

**Universidade Federal de Santa Catarina  
Programa de Pós-graduação em Engenharia Elétrica**

**Reconhecimento de Faces Humanas Através  
de Técnicas de Inteligência Artificial  
Aplicadas a Formas 3D**

Trabalho apresentado como parte dos requisitos para a obtenção do título de Doutor em  
Engenharia Elétrica na Universidade Federal de Santa Catarina,  
Florianópolis, Santa Catarina, Brasil.

**Antonio Carlos Zimmermann**

Florianópolis, 26 de março de 2003

## APROVAÇÃO

### Reconhecimento de Faces Humanas Através de Técnicas de Inteligência Artificial Aplicadas a Formas 3D

Antonio Carlos Zimmermann

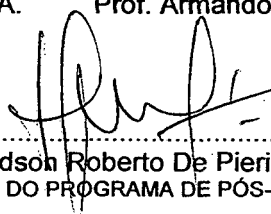
Esta Tese foi julgada adequada para a obtenção do título de

**"DOUTOR EM ENGENHARIA",**

na área de concentração de **SISTEMAS DE INFORMAÇÃO** e aprovada na sua forma final pelo PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA da UNIVERSIDADE FEDERAL DE SANTA CATARINA.

  
Prof. Jorge Muniz Barreto, D. Sc. A.  
ORIENTADOR

  
Prof. Armando Albertazzi Gonçalves Jr., Dr. Eng.  
CO-ORIENTADOR

  
Prof. Edson Roberto De Pieri, Dr.  
COORDENADOR DO PROGRAMA DE PÓS-GRADUAÇÃO

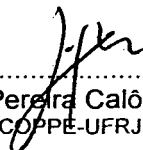
### BANCA EXAMINADORA


  
Prof. Jorge Muniz Barreto, D. Sc.A.  
INF-UFSC

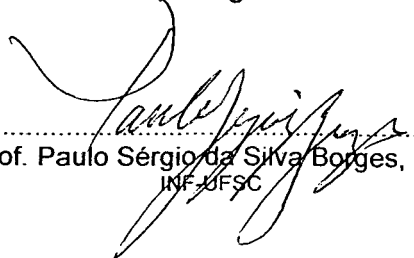
  
Prof. Armando A. Gonçalves Jr., Dr. Eng.  
EMC-UFSC

  
Profa. Ana Maria Beltran Pavani, D. Sc.  
ELE-PUC-RJ

  
Prof. Fernando Mendes de Azevedo, Dr. Sc.  
EEL-UFSC

  
Prof. Luiz Pereira Calôba, Dr. Ing.  
COPPE-UFRJ

  
Prof. Marco Antonio Martins Cavaco, Ph. D.  
EMC-UFSC

  
Prof. Paulo Sérgio da Silva Borges, Dr.  
INF-UFSC

## Resumo

Esta tese propõe uma nova metodologia de reconhecimento de faces humanas. Diferente dos tradicionais métodos que empregam imagens bidimensionais e níveis de cinza, o paradigma aqui proposto utiliza a forma geométrica da face como parâmetro de avaliação e autenticação. Destacam-se como vantagens desta técnica o melhor desempenho principalmente frente aos problemas de iluminação e posicionamento espacial.

Este trabalho engloba duas áreas de conhecimento distintas. A primeira, abordando aspectos da extração da forma tridimensional da face humana e a segunda o reconhecimento facial. A metodologia de extração da geometria baseia-se no método de Perfilometria de Fourier para obter a forma geométrica das faces. A metodologia de reconhecimento de faces divide-se em duas abordagens distintas: A primeira delas do tipo algorítmica, por Raciocínio Baseado em Casos - RBC empregando a distância de Hamming como medidor da similitude entre duas formas de faces e a segunda do tipo conexionista, baseando-se no emprego de Redes Neurais Artificiais –RNA do tipo Funções de Base Radiais - FBR, para a classificação das faces.

O modelo de reconhecimento adotado nesta tese é o da verificação, onde o indivíduo se apresenta previamente e ao sistema é atribuída a tarefa de verificação da veracidade da identidade alegada. A inspiração biológica está totalmente presente neste trabalho, em primeiro lugar porque os seres humanos, no processo de reconhecimento de seus semelhantes, utilizam-se das faces humanas e a forma da face é uma das grandezas identificadoras. Em segundo lugar porque as metodologias de reconhecimento inspiradas em Redes Neurais Artificiais são inerentemente paradigmas biológicos.

Finalmente empregando uma grande base de faces humanas tridimensionais, são mostrados os resultados da aplicação dos paradigmas de verificação desenvolvidos, comparando-se os resultados obtidos através de uma análise detalhada.

# Abstract

This thesis proposes a new methodology of human faces recognition, different from the traditional methods that use two-dimensional images and gray levels. The paradigm proposed here uses the geometric form of the face as evaluation parameter for authentication. The viabilization of the use of the tridimensional faces is being possible in the days today, due to the increase of the capacity of processing of the new computers. The advantages of this technique are the best acting mainly front to illumination problems and geometric positioning.

This work involves two different knowledge areas. The first, approaching aspects of the human face extraction in a three-dimensional way and second the facial recognition. The methodology for the extraction of the geometry bases on the method of Fourier Perfilometry to obtain the geometric form of the faces. The methodology for the recognition of faces are divided in two different approaches: The first of them, the algorithmic type by Case Base Reasoning uses the distance of Hamming as meter of the similarity between two forms of faces and the second, a conexionist model, basing in the use of Artificial Neural Networks with Radial Basis Functions, for the face's classification.

The recognition model adopted in this thesis is the verification's method, where the person previously identifies itself and to the system is attributed the task of verification the truthfulness of the alleged identity. The biological inspiration is totally present in this work, in first place because the humans, in the process of their fellow creatures' recognition uses the human faces and the form of the face is one of the identifier greatness used. In second place this is due because the inspired recognition methodologies like Artificial Neural Networks are inherently biological paradigms.

Finally using a huge base of tridimensional human faces, are shown the application's results of the developed verification paradigm, being compared the results obtained through a detailed analysis.

# Dedicatória

À minha esposa Rosana,  
meus filhos Michael e Richard,  
aos meus pais Irineu Carlos (in memorian) e Iris Maria.

## Agradecimentos

Ao professor D. Sc. A. Jorge Muniz Barreto pela orientação, ricas discussões, confiança e amizade repassadas ao longo deste profícuo período de convivência.

Ao professor Dr. Eng. Armando Albertazzi Gonçalves Jr., pela co-orientação, apoio, amizade e concordância para o desenvolvimento desta pesquisa.

Aos amigos e professores Dr. rer. nat. Frank Hrebabetzky, Dr. Eng. João da Silva Dias, Dr. João Vollert e Dr. Paulo Sérgio da Silva Borges, pelo enriquecimento advindo de discussões sobre pontos específicos do tema apresentado.

Ao estagiário Leonardo Soliz Encinas, pela sempre grande e irrestrita ajuda.

À mestranda Luciene de Oliveira Marin e ao estagiário Procópio Silveira Stein pela colaboração prestada.

A todas pessoas que integraram a base de dados de faces humanas.

Aos colegas do LABMETRO, IEB, L<sub>3</sub>C e UFSC pelo companheirismo e apoio recebido.

Aos membros da Banca Examinadora, pelas importantes contribuições sugeridas.

Ao LABMETRO, IEB e à Universidade Federal de Santa Catarina, pela oportunidade concedida.

Ao CNPq, que através do PADCT disponibilizou recursos ao projeto SORFACE, agilizando o desenvolvimento desta pesquisa.

À minha família pelo apoio incondicional.

Enfim, à todos que direta ou indiretamente atuaram para o êxito deste trabalho.

# Sumário

<b><u>APROVAÇÃO</u></b>	<b>ii</b>
<b>Resumo</b>	<b>iii</b>
<b>Abstract</b>	<b>iv</b>
<b>Dedicatória</b>	<b>v</b>
<b>Agradecimentos</b>	<b>vi</b>
<b>Sumário</b>	<b>vii</b>
<b>Lista de Figuras</b>	<b>xi</b>
<b>Lista de Tabelas</b>	<b>xiv</b>
<b>Lista de Símbolos</b>	<b>xv</b>
<b>Lista de Siglas</b>	<b>xvi</b>
<b>Capítulo 1</b>	
<b>Introdução</b>	<b>18</b>
1.1 Motivação e Justificativas .....	20
1.2 Objetivos .....	21
1.3 Concepção e Aspectos Inovadores da Tese .....	21
1.4 Aplicabilidade .....	22
1.5 Apresentação do Trabalho .....	22
<b>Capítulo 2</b>	
<b>Biometria</b>	<b>24</b>
2.1 Os Sistemas de Reconhecimento .....	25
2.2 Biometria Como Paradigma de Verificação .....	26
2.3 Características Desejáveis .....	28
2.4 Avaliação de Desempenho de Sistemas Biométricos .....	28

2.4.1	A base de dados.....	29
2.4.2	A Classificação .....	29
2.4.3	Análise do Limiar de Classificação.....	31
2.4.4	O Dilema do Limiar de Classificação.....	33
2.5	Técnicas Biométricas Clássicas .....	35
2.5.1	Impressão Digital .....	36
2.5.2	Geometria da Mão .....	36
2.5.3	Face .....	37
2.5.4	Íris .....	37
2.5.5	Retina .....	37
2.5.6	Vascularização .....	38
2.5.7	Voz .....	38
2.5.8	Assinatura.....	38
2.5.9	DNA.....	39
2.5.10	Outras Técnicas.....	39
2.6	Comparação das Técnicas .....	39
2.7	Aplicabilidade Quanto à Curva de Performance .....	41
2.8	A Importância do Protocolo de Ensaios .....	41
<b>Capítulo 3</b>		
<b>Reconhecimento da Geometria Facial</b>		<b>43</b>
3.1	Reconhecimento de Padrões.....	43
3.1.1	Abordagens em Reconhecimento de Padrões .....	45
3.1.2	Aplicações do Reconhecimento de Padrões .....	48
3.2	Reconhecimento de Faces .....	49
3.3	Sistemas de Reconhecimento de Faces.....	50
3.3.1	Métodos Baseados em Características Geométricas Locais .....	50
3.3.2	Métodos Baseados em Modelos de Representações Globais.....	52
3.3.3	Métodos de Modelos Híbridos.....	55
3.3.4	Os Resultados dos Métodos de Reconhecimento .....	56
3.4	Protocolo de Ensaios.....	58
3.5	Aplicações dos Sistemas de Reconhecimento de Faces .....	59
<b>Capítulo 4</b>		
<b>Medição da Geometria</b>		<b>60</b>
4.1	Técnicas Ópticas de Medição de Formas .....	60
4.1.1	Radar Laser.....	61
4.1.2	Topogrametria .....	61
4.1.3	Técnicas de Moiré .....	62
4.1.4	Holografia Interferométrica.....	64
4.1.5	Interferometria de Speckle .....	64
4.1.6	Medição por Pontos com Triangulação .....	65
4.1.7	Medição por Linhas com Triangulação.....	65
4.1.8	Sistemas de Varredura .....	66
4.1.9	Projeção de Franjas.....	67
4.2	O Princípio de Medição de Formas Utilizado .....	68
4.2.1	Detalhamento do Princípio de Projeção de Franjas.....	69



4.2.2	Processamento do Mapa de Franjas.....	70
4.2.3	Extração do Sinal das Franjas.....	70
4.2.4	Remoção do Salto de Fase.....	73
4.2.5	O Sistema Óptico Geométrico.....	75
4.3	A Base de Faces Humanas.....	84
4.3.1	As Imagens de Face da Base.....	85
4.3.2	As Formas de Faces da Base.....	88
4.3.3	O Tamanho da Base de Dados.....	88
4.3.4	A População da Base.....	89
<b>Capítulo 5</b>		
<b>Reconhecimento por Raciocínio Baseado em Casos - RBC</b>		<b>91</b>
5.1	O Raciocínio Baseado em Casos - RBC.....	92
5.1.1	O Processo de Recuperação.....	94
5.1.2	O Processo de Reutilização.....	95
5.1.3	O Processo de Revisão.....	95
5.1.4	O Processo de Retenção.....	95
5.2	Cálculo da Similitude.....	96
5.2.1	A Distância Euclidiana.....	96
5.2.2	A Distância de Hamming.....	97
5.2.3	A Similitude e o Limiar de Classificação.....	98
5.3	Otimização no Cálculo das Distâncias.....	98
5.3.1	Normalização Espacial.....	98
5.3.2	Minimização dos Erros de Posicionamento.....	101
5.4	A Ponderação Facial.....	103
5.4.1	Ponderação Comportamental.....	103
5.4.2	Ponderação Estatística.....	104
5.4.3	Comparação Entre os Métodos.....	105
<b>Capítulo 6</b>		
<b>Abordagem Conexionista</b>		<b>107</b>
6.1	Introdução às Redes Neurais Artificiais.....	108
6.1.1	Um Pouco da História das RNAs.....	108
6.1.2	Modelo do Neurônio Artificial.....	110
6.1.3	Funções de Ativação.....	113
6.1.4	Arquitetura das RNAs.....	114
6.1.5	Funcionamento das Redes Neurais Artificiais.....	116
6.1.6	Aprendizado das RNAs.....	117
6.1.7	Paradigmas de Aprendizado.....	119
6.1.8	Regras de Aprendizado das RNAs.....	121
6.1.9	Redes Neurais para Reconhecimento de Padrões.....	123
6.1.10	Justificativa da Escolha do Paradigma.....	127
6.2	Rede Neural do Tipo Função de Base Radial.....	128
6.2.1	Arquitetura da RNA tipo FBR.....	128
6.2.2	Relação entre a Entrada e a Saída da Rede.....	129
6.2.3	Funções de Ativação.....	130
6.2.4	Número de Neurônios.....	132

---

6.2.5	O Raio da FBR .....	133
6.2.6	Métodos de Treinamento Não Supervisionados .....	136
6.2.7	Métodos de Treinamento Supervisionados .....	139
6.3	Considerações Importantes no Projeto com RNAs .....	141
6.3.1	Invariância em RNAs .....	141
6.3.2	Redução da Dimensionalidade .....	142
<b>Capítulo 7</b>		
<b>Validação do Sistema</b> .....		<b>149</b>
7.1	A Concepção do Sistema .....	149
7.2	Extração das Formas de Faces Humanas .....	150
7.2.1	Os Sinais do Processamento da Forma .....	152
7.2.2	Ajuste do Sistema .....	153
7.2.3	Vantagens e Problemas Potenciais .....	154
7.2.4	Fontes de Erros .....	155
7.3	Reconhecimento da Forma .....	156
7.3.1	A Base de Dados dos Ensaios .....	156
7.3.2	Ensaios de Verificação Através do RBC .....	157
7.3.3	Ensaios de Verificação Através de RNAs do tipo RBF .....	161
<b>Capítulo 8</b>		
<b>Conclusões e Perspectivas</b> .....		<b>172</b>
8.1	Conclusões Sobre a Extração de Dados das Faces .....	172
8.2	Conclusões sobre o Reconhecimento das Faces .....	173
8.3	Considerações Gerais .....	174
8.4	Principais Contribuições deste Trabalho .....	175
8.5	Perspectivas para Trabalhos Futuros .....	175
<b>Referências</b> .....		<b>178</b>
<b>Índice Remissivo</b> .....		<b>186</b>
<b>Apêndice A</b>		
<b>Telas dos Softwares Desenvolvidos</b> .....		<b>193</b>
<b>Apêndice B</b>		
<b>Imagens da Bancada de Ensaio</b> .....		<b>196</b>

## Lista de Figuras

Figura 1.1	Imagem de indivíduo conhecido com baixa resolução.....	19
Figura 1.2	Diagrama esquemático do Sistema de Aquisição de Faces Humanas - SAFH....	21
Figura 2.1	Biometria e os paradigmas da verificação.....	27
Figura 2.2	Distribuição probabilística de lobos e ovelhas e suas taxas de erros. ....	30
Figura 2.3	Distribuições das populações em um processo de verificação.....	31
Figura 2.4	TFA e TFR em função do Limiar de Classificação.....	33
Figura 2.5	Curva de Performance da TFA e TFR.....	34
Figura 3.1	Representação de sistemas e classificação e interpretação [Schalkoff, 1992]. ...	45
Figura 3.2	Estrutura típica de um banco de dados de faces.....	57
Figura 4.1	Diagrama esquemático de um medidor do tipo RADAR LASER .....	61
Figura 4.2	Diagrama esquemático do sistema topogramétrico.....	62
Figura 4.3	Princípio de funcionamento do Sistema Moiré de sombras.....	63
Figura 4.4	Moiré de Projecção.....	64
Figura 4.5	Medidor de deslocamentos tipo triangulação.....	65
Figura 4.6	Diagrama esquemático do medidor por plano de luz.....	66
Figura 4.7	Diagrama esquemático de um medidor do tipo varredura.....	66
Figura 4.8	Projecção de franjas com e sem o problema de ambigüidade.....	67
Figura 4.9	Luz codificada para evitar o problema de ambigüidade das franjas.....	68
Figura 4.10	Diagrama esquemático do sistema de projecção de franjas.....	69
Figura 4.11	Representação das franjas projetadas em uma face humana.....	70
Figura 4.12	Espectro do sinal $G(f_x, y)$ .....	72
Figura 4.13	Sinal genérico de fase $\theta(x,y)$ com salto de fase.....	73
Figura 4.14	Sinal genérico de fase $\theta(x,y)$ com salto de fase removido.....	74
Figura 4.15	Geometria do sistema óptico.....	76

---

Figura 4.16	Relação entre fase e altura no objeto.....	82
Figura 4.17	Amostra de imagens típicas da base de faces. ....	86
Figura 4.18	Posicionamento espacial das faces no banco de faces.....	87
Figura 5.1	Representação da relação problema/solução em RBC. ....	93
Figura 5.2	Ciclo dos 4 RE's do RBC [Aamodt e Plaza, 2002]. ....	94
Figura 5.3	Visualização do cálculo de distâncias entre duas formas de faces. ....	98
Figura 5.4	Etapas da normalização das formas das faces. ....	99
Figura 5.5	Tempo de permanência do olho sobre as regiões da face no aprendizado [Minut et al., 2002] e [Henderson et al., 2000]. ....	104
Figura 5.6	MPP de regiões faciais segundo o critério de ponderação estatística. ....	105
Figura 5.7	Priorização de regiões faciais segundo dois métodos distintos. ....	106
Figura 6.1	Representação de um neurônio biológico [Zurada, 1992]. ....	111
Figura 6.2	Modelo de McCulloch e Pitts para o neurônio artificial [Haykin, 1994]. ....	112
Figura 6.3	Principais funções de ativação ou de transferência.....	113
Figura 6.4	Arquitetura típica de uma RNA genérica tipo direta. ....	115
Figura 6.5	Validação cruzada dos dados do conjunto de treino. ....	118
Figura 6.6	Curva típica de erro de treinamento de uma RNA. ....	119
Figura 6.7	Uma taxionomia do processo de aprendizado [Haykin, 1994]. ....	119
Figura 6.8	Padrões (a) linearmente separáveis e (b) não linearmente separáveis. ....	124
Figura 6.9	Arquitetura genérica típica de uma RNA do tipo FBR.....	129
Figura 6.10	Curvas de interpolação polinomial de funções. ....	133
Figura 6.11	Representação do Espaço de Formas de Faces - EFF como sub-espço do Espaço de Formas - EF. ....	144
Figura 6.12	Representação de dados bidimensionais segundo eixos quaisquer (a) e ortonormais (b). ....	144
Figura 7.1	Face com franjas (a) e sinal de modulação da face (b). ....	150
Figura 7.2	Face com salto de fase (a) e face sem salto de fase (b). ....	151
Figura 7.3	Dados típicos das alturas da forma da face de um manequim com escala real em milímetros. ....	151
Figura 7.4	Sinais característicos do processo de extração de formas. ....	152
Figura 7.5	Padrão de dimensional corporificado. ....	153
Figura 7.6	Comparação entre o modelo simplificado e o completo de cálculo de alturas. ...	154
Figura 7.7	Forma de face humana real com imperfeições típicas.....	156

---

Figura 7.8	Distribuição da DH para os dados originais.....	158
Figura 7.9	Distribuição da DH para os dados com ONE.....	158
Figura 7.10	Distribuição da DH para os dados com OPF.....	159
Figura 7.11	Distribuição da DH para os dados com OH.....	159
Figura 7.12	Caracterização do ponto de igualdade das taxas TFR e TFA.....	160
Figura 7.13	Distribuições das DH's para resoluções de (a) 41x29 e (b) 61x43.....	161
Figura 7.14	Comportamento do erro de treinamento em função do Raio da FBR e do número de neurônios.....	162
Figura 7.15	Comportamento do erro de validação em função do raio da FBR e do número de neurônios.....	162
Figura 7.16	Histograma das distâncias entre centros dos pontos do conjunto de treinamento.....	163
Figura 7.17	Resultado da classificação dos lobos e ovelhas usando RNAs do tipo RBF.....	164
Figura 7.18	Histograma da classificação das ovelhas e lobos.....	165
Figura 7.19	Raio da FBR e o número de neurônios utilizados no ensaio.....	165
Figura 7.20	Melhor Ajuste Linear - MAL e Coeficiente de Correlação - CC entre os dados reais e os ideais.....	166
Figura 7.21	Classificação das ovelhas e lobos para seis indivíduos (a) dados normais e (b) dados reduzidos pela ACP.....	168
Figura 7.22	Distribuição das ovelhas e lobos para seis indivíduos e suas taxas de TFR e TFA para (a) dados normais e (b) dados reduzidos pela ACP.....	168
Figura 7.23	Comportamento do erro de validação em função do número de exemplos de treinamento apresentados a rede.....	170
Figura 7.24	Comportamento do número de neurônios da camada intermediária em função do número de ovelhas no treinamento da rede.....	170
Figura 7.25	Comportamento do raio da FBR em função no número de ovelhas no treinamento da rede.....	171

## Lista de Tabelas

Tabela 2.1	Tabela de contingência genérica dos dados de uma classificação.....	35
Tabela 2.2	Comparação entre técnicas biométricas [Ruggles, 98 A.D.].....	40
Tabela 3.1	Desempenho das técnicas de reconhecimento de faces existentes. ....	56
Tabela 3.2	Desempenhos típicos de um sistema de verificação de faces.....	58
Tabela 4.1	Simplificações aplicáveis ao modelo completo de " $\beta_1$ " e " $\beta_2$ "......	81
Tabela 4.2	Composição da base de faces desenvolvida.....	88
Tabela 4.3	Classificação dos indivíduos segundo seus continentes/paises de origem.....	89
Tabela 4.4	Distribuição dos indivíduos da base segundo sua faixa etária.....	90
Tabela 4.5	Distribuição dos indivíduos da base segundo seu sexo.....	90
Tabela 6.1	Características das redes FBRs e PMCs. ....	126
Tabela 6.2	Funções de ativação do tipo base radial. ....	131
Tabela 6.3	Critérios para a escolha do raio da FBR.....	135
Tabela 6.4	Procedimento para cálculo de clusters segundo o algoritmo do k-médias.....	137
Tabela 6.5	ACP aplicada às matrizes de formas de faces. ....	148

## Lista de Símbolos

$\alpha$  – Ângulo entre os eixos ópticos da câmera e do projetor.

$\beta$  – Ângulo entre duas franjas consecutivas do ponto de vista da câmera.

$\gamma$  – Ângulo entre duas franjas consecutivas do ponto de vista do projetor.

$\theta(x,y)$  – Fase do sinal de intensidade luminosa das franjas.

$\omega$  – Velocidade angular.

$\varepsilon$  – Período das franjas, projetado sobre o plano de medição.

$\varepsilon'$  – Período das franjas, projetado sobre o plano perpendicular ao eixo óptico do projetor.

$\pi$  – Constante (3,14159..).

$\rho$  – Ângulo de rotação do objeto na etapa de normalização.

$\sigma_{n-1}$  – Desvio padrão dos dados da diferença das alturas dos objetos.

$\varphi$  – Função de ativação.

$\theta$  – Limiar de atuação da função de ativação.

$\Sigma$  – Ponto de confluência da rede neural artificial.

$\Phi$  – Matriz de interpolação.

# Lista de Siglas

ACP - Análise de Componentes Principais  
ADALINE - ADaptive LINear Elements  
ADL - Análise de Discriminantes Lineares  
AG - Algoritmos Genéticos  
AQV - Aprendizado de Quantificação Vetorial  
ARL - Army Research Laboratory  
ARSF - Algoritmos de Remoção de Saltos de Fase  
BC - Biometric Consortium  
CC - Coeficiente de Correlação  
CP - Curva de Performance  
DH - Distância de Hamming  
DVS - Decomposição em Valores Singulares  
DVS - Decomposição em valores Singulares  
EF - Espaço de Formas  
EFF - Espaço de Formas de Faces  
FBR - Função de Base Radial  
FC - Frequência de Cruzamento  
FERET - Face Recognition Technology  
FM - Frequência Modulada  
GBP - Gradient Back Propagation  
GD - Gradiente Descendente  
 $H_0$  - Hipótese Nula  
 $H_1$  - Hipótese Alternativa  
IBIA - International Biometric Industry Association  
IEEE - Institute of Electrical and Electronics Engineers  
LASER - Light Amplified Stimulated Emission of Radiation  
LC - Limiar de Classificação



LIDAR - Light Detector and Ranging  
LMTE - Valor Limite  
MADALINE - Multiple ADaptive LINear Elements  
MAL - Melhor Ajuste Linear  
MAOK - Mapas Auto-organizáveis de Kohonen  
MC - Matriz de Correlação  
MCA - Modelo de Contorno Ativo  
MCN - Método do Corte de Nós  
MMC - Máquina de Medir por Coordenadas  
MPP - Matriz de Pesos Próprios  
MPU - Matriz de Pesos Universais  
MQ - Mínimos Quadrados  
MQO - Mínimos Quadrados Ortogonais  
NIP - Número de Identificação Pessoal  
OH - Otimização Híbrida  
ONE - Otimização por Normalização Espacial  
OPF - Otimização por Ponderação Facial  
PA - Parada Antecipada  
PF - Perfilometria de Fourier  
PMC - Perceptron Multicamadas  
RBC - Raciocínio Baseado em Casos  
RNAs - Redes Neurais Artificiais  
RNP - Redes Neurais Probabilísticas  
RNRG - Rede Neural de Regressão Generalizada  
RRG - Regressão de Redes Generalizadas  
SARF - Sistema Automático de Reconhecimento de Faces  
TF - Transformada de Fourier  
TFA - Taxa de Falsa Aceitação  
TFR - Taxa de Falsa Rejeição  
TME - Taxa de Mesmo Erro  
USPS - United States Postal Service  
VC - Validação Cruzada  
VS - Valores Singulares

# Capítulo 1

## Introdução

---

*Nós alcançamos mais do que sabemos.*

*Nós sabemos mais do que entendemos.*

*Nós entendemos mais do que podemos explicar.*

- Claude Bernard, (1813 - 1878).

**O** reconhecimento de indivíduos é a base da cultura de credibilidade da atual sociedade humana. O sucesso na satisfação desta necessidade está firmemente interligado com a capacidade dos indivíduos em se identificarem com confiança. Atualmente esta tarefa é realizada por paradigmas de autenticação como assinatura, impressão digital, faces e outros. Daí a importância do reconhecimento na estrutura social em que hoje se vive.

Neste paradigma sociocultural alguns mecanismos e/ou procedimentos garantem a necessária confiança nos métodos adotados. Os principais recursos aceitos atualmente na autenticação de indivíduos estão baseados no reconhecimento indireto de assinaturas, impressões digitais, voz e face. Nenhum deles, porém pode ser considerado tecnologicamente suficientemente desenvolvido a ponto de poder ser considerado o melhor método.

O reconhecimento de faces humanas é a mais natural das técnicas biométricas. Os próprios seres humanos a utilizam na tarefa de identificação de seus pares. Eles são capazes de detectar e identificar faces com perfeição mesmo em um cenário complexo com um mínimo de esforço. Reconhecer faces é uma tarefa que os indivíduos repetem milhares de vezes diariamente e com extrema rapidez. Apesar dos indivíduos reconhecerem faces com facilidade, esta é uma tarefa extremamente complexa. Ela envolve a captura e processamento de inúmeras informações associadas ou não a informações adquiridas anteriormente e armazenadas na memória. Dos sentidos humanos, a visão é a que se constitui na principal ferramenta de captura de informações para a tarefa de reconhecimento.

Um dos principais aspectos positivos do reconhecimento de faces em humanos é sua robustez, eles são capazes de reconhecer imagens distorcidas (como no caso de caricaturas estudadas por Benson e Perrett [Benson e Perrett, 1991]), imagens com poucos detalhes como a apresentada na, figura 1.1, imagens com detalhes ocultos (indivíduos com óculos escuros p. ex.), bem como imagens invertidas (conforme reportado por Diamond e Carey [Diamond e Carey, 1986]). Somos capazes de reconhecer faces tão facilmente que alguns cientistas acreditam que no cérebro humano exista uma região dedicada e treinada para este tipo de processamento. Selos de correio e cédulas monetárias são um bom exemplo do uso da habilidade humana de reconhecimento de faces, onde falsificações são facilmente detectadas através da capacidade altamente desenvolvida de memorização e reconhecimento.

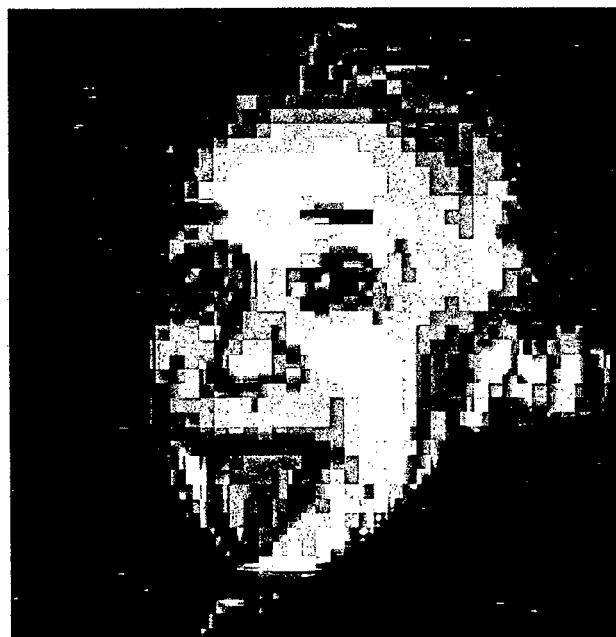


Figura 1.1 Imagem de indivíduo conhecido com baixa resolução.

A compreensão dos mecanismos que o cérebro humano emprega no reconhecimento de faces constitui ainda hoje um desafio para neurologistas e psicólogos.

Construir uma máquina capaz de reconhecer automaticamente um indivíduo com confiabilidade e eficiência, implica no conhecimento dos mecanismos biopsicológicos que o homem utiliza para resolver esta tarefa. Nesta linha de raciocínio a solução do problema proposto nesta tese remete ao uso de inspiração biológica através de técnicas computacionais modernas como a de inteligência artificial. Os principais paradigmas da inteligência artificial aplicáveis a este caso são as Redes Neurais [Cao et al., 1999], [Coimbra et al., 1999], [Cottrell e Fleming, 1990], [Edelman et al., 1992], [Garcia, 1999], [Ham e Park, 1999], Algoritmos Genéticos [Bala et al., 1998], [Bala et al., 1995], [Huang e Wechsler, 1999], Raciocínio Baseado em Casos [Aamodt e Plaza, 1994], [Barreto, 2001], [Zimmermann et al., 2002] e Árvores de Decisão [Bala et al., 1998], [Huang et al., 1996],

## **1.1 Motivação e Justificativas**

O reconhecimento de faces humanas é um capítulo aberto, apenas iniciando seu desenvolvimento. As técnicas existentes propostas para tal tarefa não atendem os requisitos metrológicos e operacionais quanto ao grau de incerteza dos resultados e velocidade de processamento.

A grande maioria dos sistemas de reconhecimento de faces existente emprega basicamente imagens 2D e processamentos estatísticos para tomadas de decisão. Estas tecnologias não oferecem ainda o grau de confiabilidade desejado para a maioria das aplicações a que se propõem, deixando espaço para que muito ainda se possa fazer com o aperfeiçoamento das técnicas existentes e/ou desenvolvimento de novas.

Acredita-se que a o princípio de reconhecimento aqui proposto, baseado na forma da face, tem mais chances de sucesso que os métodos tradicionais com imagens 2D e níveis de cinza. Isto se deve à invariância do método com a iluminação e à possibilidade de compensação do mau posicionamento espacial da face.

O sistema desenvolvido neste trabalho procurará através de uma combinação de técnicas atingir melhores desempenhos que os sistemas convencionais existentes, buscando principalmente uma maior robustez quanto a adversidades das condições de iluminação e posição geométrica da face.

Esta tese trata o termo “reconhecimento” como um termo biométrico genérico que engloba tanto o processo de “verificação” como o de “identificação” da identidades de indivíduos.

## 1.2 Objetivos

O principal objetivo deste trabalho é o de desenvolver uma técnica potencialmente robusta para o reconhecimento de faces humanas, combinando uma metodologia de aquisição e processamento de formas de faces humanas conjugada com uma metodologia de reconhecimento.

## 1.3 Concepção e Aspectos Inovadores da Tese

A abordagem de reconhecimento de faces proposta neste trabalho utiliza técnicas de iluminação estruturada [Harding, 1999], [Mass, 1999] e [Sorgel e Robert, 1999] para medir a geometria da face. Os dados da forma da face são obtidos através de uma câmera que captura a imagem da face com as franjas de luz obliquamente projetadas sobre ela. No processamento da imagem utiliza-se o método da análise de fase das franjas projetadas [Gasvik, 1995], [Robinson e Reid, 1993] e [Takeda e Mutoh, 1999]. Na figura 1.2 pode-se ver o diagrama esquemático de tal sistema. Neste trabalho, com vistas a uma intercomparação da eficiência dos métodos usados para a classificação da forma da face, o reconhecimento será realizado através de dois paradigmas distintos: (i) Raciocínio Baseado em Casos - RBC, (ii) Redes Neurais Artificiais - RNA. No método por RBC usa-se como classificador a Distância de Hamming - DH entre a face sob teste e a face padrão. No método por RNA será implementada uma rede do tipo Função de Base Radial - FBR como classificadora. Mais detalhes sobre o sistema de reconhecimento aqui proposto são descritos nos capítulos 4, 5 e 6.

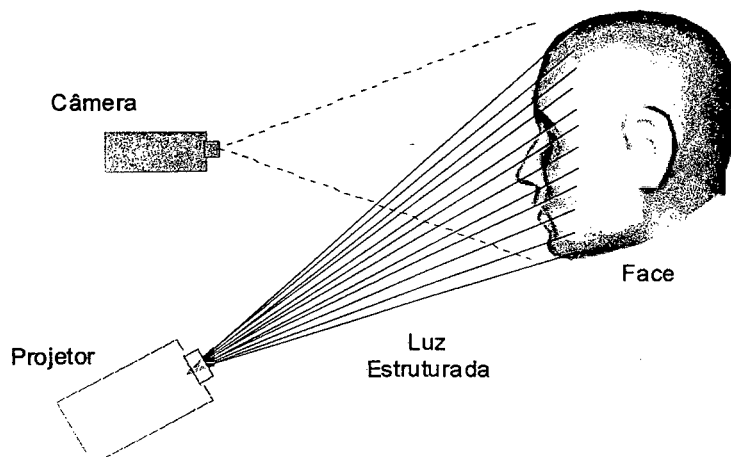


Figura 1.2 Diagrama esquemático do Sistema de Aquisição de Faces Humanas - SAFH.

---

O uso dos dados extraídos das formas de faces como entrada para um sistema de classificação é biologicamente natural e plausível. Não se tem notícias até agora da existência de sistemas de reconhecimento que utilizem especificamente estas técnicas conjugadas.

## 1.4 Aplicabilidade

A principal aplicação deste tipo de sistema de reconhecimento está associada à área de segurança. São potenciais interessados os fabricantes e usuários de sistemas de controle de acesso de indivíduos a locais e ou informações. Indiretamente a tecnologia desenvolvida tem sua aplicabilidade aumentada pois fundamentalmente é um sistema de reconhecimento de padrões podendo atender às mais diversas demandas, como por exemplo, automatização de linhas de produção industriais.

Após o final da guerra fria, grandes empresas internacionais, antes fabricantes de radares e outros grandes sistemas de segurança para as superpotências, voltaram seus interesses para as áreas de segurança mais específicas, a exemplo de uma empresa localizada em Liege<sup>1</sup> que havia desenvolvido um sistema de alarme baseado em som e estava desenvolvendo um sistema combinado com visão, usando inspiração biológica para a localização de invasores em locais protegidos [Bottemanne et al., 1994].

## 1.5 Apresentação do Trabalho

Este trabalho está dividido em mais 8 capítulos conforme descrição sucinta feita a seguir:

- i. Capítulo 2 - Apresenta os sistemas biométricos existentes incluindo uma apresentação e discussão sobre os procedimentos de avaliação metrológica e operacional destes.
- ii. Capítulo 3 - É apresentada uma revisão bibliográfica dos Sistemas de Reconhecimento de Padrões - SRP e de dos Sistemas de Reconhecimento de Faces Humanas - SRFH abordando as técnicas, discussões dos parâmetros relevantes e processos convencionais utilizados na solução do problema.
- iii. Capítulo 4 - Apresenta as técnicas existentes para medição da geometria e descreve detalhadamente a técnica utilizada neste trabalho. Aborda o princípio de funcionamento, aspectos ópticos e matemáticos envolvidos, bem como discute os aspectos práticos da montagem do experimento apresentando suas vantagens e

---

<sup>1</sup> Liege - Província Belga.

- 
- desvantagens. Apresenta também a base de dados de faces construída e as considerações relativas a este trabalho.
- iv. Capítulo 5 - Descreve o RBC como um dos paradigmas de classificação adotado. São abordados aspectos teóricos e práticos dentre eles a metodologia de minimização de erros geométricos. Também são apresentados a concepção do sistema e os algoritmos utilizados.
  - v. Capítulo 6 - Apresenta a abordagem da classificação por métodos conexionistas baseados em RNA's. O paradigma utilizado é o da FBR fazendo uma breve comparação com outros modelos. Abordam-se aspectos teóricos e práticos das RNA's e em específico da FBR, suas vantagens e desvantagens.
  - vi. Capítulo 7 - De cunho essencialmente experimental, apresenta a concepção do sistema e os resultados obtidos com os ensaios realizados pelas duas abordagens de classificação descritas nos capítulos 5 e 6. Também são discutidos os resultados dos ensaios.
  - vii. Capítulo 8 - É o fechamento do trabalho onde se discute os resultados experimentais da classificação das faces pelos paradigmas adotados e a interpretações destes. Apresentan-se sugestões de melhoria e continuidade do trabalho.

## Capítulo 2

# Biometria

---

Humanos nunca planejarão uma invenção mais bonita, mais simples  
ou mais direta do que a natureza, porque nas invenções dela,  
nada está faltando e nada é supérfluo.

- Leonardo Da Vinci, (1450-1519).

**B** iometria<sup>1</sup>, num sentido amplo, é o uso da matemática e estatística aplicadas à biologia. No domínio deste trabalho será adotada uma definição mais específica onde a “Biometria é a medição de características físicas ou comportamentais únicas em seres humanos usadas para a verificação da identidade destes” [Mansfield e Roethenbaugh, 1998]. Enquanto a palavra biometria parece ser relativamente nova, ela quantifica a mais antiga das técnicas de reconhecimento de indivíduos usada por humanos [Chandrasekaran, 1997]. Na

---

<sup>1</sup> Biometria. i. Ramo da ciência que estuda os seres vivos ii. Cálculo da duração provável da vida. iii. Biometria estatística. Parte da estatística que investiga atributos biológicos quantitativos pertinentes a uma população de seres vivos [Ferreira, 1986]. iv. Estudo das medidas e de estruturas e órgãos de seres vivos, bem como da importância funcional dessas medidas; biométrica v. Biometria ramo da estatística relacionado à ecologia, no qual se estudam as características biológicas quantitativas de uma população [Houaiss et al., 2001].



língua portuguesa, o primeiro registro desta palavra já é verificado no ano de 1913 [Houaiss et al., 2001].

Em se tratando de biometria no âmbito dos seres humanos, existem dois tipos de paradigmas biométricos que são o comportamental e o fisiológico. A biometria fisiológica deriva da medição direta de características físicas de uma parte do corpo humano como a impressão digital, íris, retina, faces e outras que serão abordadas com mais detalhes à frente neste capítulo. Por outro lado a biometria comportamental está baseada nos dados derivados das ações de um ser humano, que indiretamente refletem as características deste indivíduo. São exemplos de biometria comportamental a análise da voz, assinaturas e gestos. Geralmente as grandezas observadas na biometria comportamental possuem o tempo como fator métrico de discriminação. Na maioria dos casos a biometria fisiológica apresenta resultados com menor índices de erros.

## 2.1 Os Sistemas de Reconhecimento

Os sistemas de reconhecimento podem ser classificados em dois grandes grupos que são:

- i. Sistemas de Identificação
- ii. Sistemas de Verificação

Em sistemas de verificação [Mansfield e Roethenbaugh, 1998] as características de um indivíduo desconhecido são comparadas com as características de um indivíduo em particular<sup>2</sup> com o objetivo de verificar a identidade deste. Caso típico de quando o indivíduo se identifica previamente e o sistema simplesmente verifica a veracidade da informação com os dados reivindicados. Uma aplicação típica é a do uso no sistema bancário onde o indivíduo previamente cadastrado fornece seu Número de Identificação pessoal - NIP e após submete-se à checagem biométrica. O sistema determina o grau de coincidências existentes nas características biométricas do indivíduo que, uma vez sendo superior ao limiar de classificação previamente estipulado, considera o indivíduo positivamente identificado.

Em sistemas de identificação [Mansfield e Roethenbaugh, 1998] as características biométricas do indivíduo desconhecido são comparadas com as características dos indivíduos conhecidos pertencentes a uma base de dados<sup>3</sup> com o objetivo de achar o elemento da base que melhor se adapte a estas características, o qual será assumido ou não como elemento correspondente

---

<sup>2</sup> "One-to-one".

<sup>3</sup> "One-to-many".

dependendo dos níveis de aceitação/rejeição do sistema. Usado quando não se tem previamente nenhuma indicação de quem seria o indivíduo no universo do banco de dados. Um sistema policial de vigilância em aeroportos, por exemplo, compararia as características de um indivíduo com as de uma base de dados de elementos procurados.

Os esforços dos sistemas de identificação e verificação diferem substancialmente [Golfarelli et al., 1999]. No processo de verificação o sistema compara dois indivíduos, na identificação o indivíduo é comparado contra todos do grupo [Philips et al., 2000]. Esta é na realidade uma das maiores diferenças entre estes dois tipos de sistemas de reconhecimento.

É importante ressaltar as diferenças entre estes dois sistemas de reconhecimento, pois um sistema projetado para verificação não funciona para identificação e vice-versa. As diferenças residem no tipo de aplicação do sistema e conseqüentemente na arquitetura de implementação do mesmo. Enquanto na identificação o tempo gasto para a tarefa depende do tamanho da base de dados (quantidade de indivíduos na base), na verificação o sistema não é dependente do número de indivíduos da base. Em sistemas de identificação o desempenho da busca na base de dados é uma preocupação adicional. Nestes deve-se procurar empregar algoritmos para otimização do armazenamento e recuperação dos dados, bem como reduzir ao máximo a quantidade de informação armazenada tendo em vista o espaço a ser ocupado e o tempo de processamento das informações.

Apesar de que o escopo desta tese trata de um sistema de verificação a maioria dos conceitos e procedimentos que aqui serão firmados e discutidos valem, sem restrição, para os sistemas de identificação.

## **2.2 Biometria Como Paradigma de Verificação**

A verificação de indivíduos tem sido feita, tradicionalmente, por dois modelos que se baseiam em se “conhecer” alguma informação ou por se “ter” algum dispositivo que se qualifique para tal. A verificação pelo conhecimento é o caso clássico do uso de senhas alfanuméricas enquanto a verificação pela posse trata-se do uso de dispositivos que só o indivíduo qualificado possui, como chaves, por exemplo.

Os NIPs foram os primeiros autenticadores a serem usados em sistemas de reconhecimento automático. Contudo, deve-se ressaltar que o reconhecimento é do NIP e não do indivíduo que o possui. O mesmo se aplica a sistemas baseados em cartões magnéticos, “smart cards”, chaves eletrônicas e outros. Pode-se facilmente reconhecer o NIP mas não necessariamente o indivíduo que o está apresentando.

Estes dois métodos não são satisfatoriamente seguros, visto que “senhas” podem ser descobertas ou esquecidas e “chaves” podem ser roubadas ou copiadas. Neste sentido a biometria vem preencher uma lacuna de segurança em sistemas de verificação, pois ao invés de se basear em “conhecer” ou “ter” ela fundamenta no “ser”. Na figura 2.1 pode-se ver um diagrama desta inter-relação.

Características biométricas são inerentes ao indivíduo, sendo quase impossíveis de serem roubadas, copiadas ou esquecidas. Um indivíduo pode esquecer seu cartão ou seu NIP, mas não pode esquecer ou perder seus dedos, voz, olhos ou face.

Todavia não se deve desprezar completamente a contribuição dos paradigmas “ter” e “conhecer” visto que quando se busca um aumento da segurança em sistemas de verificação o modelo mais seguro, é na realidade uma combinação dos três. Por exemplo, cartões plásticos, mesmo que do tipo inteligentes<sup>4</sup>, quando usados sozinhos, são inseguros, visto que podem ser esquecidos, perdidos, roubados e ou tomarem-se ilegíveis. Um aumento na segurança pode ser atingido através da utilização de diferentes técnicas combinadas, onde os pontos fortes de uma determinada técnica compensem os pontos fracos de outra. Neste sentido, o uso de um sistema de verificação híbrido que utilize mais de um dos paradigmas acima é o mais indicado.

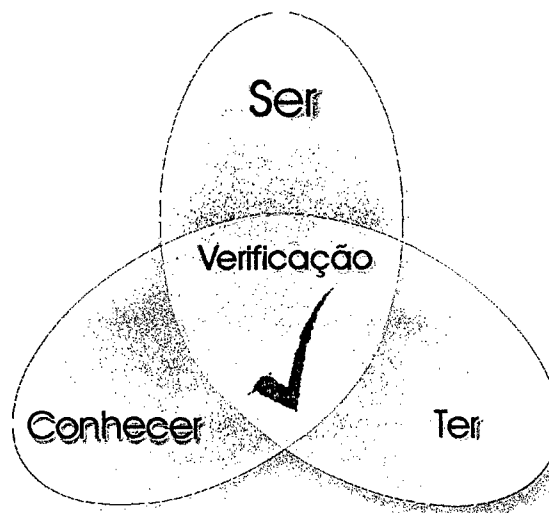


Figura 2.1 Biometria e os paradigmas da verificação.

<sup>4</sup> "Smart cards".

## 2.3 Características Desejáveis

A princípio qualquer característica biológica mensurável pode ser usada como grandeza biométrica. Porém existem alguns aspectos metrológicos e operacionais que devem ser observados na seleção destas características com o objetivo de melhorar o desempenho do processo como um todo. Quanto aos aspectos metrológicos da grandeza biométrica são propriedades desejadas:

- i. Universalidade: Cada indivíduo deve possuir a grandeza em questão.
- ii. Perenidade: A grandeza não deve mudar ao longo do tempo.
- iii. Singularidade: Cada indivíduo deve possuir um valor da grandeza diferente de outro indivíduo.
- iv. Mensurabilidade: A grandeza deve ser de fácil quantificação.

Quanto ao aspecto operacional dos sistemas de medição, são características desejáveis:

- i. Desempenho: Em termos da precisão conseguida com a medida da grandeza.
- ii. Aceitabilidade: Deve existir aceitabilidade da medição por parte dos usuários.
- iii. Segurança: Deve ser robusta contra fraudes.
- iv. Tempo de Processamento: O tempo de processamento do reconhecimento deve ser o menor possível.
- v. Facilidade de Uso: O sistema deve ser de fácil manuseio por qualquer tipo de indivíduo.
- vi. Privacidade: As informações pessoais devem permanecer seguras nos sistemas.
- vii. Custo: Baixo custo é sempre uma meta a ser atingida.

## 2.4 Avaliação de Desempenho de Sistemas Biométricos

O desempenho de um sistema de reconhecimento biométrico é sem dúvida o fator mais importante na análise da aplicabilidade destes sistemas. Os principais fatores de qualificação estão relacionados com a velocidade de processamento e os erros de classificação.

### **2.4.1 A base de dados**

A avaliação de um sistema biométrico se dá através da intercomparação dos indivíduos de uma base de dados. A base de dados é em geral composta de vários indivíduos com identidades diferentes bem como um número de variações em torno de cada um. No processo de verificação o número de indivíduos distintos presentes na base de dados está relacionado com o número de indivíduos distintos que se deseja autenticar. Já o número de variações de um mesmo indivíduo depende do algoritmo de classificação utilizado, bem como da amplitude das variações existentes para cada um.

### **2.4.2 A Classificação**

No processo de classificação de um sistema de validação biométrico o indivíduo sob teste primeiramente apresenta-se alegando determinada identidade. Assim, através de um processo de classificação propriamente dito este indivíduo é comparado, com base em determinados parâmetros, às características do indivíduo que este alega ser e que estão armazenadas na base de dados. O resultado desta comparação é um valor de pontuação entre zero (0) e um (1). Assim, se este valor estiver acima do Limiar de Classificação - LC faz com que o indivíduo seja admitido como pertencente à classe da qual ele declarava pertencer. Se o resultado da classificação produzir um valor menor que o LC o indivíduo é considerado não pertencente à classe alegada. Neste trabalho, convencionou-se associar indivíduos que pertencem a uma mesma classe a valores próximos de um (1) e indivíduos de classes diferentes a valores próximos de zero (0).

Na classificação aqui realizada, cada indivíduo será comparado contra os outros restantes da base de dados. Este procedimento é feito para todos os indivíduos da base de dados.

A população de indivíduos será separada então em duas grandes classes chamadas de classes dos lobos e classe das ovelhas. À classe dos lobos pertencerão todos os indivíduos da base de dados menos aquele que está sendo verificada a identidade. À classe das ovelhas pertencerão os indivíduos que estão sendo verificados a identidade e suas variações. É muito importante o entendimento desta concepção, pois, neste trabalho, toda a metodologia de avaliação dos sistemas biométricos será construída em cima destes conceitos. Na figura 2.2 pode-se ver uma representação típica destas distribuições.

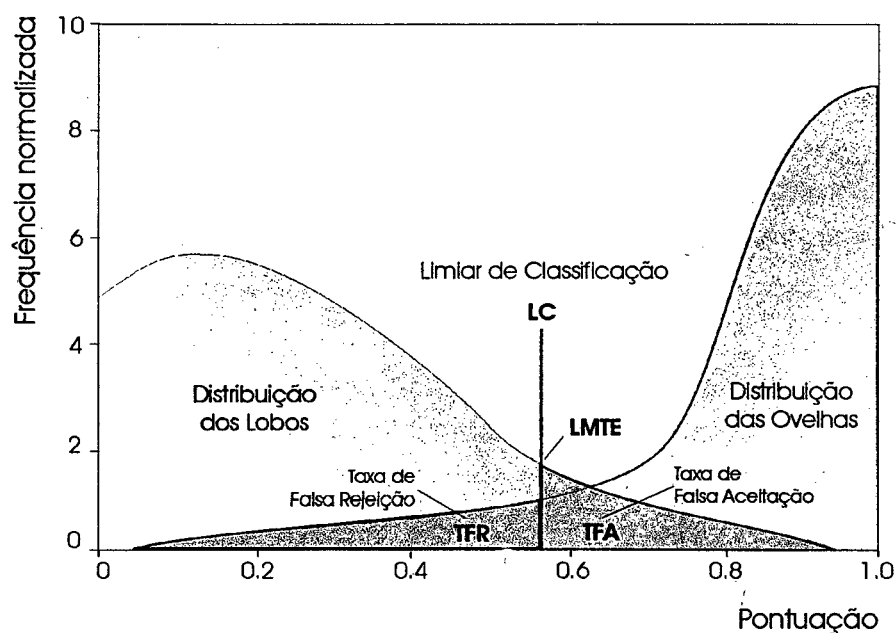


Figura 2.2 Distribuição probabilística de lobos e ovelhas e suas taxas de erros.

A figura 2.3 mostra um banco de dados típico e hipotético com 5 indivíduos diferentes onde cada um é verificado contra todos os outros gerando um total de 25 resultados. Supõe-se ainda existir alto grau de correlação entre somente dois indivíduos (B e C) podendo significar serem estes o mesmo indivíduo. O resultado dos testes será uma pontuação de 0 a 1 para cada combinação. Como possuímos  $N = 5$  temos a certeza de que pelo menos cinco elementos serão coincidentes e receberão assim uma pontuação alta. Ainda como existem dois elementos com alto grau de correlação (B e C) espera-se uma pontuação alta também para este par com valor dependendo da correlação entre elementos. Os outros  $(N * N) - (N - 2 * n)$  valores serão os que possuem baixo grau de correlação.

A distribuição dos resultados deste caso hipotético é semelhante a da figura 2.2 onde a distribuição de ovelhas representa a distribuição dos  $N + 2 * n$  elementos com alto grau de correlação e a distribuição de lobos representa os  $N * N - (N + 2 * n)$  indivíduos com baixa correlação.

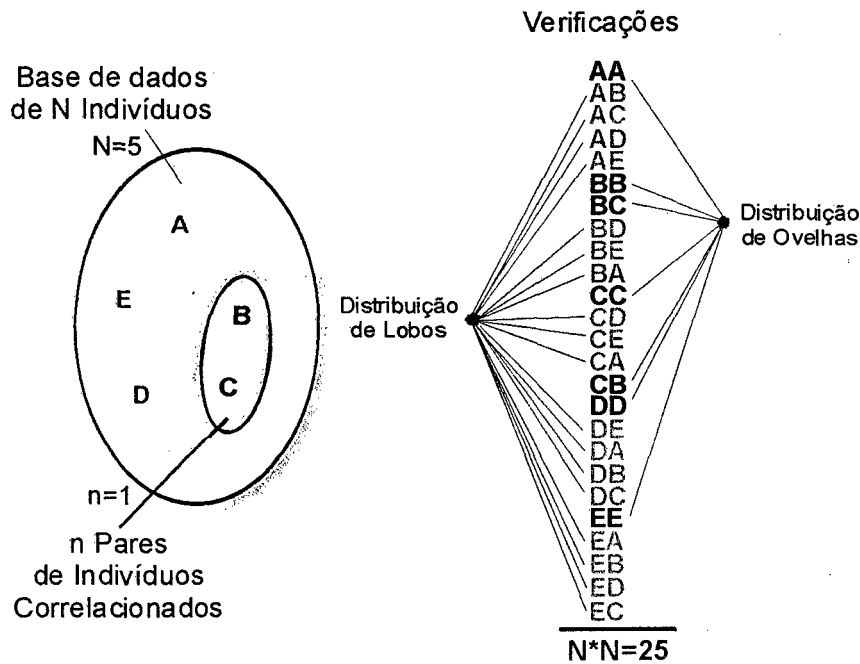


Figura 2.3 Distribuições das populações em um processo de verificação.

### 2.4.3 Análise do Limiar de Classificação

Muitos índices podem ser usados para a qualificação de sistemas de medição biométricos [Mansfield e Roethenbaugh, 1998], [Philips et al., 2000], porém os mais importantes, representativos e aceitos internacionalmente pela comunidade científica são a Taxa de Falsa Aceitação - TFA, Taxa de Falsa Rejeição - TFR e a Taxa de Mesmo Erro - TME [Golfarelli et al., 1999].

Estas taxas são obtidas aplicando-se a teoria da decisão estatística a qual auxilia a tomada de decisões acerca de populações, baseadas nas informações das amostras destas.

Supondo que a situação mais crítica seja a de que um sistema de segurança possa admitir indivíduos não credenciados, estipula-se como Hipótese Nula -  $H_0$  que os indivíduos sejam pertencentes à população dos credenciados e como Hipótese Alternativa -  $H_1$  que os indivíduos pertençam a população dos não credenciados. E assim tenta-se refutar  $H_0$ .

Se uma hipótese for rejeitada quando deveria ser verdadeira, diz-se que foi cometido um erro do tipo I e se por outro lado for aceita uma hipótese a qual deveria ter sido recusada, diz-se que

foi cometido um erro do tipo II. Nos dois casos anteriores houve uma decisão errada ou um erro de julgamento.

Na prática o que se busca é a minimização dos erros de decisão. Nem sempre isto é possível pois em geral a diminuição de um acarreta o aumento do outro. Em geral a diminuição de ambos é conseguida somente através do aumento do tamanho das amostras das populações.

### **Taxa de Falsa Aceitação - TFA**

É a probabilidade de cometer um erro do tipo II, onde indivíduos impostores são aceitos quando não deveriam ser. A TFA Representa a probabilidade dos lobos serem aceitos, assim:

$$TFA = \frac{\textit{Total de Falsas Aceitações}}{\textit{Total de Tentativas Fraudulentas}} \quad (2.1)$$

Esta taxa caracteriza-se pela situação onde o sistema erroneamente aceita um indivíduo sendo que este não pertence à população dos credenciados.

TFA's são características de sistemas com larga tolerância de classificação. Também são conhecidos por sistemas tolerantes, onde uma alta TFA significa falha de segurança.

### **Taxa de Falsa Rejeição - TFR**

É a probabilidade de cometer um erro do tipo I, onde indivíduos legítimos não são aceitos quando o deveriam ser. A TFR representa a probabilidade das ovelhas serem rejeitadas, assim:

$$TFR = \frac{\textit{Total de Falsas Rejeições}}{\textit{Total de Tentativas Legítimas}} \quad (2.2)$$

Esta taxa caracteriza-se pela situação onde o sistema erroneamente rejeita o indivíduo em questão, sendo que o indivíduo pertence à população dos credenciados.

TFR's são características de sistemas com estreito grau tolerância de classificação. Num contexto mais comum são ditos sistemas rígidos, onde uma alta TFR significa alta segurança.

### **Taxa de Mesmo Erro - TME**

Representa o ponto onde as taxas TFA e TFR possuem o mesmo valor. Na figura 2.4 vê-se a distribuição das TFA e TFR em função da variação do limiar de classificação.



Note que as duas taxas possuem comportamento antagônico. Ao se aumentar uma delas a outra diminui e vice-versa. Este comportamento muitas vezes dificulta a interpretação do desempenho dos sistemas de verificação. Em contrapartida a TME é um índice único que pode ser usado para expressar os erros de um sistema biométrico de reconhecimento.

A TME pode ser usada na avaliação de sistemas biométricos, usando-se uma mesma base de dados para todos os sistemas sob teste. Também pode ser usada na avaliação da grandeza biométrica, utilizando-se um mesmo sistema biométrico e várias grandezas sob teste.

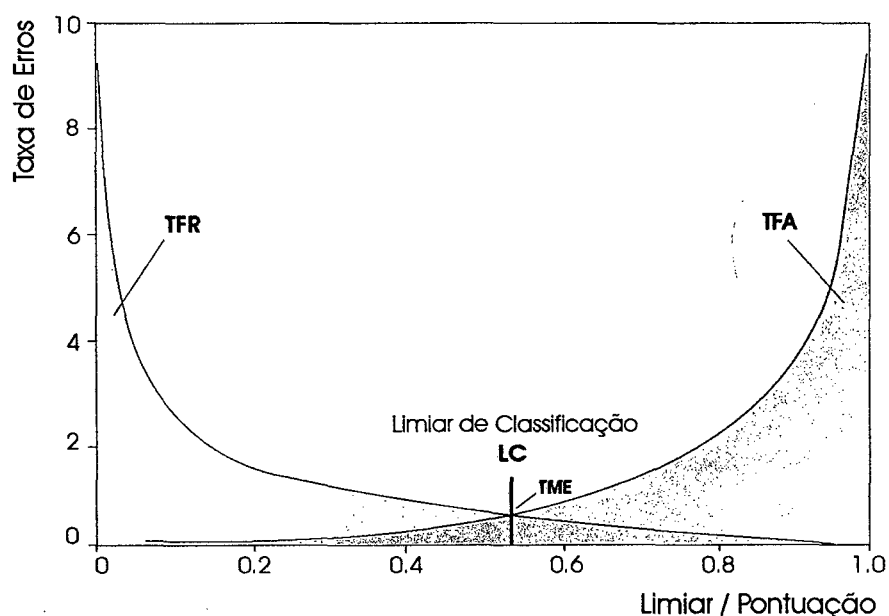


Figura 2.4 TFA e TFR em função do Limiar de Classificação.

#### 2.4.4 O Dilema do Limiar de Classificação

Somente em sistemas de validação ou identificação ideais ou reais de alto desempenho as TFA e TFR são zero. Na grande maioria dos sistemas comerciais existentes, as taxas TFA e TFR não são nulas. A figura 2.2 mostra a transposição das distribuições caracterizando uma situação em que qualquer que seja o valor do limiar de classificação escolhido não separará completamente as duas populações. Taxas de erros com valores não nulos levam sempre ao problema da escolha do limiar de classificação e suas implicações no desempenho e aplicação do sistema. A escolha do limiar deverá ser de um valor tal que não sejam aceitos todos os

indivíduos e que também não sejam rejeitados todos da base de dados. O sistema deve operar entre estes dois extremos.

O valor da TFA representa que parcela da população de lobos consegue ser autenticada positivamente quando deveriam ser rejeitados e o valor da TFR representa o quanto da população de ovelhas é rejeitada quando deveriam ser aceitas. Estes valores ocorrem aos pares, isto é para cada TFR existe uma TFA correspondente. Nestes casos o dilema reside no fato de que ao se procurar estabelecer uma baixa TFA, minimizando o aceite de lobos, também se está aumentando a TFR, ou seja se está aumentando a restrição de aceite para ovelhas e vice-versa.

### Curvas de Performance - CP

A figura 2.5 mostra o gráfico da Curva de Performance - CP<sup>5</sup> a qual expressa a relação intrínseca entre a TFA e TFR.

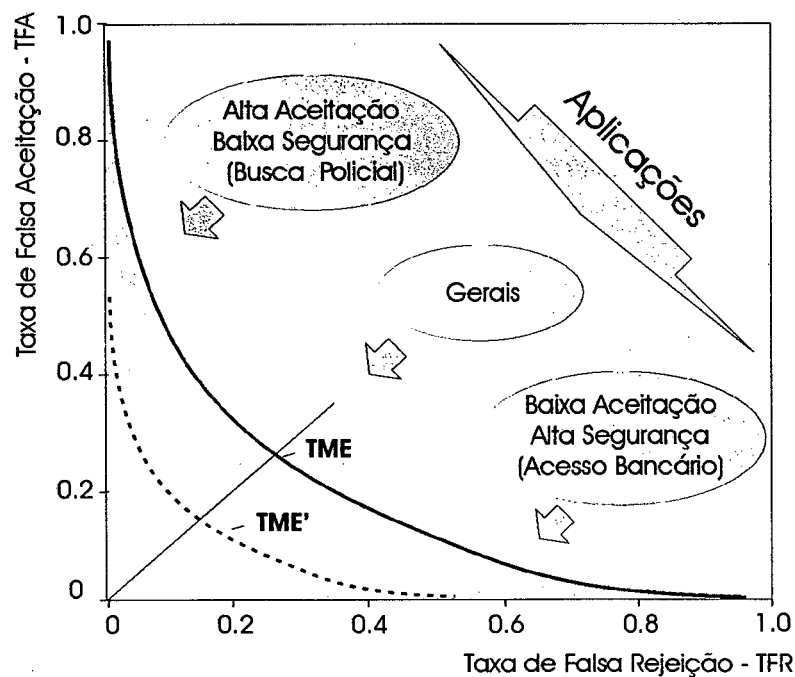


Figura 2.5 Curva de Performance da TFA e TFR.

<sup>5</sup> Também conhecida como "Receiver's Operating Characteristic" (ROC) [Gutta et al., 1995], "Detection Error-Reject Trade-off Curve" (DETC) [Golfarelli et al., 1999] ou "Performance Curve" (PC) [Philips et al., 2000].

Quanto menor o valor da TME menor será a área abaixo da CP e conseqüentemente melhor o desempenho do sistema. A figura 2.5 apresenta duas curvas correspondentes a dois sistemas com comportamentos diferentes. Pode-se ver que um sistema representado pela CP pontilhada (TME') tem menores TFA's e TFR's que um outro representado pela CP contínua (TME). Sistemas com menores taxas de erro representam sistemas com melhores características classificadoras. Também é comum definir-se a TME como sendo uma Frequência de Cruzamento - FC [Ruggles, 98 A.D.] sendo expressa na forma  $1 : X$ , onde :

$$X = \frac{1}{TME} \quad (2.3)$$

e  $X$  é geralmente arredondado para um número inteiro [Thian, 2001].

### Tabela de Contingência

Os resultados da classificação podem ser sumarizados em uma tabela de contingência como a mostrada na tabela 2.1.

Tabela 2.1 Tabela de contingência genérica dos dados de uma classificação.

		Classificação		
		Positiva	Negativa	Totais
Verdade	Ovelhas (%)	100 - TFR	TFR	100
	Lobos (%)	TFA	100 - TFA	100
	Totais	100-TFR+TFA	100-TFA+TFR	200

## 2.5 Técnicas Biométricas Clássicas

As principais medidas biológicas utilizadas, tanto fisiológicas como comportamentais, são:

- i. Impressões Digitais
- ii. Geometria da Mão
- iii. Face
- iv. Íris
- v. Retina
- vi. Vascularização

- vii. Voz
- viii. Assinatura
- ix. DNA

### **2.5.1 Impressão Digital**

É conhecido por todos que a impressão digital de um ser humano é singular. Impressões digitais são essencialmente representações gráficas do tipo fluxogramas. Suas formações dependem de condições iniciais na formação embrionária do ser. Esta é uma das mais antigas grandezas biométricas usadas em reconhecimento e de maior maturidade, sendo plenamente aceita como técnica de identificação em casos forenses.

O padrão de uma impressão digital é obtido tipicamente por uma das seguintes técnicas:

- i. Através do escaneamento de uma imagem impressa obtida pelo pressionamento do dedo envolvido por tinta fresca em uma folha de papel
- ii. Por um escâner de impressões digitais que extrai a imagem diretamente do dedo do indivíduo.

A classificação se faz através da análise das posições e quantidades de confluências e divergências do padrão digital.

Um dos primeiros relatos de sua aplicação reporta à China no século XIV, onde pais usavam a impressão digital dos pés das suas crianças para diferenciá-las de outras. A primeira classificação sistemática de impressões digitais foi estabelecida durante a ocupação da Índia pelos ingleses no século XVII que um policial de nome Henry [Ruggles, 98 A.D.]. Ainda hoje várias instituições em todo o mundo têm pesquisado a respeito inclusive o FBI<sup>6</sup>. Esta técnica é a mais usada comercialmente até esta data em bancos e em controle de acesso de áreas restritas.

### **2.5.2 Geometria da Mão**

Esta técnica é baseada no fato dos indivíduos possuírem a forma da mão bem definida. Vários são os sistemas usados para a medição que podem ser classificados em mecânicos e ópticos. Cada sistema utiliza parâmetros de controle próprios como comprimento de dedos, largura, etc. É interessante salientar que apesar do crescimento e envelhecimento do corpo humano, a geometria não muda significativamente após certa idade [Ruggles, 98 A.D.].

---

<sup>6</sup> "FBI - Federal Bureau of Investigation".

### **2.5.3 Face**

O reconhecimento de faces é uma técnica relativamente nova e em expansão devido ao recente surgimento e barateamento dos sistemas de aquisição de imagens. É uma técnica bastante passiva e natural, pois não exige que o indivíduo a ser reconhecido se posicione precisamente em relação ao sistema de câmeras. A face pode ser capturada a alguns metros de distância. É uma técnica não invasiva, não oferecendo nenhum tipo de risco a saúde do usuário. A desvantagem é que o sistema não é capaz de reconhecer gêmeos idênticos.

O reconhecimento se dá através da imagem adquirida da face, que basicamente pode ser do tipo bidimensional ou tridimensional. Medidas geométricas da face como distâncias entre olhos e nariz, curvatura da boca e outras ou como o uso de imagens da face como um todo são algumas das técnicas usadas na classificação de faces. É um sistema essencialmente inspirado na biologia visto ser esta a maneira com que os humanos reconhecem os seus semelhantes. Ao longo deste trabalho serão exaustivamente apresentados todos os detalhes a respeito desta técnica.

### **2.5.4 Íris**

Um processo caótico rege a formação do padrão da íris no indivíduo durante seu processo de formação embrionária. Este padrão é admitido como sendo singular para cada indivíduo, mesmo em gêmeos idênticos [Daugman, 1995], sendo altamente estável ao longo da vida do indivíduo.

A medição do padrão da íris fornece alta taxa de certeza de reconhecimento, contudo por ser a imagem do padrão da íris muito pequena, requer-se um sistema de aquisição com lentes de alta precisão de foco e alta resolução, o que torna o sistema óptico complexo e caro. Este processo torna-se assim muito sensível ao movimento do olho uma vez que se nestes sistemas tem-se uma estreita profundidade focal, o que compromete a qualidade na aquisição da imagem.

### **2.5.5 Retina**

Leitores de retina verificam o padrão de vascularização no fundo do olho. Isto requer que o indivíduo olhe fixo para dentro de uma lente. Enquanto isto, um feixe de luz é introduzido no olho até atingir a retina para que depois o padrão de vascularização seja adquirido. Embora esta técnica produza alta taxa de certeza de identificação [Jain et al., 1999], a maioria dos indivíduos tem sérias restrições quanto à projeção de um feixe de luz dentro de seus olhos. Por

outro lado o fato de que a retina possa refletir o estado de saúde do indivíduo aumenta a incerteza na identificação.

### **2.5.6 Vascularização**

A tecnologia de padrão vascular é semelhante a aquela do padrão vascular da retina, na qual o mapa de veias de mãos, pulsos e faces são adquiridos com o uso de luz infravermelha. Existe rejeição destas técnicas por indivíduos que se opõem ao uso de luz infravermelha sobre seu corpo.

### **2.5.7 Voz**

A fala é produzida através de pulsos glotais que incidem sobre as cordas vocais. Os pulsos glotais variam muito rapidamente e por isto é dito terem componentes de alta frequência. Por outro lado as cordas vocais variam muito lentamente e a elas é atribuída a baixa frequência da voz. Quando estas variações são capturadas, estas características são usadas como grandezas biométricas fisiológicas e comportamentais. Também é sabido que a voz não possui suficiente singularidade [Poh e Korezak, 2001] que permita o reconhecimento de indivíduos em uma grande base de dados. Em algumas situações o alto nível de ruído presente no ambiente dificulta o processamento da voz, que também costuma ser prejudicado por problemas na garganta como p. ex. rouquidão.

Estes sistemas são facilmente fraudados pois gravações com a voz de um determinado indivíduo pode enganar o sistema.

Apesar destes problemas a voz é hoje largamente empregada como grandeza biométrica para reconhecimento. Os seres humanos utilizam diariamente esta técnica, p.ex., ao falar pelo telefone. Além disto o custo do sistema de aquisição do sinal de voz é bastante baixo, o que lhe confere um futuro promissor.

### **2.5.8 Assinatura**

A assinatura de cada indivíduo é bastante singular. Embora até esta data seja largamente utilizada pela sociedade para reconhecimentos em documentos legais, possui grandes problemas quanto à segurança contra fraudes. Também é conhecido que os indivíduos mudam sua assinatura ao longo do tempo ou de acordo com sua disposição ou humor, dificultando a tarefa de reconhecimento dos sistemas automáticos. Devido a estes fatores, aliados ainda a possibilidade do uso de assinaturas falsas, a confiabilidade destes sistemas em geral é baixa.

### 2.5.9 DNA

O DNA é uma técnica que está sendo largamente utilizada na identificação de indivíduos. Os recentes avanços tecnológicos proporcionam cada vez maior rapidez neste tipo de procedimento e conseqüente barateamento de custos, viabilizando sua utilização. Suas aplicações vão desde a investigação de paternidade até a identificação de criminosos na ciência forense. Por estes motivos, apesar do tempo de resposta e os custos deste exame estarem diminuindo, esta técnica ainda não está disponível para uso em controle de acesso e ou informações.

### 2.5.10 Outras Técnicas

- i. Odor corporal
- ii. Dobras da junta dos dedos
- iii. Ritmo de datilografia

## 2.6 Comparação das Técnicas

Todas as técnicas biométricas possuem vantagens e desvantagens intrínsecas, porém em termos gerais pode-se afirmar que terá maior potencial de sucesso a técnica que reunir a maior quantidade de características desejáveis, como as listadas na seção 2.3.

A tabela 2.2 mostra uma comparação entre algumas das principais técnicas biométricas classificadas segundo a Taxa de Mesmo Erro - TME.

Quanto menor a TME mais singular é a grandeza biométrica. Apesar dos padrões de retina e iris fornecerem as menores taxas de erro, estas técnicas possuem as mais baixas taxas de aceitabilidade por parte dos usuários. A geometria da mão e a impressão digital possuem a mesma TME, enquanto as grandezas comportamentais como voz e assinatura possuem as mais baixas taxas de TME. Finalmente, a técnica da impressão digital, não obstante sua freqüente associação com assuntos relacionados a crime, possui uma boa aceitação entre os usuários.

Além da similaridade com a técnica que os próprios seres humanos utilizam para reconhecerem a si mesmos, o reconhecimento de faces possui uma das maiores aceitabilidades dentre as técnicas de reconhecimento. Enquanto outras técnicas requerem uma ação ativa do usuário, no sentido de oferecer a grandeza biométrica ao sistema de medição, o reconhecimento de faces pode ser usado de forma passiva exigindo apenas que a imagem da face seja adquirida. Esta

vantagem amplia o horizonte de aplicações uma vez que ela pode ser usada para monitoramento policial p.ex., sem que o indivíduo fique sabendo que esta sendo monitorada e/ou identificada.

Tabela 2.2 Comparação entre técnicas biométricas [Ruggles, 98 A.D.]

Técnicas Biométricas - TB -	Taxa de Mesmo Erro - TME -
Retina	1:10.000.000
Íris	1:131.000
Impressão Digital	1:500
Geometria da Mão	1:500
Assinatura	1:50
Voz	1:50
Face e Vascularização	Sem dados <sup>7</sup>

O reconhecimento de faces permite que indivíduos possam fazer visualmente a tarefa de reconhecimento em imagens pois esta é a sua maneira natural de reconhecer. Com respeito à aquisição dos dados, é mais fácil adquirir uma boa imagem de uma face do que de uma impressão digital (excluindo ainda aqueles 5% de indivíduos que não conseguem fornecer uma impressão digital suficientemente clara para uso em reconhecimento). As razões para esta dificuldade incluem cortes na pele, dedos com bandagens, calos, pele muito seca e/ou úmida, doença de pele, sensores sujos e/ou manchados. Desvantagens similares se aplicam para a identificação baseada na geometria da mão. O uso destes leitores de impressão digital e/ou de mãos pode transmitir germes devido ao contacto físico exigido. Em contrapartida, um sistema de reconhecimento de faces não requer limpeza de sua óptica, uma vez que a leitura é feita sem contato físico entre o equipamento e o indivíduo.

Outras técnicas como a do reconhecimento da íris e o da retina falham em indivíduos que usam lentes de contacto colorida pois estas podem mascarar a íris e a retina. Outro problema é a exigência de que o usuário posicione o olho de forma precisa e estática durante o processo de aquisição, o que não é tolerado por certos usuários. Alguns indivíduos têm sérias restrições a

<sup>7</sup> Dados não disponíveis por se tratar de tecnologias recentes [Ruggles, 98 A.D.] .



submeter-se a sistemas de aquisição que introduza luz dentro de seu olho. Reflexos de óculos também podem impedir a aquisição da grandeza.

Por outro lado, sistemas de reconhecimento de faces requerem somente uma câmera de baixo custo e não necessitam de precisão no posicionamento pois seu campo focal é suficientemente grande para cobrir uma ampla faixa de aquisição. Bons algoritmos de reconhecimento possuem bom desempenho mesmo com reflexos de óculos, olhos fechados e/ou expressão faciais diferentes.

## **2.7 Aplicabilidade Quanto à Curva de Performance**

Através da figura 2.5 pode-se também analisar qual a CP do sistema é mais adequada para determinadas aplicações. Um sistema que possui alta TFR é considerado um sistema recomendado para aplicações em alta segurança, visto que, apesar deste rejeitar algumas ovelhas este minimiza a aceitação de lobos. Um exemplo típico da aplicação de um sistema com este tipo de comportamento é a proteção de informações sigilosas e/ou valores onde se admite um certo grau de rejeição de indivíduos capacitados mas o aceite de um indivíduo não capacitado é falta muito grave. Ao se utilizar um limiar que aproxime os valores das TFA e TFR como a TME, se está adequando o sistema para aplicações do tipo gerais. Por outro lado, ao se buscar aumentar a TFA tentando a minimização de ocorrências de falsas rejeições, se está concomitantemente diminuindo a TFR, ou seja, aumentando a tolerância do sistema para a ocorrência de falsas aceitações. Como exemplo de aplicação para este tipo de sistema podem-se citar as áreas de controle policial, onde não é importante a retenção de alguns indivíduos inocentes, mas é falta grave a dispensa de indivíduos procurados.

As TFA e TFR e suas inter-relações são intrínsecas da concepção de cada sistema e também função do limiar de classificação. Porém pode-se melhorar o desempenho destes através de intervenções no âmbito do projeto buscando obter CP's que produzam áreas menores abaixo da curva (figura 2.5).

## **2.8 A Importância do Protocolo de Ensaio**

Um protocolo de avaliação determina como ensaiar um sistema, desde os aspectos relativos à seleção dos dados até a medição do desempenho e interpretação e publicação dos resultados. Fabricantes e pesquisadores de sistemas biométricos utilizam as TFR e TFA para qualificarem

---

os seus sistemas biométricos. Grupos independentes como "International Biometric Industry Association - IBIA"<sup>8</sup> e "Biometric Consortium - BC"<sup>9</sup> estão trabalhando neste sentido. Apesar disto, ainda não existe um bom protocolo de ensaios, comum a todos, permitindo com que os resultados alcançados individualmente sejam de difícil comparação.

---

<sup>8</sup> "IBIA - International Biometric Industry Association", defende os interesses da indústria biométrica.

<sup>9</sup> "BC - Biometric Consortium", foi criado pelo governo Americano, apoiado pelo NIST, representa seus interesses em pesquisa, desenvolvimento, ensaios, aplicações e regulamentações do setor.

## Capítulo 3

# Reconhecimento da Geometria Facial

---

*Minha face é meu passaporte.*

- Vladimir Horowitz, (1903-1989).

O reconhecimento da geometria facial representa a principal contribuição deste trabalho. Neste capítulo será apresentado um resumo das principais técnicas de reconhecimento de padrões aplicáveis ao reconhecimento de geometrias (formas) e das técnicas de reconhecimento de faces. O reconhecimento da geometria da face é um caso particular de reconhecimento de padrões onde muitas das considerações pertinentes a este assunto são válidas e se aplicam diretamente ao assunto deste trabalho.

### 3.1 Reconhecimento de Padrões

Genericamente, reconhecimento de padrões é a ciência que se preocupa com a descrição e classificação (reconhecimento) de medidas [Schalkoff, 1992]. O reconhecimento de padrões é um importante campo de aplicação de sistemas inteligentes, pois necessita constantemente tomar decisões.

O reconhecimento de padrões pode ser definido como:

- i. Segundo Ross [Ross, 1995], é um processo de identificar estruturas em dados utilizando-se de comparações com estruturas amplamente conhecidas. Estas estruturas são criadas através de métodos de classificação.
- ii. Segundo Klir [Klir e Yuang, 1995], o reconhecimento de padrões pode ser definido como um processo pelo qual buscam-se estruturas nos dados e classificam-se estas estruturas dentro de categorias tais que o grau de associação é maior entre as estruturas da mesma categoria e menor entre as categorias de estruturas diferentes. As categorias relevantes são usualmente caracterizadas por estruturas prototípicas derivadas da experiência do passado. Cada categoria pode ser caracterizada por mais de uma estrutura prototípica. A estrutura típica de um sistema de reconhecimento de padrões consiste de um sensor, um mecanismo de extração de características que pode ser um algoritmo ou não, e um algoritmo de descrição ou classificação dependendo da aplicação.
- iii. De acordo com Schalkoff [Schalkoff, 1992], o reconhecimento de padrões pode ser visto como o particionamento de dados, representados por pontos em um espaço multidimensional, em regiões indicativas das classes.

O reconhecimento de padrões pode ser caracterizado como redução da informação ou mapeamento da informação. Uma visão abstrata do problema de classificação e descrição pode ser vista na figura 3.1, onde cada classe  $w_{i,j,k}$  gera um subespaço  $P_{1,2,3,4}$  no espaço de padrões "P", através de uma relação  $G_{i,j,k}$ , que pode ser estatística p. ex.. Nota-se que podem haver sobreposição de subespaços. Outra relação "M" estabelece o vínculo entre o espaço de padrões "P" e o espaço de medidas "F". Usando deste conceito, a solução dos problemas de reconhecimento de padrões seria simplesmente encontrar as relações "M" e "G", para cada medição  $m_{i,j,k}$ . Infelizmente isto não é tão simples, pois na prática estes mapeamentos inversos não são funções.

A figura 3.1 mostra que idênticas medições  $m_i$  foram resultados de diferentes "p", que por sua vez correspondem a diferentes classes "w". Isto mostra um dos problemas potenciais do reconhecimento de padrões que é a ambigüidade.

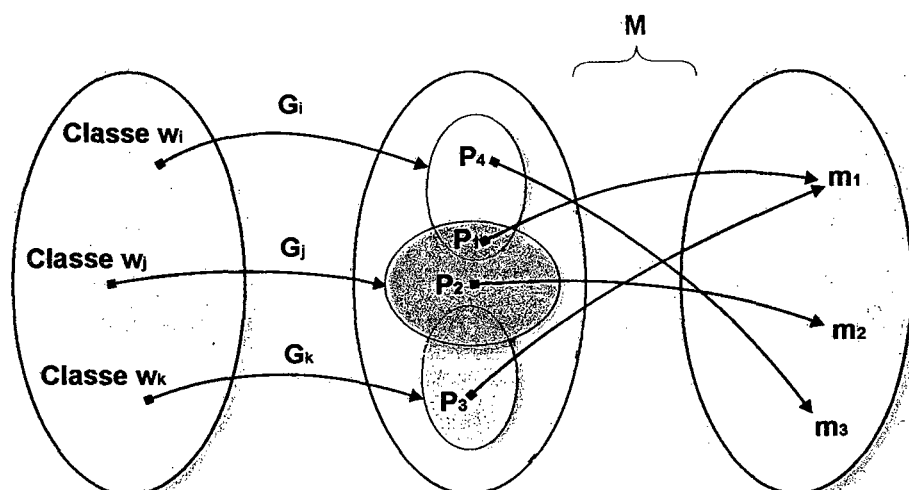


Figura 3.1 Representação de sistemas e classificação e interpretação [Schalkoff, 1992].

### 3.1.1 Abordagens em Reconhecimento de Padrões

Historicamente as duas metodologias mais importantes em reconhecimento de padrões foram a estatística e a sintática. Recentemente tecnologias emergentes como a das redes neurais e conjuntos nebulosos vem se constituindo em mais um paradigma de solução de problemas em reconhecimento de padrões [Raittz, 1997].

#### Abordagem Estatística

Os métodos estatísticos são os mais utilizados em sistemas de reconhecimento devido a sua simplicidade de implementação, por serem os métodos matemáticos bem estabelecidos e por possuírem uma ampla história constituída.

Segundo [Bishop, 1995], a forma mais geral e natural de formular soluções para o reconhecimento de padrões é o RP estatístico, através do qual é reconhecida a natureza estatística tanto da informação que se quer representar quanto dos resultados que devem ser expressos.

Robert Schalkoff [Schalkoff, 1992], diz que o reconhecimento de padrões estatístico, como sugere o próprio nome, assume uma base estatística para os algoritmos de classificação. De um conjunto de medidas, que denotam as características, é extraído dos dados de entrada e usados para associar cada vetor de características a uma das  $k$  classes. Presume-se que as características são geradas por um estado da natureza, portanto existe um modelo subordinado

a um estado com um conjunto de probabilidades e ou funções de densidade de probabilidade correspondente, passível de ser usado para representá-lo.

Funções de distância são modelos estatísticos muito usados para medir o grau de proximidade entre duas entidades. Dentre eles podemos citar como mais importantes a Distância de Hamming<sup>1</sup>, Euclidiana, de Mahalanobis, medida de Tanimoto e funções de similaridade como o classificador de Bayes [Barreto, 1999], [Shridhar e Badreldin, 1999], [Schalkoff, 1992], [Machado, 1994] [Zimmermann et al., 2000b] e [Mitchell, 1997].

### **Abordagem Sintática**

Envolvem a construção de uma linguagem formal, que representará o conjunto de objetos pertencentes a uma determinada classe, através da análise de suas características sintáticas. Os métodos sintáticos aparecem geralmente associados a outras metodologias e são indicados para problemas onde a análise da estrutura do objeto é fundamental no processo de reconhecimento [Barreto, 1999]. Algumas vezes a informação disponível não se resume somente na presença ou ausência de valores numéricos de um conjunto de características mas a interrelação ou interconexão das características também produz informações estruturais importantes, que facilitam a descrição ou classificação. Assim funciona o reconhecimento de padrões sintático. Este paradigma geralmente formula uma descrição hierárquica de padrões complexos, a partir de padrões mais simples. Estes elementos, extraídos dos dados de entrada se encontram no nível mais baixo onde são também conhecidos como primitivas.

O reconhecimento de objetos que envolvem a análise de sua estrutura requer geralmente a aplicação de métodos sintáticos. Um exemplo de problema resolvido pelo método sintático é a identificação de impressões digitais [Machado, 1994], cuja disposição de elementos segue padrões rígidos. No campo de reconhecimento de objetos tridimensionais, Lin & Fu [Lin e Fu, 1986] apresentam uma solução através da construção de gramáticas. Reconhecedores de caracteres manuscritos também utilizam métodos sintáticos, como o apresentado por Ward [Ward e Kuklenski, 1989]. Problemas de interpretação de chapas de raio-x podem ser resolvidos através da associação de métodos sintáticos da Teoria de Lógica Indivíduo [Pathak e Pal, 1986].

---

<sup>1</sup> Algumas vezes referida como distância de Manhattan.

### **Abordagem Heurística**

Métodos heurísticos possibilitam a consideração de conhecimento e experiência prévia de especialistas. Sistemas baseados em regras são utilizados em problemas de interpretação de ultrasonografias médicas, desenhos de formulação química e extração de redes rodoviárias em imagens de satélite. Árvores de decisão constituem um método simples, de baixo custo computacional. Kruger [Kruger e Thompson, 1981] as utiliza para identificar objetos a serem montados, em uma linha de montagem automatizada.

### **Abordagem Conexionalista**

A utilização da abordagem conexionalista ou por redes neurais no reconhecimento de padrões vem crescendo bastante nos últimos anos. A comparação das redes artificiais com o modelo biológico do sistema de visão de seres vivos é inevitável e fundamentou os primeiros trabalhos neste campo [Marko, 1974]. Fukushima [Fukushima, 1969] apresenta uma discussão a respeito dos diversos modelos de redes neurais e sua aplicação em problemas de reconhecimento de padrões. Trabalhos como o de [Bishop, 1995], defendem a idéia de que o reconhecimento de padrões através do uso de redes neurais é uma extensão aperfeiçoada do reconhecimento de padrões estatístico convencional. Com efeito, tradicionalmente estatística se fundamenta na teoria das probabilidades mas também se pode falar em estatística baseada em redes neurais, cada uma conservando sua identidade. Cichocki e Unbenhabem [Cichocki e Unbenhabem, 1993] definem a rede neural como um sistema de processamento de sinal ou informação composto por um grande número de elementos simples de processamento chamados neurônios artificiais ou simplesmente nós. Estes são interconectados por elos diretos chamados de conexões, que colaboram para realizar um processamento paralelo distribuído na realização da tarefa computacional desejada. O treinamento da rede é realizado através da atualização dos pesos sinápticos nas camadas intermediárias. Assim um novo padrão ao ser apresentado à rede fará com que esta indique através da camada de saída a classe que o representa.

Em 1989, Pomerleau construiu uma rede neural para controlar a direção de um carro, conseguindo resultados interessantes, embora de pouca aplicabilidade prática. No mesmo ano, o correio dos Estados Unidos (USPS<sup>2</sup>) desenvolveu um sistema de reconhecimento de códigos postal manuscritos, com alto índice de acertos.

---

<sup>2</sup> USPS - United States Postal Service.

### **Abordagem por Conjuntos Nebulosos**

O emprego de conjuntos nebulosos para o reconhecimento de padrões já foi usado com sucesso em meados da década de 60. Atualmente a literatura a respeito de reconhecimento de padrões com base em conjuntos nebulosos é bastante extensa [Klir e Yuang, 1995]. Existem duas formas clássicas de reconhecimento nebuloso de padrões: (i) os métodos de reconhecimento de padrões através de listas de pertinências e (ii) os métodos nebulosos sintáticos.

No método nebuloso sintático cada padrão é representado por uma cadeia de subpadrões concatenados chamados de primitivas. Estas primitivas são vistas como o alfabeto de uma linguagem formal. Um padrão então é uma sentença gerada por alguma gramática. Todos os padrões, cujas sentenças são geradas pela mesma gramática, pertencem a uma mesma classe. Um padrão desconhecido então é classificado como pertencente a uma classe particular, se ele puder ser gerado pela gramática correspondente a ela.

### **3.1.2 Aplicações do Reconhecimento de Padrões**

Reconhecimento de padrões engloba uma ampla faixa de problemas de processamento de informação, que vão desde o reconhecimento de voz, faces humanas e caracteres, até a detecção de erros em equipamentos ou diagnósticos médicos.

Indivíduos no dia a dia aplicam seus próprios métodos de reconhecimento de padrões em praticamente todas as áreas da atividade humana. A matemática pura não é suficiente para o bom desempenho na resolução da maioria dos problemas. Tome-se como exemplo o jogador de xadrez onde a análise de todas as possibilidades em uma dada posição complexa é virtualmente impossível. No entanto se este for hábil, desenvolverá suas próprias técnicas para identificar os padrões das várias posições com base em sua experiência e usará tais técnicas para obter vantagem no jogo.

As principais aplicações das técnicas de reconhecimento de padrões são:

- i. Reconhecimento de objetos
- ii. Reconhecimento de voz
- iii. Reconhecimento de faces
- iv. Tratamento de impressões digitais
- v. Reconhecimento de caracteres
- vi. Análise de eletrocardiogramas



- vii. Diagnósticos médicos

## 3.2 Reconhecimento de Faces

O reconhecimento de faces é uma das tarefas que o homem aprendeu a realizar com perfeição. O avanço da tecnologia de reconhecimento de padrões, visão computacional e inteligência artificial contribuem com elementos tecnológicos que possibilitam a automação deste processo com bons resultados. Humanos podem detectar faces em cenários complexos com pouco ou nenhum esforço [Hancock et al., 1995]. Esta capacidade é bastante robusta, mesmo sob grandes variações no estímulo visual devido a condições adversas de iluminação e de visualização, expressões faciais, idade, o uso de acessórios como óculos e mudanças no estilo dos cabelos. Entender o mecanismo de visão humana usada para reconhecer faces e produzir artificialmente um sistema que execute esta tarefa é difícil e complexo.

O problema global de reconhecimento de faces pode ser dividido em:

- i. Detecção de uma face padrão em uma multidão
- ii. Detecção de pontos de referência em faces
- iii. Reconhecimento de faces
- iv. Análise de expressões faciais [Samal e Iyengar, 1992]

Ao longo de aproximadamente três décadas cientistas das áreas de neurologia, psicologia, visão computacional, processamento de imagens e computação vem estudando as técnicas de reconhecimento de faces usadas por humanos e máquinas [Brunelli e Poggio, 1993], [Chellappa et al., 1995].

O reconhecimento de faces por máquinas é um assunto emergente e ativo que emprega muitas áreas de conhecimento como visão computacional, processamento de imagens, reconhecimento de padrões e inteligência artificial. Os esforços empregados nos últimos dez anos, principalmente por pesquisadores da área de visão computacional, têm gerado resultados animadores, propiciados principalmente pelo desenvolvimento de: (i) novos computadores com maior poder computacional, (ii) de técnicas e equipamentos para processamento de imagens, através principalmente de novos sensores CCD, com maior velocidade e resolução, (iii) do desenvolvimento da inteligência artificial - IA como alternativa vantajosa sobre as técnicas convencionais de classificação.

### O Primeiro Sistema Reconhecedor de Indivíduos

Apesar do reconhecimento automático de faces ter despertado um interesse maior dos cientistas e usuários somente a partir da década de 60, os primeiros trabalhos em reconhecimento de faces datam do fim do século XIX. Segundo reportado por Benson e Perret [Benson e Perrett, 1992], Sir Francis Galton [Galton, 1888] um cientista inglês, primo de Charles Darwin, em seu discurso sobre identificação pessoal no "Royal Institution" em 25 de maio de 1888, explicou que ele freqüentemente se decepcionava com a sua incapacidade de expressar verbalmente semelhanças hereditárias e tipos de traços em indivíduos. Assim para aliviar a si mesmo deste embaraço, ele desenvolveu uma técnica de reconhecimento e a aplicava no reconhecimento de prisioneiros franceses. A técnica consistia em medir quatro parâmetros biométricos principais: (i) comprimento da cabeça, (ii) largura de cabeça, (iii) comprimento do pé, (iv) comprimento do dedo médio da mão. Ele então as classificava em três possíveis valores, (grande, médio e pequeno) dando um total de 81 possíveis classes primárias. Galton sentiu que seria vantajoso se tivesse um método automático para classificação destes. Com este propósito teria desenvolvido um mecanismo, que ele mesmo chamou de seletor mecânico, o qual poderia ser usado para a medição de perfis faciais. Galton reportou que, uma vez escolhendo-se bem as medidas do perfil, a maioria das medições que ele havia tentado no processo de reconhecimento dos prisioneiros foram muito eficientes.

### **3.3 Sistemas de Reconhecimento de Faces**

As mais diferentes técnicas para a abordagem do problema vêm sendo propostas nos últimos anos [Brunelli e Poggio, 1993], [Robertson e Craw, 1993], [Huang, 1998]. No entanto, todas podem ser classificadas basicamente em dois grandes grupos:

- i. Abordagem abstrativa ou de características geométricas locais<sup>3</sup>.
- ii. Abordagem holística ou de modelos de representações globais<sup>4</sup>.

#### **3.3.1 Métodos Baseados em Características Geométricas Locais.**

Chamado de método abstrativo ou de características geométricas locais, esta técnica extrai índices geométricos das características e/ou traços das faces. Neste paradigma o processamento das informações é geralmente feito por técnicas de reconhecimento de padrões do tipo estatísticas. Os pontos de referência escolhidos na face são geralmente os de mais fácil

---

<sup>3</sup>"Geometric feature-based matching".

<sup>4</sup>"Template matching".

identificação como centro dos olhos, pico do nariz e curvatura da boca. O reconhecimento é feito comparando-se índices geométricos, geralmente distâncias, entre os pontos de referência da face sob teste com os índices das faces da base de dados. A idéia introduzida por Galton [Galton, 1888], de utilizar mais de uma medida para classificar o indivíduo, ao invés de um só índice, tem sido atualmente utilizada em diversos trabalhos.

A vantagem deste tipo de técnica é sua invariância com as condições de iluminação. Um ponto de referência geométrico não muda se a imagem for clara ou escura. A maior desvantagem é a baixa tolerância a expressões faciais que mudam os pontos de referência. Além disto, reconhecer automaticamente os pontos de referência em uma face não é uma tarefa muitas vezes tão simples. Muitos autores têm construído sistemas com este princípio, incluindo Kanade [Kanade, 1997], Craw, Tock et al. [Craw et al., 1992] e Robertson [Robertson e Sharman, 1992]. Todos eles utilizam imagens 2D com níveis de cinza e usam técnicas como análise de cantos<sup>5</sup> para processar a imagem e estabelecer os pontos de referência.

### Traços de perfil

Harmon [Harmon et al., 1981] propôs uma abordagem baseada no traço de perfil de faces humanas. Os traços são calculados a partir de algumas marcas colocadas ao longo do perfil da face. Iniciando com uma base de dados de 112 elementos, três faces de treinamento e uma imagem de teste, eles representaram a face usando um vetor de 17 traços/elementos.

Najman [Najman et al., 1993] também utilizaram traços geométricos de perfis. A linha de perfil foi construída usando de 8 a 100 pontos de controle. Testes foram feitos em uma base de dados de 10 elementos com 31 imagens de treinamento e 10 imagens de teste para cada elemento. Três métodos de classificação foram tentados: (i) Análise de Componentes Principais - ACP<sup>6</sup>, seguida de discriminador quadrático, (ii) k-vizinho mais próximo - k-VP, (iii) "Gradient Back Propagation - GBP".

Wu e Huang [Wu e Huang, 1990] usaram traços geométricos de perfil no trabalho, com seis pontos de controle, obtendo uma B-spline cúbica. A base de dados continha 18 elementos com três imagens de treinamento e uma imagem de teste para cada elemento.

---

<sup>5</sup> "Edge".

<sup>6</sup> "Principal Components Analysis - PCA".

### **Traços frontais**

Em uma das primeiras tentativas de identificação automática de faces, Kanade [Kanade, 1997] projetou um sistema que extraía 16 traços faciais frontais, mais tarde reduzidos para 13. Foram usadas uma base de dados de 20 elementos com uma imagem de treinamento e uma imagem de teste para cada elemento.

Brunelli e Poggio [Brunelli e Poggio, 1993] implementaram recentemente um identificador baseado em traços geométricos levemente baseado no trabalho de Kanade. O identificador foi testado numa grande base de dados de 47 elementos.

Outra abordagem baseada em perfis de traços faciais foi implementada por Wong [Wong et al., 1989] usando várias distâncias como Índice geométrico: (i) entre olhos, (ii) entre nariz e olho esquerdo, (iii) entre nariz e olho direito.

### **Traços frontais e de perfil**

Goldstein [Goldstein et al., 1971] em seus primeiros trabalhos usaram um conjunto de 34 traços frontais e de perfis (mais tarde reduzido para 22). As características incluíam, entre outras, o comprimento e textura dos cabelos, comprimento do nariz, largura da boca e perfil do queixo. As características eram classificadas em uma faixa de um a cinco e eram locadas manualmente. O modelo mostrava que para uma população de 255 elementos, seis características eram suficientes para o reconhecimento.

## **3.3.2 Métodos Baseados em Modelos de Representações Globais**

Também chamado de holístico, este método é baseado em modelos que usam representações globais da imagem da face. Ao contrário do método apresentado no item 3.3.1 que extrai determinadas características geométricas, nesta abordagem a face é vista como um todo. Alguns algoritmos de detecção de face são requeridos quando esta faz parte de um contexto complexo da imagem adquirida.

O processamento das informações neste caso pode ser realizado através das seguintes abordagens: (i) conexionista como a de redes neurais [Poggio e Giroi, 1999], (ii) ACP [Penev e Atick, 1999], (iii) Decomposição em Valores Singulares - DVS<sup>7</sup> usando "FACES Próprias"<sup>8</sup> [Turk e Pentland, 1991], (iv) um modelo deformável como descrito por Yuille et al. [Yuille et al., 1992] e

---

<sup>7</sup> "SVD - Singular Value Decomposition".

<sup>8</sup> "Eigenfaces".

[Lengagne et al., 1996], (v) Modelo de Contorno Ativo - MCA<sup>9</sup> como reportado por Huang e Chen [Huang e Chen, 1992] e originalmente descrito por Kass et al. [Kass et al., 1987].

### **Análise de Componentes Principais - ACP**

A mais simples versão de reconhecimento baseado em representações globais é obtida quando toda a face é usada como um simples modelo. A imagem é classificada através do cálculo das distâncias (distância de Hamming, p. ex.) entre os modelos gerados pelas imagens no conjunto de treinamento escolhendo-se a que mais se adapta (menor valor da distância).

O procedimento de Karhunen-Loève de Kirby e Sirovich [Kirby e Sirovich, 1990] e a abordagem através da Análise de Componentes Principais de Turk e Pentland [Turk e Pentland, 1991] são baseados neste método simples de verificação. A matriz e o modelo, contudo, não são as imagens originais mas sim sua projeção em um sistema de coordenadas otimizadas. Um conjunto de imagens com diferentes condições de iluminação, poses e caricaturas de cada elemento é transformado em uma única matriz de pontos chamada de "Faces Próprias"<sup>10</sup>. Turk e Pentland usaram uma base de faces com 16 elementos.

### **Mapas de Linhas de Mesma Densidade**

Uma abordagem diferente para o sistema baseado em modelos foi proposto por Nakamura [Nakamura et al., 1991]. A técnica que eles apresentaram fez uso de mapas de níveis de cinza com curvas de mesma intensidade representando as imagens de faces. Em resumo, usando suas próprias palavras, "se o brilho de uma imagem é tomado como as alturas de uma montanha o mapa de mesma intensidade corresponde à curvas de mesma altitude". Uma base de dados de 10 elementos com um elemento de treinamento e um elemento de teste foi usada. Três elementos usavam óculos, dois elementos usavam barba fina e duas mulheres tinham diferentes maquilagens e estilos de cabelo no conjunto de imagens de treinamento e de testes. Foram reportadas taxas de reconhecimento perfeitas.

### **Métodos de Correlação de Múltiplos Modelos**

Um dos primeiros estudos baseados em representação de múltiplos modelos foi apresentado por Baron. Uma base de dados de 42 elementos foi usada e cada uma foi representada por até cinco características selecionadas manualmente (toda face, olho direito, queixo e cabelo) e

---

<sup>9</sup> "Snake".

<sup>10</sup> "Eigenface".

cada característica contendo até quatro modelos distintos. Um total de 20 modelos (matriz de pontos) foi arquivado para cada elemento, com cada modelo sendo uma matriz de pontos de 15x15. A imagem para teste foi em primeiro lugar reduzida a uma matriz de 15x16 de toda a face e depois comparada com cada elemento (toda face) do conjunto de treinamento. Se o valor da correlação entre a imagem de teste reduzida estivesse dentro de certos valores aceitáveis, as outras características faciais eram chamadas e usadas para reconhecimento. Se pelo menos três de quatro características tivessem seus valores de correlação no nível de aceite, a imagem sob teste era reconhecida

Mais recentemente Brunelli e Poggio [Brunelli e Poggio, 1993] apresentaram resultados baseados uma abordagem semelhante. Usaram uma base de dados de 47 elementos, onde cada elemento era representado por uma imagem frontal total da face e um conjunto de quatro modelos (olhos, nariz, boca e toda a face). O reconhecimento da face sob teste era feito pela computação da correlação cruzada de cada modelo.

#### **Modelos de Vetores Quantizados**

Ramsay [Ramsay et al., 1992] usou uma abordagem baseada em modelos onde cada uma das oito características que eles selecionaram foram substituídas com um modelo similar desenhado de um livro de códigos através de quantização de vetores. Vários algoritmos podem ser usados para gerarem livros de códigos úteis. Dois algoritmos foram desenvolvidos por Ramsay [Ramsay et al., 1992]. Usando uma base de dados de 30 elementos com 10 elementos de treinamento e 10 imagens de teste para cada elemento.

#### **Verificação de Modelos Baseados em Redes Neurais**

Modelos e/ou padrões também têm sido usados como entradas para sistemas baseados em redes neurais. Allison [Allison et al., 1992] usou uma imagem de 32x32 pontos de toda face e duas 64x32 das regiões dos olhos e boca. Estes modelos foram usados como entradas de um mapa de característica auto-organizadas de Kohonen [Kohonen.T., 1988]. O mapa produzia uma topologia que preservava a estrutura da entrada do modelo. Os mapas eram usados como entradas de um perceptron multicamadas que processava a classificação.

Outro trabalho feito por Cottrell e Fleming [Cottrell e Fleming, 1990] estudou a performance de uma rede que automaticamente extraía as características de um modelo de 64x64 pontos (camada intermediária) e colocava este resultado em uma rede de uma camada para identificação e classificação por sexo. Imagens de teste foram perfeitamente identificadas com uma base de dados de 11 elementos.

Stonham [Stonham, 1986] detalhou experimentos de reconhecimento de faces usando uma máquina de reconhecimento de padrões de uso geral chamada de WISARD<sup>11</sup>. Uma base de dados de 16 elementos foi usada e uma imagem completa de 153x214 pontos foi introduzida em uma rede de uma só camada auto-adaptativa. O tempo de aquisição dos dados era de aproximadamente 20 segundos. Uma média de 200 a 400 imagens era adquirida para completar o treinamento. Resultados de testes em tempo real foram reportados como sem erros.

### 3.3.3 Métodos de Modelos Híbridos

Craw e Cameron [Craw e Cameron, 1992], [Craw, 1992], propuseram uma abordagem híbrida baseada parcialmente na verificação de modelos e de características geométricas. Uma face foi modelada usando uma máscara com 59 pontos de controle e cada face foi descrita por dois vetores:

- i. Vetor de forma, contendo a informação sobre a geometria dos traços da face.
- ii. Vetor de textura, contendo os níveis de cinza da textura da face.

Altos sucessos de reconhecimento foram relatados mesmo quando existiam significantes diferenças entre imagens de teste e de treinamento.

Lanitis [Lanitis et al., 1994] também usaram uma combinação de forma e níveis de cinza para codificar a aparência de faces humanas. Eles usaram uma base de dados de 30 elementos com 10 imagens de treinamento e 10 imagens de teste para cada elemento. As faces foram modeladas usando-se 3 métodos:

- i. Uma forma flexível baseada num modelo de pontos de distribuição (uma introdução a esta técnica pode ser encontrada em Cootes [Cootes et al., 1992]). Este modelo captura variações de forma e pode ser usado para localização de faces em uma imagem.
- ii. Um modelo de níveis de cinza livre de forma, obtido pela deformação e alinhamento de cada face de treinamento em relação à face média.
- iii. Um perfil de níveis de cinza local, consistindo de um grande número de perfis locais obtidos ao longo da perpendicular à fronteira da forma em cada modelo de forma de pontos.

---

<sup>11</sup> WISARD - Wilkie, Stonham and Aleksander's Recognition Device.

### 3.3.4 Os Resultados dos Métodos de Reconhecimento

Comparar e interpretar os resultados de diferentes métodos de reconhecimento é uma tarefa complexa, pois a não adoção de um protocolo de ensaios em comum, faz com que cada grupo de trabalho publique suas conclusões baseados em preceitos próprios o que na maioria das vezes dificulta a intercomparação dos resultados. Os exemplos das técnicas descritas nos itens precedentes são um bom exemplo de como cada autor publica os resultados do desempenho dos seus sistemas. Na tabela 3.1 pode-se ver o resumo acerca do desempenho dos diferentes métodos aqui descritos, inclusive com os dados da performance destes, publicados pelos autores.

A insuficiência de dados pode ser vista quando o autor fornece p.ex. a performance do sistema mas não diz com que conjunto de treinamento e validação estes resultados foram alcançados.

Note que mais uma vez, a comparação direta dos resultados da tabela 3.1 fica prejudicada pois os dos conjuntos de indivíduos "S", "I", "J" e "K" são diferentes e não existe especificação clara do desempenho através das TFA's, TFR's e TME's.

Na literatura, a comparação mais direta entre dois métodos foi feita por Brunelli e Poggio [Brunelli e Poggio, 1993], que usando uma mesma base de dados, implementaram o método baseado em características geométricas locais e em modelos de representações globais.

Tabela 3.1 Desempenho das técnicas de reconhecimento de faces existentes.

Tipo	Pesquisador(es)	S x i	j	k	r
Ab	Harmon et al. [Harmon et al., 1981]	112	3	1	96%
Ab	Najman et al [Najman et al., 1993]	10	31	10	90%
Ab	Wu and Wuang [Wu e Huang, 1990]	18	3	1	100%
Ab	Kanade [Kanade, 1997]	20	1	1	75%
Ab	Nakamura et al. [Nakamura et al., 1991]	10	1	1	100%
Ab	Sutherland et al. [Ramsay et al., 1992]	30	10	10	89%
Ab	Lanitis et al. [Lanitis et al., 1994]	30	10	10	92%
Ab	Brunelli e Poggio [Brunelli e Poggio, 1993]	47	-	-	90%
Ho	Brunelli e Poggio [Brunelli e Poggio, 1993]	47	-	-	100%
Ho	Wong et al. [Wong et al., 1989]	6	-	-	100%
Ho	Turk and Pentland [Turk e Pentland, 1991]	16	-	-	96%
Ho	Baron	42	-	-	100%
Ho	Cottrell e Fleming [Cottrell e Fleming, 1990]	11	-	-	100%
Ho	Stonham [Stonham, 1986]	16	200-400	-	100%

Ab - Abstrativo ; Ho - Holístico; S - Indivíduos distintos; i - Variações;  
j - Conjunto de treinamento ; k - Conjunto de validação; r - performance.



Robson e Craw [Robertson e Craw, 1993] discutiram os testes de sistemas de reconhecimento de faces e concluíram que sistemas que trabalham bem com faces restritas para sua aplicação específica não obtém os mesmos resultados com faces gerais de origem independente. Isto mostra a relação que pode existir entre os resultados de um experimento e a base de dados utilizada.

Para comparar sistemas de reconhecimento deve-se estabelecer nomenclaturas comuns e especificações forneçam, no mínimo, os seguintes aspectos e parâmetros [Samaria, 1994]:

- i. O número de indivíduos distintos da base de dados "S", em geral compreendendo variações de raça, sexo e idade.
- ii. A base "S" deve conter o mesmo número de variações "i" de cada indivíduo. Onde o tamanho final da base de dados é então dada por "i x S".
- iii. Das "i" variações de um mesmo indivíduo, "j" são usadas para treinamento e "k" são usadas para testes. Assim o tamanho do conjunto de treinamento é dado por "j x I" e do conjunto de validação é "k x I".
- iv. Como medida de desempenho do sistema, no mínimo as taxas TFA, TFR devem ser especificadas.

A figura 3.2 mostra a estruturação de um banco de faces típico.

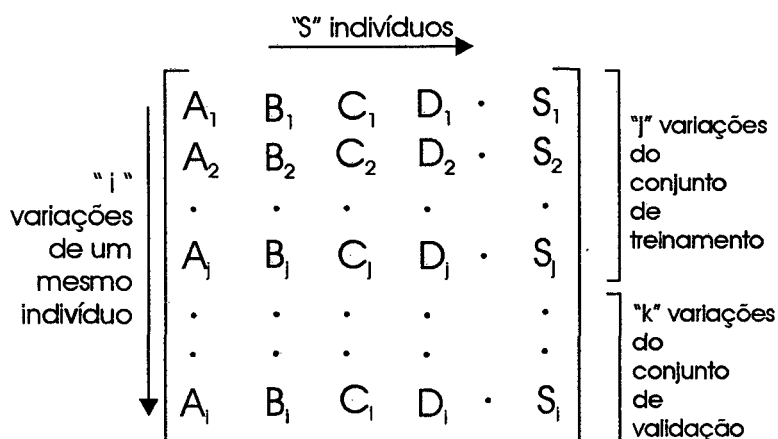


Figura 3.2 Estrutura típica de um banco de dados de faces.

### 3.4 Protocolo de Ensaio

A mais importante metodologia de avaliação de sistemas de reconhecimento de faces teve início em setembro de 1996 no "Army Research Laboratory - ARL"<sup>12</sup> nos EUA, o qual iniciou um programa de pesquisa e desenvolvimento chamado de "Face Recognition Technology - FERET" [Gutta et al., 1995], [Gutta et al., 1998] e [Phillips et al., 2000]. Este programa busca padronizar os procedimentos de teste dos sistemas reconhecedores de faces, colocando à disposição dos pesquisadores procedimentos de ensaios e uma base de dados de faces que em 1997 contava com 14.126 imagens faciais de 1199 indivíduos [Phillips et al., 2000]. FERET descreve procedimentos tanto para sistemas de verificação como de identificação. Sua base de dados é composta de imagens com diferentes condições de iluminação, tamanho e tipos de rostos que estão reunidas em grupos segundo a etnia, sexo e idade.

A título de exemplo, na tabela 3.2 pode-se observar os resultados de uma avaliação em um método de verificação de faces típico, utilizando os procedimentos e a base de dados do FERET.

Nota-se que a TFR aumenta conforme aumentam as variações das faces na base de faces. Um aumento acentuado na TFR ocorre quando a imagem tem a iluminação variada, mostrando que este é um fator importante a ser controlado nas tecnologias que utilizam reconhecimento baseado em imagens 2D. Também é nítida a influência de imagens adquiridas com espaço de tempo de até 1,5 ano.

Tabela 3.2 Desempenhos típicos de um sistema de verificação de faces

Categoria	TFA	TFR
Mesmo dia, mesma iluminação.	2	0,4
Mesmo dia, diferente iluminação.	2	9
Diferentes dias.	2	11
Diferentes dias, espaçados de 1,5 anos.	2	43

Conforme já reportado, a comparação de resultados de vários métodos de identificação é uma tarefa complexa, pois estes são obtidos utilizando-se diferentes conjuntos de imagens e procedimentos. Isto se deve por não estar completamente difundida e implantada uma

<sup>12</sup> "<http://www.arl.mil>".

terminologia comum para a elaboração dos ensaios e apresentação dos resultados. Uma comparação eficiente entre os métodos de reconhecimento de faces existentes, com base em dados publicados sem uma prévia padronização dos ensaios, fica prejudicada pela falta de elementos em comum entre estes ensaios.

A seguir, a título de orientação é apresentada uma lista de questões a serem verificadas quando se comparam resultados de identificações de faces baseados em métodos diferentes. As respostas às questões a seguir contribuem para o entendimento dos resultados alcançados:

- i. Expressões faciais, orientação da face e iluminação são controladas?
- ii. Óculos, barba e outras marcas faciais são permitidos?
- iii. Os elementos da base de dados estão separados por sexo, origem étnica e idade?
- iv. Quantos indivíduos existem na base de dados? Quantos foram usados para treinamento e quantos para testes?
- v. As características faciais são alocadas manualmente? A escala da imagem é controlada?

### **3.5 Aplicações dos Sistemas de Reconhecimento de Faces**

Um Sistema Automático de Reconhecimento de Faces - SARF pode ser utilizado nos mais diversos campos de aplicações:

- i. Na Indústria, este sistema é utilizado para o controle de acesso a áreas restritas e a sistemas de informações.
- ii. Na área governamental, o sistema atende as necessidades de vigilância de fronteiras, controle de imigração, monitoramento de segurança em portos e aeroportos.
- iii. O sistema pode ainda ser utilizado em casos forenses e em técnicas contra terrorismo por agências de inteligência ou departamentos de polícia.
- iv. As forças armadas podem utilizar o sistema para controle de tropas, monitoramento de campo de batalha e em identificação pessoal de militares.
- v. Na medicina pode ser usado no monitoramento da evolução do estado clínico de pacientes através de suas expressões faciais [Arad et al., 1994], ou no próprio diagnóstico de doenças que se refletem em mudanças na face.

## Capítulo 4

# Medição da Geometria

---

*Onde existe matéria, existe geometria.*

Johannes Kepler, (1571-1630).

**M**edir faces humanas, com um mínimo de iteração entre o indivíduo e máquina num curto espaço de tempo, requer um sistema de medição com características especiais. Nestes casos, sistemas ópticos são recomendados pela ausência de contato entre a face e o sensor de medição, e pela sua característica de aquisição de imagens quase instantâneas.

### 4.1 Técnicas Ópticas de Medição de Formas

Existem várias técnicas ópticas para medição de formas de objetos. Neste trabalho serão apresentadas, a título de visão geral, as técnicas que permitem a extração da forma a partir de uma imagem.

### 4.1.1 Radar Laser

Sistemas de medição baseado em radar LASER, algumas vezes conhecido como LIDAR<sup>1</sup>, têm se tornado comercialmente disponíveis ultimamente [Kerr, 1988], [Harding, 1999] e [Bieman, 1988]. Como pode ser visto na figura 4.1, o sistema radar LASER envia um feixe de luz modulada que é refletida sobre a superfície do objeto retornando ao receptor. A altura do objeto sob teste é avaliada a partir do tempo gasto pelo sinal desde sua geração até sua recepção ("flight-time") ou pelo deslocamento da fase ou frequência da onda. O feixe é então aplicado em forma de varredura sobre o objeto em teste e sua forma é extraída. Uma resolução de décimos de milímetros é alcançada com este sistema.

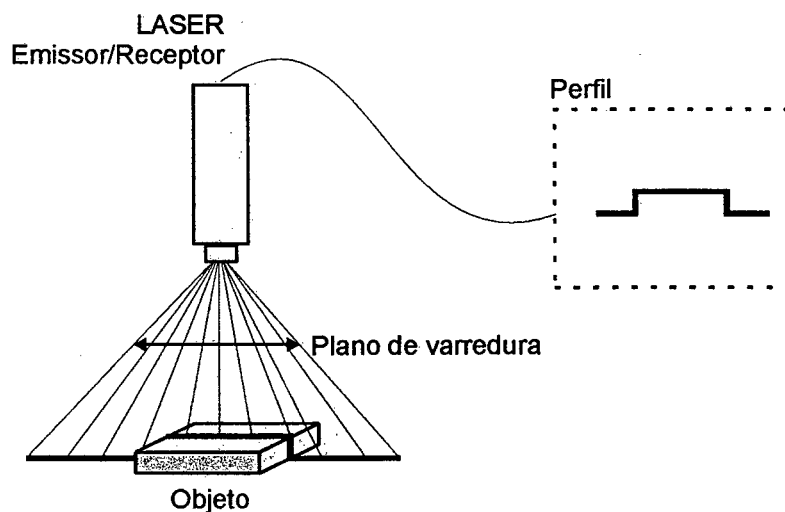


Figura 4.1 Diagrama esquemático de um medidor do tipo RADAR LASER .

### 4.1.2 Topogrametria

Quando um ser humano vê um objeto ele recebe a informação 3D baseada principalmente no efeito de paralaxe<sup>2</sup> e visualização das sombras. Topogrametria utiliza o mesmo princípio onde a informação de paralaxe é proveniente de duas observações do mesmo objeto fornecidas por duas perspectivas diferentes. Estas duas perspectivas são fornecidas por duas câmeras espacialmente distanciadas conforme mostrada na figura 4.2.

<sup>1</sup> LIDAR - Light Detector and Ranging [Photonics Spectra, 1995]

<sup>2</sup> Designação para o fenômeno óptico que causa um movimento relativo entre dois objetos quando o ponto de vista é deslocado lateralmente.

Para a extração da forma, correlacionam-se as duas imagens onde as diferenças entre os pontos fornecem a informação da profundidade do objeto. A aplicação deste método está limitada a superfícies com pontos característicos bem definidos. Medição de formas em superfícies muito suaves não é indicada para este tipo de método. Um dos maiores campos de aplicação da topogrametria tem sido no processo de visão de robôs.

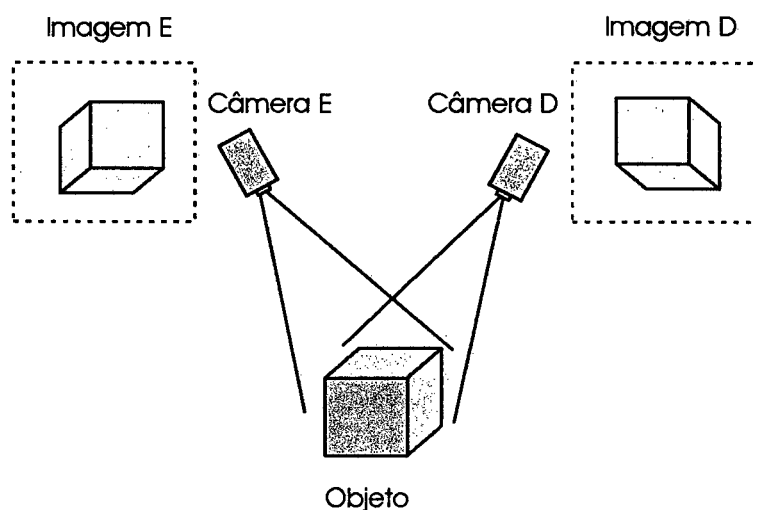


Figura 4.2 Diagrama esquemático do sistema topométrico.

### 4.1.3 Técnicas de Moiré

Técnicas de Moiré são bastante conhecidas e utilizadas em medição de formas de objetos.

#### Moiré de sombras

Este método utiliza apenas uma grade física para gerar a interferência que resulta em um mapa de curvas de níveis proporcionais às alturas dos objetos. Na figura 4.2 pode-se ver o diagrama esquemático de um sistema Moiré de sombras. A grade física projeta um padrão de sombra sobre a superfície. Este padrão de sombra quando visto através da mesma grade física pela câmera mas em uma outra posição, diferente da do projetor, provoca um "batimento" entre a sombra da grade e as suas próprias linhas que resulta em um mapa de curvas de nível proporcional às alturas. Para a transformação das curvas de nível obtidas pelo método em alturas correspondentes, necessita-se de pelo menos três imagens com deslocamento de fase, o que dificulta o uso desta técnica neste trabalho.

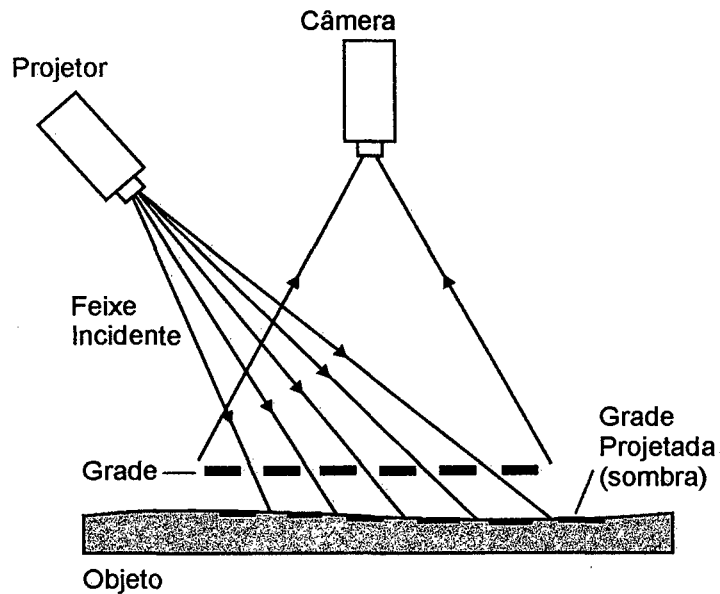


Figura 4.3 Princípio de funcionamento do Sistema Moiré de sombras.

### **Moiré de projeção**

Neste sistema franjas são projetadas sobre a superfície do objeto através da grade do projetor. A superfície do objeto deforma estas franjas que são captadas pela câmera após a interferência com uma segunda grade conforme pode ser visto na figura 4.4.

Esta técnica também fornece um mapa de curvas de nível. O processamento destas curvas é facilitado se, a exemplo do sistema de Moiré de sombras, se fizer no mínimo 3 imagens com deslocamento de fase entre elas. Este sistema também se mostra mais indicado para aplicações estáticas.

### **Interferometria de Moiré**

Este método de extração de formas necessita que a superfície do objeto seja pelo menos semi-espelhada, requisito não disponível em nosso tipo de aplicação.

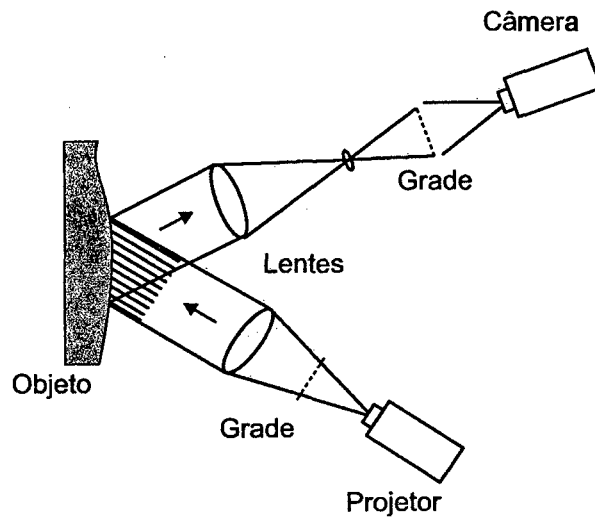


Figura 4.4 Moiré de Projeção

#### 4.1.4 Holografia Interferométrica

A holografia interferométrica, a princípio, não pode ser usada para a medição de formas, mas de deslocamentos. Porém uma modificação na concepção básica do princípio torna possível seu uso na medição de formas. Para tanto se faz necessário o uso de dois comprimentos de onda diferentes na etapa de gravação do holograma. O resultado obtido é um mapa de curvas de nível proporcionais às alturas do objeto. Este procedimento, caro, complexo e muito sensível, tem sua aplicação dificultada em nosso trabalho.

#### 4.1.5 Interferometria de Speckle

Método bastante conhecido em medição de deformações, pode ser adaptado, a exemplo da Holografia Interferométrica, para medição de formas. Para tanto, usa-se dois comprimentos de ondas diferentes para a geração do mapa de Speckle. O fato de se ter que obter duas imagens dificulta a aplicação deste princípio neste trabalho pelas mesmas razões do método do item 4.1.4.



### 4.1.6 Medição por Pontos com Triangulação

Vários medidores de formas utilizam triangulação, dentre eles o mais popular é o "range finder"<sup>3</sup> [Göpel et al., 1992] e [Harding, 1999]. A imagem do feixe de luz que incide sob a superfície do objeto é capturada e projetada pelo sistema óptico sobre um fotodetector linear. Através da relação trigonométrica de dois triângulos formados entre o objeto, lente e fotodetector do sistema óptico é calculada a distância entre o objeto e o medidor. Na figura 4.5 pode-se ver um diagrama esquemático do princípio da triangulação onde o sistema está medindo a distância entre dois pontos sobre uma mesma superfície, que se desloca para uma nova posição.

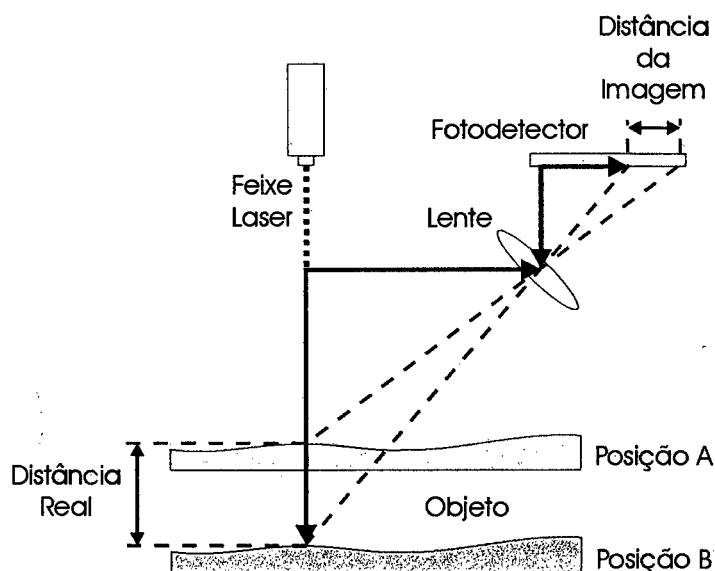


Figura 4.5 Medidor de deslocamentos tipo triangulação.

### 4.1.7 Medição por Linhas com Triangulação

O medidor de deslocamentos por plano de luz utiliza o princípio da triangulação mas usando um feixe tipo plano de luz que projeta uma linha sobre o objeto. O plano de luz projetado é também chamado de folha de luz. Na figura 4.6 pode-se ver o diagrama esquemático deste sistema de medição.

<sup>3</sup> "Rangefinders" - (i) Medidor de distância óptica que depende da triangulação de dois feixes convergentes sobre um objeto. (ii) Um sistema baseado na medição de tempo de viagem de uma onda entre o objeto e o ponto de referência, como em radar, sonar e sistemas LASER [Photonics Spectra, 1995].

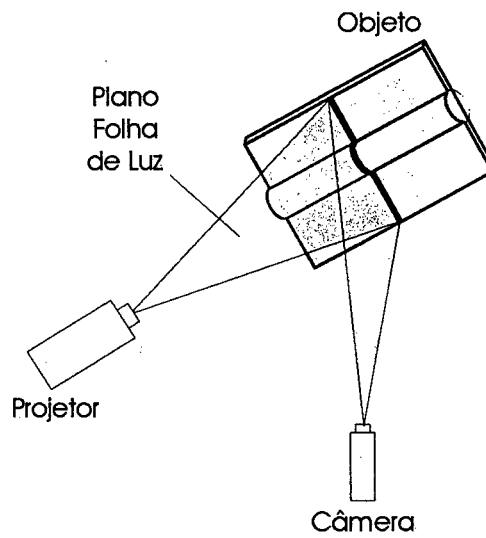


Figura 4.6 Diagrama esquemático do medidor por plano de luz.

#### 4.1.8 Sistemas de Varredura

Os sistemas de medição de forma do tipo por varredura são muito usados em engenharia reversa. Pela sua característica de rotação completa, são capazes de medir um objeto tridimensional em todo seu contorno. Na figura 4.7 pode-se ver o diagrama esquemático deste princípio de medição de formas.

Este sistema exige a necessidade de movimento relativo entre o objeto e o sistema óptico de medição, o que implica num maior tempo de aquisição, não indicado para a aplicação proposta neste trabalho, onde se deseja mínima iteração entre o objeto (face humana) e o sistema de aquisição.

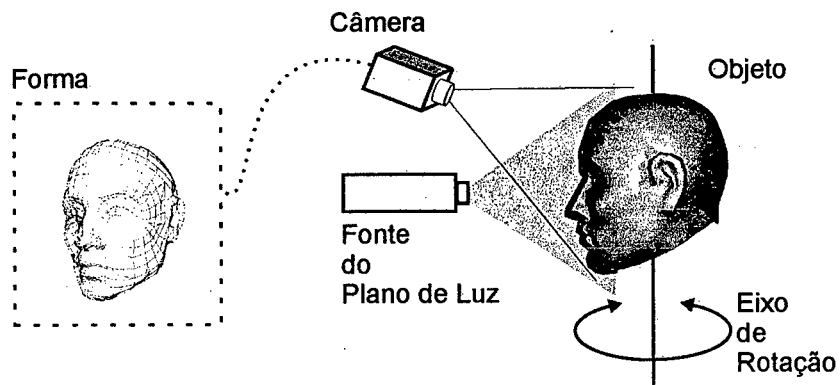


Figura 4.7 Diagrama esquemático de um medidor do tipo varredura.

### 4.1.9 Projeção de Franjas

A medição de formas através de franjas projetadas sobre a superfície de objetos tem sido muito usada em combinações com modernas técnicas de processamento de digital imagens. Diversas são as abordagens possíveis para a solução deste problema que fundamentalmente utilizam o princípio da triangulação para a determinação das alturas do objeto a partir do padrão de franjas projetadas sobre o objeto.

#### Problemas em Sistemas de Medição com Franjas

Um dos principais problemas do método de projeção de franjas é a ambigüidade destas devido a projeção sobre objetos descontínuos. Na figura 4.8 pode-se observar a diferença entre franjas projetadas sobre uma superfície contínua e sobre uma outra superfície descontínua. O problema da ambigüidade reside no fato de que algumas franjas, no mapa de franjas, dependendo da geometria do objeto podem ser confundidas.

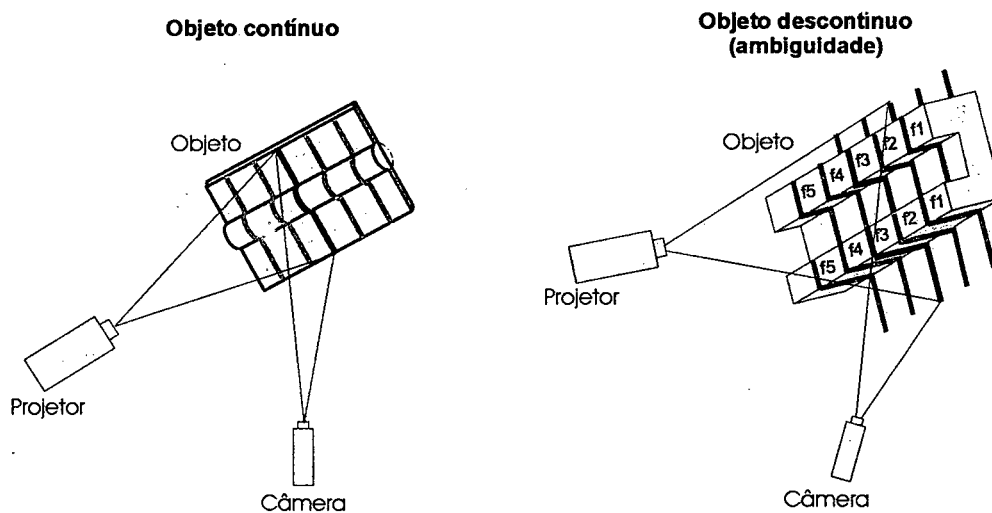


Figura 4.8 Projeção de franjas com e sem o problema de ambigüidade.

Uma maneira de se evitar este problema é pelo emprego de franjas com luz codificada. Os diferentes níveis de cinza utilizados neste método permitem que através de um algoritmo adequado se consiga evitar a ambigüidade das franjas. Na figura 4.9 pode-se ver como este princípio se aplica a objetos com descontinuidade.

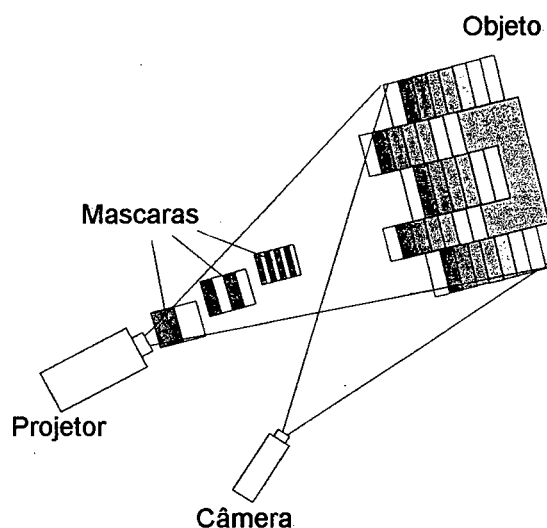


Figura 4.9 Luz codificada para evitar o problema de ambigüidade das franjas.

## 4.2 O Princípio de Medição de Formas Utilizado

O princípio de medição de formas utilizado neste trabalho é o de projeção de franjas. A escolha deste método se deve em função dos requisitos do sistema de extração da forma e das características da face humana.

A aplicação proposta neste trabalho exige que o sistema de extração de formas tenha as seguintes características:

- i. Rapidez de aquisição
- ii. Simplicidade
- iii. Robustez
- iv. Baixo Custo

Quanto ao tipo de geometria da face humana, esta aponta para o método de projeção de franjas uma vez que com uma escolha adequada da posição de incidência das franjas, ficam diminuídas as possibilidades de ambigüidade.

A faixa de abrangência de um sistema de projeção de franjas sobre uma face humana frontalmente posicionada também atende as necessidades qualitativas e quantitativas do sistema proposto de reconhecimento de faces.

### 4.2.1 Detalhamento do Princípio de Projeção de Franjas

O sistema de projeção de franjas consiste em uma fonte de luz que projeta franjas obliquamente sobre a superfície do objeto (face humana) e uma câmera com posicionamento aproximadamente perpendicular a esta mesma superfície a qual adquire a imagem do mapa de franjas projetado. Na figura 4.10 pode-se ver um diagrama esquemático simplificado de tal sistema.

A frequência espacial das franjas projetadas varia de acordo com a inclinação da superfície. Ainda na figura 4.10 vê-se que na região "c" do objeto a frequência espacial possui um valor típico e constante. Já nas regiões "a" e "b" do objeto, onde existem inclinações, o valor da frequência espacial observada varia para  $f_b$  e  $f_a$ . Note que o tipo de inclinação, se com derivada positiva ou negativa, determina o valor da frequência espacial como maior ou menor que a frequência da região plana respectivamente.

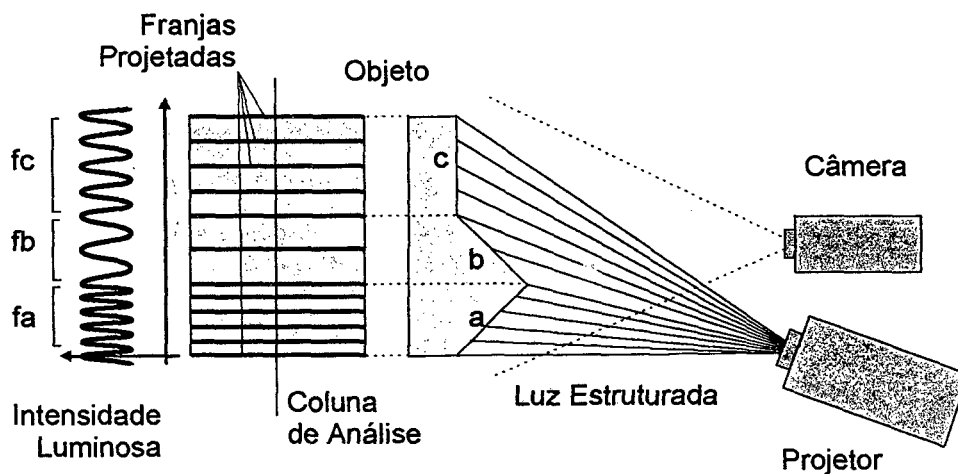


Figura 4.10 Diagrama esquemático do sistema de projeção de franjas.

O conjunto das franjas projetadas é chamado de mapa de franjas e é o resultado da observação das linhas de luz projetadas sobre a superfície do objeto. Elas contêm a informação da forma do objeto. Na figura 4.11 pode-se ver a representação gráfica de uma face humana típica com franjas projetadas. Note que as franjas sofrem deslocamentos verticais proporcionais à profundidade (altura) da face.



Figura 4.11 Representação das franjas projetadas em uma face humana.

#### **4.2.2 Processamento do Mapa de Franjas**

O processamento do mapa de franjas compreende a extração qualitativa da informação da forma da face. O cálculo quantitativo entre a forma ou alturas físicas da face é feito utilizando-se o princípio da triangulação aplicado à geometria do sistema óptico.

Diversas técnicas de processamento do mapa de franjas têm sido apresentadas na literatura especializada. A técnica utilizada neste trabalho, referida na literatura como “franjas portadoras” [Gonçalves, 1991], realiza o processamento do mapa de franjas no domínio espacial e não temporal. Isto representa uma vantagem muito importante, pois o mapa de franjas pode ser adquirido de uma só vez. Os outros métodos temporais tradicionais necessitam de uma série de imagens adquiridas ao longo do tempo para conseguirem extrair a informação da forma. Outro aspecto favorável é que a forma mais bem definida deste tipo de franja possibilita: (i) o emprego de técnicas automáticas para a aquisição do mapa, (ii) possibilidade de medição de frações da ordem de franjas e (iii) boa relação sinal/ruído.

#### **4.2.3 Extração do Sinal das Franjas**

O método usado para a extração dos dados da forma da face é conhecido como Perfilometria de Fourier - PF [Takeda et al., 1999] e [Takeda e Mutoh, 1999]. Este método permite que se extraia informação através do processamento dos níveis de cinza existentes entre as franjas claras e escuras. Para tanto considera-se o sinal de intensidade luminosa dos mapas de franjas

como se fossem um sinal de Freqüência Modulada - FM. Nele as franjas projetadas e observadas sobre uma linha perpendicular à direção da luz estruturada podem ser representadas pela equação 4.1 a seguir:

$$g(x, y) = a(x, y) + b(x, y) \cos[\theta(x, y) + 2\pi f_o x] \quad (4.1)$$

Onde:

$g(x, y)$  é o sinal de intensidade luminosa sobre a superfície do objeto.

$a(x, y)$  é a amplitude do sinal (iluminação) de fundo.

$b(x, y)$  é a amplitude da modulação do sinal (das franjas).

$\theta$  é o ângulo de modulação de fase das franjas projetadas.

$f_o$  é a freqüência espacial das franjas portadoras na direção  $y$ .

$x$  e  $y$  são as coordenadas espaciais.

O parâmetro que carrega a informação das alturas do objeto (grandeza de interesse) é o ângulo de fase  $\theta(x, y)$ . Para se obter  $\theta(x, y)$ , evitando a ambigüidade de sinais e a influência do sinal de iluminação de fundo, empregam-se as relações de Euler na equação 4.1:

$$e^{i\theta} = \cos(\theta) + i\text{sen}(\theta) \quad (4.2)$$

$$\cos(\theta) = \frac{1}{2}(e^{i\theta} + e^{-i\theta}) \quad (4.3)$$

$$g(x, y) = a(x, y) + c(x, y)e^{i2\pi f_o x} + c(x, y)e^{-i2\pi f_o x} \quad (4.4)$$

Onde :

$$c(x, y) = \frac{1}{2}b(x, y)e^{i\theta(x, y)} \quad (4.5)$$

Aplicando-se a Transformada de Fourier a este sinal, obtém-se:

$$G(f_x, y) = A(f_x, y) + C(f_x - f_0, y) + C(f_x + f_0, y) \quad (4.6)$$

Na figura 4.12 abaixo pode-se ver o espectro característico do sinal de intensidade das franjas.

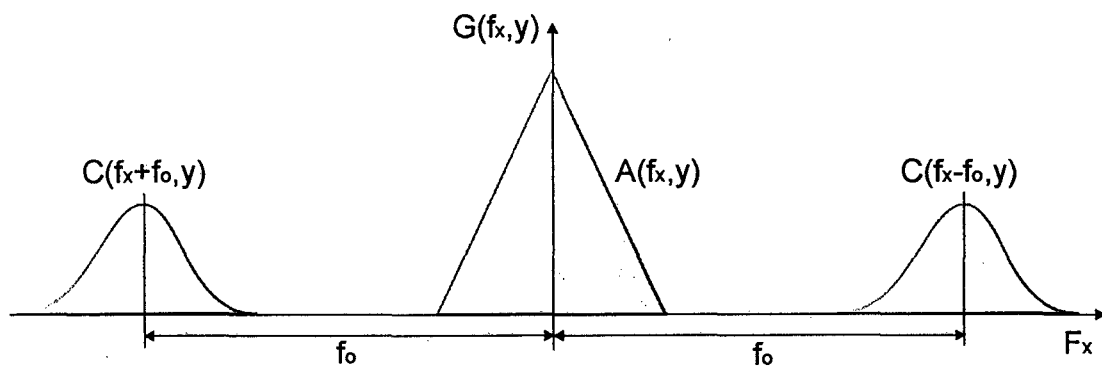


Figura 4.12 Espectro do sinal  $G(f_x, y)$

Note que o terceiro termo da equação 4.6, devido à exponencial negativa, produz uma componente no eixo negativo das freqüências, o qual não tem significado físico podendo ser desprezado. Também a componente da amplitude de iluminação de fundo  $A(f_x, y)$  pode ser eliminada, se aplicado um filtro passa altas a este sinal. O sinal resultante é então:

$$G(f_x, y) = C(f_x + f_0, y) \quad (4.7)$$

Aplicando-se a Transformada Inversa de Fourier - TIF à equação 4.7:

$$g(x, y) = c(x, y) = \frac{1}{2} b(x, y) e^{i\theta(x, y)} \quad (4.8)$$

O sinal  $g(x, y)$  pode ainda ser transformado aplicando-se as relações de Euler de maneira inversa, na equação 4.8:

$$g(x, y) = \frac{1}{2} b(x, y) \cos[\theta(x, y)] + i \frac{1}{2} b(x, y) \text{sen}[\theta(x, y)] \quad (4.9)$$



Onde  $\theta(x, y)$ , sinal de interesse que carrega a informação das alturas da face, pode então ser isolado na equação 4.9 resultando em:

$$\theta(x, y) = \operatorname{tg}^{-1} \frac{\operatorname{Im}[c(x, y)]}{\operatorname{Re}[c(x, y)]} \quad (4.10)$$

#### 4.2.4 Remoção do Salto de Fase

Note-se que a informação da forma do objeto está presente em  $\theta(x, y)$  da equação 4.10, no entanto os valores numéricos que esta equação fornece estão compreendidos entre  $+\pi$  e  $-\pi$  fornecendo uma descontinuidade nos valores das alturas correspondentes. Isto é chamado de salto de fase. Para eliminar as descontinuidades deste sinal, deve-se proceder à remoção dos saltos de fase<sup>4</sup> para que os dados da forma adquiram o formato natural do objeto. Na figura 4.13 pode-se ver o exemplo típico de um caso onde valores de fase  $\theta(x, y)$  foram calculados por uma função do tipo da equação 4.10. Note a presença das descontinuidades ou saltos de fase.

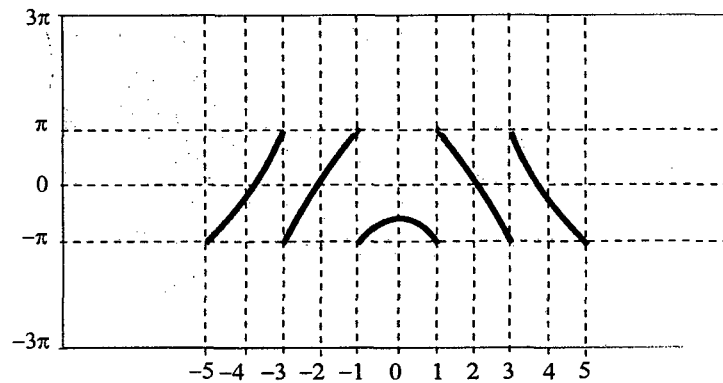


Figura 4.13 Sinal genérico de fase  $\theta(x, y)$  com salto de fase.

Para recuperar o sinal da forma natural do objeto que o originou, usam-se algoritmos específicos para esta função chamados de Algoritmos de Remoção de Saltos de Fase - ARSF [Asundi e Wensen, 1999], [Costantini, 2016], [Fritz et al., 1999], [Ghiglia et al., 1999], [Ghiglia e Romero, 1999], [Herráez et al., 1999], [Hung e Yamada, 1999], [Krämer e Loffeld, 2016],

<sup>4</sup> Phase unwrapping.

[Sandia National Laboratories, 1999], [Tarayre-Oriot e Massonnet, 1999], [Wei et al., 1999] e [Xu e Cumming, 1999].

O problema da remoção do salto de fase consiste em detectar as discontinuidades e adicionar o valor  $2\pi$  cada vez que o ângulo de fase saltar de  $\pi$  para  $-\pi$  e subtrair  $2\pi$  cada vez que o ângulo de fase saltar de  $-\pi$  para  $\pi$ . Na figura 4.14 pode-se ver o resultado da aplicação de um algoritmo de remoção do salto de fase ao sinal da figura 4.13.

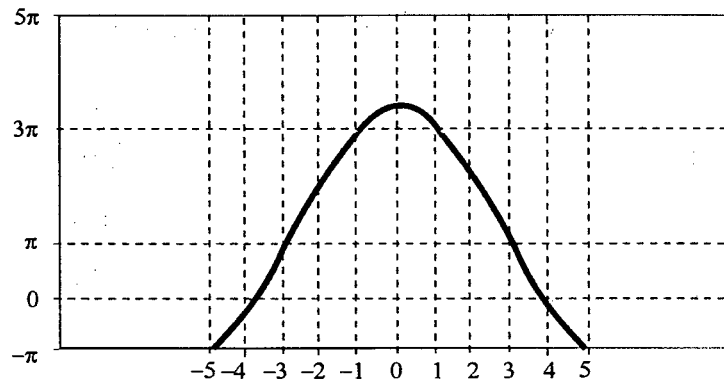


Figura 4.14 Sinal genérico de fase  $\theta(x,y)$  com salto de fase removido.

Os ARSF's processam a informação com base em um referencial inicial e calculam os próximos valores sempre em relação ao valor anterior. Neste processo, deve-se tomar cuidado com a qualidade do ponto de partida do ARSF, para que os dados subsequentes calculados a partir do ponto inicial possuam a integridade desejada.

### O ARSF Utilizado

Neste trabalho optou-se pelo ARSF do tipo bidimensional, desenvolvendo-se um algoritmo que privilegia os pontos da imagem que possuem maior grau de certeza [Asundi e Wensen, 1999]. Esta estratégia também chamada de "preenchimento por alagamento"<sup>5</sup> divide a imagem em diversos níveis de confiabilidade. O processamento se dá priorizando os pontos de mais alta confiabilidade processando seus vizinhos de mesmo nível e deixando os pontos de menor confiabilidade para o final. Apesar de se fazer a remoção do salto de fase da imagem da fase  $\theta(x,y)$ , usa-se como parâmetro de avaliação da confiabilidade dos pontos a amplitude de modulação da imagem do objeto (fator  $b(x,y)$  da equação 4.1). Os valores da modulação da imagem feitos a partir da Transformada Inversa de Fourier - TIF do sinal filtrado  $C(f_x-f_0,y)$

<sup>5</sup> "Flood Fill" - Preenchimento por alagamento.

(Figura 4.12) são diretamente proporcionais ao contraste das franjas no mapa. Portanto este parâmetro pode ser usado para a indicação das regiões de maior confiabilidade na imagem da fase  $\theta(x, y)$ .

#### 4.2.5 O Sistema Óptico Geométrico

O sistema óptico utilizado neste trabalho é de concepção simples, que utiliza elementos ópticos comuns e de baixo custo. A figura 4.15 a seguir, mostra a distribuição geométrica dos componentes do sistema óptico proposto.

Legenda:

$d$  = Distância entre o pivô óptico da câmera e o plano de projeção das franjas;

$q$  = Distância entre os pivôs ópticos da câmera e do projetor;

$O$  = Centro geométrico dos eixos ópticos da câmera e do projetor;

$C$  = Centro geométrico do pivô óptico da câmera;

$P$  = Centro geométrico do pivô óptico do projetor;

$\beta$  = Ângulo de abertura da câmera entre duas franjas consecutivas;

$\gamma$  = Ângulo de abertura entre duas franjas consecutivas no projetor;

$\varepsilon$  = Distância entre duas franjas a distância 's' devido ao ângulo " $\gamma$ ";

$\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$  = Parcelas de  $\varepsilon$ ;

$\alpha$  = Ângulo entre os eixos ópticos da câmera e do projetor;

$R$  = Plano de referência;

$h$  = Distância entre câmera e projetor, paralela ao plano de ref.;

$a$  e  $b$  = Afastamentos sobre o plano de ref. relativo ao ângulo de projeção  $\gamma$ .

Através do equacionamento da geometria óptica da figura 4.15 pode-se extrair as relações entre o ângulo " $\beta$ " e as demais variáveis de modo a se analisar o comportamento deste sistema.

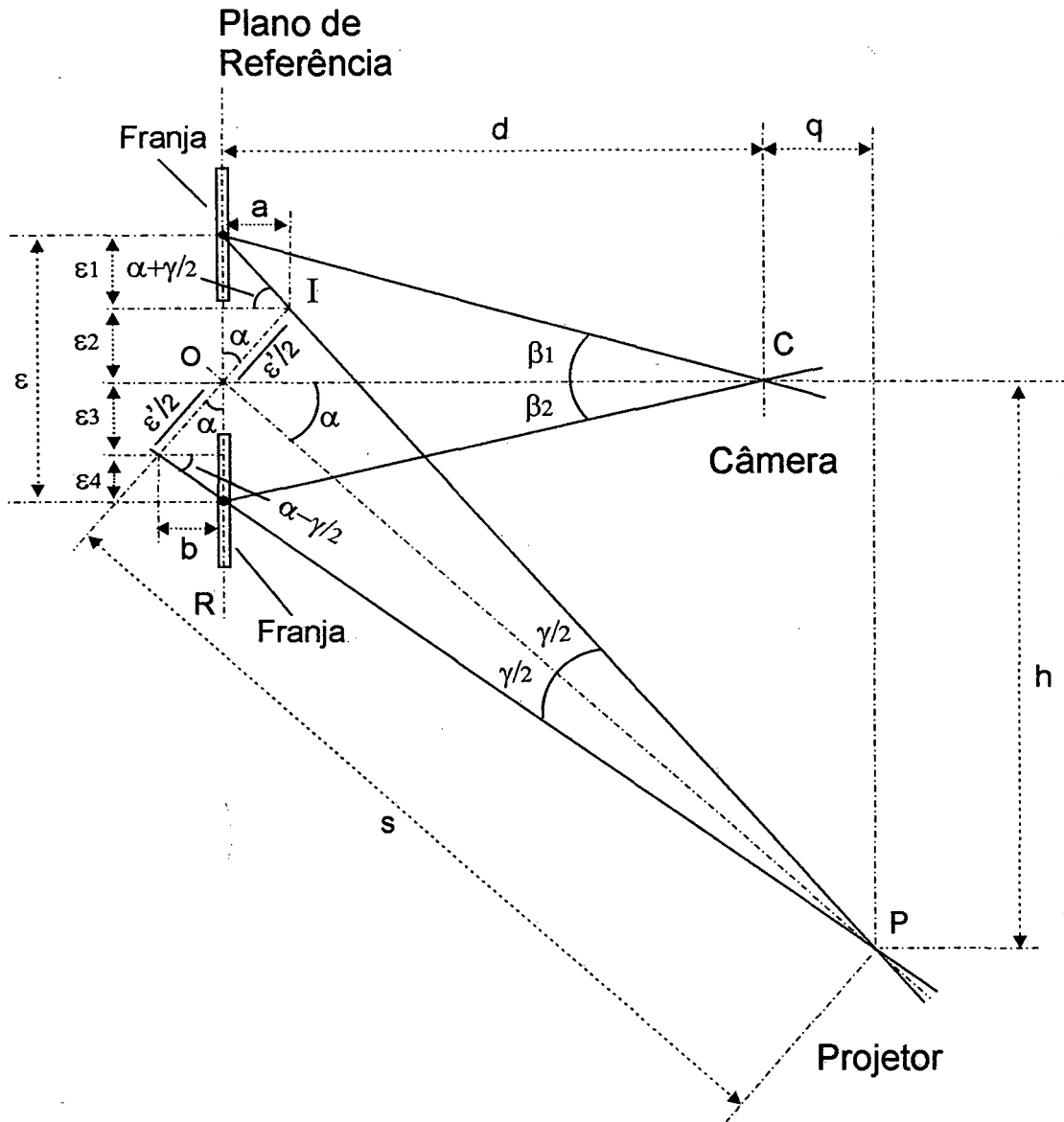


Figura 4.15 Geometria do sistema óptico.

As relações entre "β", "ε" e "d" são dadas por:

$$\beta = \beta_1 + \beta_2 \quad (4.11)$$

$$\beta_1 = a \tan\left(\frac{\epsilon_1 + \epsilon_2}{d}\right) \quad (4.12)$$

$$\beta_2 = a \tan\left(\frac{\varepsilon_3 + \varepsilon_4}{d}\right) \quad (4.13)$$

A relação entre “ $\gamma$ ”, “ $\varepsilon$ ” e “ $s$ ” é dada por:

$$\varepsilon' = 2 \cdot s \cdot \tan\left(\frac{\gamma}{2}\right) \quad (4.14)$$

Da figura 4.15 tem-se ainda:

$$\varepsilon_4 = \varepsilon_3 = \cos(\alpha) \cdot \left(\frac{\varepsilon'}{2}\right) \quad (4.15)$$

$$\varepsilon_1 = a \cdot \tan\left(\alpha + \frac{\gamma}{2}\right) \quad (4.16)$$

$$\varepsilon_4 = b \cdot \tan\left(\alpha + \frac{\gamma}{2}\right) \quad (4.17)$$

Onde “ $a$ ” e “ $b$ ” são dados por:

$$a = b = \text{sen}(\alpha) \cdot \frac{\varepsilon'}{2} \quad (4.18)$$

Substituindo-se “ $a$ ” e “ $b$ ” na equação 4.16 e 4.17:

$$\varepsilon_1 = \tan\left(\alpha + \frac{\gamma}{2}\right) \cdot \text{sen}(\alpha) \cdot \frac{\varepsilon'}{2} \quad (4.19)$$

$$\varepsilon_4 = \tan\left(\alpha - \frac{\gamma}{2}\right) \cdot \text{sen}(\alpha) \cdot \frac{\varepsilon'}{2} \quad (4.20)$$

A relação entre “ $s$ ”, “ $d$ ”, “ $q$ ” e “ $\alpha$ ” é dada por:

$$s = \frac{d+q}{\cos(\alpha)} \quad (4.21)$$

$$s = \sqrt{(d+q)^2 + h^2} \quad (4.22)$$

Finalmente combinando-se as equações 4.21 e 4.22 obtém-se a relação do ângulo, entre os dois eixos ópticos do projetor e da câmera, e os demais parâmetros geométricos:

$$\cos(\alpha) = \frac{d+q}{\sqrt{(d+q)^2 + h^2}} \quad (4.23)$$

### **A Frequência Espacial**

Idealmente a frequência espacial das franjas projetadas sobre a superfície de um objeto plano deve ser constante dentro da faixa de medição sob o ponto de vista da câmera. Erros de medição da fase  $\theta(x, y)$  e conseqüentemente erros na medição da forma serão introduzidos se dentro do volume de medição onde estará a face houver variação da frequência espacial das franjas projetadas.

Neste trabalho usa-se um projetor de franjas de projeção central. Com este equipamento obtém-se um efeito de diminuição da frequência espacial quando aumenta a distância entre o projetor e o anteparo, ao longo do eixo do projetor. Este efeito é indesejado, uma vez que os resultados da medição dependerão da distância entre objeto e projetor. Por outro lado a maioria das câmeras possuem esta mesma característica, mas em sentido inverso, o que enseja esperar que se possa compensar o problema do projetor com um efeito contrário fornecido pela câmera. E isto é realmente o que ocorre. O efeito da projeção não paralela das franjas compensa o efeito não paralelo do campo de visão da câmera. Para que isto aconteça, deve-se posicionar o pivô óptico da câmera na mesma linha do pivô óptico do projetor, ou seja pela figura 4.15 a dimensão "q" deve ser igual a zero. Este posicionamento porém não é tão simples de ser conseguido, pois os fabricantes de câmeras e projetores em geral não fornecem nenhuma informação a respeito da posição do pivô óptico de seus sistemas. Neste trabalho o pivô óptico do sistema foi determinado de maneira experimental através ensaios realizados em uma bancada óptica.

**Ângulo de Visão Entre Duas Franias Consecutivas e a Freqüência Espacial**

Da figura 4.15 pode-se obter a relação entre o ângulo “ $\beta$ ” e a freqüência espacial “ $f$ ” conforme mostrado a seguir:

$$\beta = \frac{\varepsilon}{d} \quad (4.24)$$

$$\varepsilon = \frac{1}{f} \quad (4.25)$$

$$\beta = \frac{1}{f \cdot d} \quad (4.26)$$

Onde a variação relativa de “ $\beta$ ” é dada por:

$$\Delta\beta = \beta_1 - \beta_2 = \frac{1}{d \cdot f_2} - \frac{1}{d \cdot f_1} = \frac{f_1 - f_2}{d \cdot f_1 \cdot f_2} \quad (4.27)$$

Finalmente:

$$\frac{\Delta\beta}{\beta} = -\frac{\Delta f}{f} \quad (4.28)$$

Pode-se ver através da equação 4.28 que o ângulo “ $\beta$ ” e a freqüência espacial “ $f$ ” tem a mesma relação de variação porém de sentidos contrários, como era de se esperar.

**Minimizando a Influência da Variação da Freqüência Espacial**

Minimizar a influência da freqüência espacial dentro do campo de medição significa, conforme já explanado no item acima, achar uma relação ideal de geometria óptica que diminua os efeitos da variação da freqüência espacial em função do posicionamento do objeto no campo de medição.

Conforme mostrado no item anterior na equação 4.28, a análise da influência da freqüência espacial pode ser realizada através da análise do ângulo “ $\beta$ ”. A relação entre ângulo “ $\beta$ ” e os parâmetros optogeométricos do protótipo é desenvolvida a seguir.

Substituindo-se “ $\varepsilon_1$ ”, “ $\varepsilon_2$ ”, “ $\varepsilon_3$ ” e “ $\varepsilon_4$ ” nas equações 4.12 e 4.13, obtém-se:

$$\beta_1 = a \tan \left( \frac{\cos(\alpha) \cdot \frac{\varepsilon'}{2} + \tan\left(\alpha + \frac{\gamma}{2}\right) \cdot \text{sen}(\alpha) \cdot \frac{\varepsilon'}{2}}{d} \right) \quad (4.29)$$

$$\beta_2 = a \tan \left( \frac{\cos(\alpha) \cdot \frac{\varepsilon'}{2} + \tan\left(\alpha - \frac{\gamma}{2}\right) \cdot \text{sen}(\alpha) \cdot \frac{\varepsilon'}{2}}{d} \right) \quad (4.30)$$

Substituindo-se “ $\varepsilon$ ” da equação 4.14 e “ $s$ ” da equação 4.22 nas equações 4.29 e 4.30 obtém-se os modelos completos:

$$\beta_1 = a \tan \left( \frac{\tan\left(\alpha + \frac{\gamma}{2}\right) \cdot \text{sen}(\alpha) \cdot \frac{(d+q)}{\cos(\alpha)} \cdot \tan\left(\frac{\gamma}{2}\right) + (d+q) \cdot \tan\left(\frac{\gamma}{2}\right)}{d} \right) \quad (4.31)$$

$$\beta_2 = a \tan \left( \frac{(d+q) \cdot \tan\left(\frac{\gamma}{2}\right) - \tan\left(\frac{\gamma}{2} - \alpha\right) \cdot \text{sen}(\alpha) \cdot \frac{(d+q)}{\cos(\alpha)} \cdot \tan\left(\frac{\gamma}{2}\right)}{d} \right) \quad (4.32)$$

Onde : “ $\beta$ ” é dado por “ $\beta_1$ +” $\beta_2$ ”.

Algumas simplificações podem ser feitas nas equações 4.31 e 4.32 de modo a reduzi-las, estas podem ser vistas na tabela 4.1.



Tabela 4.1 Simplificações aplicáveis ao modelo completo de "β<sub>1</sub>" e "β<sub>2</sub>".

Substituir:	Por:	Justificativa:
$\alpha\gamma$	$\alpha$	$(\alpha+\gamma)$ é de segunda ordem em relação à $\alpha$
$\tan(\beta)$	$\beta$	$\beta \ll d$
$\text{sen}(\gamma)$	$\gamma$	$\gamma \ll s$
$\text{cos}(\gamma)$	1	$\gamma \ll s$

A relação simplificada entre o ângulo "b" e os parâmetros ópticogeométricos é finalmente dada por:

$$\beta = \frac{\gamma \cdot (d + q)}{d \cdot \cos(\alpha)^2} \quad (4.33)$$

Derivando-se "β" em função de "d" obtém-se a taxa de variação deste em função da distância câmera objeto:

$$\frac{\partial \beta}{\partial d} = -\frac{\gamma \cdot q}{d \cdot \cos(\alpha)^2} \quad (4.34)$$

A equação 4.34 mostra que se "q" for igual a zero, "β" não varia com a posição "d". Isto significa que quando os dois pivôs dos sistemas ópticos da câmera e do projetor estiverem alinhados ( $q = 0$ ) o ângulo "β" não varia. O que equivale a dizer que a frequência espacial não varia nestas condições. Desta forma fica mostrada a importância do alinhamento dos pivôs para a medição.

É importante lembrar que todo este desenvolvimento matemático vale para a condição das duas primeiras franjas em torno do eixo óptico da câmera. Como o protótipo tem outras franjas além das duas equacionadas, pode-se dizer que quando houver o alinhamento entre os dois pivôs dos sistemas ópticos teremos a condição de menor erro no sistema. Um levantamento mais rigoroso, somente, faz sentido se os erros requeridos forem muito pequenos e neste caso o procedimento de levantamento dos erros seria através de uma calibração prática do sistema.

**Relação Entre Fase e Altura**

De posse da informação do ângulo de fase  $\theta(x, y)$  com o salto de fase removido procede-se a transformação dimensional desta informação com o objetivo de se obter as alturas da forma do objeto em uma unidade dimensional conhecida.

Note-se que uma vez que o sistema de aquisição de imagens seja o mesmo, isto é, possua os mesmos parâmetros ópticos geométricos durante a aquisição da imagem do banco de dados e da imagem para teste, esta transformação não é necessariamente obrigatória. A figura 4.16 mostra a relação entre a fase  $\theta(x, y)$  e a altura " $\Delta d$ " de um objetivo genérico para a geometria óptica usada neste trabalho.

Legenda:

- S = Eixo óptico da câmera;
- G = Eixo óptico do projetor;
- d = Distância entre o pivô da câmera "C" e o do plano de referência;
- h = Distância entre a câmera e o projetor;
- $\Delta d$  = Altura entre o plano de referência e o objeto;
- $\epsilon$  = Distância entre duas franjas consecutivas projetadas no plano de referência;
- $\epsilon'$  = Distância entre duas franjas consecutivas de um projetor telecêntrico.

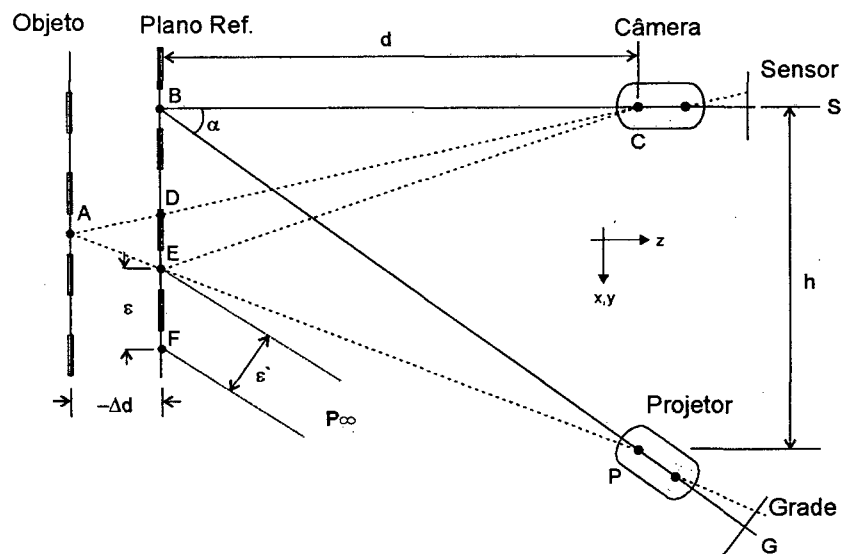


Figura 4.16 Relação entre fase e altura no objeto.

A altura " $\Delta d$ " pode ser obtida da relação existente entre os dois triângulos semelhantes  $\triangle ACP$  e  $\triangle ADE$ :

$$\frac{DE}{-\Delta d} = \frac{h}{d - \Delta d} \quad (4.35)$$

$$\Delta d = d \cdot \frac{DE}{DE - h} \quad (4.36)$$

O segmento "EF", representa a distância entre duas franjas consecutivas projetadas no plano de referência e aqui denominada de " $\varepsilon$ ". Apesar de não se utilizar um projetor telecêntrico [1], devido à distância " $\varepsilon$ " ser muitas vezes menor que " $d$ ", a relação entre o período das franjas " $\varepsilon$ " e a frequência espacial pode ser adotada para toda a área de medição, assim:

$$f_o = \frac{1}{\varepsilon} = \frac{\cos \alpha}{\varepsilon'} \quad (4.37)$$

Note que, a cada período de franja " $\varepsilon$ ", a fase " $\theta$ " varia de 0 a  $2\pi$ , podemos então calcular o sub-intervalo "DE" em função da fase " $\Delta\theta$ " através de uma proporcionalidade direta:

$$\frac{2 \cdot \pi}{\varepsilon} = \frac{\Delta\theta}{DE} \quad (4.38)$$

$$DE = \frac{1}{2} \cdot \Delta\theta \cdot \frac{\varepsilon}{\pi} \quad (4.39)$$

Substituindo-se "DE" na equação 4.36, obtém-se a altura " $\Delta d$ " em função da fase " $\Delta\theta$ ":

$$\Delta d = d \cdot \left[ -\Delta\theta \cdot \frac{\varepsilon}{(-\Delta\theta \cdot \varepsilon + 2 \cdot h \cdot \pi)} \right] \quad (4.40)$$

Substituindo-se " $\varepsilon$ " da equação 4.37 na equação 4.40, obtém-se a equação que representa as alturas do objeto em relação ao plano de referência:

$$\Delta d = d \cdot \left[ \frac{\Delta\theta}{(\Delta\theta - 2 \cdot h \cdot \pi \cdot fo)} \right] \quad (4.41)$$

Para pequenas variações de fase " $\Delta\theta$ " em relação à altura " $h$ " e à frequência " $fo$ " ( $2 \cdot \pi \cdot h \cdot fo$ ) pode-se aproximar a equação 4.41 por uma relação simplificada e linear dada por:

$$\Delta D = d \cdot \left[ \frac{-\Delta\theta}{2 \cdot h \cdot \pi \cdot fo} \right] \quad (4.42)$$

A equação 4.42 finalmente representa a relação entre as alturas " $\Delta D$ " e a fase " $\Delta\theta$ ".

#### **Máxima inclinação da forma do objeto**

A máxima inclinação ( $\delta h/\delta x$ ) que a forma do objeto pode ter, para que ainda tenha projetado sobre si as franjas, é dada pela equação 4.43 [Takeda e Mutoh, 1999] :

$$\left| \frac{d}{dx} \cdot \Delta d(x, y) \right| < \frac{1}{3} \cdot \left( \frac{d}{h} \right) \quad (4.43)$$

### **4.3 A Base de Faces Humanas**

Neste trabalho foi construída uma base de dados de faces humanas com o objetivo de permitir a avaliação dos métodos de classificação desenvolvidos. Apesar de existirem base de dados de faces humanas disponíveis na internet <sup>6, 7 e 8</sup> provenientes de outros grupos de pesquisa, estas bases de dados são constituídas de faces em duas dimensões e níveis de cinza, não servindo para nossos propósitos.

A estruturação da base de faces aqui construída obedeceu a critérios que (i) tentam contemplar o universo de possibilidades que podem ocorrer em condições normais de funcionamento do sistema, (ii) permitem sua utilização em processos de aprendizado de classificadores como os utilizados nas RNA's. É comum nestes processos reservar os casos mais extremos para treinamento das RNA's e os casos intermediários para teste.

<sup>6</sup> Manchester - Base de dados de faces humanas da Universidade de Manchester, Manchester, UK

<sup>7</sup> MIT - Base de dados de faces do Massachusetts Institut of Technology, MA, USA.

### 4.3.1 As Imagens de Face da Base

A base de dados de faces tridimensionais foi construída com um total de 52 indivíduos diferentes. Para cada indivíduo foram adquiridas 32 imagens com variações que podem ser classificadas em dois grandes grupos segundo expressões faciais e posicionamento espacial.

#### Expressões Faciais

As imagens desta categoria, num total de 12 por indivíduo, se distribuem da seguinte maneira:

(i) Olhos abertos, (ii) Olhos Semi-abertos, (iii) Olhos Fechados, (iv) Sobrancelhas Altas, (v) Sobrancelhas Baixas, (vi) Sobrancelhas Relaxadas, (vii) Boca Fechada, (viii) Boca Semi-aberta, (ix) Boca Aberta, (x) Bochecha Inflada, (xi) Bochecha Sugada e (xii) Óculos.

A figura 4.17 é uma amostra do tipo de imagens utilizadas na base de faces.

#### Posicionamento Espacial:

As imagens desta categoria se distribuem segundo os seguintes critérios:

- i. Rotação em torno do eixo X : +5°, +10°, 0°, -5° e -10°;
- ii. Rotação em torno do eixo Y \* : +5°, +10°, 0°, -5° e -10°;
- iii. Rotação em torno do eixo Z \* : +10°, +20°, 0°, -10° e -20°;
- iv. Escalonamento em torno do eixo Z : +10%, +5%, 0%, -5% e -10%.

Na figura 4.18 pode-se ver esquematicamente a distribuição espacial das faces na base de faces.

---

<sup>8</sup> FERET - Base de dados de faces do NIST, USA

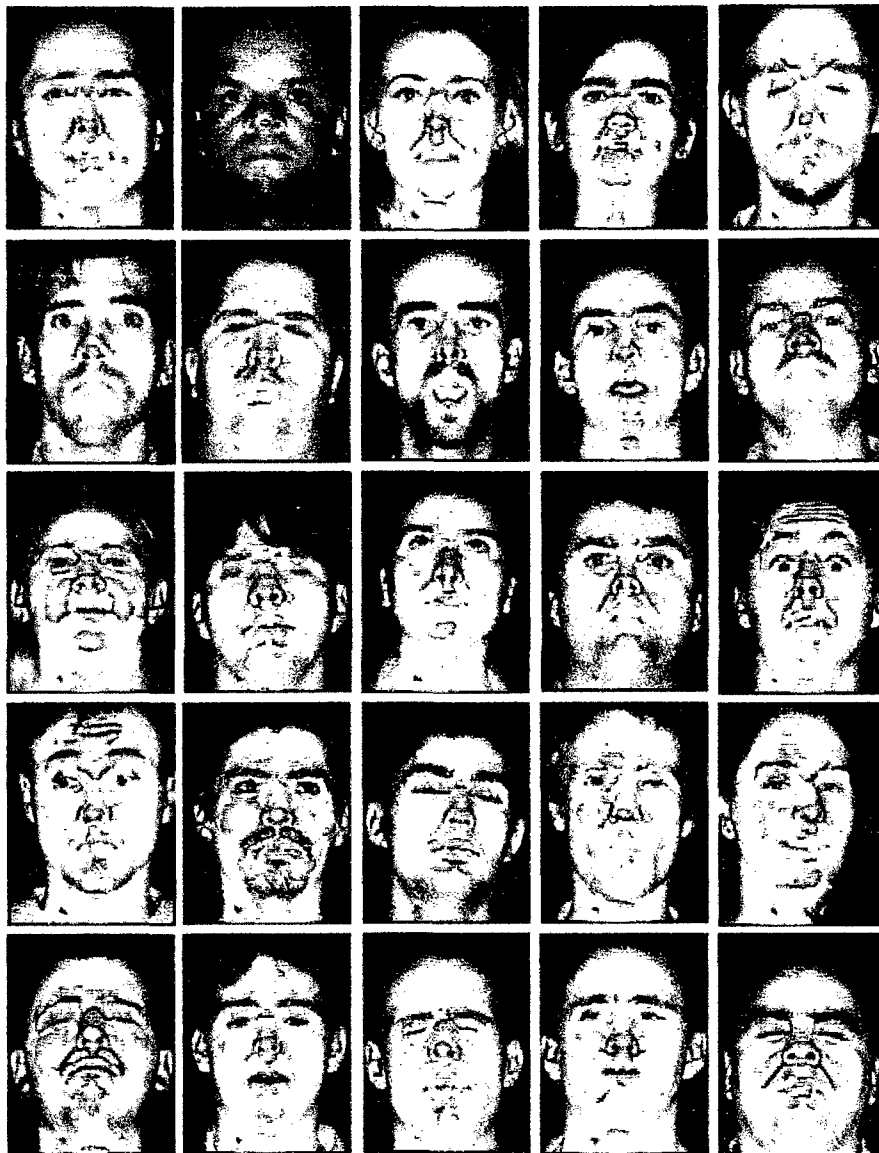


Figura 4.17 Amostra de imagens típicas da base de faces.

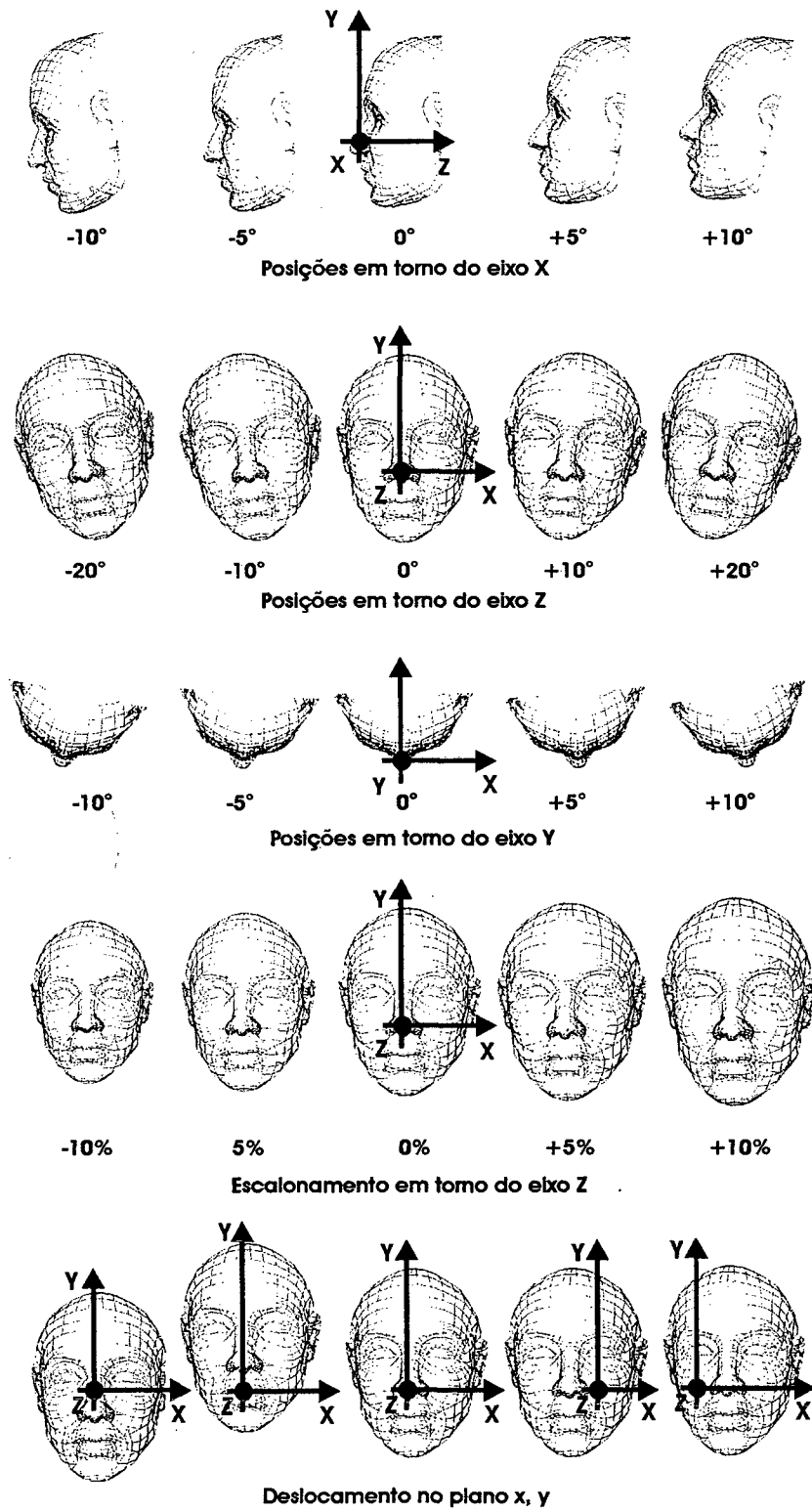


Figura 4.18 Posicionamento espacial das faces no banco de faces.

### 4.3.2 As Formas de Faces da Base

Processando-se as imagens dos 52 indivíduos, com 32 variações cada, obtiveram-se as matrizes de formas das faces. Estas formas, foram ainda processadas matematicamente para se obter resoluções, níveis de ruído e filtragem diferentes.

#### Adicionando Ruído e Filtrando as Formas de Faces

Com o objetivo de aumentar a similaridade com possíveis situações de um caso real da base de dados, foram adicionados matematicamente dois níveis de ruído randômico e dois níveis de filtragem para cada uma das formas originais da base de dados.

#### A Resolução da Matriz de Formas

Em se tratando de um processo experimental e não se conhecendo de antemão nenhuma informação de qual a resolução da matriz de formas é suficiente para se conseguir a classificação desejada, processaram-se dados com 9 resoluções diferentes: (i) 21 x 15, (ii) 31 x 22, (iii) 41 x 29, (iv) 51 x 36, (v) 61 x 43, (vi) 71 x 50, (vii) 81 x 57, (viii) 91 x 64 e (ix) 101 x 71.

### 4.3.3 O Tamanho da Base de Dados

A tabela 4.2 resume a distribuição das imagens e formas de faces da base construída.

Tabela 4.2 Composição da base de faces desenvolvida.

Tipos de variações	Quantidade
Expressões faciais	+ 12
Posicionamento geométrico	+ 20
Níveis de ruído e filtragem	+ 5
Resoluções	+ 9
Indivíduos	+ 52
Total de formas de faces	74.880



#### 4.3.4 A População da Base

A base de faces construída foi ainda distribuída segundo critérios de etnia, faixa etária e sexo dos indivíduos, conforme pode ser visto nas tabelas 4.3, 4.4 e 4.5.

Tabela 4.3 Classificação dos indivíduos segundo seus continentes/paises de origem.

Origem	Distribuição	
Africa	1	2%
Alemanha	8	15%
Áustria	0.5	1%
Bolívia	0.5	1%
Brasil	7	13%
Espanha	2	4%
França	0.5	1%
Itália	15	29%
Japão	1	2%
Líbano	0.5	1%
Paraguai	0.5	1%
Polônia	1	2%
Portugal	14	27%
Rússia	0.5	1%
Total	52	100%

Os dados étnicos da tabela 4.3 são referentes aos pais (pai + mãe) do indivíduo. Isto explica a quantidade absoluta com valor "0.5" que aparece na tabela 4.3 quando se tem um só indivíduo com aquela ascendência.

Tabela 4.4 Distribuição dos indivíduos da base segundo sua faixa etária.

Faixa etária	Distribuição	
15..19	4	7%
20..24	26	45%
25..29	9	16%
30..34	9	16%
35..39	5	9%
40..44	4	7%
Total	52	100%

Tabela 4.5 Distribuição dos indivíduos da base segundo seu sexo.

Sexo	Distribuição	
f	12	23%
m	45	87%
Total	52	100%

Ainda, 9 dos 52 indivíduos da base possuem barba ou cavanhaque.

## Capítulo 5

# Reconhecimento por Raciocínio Baseado em Casos - RBC

---

*Eu tenho uma lâmpada pela qual meus passos são guiados e esta é a lâmpada da experiência. Eu não conheço maneira de julgar o futuro a não ser pelo passado.*

- Patrick Henry, (1736-1799).

**N**a natureza, especialistas humanos aprendem e ensinam com as experiências, explicam novas regras com casos, exemplificam com base em experiências passadas, extrapolam a partir de casos existentes, mencionam experiências passadas como forma de persuasão e ensaiam novas estratégias. Sem se darem conta, na realidade o que eles estão fazendo é utilizar a memória para a solução dos problemas. Raciocínio Baseado em Casos - RBC é um tipo de raciocínio a que se é levado quando os dados acerca de um problema são exemplos resolvidos do mesmo.

## 5.1 O Raciocínio Baseado em Casos - RBC

O RBC nasceu com Roger C. Schank em 1982, época em que os conexionistas ainda estavam adormecidos devido à retração provocada pelo livro de Minski e Papert, "Os Perceptrons" de 1969. Assim nada mais natural que ele tenha desenvolvido suas técnicas esquecendo da existência das redes neurais e que elas podem ser treinadas por exemplos, sendo portanto implementadoras de raciocínio baseado em casos. Apesar disto neste trabalho será considerado somente a abordagem simbólica do RBC. O RBC utiliza analogias com casos semelhantes para a solução de problemas. A solução mais próxima adotada no passado é reutilizada para a solução do caso atual. *"É mais fácil ver um especialista resolvendo um problema do que fazê-lo explicar quais as regras de raciocínio que o guia"* [Barreto, 2001].

O uso do RBC é particularmente indicado quando se dispõe de uma grande e bem estruturada base de dados de casos resolvidos. RBC's tornam-se robustos na solução de problemas se possuírem uma boa coleção de exemplos para comparação dos casos semelhantes. RBC's não requerem nenhum modelo específico de domínio, assim a resolução de problemas toma-se uma questão de acúmulo de casos de sucesso em torno de um determinado assunto. A implementação é reduzida à identificação de características<sup>1</sup> que melhor descrevem o caso. O acesso às informações pode ser gerenciado por técnicas de indexação como as utilizadas em bases de dados.

As relações entre o espaço de problemas e o de soluções no RBC podem ser representadas através do diagrama apresentado na figura 5.3, onde para cada caso presente no espaço de estado dos problemas existe uma solução mapeada no espaço de estado das soluções. Diz-se que estes casos ou exemplos fazem parte da base de conhecimento do sistema.

A solução de um novo problema se dá através da associação deste ao seu problema vizinho mais próximo avaliado através de critérios estabelecidos por uma função de similitude que envolve dados de ambos os casos. Assim a solução adotada para o caso anterior mais próximo será também a solução adotada para o novo problema no espaço de soluções.

---

<sup>1</sup> "Features".

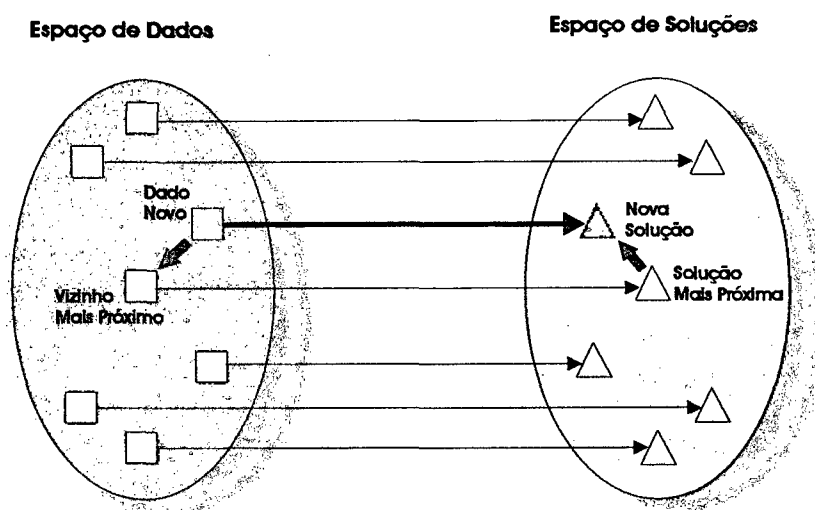


Figura 5.1 Representação da relação problema/solução em RBC.

Segundo Aamodt & Plaza [Aamodt e Plaza, 2002] o RBC pode ser definido no seu mais alto nível de generalização como um ciclo que engloba as seguintes etapas:

- i. **Recuperação** do caso mais semelhante.
- ii. **Reutilização** da informação e conhecimento naquele caso para resolver o problema.
- iii. **Revisão** da solução proposta.
- iv. **Retenção** de partes da experiência adquirida para ser útil em solução de problemas futuros.

Na figura 5.2 pode-se ver o ciclo típico de um sistema de RBC.

Nela um caso é definido como a entidade que contém lições passadas e o contexto de onde a lição pode ser usada. Ele é uma porção de conhecimento contextualizado representando uma experiência [Alterman, 1989]. O caso na sua forma geral contém: o problema, a solução e a consequência. Este ciclo, raramente, ocorre sem a intervenção humana pois a maioria dos sistemas (ferramentas) baseadas em RBC atua, primariamente, somente como recuperadores e reutilizadores de casos. Etapas como revisão e retenção geralmente são assumidas pelos gerentes (humanos) da base de dados.

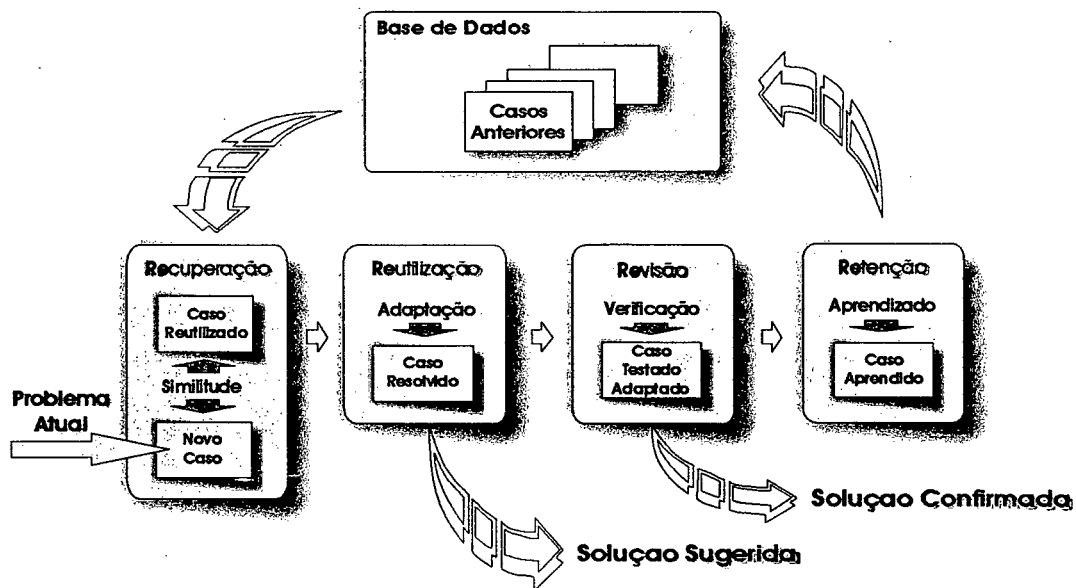


Figura 5.2 Ciclo dos 4 RE's do RBC [Aamodt e Plaza, 2002].

### 5.1.1 O Processo de Recuperação

Uma vez descrito o problema, o processo de recuperação faz uma busca nos casos com maior similitude no banco de casos. Para que esta busca seja eficaz é necessário que haja uma eficiente indexação dos casos. São características desejáveis dos índices de uma indexação: (i) previsibilidade, (ii) generalidade para permitir seu uso em outros casos e (iii) objetividade para que este possa ser associado facilmente com algum caso no futuro. Tanto processos manuais como automáticos têm sido usados para a seleção de índices. Apesar da crença de Janet Kolodner [Kolodner, 1993] de que indivíduos tendem a escolher melhor os índices do que algoritmos, os métodos de indexação automáticos são objeto de constantes pesquisas e avanços.

O processo de recuperação envolve os seguintes processos distintos:

- i. Identificação das características através da indexação.
- ii. A busca do caso similar propriamente dita, a qual em geral é realizada por um modelo de busca serial, hierárquica ou paralela.
- iii. A avaliação de similitude entre os casos.
- iv. Seleção do caso mais próximo.

Ao contrário das buscas em bases de dados onde se tem um alvo bem determinado, a recuperação de dados em CBR deve usar heurísticas capazes de fazer um casamento parcial,

visto que em geral um novo caso não obrigatoriamente será completamente semelhante ao anterior. O resultado deste processo pode ser a seleção de mais de um caso que são suficientemente próximos ao novo caso. Na seção 5.2 são abordados com detalhes os mecanismos de avaliação de similitude entre os casos.

### **5.1.2 O Processo de Reutilização**

Nesta etapa, os casos anteriores e atuais são combinados resultando em uma “solução sugerida”. A reutilização volta-se para as diferenças entre o caso anterior e o novo caso, e que partes do caso antigo poderão ser transferidas para o novo caso. Este processo envolve a cópia das características semelhantes e a adaptação do novo caso.

### **5.1.3 O Processo de Revisão**

No processo de revisão a solução sugerida é avaliada através de problemas reais e/ou analisada por um “professor”, podendo sofrer correções se seu desempenho se não atingir os níveis preestabelecidos. Em caso de sucesso o resultado é apresentado como “solução confirmada”. Este processo é composto de duas etapas:

- i. Avaliação da solução adotada com base na reutilização do caso anterior.
- ii. Adaptação da solução do caso se a avaliação da solução não satisfizer. Uma vez que um caso é recuperado, o sistema de RBC deve adaptar o caso escolhido para as necessidades do corrente problema. O processo de adaptação considera as diferenças entre os dois casos e aplica regras ou fórmulas que consideram estas diferenças junto ao novo caso.

Observa-se que quando a solução de um problema obtida através da reutilização de um caso não é satisfatória abre-se uma oportunidade para aprender com os erros.

### **5.1.4 O Processo de Retenção**

Finalmente o processo do RBC se completa pela retenção do conhecimento gerado no ciclo e inclusão destas informações na base de dados para uso futuro. O armazenamento dos casos é um fator importante para o desempenho de um sistema de RBC. A base de dados deve ser estruturada de forma a permitir uma eficiente busca durante o processo de recuperação de casos.

## 5.2 Cálculo da Similitude

Uma das condições necessárias para o uso do RBC é que o espaço de dados esteja em um espaço métrico. A escolha de um caso semelhante para a solução de problemas exige o uso de alguma ferramenta que diga quão próximo a suposta solução está do caso sob análise. A medida de proximidade entre um caso  $a$  e um caso  $b$  é chamada de medida de similitude  $S(a, b)$ . Esta função associa a todo par de dados do conjunto  $U$  um valor real não negativo tal que  $d(a, b): U \rightarrow \mathbb{R}^+$ . Assim  $\forall a, b \in U$  as propriedades da similitude são definidas como:

- i. Assertividade:  $d(a, b) = 0$  se  $a = b$ .
- ii. Simetria:  $d(a, b) = d(b, a)$ .
- iii. Desigualdade triangular:  $d(a, c) \leq d(a, b) + d(b, c)$ .

Dado um conjunto de pontos pode-se definir várias funções distância formando vários espaços métricos diferentes. Assim podem ser usadas na avaliação de similitudes entre casos as funções que calculam as seguintes distâncias: (i) Distância Euclidiana, (ii) Distância de Hamming, (iii) Distância de Mahalanobis, (iv) Distância de Minkowski, (v) Distância de Camberra e outras. Se a distância Euclideana é usada tem-se o espaço Euclidiano e é neste espaço que são válidos os teoremas apresentados no célebre livro "Os Elementos" de Euclides. Este espaço tem grande motivação para muitos por coincidir com a geometria a que estamos acostumados em nosso dia corrente. A de Minkewski define o espaço de mesmo nome e é pouco intuitivo por ser o da teoria da relatividade. O de Hamming é também pouco intuitivo mas em geral implica num mínimo de cálculos acelerando bastante algoritmos de busca em CBR simbólico. A escolha da função de distância é importante, pois cada uma delas possui características particulares que serão refletidas nos resultados do cálculo [Duda et al., 2000]. As duas mais conhecidas são a distância de Hamming ou Manhattan e a distância Euclidiana.

### 5.2.1 A Distância Euclidiana

É adequada para o tratamento de classes cujos elementos tendem a se agrupar em torno da média, ou seja, com pequena variância, é o caso de sistemas de reconhecimento de caracteres impressos em cheques [Kimura e Shridhar, 1999] e [Shridhar e Badreldin, 1999]. O cálculo da distância Euclidiana é indicado também para problemas onde as classes apresentam comportamento semelhante quanto à forma da função de distribuição de probabilidades e valores de variância [Machado, 1994]. A função da distância Euclidiana pode ser vista na equação 5.1:



$$dE = \sqrt{\sum (a_i - b_i)^2} \quad (5.1)$$

Onde:

$dE$  = distância Euclidiana;

$a_i$  = Dados da forma de face de referência;

$b_i$  = Dados da forma de face sob teste;

$i$  =  $i$ -ésimo termo dos conjuntos de dados.

## 5.2.2 A Distância de Hamming

A distância de Hamming utiliza somente a soma dos módulos da diferença ao invés da raiz quadrada da soma quadrática das distâncias como na Euclidiana. Isto lhe confere um menor custo computacional [Barreto, 2001], aliado à vantagem de que os valores extremos (outliers) não são amplificados como no caso da distância Euclidiana onde estas diferenças são elevadas ao quadrado.

A distância de Hamming é dada por:

$$dH(a, b) = \sum_i |a_i - b_i| \quad (5.2)$$

Onde:

$dH$  = distância de Hamming;

$a_i$  = dados da forma de face de referência;

$b_i$  = dados da forma de face sob teste;

$i$  =  $i$ -ésimo termo dos conjuntos de dados.

A figura 5.3 mostra esquematicamente como são calculadas as distâncias entre duas formas de faces.

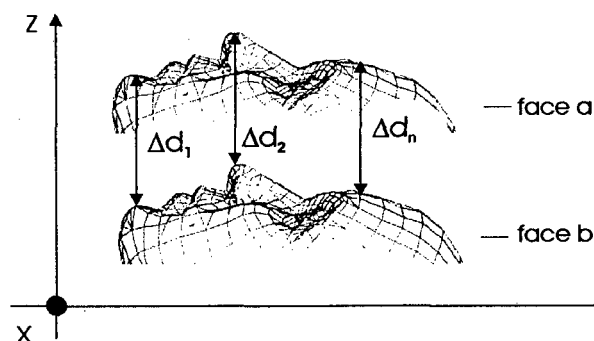


Figura 5.3 Visualização do cálculo de distâncias entre duas formas de faces.

### 5.2.3 A Similitude e o Limiar de Classificação

É esperado que as similitudes calculadas para um grupo de faces de um mesmo indivíduo e suas variações (conjunto de ovelhas) devam possuir um valor próximo a zero em uma escala de 0 a 1. Isto se deve repetir para todos os indivíduos e suas variações. Já quando calculada para indivíduos diferentes (conjunto de lobos) espera-se que o valor da similitude seja próximo a um.

Um cálculo prévio feito em uma parte da população chamada de conjunto de treinamento fornece parâmetros para a escolha do valor do limiar de classificação.

## 5.3 Otimização no Cálculo das Distâncias

Com o objetivo de diminuir a incerteza dos resultados, em se tratando de distâncias, existe a necessidade de se fazer um tratamento prévio dos dados das formas de faces antes de se proceder ao cálculo das distâncias propriamente dito. Neste sentido, deve-se fazer um pré-processamento dos dados buscando-se calcular somente as diferenças efetivas inerentes às formas das faces e minimizados demais erros envolvidos. Dentre os pré-processamentos para minimização de erros, os mais significativos são: (i) normalização espacial, (ii) compensação dos erros da normalização espacial.

### 5.3.1 Normalização Espacial

Cada novo processamento de uma imagem facial, mesmo que de um mesmo indivíduo, traz consigo diferenças devido aos próprios erros de cálculo envolvidos, mas principalmente devido ao desalinhamento espacial da face pelo próprio indivíduo. Como há interesse em interagir o mínimo possível com este indivíduo durante o processo de aquisição da imagem, a

compensação destes erros de posicionamento espacial é feita ao final do processo de extração da forma. Esta normalização espacial tem como objetivo posicionar todas as faces segundo um mesmo referencial espacial de forma que, ao se calcular as distâncias entre elas, se esteja somente contabilizando as diferenças efetivas entre as faces dos indivíduos. A normalização, como proposta neste trabalho, prevê a compensação de seis graus de liberdade das formas faciais. O processo de normalização pode ser dividido nas seguintes etapas:

- i. Estabelecimento de um ponto de referência: Este ponto de referência deve ser o mais robusto possível devendo preferencialmente ser alguma característica geométrica marcante, comum a todas as faces humanas e situado em uma região onde se tenha uma boa qualidade na forma da face. Ponto máximo do nariz, a linha vertical deste, o vale ao longo desta linha e a linha entre os olhos, são algumas das possibilidades.
- ii. Translação no plano x, y: Translação do plano x,y de modo a levar o ponto de referência da face ao ponto de referência do sistema.
- iii. Rotação no eixo z: Rotação em torno do eixo z até que haja coincidência da face com o sistema de referência.
- iv. Escalonamento no plano x, y: Escalonamento na forma da face sobre o plano x, y de modo que as formas adquiram um mesmo tamanho.
- v. Rotação no eixo y: Rotação em torno do eixo y até que haja coincidência da face com o sistema de referência.
- vi. Rotação no eixo x: Rotação em torno do eixo x até que haja coincidência da face com o sistema de referência.

A figura 5.4 mostra os movimentos durante o processo de normalização das formas das faces.

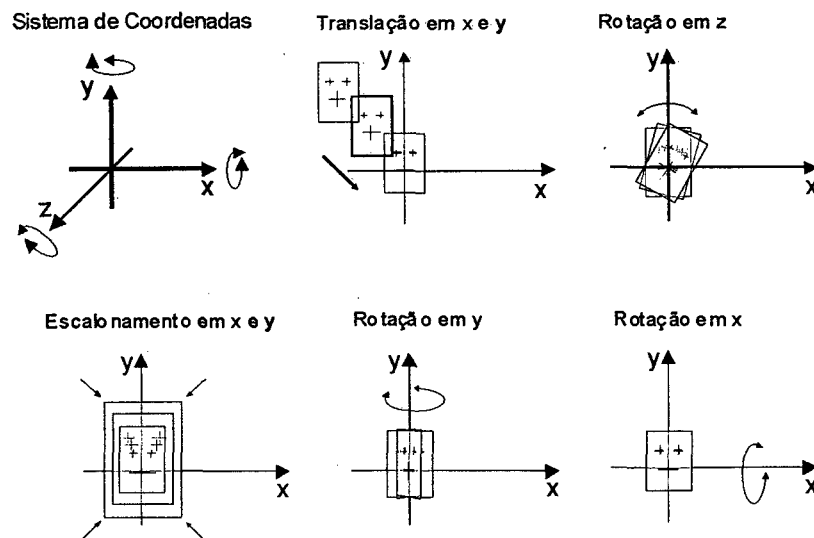


Figura 5.4 Etapas da normalização das formas das faces.

As equações que permitem a normalização espacial descrita acima são apresentadas a seguir [Foley et al., 1996] e [Spiegel, 1974]:

Os pontos do conjunto de dados no sistema original de coordenadas são dados por:

$$P = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (5.3)$$

Os pontos do conjunto de dados no novo sistema de coordenadas são dados por:

$$P' = \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} \quad (5.4)$$

A translação do conjunto de pontos é dada por:

$$P' = T \cdot P \quad (5.5)$$

$$T = \begin{bmatrix} 1 & 0 & 0 & dx \\ 0 & 1 & 0 & dy \\ 0 & 0 & 1 & dz \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5.6)$$

Onde  $dx$ ,  $dy$  e  $dz$  são os deslocamentos desejados para cada eixo.

A rotação do conjunto de pontos é dada por:

$$P' = R \cdot P \quad (5.7)$$

$$R_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \rho & -\sin \rho & 0 \\ 0 & \sin \rho & \cos \rho & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5.8)$$

$$R_y = \begin{bmatrix} \cos \rho & 0 & \sin \rho & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \rho & 0 & \cos \rho & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5.9)$$

$$R_x = \begin{bmatrix} \cos \rho & -\sin \rho & 0 & 0 \\ \sin \rho & \cos \rho & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5.10)$$

onde  $\rho$  é o ângulo de rotação desejado em torno dos eixos  $x$ ,  $y$  e  $z$ .

O escalonamento do conjunto de pontos é dado por:

$$P' = S \cdot P \quad (5.11)$$

$$S = \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5.12)$$

Onde  $s_x$ ,  $s_y$  e  $s_z$  são os fatores de escala para os eixos  $x$ ,  $y$  e  $z$ .

### 5.3.2 Minimização dos Erros de Posicionamento

Apesar do processo de normalização espacial, descrito na seção 5.3.1, ainda é possível melhorar mais os erros residuais devido à falhas no correto posicionamento das formas das faces. Isto é feito através de um ajuste espacial fino nas formas das faces. Estes erros aqui chamados de sistemáticos e aleatórios podem ser ainda minimizados da seguinte forma [Zimmermann et al., 2000a]:

- i. Erros Sistemáticos: Podem ser minimizados interpolando-se um plano que melhor se adapte (regra de mínimos quadrados bidimensional) sob a forma resultante do cálculo da distância de Hamming ponto a ponto. Este plano representa os possíveis erros sistemáticos residuais do cálculo das distâncias. A compensação deste erro é feita

através da diminuição dos valores deste plano e dos valores da distância inicialmente calculados.

- ii. Erros Aleatórios: Estes erros podem ser minimizados através da eliminação de valores atípicos da distância, p. ex. pelo critério de máximo desvio padrão permitido. Sob este critério um valor típico de desvio padrão máximo permitido pode ser o de  $3 \cdot \sigma_{n-1}$ , o que vem a garantir, para uma distribuição normal, que 99.73% dos dados se encontram dentro deste intervalo.

O método dos mínimos quadrados utilizados para a interpolação do plano dos erros sistemáticos é exposto a seguir:

A equação geral da superfície é dada por:

$$z = A_0 + A_1 \cdot x + A_2 \cdot y \tag{5.13}$$

Na forma matricial, tem-se:

$$\begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ 1 & x_n & y_n \end{bmatrix} \cdot \begin{Bmatrix} A_0 \\ A_1 \\ A_2 \end{Bmatrix} = \begin{Bmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ \vdots \\ z_n \end{Bmatrix} \tag{5.14}$$

Calculando-se os coeficientes:

$$[XY] \cdot \{A\} = \{Z\} \tag{5.15}$$

$$([XY]^T \cdot [XY]) \cdot \{A\} = [XY]^T \cdot \{Z\} \tag{5.16}$$

$$\{A\} = ([XY]^T \cdot [XY])^{-1} \cdot [XY]^T \cdot \{Z\} \tag{5.17}$$

Onde  $\{XY\}$  e  $\{Z\}$  são os conjuntos de pontos ordenados por onde deseja-se passar o plano e  $\{A\}$  é o vetor dos coeficientes  $A_0$ ,  $A_1$  e  $A_2$  do plano. Os erros em cada ponto são expressos por:

$$\{\varepsilon\} = [XY] \cdot \{A\} - \{Z\} \quad (5.18)$$

E o somatório dos erros quadráticos é dado por:

$$S_{eq} = \{\varepsilon\}^T \cdot \{\varepsilon\} \quad (5.19)$$

## 5.4 A Ponderação Facial

Com o intuito de maximizar o desempenho no cálculo de distâncias, é natural se pensar em atribuir pesos diferenciados às regiões da face. Regiões que tenham maior importância recebendo um peso maior e regiões com menor importância recebendo um peso menor.

Na natureza, os humanos, durante o processo de aprendizado de uma nova face, posicionam seus olhos com um tempo diferenciado para cada parte da face do indivíduo que está sendo memorizado. Algumas partes da face recebem mais atenção dependendo-se maior tempo nestas regiões e em outras menos.

### 5.4.1 Ponderação Comportamental

Estudos que analisam o movimento do olho de um indivíduo quando está reconhecendo outro indivíduo são descritos por Moghaddan e Pentland [Moghaddan e Pentland, 1994]. Uma clássica experiência neste sentido foi a desenvolvida por Minut e Henderson [Minut et al., 2002] e [Henderson et al., 2000], a qual monitorava a posição e o tempo que o olho humano despendia sob as regiões da face de um indivíduo desconhecido quando este era apresentado a um indivíduo que teria de memorizá-lo. O gráfico da figura 5.5 mostra o tempo médio relativo que um ser humano despende no processo de aprendizado de uma nova face.

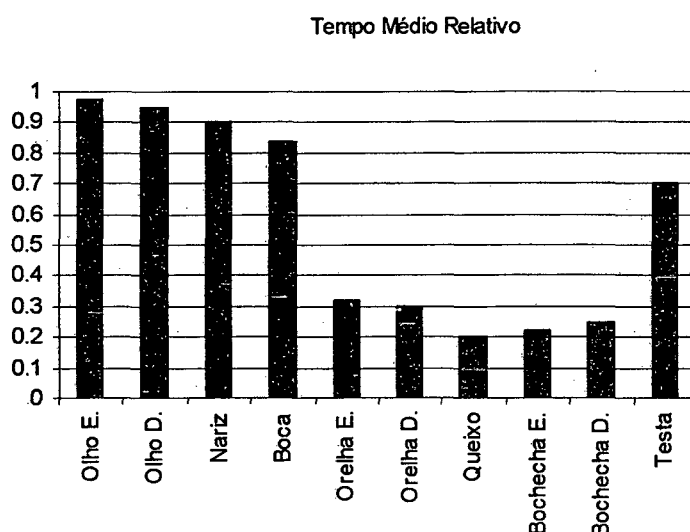


Figura 5.5 Tempo de permanência do olho sobre as regiões da face no aprendizado [Minut et al., 2002] e [Henderson et al., 2000].

## 5.4.2 Ponderação Estatística

Formas de faces de um mesmo indivíduo, apesar de possuírem diferenças devido a diferentes expressões faciais, possuem regiões que variam umas mais que as outras. Seria natural, num processo de classificação onde se busca robustez, priorizar as regiões onde se tem uma maior invariância de forma. Valorizando assim as regiões onde ocorrem pequenas variações em função das expressões faciais do indivíduo no dia a dia e depreciando regiões onde existam maiores variações.

### Matriz de Pesos Próprios

Este raciocínio leva a construção de uma matriz de pesos para cada indivíduo, neste trabalho, chamada de Matriz de Pesos Próprios - MPP. A MPP é obtida através do cálculo do desvio padrão associado com o conjunto de formas de faces de todas as variações de um mesmo indivíduo. Na figura 5.6 pode-se ver as MPP's para quatro indivíduos diferentes, onde as cores próximas ao vermelho representam regiões com menor variação de formas para um mesmo indivíduo ou regiões de maior peso. Já as regiões próximas à cor azul, representam regiões com maior variação de forma e que recebem um peso inferior.



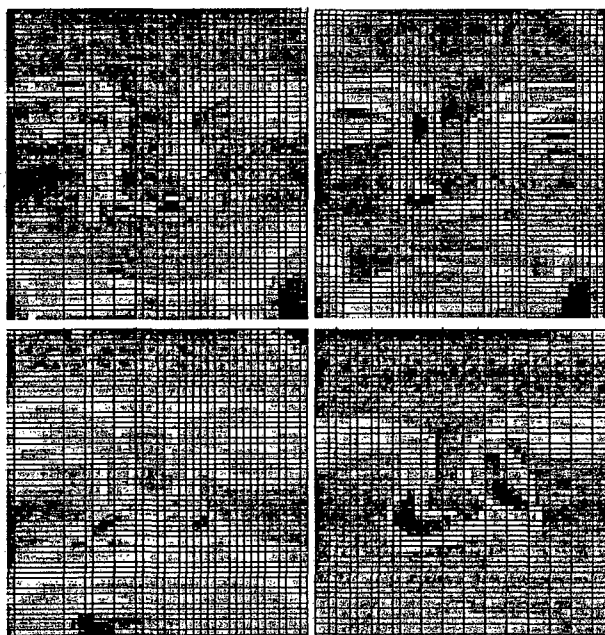


Figura 5.6 MPP de regiões faciais segundo o critério de ponderação estatística.

### **Matriz de Pesos Universais**

Estendendo o raciocínio aplicado à MPP, onde se prioriza as regiões de menor variação entre faces de um mesmo indivíduo, pode-se estabelecer uma matriz de pesos entre indivíduos diferentes, que valorize as regiões onde haja maior variação da forma. Por ser esta uma matriz composta de todos os indivíduos de uma base de faces esta é chamada de Matriz de Pesos Universais - MPU.

### **Combinando as Matrizes de Pesos**

A combinação das duas matrizes de pesos é possível usando-se algum critério que estabeleça uma relação, p. ex. linear, entre a MPP e a MPU, quando aplicada às formas de faces no processo de classificação.

### **5.4.3 Comparação Entre os Métodos**

Devido ao sucesso do reconhecimento de faces por humanos, seria de se esperar que o modelo de importância adotado por eles servisse como parâmetro para a ponderação de regiões faciais dentro do escopo deste trabalho. Alguns aspectos, no entanto devem ser observados antes de se prosseguir nesta linha de raciocínio. O principal deles diz respeito à

natureza da informação que se está processando. No âmbito do experimento de Henderson e Minut [Henderson et al., 2000, Minut et al., 2002], o reconhecimento das faces por humanos está sendo feito fundamentalmente por duas informações que são as cores e a forma da face. O experimento não separa a influência de cada uma das informações. É de se esperar portanto que existam diferenças entre a priorização das regiões faciais pelo método citado e o proposto neste trabalho que considera somente a forma da face como informação de reconhecimento. Na figura 5.7 pode-se ver a priorização de regiões faciais segundo os dois métodos.

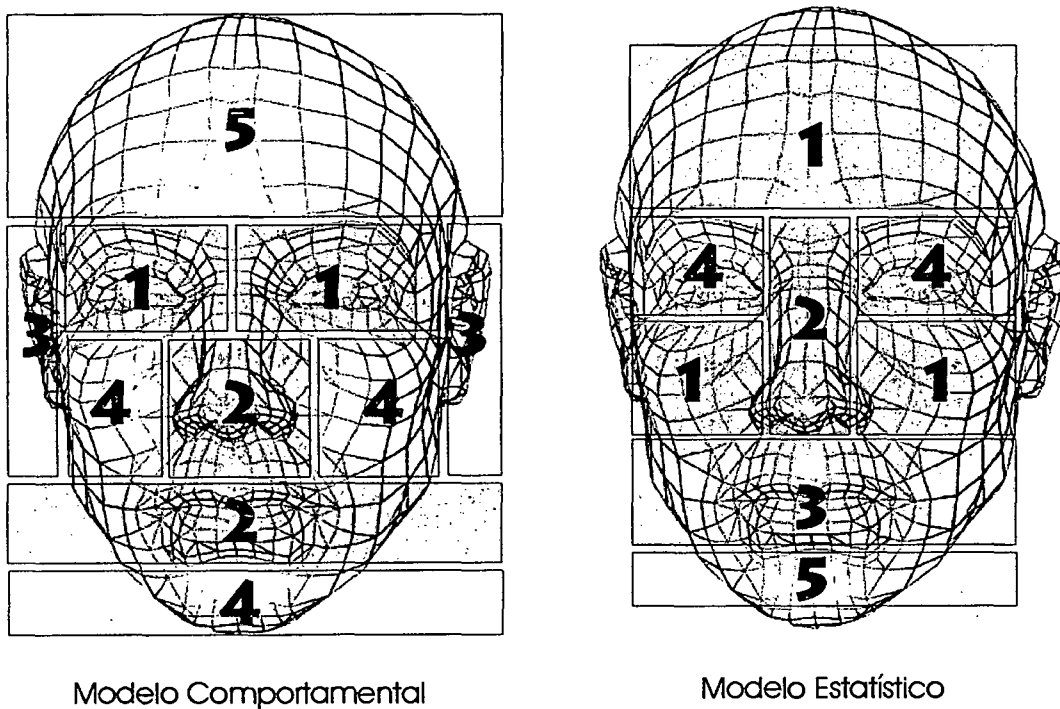


Figura 5.7 Priorização de regiões faciais segundo dois métodos distintos.

## Capítulo 6

# Abordagem Conexionista

---

*O cérebro, a obra-prima de criação, é quase um desconhecido para nós.*

- Nicolaus Steno, (1638-1686).

**E**ntender o funcionamento do cérebro humano sempre foi um desafio à comunidade científica. Sistemas conexionistas ou de redes neurais emergiram como um ramo da inteligência artificial cujo propósito principal era o de compreender o funcionamento do cérebro humano e a partir daí desenvolver uma estrutura artificial com as mesmas características de processamento, aprendizado e tomadas de decisão semelhantes à da estrutura biológica.

Redes Neurais Artificiais - RNA são de fácil implementação e têm sido aplicadas com sucesso nos mais diversos problemas. Dentre as principais áreas de aplicação, podemos citar: (i) Sistemas de Controle [Azevedo et al., 2000, Bottemanne et al., 1994, Henderson et al., 2000, Konen e Shulze-Krüger, 1995], (ii) reconhecimento de padrões [Bishop, 1995, Lippmann, 1989, Lippmann, 1999], (iii) aproximação de funções [Lawrence et al., 1996b].

Na seção 6.1 será apresentada uma breve revisão de RNAs, onde o leitor mais experimentado no assunto pode suprimir a leitura desta seção sem perda do idéia global do trabalho. Na seção 6.2 será abordado com mais detalhes o paradigma de RNA utilizado neste trabalho,

culminando com a abordagem, na seção 6.3, de tópicos importantes para o projeto de RNA com aplicação em reconhecimento de faces.

## 6.1 Introdução às Redes Neurais Artificiais

As RNAs vêm se tomando uma ferramenta cada vez mais poderosa vindo a dar uma nova perspectiva a IA. Desde seu aparecimento, muitas definições foram criadas com o intuito de explicar seu funcionamento e analogias ao sistema nervoso humano estão presentes em praticamente todas elas:

- i. As RNAs são técnicas inovadoras de processamento de informações, que têm como objetivo básico simular, em computadores, o funcionamento do sistema nervoso biológico [Nelson e Illingworth, 1991].
- ii. RNAs são sistemas de processamento de informações que possuem características de funcionamento similares às das redes neurais biológicas. Mais especificamente, generalizações de modelos matemáticos da biologia neural [Fausett, 1994].

Estas definições podem ser mais bem compreendidas pela análise da afirmativa de Haykin [Haykin, 1994]:

*"RNAs são processadores paralelos e distribuídos, que têm uma propensão natural para armazenar o conhecimento proveniente da experiência e torná-lo útil. Assim, assemelhando-se ao cérebro humano em dois aspectos:*

- i. No conhecimento adquirido pela rede através de um processo de aprendizado.*
- ii. Nas intensidades das conexões entre neurônios ou pesos sinápticos, onde se armazenam o conhecimento."*

Uma característica importante deste paradigma é o aprendizado. Ele é inerente às conexões dos neurônios, a qual ocorre similarmente na interconexão dos neurônios biológicos e não encontrados em técnicas de cálculo algoritmos convencionais.

### 6.1.1 Um Pouco da História das RNAs

Fisiologistas e filósofos ancestrais tinham diferentes opiniões sobre a função do cérebro humano. No século 4 AC Aristóteles considerava que o cérebro era um órgão secundário que serviria somente como refrigerador do sangue bombeado pelo coração. O cérebro humano continua sendo desvendado, mas as dúvidas sobre seu funcionamento residem hoje em uma escala muito menor.

O primeiro passo voltado para um modelo de neurônio artificial veio em 1943 com Warren McCulloch, neurofisiologista, e Walter Pitts, matemático, os quais escreveram um artigo sobre como os neurônios trabalham e modelaram uma rede neural usando circuitos elétricos. Reforçando este conceito, Donald Hebb escreveu um livro intitulado "Organization of Behavior" em 1949. Neste livro, Hebb sugeria que haveria um reforço das conexões de uma rede neural cada vez que esta fosse usada.

Com a modernização dos computadores na década de 50, Nathaniel Rochester da IBM Research Labs, simulou uma rede neural em computador. Ainda nessa década, mais precisamente em 1956, o "Dartmouth Summer Research Project" deu força à inteligência artificial e às redes neurais.

Nos anos seguintes, John von Neumann sugeriu imitar um neurônio usando relés ou válvulas. Também Frank Rosenblatt, um neurobiólogo, iniciou seus trabalhos com o "Perceptron", que foi uma das mais antigas redes neurais artificiais e ainda em uso hoje em dia.

Em 1959, Bernard Widrow e Marcian Hoff, de Stanford, desenvolveram o que eles chamaram de ADALINE e MADALINE. Estes nomes se originaram de "Multiple ADaptive LINear Elements". MADALINE foi a primeira rede neural a ser aplicada em um trabalho real, como um filtro adaptativo que cancela eco em linhas telefônicas. Este paradigma, ainda hoje, continua em uso comercial.

Infelizmente, em 1969, Marvin Minsky e Seymour Papert em seu livro intitulado "Perceptrons" desestimularam o uso do Perceptron alegando ser este limitado. Este ato desencorajou o uso e os investimentos em projetos com redes neurais a partir de 1970, tomando esta década conhecida como o período obscuro das redes neurais.

Somente em 1982 alguns eventos renovaram o interesse pelas redes neurais. John Hopfield apresentou um artigo para o "National Academy of Sciences" onde ele utilizava conexões bidirecionais ao invés das unidirecionais que se utilizavam antes. Também naquele mesmo ano Reilly and Cooper usaram uma rede híbrida com múltiplas camadas. Em 1982 também, em uma conferência conjunta entre EUA e Japão a "Cooperative/Competitive Neural Networks", o Japão anunciou investimentos em redes neurais, o que levou os EUA a acompanhá-lo.

Em 1985, o "American Institute of Physics" iniciou o que se tomaria um encontro anual chamado de "Neural Networks Computing". Em 1987, o IEEE, em sua "First Conference on Neural Networks", já contava com 1800 inscritos. A confiança no paradigma conexionista havia retornado.

Em 1986, com a evidência das RNAs de múltiplas camadas, o problema era estender a regra de Widrow-Hoff para estas redes. David Rumelhart, bem como outros grupos, dedicou-se a

este estudo, culminando com o mais famoso dos algoritmos de treinamento de redes o chamado "back-propagation".

Em 1989 ainda, no encontro intitulado "Neural Networks for Defence", Bernard Widrow cientificou à sua audiência que eles estavam engajados na quarta guerra mundial, referindo-se a uma guerra a ser travada na produção industrial e comércio mundial.

Hoje as RNAs estão sendo amplamente empregadas tendo sido inclusive incorporadas a produtos de alto consumo, através de chips dedicados<sup>1</sup>. A idéia fundamental atrás das RNAs continua sendo a de que se elas funcionam na natureza serão capazes de funcionar em computadores. Um histórico detalhado das RNAs pode ser encontrado em [Haykin, 1994], [Anderson e McNeill, 1992], [Barreto, 2001] e [Wasserman, 1993].

### 6.1.2 Modelo do Neurônio Artificial

O cérebro humano é constituído de uma vasta rede de elementos de computação chamados de neurônios. Em um cérebro humano típico estima-se existirem 100 bilhões de células de diversos tipos. O neurônio é um tipo especial de célula que conduz sinais elétricos. Existem cerca de 10 bilhões neurônios no cérebro humano. Os restantes 90 bilhões são chamadas de "glial" ou células de acoplamento, que funcionam como células de apoio aos neurônios, aprimorando a transmissão de impulsos elétricos, fortalecendo as conexões sinápticas [Bose e Liang, 1996].

Em termos biológicos, o fortalecimento do acoplamento dos neurônios é desejável, uma vez que, problemas como o mal de Alzheimer e perdas de memória são atribuídos a uma deficiência destes acoplamentos. Embora as funções das células "gliais" se mostrem importantes no processamento das informações cerebrais, até o momento não se tem uma modelagem matemática para elas.

Apesar dos neurônios biológicos se apresentarem de diversas formas, pode-se generalizar seu modelo conforme o apresentado na figura 6.1, onde podem ser vistos o corpo celular, as conexões de entrada ou dendritos, as conexões de saída ou axônios e as conexões sinápticas que conectam os neurônios através dos seus respectivos dendritos e axônios.

Segundo Haykin [Haykin, 1994] *"Um neurônio é a unidade fundamental de processamento de informação de uma rede neural"*. Mcculloch e Pitts, na década de 50, desenvolveram um modelo para o neurônio biológico, utilizado até hoje nos modelos de redes neurais [Haykin, 1994], [Barreto, 1999] e [Kovacs, 1996].

---

<sup>1</sup> "ASIC - Application Specific Integrated Circuit".

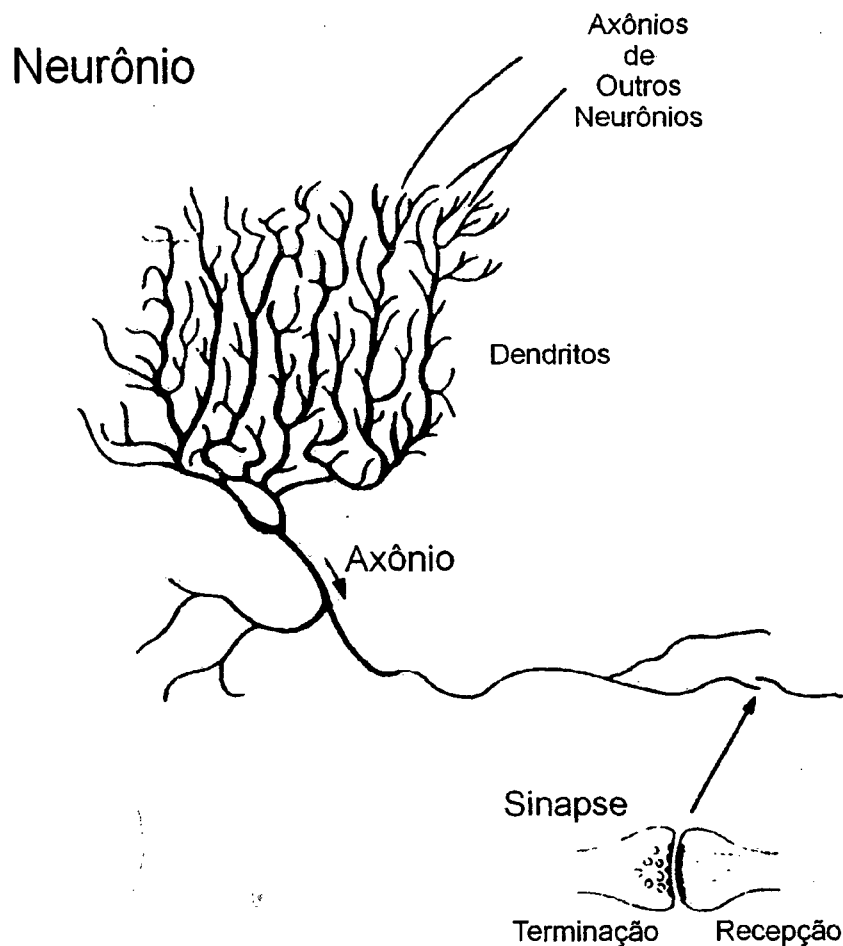


Figura 6.1 Representação de um neurônio biológico [Zurada, 1992].

O modelo do neurônio de McCulloch e Pitts pode ser visto na figura 6.2. São parâmetros importantes dos modelos de neurônios artificiais:

- i. As entradas: Representadas pelos valores  $X_1, \dots, X_n$ .
- ii. A polarização: É um valor que não varia com as entradas, usualmente de valor constante, tendo a função de adequar a saída do neurônio aos níveis desejados.
- iii. Os pesos: Estabelecidos para cada uma das entradas, são os valores que as multiplicarão escalarmente, ponderando a influência de cada uma delas no processo.
- iv. A confluência: Representa uma função matemática genérica  $u_k$  que recebe todas as entradas multiplicadas pelos pesos, usualmente representada pela função soma, conforme pode ser visto na equação 6.1:

$$u_k = \sum_{m=1}^n w_{km} x_m \tag{6.1}$$

Onde  $u_k$  representa o k-ésimo neurônio,  $x_1, x_2, \dots, x_n$  são as entradas,  $w_{k1}, w_{k2}, \dots, w_{kn}$  são os pesos sinápticos. Num sentido mais amplo, a confluência pode assumir a função de qualquer operador matemático [Brasil et al., 1999], [Barreto, 1999] e [Zimmermann et al., 1998].

- v. A função de ativação: O resultado do ponto de confluência é submetido a uma função de ativação  $\varphi(\cdot)$ , conforme mostrada na equação 6.2. Esta função confere as características da saída do neurônio artificial, que no modelo de Mcculloch e Pitts, em particular, é uma saída binária. Na seção 6.1.3 são apresentados os principais tipos de função de ativação e suas implicações nos valores de saída do neurônio artificial

$$y_k = u_k \cdot \varphi(\cdot) \tag{6.2}$$

- vi. O limiar de atuação: O parâmetro  $\theta_k$  representa o limiar de atuação da função de ativação. Normalmente é utilizado como um valor abaixo do qual a saída do neurônio é nula. Uma função semelhante pode ser conseguida através da entrada “ $x_0$ ” com um peso referente a uma entrada de valor constante igual a 1 denominada de polarização.

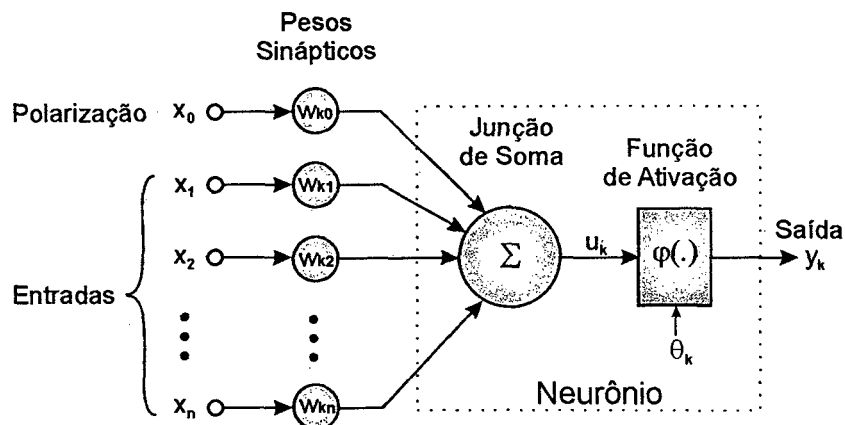


Figura 6.2 Modelo de Mcculloch e Pitts para o neurônio artificial [Haykin, 1994].



### 6.1.3 Funções de Ativação

A função de ativação  $\varphi(\cdot)$  ou função de transferência, figura 6.3, formata a amplitude da saída do neurônio, adota-se geralmente intervalos normalizados onde  $L_{y+}$  e  $L_{y-}$ , assumem valores 0 e 1, mas podem assumir qualquer valor dependendo da excursão esperada do sinal de saída.

O modelo do neurônio de McCulloch e Pitts p. ex. pode ser generalizado através da utilização de uma função de ativação contínua do tipo rampa ou tipo sigmoidal. Também outras funções que não a binária podem ser usadas como função de ativação  $\varphi(\cdot)$ . Na figura 6.3, pode-se ver 3 principais classes de funções de ativação:

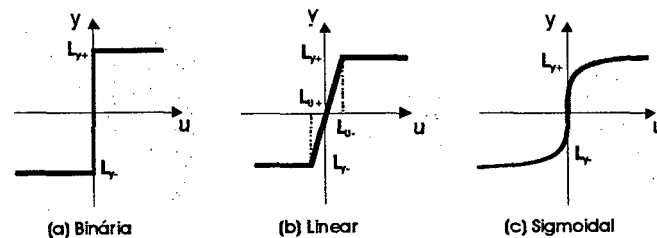


Figura 6.3 Principais funções de ativação ou de transferência.

- i. Função Binária: Utilizada quando se deseja obter apenas dois valores para os resultados, exemplo típico de utilização em classificação de dados. Para esta função, conforme ilustrada na figura 6.3a, tem-se:

$$y = \begin{cases} L_{y+}, & \text{se } u > 0 \\ L_{y-}, & \text{se } u \leq 0 \end{cases} \quad (6.3)$$

- ii. Função Linear: Empregada quando se deseja uma resposta linear dentro de uma faixa restrita de entrada  $[L_{u+}, L_{u-}]$  e de saída  $[L_{y+}, L_{y-}]$ . Para esta função linear por partes, ilustrada na figura 6.3b, tem-se:

$$y = \begin{cases} L_{y-}, & \text{se } u \leq L_{u-} \\ u, & \text{se } L_{u-} < u < L_{u+} \\ L_{y+}, & \text{se } u \geq L_{u+} \end{cases} \quad (6.4)$$

- iii. **Função Sigmoidal:** A função sigmoidal, figura 6.3c, é a função de ativação mais utilizada em redes neurais artificiais. Definida como uma função monotonicamente crescente apresenta propriedades assintóticas e de suavidade. Um exemplo de função sigmoidal é a chamada função logística, definida por:

$$y(u) = \frac{1}{1 + e^{-u}} \quad (6.5)$$

onde  $L$  é o parâmetro de inclinação da função sigmoidal. Os valores de  $L_{u+}$  e  $L_{u-}$  podem assumir valores arbitrários, mas usualmente  $[1,0]$ , bem como os valores  $L_{u+}$  e  $L_{u-}$  podem assumir quaisquer valores dependendo da faixa de resposta desejada, mas usualmente assumindo os valores  $[1,0]$ . Em particular se a faixa de  $L_{u+}$  e  $L_{u-}$  varia de  $[1,-1]$ , utiliza-se a função tangente hiperbólica, conforme a equação 6.6:

$$y(u) = \frac{1 - e^{-u}}{1 + e^{-u}} \quad (6.6)$$

#### 6.1.4 Arquitetura das RNAs

A arquitetura de uma RNA é caracterizada pela forma com que os neurônios se arranjam. Grafos são uma boa representação das redes para análise de sua topologia [Barreto, 2001]. Dentre outros parâmetros costuma-se mais freqüentemente classificar a arquitetura de redes segundo:

- i. Número de camadas intermediárias: Número de camadas entre a camada de entrada e a de saída.
- ii. Número de neurônios nas camadas intermediárias.
- iii. Número de neurônios de entrada: Número de neurônios na camada de entrada.
- iv. Número de neurônios de saída: Número de neurônios na camada de saída.
- v. Sentido de propagação dos sinais, se direta<sup>2</sup> ou recorrente<sup>3</sup>: Redes recorrentes, ou com retro-alimentação, são redes onde o sinal da camada de neurônios de saída pode ser utilizado como uma nova entrada para a rede, até que um equilíbrio seja atingido, neste

<sup>2</sup> Direta ou "feed-forward".

<sup>3</sup> Recorrente ou "Reccurent".

caso, o sinal de saída da camada de neurônios pode ser utilizado como uma nova entrada para os mesmos, até que um equilíbrio seja atingido.

Na figura 6.4 pode-se ver a estrutura de uma RNA genérica.

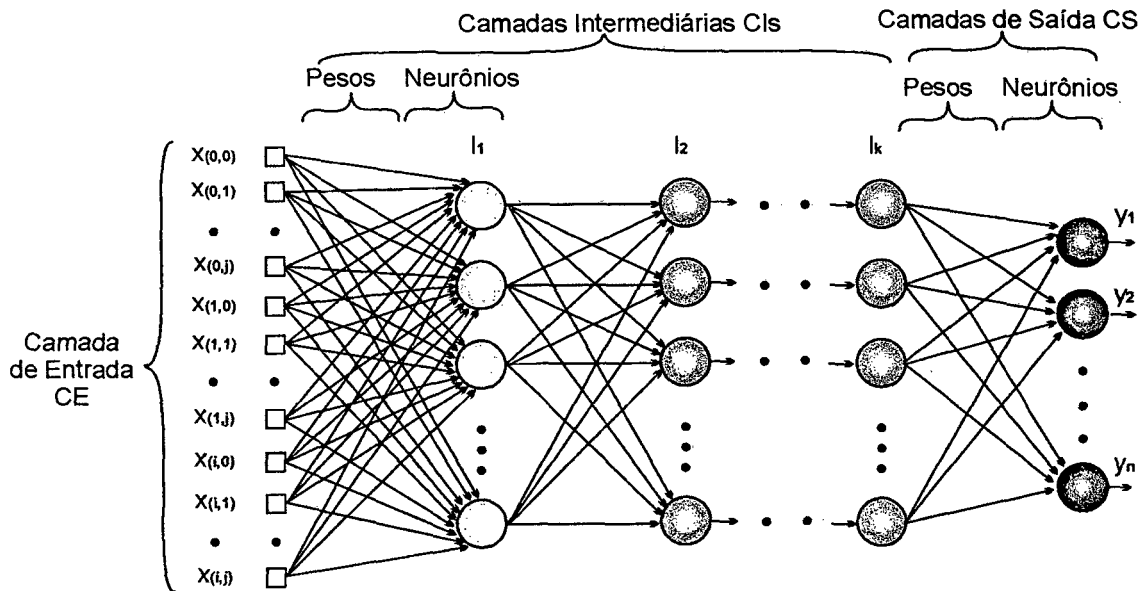


Figura 6.4 Arquitetura típica de uma RNA genérica tipo direta.

O número de camadas intermediárias em uma rede neural projetada para atender determinada aplicação não está completamente sistematizado e constitui objeto de estudo de muitos pesquisadores. O mesmo ocorre com o número de neurônios nas diversas camadas de uma estrutura.

A camada de saída, em um problema de classificação de dados, representa o número de classes em que se deseja separar os padrões de entrada. Na classificação de faces humanas, por exemplo, se os padrões de entrada estão distribuídos entre "n" indivíduos distintos (classes), deverão existir pelo menos "n" estados possíveis na saída para que se consiga a classificação desejada.

O reconhecimento de faces humanas, conforme já tratado no Capítulo 2, pode ser abordado de duas formas distintas:

- i. **Verificação:** Quando o indivíduo que se submete ao reconhecimento se apresenta previamente e tem suas características comparadas com as características do próprio indivíduo previamente armazenadas. Também chamado de comparação "um para um"<sup>4</sup>.
- ii. **Identificação:** Quando o indivíduo não se apresenta e as suas características são comparadas com as características de todos os indivíduos existentes em um banco de dados. Também chamado de comparação "um para todos"<sup>5</sup>.
- iii. Como o escopo deste trabalho é a verificação de faces, na camada de saída da RNA basta haver somente um neurônio que forneça um resultado com valores lineares entre  $[0,1]$  onde o maior índice represente uma maior proximidade entre a face sob teste com a face da base de dados. Outra possibilidade é de que a saída apresente dois níveis lógicos indicando diretamente se a face sob teste pertence à identidade alegada.

Diferentes arquiteturas têm sido desenvolvidas buscando a otimização do desempenho das RNAs nas diferentes aplicações a que se destinam. Em termos de classificação de padrões, algumas arquiteturas mais usadas são: (i) Redes de Hopfield, (ii) Memórias Associativas, (iii) Perceptron Multicamadas - PMC, (iv) ADALINE e MADALINE, (v) Mapa Auto-Organizável de Kohonen e (vi) Redes com Função de Base Radial.

### **6.1.5 Funcionamento das Redes Neurais Artificiais**

O cérebro humano constantemente recebe novas informações que são armazenadas e comparadas com informações previamente existentes. Através deste processamento de informação, as decisões requeridas para a solução do problema em questão são então tomadas. Genericamente o sistema nervoso humano, bem como o dos primatas, pode ser modelado de forma a dividir-se em três estágios fundamentais [Pandya e Macy, 1996]:

- i. **Estágio Sensor** - Onde os estímulos externos são captados, p.ex. o sistema visual, olfativo, tato, etc.
- ii. **Estágio Processador** - Onde as informações captadas são processadas (cérebro).
- iii. **Estágio de Saída** - Onde são obtidas as respostas aos estímulos (ação motora, orgânica, sonora).

Um dos maiores desafios no estudo do cérebro humano é o de entender como as informações processadas durante o seu desenvolvimento são armazenadas e aprendidas. Em RNAs a armazenagem do aprendizado é feita através dos pesos nas conexões sinápticas. Estes pesos representam o conhecimento armazenado em uma rede neural artificial.

---

<sup>4</sup> "One-to-one".

### 6.1.6 Aprendizado das RNAs

O aprendizado de uma RNA ocorre através de um processo de treinamento, onde um conjunto de casos exemplo, constituídos de entradas e saídas conhecidas, são apresentadas à rede. Neste processo iterativo, os pesos sinápticos nas conexões são ajustados dizendo-se neste momento que a rede está aprendendo. A partir deste momento, diz-se que a rede está apta a decidir de forma semelhante para novos dados que lhe sejam apresentados. Os pesos sinápticos de cada conexão, são os responsáveis pelo armazenamento do conhecimento adquirido pelas RNAs. Uma vez terminado o processo de treinamento, a rede está pronta para uso.

#### Conjuntos de Dados

A rigor o processo de treinamento de uma RNA exige que os dados disponíveis sejam divididos em três conjuntos:

- i. Conjunto de Treinamento: Usado para treinar a rede. Os elementos deste conjunto são apresentados repetidamente em ordem geralmente aleatória.
- ii. Conjunto de Validação: Usado para acompanhar a evolução do erro durante o processo de treino. Com o cálculo do erro em paralelo ao treinamento pode-se decidir o melhor momento para interromper o treinamento para melhorar a generalização do processo.
- iii. Conjunto de Teste: Usado para o teste final do desempenho da rede.

Em muitos processos de treinamento o conjunto de dados de treinamento é particionado somente em conjuntos de treinamento e teste. Em situações onde existem poucos dados disponíveis é possível usar técnicas de compartilhamento de conjuntos como a de Validação Cruzada - VC<sup>6</sup> onde os usos dos conjuntos de treinamento e teste são alternados no processo de treinamento. Na figura 6.5 pode-se ver como isto é feito para um conjunto de treinamento que é particionado em  $n$  sub-conjuntos  $D_1, D_2, \dots, D_n$ . Neste caso, faz-se  $m$  ciclos de treinamento  $tr1, tr2, \dots, trm$ , sendo que em cada um dos ciclos um novo subconjunto é usado para teste e o restante para treino.

A desvantagem desta estratégia é que o treinamento deve ser repetido várias vezes o que em algumas situações leva a exigir um grande tempo de processamento.

---

<sup>5</sup> "One-to-many".

<sup>6</sup> "Cross Validation".

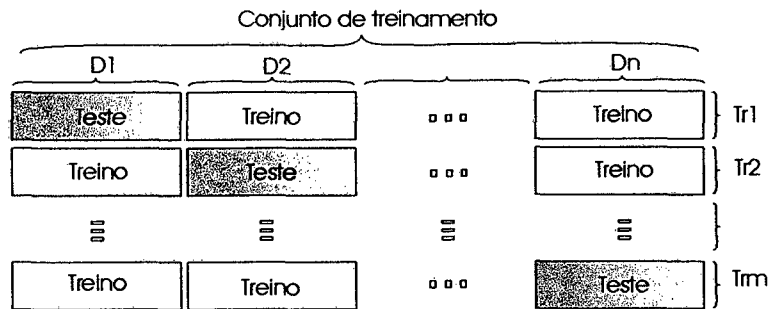


Figura 6.5 Validação cruzada dos dados do conjunto de treino.

**Generalização**

Generalização é o principal objetivo buscado pelo processo de treinamento de uma RNA. É uma característica da rede alcançada no processo de treinamento, que representa a capacidade do sistema em responder corretamente a um conjunto de dados chamado de conjunto de teste, sem que este conjunto tenha sido apresentado para a rede em sua etapa de treinamento. Diz-se que uma rede possui boa capacidade de generalização quando, após seu treinamento, esta responde de maneira correta a um conjunto pertencente à mesma família (com as mesmas características) do qual esta foi treinada, porém não contendo nenhum elemento igual.

Os principais parâmetros que devem ser observados para uma boa generalização são:

- i- Escolha do conjunto de treinamento: O conjunto de treinamento deve ser representativo do conjunto de dados total. Os dados de treinamento devem conter uma amostra de todos os casos possíveis, de preferência os casos extremos.
- ii- Parada antecipada: Uma parada antecipada no processo de treinamento determina que a rede não se especialize demais para aquele conjunto de dados de treinamento. Uma rede muito especializada faz com que esta não seja capaz de reconhecer os dados de teste como pertencentes à mesma classe dos dados de treinamento, prejudicando o desempenho desta.

Na figura 6.6 pode-se ver o comportamento típico do erro de treinamento de uma RNA. Note que a uma boa generalização se dá em uma região onde o erro para o conjunto de treinamento ainda não atingiu o seu ponto mínimo.

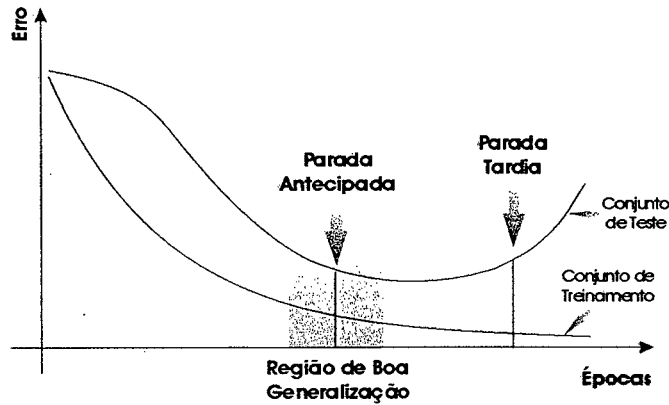


Figura 6.6 Curva típica de erro de treinamento de uma RNA.

### 6.1.7 Paradigmas de Aprendizado

Segundo Haykin, o processo do aprendizado pode ser classificado em dois grandes grupos segundo: (i) As Regras de Aprendizado e (ii) O Paradigma de Aprendizado. Esta abordagem esta esquematizada na figura 6.7.

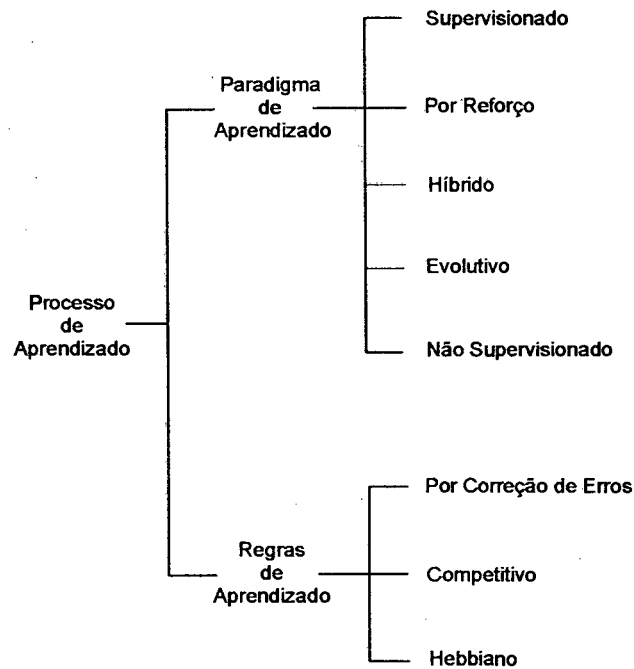


Figura 6.7 Uma taxionomia do processo de aprendizado [Haykin, 1994].

---

Quanto aos paradigmas do aprendizado, estes podem ser divididos em: (i) supervisionado, (ii) aprendizado forçado, (iii) não supervisionado ou (iv) híbrido. A seguir serão abordados com um pouco mais de detalhes os principais paradigmas de aprendizados utilizados em RNAs.

### **Aprendizado Supervisionado**

A supervisão no aprendizado de uma RNA é caracterizada pelo uso de um algoritmo chamado de “professor” que monitora o quanto as saídas atuais da rede diferem da saída esperada [Barreto, 1999]. O processo de treinamento é um processo iterativo, ou seja, apresentam-se as entradas do conjunto de treinamento a RNA, inicialmente com os pesos atribuídos por algum dos critérios citados anteriormente, e comparam-se as saídas destas com os resultados esperados. Este processo se repete até que a rede comece a apresentar resultados mais semelhantes possíveis aos do conjunto de treinamento.

A atualização dos pesos pode ser feita de duas maneiras diferentes no processo de treinamento:

- i. Por padrões. A atualização dos pesos é feita após cada elemento do conjunto de treinamento ter sido submetido à rede. Neste modo em cada ciclo de apresentação, também chamado de época, os pesos serão atualizados tantas vezes quantas forem os números de padrões do conjunto de treinamento.
- ii. Por grupos: A atualização dos pesos é feita após todos os elementos do grupo terem sido submetidos à rede. A atualização dos pesos pode ser feita basicamente de duas formas: a) Ao final da apresentação de todos os elementos do conjunto à RNA, ou seja uma atualização por épocas, ou ao final da apresentação de cada conjunto. b) Ao final da apresentação de todos os elementos de um determinado grupo, se o conjunto de elementos possuir diferentes grupos de elementos. Assim a atualização dos pesos ocorrerá tantas vezes quantas forem o número de grupos de um conjunto de treinamento.

### **Aprendizado Auto-Organizável ou Não Supervisionado**

Aprendizado não supervisionado se caracteriza pela atualização dos pesos independentemente da informação da saída da RNA. Neste caso, os pesos são ajustados na medida em que a rede vai sendo provida de padrões de entrada selecionados como representativos de cada classe. O processo consiste em se fazer a rede responder de modo semelhante para exemplos de elementos semelhantes [Barreto, 1999].



### **Aprendizado Híbrido**

O aprendizado híbrido pode ser utilizado através da adoção conjunta para uma mesma rede das duas formas de aprendizado vistas anteriormente: a supervisionada e a não supervisionada. RNAs com Função de Base Radial [Bishop, 1995] admitem este tipo de aprendizado, onde a primeira camada de conexões pode ser treinada de forma não supervisionada e a segunda, de forma supervisionada.

### **Aprendizado por Reforço**

Este aprendizado consiste no uso do método da tentativa e erro. Não existe a figura do “professor” para orientar o sistema acerca do seu direcionamento em relação à saída desejada. A realimentação acerca do sucesso do treinamento é dada através de um parâmetro chamado de coeficiente de desempenho ou sinal de reforço [Haykin, 1994].

### **Aprendizado Evolutivo**

Aprendizado evolutivo é um modelo inspirado na evolução biológica. Usando Algoritmos Genéticos - AG, é capaz de atuar na tanto na arquitetura de uma RNA quanto no seu aprendizado, alterando os valores dos pesos de modo a fazer com que a rede seja capaz de resolver determinados problemas [Barreto, 1999]. O paradigma do aprendizado evolutivo é extremamente interessante para uso em reconhecimento de padrões, visto que além do aprendizado, problema comum a todas as aplicações, ele é capaz também de otimizar a arquitetura da RNA, que invariavelmente é um problema complexo, pois reconhecer padrões no mundo real significa trabalhar com padrões não linearmente separáveis.

## **6.1.8 Regras de Aprendizado das RNAs**

Ainda segundo a taxionomia do processo de aprendizado apresentada por Haykin, figura 6.7, as regras de aprendizado podem ser classificadas em: (i) aprendizado competitivo, (ii) aprendizado Hebbiano, (iii) aprendizado por correção de erros e (iv) aprendizado por funções de base radial.

### **Aprendizado Competitivo**

O aprendizado competitivo se caracteriza pela competição entre os neurônios, de modo que a competição os leva a ter somente um deles acabar excitado. Este tipo de aprendizado tornou-se muito popular em redes de Kohonen.

### **Aprendizado Hebbiano**

A lei de Hebb é uma das mais antigas regras de aprendizado utilizadas. A regra diz basicamente: Ajuste a intensidade da conexão entre os neurônios A e B de uma quantidade proporcional ao valor da ativação simultânea dos dois neurônios [Barreto, 1999].

A lei de Hebb reduz o problema do aprendizado para um processo puramente local onde o valor dos pesos é alterado em função dos erros detectados localmente, não dependendo dos valores das demais variáveis espalhadas pelo sistema. Definindo os pesos iniciais da rede iguais a zero a regra de Hebb, foi proposta para ajustar pesos de redes lineares que têm como característica básica a capacidade de classificar apenas padrões linearmente separáveis [Kovacs, 1996].

### **Aprendizado pela Regra de Correção de Erros**

Este tipo de aprendizado é o mais usado nas RNAs em geral e divide-se em:

- i. Regra Delta, "Least Mean Square" ou "Widrow-Hoff": Criada a partir da generalização da lei de Hebb, efetua uma otimização que pode ser interpretada como o modelo matemático de um mecanismo de seleção. Esta regra se aplica a redes com apenas duas camadas de neurônios [Barreto, 2001].
- ii. Retropropagação<sup>7</sup>: É considerada a generalização da regra Delta para redes diretas com mais de duas camadas [Barreto, 1999]. Porém a retropropagação é somente aplicável à neurônios com funções de ativação diferenciáveis. Existem variações deste como os algoritmos: (i) "Quickprop", (ii) "Variable Learning Rate", (iii) "Resilient Backpropagation", (iv) "Conjugate Gradient Algorithms", (v) "Quasi-Newton Algorithms", (vi) "Reduced Memory Levenberg-Marquardt", dentre outros. Estes algoritmos diferem entre si pela capacidade de melhorar a performance do algoritmo de treinamento em termos de (i) tempo de processamento, (ii) recursos computacionais e (iii) precisão dos resultados. Este tipo de algoritmo de treinamento é sem dúvida o mais difundido entre as RNAs, encontrando-se amplamente explorado na literatura como em [Fausett, 1994], [Gallant, 1993], [Kovacs, 1996], [Poggio e Girosi, 1999], [Rich e Knight, 1994], [Haykin, 1994], [Brasil et al., 1999] e [Barreto, 1999] não sendo necessário o aprofundamento do assunto dentro deste trabalho.

---

<sup>7</sup> "Backpropagation".

### **Conclusão a Respeito das Regras de Treinamento**

Um bom treinamento da RNA resulta em uma boa generalização. Para isto a escolha de uma regra de aprendizado adequada ao tipo de rede é fundamental. O aprendizado, segundo as regras da lei de Hebb, está limitado à classificação de padrões linearmente separáveis, o que não atende os propósitos deste trabalho. No entanto, a regra de correção de erros representada pelo algoritmo da retropropagação é uma excelente escolha para o sistema de reconhecimento de faces, ainda mais se aliada a alguma estratégia de aceleração da busca do erro mínimo tal como a taxa de aprendizado variável e momentos. Outra regra que pode ser usada a qual apresenta a vantagem de ter menor tempo de treinamento é a que utiliza as FBRs, neste caso, naturalmente, a regra deve ser escolhida em função do tipo de rede utilizada.

Algumas observações devem ser feitas para o sucesso do processo de treinamento, em geral elas dizem respeito a aspectos conhecidos mas que tem fundamental importância como:

- i. **Excessivos Ciclos de Treinamento<sup>8</sup>**: Quando o número de neurônios da camada intermediária esta correto, mas o processo de treinamento avançou para além do ponto de Parada Antecipada - PA. Neste caso a rede não terá uma boa generalização.
- ii. **Excessivo Número de Neurônios<sup>9</sup>**: Quando a rede tem muitos neurônios na(s) camada(s) intermediária(s). Neste caso o sistema neural irá aprender os exemplos de treinamento e não a relação entre os conjuntos de entrada e saída. Neste caso a rede novamente não terá uma boa generalização.
- iii. **Insuficiente Número de Neurônios**: A rede não irá produzir um bom modelo do problema.
- iv. **Insuficiente Número de Exemplos**: A rede irá memorizar somente os exemplos apresentados.
- v. **Conjunto Desbalanceado**: Quando existe um número desproporcional entre a classe que se deseja reconhecer (ovelhas) em relação às outras classes presentes (lobos).
- vi. **Baixa Correlação**: Quando se tem uma relação fraca entre os parâmetros de entrada e saída da rede.

#### **6.1.9 Redes Neurais para Reconhecimento de Padrões**

Do ponto de vista das RNAs, o problema de reconhecimento de faces como proposto neste trabalho é um problema estático, o que permite o uso de redes neurais estáticas. Neste sentido

---

<sup>8</sup> "Over-training".

os principais paradigmas de RNAs para reconhecimento de padrões são: (i) Perceptron, (ii) Mapa Auto-organizável de Kohonen e (iii) Rede Neural com Função de Base Radial.

### Perceptron

O Perceptron, desenvolvido por Rosenblatt em 1958, é uma rede com apenas uma camada de neurônios. Muito citado ele é definido como uma rede linear, portanto consegue classificar apenas dados que sejam linearmente separáveis. O Perceptron também tem a característica de ser direta ou seja, seus sinais se propagam somente no sentido direto na rede. Consiste basicamente de um neurônio conectado a várias entradas através de conexões com pesos ajustáveis. Tais pesos são normalmente ajustados através de um aprendizado supervisionado.

O Perceptron, apesar do desencorajamento publicado por Minsky e Pappert citado na seção 6.1.1 deste trabalho, pode ser implementado com mais de uma camada, sendo chamado neste caso de Perceptron Multicamadas - PMC sendo capaz de classificar padrões não linearmente separáveis.

Na figura 6.8 pode-se ver os tipos de distribuições espaciais de padrões que podem ser classificados pelo Perceptron de uma só camada (a) e pelo de múltiplas camadas (b).

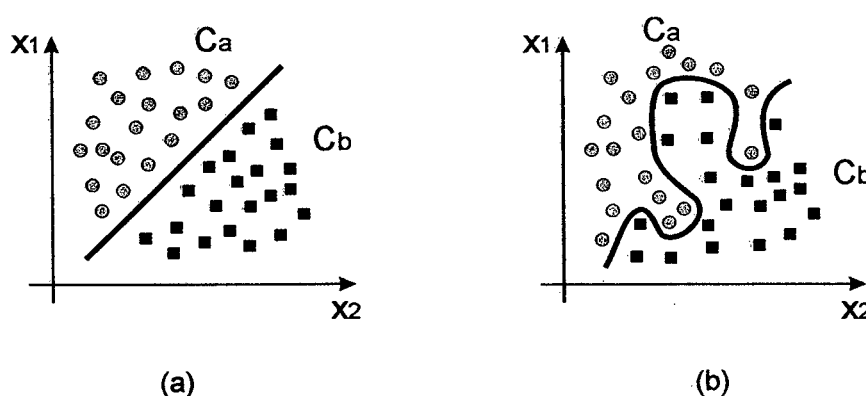


Figura 6.8 Padrões (a) linearmente separáveis e (b) não linearmente separáveis.

Sua capacidade de classificar dados não lineares confere ao PMC maior proximidade aos problemas reais. Um algoritmo de treinamento bastante utilizado para este modelo de rede é o de retropropagação<sup>10</sup>.

<sup>9</sup> "Overfitting".

<sup>10</sup> "Backpropagation".

### **Mapa Auto-organizável de Kohonen**

Mapas Auto-organizáveis de Kohonen - MAOK são utilizados desde problemas de otimização até o de classificação de padrões não linearmente separáveis. Um MAOK possui neurônios que podem ser arranjados unidimensionalmente no caso de uma estrutura linear ou bidimensional para estruturas em grade. O treinamento das MAOK é feito de forma não supervisionada. Em classificação de dados foi desenvolvida uma variação dos algoritmos de treinamento dos MAOK denominada de Aprendizado de Quantificação Vetorial - AQV<sup>11</sup> [Schalkoff, 1992]. A diferença entre estes dois paradigmas reside no fato de que no caso do AQV o número definido de clusters é igual ao número de categorias dos dados e os pesos são ajustados de forma supervisionada.

### **Rede Neural com Função de Base Radial**

Rnas com Função de Base Radial - FBR<sup>12</sup> são uma classe de RNAs que em sua forma mais reduzida possuem uma só camada intermediária. Esta camada, a qual utiliza funções de ativação tipo FBR (geralmente funções gaussianas), tem sua ativação estimulada pela distância Euclidiana entre as entradas e os centros das FBRs. Este tipo de rede será abordado com mais detalhes na seção 6.2 deste trabalho. O Método de treinamento mais utilizado para este tipo de rede é o não supervisionado para os parâmetros da camada intermediária e de saída. Este tipo de treinamento confere ao modelo um tempo de treinamento bastante reduzido. RNAs com FBRs têm sido extensivamente usadas para interpolação, regressão e classificação devido às suas propriedades de aproximadora universal.

### **Conclusões a Cerca das Arquiteturas das RNAs PMC e FBR**

RNAs do tipo PMC e FBR são exemplos de redes diretas multicamadas ambas aproximadoras universais. As principais diferenças entre estes dois paradigmas estão resumidas na tabela 6.1.

---

<sup>11</sup> "Learning Vector Quantization – LVQ".

<sup>12</sup> "Radial Basis Function - RBF"

Tabela 6.1 Características das redes FBRs e PMCs.

<b>Características</b>	<b>Função de Base Radial FBR</b>	<b>Perceptron Multicamadas PMC</b>
Linearidade das camadas: Intermediária	Não lineares	Não lineares
Camada de saída	Linear	Ambas
Confluência <sup>13</sup> na camada intermediária	Processa a distância Euclidiana entre as entradas e o centros das FBRs	Processa a soma das entradas multiplicadas pelos pesos sinápticos
Treinamento	Supervisionado, não supervisionado ou híbrido.  Geralmente: 1 - Não supervisionado para a primeira camada 2 - Supervisionado para a segunda camada	Geralmente supervisionado
Generalização	Boa	Boa
Arquitetura	Geralmente uma só camada intermediária	Várias Camadas
Tempo de treinamento	Menor	Maior
Tempo de utilização	Maior	Menor
Interpretação do funcionamento	Fácil	Difícil
Separabilidade	Hiper-elipsóides	Hiper-planos

<sup>13</sup> Operador matemático genérico, localizado entre os pesos e a função de ativação. Usualmente função de soma.

### 6.1.10 Justificativa da Escolha do Paradigma

Neste trabalho optou-se pelo uso da RNA do tipo FBR. Os principais motivos que levaram a esta opção foram:

- i. **Tempo de treinamento na etapa de desenvolvimento:** A possibilidade de treinamento em um tempo menor torna-se interessante durante a fase de desenvolvimento do sistema. Nesta fase faz-se uma série de pré-ensaios, com o objetivo de conhecer o comportamento da rede segundo os seus parâmetros intrínsecos. Estes ensaios invariavelmente são repetitivos o que exige muito tempo. Já o tempo de treinamento, quando o sistema está pronto para uso e se conhecem os parâmetros da rede e sua influência, tem menos importância tendo em vista principalmente que o treinamento para uma face/indivíduo vai ser feito uma vez só e a velocidade dos computadores cresce a cada dia.
- ii. **Fácil configuração:** Uma RNA tipo PMC possui muitas camadas intermediárias com muitos pesos podendo estar conectados de forma complexa e utilizando diferentes funções de ativação. Em contrapartida RNAs com FBRs têm uma arquitetura mais simples, geralmente constituída de uma camada intermediária com os parâmetros das FBRs e uma camada de saída que faz combinações lineares das ativações das FBRs.
- iii. **Fácil compreensão:** O aprendizado não supervisionado das RNAs tipo FBRs possibilita uma maior compreensão do processo intrínseco da rede e da influência dos seus parâmetros, o que não ocorre com os outros paradigmas em que o treinamento é feito com todos os parâmetros sendo mudados ao mesmo tempo.
- iv. **Paradigma pouco explorado:** O emprego de um paradigma de RNA, que possui fortes pré-indicações para o tipo de aplicação como a deste trabalho, mas que tem muito a ser explorado a exemplo das redes PMC, foi um dos motivos que levaram a escolha da RNA tipo FBR.

É sabido que o problema de reconhecimento de faces humanas é um problema não linear mas de complexidade desconhecida. A tarefa de separação dos padrões no espaço de entrada em classes é executada no PMC através de hiperplanos e na RNA tipo RBF através de hiperelipsóides. Acreditando que a utilização de superfícies do tipo hiperelipsóides é um método mais eficiente de separabilidade de espaços, neste trabalho optou-se por este modelo.

## 6.2 Rede Neural do Tipo Função de Base Radial

Uma RNA do tipo Função de Base Radial - FBR é uma rede em que a ativação dos neurônios da camada intermediária é dada pela diferença entre o vetor de entrada  $\mathbf{x}$  e o vetor dos centros  $\mathbf{c}$  das FBRs usadas como função de ativação desta camada. Este tipo de rede geralmente possui uma só camada intermediária, mas pode haver mais de uma conforme descrito por He e Lapedes [He e Lapedes, 1991], Orr [Orr, 1996] e Azevedo [Azevedo et al., 2000].

Uma FBR é uma função de mapeamento não linear, tipicamente gaussiana, e que eleva as entradas de um espaço  $n$ -dimensional a um outro espaço  $H$  de ordem mais elevada.

FBRs são geralmente referidas como de "mapeamento local", no sentido de que somente poucas bases radiais irão responder significativamente para cada nova entrada  $\mathbf{x}$  que for apresentada à rede. Isto se dá devido à limitada área de atuação que cada centro das FBR da rede. Por esta razão uma rede do tipo FBR usualmente requer mais centros para conseguir um bom mapeamento. Este comportamento também ajuda com que haja pouca interferência dos centros uns com os outros durante o processo de treinamento. Em contrapartida redes do tipo PMC que utilizam funções de ativação como as referidas na seção 6.1.3 são referidas como de "mapeamento global", significando que todas os neurônios das camadas intermediárias respondem para todo o espaço de entrada de  $\mathbf{x}$ . Este comportamento pode se manifestar durante a etapa de treinamento onde estas interferências podem abaixar em muito a velocidade do processo de aprendizado da rede.

### 6.2.1 Arquitetura da RNA tipo FBR

Uma RNA tipo FBR é na sua essência uma estrutura de rede direta de múltiplas camadas. Na grande maioria dos casos esta rede possui 3 camadas, cada uma delas com funções bem definidas:

- i. Camada de Entrada: É composta por unidades sensoras e conecta a rede com os dados de entrada.
- ii. Camadas Intermediárias: Geralmente constituída de apenas uma, aplica uma transformação não-linear ao espaço de entrada resultando em um espaço intermediário. Esta transformação ocorre devido às funções de ativação do tipo base radial, por serem não-lineares.
- iii. Camada de Saída: É sempre uma camada com FA linear fazendo com que a saída da rede seja uma combinação linear das respostas das camadas intermediárias [Haykin, 1994].



A arquitetura típica e genérica de uma FBR pode ser vista na figura 6.9.

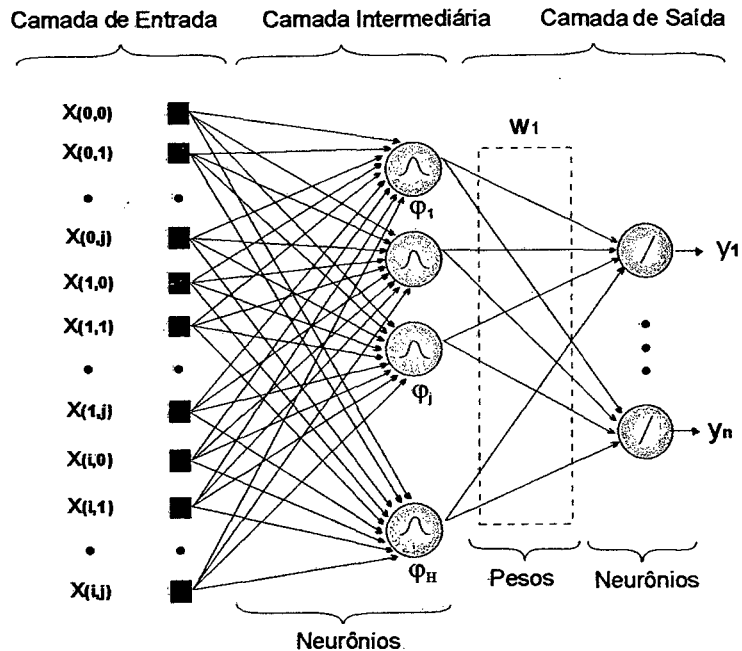


Figura 6.9 Arquitetura genérica típica de uma RNA do tipo FBR.

As principais características da RNA tipo FBR são:

- i. Possuir na confluência<sup>14</sup> da camada intermediária uma função que calcula a distância Euclidiana entre as entradas e os centros das FBR's.
- ii. A função de ativação na camada intermediária é do tipo radial e na camada de saída linear.
- iii. Possui um único conjunto de pesos ajustáveis  $w_j$  na camada de saída.

### 6.2.2 Relação entre a Entrada e a Saída da Rede

RNAs tipo FBRs são redes neurais artificiais que utilizam uma categoria de funções de ativação diferentes das funções vistas anteriormente na seção 6.1.3. Enquanto as funções de ativação

<sup>14</sup> Confluência - Operador matemático que combina as entradas do neurônio. Geralmente expressa pela soma destas, mas podendo assumir outras funções como a distância Euclidiana no caso das FBRs [Barreto, 1999].

da seção 6.1.3 são usadas em redes que utilizam produto escalar como argumento de suas camadas intermediárias como as do PMC, funções de ativação do tipo radial são utilizadas em redes que usam distâncias radiais como argumento de suas camadas intermediárias [Rögnvaldsson, 2001]. Na equação 6.7 está representada matematicamente a relação que existe entre as entradas da rede  $\mathbf{x}$ , as posições dos centros  $\mathbf{c}_j$ , o valor médio dos centros  $c_{j0}$  e onde  $\|\cdot\|$  representa uma norma, geralmente a Euclidiana definida no espaço de entrada como sendo a distância Euclidiana dos centros das FBRs  $c_j$  às entradas  $\mathbf{x}$ .

$$h_j(\mathbf{x}) = \varphi\left(\frac{\|\mathbf{c}_j - \mathbf{x}\|}{c_{j0}}\right) \quad (6.7)$$

Uma RNA tipo FBR com  $i^j$  entradas e  $n$  saídas, implementa um mapeamento do tipo  $y_n : \mathbb{R}^{i^j} \rightarrow \mathbb{R}^n$ , o qual pode ser visto na equação 6.8:

$$y(\mathbf{x}) = w_0 + \sum_{i=1}^j w_i h_i(\mathbf{x}) \quad (6.8)$$

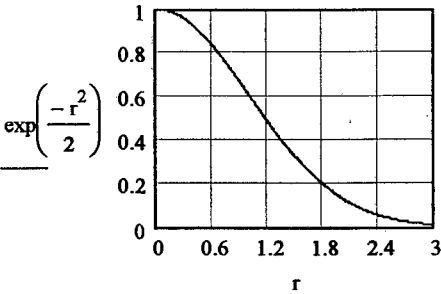
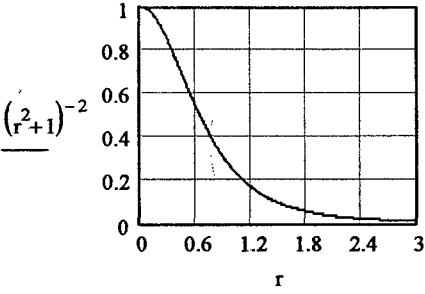
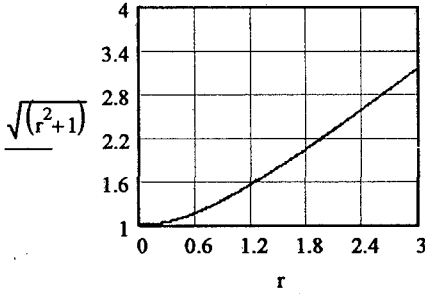
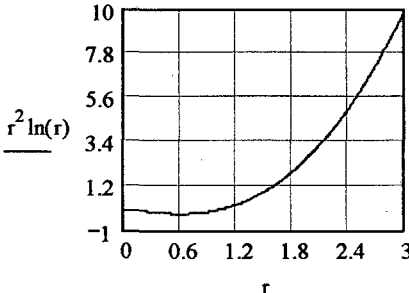
### 6.2.3 Funções de Ativação

A função de ativação mais comum para a camada intermediária da rede com FBR é do tipo gaussiana, dada pela equação 6.10. Também podem ser usadas as função de base radial multiquadrática inversa equação 6.11, multiquadrática equação 6.12 e a Thin-plate spline equação 6.13 [Chen et al., 1991, Haykin, 1994, Orr, 1996, Rögnvaldsson, 2001]. Estas funções  $\varphi(r)$  com suas respectivas curvas podem ser vistas na tabela Tabela 6.2.

Onde  $r$  é o raio da função radial conforme apresentado na equação 6.9:

$$r = \frac{\|\mathbf{c}_j - \mathbf{x}\|}{c_{j0}} \quad (6.9)$$

Tabela 6.2 Funções de ativação do tipo base radial.

Função de ativação	Curva
<p>(i) Gaussiana</p> $\varphi(r) = \exp\left(\frac{-r^2}{2}\right) \quad (6.10)$	
<p>(ii) Multiquadrática Inversa</p> $\varphi(r) = (r^2 + 1)^{-2} \quad (6.11)$	
<p>(iii) Multiquadrática</p> $\varphi(r) = \sqrt{r^2 + 1} \quad (6.12)$	
<p>(iv) "Thin-plate spline"</p> $\varphi(r) = r^2 \ln(r) \quad (6.13)$	

A função gaussiana decresce monotonicamente com a distância a partir do centro "c" enquanto a função multiquadrática aumenta. As FBRs com função de ativação do tipo gaussiana são chamadas de mapeadoras locais, pois dão uma resposta significativa para a região do centro "c" e vizinhança decaindo conforme se afasta deste, enquanto as funções do tipo da multiquadrática, pelo motivo inverso, são chamadas de mapeadoras globais. Funções de ativação do tipo gaussiana têm maior plausibilidade biológica devido à sua resposta finita [Orr, 1996]. O uso de FBR possibilita resolver simplificada o problema de se conseguir uma interpolação exata, onde uma função exata aplicada para dados com distúrbios tende a fazer com que esta se tome instável [Bishop, 1995].

Os neurônios da camada de saída podem ser modelados com uma função de ativação do tipo linear visto que a camada de saída de uma RNA tipo RBF possui como função principal a concatenação dos resultados.

Existem algumas variações da FBR tradicional, entre elas podemos citar: (i) Regressão de Redes Generalizadas - RRG<sup>15</sup> e (ii) Redes Neurais Probabilísticas - RNP<sup>16</sup> [Wasserman, 1993]

#### 6.2.4 Número de Neurônios

O raciocínio mais simples que se pode ter é o de utilizar tantos neurônios/centros "j" na camada intermediária quantos forem os exemplos de treinamentos presentes "k". Deste modo cada neurônio se especializaria em um exemplo atendendo aos requisitos do treinamento, ou seja  $j = k$ . Treinando-se com um número de neurônios igual ao número de exemplos, obtém-se uma solução exata para o problema. A figura 6.10 mostra uma analogia entre o grau de interpolação de uma função polinomial e o número de neurônios de uma RNA tipo RBF. Note-se na figura 6.10 (a) que quando se usa o grau do polinômio igual ao número de dados ocorre a interpolação exata, onde a curva interpolada passa exatamente sobre os dados. Este mesmo raciocínio pode ser estendido para uma RNA tipo FBR, a qual quando utiliza um número de centros igual ao número de exemplos no conjunto de treinamento encontra uma solução exata. Soluções exatas não são boas soluções quando os dados possuem ruído pois neste caso há uma solução fortemente baseada em dados poucos confiáveis. Nesta situação uma interpolação exata quase sempre produz uma má interpolação nas regiões entre pontos (figura 6.10 (a)). Além do que o número de neurônios neste caso é maior, podendo de alguma forma inviabilizar o projeto. Já na figura 6.10 (b) usando um grau de polinômio menor que o número de dados disponível a interpolação é na média melhor que a interpolação exata. Este mesmo

---

<sup>15</sup> "Generalized Regression Networks - GRNN".

raciocínio pode ser estendido para a análise do número de neurônios de uma RNA tipo RBF. Estas redes menores, geralmente produzem melhores resultados quando submetidas a um conjunto de teste. Isto se deve à melhor generalização que a rede com menos neurônios atinge em oposição a aquela que utiliza todos os neurônios e interpolação exata.

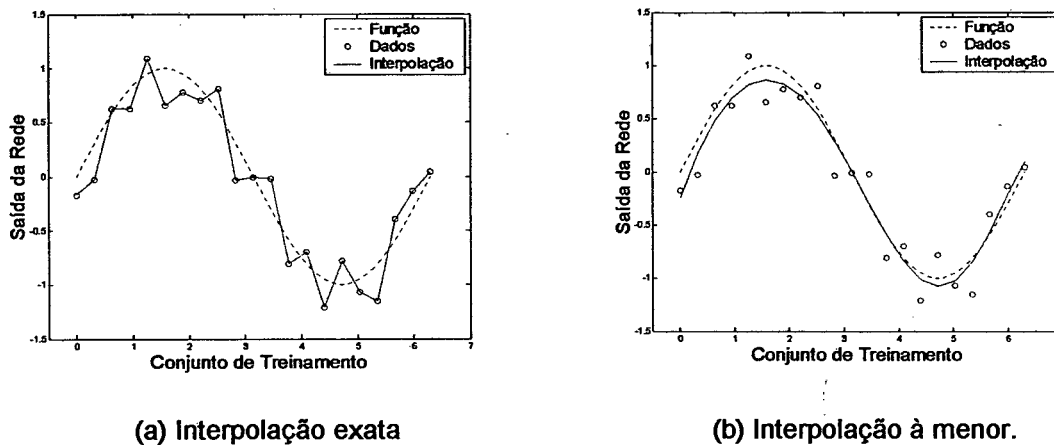


Figura 6.10 Curvas de interpolação polinomial de funções.

O conjunto de centros/neurônios pode ser reduzido basicamente através de:

- i. Escolha randômica de um subconjunto.
- ii. Seleção de clusters nos dados de entrada.
- iii. Escolha seletiva das entradas que minimizam o erro.
- iv. Estes métodos são abordados com mais detalhes nas seções que tratam exclusivamente sobre treinamento das RNAs do tipo FBRs.

### 6.2.5 O Raio da FBR

Note-se que do ponto de vista do sistema de classificação, os dados das formas de faces neste trabalho, em sua forma mais reduzida (21 x 15), possuem 315 valores por forma de face, e são assim considerados multidimensionais. Alta dimensionalidade dos dados dificulta uma representação gráfica exata dos parâmetros da rede e conseqüentemente o entendimento dos mesmos. Devido à dificuldade de se visualizar os sistemas multidimensionais recorre-se, sem perda de generalidade, a algumas simplificações na representação dos mesmos. A distribuição

<sup>16</sup> Redes Neurais Probabilísticas – RNP ou "Probabilistic Neural Network – PNN".

especial dos centros das FBRs é um exemplo deste problema, que tem sido contornado através do uso de uma representação em duas dimensões o que tem se mostrado eficiente.

A otimização do raio da FBR tem sido objeto de muitas publicações. Os dois critérios mais utilizados [Benoudjit et al., 2002] são os que utilizam:

- i. **Raio Fixo:** Considera o raio da FBR constante em todas as ocorrências [Haykin, 1994, Orr, 1999, Orr, 1996]. O algoritmo utilizado nestes casos é o mostrado na equação 6.14:

$$r = \frac{d_{\max}}{\sqrt{2M}} \quad (6.14)$$

Onde:  $d_{\max}$  é a máxima distância entre os centros e  $M$  o número de centros.

- ii. **Raio Variável:** Este critério considera que os raios podem ter valores diferentes para cada centro. Um critério neste sentido pode ser o mostrado na equação 6.15 [Moody e Darken, 1989]:

$$r_j = \frac{1}{s} \left( \sum_{i=1}^s \| \mathbf{c}_i - \mathbf{c}_j \|^2 \right)^{\frac{1}{2}} \quad (6.15)$$

Onde:  $s$  tipicamente assume o valor 2 e  $\mathbf{c}_i$  são os  $s$  centros vizinhos do centro  $\mathbf{c}_j$ .

Na tabela 6.3, pode-se ver de forma gráfica alguns critérios para a escolha do raio  $r$  da FBR, baseado em uma análise bidimensional da posição dos centros das FBRs. Cumpre ressaltar que cada aplicação tem suas características próprias merecendo um estudo preliminar para se determinar qual a melhor estratégia de escolha do raio a ser usado.

Tabela 6.3 Critérios para a escolha do raio da FBR.

Distâncias	Representação gráfica
<p>(i) Distância Mínima</p> $d = \min\left(\frac{d_1, d_2, d_3, d_4}{2}\right) \quad (6.16)$	
<p>(ii) Distância Média</p> $d = \left(\frac{d_1, d_2, d_3, d_4}{4}\right) \quad (6.17)$	
<p>(iii) Distância Máxima</p> $d = \max(d_1, d_2, d_3, d_4) \quad (6.18)$	
<p>(iv) Distância do Vizinho mais Próximo</p> $d = vmp(d_1, d_2, d_3, d_4) \quad (6.19)$	

## 6.2.6 Métodos de Treinamento Não Supervisionados

Nas RNAs tipo RBFs, as camadas intermediária e de saída desempenham funções bastante distintas, o que segundo Lowe [Lowe.D, 1991] já seria suficiente para separar o procedimento de aprendizado não supervisionado em duas etapas distintas:

- i. Primeira Etapa: Relacionada com a camada intermediária, compreende a determinação dos parâmetros (i) número de centros das FBRs, (ii) posição dos centros das FBRs e (iii) raio ou variância das FBRs.
- ii. Segunda Etapa: Relacionada com a camada de saída, compreende a determinação dos pesos desta camada.

A determinação do número de centros e do raio da FBR já foram abordados nas seções 6.2.4 e 6.2.5. A posição dos centros das FBRs no treinamento não supervisionado é em geral realizada através de dois métodos:

- i. Seleção Aleatória de Dados: Este tipo de modelo é adequado para quando se desejar respostas rápidas geralmente em processos em tempo real. A idéia é selecionar aleatoriamente um sub conjunto de centros dentre o conjunto total.
- ii. Métodos de Agrupamento: O agrupamento prévio de elementos com uma mesma característica é uma estratégia de minimização de esforços. O sistema não terá que classificar independentemente todos os exemplos de entradas, mas sim somente os grupos previamente formados por elementos com uma mesma característica em comum chamados de clusters. Desta forma emprega-se menos centros para a classificação pois estão se classificando grupos e não elementos individuais. Os principais métodos de agrupamento de dados são [Bala et al., 1995, Bishop, 1995, De Carvalho, 2001, Kohonen.T., 1988]: (i) k-médias, (ii) k-médias fuzzy, (iii) k-vizinhos mais próximos, (iv) mapa auto-organizado de Kohonen e (v) agrupamento por algoritmos genéticos [Azevedo et al., 2000]. Dentre eles o algoritmo de k-médias é um dos mais utilizados.

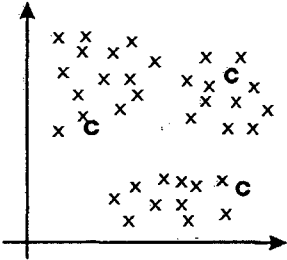
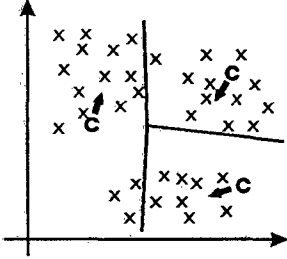
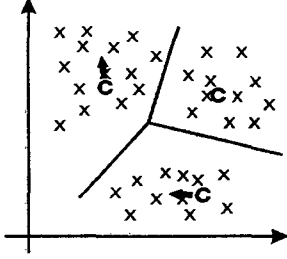
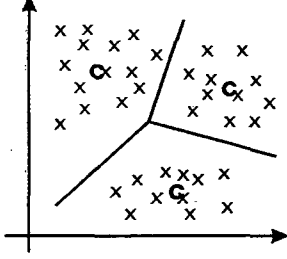
A seleção de sub-conjuntos de forma não supervisionada geralmente não generaliza tão bem quanto os métodos de regularização em modelos lineares como o PMC [Frank e Friedmann, 1993, Wettschereck e Dietterich, 1992].

### O Algoritmo de k-médias

Este algoritmo, desenvolvido por Moody e Darken [Moody e Darken, 1989] posiciona os centros das FBRs em regiões do espaço onde existam mais dados significativos presentes, veja procedimento resumido na tabela 6.4.



Tabela 6.4 Procedimento para cálculo de clusters segundo o algoritmo do k-médias

Procedimento	Visualização
<p>(i) Inicialize a posição dos "M" centros "c" aleatoriamente ou com os mesmos valores das entradas: <math>c_k = x_k, k = 1, 2, 3, \dots, M</math></p>	
<p>(ii) Calcule a distância entre cada entrada e os centros e separe os grupos de menor distâncias:</p> $c(x) = \arg \min_k \ x(n) - c_k(n)\ , k = 1, 2, \dots, M, \text{ onde } c_k(n) \text{ é o centro da } k^{\text{th}} \text{ FBR na } n^{\text{th}} \text{ iteração.}$ <p>Calcule a nova posição dos centros baseado na média das entradas de seu grupo: <math>c_k = \frac{1}{M_k} \sum_{j=1}^{M_k} x_j</math></p>	
<p>(iii) Repita o passo (ii) até que não haja mais mudanças na posição dos centros "c".</p>	
<p>(iv) Fim do procedimento com os "M" centros alocados no centro de seus clusters.</p>	

Do ponto de vista dos dados de entrada, cada exemplo terá associado a ele um novo centro pertencente ao cluster mais próximo. Este processo é iterativo e o algoritmo que parte de um número previamente especificado de centros de FBRs vai readaptando os valores da alocação dos centros de forma a conseguir otimizar o erro de alocação. O algoritmo do k-médias para  $M$  clusters com um número  $N$  de centros, constituídos de um grupo de entradas  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ .

### Determinação dos Pesos da Camada de Saída

A finalização do processo de treinamento se dá através da determinação dos pesos da camada de saída através da obtenção de  $\mathbf{w}$  na equação 6.20.

$$y(\mathbf{x}) = \mathbf{w}\Phi \quad (6.20)$$

Onde  $y(\mathbf{x})$  é a saída da rede,  $\mathbf{w}$  é a matriz de pesos requerida e  $\Phi$  é a matriz de interpolação com colunas proporcionais ao número de camadas intermediárias "j" e linhas "k" com os valores das respostas para cada exemplo de observado é dada pela equação 6.21.

$$\Phi = \begin{pmatrix} \varphi_1(1) & \varphi_2(1) & \dots & \dots & \varphi_j(1) \\ \varphi_1(2) & \varphi_2(2) & \dots & \dots & \varphi_j(2) \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \varphi_1(n) & \varphi_2(n) & \dots & \dots & \varphi_j(k) \end{pmatrix} \quad (6.21)$$

Uma solução do tipo dado pela equação 6.22 seria a saída mais natural para este problema.

$$\mathbf{w} = \Phi^{-1}y(\mathbf{x}) \quad (6.22)$$

Esta operação embora simples, só é possível se  $\Phi$  obedecer ao teorema de Michellis [Michelli, 1986] o qual estabelece que uma matriz é não singular se os elementos de  $\mathbf{x}$  forem distintos e  $\varphi$  for monotônico no intervalo de  $[0, \infty)$ . No entanto em algumas situações, a matriz  $\Phi$  pode ser singular, impedindo que se obtenha matriz inversa da mesma. Na prática trabalhando-se com grandes volumes de dados é muito comum obter matrizes esparsas ou com dados linearmente dependentes o que a torna não inversível. Uma solução para contornar tais

problemas é a utilização do algoritmo da Pseudo-Inversa de More e Penrose. Este artifício é freqüentemente empregado para contornar tais situações e pode ser visto na equação 6.23 sendo utilizado para o cálculo da matriz de pesos  $w$  :

$$w = (\Phi^T \Phi)^{-1} \Phi^T y \quad (6.23)$$

Ainda segundo Broomhead e Lowe [Lowe e Broomhead, 1988] pode ocorrer uma má generalização para alguns tipos de problemas devido ao método matemático aqui descrito. Isto é em parte explicado pela interpolação exata e seus problemas com dados em presença de ruído como citado na seção 6.2.4 deste trabalho. Para casos onde o ruído não pode ser evitado, existem paradigmas mais adaptados como o caso do algoritmo conhecido como Rede Neural de Regressão Generalizada - RNRG<sup>17</sup>.

### 6.2.7 Métodos de Treinamento Supervisionados

O método de treinamento não supervisionado, visto na seção 6.2.6, é sem dúvida um procedimento eficiente e rápido mais que não aborda o problema visando sua otimização. É possível treinar uma RNA tipo RBF com um algoritmo supervisionado p. ex. baseado no Gradiente Descendente - GD<sup>18</sup>. Haykin [Haykin, 1994] apresenta um conjunto de algoritmos baseados no GD que determinam (i) o valor dos pesos lineares, (ii) a posição dos centros e (iii) o raio da FBR através do método do gradiente descendente. Apesar de buscar uma otimização do problema, este procedimento representa um caso de otimização não linear, o qual exige um alto custo computacional do sistema. Deste ponto de vista, o método supervisionado perde uma de suas maiores vantagens que é a o tempo de treinamento reduzido. Também a fácil interpretabilidade das funções da camada intermediária são perdidas quando se usa o aprendizado supervisionado [Bishop, 1995].

Melhorias na velocidade de treinamento podem ser conseguidas através de técnicas que identificam previamente quais são os centros de maior importância como as apresentadas por Omohundro [Omohundro, 1987]. Um outro método que seleciona os centros de forma supervisionada e chamado de Método dos Mínimos Quadrados Ortogonais<sup>19</sup>, que segundo Bishop [Bishop, 1995] é um dos métodos mais eficazes para o treinamento de FBRs.

---

<sup>17</sup> "Reneralized Regression Neural Network – GRNN".

<sup>18</sup> "Least Mean Square", Regra Delta, "Widrow-Hoff" ou Retropropagação.

<sup>19</sup> "Orthogonal Least Squares".

### **Método dos Mínimos Quadrados Ortogonais**

Este método desenvolvido por Chen [Chen et al., 1991] consiste na seleção da posição e número dos centros de forma supervisionada, sendo a escolha do raio da FBR de atribuição prévia ou não supervisionada. Ao contrário do que em geral ocorre com os treinamentos supervisionados onde os parâmetros são ajustados de forma contínua, o algoritmo dos Mínimos Quadrados Ortogonais - MQO utiliza uma forma combinatória de otimização do número de centros.

O método de MQO tem propriedades numéricas superiores quando comparado ao método dos Mínimos Quadrados - MQ [Chen et al., 1991]. Este algoritmo permite o uso de várias estratégias de seleção de centros [Miller, 1990, Myers, 1986], das quais as mais usadas são:

- i. Seleção Crescente<sup>20</sup>: Inicia-se com um centro e passo a passo incrementa-se este número até que o erro mínimo e/ou o número de centros máximo sejam atingidos.
- ii. Seleção Decrescente: Inicia com o número de centros máximo e passo a passo decresce um centro até que o erro máximo seja atingido.
- iii. Seleção Passo a Passo: Inicia-se com um centro e a cada passo um centro é acrescido ou subtraído submetendo-os a testes estatísticos.
- iv. Seleção por Saltos e Restrições: Este algoritmo, seleciona sub-conjuntos de centros que minimizam a função de erro, este sub-conjunto pode ser encontrado sem examinar todo o conjunto, sendo vantajoso para conjuntos de até 50 centros candidatos [Fumival e Wilson, 1974].

O método de treinamento supervisionado com MQOs com seleção de centros crescente é um dos mais utilizados. Seu algoritmo consiste dos seguintes passos:

- i. Atribuir um valor de raio  $r_0$  comum para todos os centros.
- ii. Escolher um primeiro centro para a RNA com o valor retirado do conjunto de treino, tal que resulte no menor erro na saída da rede para todo o conjunto de treino.
- iii. Adicionar mais um centro escolhido entre os restantes do conjunto de treino.
- iv. Escolher aquele centro que introduzir o menor erro na saída da rede para todo o conjunto de treino.
- v. Repetir os passos (iii) e (iv) até que ou o erro mínimo ou o número de centros máximo sejam alcançados.

---

<sup>20</sup> Algoritmo MQO desenvolvido por Chen [Chen et al., 1991] e utilizado pela Mathworks no software Matlab.

É fácil compreender porque um algoritmo que desempenhe esta função tem um alto custo computacional, pois em cada passo é necessário calcular uma solução para as combinações de todas as entradas, p. ex. calculada com o método da pseudo-inversa. No entanto o método dos MQO de Chen, ao se utilizar da Ortogonalização de Gram Schmidt - OGS [Horn e Johnson, 1985], permite que a busca pelos melhores centros se faça de maneira otimizada economizando tempo de processamento. O algoritmo de treinamento supervisionado dos MQO pode ser encontrado em detalhes em [Chen et al., 1991, Haykin, 1994, Orr, 1996, Rögnvaldsson, 2001, Sherstinsky e Picard, 1993].

### **6.3 Considerações Importantes no Projeto com RNAs**

Nesta seção são abordados alguns aspectos importantes do projeto de RNAs relativas ao desempenho e robustez aplicado ao reconhecimento de padrões.

#### **6.3.1 Invariância em RNAs**

Em muitas situações práticas, os dados de entrada se apresentam sob ação de transformações que podem comprometer o desempenho do sistema classificador. Estas transformações produzem um conjunto de entrada sem uma desejada padronização. No caso do reconhecimento de padrões este é em geral um problema de ordem geométrica onde os dados de entrada podem apresentar variações de sua posição espacial como (i) translação, (ii) rotação e (iii) escalonamento [Barnard e Casasent, 1991, Fukumi et al., 1997, Perantonis e Lisboa, 1992]. Também as variações das condições de iluminação são outra fonte importante de transformação dos dados de entrada, porém no escopo deste trabalho, devido à natureza da extração dos dados, este não é um problema.

A consideração destes aspectos no projeto da rede neural traz consideráveis melhorias no desempenho do sistema. Diz-se que um sistema é invariante a uma determinada transformação quando sua saída permanece inalterada mesmo que suas entradas sofram ação destas transformações [Bishop, 1995]. Existem porém métodos para contornar estes problemas de falta de padronização. Em geral as invariâncias estão divididas em três categorias:

- i. **Invariância por Treinamento:** Este método usa uma quantidade bastante grande de exemplos, em várias posições espaciais de forma a suprir uma faixa grande de possibilidades de situações de posicionamento espacial que os dados possam vir a apresentar. A RNA interpola entre casos de exemplos, portanto consegue-se uma boa invariância selecionando-se para o conjunto de treino os valores extremos do conjunto de dados.

- ii. **Invariância por Seleção de Características:** Características especiais, invariantes segundo certas transformações podem ser usadas para a classificação dos dados. A extração destas características pode envolver um pré-processamento dos dados.
- iii. **Invariância por Estrutura:** A invariância é obtida através da escolha de uma adequada arquitetura de rede. Nela podem ser acrescentadas conexões específicas entre neurônios de modo a não se obter mudanças na saída devido às transformações na entrada. Estes sistemas geralmente são de ordem elevada e os neurônios possuem representação local.
- iv. **Invariância por Treinamento:** O método de invariância por treinamento é o único que exige muitos exemplos de treinamento. As vantagens e desvantagens de cada método podem ser vistas com mais detalhes em [Barnard e Casasent, 1991, Bishop, 1995].

### **6.3.2 Redução da Dimensionalidade**

Dimensionalidade é usada aqui como uma medida do número de características de uma representação de padrões, em outras palavras, ela representa a dimensão do espaço de características. A principal razão para que a dimensionalidade de um sistema classificador seja a menor possível está relacionada com o custo computacional. RNAs com menos neurônios possuem algumas vantagens como: (i) ter menos parâmetros para serem determinados, (ii) generalizar mais facilmente e (iii) ter um tempo de treinamento menor.

Neste trabalho, as matrizes com informações das formas de faces (seção 4.3), possuem tamanhos que variam de 101 x 71 (7171 valores) até 21 x 15 (315 valores). Note que mesmo utilizando-se uma matriz de formas com menor resolução (21 x 15), teremos uma dimensionalidade de 315. Esta taxa é considerada alta e com conseqüências diretas no desempenho do sistema, uma vez que o processo de classificação pode exigir pré-processamentos onde são feitas rotações, translações, escalonamentos.

A alta dimensionalidade também afeta diretamente a arquitetura da RNA uma vez que mais neurônios e conseqüentemente mais funções de ativação serão calculadas a cada processamento. Watanabe [Watanabe, 1985] lembra da importância de se caracterizar os padrões com poucas características fortes, citando que dois padrões apesar de diferentes, se representados por uma quantidade grande de características similares, podem vir a ser considerados similares. Por outro lado um classificador pode ter seu poder de discriminação reduzido se lhe forem apresentadas poucas características.

A redução da dimensionalidade pode ocorrer basicamente através de duas maneiras [de Campos, 2001]:

- i. Seleção de Características <sup>21</sup>: A redução se dá através do uso de subconjuntos com características fundamentais retirados do conjunto principal das características como o (i) Método do Corte de Nós - MCN <sup>22</sup> e (ii) Algoritmos Genéticos - AG.
- ii. Extração de Características: Este método cria um novo espaço a partir do espaço original. Nesta categoria as principais técnicas são as de: (i) ACP, (ii) Transformada de Fourier - TF e a (iii) Análise de Discriminantes Lineares – ADL <sup>23</sup>.

Geralmente a redução de dimensionalidade vem acompanhada de perda de informação. O mérito dos métodos de redução de dimensionalidade reside no fato de promover uma redução efetiva com um mínimo de perda de informação.

O objetivo deste trabalho não é o de fazer uma análise profunda nos tipos de algoritmos e/ou métodos disponíveis para redução de dimensionalidade, mas sim mostrar que o uso de um método de redução de dimensionalidade traz grandes benefícios quanto à redução do custo computacional do sistema, conseguido com mínimas perdas de informação. Aqui, optou-se por empregar o método da ACP, um dos métodos mais conhecidos e de reconhecida eficácia [Jain et al., 2000].

### **Análise dos Componentes Principais - ACP**

Este método estatístico parte do princípio que um sistema pode ser reduzido desde que seus dados possuam algum grau de correlação entre si [Jolliffe, 1986]. Vetores de formas de face podem ser considerados de alta dimensionalidade, pois estes pertencem a um grande espaço vetorial não otimizado.

A figura 6.11 mostra uma distribuição espacial típica para a representação das formas de faces. Formas de faces possuem características comuns como (i) contornos elípticos, dois olhos, (iii) um nariz, (iv) o nariz entre os olhos, etc. Sob este enfoque pode-se admitir que os dados das formas de face possam estar correlacionados e conseqüentemente o Espaço de Formas - EF utilizado estar sobredimensionado. Se isto ocorrer o Espaço de Formas de Faces - EFF pode ser representado por um subconjunto reduzido deste.

A ACP vem desta forma auxiliar na análise dos dados verificando se existem dados correlacionados que podem ser desprezados para a redução da dimensionalidade do sistema.

---

<sup>21</sup> "Feature selection".

<sup>22</sup> "Node Pruning".

<sup>23</sup> "Linear Discriminant Analysis – LDA".

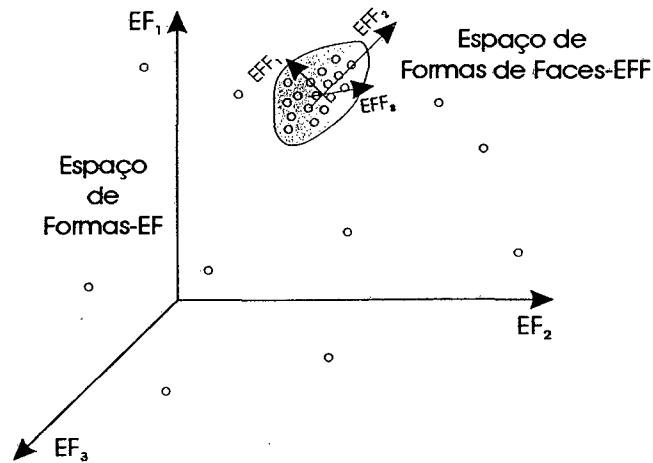


Figura 6.11 Representação do Espaço de Formas de Faces - EFF como sub-espaço do Espaço de Formas - EF.

Na figura 6.12 (a), pode-se ver uma representação típica de dados pertencentes a uma base vetorial  $X$ . Note que estes dados representados segundo a base  $X$  possuem uma correlação maior do que se os mesmos fossem representados sob uma nova base  $Y$ , conforme apresentado na figura 6.12 (b). A grandeza que representa a correlação dos dados é a matriz de covariância. Note que os novos eixos desta base estão sob a direção da maior variação dos dados.

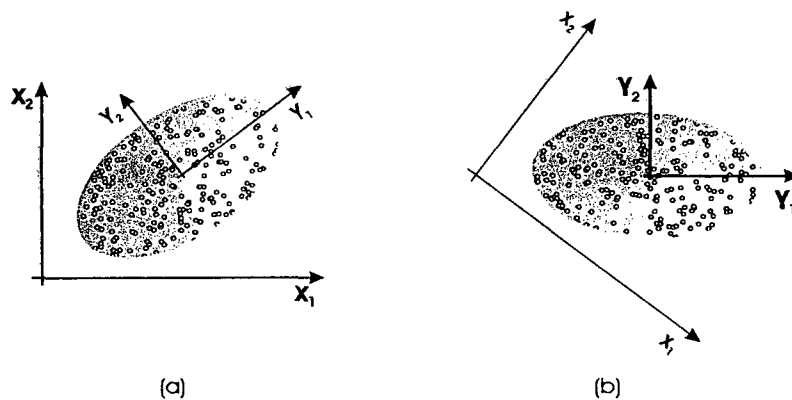


Figura 6.12 Representação de dados bidimensionais segundo eixos quaisquer (a) e ortonormais (b).

A ACP consiste no mapeamento de um espaço vetorial  $X$  com  $n$  padrões de dimensão  $d$  dado por:



$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,n} \\ x_{2,1} & x_{2,2} & \dots & x_{2,n} \\ \dots & \dots & \dots & \dots \\ x_{d,1} & x_{d,2} & \dots & x_{d,n} \end{bmatrix} \quad (6.24)$$

em um novo espaço vetorial  $Y$  ( $y_i \in \mathbb{R}^p$ ), com  $p < d$ :

$$Y = \begin{bmatrix} y_{1,1} & y_{1,2} & \dots & y_{1,n} \\ y_{2,1} & y_{2,2} & \dots & y_{2,n} \\ \dots & \dots & \dots & \dots \\ y_{d,1} & y_{d,2} & \dots & y_{p,n} \end{bmatrix} \quad (6.25)$$

Existem basicamente dois métodos que podem ser usados para se extrair os componentes principais de um sistema que são:

- i. Matriz de Correlação - MC
- ii. Decomposição em valores Singulares - DVS

### **ACP através da Através da Matriz de Correlação - MC**

Apresentado de forma mais detalhada [Martinez, 2001], o processo consiste em se calcular os autovetores  $e_i$  e os correspondentes autovalores  $v_i$  de  $X$ . A seleção dos componentes principais se faz mediante a escolha de autovetores  $u_i$  que correspondam aos autovalores  $v_i$  mais significativos. Isto se dá através da obtenção de uma transformação linear  $W$  que mapeie o espaço original  $d$ -dimensional  $X$  para o novo espaço  $Y$   $p$ -dimensional, tal que  $y_i = W^T \cdot x_i$  para  $i = 1, 2, \dots, N$ .

A matriz  $W$  é formada pelos autovetores  $u_i$  obtidos pela solução de  $v_i u_i = C u_i$ , onde  $C = X \cdot X^T$  é a matriz de covariância e  $v_i$  os autovalores associados com os autovetores  $v_i$ . Antes do cálculo da matriz de covariância  $C$  os vetores são normalizados ou seja  $\|x_i\| = 1$  e subtraída a média de todos os vetores de forma a assegurar que o autovetor com o maior autovalor represente a nova base em que a variância dos vetores é máxima.

O melhor espaço em baixa dimensão é conseguido através da seleção dos melhores autovetores da matriz de covariância  $C$  de  $X$ , os quais correspondem aos maiores autovalores.

Mais detalhes sobre a ACP podem ser encontrados em [Jolliffe, 1986] [Abdi et al., 1998, Bishop, 1995, Feraud, 1999, Hancock et al., 1996, Haykin, 1994, ROMDHANI, 1996, Valentin et al., 1998].

Devido ao fato que as imagens resultantes deste processamento se assemelhem muito a imagens de faces, os autovetores da ACP são chamados de Autofaces<sup>24</sup>. Os vetores autofaces podem ser usados diretamente para o reconhecimento de faces através de um classificador qualquer [Belhumeur et al., 1999, Lawrence et al., 1996a, Li et al., 1999, Moghaddam et al., 1999, Pentland et al., 1994, Turk e Pentland, 1991, Valentin et al., 1998]. De maneira análoga, a aplicação da ACP nas faces tridimensionais deste trabalho resulta em autovetores que são aqui denominados de Autoformas<sup>25</sup>.

### **ACP através da Decomposição em Valores Singulares - DVS**

A Decomposição em Valores Singulares - DVS é uma fatoração que também revela os componentes principais de uma matriz. DVS é uma ferramenta que permite a redução de matrizes baseada na obtenção de seus Valores Singulares - VS. Na ACP via Matriz de Correlação os autovetores são obtidos a partir da matriz de covariância dos dados enquanto que na DVS os autovetores são obtidos diretamente da matriz de dados. Uma DVS equivale a uma ACP via Matriz de Correlação quando os dados de entrada possuírem média igual a zero.

A DVS permite que uma matriz  $X$  com  $X \in \mathcal{R}^{m \times n}$  possa ser decomposta em termos de três outras matrizes [Anderson et al., 1999]:  $U \in \mathcal{R}^{m \times n}$ ,  $D \in \mathcal{R}^{n \times n}$  e  $V \in \mathcal{R}^{n \times n}$  conforme a equação 6.26:

$$X_{m \times n} = U_{m \times n} D_{n \times n} V_{n \times n}^T \quad (6.26)$$

onde:  $X_{m \times n}$  é a matriz original de alta dimensionalidade,  $U_{m \times n}$  é uma matriz cujas colunas são os autovetores de  $XX^T$ , dada pela equação 6.27:

---

<sup>24</sup> "Eigenfaces".

<sup>25</sup> "Eigenshapes".

$$XX^T = UDV^TVDU^T = UD^2U^T \quad (6.27)$$

$V_{n \times n}$  é uma matriz cujas colunas são os autovetores de  $X^T X$ , dada pela equação 6.28:

$$X^T X = VDU^TUDV^T = VD^2V^T \quad (6.28)$$

e  $D_{n \times n}$  é a matriz dos VS, dada pela equação 6.29:

$$D = \text{diag}(\sigma_1 \ \sigma_2 \ \dots \ \sigma_n) \quad (6.29)$$

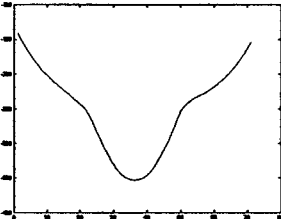
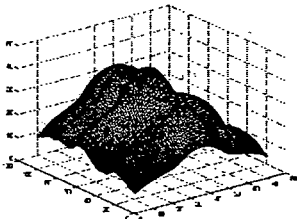
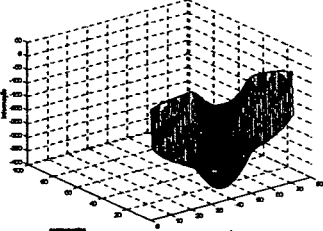
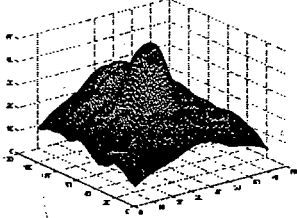
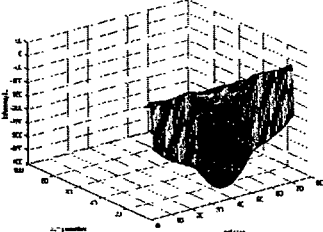
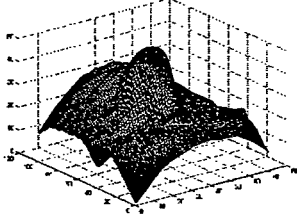
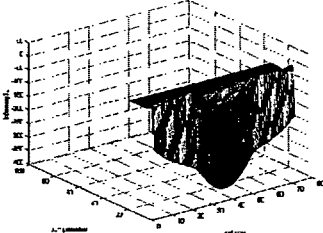
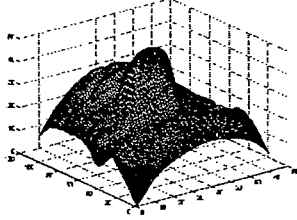
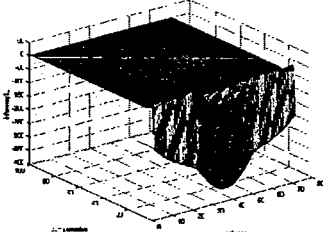
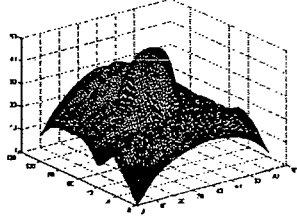
ordenada tal que  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$ . Ainda se  $U = (u_1 \ u_2 \ \dots \ u_n)$  e  $V = (v_1 \ v_2 \ \dots \ v_n)$ , a matriz reconstruída pode ser obtida para  $k$  valores singulares ("rank") através da equação 6.30:

$$Y = \sum_{i=1}^n u_i \cdot \sigma_i \cdot v_i^T \quad (6.30)$$

A tabela 6.5 ilustra a redução de dimensionalidade de uma matriz de forma de face originalmente com tamanho 101 x 71 (7171 valores) [Encinas et al., 2002]. Observa-se nesta tabela que os valores singulares significativos estão presentes apenas nas primeiras posições. Isto equivale a dizer que a matriz de formas de faces é representada eficazmente apenas pelos primeiros valores singulares como era de se esperar.

Isto pode ser visualmente verificado através da análise da forma reconstruída da face. Note-se que visualmente não existam diferenças significativas entre a face representada pelos oito primeiros valores singulares da representação máxima com os 101 valores singulares.

Tabela 6.5 ACP aplicada às matrizes de formas de faces.

Figura	Número de Valores Singulares	Distribuição dos Valores Singulares (Espaço Reduzido)	Forma da Face Reconstruída Autoforma - AF (Espaço Normal)
(a)	1		
(b)	2		
(c)	8		
(d)	20		
(e)	101		

## Capítulo 7

# Validação do Sistema

---

*Nos anos sessenta e setenta, estudantes me perguntavam freqüentemente, "Qual tipo de representação é melhor?" e eu normalmente respondia que nós precisaríamos de mais pesquisa. Mas agora eu responderia: Para resolver problemas realmente difíceis, nós teremos que usar várias representações diferentes. Isto é porque cada tipo particular de estrutura de dados tem suas próprias virtudes e deficiências, e nenhum por si só pareceria adequado para todas as funções diferentes envolvidas como o que nós chamamos de bom senso.*

- Marvin Minsky, (1927).

**N** Neste capítulo são apresentadas, as concepções do sistema de reconhecimento de faces humanas proposto, os resultados dos ensaios de qualificação do mesmo pelo método do Raciocínio Baseado em Casos - RBC e Por Redes Neurais Artificiais - RNAs.

### 7.1 A Concepção do Sistema

O sistema de extração da forma da face usado neste trabalho é de projeção de franjas, com processamento de alturas através do método de perfilometria de Fourier [Takeda et al., 1999] e [Takeda e Mutoh, 1999]. A este processamento são aplicados algoritmos de interpolação e

normalização com o objetivo de obter uma base de formas de faces adequada à proposição deste trabalho.

Com objetivo de poder fazer uma comparação de desempenho, a verificação das faces se dá através de dois paradigmas distintos, o de (i) Raciocínio Baseado em Casos - RBC e o (ii) Conexionista, este último utilizando uma Rede Neural Artificial - RNA do tipo Função de Base Radial - FBR como classificador.

## 7.2 Extração das Formas de Faces Humanas

O processamento de extração da forma de faces humanas<sup>1</sup> obedece exatamente a seqüência descrita na seção 4.2 para extração de formas de objetos genéricos. Nas figuras Figura 7.1 e 7.2 pode-se ver as imagens resultantes da aplicação do método em uma face de manequim com as mesmas características dimensionais de uma face humana típica.

A figura 7.1 (a) mostra a face com as franjas projetadas. É a partir desta imagem que se inicia o processo de obtenção da forma da face. A qualidade da medição da forma da face depende muito da qualidade desta imagem inicial. Bom contraste e continuidade das franjas são os principais parâmetros a serem otimizados para se atingir um bom resultado. Na figura 7.1 (b) pode-se ver o sinal da fase  $\theta(x, y)$  das franjas sobre a face. Note-se que nesta imagem já existe a informação das alturas da forma da face. Esta informação, conforme já explicado anteriormente na seção 4.2.4, deve ser modificada para a remoção do salto de fase e obtenção da forma da face no formato desejado.

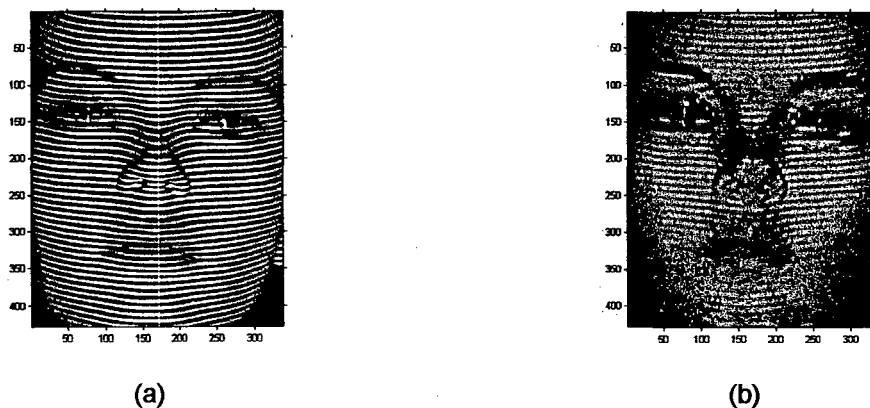


Figura 7.1 Face com franjas (a) e sinal de modulação da face (b).

<sup>1</sup> Desenvolvido em ambiente Matlab.

A remoção do salto de fase se faz através de um algoritmo que utiliza a informação do sinal de modulação (seção 4.2.4). Na figura 7.2 (a) pode-se ver o sinal de modulação da face do manequim.

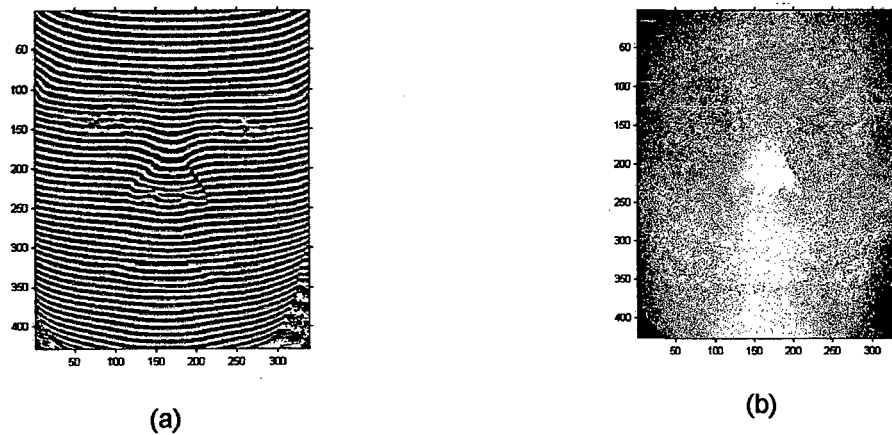


Figura 7.2 Face com salto de fase (a) e face sem salto de fase (b) .

Na figura 7.2 (b) pode-se ver a imagem da fase  $\theta(x, y)$  das franjas com o salto de fase removido. A transformação deste sinal de fase em alturas correspondentes é obtida através da equação 4.42.

Finalmente a figura 7.3 mostra os dados da forma do da face do manequim já transformados em alturas correspondentes. Por motivo de visualização estes dados estão apresentados com uma resolução de apenas 20% da resolução das imagens das figuras 7.1 e 7.2 que os geraram.

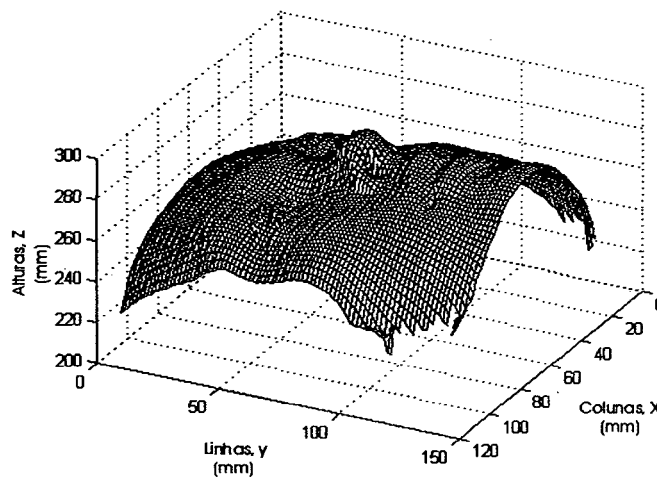


Figura 7.3 Dados típicos das alturas da forma da face de um manequim com escala real em milímetros.

A tela do software desenvolvido para a extração das formas das faces encontra-se no Apêndice A.

### 7.2.1 Os Sinais do Processamento da Forma

Na figura 7.4 pode-se observar sinais típicos provenientes do processamento das imagens das figuras 7.1 e 7.2.

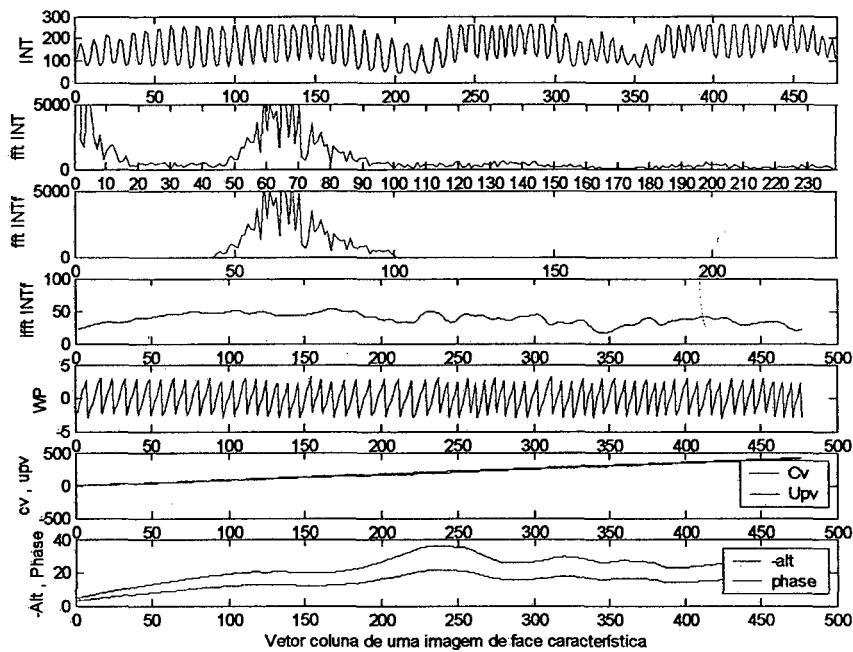


Figura 7.4 Sinais característicos do processo de extração de formas.

Os sinais da figura 7.4 se referem a uma coluna de pontos da matriz de dados da imagem da face com franjas, a qual pode ser visualizada na figura 7.1 através da coluna em branco no centro da face. Os diversos sinais da figura 7.4 são comentados abaixo:

- i. INT - Intensidade luminosa da coluna da face.
- ii. fft INT - Transformada de Fourier do sinal de intensidade luminosa.
- iii. fff INT f - Filtragem do sinal "fft INT".
- iv. ifft INT f - Transformada de Fourier Inversa do sinal "fff INT f".
- v. WP - Sinal de fase com salto.



- vi. cv, upv - Frequência portadora, fase portadora e salto removido.
- vii. Alt, Phase - Alturas em mm, fase sem portadora e com salto removido.

## 7.2.2 Ajuste do Sistema

### A Relação Fase Altura

A relação fase altura para este sistema foi verificada através do uso de uma forma geométrica conforme a figura 7.5.

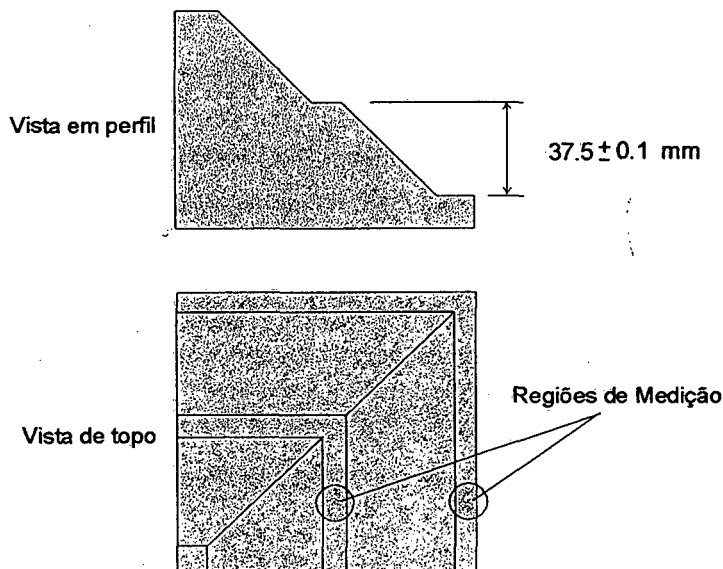


Figura 7.5 Padrão de dimensional corporificado.

Para este ensaio preliminar os parâmetros optogeométricos da figura 4.15 utilizados foram: (i)  $d = 1700$  mm, (ii)  $h = 330$  mm e (iii)  $f_0 = 0,5$ . A variação de fase verificada entre as duas regiões de medição mostradas na figura 4.15 foi de  $\Delta\theta = 22,0 \pm 0,3$ . A altura calculada, utilizando-se a equação 4.41, foi de  $\Delta d(\Delta\theta) = 36,9 \pm 0,5$  mm enquanto pela equação 4.42 (equação simplificada)  $\Delta D(\Delta\theta) = 36,1 \pm 0,5$  mm. Nesta posição o valor verdadeiro convencional do padrão é de  $37,5 \pm 0,1$  mm.

Note-se que esta calibração é bastante rudimentar para um sistema tridimensional pois retrata apenas a diferença de altura entre duas regiões bastante restritas do padrão. Porém, aceitando que um resultado semelhante pode ser obtido para todo o plano x, y de medição, a incerteza obtida é plenamente aceitável para o escopo deste trabalho.

A figura 7.6 mostra o comportamento dos resultados de  $\Delta d(\Delta\theta)$  segundo a equação 4.41 (modelo completo) e de  $\Delta D(\Delta\theta)$  segundo a equação 4.42 (modelo simplificado). Note que dependendo da faixa de medição ( $\Delta\theta$ ) a simplificação pode ser usada com sucesso.

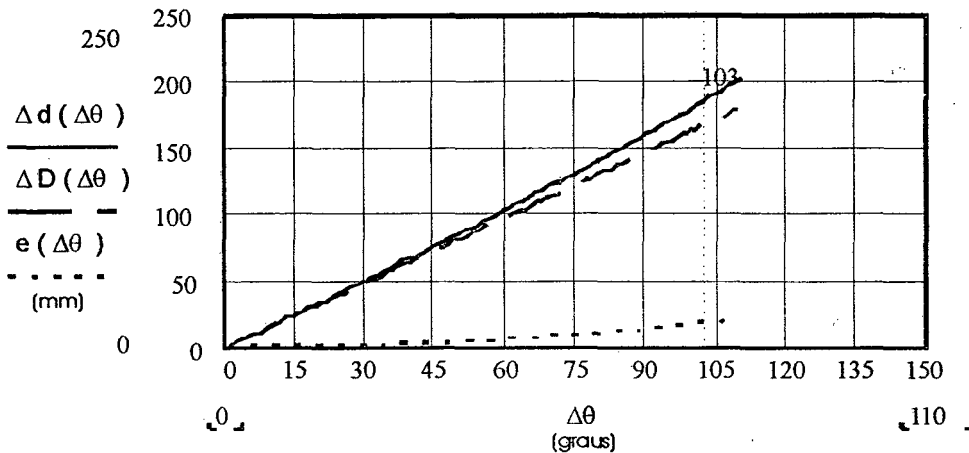


Figura 7.6 Comparação entre o modelo simplificado e o completo de cálculo de alturas.

### Faixa de Medição segundo a Freqüência Portadora

O sistema deve sofrer uma calibração para a determinação e ajuste da freqüência portadora para cada montagem. Esta freqüência depende dos parâmetros ópticos geométricos do experimento. O uso da freqüência portadora correta possibilita a obtenção de menores erros na conversão entre fase e alturas da forma a ser medida. A calibração se faz da seguinte forma:

- i. Ajustam-se preliminarmente os elementos do sistema de aquisição de imagens, projetor, câmera e faixa de medição.
- ii. Procura-se experimentalmente a posição relativa ótima entre câmera e projetor de forma a fazer coincidir os dois pivôs ópticos através do monitoramento da freqüência espacial dentro da faixa de medição desejada. A coincidência dos pivôs resulta na menor variação da freqüência portadora dentro do campo de medição.

### 7.2.3 Vantagens e Problemas Potenciais

A principal vantagem da medição de formas pelo princípio de projeção de franjas é sua simplicidade e velocidade de processamento da informação da forma da face. Enquanto em outros sistemas necessita-se da projeção de duas grades (Moiré), de duas câmeras para

captura da imagem (topogrametria) ou da rotação relativa entre câmera e objeto (Varredura e LASER RADAR), o Sistema proposto neste trabalho possui uma só câmera, um projetor de luz estruturada e a imagem é capturada de uma só vez.

Em termos de esforço computacional o processamento da imagem não compromete o desempenho do sistema. Na sua versão experimental o software de processamento de formas feito em MATLAB demora cerca de 30 s para processar a forma de uma face de 400 x 200 pontos<sup>2</sup>. Note que o software desenvolvido mostra graficamente todas as fases do processo (figuras 7.1, 7.2, 7.3, 7.4 e 7.7, ) o que aumenta o tempo de processamento. Em uma versão definitiva com imagens de menor tamanho e com linguagem de mais baixo nível estima-se que este tempo diminua pelo menos cerca de 10 vezes.

#### 7.2.4 Fontes de Erros

Faces humanas reais apresentam algumas peculiaridades que muitas vezes dificultam a extração de uma boa representação da geometria da face. Os principais parâmetros responsáveis por possíveis defeitos nas formas das faces humanas são:

- i. Pelos na face como barba, sobrancelha e cabelos dificultam a visualização das franjas, fornecendo dados de forma com baixa confiabilidade.
- ii. A pupila dos olhos é uma região de difícil processamento.
- iii. O nariz, dependendo do seu formato, pode ocasionar dificuldades para medir a forma da face.
- iv. A cor da face, que pode se apresentar dentro de uma grande faixa de valores desde escura até clara ou até reflexiva devido a gorduras ou suor, exige sistemas de projeção e câmeras com um certo grau de adaptabilidade.

Na figura 7.7, a título de exemplificação das possíveis causas de erros devido às características da imagem, pode-se ver uma imagem de forma de face humana típica e suas imperfeições devido ao método utilizado nas regiões dos olhos e de pelos como bigode e cavanhaque.

---

<sup>2</sup> Micro Computador PC Pentium I, 233MHz, 48Mb RAM.

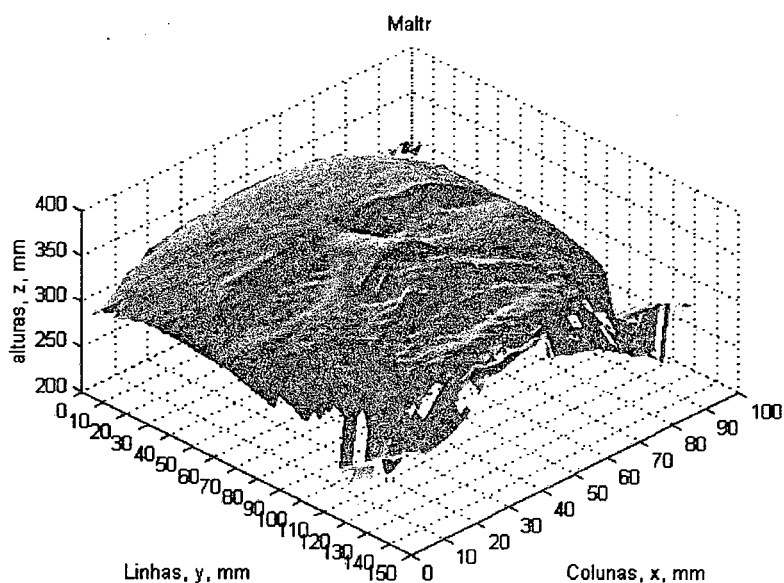


Figura 7.7 Forma de face humana real com imperfeições típicas.

## 7.3 Reconhecimento da Forma

Serão apresentados aqui resultados experimentais de diversos ensaios preliminares de verificação da forma das faces através de RBC e por RNAs do tipo RBF.

### 7.3.1 A Base de Dados dos Ensaio

As formas de faces utilizadas nos ensaios são as apresentadas na seção 4.3.1 deste trabalho. Devido à grande quantidade de formas de faces desta base, nos ensaios realizados, usaram-se subconjuntos desta.

Para caracterizar a capacidade do classificador em diferenciar faces, tanto para os cálculos segundo o paradigma do RBC ou o de RNAs, usou-se uma relação cruzada entre os indivíduos, onde uma face de cada indivíduo é testada contra uma face de cada um dos indivíduos. A estes resultados dá-se o nome de conjunto de lobos. Já para a caracterização da capacidade do classificador em associar faces de mesmos indivíduos, usaram-se variações de faces do mesmo indivíduo, de modo a que cada variação seja confrontada com outra variação do mesmo indivíduo. A estes resultados dá-se o nome de conjunto de ovelhas.

### 7.3.2 Ensaios de Verificação Através do RBC

O objetivo destes ensaios é o de verificar o desempenho do método de RBC na separação das populações de ovelhas e lobos. Esta análise é feita através do cálculo da DH. Para tanto, foram utilizados 52 indivíduos da base de formas de faces e calculou-se a Distância de Hamming - DH entre eles:

- i. DH da população de ovelhas: Foram utilizadas as primeiras 22 variações de cada indivíduo e calculando a DH destes entre si, totalizando 1.092 (22x52) DH's.
- ii. DH da população de lobos: Foi utilizada uma única variação de cada indivíduo e calculado a DH contra todos os indivíduos da base, totalizando 2.652 (51x52) DH's.

#### Aplicando Otimizações às Faces da Base

Conforme abordado na seção 5.3, através de pré-processamentos dos dados de faces consegue-se melhorar o desempenho da classificação por RBC. A aplicação da normalização e minimização dos erros de posicionamento espacial, bem como o uso da ponderação estatística, otimiza os dados melhorando o desempenho do classificador.

Neste trabalho foram aplicados três tipos de otimização nos dados de forma de faces:

- i. Otimização por Normalização Espacial - ONE: A otimização se dá através da normalização espacial das formas de faces (seção 5.31) e da correção de erros de posicionamento (seção 5.32).
- ii. Otimização por Ponderação Facial - OPF: Uma matriz de pesos é aplicada a cada matriz de formas de faces. Esta matriz de pesos, chamada de matriz de pesos próprios (seção 5.42), realça as regiões da face que são comuns entre as faces de treinamento de um mesmo indivíduo.
- iii. Otimização Híbrida - OH: Esta otimização emprega simultaneamente a ONE e a OPF.

São mostrados a seguir, os resultados da aplicação dos pré-processamentos dos dados. A figura 7.8 mostra a distribuição das DH das populações de lobos e ovelhas para os dados originais.

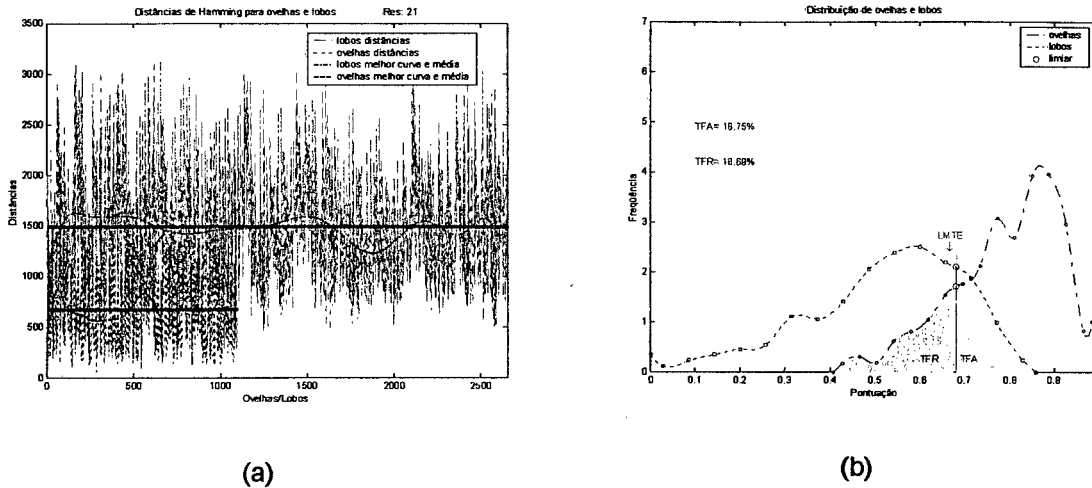


Figura 7.8 Distribuição da DH para os dados originais.

A figura 7.9 mostra a distribuição das DH das populações de lobos e ovelhas para os dados originais normalizados espacialmente (ONE).

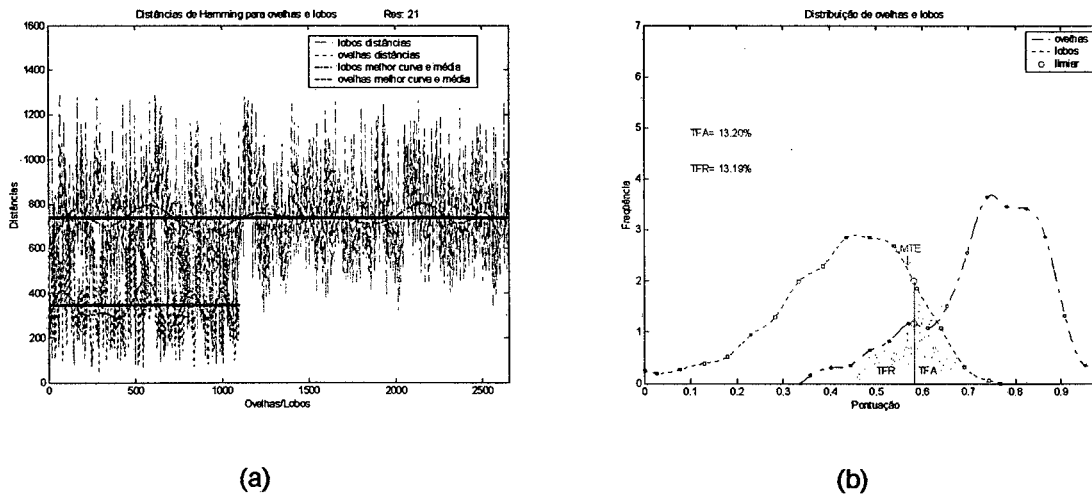


Figura 7.9 Distribuição da DH para os dados com ONE.

A figura 7.10 mostra a distribuição das DH das populações de lobos e ovelhas para os dados originais submetidos a ponderação facial (OPF).

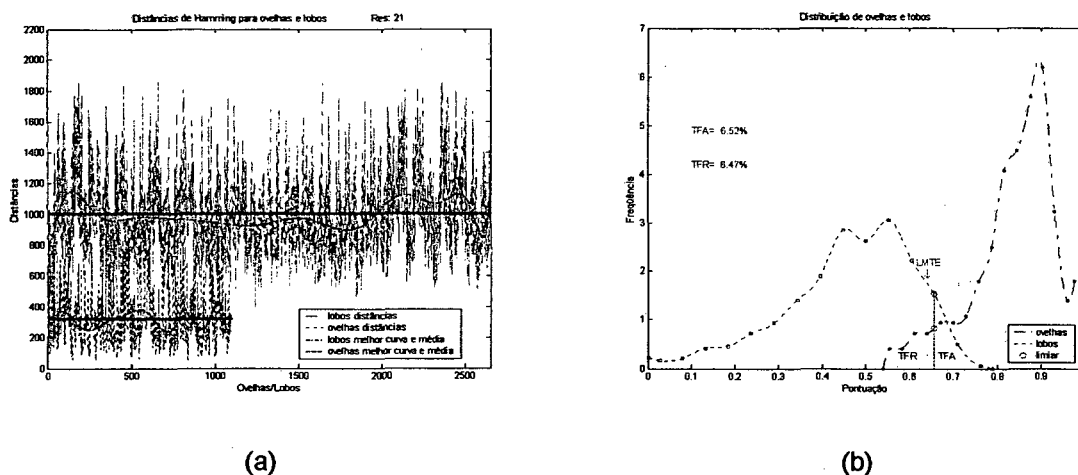


Figura 7.10 Distribuição da DH para os dados com OPF

A figura 7.11 mostra a distribuição das DH das populações de lobos e ovelhas para os dados originais submetidos normalização espacial e à ponderação facial (OH).

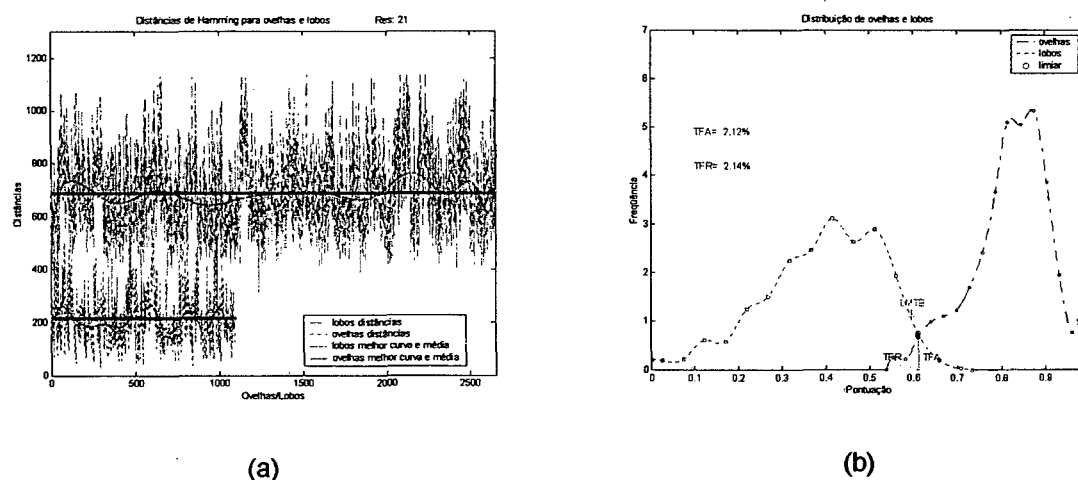


Figura 7.11 Distribuição da DH para os dados com OH.

Note-se que, a cada pré-processamento introduzido, as taxas TFR e TFA diminuem, culminando com a obtenção do melhor desempenho do sistema através do uso da Otimização Híbrida.

Apesar dos dados das DH's possuírem pontos que pertencem a uma faixa comum, pode-se ver na figura 7.11 que as médias das duas distribuições de DH's estão bastante separadas e as taxas TFR e TFA com valores em torno de 2% viabilizam o método como classificador de faces.

O Valor Limite - LMTE pode ser visto na figura 7.12(a), o qual para a situação estabelecida neste trabalho de taxas iguais (TFR=TFA), possui um valor aproximado de 0,68. A curva de performance, que estabelece o ponto de operação do sistema, pode ser vista na figura 7.12(b).

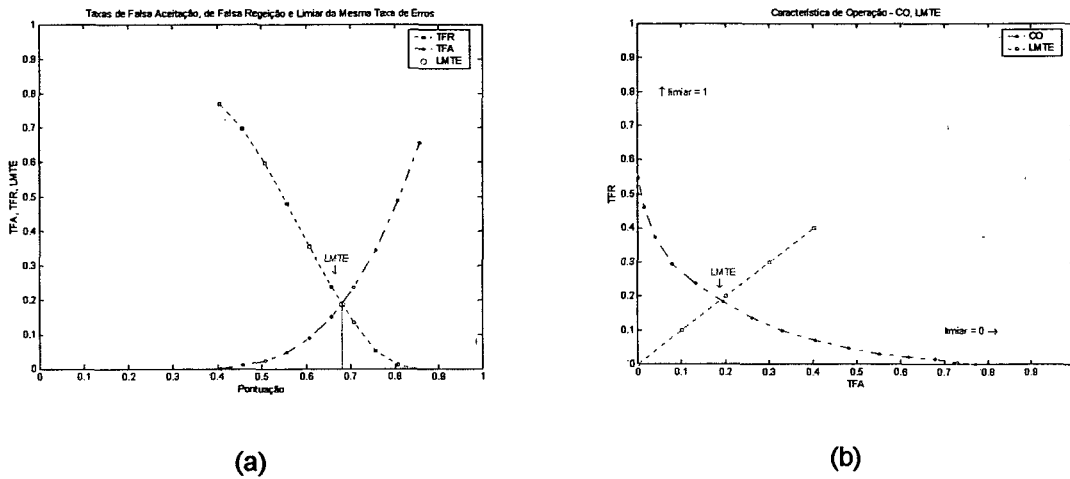


Figura 7.12 Caracterização do ponto de igualdade das taxas TFR e TFA.

**Análise da Resolução das Formas de Faces Ótima**

Nas figuras 7.11(a), 7.13(a) e 7.13(b) pode-se ver o comportamento dos dados das DH's em função da resolução da matriz de dados das faces.

A escolha do tipo de resolução da matriz de formas de faces se dá em função do (i) custo computacional e (ii) das taxas de erro TFR e TFA. Para os dados de formas de faces deste ensaio, tanto em termos de custo computacional como em termos das taxas de erros que representam a capacidade do sistema em separar corretamente as duas classes, a resolução 21x15, figura 7.11(a), mostrou possuir melhores características em ambos requisitos, sendo assim escolhida e utilizada neste trabalho em termos de RBC.



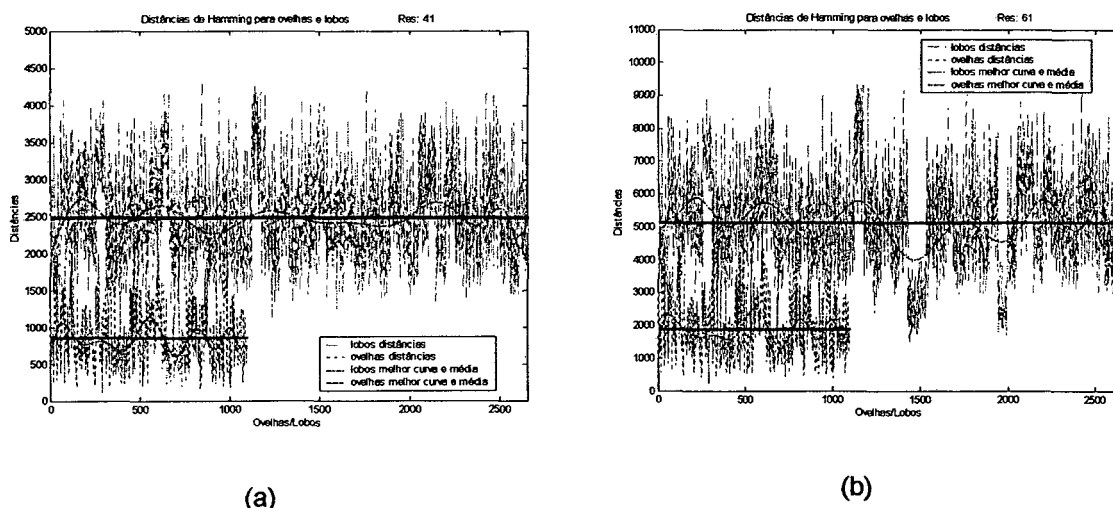


Figura 7.13 Distribuições das DH's para resoluções de (a) 41x29 e (b) 61x43.

### 7.3.3 Ensaios de Verificação Através de RNAs do tipo RBF

Estes ensaios objetivam verificar o desempenho das RNAs do tipo RBF na separação das populações de ovelhas e lobos. O modelo de treinamento utilizado neste trabalho é o dos Mínimos Quadrados Ortogonais - MQO de Chen [Chen et al., 1991].

Neste caso de classificação por RNA está implícita a utilização de uma parte dos dados como conjunto de treinamento. Desta forma o conjunto de validação diminui em relação ao paradigma anterior por RBC.

A figura 7.14 mostra uma curva típica de treinamento de uma RNA tipo RBF por MQO obtida para um dos 52 indivíduos. Note-se que o erro de treinamento decai com o aumento do número de neurônios na camada intermediária, atingindo um valor bastante baixo quando se usa um número de neurônios igual ao número de exemplos do treinamento. Nesta mesma figura pode ser observado o comportamento do erro de treinamento em função do raio da FBR. O erro de treinamento é tanto menor quanto menor for o raio utilizado.

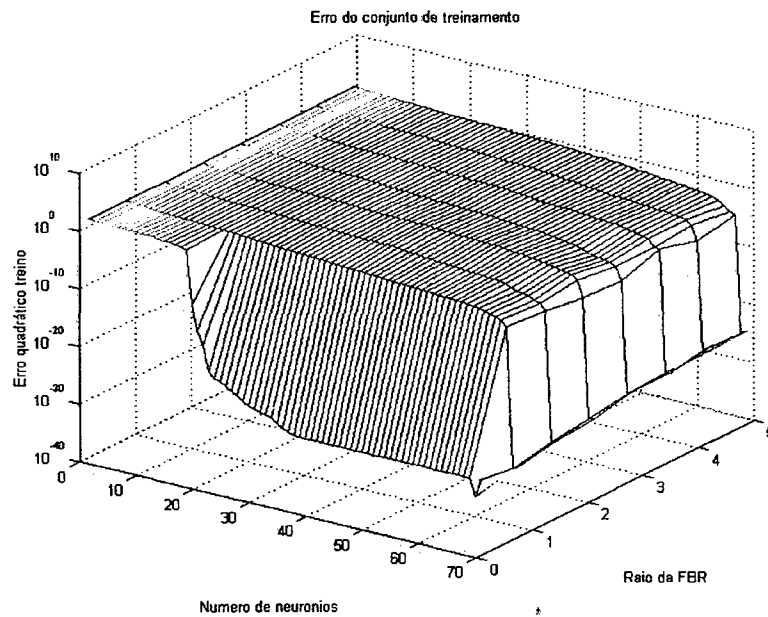


Figura 7.14 Comportamento do erro de treinamento em função do Raio da FBR e do número de neurônios.

Na figura 7.15 pode ser visto o comportamento do erro de validação em função do raio da FBR e do número de neurônios utilizados.

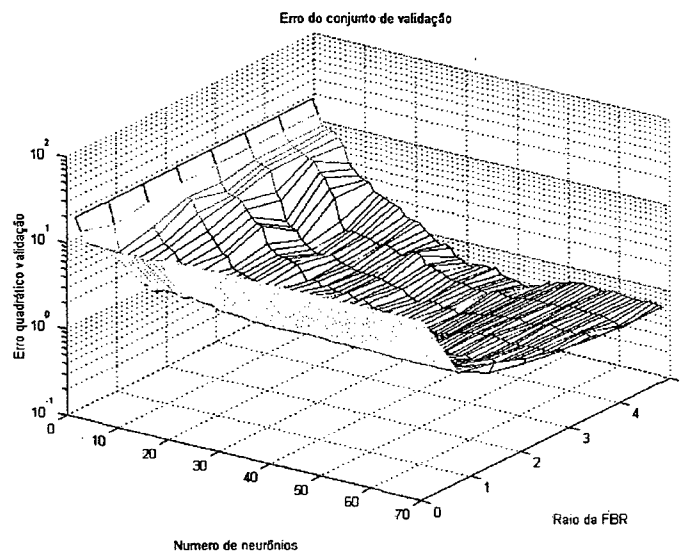


Figura 7.15 Comportamento do erro de validação em função do raio da FBR e do número de neurônios.

Comparando as figuras 7.14 e 7.15 pode-se observar que a melhor resposta da rede para o conjunto de validação se dá com valores do raio da FBR maiores que os de melhor resultado no treinamento. Também em termos de números de neurônios, a melhor resposta utiliza valores menores que o número máximo igual ao número de exemplos. Este comportamento é esperado, conhecido e ilustra a capacidade de generalização da RNA, já tratada na seção 6.2.4.

Enquanto que o número de neurônios pode ser previamente limitado como sendo igual ao número máximo de exemplos de treinamento, o raio da FBR pode ser estimado calculando-se as distâncias Euclidianas entre os centros das FBRs, conforme pode ser visto na figura 7.16 para um indivíduo típico. Uma distância menor que a metade da distância com maior frequência pode ser utilizada para o valor do raio da FBR.

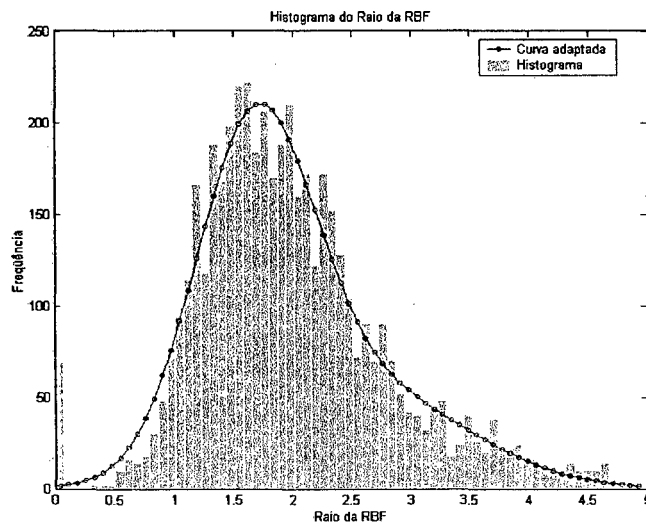


Figura 7.16 Histograma das distâncias entre centros dos pontos do conjunto de treinamento.

Com o objetivo de melhorar a invariância da RNA (seção 4.3), em relação ao conjunto das ovelhas, foram escolhidas como conjunto de treinamento as variações com valores mais extremos, ficando as de valores intermediários para o conjunto de validação. Com este procedimento resgata-se na etapa de treinamento, a propriedade de interpolação entre casos das RNAs.

### **A Classificação com Todas Variações**

Com o objetivo de se fazer uma comparação direta com os resultados da classificação por RBC, vistos na seção 7.3.2, utilizou-se aqui os mesmos elementos da base de dados daqueles ensaios. Neste ensaio, porém, foram acrescentadas mais dez variações de formas de faces.

Estas variações são correspondentes às de número [23 24 25 26 27 28 29 30 31 32] da base de dados e correspondem às variações de rotação em torno do eixo "z" e escalonamento das formas, (seção 4.3). Neste paradigma de classificação existe a necessidade da utilização de um conjunto de treinamento o que faz com que o conjunto de validação fique diminuído.

Neste ensaio, utilizaram-se os 52 indivíduos da base de formas de faces original, dividindo-se as populações em:

- i. **Conjunto de Treinamento da População de Ovelhas:** Foram utilizados 18 variações [1 3 5 7 9 10 11 13 15 17 18 20 22 23 25 27 28 30 32] e nível de ruído zero, para cada indivíduo (seção 4.3). O número total de indivíduos deste conjunto totaliza 936 (18x52) classificações.
- ii. **Conjunto de Treinamento da População de Lobos:** Foi utilizada uma única variação de cada indivíduo, variação [1] (seção 4.3), e utilizados todos os indivíduos da base, totalizando 2.652 (51x52) classificações.
- iii. **Conjunto de Validação da População de Ovelhas:** Foram utilizados 12 variações [2 6 8 12 14 16 19 21 24 26 29 31] e nível de ruído zero, para cada indivíduo (seção 4.3). O número total de indivíduos classificados totaliza 624 (12x52).
- iv. **Conjunto de Validação da População de Lobos:** Foi utilizada uma única variação de cada indivíduo, variação [8] (seção 4.3), e utilizados todos os indivíduos da base, totalizando 2.652 (51x52) classificações.

A seguir na figura 7.17 são mostrados os gráficos dos resultados da classificação.

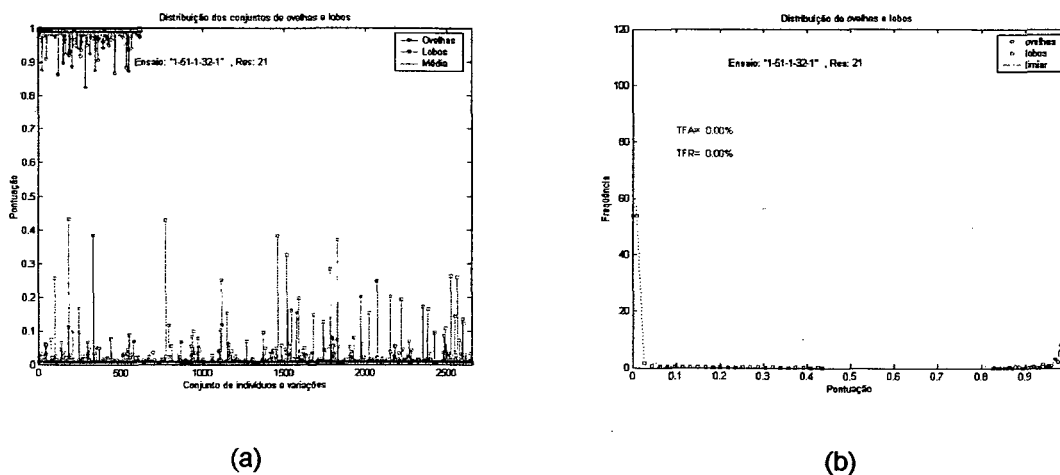


Figura 7.17 Resultado da classificação dos lobos e ovelhas usando RNAs do tipo RBF

A figura 7.18 mostra a distribuição da pontuação da classificação das ovelhas e lobos.

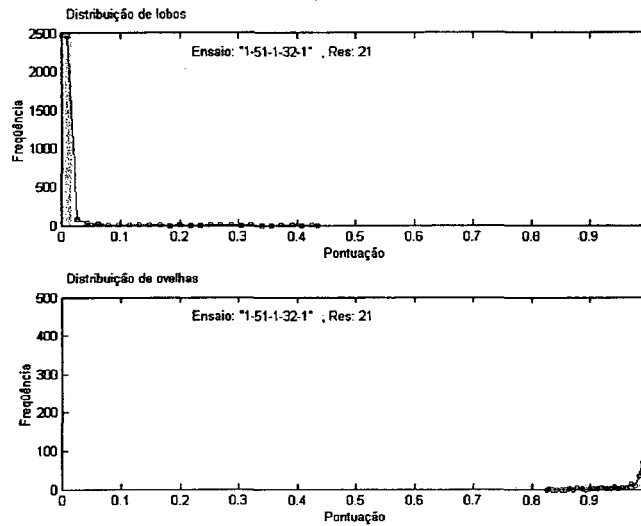


Figura 7.18 Histograma da classificação das ovelhas e lobos.

Na figura 7.19 podem-se ver os valores do raio da FBR e o número de neurônios utilizados neste ensaio. Note que os raios das FBRs, apesar de constante para um mesmo indivíduo, variaram de indivíduo para indivíduo numa faixa que em média varia entre 1 e 6.

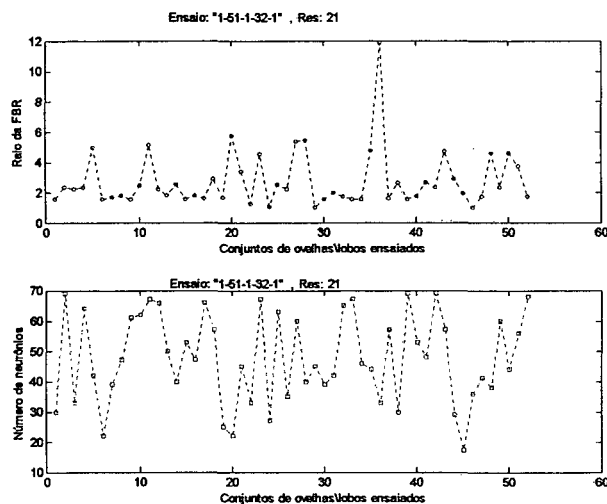


Figura 7.19 Raio da FBR e o número de neurônios utilizados no ensaio.

O número de neurônios utilizado é em média menor que o número máximo de neurônios dados pelo número de exemplos apresentados à rede que foi de 69 (18 ovelhas + 51 lobos). Isto

mostra que o erro requerido pode ser atingido com um número menor de neurônios, situação que diminui o custo computacional e favorece a generalização das RNAs do tipo RBFs.

A figura 7.20 mostra uma forma alternativa de mostrar a distribuição dos resultados da classificação das ovelhas e lobos que permite o cálculo da reta do Melhor Ajuste Linear - MAL e o respectivo Coeficiente de Correlação - CC entre esta reta e a reta ideal de 45 graus.

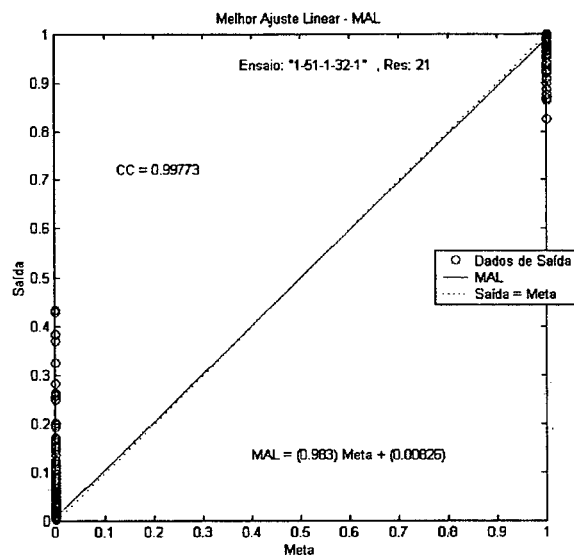


Figura 7.20 Melhor Ajuste Linear - MAL e Coeficiente de Correlação - CC entre os dados reais e os ideais.

Conforme pode ser visto na figura 7.17 os valores das TFR e TFA são zero, demonstrando o bom desempenho da rede como classificadora para as formas de faces utilizadas.

### **A Classificação com Dimensionalidade Reduzida**

No item anterior, mostraram-se os resultados da classificação para todas as variações de cada indivíduo, mas com a dimensionalidade normal de 315 (21x 15). Conforme já abordado na seção 6.3.2 a importância da redução da dimensionalidade reside no fato de se diminuir o custo computacional do processo de classificação. Neste trabalho aplicou-se a ACP como paradigma de redução da dimensionalidade dos dados.

Este ensaio teve por objetivo avaliar os resultados do emprego da ACP no processo e classificação das formas das faces. Para isto, escolheu-se 6 indivíduos (os seis primeiros por ordem alfabética), sendo que as conclusões dos resultados obtidos para estes indivíduos podem ser estendidas a todos os 52 indivíduos da base de dados. Para os dados normais, sem

---

aplicação da ACP figuras 7.21(a) e 7.20(a), os conjuntos de treinamento e validação deste ensaio ficaram assim constituídos:

- i. **Conjunto de Treinamento da População de Ovelhas:** Foram utilizadas 18 variações e nível de ruído zero, para cada um dos seis indivíduos (seção 4.3). O número total de indivíduos-ovelhas deste conjunto totaliza 108 (18x6) classificações.
- ii. **Conjunto de Treinamento da População de Lobos:** Foi utilizada uma única variação de cada um dos 52 indivíduos, variação [1] (seção 4.3), e utilizados todos os seis indivíduos da base, totalizando 306 (51x6) classificações.
- iii. **Conjunto de Validação da População de Ovelhas:** Foram utilizadas 12 variações e nível de ruído zero, para cada um dos seis indivíduos (seção 4.3). O número total de indivíduos classificados totaliza 72 (12x6).
- iv. **Conjunto de Validação da População de Lobos:** Foi utilizada uma única variação de cada indivíduo, [8] (seção 4.3), e utilizados todos os seis indivíduos da base, totalizando 306 (51x6) classificações.

Para os dados com aplicação da ACP figuras 7.21(b) e 7.20(b), os conjuntos de treinamento e validação ficaram assim constituídos:

- i. **Conjunto de Treinamento da População de Ovelhas:** Foram utilizadas 22 variações e nível de ruído zero, para cada um dos seis indivíduos. O número total de indivíduos deste conjunto totaliza 132 (22x6) classificações.
- ii. **Conjunto de Treinamento da População de Lobos:** Foram utilizadas 52 indivíduos [1] e utilizados todos os seis indivíduos da base, totalizando 372 (62x6) classificações.
- iii. **Conjunto de Validação da População de Ovelhas:** Foram utilizadas 12 variações e nível de ruído zero, para cada um dos seis indivíduos. O número total de indivíduos-ovelhas classificados totaliza 72 (12x6).
- iv. **Conjunto de Validação da População de Lobos:** Foi utilizada uma única variação de cada um dos 52 indivíduos [8], e utilizados todos os seis indivíduos da base, totalizando 306 (51x6) classificações.

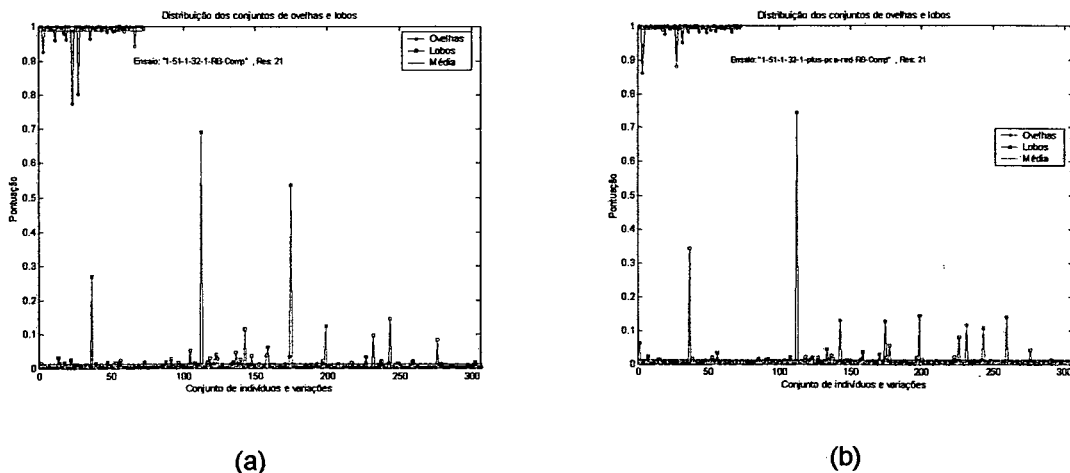


Figura 7.21 Classificação das ovelhas e lobos para seis indivíduos (a) dados normais e (b) dados reduzidos pela ACP.

Note que existem pequenas diferenças entre as classificações apresentadas pelas figuras 7.17(a) e 7.21(a). Estas diferenças residem no(a):

- i. Número de indivíduos: Os dois ensaios têm um número de indivíduos diferentes.
- ii. Número de neurônios da camada intermediária: O primeiro ensaio possui um número de neurônios específico para cada indivíduo. Neste ensaio o número de neurônios foi limitado em 30.
- iii. Raio da FBR: O raio da FBR possui valores diferentes entre os dois ensaios.

A figura 7.20 mostra as taxas TFR e TFA dos dois ensaios.

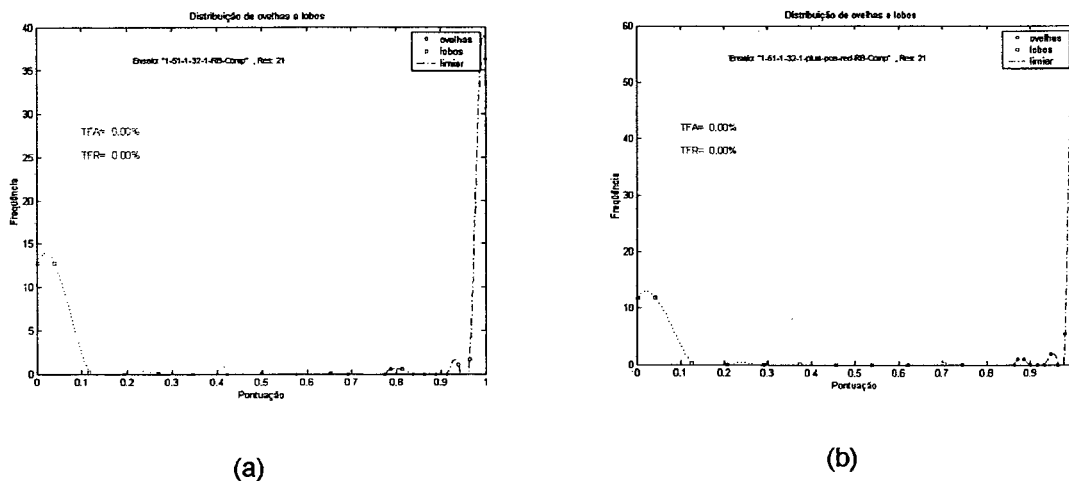


Figura 7.22 Distribuição das ovelhas e lobos para seis indivíduos e suas taxas de TFR e TFA para (a) dados normais e (b) dados reduzidos pela ACP.



Apesar de não haver significativas diferenças entre as taxas e erros nas classificações das populações normais e com ACP, a principal diferença está no custo computacional. Enquanto para a situação dos dados normais, onde os dados de entrada eram matrizes com dimensionalidade 315, nos dados com a aplicação de ACP a dimensionalidade dos dados se reduziu para 36, portando uma redução significativa de aproximadamente de 8 vezes.

### **Influência do Tamanho do Conjunto de Treinamento**

O objetivo deste ensaio é o de analisar qual o tamanho mínimo do conjunto de treinamento para se conseguir o erro requerido. Neste ensaio, a exemplo do ensaio anterior, foram utilizados os 6 primeiros indivíduos da base de forma de faces (ordem, alfabética). As populações de ovelhas e lobos utilizadas neste ensaio ficaram distribuídas como:

- i. **Conjunto de Treinamento da População de Ovelhas:** Foram utilizados 18 variações [3 6 9 12 15 18] e nível de ruído zero, para cada um dos seis indivíduos.
- ii. **Conjunto de Treinamento da População de Lobos:** Foram utilizadas 52 indivíduos, variação [1], e utilizados todos os seis indivíduos da base.
- iii. **Conjunto de Validação da População de Ovelhas:** Foram utilizados 12 variações [2 6 8 12 14 16 19 21 24 26 31] e nível de ruído zero, para cada um dos seis indivíduos.
- iv. **Conjunto de Testes da População de Lobos:** Foram utilizados uma única variação de cada um dos 52 indivíduos, variação [8], e utilizados todos os seis indivíduos da base.

Na figura 7.23 pode-se ver que o comportamento do erro quadrático para o conjunto de validação, na saída da rede, em função do número de ovelhas usadas no treinamento. A figura 7.23 mostra o comportamento do erro para cada um dos seis indivíduos e também a curva média entre eles.

É importante notar que, neste trabalho, as variações utilizadas para treinamento foram selecionadas de modo a situarem-se como casos extremos, ficando as variações intermediárias para validação. Desta forma a rede aprende com os casos extremos e é validada com os casos intermediários. Assim na figura 7.23 o erro de validação diminui não só pelo aumento do número de exemplos no treinamento, mas principalmente pela inclusão gradativa no treinamento de formas de faces que situam-se como variações ou exemplos extremos os quais proporcionam à rede a possibilidade de desempenhar seu papel de interpoladora intercasos.

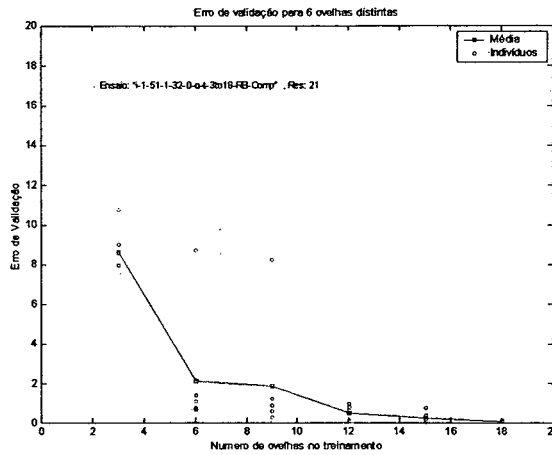


Figura 7.23 Comportamento do erro de validação em função do número de exemplos de treinamento apresentados a rede.

Nas figuras 7.24 e 7.25 pode-se observar o comportamento da rede em termos do número de neurônios necessários para atingir os erros apresentados na figura 7.23. Note que na fase inicial da curva média, à medida que o número de exemplos, cresce o número de neurônios também cresce. O decaimento do número de neurônios na parte final da curva pode ser explicado pelo aumento do número de exemplos que contemplam o fechamento dos casos extremos o que beneficia a interpolação acima citada.

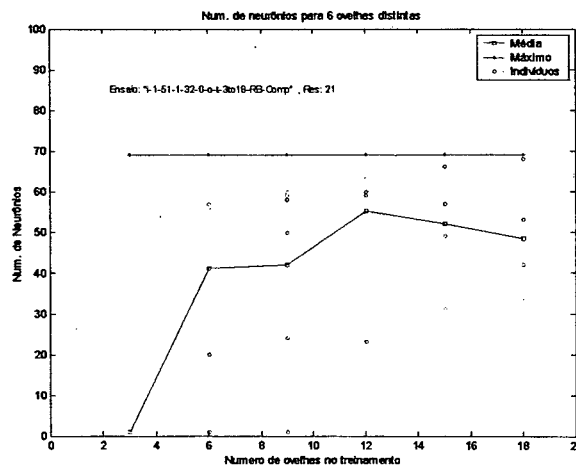


Figura 7.24 Comportamento do número de neurônios da camada intermediária em função do número de ovelhas no treinamento da rede.

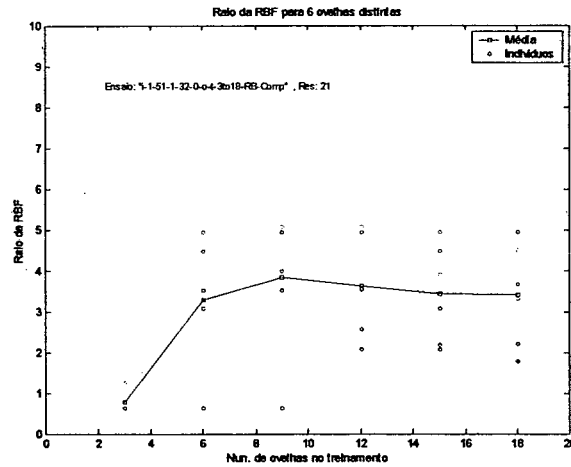


Figura 7.25 Comportamento do raio da FBR em função no número de ovelhas no treinamento da rede.

Na figura 7.25 pode-se ver o comportamento dos valores dos raios das FBRs em função do número de ovelhas no treinamento. O objetivo desta observação é o de conhecer a faixa de valores do raio das FBRs deste ensaio. Nota-se que, com poucos exemplos de treinamento, o valor do raio é pequeno para os requisitos de erro e, à medida que o número de exemplos cresce o raio da FBR cresce até atingir uma faixa entre 3 e 4, permanecendo assim até o número máximo de exemplos deste ensaio.

## Capítulo 8

# Conclusões e Perspectivas

---

*Os Problemas que existem no mundo de hoje não podem ser resolvidos pelo nível de pensamento que os criou.*

- Albert Einstein, (1879-1955).

**E**ste trabalho teve por objetivo a validação de um paradigma para o reconhecimento de faces humanas. Numa visão sistemática podem-se dividir os esforços despendidos neste trabalho como os de desenvolvimento de duas tecnologias distintas: (i) a de extração dos dados da face humana e (ii) a de reconhecimento das faces.

### 8.1 Conclusões Sobre a Extração de Dados das Faces

O modelo de dados de faces escolhido foi o tridimensional, por apresentar vantagens sobre o modelo bidimensional como a possibilidade de correção do mau posicionamento espacial, imunidade às variações de iluminação sobre a face e principalmente por ser a tecnologia de

Medição de formas não é uma tecnologia nova, mas sua aplicação está sendo cada vez mais viabilizada tendo em vista o advento de computadores com maior capacidade de processamento.

O paradigma de extração de formas utilizado, conhecido como "perfilometria de Fourier" [Takeda e Mutoh, 1999], oferece como principal vantagem a simplicidade. A aquisição da forma é realizada através de uma única câmera através de uma só imagem, conferindo rapidez ao sistema.

Do ponto de vista biométrico, o paradigma de reconhecimento de faces tem o maior índice de aceitabilidade pelos usuários pois além de não ser agressivo, extremamente passivo, é o método que os humanos utilizam para o reconhecimento de faces.

No processo de extração das formas de faces o algoritmo de processamento das franjas necessita ser aperfeiçoado no sentido de se tornar mais robusto diante de casos de obstrução das franjas, como ocorre em algumas situações na região em torno do nariz e também em casos onde ocorre mesclagem das franjas escuras com regiões escuras da face como por exemplo sobrancelhas pretas.

Apesar do sistema de medição de formas desenvolvido (seção 7.2.2) ter sido calibrado de forma parcial e não exaustivamente em todo o campo de medição, o que seria bastante trabalhoso, os resultados da classificação obtidos neste trabalho, mostraram que o sistema atende bem as especificações dos ensaios propostos.

A base de dados foi construída com 52 indivíduos diferentes, cada um com 32 variações entre expressões e posições espaciais, com raças, sexo, cor e idade variadas, o que forneceu uma base de faces suficientemente rica para a proposição deste trabalho.

A resolução de 21 x 15 da matriz de formas utilizada mostrou-se suficiente para uma completa separação de todas as ovelhas em relação à população de lobos no ensaio apresentado no capítulo 7.

## **8.2 Conclusões sobre o Reconhecimento das Faces**

Neste trabalho, implementaram-se dois paradigmas de inteligência artificial para a tarefa de classificação das faces: (i) o RBC e o por (ii) RNAs tipo FBR. O objetivo desta dupla implementação foi o de poder avaliar de forma relativa o desempenho das duas técnicas.

A classificação por RBC é a técnica mais simples e natural exigindo porém que sejam realizados pré-processamentos nos dados da forma das faces para que os resultados da

classificação sejam maximizados. Dentre os pré-processamentos necessários pode-se citar a otimização por normalização espacial e a otimização por ponderação facial. A contribuição de cada uma das otimizações fica clara observando-se os resultados das classificações nas figuras 7.8, 7.9, 7.10, 7.11, onde se vê que o melhor desempenho é atingido com a implementação híbrida de ambas otimizações. As menores taxas de erro TFA e TFR, para os ensaios realizados aplicando-se a otimização híbrida, atingiram valores em torno de 2% os quais representam desempenhos razoáveis para sistemas biométricos.

Já a classificação por RNAs do tipo RBFs, com treinamento por mínimos quadrados ortogonais, conforme mostra a figura 7.17, mostrou desempenho bastante superior ao da classificação por RBC, tendo para uma ampla faixa de limites possíveis taxas de erro do tipo TFA e TFR com valores iguais a zero. Note que este desempenho foi atingido com as formas de faces originais, isto é, sem a aplicação de qualquer otimização, seja de posição espacial ou por ponderação facial. Fato que vem a demonstrar o bom desempenho, já esperado, do paradigma de RNAs em reconhecimento de padrões, mais especificamente na classificação das formas de faces humanas.

Uma melhoria no desempenho das RNAs poderia ser conseguida aplicando-se algum tipo de otimização nos dados de formas de faces, a exemplo do que foi aplicado nas formas de faces no paradigma do RBC. Porém neste caso deve-se avaliar a relação custo / benefício envolvido tendo em vista que a implementação de uma otimização de posição espacial e ou ponderação facial acarreta aumento do custo computacional no sistema.

A invariância dos resultados da RNA é assegurada através de um treinamento controlado, onde se escolheu os casos extremos de posições espaciais e expressões faciais para treinamento, deixando os intercasos para o conjunto de validação. Isto pode ser observado na figura 7.23 onde o erro diminui, atingindo o seu valor mínimo quando são usados todos os exemplos de treinamento (casos extremos).

A redução de dimensionalidade proporcionada através da aplicação da ACP, de 315 para 36, proporcionou uma redução expressiva nos vetores de entrada das formas de faces de aproximadamente 8 vezes, sem alteração significativa dos resultados da classificação, vide figura 7.21. Isto se traduz em maior velocidade e diminuição dos recursos de máquina necessários ao processamento.

### 8.3 Considerações Gerais

Para uma comparação de desempenho entre sistemas de reconhecimento de faces humanas, fica evidente a necessidade do(a):

- i. Estabelecimento de um amplo protocolo que permita a aplicação das mesmas regras de ensaio, parâmetros e avaliação de desempenho, tanto para sistemas bidimensionais quanto para sistemas tridimensionais.
- ii. Criação de uma base de dados bivalente comum as duas tecnologias que use os mesmos indivíduos e os mesmos graus de dificuldade, tanto para sistemas bidimensionais como para tridimensionais.

## 8.4 Principais Contribuições deste Trabalho

As principais contribuições deste trabalho foram:

- i. Chamar a atenção para a importância de um protocolo geral de ensaios que permita a intercomparação de sistemas de reconhecimento de tecnologias diversas.
- ii. Analisar e sistematizar os principais métodos de reconhecimento de padrões e de reconhecimento facial existentes bem como suas aplicações.
- iii. Analisar os principais sistemas de medição de geometria existentes e adaptar o sistema de extração de formas para a aquisição de formas de faces humanas.
- iv. Construir uma base de dados de formas de faces robusta com 52 indivíduos diferentes, com 32 variações de cada um, em 9 resoluções diferentes e 5 níveis de ruído, totalizando 74.880 formas de faces.
- v. Apresentar o conceito de RBC, a otimização espacial desenvolvida para o cálculo da similitude entre faces e a introdução do conceito de ponderação facial.
- vi. Apresentar os parâmetros importantes das RNAs para o reconhecimento de padrões e/ou faces como: as otimizações de projeto quanto à invariância das RNAs e a dimensionalidade do sistema.
- vii. Comparar o desempenho entre os métodos de classificação por RBC e por RNA, através de uma implementação prática com dados reais, da teoria desenvolvida.
- viii. Propor assuntos afins como perspectivas de trabalhos futuros.

## 8.5 Perspectivas para Trabalhos Futuros

São apresentados a seguir tópicos de estudos e ensaios a serem realizados com o objetivo de complementar e aprimorar o assunto abordado neste trabalho:

- i. Um estudo mais detalhado e aprofundado sobre os erros de medição do sistema proposto neste trabalho, tomando-se p. ex. como padrão uma máquina de medir por coordenadas permitiria uma maior confiabilidade quanto aos aspectos metrológicos e operacionais do processo de extração das formas e poderia ajudar na interpretação dos

- resultados do classificador, quando este, p. ex. apresentar resultados inconsistentes ou atípicos.
- ii. Melhorar o desempenho do sistema de extração de formas de faces, quanto a problemas relacionados com a remoção do salto de fase, de modo que o mesmo possa operar de forma totalmente automática.
  - iii. Implementar outros métodos de treinamento das RNAs do tipo FBRs, como por exemplo o totalmente supervisionado, ou outro que atribua um valor diferenciado aos Raios da FBR, com o objetivo de comparar com o método utilizado neste trabalho.
  - iv. Análise de outros paradigmas de RNAs para que os resultados sejam comparados aos obtidos neste trabalho através da RNA do tipo FBR.
  - v. Outras técnicas de redução de dimensionalidade como LDA podem ser usadas, com o objetivo de redução de custos computacionais, e ter seus resultados comparados com os da ACP utilizados neste trabalho.
  - vi. A técnica de redução de dimensionalidade por ACP aplicadas a faces de duas dimensões chamada de "faces próprias" pode ser aplicada às formas de faces, aqui chamada de "formas próprias"<sup>2</sup>, tanto para a redução de dimensionalidade como para a classificação em si, buscando-se minimizar os custos computacionais.
  - vii. Utilização de uma técnica híbrida para a obtenção dos dados de faces, que combine os dados de forma e níveis de cinza destas, com o objetivo de aumentar a singularidade dos dados e aproximar o modelo de extração a aquele que os humanos utilizam.
  - viii. Analisar o impacto nas taxas de erro e custo computacional devido à utilização de resoluções menores que a de 21x15 utilizada neste trabalho.
  - ix. Uso de técnicas não convencionais para a determinação da arquitetura de redes como o uso de algoritmos genéticos e de computação evolucionária.
  - x. Implementação de um eficiente mecanismo de busca em base de dados com o objetivo de transformar o sistema de verificação desenvolvido neste trabalho em um sistema de identificação.
  - xi. Análise de uma aplicação híbrida, onde a técnica de reconhecimento de faces humanas é acrescida de mais uma técnica biométrica, como p.ex. voz, com o objetivo aumentar a confiabilidade do sistema, diminuindo ainda mais as taxas TFA e TFR.

---

<sup>2</sup> "Eigenshapes".



---

E assim, através dos estudos desenvolvidos neste trabalho, acredita-se ter conseguido dar uma pequena colaboração no sentido de esclarecer a importância das principais variáveis envolvidas no processo de verificação de faces humanas tridimensionais, com ênfase na extração da forma e da classificação destas através de métodos de inteligência artificial.

## Referências

1. AAMODT, A. & PLAZA, E. Case Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. 2002.
2. AAMODT, A. & PLAZA, E. Case Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches. vol. 7, p. 39-59. 1994.
3. ABDI, H., VALENTIN, D. & O'TOOLE, A. J. A generalized autoassociator model for face processing and sex categorization: From principal components to multivariate analysis. *Optimality in biological and artificial networks?* Richardson, U.S.A. 1998.
4. ALLISON, N. M., ELLIS, A. W., FLUDE, B. M. & et al. A connectionist model of familiar face recognition. In: *IEE Colloquium on Machine Storage and Recognition of Faces*. v.017, no. 5. 1992. p. 1-10.
5. ALTERMAN, R. Panel discussion on case representation. Pensacola Beach, FL, US. 1989.
6. ANDERSON, D. & MCNEILL, G. *Artificial Neural Networks Technology*. Utica, NY, USA, Rome Laboratory, Aug/1992.
7. ANDERSON, E., BAI, Z., BISHOP, C. & et al. *LAPACK Users' Guide Third Edition*. Philadelphia, PA, USA, SIAM, 1999. ISBN: 0-89871-447-8.
8. ARAD, N., DYN, N., REYSFELD, D. & et al. Image Warping by Radial Basis Functions: Application to Facial Expressions. vol. 2, no. 56, p. 161-172. 1994.
9. ASUNDI, A. & WENSEN, Z. Fast phase-unwrapping algorithm based on a gray-scale mask and flood fill. 1999.
10. AZEVEDO, F. M. D., BRASIL, L. M. & OLIVEIRA, R. C. L. d. *Redes neurais com aplicações em controle e em sistemas especialistas*. 1ª ed. Florianópolis, SC, Brasil, Visual Books, Sep/2000. ISBN: 85-7502-005-6.
11. BALA, J., DE JONG, K., HUANG, J. & et al. Using Learning to Facilitate the Evolution of Features for Recognizing Visual Concepts. In: *ECJ97*. Fairfax, VA, USA, 1998.
12. BALA, J., HUANG, J. & VAFAIE, H. Hybrid Learning Using Genetic Algorithms and Decision Trees for Pattern Classification. In: *International Joint Conference on Artificial Intelligence - IJCAI95*. Fairfax, VA, USA, Aug/1995.
13. BARNARD, E. & CASASANT, D. Invariance and Neural Nets. Etienne Barnard and David Casasent. p. 498-508. 1991.
14. BARRETO, J. M. *Inteligência Artificial no limiar do século XXI - Abordagem híbrida*. 2ª ed. Florianópolis, Sc, Brasil, Duplic Prestação de Serviços, 1999. ISBN: 85-900382-4-6.
15. BARRETO, J. M. *Inteligência Artificial no limiar do século XXI - Abordagem Híbrida Simbólica Conexionista e Evolutiva*. 2ª ed. Florianópolis, Sc, Brasil, Duplic Prestação de Serviços, 2001. ISBN: 85-900382-5-4.

16. BELHUMEUR, P.N., HESPANHA, J.P. & KRIEGMAN, D.J. Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. 1999.
17. BENOUDJIT, N., ARCHAMBEAU, C., LENDASSE, A. & et al. Width optimization of the Gaussian kernels in Radial Basis Function Networks. In: *European Symposium on Artificial Neural Networks*. Bruges, Belgica, Apr/2002. p. 425-432.
18. BENSON, P. & PERRETT, D.I. Face to Face with the perfect image. p. 32-35. Feb 22, 1992.
19. BENSON, P.J. & PERRETT, D.I. Perception and recognition of photographic quality facial caricatures: implications for the recognition of natural images. vol. *Face Recognition*, p. 105-135. 1991.
20. BIEMAN, L. Three-Dimensional Machine Vision. no. May, p. 81- 1988.
21. BISHOP, C. M. *Neural Networks for pattern recognition*. Oxford University press, 1995.
22. BOSE, N. K. & LIANG, P. *Neural network fundamentals with graphs, algorithms and applications*. McGraw-Hill Inc, 1996. ISBN: 0 07 006618 3.
23. BOTTEMANNE, BARRETO, J. M. & ROUCOUX, A. Control of Visual Sensors with a Physiologically Inspired Neural Network. In: *IMACS Symposium on Mathematical Modeling Technical University of Vienna*. Austria, Feb/1994.
24. BRASIL, L.M., AZEVEDO, F.M.D. & BARRETO, J.M. Learning Algorithm for Connectionist Systems. 1999.
25. BRUNELLI, R. & POGGIO, T. Face Recognition: Features versus Templates. Trento, ITALY. vol. 10, no. 15, p. 1024-1052. 1993.
26. CAO, J., AHMADI, M. & SHRIDHAR, M. A Hierarchical Neural Network Architecture For Handwritten Numeral Recognition. Windsor, Ontario, Canada. 1999.
27. CHANDRASEKARAN, R. ID Systems using the human body are here, but privacy issues persist. Mar 30, 1997.
28. CHELLAPPA, R., WILSON, C.L. & SIROHEY, S. Human and Machine Recognition of Faces: A Survey. no. 83, p. 704-740. 1995.
29. CHEN, S., COWAN, C.F.N. & GRANT, P.M. Orthogonal least squares learning algorithm for radial basis function networks. vol. 2, no. 2, p. 302-309. Mar, 1991.
30. CICHOCKI, A. & UNBENHABEN, R. *Neural networks for optimization and signal processing*. John Wiley & Sons inc, 1993.
31. COIMBRA, A.J.F., MARINO-NETO, J., FREITAS, C.G. & et al. Automatic Detection Of Sleep-Waking States Using Kohonen Neural Networks. Florianópolis, SC, Brazil. 1999.
32. COOTES, T.F., TAYLOR, C.J., COOPER, D.H. & et al. Training models of shape from sets of examples. p. 9-18. 1992.
33. COSTANTINI, M. A Phase Unwrapping Method Based on Network Programming. Frascati, Roma, Italy. 2016.
34. COTTRELL, G. W. & FLEMING, M. Face recognition using unsupervised feature extraction. In: *International Neural Network conference*. v.1, 1990. p. 322-325.
35. CRAW, I. Recognising face features and faces. vol. 017, no. 7, p. 1-4. 1992.
36. CRAW, I. & CAMERON, P. Face recognition by computer. In: *British Machine Vision Conference*. 1992. p. 488-507.
37. CRAW, I., TOCK, D. & BENNETT, A. Finding face features. Berlin. vol. *Proceedings ECCV92*, p. 92-96. 1992.
38. DAUGMAN, J. Face Recognition by Feature Demodulation. In: *IWAFGR95*. Zurich, Swiss, no. *Proceedings of the International Conference on Automatic Face and Gesture Recognition*. 1995. p. 110-115.

39. DE CAMPOS, T. E. Técnicas de Seleção de Características com Aplicações em Reconhecimento de Faces. São Paulo, SP, Brasil. Feb 25, 2001.
40. DE CARVALHO, A. P. D. L. F. Reconhecimento de Padrões. São Paulo, SP, Brasil. Aug 31, 2001.
41. DIAMOND, R. & CAREY, S. Why faces are and are not special: an effect of expertise. no. 115, p. 107-117. 1986.
42. DUDA, R. O., HART, P. E. & STORK, D. G. *Pattern Classification*. 2<sup>aa</sup> ed. Wiley-Interscience, Oct/2000. ISBN: 0471056693.
43. EDELMAN, S., REISFELD, D. & YESHURUN, Y. Learning to Recognize Faces from Examples. *Proceedings 2nd European Conference on Computer Vision*. Italy. p. 787-791. May, 1992.
44. ENCINAS, LEONARDO SOLIZ & ZIMMERMANN, A. C. *PCA Aplicado à Formas de Faces*. Florianópolis, SC, Brasil, 2002. Relatório de estudo/pesquisa.
45. FAUSETT, L. V. *Fundamentals of neural networks - Architectures, algorithms and applications*. New Jersey, USA. 1994.
46. FERAUD, R. PCA, Neural Networks and Estimation for Face Detection. In: *NATO*. Lannion, France, 1999.
47. FERREIRA, A. B. d. H. *Novo Dicionário da Língua Portuguesa*. 2a (18a)<sup>a</sup> ed. Rio De Janeiro, RJ, Nova Fronteira, 1986.
48. FOLEY, J. D., DAM, A., FEINER, S. K. & et al. *Computer graphic principles and practice*. Addison-Wesley Publishing Company, 1996. ISBN: 0 237 251 1.
49. FRANK, I. E. & FRIEDMANN, J. H. A statistical view of some chemometrics regression tools. vol. 35, p. 109-148. 1993.
50. FRITZ, H., HAUBKNECHT, Th. & NYÁRSIK, F. H. A Phase Unwrapping Algorithm with reduced Error or Propagation by Means of Pixel Validation with a Modulation Criterion. 1999.
51. FUKUMI, M., OMATU, S. & NISHIKAWA, Y. Rotation-invariant Neural Pattern Recognition System Estimating a Rotation Angle. In: *IEEE Transactions on Neural Networks*. 581, v.8, no. 3. 1997. -568.
52. FUKUSHIMA, K. Visual Feature extraction by a multilayered network of analog threshold elements. Oct, 1969.
53. FURNIVAL, G. M. & WILSON, R. W. Regression by Leaps and Bounds. vol. 16, p. 499-511. 1974.
54. GALLANT, S. L. *Neural Network Learning and Expert Systems*. Cambridge, Massachusetts, MIT Press, 1993.
55. GALTON, F. Personal identification and description. p. 173-177. Jun 21, 1888.
56. GARCIA, O. N. A New Method on Human Face Recognition - Neural Network Recognition Technique Based on Image Eigenvalues. 1999.
57. GASVIK, K. J. *Optical Metrology*. 2nd Edition<sup>a</sup> ed. Chichester, West Sussex, England, John Wiley & Sons, Inc., 1995. ISBN: 0471954748.
58. GHIGLIA, D. C., MASTIN, G. A. & ROMERO, L. A. Cellular-automata Method for Phase Unwrapping. Albuquerque, New Mexico, USA. 1999.
59. GHIGLIA, D. C. & ROMERO, L. A. Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods. 1999.
60. GOLDSTEIN, A. J., HARMON, L. D. & LESK, A. B. Identification of human faces. In: *Proceedings of the IEEE*. v.5, no. 59. 1971. p. 748-760.
61. GOLFARELLI, M., MAIO, D. & MALTONI, D. On the Error-Reject Trade-Off in Biometric Verification Systems. 1999.
62. GONÇALVES, A. A. J. *Análise de Mapas de Franjas*. Florianópolis, SC, Brasil, LABmetro, UFSC, 1991.

63. GÖPEL, W., HESSE, J. & ZEMEL, J. N. *Optical Sensors*. VHC, 1992. vol. 6, Sensors.
64. GUTTA, S., HUANG, J., SINGH, D. & et al. Benchmark Studies on Face Recognition. In: *IWAFGR95*. Fairfax, VA, USA, 1995.
65. GUTTA, S., HUANG, J., SINGH, D. & et al. The FERET Facial Database. In: *NIST95workshop*. Fairfax, VA, USA, 1998.
66. HAM, Y.K. & PARK, H.-H. 3D Object Recognition in Range Images Using Hidden Markov Models and Neural Networks. Seoul, South Korea. 1999.
67. HANCOCK, P. J. B., BURTON, A. M. & BRUCE, V. Preprocessing Images of Faces: Correlations with Human Perceptions of Distinctiveness and Familiarity. In: *IEE95*. Stirling, UK, 1995.
68. HANCOCK, P. J. B., BURTON, A. M. & BRUCE, V. Face processing: human perception and principal components analysis. In: *MC96*. Stirling, UK, 1996.
69. HARDING, K. Overview of 3D Contouring Systems for Reverse Engineering Applications. Ann Arbor, MI, USA. 1999.
70. HARMON, R.D., KHAN, M.K., LASH, R. & et al. Machine Identification of Human Faces. *Pattern Recognition*. vol. 13, no. 2, p. 97-110. 1981.
71. HAYKIN, S. *Neural Networks*. Englewood Cliffs, NJ, Macmillan College Publishing Company, Inc., 1994. ISBN: 0023527617.
72. HE, X. & LAPEDES, A. *Nonlinear modeling and prediction by successive approximation using radial basis functions*. Technical Report LA-UR-91-1375<sup>a</sup> ed. Los Alamos, NM, National Laboratory, 1991.
73. HENDERSON, J. M., FALK, R., MINUT, S. & et al. *Gaze Control for Face Learning and Recognition by Humans and Machines*. Eye Movement Laboratory Technical Report<sup>a</sup> ed. Michigan, EUA, Michigan State University, 2000.
74. HERRÁEZ, M.A., BURTON, D.R., LALOR, M.J. & et al. Robust, simple, and fast algorithm for phase unwrapping. Miguel Arevalillo Herráez David R. Burton, Michael J. Lalor, and David B. Clegg. 1999.
75. HORN, R. A. & JOHNSON, C. R. *Matrix Analysis*. Cambridge, UK, Cambridge University Press, 1985.
76. HOUAISS, A., VILLAR, M. d. S. & ET ALL *Dicionário Eletrônico Houaiss da Língua Portuguesa*. Editora Objetiva Ltda., Dec/2001.
77. HUANG, C.L. & CHEN, C.W. Human facial feature extraction for face interpretation and recognition. vol. 12, no. 25, p. 1435-1444. 1992.
78. HUANG, J., GUTTA, S. & WECHSLER, H. Detection of Human Faces Using Decision Trees. In: *2nd International Conference on Automated Face and Gesture Recognition (ICAFGR'96)*. Fairfax, VA, USA, 1996.
79. HUANG, J. & WECHSLER, H. Eye Location Using Genetic Algorithm. Fairfax, VA, USA. 1999.
80. HUANG, R.-H. J. *Detection Strategies For Face Recognition Using Learning And Evolution*. Fairfax, Virginia, USA, Department of computer science, George Mason University, May/1998. Dissertation of thesis degree.
81. HUNG, K.M. & YAMADA, T. Phase Unwrapping By Regions Using Least-Squares Approach. Singapore. 1999.
82. JAIN, A. D., DUIN, R. P. & FLYM, P. J. Statistical Pattern Recognition: A review. In: *IEE Transactions on Pattern Analysis and Machine Intelligence*. v.22, no. 1. 2000. p. 4-37.
83. JAIN, A. K., BOLLE, R. & PANKANTI, S. *Biometrics: Personel Identification in Networked Society*. Kluwer Academic Pub., 1999. ISBN: 0-7923-8345-1.
84. JOLLIFFE, I. T. *Principal Components Analysis*. New York, Springer-Verlag, 1986.
85. KANADE, T. Computer recognition of human faces. 1997.

86. KASS, M., WITKIN, A. & TERZOPOULOS, D. SNAKES: Active contour models. In: *Proceedings of the International Conference on Computer Vision*. 1987. p. 259-268.
87. KERR, J.R. Real Time Imaging Rangefinder for Autonomous Land Vehicles. vol. 1007, p. 349-1988.
88. KIMURA, F. & SHRIDHAR, M. Handwritten Numerical Recognition Based on Multiple Algorithms. 1999.
89. KIRBY, M. & SIROVICH, L. Application of the Karhunen-Loève procedure for the characterisation of human faces. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*. v.1, no. 12. 1990. p. 103-108.
90. KLIR, G. J. & YUANG, B. *Fuzzy sets and fuzzy logic - theory and applications*. Prentice Hall, 1995.
91. KOHONEN, T. *Self organization and associative memory*. 2ª ed. Springer verlag, 1988.
92. KOLODNER, J.L. *Case Base Reasoning*. 1993.
93. KONEN, W. & SHULZE-KRÜGER, E. ZN-Face: A system for access control using automated face recognition. In: *International Workshop on Automatic Face- and Gesture-Recognition - WAFGR95*. Bochum, Germany, 1995.
94. KOVACS, Z. L. *Redes Neurais Artificiais Fundamentos e Aplicações*. 1ª ed. São Paulo, OESP Grafica SA, 1996. vol. 1,
95. KRÄMER, R. & LOFFELD, O. Presentation of an improved Phase Unwrapping Algorithm based on Kalman filters combined with local slope estimation. Siegen, Germany. 2016.
96. KRUGER, R. & THOMPSON, W. A technical and economic assesment of computer vision for industrial inspection and robotic assembly. vol. 69, no. 12, 1981.
97. LANITIS, A., TAYLOR, C. J. & COTTES, T. F. An Automatic Face Identification System Using Flexible Appearance Models. In: *IVC94*. Manchester, UK, 1994.
98. LAWRENCE, S., GILES, C.L., TSOI, A.C. & et al. Face Recognition: A Hybrid Neural Network Approach. St. Lucia, Australia. 1996a.
99. LAWRENCE, S., TSOI, A. C. & ANDREW, D. Function Approximation with Neural Networks and Local Methods: Bias, Variance and Smoothness. In: *Australian Conference on Neural Networks*. St. Lucia, Australia, 1996b. p. 16-21.
100. LENGAGNE, R., TAREL, J.-P. & MONGA, O. From 2D Images to 3D Face Geometry. In: *Second International Conference on Automatic Face and Gesture Recognition*. Killington, USA, Oct/1996.
101. LI, S.Z., LU, J., CHAN, K.L. & et al. Hierarchical Linear Combinations for Face Recognition. Singapore. 1999.
102. LIN, W. & FU, K.-S. A syntatic approach to three-dimensional object recognition. 1986.
103. LIPPMANN, R.P. Pattern classification using neural networks. p. 47-64. Nov, 1989.
104. LIPPMANN, R.P. *An Introduction to Computing with Neural Nets*. 1999.
105. LOWE, D. & BROOMHEAD, D. S. Multi-variable functional interpolation and adaptive networks. In: *Complex Systems*. v.2, 1988. p. 321-355.
106. LOWE, D. What have neural networks to offer statistical pattern processing? San Diego, CA, USA. p. 460-471. 1991.
107. MACHADO, A. M. C. *Metodologias para Reconhecimento de Padrões em Visão Computacional*. Belo Horizonte, MG, Brasil, UNIVERSIDADE FEDERAL DE MINAS GERAIS, Dec/1994.
108. MANSFIELD, T. & ROETHENBAUGH, G. *1998 Glossary of Biometric Terms*. Association for Biometrics & International Computer Security Association, 1998.
109. MARKO, H. A biological approach to pattern recognition. Jan, 1974.

110. MARTINEZ, A. M. PCA versus LDA. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*. v.23, no. 2. Feb/2001.
111. MASS, H.-G. Robust Automatic Surface Reconstruction with Structured Light. Zurich, Swiss, 1999.
112. MICHELLI, C. A. Interpolation of scattered data *Distance matrices and conditionally positive definite functions* Constructive Approximation. v.Constructive Aproximation, no. 2. 1986. p. 11-22.
113. MILLER,A.J. Subset Selection in Regression. 1990.
114. MINUT, S., MAHADEVAN, S., HENDERSON, J. M. & et al. Face Recognition Using Fovela Vision. Michigan, USA, no. Department of Psychology of Michigan State University. 2002.
115. MITCHELL, T. M. *Machine Learning*. WCB/McGraw-Hill, 1997. ISBN: 0070428077.
116. MOGHADDAM,B., WAHID,W. & PENTLAND,A. Beyond Eigenfaces: Probabilistic Matching for Face Recognition. Cambridge, MA, USA. 1999.
117. MOGHADDAN, B. & PENTLAND, A. P. Face recognition using view-based and modular eigenspaces. In: *Automatic systems for the identification and inspection of humans*. v.2277, 1994.
118. MOODY,J. & DARKEN,C.J. Fast learning in networks of locally-tuned learning units. vol. 1, no. 2, p. 281-294. 1989.
119. MYERS,R.H. Classical and Modern Regression with Applications. Boston, MA, USA. 1986.
120. NAJMAN,L., VAILLAN,R. & PERNOT.E. Face from sideview to identification. 1993.
121. NAKAMURA,O., MATHUR,S. & MINAMI,T. Identification of human faces based on isodensity maps. vol. 3, no. 24, p. 263-272. 1991.
122. NELSON, M. M. & ILLINGWORTH, W. T. *Practical Guide to Neural Nets*. Woburn, MA, USA, Addison -Wesley Publishing Ciompany, Inc., 1991. ISBN: 0201523760.
123. OMOHUNDRO,S.M. Efficient algorithms with neural network behavior. vol. 1, p. 273-347. 1987.
124. ORR,M. Optimising the Widths of Radial Basis Functions. Rdinburg, Scotland, UK. 1999.
125. ORR, M. J. L. *Introduction to Radial Basis Functions Networks*. Edinburgh, Scotland, Center for Cognitive Science, University of Edinburgh, 1996.
126. PANDYA, A. S. & MACY, R. B. *Pattern Recognition with Neural Networks in C++*. 1ª ed. Boca Raton, FL, USA, CRC Press Inc., 1996. vol. 1,
127. PATHAK,A. & PAL,S.K. Fuzzy grammar in syntati recognition of skeletal maturity from x-rays. Sep, 1986.
128. PENEV,P.S. & ATICK,J.J. Local Feature Analysis: A General Statistical Theory For Object Representation. New York, NY, USA. 1999.
129. PENTLAND, A., MOGHADDAM, B. & STARNER, T. View-Based and Modular Eigenspaces for Face Recognition. In: *IEEE Conference on Computer Vision & Pattern Recognition*. v.Technical Report No. 245, 1994.
130. PERANTONIS,S.J. & LISBOA,P.J.G. Translation, Rotation, and Scale Invariant Pattern Recognition by High-Order Neural Networks and Moment Classifiers. p. 241-251. 1992.
131. PHILIPS,P.J., MARTIN,A., WILSON,C.L. & et al. An Introduction to Evaluating Biometrics Systems. USA. no. FEB, p. 56-63. Feb, 2000.
132. PHILLIPS, P. J., MOON, H., RAUSS, P. & et al. The FERET Evaluation Methodology for Face-Recognition Algorithms. New York, NY, USA, 2000.
133. PHOTONICS SPECTRA *The Photonics Dictionary*. Pittsfield, MA, Laurin Publication, 1995. vol. 4,
134. POGGIO,T. & GIROSI,F. Networks for Approximation and Learning. 1999.

135. POH, N. & KOREZAK, J. Hybrid Biometric Person Authentication Using Face and Voice Features. Halmstad, Sweden, no. Third International Conference, Audio and Video-Based Biometric Person Authentication AVBPA 2001. Jun/2001. p. 348-353.
136. RAITTZ, R. T. *Uma abordagem para reconhecimento de padrões*. Florianópolis, SC, Brasil, Engenharia de produção e sistemas, UFSC, Dec/1997. Dissertação de mestrado.
137. RAMSAY, C. S., SUTHERLAND, K., RENSHAW, D. & et al. A comparison vector quantization codebook generation algorithms applied to face recognition. In: *British Machine Vision Conference*. 1992. p. 508-517.
138. RICH, E. & KNIGHT, K. *Inteligência Artificial*. 2ª ed. São Paulo, MAKRON Books do Brasil Editora Ltda., 1994. vol. 1,
139. ROBERTSON, G. & SHARMAN, K. C. Object Locaton using proportions of the directions of intensity gradient-prodigy. In: *Proceedings Vision Interface 92*. Vancouver, CANadá, 1992. p. 189-195.
140. ROBERTSON, G. & CRAW, I. Testing face recognition systems. In: *British Machine Vision Conference*. v.1, 1993. p. 25-34.
141. ROBINSON, D. W. & REID, G. T. *Interferogram Analysis: Digital Fringe Patter Measurement Techniques*. Bristol, England, Institute of Physics Publisshing, 1993. ISBN: 075030179x.
142. RÖGNVALDSSON, T. S. *Radial Basis Functions Networks*. Halmstad, Swedish, Halmstad University, Feb/2001. vol. Learning & Adaptive Systems,
143. ROMDHANI, S. Face Recognition Using Principal Components Analysis. Glasgow, UK. no. Master Thesis, 1996.
144. ROSS, T. J. *Fuzzy Logic With Engineering Applications*. McGraw-Hill, Inc., 1995. ISBN: 0070539170 or 00711363771.
145. RUGGLES, T. G. *Comparison of Biometric Techniques*. CA, USA, The Biometric Consulting Group, Mar/98 A.D. Comparison of Biometrics.
146. SAMAL, A. & IYENGAR, P. A. Automatic Recognition and Analysis of Human Faces and Facial Expressions: Survey. *Pattern Recognition*. vol. 25, no. 1, p. 65-77. 1992.
147. SAMARIA, F. S. *Face Recognition Using Markov Models*. Tese de Doutorado. Cambridge, UK, University of Cambridge, 1994.
148. SANDIA NATIONAL LABORATORIES New Algorithm for Phase Unwrapping. 1999.
149. SCHALKOFF, R. *Pattern Recognition: Statistical, Structural and Neural Approaches*. USA, John Wiley & Sons, Inc., 1992. ISBN: 0471529745.
150. SHERSTINSKY, A. & PICARD, R. W. On the efficiency of the orthogonal least squares training method for radial basis function networks. no. Tecnical Report # 271, Oct, 1993.
151. SHRIDHAR, M. & BADRELDIN, A. High Accuracy Character Recognition Algorithm Using Fourier And Topological Descriptors. Windsor, Ontario, Canada. 1999.
152. SORGEL, W. & ROBERT, S. An Efficient and Accurate Structured Light Algorithm for Direct Measurement of Cylindrical Surface Parameters. Clemson, SC, USA, 1999.
153. SPIEGEL, M. R. *Manual de fórmulas e tabelas matemáticas*. Coleção Schaunª ed. McGraw-Hill do Brasil, 1974.
154. STONHAM, T. J. Practical face recognition and verification with WISARD. p. 426-441. 1986.
155. TAKEDA, M., INA, H. & KOBAYASHI, S. Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry. 1999.
156. TAKEDA, M. & MUTOH, K. Fourier transform profilometry for the automatic measurement of 3-D object shapes. 1999.
157. TARAYRE-ORIOT, H. & MASSONNET, D. New Methods of Phase Unwrapping in SAR Interferometry. Cedex, Chatillon, France. 1999.



158. THIAN, N.P.H. Biometric Authentication System. Penang, Malaysia. no. Master of Science Thesis, May, 2001.
159. TURK, M. & PENTLAND, A. Eigenfaces for recognition. In: *Journal of Cognitive Neuroscience*. v.1, no. 3. 1991. p. 71-86.
160. VALENTIN, D., ABDI, H. & O'TOOLE, A.J. Principal component and neural network analyses of face images: Explorations into the nature of information available for classifying faces by sex. Richardson, U.S.A. 1998.
161. WARD, J.R. & KUKLENSKI, T. A model for variability effects in hanprinting with implications for the design of handwriting character recognition systems. vol. 18, no. 3, May, 1989.
162. WASSERMAN, P. D. *Advanced Methods in Neural Computing*. New York, USA, Van Nostrand Reinhold, 1993.
163. WATANABE, S. *Pattern Recognition: Human and Mechanical*. New York, NJ, USA, Wiley and Sons, 1985.
164. WEI, C., CHEN, M., HOU, W. & et al. Phase measurement algorithm without phase-unwrapping problem for phase-stepping interferometry. p. Chunlong Wei', Mingyi Chen', Weidong Hou Zhijiang Wang'- 1999.
165. WETTSCHERECK, D. & DIETTERICH, T. Improving the performance of radial basis function networks by learning center locations. San Mateo, CA. p. 1133-1140. 1992.
166. WONG, K.H., LAW, H.H.M. & TSANG, P.W.M. A system for recognising human faces. p. 1638-1642. 1989.
167. WU, C.J. & HUANG, J.S. Human face profile recognition by computer. vol. 3/4, no. 23, p. 255-259. 1990.
168. XU, W. & CUMMING, I. Unwrapping the Difficult Sardegna Interferogram. Vancouver, BC, Canada. 1999.
169. YUILLE, A.L., HALLINAN, P.W. & COHEN, C.S. Feature Extraction from faces using deformable templates. vol. 8, no. 2, p. 99-111. 1992.
170. ZIMMERMANN, A. C., BARRETO, J. M., DIAS, J. S. & et al. Algoritmo de Redes Neurais como Filtro de Imagens Radiológicas. In: *IV Fórum Nacional de Ciência e Tecnologia em Saúde 1998*. Curitiba, PR, 1998.
171. ZIMMERMANN, A. C., GONÇALVES, A. A., Jr. & BARRETO, J. M. A 3D Object Extraction and Recognition Method. In: *IEEE - Sixth International Conference on Control, Automation, Robotics and Vision - ICARV 2000*. Singapore, Republic of Singapore, 2000a.
172. ZIMMERMANN, A. C., GONÇALVES, A. A., Jr. & BARRETO, J. M. General Non-Invasive Shape Reconstruction and Recognition Method Applied to 3D External Biologic Morphologies. In: *IEEE International Symposium on Bio-Informatics & Biomedical Engineering 2000 - BIBE 2000*. Washington, USA, 2000b.
173. ZIMMERMANN, A. C., BARRETO, J. M., ENCINAS, L. S. & et al. UMA PROPOSTA DE UTILIZAÇÃO DE RACIOCÍNIO BASEADO EM CASOS - RBC EM RECONHECIMENTO FACES HUMANAS TRI-DIMENSIONAIS. In: *II Congresso Brasileiro de Computação*. Itajaí, SC, Brazil, no. Inteligência Artificial. Aug/2002.
174. ZURADA, J. M. *Introduction to Artificial Neural Networks*. 8ª ed. St. Paul, MN, USA, West Publishing Company, 1992. ISBN: 0314933913.

# Índice Remissivo

- 3D, 22, 61
- Aamodt & Plaza, 93
- Abordagem Determinística, 47
- Abordagem Estatística, 45
- Abordagem Heurística, 46, 47
- Abordagem por Conjuntos Difusos, 47
- Abordagem Sintática, 46
- abstrativa, 50
- Aceitabilidade, 28
- Algoritmos de Remoção de Saltos de Fase, 73
- algoritmos genéticos, 20
- Algoritmos Genéticos, 176
- Allison, 54
- Alzheimer, 110
- Análise da Resolução das Formas, 160
- Análise de Discriminantes Lineares —, 143
- Análise de eletrocardiogramas, 48
- Análise pelos Componentes Principais, 52
- Ângulo de Visão Entre Duas Franjas, 79
- Aprendizado Auto-Organizável, 120
- Aprendizado Competitivo, 121
- Aprendizado das RNAs, 117
- Aprendizado de Quantificação Vetorial, 125
- Aprendizado Evolutivo, 121
- Aprendizado Hebbiano, 122
- Aprendizado Híbrido, 121
- Aprendizado por Reforço, 121
- Aprendizado Supervisionado, 120
- aproximação de funções, 107
- Army Research Laboratory, 58
- Arquitetura da RNA tipo FBR, 128
- Arquitetura das RNAs, 114
- árvores de decisão, 20
- Assinatura, 36, 38
- audição, 19
- autovetores, 146
- avaliação metrológica, 22
- Baron, 53, 56
- base de dados, 25, 29, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59
- Bayes, 46
- Benson, 19, 49
- Bernard Widrow, 109

- 
- Biometria, 24
- biometria fisiológica, 25
- Biometric Consortium, 42
- bio-psicológicos, 20
- Bishop, 139
- Broomhead & Lowe, 139
- Brunelli, 52, 54, 56
- Brunelli, 49
- B-spline cúbica, 51
- Cálculo das Distâncias, 98
- camada de entrada, 114
- camada de saída, 47, 114, 115, 116, 132
- camadas intermediárias, 47, 114, 129
- Cameron, 55
- Característica de Operação, 41
- características geométrica locais, 50
- Características Geométricas Locais, 50
- Carey, 19
- caricaturas, 19, 53
- cartões magnéticos, 26
- chaves eletrônicas, 26
- Chellappa, 49
- Chen, 52, 140
- Cichocki, 47
- classificação, 43, 44, 45, 46, 50, 51, 54, 115, 116, 123, 125
- Classificação, 29, 163
- Classificação com Dimensionalidade Reduzida, 166
- conexionista, 52
- confluência, 129
- Conjugate Gradient, 122
- Conjuntos de Dados, 117
- conjuntos difusos, 45, 47
- Cootes, 55
- Cottrell, 54, 56
- Craw, 51, 55, 56
- Custo, 28
- Decomposição em Valores Singulares, 52, 146
- Desempenho, 28
- Diamond, 19
- Dimensionalidade Reduzida, 166
- Distância de Hamming, 97
- distância Euclidiana, 53
- Distância Euclidiana, 21
- DNA, 36, 39
- Dobras da junta dos dedos, 39
- e Poggio, 49, 52, 54, 56
- edge, 51
- eigenface, 53
- Erro tipo 1, 32
- Erro tipo 2, 32
- Erros Aleatórios, 102
- Erros de Posicionamento, 101
- Erros Sistemáticos, 101
- Estágio de Saída, 116
- Estágio Processador, 116

- 
- Estágio Sensor, 116
- Euclidiana, 46, 96, 97, 98, 101, 126, 129
- Euler, 71, 72
- Evolucionária, 176
- Expressões faciais, 59
- Expressões Faciais, 85
- Face, 35, 37, 58, 150, 151
- faces humanas, 21, 55, 60, 150, 155
- Facilidade de Uso, 28
- Faixa de Medição, 154
- falsificações, 19
- FBI, 36
- FERET, 58, See Facial Recognition Technology
- Filtrando as Formas de Faces, 88
- Fleming, 54, 56
- fligh-time, 61
- Fontes de Erros, 155
- forenses, 59
- forma da face, 21, 70, 150, 151, 154, 155
- Formas de Faces da Base, 88
- franjas, 21, 63, 67, 68, 69, 70, 71, 72, 75, 78, 82, 83, 84, 150, 151, 152, 154, 155
- Franjas, 67, 70
- Frank Rosenblatt, 109
- freqüência espacial, 69, 71, 78, 83
- Freqüência Espacial, 78, 79
- Freqüência Modulada, 71
- Fukushima, 47
- Função de Base Radial, 121, 125, 126
- função gaussiana, 132
- função multiquadrática, 132
- funções de ativação, 113, 122
- Funções de Ativação, 113, 130
- Galton, 50
- Gaussiana, 131
- Generalização, 118
- Geometria da Mão, 35, 36
- gliais, 110
- glial, 110
- Goldstein, 52
- Gradient Back Propagation, 51
- Gradiente Descendente, 139
- Harmon, 51, 56
- Haykin, 108
- Hebb, 109, 122, 123
- Henderson e Minut, 106
- holística, 50
- Holografia, 64
- Holografia Interferométrica, 64
- Hopfield, 116
- Huang, 51, 52
- identificação, 25
- Identificação, 116
- iluminação estruturada, 21
- Impressão Digital, 36
- Impressões Digitais, 35
- Influência do Tamanho do Conjunto de Treinamento, 169

- 
- inspiração biológica, 20
- inteligência artificial, 20, 22, 49
- Interferometria de Speckle, 64
- International Biometric Industry Association, 42
- Invariância em RNAs, 141
- Invariância por Estrutura, 142
- Invariância por Seleção de Características, 142
- Invariância por Treinamento, 142
- Íris, 35, 37
- John Hopfield, 109
- Kanade, 51, 52, 56
- Karhunen-Loève, 53
- Kass, 52
- Kirby, 53
- Klir, 44
- k-médias, 136
- Kohonen, 54, 116, 125
- Kruger, 47
- Lanitis, 55, 56
- LASER, 61, 65, 155
- Laser Radar, 61
- LASER RADAR, 61, 155
- LIDAR, 61
- Limiar de Classificação, 29, 31, 98
- Lin & Fu, 46
- Lógica Nebulosa, 46
- MADALINE, 109
- Mahalanobis, 46
- manequim, 150, 151
- Manhattam, 46
- Mapa de Franjas, 70
- Mapas de Linhas de Mesma Densidade, 53
- Marcian Hoff, 109
- Marvin Minsky, 109
- MATLAB, 155
- matriz, 53, 152
- Matriz de Pesos Próprios, 104
- Matriz de Pesos Universais, 105
- máxima inclinação, 84
- Máxima inclinação da forma do objeto, 84
- Mcculloch, 110
- medição de formas, 60, 62, 64, 67, 68, 154
- memorização, 19
- Mensurabilidade, 28
- Método sintático difuso, 48
- Métodos de Correlação de Múltiplos Modelos, 53
- Métodos de Modelos Híbridos, 55
- Métodos de Reconhecimento, 55
- Métodos de Treinamento Não Supervisionados, 136
- Mínimos Quadrados Ortogonais, 139
- Minsky, 124
- Modelos de Vetores Quantizados, 54
- Moiré, 62, 63, 64, 154
- Multiquadrática, 131
- Multiquadrática Inversa, 131

- 
- Najman, 51, 56
- Nakamura, 53, 56
- Nathanial Rochester, 109
- neurologia, 49
- neurologistas, 20
- Normalização Espacial das Formas das Faces, 98
- número de identificação pessoal. See NIP
- Número de Neurônios, 132
- Odor corporal, 39
- olfato, 19
- one-to-many, 25
- Ortogonalização de Gran Schmidt, 141
- Otimização Híbrida, 157
- Otimização por Normalização Espacial, 157
- Otimização por Ponderação Facial, 157
- Otimizações, 157
- Pappert, 124
- Paradigmas de Aprendizado, 119
- Paradigmas de Redes Neurais, 123
- Pentland, 53, 56
- Perceptron, 109, 116, 124, 126
- Perceptron Multicamadas, 124
- Perenidade, 28
- perfilometria de Fourier, 149
- Perfilometria de Fourier, 70
- Perret, 49
- Perrett, 19
- pesos sinápticos, 47, 108, 117, 121, 126, 129
- Pitts, 110, 112, 113
- pivô, 75, 78, 82
- pivôs, 75
- Plano de Luz, 65
- Pomerleau, 47
- Ponderação Comportamental, 103
- Ponderação Estatística, 104
- Ponderação Facial, 103
- ponto de referência, 51, 65, 99
- População da Base, 89
- Posicionamento Espacial, 85
- Princípio da Triangulação, 65
- Privacidade, 28
- processamento de imagens, 49
- Processo de Recuperação, 94
- processo de Retenção, 95
- processo de Reutilização, 95
- processo de Revisão, 95
- Projeção de Franjas, 67, 69
- protocolo de avaliação, 41
- Protocolo de Ensaios, 41
- psicologia, 49
- psicólogos, 20
- pupila dos olhos, 155
- Quasi-Newton Algorithms, 122
- Quickprop, 122
- Raciocínio Baseado em Casos, 92
- Raio da FBR, 133

- 
- Ramsay, 54
- Rangefinders, 65
- reconhecimento, 19, 20, 21, 24, 25, 37, 40, 49, 50, 52, 53, 54, 55, 56, 57, 58, 59, 68, *See* reconhecimento de faces
- Reconhecimento de caracteres, 48
- reconhecimento de faces, 19, 39, 40, 41, 49, 56
- Reconhecimento de faces, 48
- Reconhecimento de Faces, 49
- Reconhecimento de faces humanas, 19
- Reconhecimento de objetos, 48
- Reconhecimento de padrões, 43, 48
- Reconhecimento de Padrões, 45
- abordagens do, 45
- Reconhecimento de voz, 48
- Rede Neural do Tipo Função de Base Radial, 128
- redes neurais, 20, 23, 52, 54
- Redes Neurais Probabilísticas, 132, 133
- Reduced Memory Levenberg-Marquardt, 122
- Regra de Correção de Erros, 122
- regra Delta, 122
- Regras de Aprendizado, 121
- Regressão de Redes Generalizadas, 132
- Relação entre a Entrada e a Saída da Rede, 129
- Relação Entre Fase e Altura, 82
- Relação Fase Altura, 153
- representações globais, 50, 52
- Representações Globais, 52
- Resilient Backpropagation, 122
- Resolução da Matriz de Formas, 88
- Retina, 35, 37
- Retropropagação, 122
- Ritmo de datilografia, 39
- Robertson, 51
- robôs, 62
- robustez, 19, 20
- Rosenblatt, 124
- Ross, 44
- Royal Institution, 50
- Rumelhart, 109
- salto de fase, 73, 74, 82, 150, 151
- Salto de Fase, 73
- saltos de fase, 73
- Schalkoff, 44, 45
- Segurança, 28
- sentidos humanos, 19
- Seymour Papert's, 109
- Similitude, 96
- Singularidade, 28
- Sirovich, 53
- Sistema Automático de Reconhecimento de Faces, 59
- sistema de decisão, 22
- Sistema Óptico Geométrico, 75
- Sistemas conexionistas, 107
- Sistemas de Identificação, 25
- Sistemas de Varredura, 66

- 
- Sistemas de Verificação, 25
- Sistemas ópticos, 60
- smart cards, 26, 27
- snake, 52
- Speckle, 64
- Stonhan, 54
- Tabela de Contingência, 35
- Tamanho da Base de Dados, 88
- Tanimoto, 46
- Taxa de Falsa Aceitação, 32
- Taxa de Falsa Rejeição, 32
- Taxa de Mesmo Erro, 32
- Técnicas biométricas. See biometria
- Técnicas Biométricas Clássicas, 35
- telecêntrico, 82, 83, See projetor telecêntrico
- Tempo de Processamento, 28
- Thin-plate spline, 131
- Tock, 51
- Topogrametria, 61, 62, 155
- Traços de perfil, 51
- Traços frontais, 51, 52
- Transformada de Fourier, 143
- Transformada Inversa de Fourier, 72, 74
- Translação da forma da face, 99
- Tratamento de impressões digitais, 48
- Turk, 53, 56
- Unbenhabem, 47
- Universalidade, 28
- Valores Singulares, 148
- Variable Learning Rate, 122
- Vascularização, 35, 38
- verificação, 25
- Verificação, 54, 116
- Verificação de Modelos Baseados em Redes Neurais, 54
- visão, 19, 49, 60, 62, 78
- Voz, 36, 38
- Walter Pitts, 109
- Ward, 46
- Warren McCulloch, 109
- WISARD, 54
- Wong, 52, 56
- Wu, 51, 56
- Yuile, 52

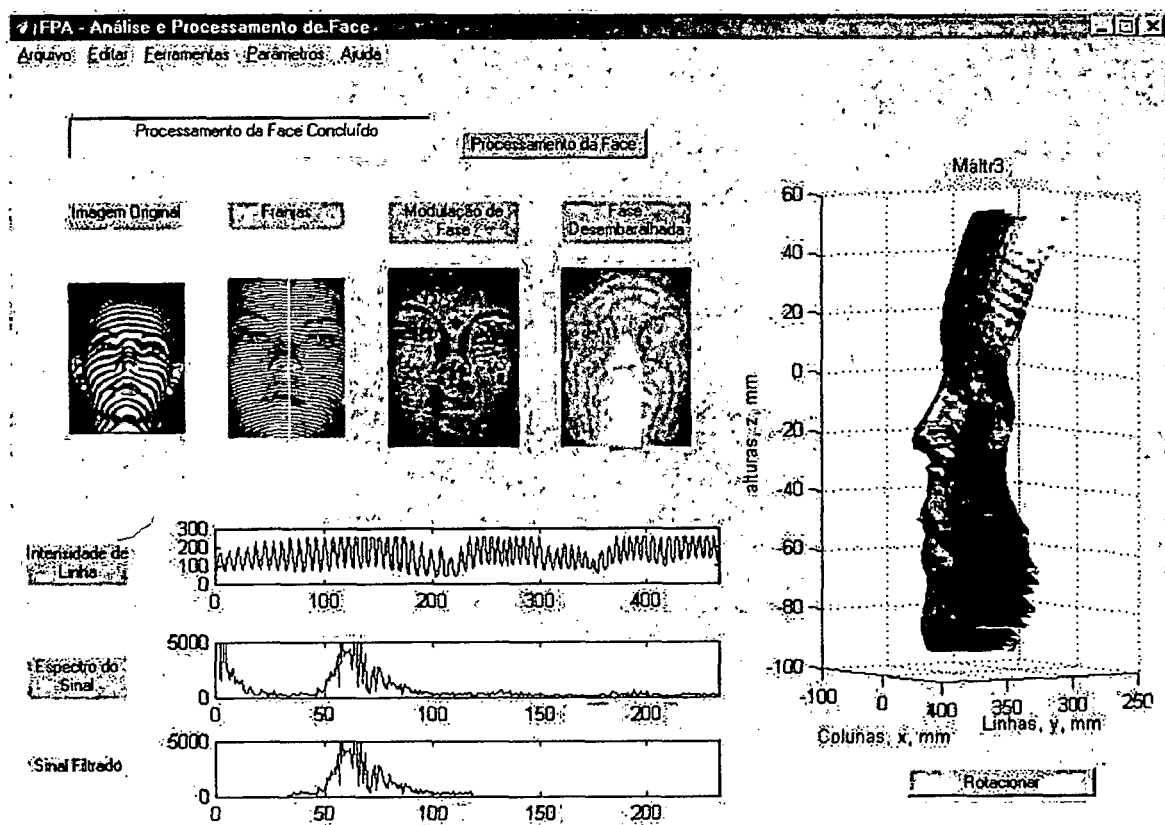


## **Apêndice A**

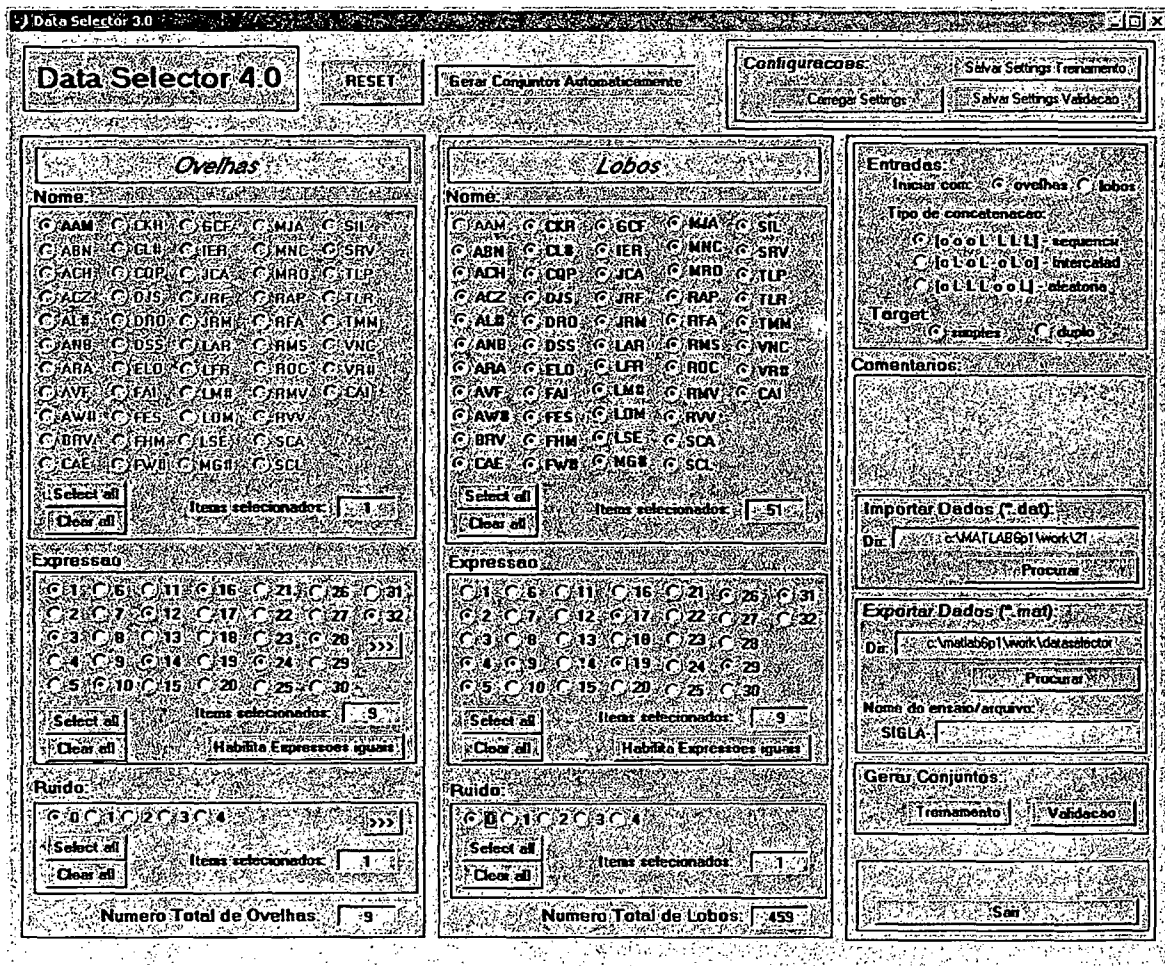
### **Telas dos Softwares Desenvolvidos**

---

a) Tela principal do software desenvolvido para a extração da forma da face:



b) Tela principal do software para seleção de indivíduos na composição dos conjuntos de treinamento e validação das redes neurais:

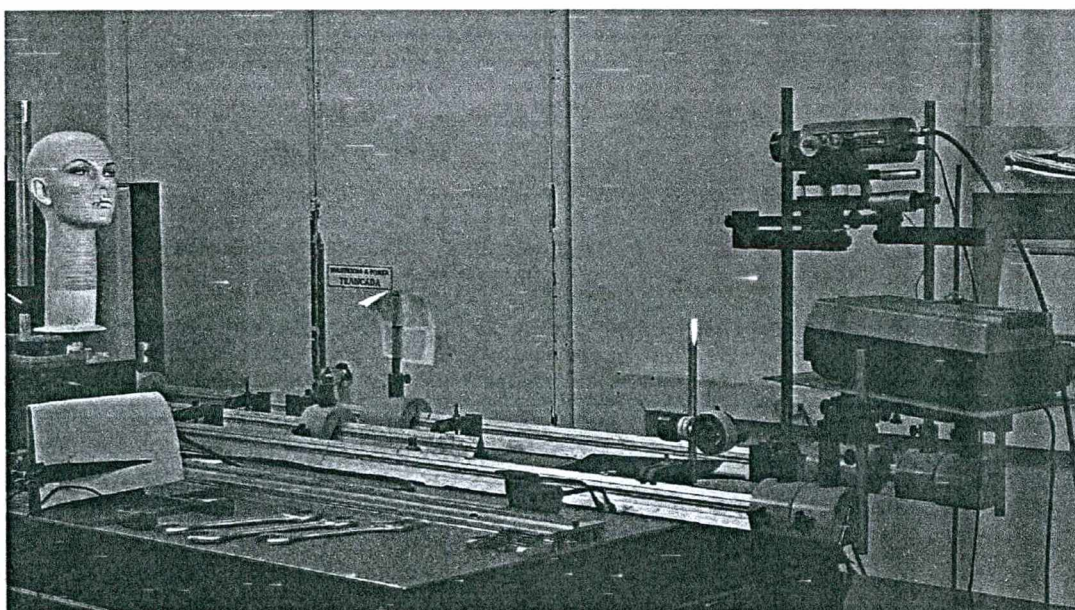


## **Apêndice B**

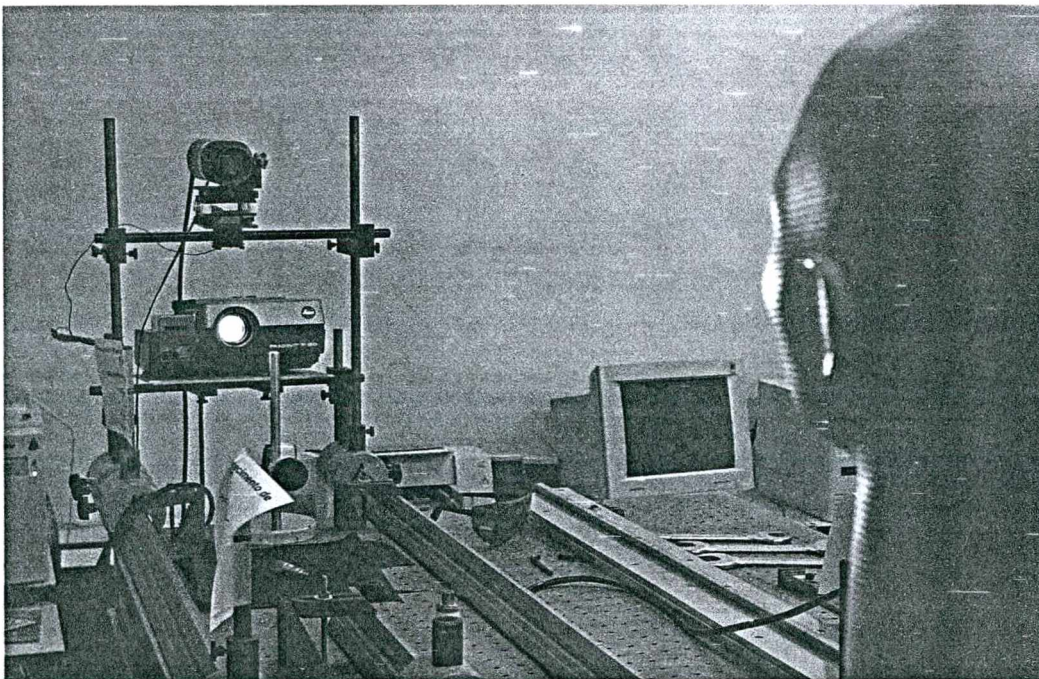
### **Imagens da Bancada de Ensaio**

---

a) Detalhes da bancada de ensaios:



Vista do projetor para o manequim



Vista do manequim para o projetor.

b) Detalhes da projeção das franjas sobre o manequim:



Manequim com projeção de franjas