

UNIVERSIDADE DE ÉVORA
Departamento de Informática

**RECONHECIMENTO DE VOZ UTILIZANDO UM MODELO
DECISÃO MÚLTIPLO ATRIBUTO DIFUSO**

Por
Tiago Manuel Louro Machado de Simas

Dissertação apresentada na
Universidade de Évora para obtenção de
grau de Mestre em Inteligência
Artificial Aplicada.

Esta dissertação não inclui as críticas e
sugestões feitas pelo Júri.

Esta dissertação foi realizada sob
orientação de Rita A. Ribeiro.

LISBOA
(2001)

Aos Meus Pais e Irmãos,
pela Motivação que me deram e Paciência
que tiveram neste período da minha Vida.

AGRADECIMENTOS

Gostaria de Agradecer à minha Orientadora de Mestrado, Rita Almeida Ribeiro, por ter-me mostrado como era o Mundo Científico.

Gostaria de Agradecer ao Luís Correia por todo o apoio que me deu durante o mestrado.

SUMÁRIO

Esta tese de Mestrado em Inteligência Artificial Aplicada, tem como objectivo propor uma alternativa a técnicas usadas no reconhecimento de padrões usando técnicas da tomada de decisão múltiplo atributo difusas. De forma a ilustrar a flexibilidade do método um exemplo prático do reconhecimento de dígitos 0-9 é apresentado no capítulo 5.

Este trabalho teve por base o estudo de três áreas fundamentais, análise e processamento de sinal, a teoria dos conjuntos difusos e as teorias da decisão múltiplo critério. Esta última focando especificamente a subárea denominada tomada de decisão múltiplo atributo.

ABSTRACT

This dissertation proposes a new technique for the pattern recognition based in Fuzzy Multiple Attribute Decision techniques. In chapter 5 we present a practical example to show the flexibility of this technique.

This work had as base of study three main areas, the Discrete-Time Signal Processing, the theory of Fuzzy Sets and the theories of multiple criteria decision-making in special the sub-area multiple attribute decision-making.

ÍNDICE DE MATÉRIAS

CAPÍTULO 1	
1.1. INTRODUÇÃO GERAL	1
CAPÍTULO 2	
2.1. INTRODUÇÃO	3
2.2. SISTEMA FISIOLÓGICO HUMANO	3
2.3. MODELO FÍSICO	4
2.3.1. ONDAS SONORAS A UMA DIMENSÃO NUM TUBO	5
2.3.2. ONDAS SONORAS A UMA DIMENSÃO NUM SISTEMA MULTI-TUBULAR	9
2.3.2.1. VÁRIOS EFEITOS A CONSIDERAR NA TRANSFERÊNCIA DE UMA ONDA SONORA ENTRE 2 SEGMENTOS	10
2.3.2.2. FUNÇÃO DE TRANSFERÊNCIA PARA UM SISTEMA VOCAL DE 2 SEGMENTOS	12
2.3.2.3. FUNÇÃO DE TRANSFERÊNCIA PARA UM SISTEMA VOCAL DE p SEGMENTOS	14
2.4. CONCLUSÕES	15
CAPÍTULO 3	
3.1. INTRODUÇÃO	17
3.2. BREVE DESCRIÇÃO DA ANÁLISE DE SINAL E SISTEMAS	17
3.3. MODELO LINEAR PREDICTIVE CODING (LPC)	18
3.4. DETERMINAÇÃO DOS COEFICIENTES $\{a_i\}$	20
3.4.1. MÉTODO AUTOREGRESSIVO DE BURG	21
3.4.2. ESTABILIDADE DO MÉTODO DE BURG	25
3.4.3. CONSIDERAÇÕES DE APLICABILIDADE DO MÉTODO AUTOREGRESSIVO DE BURG	26
3.4.4. MÉTODO AUTOREGRESSIVO DE GHAEL-SANDGATE	26
3.4.4.1. DESCRIÇÃO DO MÉTODO AUTOREGRESSIVO DE GHAEL-SANDGATE	27
3.5. CONCLUSÕES	28

CAPÍTULO 4	
4.1. INTRODUÇÃO	30
4.2. CONJUNTOS DIFUSOS	30
4.2.1. NOÇÃO DE CONJUNTO DIFUSO	30
4.2.2. OPERAÇÕES SOBRE CONJUNTOS DIFUSOS	38
4.2.2.1. COMPLEMENTO DIFUSO	38
4.2.2.2. UNIÃO DIFUSA	41
4.2.2.3. INTERSECÇÃO DIFUSA	42
4.2.2.4. CONSIDERAÇÕES SOBRE A UNIÃO E A INTERSECÇÃO DIFUSA	43
4.2.2.5. OPERAÇÕES DE AGREGAÇÃO GERAIS	44
4.3. MODELOS DE TOMADA DECISÃO MULTICRITÉRIO	48
4.3.1. BREVE DESCRIÇÃO DOS MODELOS USADOS NA TOMADA DE DECISÃO	48
4.3.2. MODELOS DE DECISÃO MÚLTIPLO ATRIBUTO DIFUSOS	50
4.3.3. MODELO AHP ESTENDIDO COM OWA	52
4.3.3.1. MÉTODO AHP (ANALYTICAL HIERARCHY PROCESS)	52
4.3.3.2. OPERADOR OWA	56
4.3.3.2.1. OWA E QUANTIFICADORES LINGUÍSTICOS	57
4.3.3.2.2. OWA QUANTIFICADOS E HISTOGRAMAS DE VARIÁVEIS	61
4.3.3.3. EXEMPLO DA EXTENSÃO AO AHP COM OWA+QUANTIFICADORES	64
4.4. CONCLUSÕES	69
CAPÍTULO 5	
5.1. INTRODUÇÃO	70
5.2. CONSIDERAÇÕES GERAIS	70
5.3. RECONHECIMENTO DOS DÍGITOS NA LÍNGUA PORTUGUESA UTILIZANDO UM MÉTODO DE TOMADA DE DECISÃO MÚLTIPLO ATRIBUTO DIFUSO	72
5.3.1. EXTRACÇÃO DE PARÂMETROS	73
5.3.2. COMPARAÇÃO DE PADRÕES	75
5.3.3. REGRA DE DECISÃO	77
5.3.4. RESULTADOS	81
5.4. RECONHECIMENTO DOS DÍGITOS NA LÍNGUA PORTUGUESA UTILIZANDO UM MÉTODO DE TOMADA DE DECISÃO MÚLTIPLO ATRIBUTO	81
5.5. RECONHECIMENTO DOS DÍGITOS NA LÍNGUA	83

PORTUGUESA UTILIZANDO UM MÉTODO DE QUANTIZAÇÃO VECTORIAL E COMPARAÇÃO DE PADRÕES SEM ALINHAMENTO NO TEMPO	
5.5.1. CALCULO DOS COEFICIENTES CEPSTRAIS E DELTA- CEPSTRAIS	83
5.5.2. QUANTIZAÇÃO VECTORIAL	84
5.5.3. RECONHECIMENTO DE PADRÕES BASEADO NA QUANTIZAÇÃO VECTORIAL SEM ALINHAMENTO NO TEMPO	85
5.5.4. RESULTADOS	86
5.6. BREVES CONCLUSÕES AOS MODELOS	86
5.7. CONCLUSÕES	87
 CAPÍTULO 6	
6.1. INTRODUÇÃO	88
6.2. CONCLUSÕES E TRABALHO FUTURO	88
 BIBLIOGRAFIA	90

ÍNDICE DE FIGURAS

Figuras	Pagina
Figura 2.1- Raio-X do aparato vocal do ser Humano [Rabiner, 1978]	3
Figura 2.2- Diagrama do aparato Vocal Humano [Rabiner, 1978]	4
Figura 2.3- Sistema Multitubular [Titze, 1994]	5
Figura 2.4- Sistema tubular	6
Figura 2.5 – Sistema multitubular	9
Figura 2.6 – Ondas Sonoras num tubo	10
Figura 2.7 – Atrasos por segmento	10
Figura 2.8 – Junção de segmentos	11
Figura 2.9 – Sistema vocal de 2-segmentos	12
Figura 2.10 – Envoltente espectral, os picos correspondem às formantes	15
Figura 3.1- Dois picos detectados pelo método 2, de Ghael-Sandgathe (envoltente superior) que não são detectados pelo método 1, método de Burg (envoltente inferior).	27
Figura 4.1- O número está perto de 2	33
Figura 4.2- Conjunto difuso convexo e função de pertença não convexa	36
Figura 4.3- Conjunto difuso não convexo e função de pertença não convexa	36
Figura 4.4- Classe de Sugeno [Klir, 1988]	40
Figura 4.5- Classe de Yager [Klir, 1988]	41
Figura 4.6- Classificação dos operadores de agregação [Klir, 1988].	46
Figura 4.7- Modelos de Tomada de Decisão	49
Figura 4.8- Hierarquia do objectivo escolha de casa	53
Figura 4.9 – Quantificadores relativos “jovem” e “muito jovem”	58
Figura 4.10- “Tabela” de veracidade difusa[Klir, 1988]	58
Figura 4.11- Quantificadores RIM	59
Figura 4.12- Exemplo de histograma de uma variável Física	61
Figura 4.13 – Conjuntos difusos para os primeiros rectângulos (duas primeiras partições) do histograma	62
Figura 4.14- Hierarquia para a escolha de um engenheiro informático [Yager, 1999]	64
Figura 5.1- Limites de percepção do ouvido Humano. [http://www.iurc.montp.insem.fr/cric/audition/english/sound/sound.htm]	71

Figura 5.2- Parte da sequência de quatro capturados por uma pessoa	71
Figura 5.3- Forma como foram cortados os sinais para a base de dados	72
Figura 5.4- Modelo usado para o reconhecimento de padrões	73
Figura 5.5- Modelo LPC utilizado para extracção de parâmetros	74
Figura 5.6 – Construção dos conjuntos difusos	76
Figura 5.7- Hierarquia para as alternativas 0-9	77
Figura 5.8- Funções com uma certa incerteza mas não normalizadas	82
Figura 5.9- Sistema de reconhecimento de padrões baseado na Quantização Vectorial [Rabiner, 1993].	85

ÍNDICE DE QUADROS

Quadros	Pagina
Tabela 4.1- t-normas e t-conormas [Zimmermann, 1991]	42
Tabela 4.2- Classificação dos principais métodos de TDMA [Ribeiro,1996].	51
Tabela 4.3 – Relação par-a-par dos atributos	52
Tabela 4.4 – Vector dos pesos	55
Tabela 4.5 – Classificação atribuída pelo decisor	55
Tabela 5.1- Valor de pertença, CD-Conjunto Difuso	79
Tabela 5.2- Importâncias por Janela e Alternativa, CD-Conjunto Difuso	79
Tabela 5.3- Resultado da agregação do nível 5	79
Tabela 5.4- Resultado da agregação do nível 5 Normalizada	79
Tabela 5.5- Resultado da agregação de Mb com Mgs	80
Tabela 5.5- Percentagem de resultados correctos	81
Tabela 5.6- Percentagem de resultados correctos	82
Tabela 5.7- Percentagem de resultados correctos	86

“Só Sei que Nada Sei...”
Sócrates

“To Be or Not To Be, This is the Question.”
Shakespeare

CAPÍTULO 1

Introdução

1.1. INTRODUÇÃO GERAL

Esta tese de Mestrado em Inteligência Artificial Aplicada, tem como objectivo propor uma alternativa a técnicas usadas no reconhecimento de padrões usando técnicas da tomada de decisão múltiplo atributo difusas. De forma a ilustrar a flexibilidade deste método usou-se como exemplo prático o reconhecimento dos dígitos 0-9.

Este trabalho teve por base o estudo de três áreas fundamentais, análise e processamento de sinal, a teoria dos conjuntos difusos e as teorias da decisão múltiplo critério. Esta última focando especificamente a subárea denominada tomada de decisão múltiplo atributo.

O capítulo 2 começa por descrever o nosso sistema vocal a partir da fisiologia Humana. A seguir descreve-se um modelo físico para modular a propagação do som (voz) através do nosso sistema vocal, onde é sugerido um sistema multitubular. Começa-se por se deduzir as equações de onda num sistema de um só tubo uniforme e depois generaliza-se para um caso multitubular onde é apresentada a função de transferência para esse sistema. São apresentados como parâmetros para caracterizar a voz as frequências de ressonância denominadas por formantes.

No capítulo 3 vamos ver como a teoria do sinal e sistemas se pode relacionar com o modelo físico multitubular apresentado no capítulo anterior. Iniciamos com uma breve descrição da análise do sinal e sistemas, onde se faz uma breve introdução a conceitos básicos da teoria do sinal. Depois introduz-se o modelo LPC (Linear Predictive Coding) que relaciona o modelo físico multitubular do capítulo anterior com as técnicas do processamento do sinal. A seguir explicam-se dois modelos autoregressivos como resolução das equações apresentadas no modelo LPC. O método autoregressivo de Burg e o modelo autoregressivo de Ghael-Sandgate. São apresentadas também as condições de aplicabilidade dos métodos tais como também a estabilidade destes métodos.

O capítulo 4 está dividido em duas partes, uma em que se faz um resumo geral da teoria dos conjuntos difusos e uma segunda parte onde se faz um resumo da teoria da Tomada de Decisão Múltiplo Critério Difusa. Na

primeira parte deste capítulo far-se-á uma breve introdução ao conceito de conjunto difuso. Depois serão definidas as estruturas matemáticas para as operações entre conjuntos difusos (negação, intersecção e união). Depois será definida uma estrutura matemática mais generalizada de operações de agregação entre conjuntos difusos.

Na segunda parte deste capítulo far-se-á uma abordagem breve e geral sobre os vários modelos utilizados na Tomada de Decisão Múltiplo Critério Difuso: (a) Tomada de Decisão Múltiplo Objectivo Difusa e (b) Tomada de Decisão Múltiplo Atributo Difusa. Serão apresentados de uma forma mais detalhada os métodos mais usados na Tomada de Decisão Múltiplo Atributo Difusa. Apresenta-se em particular o método AHP (Analytical Hierarchy Process), introduzido por Saaty, e o seu modelo estendido por Yager, utilizando os operadores de agregação OWA (Ordered Weighted Averaging).

Terminamos o capítulo apresentando algumas das vantagens de se usar este modelo estendido AHP.

No capítulo 5 são apresentados três métodos para o reconhecimento dos dígitos, números de 0 a 9, na língua Portuguesa. Primeiro apresenta-se um método de tomada de decisão múltiplo atributo difuso que foi escolhido, o AHP estendido. Em segundo apresenta-se o método de tomada de decisão múltiplo atributo clássico, baseado em médias ponderadas. Em terceiro apresenta-se um método já usado no reconhecimento de dígitos, o método de quantização vectorial e comparação de padrões sem alinhamento temporal. Por último será feita uma breve discussão acerca dos métodos aplicados.

No capítulo 6 apresentam-se as conclusões com respeito ao trabalho da tese e apresentam-se algumas sugestões para trabalho futuro.

**Sistema Fisiológico e Físico e Modelos Matemáticos Aplicados
ao processamento de voz**

2.1. INTRODUÇÃO

Neste capítulo começa-se por se descrever o nosso sistema vocal a partir da fisiologia Humana. A seguir descreve-se um modelo físico para modular a propagação do som (voz) através do nosso sistema vocal, onde é sugerido um sistema multitubular. Começa-se por se deduzir as equações de onda num sistema de um só tubo uniforme e depois generaliza-se para um caso multitubular onde é apresentada a função de transferência para esse sistema. São ainda apresentados como parâmetros para caracterizar a voz as frequências de ressonância denominadas por formantes.

2.2. SISTEMA FISIOLÓGICO HUMANO

Na figura 2.1 podemos ver um Raio-X do sistema fisiológico Humano por onde a voz é processada [Rabiner, 1978], o nosso sistema vocal.

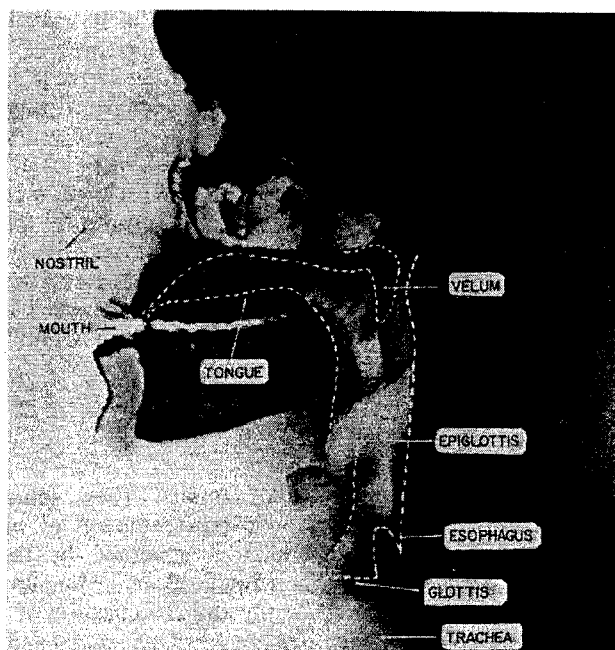


Figura 2.1- Raio-X do aparato vocal do ser Humano [Rabiner, 1978]

Na figura 2.1 está demarcado a tracejado as zonas por onde a voz é processada e como se pode ver começa na glote (ou cordas vocais) e termina na boca (lábios).

Na figura 2.2 podemos ver em detalhe o diagrama pelo qual a voz é processada no nosso sistema fisiológico.

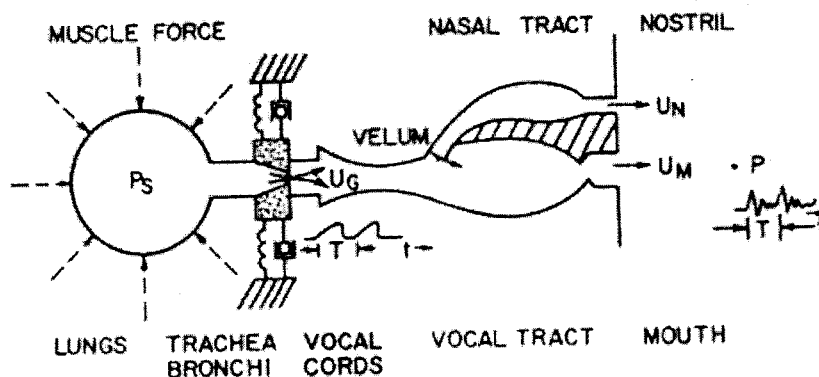


Figura 2.2- Diagrama do aparato Vocal Humano [Rabiner, 1978]

A voz é um conjunto de ondas sonoras que são geradas na Glote ou Cordas Vocais que se propaga pela Laringe, Faringe, Boca e zona Nasal. O som é produzido pela contração dos músculos na caixa torácica que cria uma pressão e faz com que o ar seja expelido para fora do sistema fisiológico. O ar passa pela traqueia e ao passar pela glote (cordas vocais) provoca uma vibração nestas onde são produzidas ondas sonoras (impulso inicial). Estas ondas sonoras estão confinadas ao sistema fisiológico (sistema vocal) apresentado nas figuras anteriores, e ao propagarem-se pela laringe, faringe, boca e zona nasal, são alteradas de acordo com a forma que estas zonas têm, saindo assim o som pela saída da boca e do nariz (sinal de saída). O som assim produzido depende da forma como vamos modulando estas zonas, boca, “velum” véu palatino. Isto é, quando produzimos um som contraímos ou descontraímos mais as cordas vocais; fechamos ou abrimos mais/menos o véu palatino para que saia mais/menos som pelas vias nasais e ao abrimos a boca alteramos esta zona,, mais ou menos, modulando assim o som ao longo do nosso sistema vocal.

2.3. MODELO FÍSICO

Do ponto de vista físico podemos modelar o sistema fisiológico como um sistema multitubular, como podemos ver na figura 2.3. As ondas sonoras são produzidas na glote e depois propagam-se por um sistema de tubos que representam a faringe e a boca. Aqui não estamos a ter em consideração a cavidade nasal.

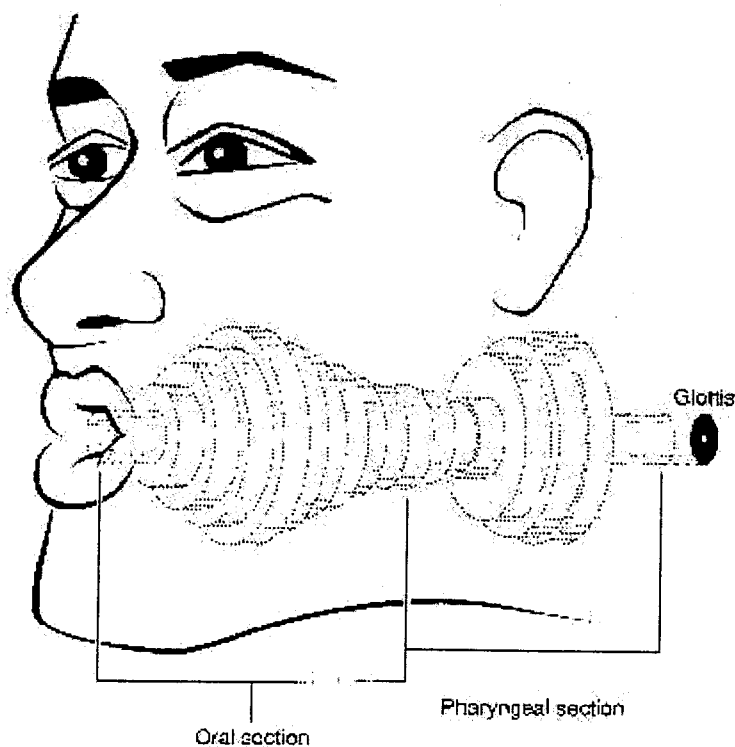


Figura 2.3- Sistema Multitubular [Titze, 1994]

Em resumo o problema de modelar o sistema fisiológico consiste em determinar como é que as ondas sonoras se comportam num sistema multitubular.

2.3.1. ONDAS SONORAS A UMA DIMENSÃO NUM TUBO

Nesta secção vamos deduzir matematicamente as equações de onda confinadas a um tubo.

Por definição, ondas sonoras são ondas mecânicas que se podem propagar através dum gás, líquido ou sólido [Halliday, 1988]. No ar provocam perturbações de pressão (compressão ou descompressão) que se propagam em volume pelo meio que é o ar.

Onda Sonora confinada a um Tubo:

Consideremos uma pequena perturbação de ar confinado a um tubo. Admitindo que a secção transversal do tubo é uniforme e de área A , isto é, que o percurso fisiológico é constante em x e que A é invariante em relação ao tempo t , conforme se pode ver na figura 2.4:

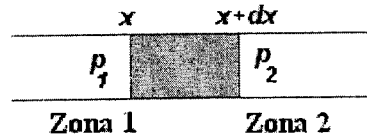


Figura 2.4- Sistema tubular

Temos assim na zona 1 (inicial) uma pressão p_1 e uma velocidade v_1 e na zona 2 uma pressão p_2 e uma velocidade v_2 que se pode traduzir nas equações:

$$\text{Zona 1} \begin{cases} p_1 = p \\ v_1 = v = \frac{u}{A} \end{cases} \quad (2.3.1)$$

$$\text{Zona 2} \begin{cases} p_2 = p + \frac{\partial p}{\partial x} dx \\ v_2 = v + dv = \frac{1}{A} \left(u + \frac{\partial u}{\partial x} \right) \end{cases} \quad (2.3.2)$$

Onde u é a variação de velocidade volumétrica do fluxo na posição x e no instante t , isto é, a quantidade de volume deslocado por unidade de tempo. Por outro lado temos que volume de ar deslocado V , é dado por:

$$V = A \times dx$$

Logo a sua variação ao longo do tempo é dada por,

$$\frac{\partial V}{\partial t} = A \times dv = A \times \frac{dx}{A} \cdot \frac{\partial u}{\partial x} = \frac{V}{A} \times \frac{\partial u}{\partial x} \quad (2.3.3)$$

onde, $dv = \frac{dx}{dt}$ e $u = \frac{d(A \cdot x)}{dt}$ (velocidade volúmica).

Estamos assim nas condições de determinar a força resultante que actua neste volume V :

$$F = A \cdot p - A \cdot \left(p + \frac{\partial p}{\partial x} dx \right) = -A \frac{\partial p}{\partial x} dx \quad (2.3.4)$$

Por outro lado temos pela lei dos gases ideais:

$$pV = nRT$$

onde [Dias de Deus, 1992]:

$$n = \text{moles de ar} = \frac{\text{moleculas}}{6 \times 10^{23}}$$

$$R = \text{Constante dos gases ideais} = 8.314 \text{ J/(K mol)}$$

$$T = \text{Temperatura (}^\circ\text{K)}$$

por outro lado temos $n = \frac{\rho \cdot V}{M}$

Onde:

$$\rho = \text{Densidade do ar } (\approx 1.225 \text{ kg/m}^3)$$

$$M = \text{Massa molecular do ar} = 0.029 \text{ kg/mol}$$

Substituindo este ultimo resultado na equação dos gases ideais obtemos:

$$p = \rho \times \frac{RT}{M} \quad (2.3.5)$$

Definindo a velocidade do som como:

$$c^2 = \frac{\gamma \cdot RT}{M} \approx (340 \text{ m/s})^2$$

Onde $\gamma = \frac{C_p}{C_v}$ é a razão dos calores específicos do ar a pressão constante e

volume constante, considerando um gás diatômico $\gamma=1.4$ [Fermi, 1937] e em consequência a velocidade do som depende unicamente da temperatura. Temos por substituição em (2.3.5):

$$p \cdot \gamma = \rho \cdot c^2 \quad (2.3.6)$$

Diz-se que uma transformação de um sistema termodinâmico é adiabática se é reversível e se o sistema se encontra termicamente isolado, de modo que durante a transformação nenhum calor pode ser permutado entre ele e a vizinhança [Fermi, 1937]. No nosso caso, propagação do ar pelo sistema tubular, podemos considerar que as diferenças de pressão que se propagam ao longo do tubo são muito rápidas de modo a não trocarem calor com o tubo e assim as transformações podem-se considerar adiabáticas. Neste caso é válida a seguinte lei:

$$\frac{\partial}{\partial t}(pV^\gamma) = 0 \Rightarrow V^\gamma \frac{\partial p}{\partial t} + p\gamma V^{\gamma-1} \frac{\partial V}{\partial t} = 0 \quad (2.3.7)$$

Usando as equações (2.3.3) e (2.3.6) e substituindo nesta última equação obtemos:

$$\begin{aligned} V^\gamma \frac{\partial p}{\partial t} &= -\rho c^2 \frac{V^\gamma}{A} \frac{\partial u}{\partial x} \Leftrightarrow \\ A \frac{\partial p}{\partial t} &= -\rho c^2 \frac{\partial u}{\partial x} \end{aligned} \quad (2.3.8)$$

Aplicando a segunda lei de Newton ao sistema:

$$F = m \times a$$

onde a é a aceleração $a = \frac{dv}{dt} = \frac{1}{A} \frac{\partial u}{\partial t}$ e $m = \rho V$.

Obtemos, utilizando a (2.3.4),

$$\begin{aligned} -A \frac{\partial p}{\partial x} dx &= \rho V \frac{1}{A} \frac{\partial u}{\partial t} \Leftrightarrow -V \frac{\partial p}{\partial x} = \rho V \frac{1}{A} \frac{\partial u}{\partial t} \Leftrightarrow \\ -A \frac{\partial p}{\partial x} &= \rho \frac{\partial u}{\partial t} \end{aligned} \quad (2.3.9)$$

As equações (2.3.8) e (2.3.9) são conhecidas como as equações de onda:

$$A \frac{\partial p}{\partial t} = -\rho c^2 \frac{\partial u}{\partial x} \quad \text{e} \quad \rho \frac{\partial u}{\partial t} = -A \frac{\partial p}{\partial x}$$

Cujas soluções são ondas estacionárias, isto é, a soma de duas ondas progressivas u^- e u^+ que se deslocam em sentidos opostos, com uma velocidade $-c$ e c .

$$u(x,t) = u^+ \left(t - \frac{x}{c} \right) - u^- \left(t + \frac{x}{c} \right) \quad (2.3.10)$$

$$p(x,t) = \frac{\rho c}{A} \left\{ u^+ \left(t - \frac{x}{c} \right) + u^- \left(t + \frac{x}{c} \right) \right\} \quad (2.3.11)$$

É fácil provar que para as funções diferenciáveis u^- e u^+ as equações (2.3.10 e 2.3.11) são soluções das equações de onda, conforme o Teorema D'Alembert [Dilão, 1993].

Note-se que existem outras formas de dedução matemática das equações de onda, como por exemplo a dada por [Dias de Deus, 1992].

2.3.2. ONDAS SONORAS A UMA DIMENSÃO NUM SISTEMA MULTI-TUBULAR

Vimos na secção anterior as equações que regem a propagação de uma onda sonora num sistema tubular de secção constante. Nesta secção vamos ver como é que uma onda se modifica ao longo de um sistema multitubular. Atendendo às equações (2.3.8) e (2.3.9) e as suas soluções, equações (2.3.10) e (2.3.11), basta-nos considerar os seguintes efeitos que acontecem à onda quando se transmite de um tubo para outro. Temos três efeitos, *atraso da onda por segmento de tubo*, condições fronteira entre segmento de tubos que se divide na *continuidade das funções que representam as soluções* das equações de onda e a *reflexão das ondas ao passar de um meio para outro*, isto é, de um tubo para outro.

Podemos modular o sistema fisiológico (Sistema Vocal) como um sistema multi-tubular, isto é, um tubo com p segmentos, figura 2.5 [Rabiner, 1978]:

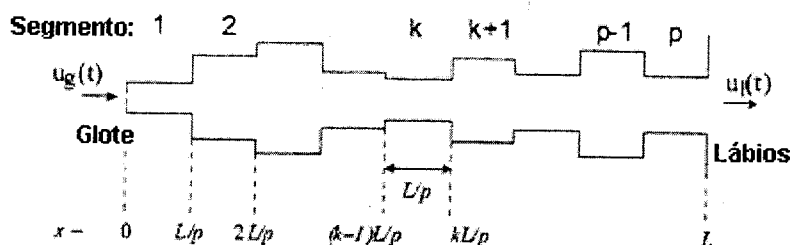


Figura 2.5 – Sistema multitubular

Onde u_g e u_l são os fluxos de ar na glote e lábios respectivamente (medidos em litros por segundo). Nota-se que em adultos este sistema tem um comprimento L de cerca de 15 a 17 cm [Rabiner, 1978].

Considera-se que o comprimento de cada segmento corresponde a metade do comprimento onda, frequência de Nyquist, temos para uma frequência de amostragem de $f_{amost}=11\text{kHz}$ e um comprimento $L=15\text{cm}$, temos $\frac{L}{p} = \frac{\lambda}{2} = \frac{1}{2}cT$, onde $c=340\text{ m/s}$ à temperatura ambiente e T é o período de

amostragem $=1/f_{amost}$. Temos assim,

$$p = \frac{2Lf_{amost}}{c} = \frac{2 \times 0.15}{340} \cdot f_{amost} \approx 0.001 f_{amost} \approx 11,$$

onde p é o número de tubos considerados para representar o nosso sistema vocal.

2.3.2.1. VÁRIOS EFEITOS A CONSIDERAR NA TRANSFERÊNCIA DE UMA ONDA SONORA ENTRE 2 SEGMENTOS

Ondas sonoras num tubo:

Tínhamos visto que num tubo as soluções da equação de onda são a sobreposição de duas ondas progressivas de velocidade $-c$ e c , isto é, uma onda estacionária. Consideremos $u^+ = u$ e $u^- = v$, conforme as equações (2.3.10) e (2.3.11), densidade do ar ρ e velocidade do som c figura 2.6:

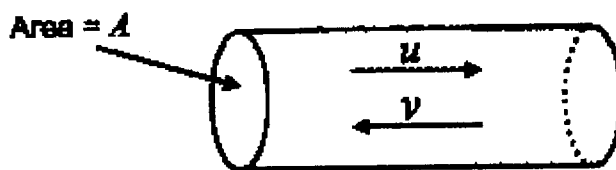


Figura 2.6 – Ondas Sonoras num tubo

Tínhamos visto que das equações (2.3.10) e (2.3.11):

$$\text{Fluxo de ar Total} = u - v$$

$$\text{Pressão acústica Total} = (u + v) \times \rho c / A$$

Vamos agora ver os diversos efeitos que ocorrem na propagação da onda sonora pelos diferentes segmentos, isto é, como é que a onda se vai transformando ao longo do sistema vocal (fisiológico).

Atrasos por segmento:

Consideremos um segmento representado na figura 2.7:

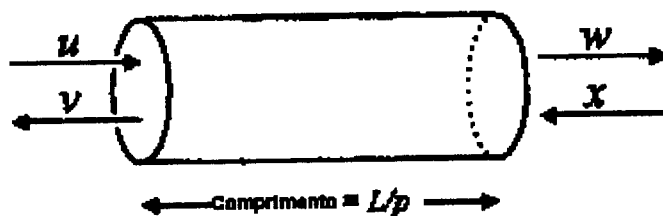


Figura 2.7 – Atrasos por segmento

Temos:

$$v(t) = x\left(t - \frac{L}{cp}\right)$$

$$u(t) = w\left(t + \frac{L}{cp}\right)$$

Onde L/cp é o tempo que o som leva a percorrer um segmento. Se o comprimento do segmento corresponde a metade do período de amostragem, temos pela transformada-Z que um atraso da onda num segmento (tempo que a onda leva a percorrer o segmento) corresponde a:

$$V(z) = z^{-1/2}X(z)$$

$$U(z) = z^{+1/2}W(z)$$

Ou na forma matricial:

$$\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} z^{+1/2} & 0 \\ 0 & z^{-1/2} \end{pmatrix} \begin{pmatrix} W \\ X \end{pmatrix}$$

Vejam agora as condições fronteira, isto é, como é que a onda se modifica ao passar de um tubo para outro.

Junção entre segmentos:

Consideremos a figura 2.8:

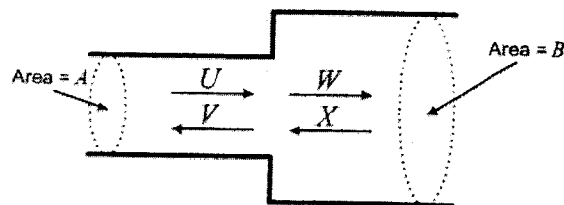


Figura 2.8 – Junção de segmentos

Condições fronteira:

$$\text{Continuidade de Fluxo de ar: } (U-V)=(W-X)$$

$$\text{Continuidade de pressão: } \frac{\rho c}{A}(U+V) = \frac{\rho c}{B}(W+X)$$

Na forma matricial temos:

$$\begin{pmatrix} 1 & -1 \\ B & B \end{pmatrix} \begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ A & A \end{pmatrix} \begin{pmatrix} W \\ X \end{pmatrix}$$

Logo:

$$\begin{pmatrix} U \\ V \end{pmatrix} = \frac{1}{2B} \begin{pmatrix} A+B & A-B \\ A-B & A+B \end{pmatrix} \begin{pmatrix} W \\ X \end{pmatrix}$$

Coefficiente de reflexão:

Uma onda sonora ao passar de um segmento para outro, figura 2.8, altera a sua densidade logo tem-se um coeficiente de reflexão [Dilão, 1993],

$$r = \frac{B-A}{B+A}$$

Logo:

$$\begin{pmatrix} U \\ V \end{pmatrix} = \frac{1}{2B} \begin{pmatrix} A+B & A-B \\ A-B & A+B \end{pmatrix} \begin{pmatrix} W \\ X \end{pmatrix} = \frac{1}{1+r} \begin{pmatrix} 1 & -r \\ -r & 1 \end{pmatrix} \begin{pmatrix} W \\ X \end{pmatrix}$$

Onde $-1 \leq r \leq 1$.

2.3.2.2. FUNÇÃO DE TRANSFERÊNCIA PARA UM SISTEMA VOCAL DE 2 SEGMENTOS

Nesta secção vamos determinar a função de transferência para um sistema vocal de dois segmentos e na secção 2.2.2.3 vamos generalizar para um sistema de p segmentos.

Consideremos que o nosso sistema vocal era constituído por 2 segmentos, figura 2.9:

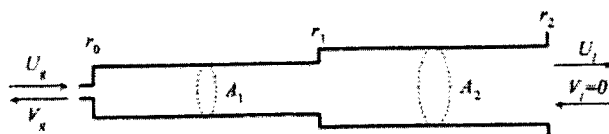


Figura 2.9 – Sistema vocal de 2-segmentos

Tendo em consideração todos os efeitos descritos na secção anterior tem-se para as diversas zonas r_0 , r_1 e r_2 onde r_0 e r_2 correspondem às zonas da glote e lábios, respectivamente, temos:

Zona r_0 :

$$\begin{pmatrix} U_g \\ V_g \end{pmatrix} = \frac{1}{1+r_0} \begin{pmatrix} 1 & -r_0 \\ -r_0 & 1 \end{pmatrix} \begin{pmatrix} z^{+1/2} & 0 \\ 0 & z^{-1/2} \end{pmatrix} \begin{pmatrix} W_0 \\ X_0 \end{pmatrix}$$

Zona r_1 :

$$\begin{pmatrix} W_0 \\ X_0 \end{pmatrix} = \frac{1}{1+r_1} \begin{pmatrix} 1 & -r_1 \\ -r_1 & 1 \end{pmatrix} \begin{pmatrix} z^{+1/2} & 0 \\ 0 & z^{-1/2} \end{pmatrix} \begin{pmatrix} W_1 \\ X_1 \end{pmatrix}$$

Zona r_2 :

$$\begin{pmatrix} W_1 \\ X_1 \end{pmatrix} = \frac{1}{1+r_2} \begin{pmatrix} 1 & -r_2 \\ -r_2 & 1 \end{pmatrix} \begin{pmatrix} U_l \\ V_l \end{pmatrix}$$

Logo fazendo a composição de todas as Zonas:

$$\begin{pmatrix} U_g \\ V_g \end{pmatrix} = \frac{z}{\prod_{k=0}^2 (1+r_k)} \begin{pmatrix} 1+(r_0r_1+r_1r_2)z^{-1}+r_0r_2z^{-2} \\ -r_0-(r_1+r_0r_1r_2)z^{-1}-r_2z^{-2} \end{pmatrix} U_l$$

Podemos desprezar V_g , pois esta é absorvida pelos pulmões. Assim obtemos a equação de transferência para o sistema vocal de 2 segmentos da seguinte forma:

$$H(z) = \frac{U_l}{U_g} = \frac{\prod_{k=0}^2 (1+r_k)z^{-1}}{1+(r_0r_1+r_1r_2)z^{-1}+r_0r_2z^{-2}}$$

$$H(z) = \frac{Gz^{-1}}{1+(r_0r_1+r_1r_2)z^{-1}+r_0r_2z^{-2}} = \frac{Gz^{-1}}{1-a_1z^{-1}-a_2z^{-2}}$$

2.3.2.3. FUNÇÃO DE TRANSFERÊNCIA PARA UM SISTEMA VOCAL DE p SEGMENTOS

Para um sistema vocal de p segmentos temos:

$$\begin{pmatrix} U_g \\ V_g \end{pmatrix} = \frac{z^{-\frac{1}{2}p}}{\prod_{k=0}^p (1+r_k)} \prod_{k=0}^{p-1} \begin{pmatrix} 1 & -r_k z^{-1} \\ -r_k & z^{-1} \end{pmatrix} \begin{pmatrix} 1 \\ -r_p \end{pmatrix} U_l$$

Donde obtemos para a função de transferência de um sistema vocal multi-tubular de p segmentos:

$$H(z) = \frac{U_l}{U_g} = \frac{Gz^{-\frac{1}{2}p}}{1 - \sum_{k=1}^p a_k z^{-k}} \quad (2.3.12)$$

Onde:

- G é o Ganho;
- $z^{-1/2p}$ é atraso de tempo acústico ao longo do sistema vocal;
- O denominador representa a ordem p dos pólos;

Podemos ver na equação 2.3.12, função de transferência, que esta tem p pólos. Se calcularmos a resposta em frequência dos coeficientes a_1, a_2, \dots, a_p , pela aplicação da FFT, determinamos a envolvente espectral da onda sonora, onde os picos dessa envolvente correspondem às formantes, isto é, frequências de ressonância, que correspondem aos pólos que estão mais próximos do raio unitário do plano complexo, ver figura 2.10.

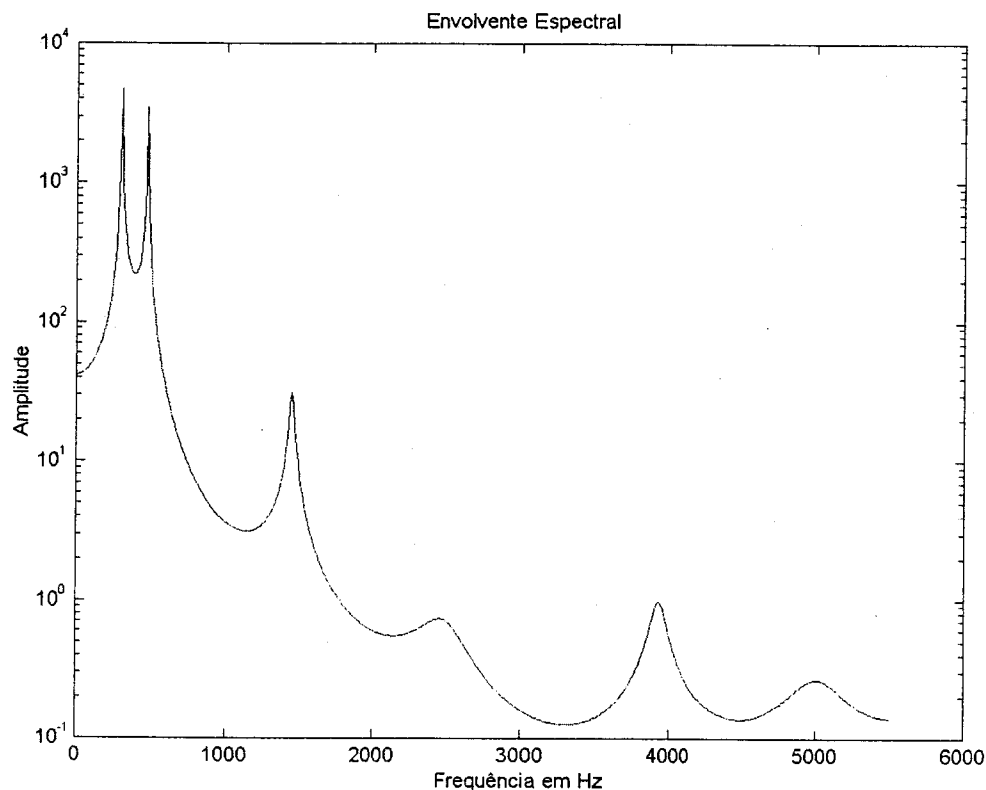


Figura 2.10 – Envolvente espectral, os picos correspondem às formantes

Podemos notar que os coeficientes a_1, a_2, \dots, a_p dependem dos diâmetros dos diversos tubos, isto é da fisiologia do sistema vocal, e assim de orador para orador, por exemplo uma criança e um adulto, fazendo assim variar as frequências de ressonância (Formantes). Nos sistemas de reconhecimento estudados tentamos caracterizar o sinal pelas duas primeiras formantes, pois estas transportam as características do som a ser analisado [Peterson & Barney, 1952], [Bogert & Peterson, 1957], [Kent, 1978] e [Kent, 1979].

No próximo capítulo vamos estudar como obter os valores dos coeficientes a_1, a_2, \dots, a_p a partir de técnicas de processamento de sinal.

Um leitor mais interessado poderá ver com melhor detalhe estas deduções no livro de Rabiner [Rabiner, 1978].

2.4. CONCLUSÕES

Neste capítulo descreveu-se o nosso sistema vocal a partir da fisiologia humana. A seguir descreveu-se um modelo físico que modula a propagação do som (voz) através do nosso sistema vocal, onde foi sugerido um sistema multitubular. Primeiro começou-se por se deduzir as equações de onda num sistema de um só tubo uniforme e depois generalizar para um caso multitubular onde foi apresentada a função de transferência para esse

2- Sistema Fisiológico e Físico e Modelos Matemáticos Aplicados ao processamento de voz

sistema. Foram apresentados como parâmetros para caracterizar a voz as frequências de ressonância denominadas por formantes.

Análise de Sinal e Processamento de Voz

3.1. INTRODUÇÃO

Neste capítulo vamos ver como a teoria do sinal e sistemas se pode relacionar com o modelo físico multitubular apresentado no capítulo anterior.

Iniciamos assim com uma breve descrição da análise do sinal e sistemas, onde se tenta dar uma breve introdução a conceitos básicos da teoria do sinal. Depois introduz-se o modelo LPC (Linear Predictive Coding) que relaciona o modelo físico multitubular do capítulo anterior com as técnicas do processamento do sinal. A seguir explicam-se dois modelos autoregressivos como resolução das equações apresentadas no modelo LPC. O método autoregressivo de Burg e o modelo autoregressivo de Ghael-Sandgate. São apresentadas também as condições de aplicabilidade dos métodos bem como a estabilidade deles.

3.2. BREVE DESCRIÇÃO DA ANÁLISE DE SINAL E SISTEMAS

No capítulo anterior vimos como se poderia modelar fisicamente a propagação de uma perturbação acústica pelo sistema vocal Humano.

Para estudar esse fenómeno há que adquirir dados, isto é, captar por sensores (microfones, etc.) uma medida de grandeza que nos permita estudar esse fenómeno. Assim utilizando por exemplo um microfone à saída da boca podemos captar um sinal (analógico ou digital) que é a conversão da pressão acústica provocada num sensor (microfone) uma diferença de potencial eléctrica, isto é, o microfone ao sentir variações de pressão acústica (mecânica) associa através de uma relação linear a intensidade de pressão com uma determinada tensão ao longo do tempo. Esses valores de tensão podem ser registados de uma forma contínua e assim trata-se de um sinal analógico, ou de uma forma discreta, isto é, são registados valores em intervalos discretos ao longo do tempo e obtendo-se assim um sinal digital.

Um sinal analógico é um registo de valores ao longo do tempo de uma forma continua. Um sinal digital é uma sucessão de valores espaçados ao longo do tempo em intervalos de tempo bem definidos.

Da mesma forma que se definiu um sinal digital podemos definir um sistema, isto é, um sistema diz-se contínuo no tempo se tanto o sinal de entrada como o sinal de saída do sistema forem contínuos no tempo, um sistema diz-se discreto no tempo se tanto o sinal de entrada como o sinal de saída do sistema forem discretos no tempo. Assim podemos definir um sistema digital como um sistema cuja o sinal de entrada é digital e o sinal de saída também é digital.

Neste capítulo vamos tratar de sinais e sistemas digitais. Caso o leitor esteja interessado numa leitura mais detalhada ver [Chen, 1994] [Oppenheim, 1999].

3.3. MODELO LINEAR PREDICTIVE CODING (LPC)

Vimos no capítulo 2 que do ponto de vista Físico uma palavra é uma propagação do som pelo nosso sistema Fisiológico (meio), Glote, Boca, Vias Nasais e Garganta. É gerado um impulso (através dos músculos da glote) que provoca uma perturbação no ar e esta propaga-se da glote à boca, nariz e garganta.

Vimos no ainda capítulo 2 uma forma de modular esta propagação a partir de conceitos físicos e verificou-se que se pode modelar o meio de propagação como sendo um modelo físico de pequenos tubos com determinada impedância. Estas considerações levam-nos a propor um modelo físico linear como aproximação. Assim, poder-se-à considerar o sistema como sendo linear e invariante com relação ao tempo (LTI) [Shaughnessy, 1990] e [Rabiner, 1978].

Considerando o nosso sistema como sendo LTI, este sistema fica descrito pela convolução de duas funções:

$$y(k) = h(k) * u(k) \quad (3.3.1)$$

onde $y(k)$ é o “output” (resposta do sistema à perturbação inicial $u(k)$), sendo $h(k)$ a função de transferência, normalmente inferida de parâmetros físicos conhecidos. Esta função de transferência dá-nos a relação entre o sinal captado (output) e o sinal inicialmente gerado (impulso). O facto deste impulso passar por um meio físico provoca uma alteração e essa alteração é dada por $h(k)$.

No caso em estudo só é conhecido $y(k)$ à saída da boca, pois só usamos um microfone à saída da boca, não conhecemos nem $h(k)$ nem $u(k)$. Existem trabalhos em que se tenta medir o $u(k)$ e a partir de conceitos físicos inferir $h(k)$, [<http://speech.llnl.gov/thesis/index.htm>]. Como os sinais recolhidos (voz) são digitais as nossas funções são séries temporais.

Assim, podemos seguir vários caminhos para a determinação de $h(k)$. Um deles é obter a função de transferência a partir de propriedades físicas inerentes ao sistema, [<http://speech.llnl.gov/thesis/index.htm>]. Outro método usado em muitos trabalhos [Rabiner, 1993] é o método Linear Predictive Coding (LPC) onde se aproxima o sinal $y(k)$ à saída da boca como uma combinação linear dos valores de $y(k)$ passados, como podemos ver em (3.3.2).

$$y(k) = \sum_{i=1}^p a_i \cdot y(k-i) + G \cdot u(k) \quad (3.3.2)$$

Onde o parâmetro G é o ganho e os coeficientes a_i são os coeficientes do filtro digital aplicado ao sinal.

Podemos assim determinar $h(k)$ com base nas equações (3.3.1) e (3.3.2) no domínio das frequências aplicando a transformada-Z (para uma descrição mais detalhada ver [Chen, 1994] e [Shaughnessy, 1990]).

Assim a transformada-z para o método LPC (3.3.2) é:

$$Y(z) = Z(y(k)) = Y(z) \cdot \sum_{i=1}^p a_i \cdot z^{-i} + G \cdot U(z) \quad (3.3.3)$$

e atendendo que a transformada-Z da convolução (3.3.1) é:

$$Z(h(k) * u(k)) = Z(h(k)) \cdot Z(u(k)) \quad (3.3.4)$$

podemos obter:

$$Y(z) = H(z) \cdot U(z) \quad (3.3.5)$$

onde $Y(z)=Z(y(k))$, $U(z)=Z(u(k))$ e $H(z)=Z(h(k))$ e :

$$H(z) = \frac{G}{1 - \sum_{i=1}^p a_i \cdot z^{-i}} \quad (3.3.6)$$

Considerando que os coeficientes a_i podem ser determinados por métodos Autoregressivos (AR) [Shaughnessy, 1990] e [Rabiner, 1978], Pode-se assim escrever $H(z)$ como a divisão de um polinómio complexo;

$$H(z) = \frac{G}{A(z)} \quad (3.3.7)$$

Onde $A(z) = 1 - \sum_{i=1}^p a_i \cdot z^{-i}$.

Vemos que as equações (2.3.12) e (3.3.7) são iguais na sua forma. Se os coeficientes a_i de ambas as equações forem idênticos podemos assim dizer que os modelos são equivalentes [Rabiner, 1978] e calcular os coeficientes por métodos Autoregressivos, para assim definir a função de transferência $H(z)$ (filtro digital).

A partir das condições físicas consideradas no capítulo anterior vemos que este filtro (3.3.6 e 7) modela a geração e propagação de sons audíveis, no estado estacionário, que se propagam pelo sistema vocal apresentado no capítulo anterior. Ficam excluídos os sons nasais.

A vantagem deste modelo está em facilmente podermos calcular os coeficientes a_i e assim poder-se fazer uma análise espectral do sinal, que nos permite por um lado calcular as formantes de um determinado som, e por outro classificar o som a partir deste tipo de parâmetros.

3.4. DETERMINAÇÃO DOS COEFICIENTES $\{a_i\}$

Consideremos que temos uma amostra de uma palavra que poderá ser uma vogal ou um número falado para um microfone (um sinal digital), isto é, uma sucessão de pontos que foram recolhidos por um microfone à saída da boca a uma frequência f_s .

$$y_1, y_2, \dots, y_k, \dots, y_n$$

Vimos que um sistema LPC pode ser representado por uma equação às diferenças (3.3.2):

$$y(k) = \sum_{i=1}^p a_i \cdot y(k-i) + G \cdot u(k)$$

que nos diz que podemos saber o elemento de ordem k da sucessão a partir dos p elementos anteriores a k mais o impulso $u(k)$ inicial gerado pela glote.

Na verdade no caso geral não temos sensores para determinar a forma de $u(k)$, e assim temos de usar um *Linear Predictor*¹, isto é, temos de considerar que $y(k)$ se escreve como uma combinação linear dos p elementos anteriores a k

$$\tilde{y}(k) = \sum_{i=1}^p \alpha_i \cdot y(k-i) \quad (3.4.1)$$

onde α_i são os *Prediction coefficients*.

Acresce que, estamos a cometer um erro $e(k)$ denominado *prediction error* e que é definido por:

$$e(k) = y(k) - \tilde{y}(k) = y(k) - \sum_{i=1}^p \alpha_i \cdot y(k-i) \quad (3.4.2)$$

Da equação (3.4.2) podemos deduzir a função de transferência para a sucessão onde $e(k)$ é o termo:

$$A(z) = 1 - \sum_{i=1}^p \alpha_i \cdot z^{-i} \quad (3.4.3)$$

Se considerarmos que $\alpha_i = a_i$ então temos um erro da ordem $e(k) = G \cdot u(k)$. Para sons audíveis à saída da boca, $u(k)$ representa uma série de impulsos, o que faz com que $e(k)$ seja pequeno na maioria dos casos.

O problema básico restringe-se a determinar os coeficientes α_i de forma a minimizar $e(k)$.

Mais detalhes sobre esta matéria podem ser vistos em [Rabiner, 1978].

3.4.1. MÉTODO AUTOREGRESSIVO DE BURG

Consideremos o método autoregressivo de Burg de ordem p , isto é, de acordo com a equação (2.3.12) podemos dizer que a ordem do método corresponde ao número p de segmentos do sistema tubular, e consideremos a sucessão de valores da amostra do sinal

$$y_1, y_2, \dots, y_{n-p}, \dots, y_n, \dots, y_{n+p}, \dots, y_m$$

¹ Optou-se por usar os termos em Inglês pelo facto de não encontrar uma tradução para o termo predictor.

Convém aplicar ao sinal uma janela para minimizar as descontinuidades nas extremidades do sinal. Normalmente usa-se a janela de Hamming

$$w(k) = \begin{cases} 0.54 - 0.46 \cdot \cos\left(\frac{2\pi \cdot k}{N-1}\right) & 0 \leq k \leq N-1 \\ 0 & \text{fora} \end{cases}$$

Aplicando uma janela em redor do ponto y_n temos

$$0,0,0,\dots,0,y_n,\dots,y_{n+p},\dots,y_{n+N-1},0,\dots,0,0,0$$

isto é,

$$y(k) = y(n+k) \cdot w(k)$$

Por outro lado temos da equação (3.4.3)

$$A^{(p)}(z) = 1 - \sum_{i=1}^p \alpha_i^{(p)} \cdot z^{-i} \quad (3.4.4)$$

onde o índice (p) representa a ordem do método.

Adaptando também o *prediction error* dado pela equação (3.4.2)

$$e^{(p)}(k) = y(k) - \sum_{i=1}^p \alpha_i^{(p)} \cdot y(k-i) \quad (3.4.5)$$

Ou aplicando a transformada-Z,

$$E^{(p)}(z) = A^{(p)}(z) \cdot Y(z) \quad (3.4.6)$$

Dadas as equações anteriores (3.4.4-6) pretendemos calcular os coeficientes α_i recursivamente, onde k_p é o parâmetro recursivo calculado a partir da equação (3.4.16). Para isso definimos α_i da seguinte forma recursiva [Rabiner, 1978],

$$\begin{cases} \alpha_p^{(p)} = k_p \\ \alpha_j^{(p)} = \alpha_j^{(p-1)} - k_p \cdot \alpha_{p-j}^{(p-1)} \quad 1 \leq j \leq p-1 \end{cases} \quad (3.4.7)$$

Substituindo a equação (3.4.7) na equação (3.4.4) e aplicando a transformada-Z obtemos a seguinte formula recorrente

$$A^{(p)}(z) = A^{(p-1)}(z) - k_p \cdot z^{-p} \cdot A^{(p-1)}(z^{-1}) \quad (3.4.8)$$

e substituindo esta ultima equação (3.4.8) na equação (3.4.6) ficamos com

$$E^{(p)}(z) = A^{(p-1)}(z) \cdot Y(z) - k_p \cdot z^{-p} \cdot A^{(p-1)}(z^{-1}) \cdot Y(z) \quad (3.4.9)$$

Vemos que o primeiro termo da equação (3.4.9) é a transformada-Z do *predictor error* de ordem $(p-1)$. Quanto ao segundo termo pode ser dada uma interpretação semelhante

$$B^{(p)}(z) = z^{-p} \cdot A^{(p)}(z^{-1}) \cdot Y(z) \quad (3.4.10)$$

Notando que a transformada-Z inversa da equação (3.4.10) é

$$b^{(p)}(k) = y(k-p) - \sum_{i=1}^p \alpha_i^{(p)} \cdot y(k-p+i) \quad (3.4.11)$$

Vemos que na equação (3.4.5) temos o *predictor error* para a sucessão $y(k)$ e que na equação (3.4.11) temos o *predictor error* para a sucessão $y(k-p)$. Se repararmos nos índices de ambas as equações notamos que são utilizados exactamente os mesmos termos de $y(k)$, isto é, os termos $\{y(k-p), \dots, y(k-1), y(k)\}$ para prever $y(k)$ e $y(k-p)$. Como na equação (3.4.5) usamos os termos de trás para a frente chamamos a esse erro de *forward error* e como a equação (3.4.11) usa os termos da frente para trás chamamos de *backward error*.

Podemos assim exprimir o erro total como função destes dois erros,

$$e^{(p)}(k) = e^{(p-1)}(k) - k_p \cdot b^{(p-1)}(k-1) \quad (3.4.12)$$

Substituindo a equação (3.4.8) na equação (3.4.10) obtemos,

$$B^{(p)}(z) = z^{-p} \cdot A^{(p-1)}(z^{-1}) \cdot Y(z) - k_p \cdot A^{(p-1)}(z) \cdot Y(z)$$

ou

$$B^{(p)}(z) = z^{-1} \cdot B^{(p-1)}(z) - k_p \cdot E^{(p-1)}(z) \quad (3.4.13)$$

Ficando assim com o *backward prediction error*

$$b^{(p)}(k) = b^{(p-1)}(k-1) - k_p \cdot e^{(p-1)}(k) \quad (3.4.14)$$

Atendendo aos resultados obtidos Burg [Burg, 1968] desenvolveu um procedimento que nos permite calcular os α_i a partir do *forward* e *backward errors*, minimizando a soma dos quadrados destes erros. Temos assim,

$$Er^{(p)} = \sum_{k=0}^{N-1} \left[\left(e^{(p)}(k) \right)^2 + \left(b^{(p)}(k) \right)^2 \right] \quad (3.4.15)$$

Substituindo as equações (3.4.12) e (3.4.14) na equação (3.4.15) e diferenciando em ordem a k_p e igualando a zero,

$$\frac{\partial Er^{(p)}}{\partial k_p} = 0$$

obtemos,

$$k_p = \frac{2 \cdot \sum_{k=0}^{N-1} \left[e^{(p-1)}(k) \cdot b^{(p-1)}(k-1) \right]}{\sum_{k=0}^{N-1} \left[e^{(p-1)}(k) \right]^2 + \sum_{k=0}^{N-1} \left[b^{(p-1)}(k-1) \right]^2} \quad (3.4.16)$$

Pode-se demonstrar que [Rabiner, 1978] os k_p estimados a partir da equação (3.4.16) estão compreendidos no intervalo,

$$-1 \leq k_p \leq 1 \quad (3.4.17)$$

O procedimento proposto consiste no seguinte algoritmo:

1. Inicializar o conjunto de valores $e^{(0)}(k) = y(k) = b^{(0)}(k)$.
2. Calcular $k_1 = \alpha_1^{(1)}$ pela equação (3.4.16).



3. Determinar os erros *forward e backward prediction errors* $e^{(1)}(k)$ e $b^{(1)}(k)$ pelas equações (3.4.12) e (3.4.14).
4. Faça $p=2$.
5. Calcular $k_p = \alpha_{p1}^{(p)}$ pela equação (3.4.16).
6. Determinar $\alpha_j^{(p)}$ para $j = 1, 2, \dots, p-1$ pela equação (3.4.7).
7. Determinar os erros *forward e backward prediction errors* $e^{(p)}(k)$ e $b^{(p)}(k)$ pelas equações (3.4.12) e (3.4.14).
8. Faça $p = p + 1$.
9. Se $p \leq \text{ordem do método}$ vá para o passo 5.
10. Fim do procedimento.

3.4.2. ESTABILIDADE DO MÉTODO DE BURG

Diz-se que [Rabiner, 1978] o método Autoreressivo de Burg é estável se os pólos da equação (3.3.7) estiverem dentro do raio de convergência unitário do plano complexo da transformada-Z.

$$H(z) = \frac{G}{A(z)} \quad (3.3.7)$$

Isto é, se os zeros do polinómio

$$A(z) = 1 - \sum_{i=1}^p \alpha_i \cdot z^{-i} \quad (3.4.3)$$

Estiverem dentro do raio de convergência unitário.

Existe um teste neste método que nos permite ver se o sistema é estável. Basta-nos verificar se os k_p estimados a partir da equação (3.4.16) estão compreendidos no intervalo,

$$-1 \leq k_p \leq 1 \quad (3.4.17)$$

Caso contrário o sistema é instável, encontrando-se zeros fora do raio de convergência.

3.4.3. CONSIDERAÇÕES DE APLICABILIDADE DO MÉTODO AUTOREGRESSIVO DE BURG

Existem duas considerações a fazer para a utilização do método de Burg. A ordem p que devemos usar para estimar os coeficientes α_i e o comprimento da janela utilizada.

A escolha de p depende da frequência de amostragem utilizada. Assim devemos escolher uma ordem p de forma a termos uma densidade média de 2 pólos (pólos complexos) por 1Khz para representar a contribuição do sistema vocal [Rabiner, 1978]. A título de exemplo, se tivermos uma frequência de amostragem de 11025 Hz devemos usar $p = \frac{11025}{1000} \approx 11$.

Normalmente fazem-se algumas correcções usando para o caso anterior um p entre 11 a 13 [Rabiner, 1978].

No caso do comprimento da janela, N (numero de pontos da sucessão), devem usar-se comprimentos suficientemente grandes de forma a suavizar os efeitos da janela imposta, de forma a não prejudicar os resultados e suficientemente pequena para minimizar o erro cometido [Rabiner, 1978]. Assim devem usar-se janelas da ordem de $N=400$ amostras para uma frequência de amostragem de 11025Hz.

3.4.4. MÉTODO AUTOREGRESSIVO DE GHAEL-SANDGATE

O método de Ghael-Sandgate consiste em calcular os coeficientes α_i a partir do método de Burg e depois usa uma medida de escala para afastar os pólos consecutivos que se encontram mais próximos, recalculando novamente esses coeficientes de acordo com uma medida de escala.

O método torna-se útil quando temos pólos muito próximos um do outro fazendo com que não seja detectado uma formante, como podemos ver na figura 3.1. o método de Burg não consegue detectar a formante F1 nem F2 enquanto que o método de Ghael-sandgate detecta essas duas formantes no local certo.

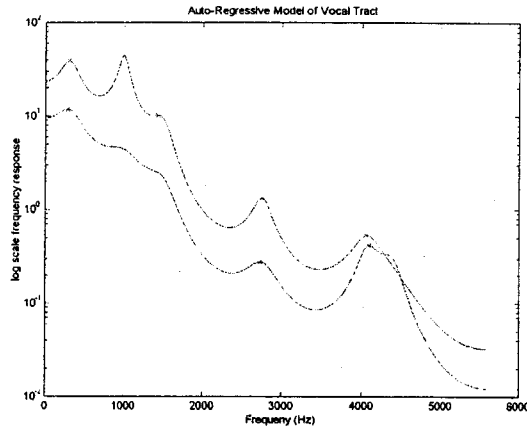


Figura 3.1- Dois picos detectados pelo método 2, de Ghael-Sandgate (envolvente superior) que não são detectados pelo método 1, método de Burg (envolvente inferior).

3.4.4.1. DESCRIÇÃO DO MÉTODO AUTOREGRESSIVO DE GHAEL-SANDGATE

O método de Ghael-Sandgate é uma variante do método de Burg. Vejamos as diferenças a partir da descrição do método.

1º Sejam α_i os coeficientes LPC determinados pelo método de Burg. Temos assim:

$$A(z) = 1 - \sum_{i=1}^p \alpha_i z^{-i}$$

Determinemos os zeros de $A(z)$,

$$1 - \sum_{i=1}^p \alpha_i z^{-i} = 0$$

$$z^p \cdot \left(1 - \sum_{i=1}^p \alpha_i z^{-i} = 0 \right)$$

$$z^p - \sum_{i=1}^p \alpha_i z^{p-i} = 0$$

Seja $Rz = \{z_1, z_2, \dots, z_p\}$ o conjunto que contem os p-pólos de $A(z)$, e seja

$Pz = \{z'_1, z'_2, z'_3, \dots, z'_n\}$ com $n \leq \frac{p}{2}$ o conjunto de pólos da $A(z)$ cujos

zeros têm a componente imaginária superior a zero. Os pólos aparecem conjugados, pela teoria da análise complexa, não sendo necessário analisar todos os pólos.

2º Calcula-se agora o conjunto $D = \{d_1, d_2, \dots, d_{n-1}\}$ o conjunto das distâncias entre pólos consecutivos, isto é, $d_i = |z'_{i+1} - z'_i|$.

3º seja $m = \min(D) = d_i$.

4º Identificado o d_i mínimo sabemos quais os pólos que se encontram mais próximos, estabelecemos assim um factor escala s , da seguinte forma:

$$R1 = |z'_i|$$

$$R2 = |z'_{i+1}|$$

$$s = \frac{1}{\left(\frac{R2 - R1}{2}\right) + R1}$$

5º Multipliquemos todas as raízes de $A(z)$, isto é, todos os elementos de Rz por esse factor escala obtendo um novo conjunto de pólos $s \cdot Rz = \{w_1, w_2, \dots, w_p\}$ onde $w_i = s \cdot z_i$. Obtemos assim

$$P'(z) = (z - w_1)(z - w_2) \cdots (z - w_p) = z^p - \sum_{i=1}^p \alpha'_i z^{p-i} \quad \text{ou de uma outra}$$

forma:

$$A'(z) = 1 - \sum_{i=1}^p \alpha'_i z^{-i}$$

Onde os α'_i são os novos coeficientes LPC.

As condições de aplicabilidade e de estabilidade do modelo de Ghael-Sandgate obedecem as mesmas regras do modelo de Burg.

3.5. CONCLUSÕES

Neste capítulo viu-se como a teoria do sinal se pode relacionar com o que foi apresentado no capítulo anterior.

Iniciou-se com uma breve descrição da análise do sinal e sistemas, onde se tentou dar uma breve introdução a conceitos básicos da teoria do sinal e sistemas. Depois introduziu-se o modelo LPC (Linear Predictive Coding) que relaciona o modelo físico multitubular do capítulo anterior com as técnicas do processamento do sinal. A seguir explicaram-se dois modelos autoregressivos como resolução das equações apresentadas no modelo

LPC. O método autoregressivo de Burg e o modelo autoregressivo de Ghael-Sandgate. Foram apresentadas também as condições de aplicabilidade dos métodos tais como também a estabilidade destes métodos.

Conjuntos Difusos e Modelos de Decisão Múltiplo Atributo Difusos

4.1. INTRODUÇÃO

Este capítulo está dividido em duas partes, uma em que se faz um resumo geral da teoria dos conjuntos difusos [Zadeh, 1973] e outra parte onde se faz um resumo da teoria da Tomada de Decisão com critérios múltiplos [Ribeiro, 1996].

Na primeira parte deste capítulo far-se-á uma breve introdução ao conceito de conjunto difuso. Depois serão definidas as estruturas matemáticas básicas das operações entre conjuntos difusos (negação, intersecção e união) e finalmente será definida uma estrutura matemática mais generalizada para as operações de agregação entre conjuntos difusos.

Na segunda parte deste capítulo far-se-á uma abordagem breve e geral sobre a Tomada de Decisão Múltiplo Critério: com especial ênfase na Tomada de Decisão Múltiplo Atributo. Serão apresentados detalhes sobre os métodos mais usados na Tomada de Decisão Múltiplo Atributo, em particular o método AHP (Analytical Hierarchy Process), [Saaty, 1978] introduzido por Saaty, e a sua extensão, utilizando os operadores de agregação OWA (Ordered Weighted Averaging) [Yager, 1999]. Terminamos apresentando algumas das vantagens da escolha e uso do método do AHP estendido com operadores OWA.

4.2. CONJUNTOS DIFUSOS

Nesta secção apresentamos a noção de conjunto difuso, [Klir, 1988] fazendo o seu paralelismo com os conjuntos “crisp” (secção 4.2.1) e como podemos fazer operações com conjuntos difusos (secção 4.2.2).

4.2.1. NOÇÃO DE CONJUNTO DIFUSO

Quando tentamos quantificar ou qualificar algo, trabalhamos dentro de um universo de objectos. No caso matemático ensinaram-nos que na parte quantitativa poderíamos trabalhar com Universos de conjuntos de objectos tais como os números Naturais, Inteiros; Racionais, Irracionais, Reais, Complexos, e outros. Dentro de cada universo poderíamos trabalhar com

subconjuntos desses conjuntos, e esses subconjuntos poderiam ser definidos numa forma compreensiva, isto é, $C1=\{x\in\text{Universo: Proposição lógica}\}$, ou de uma forma extensiva, $C2=\{1,2,3,4,5,6,7,8,9\}$ e dentro desse universo, seja denotado por U , havia elementos que pertenciam a $C1$ e a $C2$ e elementos que pura e simplesmente não pertenciam a $C1$ e $C2$, isto é, dentro de todos os elementos de U ou é *verdadeiro* ou *falso* que um determinado elemento pertença ou não a um subconjunto. Consideremos o seguinte exemplo:

Exemplo 4.1: seja $U=\{1,2,3,4,5,6,7,8,9,10,11,12,13,14,15\}$ o nosso universo e seja $C1=\{x\in U: x \text{ é menor que } 10\}$ então temos na realidade o seguinte conjunto:

$$C1=\{1/1,2/1;3/1,4/1,5/1,6/1,7/1,8/1,9/1,10/0,11/0,12/0,13/0,14/0,15/0\},$$

onde $x/1$ significa que x pertence ao conjunto $C1$ e $x/0$ significa que x não pertence ao conjunto $C1$.

Podemos definir os conjuntos “crisp” da seguinte forma:

Definição 4.1(a): Seja x um objecto genérico de uma colecção de objectos de U . Então um conjunto “Crisp” C sobre U define-se como um conjunto de pares ordenados:

$$C=\{(x,\mu_C(x)): x\in U\}$$

onde $\mu_C(x)$ se designa por função de pertença e representa o grau de pertença de x em C , que no presente caso tem valores $\mu_C(x):U\rightarrow\{0,1\}$.

A definição de conjunto difuso pode-se considerar como uma generalização da de conjunto “crisp” e uma das formas de definir um conjunto difuso é:

definição 4.1(b): [Pires, 1991] Seja x um objecto genérico de uma colecção de objectos de U . Então um conjunto difuso C sobre U define-se como um conjunto de pares ordenados:

$$C=\{(x,\mu_C(x)): x\in U\}$$

onde $\mu_C(x)$ se designa por função de pertença e representa o grau de pertença de x em C ou ainda o grau de compatibilidade entre o elemento x e o conceito representado pelo conjunto C . Esta função é definida de U em M , $\mu_C(x):U\rightarrow M$, onde M se designa por espaço de pertença e é um subconjunto de números reais não negativos, com supremo finito no intervalo $[0,1]$. O conjunto U é designado por conjunto universal.

No caso de o conjunto C ser “crisp” $\mu_C: U \rightarrow \{0,1\}$, isto é, como apresentamos nos exemplo 1 o elemento x de U pertencia ou não pertencia ao conjunto C , no caso difuso podemos dizer que um determinado elemento $x \in U$ pode pertencer a C com um grau de verdade entre $[0,1]$. No fundo atribuímos uma gama de valores de veracidade a um determinado elemento de U , respectivamente a um conjunto C difuso definido na forma anterior.

Tal como acontecia com os conjuntos “crisp”, estes podem ser discretos ou contínuos. Para diferenciar os discretos dos contínuos usa-se a seguinte notação.

Notação 1- se o conjunto A for discreto podemos representar os seus membros da seguinte forma:

$$A = \sum (x_i / \mu_i). \quad (4.2.1)$$

Notação 2- se o conjunto A for contínuo podemos representar os seus membros da seguinte forma:

$$A = \int_U x / \mu_A(x). \quad (4.2.2)$$

Temos no exemplo 4.2 um conjunto difuso discreto e no exemplo 4.3 um conjunto difuso contínuo.

Exemplo 4.2: seja $U = \{1,2,3,4,5,6,7,8,9,10,11,12,13,14,15\}$ o nosso universo e seja C_1 o nosso conjunto difuso com $\mu_{C_1}: U \rightarrow [0,1]$ definido da seguinte forma

$$C_1 = \{1/0, 2/0.2, 3/0.3, 4/0.5, 5/0.8, 6/1, 7/0.7, 8/0.4, 9/0.2, 10/0, 11/0, 12/0, 13/0, 14/0, 15/0\},$$

Onde por exemplo $\mu_{C_1}(4) = 0.5$, o que quer dizer que o grau de veracidade com que o elemento 4 pertence a C_1 é de 0.5.

Exemplo 4.3: Vejamos outro exemplo, seja U o conjunto dos números Reais e $C_1 = \{(x, \mu_{C_1}) \in U \times [0,1]: \text{o numero está perto de } 2\}$, onde μ_{C_1} representa o grau de pertença do objecto x a U , isto é, onde $\mu_{C_1}: U \rightarrow [0,1]$. Temos assim o conjunto difuso C_1 definido por $\mu_{C_1}(x)$ da seguinte forma:

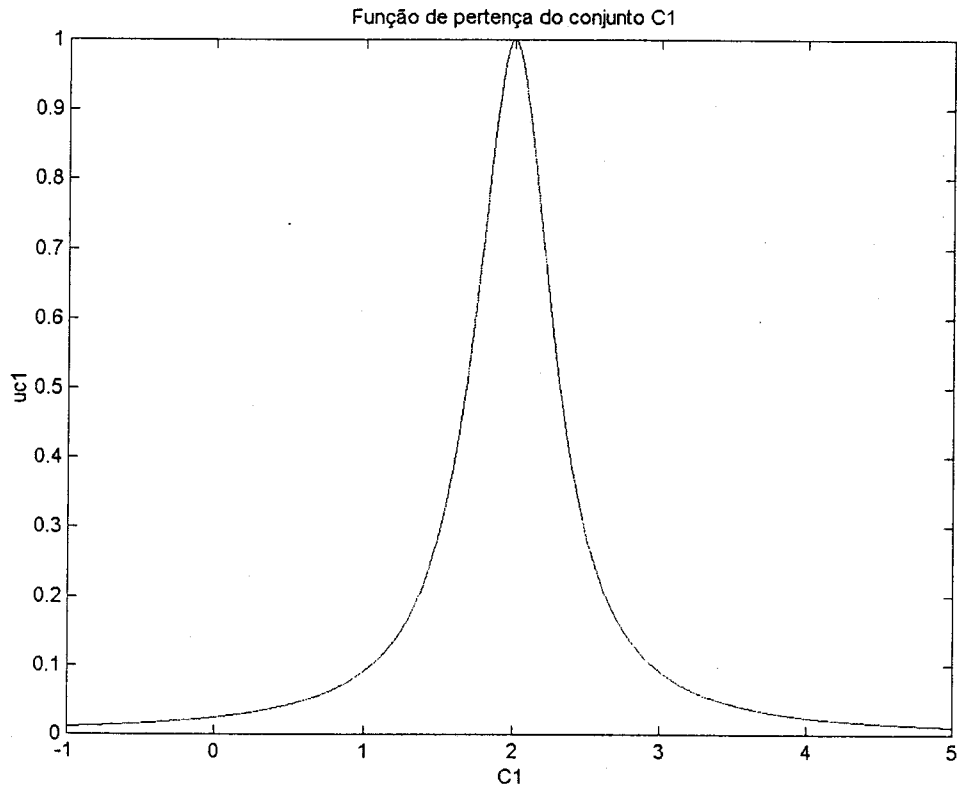


Figura 4.1- O número está perto de 2

Consideremos um último exemplo retirado de [Klir, 1988] que nos será útil para as definições que iremos fazer a seguir sobre conjuntos difusos.

Exemplo 4.4: Consideremos o seguinte universo das idades $U=\{5,10,20,30,40,50,60,70,80\}$ e os seguintes conjuntos difusos, Criança, Adulto, Jovem e Idoso, tabelados na seguinte tabela:

Elementos	Criança	Adulto	Jovem	Idoso
5	0	0	1	0
10	0	0	1	0
20	0	0.8	0.8	0.1
30	0	1	0.5	0.2
40	0	1	0.2	0.4
50	0	1	0.1	0.6
60	0	1	0	0.8
70	0	1	0	1
80	0	1	0	1

Onde por exemplo o conjunto difuso discreto adulto é definido por,

$$\text{Adulto}=\{5/0,10/0,20/0.8,30/1,40/1,50/1,60/1,70/1,80/1\}.$$

Passemos agora a descrever alguns conceitos básicos e propriedades dos conjuntos difusos. Usou-se como base desta descrição o livro de Klir [Klir, 1988].

Começemos por definir o conjunto de todos os subconjuntos de um conjunto “crisp”, denominado em inglês por “*power set*”:

Definição 4.2- Power set: Seja X um conjunto “crisp” então ao conjunto de todos os subconjuntos de X chama-se *power set* e designa-se por $P(X)$.

Definição 4.3- Definimos cardinalidade de um conjunto X pelo numero de elementos que esse conjunto tem e designamos por $\#X$.

Exemplo 4.5: Consideremos $X=\{1,2,3\}$, então o *power set* de X é o conjunto $P(X)=\{\emptyset,\{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\},\{1,2,3\}\}$ com a cardinalidade $\#P(X)=8$.

Definição 4.4- Power set difuso: Seja X um conjunto “crisp” então ao conjunto de todos os subconjuntos de X difusos chama-se *power set difuso* do conjunto X e designa-se por $\tilde{P}(X)$.

Notar que nesta ultima *definição 4.4* o conceito de *power set difuso* de um conjunto X é um pouco mais “qualitativa” de que “quantitativa” pois representa todos os conjuntos difusos que podemos formar a partir do conjunto X , (o nosso universo para o conjunto difuso), e a função de pertença já não é agora simplesmente $\mu_X: U \rightarrow \{0,1\}$ mas sim $\mu_X: U \rightarrow [0,1]$. De seguida vamos definir o suporte de um conjunto difuso C .

Definição 4.5- Chama-se suporte de um conjunto difuso ao conjunto definido da seguinte forma:

$$\text{sup} : \tilde{P}(X) \rightarrow P(X), \text{ tal que } \text{sup}(C) = \{x \in X : \mu_c(x) > 0\}.$$

Exemplo 4.6: Os suportes dos conjuntos difusos Criança e Jovem do exemplo 4.4 são:

$$\text{sup}(\text{Criança})=\emptyset \text{ e } \text{sup}(\text{Jovem})=\{5,10,20,30,40,50\}.$$

Definição 4.6- Chama-se *altura* de um conjunto difuso C ao valor de pertença mais alto dos elementos de C .

Definição 4.7- Um conjunto difuso C diz-se normalizado se $altura(C)=1$.

Definição 4.8- Chama-se α -cut de um conjunto difuso C ao conjunto definido da seguinte forma $C_\alpha = \{x \in U : \mu_C(x) \geq \alpha\}$.

Exemplo 4.7: Consideremos de novo o conjunto Jovem do exemplo 4.4, temos que $Jovem_{0,2} = \{5,10,20,30,40\}$ e $Jovem_{0,8} = \{5,10,20\}$, respectivamente para os α -cuts de 0.2 e 0.8.

Definição 4.9- Ao conjunto de todos os níveis $\alpha \in [0,1]$ de um conjunto difuso C , chama-se conjunto dos níveis de C e define-se da seguinte forma:

$$\Lambda_C = \{\alpha : \mu_C(x) = \alpha \text{ para algum } x \in U\}.$$

Definição 4.10- Um conjunto $X \subset \mathcal{R}^n$ "crisp" diz-se convexo se para qualquer par de pontos x e y de X e para todo o real λ satisfazendo, $0 < \lambda < 1$, temos $\lambda x + (1-\lambda)y$ pertence a X . [Apostol, 1978]. Isto é todos os pontos pertencentes à recta que une os pontos x e y de X também pertencem ao conjunto X .

Definição 4.11- Um conjunto C difuso diz-se convexo se e só se cada um dos conjuntos α -cut's for um conjunto convexo. Equivalentemente um conjunto C é convexo se e só se se para qualquer par de pontos x e y de C e para todo o real λ satisfazendo, $0 < \lambda < 1$, temos,

$$\mu_C(\lambda x + (1-\lambda)y) \geq \min[\mu_C(x), \mu_C(y)].$$

É de notar que esta ultima definição não obriga que a função de pertença seja convexa. Por exemplo o conjunto difuso definido pela função de pertença definida na figura 4.2 não é convexa mas o conjunto difuso é, no caso figura 4.3 a função de pertença não é convexa e o conjunto difuso também não o é.

4- Conjuntos Difusos e Modelos de Decisão Múltiplo Atributo Difusos

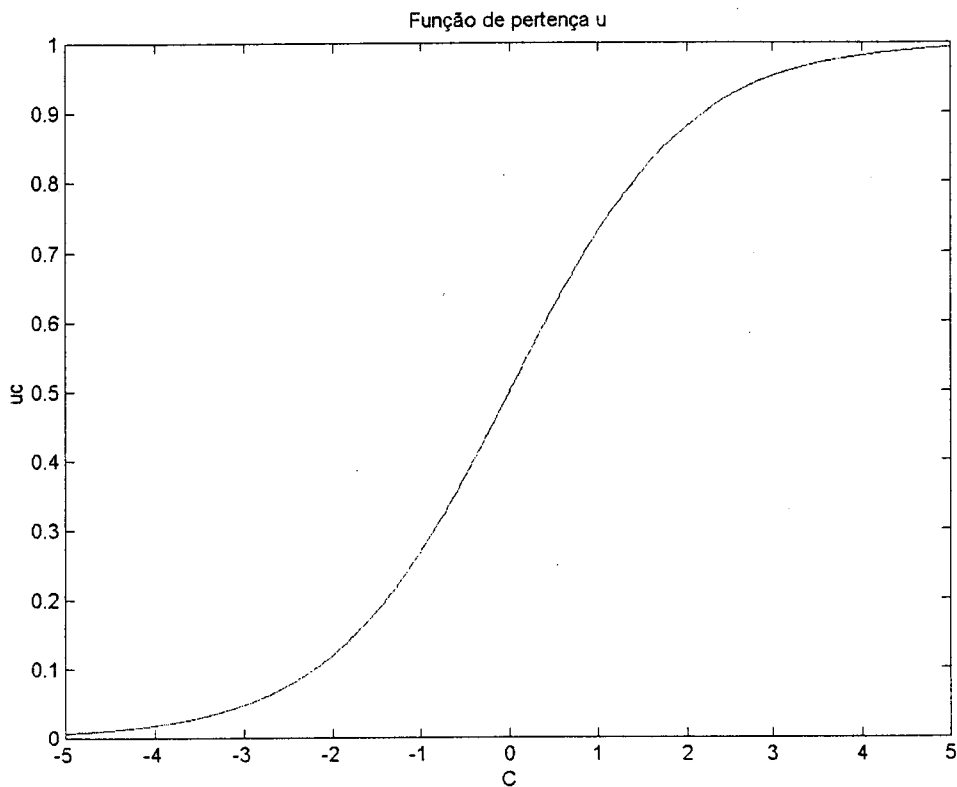


Figura 4.2- Conjunto difuso convexo e função de pertinência não convexa

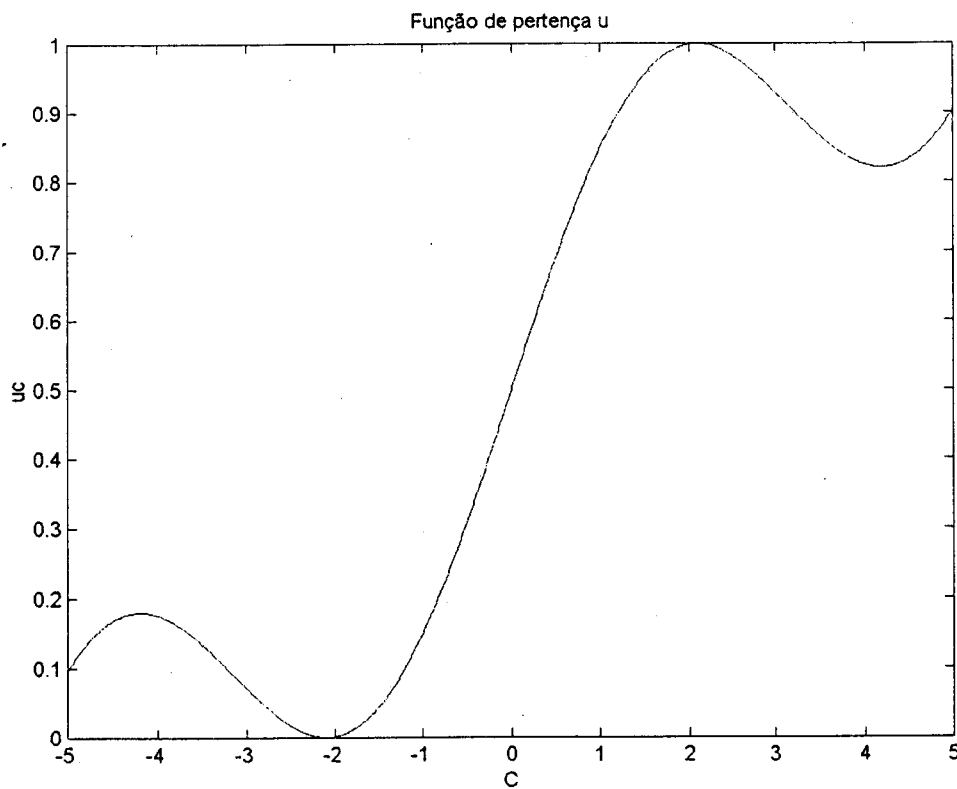


Figura 4.3- Conjunto difuso não convexo e função de pertinência não convexa

Com esta ultima definição podemos definir um conceito importante na teoria dos conjuntos difusos, que é o conceito de número difuso.

Definição 4.12- Um número diz-se difuso se for representado por um conjunto difuso normalizado e convexo.

Um exemplo de um número difuso é o conjunto representado na figura 4.1, que pode representar o número dois.

De seguida vão-se definir alguns conceitos relacionados com operações entre conjuntos difusos que serão alvo de maior aprofundamento na secção 4.2.2.

Definição 4.13- Diz-se que $A \subseteq B$ sendo A e B dois conjuntos difusos se para qualquer $x \in U$ temos, $\mu_A \leq \mu_B$.

Definição 4.14- Diz-se que $A = B$ sendo A e B dois conjuntos difusos se para qualquer $x \in U$ temos, $\mu_A = \mu_B$.

Definição 4.15- Diz-se que $A \neq B$ sendo A e B dois conjuntos difusos se existe pelo menos um $x \in U$ temos, $\mu_A \neq \mu_B$.

Definição 4.16- Assim temos que $A \subset B$, sendo A e B difusos se e só se $A \subseteq B$ e $A \neq B$.

A seguir apresenta-se o *princípio da extensibilidade* [Zadeh, 1965][Zimmermann, 1991] que nos permite generalizar o conceito de função para o campo dos conjuntos difusos e aritmética difusa.

Princípio da extensibilidade – Seja X o produto cartesiano dos universos $X = X_1 \times X_2 \times \dots \times X_r$, e A_1, A_2, \dots, A_r são r conjuntos difusos em X, respectivamente. Seja f uma função de X para o universo Y, $y = f(x_1, x_2, \dots, x_r)$. Então o princípio da extensibilidade permite-nos definir um conjunto difuso B em Y por:

$$B = \{(y, \mu_B(y)) : y = f(x_1, x_2, \dots, x_r), (x_1, x_2, \dots, x_r) \in X\}$$

onde,

$$\mu_B(y) = \begin{cases} \sup_{(x_1, x_2, \dots, x_r) \in X} \min\{\mu_{A_1}(x_1), \dots, \mu_{A_r}(x_r)\} & , f^{-1}(y) \neq \emptyset \\ 0 & f^{-1}(y) = \emptyset \end{cases}$$

onde f^{-1} é a inversa de f .

Este princípio é muito importante para a extensão das operações aritméticas com números difusos tais como a extensão à análise matemática dos números difusos.

4.2.2. OPERAÇÕES SOBRE CONJUNTOS DIFUSOS

Os conjuntos difusos servem para exprimir um certo tipo de incerteza, como por exemplo se tivermos um atributo como “bonito” existe associado a este tipo de atributo uma certa incerteza pois não está bem definido. Outro exemplo é o caso de termos uma variável (atributo) de um determinado sistema físico e devido aos vários erros cometidos pelos instrumentos de medida, algoritmos, humanos, etc., termos uma determinada incerteza em relação ao valor a considerar.

Quando temos diversos atributos num sistema a estudar é necessário combinar os diversos atributos difusos e para isso usamos operações entre conjuntos difusos.

Os operadores podem ser definidos em três classes, operadores de união, intersecção e operadores de agregação. Aos operadores de união designam-se na literatura por conormas triangulares ou t-conormas, aos de intersecção por normas triangulares ou t-normas [Zimmermann, 1991]. Estes operadores podem ser classificados em paramétricos ou não paramétricos [Zimmermann, 1991], isto é, se na sua definição dependem de um parâmetro ou não, respectivamente. Veremos nas secções 4.2.2.2. e 4.2.2.3. exemplos destes operadores.

Nesta secção vamos rever brevemente os diversos tipos de operações e seguiremos principalmente o livro de Klir [Klir, 1988].

4.2.2.1. COMPLEMENTO DIFUSO

O complemento de um conjunto difuso A é uma função c que associa a cada elemento de A uma nova função de pertença dada por $c(\mu_A(x))$. Em geral, temos;

Definição 4.17- A função complemento de um conjunto difuso A é definida por: $c:[0,1] \rightarrow [0,1]$ tal que $c(\mu_A(x)) = \mu_{\bar{A}}(x)$. Onde \bar{A} representa o complemento de A .

Resta-nos agora definir a classe de funções c que trazem alguma coerência à noção de complemento de um conjunto difuso. Para isso vamos definir uma axiomática desta classe de funções. Assim para que uma determinada função possa ser considerada complemento de um conjunto difuso basta obedecer aos dois seguintes axiomas.

Consideremos c uma função de acordo com a definição 4.17 tal que obedeça aos seguintes axiomas:

Axioma c1- $c(0)=1$ e $c(1)=0$.

Axioma c2- $\forall a,b \in [0,1]$, se $a < b$ então $c(a) \geq c(b)$, isto é c é monótona decrescente.

A classe de funções definidas pela definição 4.17 e pelos axiomas $c1$ e $c2$ definem a classe mais geral de funções complemento, seja designada por C_1 essa classe (nota não fazer confusão com a classe das funções contínuas diferenciáveis de 1º ordem). Podemos acrescentar mais dois axiomas de forma a definir sub classes de C_1 .

Axioma c3- c é uma função contínua.

Estes 3 axiomas definem uma subclasse de C_1 que designaremos por C_2 , tal que $C_2 \subset C_1$.

Axioma c4- $\forall a \in [0,1]$, $c(c(a))=a$.

Os quatro axiomas definem uma classe de funções designada por C_3 , tal que $C_3 \subset C_2 \subset C_1$.

Exemplo 4.8- Um exemplo de uma função que obedece aos quatro axiomas é a definição inicial de complemento, $c(x)=1-c(x)$, que facilmente se verifica que obedece aos axiomas.

Exemplo 4.9- Um exemplo de uma função que obedece aos três primeiros axiomas é a função $c(x) = \frac{1}{2}(1 + \cos(\pi x))$. Facilmente se vê que não obedece ao axioma $c4$.

Exemplo 4.10- Um exemplo de uma função que só obedece aos axiomas fundamentais c1 e c2 é a função $c(x) = \begin{cases} 1 & x \leq t \\ 0 & x \geq t \end{cases}$.

Existem 2 classes de funções que obedecem aos quatro axiomas anteriores e são importantes de considerar que é a classe de Sugeno e a classe de Yager.

Classe de Sugeno: seja $\lambda \in (-1, \infty)$ tal que $c_\lambda(x) = \frac{1-x}{1+\lambda \cdot x}$.

Classe de Yager: seja $w \in (0, \infty)$ tal que $c_w(x) = (1-x^w)^{1/w}$.

Podemos ver na figura 4.4 e 4.5 representadas as classes de Sugeno e de Yager respectivamente.

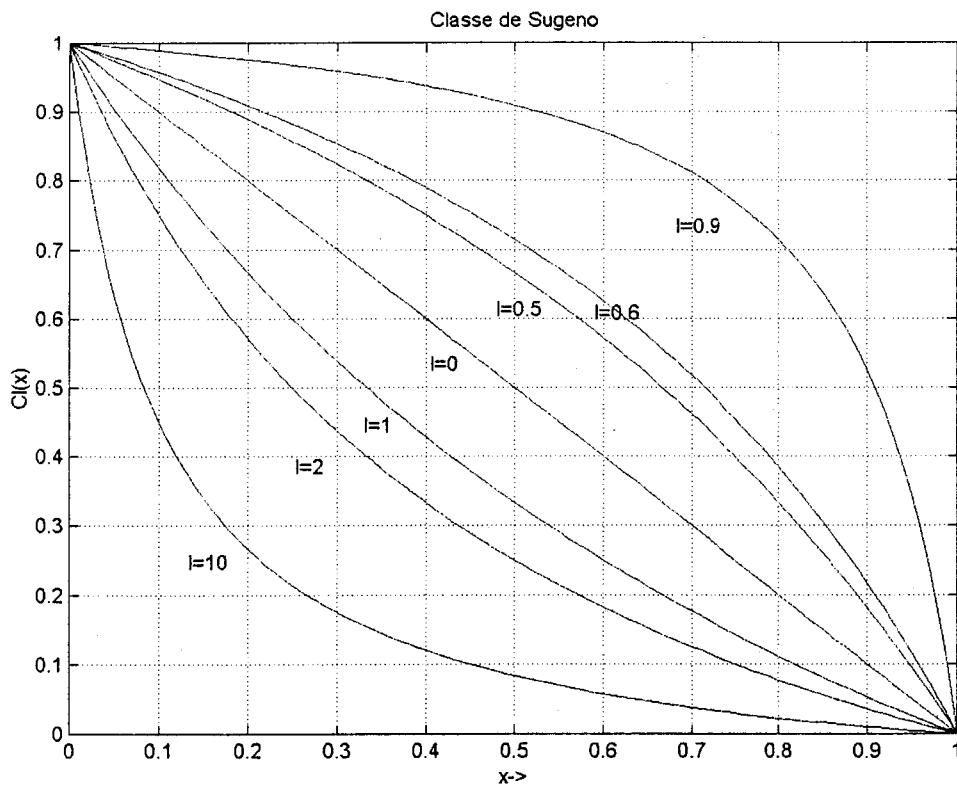


Figura 4.4- Classe de Sugeno [Klir, 1988]

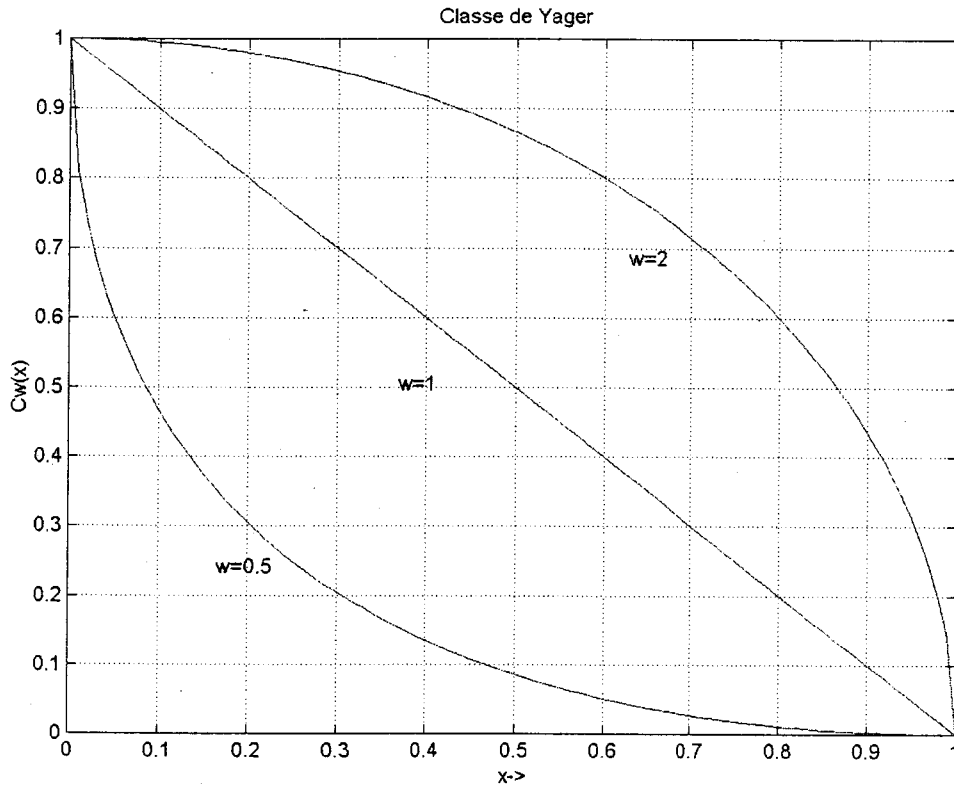


Figura 4.5- Classe de Yager [Klir, 1988]

Se o leitor desejar um estudo mais aprofundado sobre a função complementar pode consultar [Klir, 1988].

4.2.2.2. UNIÃO DIFUSA

A união de dois conjuntos difusos é definida por uma função u de acordo com a definição:

Definição 4.18- A função união de dois conjuntos difusos A e B é definida por: $u: [0,1] \times [0,1] \rightarrow [0,1]$ tal que $u(\mu_A(x), \mu_B(x)) = \mu_{A \cup B}(x)$. A este tipo de função chama-se também de t-conorma ou conormas triangulares.

Para que uma função u seja considerada como uma função de união ou t-conorma, de dois conjuntos difusos tem de obedecer aos quatro axiomas que se seguem.

Axioma u1- $u(0,0)=0$ e $u(0,1)=u(1,0)=u(1,1)=1$.

Axioma u2- u é comutativa, isto é, $u(a,b)=u(b,a)$.

Axioma u3- u é monótona, isto é, se $a \leq a'$ e $b \leq b'$ então $u(a,b) \leq u(a',b')$.

Axioma u4- u é associativa, isto é, $u(u(a,b),c)=u(a,u(b,c))$.

À definição 4.18 e ao conjunto dos axiomas $u1$ - $u4$ definem a classe principal de funções de união entre conjuntos difusos ou t -conormas. Podemos adicionar mais dois axiomas a esta classe obtendo assim subclasses desta última, que são.

Axioma u5- u é contínua.

Axioma u6- u é idempotente, $u(a,a)=a$.

Na tabela 4.1 podemos ver algumas funções união paramétricas e não paramétricas obedecem aos axiomas $u1$ - $u4$ mas poderão não obedecer a $u5$ - $u6$.

References	Intersection operators <i>t</i> -norms	Averaging operators	Union operators <i>t</i> -conorms
	non-parametrized		
Zadeh 1965	minimum algebraic product		maximum algebraic sum
Giles 1976	bounded sum		bounded difference
Hamacher 1978	Hamacher product		Hamacher sum
Mizumoto 1982	Einstein product		Einstein sum
Dubois and Prade 1980, 1982	drastic product		drastic sum
Dubois and Prade 1984		arithmetic mean geometric mean	
Silvert 1979		symmetric summation and differences	
parametrized families			
Hamacher 1978	Hamacher-intersection- operators		Hamacher-union-operators
Yager 1980	Yager-intersection- operators		Yager-union-operators
Dubois and Prade 1980a, 1982a, 1984	Dubois-intersection- operators		Dubois-union-operators
Werners 1984		"fuzzy and", "fuzzy or"	
Zimmermann and Zysno 1980		"compensatory and", convex comb. of maximum and minimum, or algebraic product and algebraic sum	

Tabela 4.1- t -normas e t -conormas [Zimmermann, 1991]

4.2.2.3. INTERSECÇÃO DIFUSA

A intersecção de dois conjuntos difusos é definida por uma função i de acordo com a definição:

Definição 4.19- A função intersecção de dois conjuntos difusos A e B é definida por: $i:[0,1] \times [0,1] \rightarrow [0,1]$ tal que $u(\mu_A(x), \mu_B(x)) = \mu_{A \cap B}(x)$. A este tipo de função chama-se também de t-norma ou normas triangulares.

Para que uma função i seja considerada como uma função de intersecção ou t-norma de dois conjuntos difusos tem de obedecer aos quatro axiomas que se seguem.

Axioma i1- $i(1,1)=1$ e $i(0,1)=i(1,0)=i(1,1)=0$.

Axioma i2- i é comutativa, isto é, $i(a,b)=i(b,a)$.

Axioma i3- i é monótona, isto é, se $a \leq a'$ e $b \leq b'$ então $i(a,b) \leq i(a',b')$.

Axioma i4- i é associativa, isto é, $i(i(a,b),c)=i(a,i(b,c))$.

A definição 4.19 e o conjunto dos axiomas i1-i4 definem a classe principal de funções de intersecção entre conjuntos difusos.

Podemos adicionar mais dois axiomas a esta classe obtendo assim subclasses desta última, que são.

Axioma i5- i é contínua.

Axioma i6- i é idempotente, $i(a,a)=a$.

Na tabela 4.1 podemos ver algumas funções intersecção paramétricas e não paramétricas obedecem aos axiomas i1-i4 mas poderão não obedecer a i5-i6.

4.2.2.4. CONSIDERAÇÕES SOBRE A UNIÃO E A INTERSECÇÃO DIFUSA

Nesta secção vão-se fazer algumas considerações gerais sobre as operações de união e de intersecção de conjuntos difusos que são importantes para se perceber o contexto em que os diversos operadores operam sobre os conjuntos difusos.

consideremos as funções,

$$u_{\max}(a,b) = \begin{cases} a & \text{se } b = 0 \\ b & \text{se } a = 0 \\ 1 & \text{caso contrário} \end{cases}$$

para a união de conjuntos difusos, e

$$i_{\min}(a,b) = \begin{cases} a & \text{se } b = 1 \\ b & \text{se } a = 1 \\ 0 & \text{caso contrário} \end{cases}$$

para a intersecção de conjuntos difusos.

Pode-se provar que, ver [Klir, 1988] pag. 52-57,

Teorema 4.2.1- $\forall a,b \in [0,1]$ temos $\max(a,b) \leq u(a,b) \leq u_{\max}(a,b)$, onde u é uma função de união que obedece aos axiomas $u1-u4$.

da mesma forma para a intersecção de conjuntos difusos,

Teorema 4.2.2- $\forall a,b \in [0,1]$ temos $i_{\min}(a,b) \leq i(a,b) \leq \min(a,b)$, onde i é uma função de intersecção que obedece aos axiomas $i1-i4$.

Podemos ver que entre $\min(a,b)$ e $\max(a,b)$ poderão existir uma série de operadores de agregação que será tema da próxima secção.

Convém ainda observar que,

Teorema 4.2.3- $\forall a,b \in [0,1]$, $u(a,b) = \max(a,b)$ é o único operador de união de conjuntos difusos que é contínuo e idempotente, isto é, a única função que obedece aos axiomas $u1-u6$.

Da mesma forma para a intersecção, temos,

Teorema 4.2.4- $\forall a,b \in [0,1]$, $i(a,b) = \min(a,b)$ é o único operador de intersecção de conjuntos difusos que é contínuo e idempotente, isto é, a única função que obedece aos axiomas $i1-i6$.

deve-se notar ainda que,

Teorema 4.2.5- As operações de união e intersecção e de complemento contínuo que satisfazem a lei do terceiro excluído e a lei da contradição, não são, ou idempotentes ou distributivas.

4.2.2.5. OPERAÇÕES DE AGREGAÇÃO GERAIS

Consideremos agora o problema mais geral de queremos agregar n conjuntos difusos, A_1, A_2, \dots, A_n , de forma a obter uma função de pertença representante da agregação:

Definição 4.20- seja h uma função de agregação de $n \geq 2$ conjuntos difusos, A_1, A_2, \dots, A_n , definida da seguinte forma $h: [0,1]^n \rightarrow [0,1]$, onde $\mu_A(x) = h(\mu_{A_1}(x), \mu_{A_2}(x), \dots, \mu_{A_n}(x))$, para cada $x \in U$.

As funções h que obedecem aos axiomas $h1-h2$ chamam-se funções de agregação de conjuntos difusos [Klir, 1988].

Axioma $h1$ - condições de fronteira, $h(0, 0, \dots, 0) = 0$ e $h(1, 1, \dots, 1) = 1$.

Axioma $h2$ - h é monótona crescente para todos os seus argumentos, isto é, para todos os pares $(a_i : i \in N_n)$ e $(b_i : i \in N_n)$, onde $a_i, b_i \in [0,1]$, se $a_i \geq b_i$ para todo $i \in N_n$, então $h((a_i : i \in N_n)) \geq h((b_i : i \in N_n))$.

Estes dois axiomas definem o conjunto de funções de agregação de conjuntos difusos. Podemos restringir um pouco mais este conjunto adicionando mais dois axiomas. Na verdade podemos restringir este conjunto adicionando axiomas de forma a obter classes de funções de agregação que obedecem aos axiomas $h1-h2$.

Axioma $h3$ - h é continua.

Axioma $h4$ - h é uma função simétrica em todos os seus argumentos, isto é, $h(a_i : i \in N_n) = h(a_{p(i)} : i \in N_n)$ onde p é uma permutação qualquer de N_n .

Vemos com facilidade que os operadores de agregação união e intersecção são uma subclasse dos operadores de agregação h , que obedecem aos axiomas $h1, h2$ e $h4$ e com mais restrições nas condições fronteira. Estes operadores de união e intersecção também estão restringidos aos teoremas 4.2.1 e 4.2.2. por isso existe uma gama de operadores de agregação entre $\min(a_1, a_2, \dots, a_n) \leq h(a_1, a_2, \dots, a_n) \leq \max(a_1, a_2, \dots, a_n)$ que esses operadores de união e intersecção não abrangem. Chamamos a esse tipo de operadores de agregação, de operadores de média (Averaging Operators), ver figura 4.6.

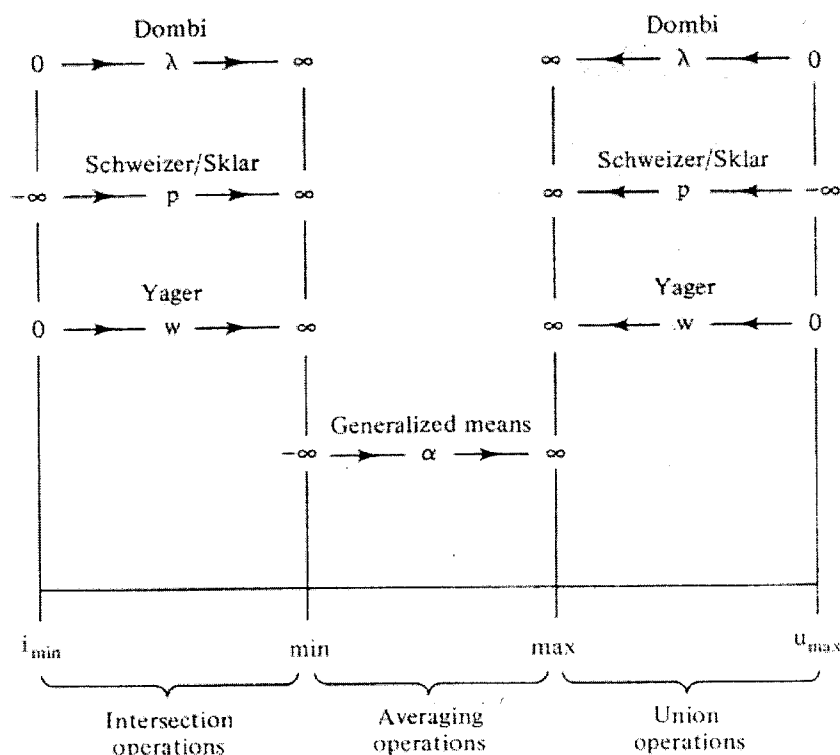


Figura 4.6- Classificação dos operadores de agregação [Klir, 1988].

Adicionemos assim aos axiomas h1-h4 o seguinte axioma,

Axioma h5- Aos operadores h que obedecem a h1-h4 e $\min(a_1, a_2, \dots, a_n) \leq h(a_1, a_2, \dots, a_n) \leq \max(a_1, a_2, \dots, a_n)$, chamamos de operadores de agregação de média.

Os operadores de agregação de média poderão também obedecer a mais propriedades, restringindo assim a classe a mais subclasses dessa classe de operadores, ver [Grabisch, 1998].

Os operadores de agregação de média estão incluídos numa classe mais geral de operadores que se chama Operadores Compensatórios [Grabisch, 1998]. Estes operadores obedecem ao axioma h5, mas não são operadores conjuntivos nem disjuntivos [Grabisch, 1998].

Os operadores de média são apropriados para combinar pontuações “Scores” na avaliação de problemas de múltiplo critério.

Uma classe importante de operadores de média são as médias generalizadas, estas são definidas por:

Definição 4.21- seja h definida por:

$$h_{\alpha}(a_1, a_2, \dots, a_n) = \left(\frac{a_1^{\alpha} + a_2^{\alpha} + \dots + a_n^{\alpha}}{n} \right)^{1/\alpha} \quad (4.2.3)$$

onde $\alpha \in \mathcal{R} \setminus \{0\}$.

Vemos com facilidade que os operadores de média generalizados obedecem aos axiomas h1-h5. Podemos ver por exemplo para h5 que,

$$h_{-\infty}(a_1, a_2, \dots, a_n) = \min(a_1, a_2, \dots, a_n),$$

$$h_{+\infty}(a_1, a_2, \dots, a_n) = \max(a_1, a_2, \dots, a_n).$$

Da definição 4.21 resulta:

Definição 4.22- Média Geométrica:

$$h_0(a_1, a_2, \dots, a_n) = (a_1 \cdot a_2 \cdots a_n)^{1/n}. \quad (4.2.4)$$

Definição 4.23- Média Harmónica:

$$h_{-1}(a_1, a_2, \dots, a_n) = \left(\frac{n}{\frac{1}{a_1} + \frac{1}{a_2} + \dots + \frac{1}{a_n}} \right) \quad (4.2.5)$$

Definição 4.24- Média Aritmética:

$$h_1(a_1, a_2, \dots, a_n) = \frac{1}{n}(a_1 + a_2 + \dots + a_n). \quad (4.2.6)$$

Note-se que em muitas situações de decisão múltiplo atributo é necessário pesos ou importâncias relativas nos critérios/atributos. Assim surge uma nova classe de operadores que são os Weighted Operators ou Operadores de Peso, dos quais sobressaem os operadores de média com pesos, definidos da seguinte forma:

Definição 4.25- Operadores de Média com Pesos:

$$h_\alpha(a_1, a_2, \dots, a_n, w_1, w_2, \dots, w_n) = \left(\sum_{i=1}^n w_i a_i^\alpha \right)^{1/\alpha} \quad (4.2.7)$$

onde $w_i \geq 0$ ($i \in N_n$), são os pesos que exprimem a importância relativa de cada conjunto difuso a agregar e têm de obedecer ainda à seguinte propriedade,

$$\sum_{i=1}^n w_i = 1. \quad (4.2.8)$$

Outro exemplo de Operadores de Peso são os OWA (Ordered Weighted Averaging) introduzidos por [Yager, 1988], estes estão incluídos na classe dos operadores de média, eles são idempotentes, monótonos e comutativos,

isto é, obedecem aos axiomas h1-h5. Estes operadores por serem utilizados no trabalho prático, serão detalhados na secção 4.3.3.2..

Nesta secção foram apresentadas algumas das classes de operadores de agregação, para mais detalhes remeto para [Grabisch, 1998] e [Grabisch, 1995].

4.3. MODELOS DE TOMADA DECISÃO MULTICRITÉRIO

Nesta secção começa-se por falar dos principais modelos de Tomada de Decisão (secção 4.3.1.) e depois na secção 4.3.2. aprofunda-se especificamente o modelo de decisão múltiplo atributo que é o modelo usado no trabalho experimental desta tese.

4.3.1. BREVE DESCRIÇÃO DOS MODELOS USADOS NA TOMADA DE DECISÃO

Esta secção baseia-se no artigo de [Ribeiro, 1996].

Primeiro que tudo tem de se definir o que é entendido por Tomada de Decisão. Segundo Turban “A Tomada de Decisão é um processo de escolha entre cursos de alternativas de acção com o propósito de alcançar um objectivo ou vários objectivos”[Turban, 1998].

Segundo Bellman e Zadeh a Tomada de Decisão num ambiente Difuso é “um processo de decisão no qual os objectivos e/ou os constrangimentos, mas não necessariamente do sistema de controlo, são difusos na sua natureza. Quer isto dizer que os constrangimentos constituem classes de alternativas cujas as fronteiras não são bem definidas.”[Bellman & Zadeh, 1970].

O processo de Tomada de Decisão Múltiplo Critério Difusa pode dividir-se em dois campos, tomada de decisão múltiplo objectivo, e tomada de decisão múltiplo atributo, que se pode hierarquizar conforme a figura 4.7.

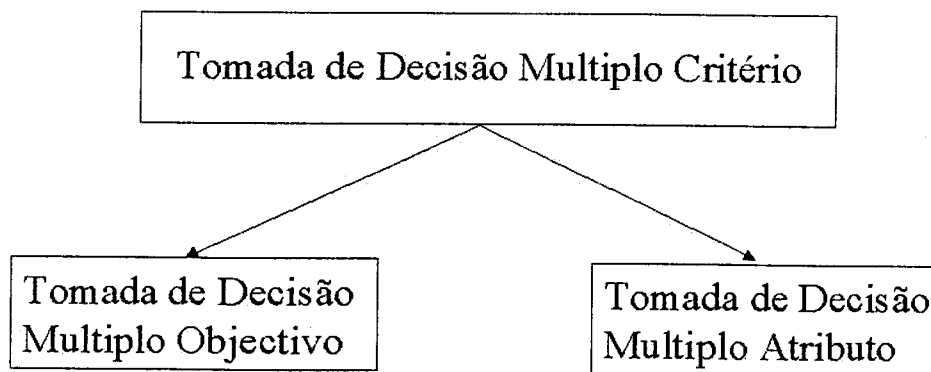


Figura 4.7- Modelos de Tomada de Decisão

A tomada de decisão múltiplo objectivo consiste num conjunto de objectivos quantitativos em conflito que não podem ser alcançados simultaneamente [Zimmermann, 1987].

A tomada de decisão múltiplo atributo envolve problemas em que temos um conjunto de alternativas finitas e pré-definidas que são avaliadas através de atributos e pretende-se seleccionar a melhor alternativa.

Convém antes de mais definir algumas entidades usadas em linguagem da tomada de decisão múltiplo critério.

1. *Alternativas*: Um conjunto de objectos, produtos, acções, ou estratégias de escolha [Starr, 1977]. Por exemplo, uma lista de carros a escolher ou a selecção de um dígito de entre um conjunto de dígitos.
2. *Atributos*: Cada alternativa é definida por um conjunto de características (parâmetros). Estas podem representar características físicas, tal como peso ou cor [Starr, 1977].
3. *Preferências/pesos*: As preferências ou pesos representam a importância relativa de cada atributo entre atributos ou a importância relativa duma instância de um atributo, que respectivamente se denominam por preferência inter-atributo e preferência intra-atributo [MacCrimmon, 1973].

As entidades, atributos e pesos, podem ser representados por conjuntos difusos quando estamos em ambiente difuso. Alias, o primeiro trabalho de extensão difusa dos modelos de TDMC foi apresentado por [Bellman & Zadeh, 1970]. Este artigo seminal permitiu que se desenvolvessem as duas sub-áreas, tomada de decisão múltiplo objectivo difuso e tomada de decisão múltiplo atributo difuso [Chen & Hwang, 1992]. Nesta tese será detalhada a TDMA difusa pois é uma destas técnicas que será usada no reconhecimento de padrões.

4.3.2. MODELOS DE DECISÃO MÚLTIPLO ATRIBUTO DIFUSOS

Segundo R. A. Ribeiro a “Tomada de Decisão Múltiplo Atributo (TDMA) trata de problemas onde se pretende escolher uma alternativa de um conjunto de alternativas que são caracterizadas em termos de atributos” [Ribeiro, 1996].

Na TDMA podemos ter dois tipos de objectivos:

- (a) pretendemos escolher uma alternativa de um conjunto de alternativas baseadas em valores e importâncias dos atributos de cada alternativa;
- (b) classificar as alternativas por ordem ou por similaridade.

De uma forma geral temos a seguinte formulação:

- (a) sejam A_1, A_2, \dots, A_n um conjunto de alternativas avaliadas de acordo com os atributos C_1, C_2, \dots, C_m (critérios);
- (b) seja R_{ij} o valor numérico atribuído ao atributo(critério) C_j para a alternativa A_i ;
- (c) a classificação para a alternativa A_i resulta da agregação dos vários valores numéricos $R_{i1}, R_{i2}, \dots, R_{im}$, isto é, $D(A_i) = (R_{i1} \circ R_{i2} \circ \dots \circ R_{im})$, onde \circ representa o operador de agregação.

Normalmente nos problemas de TDMA difusa, existem duas fases no processo de selecção da melhor alternativa:

Fase I- Agregação do grau de satisfação de todos os critérios por alternativa;

Fase II- Classificar as alternativas de acordo com o grau de satisfação obtido na fase I.

Existem vários métodos para a resolução destas duas fases da TDMA, que se apresentam na tabela 4.2.

4- Conjuntos Difusos e Modelos de Decisão Múltiplo Atributo Difusos

		Phase	Criteria	Weights	Aggregation rule	Solution
G1	M1. Kahne (1975) [16]	I + II	Stochastic	Intervals	Monte Carlo simulation	Crisp
	M2. Saaty (1978) [24]	I	Pairwise comparison	Crisp	Hierarchical aggregation	Crisp
G2	M3. Laarhoven and Pedrycz (1983) [19]	I	Crisp	Crisp	Hierarchical aggregation	Fuzzy
G3	M4. Bellman and Zadeh (1970) [8]	I + II	Fuzzy	Crisp	Max-min	Crisp
	M5. Yager (1977, 1978, 1981) (3 methods) [33, 34, 35]	I + II	Fuzzy	Crisp/Fuzzy	Max-min	Crisp
G4	M6. Baas and Kwakernaak (1977) [2]	I + II	Fuzzy ratings	Fuzzy	Weighted average	Fuzzy
	M7. Dong, Shah and Wong (1985, 1987) (2 methods) [10, 11]	I	Fuzzy ratings	Fuzzy	Extension principle + α -cuts + intervals	Fuzzy
	M8. Dubois and Prade (1980) [14]	—	Fuzzy	Fuzzy	Approximate extension principle	Fuzzy
	M9. Schmucker (1985) [10]	—	Fuzzy	Fuzzy	Extension principle	Fuzzy
	M10. Tseng and Klein (1992) [27]	I	Linguistic variables	Fuzzy	Weighted average	Crisp
G5	M11. Yager (1988) [36]	I + II	Fuzzy	Crisp	OWA operators	Crisp
	M12. Baldwin (1992) [3]	I	Fuzzy	Crisp	Evidential logic rule	Crisp

Tabela 4.2- Classificação dos principais métodos de TDMA difusa [Ribeiro,1996].

Segundo R. A. Ribeiro a classificação dos diversos métodos nos grupos de G1-G5 deve-se ao seguinte.

“Grupo G1- Métodos alternativos à teoria dos conjuntos difusos. Os métodos deste grupo não representam métodos não-difusos, simplesmente servem de base para os métodos seguintes.”.

“Grupo G2- corresponde à “fuzzificação” do método de Saaty do grupo G1.”.

“Grupo G3- correspondem aos métodos conjuntivos.”.

“Grupo G4- determina os valores para os critérios usando classificações difusas linguísticas e variáveis linguísticas para os pesos, que são usados com a equação clássica de média ponderada.”.

“Grupo G5- apresentam métodos que trabalham com critérios que são conjuntos difusos e que não são avaliados por variáveis linguísticas.” [Ribeiro, 1996].

Nesta tese estamos interessados no grupo 2 e especificamente numa das extensões do método AHP (Analytical Hierarchy Process)[Saaty, 1978] apresentado por Yager em [Yager, 1999], que combina o método do grupo G1, AHP, com o método do grupo G5, OWA. Ou seja temos o método AHP com a agregação feita pelo operador OWA já introduzido atrás. Na secção 4.3.3 esta extensão será apresentada em mais detalhe.

Para um leitor mais interessado no aprofundamento de cada um dos métodos apresentados na tabela 4.2 sugere-se [Ribeiro, 1996] e [Chen & Hwang, 1992].

4.3.3. MODELO AHP ESTENDIDO COM OWA

Nesta secção vamos introduzir a partir de um exemplo prático o método AHP de Saaty [Saaty, 1978] e rever e introduzir novos conceitos ao operador de agregação OWA de Yager [Yager, 1999] e finalmente descrever o método AHP com agregação OWA.

4.3.3.1. MÉTODO AHP (ANALYTICAL HIERARCHY PROCESS)

Consideremos o seguinte problema [Yager, 1999]: Pretendemos escolher uma casa de acordo com os seguintes critérios (atributos ou parâmetros), na zona existirem boas escolas (Boas Escolas), a distância ao local de trabalho (Distância ao Trabalho) ser perto, e o preço da casa (Preço) ser razoável. Podemos assim definir o nosso objectivo como sendo a escolha de uma casa, tendo como alternativas as casas que se encontram à venda (casa A e casa B) e como atributos (boas escolas, distância ao trabalho, preço). Podemos hierarquizar este problema por níveis de decisão entre objectivo (primeiro nível) e os diversos atributos que se podem indo hierarquizando em subníveis. No nosso caso em particular só temos dois níveis, mas por exemplo podíamos ter mais níveis dizendo que boas escolas se podiam dividir em novos atributos, como preço das propinas, segurança para as crianças, grau de sucesso da escola, e o mesmo para os outros atributos. Assim para o presente caso temos a hierarquia apresentada na figura 4.8.

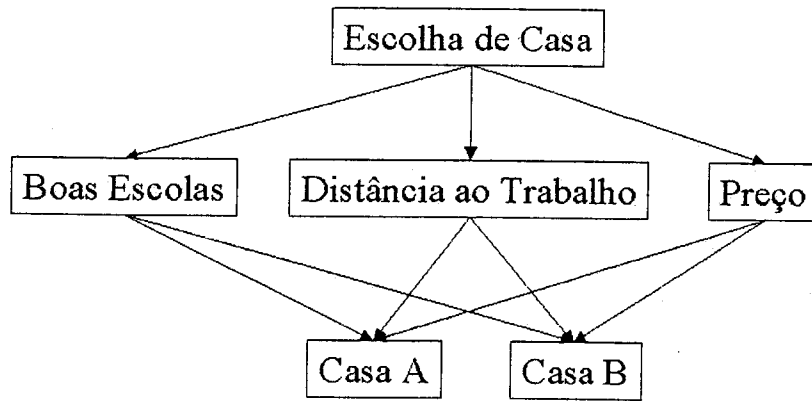


Figura 4.8- Hierarquia do objectivo escolha de casa

A representação da hierarquia de um problema corresponde ao primeiro passo do método AHP.

2º Passo: para cada nível fazer em cada nó da hierarquia uma comparação par-a-par de cada atributo para determinar as preferências relativamente aos atributos. Assim para o presente caso temos para o nível dois, primeiro nó a seguinte matriz de comparações (Matriz M):

	Boas Escolas	Distância ao Trabalho	Preço
Boas Escolas	1	3	-
Distância ao Trabalho	-	1	-
Preço	5	6	1

Tabela 4.3 – Relação par-a-par dos atributos

Estas comparações poderão não ser consistentes, mas na vida real é difícil criar relações completamente consistentes, e por isso vai-se medir o nível de inconsistência, como veremos no 3º passo.

Agora ainda temos de completar a matriz par-a-par para obter a matriz dos recíprocos. Ou seja temos de preencher a matriz M com os seus recíprocos,

segundo a propriedade $a_{ij} = \frac{1}{a_{ji}}$.

3º Passo: Temos a matriz M. definida com todas as comparações par-a-par,

$$M = \begin{pmatrix} 1 & 3 & 1/5 \\ 1/3 & 1 & 1/6 \\ 5 & 6 & 1 \end{pmatrix}$$

vamos aplicar o método dos valores próprios [Saaty,1977], que consiste no processo iterativo:

1. comecemos com a matriz M e normalizemos cada coluna pelo valor, $\|\vec{w}\| = \sum_i |w_i|$, onde w_i são os valores de cada elemento de cada coluna, obtemos assim a matriz M_1 ;
2. Repetir o passo 1 até todas as colunas serem iguais (quase iguais), isto é, atingir-se na iteração um ponto fixo.

Estas colunas dão-nos assim o vector próprio normalizado da matriz M , associado ao maior valor próprio da matriz M . Podemos assim calcular o coeficiente de inconsistência associado a M através da formula $CI = \frac{\lambda_{\max} - n}{n - 1}$, onde $\lambda_{\max} = \max(|\lambda_i|)$ em que λ_i são as soluções de $\det(M - \lambda I) = 0$ e n o numero de colunas. O coeficiente de inconsistência deve ser inferior a 10% para garantir que existe consistência lógica nas preferências do decisor.

Temos assim para a primeira iteração:

$$M_1 = \begin{pmatrix} 0.158 & 0.3 & 0.147 \\ 0.053 & 0.1 & 0.122 \\ 0.789 & 0.6 & 0.731 \end{pmatrix}$$

segunda iteração;

$$M_2 = \begin{pmatrix} 0.156 & 0.166 & 0.167 \\ 0.110 & 0.099 & 0.109 \\ 0.734 & 0.735 & 0.723 \end{pmatrix}$$

terceira iteração;

$$M_3 = \begin{pmatrix} 0.165 & 0.165 & 0.165 \\ 0.108 & 0.108 & 0.108 \\ 0.726 & 0.726 & 0.726 \end{pmatrix}$$

Para determinarmos o grau de inconsistência temos $\det(M - \lambda I) = 0$ cuja a solução é $\lambda = [3.094 \quad -0.047 + i \cdot 0.537 \quad -0.047 - i \cdot 0.537]$ logo $\lambda_1 = 3.094$ e $\lambda_2 = \lambda_3 = 0.539$ logo $\lambda_{\max} = 3.094$ e temos assim $CI=0.047$ que é aceitável por ser inferior a 10%.

Obtemos assim o vector $w=[0.165 \ 0.108 \ 0.726]$ que corresponde à seguinte tabela 4.4.

Boas Escolas	Distância ao Trabalho	Preço
0.165	0.108	0.726

Tabela 4.4 – Vector dos pesos

Que nos diz que damos mais valor ao preço 72,6% depois às boas escolas 16,5% e finalmente a distância ao trabalho de 10,8%.

4º Passo: fazia-se exatamente o mesmo que no passo 2 e 3 para as alternativas, em relação a cada atributo (nó inferior) e obtinham-se, primeiro as matrizes de comparação e depois os vectores de peso. Para simplificar, neste trabalho vamos considerar para cada alternativa, empiricamente uma importância relativamente a cada atributo, isto é, para a casa A temos $x_A=[0.7 \ 0.3 \ 0.5]$ e para a casa B temos $x_B=[0.5 \ 0.8 \ 0.7]$. Estes valores representam a classificação atribuída pelo decisor para cada alternativa em relação a todos os atributos. Por exemplo, neste caso, a casa A tem a classificação de 70% relativamente a ter boas escolas perto, enquanto a casa B apenas é classificada com 50%.

A tabela 4.5 apresenta esses valores em formato matricial.

	Boas Escolas	Distância ao Trabalho	Preço
Casa A	0.7	0.3	0.5
Casa B	0.5	0.8	0.7

Tabela 4.5 – Classificação atribuída pelo decisor

5º Passo: Agora temos de fazer a agregação para obter a classificação de cada alternativa. No caso do AHP, o autor propôs a resolução completa através dos vectores e valores próprios, usando uma média ponderada destes. Para mais detalhes ver [Satty, 1977]. Para simplificar a explicação vamos ilustrar este passo com a média ponderada,

$$D(A_j) = \sum_{i=1}^q w_i x_{ji} = w \cdot x_j^T, \text{ onde } q \text{ é o número de alternativas e } x_{ji} \text{ os}$$

valores da matriz da tabela 4.4 e os w_i os pesos determinados no 3º passo, apresentados na tabela 4.3. Depois a alternativa que se escolhe é a que tiver maior grau de satisfação. Assim para o nosso exemplo temos,

$$D(A) = w \cdot x_A^T = 0.7 \times 0.165 + 0.3 \times 0.108 + 0.5 \times 0.726 = 0.511$$

$$D(B) = w \cdot x_B^T = 0.5 \times 0.165 + 0.8 \times 0.108 + 0.7 \times 0.726 = 0.677$$

Logo a alternativa B era a escolhida, isto é, escolher-se-ia a casa B.

4.3.3.2. OPERADOR OWA

Nesta secção vamos apresentar a definição dos operadores OWA introduzidos por Yager [Yager, 1999].

Neste trabalho usamos estes operadores com quantificadores e por esta razão ambos serão objecto de detalhe.

Definição 4.25- Operador OWA de dimensão n é uma função $OWA_W: [0,1]^n \rightarrow [0,1]$, que tem associado um vector de dimensão n $W=[w_1, w_2, \dots, w_n]^T$ tal que,

1. $w_i \in [0,1]$;
2. $\sum_{i=1}^n w_i = 1$.

onde $OWA_W(a_1, a_2, \dots, a_n) = \sum_{j=1}^n w_j b_j$, onde b_j é o j -elemento maior de a_i ,

isto é, consideremos a permutação p tal que $a_{(1)} \geq a_{(2)} \geq \dots \geq a_{(n)}$, assim $b=[a_{(1)}, a_{(2)}, \dots, a_{(n)}]^T$, logo $OWA_W(a_1, a_2, \dots, a_n) = W^T \cdot b$.

Exemplo 4.11- Seja $W=[0.4 \ 0.3 \ 0.2 \ 0.1]^T$. Considere a agregação $OWA_W(0.7 \ 1 \ 0.3 \ 0.6)$. Neste caso temos $b=[1 \ 0.7 \ 0.6 \ 0.3]^T$ de onde temos que a agregação,

$$OWA_W(0.7 \ 1 \ 0.3 \ 0.6) = W^T \cdot b = (0.4)(1) + (0.3)(0.7) + (0.2)(0.6) + (0.1)(0.3) = 0.76.$$

Podemos facilmente ver que:

- se $w_1=1$ e $w_j=0$ para $(j \neq 1)$ então
 $OWA_W(a_1, a_2, \dots, a_n) = \max(a_1, a_2, \dots, a_n)$.
- se $w_n=1$ e $w_j=0$ para $(j \neq n)$ então
 $OWA_W(a_1, a_2, \dots, a_n) = \min(a_1, a_2, \dots, a_n)$.
- se $w_j=1/n$, $(\forall j \in N_n)$ então $OWA_W = \frac{1}{n} \sum_{j=1}^n a_j$.

Vamos agora apresentar os operadores OWA com quantificadores linguísticos.

4.3.3.2.1. OWA E QUANTIFICADORES LINGUÍSTICOS

Agora vamos ver como podemos quantificar os OWA, através de quantificadores linguísticos e como a partir desta combinação podemos extrair informação de histogramas não normalizados.

Na lógica bivalente $\{0,1\}$ podemos introduzir o conceito de quantificadores, tais como o quantificador existencial e o quantificador universal, isto é, existem alguns elementos de um conjunto X que obedecem a uma determinada propriedade ou todos os elementos de um conjunto X obedecem a uma determinada propriedade, respectivamente. Em lógica difusa podemos ter um número mais vasto de quantificadores linguísticos [Zadeh, 1983]. Tal como os operadores variam entre um u_{\max} e um i_{\min} os quantificadores linguísticos variam entre “todos os critérios têm de ser satisfeitos”, quantificador universal e “pelo menos um critério tem de ser satisfeito”, quantificador existencial. Alguns exemplos de quantificadores linguísticos propostos por Zadeh [Zadeh, 1983], são “a maioria”, “alguns”, “poucos” e “pelo menos 10”. Segundo Yager [Yager, 1999] os quantificadores podem dividir-se em duas classes gerais distintas, os quantificadores absolutos e os quantificadores relativos. Os quantificadores absolutos são subconjuntos difusos de $[0,+\infty[$, e os quantificadores relativos são subconjuntos difusos do intervalo $[0,1]$. exemplos de quantificadores absolutos são “pelo menos 10”, “cerca de 14”, exemplos de quantificadores relativos são “a maioria”, “poucos”, “metade”. Consideremos um exemplo:

Exemplo 4.12- Considere 2 quantificadores relativos “jovem” e “muito jovem” (figura 4.9) e a seguinte “tabela” de veracidade difusa (figura 4.10). Sabendo que o João tem 30 anos podemos quantificar as seguintes proposições;

- O João é muito jovem;
- O João é jovem.

4- Conjuntos Difusos e Modelos de Decisão Múltiplo Atributo Difusos

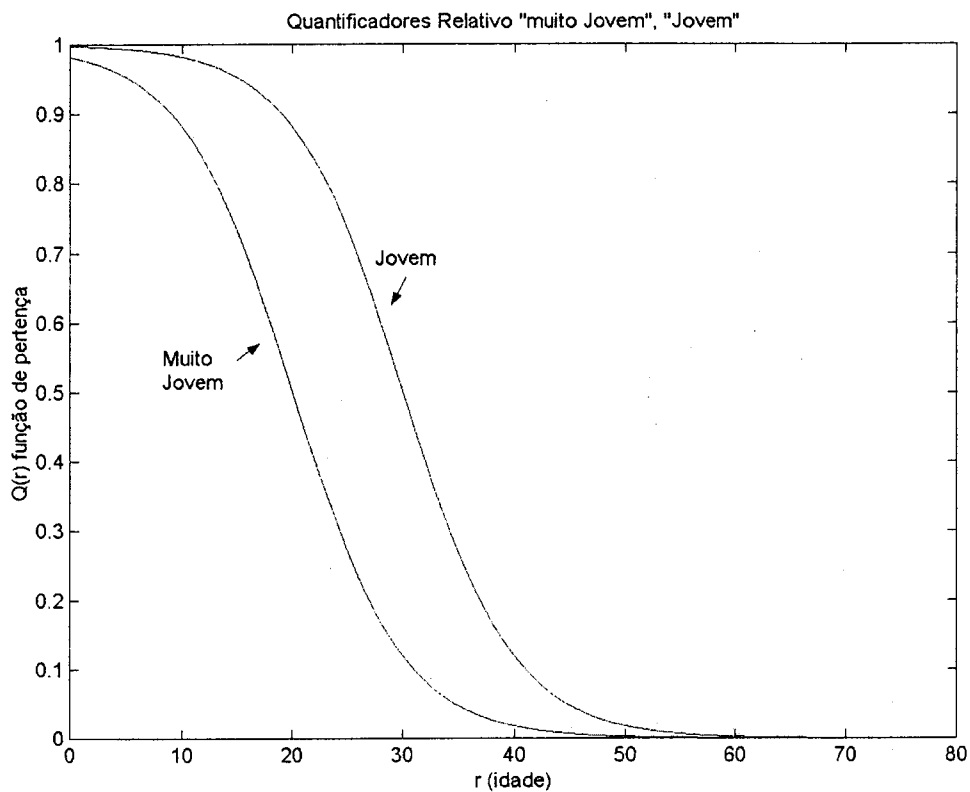


Figura 4.9 – Quantificadores relativos “jovem” e “muito jovem”

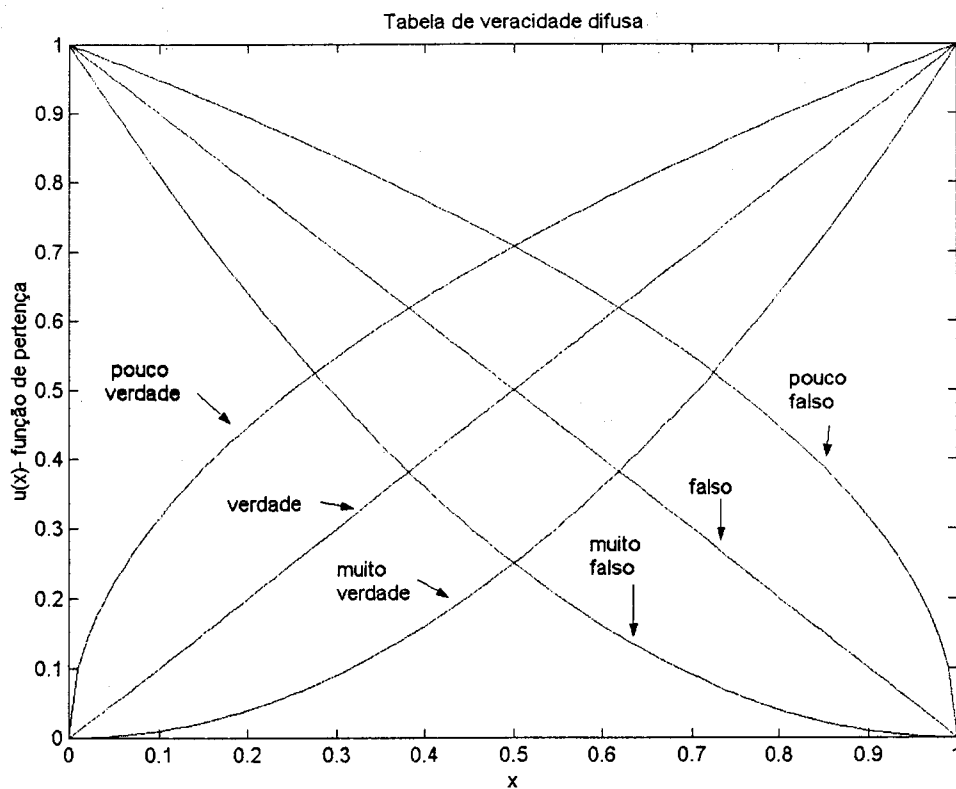


Figura 4.10- “Tabela” de veracidade difusa[Klir, 1988]

Podemos ver que a primeira quantificação “O João é muito jovem”, o grau de pertença é de 0.15 o que corresponde na figura 4.10 a ser pouco verdade com um grau de veracidade de cerca 0.35, é verdade com um grau de veracidade de 0.15 e é muito verdade com um grau de veracidade de 0.03, podíamos fazer o mesmo para a falsidade. Podemos ver que a segunda quantificação “O João é jovem” o grau de pertença é de 0.5 o que corresponde na figura 4.10 a ser é pouco verdade com um grau de veracidade de cerca 0.7, é verdade com um grau de veracidade de 0.5 e é muito verdade com um grau de veracidade de 0.22, podíamos fazer o mesmo para a falsidade.

Consideremos uma subclasse dos quantificadores relativos chamada de “Regular Increasing Monotone” (RIM) [Yager, 1999].

Definição 4.26- Um quantificador Q é “Regular Increasing Monotone” (RIM) se obedecer às seguintes propriedades:

- 1- $Q(0)=0$;
- 2- $Q(1)=1$;
- 3- se $x>y$ então $Q(x)\geq Q(y)$.

Vemos como exemplo que os quantificadores da figura 4.11 são RIM.

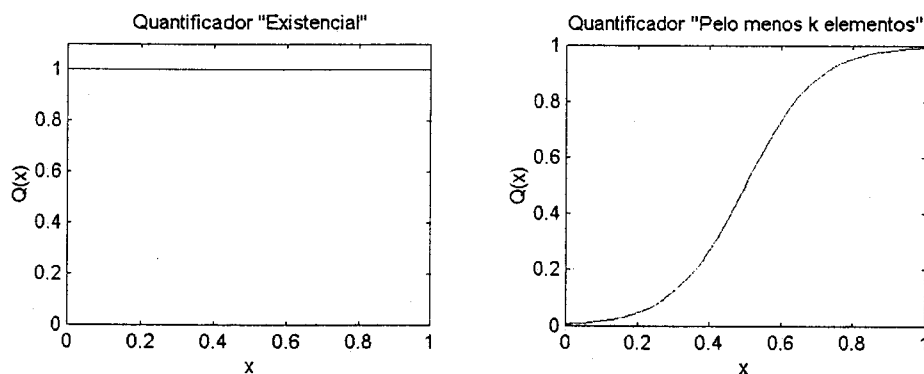


Figura 4.11- Quantificadores RIM

Estes quantificadores são importantes para a análise múltiplo atributo quando pretendemos maximizar a satisfação de uma determinada alternativa, pois podemos definir que pretendemos satisfazer a “maioria” dos critérios ou apenas “alguns”, isto é, os melhores.

Um exemplo de aplicação dos quantificadores RIM é a sua conjugação com os operadores OWA [Yager, 1999]. Seja F um operador OWA de dimensão n de acordo com a definição 4.25, onde o vector peso w_i é definido da seguinte forma,

$$\forall i \in N_n : w_i = Q\left(\frac{i}{n}\right) - Q\left(\frac{i-1}{n}\right) \quad (4.3.9)$$

Onde $Q(x)$ é um quantificador RIM. Sendo o grau de satisfação para a alternativa x segundo os atributos A_1, A_2, \dots, A_n , dada por;

$$D(x) = F(A_1(x), A_2(x), \dots, A_n(x)). \quad (4.3.10)$$

Vejamos agora um método sugerido por Yager [Yager et al.,1999], para a introdução de quantificadores, nesta abordagem de OWA+RIM.

Consideremos um problema com $\{A_1, A_2, \dots, A_n\}$ atributos, e que a cada atributo atribuímos uma importância $v_j \in [0, +\infty[$, e seja o vector $b = [A_{(1)}, A_{(2)}, \dots, A_{(n)}]$ onde $A_{(1)} \geq A_{(2)} \geq \dots \geq A_{(n)}$, isto é, ordenemos os atributos por ordem crescente e seja $u = [v_{(1)}, v_{(2)}, \dots, v_{(n)}]$. Construimos assim o operador OWA, F de dimensão n , da seguinte forma,

$$\forall j \in N_n : w_j(x) = Q\left(\frac{\sum_{k=1}^j u_k}{T}\right) - Q\left(\frac{\sum_{k=1}^{j-1} u_k}{T}\right) \quad (4.3.11)$$

onde $T = \sum_{k=1}^n u_k$ e o grau de satisfação para a alternativa x é dada por,

$$D(x) = \sum_{j=1}^n b_j w_j(x), \quad (4.3.12)$$

Onde b_j é o j elemento do vector $b = [A_{(1)}, A_{(2)}, \dots, A_{(n)}]$ e w_j é definido por (4.3.11).

Por sua vez (4.3.11) e (4.3.12) definem o operador OWA, F de dimensão n , de acordo com a definição 4.25.

4.3.3.2.2. OWA QUANTIFICADOS E HISTOGRAMAS DE VARIÁVEIS

Consideremos a seguinte distribuição de uma variável física visualizada na figura 4.12.

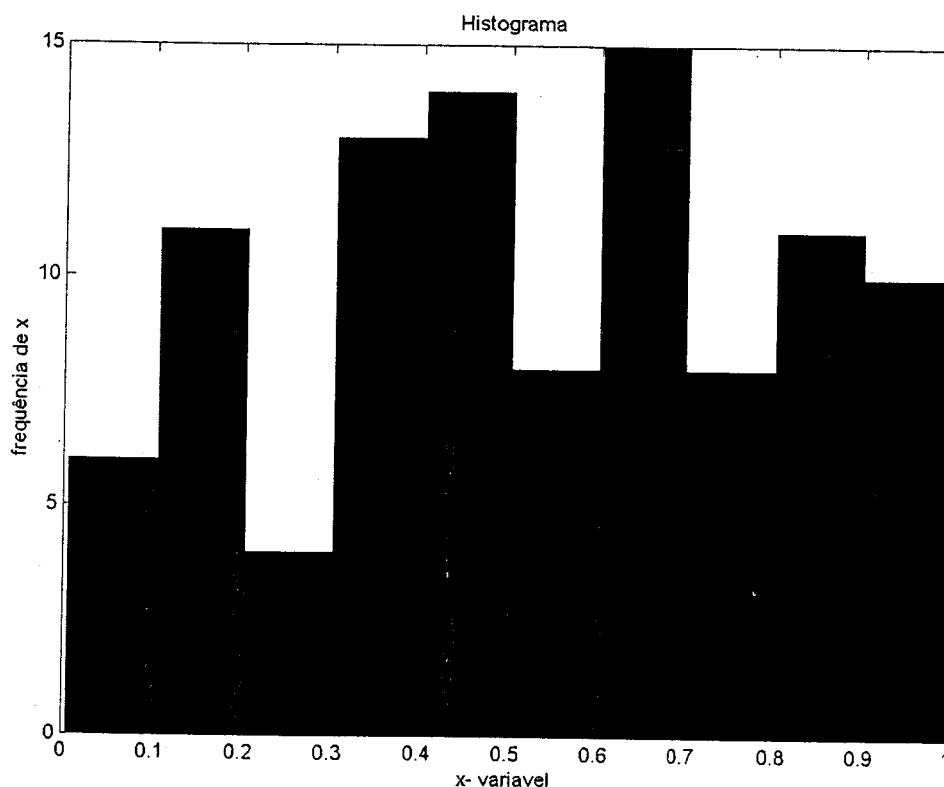


Figura 4.12- Exemplo de histograma de uma variável Física

Ao criarmos o histograma escolhemos sempre uma partição no domínio discreta, que no limite nos dá uma distribuição contínua. Aí tomamos normalmente duas atitudes ou utilizamos o histograma que é uma função dos degraus como representação da nossa variável física ou inferimos uma determinada distribuição estatística contínua que julgamos representar melhor a nossa variável física. Sugiro aqui uma nova aproximação. A utilização de conjuntos difusos com uma agregação OWA+Importâncias+Quantificadores. Assim poder-se-á por exemplo representar um determinado intervalo do histograma como um conjunto difuso, pois no histograma normal estamos a atribuir por exemplo para o primeiro intervalo sempre a mesma importância, e o mesmo para os outros intervalos. Quanto maiores forem estes intervalos, maior a incerteza na distribuição. Por exemplo para os dois primeiros intervalos podemos atribuir os seguintes conjuntos difusos, figura 4.13.

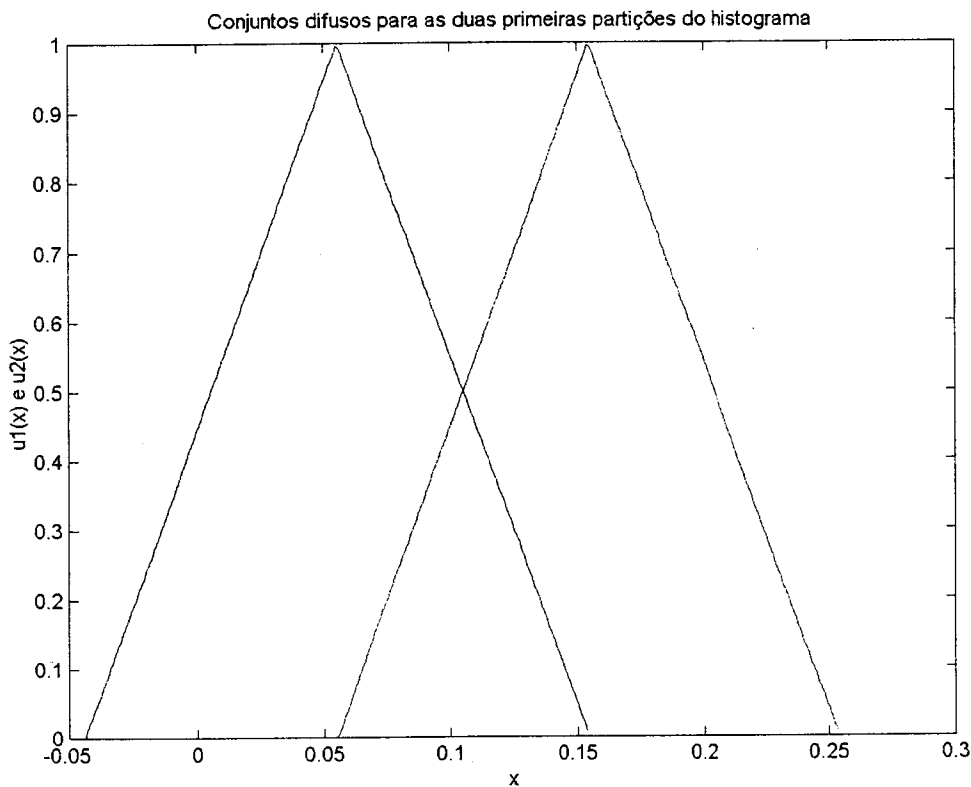


Figura 4.13 – Conjuntos difusos para os primeiros rectângulos (duas primeiras partições) do histograma

Com as respectivas importâncias $v_1=6$ e $v_2=11$ retiradas da frequência de x dos 2 intervalos. Vejamos um exemplo.

Na figura 4.13 apresentamos dois conjuntos difusos que representam e aproximam os dois primeiros intervalos do histograma, com funções de pertinência triangulares. Uma função de pertinência triangular é definida por três pontos valor mais à esquerda, valor mais à direita e vértice. Neste caso as funções de pertinência são definidas por forma ao vértice corresponder ao valor médio do intervalo considerado do histograma e os extremos corresponderem aos valores médios do intervalo anterior e posterior, respectivamente.

Consideremos outra vez o exemplo das figuras 4.12-13. Temos como vector das importâncias $v=[6 \ 11 \ 4 \ 13 \ 14 \ 8 \ 15 \ 8 \ 11 \ 10]$ dadas pelas respectivas frequências e temos como domínio da variável física o intervalo $x \in [-0.04, 1.05]$, que pode ser por exemplo uma gama de temperaturas num determinado sistema físico. Suponhamos que experimentalmente obtínhamos o valor de $x=0.12$ e queríamos ver qual a sua importância no sistema. de acordo com o histograma obtínhamos uma importância de 11, que normalizada pelo método “crisp” daria

$11/15=0.733$. Utilizando os RIM teríamos, $Q(x) = \sqrt[4]{x}$ como quantificador “poucos”, que neste caso deve ser interpretado como um quantificador que vai seleccionar os melhores valores de x . Para os diversos atributos que no nosso caso são o número de intervalos considerados no histograma temos os seguintes valores de pertença para cada atributo, $\{0.348, 0.6512, 0, 0, 0, 0, 0, 0, 0, 0\}$. Ordenando os vectores conforme prescrito nos OWA obtemos $b=[0.651 \ 0.349 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$, vem $u=[11 \ 6 \ 4 \ 13 \ 14 \ 8 \ 15 \ 8 \ 11 \ 10]$ vector de pesos.

Conforme referido em 4.3.2.2.1, temos,

$$T = \sum_{k=1}^{10} u_k = 11 + 6 + 4 + 13 + 14 + 8 + 15 + 8 + 11 + 10 = 100;$$

$$w_1 = \sqrt[4]{\frac{11}{100}} - \sqrt[4]{\frac{0}{100}} = 0.576;$$

$$w_2 = \sqrt[4]{\frac{(11+6)}{100}} - \sqrt[4]{\frac{11}{100}} = 0.642 - 0.576 = 0.066;$$

$$w_3 = \sqrt[4]{\frac{21}{100}} - \sqrt[4]{\frac{17}{100}} = 0.677 - 0.642 = 0.035;$$

$$w_4 = \sqrt[4]{\frac{34}{100}} - \sqrt[4]{\frac{21}{100}} = 0.764 - 0.677 = 0.087;$$

Podíamos continuar a fazer os cálculos mas não é necessário pois os outros pesos são zero e quando agregamos obtemos,

$$D(0.12) = \sum_{j=1}^{10} b_j w_j(0.12) = 0.375 + 0.023 + 0 + \dots + 0 = 0.398$$

Assim a classificação apresentada pelo método difuso é de 0.398 enquanto pelo método “crisp” é de 0.733. O método difuso poderá ser interessante para aplicações onde nos interessa saber um pouco o que acontece na vizinhança do nosso intervalo, pois representa sempre uma média entre os três vizinhos. Assim sendo as transições entre intervalos são difusas e não “crisp” como acontece num histograma.

Como vimos no exemplo anterior podemos obter algumas vantagens em utilizar conjuntos difusos ao invés de histogramas para estimar a importância de uma determinada grandeza, pois temos um número muito vasto de quantificadores RIM à nossa disposição.

4.3.3.3. EXEMPLO DA EXTENSÃO AO MÉTODO AHP COM OWA+QUANTIFICADORES

A extensão do método AHP introduzida por Yager [Yager, 1999], baseia-se na aplicação de operadores de agregação OWA ao nosso conjunto de atributos (critérios), isto é, a extensão não retira nenhuma propriedade ao método AHP, mas sim, enriquece o método. Vejamos através de um exemplo apresentado em [Yager, 1999].

Consideremos o problema de escolher um engenheiro informático “Choice of a developer” [Yager, 1999]. A modelização deste problema e a sua representação hierárquica estão descritas na figura 4.14.

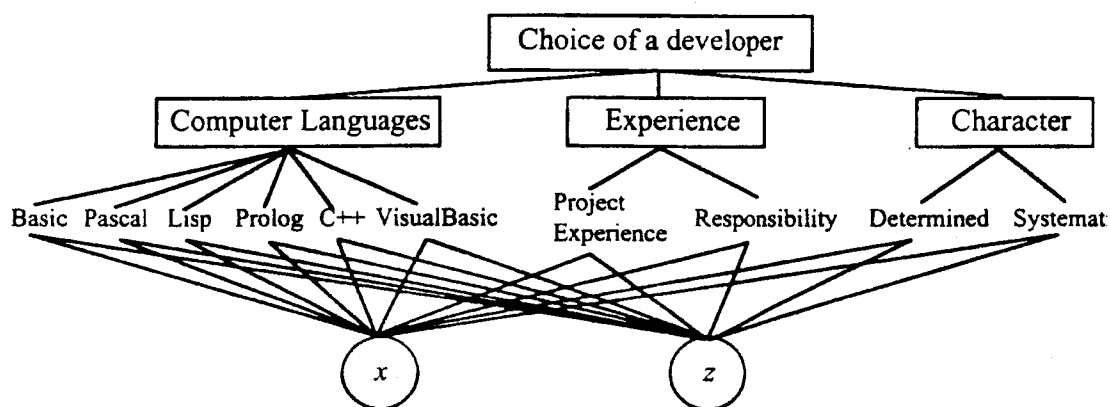


Figura 4.14- Hierarquia para a escolha de um engenheiro informático [Yager, 1999]

Consideremos as seguintes comparações par-a-par e os seus respectivos recíprocos, para os nós “Choice of a developer”, “Experience” e “Character”,

Choice of a developer			
	CL	Exp	Ch
CL	1	3	2
Exp	1/3	1	2
Ch	1/2	1/2	1

Experience		
	Project	Responsibility
Project	1	1/3
Responsibility	3	1

Character		
	Determined	Systematic
Determined	1	1/2
Systematic	2	1

Aplicando o método dos valores próprios já exemplificado na secção 4.3.3.1. às comparações feitas anteriormente obtemos os seguintes vectores de importâncias:

“Choice of a developer”:

CL	Exp	Ch
0.542	0.240	0.218

“Experience”:

Project	Responsibility
0.25	0.75

“Character”:

Determined	Systematic
0.33	0.67

e para “Computer Languages” consideremos as seguintes importâncias dadas por um Decisor que atribui as seguintes importâncias:

Basic	Pascal	Lisp	Prolog	C++	Visual Basic
1	3	2	2	5	5

Os graus de satisfação entre atributos e alternativas apresentam-se nas tabelas seguintes:

	Basic	Pascal	Lisp	Prolog	C++	Visual Basic
x	0.8	0.9	0.2	0.1	0.5	0.6
z	0.5	0.5	0.1	0.2	0.7	0.6

	Project	Responsability	Determined	Systematic
x	0.8	0.5	0.5	0.7
z	0.4	0.8	0.8	0.6

Agora o próximo passo será a agregação com o operador OWA de baixo para cima da hierarquia. Assim temos,

“Computer Languages”:

Quantificador: $Q(t) = t^2$, “most”- valoriza mais os últimos atributos

$$b(x) = [0.9 \quad 0.8 \quad 0.6 \quad 0.5 \quad 0.2 \quad 0.1]$$

$$u(x) = [3 \quad 1 \quad 5 \quad 5 \quad 2 \quad 2];$$

$$T(x) = \sum_{j=1}^6 u_j(x) = 18.$$

$$w_1(x) = Q\left(\frac{3}{18}\right) - Q\left(\frac{0}{18}\right) = 0.17^2 - 0 = 0.028;$$

$$w_2(x) = Q\left(\frac{4}{18}\right) - Q\left(\frac{3}{18}\right) = 0.022;$$

$$w_3(x) = Q\left(\frac{9}{18}\right) - Q\left(\frac{4}{18}\right) = 0.200;$$

$$w_4(x) = Q\left(\frac{14}{18}\right) - Q\left(\frac{9}{18}\right) = 0.355;$$

$$w_5(x) = Q\left(\frac{16}{18}\right) - Q\left(\frac{14}{18}\right) = 0.185;$$

$$w_6(x) = Q\left(\frac{18}{18}\right) - Q\left(\frac{16}{18}\right) = 0.210;$$

Logo o grau de satisfação para o atributo “Computer Languages” e alternativa x é dado pela agregação:

$$D(x) = \sum_{j=1}^6 b_j w_j(x) = 0.398.$$

Da mesma forma podíamos fazer o calculo para a alternativa z obtendo um grau de satisfação de:

$$D(z) = \sum_{j=1}^6 b_j w_j(z) = 0.399.$$

“Experience”:

$$\text{Quantificador: } Q(t) = \begin{cases} Q(0) = 0 \\ Q(t) = 1 \quad \text{se } t \neq 0 \end{cases}, \text{ “Existencial”}$$

$$b(x) = [0.8 \quad 0.5];$$

$$u(x) = [0.25 \quad 0.75];$$

$$T(x) = \sum_{j=1}^2 u_j(x) = 1.$$

$$w_1(x) = Q(0.25) - Q(0) = 1 - 0 = 1;$$

$$w_2(x) = Q(1) - Q(0.25) = 0;$$

Logo o grau de satisfação para o atributo “Experience” e alternativa x é dado pela agregação:

$$D(x) = \sum_{j=1}^2 b_j w_j(x) = 0.8.$$

Da mesma forma podíamos fazer o calculo para a alternativa z obtendo um grau de satisfação de:

$$D(z) = \sum_{j=1}^2 b_j w_j(z) = 0.8.$$

“Character”:

$$\text{Quantificador: } Q(t) = \begin{cases} Q(1) = 1 \\ Q(t) = 0 \quad \text{se } t \neq 1 \end{cases}, \text{ “Universal”}$$

$$b(x) = [0.7 \quad 0.5];$$

$$u(x) = [0.67 \quad 0.33];$$

$$T(x) = \sum_{j=1}^2 u_j(x) = 1.$$

$$w_1(x) = Q(0.67) - Q(0) = 0;$$

$$w_2(x) = Q(1) - Q(0.67) = 1;$$

Logo o grau de satisfação para o atributo “Character” e alternativa x é dado pela agregação:

$$D(x) = \sum_{j=1}^2 b_j w_j(x) = 0.5.$$

Da mesma forma podíamos fazer o calculo para a alternativa z obtendo um grau de satisfação de:

$$D(z) = \sum_{j=1}^2 b_j w_j(z) = 0.6.$$

Agora fazemos a agregação para o atributo que está no topo da hierarquia:

“Choice of a developer”:

Quantificador: $Q(t) = t^2$, “most”

$$b(x) = [0.8 \quad 0.5 \quad 0.398];$$

$$u(x) = [0.240 \quad 0.218 \quad 0.542];$$

$$T(x) = \sum_{j=1}^3 u_j(x) = 1.$$

$$w_1(x) = Q(0.240) - Q(0) = 0.058;$$

$$w_2(x) = Q(0.458) - Q(0.240) = 0.152;$$

$$w_3(x) = Q(1) - Q(0.458) = 0.790;$$

Logo o grau de satisfação para o atributo “Character” e alternativa x é dado pela agregação:

$$D(x) = \sum_{j=1}^3 b_j w_j(x) = 0.437.$$

e para a alternativa z obtiamos,

$$D(z) = \sum_{j=1}^3 b_j w_j(z) = 0.452.$$

Donde podemos concluir que o candidato z é o melhor pois tem um grau de satisfação maior.

O método AHP estendido, tem a vantagem de permitir manipular os quantificadores tais como as importâncias atribuídas aos atributos e desta forma ajustar os parâmetros para obter um valor óptimo. por todas as razões apresentadas este foi o método escolhido para o reconhecimento dos dígitos de língua Portuguesa.

4.4. CONCLUSÕES

Na primeira parte deste capítulo fez-se uma introdução ao conceito de conjunto difuso. Depois foram definidas as estruturas matemáticas operações entre conjuntos difusos (negação, intersecção e união). Finalmente foi definida uma estrutura matemática mais generalizada de operações de agregação entre conjuntos difusos, terminando assim uma breve introdução à teoria dos conjuntos difusos.

Na segunda parte deste capítulo fez-se uma abordagem geral à problemática da tomada de decisão múltiplo critério, e foram apresentados em mais detalhe os métodos mais usados na Tomada de Decisão Múltiplo Atributo difuso. Em particular detalhou-se o método AHP (Analytical Hierarchy Process) introduzido por Saaty, e o seu modelo estendido por Yager, utilizando os operadores de agregação OWA (Order Weighted Averaging). O aprofundamento deste método AHP estendido com o OWA deve-se a que foi escolhido como método de reconhecimento de padrões para o trabalho prático apresentado nesta tese.

Por fim terminamos o capítulo apresentando algumas das vantagens de se usar o método estendido AHP.

Reconhecimento de Dígitos na língua Portuguesa.

5.1. INTRODUÇÃO

Neste capítulo vamos apresentar três métodos para o reconhecimento dos dígitos, números de 0 a 9, na língua Portuguesa.

Primeiro apresentamos o método de tomada de decisão múltiplo atributo difuso.

Em segundo apresentamos o método de tomada de decisão múltiplo atributo.

Em terceiro apresentamos um método já usado no reconhecimento de dígitos, o método de quantização vectorial e comparação de padrões sem alinhamento temporal.

E por ultimo será feita uma breve conclusão acerca dos métodos aplicados. Este trabalho teve por base uma aplicação de reconhecimento de vogais [Simas & Ribeiro, 2001a]. No reconhecimento das vogais usou-se o método multiatributo difuso. Esta primeira experiência provou que este tipo de métodos flexíveis podem ser úteis no reconhecimento de padrões. Nesta tese este trabalho não será detalhado pois apenas serviu como primeira experiência para o desenvolvimento da aplicação apresentada neste capítulo, reconhecimento dos dígitos de 0 a 9.

5.2. CONSIDERAÇÕES GERAIS

Neste trabalho de reconhecimento de dígitos 0 a 9 na língua Portuguesa teve-se como objectivo criar um sistema, o mais real possível. Uma primeira experiência deste trabalho foi apresentada em França [Simas & Ribeiro, 2001b].

Nesta aplicação, usou-se hardware e software comuns e um conjunto heterogéneo de pessoas para se criar uma base de dados de sons dos dígitos 0 a 9. Recolheram-se 900 sons, dando 90 sons por dígito. Estes sons ou sinais foram recolhidos de um total de 14 pessoas 8 masculinos e 6 femininos, com idades compreendidas entre os 10 e 64 anos.

Usou-se um microfone Normal unidireccional, com uma resposta em frequência de 20Hz a 20kHz e um software de captura "Sound Forge v4.5". O sinal foi capturado com uma frequência de amostragem de 11025Hz e uma amplitude de 16 bit's. Por exemplo numa linha telefónica os sinais são processados a uma frequência de amostragem de 8000Hz e uma amplitude

de 8 bit's. Podemos ver na figura 5.1 os limites de percepção do sistema auditório.

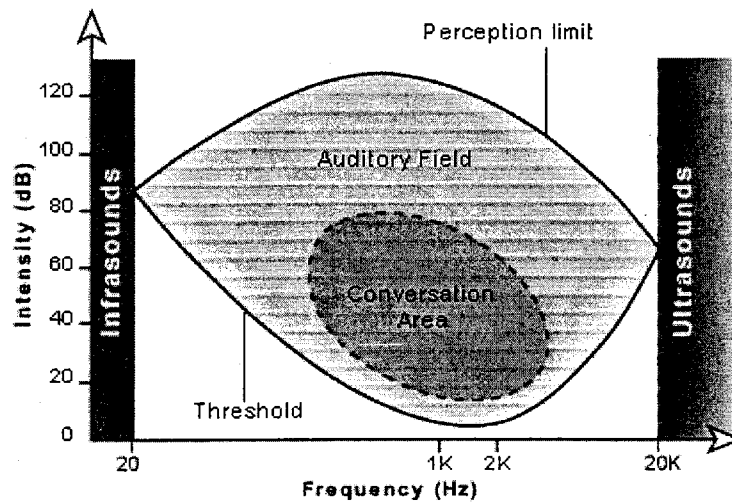


Figura 5.1- Limites de percepção do ouvido Humano.

[<http://www.iurc.montp.insem.fr/cric/audition/english/sound/sound.htm>]

Numa linha telefónica a banda de frequências é limitada a 300-3400Hz por razões económicas, por exemplo as emissões de rádio em AM têm uma banda de 100-4000Hz. e como podemos ver na figura 5.1 todas estas bandas são audíveis dentro da gama de conversação. As duas primeiras formantes normalmente encontram-se dentro dos 300Hz a 4000Hz de forma que uma banda neste intervalo apanharia as duas primeiras formantes, e como estas caracterizam, como já vimos nos capítulos 2 e 3, o processamento da voz Humana, podemos trabalhar dentro desta banda de 300Hz a 4000Hz. Nos ensaios feitos foi utilizado um filtro para limitar as frequências de 300Hz a 4000Hz, mas como os resultados foram inferiores aos da banda completa, aqui isso só se apresentam os resultados sem a utilização de filtros de banda.

Uma consideração importante a tomar é que os sons foram adquiridos de uma forma contínua e depois “cortados” manualmente, de acordo com o critério de que zonas iniciais e terminais de pequena amplitude não seriam consideradas. Isto porque o sistema informático apenas capturava o sinal total. Por exemplo na figura 5.2, mostra-se uma sequência de sons relativamente ao dígito, temos representados dois quatros. De forma a isolar os quatros estes foram “cortados” como se mostra na figura 5.3.

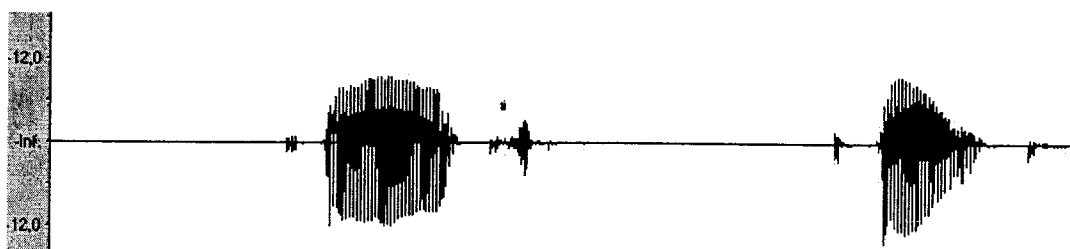


Figura 5.2- Parte da sequência de quatro capturados por uma pessoa

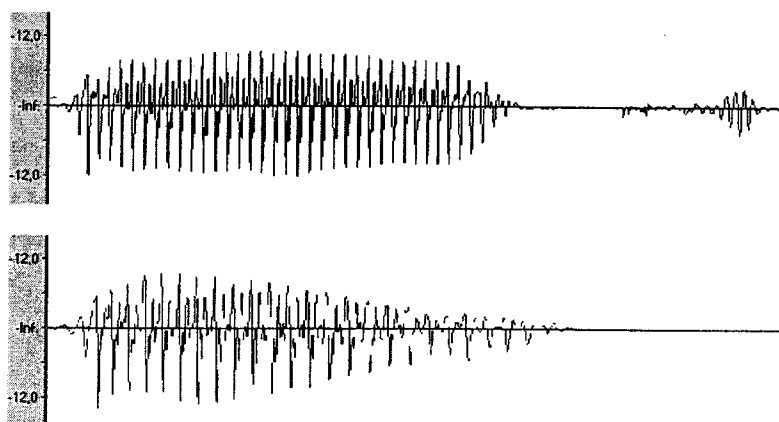


Figura 5.3- Forma como foram cortados os sinais para a base de dados.
Exemplo do dígito quatro.

Como se pode ver na figura 5.3 a parte inicial foi ignorada em ambos os sinais, e quanto ao segundo sinal a parte final do quatro foi ignorada, isto é, a parte do “tro”, pois o sinal quatro composto pelas sílabas “qu”+”á”+”tro”. Penso que na realidade um sistema de captura, apenas captura parte dos sinais, isto é, não consegue detectar todas as suas sílabas. Os sinais (Sons) foram adquiridos em diversos ambientes, silêncio a ruidoso, tais como salas de aula, quartos, etc. Assim criou-se uma base de dados de acordo com estes critérios que me parecem ser os mais realísticos para o reconhecimento de dígitos.

5.3. RECONHECIMENTO DOS DÍGITOS NA LÍNGUA PORTUGUESA UTILIZANDO UM MÉTODO DE TOMADA DE DECISÃO MÚLTIPLO ATRIBUTO DIFUSO

Os sistemas de reconhecimento de voz compreendem uma colecção variada de algoritmos vindos de uma grande variedade de disciplinas, desde o reconhecimento de padrões estatístico à teoria da comunicação e processamento de sinal, entre outros [Rabiner, 1993].

A aproximação feita neste trabalho para o reconhecimento de padrões foi: Sinal→Extracção de parâmetros→Comparação de padrões com uma

referência de padrões constituída por conjuntos difusos por alternativa→Tomada de decisão múltiplo atributo difusa→reconhecimento do padrão. Esta aproximação pode ser vista na figura 5.4.

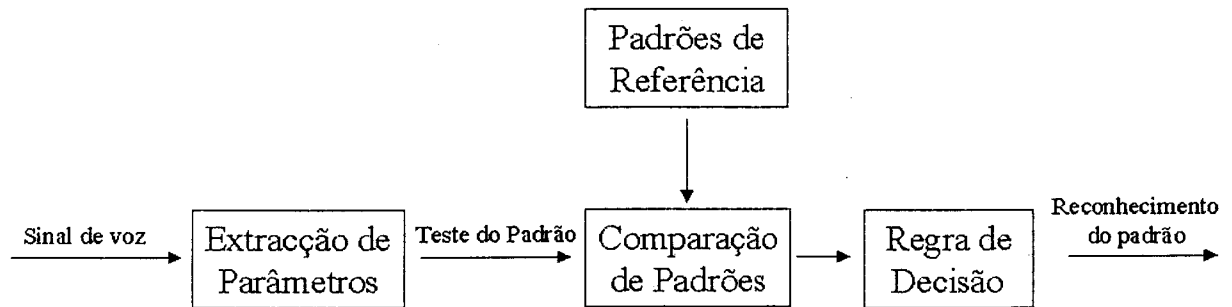


Figura 5.4- Modelo usado para o reconhecimento de padrões

Para a Extracção de parâmetros é usada a teoria do processamento de sinal e análise espectral, estudada no capítulo 3, extraindo os parâmetros F1 e F2; F1' e F2' que representam respectivamente primeira formante e segunda formante pelo método de Burg e pelo método de Ghael-Sandgathe. Para a comparação de padrões é usado o método apresentado em 4.3.3.2.2. OWA Quantificados e Histogramas de variáveis do capítulo 4, onde cada padrão de referência é um histograma não normalizado que é transformado num conjunto de conjuntos difusos com pesos, que representam o padrão de referência. Para a Regra de decisão é usada a extensão do método AHP com o operador de agregação OWA apresentado na secção 4.3.3.3 do capítulo 4.

5.3.1. EXTRACÇÃO DE PARÂMETROS

Do ponto de vista Físico, a pronúncia de um número (Dígito) origina a propagação de uma onda sonora através do nosso sistema fisiológico (sistema vocal), Glote, laringe, faringe língua, cavidade bocal, cavidade nasal e lábios. Uma força muscular causa uma perturbação de ar que gera um impulso e este propaga-se através do sistema vocal onde se transforma e sai pelos lábios e nariz.

No nosso caso só conhecemos o sinal à saída dos lábios $y(n)$, não conhecemos a função de transferência $h(n)$, nem o impulso inicial $x(n)$. Como foi visto no capítulo 2 e 3 podemos usar um modelo físico onde podemos estimar a função de transferência $h(n)$, através de um modelo LPC (Linear Predictive Coding), que no domínio das frequências vimos que se podia escrever como:

$$H(z) = \frac{1}{1 - \sum_{i=1}^p a_i \cdot z^{-i}} \quad (5.2.1)$$

Onde p é a ordem do modelo LPC. Depois eram calculados os coeficientes LPC, a_i a partir dos métodos de Burg e Ghael-Sandgate e calculando a resposta em frequência deste conjunto de $p+1$ coeficientes, determinávamos a envolvente espectral e por sua vez os picos correspondiam às formantes, que por sua vez são os zeros do polinómio complexo

$$P(z) = 1 - \sum_{i=1}^p a_i \cdot z^{-i}, \text{ mais próximos do raio unitário do plano complexo.}$$

Como veremos no ponto 5.5. através da análise LPC podemos extrair outro tipo de parâmetros que se chamam de coeficientes ceptrais e os coeficientes delta-ceptrais que nos permitem classificar sons [Rabiner, 1993].

Normalmente para reduzir os erros no cálculo dos coeficientes a_i da função de transferência usam-se janelas da ordem dos 20ms [Rabiner, 1978], pois estamos a considerar que o sinal sonoro é invariante no tempo, isto é, fisicamente é estacionário dentro dessa janela. Por isso se o sinal variar muito com o tempo dentro dessa janela o modelo corre o risco de cometer erros na previsão dos coeficientes a_i . Assim é assumido um compromisso quanto ao comprimento da janela e a frequência de amostragem, assim para frequências da ordem dos 10kHz usam-se janelas de 20ms [Rabiner, 1978]. No presente caso como se está a lidar com um problema onde existe uma certa incerteza quanto aos coeficientes a_i e consequentemente no cálculo das formantes, resolveu-se não usar uma janela de 20ms mas sim dividir o sinal em 10 janelas independentemente da duração do sinal. Ou seja, para um sinal de 0.2s estamos a considerar janelas de 0.02s=20ms e no caso do sinal ser da ordem de 0.8s estamos a considerar janelas de 0.08=80ms. Assim tendo uma incerteza maior para sinais mais compridos. Os sinais neste trabalho variam entre os 0.2s e os 0.9s.

O modelo LPC utilizado neste trabalho é representado na figura 5.5.

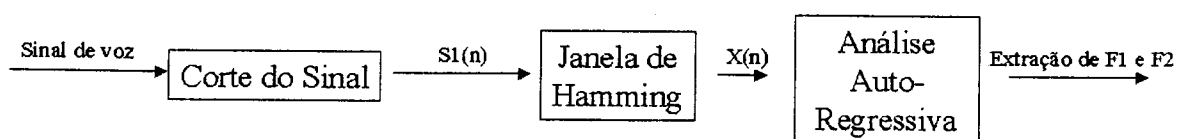


Figura 5.5- Modelo LPC utilizado para extração de parâmetros

Os passos representados na figura 5.5 podem resumir-se da seguinte forma:

1. Corte do sinal – Aqui foi cortado o sinal de acordo com os critérios descritos no ponto 5.2;

2. Janela de Hamming – aqui podia ter sido usado outro tipo de janela, mas foi escolhida a janela de Hamming de forma a minimizar as descontinuidades no início e no final do sinal quando neste é aplicado uma janela. O comprimento da janela de Hamming foi considerado igual ao tamanho da janela isto é $x(n) = s_1(n) \cdot w(n)$, onde a janela de Hamming $w(n)$ é definida matematicamente pela seguinte expressão: $w(n) = 0.54 - 0.46 \cdot \cos\left(\frac{2\pi \cdot n}{N-1}\right)$, $0 \leq n \leq N-1$.
3. Análise Auto-regressiva: Aqui foram usados os métodos de Burg e Ghael-Sandgate para o cálculo dos coeficientes LPC que por sua vez nos permitem calcular as Formantes F1 e F2.

5.3.2. COMPARAÇÃO DE PADRÕES

A comparação de padrões é dividida em duas partes. Uma primeira que consiste em criar um conjunto de padrões de referência ou “codebook” e uma segunda parte que consiste na comparação de um determinado padrão com esse conjunto de padrões de referência.

Como foi dito na secção 5.3.1, no processamento do sinal são aplicadas janelas de forma a reduzir os erros de cálculo na determinação dos atributos. No nosso caso tínhamos visto que o sinal é dividido em 10 janelas e para cada uma dessas janelas eram calculados os atributos/parâmetros, que neste caso são as duas primeiras formantes por cada um dos métodos autoregressivos. Para construirmos o nosso conjunto de padrões de referência usou-se um conjunto de treino com 650 sons (65 por alternativa).

Para cada janela determinamos 4 atributos, obtendo para cada atributo uma distribuição de valores, que são representados inicialmente por um histograma, como por exemplo o dígito zero apresentado na figura 5.6(a). Para cada um desses histogramas construímos os nossos conjuntos difusos de acordo com o método explicado na secção 4.3.3.2.2. OWA Quantificados e Histogramas de variáveis. Assim temos dez janelas por atributo e para cada uma dessas janelas temos uma representação conforme a figura 5.6.

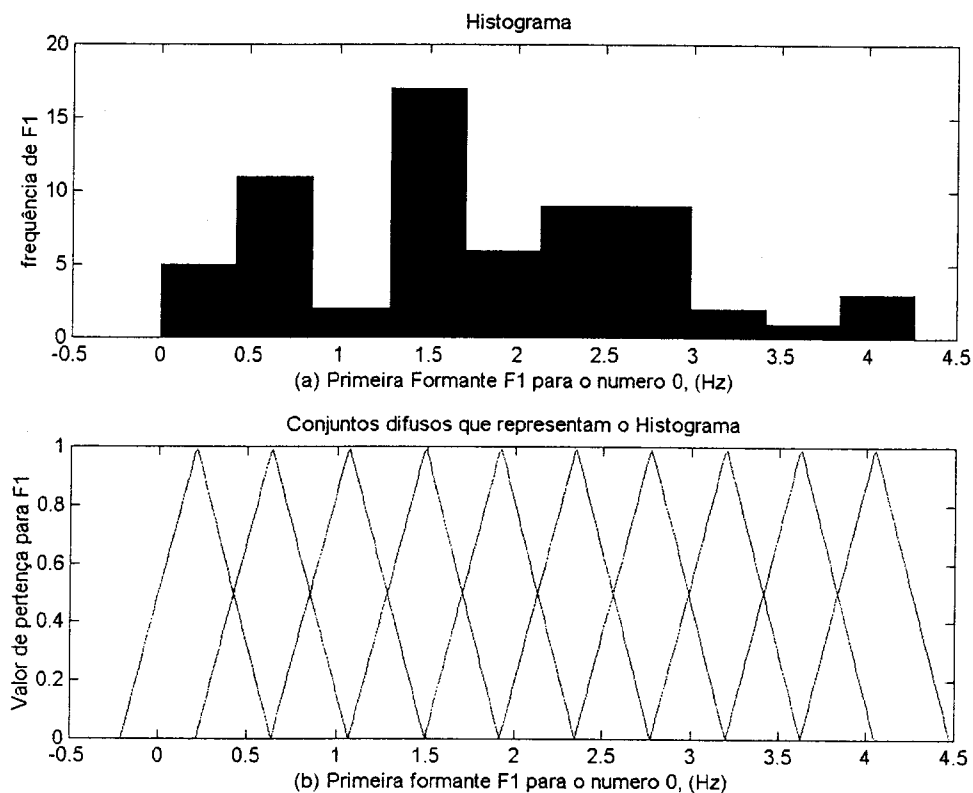


Figura 5.6 – Construção dos conjuntos difusos

Assim cada atributo é representado por 10 vectores (um vector por janela) cujas as coordenadas são pares de (conjunto difuso, importância atribuída a esse conjunto difuso), onde a título de exemplo na figura 5.6 temos o vector representativo $((CD_1,5);(CD_2,11);...;(CD_{10},3))$ que representa a primeira formante F1 para o numero zero na primeira janela. Temos assim para cada alternativa um conjunto de 40 vectores representativos do padrão, isto é, $(10 \text{ janelas}) \cdot (4 \text{ atributos}) = 40$ vectores. Como temos 10 alternativas temos assim $10 \cdot 40 = 400$ vectores no total que formam o nosso “codebook” ou padrões de referência.

Devemos notar que os padrões são todos diferentes, ao contrário do que se poderia pensar, pois na verdade em cada janela de cada atributo temos domínios diferentes, isto é, por exemplo para a janela 1 do atributo F1 temos um domínio para F1 que varia entre 800Hz e 3000Hz e por exemplo para a janela 2 do mesmo atributo F1, varia de 900Hz a 2500Hz. Acresce que como dividimos este domínio em dez partes iguais os conjuntos difusos serão todos diferentes de janela para janela para um mesmo atributo de atributo para atributo para janelas iguais, etc.

A comparação de padrões é feita a partir de um vector com $4 \cdot 10 = 40$ coordenadas que é gerado pela extracção de parâmetros que é comparado

com os padrões de referência de uma forma hierarquizada (ver figura 5.7), isto é, para um determinado sinal é extraído o vector;

$$\left[\{F_1^1, F_1^2, \dots, F_1^{10}\} \quad \{F_2^1, F_2^2, \dots, F_2^{10}\} \quad \{F_1'^1, F_1'^2, \dots, F_1'^{10}\} \quad \{F_2'^1, F_2'^2, \dots, F_2'^{10}\} \right]$$

onde F_1^2 representa a primeira formante extraída pelo método de Burg para a segunda janela, onde F_1^2 vai ser comparado com todos os vectores correspondentes à primeira formante do método de Burg com respeito à segunda janela de acordo o método explicado na secção 4.3.3.2.2. OWA Quantificados e Histogramas de variáveis do capítulo 4.

Na secção seguinte será apresentado um exemplo que mostra com mais detalhe esta comparação de padrões.

5.3.3. REGRA DE DECISÃO

O método de decisão usado (ver figura 5.4) foi o método estendido AHP, isto é, o método descrito na secção 4.3.3.3 do capítulo 4. A hierarquia de conceitos está representada na figura 5.7.

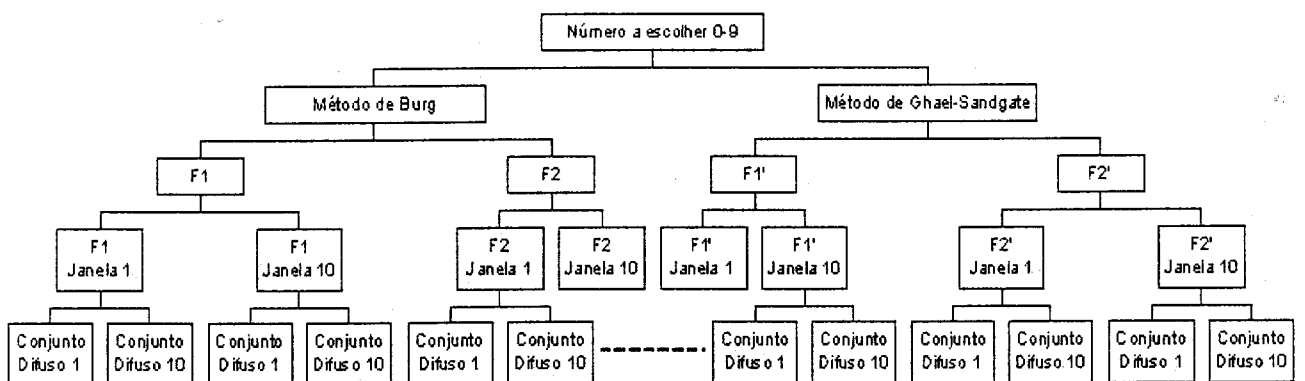


Figura 5.7- Hierarquia para as alternativas 0-9

1º nível: temos o objectivo “Número a escolher 0-9”;

2º nível: temos o vector [“Método de Burg”, “Método de Ghael-Sandgate”] correspondente aos dois métodos usados no reconhecimento;

3º nível: temos o vector [F1, F2, F1', F2'] das duas formantes de cada método;

4º nível: temos o vector:

$$\left[\{F_1^1, F_1^2, \dots, F_1^{10}\} \quad \{F_2^1, F_2^2, \dots, F_2^{10}\} \quad \{F_1'^1, F_1'^2, \dots, F_1'^{10}\} \quad \{F_2'^1, F_2'^2, \dots, F_2'^{10}\} \right]$$

5º nível: para cada janela temos 10 conjuntos difusos definidos como descrito na secção 4.3.3.2.2..

Vamos agora apresentar os vários passos do reconhecimento dos dígitos, que constitui o que se denomina por regra de decisão.

Como vimos na extracção de parâmetros obtemos um vector de dimensão 40.

1º Passo: Extracção de parâmetros

$$V = \left[\{F_1^1, F_1^2, \dots, F_1^{10}\} \quad \{F_2^1, F_2^2, \dots, F_2^{10}\} \quad \{F_1^1, F_1^2, \dots, F_1^{10}\} \quad \{F_2^1, F_2^2, \dots, F_2^{10}\} \right]$$

onde F_1^2 representa a primeira formante extraída pelo método de Burg para a segunda janela.

2º Passo: Grau de pertença ou comparação de padrões

Por exemplo, para o valor F_1^1 , primeira formante da 1ª janela temos no nível 4 da hierarquia (ver figura 5.7) uma representação de 10 conjuntos difusos, isto é, esta janela é representada por um vector de 10 pares ordenados (conjunto difuso, importância). Para o valor F_1^1 extraído determinamos o grau de pertença de F_1^1 nesses conjuntos difusos (ver tabela 5.1). Depois agregamos de acordo com o exemplo da secção 4.3.3.2.2.. Deve-se notar que os conjuntos difusos são diferentes de janela para janela pois os domínios dos parâmetros variam de janela para janela, isto é, todos os histogramas têm domínios diferentes. Fazemos isto para cada alternativa, pois cada alternativa corresponde um histograma com dez conjuntos difusos, obtemos assim após a agregação um vector de dimensão 10 pois temos 10 alternativas.

Fazemos isto para cada valor do vector do 1º passo, para a respectiva janela, obtendo assim por janela um vector de dimensão 10. Finalmente normaliza-se cada vector pelo valor máximo do vector, para obter agregações consistentes.

De forma a melhor ilustrar o processo consideremos por exemplo os cálculos para o valor F_1^1 :

5.3.4. RESULTADOS

Dentro do conjunto de sons (sinais) recolhidos foram usados 650 sons (65 por alternativa) para treino, isto é, para a construção dos padrões de referência e 250 sons (25 por alternativa) foram usados para teste. Vemos que a base de dados é pequena, e o ideal seria ter pelo menos 100 oradores cada um com 10 sons por alternativa (10 alternativas) o que daria $100 \cdot 10 \cdot 10 = 10000$ sons com 1000 por alternativa. Estamos muito aquém daqueles valores pois não se conseguiu construir uma base de dados tão grande. No entanto pensamos que a dimensão da base de dados recolhida já é suficiente para inferir resultados representativos do uso do método proposto para reconhecimento de dígitos.

Assim os resultados em percentagem de sucesso para a base de dados usada apresentam-se na tabela 5.5,

	0	1	2	3	4	5	6	7	8	9
Resultados	32	24	40	60	48	80	28	52	56	52

Tabela 5.5- Percentagem de resultados correctos

Vemos que a precisão máxima é de 80% para o dígito 5 e a mínima é de 24% para o dígito 1, com um erro da ordem de $100/25 = \pm 4\%$. Temos assim um resultado global de **47.2%** de casos certos com um erro da ordem $100/250 = \pm 0.4\%$.

5.4. RECONHECIMENTO DOS DÍGITOS NA LÍNGUA PORTUGUESA UTILIZANDO UM MÉTODO DE TOMADA DE DECISÃO MÚLTIPLO ATRIBUTO

Este método é igual ao anterior só que não é difuso pois não se usam conjuntos difusos, mas sim um outro tipo de funções que representam uma determinada incerteza como nos conjuntos difusos mas não obedecem à axiomática apresentada no capítulo 4.

O método é igual ao anterior, só que ao invés de usarmos o operador OWA com os histogramas de acordo com a secção 4.3.3.2.2. são usadas as funções triangulares de acordo com a figura 5.8 (a) e (b) ou seja, definem-se funções aproximativas e “overlapping” para cada máximo do histograma.

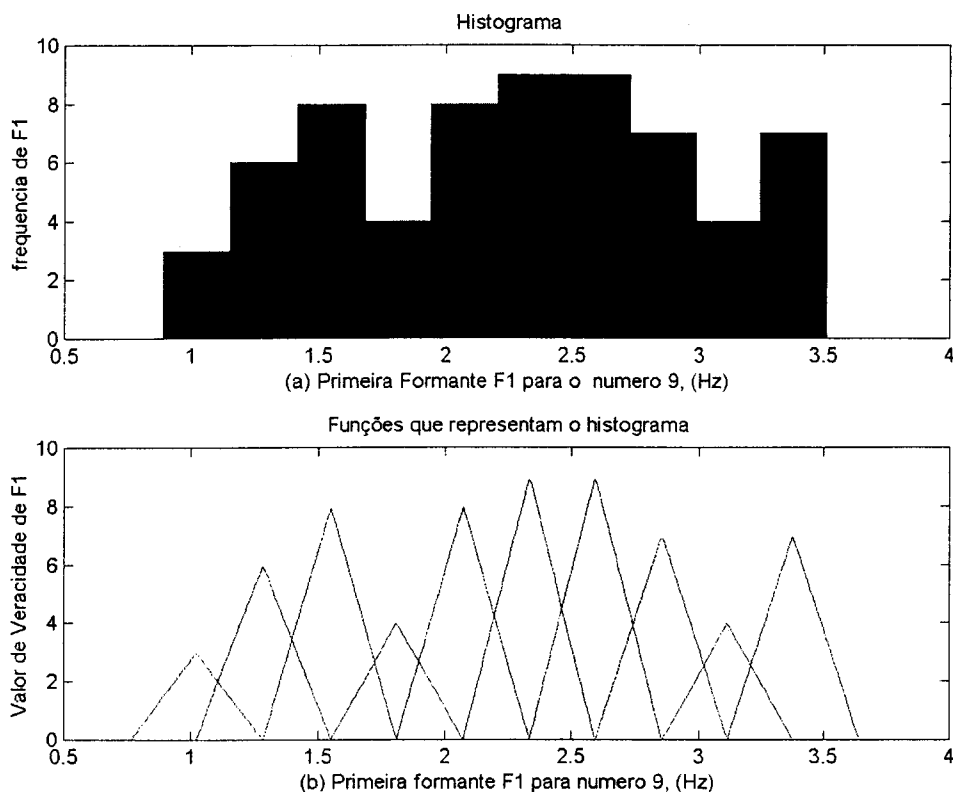


Figura 5.8- Funções com incerteza mas não normalizadas

Portanto, o sistema de agregação só difere do anterior, secção 5.3.4., no segundo passo, pois ao invés de termos 10 conjuntos difusos por janela temos 10 funções triangulares não normalizadas que representam uma determinada incerteza e a agregação no nível 4 da hierarquia é feita com o operador “max” ao invés do OWA.

Os resultados obtidos com esta pequena diferença estão apresentados na tabela 5.6.

	0	1	2	3	4	5	6	7	8	9
Resultados	32	40	52	72	64	80	44	52	44	60

Tabela 5.6- Percentagem de resultados correctos

Novamente a precisão máxima é de 80%, mas a mínima aumenta para 32%, com um erro da ordem de $100/25=\pm 4\%$. Temos assim um resultado global de **54%** de casos certos com um erro da ordem $100/250=\pm 0.4\%$.

5.5. RECONHECIMENTO DOS DÍGITOS NA LÍNGUA PORTUGUESA UTILIZANDO UM MÉTODO DE QUANTIZAÇÃO VECTORIAL E COMPARAÇÃO DE PADRÕES SEM ALINHAMENTO NO TEMPO

Face aos resultados anteriores entendeu-se que se deveria comparar estes resultados com outras técnicas que garantem o sucesso ao problema do reconhecimento de dígitos. Por exemplo o trabalho de [Rodrigues & Trancoso, 1996] garante para este tipo de problema, aplicado às linhas telefónicas um sucesso da ordem dos 99%, utilizando uma base de dados SPEECHDAT, que contem 4000 oradores. Como não tive acesso à base de dados SPEECHDAT foi construída uma base de dados com os critérios referidos em 5.2. Como não se podem comparar os resultados apresentados nesta tese com os de uma outra base de dados, pois os critérios de construção da base de dados são diferentes, optou-se por aplicar um método apresentado por Rabiner [Rabiner, 1993] à base de dados construída segundo 5.2. Assim teremos uma plataforma de comparação entre os dois métodos introduzidos nas secções 5.4. e 5.5..

5.5.1. CALCULO DOS COEFICIENTES CEPSTRAIS E DELTA-CEPSTRAIS

O método consiste em usar-se a análise LPC mas ao invés de se extraírem as formantes extraem-se os coeficientes Cepstrais e os coeficientes delta-cepstrais. Após o calculo dos coeficientes LPC, que neste caso foi usado o método autorecurssivo de Durbin que pode ser visto em [Rabiner, 1978] pag. 411, calculam-se os coeficientes cepstrais e delta-cepstrais de acordo com as fórmulas:

Calculo dos coeficientes Cepstrais:

$$c_m = a_m + \sum_{k=1}^{m-1} \left(\frac{k}{m} \right) \cdot c_k a_{m-k} , \text{ se } 1 \leq k \leq p . \quad (5.5.2)$$

onde p é a ordem do método LPC.

Devido à sensibilidade dos coeficientes de ordem mais baixa com respeito à variação espectral total e à dos coeficientes de ordem superior ao ruído, costuma-se usar uma técnica dos coeficientes cepstrais ponderados [Rabiner, 1993], isto é,

$$c'_m = w_m c_m , \text{ para } 1 \leq m \leq p$$

onde,

$$w_m = \left[1 + \frac{p}{2} \sin\left(\frac{\pi m}{p}\right) \right], \text{ para } 1 \leq m \leq p. \quad (5.5.3)$$

Daqui em diante falemos por abuso de linguagem que os coeficientes cepstrais ponderados são os coeficientes cepstrais.

Calculo dos coeficientes delta-Cepstrais:

$$\Delta c_m(t) = \mu \cdot \sum_{k=-K}^K k \cdot c_m(t+k) \quad (5.5.4)$$

onde μ é uma constante de normalização e $(2K+1)$ é número de janelas sobre a qual deverá ser feito o cálculo.

Tipicamente usa-se $K=3$ [Rabiner, 1993].

Assim para o nosso caso como temos uma ordem de $p=(fs/1000)+1$ para o nosso sistema LPC de acordo [Rabiner, 1978] com $fs=11025$ logo $p=12$ o que implica termos 13 coeficientes cepstrais mais outros 13 coeficientes delta-cepstrais, ficamos assim com um vector Q de dimensão 26.

$$Q(t) = (c'_1(t), c'_2(t), \dots, c'_p(t), \Delta c_1(t), \Delta c_2(t), \dots, \Delta c_p(t)) \quad (5.5.5)$$

5.5.2. QUANTIZAÇÃO VECTORIAL

Usaram-se janelas de 20ms para cada sinal, obtendo assim para um sinal de 0.2s um conjunto de 10 vectores Q . Assim a partir de um conjunto de treino gera-se um conjunto de vectores Q por alternativa, indiscriminadamente do tempo, isto é, a ordem temporal não importa o que importa é que para uma série de sinais para uma determinada alternativa se obtiveram n vectores Q . Para cada alternativa pegamos nesse conjunto de vectores e usamos o algoritmo k -means [Rabiner, 1993] para quantizar o conjunto. O valor de k depende do número do número de vectores que pretendemos quantificar que são n , em geral $n \approx 10k$ [Rabiner, 1993]. Assim dizemos que temos uma precisão de B bit's se $k=2^B$. Para termos uma precisão de 6 Bit's $k=64$, logo $n=640$.

Temos assim um conjunto de k vectores representativos do conjunto n de vectores para cada alternativa. Esses são considerados os nossos padrões. No caso específico usou-se o algoritmo k -means descrito em [Rabiner, 1993] pag.125.

Treinou-se 650 sinais isto é 65 sinais por alternativa. Como por alternativa se obteve um mínimo para $n=877$, temos que $k \approx n/10$ e $k=2^B$ e que num sistema de 6 Bit's é o máximo que obtemos com os dados que temos, logo $k=64$, isto quer dizer que cada padrão tem 64 vectores representativos por alternativa.

5.5.3. RECONHECIMENTO DE PADRÕES BASEADO NA QUANTIZAÇÃO VECTORIAL SEM ALINHAMENTO NO TEMPO

O reconhecimento de um sinal é feito através de medidas de distorção, que comparam o sinal que é desconhecido e é comparado com cada padrão de cada alternativa. Essa medida de distorção dá-nos o grau de insatisfação que determinado sinal tem com o padrão. Pois quanto menor for a medida de distorção para um determinado padrão maior é a semelhança com esse padrão. Podemos ver o modelo de classificação na figura 5.9.

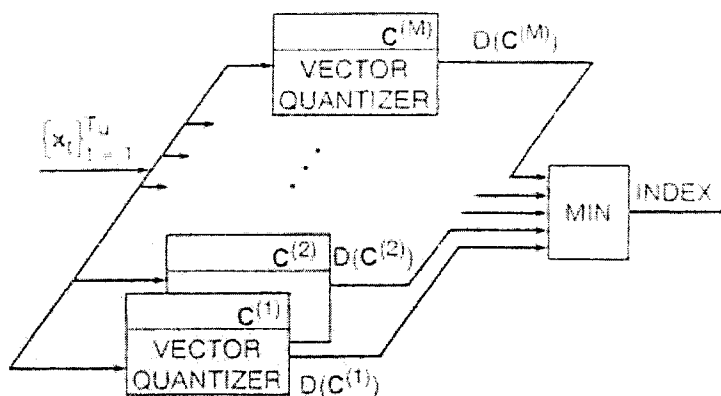


Figura 5.9- Sistema de reconhecimento de padrões baseado na Quantização Vectorial [Rabiner, 1993].

Seja $\{x_t\}_{t=1}^T$ o número de vectores do sinal a reconhecer e seja $C^{(M)} = \{y_i\}_{i=1}^K$ o número de vectores quantizados (k-means) para a alternativa M. O conjunto $C = \{C^{(1)}, C^{(2)}, \dots, C^{(M)}\}$ é designado por Codebook [Rabiner, 1993].

Usamos no nosso reconhecimento a medida de distorção D, designada por medida de distorção sem alinhamento no tempo e é definida da seguinte forma:

$$D(C^{(M)}) = \frac{1}{T} \sum_{t=1}^T d(x_t, z_t) \quad (5.5.6)$$

$$z_t = \arg \min_{y_i \in C^{(M)}} d(x_t, y_i)$$

Onde $d(.)$ é uma medida de distância. No nosso caso foi usada a medida de distância “city block distance”;

$$d(x_i, y) = \sum_{k=1}^K |x_{ik} - y_k| \quad (5.5.7)$$

Assim a alternativa escolhida para classificar o sinal de entrada é dada por;

$$D_{\min} = \min_C D(C). \quad (5.5.8)$$

5.5.4. RESULTADOS

Os resultados obtidos por este método dão-nos uma medida de conformidade com os restantes resultados obtidos com os outros métodos. Estes resultados são os piores dos três modelos testados conforme se pode ver na tabela 5.7.

	0	1	2	3	4	5	6	7	8	9
Resultados	56	60	52	60	32	56	12	40	40	32

Tabela 5.7- Percentagem de resultados correctos

Com este método vemos que a precisão máxima é de apenas 60% e a mínima é de 12%, com um erro da ordem de $100/25 = \pm 4\%$. Temos assim um resultado global de 44% de casos certos com um erro da ordem $100/250 = \pm 0.4\%$. O resultado global é da mesma ordem de grandeza do primeiro método, apesar de ser inferior.

5.6. BREVES CONCLUSÕES AOS MODELOS

Como referimos vemos que a aproximação feita na secção 5.5 dá-nos os piores resultados dos métodos propostos, foi muito importante para ver a coerência da base de dados. Assim poderemos talvez concluir que os resultados baixos obtidos, comparativamente aos 99% obtidos por [Rodrigues & Trancoso, 1996] não são comparáveis, pois dependem da

base de dados. Não quer dizer que usando os dois métodos anteriores se obtenham bons resultados com a base de dados SPEECHDAT, mas seria interessante experimentar os outros métodos utilizados.

Os métodos apresentados nas secções 5.3 e 5.4 parecem ser uma boa alternativa aos métodos “crisp” clássicos. Estes têm vantagens como a de se poder usar vários tipos de atributos de natureza diferente que nos permitem melhorar a classificação do padrão. Neste tipo de problemas é também interessante poder-se dar mais ou menos importância a determinadas zonas do sinal, pois permite fazer o “tuning” do sistema.

5.7. CONCLUSÕES

Neste capítulo apresentamos três métodos para o reconhecimento dos dígitos, números de 0 a 9, na língua Portuguesa.

Primeiro apresentamos o método de tomada de decisão múltiplo atributo difuso.

Em segundo apresentamos o método de tomada de decisão múltiplo atributo.

Em terceiro apresentamos um método já usado no reconhecimento de dígitos, o método de quantização vectorial e comparação de padrões sem alinhamento temporal.

E por ultimo foi feita uma breve conclusão acerca dos métodos aplicados.

Conclusões e trabalho Futuro

6.1. INTRODUÇÃO

Neste capítulo apresentam-se as conclusões com respeito ao trabalho da tese e apresentam-se algumas sugestões para trabalho futuro.

6.2. CONCLUSÕES E TRABALHO FUTURO

Esta tese de Mestrado em Inteligência Artificial Aplicada, teve como objectivo propor uma alternativa a técnicas usadas no reconhecimento de padrões usando técnicas da tomada de decisão múltiplo atributo difusas, tendo como exemplo ilustrativo o reconhecimento dos dígitos 0-9 apresentado no capítulo 5.

Este trabalho teve por base o estudo de três áreas fundamentais, análise e processamento de sinal, a teoria dos conjuntos difusos e as teorias da decisão múltiplo critério. Esta última focando especificamente a subárea denominada tomada de decisão múltiplo atributo.

Quanto à parte prática, capítulo 5, viu-se que foram aplicadas várias técnicas para validar os resultados. Dos resultados obtidos podemos inferir que uma base de dados consistente é fundamental para o estudo deste tipo de problemas. Não deixa de ser interessante que a base de dados construída refere-se ao caso real do dia a dia, com o reconhecimento de voz, pois eu pessoalmente não acredito em 99% de acerto nos casos reais, como podemos ter a experiência com os nossos telemóveis que usam este tipo de técnicas e quando agora recentemente telefonamos para a Portugal Telecom, que tem um sistema de reconhecimento de voz, mas certamente não parece garantir valores de reconhecimento tão altos.

Poderia ainda fazer-se um estudo de comparação com outras técnicas de sucesso no reconhecimento de voz tais como por exemplo as de Cadeias de Markov Escondidas associadas a sistemas de extracção de parâmetros LPC que é baseado no sistema de produção de voz ou usar outras técnicas baseadas no nosso sistema auditivo.

Penso que a abordagem múltiplo atributo para este tipo de problema provou ser uma abordagem flexível para o reconhecimento com ruído (incerteza). Será interessante ver com mais detalhe outras aplicações de reconhecimento de voz. Outro aspecto interessante deste trabalho é o facto

de se estar a usar um modelo múltiplo atributo difuso, com operadores OWA quantificados para estender o modelo hierárquico (AHP) para a resolução deste tipo de problema, pois em toda a literatura que li não encontrei nenhuma abordagem semelhante. Esta técnica poderá ter vantagens pois a variação das importâncias no método AHP estendido, permite-nos manipular os resultados, isto é, ajustar os resultados, à semelhança de uma rede neuronal.

Como trabalho futuro poder-se-ia tentar estudar uma regra de aprendizagem para otimizar as importâncias ressalva-se que temos de garantir aplicação desta técnica a uma base de dados comparável com a de outros métodos.

Acho que seria interessante aplicar as técnicas aqui usadas à base de dados SPEECHDAT ou outra equivalente.

BIBLIOGRAFIA

- [Apostol, 1978] T. M. Apostol, Calculus II, John Wiley & Sons, (1978).
- [Bellman & Zadeh, 1970] R. E. Bellman and Zadeh, Decision-making in a fuzzy environment, *Management Sci.* 17(4) (1970) 141-164.
- [Bogert & Peterson, 1957] B. P. Bogert and G.E. Peterson, The acoustics of speech, In L.E. Travis (ed) *Handbook of Speech Pathology*, Appleton-Century-Crofts, New York (1957) 109-173.
- [Burg, 1968] J. Burg, A New Analysis technique for time series data, *Proc. NATO Advanced Study Institute on Signal Proc.*, Enschede Netherlands (1978).
- [Chen & Hwang, 1992] S. J. Chen, and C. L. Hwang, *Fuzzy Multiple Attribute Decision Making*, Springer-Verlag (1992).
- [Chen, 1994] C. Chen, *System and Signal analysis*, Saunders College Publishing, 2ªEd (1994).
- [Dias de Deus, 1992] J. Dias de Deus, M. Pimenta, A. Noronha, T. Peña and P. Brogueira, *Introdução à Física*, MacGraw-Hill (1992).
- [Dilão, 1993] R. Dilão, *Introdução aos Sistemas Dinâmicos*, Sebenta Instituto Superior Técnico (1993).
- [Fermi, 1937] E. Fermi, *Thermodynamics*, Dover Publications, New York (1937).
- [Grabisch, 1995] M. Grabisch, Fuzzy integral in multicriteria decision making, *Fuzzy Sets and Systems* 69 (1995) 279-298.
- [Grabisch, 1998] M. Grabisch, S. A. Orlovski and R. R. Yager, Fuzzy aggregation of numerical preferences, edited by Roman Slowinski (1988) 31-68.
- [Halliday, 1988] D. Halliday and R. Resnick, *Fundamentals of physics*, Third edition, John Wiley & Sons (1988).
- [<http://speech.llnl.gov/thesis/index.htm>], Tese de Doutoramento.

- [<http://www.iurc.montp.insem.fr/cric/audition/english/sound/sound.htm>], método de Ghael-Sandgate.
- [Kent, 1978] R. D. Kent, Imitation of synthesized vowel by preschool children, *J. Acoust. Soc. Am.* 63 (1978) 1193-1198.
- [Kent, 1979] R. D. Kent, Isovowel lines for the evaluation of vowel formant structure in speech disorders, *J. of Speech and Hearing Disorders* 44 (1979) 513-521.
- [Klir, 1988] G.J. Klir and T.A. Folger, *Fuzzy Sets, Uncertainty, and Information*, Prentice-Hall International Editions (1988).
- [MacCrimmon, 1973] K. R. MacCrimmon, An overview of multiple objective decision making, in: Cochrane and Zeleny, Eds., *Multiple Criteria Decision Making*, University of South Carolina Press (1973).
- [Oppenheim, 1999] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time signal processing*, second edition, Prentice Hall (1999).
- [Peterson & Barney, 1952] G. E. Peterson and H. L. Barney, Control methods used in a study of vowels, *J. Acoust. Soc. Am.* 24 (1952) 175-184.
- [Pires, 1991] J. M. Pires, Modelos de raciocínio com incerteza e imprecisão, Tese de Mestrado, Universidade Nova de Lisboa (Faculdade de Ciências e Tecnologia) (1991).
- [Rabiner, 1978] L. Rabiner and R. Schaffer, *Digital Processing of speech signals*, Prentice Hall Signal Processing Series (1974).
- [Rabiner, 1993] L. Rabiner and B. Juang, *Fundamentals of Speech Recognition*, Prentice Hall Signal Processing Series (1993).
- [Ribeiro, 1996] R.A. Ribeiro, Fuzzy Multiple Attribute Decision Making: A Review and New Preference Elicitation Techniques. *Fuzzy Sets and Systems*, 78 (1996) 155-181.
- [Rodrigues & Trancoso, 1996] F. Rodrigues and I. Trancoso, Digit Recognition Using the SPEECHDAT Corpus, CONFTELE'99-2nd conference on telecommunications (1999).
- [Satty, 1977] T. L. Satty, *Modeling Unstructured Decision Problems - The*

theory of Analytical Hierarchies, *Mathematics and Computers in Simulation* 20 (1977) 147-158.

[Satty, 1978] T. L. Satty, Exploring the interface between hierarchies, multiple objectives and fuzzy sets, *Fuzzy Sets and Systems* 1 (1978) 57-68.

[Shaughnessy, 1990] D. O. Shaughnessy, Speech communication human and machine, Addison-Wesley, New York (1990).

[Simas & Ribeiro, 2001a] T. Simas, R. A. Ribeiro, Reconhecimento de vogais utilizando um método de decisão difuso, *Investigação Operacional* 21 (2001) 93-105.

[Simas & Ribeiro, 2001b] T. Simas, R. A. Ribeiro, Number Recognition Using a Fuzzy Decision Method, EURO Summer Institute (ESI) XIX - Decision Analysis and Artificial Intelligence- Toulouse, France, (2001).

[Starr, 1977] M. K. Starr and M. Zeleny, MCDM – state and future of the arts, *TIMS Stud. Management Sci.* 6 (1977) 5-29.

[Titze, 1994] R. I. Titze, Principles of voice production, Prentice-Hall, Englewood Cliffs, NJ (1994).

[Turban, 1998] E. Turban, Decision Support and Expert Systems, MacMillan, New York (1998).

[Yager, 1988] R. R. Yager, On Ordered Weighted Averaging Aggregation Operators in Multicriteria Decision Making, *IEEE Transactions on Systems, Man, and Cybernetics* 18(1) (1988) 183-190.

[Yager, 1999] R. R. Yager, and A. Kelman, An Extension of the Analytical Hierarchy Process Using OWA Operators, *Journal of Intelligent and Fuzzy Systems* 7 (1999) 401-417.

[Zadeh, 1973] L. A. Zadeh, Outline of a new approach to the analysis of complex systems and decision process, *IEEE Trans. Systems Man Cybernet*, SMC-3(1) (1973) 28-44.

[Zadeh, 1983] L. A. Zadeh, A computational approach to fuzzy quantifiers in natural languages, *Computing and Mathematics Applications* 9 (1983) 149-184.



[Zimmermann, 1987] H. J. Zimmermann, Fuzzy Sets, Decision Making, An Expert Systems, International Series in Management Science/Operational Research, Kluwer Academic, Dordrecht (1987).

[Zimmermann, 1991] H. J. Zimmermann, Fuzzy Set Theory and Its Applications, second edition, Kluwer Academic, Dordrecht (1991).

of