	2	3	5
51	3	1	2	3	2	2	4	5

Passo 4: Para o conjunto de treinamento 1, e tendo como base os algoritmos apresentados no capítulo 7, não se identificou inicialmente a existência de atributos dispensáveis, ou atributos que quando eliminados um a um, não interferem na classificação de padrões. Esse fato foi observado a partir dos valores dos antecedentes juntamente com os conseqüentes.

Passo 5: Após tratar a tabela que contém o Sistema de Representação de Conhecimento, a fase de treinamento realiza uma busca de subconjuntos (ou redutos) que possuam o mesmo poder discriminatório que conjunto original.

Tomando como base a Tabela 8.3, e aplicando-se o algoritmo nos seus diversos passos, foram obtidos os conjuntos de atributos que formam os primeiros conjuntos de redução, ou redutos, apresentados pela tabela 8.4.

Tabela 8.4 - Redutos iniciais do conjunto de treinamento 1

Conjunto Redução	Atributos
R1	F2, F3, F4, F5, F6, F7
R2	F1, F3, F4, F5, F6, F7
R3	F1, F2, F4, F5, F6, F7
R4	F1, F2, F3, F5, F6, F7
R5	F1, F2, F3, F4, F6, F7
R6	F1, F2, F3, F4, F5, F7
R7	F1, F2, F3, F4, F5, F6

A partir dos redutos encontrados, foram realizadas novas buscas até se encontrar os conjuntos mínimos. Na aplicação tomada como exemplo foram encontrados os conjuntos mínimos apresentados pela tabela 8.5.

Tabela 8.5 - Conjuntos mínimos do conjunto de treinamento 1

Conjunto Redução	Atributos
R1	F5, F7
R2	F4, F7
R3	F4, F6
R4	F3, F7
R5	F3, F6
R6	F3, F5
R7	F3, F4
R8	F2, F7
R9	F2, F6
R10	F2, F5
R11	F2, F4
R12	F2, F3
R13	F1, F2

Passo 6: Encontrado os conjuntos mínimos, passa-se para o último passo caracterizado pela geração ou composição de um conjunto de regras baseadas nos conjuntos mínimos, e que irão representar o estado operacional do sistema.

Nessa etapa cada reduto (conjunto mínimo) agrupado ao atributo de decisão (coluna da saída), representará um novo Sistema de Representação de Conhecimento, tornando possível a verificação de regras ou relações de equivalência obtidas a partir do reduto.

Tomando como exemplo o reduto R13 que consiste dos atributos F1, F2, cria-se um novo Sistema de Representação de Conhecimento (SRC) apresentado pela tabela 8.6.

Tabela 8.6 - SRC agrupado por F1, F2, Saída (sem padrões repetidos)

Nr.	F1	F2	Saída
1	7	8	3
2	7	9	3
3	8	10	3
4	7	10	3
5	8	9	3
6	9	10	3
7	9	11	3
8	4	2	4
9	3	2	4
10	4	3	4
11	4	4	4
12	2	1	5
13	3	1	5
14	4	1	5

Nesse momento cada linha da tabela 8.6 pode ser identificada como uma regra.

Para verificar se existem regras mínimas e consistentes basta seleccionar os atributos separadamente, ou seja, F1 e F2 devem ser observados separadamente, levando em consideração a saída. Para cada valor associado a F1 (como antecedente da regra) verifica-se o valor da Saída (conseqüente da regra), a fim de identificar se a regra é consistente ou não.

É possível observar uma regra inconsistente quando se observa o atributo F1 separadamente, e considerando os exemplos identificados por *Nr.* igual a 8 e 14. Para esses exemplos o atributo *F1* possui o valor 4, porém para o atributo *Saída* possuem valores diferentes, são eles: 4 e 5. Isso caracteriza duas regras que possuem os mesmos valores para seus antecedentes, mas com valores diferentes para seus conseqüentes. Dessa forma a regra mínima $F1=4$ não é considerada, pois não tem poder de classificação quando esta é considerada isoladamente, deixando evidente a necessidade de um segundo atributo, ou mais atributos, para a classificação de padrões, quando $F1=4$.

Apenas as regras consistentes estruturam o conjunto de regras final.

Após verificar a consistência das regras mínimas da tabela 8.6 foram encontradas as seguintes regras, as quais representam as regras mínimas para o reduto que agrupa os atributos F1 e F2:

Se F1=2	então	Saída=5
Se F1=7	então	Saída=3
Se F1=8	então	Saída=3
Se F1=9	então	Saída=3

As regras mencionadas acima podem ser encontradas na tabela C.1 do apêndice C, sendo identificadas como regras de número 290, 291, 292 e 293.

Assim, pode-se perceber que o número total de regras está associado à quantidade de redutos encontrados. A tabela C.1 apresenta todas as regras encontradas pelo conjunto de treinamento 1. Entretanto, é possível perceber a existência de inúmeras regras que poderiam ser eliminadas da tabela, como é o caso da regra 287, onde o valor 2 é atribuído ao atributo F1, sendo assim nenhum outro atributo é necessário para a classificação de padrões cujo atributo F1 é igual a 2. A tabela D.1 apresenta o conjunto de regras mínimas.

b) Conjunto de Treinamento 2

O segundo conjunto de treinamento incluiu sinais de voz de dois locutores do sexo feminino, identificados pelos códigos 1 e 2, conforme tabela 5.1.

Para estruturar o SRC, foram usadas as tabelas 5.2 e 5.3, somando assim um total de 80 padrões.

Classificação dos Dados em Faixas

O processo usado no conjunto de treinamento 1, para a classificação dos dados, também foi aplicado para o conjunto de treinamento 2, obtendo-se a tabela B.2.

Simulação 2

A simulação do conjunto de treinamento 2 segue o mesmo procedimento usado pela simulação 1.

Após realizar os passos 1, 2, 3 e 4, obteve-se a tabela 8.7 que contém apenas os padrões não semelhantes. A redução obtida foi de 55%, restando apenas 45% da tabela inicial, o que corresponde aos 36 padrões da tabela 8.7.

Tabela 8.7 - SRC do conjunto de treinamento 2 após passos 1, 2, 3 e 4

U	F1	F2	F3	F4	F5	F6	F7	Saída
1	7	8	12	17	19	22	28	1
2	7	8	12	17	19	23	28	1
3	9	11	17	23	27	32	38	1
4	9	11	17	23	27	31	37	1
5	9	11	17	24	28	33	38	1
6	9	11	16	22	26	30	35	1
7	9	11	16	22	25	30	36	1
8	9	11	16	22	26	30	36	1
9	9	11	17	22	26	31	36	1
10	9	11	17	23	26	31	36	1
11	8	10	15	21	24	28	33	1
12	8	10	16	21	22	29	34	1
13	8	10	16	21	25	29	35	1
14	8	10	16	21	24	28	34	1
15	9	11	18	24	28	33	39	1
16	9	12	18	24	28	33	39	1
17	9	12	18	24	28	33	40	1
18	9	12	18	24	29	34	40	1
19	9	12	18	24	28	34	40	1
20	9	12	18	25	29	34	40	1
21	7	9	14	18	21	25	30	2
22	7	9	13	18	21	24	29	2
23	7	9	13	18	20	24	28	2
24	7	9	14	18	22	25	30	2
25	7	9	14	19	21	25	30	2
26	7	9	14	18	21	25	29	2
27	7	9	14	19	22	26	31	2
28	8	9	14	19	22	26	31	2
29	8	10	15	20	23	27	32	2
30	8	9	14	19	22	26	32	2
31	8	10	14	19	22	26	31	2
32	8	9	14	19	22	27	31	2
33	8	10	14	19	22	27	31	2
34	8	10	15	20	23	27	33	2
35	8	10	15	20	24	28	33	2
36	8	10	15	19	22	27	32	2

O passo 5 permitiu encontrar redutos iniciais que representam os primeiros conjuntos redução encontrados pelo sistema. Os redutos iniciais foram utilizados para a realização de novas buscas, de forma que foram encontrados os conjuntos mínimos apresentados na tabela 8.8, os quais permitiram a elaboração de regras a partir do passo 6. Tais regras podem ser verificadas na tabela C.2, que contém o total de regras do conjunto de treinamento 2. A tabela D.2 apresenta o conjunto de regras mínimas desse treinamento.

Tabela 8.8 - Conjuntos mínimos do conjunto de treinamento 2

Conjunto Redução	Atributos
R1	F4, F7
R2	F4, F6
R3	F4, F5
R4	F3, F4
R5	F2, F4
R6	F1, F4

c) *Conjunto de Treinamento 3*

O conjunto de treinamento 3 incluiu sinais de voz de vários locutores. Porém, para os padrões referentes aos sinais de voz, usou-se apenas os códigos 1, 2 e 3 para o atributo saída. É importante ressaltar que as saídas com valores correspondentes aos números 1 e 2, representam os reais locutores identificados por tais códigos, ou seja, os locutores identificados por 1 e 2 da tabela 5.1. Para os demais padrões, os quais correspondem aos dados dos sinais de voz dos locutores identificados por 4, 5, 6 e 7 da tabela 5.1, foi considerado apenas um valor para o atributo saída; dessa forma foi associado o valor 3 para o atributo saída desses padrões.

Considerando as observações anteriores foi possível estruturar o Sistema de Representação de Conhecimento do conjunto de treinamento 3, o qual foi obtido a partir das tabelas 5.2, 5.3, 5.5, 5.6, 5.7, e 5.8, com a escolha aleatória dos padrões, somando um total de 229 padrões.

Classificação dos Dados em Faixas

Assim como nos treinamentos anteriores, o conjunto de treinamento 3 inicia fazendo uso do procedimento de redefinição de faixas para os atributos observados. A tabela B.3 representa o SRC com seus valores redefinidos.

Simulação 3

A simulação do conjunto de treinamento 3 segue os mesmos passos da simulação 1, de forma que a partir dos quatro primeiros passos foi possível obter a tabela 8.9, que apresenta 47% do SRC inicial. A redução inicial obtida nesse conjunto de treinamento foi de 53% do total de padrões do SRC.

Tabela 8.9 - SRC do conjunto de treinamento 3 após passos 1, 2, 3 e 4

U	F1	F2	F3	F4	F5	F6	F7	Saída
1	7	8	12	17	19	22	28	1
2	7	8	12	17	19	23	28	1
3	9	11	17	23	27	32	38	1
4	9	11	17	23	27	31	37	1
5	9	11	17	24	28	33	38	1
6	9	11	16	22	26	30	35	1
7	9	11	16	22	25	30	36	1
8	9	11	16	22	26	30	36	1
9	9	11	17	22	26	31	36	1
10	9	11	17	23	26	31	36	1
11	8	10	15	21	24	28	33	1
12	8	10	16	21	22	29	34	1
13	8	10	16	21	25	29	35	1
14	8	10	16	21	24	28	34	1
15	9	11	18	24	28	33	39	1
16	9	12	18	24	28	33	39	1
17	9	12	18	24	28	33	40	1
18	9	12	18	24	29	34	40	1
19	9	12	18	24	28	34	40	1
20	9	12	18	25	29	34	40	1
21	7	9	14	18	21	25	30	2
22	7	9	13	18	21	24	29	2
23	7	9	13	18	20	24	28	2
24	7	9	14	18	22	25	30	2
25	7	9	14	19	21	25	30	2
26	7	9	14	19	22	26	31	2
27	8	9	14	19	22	26	31	2
28	8	10	15	20	23	27	32	2
29	8	9	14	19	22	26	32	2
30	8	10	14	19	22	26	31	2
31	8	9	14	19	22	27	31	2
32	8	10	14	19	22	27	31	2
33	8	10	15	20	23	27	33	2
34	8	10	15	20	24	28	33	2
35	8	10	15	19	22	27	32	2
36	4	2	4	5	5	5	7	3
37	3	2	3	5	4	4	6	3
38	3	2	3	4	4	4	5	3
39	3	2	3	4	3	4	5	3
40	3	2	4	5	4	5	6	3
41	4	3	5	6	6	6	8	3
42	4	3	4	6	6	6	8	3

Tabela 8.9 - Continuação

43	4	2	4	6	5	6	7	3
44	4	3	4	5	5	6	7	3
45	4	4	6	8	8	9	11	3
46	4	4	6	8	8	9	12	3
47	4	4	6	8	8	9	10	3
48	2	1	2	2	1	1	1	3
49	2	1	2	3	2	1	2	3
50	2	1	2	3	2	2	2	3
51	3	1	2	3	2	1	2	3
52	2	1	2	3	1	1	2	3
53	4	1	2	8	6	6	7	3
54	3	1	2	3	2	2	2	3
55	2	1	2	3	2	2	3	3
56	3	1	2	3	2	2	3	3
57	3	1	4	4	3	2	4	3
58	3	1	2	4	3	3	4	3
59	3	1	3	3	2	2	3	3
60	3	1	2	3	2	2	4	3
61	3	2	3	5	4	5	6	3
62	3	2	2	5	4	4	5	3
63	3	2	2	4	2	3	4	3
64	3	1	3	4	3	3	4	3
65	3	2	3	4	3	3	5	3
66	3	2	3	4	3	3	4	3
67	3	2	3	5	5	5	6	3
68	3	2	4	5	5	5	7	3
69	5	4	4	3	4	5	8	3
70	5	4	7	9	10	11	14	3
71	5	4	7	9	9	11	13	3
72	4	4	7	9	9	11	13	3
73	5	5	7	10	10	12	15	3
74	5	5	7	10	10	12	14	3
75	5	5	7	9	10	12	14	3
76	5	4	7	9	10	11	13	3
77	6	6	10	13	15	17	20	3
78	6	6	10	13	14	17	20	3
79	5	6	9	13	14	16	20	3
80	5	5	7	10	11	12	15	3
81	5	5	7	10	11	13	16	3
82	5	5	8	11	11	13	16	3
83	5	5	8	10	11	12	15	3
84	5	6	9	12	13	15	18	3
85	5	6	8	12	13	22	25	3
86	5	6	8	12	12	22	25	3
87	5	6	8	11	12	21	24	3
88	7	7	12	15	17	21	24	3
89	6	7	12	15	17	20	24	3
90	7	7	11	15	17	20	24	3
91	6	7	11	15	16	20	24	3
92	6	7	11	15	17	20	24	3
93	7	7	11	15	18	20	24	3
94	7	7	12	16	17	21	25	3
95	7	8	12	16	17	21	25	3
96	7	8	12	16	18	21	25	3
97	7	8	13	17	17	23	27	3
98	7	8	12	17	17	23	27	3
99	7	8	13	17	18	23	27	3

Tabela 8.9 - Continuação

100	7	8	12	16	18	22	26	3
101	7	8	12	16	19	22	26	3
102	6	7	12	15	18	20	24	3
103	6	7	11	15	18	20	24	3
104	7	8	13	17	19	23	27	3
105	6	7	11	14	16	19	23	3
106	6	7	10	14	16	18	22	3
107	6	7	11	14	16	19	22	3
108	6	7	11	15	17	20	23	3

Os conjuntos redução que possuem o mesmo poder discriminatório que o conjunto inicial, ou redutos finais, são encontrados após a realização dos passos 5 e 6. A tabela 8.10 apresenta tais conjuntos redução, que foram usados para estruturar um conjunto de regras usadas para classificar novos padrões de informações. O conjunto total de regras pode ser verificado pela tabela C.3. A tabela D.3 apresenta o conjunto de regras mínimas do treinamento 3.

Tabela 8.10 - Conjuntos mínimos do conjunto de treinamento 3

Conjunto Redução	Atributos
R1	F3, F4, F5
R2	F4, F7

8.4 – Testes e Resultados

Após a fase de treinamento passou-se para a fase de testes, na qual foram utilizados novos padrões de informação referentes a alguns locutores; o que permitiu obter resultados para avaliação do classificador proposto.

Os padrões usados nos testes podem ser observados na tabela 5.9, e apresentam as mesmas características dos padrões usados na fase de treinamento.

Cabe ressaltar que, antes de aplicar os padrões aos conjuntos de regras, se faz necessário redefinir os valores de seus atributos, lembrando ainda que para cada conjunto de treinamento foram utilizados apenas os padrões referentes aos locutores envolvidos em cada treinamento, o que permite ao sistema identificar ou não, à qual dos locutores pertence os novos padrões de informação.

Assim, para os testes realizados, os quais são apresentados a seguir, foram utilizadas novas tabelas de dados contendo os padrões, sendo que os valores dos atributos foram redefinidos. Além disso, não há repetição de padrões, e o atributo Ffund não é considerado

devido ser um atributo dispensável, onde esse aspecto foi verificado na etapa de passos prévios.

a) *Testes e Resultados 1*

Os padrões utilizados para os testes relacionados ao conjunto de treinamento 1, incluiu informações sobre dois locutores, um do sexo feminino e identificado pelo código 3, e um do sexo masculino, identificado pelo código 4.

Após obter os padrões referentes aos locutores mencionados, foi possível aplicar as informações no conjunto de regras geradas pelo treinamento 1.

Os conjuntos de regras representam seqüências que são executadas condicionalmente, de forma que cada característica ou atributo do novo padrão é verificado, tanto independentemente quanto combinado a outros atributos; essa verificação é feita até que se encontre um valor que represente a saída do sistema. É importante lembrar que, a verificação de atributos combinados irá depender das regras geradas.

A tabela 8.11 inclui os padrões usados para alimentar as regras geradas pelo conjunto de treinamento 1, e apresenta os resultados obtidos.

Tabela 8.11 - Teste de classificação para locutores do conjunto de treinamento 1

Índice	F1	F2	F3	F4	F5	F6	F7	Referência do Locutor	Saída do Sistema	Resultado
1	7	8	13	17	19	23	27	3	3	Verdadeiro
2	7	9	13	18	21	24	29	3	3	Verdadeiro
3	7	8	13	17	20	23	28	3	3	Verdadeiro
4	7	8	12	17	19	22	27	3	3	Verdadeiro
5	7	8	12	17	19	23	27	3	3	Verdadeiro
6	7	8	12	16	18	22	26	3	3	Verdadeiro
7	7	9	13	18	20	24	28	3	3	Verdadeiro
8	8	9	14	19	22	26	31	3	3	Verdadeiro
9	7	9	14	19	22	25	30	3	3	Verdadeiro
10	4	2	4	5	5	5	7	4	4	Verdadeiro
11	3	2	3	5	4	5	6	4	4	Verdadeiro
12	3	2	3	4	4	4	5	4	4	Verdadeiro
13	3	2	3	4	3	4	5	4	4	Verdadeiro
14	3	2	4	5	4	5	6	4	4	Verdadeiro
15	3	2	4	5	5	5	7	4	4	Verdadeiro
16	4	2	4	6	6	6	8	4	4	Verdadeiro
17	4	3	4	6	6	6	8	4	4	Verdadeiro
18	4	2	4	6	5	5	7	4	4	Verdadeiro
19	3	2	4	5	4	5	5	4	4	Verdadeiro
20	3	2	4	5	5	5	6	4	4	Verdadeiro
21	3	2	3	5	4	4	6	4	4	Verdadeiro

Tabela 8.12 - Resumo da tabela 8.11

	Quantidade	Taxa
Número de Acertos	21	100%
Número de Erros	0	0%
Total de Padrões	21	100%

Para o primeiro conjunto de treinamento, o sistema obteve uma taxa de acerto de 100% em 21 testes, como mostra a figura 8.1, atribuindo à saída do sistema os mesmos valores de referência do locutor. Dessa forma o sistema foi capaz de classificar todos os padrões, após a verificação das regras incluídas na tabela C.1 e tabela D.1.

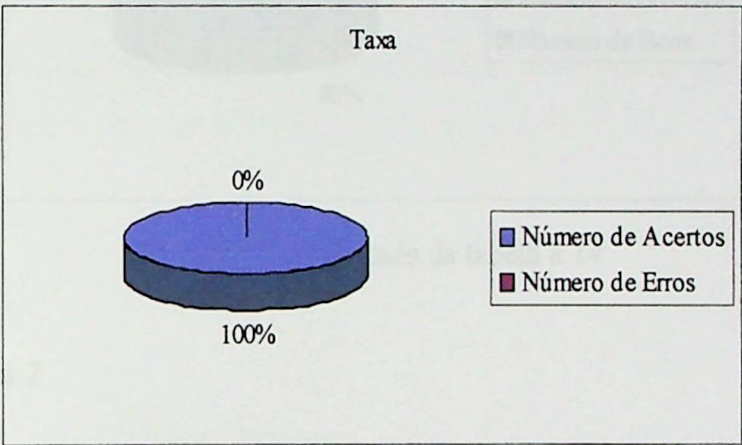


Figura 8.1 - Resultado da tabela 8.12

A tabela 8.13 inclui novos testes para o conjunto de regras, onde as gravações dos sinais de voz foram realizadas em ambiente diferente dos adquiridos e apresentados pela tabela 8.11. Para esse caso, o sistema apresentou uma taxa de erro quando na classificação de padrões contendo o novo sinal de voz, como mostra a figura 8.2.

Tabela 8.13 - Novo teste de classificação para locutores do conjunto de treinamento 1

Índice	F1	F2	F3	F4	F5	F6	F7	Referência do Locutor	Saída do Sistema	Resultado
1	3	2	2	3	3	2	3	4	4	Verdadeiro
2	3	2	2	3	2	2	2	4	4	Verdadeiro
3	3	2	2	3	3	2	5	4	4	Verdadeiro
4	2	1	2	3	2	2	2	4	5	Falso
5	3	2	3	4	3	3	4	4	4	Verdadeiro
6	3	2	3	4	3	4	5	4	4	Verdadeiro

Tabela 8.14 - Resumo da tabela 8.13

	Quantidade	Taxa
Número de Acertos	5	83%
Número de Erros	1	17%
Total de Padrões	6	100%

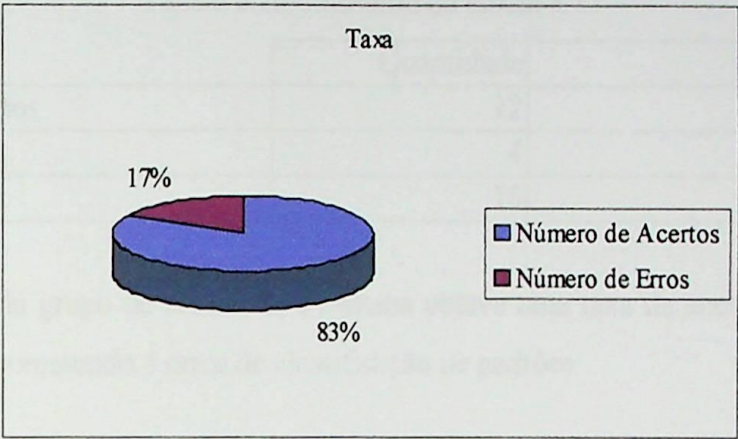


Figura 8.2 - Resultado da tabela 8.14

b) Testes e Resultados 2

Os testes para o conjunto de treinamento 2 incluíram padrões contendo informações sobre dois locutores do sexo feminino, os quais são identificados pelos códigos 1 e 2, conforme tabela 5.1.

Os padrões foram aplicados às regras obtidas na simulação, onde os resultados obtidos são apresentados pela tabela 8.15.

Tabela 8.15 - Teste de classificação para locutores do conjunto de treinamento 2

Índice	F1	F2	F3	F4	F5	F6	F7	Referência do Locutor	Saída do Sistema	Resultado
1	8	10	15	21	24	28	33	1	1	Verdadeiro
2	7	9	14	19	22	26	31	1	2	Falso
3	8	10	15	20	22	27	32	1	2	Falso
4	9	10	16	22	25	30	35	1	1	Verdadeiro
5	9	11	17	23	26	31	37	1	1	Verdadeiro
6	9	12	18	24	28	34	40	1	1	Verdadeiro
7	9	11	17	23	27	31	37	1	1	Verdadeiro
8	9	12	19	25	29	35	37	1	1	Verdadeiro
9	7	8	12	17	19	22	28	2	2	Verdadeiro
10	8	10	15	20	23	27	32	2	1	Falso

Tabela 8.15 - Continuação

11	7	9	13	18	21	24	29	2	2	Verdadeiro
12	7	8	13	18	20	23	28	2	2	Verdadeiro
13	7	9	14	19	21	25	30	2	2	Verdadeiro
14	7	8	13	18	21	25	28	2	2	Verdadeiro
15	7	9	14	19	22	25	30	2	2	Verdadeiro
16	9	8	13	19	27	31	37	2	1	Falso

Tabela 8.16 – Resumo da tabela 8.15

	Quantidade	Taxa
Número de Acertos	12	75%
Número de Erros	4	25%
Total de Padrões	16	100%

Com o segundo grupo de locutores, o sistema obteve uma taxa de acerto de 75% em 16 testes realizados, cometendo 4 erros de classificação de padrões.

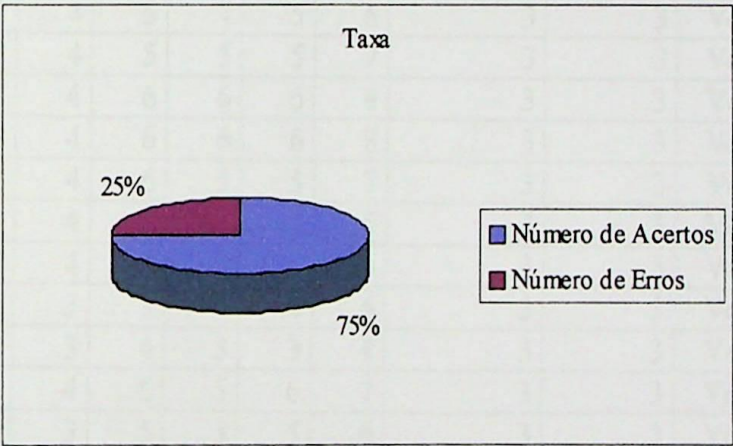


Figura 8.3 - Resultado da tabela 8.16

c) Testes e Resultados 3

Os testes realizados para o conjunto de treinamento 3 incluiu sinais de voz de vários locutores, porém os locutores são identificados por apenas três códigos de referência, como conjunto de treinamento 3.

Esse treinamento foi elaborado a fim de medir a capacidade de classificação do sistema quando se consideram padrões um tanto quanto duvidosos; isso ocorre devido as características dos padrões possuírem valores muito diferentes, porém com a mesma informação de saída.

A tabela 8.17 apresenta os padrões que foram aplicados às regras obtidas na simulação

3.

Tabela 8.17 - Teste de classificação para locutores do conjunto de treinamento 3

Índice	F1	F2	F3	F4	F5	F6	F7	Referência do Locutor	Saída do Sistema	Resultado
1	8	10	15	21	24	28	33	1	1	Verdadeiro
2	7	9	14	19	22	26	31	1	2	Falso
3	8	10	15	20	22	27	32	1	2	Falso
4	9	10	16	22	25	30	35	1	1	Verdadeiro
5	9	11	17	23	26	31	37	1	1	Verdadeiro
6	9	12	18	24	28	34	40	1	1	Verdadeiro
7	9	11	17	23	27	31	37	1	1	Verdadeiro
8	9	12	19	25	29	35	37	1	1	Verdadeiro
9	7	8	12	17	19	22	28	2	1	Falso
10	8	10	15	20	23	27	32	2	2	Verdadeiro
11	4	2	4	5	5	5	7	3	3	Verdadeiro
12	3	2	3	5	4	5	6	3	3	Verdadeiro
13	3	2	3	4	4	4	5	3	3	Verdadeiro
14	3	2	3	4	3	4	5	3	3	Verdadeiro
15	3	2	4	5	4	5	6	3	3	Verdadeiro
16	3	2	4	5	5	5	7	3	3	Verdadeiro
17	4	2	4	6	6	6	8	3	3	Verdadeiro
18	4	3	4	6	6	6	8	3	3	Verdadeiro
19	4	2	4	6	5	5	7	3	3	Verdadeiro
20	3	2	4	5	4	5	5	3	3	Verdadeiro
21	3	2	4	5	5	5	6	3	3	Verdadeiro
22	3	2	3	5	4	4	6	3	3	Verdadeiro
23	3	2	3	4	3	3	4	3	3	Verdadeiro
24	4	2	4	6	5	6	7	3	3	Verdadeiro
25	3	2	3	5	5	5	6	3	3	Verdadeiro
26	5	4	4	3	4	5	8	3	3	Verdadeiro
27	5	4	7	9	10	11	14	3	3	Verdadeiro
28	5	4	7	9	9	11	13	3	3	Verdadeiro
29	4	4	7	9	9	11	13	3	3	Verdadeiro
30	5	5	7	10	10	12	15	3	3	Verdadeiro
31	2	1	2	3	2	1	2	3	3	Verdadeiro
32	4	1	2	8	6	6	7	3	3	Verdadeiro
33	3	1	2	3	2	2	3	3	3	Verdadeiro
34	9	11	17	23	27	32	38	1	1	Verdadeiro
35	9	12	18	25	29	35	40	1	1	Verdadeiro
36	9	12	19	25	29	35	40	1	1	Verdadeiro
37	9	12	19	25	30	36	40	1	1	Verdadeiro
38	9	11	16	22	25	30	35	1	1	Verdadeiro
39	9	11	16	22	25	29	35	1	1	Verdadeiro
40	9	11	16	22	26	30	36	1	1	Verdadeiro
41	7	9	13	18	21	24	29	2	2	Verdadeiro
42	7	8	13	18	20	23	28	2	2	Verdadeiro

Tabela 8.17 - Continuação

43	7	9	14	19	21	25	30	2	2	Verdadeiro
44	7	8	13	18	21	25	28	2	2	Verdadeiro
45	4	2	4	5	5	5	5	3	3	Verdadeiro
46	3	2	2	3	3	2	3	3	3	Verdadeiro
47	3	2	2	3	2	2	2	3	3	Verdadeiro
48	3	2	2	3	3	2	5	3	3	Verdadeiro
49	2	1	2	3	2	2	2	3	3	Verdadeiro
50	3	2	3	4	3	3	4	3	3	Verdadeiro
51	3	2	3	4	3	4	5	3	3	Verdadeiro

Tabela 8.18 - Resumo da tabela 8.17

	Quantidade	Taxa
Número de Acertos	48	94%
Número de Erros	3	6%
Total de Padrões	51	100%

Nos testes realizados para o conjunto de treinamento 3, o sistema obteve uma taxa de acerto de 94%. Cabe ressaltar que os 3 erros ocorreram em virtude da semelhança dos padrões de voz de diferentes locutores.

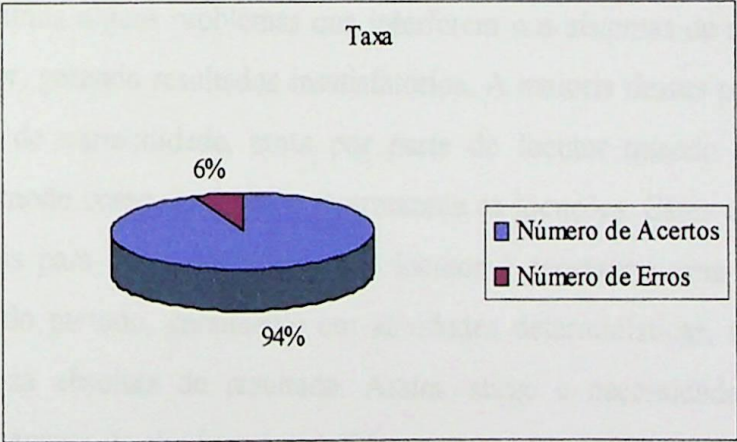


Figura 8.4 - Resultado da tabela 8.18

CONCLUSÃO

9.1 – Principais Considerações

O objetivo principal desse trabalho foi o desenvolvimento de um classificador baseado na Teoria de Conjuntos Aproximados para sinais de voz, a fim de testar a capacidade dessa ferramenta matemática e auxiliar no reconhecimento de locutor. O classificador foi desenvolvido e testado, sendo utilizado o MATLAB como plataforma de desenvolvimento. O apêndice E mostra a tela principal do sistema.

Assim, o MATLAB foi usado para desenvolver programas referentes a algumas técnicas de processamento de sinais, os quais foram capazes de extrair parâmetros que caracterizassem os sinais de voz. O MATLAB foi utilizado também para o desenvolvimento dos programas baseados nos algoritmos apresentados pelo capítulo 7.

Com a execução deste trabalho foram levantados os pontos relevantes para o desenvolvimento de um classificador de sinais de voz.

Cabe ressaltar que, apesar do grande número de pesquisas na área de reconhecimento de locutor, existem ainda alguns problemas que interferem nos sistemas de reconhecimento automático de locutor, gerando resultados insatisfatórios. A maioria desses problemas surge devido à existência de variabilidade, tanto por parte do locutor quando na geração de locuções, quanto ao modo como se obtém e se armazena as locuções. Esses aspectos devem ser considerados, pois para o reconhecimento de locutor é necessária uma total precisão; relativo a aplicação do método, geralmente em atividades determinísticas, na maioria das vezes obriga a certeza absoluta do resultado. Assim, surge a necessidade de busca de melhores algoritmos capazes de eliminar este problema.

Por isso percebeu-se que além da necessidade do uso de uma técnica de Inteligência Artificial robusta, foi importante utilizar parâmetros característicos estáveis em um certo período ao longo da locução, os quais não são afetados pela variação da maneira de falar, tal como imitações como também variações físicas como gripe.

Com relação ao uso de uma técnica de Inteligência Artificial, no caso a Teoria dos Conjuntos Aproximados, a mesma permitiu o estudo e aplicação de novos algoritmos na área de reconhecimento de locutor.

Foram realizadas várias simulações, a fim de medir a capacidade do classificador quando na identificação dos locutores. Os resultados são satisfatórios quando o novo padrão

de informação está relacionado a um locutor que esteja envolvido no conjunto de treinamento. Isso ocorre devido a semelhanças de informações sobre os locutores. Entretanto, a medida em que se aumenta o número de locutores, também aumenta a probabilidade de erro. Após as simulações verificou-se que o classificador permitiu eliminar o atributo Ffund, encontrado no Sistema de Informação inicial, devido o mesmo ser um atributo dispensável.

Um problema surgiu em treinamentos que envolviam vozes femininas. Após redefinir os valores das características para as faixas de valores, obteve-se padrões com características muito parecidas, porém com diferentes valores para a saída do sistema. Assim, constatou-se que a escolha para as faixas de valores deveria ser alterada a fim de evitar resultados insatisfatórios.

Um aspecto que pode ser considerado para primeiros testes quando no uso de novas técnicas como é o caso de “Rough Sets”, é a divisão de tarefas em módulos, o que permite diminuir a complexidade do problema. Como exemplo disso tem-se a separação de locutores em função de um determinado critério que diferencie os locutores (por exemplo: pelo sexo), o que pode vir a facilitar a classificação.

O sistema de classificação, baseado na Teoria de Conjuntos Aproximados, pode operar valores quantizados. A essência do procedimento de quantização aplicado às características é substituir valores discretos por faixas de valores. Devido o caráter não linear da maioria dos parâmetros da voz, a quantização consiste em dividir os parâmetros em um conjunto de faixas de valores, escolhidos arbitrariamente, que podem não ser particularmente eficazes, porém uma aproximação simplificada pode levar a uma porcentagem de reconhecimento aceitável quando o classificador “Rough Set” é empregado. A quantização dos reais valores dos atributos pode fornecer um problema central relacionado aos sistemas de decisão baseados em regras.

O número de faixas de valores pode ser limitado arbitrariamente. A determinação desse número pode ainda depender de observações relacionadas ao tempo e custo computacional.

O classificador permite identificar informações ou características supérfluas, além de tratar bases de conhecimento muito grande. A manipulação de uma grande massa de dados sempre foi um problema para os algoritmos convencionais, por isso pensou-se na utilização da Teoria dos Conjuntos Aproximados.

Além disso, o classificador permite transformar dados em conhecimento. Outro importante aspecto relacionado à classificação de sinais de voz é que cada condição da base de conhecimento está relacionada aos valores obtidos pelo sistema quando na etapa de

processamento dos sinais de voz. Assim, a eliminação de condições supérfluas da base de conhecimento pode resultar em benefícios para a análise final realizada pelo sistema a fim de reconhecer um locutor.

Foram utilizadas algumas características, tais como a frequência do “pitch” e as frequências de maior relevância, para o reconhecimento de locutor, entretanto é possível aumentar o número de características a fim de buscar resultados cada vez mais satisfatórios.

9.2 – Sugestões para Trabalhos Futuros

Como sugestões para trabalhos futuros tem-se a aplicação de diferentes locuções, como é o caso de testes com outras vogais, bem como testar a robustez do classificador para um número significativo de locutores, ou testar se o classificador possui efetividade e robustez quando no seu uso em tarefas mais desafiadores. Um exemplo disso seria o uso de Coeficientes Mel Cepstrais como parâmetro característico dos sinais de voz, bem como a identificação de locuções que apresentem variações como o regionalismo.

Outro importante trabalho seria o desenvolvimento do classificador em tempo real como também elaborar uma comparação de diferentes ferramentas, como por exemplo “Rough Sets” e Redes Neurais, a fim de comparar os resultados obtidos quando no reconhecimento de padrões.

APÊNDICE A

FAIXAS USADAS PARA REDEFINIR OS ATRIBUTOS OBSERVADOS

O uso de faixas de valores para a classificação de dados é de extrema importância para a Tercia das Computações Aproximadas, pois permite simplificar as análises usadas em Sistemas de Representação de Conhecimento. Como foi mencionada no item 2.3, tal simplificação permite uma redução inicial do SRC. A redução proposta quando os atributos observados de um dado duplicados, partindo de dados que se repetem, ou atributos em atributos. Assim, ao identificar os atributos e pontos eliminados do SRC.

É preciso ressaltar ainda que a quantidade de faixas de valores e a maneira escolhida para definir cada uma variável irá depender da necessidade da aplicação.

Os SRCs criados, a serem classificados para reduzir a primeira estrutura, os FI, se baseiam nas relações de faixas de frequência de frequência fundamental associadas.

Assim sendo, que se define a seguinte tabela para a partir de valores baseados em

de acordo com os gráficos apresentados nas figuras 2.4 e 2.5, que demonstram o

de acordo com as frequências, quando se compararmos locustos distintos. A comparação

de acordo com o mesmo método permite identificar um certo padrão de valores para as

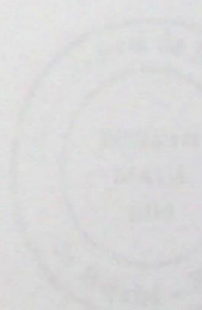
de acordo com o mesmo método, as quais apresentam algumas modificações e pouco desenvolvimento.

Os itens A.1, A.2, A.3, A.4, A.5, A.6, A.7 e A.8 apresentam as faixas de valores para

atributos criados.

A.1 - Faixas de Valores para FI

FI
1, representa a faixa para valores inferiores a 100
2, representa a faixa para valores de 101 a 120
3, representa a faixa para valores de 121 a 140
4, representa a faixa para valores de 141 a 160
5, representa a faixa para valores de 161 a 180
6, representa a faixa para valores de 181 a 200
7, representa a faixa para valores de 201 a 220
8, representa a faixa para valores de 221 a 240
9, representa a faixa para valores superiores a 240



APÊNDICE A

FAIXAS USADAS PARA REDEFINIR OS ATRIBUTOS OBSERVADOS

O uso de faixa de valores para a classificação de dados é de extrema importância para a Teoria dos Conjuntos Aproximados, pois permite simplificar as tabelas usadas como Sistemas de Representação de Conhecimento. Como foi mencionada no item 7.3, tal simplificação permite uma redução inicial do SRC. A redução acontece quando são identificados atributos de condição duplicados, padrões de dados que se repetem, ou atributos supérfluos. Assim, ao identificar tais casos é possível eliminá-los do SRC.

É preciso ressaltar ainda que a quantidade de faixa de valores e a métrica escolhida podem variar, onde essa variação irá depender da necessidade da aplicação.

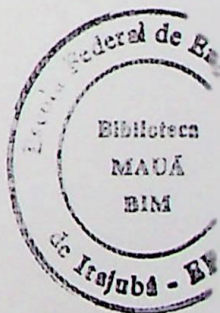
Nos SRCs em estudo, a métrica escolhida para redefinir o primeiro atributo, ou F1, se baseia em estudos relacionados às faixas de frequência da frequência fundamental masculina e feminina, sendo que as demais faixas foram escolhidas a partir de estudos baseados em gráficos tais como os gráficos apresentados pelas figuras 5.4 e 5.5, que demonstram o espectro dos sinais de voz dos locutores. Os gráficos permitem ainda identificar o deslocamento das frequências, quando se comparam locutores diferentes. A comparação de vários gráficos de um mesmo locutor permite identificar um certo padrão de valores para as frequências do locutor, as quais apresentam poucas modificações e pouco deslocamento.

Os itens A.1, A.2, A.3, A.4, A.5, A.6, A.7, e A.8 apresentam as faixas de valores para os atributos em estudo.

A.1 – Faixas de Valores para F1

Tabela A.1: Faixa de valores para o atributo F1

F1
1: representa a faixa para valores inferiores a 100
2: representa a faixa para valores de 101 a 120
3: representa a faixa para valores de 121 a 140
4: representa a faixa para valores de 141 a 160
5: representa a faixa para valores de 161 a 180
6: representa a faixa para valores de 181 a 200
7: representa a faixa para valores de 201 a 220
8: representa a faixa para valores de 221 a 240
9: representa a faixa para valores superiores a 241



A.2 – Faixas de Valores para F2

Tabela A.2: Faixa de valores para o atributo F2

F2
1 : representa a faixa para valores inferiores a 250
2 : representa a faixa para valores de 251 a 275
3 : representa a faixa para valores de 276 a 300
4 : representa a faixa para valores de 301 a 325
5 : representa a faixa para valores de 326 a 350
6 : representa a faixa para valores de 351 a 375
7 : representa a faixa para valores de 376 a 400
8 : representa a faixa para valores de 401 a 425
9 : representa a faixa para valores de 426 a 450
10 : representa a faixa para valores de 451 a 475
11 : representa a faixa para valores de 476 a 500
12 : representa a faixa para valores de 501 a 525
13 : representa a faixa para valores superiores a 526

A.3 – Faixas de Valores para F3

Tabela A.3: Faixa de valores para o atributo F3

F3
1 : representa a faixa para valores inferiores a 350
2 : representa a faixa para valores de 351 a 375
3 : representa a faixa para valores de 376 a 400
4 : representa a faixa para valores de 401 a 425
5 : representa a faixa para valores de 426 a 450
6 : representa a faixa para valores de 451 a 475
7 : representa a faixa para valores de 476 a 500
8 : representa a faixa para valores de 501 a 525
9 : representa a faixa para valores de 526 a 550
10 : representa a faixa para valores de 551 a 575
11 : representa a faixa para valores de 576 a 600
12 : representa a faixa para valores de 601 a 625
13 : representa a faixa para valores de 626 a 650
14 : representa a faixa para valores de 651 a 675
15 : representa a faixa para valores de 676 a 700
16 : representa a faixa para valores de 701 a 725
17 : representa a faixa para valores de 726 a 750
18 : representa a faixa para valores de 751 a 775
19 : representa a faixa para valores superiores a 776

A.4 – Faixas de Valores para F4

Tabela A.4: Faixa de valores para o atributo F4

F4
1 : representa a faixa para valores inferiores a 450
2 : representa a faixa para valores de 451 a 475
3 : representa a faixa para valores de 476 a 500
4 : representa a faixa para valores de 501 a 525
5 : representa a faixa para valores de 526 a 550
6 : representa a faixa para valores de 551 a 575
7 : representa a faixa para valores de 576 a 600
8 : representa a faixa para valores de 601 a 625
9 : representa a faixa para valores de 626 a 650
10 : representa a faixa para valores de 651 a 675
11 : representa a faixa para valores de 676 a 700
12 : representa a faixa para valores de 701 a 725
13 : representa a faixa para valores de 726 a 750
14 : representa a faixa para valores de 751 a 775
15 : representa a faixa para valores de 776 a 800
16 : representa a faixa para valores de 801 a 825
17 : representa a faixa para valores de 826 a 850
18 : representa a faixa para valores de 851 a 875
19 : representa a faixa para valores de 876 a 900
20 : representa a faixa para valores de 901 a 925
21 : representa a faixa para valores de 926 a 950
22 : representa a faixa para valores de 951 a 975
23 : representa a faixa para valores de 976 a 1000
24 : representa a faixa para valores de 1001 a 1025
25 : representa a faixa para valores superiores a 1026

A.5 – Faixas de Valores para F5

Tabela A.5: Faixa de valores para o atributo F5

F5
1 : representa a faixa para valores inferiores a 600
2 : representa a faixa para valores de 601 a 625
3 : representa a faixa para valores de 626 a 650
4 : representa a faixa para valores de 651 a 675
5 : representa a faixa para valores de 676 a 700
6 : representa a faixa para valores de 701 a 725
7 : representa a faixa para valores de 726 a 750
8 : representa a faixa para valores de 751 a 775
9 : representa a faixa para valores de 776 a 800
10 : representa a faixa para valores de 801 a 825
11 : representa a faixa para valores de 826 a 850
12 : representa a faixa para valores de 851 a 875
13 : representa a faixa para valores de 876 a 900
14 : representa a faixa para valores de 901 a 925
15 : representa a faixa para valores de 926 a 950
16 : representa a faixa para valores de 951 a 975
17 : representa a faixa para valores de 976 a 1000
18 : representa a faixa para valores de 1001 a 1025
19 : representa a faixa para valores de 1026 a 1050
20 : representa a faixa para valores de 1051 a 1075
21 : representa a faixa para valores de 1076 a 1100
22 : representa a faixa para valores de 1101 a 1125
23 : representa a faixa para valores de 1126 a 1150
24 : representa a faixa para valores de 1151 a 1175
25 : representa a faixa para valores de 1176 a 1200
26 : representa a faixa para valores de 1201 a 1225
27 : representa a faixa para valores de 1226 a 1250
28 : representa a faixa para valores de 1251 a 1275
29 : representa a faixa para valores superiores a 1276

A.6 – Faixas de Valores para F6

Tabela A.6: Faixa de valores para o atributo F6

F6
1 : representa a faixa para valores inferiores a 725
2 : representa a faixa para valores de 726 a 750
3 : representa a faixa para valores de 751 a 775
4 : representa a faixa para valores de 776 a 800
5 : representa a faixa para valores de 801 a 825
6 : representa a faixa para valores de 826 a 850
7 : representa a faixa para valores de 851 a 875
8 : representa a faixa para valores de 876 a 900
9 : representa a faixa para valores de 901 a 925
10 : representa a faixa para valores de 926 a 950
11 : representa a faixa para valores de 951 a 975
12 : representa a faixa para valores de 976 a 1000
13 : representa a faixa para valores de 1001 a 1025
14 : representa a faixa para valores de 1026 a 1050
15 : representa a faixa para valores de 1051 a 1075
16 : representa a faixa para valores de 1076 a 1100
17 : representa a faixa para valores de 1101 a 1125
18 : representa a faixa para valores de 1126 a 1150
19 : representa a faixa para valores de 1151 a 1175
20 : representa a faixa para valores de 1176 a 1200
21 : representa a faixa para valores de 1201 a 1225
22 : representa a faixa para valores de 1226 a 1250
23 : representa a faixa para valores de 1251 a 1275
24 : representa a faixa para valores de 1276 a 1300
25 : representa a faixa para valores de 1301 a 1325
26 : representa a faixa para valores de 1326 a 1350
27 : representa a faixa para valores de 1351 a 1375
28 : representa a faixa para valores de 1376 a 1400
29 : representa a faixa para valores de 1401 a 1425
30 : representa a faixa para valores de 1426 a 1450
31 : representa a faixa para valores de 1451 a 1475
32 : representa a faixa para valores de 1476 a 1500
33 : representa a faixa para valores de 1501 a 1525
34 : representa a faixa para valores de 1526 a 1550
35 : representa a faixa para valores de 1551 a 1575
36 : representa a faixa para valores superiores a 1576

A.7 – Faixas de Valores para F7

Tabela A.7: Faixa de valores para o atributo F7

F7
1 : representa a faixa para valores inferiores a 825
2 : representa a faixa para valores de 826 a 850
3 : representa a faixa para valores de 851 a 875
4 : representa a faixa para valores de 876 a 900
5 : representa a faixa para valores de 901 a 925
6 : representa a faixa para valores de 926 a 950
7 : representa a faixa para valores de 951 a 975
8 : representa a faixa para valores de 976 a 1000
9 : representa a faixa para valores de 1001 a 1025
10 : representa a faixa para valores de 1026 a 1050
11 : representa a faixa para valores de 1051 a 1075
12 : representa a faixa para valores de 1076 a 1100
13 : representa a faixa para valores de 1101 a 1125
14 : representa a faixa para valores de 1126 a 1150
15 : representa a faixa para valores de 1151 a 1175
16 : representa a faixa para valores de 1176 a 1200
17 : representa a faixa para valores de 1201 a 1225
18 : representa a faixa para valores de 1226 a 1250
19 : representa a faixa para valores de 1251 a 1275
20 : representa a faixa para valores de 1276 a 1300
21 : representa a faixa para valores de 1301 a 1325
22 : representa a faixa para valores de 1326 a 1350
23 : representa a faixa para valores de 1351 a 1375
24 : representa a faixa para valores de 1376 a 1400
25 : representa a faixa para valores de 1401 a 1425
26 : representa a faixa para valores de 1426 a 1450
27 : representa a faixa para valores de 1451 a 1475
28 : representa a faixa para valores de 1476 a 1500
29 : representa a faixa para valores de 1501 a 1525
30 : representa a faixa para valores de 1526 a 1550
31 : representa a faixa para valores de 1551 a 1575
32 : representa a faixa para valores de 1576 a 1600
33 : representa a faixa para valores de 1601 a 1625
34 : representa a faixa para valores de 1626 a 1650
35 : representa a faixa para valores de 1651 a 1675
36 : representa a faixa para valores de 1676 a 1700
37 : representa a faixa para valores de 1701 a 1725
38 : representa a faixa para valores de 1726 a 1750
39 : representa a faixa para valores de 1751 a 1775
40 : representa a faixa para valores de 1776 a 1800

A.8 – Faixas de Valores para Ffund

Tabela A.8: Faixa de valores para o atributo Ffund

Ffund
1: representa a faixa para valores inferiores a 100
2: representa a faixa para valores de 101 a 120
3: representa a faixa para valores de 121 a 140
4: representa a faixa para valores de 141 a 160
5: representa a faixa para valores de 161 a 180
6: representa a faixa para valores de 181 a 200
7: representa a faixa para valores de 201 a 220
8: representa a faixa para valores de 221 a 240
9: representa a faixa para valores superiores a 241

APÊNDICE B

SISTEMAS DE REPRESENTAÇÃO DE CONHECIMENTO COM ATRIBUTOS REDEFINIDOS

As tabelas apresentadas aqui representam os Sistemas de Representação de Conhecimento, porém com os valores dos atributos redefinidos segundo as tabelas incluídas no apêndice A e que contém as faixas de valores para os atributos.

Tabela B.1: SRC do conjunto de treinamento 1 com valores redefinidos

U	F1	F2	F3	F4	F5	F6	F7	Ffund	Saída
1	7	8	13	17	20	23	28	7	3
2	7	8	13	17	20	23	28	7	3
3	7	8	13	17	20	23	28	7	3
4	7	8	13	17	19	23	28	7	3
5	7	9	13	18	21	24	29	7	3
6	7	9	13	18	21	24	28	7	3
7	7	8	12	17	19	22	27	7	3
8	7	8	12	17	19	23	27	7	3
9	7	8	13	17	20	23	28	7	3
10	7	8	13	18	20	23	28	7	3
11	7	9	13	18	20	24	28	7	3
12	7	9	13	18	20	24	28	7	3
13	7	9	14	19	22	25	30	7	3
14	7	9	14	19	22	25	30	7	3
15	7	9	13	18	21	25	29	7	3
16	7	9	14	18	21	25	29	7	3
17	7	9	13	18	20	24	28	7	3
18	7	9	13	18	20	24	28	7	3
19	7	9	14	18	21	24	29	7	3
20	7	9	13	18	20	23	28	7	3
21	7	9	14	19	22	26	31	7	3
22	8	10	15	20	23	27	32	8	3
23	8	10	15	20	24	28	33	8	3
24	7	10	14	19	22	26	31	7	3
25	8	10	15	20	23	27	32	8	3
26	7	9	14	18	21	25	29	7	3
27	8	9	14	19	22	27	31	8	3
28	8	10	15	20	24	28	33	8	3
29	8	9	14	19	22	27	31	8	3
30	9	10	16	21	24	29	34	9	3
31	8	10	15	21	24	28	34	8	3
32	9	11	16	22	26	30	36	9	3
33	7	9	14	18	21	25	30	7	3
34	8	10	15	20	23	27	32	8	3
35	8	10	15	20	24	28	33	8	3
36	7	9	14	19	22	26	31	7	3
37	7	9	13	18	21	24	29	7	3
38	8	9	14	19	22	26	31	8	3
39	8	9	14	19	22	26	30	8	3

Tabela B.1: Continuação

40	7	9	14	18	21	25	29	7	3
41	4	2	4	5	5	5	7	4	4
42	3	2	3	5	4	4	6	3	4
43	3	2	3	5	4	4	6	3	4
44	3	2	3	4	4	4	5	3	4
45	3	2	3	4	4	4	5	3	4
46	3	2	3	4	3	4	5	3	4
47	3	2	4	5	4	5	6	3	4
48	4	2	4	5	5	5	7	4	4
49	4	3	5	6	6	6	8	4	4
50	4	3	4	6	6	6	8	4	4
51	4	3	4	6	6	6	8	4	4
52	4	2	4	6	5	6	7	4	4
53	4	3	4	5	5	6	7	4	4
54	4	2	4	5	5	5	7	4	4
55	3	2	4	5	4	5	6	3	4
56	3	2	3	5	4	4	6	3	4
57	4	4	6	8	8	9	11	4	4
58	4	4	6	8	8	9	12	4	4
59	4	4	6	8	8	9	11	4	4
60	4	4	6	8	8	9	10	4	4
61	3	2	3	5	4	4	6	3	4
62	3	2	3	4	4	4	5	3	4
63	3	2	4	5	4	5	6	3	4
64	3	2	3	5	4	4	6	3	4
65	4	4	6	8	8	9	11	4	4
66	4	4	6	8	8	9	12	4	4
67	3	2	3	5	4	4	6	3	4
68	3	2	3	4	4	4	5	3	4
69	3	2	3	4	4	4	5	3	4
70	3	2	3	4	4	4	5	3	4
71	3	2	3	4	3	4	5	3	4
72	3	2	4	5	4	5	6	3	4
73	4	2	4	5	5	5	7	4	4
74	4	3	5	6	6	6	8	4	4
75	4	3	4	6	6	6	8	4	4
76	4	3	4	6	6	6	8	4	4
77	3	2	3	4	4	4	5	3	4
78	3	2	3	4	4	4	5	3	4
79	4	2	4	5	5	5	7	4	4
80	4	3	5	6	6	6	8	4	4
81	2	1	2	2	1	1	1	2	5
82	2	1	2	3	2	1	2	2	5
83	2	1	2	3	2	2	2	2	5
84	3	1	2	3	2	1	2	3	5
85	2	1	2	3	1	1	2	2	5
86	2	1	2	2	1	1	2	2	5
87	2	1	2	2	1	1	6	2	5
88	2	1	2	3	2	1	2	2	5
89	4	1	2	8	6	6	7	4	5
90	3	1	2	3	2	2	2	3	5
91	2	1	2	3	2	2	3	2	5
92	2	1	2	3	1	1	2	2	5
93	3	1	2	3	2	2	3	3	5
94	3	1	2	3	2	2	3	3	5
95	3	1	2	3	2	2	3	3	5
96	3	1	4	4	3	2	4	3	5

Tabela B.1: Continuação

97	3	1	2	3	2	2	3	3	5
98	3	1	2	4	3	3	4	3	5
99	3	1	3	3	2	2	3	3	5
100	3	1	2	3	2	2	4	3	5
101	2	1	2	3	2	2	2	2	5
102	3	1	2	3	2	1	2	3	5
103	2	1	2	3	1	1	2	2	5
104	2	1	2	2	1	1	2	2	5
105	2	1	2	3	1	1	2	2	5
106	3	1	2	3	2	2	3	3	5
107	2	1	2	3	2	1	2	2	5
108	2	1	2	3	2	2	2	2	5
109	2	1	2	3	2	2	3	2	5
110	2	1	2	3	1	1	2	2	5
111	3	1	2	3	2	2	3	3	5
112	3	1	2	3	2	2	3	3	5
113	2	1	2	2	1	1	1	2	5
114	2	1	2	3	2	1	2	2	5
115	2	1	2	3	2	2	2	2	5
116	3	1	2	3	2	1	2	3	5
117	2	1	2	3	2	1	2	2	5
118	2	1	2	3	2	2	2	2	5
119	2	1	2	3	1	1	2	2	5
120	3	1	2	3	2	1	2	3	5

Tabela B.2: SRC do conjunto de treinamento 2 com valores redefinidos

U	F1	F2	F3	F4	F5	F6	F7	Ffund	Saída
1	7	8	12	17	19	22	28	7	1
2	7	8	12	17	19	22	28	7	1
3	7	8	12	17	19	23	28	7	1
4	7	8	12	17	19	22	28	7	1
5	9	11	17	23	27	32	38	9	1
6	9	11	17	23	27	31	37	9	1
7	9	11	17	23	27	32	38	9	1
8	9	11	17	24	28	33	38	9	1
9	7	8	12	17	19	23	28	7	1
10	7	8	12	17	19	22	28	7	1
11	9	11	17	23	27	32	38	9	1
12	9	11	17	23	27	31	37	9	1
13	9	11	16	22	26	30	35	9	1
14	9	11	16	22	25	30	36	9	1
15	9	11	16	22	25	30	36	9	1
16	9	11	16	22	26	30	36	9	1
17	9	11	16	22	26	30	36	9	1
18	9	11	17	22	26	31	36	9	1
19	9	11	17	23	26	31	36	9	1
20	9	11	17	22	26	31	36	9	1
21	8	10	15	21	24	28	33	8	1
22	8	10	16	21	22	29	34	8	1
23	8	10	16	21	25	29	35	8	1
24	8	10	16	21	24	28	34	8	1
25	9	11	16	22	26	30	36	9	1

Tabela B.2: Continuação

26	9	11	17	23	27	31	37	9	1
27	9	11	17	23	27	31	37	9	1
28	9	11	17	23	27	32	38	9	1
29	9	11	18	24	28	33	39	9	1
30	9	12	18	24	28	33	39	9	1
31	9	11	18	24	28	33	39	9	1
32	9	12	18	24	28	33	40	9	1
33	9	12	18	24	29	34	40	9	1
34	9	12	18	24	29	34	40	9	1
35	9	12	18	24	28	34	40	9	1
36	9	12	18	25	29	34	40	9	1
37	9	12	18	24	29	34	40	9	1
38	9	12	18	24	28	34	40	9	1
39	9	12	18	24	29	34	40	9	1
40	9	12	18	24	28	34	40	9	1
41	7	9	14	18	21	25	30	7	2
42	7	9	13	18	21	24	29	7	2
43	7	9	13	18	20	24	28	7	2
44	7	9	13	18	21	24	29	7	2
45	7	9	14	18	21	25	30	7	2
46	7	9	14	18	22	25	30	7	2
47	7	9	14	19	21	25	30	7	2
48	7	9	13	18	21	24	29	7	2
49	7	9	14	18	21	25	30	7	2
50	7	9	14	18	21	25	29	7	2
51	7	9	14	18	21	25	29	7	2
52	7	9	14	18	21	25	30	7	2
53	7	9	14	19	22	26	31	7	2
54	8	9	14	19	22	26	31	8	2
55	7	9	14	19	22	26	31	7	2
56	8	9	14	19	22	26	31	8	2
57	8	10	15	20	23	27	32	8	2
58	8	9	14	19	22	26	32	8	2
59	8	10	14	19	22	26	31	8	2
60	8	10	15	20	23	27	32	8	2
61	8	9	14	19	22	26	31	8	2
62	8	9	14	19	22	26	31	8	2
63	8	9	14	19	22	27	31	8	2
64	8	10	15	20	23	27	32	8	2
65	8	10	15	20	23	27	32	8	2
66	8	10	14	19	22	27	31	8	2
67	8	10	15	20	23	27	33	8	2
68	8	10	15	20	23	27	32	8	2
69	8	10	15	20	23	27	32	8	2
70	8	10	15	20	23	27	32	8	2
71	8	10	15	20	23	27	32	8	2
72	8	10	15	20	23	27	32	8	2
73	8	10	15	20	24	28	33	8	2
74	8	10	15	20	23	27	33	8	2
75	8	10	15	20	23	27	32	8	2
76	8	10	15	19	22	27	32	8	2
77	7	9	14	18	21	25	30	7	2

Tabela B.2: Continuação

78	7	9	13	18	21	24	29	7	2
79	7	9	13	18	20	24	28	7	2
80	7	9	13	18	21	24	29	7	2

Tabela B.3: SRC do conjunto de treinamento 3 com valores redefinidos

U	F1	F2	F3	F4	F5	F6	F7	Ffund	Saída
1	7	8	12	17	19	22	28	7	1
2	7	8	12	17	19	22	28	7	1
3	7	8	12	17	19	23	28	7	1
4	7	8	12	17	19	22	28	7	1
5	9	11	17	23	27	32	38	9	1
6	9	11	17	23	27	31	37	9	1
7	9	11	17	23	27	32	38	9	1
8	9	11	17	24	28	33	38	9	1
9	7	8	12	17	19	23	28	7	1
10	7	8	12	17	19	22	28	7	1
11	9	11	17	23	27	32	38	9	1
12	9	11	17	23	27	31	37	9	1
13	9	11	16	22	26	30	35	9	1
14	9	11	16	22	25	30	36	9	1
15	9	11	16	22	25	30	36	9	1
16	9	11	16	22	26	30	36	9	1
17	9	11	16	22	26	30	36	9	1
18	9	11	17	22	26	31	36	9	1
19	9	11	17	23	26	31	36	9	1
20	9	11	17	22	26	31	36	9	1
21	8	10	15	21	24	28	33	8	1
22	8	10	16	21	22	29	34	8	1
23	8	10	16	21	25	29	35	8	1
24	8	10	16	21	24	28	34	8	1
25	9	11	16	22	26	30	36	9	1
26	9	11	17	23	27	31	37	9	1
27	9	11	17	23	27	32	38	9	1
28	9	11	18	24	28	33	39	9	1
29	9	12	18	24	28	33	39	9	1
30	9	11	18	24	28	33	39	9	1
31	9	12	18	24	28	33	40	9	1
32	9	12	18	24	29	34	40	9	1
33	9	12	18	24	29	34	40	9	1
34	9	12	18	24	28	34	40	9	1
35	9	12	18	25	29	34	40	9	1
36	9	12	18	24	29	34	40	9	1
37	9	12	18	24	28	34	40	9	1
38	9	12	18	24	29	34	40	9	1
39	9	12	18	24	28	34	40	9	1
40	7	9	14	18	21	25	30	7	2
41	7	9	13	18	21	24	29	7	2
42	7	9	13	18	20	24	28	7	2
43	7	9	13	18	21	24	29	7	2
44	7	9	14	18	21	25	30	7	2
45	7	9	14	18	22	25	30	7	2
46	7	9	14	19	21	25	30	7	2
47	7	9	13	18	21	24	29	7	2
48	7	9	14	18	21	25	30	7	2

Tabela B.3: Continuação

49	7	9	14	18	21	25	29	7	2
50	7	9	14	18	21	25	29	7	2
51	7	9	14	18	21	25	30	7	2
52	8	9	14	19	22	26	31	8	2
53	8	9	14	19	22	26	31	8	2
54	8	10	15	20	23	27	32	8	2
55	8	9	14	19	22	26	32	8	2
56	8	10	14	19	22	26	31	8	2
57	8	10	15	20	23	27	32	8	2
58	8	9	14	19	22	26	31	8	2
59	8	9	14	19	22	26	31	8	2
60	8	9	14	19	22	27	31	8	2
61	8	10	15	20	23	27	32	8	2
62	8	10	15	20	23	27	32	8	2
63	8	10	14	19	22	27	31	8	2
64	8	10	15	20	23	27	33	8	2
65	8	10	15	20	23	27	32	8	2
66	8	10	15	20	23	27	32	8	2
67	8	10	15	20	23	27	32	8	2
68	8	10	15	20	23	27	32	8	2
69	8	10	15	20	23	27	32	8	2
70	8	10	15	20	24	28	33	8	2
71	8	10	15	20	23	27	33	8	2
72	8	10	15	20	23	27	32	8	2
73	8	10	15	19	22	27	32	8	2
74	7	9	14	18	21	25	30	7	2
75	7	9	13	18	21	24	29	7	2
76	7	9	13	18	20	24	28	7	2
77	7	9	13	18	21	24	29	7	2
78	4	2	4	5	5	5	7	4	3
79	3	2	3	5	4	4	6	3	3
80	3	2	3	5	4	4	6	3	3
81	3	2	3	4	4	4	5	3	3
82	3	2	3	4	4	4	5	3	3
83	3	2	3	4	3	4	5	3	3
84	3	2	4	5	4	5	6	3	3
85	4	2	4	5	5	5	7	4	3
86	4	3	5	6	6	6	8	4	3
87	4	3	4	6	6	6	8	4	3
88	4	3	4	6	6	6	8	4	3
89	4	2	4	6	5	6	7	4	3
90	4	3	4	5	5	6	7	4	3
91	4	2	4	5	5	5	7	4	3
92	3	2	4	5	4	5	6	3	3
93	3	2	3	5	4	4	6	3	3
94	4	4	6	8	8	9	11	4	3
95	4	4	6	8	8	9	12	4	3
96	4	4	6	8	8	9	11	4	3
97	4	4	6	8	8	9	10	4	3
98	3	2	3	5	4	4	6	3	3
99	3	2	3	4	4	4	5	3	3
100	3	2	4	5	4	5	6	3	3
101	3	2	3	5	4	4	6	3	3
102	4	4	6	8	8	9	11	4	3
103	4	4	6	8	8	9	12	4	3
104	3	2	3	4	4	4	5	3	3
105	3	2	3	4	4	4	5	3	3

Tabela B.3: Continuação

106	3	2	3	4	3	4	5	3	3
107	3	2	4	5	4	5	6	3	3
108	4	2	4	5	5	5	7	4	3
109	4	3	5	6	6	6	8	4	3
110	4	3	4	6	6	6	8	4	3
111	4	3	4	6	6	6	8	4	3
112	3	2	3	4	4	4	5	3	3
113	3	2	3	4	4	4	5	3	3
114	4	2	4	5	5	5	7	4	3
115	4	3	5	6	6	6	8	4	3
116	2	1	2	2	1	1	1	2	3
117	2	1	2	3	2	1	2	2	3
118	2	1	2	3	2	2	2	2	3
119	3	1	2	3	2	1	2	3	3
120	2	1	2	3	1	1	2	2	3
121	2	1	2	3	2	1	2	2	3
122	4	1	2	8	6	6	7	4	3
123	3	1	2	3	2	2	2	3	3
124	2	1	2	3	2	2	3	2	3
125	2	1	2	3	1	1	2	2	3
126	3	1	2	3	2	2	3	3	3
127	3	1	2	3	2	2	3	3	3
128	3	1	2	3	2	2	3	3	3
129	3	1	4	4	3	2	4	3	3
130	3	1	2	3	2	2	3	3	3
131	3	1	2	4	3	3	4	3	3
132	3	1	3	3	2	2	3	3	3
133	3	1	2	3	2	2	4	3	3
134	2	1	2	3	2	2	2	2	3
135	3	1	2	3	2	1	2	3	3
136	2	1	2	3	1	1	2	2	3
137	2	1	2	3	1	1	2	2	3
138	3	1	2	3	2	2	3	3	3
139	2	1	2	3	2	1	2	2	3
140	2	1	2	3	2	2	2	2	3
141	2	1	2	3	2	2	3	2	3
142	2	1	2	3	1	1	2	2	3
143	3	1	2	3	2	2	3	3	3
144	3	1	2	3	2	2	3	3	3
145	2	1	2	2	1	1	1	2	3
146	2	1	2	3	2	1	2	2	3
147	2	1	2	3	2	2	2	2	3
148	3	1	2	3	2	1	2	3	3
149	2	1	2	3	2	1	2	2	3
150	2	1	2	3	2	2	2	2	3
151	2	1	2	3	1	1	2	2	3
152	3	1	2	3	2	1	2	3	3
153	3	2	3	5	4	5	6	3	3
154	3	2	3	5	4	4	6	3	3
155	3	2	4	5	4	5	6	3	3
156	3	2	2	5	4	4	5	3	3
157	3	2	2	4	2	3	4	3	3
158	3	1	3	4	3	3	4	3	3
159	3	2	3	4	3	3	5	3	3
160	3	2	3	4	3	3	4	3	3
161	4	2	4	6	5	6	7	4	3
162	4	2	4	6	5	6	7	4	3

Tabela B.3: Continuação

163	3	2	3	5	5	5	6	3	3
164	3	2	4	5	5	5	7	3	3
165	5	4	4	3	4	5	8	5	3
166	5	4	7	9	10	11	14	5	3
167	5	4	7	9	9	11	13	5	3
168	4	4	7	9	9	11	13	4	3
169	5	5	7	10	10	12	15	5	3
170	5	5	7	10	10	12	14	5	3
171	5	5	7	9	10	12	14	5	3
172	5	4	7	9	10	11	13	5	3
173	6	6	10	13	15	17	20	6	3
174	6	6	10	13	14	17	20	6	3
175	5	6	9	13	14	16	20	5	3
176	5	6	9	13	14	16	20	5	3
177	5	5	7	10	11	12	15	5	3
178	5	5	7	10	11	13	16	5	3
179	5	5	8	11	11	13	16	5	3
180	5	5	7	10	11	12	15	5	3
181	3	2	3	5	4	5	6	3	3
182	3	2	3	5	4	4	6	3	3
183	3	2	4	5	4	5	6	3	3
184	3	2	2	5	4	4	5	3	3
185	5	5	7	10	10	12	15	5	3
186	5	5	7	10	11	12	15	5	3
187	5	5	8	10	11	12	15	5	3
188	5	5	7	10	11	12	15	5	3
189	5	6	9	12	13	15	18	5	3
190	5	6	8	12	13	22	25	5	3
191	5	6	8	12	12	22	25	5	3
192	5	6	8	11	12	21	24	5	3
193	7	7	12	15	17	21	24	7	3
194	6	7	12	15	17	20	24	6	3
195	7	7	11	15	17	20	24	7	3
196	6	7	11	15	16	20	24	6	3
197	6	7	11	15	17	20	24	6	3
198	7	7	11	15	18	20	24	7	3
199	7	7	12	16	17	21	25	7	3
200	7	7	11	15	17	20	24	7	3
201	7	8	12	16	17	21	25	7	3
202	7	8	12	16	18	21	25	7	3
203	7	7	11	15	17	20	24	7	3
204	6	7	11	15	17	20	24	6	3
205	7	8	12	16	17	21	25	7	3
206	7	8	12	16	17	21	25	7	3
207	7	8	12	16	18	21	25	7	3
208	7	8	12	16	17	21	25	7	3
209	7	8	13	17	17	23	27	7	3
210	7	8	13	17	17	23	27	7	3
211	7	8	12	17	17	23	27	7	3
212	7	8	13	17	18	23	27	7	3
213	7	8	12	16	18	22	26	7	3
214	7	8	12	16	18	21	25	7	3
215	7	8	12	16	19	22	26	7	3
216	7	8	12	16	18	22	26	7	3
217	7	7	11	15	17	20	24	7	3
218	6	7	12	15	18	20	24	6	3
219	6	7	11	15	17	20	24	6	3

Tabela B.3: Continuação

220	6	7	11	15	18	20	24	6	3
221	7	8	13	17	19	23	27	7	3
222	6	7	11	14	16	19	23	6	3
223	6	7	10	14	16	18	22	6	3
224	6	7	11	14	16	19	22	6	3
225	6	7	11	14	16	19	23	6	3
226	6	7	11	15	17	20	24	6	3
227	6	7	11	15	17	20	23	6	3
228	6	7	11	15	17	20	24	6	3
229	6	7	11	15	17	20	24	6	3

APÊNDICE C

REGRAS OBTIDAS PELOS CONJUNTOS DE TREINAMENTO

As tabelas apresentadas aqui representam as regras encontradas após a simulação dos conjuntos de treinamento. As tabelas apresentam a mesma estrutura, onde a primeira coluna representa o índice de identificação da regra, as colunas identificadas por F1, F2, F3, F4, F5, F6 e F7 representam as características dos padrões e a última coluna, identificada por Saída, representa o código de referência do locutor.

A leitura das regras, que representam seqüências executadas condicionalmente, é realizada como mostra o exemplo abaixo.

Exemplo C.1: tomando como exemplo a regra de índice igual a 1 da tabela C.1, tem-se como seqüência a seguinte condição:

$$\text{Se } F5 = 20 \text{ e } F7 = 28 \text{ Então Saída} = 3$$

Tabela C.1: Total de regras do conjunto de treinamento 1

Índice	F1	F2	F3	F4	F5	F6	F7	Saída
1					20		28	3
2					19		28	3
3					21		29	3
4					21		28	3
5					19		27	3
6					22		30	3
7					22		31	3
8					23		32	3
9					24		33	3
10					24		34	3
11					26		36	3
12					21		30	3
13					5		7	4
14					4		6	4
15					4		5	4
16					3		5	4
17					6		8	4
18					8		11	4
19					8		12	4
20					8		10	4
21					1		1	5
22					2		2	5
23					1		2	5
24					1		6	5
25					6		7	5
26					2		3	5
27					3		4	5
28					2		4	5
29				17			28	3
30				18			29	3

Tabela C.1: Continuação

31			18		28	3
32			17		27	3
33			19		30	3
34			19		31	3
35			20		32	3
36			20		33	3
37			21		34	3
38			22		36	3
39			18		30	3
40			5		7	4
41			5		6	4
42			4		5	4
43			6		8	4
44			6		7	4
45			8		11	4
46			8		12	4
47			8		10	4
48			2		1	5
49			3		2	5
50			2		2	5
51			2		6	5
52			8		7	5
53			3		3	5
54			4		4	5
55			3		4	5
56			17	23		3
57			18	24		3
58			17	22		3
59			18	23		3
60			19	25		3
61			18	25		3
62			19	26		3
63			20	27		3
64			20	28		3
65			19	27		3
66			21	29		3
67			21	28		3
68			22	30		3
69			5	5		4
70			5	4		4
71			4	4		4
72			6	6		4
73			5	6		4
74			8	9		4
75			2	1		5
76			3	1		5
77			3	2		5
78			8	6		5
79			4	2		5
80			4	3		5
81		13			28	3
82		13			29	3
83		12			27	3
84		14			30	3
85		14			29	3
86		14			31	3
87		15			32	3

Tabela C.1: Continuação

88			15				33	3
89			16				34	3
90			15				34	3
91			16				36	3
92			4				7	4
93			3				6	4
94			3				5	4
95			4				6	4
96			5				8	4
97			4				8	4
98			6				11	4
99			6				12	4
100			6				10	4
101			2				1	5
102			2				2	5
103			2				6	5
104			2				7	5
105			2				3	5
106			4				4	5
107			2				4	5
108			3				3	5
109			13			23		3
110			13			24		3
111			12			22		3
112			12			23		3
113			14			25		3
114			13			25		3
115			14			24		3
116			14			26		3
117			15			27		3
118			15			28		3
119			14			27		3
120			16			29		3
121			16			30		3
122			4			5		4
123			3			4		4
124			5			6		4
125			4			6		4
126			6			9		4
127			2			1		5
128			2			2		5
129			2			6		5
130			4			2		5
131			2			3		5
132			3			2		5
133			13		20			3
134			13		19			3
135			13		21			3
136			12		19			3
137			14		22			3
138			14		21			3
139			15		23			3
140			15		24			3
141			16		24			3
142			16		26			3
143			4		5			4
144			3		4			4

Tabela C.1: Continuação

145			3		3			4
146			4		4			4
147			5		6			4
148			4		6			4
149			6		8			4
150			2		1			5
151			2		2			5
152			2		6			5
153			4		3			5
154			2		3			5
155			3		2			5
156			13	17				3
157			13	18				3
158			12	17				3
159			14	19				3
160			14	18				3
161			15	20				3
162			16	21				3
163			15	21				3
164			16	22				3
165			4	5				4
166			3	5				4
167			3	4				4
168			5	6				4
169			4	6				4
170			6	8				4
171			2	2				5
172			2	3				5
173			2	8				5
174			4	4				5
175			2	4				5
176			3	3				5
177		8					28	3
178		9					29	3
179		9					28	3
180		8					27	3
181		9					30	3
182		9					31	3
183		10					32	3
184		10					33	3
185		10					31	3
186		10					34	3
187		11					36	3
188		2					7	4
189		2					6	4
190		2					5	4
191		3					8	4
192		3					7	4
193		4					11	4
194		4					12	4
195		4					10	4
196		1					1	5
197		1					2	5
198		1					6	5
199		1					7	5
200		1					3	5
201		1					4	5

Tabela C.1: Continuação

202		8				23		3
203		9				24		3
204		8				22		3
205		9				25		3
206		9				23		3
207		9				26		3
208		10				27		3
209		10				28		3
210		10				26		3
211		9				27		3
212		10				29		3
213		11				30		3
214		2				5		4
215		2				4		4
216		3				6		4
217		2				6		4
218		4				9		4
219		1				1		5
220		1				2		5
221		1				6		5
222		1				3		5
223		8			20			3
224		8			19			3
225		9			21			3
226		9			20			3
227		9			22			3
228		10			23			3
229		10			24			3
230		10			22			3
231		11			26			3
232		2			5			4
233		2			4			4
234		2			3			4
235		3			6			4
236		3			5			4
237		4			8			4
238		1			1			5
239		1			2			5
240		1			6			5
241		1			3			5
242		8		17				3
243		9		18				3
244		8		18				3
245		9		19				3
246		10		20				3
247		10		19				3
248		10		21				3
249		11		22				3
250		2		5				4
251		2		4				4
252		3		6				4
253		2		6				4
254		3		5				4
255		4		8				4
256		1		2				5
257		1		3				5
258		1		8				5

Tabela C.1: Continuação

259		1	4			5
260		8	13			3
261		9	13			3
262		8	12			3
263		9	14			3
264		10	15			3
265		10	14			3
266		10	16			3
267		11	16			3
268		2	4			4
269		2	3			4
270		3	5			4
271		3	4			4
272		4	6			4
273		1	2			5
274		1	4			5
275		1	3			5
276	7	8				3
277	7	9				3
278	8	10				3
279	7	10				3
280	8	9				3
281	9	10				3
282	9	11				3
283	4	2				4
284	3	2				4
285	4	3				4
286	4	4				4
287	2	1				5
288	3	1				5
289	4	1				5
290	2					5
291	7					3
292	8					3
293	9					3
294		1				5
295		2				4
296		3				4
297		4				4
298		8				3
299		9				3
300		10				3
301		11				3
302			2			5
303			5			4
304			6			4
305			12			3
306			13			3
307			14			3
308			15			3
309			16			3
310			2			5
311			3			5
312			5			4
313			6			4
314			17			3
315			18			3

Tabela C.1: Continuação

316				19				3
317				20				3
318				21				3
319				22				3
320					1			5
321					2			5
322					4			4
323					5			4
324					8			4
325						1		5
326						2		5
327						3		5
328						4		4
329						5		4
330						9		4
331						22		3
332						23		3
333						24		3
334						25		3
335						26		3
336						27		3
337						28		3
338						29		3
339						30		3
340							1	5
341							2	5
342							3	5
343							4	5
344							5	4
345							8	4
346							10	4
347							11	4
348							12	4
349							27	3
350							28	3
351							29	3
352							30	3
353							31	3
354							32	3
355							33	3
356							34	3
357							36	3

Tabela C.2: Total de regras do conjunto de treinamento 2

Índice	F1	F2	F3	F4	F5	F6	F7	Saída
1				18			30	2
2				18			29	2
3				18			28	2
4				19			30	2
5				19			31	2
6				20			32	2
7				19			32	2
8				20			33	2
9				17			28	1
10				23			38	1

Tabela C.2: Continuação

11				23			37	1
12				24			38	1
13				22			35	1
14				22			36	1
15				23			36	1
16				21			33	1
17				21			34	1
18				21			35	1
19				24			39	1
20				24			40	1
21				25			40	1
22				18		25		2
23				18		24		2
24				19		25		2
25				19		26		2
26				20		27		2
27				19		27		2
28				20		28		2
29				17		22		1
30				17		23		1
31				23		32		1
32				23		31		1
33				24		33		1
34				22		30		1
35				22		31		1
36				21		28		1
37				21		29		1
38				24		34		1
39				25		34		1
40				18	21			2
41				18	20			2
42				18	22			2
43				19	21			2
44				19	22			2
45				20	23			2
46				20	24			2
47				17	19			1
48				23	27			1
49				24	28			1
50				22	26			1
51				22	25			1
52				23	26			1
53				21	24			1
54				21	22			1
55				21	25			1
56				24	29			1
57				25	29			1
58			14	18				2
59			13	18				2
60			14	19				2
61			15	20				2
62			15	19				2
63			12	17				1
64			17	23				1
65			17	24				1
66			16	22				1
67			17	22				1



Tabela C.2: Continuação

68			15	21			1
69			16	21			1
70			18	24			1
71			18	25			1
72		9		18			2
73		9		19			2
74		10		20			2
75		10		19			2
76		8		17			1
77		11		23			1
78		11		24			1
79		11		22			1
80		10		21			1
81		12		24			1
82		12		25			1
83	7			18			2
84	7			19			2
85	8			19			2
86	8			20			2
87	7			17			1
88	9			23			1
89	9			24			1
90	9			22			1
91	8			21			1
92	9			25			1
93	9						1
94		9					2
95		8					1
96		11					1
97		12					1
98			13				2
99			14				2
100			12				1
101			16				1
102			17				1
103			18				1
104				17			1
105				18			2
106				19			2
107				20			2
108				21			1
109				22			1
110				23			1
111				24			1
112				25			1
113					19		1
114					20		2
115					21		2
116					22		2
117					23		2
118					25		1
119					26		1
120					27		1
121					28		1
122					29		1
123						22	1
124						23	1

Tabela C.2: Continuação

125						24		2
126						25		2
127						26		2
128						27		2
129						29		1
130						30		1
131						31		1
132						33		1
133						34		1
134							29	2
135							30	2
136							31	2
137							32	2
138							33	1
139							34	1
140							35	1
141							36	1
142							37	1
143							38	1
144							39	1
145							40	1

Tabela C.3: Total de regras do conjunto de treinamento 3

Índice	F1	F2	F3	F4	F5	F6	F7	Saída
1			12					1
2			17					1
3			16					1
4			18					1
5			14					2
6			4					3
7			3					3
8			5					3
9			6					3
10			2					3
11			7					3
12			10					3
13			9					3
14			8					3
15			11					3
16				17				1
17				23				1
18				24				1
19				22				1
20				21				1
21				25				1
22				18				2
23				19				2
24				20				2
25				5				3
26				4				3
27				6				3
28				8				3
29				2				3
30				3				3
31				9				3

Tabela C.3: Continuação

32			10				3
33			13				3
34			11				3
35			12				3
36			15				3
37			16				3
38			14				3
39				19			1
40				27			1
41				28			1
42				26			1
43				25			1
44				22			1
45				29			1
46				21			2
47				20			2
48				22			2
49				23			2
50				5			3
51				4			3
52				3			3
53				6			3
54				8			3
55				1			3
56				2			3
57				10			3
58				9			3
59				15			3
60				14			3
61				11			3
62				13			3
63				12			3
64				17			3
65				16			3
66				18			3
67					28		1
68					38		1
69					37		1
70					35		1
71					36		1
72					33		1
73					34		1
74					39		1
75					40		1
76					30		2
77					29		2
78					31		2
79					32		2
80					33		2
81					7		3
82					6		3
83					5		3
84					8		3
85					11		3
86					12		3
87					10		3
88					1		3

Tabela C.3: Continuação

89						2	3
90						3	3
91						4	3
92						14	3
93						13	3
94						15	3
95						20	3
96						16	3
97						18	3
98						25	3
99						24	3
100						27	3
101						26	3
102						23	3
103						22	3
104			17			28	1
105			23			38	1
106			23			37	1
107			24			38	1
108			22			35	1
109			22			36	1
110			23			36	1
111			21			33	1
112			21			34	1
113			21			35	1
114			24			39	1
115			24			40	1
116			25			40	1
117			18			30	2
118			18			29	2
119			18			28	2
120			19			30	2
121			19			31	2
122			20			32	2
123			19			32	2
124			20			33	2
125			5			7	3
126			5			6	3
127			4			5	3
128			6			8	3
129			6			7	3
130			8			11	3
131			8			12	3
132			8			10	3
133			2			1	3
134			3			2	3
135			8			7	3
136			3			3	3
137			4			4	3
138			3			4	3
139			5			5	3
140			3			8	3
141			9			14	3
142			9			13	3
143			10			15	3
144			10			14	3
145			13			20	3

Tabela C.3: Continuação

146				10			16	3
147				11			16	3
148				12			18	3
149				12			25	3
150				11			24	3
151				15			24	3
152				16			25	3
153				17			27	3
154				16			26	3
155				14			23	3
156				14			22	3
157				15			23	3
158			12	17	19			1
159			17	23	27			1
160			17	24	28			1
161			16	22	26			1
162			16	22	25			1
163			17	22	26			1
164			17	23	26			1
165			15	21	24			1
166			16	21	22			1
167			16	21	25			1
168			16	21	24			1
169			18	24	28			1
170			18	24	29			1
171			18	25	29			1
172			14	18	21			2
173			13	18	21			2
174			13	18	20			2
175			14	18	22			2
176			14	19	21			2
177			14	19	22			2
178			15	20	23			2
179			15	20	24			2
180			15	19	22			2
181			4	5	5			3
182			3	5	4			3
183			3	4	4			3
184			3	4	3			3
185			4	5	4			3
186			5	6	6			3
187			4	6	6			3
188			4	6	5			3
189			6	8	8			3
190			2	2	1			3
191			2	3	2			3
192			2	3	1			3
193			2	8	6			3
194			4	4	3			3
195			2	4	3			3
196			3	3	2			3
197			2	5	4			3
198			2	4	2			3
199			3	5	5			3
200			4	3	4			3
201			7	9	10			3
202			7	9	9			3

Tabela C.3: Continuação

203			7	10	10			3
204			10	13	15			3
205			10	13	14			3
206			9	13	14			3
207			7	10	11			3
208			8	11	11			3
209			8	10	11			3
210			9	12	13			3
211			8	12	13			3
212			8	12	12			3
213			8	11	12			3
214			12	15	17			3
215			11	15	17			3
216			11	15	16			3
217			11	15	18			3
218			12	16	17			3
219			12	16	18			3
220			13	17	17			3
221			12	17	17			3
222			13	17	18			3
223			12	16	19			3
224			12	15	18			3
225			13	17	19			3
226			11	14	16			3
227			10	14	16			3

APÊNDICE D

REGRAS MÍNIMAS DOS CONJUNTOS DE TREINAMENTO

Tabela D.1: Regras mínimas do conjunto de treinamento 1

Índice	F1	F2	F3	F4	F5	F6	F7	S
1	2							5
2	3	1						5
3	3	2						4
4	4	1						5
5	4	2						4
6	4	3						4
7	4	4						4
8	7							3
9	8							3
10	9							3
11		1						5
12		2						4
13		3						4
14		4						4
15		8						3
16		9						3
17		10						3
18		11						3
19			2					5
20			3				6	4
21			3		3			4
22			3	4				4
23			4				7	4
24			4				6	4
25			4			6		4
26			4		6			4
27			4		3			5
28			4	4				5
29			5					4
30			6					4
31			12					3
32			13					3
33			14					3
34			15					3
35			16					3
36					6		7	5
37				8			7	5
38				8		6		5
39				2				5
40				3				5
41				5				4
42				6				4
43				17				3
44				18				3
45				19				3
46				20				3
47				21				3
48				22				3
49					1			5

Tabela D.1: Continuação

50					2			5
51					4			4
52					5			4
53					8			4
54						1		5
55						2		5
56						3		5
57						4		4
58						5		4
59						9		4
60						22		3
61						23		3
62						24		3
63						25		3
64						26		3
65						27		3
66						28		3
67						29		3
68						30		3
69							1	5
70							2	5
71							3	5
72							4	5
73							5	4
74							8	4
75							10	4
76							11	4
77							12	4
78							27	3
79							28	3
80							29	3
81							30	3
82							31	3
83							32	3
84							33	3
85							34	3
86							36	3

Tabela D.2: Regras mínimas do conjunto de treinamento 2

Índice	F1	F2	F3	F4	F5	F6	F7	Saída
1				17				1
2				18				2
3				19				2
4				20				2
5				21				1
6				22				1
7				23				1
8				24				1
9				25				1
10							29	2
11							30	2
12							31	2
13							32	2
14							34	1
15							35	1

Tabela D.2: Continuação

16							36	1
17							37	1
18							38	1
19							39	1
20							40	1
21						22		1
22						23		1
23						32		1
24						31		1
25						33		1
26						30		1
27						29		1
28						34		1
29						25		2
30						24		2
31						26		2
32						27		2
33					19			1
34					20			2
35					21			2
36					23			2
37					25			1
38					26			1
39					27			1
40					28			1
41					29			1
42			12					1
43			13					2
44			14					2
45			16					1
46			17					1
47			18					1
48		8						1
49		9						2
50		11						1
51		12						1
52	9							1

Tabela D.3: Regras mínimas do conjunto de treinamento 3

Índice	F1	F2	F3	F4	F5	F6	F7	Saída
1			2					3
2			3					3
3			4					3
4			5					3
5			6					3
6			7					3
7			8					3
8			9					3
9			10					3
10			11					3
11			12	17	19			1
12			12	15	17			3
13			12	16	17			3
14			12	16	18			3
15			12	17	17			3

Tabela D.3: Continuação

16			12	16	19			3
17			12	15	18			3
18			13					2
19			14					2
20			15	21	24			1
21			15	20	23			2
22			15	20	24			2
23			15	19	22			2
24			16					1
25			17					1
26			18					1
27				2				3
28				3				3
29				4				3
30				5				3
31				6				3
32				8				3
33				9				3
34				10				3
35				11				3
36				12				3
37				13				3
38				14				3
39				15				3
40				16				3
41			13	17	17			3
42			13	17	18			3
43			13	17	19			3
44				17		28		1
45				17		27		3
46				19				2
47				20				2
48				21				1
49				22				1
50				23				1
51				24				1
52				25				1
53					1			3
54					2			3
55					3			3
56					4			3
57					5			3
58					6			3
59					8			3
60					9			3
61					10			3
62					11			3
63					12			3
64					13			3
65					14			3
66					15			3
67					16			3
68					17			3
69					18			3
70					20			2
71					21			2
72			16	21	22			1

Tabela D.3: Continuação

73			14	18	22			2
74			14	19	22			2
75					23			2
76			16	21	24			1
77					25			1
78					26			1
79					27			1
80					28			1
81					29			1
82							1	3
83							2	3
84							3	3
85							4	3
86							5	3
87							6	3
88							7	3
89							8	3
90							10	3
91							11	3
92							12	3
93							13	3
94							14	3
95							15	3
96							16	3
97							18	3
98							20	3
99							22	3
100							23	3
101							24	3
102							25	3
103							26	3
104							27	3
105				18			28	2
106							29	2
107							30	2
108							31	2
109							32	2
110				21			33	1
111				20			33	2
112							34	1
113							35	1
114							36	1
115							37	1
116							38	1
117							39	1
118							40	1

APÊNDICE E

O SISTEMA

O sistema desenvolvido possui uma tela principal, ilustrada pela figura E.1, com opções que permitem a extração das características dos sinais de voz assim como a redefinição de faixas da tabela de dados, e a eliminação de informação redundante, que devem ser acionadas antes da aplicação dos Conjuntos Aproximados.

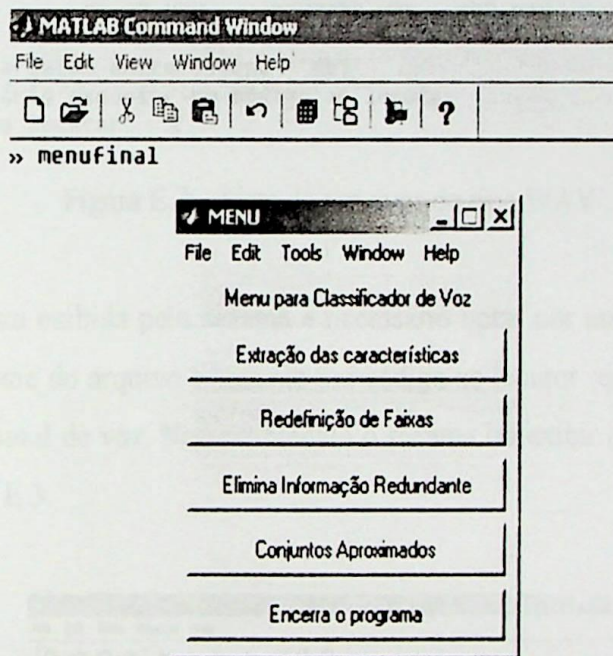


Figura E.1 – Tela Principal do Sistema

Ao optar pela extração das características, o sistema irá listar os nomes dos arquivos do tipo WAV que contém os sinais de voz, como mostra a figura E.2.

Deseja armazenar um banco de dados existente? (s/n) : n
 Deseja incluir informações no banco de dados? (s/n) : s
 Cálculo do espectro do sinal de voz

andso1.wav	j5.wav	marilia1.wav	mauro55.wav	zz1.wav
andso10.wav	j6.wav	marilia10.wav	mauro66.wav	zz10.wav
andso2.wav	j7.wav	marilia2.wav	mauro77.wav	zz2.wav
andso3.wav	j8.wav	marilia3.wav	mauro88.wav	zz3.wav
andso4.wav	j9.wav	marilia4.wav	mauro99.wav	zz4.wav
andso5.wav	marc1.wav	marilia5.wav	sb1.wav	zz5.wav
andso6.wav	marc10.wav	marilia6.wav	sb10.wav	zz6.wav
andso7.wav	marc2.wav	marilia7.wav	sb2.wav	zz7.wav
andso8.wav	marc3.wav	marilia8.wav	sb3.wav	zz8.wav
andso9.wav	marc4.wav	marilia9.wav	sb4.wav	zz9.wav
j1.wav	marc5.wav	mauro1010.wav	sb5.wav	
j10.wav	marc6.wav	mauro11.wav	sb6.wav	
j2.wav	marc7.wav	mauro22.wav	sb7.wav	
j3.wav	marc8.wav	mauro33.wav	sb8.wav	
j4.wav	marc9.wav	mauro44.wav	sb9.wav	

Nome do arquivo entre aspas: 'sb1'
 É necessário associar um código ao locutor
 Código do locutor: 1|

Figura E.2 – Lista de arquivos do tipo WAV

A partir da lista exibida pelo sistema é necessário optar por um dos arquivos. Para isso basta mencionar o nome do arquivo e associar um código ao locutor responsável pela pronuncia do som referente ao sinal de voz. Nesse momento o sistema irá exibir o gráfico do sinal de voz, como mostra a figura E.3.

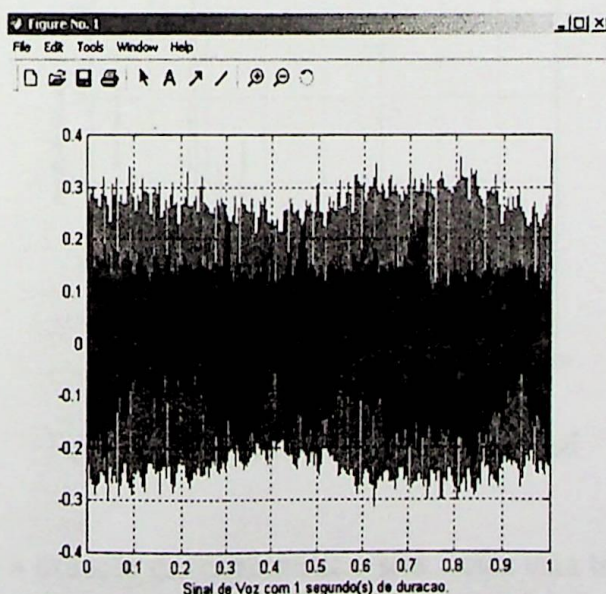


Figura E.3 – Exibição do sinal de voz

Para cada sinal haverá uma segmentação, como mostra a figura E.4, para a extração das características contidas no mesmo, como ilustra a figura E.5.

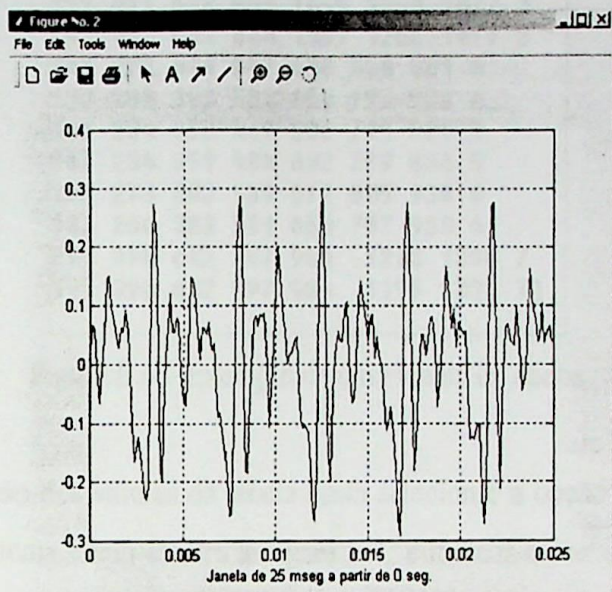


Figura E.4 – Sinal segmentado

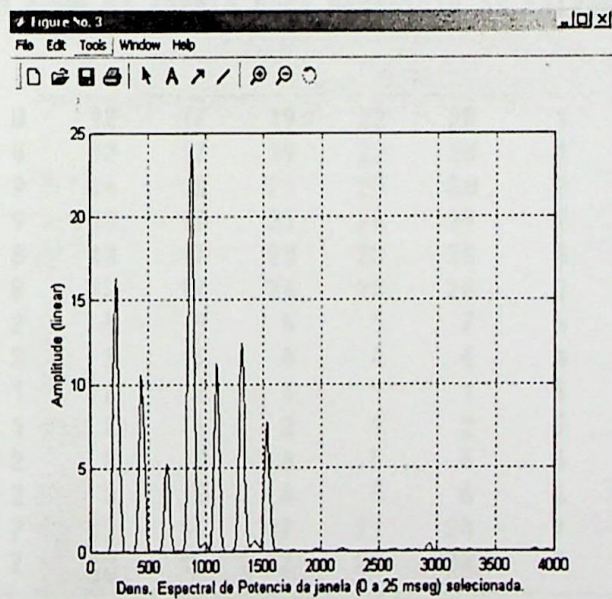


Figura E.5 – Características contida no sinal

Assim, durante a extração das características será criada uma tabela de dados contendo um conjunto de informações que representam as características contidas no sinal de voz. Um exemplo de tabela de dados é mostrado pela figura E.6.


```

>> dados=[
211 414 617 828 1039 1242 1484 1
211 414 625 836 1047 1250 1484 1
219 438 656 875 1094 1313 1531 2
219 430 648 867 1078 1289 1508 2
211 422 633 844 1055 1258 1477 3
211 422 633 844 1055 1266 1477 3
141 273 406 547 688 820 961 4
133 266 398 530 664 797 938 4
117 234 352 469 586 711 820 5
117 234 359 484 602 719 836 5
133 273 383 539 672 805 938 6
133 266 383 531 664 797 930 6
203 398 602 797 984 1203 1398 7
195 398 602 797 984 1195 1391 7]

```

Figura E.6 – Exemplo de uma tabela de dados

Para a redefinição dos valores da tabela basta seleccionar a opção Redefinição de Faixas, da tela principal do sistema como mostra a Figura E.1, e indicar o nome da tabela contendo os dados. A figura E.7 mostra os valores redefinidos da tabela ilustrada pela figura E.6.

Entre com o Nome da Tabela para Redefinir as Faixas: dados

d =

7	8	12	17	19	22	28	1
7	8	12	17	19	22	28	1
7	9	14	18	21	25	30	2
7	9	13	18	21	24	29	2
7	8	13	17	20	23	28	3
7	8	13	17	20	23	28	3
4	2	4	5	5	5	7	4
3	2	3	5	4	4	6	4
2	1	2	2	1	1	1	5
2	1	2	3	2	1	2	5
3	2	3	5	4	5	6	6
3	2	3	5	4	4	6	6
7	7	12	15	17	21	24	7
6	7	12	15	17	20	24	7

Figura E.7 – Exemplo de valores redefinidos de uma tabela

Após obter as características e conseqüentemente as tabelas de dados com valores redefinidos e sem informação redundante, pode-se optar pela técnica dos Conjuntos Aproximados, o qual irá realizar a busca dos redutos. A figura E.8 ilustra um exemplo com o procedimento inicial da busca de reduto. Dessa forma o sistema será finalizado ao encontrar os redutos mínimos.

Your variables are:

banconovo

Digite o Nome do Sistema de Informação ou Banco de Dados : banconovo

entra =

7	9	14	18	21	25	30	2
7	9	13	18	21	24	29	2
7	9	13	18	20	24	28	2
4	2	4	5	5	5	7	4
3	2	3	4	4	4	10	4
3	2	3	5	4	4	10	4
3	2	3	4	3	4	10	4
3	2	4	5	4	5	10	4

Subconjuntos que possuem o mesmo poder discriminatório

subconjuntos =

subconjuntos =

subconjuntos =

F2 F3 F4 F5 F6 F7	F3 F4 F5 F6 F7	F3 F4 F5 F6 F7
F1 F3 F4 F5 F6 F7	F2 F4 F5 F6 F7	F1 F4 F5 F6 F7
F1 F2 F4 F5 F6 F7	F2 F3 F5 F6 F7	F1 F3 F5 F6 F7
F1 F2 F3 F5 F6 F7	F2 F3 F4 F6 F7	F1 F3 F4 F6 F7
F1 F2 F3 F4 F6 F7	F2 F3 F4 F5 F7	F1 F3 F4 F5 F7
F1 F2 F3 F4 F5 F7	F2 F3 F4 F5 F6	F1 F3 F4 F5 F6
F1 F2 F3 F4 F5 F6		

Figura E.8 – Busca de redutos

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] LEE, Chin-Hui, SOONG, Frank K., PALIWAL, Kuldip K.: "Automatic Speech and Speaker Recognition Advanced Topics", Kluwer Academic Publishers, USA, 1996.
- [2] CORSI, P. "Speaker Recognition: a survey. In: Proceedings of the Nato Advanced Study Institute", Bonas, 1981. "Automatic Speech Analysis and Recognition", Dordrecht, D. Deitel Publishing company, 1982. pp. 277-308.
- [3] SMITH, Steven W.: "The Scientist and Engineer's Guide to Digital Signal Processing", Second Edition, California Technical Publishing, San Diego, California, 1999. [<http://www.dspguide.com>]
- [4] MAGNI, A. B.: "Reconhecimento Automático de Locutor com Coeficientes Mel Cepstrais e Redes Neurais Artificiais", São Paulo, 1998, Dissertação de Mestrado, Escola Politécnica da Universidade de São Paulo.
- [5] BEZERRA, M. R., "Reconhecimento Automático de Locutor Para Fins Forenses, Utilizando Técnicas de Redes Neurais", Rio de Janeiro, 1994, Dissertação de Mestrado, Instituto Militar de Engenharia.
- [6] CZYZEWSKI, Andrzej : "Speaker-Independent Recognition of Isolated Words Using Rough Sets", In: Sound Engineering Department, Faculty of Electronics, Technical University of Gdansk, 80-952 Gdansk, Poland. Elsevier Science Inc., 1998.
- [7] FERREIRA, A., B. H.: "Novo dicionário da língua portuguesa". Rio de Janeiro: Nova Fronteira, 1975.
- [8] KELLER, E. "Fundamentals of Speech Synthesis and Speech Recognition Basic Concepts, State-of-the-Art and Future Challenges", John Wiley & Sons Ltd., 1994.
- [9] RABINER, L.R. and SCHAFER, R.W., "Digital Processing Of Speech Signals", Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1978.

- [10] CASAGRANDE, R.: "Redes Neurais Artificiais com retardos temporais aplicadas ao Reconhecimento Automático de Locutor", São Paulo, 1997, Dissertação de Mestrado, Escola Politécnica da Universidade de São Paulo.
- [11] MARVEN, C. and EWERS G.: "A Simple Approach to Digital Signal Processing", John Wiley & Sons, Inc., New York, 1996.
- [12] BRIGHAM, E. O.: "The Fast Fourier Transform", Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1974.
- [13] EMBREE, P.M. and KIMBLE B.: "C Language Algorithms for Digital Signal Processing", Prentice Hall, Englewood Cliffs, New Jersey 1991.
- [14] OPPENHEIM, A. V. and SCHAFER, R. W.: "Digital Signal Processing", Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1975.
- [15] TIMOSZCZUCK, A.P.: "Reconhecimento Automático de Locutor com Redes Neurais Artificiais do Tipo Radial Basis Function (RBF) e Minimal Temporal Information (MTI)", São Paulo, 1998, Dissertação de Mestrado, Escola Politécnica da Universidade de São Paulo.
- [16] PAPAMICHALIS, P. E.: "Practical Approaches To Speech Coding", Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1987.
- [17] DAVIS, S. and MERLMESTEIN, P.: "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences", IEEE Trans. on ASSP, Aug, 1980. pp. 357-366.
- [18] OPPENHEIM, A.V. and SCHAFER: "Homomorphic Analysis of Speech", IEEE Trans. Audio Electroacoust., vol. AU-16, June, 1968.
- [19] NOLL, A. M. "Cepstrum Pitch Determination", J. Acoust. Soc. Am., vol.41, pp.293-309, Feb. 1967.

[20] BRÜEL & Kjær: "Digital Signal Analysis Using Digital Filters and FFT Techniques", Selected Reprints from Technical Review, August, 1981.

[21] RORABAUGH, C. B.: "DSP Primer", McGraw-Hill, New York, 1999. pp. 469-488

[22] STEVENS, S. S. and VOLKMAN, J.: "The Relation of Pitch to Frequency", American Journal of Psychology, vol 53, pp. 329, 1940.

[23] DELLER, J. R. et al. 'Discrete- Time Processing of Speech Signals', Macmillan Publishing Company, New York, 1993

[24] RABINER, L. R. and JUANG, B. H., 'Fundamentals of Speech Recognition', Prentice Hall, New Jersey, 1993.

[25] PICONE, J. W., 'Signal Modeling Techniques in Speech Recognition', Proc. of the IEEE, Vol 81, No 9, pp. 1215-1247, setembro 1993.

[26] OLIVEIRA, E. M.: "Eletrônica Digital" Editora Campus, Rio de Janeiro, 1987.

[27] LU, Guojun. "Communication and Computing for Distributed Multimedia Systems." Artech House, 1996.

[28] HENRIQUES, S. B.: "Transporte de Áudio e Vídeo MPEG em Redes ATM", Pará, 1998, Monografia de Especialização, Universidade Federal do Pará.

[29] PAWLAK, Z.: "Rough Sets, Theoretical Aspects Of The Reasoning About Data", Kluwer, Dordrecht, 1991.

[30] JAGIELSKA, I.; MATTHEWS, C. and WHITFORT, T.: "An Investigation into the Application of Neural Networks, Fuzzy Logic, Genetic Algorithms, and Rough Sets to Automated Knowledge Acquisition for Classification Problems", Neurocomputing 24, Elsevier Science B.V., 1999.

- [31] LIPSCHUTZ, S.: "Teoria dos Conjuntos", Coleção Schaum Editora McGraw-Hill do Brasil, Ltda., 1972.
- [32] WALCZAK, B. and MASSART, D.L.: "Rough Sets Theory", Chemometrics and Intelligent Laboratory Systems 47, Elsevier Science B.V., 1999.
- [33] KRYSZKIEWICZ, M.: "Rough Set Approach to Incomplete Information Systems", Information Sciences 112, Elsevier Science B.V., 1998.
- [34] KRYSZKIEWICZ, M.: "Rules in Incomplete Information Systems", Information Sciences 113, Elsevier Science B.V., 1999.
- [35] NICOLETTI, M. e UCHÔA, J.: "O uso da Teoria de Conjuntos Aproximados na Determinação de Redutos de Conjuntos de Atributos", Relatório Técnico do Departamento de Computação 001/98. São Carlos, DC-UFSCar, 1998.
- [36] SZLADOW, A. and ZIARKO, W.: "Rough Sets: Working With Imperfect Data", AI Expert, pp. 36-41, July 1993.
- [37] ZIARKO, W. P.: "The Discovery, Analysis, and Representation of Data Dependencies in Databases", In: Piatestsky-Shapiro, G. & FRAWLEY, W. J., Knowledge Discovery in Databases, Boston, The MIT press, pp.195-209, 1991.
- [38] SILVA, L. E. B., TORRES, G. L., HENRIQUES, S. B., PINTO, J. O. P.: "Induction Motors Fault Diagnosis Based On Rough Set Classification", Marcon 2001 – Maintenance And Reliability Conference, May – 6-9, Gatlinburg, Tennessee, USA, 2001.
- [39] NICOLETTI, M. e UCHÔA, J.: "ILROS: Um Sistema de Aprendizado Indutivo de Máquina Baseado em Conjuntos Aproximados", Relatório Técnico do Departamento de Computação 001/98. São Carlos, DC-UFSCar, 1998.
- [40] CARVALHO, J. C. S. : "Aplicação de Conjuntos Aproximados Utilizando Banco de Dados Relacionais", Itajubá-MG, 2000, Dissertação de Mestrado, Escola Federal de Engenharia de Itajubá.

[41] TORRES, G. L., ROSSI, R., JARDINI, J. A., SILVA, A. P. A., QUINTANA, V. H.:
“Power System Security Analysis based on Rough Classification”, pp. 263-274, Rough Fuzzy
Hybridization – A New Trend in Decision Making, S. K. Pal, A. Skowron – Springer-Verlag,
Singapore, 1999.