UNIVERSIDADE DE SÃO PAULO

ESCOLA DE ENGENHARIA DE SÃO CARLOS

DEPARTAMENTO DE ENGENHARIA ELÉTRICA

RECONHECIMENTO DE FACES HUMANAS USANDO REDES NEURAIS MLP

THIAGO LOMBARDI GASPAR

Dissertação apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo, como parte dos requisitos para a obtenção do título de Mestre em Engenharia Elétrica.

Orientadora: Prof[a]. Dr[a]. Maria Stela Veludo de Paiva

São Carlos

2006

Aos meus pais *Florisvaldo* e *Eliana*,

à irmã *Thiele*,

e à namorada *Alini*,

com muito carinho e amor...

Agradecimentos

À Profa. Dra. Maria Stela Veludo de Paiva pelos ensinamentos e orientações, principalmente pela amizade e confiança para a realização deste trabalho.

Ao Prof. Dr. Ivan Nunes da Silva e a Profa. Dra. Agma Juci Machado Traina pelas orientações para a realização e finalização deste trabalho.

Ao todos os amigos do Laboratório de Visão Computacional (LAVI), especialmente ao Anderson e o Júnior, pelo aprendizado diário e pelos bons momentos de amizade.

Aos funcionários do Departamento de Engenharia Elétrica (EESC-USP) que sempre estiveram dispostos a colaborar e, em especial, às secretárias Denise e Marisa pela atenção e amizade.

Aos amigos da república, Anderson, Alexandre e Paulo, pelas conversas sempre animadas e principalmente pela amizade.

Em especial, a minha namorada Alini, que se propôs a ajudar-me e orientar-me para a realização deste trabalho. Minha sincera gratidão com muito carinho e amor.

Aos meus pais, que me ensinaram sempre a vencer todos os obstáculos da vida com muita paciência, humildade e sempre acreditando no meu objetivo profissional. Agradeço pelo apoio constante e pelas orações. E aos meus avós, que sempre acreditaram em minha perseverança.

E, principalmente a Deus, pela proteção e sabedoria dada para desenvolver toda esta pesquisa.

Resumo

GASPAR, T. L. (2006), *Reconhecimento de faces humanas usando redes neurais MLP*. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2006.

O objetivo deste trabalho foi desenvolver um algoritmo baseado em redes neurais para o reconhecimento facial. O algoritmo contém dois módulos principais, um módulo para a extração de características e um módulo para o reconhecimento facial, sendo aplicado sobre imagens digitais nas quais a face foi previamente detectada. O método utilizado para a extração de características baseia-se na aplicação de assinaturas horizontais e verticais para localizar os componentes faciais (olhos e nariz) e definir a posição desses componentes. Como entrada foram utilizadas imagens faciais de três bancos distintos: PICS, ESSEX e AT&T. Para esse módulo, a média de acerto foi de 86.6%, para os três bancos de dados. No módulo de reconhecimento foi utilizada a arquitetura *perceptron* multicamadas (MLP), e para o treinamento dessa rede foi utilizado o algoritmo de aprendizagem *backpropagation*. As características faciais extraídas foram aplicadas nas entradas dessa rede neural, que realizou o reconhecimento da face. A rede conseguiu reconhecer 97% das imagens que foram identificadas como pertencendo ao banco de dados utilizado. Apesar dos resultados satisfatórios obtidos, constatou-se que essa rede não consegue separar adequadamente características faciais com valores muito próximos, e portanto, não é a rede mais eficiente para o reconhecimento facial.

Palavras – chave: reconhecimento facial, extração de características, redes neurais, perceptron multicamadas, backpropagation.

Abstract

GASPAR, T. L. (2006), *Human face recognition using MLP neural networks*. M.Sc. Dissertation - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2006.

This research presents a facial recognition algorithm based in neural networks. The algorithm contains two main modules: one for feature extraction and another for face recognition. It was applied in digital images from three database, PICS, ESSEX and AT&T, where the face was previously detected. The method for feature extraction was based on previously knowledge of the facial components location (eyes and nose) and on the application of the horizontal and vertical signature for the identification of these components. The mean result obtained for this module was 86.6% for the three database. For the recognition module it was used the multilayer perceptron architecture (MLP), and for training this network it was used the backpropagation algorithm. The extracted facial features were applied to the input of the neural network, that identified the face as belonging or not to the database with 97% of hit ratio. Despite the good results obtained it was verified that the MLP could not distinguish facial features with very close values. Therefore the MLP is not the most efficient network for this task.

Keywords: face recognition, feature extraction, neural networks, multilayer perceptron, backpropagation.

Lista de Figuras

Figura 1: Exemplo de *eigenfaces*. 16
Figura 2: Grafos das faces adaptadas para diferentes poses. 17
Figura 3: Uma rede neural híbrida EPP/PPR. A regra de aprendizagem não-supervisionada é aplicada apenas na camada oculta. 22
Figura 4: Uma imagem normalizada para cada classe (esquerda) e a variação de apenas uma imagem facial normalizada (direita). 23
Figura 5: Extração de características usando a rede MLP (esquerda) e extração de características usando uma rede híbrida BCM / MLP (direita). 24
Figura 6: Diagrama do sistema de reconhecimento de face. 25
Figura 7: Resultado do detector de face. Baseado nisto, é aplicado o localizador dos olhos. 26
Figura 8: Algoritmo básico usado para detecção de face. 29
Figura 9: (a) Exemplo de imagens faciais, randomicamente rotacionadas, transladadas e pela menor soma escalar; (b) Exemplo de imagens não-faciais. 30
Figura 10: (a) Saída obtida pelo sistema de detecção com inúmeras pessoas na foto; (b) Exemplo de resultado obtido com o reconhecimento de um desenho e uma caricatura; (c) Detecção de uma não-face, ou seja, a face foi detectada corretamente, porém o sistema localizou um falso-positivo entre o pescoço e o tórax. 30
Figura 11: Exemplo do banco de dados *Yale Face*. São apresentadas 4 faces com diferentes expressões faciais (neutra, feliz, triste e surpreso). 31
Figura 12: Esquema estrutural da arquitetura de rede usada para executar o reconhecimento de expressão facial. 32
Figura 13: Estrutura modular do sistema *FADER*. 33
Figura 14: Exemplo de imagens adquiridas por uma webcam. 34
Figura 15: Exemplo do sistema de detecção facial: (a) imagem original e (b) imagem com as bordas detectadas e com as faces detectadas. 34
Figura 16: Classificação usando hiperplano entre duas classes: (a) Hiperplano traçado arbitrariamente entre as classes; (b) o OSH com a grande margem identificada pelas linhas tracejadas, passando pelos dois vetores de suporte. 36

Figura 17: Nesta figura os círculos representam os padrões face e os quadrados os padrões não-face. Na borda entre as duas classes se encontra um vetor de suporte do sistema. 37
Figura 18: Resultado do sistema de detecção de faces usando SVM. 38
Figura 19: Exemplo de uma árvore binária de uma imagem facial. 39
Figura 20: Exemplo do banco de dados faciais usadas no primeiro experimento. 39
Figura 21: Diagrama esquemático do Sistema Neuro-Fuzzy. 42
Figura 22: Projeção *Y* e *X* para uma simples imagem. 46
Figura 23: Projeção *Y* e *X* das características da boca. 47
Figura 24: Resultados do banco de dados AR. 48
Figura 25: Relação entre as características faciais. 48
Figura 26: Nove pontos característicos vitais na face. 49
Figura 27: Imagem modelo da face humana. 50
Figura 28: Projeção horizontal na região dos olhos. 50
Figura 29: Localizando a íris e as extremidades dos olhos. 50
Figura 30: Processamento da área do nariz. 51
Figura 31: Localizando as extremidades da boca. 51
Figura 32: Parâmetros para a representação do conhecimento de uma imagem facial... 53
Figura 33: Resultado da extração de característica facial: (a) e (b) resultado da localização dos componentes faciais usando MBR. 54
Figura 34: Diagrama em blocos do algoritmo automático de extração de características. 55
Figura 35: Três sub-regiões selecionadas da região facial para extrair as principais características no primeiro estágio de extração. 56
Figura 36: A sub-região R_c é excluída para extrair o ponto característico do nariz. 57
Figura 37: Sub-regiões especificas R'_a, R'_b e R'_c. 57
Figura 38: A imagem com o crescimento de região é usada para estimar as áreas características. (a) a versão binária da imagem e (b) o resultado correspondendo às áreas características. 58
Figura 39: Diagrama em blocos da evolução genética adotada. 59
Figura 40: Resultados da simulação a partir da seqüência de Clarie. 60
Figura 41: Sistema de Codificação Baseado em Modelo. 62
Figura 42: Um simples grafo com arcos direcionados. 62
Figura 43: Grafo deformável para extração de característica facial. 64

Figura 44: Resultados experimentais do banco de imagens “Clarie”, “Miss América” e quatro imagens do bando de imagens M2VTS. 64
Figura 45: Rede *feedforward*. Com uma única camada de neurônios. 68
Figura 46: Rede *feedforward*. Totalmente conectada com uma camada oculta e uma camada de saída. 68
Figura 47: Rede recorrente com retroalimentação entre neurônios. 69
Figura 48: Estrutura reticulada. 69
Figura 49: Modelo não-linear de um neurônio. 71
Figura 50: Rede neural de camada única. (a) Arquitetura da rede; (b) Direção da propagação da função sinal. 72
Figura 51: Função de *Heaviside*. 73
Figura 52: Curva de aprendizagem idealizada do algoritmo LMS. 75
Figura 53: Rede neural multicamadas. (a) Arquitetura da rede. (b) Direção da propagação da função sinal. 76
Figura 54: Gráfico da função sigmóide. 78
Figura 55: Diagrama em blocos da aprendizagem supervisionada. 83
Figura 56: Diagrama em blocos da aprendizagem não-supervisionada. 84
Figura 57: Módulos desenvolvidos no algoritmo para reconhecimento de faces. 87
Figura 58: Exemplo dos bancos de dados: (a) Banco de dados da Universidade de *Stirling* (PICS); (b) Banco de dados ESSEX (ESSEX); (c) Banco de dados da AT&T (AT&T). 88
Figura 59: Resultado da detecção de face usando o *Face Detector Software Web* (FDSW). 88
Figura 60: Diagrama em blocos do pré-processamento da imagem. 89
Figura 61: Diagrama em blocos da segmentação da imagem. 90
Figura 62: Máscara 3 x 3 do operador de Sobel. 91
Figura 63: Diagrama do processo de aplicação da assinatura horizontal e vertical. 92
Figura 64: Proporção do cálculo da Distância Euclidiana entre os componentes faciais. 93
Figura 65: Topologia da rede neural proposta para o reconhecimento de faces. 95
Figura 66: Resultado da aplicação do filtro da mediana e do alargamento de contraste: (a) imagem original; (b) resultado da filtragem mediana; (c) resultado do alargamento de contraste. 97

Figura 67: Resultado da segmentação facial: (a) imagem pré-processada; (b) binarização global simples e operação morfológica de abertura de área; (c) detector de bordas de Sobel com a máscara 3 x 3; (d) operação de preenchimento de *pixel*. 98
Figura 68: Espectro da contagem no número de *pixels* brancos da imagem: (a) Assinatura horizontal do lado direito facial; (b) Assinatura horizontal do lado esquerdo facial; (c) Assinatura vertical. 99
Figura 69: Resultado das assinaturas horizontais e verticais. 99
Figura 70: Localização dos pontos característicos faciais (olho direito, olho esquerdo, nariz e o centro da face na região dos olhos). 100
Figura 71: Problemas encontrados usando as assinaturas horizontais e verticais na região dos olhos: (a) e (b) Banco de dados PICS (PICS); (c) e (d) Banco de dados ESSEX (ESSEX); (e) Banco de dados AT&T (AT&T). 102
Figura 72: Problemas encontrados para localizar os olhos e o nariz usando assinatura horizontal: (a) Erro na localização dos olhos e do nariz; (b) e (c) Localização das sobrancelhas, ao invés dos olhos; (d) Localização correta do nariz e incorreta dos olhos. 103
Figura 73: Problema na localização dos olhos usando assinatura vertical..... 103
Figura 74: Problema na localização de componentes faciais em imagens inclinadas. . 104
Figura 75: Problema na localização de componentes faciais em imagens de baixa luminosidade. 104
Figura 76: Problema na localização dos olhos em imagens faciais usando óculos. 105
Figura 77: Problema na localização do nariz em imagens faciais. 105
Figura 78: Curva de convergência da rede neural MLP para os dados de treinamento. 107

Lista de Tabelas

Tabela 1: Avaliação Comercial dos Sistemas de Reconhecimento de Face (ZHAO et al., 2003). 8
Tabela 2: Taxa de reconhecimento (%) para o conjunto de faces do conjunto de dados. 21
Tabela 3: Resultado da classificação no conjunto de dados. 23
Tabela 4: Desempenho do sistema de reconhecimento de faces usando um banco de dados com grandes orientações faciais. 27
Tabela 5: Desempenho dos diferentes sistemas de reconhecimento de faces aplicado em 200 pessoas do banco de dados FERET..... 28
Tabela 6: Comparação da execução do SVM no sistema de detecção de faces. 37
Tabela 7: Taxa de reconhecimento por algoritmo EBP e o proposto algoritmo neuro-fuzzy..... 43
Tabela 8: Tabela da localização dos pontos característicos..... 52
Tabela 9: Média proporcional (%) do tempo de execução nos diferentes estágios de processamento, quando GA não é aplicado no Segundo Estágio de Extração de Característica..... 60
Tabela 10: Média proporcional (%) do tempo de execução nos diferentes estágios de processamento, quando GA é aplicado no Segundo Estágio de Extração de Características. 60
Tabela 11: Resultado da aplicação das assinaturas horizontais e verticais..... 101
Tabela 12: Resultado do reconhecimento de faces usando redes neurais MLP..... 107

Lista de Siglas

ADALINE	Adaptive Linear Neuron (em português: Neurônio Linear Adaptativo)
BCM	Bienenstock, Cooper e Munro
DBNN	Decision Based Neural Network (em português: Decisão Baseada em Rede Neural)
DFFS	Distance to Feature Space (em português: Distância para o Espaço Característico)
EBP	Error-backpropagation
FADER	Face Detection and Recognition
FDSW	Face Detector Software Web
FERET	Face Recognition Technology (em português: Tecnologia no Reconhecimento Facial)
FLD	Fisher's Linear Discriminant (em português: Discriminante Linear de Fishers)
GA	Genetic Algorithms (em português: Algoritmos Genéticos)
ICA	Independent Component Analysis (em português: Análise do Componente Independente)
LDA	Linear Discriminant Analysis (em português: Análise do Discriminante Linear)
LMS	Least Mean Squared (em português: Mínimo Quadrado Médio)
MATLAB	Matrix Laboratory
MBR	Minimun Bounding Retangle (em português: Retângulo de Fronteira Mínima)
MLP	Multilayer Perceptron (em português: Perceptron Multicamadas)
MSE	Mean Squared Error (em português: Erro Quadrado Médio)
MSG	Most Significant Genes (em português: Gene Mais Significante)
NCC	Nearest Center Classification (em português: Classificação do Centro Mais Próximo)
ORL	Oliver Research Laboratory (em português: Laboratório de Pesquisa Oliver)

OSH	Optimal Separating Hyperplane (em português: Hiperplano de Separação Ótima)
PCA	Principal Component Analysis (em português: Análise do Componente Principal)
PDBNN	Probabilistic Decision Based Neural Network (em português: Decisão Probabilística Baseada em Rede Neural)
PIN	Personal Identification Number (em português: Número de identificação pessoal)
RBFN	Radial Based Functions Network (em português: Rede de Funções de Base Radial)
RBFs	Radial Based Functions (em português: Funções de Base Radial)
RG	Registro Geral
RNAs	Redes Neurais Artificiais
SOM	Self Organize Map (em português: Mapa Auto-organizável)
SUSAN	Smallest Univalue Segment Assimilation Nucleus
SV	Support Vector (em português: Vetores de Suporte)
SVM	Support Vector Machines (em português: Máquinas de Vetor de Suporte)

Sumário

RESUMO ..... **i**

ABSTRACT ..... **ii**

LISTA DE FIGURAS ..... **iii**

LISTA DE TABELAS ..... **vii**

LISTA DE SIGLAS ..... **viii**

1 INTRODUÇÃO **1**

1.1 SISTEMA AUTOMÁTICO PARA O RECONHECIMENTO DE FACES 2
1.2 ESTRUTURA DO TRABALHO 5

2 RECONHECIMENTO DE FACES HUMANAS **6**

2.1 CONSIDERAÇÕES INICIAIS 6
2.2 INTRODUÇÃO 6
2.3 CARACTERÍSTICAS RELEVANTES PARA O SISTEMA DE RECONHECIMENTO DE FACES 7
2.4 MÉTODOS HOLÍSTICOS 11
2.4.1 Reconhecimento de faces aplicando PCA *13*
2.4.2 Análise Discriminante Linear (Linear Discriminant Analysis – LDA) *19*
2.5 REDES NEURAIS APLICADAS AO RECONHECIMENTO DE FACES 20
2.6 OUTRAS TÉCNICAS PARA O RECONHECIMENTO DE FACES 35
2.6.1 Máquinas de Vetores de Suporte (SVM) *35*
2.6.2 Sistema Neuro-Fuzzy *40*
2.7 CONSIDERAÇÕES FINAIS 44

3 TÉCNICAS DE EXTRAÇÃO DE CARACTERÍSTICAS PARA O RECONHECIMENTO DE FACES **45**

3.1 CONSIDERAÇÕES INICIAIS 45
3.2 INTRODUÇÃO 45
3.2.1 Método baseado em projeção horizontal e vertical *46*
3.2.2 Extração de pontos característicos da face *48*
3.2.3 Extração de características baseado em conhecimento *52*
3.2.4 Extração de características através de Crescimento de Região e Algoritmos Genéticos *54*
3.2.5 Extração de características usando auto-espaço e grafos deformáveis *61*
3.3 CONSIDERAÇÕES FINAIS 65

4 REDES NEURAIS ARTIFICIAIS 66
4.1 CONSIDERAÇÕES INICIAIS 66
4.2 INTRODUÇÃO 66
4.2.1 Modelo de um neurônio 71
4.3 PERCEPTRON DE CAMADA ÚNICA 72
4.3.1 Algoritmo LMS (regra delta) 74
4.4 PERCEPTRON MULTICAMADAS (MLP) 76
4.4.1 Algoritmo de treinamento backpropagation 78
4.5 ALGORITMO DE APRENDIZAGEM 82
4.5.1 Introdução 82
4.5.2 Aprendizagem supervisionada 83
4.5.3 Aprendizagem não-supervisionada 84
4.6 CONSIDERAÇÕES FINAIS 85

5 ALGORITMO PARA RECONHECIMENTO DE FACES HUMANAS: METODOLOGIA 86
5.1 CONSIDERAÇÕES INICIAIS 86
5.2 MÓDULOS DESENVOLVIDOS NO ALGORITMO 86
5.3 METODOLOGIA 87
5.3.1 Pré-processamento 88
5.3.2 Segmentação 89
5.3.3 Métodos de extração de características 91
5.3.3.1 Assinaturas horizontais e verticais para extração de características ... 91
5.3.3.2 Distância Euclidiana 92
5.3.4 Reconhecimento de faces usando redes neurais 93
5.3.5 Software de desenvolvimento 95

6 RESULTADOS E DISCUSSÕES 96
6.1 CONSIDERAÇÕES INICIAIS 96
6.2 RESULTADOS FINAIS 96
6.2.1 Pré-processamento da imagem 96
6.2.2 Segmentação 97
6.2.3 Assinaturas horizontais e verticais para localização dos componentes faciais 98
6.2.3.1 Problemas apresentados nas imagens faciais usando assinaturas horizontais e verticais 102
6.2.3.1.1 Assinatura horizontal e vertical na região dos olhos 102
6.2.3.1.2 Assinatura horizontal na região dos olhos e do nariz 102
6.2.3.1.3 Assinatura vertical na região dos olhos 103
6.2.3.1.4 Assinatura horizontal e vertical em imagens inclinadas 103
6.2.3.1.5 Assinatura horizontal e vertical em imagens faciais com baixa luminosidade 104
6.2.3.1.6 Assinatura horizontal na região dos olhos em imagens faciais com óculos 104
6.2.3.1.7 Assinatura horizontal para localização do nariz 105
6.2.4 Reconhecimento de faces usando rede neural MLP 105

7 CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS 108

7.1 CONCLUSÕES 108
7.2 CONTRIBUIÇÕES 109
7.3 PROPOSTAS PARA TRABALHOS FUTUROS 110

REFERÊNCIAS BIBLIOGRÁFICAS 111

A Apêndice – Imagens do banco de dados PICS 124

B Apêndice – Imagens do banco de dados ESSEX 125

C Apêndice – Imagens do banco de dados AT&T 127

D Apêndice – Resultado do pré-processamento das imagens dos bancos de dados PICS, ESSEX e AT&T 128

E Apêndice – Resultado da segmentação das imagens dos bancos de dados PICS, ESEX e AT&T 131

F Apêndice – Resultado das assinaturas horizontais e verticais dos bancos de dados PICS, ESSEX e AT&T 134

1 Introdução

Sistemas biométricos são sistemas automáticos que permitem a identificação de indivíduos através de suas características físicas e comportamentais. As características físicas incluem impressão digital, geometria da mão, identificação da retina, identificação da íris e reconhecimento de faces. As características comportamentais se referem ao reconhecimento da dinâmica da digitação, reconhecimento de voz e reconhecimento de assinatura (VIGLIAZZI, 2003). Essas características são únicas para cada indivíduo e não apresentam alterações significativas no decorrer da vida do mesmo.

Métodos seguros para reconhecimento automático de indivíduos ou confirmação de identidade, baseados em características biométricas, têm se tornado de extremo interesse, principalmente quando o fator segurança é importante, seja em ambientes físicos ou virtuais. A vantagem da identificação biométrica é a de não poder ser transferida de uma pessoa para outra, uma vez que as características biométricas são únicas para cada indivíduo.

Uma grande contribuição dos sistemas biométricos automatizados é a de evitar fraudes, pois com os sistemas utilizados atualmente não é possível afirmar que aquele indivíduo é realmente ele ou não, utilizando por exemplo, um simples documento de identidade.

Dentre os métodos baseados em características biométricas, tem se destacado aqueles que se baseiam em características faciais para a identificação de indivíduos. Sistemas baseados no reconhecimento facial vêm ganhando grande importância não só na identificação, como no controle de acesso de indivíduos em diversos ambientes, podendo-se destacar as seguintes aplicações (CHELLAPPA et al., 1995):

- identificação pessoal para banco, passaporte e fichas criminais;
- monitoramento de multidões em shopping centers, estações de metrô, lojas, aeroportos, bancos, empresas, etc.;
- elaboração de retrato falado e busca em fichas criminais;
- reconhecimento de expressões faciais e envelhecimento computadorizado para busca de desaparecidos, etc;
- transferência de informações nas redes de computadores.

No caso de sistemas de segurança, o uso do PIN (Número de Identificação Pessoal) que foi um dos primeiros identificadores a oferecer um reconhecimento automático, não necessariamente identifica a pessoa que o está utilizando. O mesmo caso acontece com cartões para acesso, que apresentam a desvantagem de serem apresentados por indivíduos não legalmente autorizados. Utilizando-se códigos numéricos PIN em associação com cartões de acesso é possível se obter um grau de segurança maior, mas mesmo assim este método é fácil de ser fraudado. A substituição desses sistemas por sistemas biométricos, traria maior segurança, bloqueando o acesso de pessoas não autorizadas.

1.1 Sistema automático para o reconhecimento de faces

O reconhecimento de faces humanas é usado por todos nós diariamente, quer seja para identificar pessoas ou para identificar qual o comportamento da pessoa (surpresa, triste, chorando, sorrindo) num determinado instante. Nos seres humanos, a capacidade de reconhecer começa muito cedo, pois as crianças com poucos meses de vida já são capazes de reconhecerem seus familiares mais próximos. Com o passar dos anos, essa capacidade é desenvolvida pelo cérebro humano e se aperfeiçoa, criando uma poderosa ferramenta de reconhecimento. O reconhecimento de um indivíduo pelo ser humano acontece normalmente, mesmo que a face apresente alterações significativas causadas pelo envelhecimento, ou pela presença de óculos, barba e corte de cabelo diferente.

Em um sistema automático de reconhecimento de faces, deve-se levar em consideração alguns aspectos importantes da visão humana a qual utiliza para o reconhecimento primeiramente formas e contornos e somente depois outros tipos de informação.

Os sistemas de reconhecimento de faces usam como entrada uma imagem digital da qual são extraídas características tais como, a distância entre os olhos, nariz e boca para a detecção e o reconhecimento de uma face.

As técnicas empregadas para classificar as informações extraídas das imagens faciais são as mais variadas e com o desenvolvimento das técnicas computacionais inteligentes, como Sistemas Especialistas, Lógica Nebulosa, Redes Neurais, Algoritmos Genéticos e PCA (Principal Component Analysis), novas abordagens tem surgido para o reconhecimento de faces com mais precisão na identificação e maior abrangência de aplicação (CHIANG et al., 2003).

As etapas envolvidas no reconhecimento de faces são as seguintes (PERELMUTER et al., 1995):

- **Detecção:** dada uma imagem digital qualquer, identificar a posição onde aparece a face;
- **Segmentação:** localizar as informações mais relevantes para o reconhecimento da face;
- **Representação:** selecionar as informações da face que serão utilizadas para representá-las;
- **Reconhecimento:** verificar se a face pertence ou não ao conjunto de faces.

Em um sistema automático para reconhecimento de faces como nos demais sistemas biométricos, deve-se inicialmente fazer a aquisição das características biométricas do indivíduo.

Um dos métodos usados para o reconhecimento facial, baseia-se na extração de algumas características da imagem digital, que são medidas do rosto que nunca se alteram (LEE et al., 1996). Essas características são as seguintes:

- distância entre os olhos;
- distância entre olhos, nariz e boca;
- distância entre olhos, queixo, boca e contorno dos cabelos.

Os seguintes problemas podem dificultar o processo de extração de características (YANG et al., 2002):

- posição da face;
- face parcialmente ou totalmente obstruída por outros objetos;
- expressão facial;
- complexidade no fundo da imagem.

A extração de características faciais desempenha um papel importante na interpretação visual e no reconhecimento de faces, podendo-se extrair características faciais do perfil ou frontais. Para se obter características mais confiáveis, é necessário que as imagens sejam segmentadas, de forma a eliminar os objetos não-faces presentes na imagem, os quais podem influenciar no resultado da extração.

O desempenho do algoritmo de reconhecimento de faces é avaliado por diversos fatores, tais como: precisão na resposta (baixa taxa de falsos negativos e falsos positivos), robustez contra fatores adversos, resposta em tempo real e baixo custo dos equipamentos. Os fatores adversos incluem condições de iluminação, ruídos na imagem, variação na expressão, presença de óculos, mudanças no cabelo e postura. O reconhecimento de um indivíduo pelo ser humano ocorre mesmo com a presença destes fatores. Como os sistemas atuais de identificação baseiam-se na comparação visual da face do indivíduo com uma face do documento, a presença destes fatores dificilmente afetam a identificação (BARRETT, 1997).

O presente trabalho faz uma revisão da literatura sobre os algoritmos para a extração de características e reconhecimento facial e propõe um algoritmo para o reconhecimento facial utilizando redes neurais do tipo MLP com o algoritmo de aprendizagem *backpropagation*.

1.2 Estrutura do trabalho

O presente trabalho é dividido em cinco capítulos.

No primeiro capítulo foi apresentada a introdução destacando sua aplicabilidade e importância, e apresentando o objetivo dessa pesquisa.

No capítulo 2 é apresentada uma breve introdução sobre o reconhecimento de faces humanas, bem como técnicas encontradas na literatura para realizar o reconhecimento.

No capítulo 3 são apresentadas algumas técnicas de extração de características.

No capítulo 4 é apresentada a rede MLP e o processo de aprendizagem *backpropagation*.

No capítulo 5 é descrita a metodologia utilizada para implementação do algoritmo para reconhecimento de faces.

No capítulo 6 são descritos os resultados obtidos e apresentadas as discussões.

No capítulo 7 são apresentas as conclusões e sugestões para trabalhos futuros.

2 Reconhecimento de Faces Humanas

2.1 Considerações iniciais

Nesse capítulo são descritos os principais trabalhos pesquisados na literatura, para o reconhecimento automático de faces humanas, utilizando técnicas de redes neurais artificiais, métodos holísticos (PCA), neuro-fuzzy e máquinas de vetores de suporte.

2.2 Introdução

O reconhecimento de faces humanas é um problema bastante complexo para ser implementado de forma automática, devido a diversos fatores tais como: as diferentes variações de orientação e tamanho da imagem, condições de iluminação do ambiente, diferenças na aparência, na expressão facial, e na cor da pele, dentre outros fatores que influenciam na extração de característica (SUNG e POGGIO, 1998).

Segundo Yang et. al. (2002), os seguintes fatores podem influenciar na detecção de faces:

- **Posição:** a face pode ser apresentada em diferentes maneiras dependendo da posição da câmera (frontal, 45 graus, perfil, de cabeça para baixo), e dessa forma as características faciais podem estar ocultas;

- **Presença ou ausência de componentes faciais:** alguns componentes faciais como a barba, o bigode e os óculos podem ou não estar presentes, influenciando na extração de características;
- **Expressão facial:** a aparência da face é afetada por uma expressão facial pessoal (triste, alegre, fazendo careta, boca aberta ou fechada, olhos abertos ou fechados, etc.);
- **Oclusão:** as faces podem estar parcialmente ou totalmente ocultas por outro objeto. Em algumas imagens, um grupo de pessoas pode ocultar parcialmente outras faces;
- **Condições da imagem:** Quando a imagem é formada, alguns fatores como as condições de iluminação (distância da fonte de luz, distribuição e intensidade luminosa) e as características da câmera (resposta do sensor e às lentes) podem afetar a aparência da face.

O algoritmo de detecção de face deve ser implementado com muito cuidado, pois poderá influenciar na tomada de decisão nas demais etapas do processo de reconhecimento, possibilitando ou não um reconhecimento mais preciso.

Dentre as diversas características faciais disponíveis, os olhos representam um papel importante no processo de detecção de faces, ou seja, a partir da localização dos olhos é possível determinar a localização dos demais componentes: nariz e boca (KAWAGUCHI et al., 2000; TOLBA et al., 2000; LAM e YAN, 1998).

2.3 Características relevantes para o sistema de reconhecimento de faces

A compreensão dos princípios e dos processos envolvidos no reconhecimento de faces envolve pesquisas no domínio da neurociência, psicologia e ciência da computação. O objetivo principal da neurociência e da psicologia é compreender como o sistema visual humano converte imagens da retina, de faces humanas, em representações mentais que possam ser usadas em sistemas de reconhecimento de faces. O principal objetivo da pesquisa computacional tem sido desenvolver algoritmos

capazes de reconhecer faces humanas automaticamente (O'TOOLE et al., 2000; WENG e SWETS, 1999).

Apesar dos avanços na área de reconhecimento de faces e de já existirem sistemas comerciais (Tabela 1) para esse fim, esses sistemas ainda estão distantes da potencialidade do sistema de visão humana (ZHAO et al., 2003). Por exemplo, imagens faciais adquiridas em um ambiente ao ar livre com mudanças de iluminação ou na posição da face é um problema que ainda está sendo pesquisado.

Há evidências de que o sistema visual humano seja formado por módulos de processamento dedicados ao reconhecimento facial, e não meramente uma aplicação do processo de reconhecimento de objetos gerais, onde esse processamento envolve informações holísticas (características globais) e traços do rosto (características locais) (ELLIS, 1986). Isso encoraja as aplicações em reconhecimento de faces a trabalharem apenas com o estímulo visual caracterizado por imagens 2-D (em escala de cinza) únicas ou em múltiplas vistas de uma face.

Tabela 1: Avaliação Comercial dos Sistemas de Reconhecimento de Face (ZHAO et al., 2003).

Produtos comerciais	Websites
FaceIt from Visionics	http://www.FaceIt.com
Viisage Technology	http://www.viisage.com
FaceVACS from Plettac	http://www.plettac-eletronics.com
FaceKey Corp.	http://www.facekey.com
Cognitec Systems	http://www.cognitec-systems.de
Keyware Technologies	http://www.keywareusa.com
Passfaces from ID-arts	http://www.id-arts.com
ImageWare Software	http://www.iwsinc.com
Eyematic Interfaces Inc.	http://www.eyematic.com
BioID sensor fusion	http://www.bioid.com
Visionshere Technologies	http://www.visionsheretech.com/menu.htm
Biometric Systems, Inc.	http://www.biometrica.com
FaceSnap Recoder	http://www.facesnap.de/htdocs/english/index2.html
SpotIt for face composite	http://spotit.itc.it/SpotIt.html

Segundo Zhao et al. (2003), algumas características potencialmente relevantes para o desenvolvimento de um sistema de reconhecimento de faces são apresentas a seguir:

- **O reconhecimento de faces é um processo dedicado?** (BIEDERMAN e KALOCSALK, 1998; ELLIS, 1986; GAUTHIER e LOGOTHETIS, 2000): Tradicionalmente acredita-se que o reconhecimento de faces seja um processo dedicado diferente de outra tarefa de reconhecimento de objetos. As faces são mais fáceis de serem lembradas pelos seres humanos do que outros objetos quando apresentadas em uma determinada orientação.

- **O reconhecimento de faces é o resultado de uma análise holística ou análise de características?** (BRUCE et al., 1998): Ambas informações, holísticas e de características, são cruciais para a percepção e o reconhecimento de faces. Os estudos sugerem a possibilidade de descrições globais servindo como uma aproximação ou redução de possibilidades, seguida de uma descrição baseada em característica de uma percepção mais fina. Mas se as características forem dominantes, então as descrições holísticas não serão usadas. Por exemplo, em um estudo para recordação de faces, os seres humanos facilmente focalizam em características ímpares tais como: orelhas grandes, nariz curvado e olhos com características diferentes do normal.

- **Classificação das características faciais mais significantes.** (BRUCE, 1988b), (SHEPHERD et al., 1981). O cabelo, o contorno da face, os olhos e a boca (não necessariamente nesta ordem) foram determinados como sendo importantes para perceber e reconhecer as faces. Diversos estudos mostraram que o nariz apresenta um papel insignificante; isto pode ser devido ao fato que quase todos estes estudos foram feitos usando imagens frontais. No reconhecimento de faces usando os perfis, um nariz não convencional pode ser mais importante do que os olhos ou a boca. Também foi verificado que a parte superior da face é mais útil para o reconhecimento de faces do que a parte mais inferior. Os atributos relacionados à estética, tais como a beleza, atração e o charme também foram

estudados, com a conclusão de que quanto mais atraentes forem as faces, melhores serão suas taxas de reconhecimento; e as menos atrativas vêm em seguida, deixando as faces habituais por último, em termos de facilidade para serem reconhecidas.

- **Caricaturas** (BRUCE, 1988a; PERKINS, 1975): Uma caricatura pode formalmente ser definida como "um símbolo que exagera algumas medidas relativas a um padrão (características que chamam a atenção e destacam as pessoas das demais)." Assim o comprimento de um nariz é uma medida que varia de pessoa para pessoa, e poderia ser útil como um símbolo para caricaturar alguém. As caricaturas não contêm tanta informação quanto fotografias, mas podem representar as características mais importantes de uma face.

- **Peculiaridade** (BRUCE et al., 1994): Os estudos mostram que as faces que apresentam alguma particularidade são melhores retidas pela memória e tem um melhor e mais rápido reconhecimento do que as faces convencionais. Entretanto, se uma decisão tiver que ser feita a respeito se um objeto é uma face ou não, é notado um maior tempo para reconhecer uma face do que uma face convencional. Isso pode ser explicado pelos diferentes mecanismos que estão sendo usados para a detecção e para a identificação.

- **O reconhecimento é invariante ao ponto de vista?** (BIEDERMAN, 1987; TARR e BURTHOFF, 1993): Muitos trabalhos no reconhecimento visual de objetos foram realizados com base na estrutura teórica introduzida por (MARR, 1982) em que as diferentes vistas dos objetos são analisadas para se chegar às descrições de pontos de vista invariantes. Alguns experimentos sugerem que a memória é altamente dependente do ponto de vista (HILL et al., 1997).

- **Expressões faciais.** (BRUCE, 1988a): Estudos neuropsicológicos sugerem que a análise de expressões faciais é realizada em paralelo com o reconhecimento facial.

As pesquisas em reconhecimento de faces através de imagens 2-D em escala de cinza apresentam-se geralmente em três categorias: métodos baseados em características (locais), métodos holísticos (globais), métodos híbridos (soma dos dois métodos anteriores). Os métodos baseados em características apoiam-se na identificação de determinados pontos fiduciais na face tais como os olhos, o nariz, a boca, etc. A posição destes pontos pode ser determinada e usada para computar as relações geométricas entre os pontos e pela análise local da região circunvizinha. Assim, o processamento separado para os olhos, nariz e outros pontos fiduciais são executados e combinados em um classificador para então produzir o reconhecimento. Desde que a detecção de pontos das características precede a análise, tal sistema é robusto a variações de translação da face na imagem (ZHAO et al., 2003).

Os métodos holísticos tratam os dados da imagem simultaneamente sem tentar localizar pontos individuais. A face é reconhecida como uma entidade sem regiões diferentes explicitamente isoladas nela. As técnicas holísticas utilizam a análise estatística, redes neurais e transformadas de Gabor. Também requerem geralmente amostras grandes de dados de treinamento. A vantagem dos métodos holísticos é que eles utilizam a face como um todo e não destroem nenhuma informação, processando exclusivamente determinados pontos fiduciais. Assim, geralmente fornecem resultados mais exatos de reconhecimento. Entretanto, tais técnicas são sensíveis às variações em posição e escala o que restringe seu uso às imagens com ambientes controlados.

Uma fusão dos métodos baseados em características e holísticos é chamada de método híbrido, onde as idéias consolidadas nestes sistemas estão embasadas no sistema de percepção humana, e que usa características locais e a região inteira da face para efetuar o reconhecimento. Portanto um sistema de reconhecimento artificial deveria usar ambos.

2.4 Métodos Holísticos

Os métodos holísticos aplicados ao reconhecimento de faces utilizam todos os *pixels* da imagem ou de regiões características da face para executar o reconhecimento. Neste tipo de abordagem, a dimensionalidade dos dados é igual ao número de *pixels* da imagem considerada. Devido ao grande volume de dados, existem métodos estatísticos

que reduzem a dimensionalidade, tais como o PCA, discriminantes lineares (BELHUMEUR et al., 1997) e redes neurais (LAWRENCE et al., 1996). O método de reconhecimento utilizando PCA tem sido aplicado durante o pré-processamento, para se fazer a normalização de imagens, com o objetivo de melhorar o desempenho do sistema de reconhecimento.

Em Brunelli e Poggio (1993), é realizado um estudo comparando o desempenho de métodos por atributos (locais) com um método holístico. O método holístico testado foi o de *template matching*, testando o desempenho de imagens de olhos, nariz, boca e também de toda a face. O método holístico proporcionou melhores resultados que os locais. A vantagem de abordagens globais está no fato de que pequenas variações locais não prejudicam muito o reconhecimento. A principal desvantagem está nos problemas de variação de iluminação e em alguns casos, no custo computacional. Uma melhor avaliação sobre os efeitos da iluminação, orientação e escala podem ser encontrados em Luo e Yuwapat (2003).

Em PCA os dados da imagem são representados como vetores, isto é, como pontos em um espaço de alta dimensionalidade. Por exemplo, uma imagem 2-D em escala de cinza descrita pela matriz P(m,n) pode ser representada por um vetor $X \in R^{m.n}$, com seus pixels concatenados em linhas e colunas para formar um só vetor no espaço $R^{m.n}$. Apesar do espaço de alta dimensionalidade representar as imagens faciais reais, ele não é adequado para efetuar o reconhecimento. Pode-se então caracterizar as imagens num espaço de menor dimensão, onde apenas as informações mais relevantes para o reconhecimento estarão presentes. O objetivo desta análise é identificar, representar e parametrizar uma transformação linear para viabilizar o reconhecimento facial.

Um dos fatores mais importantes encontrados nos métodos estatísticos é a compressão das informações contidas nas imagens (redução de dimensionalidade) para posteriormente se fazer a classificação. Os métodos mais relevantes na área são a Análise de Componentes Principais (PCA) (TURK e PENTLAND, 1991), Análise de Componente Independente (ICA) (BARTLETT et al., 1998), e Análise Discriminante Linear (LDA) (BELHUMEUR et al., 1997). Estes métodos são baseados em análises estatísticas dos vetores de face, e em transformações lineares e não lineares dos sistemas de coordenadas onde se encontram estes vetores. Cada método tem sua própria forma de atuação e representação para formar um espaço dimensional menor baseado em

conceitos estatísticos diferentes. Projetando o vetor de face nos vetores da base do sistema, têm-se os coeficientes da projeção definidos como a nova representação da característica de cada imagem de face. O reconhecimento da face teste é efetuado através de um classificador, que deve analisar os coeficientes dos protótipos das faces de treinamento comparando-os com os coeficientes da face teste, e de acordo com um *threshold* previamente estabelecido, definir se a face teste está no banco de treinamento ou não. Uma das formas mais simples de efetuar a classificação é utilizar as medidas de distâncias entre vetores, como por exemplo, distância Euclidiana, do cosseno, Mahalanobis, etc.

Nos próximos itens será apresentada uma breve descrição de cada um desses métodos.

2.4.1 Reconhecimento de faces aplicando PCA

A análise de componentes principais (PCA) foi introduzida primeiro por Pearson (1901), que a usou em um contexto biológico para dispor a análise de regressão linear em uma nova forma. Ela foi então desenvolvida por Hotteling (1933) em um trabalho sobre psicometria, e surgiu novamente na formulação da teoria das probabilidades, como considerada por Kahunen (1947), sendo generalizada posteriormente por Loève (1963).

O reconhecimento das *eigenfaces* foi proposto por Sirovich e Kirby (1987), com uma aplicação de análise de componentes principais (PCA) em uma matriz *n*-dimensional, com o intuito de investigar os problemas de caracterização, identificação e distinção das faces humanas, com uma abordagem estatística para o reconhecimento de padrões referentes às classes de faces. Para a experiência foi usado um banco de imagens contendo 115 faces frontais de indivíduos do sexo masculino, digitalizadas e pré-processadas de uma maneira controlada e encontrou aproximadamente 40 *eigenfaces* que foram suficientes para uma descrição muito boa do seu conjunto de imagens faciais.

O método de reconhecimento de faces utilizando a transformada de Karhunen Loève foi proposto em Kirby e Sirovich (1990) e está descrito com maiores detalhes nos próximos parágrafos.

Kirby e Sirovich (1990), estenderam sua investigação anterior (SIROVICH e KIRBY, 1987), analisando o comportamento da expansão de Karhunen-Loève usando faces frontais completas e faces refletidas, explorando as simetrias naturais presentes na face humana. Motivado pelos trabalhos de Sirovich e Kirby (1987) (1990), Turk e Pentland (1991) aplicaram a expansão de Karhunen-Loève para fazer o reconhecimento de faces. Turk e Pentland buscaram executar o reconhecimento de faces através de uma análise global e holística, em contraposição aos outros métodos que estavam sendo implementados, visando aproveitar as características faciais individuais e locais. Este trabalho teve grande contribuição para a área, pois descreveu uma metodologia visando definir um mapa de faces para viabilizar a detecção, e um espaço de faces para efetuar o reconhecimento, utilizando um subespaço *eigenfaces*. Para avaliar o método, foi criado um banco de imagens de faces estáticas e também foi desenvolvido um sistema para localizar e reconhecer as faces em um ambiente dinâmico.

Para aumentar o desempenho do método *eigenfaces* (autofaces) proposto por Turk e Pentland (1991) em banco de dados grandes, os autores generalizaram a aplicação para um espaço de faces baseado em vistas e em módulos:

- **Vistas:** permite o reconhecimento baseado em várias orientações da cabeça e a descrição modular permite a incorporação das características faciais como olhos, nariz e boca de uma forma isolada para a análise. A arquitetura baseada em vistas define um processo paralelo de observação, portanto cada subespaço de autofaces possui os seus autovetores, que explanam uma orientação e escala de acordo com o treinamento. No teste, para a imagem de entrada é calculada a distância desta para os subespaços de autofaces e o subespaço mais próximo é usado para efetuar o reconhecimento.

- **Módulos:** é baseado na formulação de vários subespaços de autofaces definidos pelas características individuais da face (olhos, nariz e boca) ou sub-espaços *eigeneyes* (autoolhos), *eigennose* (autonariz) e *eigenmouth* (autoboca), onde as faces são recortadas formando regiões adequadas ao reconhecimento. Estes subespaços foram implementados para reconhecer indivíduos holisticamente, e

por suas características faciais locais (sistema conhecido como representação modular combinada).

Os passos a seguir, resumem o processo de reconhecimento na visão de Turk e Pentland (1991):

1. Adquirir um conjunto de treinamento de imagens de faces e calcular as *eigenfaces* (autofaces), para definir o espaço da face;
2. Determine se a imagem é uma face, observando se a sua projeção no subespaço de faces aproxima-se dos pontos que representam as faces de treinamento;
3. Se a imagem for uma face, classificar os seus padrões de pesos como uma face conhecida ou desconhecida de acordo com as faces de treinamento;
4. Atualizar as autofaces e os pesos, recalculando o subespaço de faces para um conjunto de treinamento com as novas faces inseridas;
5. Se a mesma face desconhecida for apresentada várias vezes ao sistema, calcular os seus padrões de pesos e inseri-los como faces conhecidas.

O método de análise de componentes principais é também conhecido como a expansão de Karhunen-Loève e tem como principal objetivo reduzir a dimensionalidade de um conjunto ou espaço de dados e criar um novo subespaço dimensional para o conjunto de dados. Quando a aplicabilidade de PCA é referenciada a faces, recebe o nome de *eigenfaces* (autofaces), que é derivado do prefixo alemão *eigen* (auto), que significa próprio ou individual.

Analisando um conjunto de dados multidimensional, como no caso das faces, um dos problemas encontrados é a representação dos dados e a sua estruturação em um subespaço de menor dimensão. A maneira de resolver este problema é representar os dados de modo que as distâncias entre os pontos das amostras no subespaço de dimensão menor correspondam o quanto possível às heterogeneidades entre os pontos do espaço dimensional original. O PCA é um dos métodos mais famosos para resolver este tipo de problema, pois tem como objetivo encontrar um subespaço contendo como

bases ortonormais, os vetores que definem as principais direções de distribuição dos dados no espaço original, otimizando o critério de correlação cruzada entre as amostras dos dados (GROB, 1994).

A aplicabilidade de PCA em imagens faciais tem o interesse de extrair informações relevantes contidas nessas imagens faciais, codificando a imagem de forma eficiente e comparando uma imagem codificada com um banco de dados similar. Um exemplo é a obtenção da variação dos níveis de cinza em uma imagem facial, independente dos traços nela contidos. Estatisticamente pode-se aplicar a técnica de PCA, encontrando os autovetores que constituem as bases ortonormais do subespaço de faces e seus respectivos autovalores, obtidos através da matriz de covariância formada pela correlação entre os *pixels*. Cada imagem de face no conjunto de treinamento pode ser exatamente representada em termos de uma combinação linear dos autovetores e os coeficientes desta combinação serão os novos representantes da face no subespaço dimensional.

Em termos matemáticos, o interesse é encontrar o componente principal da distribuição das faces ou os autovetores da matriz de covariância do conjunto de imagens faciais. Os autovetores podem ser pensados como um conjunto de características e juntos representam as variações entre as imagens faciais e a imagem média do conjunto. Cada imagem local contribui mais ou menos para cada autovetor, de forma a se poder exibir o autovetor com um tipo de face disforme, como ilustrado na Figura 1, denominado de autoface.

Figura 1: Exemplo de *eigenfaces*.

Turk e Pentland (1991) refinaram o método, acrescentando pré-processamento e expandindo a base de dados estatística. Eles concluíram que um número menor de

eigenfaces de uma população diversa de imagens frontais é suficiente para descrever uma face com ótima precisão.

A construção de um conjunto de *eigenfaces* é totalmente computacional. Um conjunto poderia em teoria ser desenvolvido uma só vez e por todo tempo descrevendo adequadamente todos os indivíduos do sexo masculino ou feminino (BARRETT, 1997).

A partir de 1993, surgiram vários outros sistemas robustos de reconhecimento de imagens faciais. Segundo Pentland (2000), os dois melhores métodos são:

- **Aprendizado Visual Probabilístico** (MOGHADDAM e PENTLAND, 1997; MOGHADDAM et al., 1998): utiliza a transformada de *eigenface* seguida por um discriminante linear;

- ***Elastic Graph Matching*** (WISKOTT et al., 1997): a face é representada por um grafo retangular, onde cada nó é rotulado por um conjunto de coeficientes complexos da *Wavelet* de Gabor. Apenas as magnitudes dos coeficientes são usadas para o reconhecimento. Quando identifica uma face em uma nova imagem, cada grafo do banco de dados será semelhante para cada imagem separadamente e o mais parecido indicará a pessoa reconhecida. A rotação será compensada pela deformação elástica dos grafos (Figura 2).

Figura 2: Grafos das faces adaptadas para diferentes poses.

Os nós são posicionados automaticamente, através de *elastic graph matching*. Observando as imagens (Figura 2), no geral as adaptações encontram os pontos característicos fundamentais da face.

Os métodos de análise de componentes principais não-lineares podem ser agrupados em três classes (DIAMANTARAS e KUNG, 1996):

- **Redes hebbianas:** são obtidas substituindo-se os neurônios lineares nos algoritmos PCA baseados na aprendizagem hebbiana por neurônios não-lineares (KARHUNEN e JOUTSENSALO, 1995);

- **Redes replicadoras ou auto-codificadoras:** são construídas em torno de perceptrons de múltiplas camadas;

- **Curvas principais**: são baseadas em uma estimação iterativa de uma curva ou superfície relacionada com a estrutura dos dados (HASTIE e STUETZLE, 1989). Em Ritter et al. (1992) é mostrado que o mapa auto-organizável de Kohonen pode ser visto como um procedimento computacional para encontrar uma aproximação discreta das curvas principais.

Segundo Pentland (2000), o primeiro sistema famoso de reconhecimento automático de faces provavelmente foi o de Kohonen, proposto em 1989. Kohonen demonstrou que uma simples rede neural pode realizar o reconhecimento de faces usando imagens de faces registradas (normalizadas e alinhadas). Foi empregada uma rede que computa a descrição das faces através da aproximação dos autovetores da matriz de auto-correlação das imagens de face, que mais tarde ficaram conhecidos como *eigenfaces*. Apesar de famoso, o sistema de Kohonen não foi um sucesso prático, pelo fato de depender do alinhamento e da normalização das faces. Posteriormente muitos pesquisadores buscaram técnicas de reconhecimento de faces baseadas em características locais, como bordas, distâncias entre pontos característicos e outras abordagens, com o emprego de redes neurais. Resultados positivos foram obtidos com bases de dados pequenas com faces alinhadas. Por outro lado, os problemas mais realísticos de reconhecimento de faces associados com grandes bases de dados, com localização, orientação e escala da face desconhecidas, ainda não conseguiram ser resolvidas (PENTLAND, 2000).

2.4.2 Análise Discriminante Linear (Linear Discriminant Analysis – LDA)

Existem muitas técnicas para a classificação e redução de dimensionalidade dos dados na literatura estatística, mas duas delas são muito eficientes na compressão e reconhecimento de faces: PCA e Análise Discriminante Linear (LDA). O PCA já foi anteriormente discutido, e Análise Discriminante Linear pode ser entendida como um método, que busca maximizar o raio de variância entre as classes, e paralelamente minimizar a variância dentro das classes garantindo a máxima separabilidade entre as classes de um conjunto particular de dados. Para isso aplica uma transformação linear visando encontrar um sistema de coordenadas ótimas, que melhor represente os dados com máxima separabilidade. O uso do LDA para classificação de dados foi primeiramente aplicado ao problema de classificação e reconhecimento da fala, e posteriormente aplicado em faces visando prover maior performance que o PCA.

LDA é uma técnica clássica em reconhecimento de padrões desenvolvida primeiramente por Robert Fisher em 1936 (FISHER, 1938), por isso também é chamada de *Fisher's Linear Discriminant* (FLD).

A principal diferença entre PCA e LDA é que a aproximação PCA proporciona características que capturam as direções principais com relação ao espalhamento dos dados, observando as diferenças significantes entre as imagens, mas não tenta reduzir o espalhamento dos pontos característicos dentro da classe (intra-classe), pois trata cada imagem como única, sem agregar informações sobre a sua classe. Desde que nenhuma informação de agrupamento das classes é usada, exemplos da mesma classe e de classes diferentes são tratados do mesmo modo. Porém, LDA usa a informação de agrupamentos das classes e nos permite achar autovetores, ou *eigenfeatures* (autocaracterísticas) nas quais são enfatizadas as variações entre faces de indivíduos diferentes, enquanto as variações na mesma face do indivíduo provocada pelas mudanças de iluminação, expressão facial, orientação etc., é minimizada. Portanto PCA faz classificação de características e LDA faz classificação de dados. Em PCA, a forma e local do conjunto de dados originais mudam para um subespaço diferente, enquanto que o LDA não muda o local, mas apenas tenta prover mais separabilidade entre as classes e estabelecer uma melhor região de decisão entre elas (CHELLAPPA e ETEMAD, 1996).

Em Belhumeur et al. (1997) mencionam que se o conjunto de treinamento está separado por classe, faz sentido usar esta informação para construir um método mais seguro na redução da dimensionalidade do espaço de características. Portanto, usando métodos lineares embasados em informações das classes para redução de dimensionalidade, e classificadores simples no espaço de características reduzido, pode-se melhorar as taxas de reconhecimento obtidas pelo método de autofaces.

Enquadrando o problema de reconhecimento de faces com a técnica LDA pode-se dizer que a análise é similar ao método PCA proposto por Pentland et al. (1994), mas a diferença fundamental está presente no cálculo dos autovalores executada através das matrizes de dispersão entre-classes e intra-classe. Lembrando que as faces humanas vistas como padrões em três dimensões (x, y, níveis de cinza), apresentam muitas coisas em comum, são espectralmente muito similares e portanto considerando a imagem de face como um todo, tem-se alguns aspectos importantes para a discriminação, presentes nas características (autovalores) de baixa energia. Essas características podem não ser capturadas em uma redução de dimensionalidade e extração de características baseada em PCA (CHELLAPPA e ETEMAD, 1996).

2.5 Redes neurais aplicadas ao reconhecimento de faces

Muitas tarefas cognitivas são tarefas fáceis de serem executadas pelos seres humanos, demonstrando dificuldade para serem emuladas em sistemas computadorizados. Combinando as disciplinas de psicologia e engenharia pode-se encontrar soluções para alguns destes problemas. Empregando o mesmo estímulo para ambos os modelos, humanos e os computadores, é possível ampliar não apenas a compreensão do conhecimento humano, mas também o desenvolvimento de sistemas automáticos para o reconhecimento de faces. Buscando o aprimoramento de sistemas computadorizados para o reconhecimento de faces humanas, diferentes técnicas matemáticas têm sido pesquisadas que simulem o funcionamento do cérebro humano em suas funções de aprendizagem.

Redes neurais são modelos computacionais que têm sido aplicados tanto em sistemas de detecção como para o reconhecimento de faces (ALLINSON e ELLIS, 1992). Algumas regras pré-definidas para reconhecer uma face são baseadas em um

conjunto de regras utilizadas pelos seres humanos, formando assim exemplos representativos para as redes neurais. A habilidade da aprendizagem automática, sua robustez e a capacidade de se auto-adaptar torna a pesquisa em rede neural bastante atrativa (LIN e KUNG, 1997).

A seguir serão apresentados alguns trabalhos relevantes na área.

Tamura et al. (1996) aplica a rede neural MLP para identificar o sexo da pessoa, por meio de uma imagem facial, obtendo uma taxa de acerto de aproximadamente 90% para ambos os sexos. As três camadas da rede neural foram treinadas para as três resoluções de um conjunto de imagens composta de 15 homens e 15 mulheres de forma a identificar o sexo. Os resultados do teste de identificação para 30 imagens de treinamento e 30 imagens teste são mostrados na Tabela 2. Pode ser observado que o treinamento de imagens de alta resolução é mais efetivo do que o treinamento com imagens de baixa resolução.

Tabela 2: Taxa de reconhecimento (%) para o conjunto de faces do conjunto de dados.

Resolução do treinamento	**Resolução dos dados de teste**		
	32 x 32	**16 x 16**	**8 x 8**
32 x 32	90% (--)	90% (100)	87% (90)
16 x 16	83% (97)	83% (--)	83% (97)
8 x 8	67% (90)	70% (93)	87% (--)

Na Tabela 2 o primeiro número em cada coluna refere-se à taxa de reconhecimento para imagens testes não usadas para o treinamento. Os valores dentro dos parênteses representam as imagens de treinamento com diferentes resoluções.

Em Intrator et al. (1996), são descritas diferentes pesquisas para a redução da dimensionalidade e classificação, baseado na combinação do aprendizado supervisionado e não-supervisionado (INTRATOR, 1993). Primeiramente, são detectados automaticamente os olhos e a boca na imagem facial, usando a Transformada Simétrica Generalizada. O aprendizado supervisionado busca as projeções que minimizem o erro quadrático médio entre a saída da rede *feedforward* e a saída desejada. O aprendizado não-supervisionado busca as projeções que demonstrem

alguma estrutura de interesse nos dados, medindo a derivação da distribuição Gaussiana na modalidade multimodal.

O desempenho de uma rede neural aplicando o algoritmo de treinamento *backpropagation* pode ser facilmente melhorado executando inúmeros treinamentos e através do cálculo da média dos resultados (KIRBY e SIROVICH, 1990). A combinação de ambos os métodos pode ter melhor desempenho na generalização dos dados. A aplicação do treinamento híbrido em uma rede neural *feedforward* é realizada pela modificação na regra de aprendizagem das unidades ocultas, que refletirá nas demais camadas, conforme mostra a Figura 3.

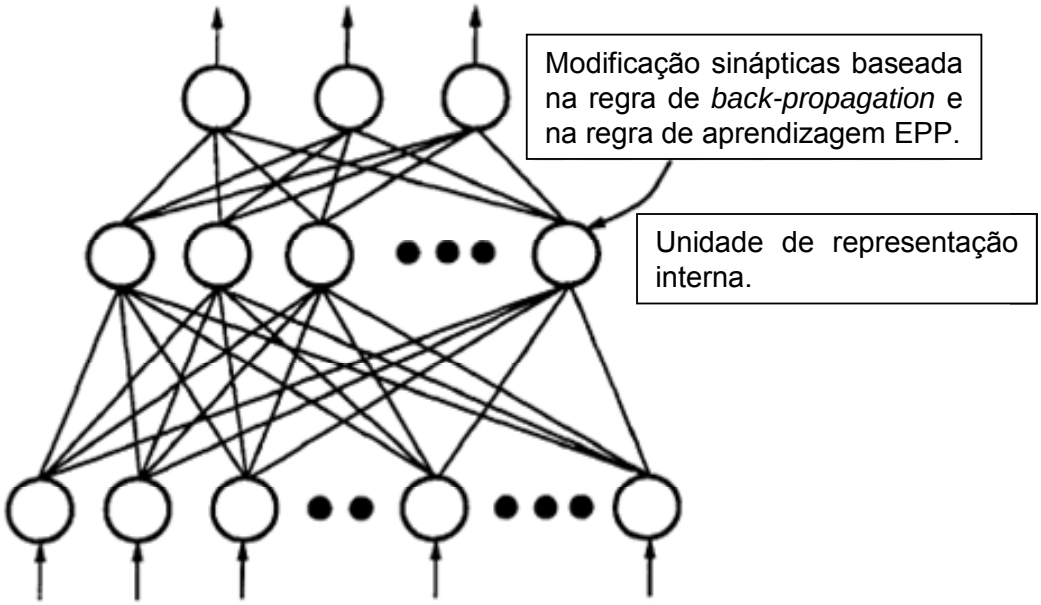


Figura 3: Uma rede neural híbrida EPP/PPR. A regra de aprendizagem não-supervisionada é aplicada apenas na camada oculta.

Há uma interação entre os neurônios adjacentes na camada oculta. Nos resultados apresentados, em todos os experimentos foi usada uma arquitetura *feedforward* formada por 12 neurônios na camada oculta, gerando os resultados apresentados na Tabela 3. O treinamento foi feito aplicando-se o algoritmo *backpropagation* para a parte supervisionada e o aprendizado através da projeção para a parte não-supervisionada.

As imagens faciais foram retiradas do banco de dados do MIT *Media Lab*. Esse banco de dados contém 27 exemplos de 16 pessoas diferentes. As imagens possuem diferentes condições de iluminação e localização da câmera. Das 27 imagens, 17 foram

escolhidas randomicamente para cada pessoa para ser usada no treinamento, enquanto as 10 restantes foram usadas para o teste (Figura 4).

Figura 4: Uma imagem normalizada para cada classe (esquerda) e a variação de apenas uma imagem facial normalizada (direita).

Os resultados da classificação são mostrados na Tabela 3. O fundo de cada imagem foi eliminado, resultando numa imagem de 60 x 40 *pixels* contendo somente a face.

Tabela 3: Resultado da classificação no conjunto de dados.

Método	**Sucesso (%)**
MLP	96.72 $\pm$ 0.31
Híbrido BCM / MLP	96.04 $\pm$ 0.96
Média MLP	**98.75**
Média Híbrido BCM / MLP	**98.38**

A rede MLP classifica de forma mais satisfatória. Os resultados do treinamento das redes com a adição do *bias* no BCM (BIENENSTOCK et al., 1982) não apresentaram resultados satisfatórios, ou seja, foi possível obter um melhor desempenho da rede em relação aos treinamentos efetuados sem este adicional. Um exemplo sobre a ação do *bias* sobre as redes pode ser visto na Tabela 3, pelo aumento do desvio padrão do resultado para o método híbrido BCM / MLP.

Figura 5: Extração de características usando a rede MLP (esquerda) e extração de características usando uma rede híbrida BCM / MLP (direita).

A Figura 5 mostra as representações da unidade oculta da rede MLP (esquerda) e da rede híbrida BCM / MLP (direita). Um total de 12 características foram extraídas, sendo que apenas 7 das projeções foram diferentes. Os resultados foram surpreendentes, apresentando uma eficiente redução de dimensionalidade, resultando em um ótimo desempenho na classificação das 16 diferentes faces usando apenas 7 características.

Lin e Kung (1997) desenvolveram um sistema de reconhecimento de faces baseado em decisões probabilísticas e em redes neurais (PDBNN). O sistema consiste de três módulos: o detector de faces que tem a função de encontrar a localização da face humana em uma imagem; o localizador de olhos determina a posição dos olhos para gerar um vetor de características significante; e finalmente o terceiro módulo é o de reconhecimento de face.

A rede neural baseada na decisão probabilística (PDBNN) tem o mérito das redes neurais e das aproximações estatísticas, sendo aplicada na detecção de faces, localização de olhos e no reconhecimento de faces em geral. Duas características do PDBNN fazem dela a execução apropriada não apenas para os sistemas de reconhecimento de faces, mas também em outros sistemas de identificação biométrica. Essas características são:

- **Característica arquitetural:** o PDBNN herda a estrutura modular de seu predecessor, Rede Neural Baseada em Decisão (DBNN) (KUNG e TAUR, 1995). A DBNN é uma eficiente rede neural para classificação. Para cada pessoa

ser reconhecida, o PDBNN dedica uma das sub-redes para representar uma determinada pessoa. Este tipo de estrutura é útil não apenas para a execução do treinamento e/ou reconhecimento, mas também para implementações em hardware.

- **Característica do desempenho:** a função discriminante do PDBNN é um comportamento da densidade probabilística. A falsa rejeição e as falsas taxas de aceitação são importantes não apenas para problemas de detecção, mas também para vários problemas de reconhecimento. O PDBNN é adequado para resolver o problema de falsa rejeição e falsa aceitação. A justificativa é que o PDBNN focaliza apenas uma distribuição de densidade em classes individuais, onde nenhuma sub-rede adicional da extra-classe é necessária.

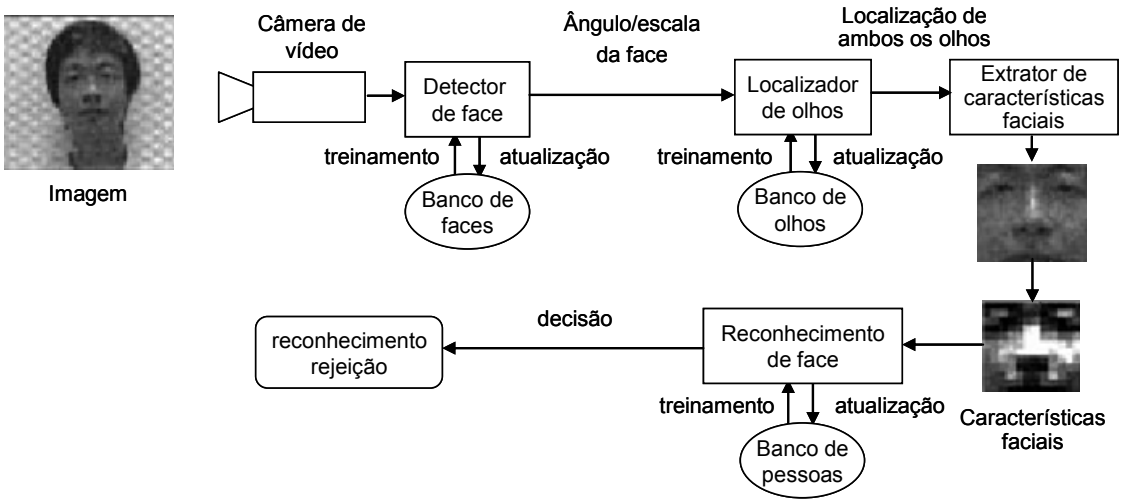


Figura 6: Diagrama do sistema de reconhecimento de face.

O sistema de reconhecimento de faces baseado no PDBNN foi desenvolvido com a colaboração entre a *Siemens Corporate Research, Princeton* e a *Princeton University*. O diagrama do sistema, descrito na Figura 6 mostra os quatro módulos: o detector de face, o localizador de olhos, o extrator de características e o reconhecimento de faces. As imagens são capturadas por meio de uma câmera, com resolução de 640 x 480 *pixels* de 8 bits em escala de cinza e posteriormente são redimensionadas para 320 x 240 *pixels* para prosseguir o processamento. Um detector de faces PDBNN é então

ativado para determinar se a imagem é uma face humana. Caso seja positivo, o localizador de olhos é ativado para localizar os olhos e a partir disso, uma sub-imagem de aproximadamente 140 x 100 *pixels* correspondendo à região facial será extraída, conforme ilustra a Figura 7. Por fim, o vetor de características é dado ao reconhecedor facial PDBNN para realizar o reconhecimento e uma subseqüente verificação.

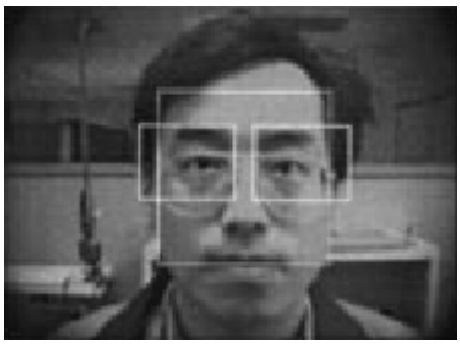


Figura 7: Resultado do detector de face. Baseado nisto, é aplicado o localizador dos olhos.

Os resultados empíricos do detector de faces PDBNN e da localização dos olhos são mostrados abaixo:

- **Detector de faces:** para a detecção de faces, 92 imagens foram usadas para o treinamento e 473 imagens para o teste. O tamanho da imagem foi de 320 x 240 *pixels*, e o tamanho da face foi de aproximadamente de 140 x 100 *pixels*. A variação da orientação da cabeça foi de aproximadamente 15 graus para as quatro direções (acima, abaixo, esquerda, direita). O desempenho da detecção foi medido pelo erro (*pixels*) entre a localização da face detectada e a verdadeira localização. Testando todas as imagens, o sistema detectou entre 98,5% a 100% das imagens de alta resolução (Figura 7).

- **Localização dos olhos:** a detecção dos olhos talvez seja o problema mais difícil, visto que:

1) os olhos são áreas muito pequenas;

2) os olhos podem estar abertos, fechados ou semi-fechados;

3) os óculos frequentemente ocultam as áreas dos olhos.

Isso requer mais imagens de treinamento, e nos experimentos, foram usadas 250 imagens detalhadas para o treinamento e 323 imagens para o teste. O sistema detectou entre 96,4% a 98,9%.

Tabela 4: Desempenho do sistema de reconhecimento de faces usando um banco de dados com grandes orientações faciais.

PDBNN	Conjunto de treinamento original	Conjunto de validação do treinamento
Reconhecimento	84.64%	98.34%
Falsa rejeição	10.03%	0.97%
Desclassificação	5.33%	0.69%

A Tabela 4 mostra o desempenho do reconhecimento de face PDBNN na validação do conjunto de teste de 40 pessoas. O conjunto de imagens faciais usadas para o treinamento foi selecionado do conjunto original de dados. Isto é, imagens faciais ruins foram também incluídas no conjunto de treinamento. Com isso, é possível afirmar que a taxa de reconhecimento é reduzida devido aos padrões de treinamento de baixa qualidade.

O resultado dos testes efetuados no sistema de reconhecimento de faces PDBNN utilizando o banco de imagens com 200 pessoas retiradas do banco de dados FERET (PHILIPS et al., 2000), podem ser observadas na Tabela 5.

Tabela 5: Desempenho dos diferentes sistemas de reconhecimento de faces aplicado em 200 pessoas do banco de dados FERET.

	Treinamento	Teste
Probabilística DBNN	97%	99%
Tradicional DBNN	100%	96%
Perceptron Multi-Camadas	99.5%	87.5%

Uma imagem de cada pessoa foi usada para o treinamento e a outra para o teste. Nos resultados do reconhecimento de faces, pode-se observar uma razoável precisão no treinamento do PDBNN (97%), porém durante o teste, o PDBNN alcança alta taxa de reconhecimento (99%) maior do que o tradicional DBNN (96%). A rede MLP usando o algoritmo de treinamento *backpropagation* foi implementada no sistema de reconhecimento de faces.

Ranganath e Arun (1997) desenvolveu um sistema de reconhecimento de faces, usando duas arquiteturas de redes de função de base radial (RBFN), e comparou o desempenho com outros algoritmos. A RBFN possui uma arquitetura *feedforward* com uma camada de entrada, uma camada oculta e uma camada de saída. O desempenho também foi avaliado para a extração dos vetores característicos das imagens faciais usando a análise de componentes principais (PCA) bem como a transformada de *wavelet*.

Hjelmas e Wroldsen (1997) e Huang e Wechsler (1999) aplicaram PCA para reduzir a dimensionalidade das imagens no trabalho de reconhecimento de faces que também usa redes neurais do tipo *perceptron* multicamadas e funções de base radial (RBFs) respectivamente. Usando PCA desta forma, foi possível distinguir características das faces com o propósito de se fazer a identificação.

Em Rowley et al. (1998) é apresentado um sistema de detecção de faces usando redes neurais. O algoritmo aplica uma ou mais redes neurais diretamente sobre imagens de entrada, onde cada rede é treinada para mostrar na saída, se o resultado é face ou não-face. O sistema é ilustrado na Figura 8 opera em dois estágios: primeiro é aplicado um conjunto de filtros baseados em redes neurais para uma imagem, e então é usado um arbitrador para combinar as saídas. Os filtros examinam cada local na imagem em

muitas escalas, procurando por locais que possam conter uma face. O arbitrador então agrupa as detecções dos filtros individualmente e elimina as detecções sobrepostas. O primeiro componente do sistema é um filtro que recebe como entrada uma região 20 x 20 *pixels* da imagem e gera uma saída associando os valores 1 ou -1, que representa a presença ou ausência de uma face, respectivamente.

O treinamento da rede neural para detecção de face é uma tarefa desafiadora, por causa da dificuldade na caracterização do protótipo da imagem "não-face". Diferente do reconhecimento de faces, no qual as classes estão discriminando as diferentes faces, ou seja, na detecção de faces, as classes estão discriminadas em "imagens contendo faces" e "imagens não contendo faces".

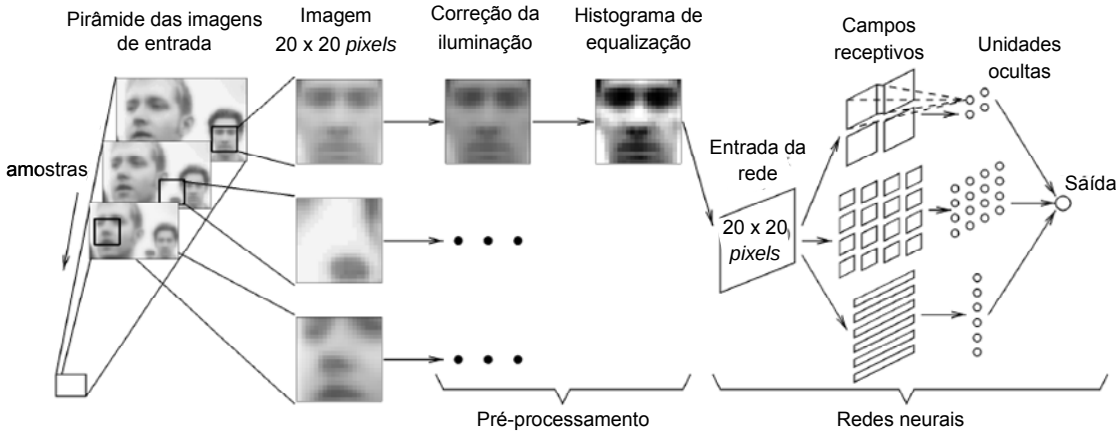


Figura 8: Algoritmo básico usado para detecção de face.

O algoritmo de filtragem é mostrado na Figura 8. Primeiro, um passo de pré-processamento é aplicado para um quadro da imagem. O quadro é então utilizado pela rede neural, que decide se o quadro contém uma face. Posteriormente, este quadro, contendo os *pixels* da imagem facial, é usado pelos campos receptivos como entrada da camada de entrada da rede neural (Figura 8). Há três tipos de unidades ocultas: 4 sub-regiões de 10 x 10 *pixels*, 16 sub-regiões de 5 x 5 *pixels*, e 6 tiras horizontais de 20 x 5 *pixels*. Todas as unidades ocultas são responsáveis por detectar as características locais que podem ser importantes para a detecção de faces.

Para treinar a rede neural usada no primeiro estágio, de forma a operar como um filtro preciso, um grande número de imagens face e não-face foram necessários. Aproximadamente 1.050 exemplos de faces foram retirados do banco de dados de

CMU, *Harvard*. As imagens faciais continham vários tamanhos, orientações, posições e intensidades (Figura 9).

Figura 9: (a) Exemplo de imagens faciais, randomicamente rotacionadas, transladadas e pela menor soma escalar; (b) Exemplo de imagens não-faciais.

O sistema foi testado com dois grandes conjuntos de imagens conhecidos como conjunto de teste 1. O conjunto de teste 1 é formado por uma coleção de 130 imagens do banco de dados CMU, fotografias escaneadas e figuras de jornais. O conjunto teste 2 é um subconjunto do banco de dados FERET, onde cada imagem contém uma face e tem um *background* uniforme e com boa iluminação. A taxa de detecção do algoritmo foi entre 77.9% e 90.3% das faces no conjunto de 130 imagens testadas, com um número tolerável de falsas detecções. Os resultados podem ser observados na Figura 10.

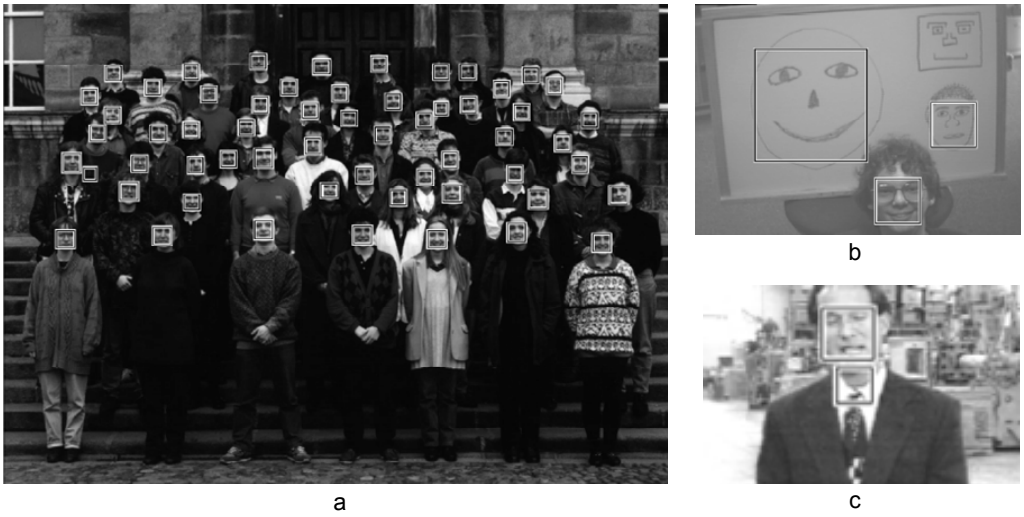


Figura 10: (a) Saída obtida pelo sistema de detecção com inúmeras pessoas na foto; (b) Exemplo de resultado obtido com o reconhecimento de um desenho e uma caricatura; (c) Detecção de uma não-face, ou seja, a face foi detectada corretamente, porém o sistema localizou um falso-positivo entre o pescoço e o tórax.

Em Talukder e Casasent (2001), é apresentada uma rede neural com funções de ativação adaptativa capazes de reconhecer as faces, usando apenas um exemplo de cada face original para compor a informação do autovetor.

Neste mesmo ano, Franco e Treves (2001), apresentaram um sistema de reconhecimento de expressões faciais usando redes neurais, aplicando o algoritmo de aprendizagem não-supervisionado. A rede é treinada com o algoritmo *backpropagation*, classificando os dados em categorias com diferentes expressões. Para o treinamento e os testes foram usadas imagens do banco de dados *Yale Face* (BELHUMEUR e KRIEGMAN, 1997), contendo 165 imagens em escala de cinza formada de 15 homens de diferentes raças e características (óculos, barba, bigode, etc) (Figura 11).

Figura 11: Exemplo do banco de dados *Yale Face*. São apresentadas 4 faces com diferentes expressões faciais (neutra, feliz, triste e surpreso).

Na Figura 11, o retângulo branco da imagem com expressão facial "surpresa" corresponde à área recortada, e usada como entrada para a rede neural. A arquitetura da rede consiste de quatro camadas de neurônios composta de unidades sigmoidais. A rede é composta por 192 entradas na camada de entrada, correspondendo a 24 x 8 *pixels* da área recortada das imagens originais. Cada neurônio de entrada transmite informações através de um simples peso hebbiano, projetando para um neurônio específico na primeira camada oculta, selecionado conforme o processo auto-organizado. Desse modo, tem neste nível uma nova redução da representação das imagens, 48 neurônios, que conservam os mesmos aspectos topológicos das entradas originais. A arquitetura da rede é mostrada na Figura 12, onde no topo é mostrada a imagem de entrada seguida pela estrutura dos campos receptivos correspondendo aos neurônios da primeira camada oculta. Nos resultados, a rede apresentou uma taxa de generalização de 84.5%.

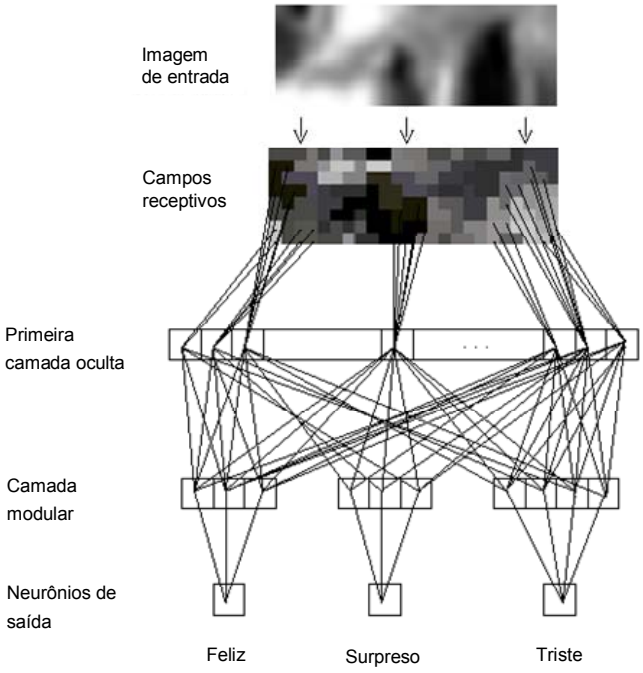


Figura 12: Esquema estrutural da arquitetura de rede usada para executar o reconhecimento de expressão facial.

A principal diferença entre as funções de redes neurais e outros métodos é que redes neurais podem ser usadas de uma maneira hierárquica, não apenas detectando os padrões mais básicos, mas também mais padrões de padrões e assim por diante (AITKENHEAD e MCDONALD, 2003).

Aitkenhead e McDonald (2003) desenvolveram um sistema de reconhecimento de faces humanas usando redes neurais, denominado FADER *(Face Detection and Recognition)*. O FADER é formado de três redes distintas, com a saída da rede 1 suprindo a entrada para a rede 2, e a saída da rede 2 suprindo a entrada para a rede 3. A rede 1 foi usada para detectar as faces nas imagens. A rede 2 detectou as subestruturas dentro das imagens faciais, ou seja, é aplicada para detectar as características salientes, de forma a construir uma descrição hierárquica da imagem. A rede 3 executou o reconhecimento facial. O diagrama em blocos apresentado na Figura 13 ilustra o funcionamento do sistema FADER:

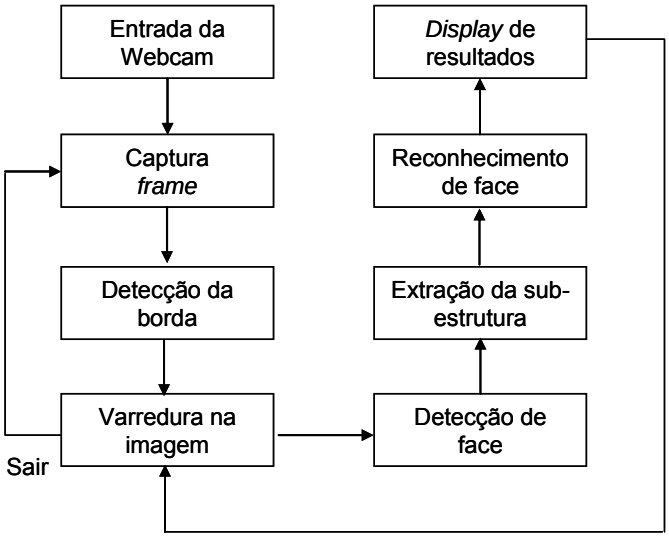


Figura 13: Estrutura modular do sistema *FADER*.

A rede neural 1 é composta de uma camada de entrada de tamanho 32 x 32, duas camadas ocultas de tamanho variável (com um limite entre 100 a 1000 neurônios) e a camada de saída formada por dois neurônios, correspondendo à presença ou ausência da face. Cada camada está totalmente conectada com seu sucessor, e o algoritmo de treinamento *backpropagation* foi usado para ajustar os pesos das conexões.

Um conjunto original de 200 faces foi usado para produzir um outro conjunto de 1000 faces, onde cada imagem foi rotacionada no sentido horário e anti-horário, entre 15º e 30º. Um conjunto de 1000 imagens não-faces com resolução 32 x 32 *pixels* foi criado através de 20 imagens capturadas com uma webcam em diferentes orientações, em um determinado ambiente, e extraindo 50 imagens com resolução 32 x 32 *pixels* para cada (Figura 14). As imagens dadas à rede neural alternam-se entre imagens contendo face e não-face. Um algoritmo detector de borda foi usado para transformar cada imagem de treinamento. O treinamento foi realizado para um total de 10.000 passos, com a rede controlando as variáveis (número de neurônios nas duas camadas ocultas, diferença entre a face positiva e negativa presente nos valores de saída).

Figura 14: Exemplo de imagens adquiridas por uma webcam.

A topologia utilizada foi uma camada de entrada com 1024 unidades de entradas, 453 neurônios na primeira camada oculta, 118 neurônios na segunda camada oculta e 2 neurônios na camada de saída da rede 1. A rede utilizada foi a MLP com o algoritmo de treinamento *backpropagation*, obtendo-se bons resultados na etapa da detecção facial (Figura 15). O maior valor alcançado usando 1000 faces obtidas de 126 imagens foi 94.7%, com 6 falsos positivos do conjunto de 1000 imagens geradas randomicamente.

a b

Figura 15: Exemplo do sistema de detecção facial: (a) imagem original e (b) imagem com as bordas detectadas e com as faces detectadas.

É aceito que o nível de acerto dado pelo sistema FADER neste trabalho, não é bom o suficiente para ser considerado para o uso em aplicações práticas. Contudo, este é um exemplo de reconhecimento facial usando um sistema inspirado no aprendizado

biológico, muito melhor que as técnicas matemáticas e que usa uma rede neural muito mais simples que o cérebro humano.

2.6 Outras técnicas para o Reconhecimento de Faces

2.6.1 Máquinas de Vetores de Suporte (SVM)

Support Vector Machines (SVM) é uma técnica de aprendizado de máquina, fundamentada nos princípios indutivos da Minimização do Risco Estrutural. Estes princípios são provenientes da Teoria do Aprendizado Estatístico, a qual está baseada no fato de que o erro da técnica de aprendizagem junto aos dados de validação (erro de generalização) é limitado pelo erro de treinamento. SVM pode ser visto com um novo método para o treinamento polinomial, redes neurais e classificadores de funções de base radial. O SVM implementa um mapeamento não-linear dos dados de entrada para um espaço característico de alta dimensionalidade, em que um hiperplano ótimo é construído para separar os dados linearmente em duas classes (VAPNIK, 1999).

Quando os dados de treinamento são separáveis, o hiperplano ótimo no espaço característico é aquele que apresenta a máxima margem de separação (Figura 16). Para dados de treinamento em que as amostras das duas classes apresentam superposição (dados não separáveis), uma generalização deste conceito é utilizada.

A técnica SVM foi introduzida por Vapnik em 1992 (BOSER et al., 1992), formulada com todas as demonstrações matemáticas (VAPNIK, 1995) e, em 1998 com maiores detalhes (VAPNIK, 1998). Outras boas referências sobre o assunto podem ser encontradas em Burges (1998) e Haykin (2001).

As áreas mais tradicionais com aplicações de SVM estão citadas a seguir, porém muitas outras aplicações podem ser encontradas na literatura:

- Reconhecimento de Dígitos Escritos à Mão (BOSER et al., 1992);
- Reconhecimento de Imagens (PONTIL e VERRI, 1998);
- Classificação de Textos (JOACHIMS, 1998 e DUMAIS et al., 1998);
- Bioinformática – Análise de dados de expressão gênica (BROWN et al., 2000);

- Detecção de imagens da Face Humana, perante qualquer outra imagem (OSUMA et al., 1997);
- Database Marketing: Prospecção de clientes para a aquisição de novos produtos (BENNETT et al., 1998).

SVM usa uma função de decisão, que é definida pela Equação 2.1:

$$f(x) = \operatorname{sgn}\left(\sum_{x_i \in SV} \lambda_i K(x, y) + b \right), \tag{2.1}$$

onde, *K(x,y)* é um *kernel* e SV é um subconjunto de exemplos de treinamento, conhecido como vetores de suporte (*support vectors*). Os vetores de suporte são representados por x_i e os pesos por λ_i .

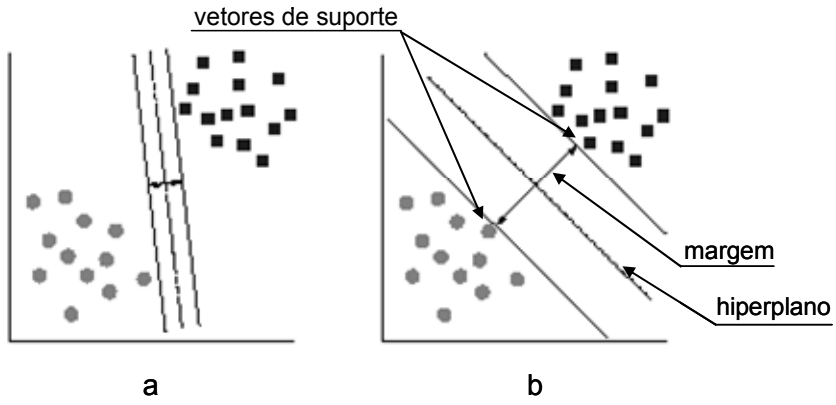


Figura 16: Classificação usando hiperplano entre duas classes: (a) Hiperplano traçado arbitrariamente entre as classes; (b) o OSH com a grande margem identificada pelas linhas tracejadas, passando pelos dois vetores de suporte.

De acordo com Vapnik, este hiperplano é chamado de *Optimal Separating Hyperplane* (OSH), o qual minimiza os riscos de desclassificação não apenas dos exemplos de treinamento, mas também dos exemplos do conjunto de teste. O SVM tem sido proposto para solucionar problemas em Visão Computacional e alguns trabalhos dessa área serão apresentados a seguir.

Osuma et al. (1997) treinou um SVM para detecção de faces, diferenciando as classes entre face e não-face, conforme mostra a Figura 17.

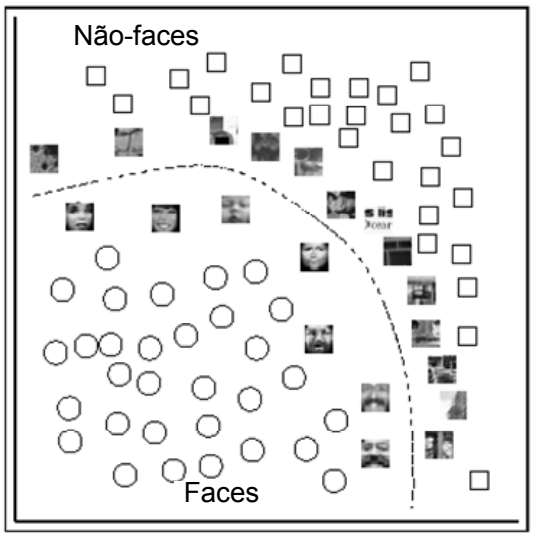


Figura 17: Nesta figura os círculos representam os padrões face e os quadrados os padrões não-face. Na borda entre as duas classes se encontra um vetor de suporte do sistema.

Foram usados dois conjuntos de imagens. O conjunto A composto por 313 imagens de alta resolução, com uma face por imagem. O conjunto B contendo 23 imagens com diferentes resoluções, totalizando 155 faces. Ambos os conjuntos foram testados usando o detector facial proposto e o de Sung e Poggio (1994).

A Tabela 6 mostra uma comparação entre os dois sistemas. Em tempo real, o sistema de detecção de faces usando SVM é aproximadamente 30 vezes mais rápido do que o sistema de Sung e Poggio (1994).

Tabela 6: Comparação da execução do SVM no sistema de detecção de faces.

	Conjunto A		**Conjunto B**	
	Taxa de detecção	Alarmes falsos	Taxa de detecção	Alarmes falsos
SVM	97.1%	4	74.2%	20
SUNG e POGGIO (1994)	94.6%	2	74.2%	11

A Figura 18 mostra os resultados do sistema, para algumas imagens testadas. Nota-se que o sistema é capaz de detectar as faces, mesmo onde algumas faces não são frontais.

Figura 18: Resultado do sistema de detecção de faces usando SVM.

Fernández e Viennet (1999) usaram o *Support Vector Machine* (SVM) para resolver o problema de identificação de face e desenvolveram dois módulos. O primeiro usa o *Principal Component Analysis* (PCA) para a extração das características faciais e cria os vetores de entrada do SVM. O segundo módulo é um classificador SV com invariância. Para o treinamento e o teste do sistema, foram utilizadas 200 imagens do banco de dados ORL (ORL). O resultado do trabalho, mostrou que bons resultados foram alcançados com a aplicação do SVM no reconhecimento de faces, obtendo-se uma taxa de erro de 1,5%.

Em Guo et al. (2000), foi proposto o uso do SVM com uma árvore binária (Figura 19), para resolver o problema do reconhecimento facial. Foi realizada também uma comparação entre o reconhecimento de face usando SVM, com as autofaces (eigenfaces) padrões, usando o critério de Classificação do Centro Mais Próximo (NCC). Para testar este sistema, foram efetuados dois experimentos. No primeiro experimento é usado um banco de dados ORL (ORL), o qual consiste de 400 imagens de 40 faces individuais, contendo variações em expressões faciais, posição facial e detalhes faciais (Figura 20). No segundo experimento, foi usando um grande conjunto

de dados contendo 1079 imagens de 137 faces individuais, do banco de dados de *Cambridge*, *Bern*, *Yale*, *Harvard* e um banco de dados próprio.

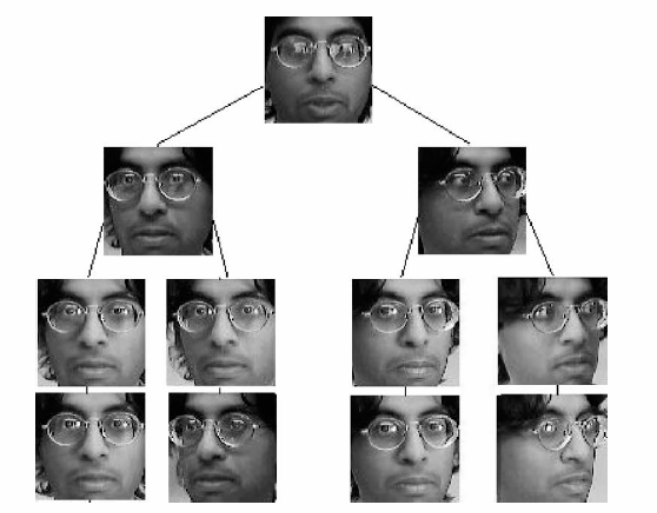


Figura 19: Exemplo de uma árvore binária de uma imagem facial.

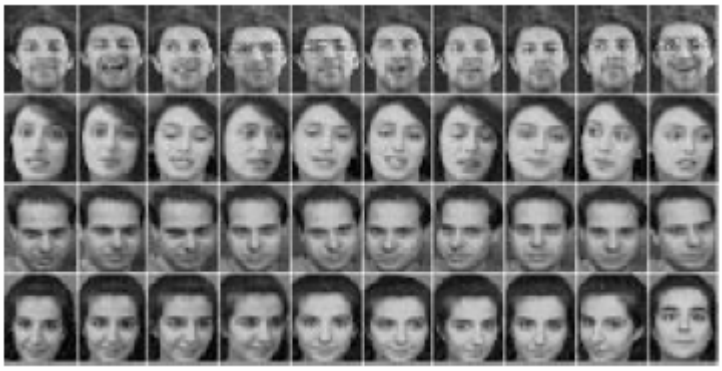


Figura 20: Exemplo do banco de dados faciais usadas no primeiro experimento.

Para os testes do reconhecimento de faces usando o banco de dados ORL (ORL), foram selecionados randomicamente 200 exemplos de imagens para o conjunto de treinamento, no qual foram calculadas as autofaces (*eigenfaces*) e o treinamento do SVM. Os 200 exemplos remanescentes foram usados para o conjunto teste. A média da taxa de erro do SVM é 3.0%, enquanto o NCC é de 5.25%.

No segundo teste, foram usados cinco bancos de dados diferentes, onde cada um era composto por: (1) o banco de dados facial ORL, descrito anteriormente; (2) o banco de dados *Bern*, contendo imagens faciais frontais de 30 pessoas; (3) o banco de dados

Yale, contendo 15 pessoas; (4) o banco de dados *Harvard*, onde foram selecionadas 5 pessoas e (5) um banco de dados próprio, contendo 179 imagens faciais frontais de 47 estudantes chineses, onde cada pessoa tinha três ou quatro imagens com diferentes expressões faciais, posição facial e detalhes faciais. As taxas de erro foram calculadas em função do número de autofaces, isto é, as dimensões das características. O resultado mostrou que a taxa de erro mínimo do SVM foi de 8.79%, enquanto a do NCC foi muito maior, 15.14%.

Nos experimentos de Guo et al. (2000), os resultados demonstram que o SVM é o melhor algoritmo de aprendizado do que o NCC para o reconhecimento de faces.

Kumar e Poggio (2000) descrevem um sistema de treinamento, capaz de rastrear faces e características faciais através dos olhos, das narinas e estimando as características básicas da boca. O sistema foi dividido em diferentes estágios, iniciando com a localização da face, seguindo com a localização das características faciais e a estimação dos parâmetros da boca. A classificação é realizada através da técnica de *Support Vector Machine* (SVM) que é aplicada após as etapas de segmentação facial, detecção facial e detecção dos olhos.

Heisele et al. (2003) desenvolveu um componente e dois sistemas globais para o reconhecimento de faces. No componente do sistema, primeiro são localizados os componentes faciais, as características são extraídas e são armazenadas em um único vetor de características, o qual é classificado por um SVM. Os dois sistemas globais reconhecem as faces, classificando um simples vetor de características formado de valores cinza de toda a imagem. No primeiro sistema global, é treinado um único classificador SVM para cada pessoa no banco de dados. O segundo sistema consiste do conjunto dos classificadores SVM, gerando agrupamentos durante o treinamento, envolvendo agrupamentos durante o treinamento.

2.6.2 Sistema Neuro-Fuzzy

Sistema neuro-fuzzy é um sistema resultante da combinação entre redes neurais e lógica fuzzy, reunindo as vantagens e evitando as desvantagens de cada método (NAUCK et al., 1992).

Os modelos fuzzy são capazes de compartilhar informações precisas e modelar o conhecimento humano, mas falta habilidade em aprender. Por outro lado, as redes neurais possuem a vantagem da computação paralela, tolerância de erro e a capacidade de aprendizagem auto-adaptativa, porém sua desvantagem é a falta de interpretação física dos pesos de conecção (ZHANG et al., 2004).

Um sistema neuro-fuzzy significa empregar estratégias de aprendizado heurístico através das teorias de redes neurais para sustentar o desenvolvimento de um sistema fuzzy. Então, neuro-fuzzy pode ser considerada como uma técnica para derivar um sistema fuzzy dos dados ou para elevar a aprendizagem dos exemplos (NAUCK e KRUSE, 1997b).

Nos modelos neuro-fuzzy, a conecção dos pesos, a propagação e a função de ativação diferem das redes neurais. O método neuro-fuzzy pode ser considerado como uma ferramenta para produzir um sistema fuzzy (Nauck e Kruse, 1997).

Segundo Kruse e Nauck (1995) existem várias aproximações para combinar redes neurais e sistemas fuzzy:

- **Processo Neural / Modelos Fuzzy:** Esta é a combinação mais fraca do sistema neuro-fuzzy. A rede neural e o sistema fuzzy são usados no processamento da mesma tarefa. A rede neural pode ser usada no pré-ou pós-processamento no sistema fuzzy, porém a rede neural não altera nenhum parâmetro do sistema fuzzy;

- **Modelos Neuro-Fuzzy Cooperativos:** Neste caso a rede neural, é usada para aprender certos parâmetros (conjuntos fuzzy, regras fuzzy ou pesos) do sistema fuzzy, através do algoritmo de aprendizagem neural. Após o aprendizado, a rede neural não está mais presente, apenas os resultados do sistema fuzzy são usados;

- **Modelos Neuro-Fuzzy Híbridos:** Uma nova arquitetura é criada utilizando rede neural e lógica fuzzy.

Os conjuntos fuzzy são geralmente exemplos para determinar um número de valores sobre o domínio deles e representar através de vetores o grau de liberdade. Para cada grau de liberdade, é usada uma entrada e um simples neurônio de saída. O grau de liberdade é geralmente parametrizado através das funções trapezoidais ou funções triangulares. A regra do algoritmo de aprendizado seleciona regras fuzzy baseadas numa seleção predefinida do espaço de entrada dado por um conjunto fuzzy.

Em 1996, Lee et al. (1996) propôs o uso do algoritmo neuro-fuzzy para executar a tarefa de reconhecimento de faces. Este algoritmo combina conceitos de redes neurais e a lógica fuzzy, onde as redes neurais têm a capacidade de aprender e o fuzzy de absorver variações em valores característicos. Por ser um método eficiente, com o algoritmo neuro-fuzzy é possível aumentar a taxa de reconhecimento e reduzir a complexibilidade computacional. O conceito fuzzy foi aplicado convertendo os valores característicos em dados fuzzificados, nas entradas do algoritmo EBP, conforme mostra a Figura 21. No algoritmo neuro-fuzzy foram usados dados fuzzy como entradas para a rede neural. Às vezes, a variação dos valores caracteristicos é muito grande, o que dificulta o reconhecimento na base de faces com estes valores caracteristicos. Para resolver este problema, cada valor de característica facial foi convertido em três dados fuzzy, onde o processo de aprendizagem é executado usando o algoritmo EBP. Após este processo, finalmente as faces são identificadas usando o algoritmo EBP. Para reduzir a complexibilidade computacional e o problema do mínimo local, foi usado o algoritmo EBP modificado.

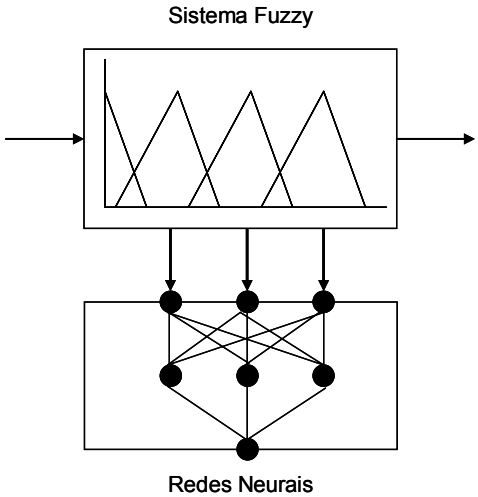


Figura 21: Diagrama esquemático do Sistema Neuro-Fuzzy.

Nesta pesquisa, Lee et. al. (1996) usou um banco com 80 imagens faciais frontais, organizado em quatro bancos de imagens, onde cada banco é constituído por 20 pessoas, sendo 14 homens e 6 mulheres. As imagens foram obtidas em diferentes condições de iluminação e quantizadas em 8 *bits*. Concluiu-se que a proposta de se aplicar o algoritmo neuro-fuzzy resultou em um eficiente aprendizado, e que o algoritmo de reconhecimento pode ser aplicado a um grande conjunto de imagens faciais.

A Tabela 7 mostra os resultados do reconhecimento usando o algoritmo EBP, e o algoritmo neuro-fuzzy proposto. Nesta tabela é possível observar que a taxa de reconhecimento do algoritmo EBP é menor que a do algoritmo neuro-fuzzy. A variável *N* representa o número de imagens faciais utilizadas por cada banco.

Tabela 7: Taxa de reconhecimento por algoritmo EBP e o proposto algoritmo neuro-fuzzy.

		Taxa de reconhecimento (%)				
Método de reconhecimento	**N**	**db1**	**db2**	**db3**	**db4**	**Média**
	10	90.0	92.5	92.5	90.0	91.3
Algoritmo EBP	15	90.0	90.0	91.7	86.7	89.6
	20	90.0	88.8	91.3	82.5	88.2
	10	92.5	95.0	97.5	95.0	95.0
Algoritmo Neuro-Fuzzy	15	93.3	93.3	96.7	93.3	94.2
	20	93.8	95.0	93.8	93.8	94.1

As simulações com o algoritmo neuro-fuzzy proposto, usando uma função de inferência trapezoidal com alterações em vários parâmetros, mostraram melhores resultados de reconhecimento do que o método convencional (LEE et al., 1996).

Muitos pesquisadores têm proposto vários modelos neuro-fuzzy adicionando complexos algoritmos de treinamento. Dentre eles, Horikawa et al. (1992) propôs um modelo neuro-fuzzy usando funções sigmóides para gerar as entradas das funções de inferência e treiná-lo com o algoritmo *backpropagation*. Exemplo de outros modelos neuro-fuzzy incluem redes neurais baseada em sistemas fuzzy (WANG e RONG, 1999) e redes neurais *perceptron* fuzzy (CHEN e CHANG, 2000). Um *perceptron* fuzzy

genérico possui a arquitetura de um perceptron multicamadas, porém os pesos são modelados com conjuntos fuzzy e a ativação, a saída e as funções de propagação são alteradas de acordo com a implementação do caminho de inferência fuzzy (NAUCK e KRUSE, 1997a). Por ser uma área de pesquisa muito recente, existe poucos trabalhos aplicados especificamente ao reconhecimento de faces humanas usando neuro-fuzzy.

2.7 Considerações finais

A revisão bibliográfica sobre reconhecimento facial apresentada neste capítulo, destaca alguns dos principais métodos da área.

Os métodos holísticos utilizam à face como um todo, e não perdem nenhuma informação. Devido a isso fornecem resultados mais precisos no reconhecimento. Por outro lado, apresentam o problema da alta dimensionalidade dos dados, e são sensíveis à variação de posição e escala.

Redes neurais têm sido aplicadas tanto para detecção como para o reconhecimento de faces. Características como aprendizagem automática, robustez e capacidade de se auto-adaptar, tornam essas redes de grande interesse para aplicações em reconhecimento facial. A desvantagem que apresentam é que a aprendizagem necessita de um elevado esforço computacional em processamentos seqüenciais. Isso depende de vários fatores, tais como: o algoritmo de treinamento, a plataforma computacional, etc.

Máquinas de vetores de suporte é um método de aprendizado estatístico, tendo apresentado excelentes resultados para o reconhecimento facial. A vantagem se utilizar este método é que ele fornece um bom desempenho de generalização em problemas de classificação de padrões. A desvantagem desse método é o tempo de execução gasto pela generalização, que é mais lenta que as redes neurais (HAYKIN, 2001).

Sistemas neuro-fuzzy combinam redes neurais com lógica fuzzy, de forma a reunir as vantagens de cada método. Por ser uma área de pesquisa mais recente, há poucos trabalhos em reconhecimento de faces.

3 Técnicas de extração de características para o reconhecimento de faces

3.1 Considerações iniciais

Neste capítulo é realizada uma revisão dos trabalhos mais recentes na literatura sobre algoritmos de extração de características a partir de imagens faciais.

3.2 Introdução

Como já mencionado na Seção 1.1, a extração de características faciais, refere-se à determinação das medidas do rosto que nunca se alteram tais como: a distância entre os olhos, a distância entre olhos, nariz e boca, e a distância entre olhos, queixo, boca e contorno dos cabelos (LEE et al., 1996).

Existem muitos métodos para localizar e extrair essas características faciais numa imagem. Tradicionalmente, o reconhecimento de face está baseado na extração de características (BLEDSOE, 1996). Estas características são detectadas manualmente ou através de algoritmos automáticos (BRUNELLI e POGGIO, 1992).

Dentre os métodos existentes serão abordados aqueles de maior interesse para esta pesquisa e serão apresentados nos itens seguintes.

3.2.1 Método baseado em projeção horizontal e vertical

Baskan et al. (2002) usa a projeção horizontal e vertical dos *pixels* de valor cinza para detectar e localizar características faciais. O conjunto de imagens de faces usadas neste sistema é constituído por imagens coloridas com diferentes condições de iluminação. Baseado nas projeções horizontais e verticais dos *pixels* de valor cinza, o sistema limita cada característica facial com o menor retângulo, o qual é comparado com o retângulo extraído manualmente ao redor da característica facial correspondente.

A projeção horizontal e vertical baseada nos *pixels* de valor cinza, é um dos métodos mais populares baseado em características geométricas. Este tipo de método tem sido usado particularmente para encontrar a posição das características faciais, acompanhado por um algoritmo *template matching* ou outro método de refinamento do resultado (CHOW e LIN, 1993; KAMPMANN e ZHANG, 1998; YIN e BASU, 1999).

O trabalho de Baskan et al. (2002) baseia-se na observação que as características faciais diferem do resto da face pelos valores de intensidade dos *pixels* e são determinadas pela localização do mínimo valor de cinza nas projeções dos *pixels* da imagem da face. Nesse método foram computados dois tipos de projeções *X* e *Y*, considerando a média dos valores dos *pixels* da região da face segmentada ao longo da direção vertical (coluna) e horizontal (linha), respectivamente. Assim, cada característica facial gera um valor mínimo na projeção *Y*, especificando as características na projeção *X* (Figura 22).

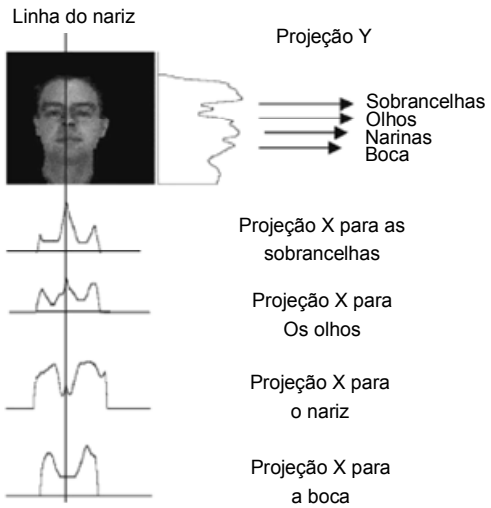


Figura 22: Projeção *Y* e *X* para uma simples imagem.

A linha que passa pelo nariz é considerada como o eixo simétrico para a face. Analisando a projeção *Y*, o primeiro mínimo corresponde à sobrancelha, o segundo mínimo aos olhos, o terceiro às narinas, o quarto à boca e o último ao queixo. A análise da posição mínima da projeção *Y* correspondente à projeção *X* é usada para localizar as características faciais. Com a projeção *Y* detecta-se o limite superior e inferior das características faciais e com a projeção *X* determina-se o limite esquerdo e direito. Estes limites determinam o retângulo para a característica facial. Requerimentos similares são usados para detectar as sobrancelhas, os olhos e o nariz (BASKAN et al., 2002). A Figura 23 ilustra este procedimento para a região da boca.

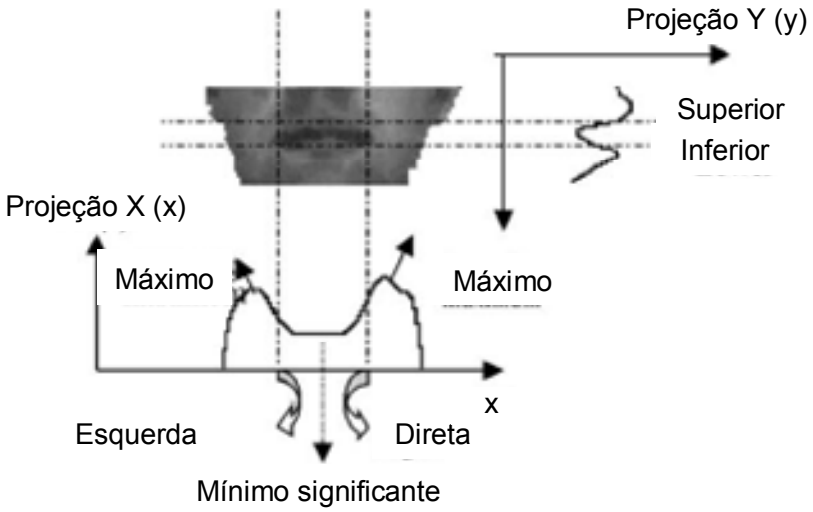


Figura 23: Projeção *Y* e *X* das características da boca.

Este algoritmo foi aplicado sob o banco de dados AR (MARTINEZ e BENAVENTE, 1998) contendo faces frontais de 125 indivíduos, com diferentes expressões faciais, totalizando 375 imagens faciais coloridas. Este método de extração de características detecta as seguintes características faciais: sobrancelhas, olhos, nariz, boca e orelhas e localiza cada característica facial com o menor retângulo possível (Figura 24) e com um alto grau de acerto, exceto para as orelhas. A detecção da orelha é muito mais difícil, pois em algumas imagens frontais somente parte da mesma pode ser observada, devido à região do cabelo.

Figura 24: Resultados do banco de dados AR.

3.2.2 Extração de pontos característicos da face

Gu et al. (2003), desenvolveu um método para extrair pontos característicos de imagens faciais, automaticamente. Baseia-se nas características visuais humanas e usa geometria e simetria facial, de forma a extrair características com as propriedades de escala, translação e invariância na rotação. O algoritmo desenvolvido é capaz de direcionar um provável caminho para localizar as posições dos olhos, as extremidades dos olhos, ponto central da narina e as extremidades da boca em uma imagem facial (Figura 25).

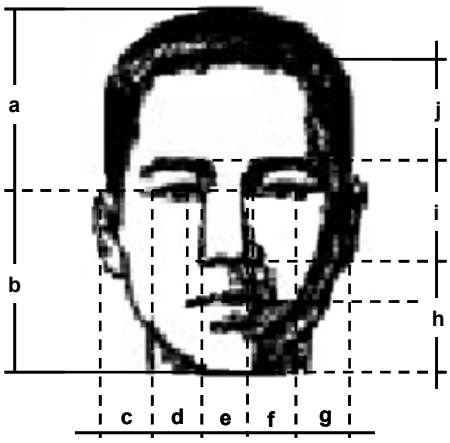


Figura 25: Relação entre as características faciais.

O número de pontos característicos deve ser suficiente para permitir a identificação da face. Se a base de dados tiver diferentes posturas de cada pessoa a ser reconhecida, a invariância do ângulo da característica geométrica é muito importante. O

método de Gu et al. (2003), localiza nove pontos característicos selecionados com invariâncias no ângulo, incluindo dois pontos centrais nos olhos, quatro pontos das extremidades dos olhos, um ponto central do nariz e dois pontos das extremidades da boca, como mostra a Figura 26.

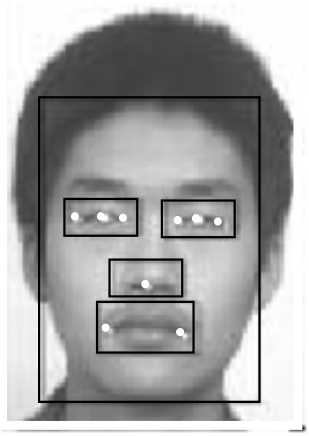


Figura 26: Nove pontos característicos vitais na face.

O operador SUSAN (*Smallest Univalue Segment Assimilating Nucleus*) é usado para definir as bordas dos componentes faciais e determinar os pontos característicos vitais na face.

O algoritmo localiza inicialmente a área facial na imagem. Conforme um prévio conhecimento da face humana, a área facial pode ser identificada através de um modelo facial normalizado (Figura 27). O modelo facial é baseado relativamente na distribuição da iluminação por toda a face. Antes de gerar o modelo facial, é aplicada a simetria, a distribuição de iluminação e o gradiente local, para certificar que as imagens contêm faces. A sub-imagem e o modelo usado para identificar é relativamente menor, aproximadamente um quarto da imagem original, para assim, fazê-lo robusto às mudanças da escala, rotação da imagem e na velocidade da identificação do modelo facial.

Figura 27: Imagem modelo da face humana.

A projeção direcional é usada para localizar os olhos, pela análise do método do vale que analisa a intensidade dos *pixels* da imagem. Um exemplo da localização dos globos oculares é mostrado na Figura 28, onde é mostrada a projeção horizontal aplicada na imagem facial.

Figura 28: Projeção horizontal na região dos olhos.

Com o operador SUSAN são localizadas as extremidades dos olhos (Figura 29).

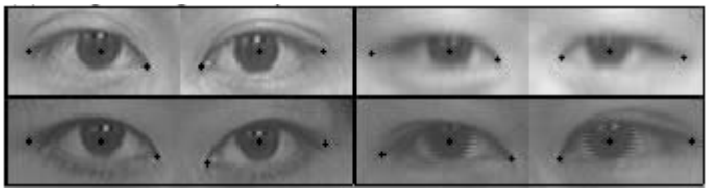


Figura 29: Localizando a íris e as extremidades dos olhos.

Para identificar a região do nariz é definida uma região retangular que passa pelas extremidades internas dos olhos e em seguida através da projeção horizontal e vertical dessa região, é identificada a narina esquerda e direita (Figura 30).

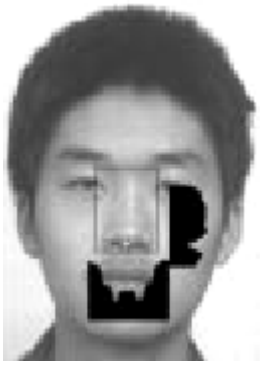


Figura 30: Processamento da área do nariz.

A boca é importante em pesquisas relacionadas ao reconhecimento de expressões faciais, pois nesse caso, a sua forma e tamanho se alteram muito. Outro fator que pode influenciar é a presença de bigode na região da boca.

Para localizar as extremidades da boca, é definida uma região retangular simétrica em relação ao ponto central do nariz e limitada pelos segmentos de reta que passam pelo centro dos olhos. Utilizando a projeção horizontal, a extremidade da boca é identificada pelo primeiro ponto do vale, como ilustrado na Figura 31.

Figura 31: Localizando as extremidades da boca.

Nos experimentos de Gu et al. (2003), foram extraídas nove pontos característicos de um conjunto com 270 pessoas, resultando em uma média de acerto de 95%, como é mostrado na Tabela 8.

Tabela 8: Tabela da localização dos pontos característicos.

Pontos característicos	Porcentagem de acertos
Globo ocular esquerdo	99,63%
Globo ocular direito	100%
Extremidade esquerda do olho esquerdo	98,52%
Extremidade direita do olho esquerdo	97,41%
Extremidade esquerda do olho direito	95,93%
Extremidade direita do olho direito	97,41%
Ponto central das narinas	98,15%
Extremidade esquerda da boca	84,81%
Extremidade direita da boca	84,44%
Média	**95,14%**

O algoritmo apresentou robustez em imagens com rotação da face em até 15º nos dois sentidos, direito e esquerdo, acima deste valor apresentou uma redução na porcentagem de acerto para a localização dos pontos característicos. Essa redução foi também verificada em imagens com forte iluminação lateral.

3.2.3 Extração de características baseado em conhecimento

Lee et al. (1996) usa regras biométricas para a extração das características faciais. Essas regras biométricas foram definidas da seguinte maneira (Figura 32):

- **Conhecimento 1:** o valor de maior projeção horizontal determina a posição dos olhos;
- **Conhecimento 2:** a distância vertical entre o olho e o nariz é maior do que a do nariz para a boca ($\overline{y_1 y_2} > \overline{y_2 y_3}$);
- **Conhecimento 3:** a face é simétrica com o nariz no centro da imagem ($\overline{x_1 x_2} \cong \overline{x_2 x_3}$);
- **Conhecimento 4:** define a posição relativa entre o olho esquerdo, o nariz e o olho direto ($x_1 < x_2 < x_3$);

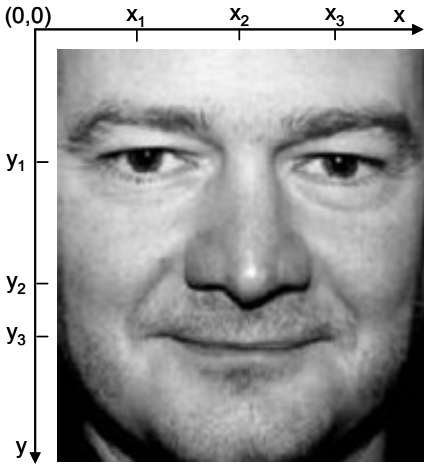


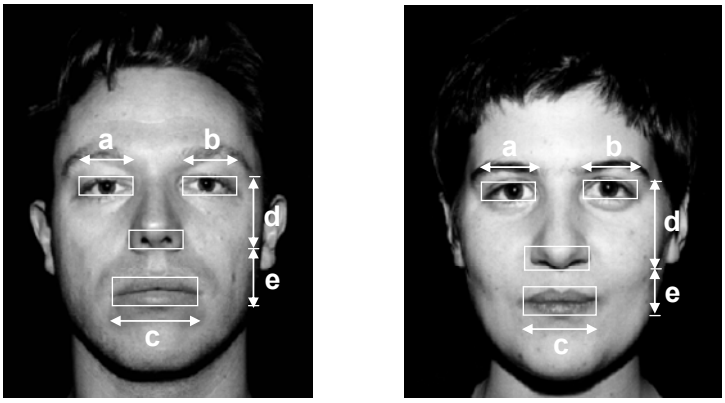
Figura 32: Parâmetros para a representação do conhecimento de uma imagem facial.

A Figura 32 exibe os parâmetros que representam as projeções horizontais e verticais, definidas respectivamente pelos eixos x e y onde, os pontos x_1 e x_3 representam a projeção horizontal dos olhos; x_2 denota a projeção horizontal que indica a localização do nariz e da boca; y_1, y_2 e y_3 é a projeção vertical que denotam as posição dos olhos, nariz e boca, respectivamente.

Este algoritmo extrai as características usando o seguinte procedimento: inicialmente calcula a projeção horizontal em cada linha da imagem facial e determina a posição *y* dos olhos baseado no fato que o valor da projeção horizontal dos olhos é maior. Então, abaixo da posição dos olhos são extraídas duas projeções horizontais que correspondem ao nariz e à boca. Se a distância vertical entre o nariz detectado e o olho for menor do que entre o nariz e a boca, o processamento é reiniciado após deletar os valores das projeções que estão próximas ao olho. Neste caso, é definido que a posição do olho detectado está incorreta por causa dos cabelos ou da sobrancelha. A posição *x* dos olhos é determinada localizando duas regiões que possuem grandes valores na projeção vertical computados sobre o intervalo limite de ± 8 pixels do eixo y da posição do olho detectado. Por fim, é determinado o valor do eixo y de cada parte, selecionando o maior valor da projeção horizontal computado sobre o intervalo limite de ± 10 pixels do valor do eixo y do componente facial detectado (LEE et al., 1996).

Posteriormente, cada componente facial é definido por um retângulo de fronteira mínima (MBR). Os resultados da extração de característica facial (os retângulos dos olhos, nariz e boca) são mostrados na Figura 33. Foram determinadas:

- as larguras (a e b) dos olhos;
- a largura (c) da boca;
- a distância vertical (d) que refere-se à distância entre a borda superior do retângulo dos olhos e a borda inferior do retângulo do nariz;
- a distância vertical (e) que refere-se a distância entre a borda inferior do retângulo do nariz e a borda inferior do retângulo da boca.



(a) (b)

Figura 33: Resultado da extração de característica facial: (a) e (b) resultado da localização dos componentes faciais usando MBR.

Usando as distâncias computadas entre os componentes faciais (olhos, nariz e boca) pode-se criar um vetor de características faciais.

3.2.4 Extração de características através de Crescimento de Região e Algoritmos Genéticos

Lin e Wu (1999) usaram uma nova metodologia para extrair características faciais aplicando Algoritmo Genéticos (GA). O algoritmo é composto de dois estágios principais: no primeiro é feita uma avaliação da região facial e no segundo são extraídas as características faciais.

No estágio da avaliação da região facial, a região facial é extraída de uma imagem. Para uma imagem com cabeça e ombros, a região facial de interesse está localizada aproximadamente no centro da imagem e geralmente destaca-se por regiões

que representam alto contraste, como por exemplo, a região dos olhos. Essas regiões podem ser extraídas através de técnicas de detecção de bordas.

No estágio de extração de pontos característicos, as características faciais tais como, os olhos, nariz e boca, são extraídas pelo método de *template matching*. Um ou vários *templates* característicos podem ser definidos de acordo com a característica analisada da face humana. Usando *templates* característicos, todos os pontos candidatos na região da face são avaliados na pré-definição das funções de custo. Aqueles pontos que tem os valores mais adequados são extraídos como pontos característicos.

O custo computacional da extração de característica facial é determinado pelos três seguintes princípios: 1) a estimação da região facial; 2) a complexidade da avaliação das funções custo; 3) a busca dos pontos característicos.

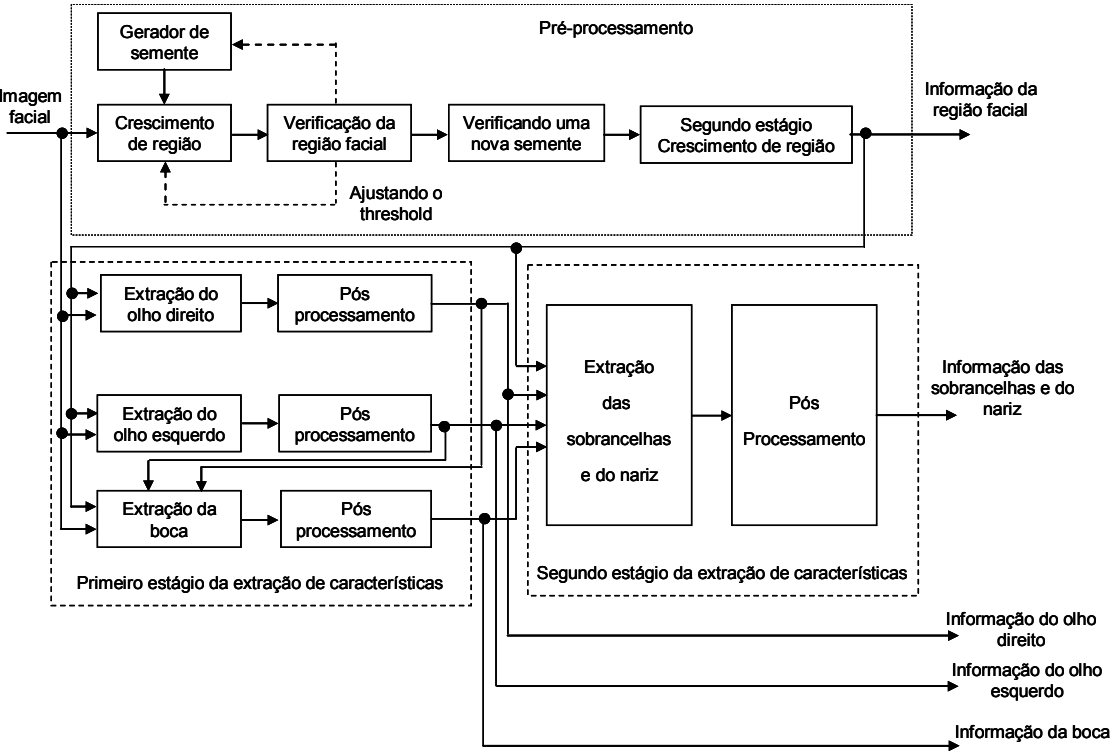


Figura 34: Diagrama em blocos do algoritmo automático de extração de características.

A Figura 34 mostra o diagrama em blocos do algoritmo de extração de característica facial. O sistema é composto de três estágios principais: 1) pré-processamento, 2) primeiro estágio de extração de características e 3) segundo estágio de extração de características. No estágio de pré-processamento, a região da face é

encontrada pela segunda chance através do método de crescimento de região. No primeiro estágio de extração de características, as características principais (olhos e boca) são extraídas da região facial. No segundo estágio de extração de características, são extraídas as características das sobrancelhas e do nariz com base nas informações intrínsecas dos olhos e boca.

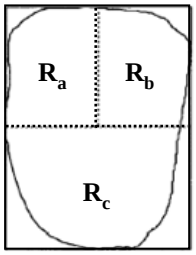


Figura 35: Três sub-regiões selecionadas da região facial para extrair as principais características no primeiro estágio de extração.

Para a extração das características principais, a região facial foi dividida em três sub-regiões, como mostra a Figura 35, onde se espera que na sub-região R_a contenha o ponto característico do olho direito, a sub-região R_b contenha o ponto característico do olho esquerdo e o ponto característico da boca esteja na sub-região R_c.

Na sub-região R_a e R_b, os valores adequados do ponto característico do nariz, estão muito mais baixos que os pontos característicos dos olhos. Contudo, na sub-região R_c, os valores adequados do ponto característico do nariz e da boca são às vezes muito similares. Portanto, é possível que o ponto característico do nariz seja extraído da sub-região R_c ao invés da característica da boca.

Para excluir o ponto característico do nariz de R_c, a sub-região deveria ser modificada após ambos pontos característicos terem sido encontrados. Como mostra a Figura 36, o ponto e_c é o ponto central característico dos dois olhos (e_r e e_l), e R_e é uma região circular com o ponto central e_c e um raio r, onde r é igual à distância entre e_r e e_l (Figura 36).

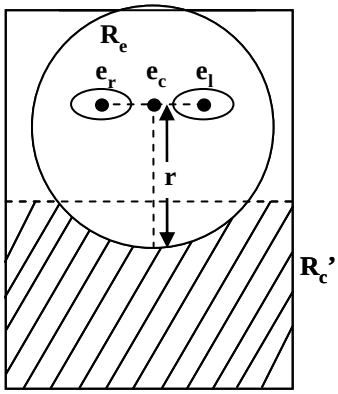


Figura 36: A sub-região R_c é excluída para extrair o ponto característico do nariz.

O ponto característico da boca é então localizado na sub-região $R_{c'}$.

Aplicando o Algoritmo Genético para encontrar os possíveis pontos característicos e encontrar o valor mais adequado em cada sub-região, os três principais pontos característicos da face podem sempre ser extraídos.

Após a extração das características principais, é aplicado o segundo estágio da extração de características, onde os outros pontos característicos são encontrados em três sub-regiões específicas R'_a, R'_b e R'_c, como mostra a Figura 37.

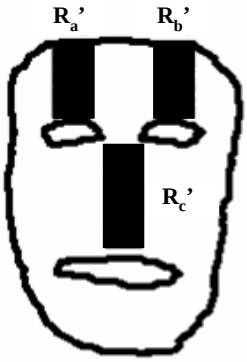


Figura 37: Sub-regiões especificas R'_a, R'_b e R'_c.

Apesar dos pontos característicos terem sido extraídos, é preciso determinar um retângulo com diferentes tamanhos para especificar as áreas de localização dos componentes característicos. O ajuste desta janela é realizado no estágio de pós-processamento, como mostra a Figura 38.

(a) (b)

Figura 38: A imagem com o crescimento de região é usada para estimar as áreas características. (a) a versão binária da imagem e (b) o resultado correspondendo às áreas características.

Com o uso do Algoritmo Genético é possível reduzir a complexidade computacional, encontrando o valor máximo global em S, onde S representa um espaço solução. Uma população P é mantida, a qual consiste de N elementos, onde N é o tamanho da população. Cada elemento em P é conhecido como um cromossomo que é formado por uma lista de genes. A população P envolve uma outra população P' que executa algumas operações genéticas. Os cromossomos que possuem altos valores adaptativos têm maior probabilidade de continuarem na população na próxima geração e de se propagarem. Por outro lado, os cromossomos fracos, são repostos por outros cromossomos fortes. Após um adequado número de gerações, a população adulta tem a expectativa de conter o elemento com o máximo valor global.

Nessa aplicação, o espaço solução S é uma sub-região de uma imagem bi-dimensional (2-D), e cada pixel da imagem na sub-região é um elemento em S.

A Figura 39 mostra o diagrama em blocos implementado no sistema, onde uma população inicial é formada antes da evolução. Na maioria das aplicações baseadas em Algoritmos Genéticos, a população inicial é criada randomicamente selecionando os elementos do espaço de saída. Os cromossomos com altos valores adaptativos na população atual têm maior probabilidade de serem selecionados como sementes da próxima geração. Este esquema probabilístico para selecionar as sementes da nova geração é conhecido como reprodução probabilística.

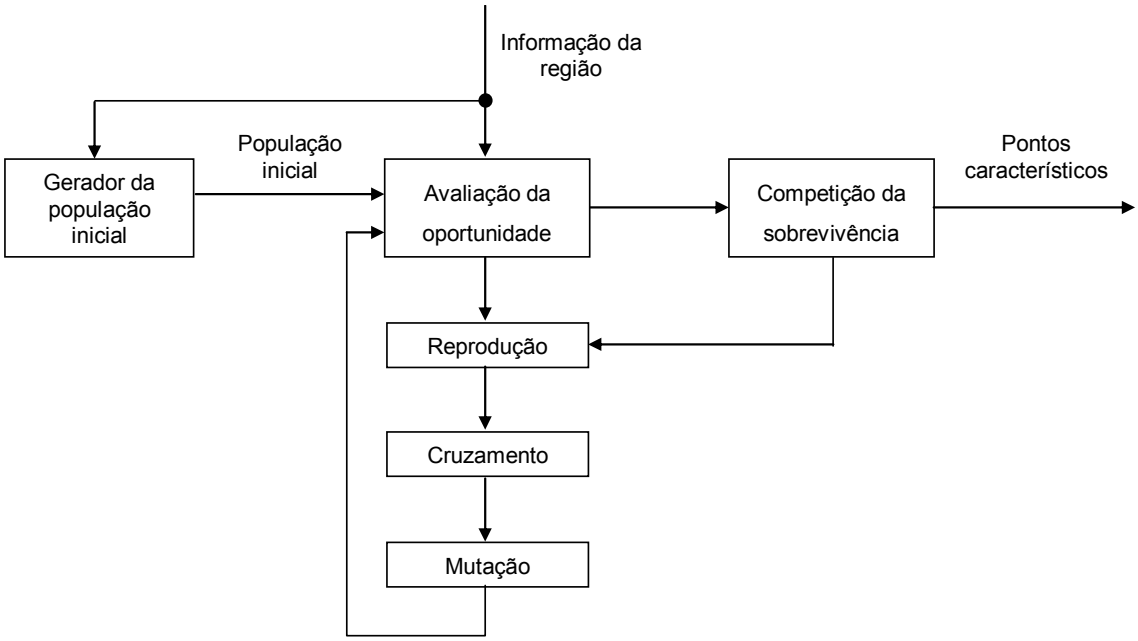


Figura 39: Diagrama em blocos da evolução genética adotada.

A operação de cruzamento é aplicada para explorar novos elementos de um espaço pesquisado. A operação de mutação é aplicada apenas quando são executados os genes mais significantes (MSG) dos cromossomos, porque a evolução aplicada na operação de cruzamento não é apropriada quando se espera a convergência dos cromossomos para o máximo. Após o estágio do cruzamento, cada cromossomo emparceirado seria processado e transferido como um cromossomo candidato da nova geração.

Há N cromossomos emparceirados após a execução da operação genética. Junto com o cromossomo original na geração atual, N cromossomos são selecionados para este cromossomo 2N de acordo com os valores adequados. Cada cromossomo pode ser selecionado apenas uma vez. Os cromossomos com alto valor adaptativo se tornarão membro da população da próxima geração, e poderão caminhar através das próximas iterações da evolução genética. A operação de classificação é necessária no estágio da competição da sobrevivência. Este estágio é adicionado no algoritmo proposto para prevenir que os cromossomos sejam substituídos pelos novos com valores de adaptação menores, porque não há garantia que o novo cromossomo gerado tenha um alto valor adaptativo em GA.

Os resultados da simulação de Lin e Wu (1999), usa um banco de imagens contendo duas seqüências de imagem CIF, Miss América e Clarie, com resolução 360 x

288 *pixels*. Cada imagem dessas duas seqüências é processada aplicando o algoritmo proposto e as características faciais foram extraídas, conforme mostra a Figura 40.

Figura 40: Resultados da simulação a partir da seqüência de Clarie.

A Tabela 9 e Tabela 10 mostra a média proporcional do tempo de execução nos diferentes estágios de processamento.

Tabela 9: Média proporcional (%) do tempo de execução nos diferentes estágios de processamento, quando GA não é aplicado no Segundo Estágio de Extração de Característica.

Estágio	Claire	Miss América	Chou	Tseng
Pré-Processamento	15.2	16.8	14.4	17.4
1º estágio de extração	65.2	50.5	39.6	36.1
2º estágio de extração	19.6	32.7	46.0	46.5

Tabela 10: Média proporcional (%) do tempo de execução nos diferentes estágios de processamento, quando GA é aplicado no Segundo Estágio de Extração de Características.

Estágio	Claire	Miss América	Chou	Tseng
Pré-Processamento	14.9	18.9	18.8	23.1
1º estágio de extração	63.8	56.7	52.0	47.7
2º estágio de extração	21.3	24.4	29.2	29.2

Comparando os resultados das duas tabelas, é possível verificar que na Tabela 10 o custo computacional no segundo estágio de extração de características é reduzido aplicando o Algoritmo Genético.

Quando o algoritmo de extração de características é aplicado em imagens com baixo *bit- rate*, como aplicações de codificação de vídeo, a relação temporal entre imagens vizinhas em uma seqüência de vídeo pode ser explorada. Como o movimento da face humana em uma cena é normalmente lento, a posição das características faciais nas imagens vizinhas são independentes uma das outras. Por meio da informação temporal, novas técnicas podem ser exploradas para tornar o algoritmo de extração de característica até mais eficiente.

3.2.5 Extração de características usando auto-espaço e grafos deformáveis

Em Ahlberg (1999) é apresentado um sistema de codificação baseado em modelo de uma seqüência de imagens, onde a extração de características de faces humana é realizada através de grafos. Segundo Ahlberg (1999), a extração de características faciais pode ser vista como um problema de otimização, buscando os parâmetros ótimos de adaptação para um modelo facial, onde o ótimo é definido como o conjunto de parâmetros descrevendo a face com a mínima distância para um espaço de faces. A aplicação de grafos deformáveis reduz a alta complexidade computacional.

A idéia básica da codificação baseada em modelo de seqüências de vídeo é ilustrada na Figura 41. A imagem da câmera é inicialmente analisada, usando técnicas de visão computacional, e a localização de objetos de interesse, por exemplo, a face humana, é identificada. Um modelo de estrutura de *wireframe* é associado a um objeto face, descrevendo a sua forma em 3-D.

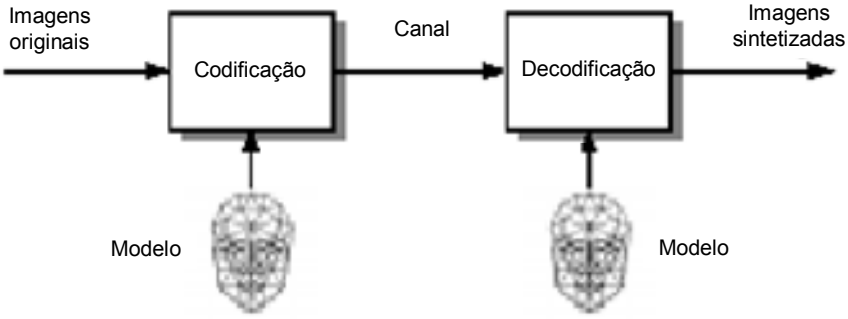


Figura 41: Sistema de Codificação Baseado em Modelo.

Segundo Ahlberg (1999), o problema de encontrar o conjunto de parâmetros que definem os parâmetros da forma do modelo *wireframe*, é a alta complexidade computacional envolvida. Existem diferentes caminhos para reduzir a complexidade computacional:

- Reduzir o espaço encontrado.
- Usar um método de busca heurística.
- Reduzir o tempo do cálculo da medida de discrepância.

Inspirado em Wiberg (1996), um grafo deformável foi implementado. Um simples grafo com seis nós é mostrado na Figura 42 e é definido da seguinte maneira:

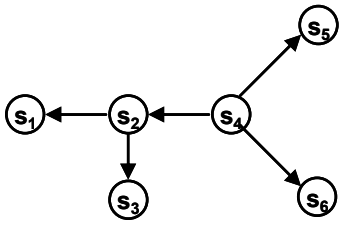


Figura 42: Um simples grafo com arcos direcionados.

- Um **grafo** é um conjunto G=[S,A] que consiste de um conjunto de regiões S = $\{s_1, s_2, ..., s_n\}$ e um conjunto de arcos direcionados $A = \{a_1, a_2, ..., a_m\}$.

- Uma **região** é um conjunto s_i=[N,m], onde N é um vetor 2-D de nós, e m é um atrator ou uma função adequada m.
- Um **nó** n_i=[x,y,v] correspondendo a uma posição na imagem associada e tendo um valor v dado pela região atrativa.
- Um **atrator** m = m(x,y) é uma função que indica a posição das regiões (x,y).
- Um **arco** é um conjunto e_i=[sj, sk, x, y, ξ, u], onde s_j e s_k são duas regiões conectadas por um arco, x e y são as distâncias por definição entre s_j e s_k na imagem, e ξ e u são as derivadas máximas para estas distâncias.
- Um **grafo avaliador** é um grafo onde os atratores têm sido computados em todas as posições e o conjunto de valores dos nós consequentemente.

O grafo é primeiramente avaliado, ou seja, os atratores são computados. Então, a ótima deformação é calculada usando programação dinâmica, conhecida como algoritmo de Viterbi (WIBERG, 1996). O algoritmo de Viterbi na versão básica é aplicado numa seqüência de regiões, mas pode ser aplicado para um grafo de círculo livre também. A ordem é então indicada pela relação pai-filho; uma região pai não pode ser processada até que todos os seus filhos tenham sido processados; a região principal (a região S_4 na Figura 42) será processada conseqüentemente por último.

O grafo ilustrado na Figura 43 foi projetado para extração das características faciais. As regiões com atratores de padrões estatísticos para o *matching* com os componentes principais da face, define a função para o atrator como sendo a distância para o espaço característico (DFFS), ou seja, a distância Euclidiana ao quadrado no espaço das componentes principais para a área facial específica (auto-bocas, auto-olhos, etc). Juntas, estas áreas cobrem importantes áreas da face. Para as regiões dos olhos, o detector de íris descrito em Daugman (1993) é usado, visto que demonstrou muita eficiência. Algumas regiões são atratores para as bordas com o intuito de encontrar os contornos da face. Detecções de bordas horizontal e vertical são executadas (simplesmente usando o filtro de Sobel), e sobrecarregando ao mesmo tempo cada região conforme a direção da borda. Há então algumas poucas regiões sem atratores, usado para sustentar outras regiões e conectar as regiões sem atratores.

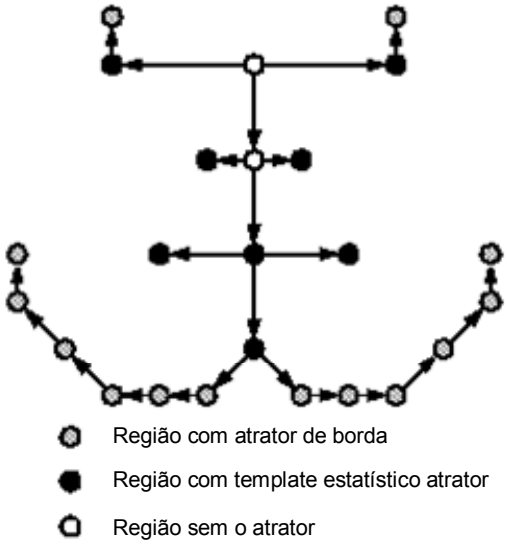


Figura 43: Grafo deformável para extração de característica facial.

Para testar o sistema, Ahlberg (1999) usou o banco de imagens M2VTS (M2VTS) bem como seqüências de imagens clássicas "Claire" e "Miss America" são usadas. Após a localização das faces, é aplicada a deformação do grafo sobre a imagem para localizar os pontos característicos faciais, conforme mostra a Figura 44.

Figura 44: Resultados experimentais do banco de imagens "Clarie", "Miss América" e quatro imagens do bando de imagens M2VTS.

3.3 Considerações finais

O objetivo deste capítulo foi não só de apresentar alguns dos algoritmos de extração de características faciais, como também de fornecer uma contribuição no que refere ao estado da arte, para novas pesquisas nessa área.

Dentre as técnicas apresentadas, a que se destacou pela simplicidade, no que se refere à estrutura de implementação, baixo custo computacional, e bons resultados na extração das características faciais foi a técnica de projeção horizontal e vertical (Seção 3.2.1).

4 Redes Neurais Artificiais

4.1 Considerações iniciais

Neste capítulo, é realizada uma revisão dos conceitos sobre redes neurais artificiais, detalhando a arquitetura de rede neural *perceptron* multicamadas, que será usada para o reconhecimento de faces, e o algoritmo de treinamento *backpropagation*.

4.2 Introdução

O cérebro humano é considerado o mais fascinante processador existente, sendo composto por aproximadamente 10 bilhões de neurônios. Todas as funções e movimentos do organismo humano estão relacionados ao funcionamento dessas pequenas células. Os neurônios são conectados uns aos outros através de sinapses, e juntos formam uma grande rede, chamada rede neural.

Pode-se considerar uma rede neural como um processador maciçamente paralelo distribuído, formado de unidades de processamento simples, que tem a função de armazenar conhecimento experimental e torná-los disponíveis para uso (HAYKIN, 2001). As redes neurais possuem uma boa velocidade de processamento, devido ao seu paralelismo interno, possibilitando o desenvolvimento de operações em tempo real.

As pesquisas sobre redes neurais artificiais iniciaram em 1943, quando Warrerm McCulloch e Walter Pitts estabeleceram as bases da neurocomputação, concebendo procedimentos matemáticos análogos ao funcionamento dos neurônios biológicos. Esta

contribuição foi puramente conceitual já que os autores não sugeriram aplicações práticas para o seu trabalho, mesmo porque os sistemas propostos por eles não tinham capacidade de aprender.

Em 1949, Donald Hebb deu um passo muito importante na história das redes neurais. Ele propôs um modelo capaz de proporcionar capacidade de aprendizado às redes neurais artificiais. Sua proposta, traduzida matematicamente, viabilizou o desenvolvimento de redes neurais eficazes.

Nos anos 50 já existiam várias implementações de redes neurais através de circuitos analógicos e, naquela época, acreditou-se que o caminho para o entendimento da inteligência humana havia sido descoberto (NIELSEN, 1991). Para reproduzir o comportamento do cérebro humano, pensava-se que bastaria construir uma rede neural suficientemente grande. Uma rigorosa análise matemática, no entanto, comprovou o pouco poder computacional dos modelos de rede neural utilizados na época, levando as pesquisas neste campo a ficarem esquecidas de meados dos anos 60 até o início da década de 80.

Nos anos 80, o interesse pela área retornou devido, em grande parte, ao surgimento de novos modelos neurais, como o proposto por Hopfield (HOPFIELD, 1982) e Kohonen (KOHONEN, 1982) e de algoritmos de aprendizado mais poderosos. Além disso, nesta mesma época, ocorreu o surgimento de computadores mais rápidos e poderosos, facilitando a implementação das redes neurais artificiais.

As redes neurais artificiais consistem num modelo computacional para solucionar problemas de ciências e engenharia, inspirado no funcionamento do cérebro humano, inclusive no seu comportamento, isto é, adquirindo conhecimento através de experiências, errando e adaptando-se às novas condições de cada problema.

Existem várias arquitetura de redes neurais e algoritmos de aprendizagem, modo como é conhecido o procedimento capaz de “ensiná-las” (CELLI, 1999). Atualmente, as redes neurais são aplicadas em sistema para a identificação de suspeitos, verificação de cartão de crédito, reconhecimento de faces, segurança, entre outros.

A arquitetura de uma rede neural é formada basicamente por uma ou mais entradas na primeira camada, uma ou mais camadas ocultas e uma camada de saída. A arquitetura da rede neural está relacionada com o algoritmo de aprendizagem usado para o seu treinamento (CASTRO, 2003). As principais arquiteturas são:

- **redes *feedforward*** (camada única): possui uma camada de entrada e uma única camada de neurônios que é a camada de saída (Figura 45).

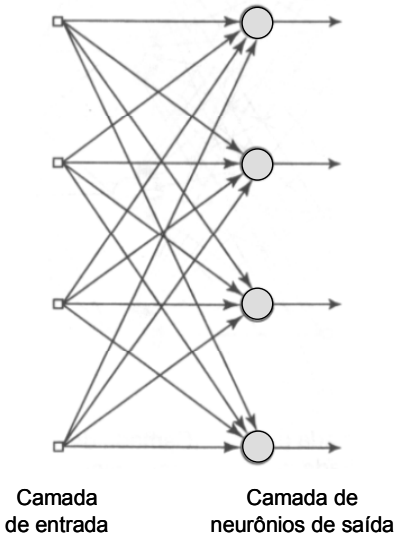


Figura 45: Rede *feedforward*. Com uma única camada de neurônios.

- **redes *feedforward*** (multicamadas): é similar a anterior, exceto pelo fato que possui uma ou mais camadas ocultas de neurônios, como ilustra a Figura 46.

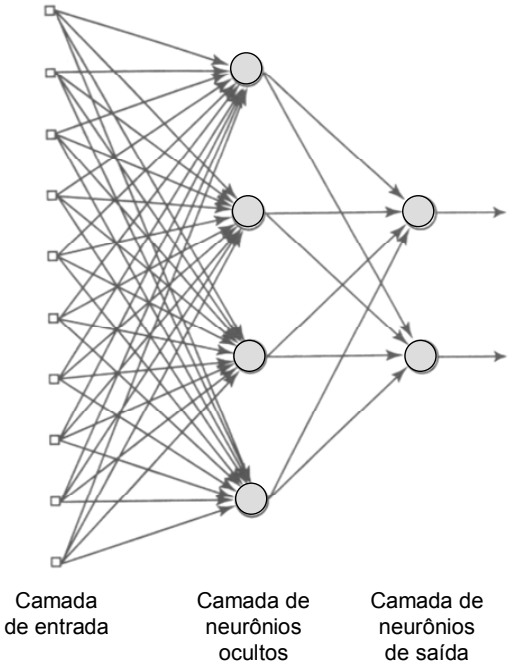


Figura 46: Rede *feedforward*. Totalmente conectada com uma camada oculta e uma camada de saída.

- **redes recorrentes**: são redes que contém retroalimentação entre neurônios de camadas diferentes (Figura 47).

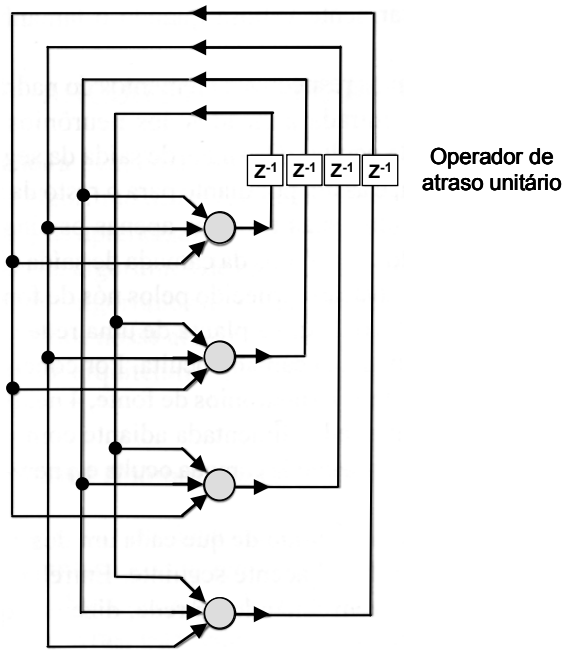


Figura 47: Rede recorrente com retroalimentação entre neurônios.

- **estrutura reticulada** (Self Organize Map (SOM)): é uma rede *feedforward* cujos neurônios são arranjados em linhas e colunas. Os sinais de entrada são os mesmos para todos os neurônios (Figura 48).

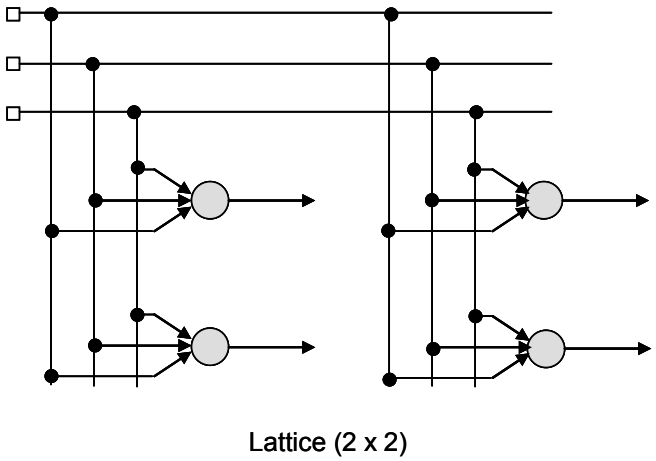


Figura 48: Estrutura reticulada.

Devido às características próprias das redes neurais, estas podem realizar alguns tipos de tarefas que não são executadas de forma satisfatória em sistemas computacionais tradicionais, mas que para o ser humano são tarefas triviais. Elas possuem a característica de se potencializar nas seguintes aplicações:

- **Reconhecimento de padrões:** atribuir um padrão de entrada a uma das várias classes predefinidas. Exemplo: reconhecimento de imagens, de voz e de escrita, etc.;
- ***Clustering* e Categorização:** explorar semelhança entre padrões e agrupar padrões parecidos. Exemplo: compressão e garimpagem de dados, identificação de semelhanças, etc.;
- **Aproximação de funções:** encontrar uma estimativa y de uma função desconhecida f. Exemplo: problemas de modelagem científica e de engenharia;
- **Previsão e Estimação:** dado um conjunto de exemplos $\{y(t_1), y(t_2), \ldots y(t_n)\}$, prever o valor $y(t_{n+1})$ no instante de tempo t_{n+1}. Exemplo: previsão de tempo, séries temporais, mercado financeiro, etc.;
- **Otimização:** otimizar e maximizar uma função sujeita ou não a restrições. Exemplo: programação linear, programação não-linear, otimização combinatorial (telefonia para tráfego óptico), programação dinâmica,
- **Memórias associativas:** recuperar item correto mesmo que a entrada seja parcial ou distorcida. Exemplo: processamento de imagens, de gráficos, caracteres, etc.;
- **Controle:** gerar entrada de controle para que o sistema siga a trajetória especificada pela referência. Exemplo: controle de processos, de robôs, etc.

Nas seções seguintes, serão apresentados o modelo de um neurônio, o *perceptron* de camada única, o *perceptron* multicamadas (MLP) e os algoritmos de treinamento.

4.2.1 Modelo de um neurônio

Um neurônio é uma unidade de processamento fundamental para a rede neural. A Figura 49 mostra o modelo de um neurônio, onde é possível identificar três elementos básicos do modelo neural (HAYKIN, 2001).

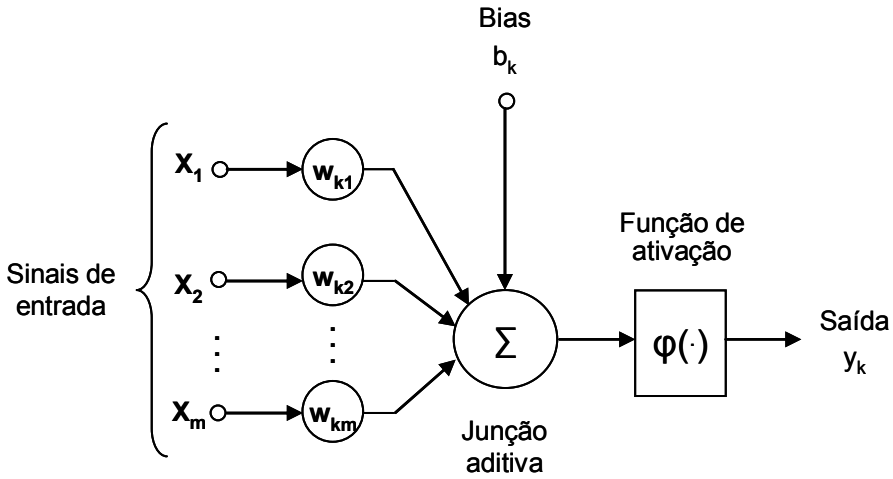


Figura 49: Modelo não-linear de um neurônio.

O neurônio é formado por elos de conexão ou sinapses, cada uma caracterizada por um peso sináptico w_{kj}. Os **sinais de entrada** x_j apresentados à rede, são multiplicados pelo peso sináptico w_{kj}, onde k representa o índice do neurônio em questão e j se refere ao terminal de entrada da sinapse. O **somador** é responsável por efetuar a soma dos sinais de entrada, ponderados pelas respectivas *sinapses* do neurônio e é também conhecido como combinador linear. A **função de ativação** tem por objetivo restringir o valor de saída de um neurônio, isto é, processa o conjunto de entradas recebidas e o transforma em estado de ativação. Normalmente, o estado de ativação dos neurônios pode assumir os seguintes valores: binários (0 e 1), bipolares (- 1 e 1) e reais ($-1 \leq \varphi(.) \leq 1$ ou $0 \leq \varphi(.) \leq 1$). O *bias* é representado por b_k e é um parâmetro externo do neurônio artificial *k*, pois permite uma melhor adaptação, por parte da rede neural, ao conhecimento a ela fornecido.

Matematicamente, um neurônio é representado conforme mostrado na Equação (4.1) e Equação (4.2):

$$u_k = \sum_{j=1}^{m} w_{kj} x_j \tag{4.1}$$

$$y_k = \varphi(u_k + b_k), \tag{4.2}$$

onde, u_k é a saída do combinador linear e y_k é o sinal de saída do neurônio.

4.3 Perceptron de Camada Única

Em 1958, Frank Rosenblatt propôs o perceptron como o primeiro modelo para aprendizagem supervisionada. Esta rede é composta por uma camada de entrada cuja saída alimenta os neurônios da camada de saída. Geralmente, os sinais de entrada se propagam para os neurônios da camada escondida, cujas saídas alimentam os neurônios da camada de saída. A propagação do sinal neste tipo de rede é puramente positiva ou *feedforward*, isto é, os sinais são propagados da entrada da rede para suas saídas. Esta arquitetura é ilustrada na Figura 50 (a) e o sentido da propagação do sinal é mostrado na Figura 50 (b).

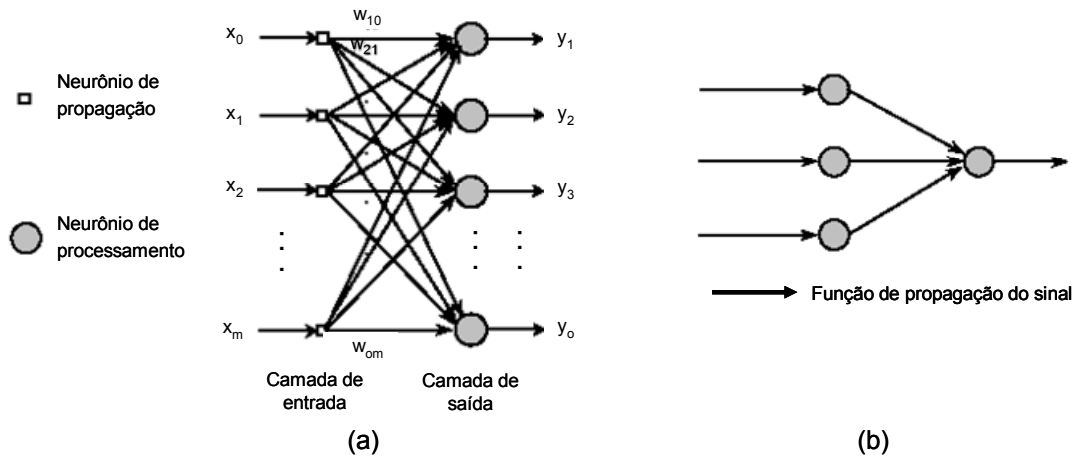


Figura 50: Rede neural de camada única. (a) Arquitetura da rede; (b) Direção da propagação da função sinal.

Para modelar da forma mais simples o funcionamento de um neurônio, são necessárias as seguintes características:

- Algumas entradas, com as respectivas constantes de multiplicação (pesos) para simular as sinapses;
- A saída da célula, que pode ser ativa ou não, dependendo se a soma das entradas atingir ou não um limiar.

O *perceptron* funciona matematicamente da seguinte maneira: para cada neurônio, os sinais de entrada são multiplicados pelos pesos, gerando um valor de saída do combinador linear (Equação 4.1). Para que os neurônios sejam ativados na saída da rede (Equação 4.2), é usada uma função de ativação "Heaviside" ou "Limiar", após a soma das entradas (Figura 51). Esta função de ativação ativa o neurônio no intervalo entre [0, 1], onde:

$$\varphi(v) = \begin{cases} 1 & se\ v \geq 0 \\ 0 & se\ v < 0 \end{cases} \tag{4.3}$$

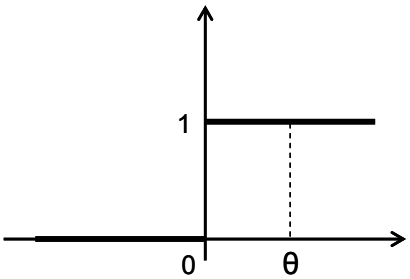


Figura 51: Função de *Heaviside*.

Posteriormente, foi proposta a ADALINE (*Adaptive Linear Neuron*) por B. Widrow e M. Hoff em 1960. Apesar de ela ser muito similar ao *perceptron*, esta rede usa uma função de ativação linear, ao invés de um *threshold*, sendo capaz de resolver problemas linearmente separáveis.

4.3.1 Algoritmo LMS (regra delta)

Em 1960, B. Widrow e M. Hoff criaram a regra de aprendizagem, também conhecida como algoritmo do erro quadrático médio (MSE) ou regra delta. O objetivo do algoritmo de aprendizagem é minimizar o erro quadrático médio (MSE) entre a saída da rede e a saída desejada.

O algoritmo LMS ou regra delta permite que a rede execute uma aprendizagem contínua mesmo após um dado padrão de entrada ter sido treinado. Uma vantagem do algoritmo LMS sobre a regra de aprendizado do perceptron é que torna a rede mais robusta a ruídos. Por este motivo, têm surgido várias aplicações para este algoritmo na área de processamento de sinais.

A regra delta é um outro exemplo de aprendizado supervisionado no qual a regra de aprendizagem é determinada por um conjunto de entradas e saídas desejadas $\{x_1, d_1\}$, $\{x_2, d_2\}$, ..., $\{x_n, d_n\}$, onde x_j é o vetor do padrão j e d_j corresponde ao alvo ou vetor de saída desejada. Esta regra ajusta os pesos da rede neural e o bias para minimizar a diferença (erro) entre a saída da rede e resposta desejada após o treinamento de todos os padrões. Isto é executado para reduzir o erro de cada padrão, um a um, através da correção dos pesos após um determinado número de padrões de treinamento. O sinal de erro é representado por e(n), como mostra a Equação 4.4:

$$e(n) = d(n) - x(n).w(n) \tag{4.4}$$

Após o cálculo do erro, são aplicadas as equações (Equação 4.5 e 4.6) de atualização dos pesos:

$$\Delta w(n) = \eta.x(n).e(n) \tag{4.5}$$

$$w(n+1) = w(n) + \Delta w(n)\,, \tag{4.6}$$

Na Equação 4.5, o η é o parâmetro da taxa de aprendizagem. Este parâmetro é responsável pela velocidade de treinamento da rede e determina o desempenho da aprendizagem por correção de erro. A correta escolha do parâmetro da taxa de aprendizagem deve ser feito com cautela, para assegurar que seja alcançada a estabilidade ou convergência do processo de aprendizagem (HAYKIN, 2001).

Uma maneira de analisar o comportamento da convergência do algoritmo LMS para o ADALINE é traçar a curva de aprendizagem. A curva de aprendizagem é um gráfico do valor médio quadrático do erro de estimação, em função do número de iterações n, conforme ilustrado na Figura 52:

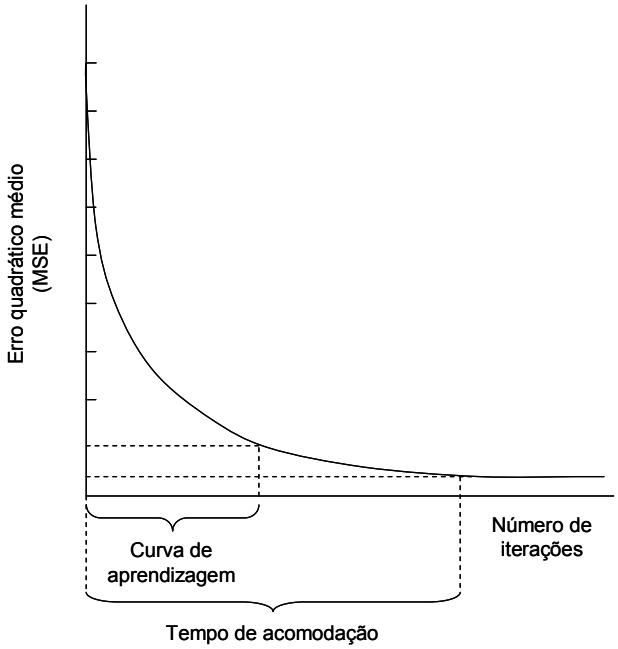


Figura 52: Curva de aprendizagem idealizada do algoritmo LMS.

Uma outra característica muito importante no algoritmo LMS é o tempo de acomodação, que é o tempo para se atingir uma condição estável (a convergência do algoritmo LMS). Quanto menor for o parâmetro da taxa de aprendizagem η, menor serão as variações dos pesos sinápticos da rede, de uma iteração para a outra, e mais suave será a trajetória no espaço de pesos. Porém, desta maneira a taxa de aprendizagem é lenta. Por outro lado, se fizermos o parâmetro da taxa de aprendizagem η muito grande, para acelerar a taxa de aprendizagem, as grandes modificações nos pesos sinápticos resultantes podem tornar a rede instável (oscilatória). Um método simples de aumentar a taxa de aprendizagem evitando, no entanto, o perigo de instabilidade, é

modificar a regra delta, incluindo um termo de momento, como mostrado por Rumelhart et al. (1986).

4.4 Perceptron multicamadas (MLP)

O *perceptron* multicamadas é uma rede do tipo *feedforward*, contendo uma ou mais camadas internas ou escondidas. As principais aplicações desta rede incluem: reconhecimento de padrões, aproximador universal de funções, processamento de imagens e de sinais e previsão de séries temporais. O processo de treinamento destas redes é realizado através do algoritmo *backpropagation*, cujo algoritmo será apresentado na Seção 4.4.1.

O MLP aprende a transformar os dados de entrada na resposta desejada, sendo amplamente utilizado na classificação de padrões. Composto por uma, duas ou várias camadas escondidas (*hidden layers*), a MLP pode implementar, qualquer mapeamento de entrada e saída, como ilustrado na Figura 53 (a):

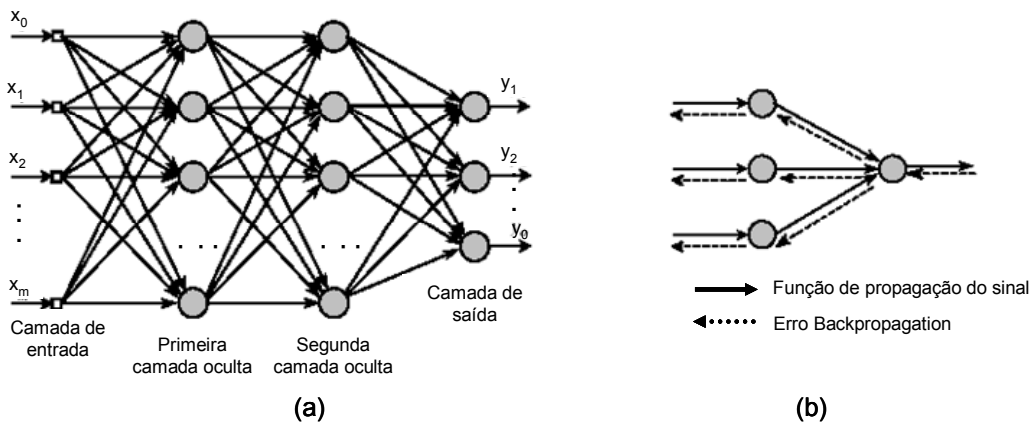


Figura 53: Rede neural multicamadas. (a) Arquitetura da rede. (b) Direção da propagação da função sinal.

A Figura 53 (a) mostra um grafo arquitetural de uma rede *perceptron* multicamadas com duas camadas ocultas e uma camada de saída, onde o vetor de entrada é representado pelas variáveis x_i, os pesos sinápticos são denotados pela letra w_k, onde k indica a camada, e a saída da rede é representada por y_j.

A Figura 53 (b) ilustra as direções de dois fluxos de sinal básico em um *perceptron* multicamadas. Dois tipos de sinais são identificados nesta rede:

1. **Sinais Funcionais**. São estímulos recebidos pelos sinais de entrada (estímulo), que se propagam para frente (neurônio por neurônio) através da rede e manifesta-se no terminal de saída como um sinal de saída;
2. **Sinais de Erro**. O neurônio de saída gera um sinal de erro e se propaga para trás (camada por camada) através da rede.

As camadas ocultas são referidas desta maneira, porque não fazem parte da saída e nem da entrada da rede, sendo a primeira camada oculta alimentada pela camada de entrada, e constituída por várias unidades sensoriais; as saídas da primeira camada são aplicadas às entradas da próxima camada oculta e assim por diante para o resto da rede (Figura 53 (a)).

Essa rede apresenta três características principais:

1. O modelo de cada neurônio da rede inclui uma função de ativação não-linear. Uma forma normalmente utilizada de não-linearidade que satisfaz esta exigência é uma não-linearidade sigmóide definida pela função logística:

$$y_j = \frac{1}{1+\exp(v_j)}, \tag{4.7}$$

onde v_j é a soma ponderada de todas as entradas sinápticas do neurônio j, y_j é a saída do neurônio. A Figura 54 esboça o gráfico da função sigmóide, onde essa função é sempre positiva e limitada superiormente por 1 e inferiormente por 0.

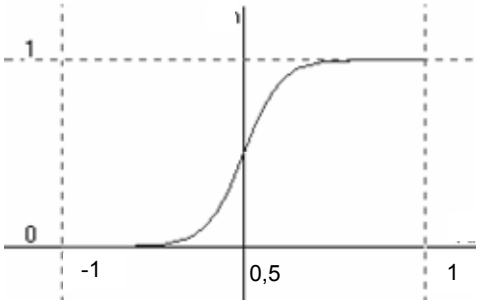


Figura 54: Gráfico da função sigmóide.

2. A rede é composta de uma ou mais camadas de neurônios ocultos, que não são parte da entrada ou da saída da rede. Estes neurônios capacitam a rede a aprender tarefas complexas extraindo progressivamente as características mais significativas dos padrões (vetores) de entrada.
3. A rede possui um alto grau de conectividade, devido às sinapses da rede. Uma modificação na conectividade da rede requer uma mudança na população das conexões sinápticas ou de seus pesos.

4.4.1 Algoritmo de treinamento *backpropagation*

A MLP utiliza para o seu treinamento o algoritmo de retropropagação ou *backpropagation*. Este algoritmo é constituído por dois passos principais: o passo para frente (*forward*) e o passo trás (*backward*). No passo *forward* é aplicado à rede um padrão de dados (vetor de entrada) às entradas da rede e este efeito se propaga através da rede, camada por camada. Durante o processo de propagação, todos os pesos sinápticos da rede não se alteram. No passo *backward* os pesos sinápticos são todos ajustados de acordo com a regra de correção de erro. Contudo, a resposta gerada pela rede é subtraída da resposta desejada (alvo) para produzir um sinal de erro. Os pesos sinápticos da rede são atualizados para que a resposta gerada pela rede aproxime-se da resposta desejada (HAYKIN, 2001). Este processo é repetido diversas vezes, até que para todas as saídas e padrões de treinamento, o erro seja menor do que o especificado.

O principal objetivo do processo de treinamento é minimizar o sinal de erro entre a resposta desejada (alvo) e a atual resposta produzida pela rede. Este sinal de erro da saída do neurônio y_i na interação n, é definido por:

$$e_i(n) = d_i(n) - y_i(n)\,, \tag{4.8}$$

onde d_i é a resposta desejada do *i-th* neurônio de saída.

Adicionando todos os erros quadráticos produzidos pelos neurônios de saída da rede com a respectiva n-th iteração, obtêm-se a seguinte função custo ou o erro para ser minimizado:

$$E(\boldsymbol{w}) = \frac{1}{2}\sum_i e_i^2 = \frac{1}{2}\sum_i (d_i - y_i)^2\,, \tag{4.9}$$

A função custo é uma função continuamente diferenciada para cada peso. Onde, um algoritmo do gradiente descendente pode ser usado para ajustar os pesos. Para as conecções w_{ij}, que ligam a camada oculta e a camada de saída, a regra do gradiente descendente é dada por:

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} = \eta \cdot \delta_i \cdot v_j\,, \tag{4.10}$$

onde η é a taxa de aprendizagem do algoritmo de *backpropagation* e δ_i é o gradiente local definido por:

$$\delta_i = \varphi'(c_i) \cdot (d_i - y_i)\,, \tag{4.11}$$

onde c_i é o sinal de entrada dos pesos aplicado para o neurônio y_i e $\varphi(.)$ é a função de ativação do neurônio. As conexões w_{ij} são iterativamente atualizadas aplicando a seguinte equação:

$$w_{ij}^{novo} = w_{ij}^{velho} + \Delta w_{ij} \,. \tag{4.12}$$

Para as conecções w_{jk}, que ligam a camada de entrada e a camada oculta, a função custo é diferenciada com o respectivo w_{jk} usando a regra da cadeia, que é dada por:

$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}} = -\eta \cdot \sum \frac{\partial E}{\partial v_j} \cdot \frac{\partial v_j}{\partial w_{jk}} = \eta \cdot \delta_j \cdot x_k \,. \tag{4.13}$$

Na Equação 4.13, o gradiente local δ_j em relação ao w_{jk} é expresso pela seguinte equação:

$$\delta_j = \varphi'(h_i) \sum w_{ij} \cdot \delta_i \,, \tag{4.14}$$

onde h_i é o peso do sinal de entrada aplicado para o neurônio v_j. Finalmente, todas as conecções w_{jk} são iterativamente atualizadas de acordo com a seguinte equação:

$$w_{jk}^{novo} = w_{jk}^{velho} + \Delta w_{jk} \,. \tag{4.15}$$

Após o processo de treinamento, a rede é capaz de fornecer a relação funcional entre as variáveis de entrada e saída do processo.

Existem algumas técnicas que auxiliam a melhorar o desempenho do treinamento da MLP:

- **Validação cruzada:** metodologia utilizada para testar o poder de generalização de uma MLP. Neste contexto, deve-se seguir os passos adiante:
    - Particionar o conjunto de dados disponível em dois subconjuntos:

        - Conjunto de treinamento: utilizado para treinar a rede (80% a 90% do total);
        - Conjunto de teste: utilizado para avaliar se a rede está generalizando de forma satisfatória (10% a 20% do total).
    - Utilizar o mesmo conjunto de treinamento para ajustar todas as topologias candidatas;
    - Após o treinamento, utilizar o conjunto de teste para avaliar e escolher (validar) o melhor modelo (melhor generalização):
        - Em certos casos, uma determinada topologia pode ter desempenho satisfatório no treinamento, mas durante a fase de validação (aplicação do conjunto teste), a rede produz resultados insatisfatórios (não consegue generalizar);
        - O aumento de neurônios e de camadas não significa que a rede irá generalizar melhor;
        - Para duas topologias que estão generalizando como o mesmo grau de precisão, deve-se optar por aquela com menor número de neurônios.

- **Inserção do termo de momento:** a velocidade do algoritmo *backpropagation* pode ser aumentada (sem perigo de instabilidade) através da inclusão de um termo de momento α , como mostrada na Equação 4.8:

$$\Delta w_{ji}(t+1) = w_{ji}(t) + \alpha(w_{ji}(t) - \Delta w_{ji}(t-1)) + \eta * d_j * y_i \qquad (4.8)$$

- **Normalização dos dados:** as variáveis referentes às entradas da rede devem ser normalizadas para a faixa [0,1] se estiver sendo utilizada a função sigmóide, ou então, para [-1,1] se for utilizada a tangente hiperbólica.

4.5 Algoritmo de Aprendizagem

4.5.1 Introdução

A rede neural se destaca pela sua capacidade de aprender, e isto é uma característica primordial de sua habilidade. As mudanças que ocorrem na estrutura de interconexão ou nos valores dos pesos sinápticos geram alterações no comportamento de ativação da rede e permite que seja realizado um novo aprendizado de um determinado comportamento. Desta forma, pode-se modificar o estado de ativação na saída da rede em resposta a um certo "estímulo" recebido através do padrão de entrada. Contudo, a rede é capaz de estabelecer associações de entrada e saída (estímulo e resposta), a fim de se adaptar a uma situação proposta. No caso de uma rede com aprendizado supervisionado, a rede deve adaptar seus pesos de maneira a passar a responder de acordo com a resposta desejada. O método utilizado para modificar o comportamento de uma rede é conhecido como regra de aprendizado.

A definição de processo de aprendizagem em redes neurais pode ser definida da seguinte maneira:

> Aprendizagem é um processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um processo de estimulação pelo ambiente no qual a rede está inserida. O tipo de aprendizagem é determinado pela maneira pela qual a modificação dos parâmetros ocorre. (HAYKIN, 2001).

Um conjunto de regras bem definidas para a solução de um problema de aprendizagem é denominado um algoritmo de aprendizagem.

Esta definição de aprendizagem implica nos seguintes eventos:

1. A rede neural é **estimulada** por um ambiente;
2. A rede neural **sofre modificações** nos seus parâmetros, como resultado das estimulações;

3. A rede neural **responde de uma maneira nova** ao ambiente, devido às modificações sofridas em sua estrutura.

Com um algoritmo de aprendizagem padrão uma rede neural aprende através de um processamento iterativo do ajuste dos pesos. O tipo de aprendizagem é definido pela forma como os pesos são ajustados. Os dois principais métodos de aprendizagem são: aprendizagem supervisionada e aprendizagem não-supervisionada.

4.5.2 Aprendizagem supervisionada

Essa regra de aprendizado incorpora o conceito de um supervisor ou um professor, que tem o conhecimento sobre o ambiente em que a rede está operando. Este conhecimento é representado por um conjunto de exemplos de entrada-saída ou padrões. Sendo assim, a rede deve ser capaz de medir a diferença entre seu comportamento atual e o comportamento de referência e então corrigir os pesos de maneira a reduzir este erro. O aprendizado supervisionado utiliza conhecimentos empíricos, que são representados por um conjunto de exemplos rotulados, ou seja, exemplos com pares de dados de entrada com a respectiva saída associada $\{x_i,d_i\}$.

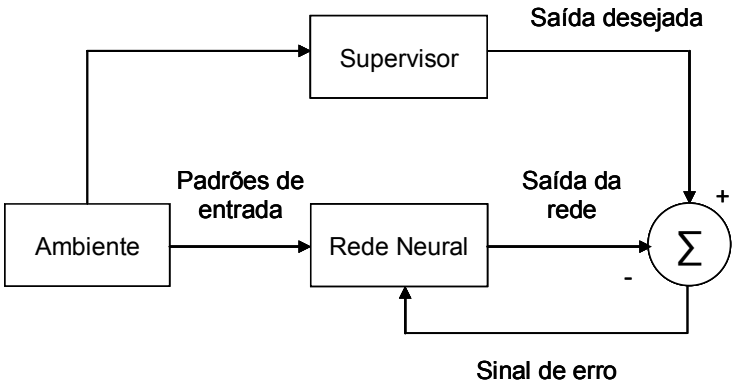


Figura 55: Diagrama em blocos da aprendizagem supervisionada.

A rede ajusta os parâmetros livres através da combinação das entradas e sinais de erro, sendo o sinal de erro a diferença entre a resposta desejada e a resposta real da rede (Figura 55). Este ajuste é realizado passo a passo com o objetivo de fazer a rede

"emular" um professor, assim se a emulação for ótima, o conhecimento do ambiente disponível ao professor é transferido para a rede neural através do treinamento. Quando esta condição é alcançada, pode-se liberar o professor e deixar a rede neural lidar com o ambiente (HAYKIN, 2001).

4.5.3 Aprendizagem não-supervisionada

Ao contrário da aprendizagem supervisionada, como o próprio nome sugere, não há um professor externo para supervisionar o processo de aprendizagem. Isto significa que não há exemplos rotulados da função a ser aprendida pela rede, conforme ilustra a Figura 56. Nessa figura não há informação sobre o sinal de erro produzido pela rede, as classes de dados são desconhecidas ou não-rotuladas e a presença de um supervisor não existe. A rede adapta seus dados de entrada estatisticamente, desenvolvendo uma habilidade para criar representações internas que codificam as características dos dados de entrada e gera novas classes automaticamente (BECKER, 1991).

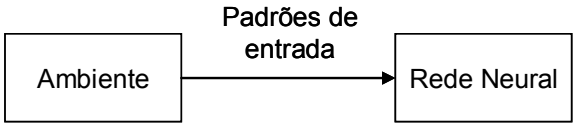


Figura 56: Diagrama em blocos da aprendizagem não-supervisionada.

A rede modifica os pesos em função de critérios internos, por exemplo, a repetição de padrões de ativação em paralelo, de vários neurônios, conhecida como aprendizagem competitiva. O comportamento resultante deste tipo de aprendizado é usualmente comparado com técnicas de análise de dados empregadas na estatística. Exemplo: separar faces humanas, sem, no entanto ter as características com a sua respectiva classe rotulada.

4.6 Considerações finais

As redes neurais artificiais são modelos computacionais que auxiliam em tarefas de classificação e reconhecimento de um determinado conjunto de dados.

Vários trabalhos têm sido desenvolvidos aplicando redes neurais artificiais para o reconhecimento automáticos de faces, onde uma estrutura exata é difícil de ser definida, face à grande variabilidade, provocada por expressões faciais, cor da pele, objetos presentes na face, problemas de iluminação, etc. Visando diminuir esta variabilidade, algumas restrições são impostas às aplicações, procurando torná-las menos complexas.

A rede neural *perceptron* multicamadas é uma das redes mais utilizadas para o reconhecimento de padrões, apresentando bons resultados no reconhecimento facial conforme apresentado no Capítulo 2 (TAMURA et. al., 1996; INTRATOR et al., 1996; LIN e KUNG, 1997; FRANCO e TREVES, 2001; AITKENHEAD e MCDONALD, 2003).

5 Algoritmo para Reconhecimento de Faces Humanas: Metodologia

5.1 Considerações iniciais

Nesse capítulo é apresentada a metodologia utilizada na implementação do algoritmo para reconhecimento de faces, baseado em redes neurais.

5.2 Módulos desenvolvidos no algoritmo

A implementação do algoritmo para reconhecimento de faces foi divida em quatro módulos, conforme ilustra a Figura 57. O primeiro módulo, o de pré-processamento é aplicado sobre as imagens de entrada, com o objetivo de reduzir o ruído, melhorar a qualidade da imagem e padronizar suas dimensões. O segundo módulo, o de segmentação, destaca as características de interesse. As características são então extraídas pelo terceiro módulo e aplicadas a uma rede neural *perceptron* multicamadas (MLP) para o reconhecimento, que constitui o quarto módulo.

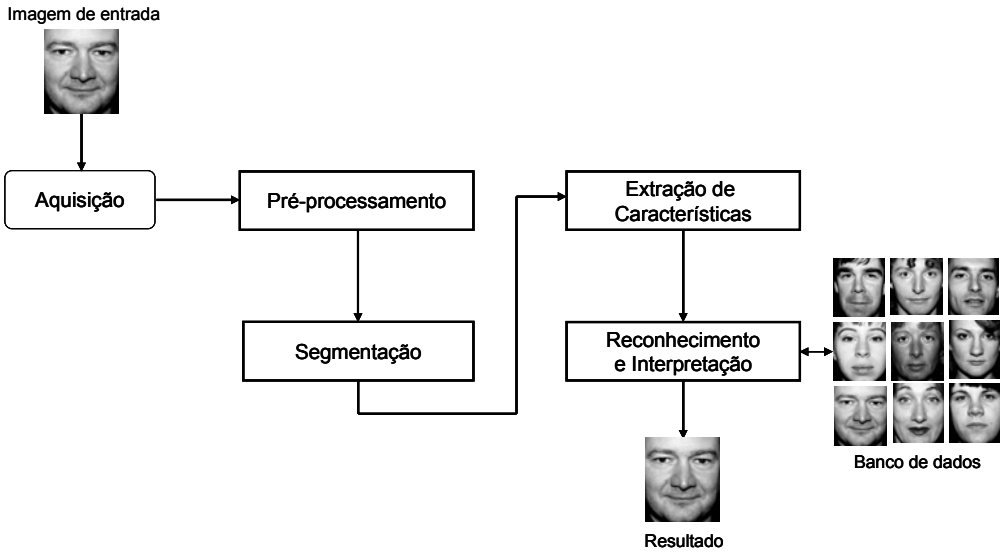


Figura 57: Módulos desenvolvidos no algoritmo para reconhecimento de faces.

5.3 Metodologia

Como imagens de entrada, foram utilizados três bancos de dados diferentes com aproximadamente 200 imagens faciais, obtidos do Departamento de Psicologia da Universidade de *Stirling* (PICS), do ESSEX (ESSEX) e do AT&T (AT&T). Os bancos contêm imagens de pessoas de ambos os sexos, dimensões de aproximadamente 350 *pixels* de altura por 274 *pixels* de largura, em escala de cinza, com 8 *bits/pixel*, porém as imagens foram redimensionadas para 189 x 168 *pixels*, de modo a se adequarem no algoritmo.

Os bancos contêm diferentes tipos de imagens, tais como, frontais, de perfil, com expressões faciais diversas, mas foram utilizadas somente as imagens frontais (Figura 58). Uma das razões é que as imagens frontais são utilizadas pela maioria dos documentos (passaportes e registro geral (RG)). Um detalhe importante a ser mencionado, é que nessas imagens, as pessoas têm bigode, barba ou óculos.

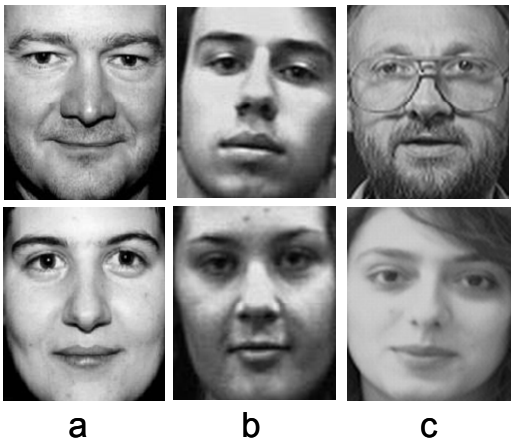


Figura 58: Exemplo dos bancos de dados: (a) Banco de dados da Universidade de *Stirling* (PICS); (b) Banco de dados ESSEX (ESSEX); (c) Banco de dados da AT&T (AT&T).

As imagens foram redimensionadas usando o *Face Detector Software Web* (FDSW). Este *free software on-line* permite que o usuário faça um *upload* da imagem para a qual se deseja realizar a detecção da face humana em uma determinada cena. Como resultado, o *software* retorna ao usuário somente a região facial, conforme mostra a Figura 59.

Figura 59: Resultado da detecção de face usando o *Face Detector Software Web* (FDSW).

5.3.1 Pré-processamento

A tarefa de pré-processamento é uma das mais importantes em um sistema de reconhecimento facial, pois a imagem resultante do processo de aquisição pode apresentar diversas imperfeições (*pixels* ruidosos, contraste ou brilho inadequado, traços interrompidos ou indevidamente conectados) que influirão de forma negativa no processo de extração de características. Sua função é de aplicar um processamento inicial sobre a imagem adquirida, para aprimorar sua qualidade.

A etapa do processo de pré-processamento pode ser visualizada no diagrama em blocos da Figura 60.

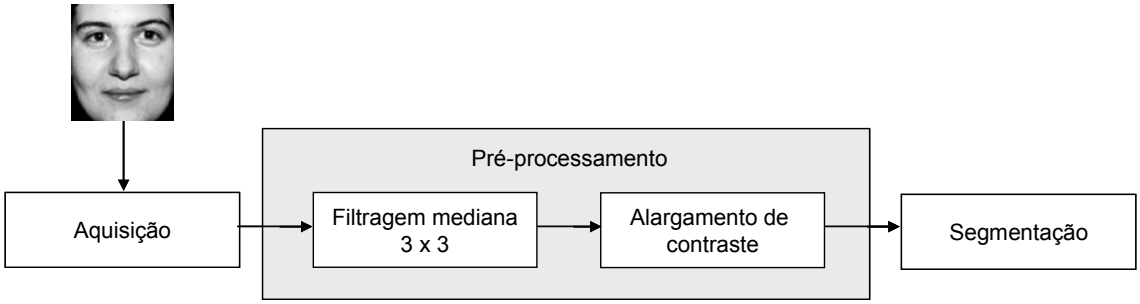


Figura 60: Diagrama em blocos do pré-processamento da imagem.

Conforme mostrado nessa figura, será utilizada a filtragem mediana 3 x 3 para remoção dos *pixels* ruidosos e suavização da imagem (GONZALEZ e WOODS, 1992).

O alargamento de contraste foi utilizado para ajustar os níveis de cinza da imagem. Esta função tem a finalidade de aumentar a escala dinâmica dos níveis de cinza da imagem sendo processada. Essa condição preserva a ordem dos níveis de cinza, impedindo artefatos na intensidade da imagem processada.

5.3.2 Segmentação

A tarefa básica da etapa de segmentação é dividir uma imagem em suas unidades significativas, ou seja, nos objetos de interesse que a compõem. A Figura 61 mostra as etapas da segmentação utilizadas.

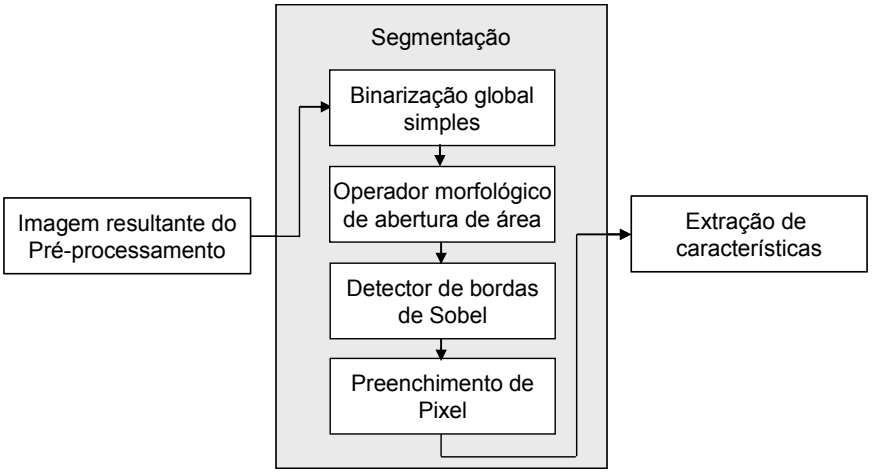


Figura 61: Diagrama em blocos da segmentação da imagem.

Cada etapa será descrita a seguir:

- **Binarização global simples:** é uma técnica, onde um limiar é definido de forma tal que, a imagem de saída tem valor binário "1", para todos os valores de *pixel* da imagem de entrada, acima do limiar e valor binário "0", caso contrário. Após a aplicação desse procedimento, tem-se uma imagem binária. O objetivo da binarização é facilitar o procedimento de extração de características.

- **Operador morfológico de abertura de área:** é responsável por remover um determinado número de agrupamentos de *pixels* que estão interconectados. Esta operação remove, a partir de uma imagem binária, todos componentes conectados que tiverem tamanho menor do que o número de *pixels* determinado, proporcionando uma imagem binária com menos ruídos. Auxilia na remoção de alguns *pixels* que possam dificultar o processo de extração de características.

- **Detector de bordas de Sobel:** este operador é aplicado correlacionando os valores da máscara com os *pixels* da imagem original (GONZALEZ e WOODS, 1992). Sua função é ressaltar as bordas Com esse operador consegue-se um desempenho satisfatório na detecção das bordas da imagem, ressaltando o contorno dos principais elementos da face (olhos e nariz), assim como todo o contorno da face. Os valores da máscara estão especificados na Figura 62:

$$\begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

Figura 62: Máscara 3 x 3 do operador de Sobel.

- **Preenchimento de *Pixel*:** este operador executa um preenchimento dos *pixels* interconectados, começando a partir das posições especificadas.

5.3.3 Métodos de extração de características

A etapa de extração de características no sistema de reconhecimento facial é responsável por extrair características das imagens resultantes do processo de segmentação, através de descritores que permitam caracterizar com precisão cada face, e que apresentem um ótimo poder de caracterização entre as faces parecidas. Estes descritores devem gerar uma estrutura de dados adequada ao algoritmo de reconhecimento. Nesse processo, a entrada ainda é uma imagem, mas a saída é um conjunto de dados correspondente àquela imagem. A técnica de extração de características é fundamental para diminuir a quantidade de dados de entrada para o processo de reconhecimento.

Nesse trabalho foi utilizado o método de assinatura horizontal e vertical para a extração de características, e esse método será descrito na próxima seção.

5.3.3.1 Assinaturas horizontais e verticais para extração de características

O método de assinaturas horizontais e verticais implementado nesse trabalho, baseia-se no artigo de Baskan et al. (2002). Como já mencionado na Seção 3.3, esse é um método de baixo custo computacional, fácil implementação e que consegue localizar os componentes faciais de forma aceitável. A Figura 63 mostra um diagrama em blocos dessa etapa.

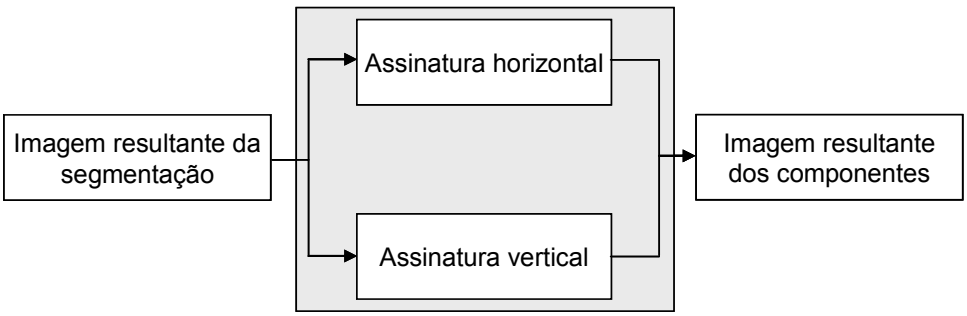


Figura 63: Diagrama do processo de aplicação da assinatura horizontal e vertical.

A imagem facial resultante da segmentação é submetida à etapa de extração de características, onde posteriormente serão localizadas as posições dos componentes faciais. Nesse método, os *pixels* brancos são contados e são criados vetores contendo o número total de *pixels* de cada linha e coluna da imagem. Esses vetores são plotados em um espectrograma para que sejam analisadas as posições onde estão localizados os componentes faciais (GASPAR e PAIVA, 2005).

Após analisar as posições dos componentes faciais, são traçados segmentos de reta coloridos na imagem facial, definindo a posição exata dos componentes faciais e gerando os cruzamentos coloridos para determinar as coordenadas dos componentes faciais.

Com base nesse resultado, identifica-se a posição dos olhos e do nariz, através do cálculo das distâncias euclidianas entre as coordenadas dos componentes faciais (Seção 5.3.3.2), determina-se as características desejadas, conforme apresentado na seção seguinte.

5.3.3.2 Distância Euclidiana

A Distância Euclidiana é um método usado para medir distâncias entre vetores, onde à distância D_e é definida pela Equação 5.1:

$$D_e(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}\ , \tag{5.1}$$

onde *x* e *y* representam respectivamente a linha e a coluna da matriz da imagem.

No módulo de extração de características, a Distância Euclidiana foi usada para calcular a distância entre os componentes faciais (olhos, nariz e centro da face) e extrair as seguintes características faciais ilustradas na Figura 64:

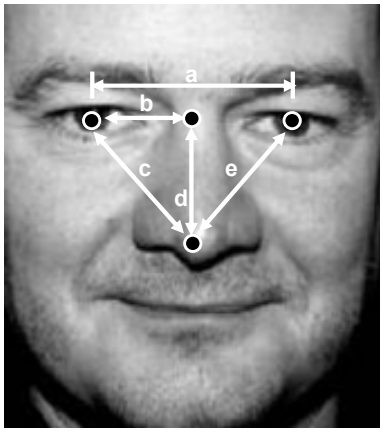


Figura 64: Proporção do cálculo da Distância Euclidiana entre os componentes faciais.

- distância entre olho esquerdo e olho direito (a);
- distância entre olho direito e centro da face (b);
- distância entre o olho direito e nariz (c);
- distância entre o centro da face e nariz (d);
- distância entre o olho esquerdo e nariz (e);

Estas características foram usadas no processo de treinamento da rede neural.

5.3.4 Reconhecimento de faces usando redes neurais

As redes neurais artificiais (RNAs) tem um papel fundamental no processo de classificação e reconhecimento de padrões, ou seja, a rede identifica o padrão de entrada em categorias pré-definidas ou não.

O uso das RNAs em reconhecimento de faces vem sendo aplicado para a solução de muitos problemas como: classificação por sexo, reconhecimento de faces e classificação de expressões faciais. O sucesso de uma aplicação em reconhecimento de faces depende de como a informação é apresentada à rede neural.

A maioria das aplicações em reconhecimento de faces são implementadas usando a rede neural *perceptron* multicamadas (BANARSE et al., 1999; YOON et al., 1998; LAWRENCE et al., 1997).

Outra rede bastante utilizada é o Mapa Auto-Organizável (SOM) ou rede de Kohonen (KOHONEN, 1990), que tem sido aplicada principalmente para aglomeração de dados e mapeamento de características.

Neste trabalho foi implementada a rede neural do tipo *perceptron* multicamadas utilizando o algoritmo de treinamento *backpropagation*. A estratégia utilizada para alcançar a topologia desejada foi a *cross-validation*. A topologia da rede é constituída por:

- uma camada de entrada constituída por 5 entradas, onde são aplicados os padrões com as características faciais;
- três camadas ocultas: a primeira camada oculta é composta por 30 neurônios, a segunda camada oculta por 60 neurônios e a terceira camada oculta por 20 neurônios;
- uma camada de saída, contendo um único neurônio para validação da face.

A camada de saída é responsável por gerar a resposta, identificando ou não a face dentro do conjunto de faces para o qual a rede foi treinada (Figura 65) (AITKENHEAD e MCDONALD, 2003).

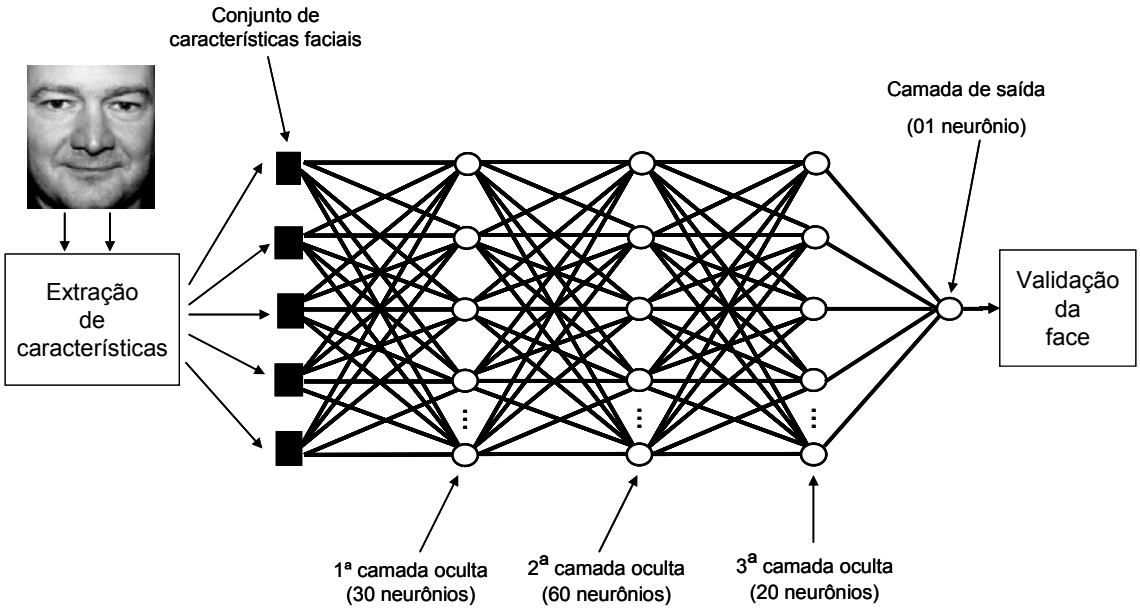


Figura 65: Topologia da rede neural proposta para o reconhecimento de faces.

O tipo de aprendizado utilizado nessa rede é supervisionado, onde o conjunto de características faciais é rotulado, isto é, a cada padrão de entrada será associado o valor "1" que representará a face pertencente ao banco de dados. Como conjunto de testes foram utilizados padrões pertencentes ao banco e padrões não pertencentes ao banco, denominado respectivamente como banco de dados de treinamento e teste, permitindo comprovar a generalização adequada da rede.

5.3.5 Software de desenvolvimento

Para o desenvolvimento do algoritmo, foi utilizado o MATLAB (*Matrix Laboratory*) na versão 6.5, com os *toolbox* de processamento de imagens e redes neurais. O MATLAB é um ambiente de programação que permite o uso da linguagem de programação própria dele, com chamadas de funções de processamento de imagens e de redes neurais, agilizando o desenvolvimento do algoritmo e a análise dos resultados.

6 Resultados e Discussões

6.1 Considerações iniciais

Nesse capítulo são apresentados os resultados obtidos e as discussões.

6.2 Resultados finais

O algoritmo foi avaliado utilizando como entrada, imagens de três bancos de dados distintos: PICS (PICS), ESSEX (ESSEX) e AT&T (AT&T) que possuem resolução de 189 x 168 *pixels*, de 8 *bits/pixel* (APÊNDICE A, B e C), como já mencionado na Seção 5.3.

Os resultados obtidos para essas imagens são apresentados nas próximas seções.

6.2.1 Pré-processamento da imagem

Nessa etapa, é aplicada a filtragem mediana sobre a imagem de entrada. Como já mencionado na Seção 5.3.1, esse filtro é utilizado para a remoção de ruído e suavização da imagem.

O resultado dessa operação pode ser observado na Figura 66.

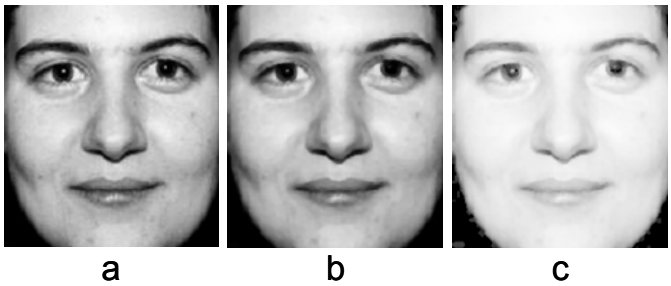


Figura 66: Resultado da aplicação do filtro da mediana e do alargamento de contraste: (a) imagem original; (b) resultado da filtragem mediana; (c) resultado do alargamento de contraste.

O taxa do alargamento de contraste foi definido com valores diferentes para cada banco de dados, conforme especificado a seguir:

- 0.3 para o banco de dados PICS e ESSEX;
- 0.55 para o banco de dados AT&T.

O Apêndice D mostra os resultados da aplicação do pré-processamentos nos bancos de dados.

6.2.2 Segmentação

Nessa etapa, foi aplicada a binarização global simples, seguida da operação morfológica de abertura de área e do detector de bordas de Sobel com máscara 3 x 3, seguido do preenchimento de *pixels*.

O resultado desse processamento é apresentado na Figura 67. No Apêndice E são apresentados os resultados para as demais imagens dos bancos utilizados.

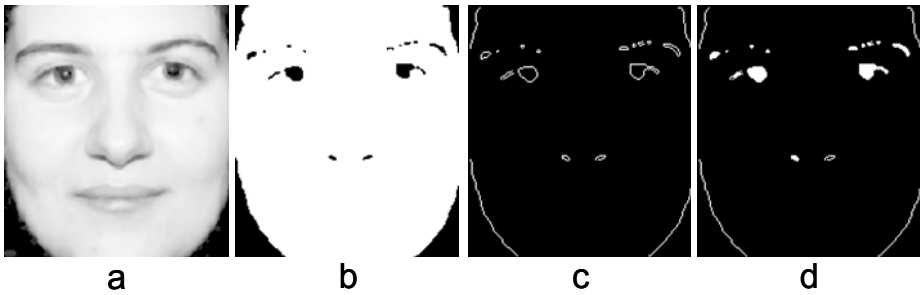


Figura 67: Resultado da segmentação facial: (a) imagem pré-processada; (b) binarização global simples e operação morfológica de abertura de área; (c) detector de bordas de Sobel com a máscara 3 x 3; (d) operação de preenchimento de *pixel*.

A operação morfológica de abertura de área é aplicada após a binarização, para a remoção de todos os agrupamentos menores que 10 *pixels* interconectados.

6.2.3 Assinaturas horizontais e verticais para localização dos componentes faciais

O algoritmo das assinaturas é bem simples e funciona da seguinte maneira: a localização dos olhos e do nariz é identificada através de picos que representam valores das regiões com maior concentração de *pixels* brancos. Essas regiões foram definidas analisando os espectrogramas através dos máximos valores apresentados (Figura 68). Em seguida, são aplicadas as assinaturas horizontais e verticais sobre a imagem binária para localizar os olhos e o nariz. Na região dos olhos, a assinatura horizontal é realizada de maneira simétrica, tendo como ponto inicial o centro da face, onde a varredura é realizada à direita e à esquerda do centro. Esse procedimento é realizado com a finalidade de localizar os olhos em faces com uma inclinação de até 25 graus.

Para analisar a contagem do número de *pixels* brancos em cada linha e em cada coluna, são criados vetores acumuladores de *pixels*. Esses vetores são então analisados pelo algoritmo para a identificação dos pontos de máximo e localização dos componentes faciais. Os componentes faciais estão localizados na região de coincidência dos máximos das assinaturas horizontais e verticais. Na região identificada é traçada um segmento de reta colorido para visualização dos resultados (Figura 69).

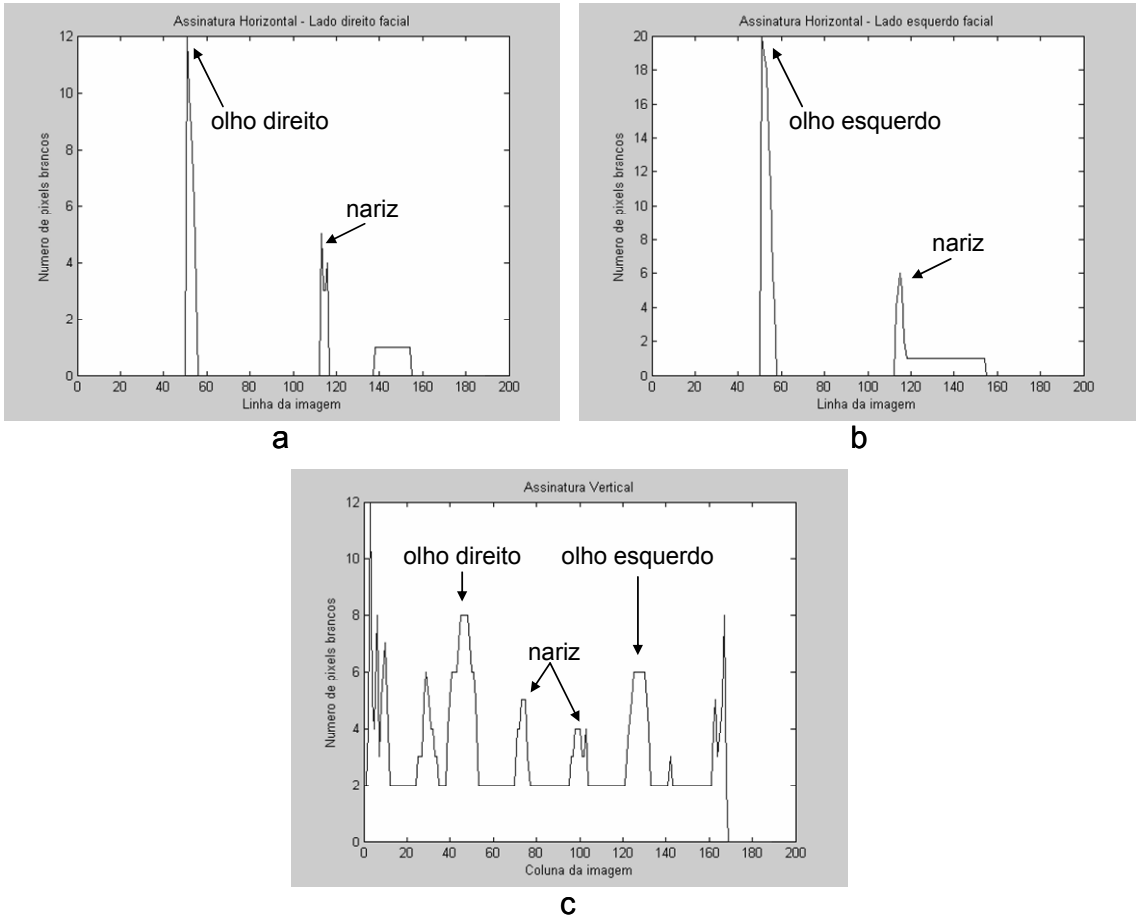


Figura 68: Espectro da contagem no número de *pixels* brancos da imagem: (a) Assinatura horizontal do lado direito facial; (b) Assinatura horizontal do lado esquerdo facial; (c) Assinatura vertical.

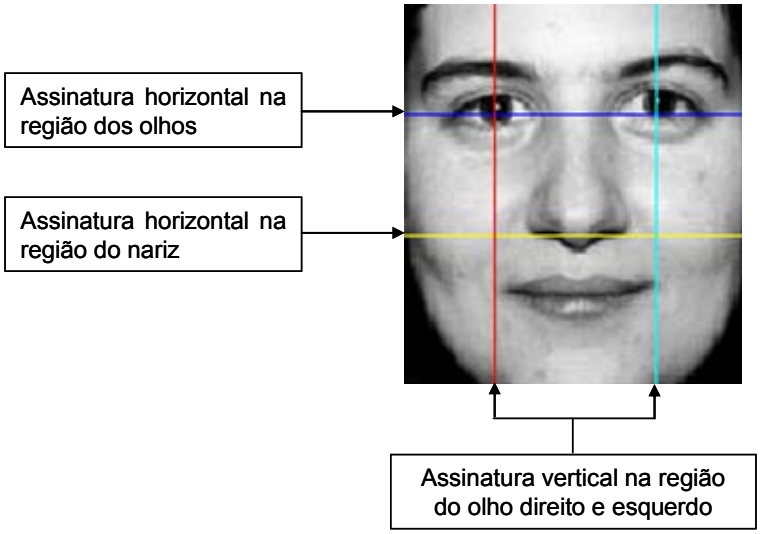


Figura 69: Resultado das assinaturas horizontais e verticais.

A Figura 69 mostra os cruzamentos coloridos sobre os componentes faciais determinados pelas regiões de máximo, das assinaturas horizontais e verticais. Uma vez identificada a região dos olhos, determina-se através da média entre o olho esquerdo (cruzamento azul escuro e vermelho) e o olho direito (cruzamento azul escuro e azul claro) a posição central da face na região dos olhos. A média entre os cruzamentos amarelo e vermelho, e amarelo e azul claro, determina a posição do nariz (Figura 70). Os resultados desse algoritmo são mostrados no Apêndice F.

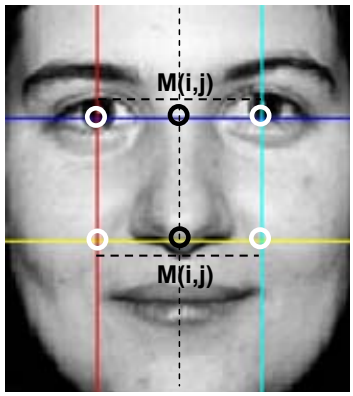


Figura 70: Localização dos pontos característicos faciais (olho direito, olho esquerdo, nariz e o centro da face na região dos olhos).

A Expressão 5.2, determina a média M(i,j) entre os pixels da coluna da posição do olho esquerdo (ColOE(i,j)) e do olho direito (ColOD(i,j)) para localizar o centro da face. Para determinar a posição do nariz, é utilizada a mesma expressão.

$$M(i,j) = \frac{ColOE(i,j) + ColOD(i,j)}{2} \tag{5.2}$$

A partir da localização da posição dos *pixels* de cada componente são extraídas as características utilizando o cálculo da distância euclidiana (Seção 5.3.3.2) entre os componentes faciais, criando assim o vetor de características faciais.

Foram testadas 178 imagens, sendo 61 imagens do banco de dados PICS, 80 imagens do banco de dados ESSEX e 37 imagens do banco de dados AT&T. O

resultado da localização dos componentes faciais (olhos e nariz) é mostrado na tabela abaixo (Tabela 11):

Tabela 11: Resultado da aplicação das assinaturas horizontais e verticais.

COMPONENTES FACIAIS	BANCO DE DADOS (178 IMAGENS)		
	BD PICS	BD ESSEX	BD AT&T
OLHOS E NARIZ	97%	65%	98%

A Tabela 11 apresenta a porcentagem de acerto obtida com a aplicação das assinaturas horizontais e verticais sobre a imagem facial segmentada. A taxa de erro encontrada para as imagens dos três bancos de dados, está relacionada com a inclinação e rotação da face, e também com problemas de iluminação no momento da aquisição. No caso do banco de dados ESSEX, o erro foi maior devido à baixa resolução das imagens e a não padronização da posição facial no momento da aquisição (faces com alto grau de inclinação). As imagens contendo faces com a expressão facial sorrindo, apresentam alta taxa de erro na localização dos olhos e do nariz, que se deve ao problema do alto grau de inclinação e rotação facial, e acrescido da expressão facial. Em algumas dessas, os olhos estavam parcialmente ou totalmente fechados, dificultando a segmentação na região dos olhos. Nessas mesmas imagens, devido ao fato da face estar com a expressão facial sorrindo, o nariz foi confundido com a boca, pelo fato desta estar muito aberta.

Em imagens que apresentam uma inclinação de até 25 graus, do lado direito ou do lado esquerdo, o algoritmo de assinatura horizontal e vertical consegue encontrar os olhos com resultados satisfatórios.

A próxima seção discute os problemas encontrados na etapa de extração de características de forma mais detalhada.

6.2.3.1 Problemas apresentados nas imagens faciais usando assinaturas horizontais e verticais

A seguir são apresentados alguns problemas encontrados no módulo de extração de características:

6.2.3.1.1 Assinatura horizontal e vertical na região dos olhos

A Figura 71 apresenta um problema encontrado na aplicação das assinaturas horizontais e verticais na região dos olhos. Na Figura 71 (a), Figura 71 (d) e Figura 71 (e) é observado que os olhos não foram localizados devido a alguns fatores, tais como: falta de iluminação e padronização em relação ao tamanho das imagens faciais. Observa-se que na Figura 71 (b) foi localizado somente o canto do olho. Isto se deve ao fato dos olhos estarem parcialmente fechados, ou seja, o olho direito foi confundido pela assinatura vertical (linha vermelha) como sendo o supercílio.

Um outro problema encontrado é a rotação dos olhos, como pode ser observado na Figura 71 (c), o olho esquerdo não foi localizado, devido a sua rotação.

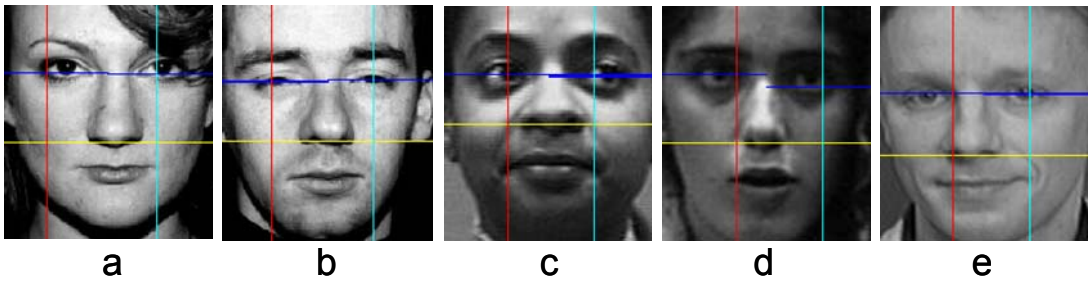


Figura 71: Problemas encontrados usando as assinaturas horizontais e verticais na região dos olhos: (a) e (b) Banco de dados PICS (PICS); (c) e (d) Banco de dados ESSEX (ESSEX); (e) Banco de dados AT&T (AT&T).

6.2.3.1.2 Assinatura horizontal na região dos olhos e do nariz

Analisando a região dos olhos, é possível verificar que em algumas imagens, a sobrancelha está muito próxima da região dos olhos, fazendo com que o algoritmo confunda os olhos com as sobrancelhas.

Nas imagens Figura 72 (a) e Figura 72 (c), verifica-se um problema na localização do nariz, devido ao fato da face não estar bem posicionada no momento da aquisição da imagem.

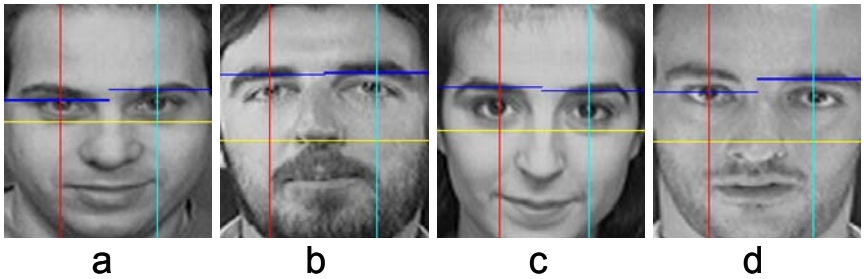


Figura 72: Problemas encontrados para localizar os olhos e o nariz usando assinatura horizontal: (a) Erro na localização dos olhos e do nariz; (b) e (c) Localização das sobrancelhas, ao invés dos olhos; (d) Localização correta do nariz e incorreta dos olhos.

6.2.3.1.3 Assinatura vertical na região dos olhos

A Figura 73 mostra alguns problemas encontrados pela assinatura vertical para localizar a posição dos olhos (linha vermelha e azul claro). O algoritmo não consegue localizar a posição dos olhos, pois naquela região há uma grande quantidade de *pixels* brancos encontrados nas sobrancelhas e nos cantos dos olhos.

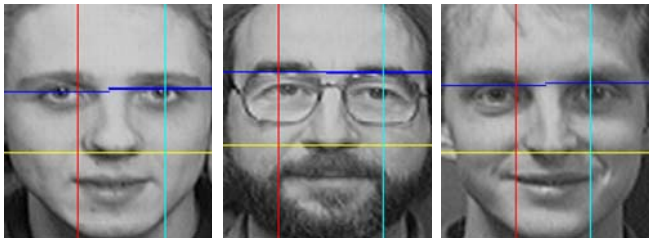


Figura 73: Problema na localização dos olhos usando assinatura vertical.

6.2.3.1.4 Assinatura horizontal e vertical em imagens inclinadas

A Figura 74 apresenta alguns problemas encontrados na localização dos olhos em imagens com as faces inclinadas para o lado direito e esquerdo.

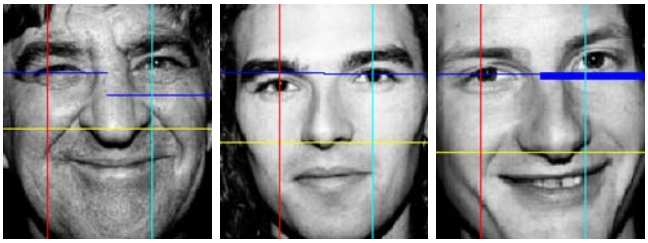


Figura 74: Problema na localização de componentes faciais em imagens inclinadas.

6.2.3.1.5 Assinatura horizontal e vertical em imagens faciais com baixa luminosidade

A Figura 75 mostra o problema das assinaturas horizontais e verticais em imagens com baixa luminosidade, ou seja, pouca iluminação no momento da aquisição da imagem facial. É observado, que o algoritmo é capaz de localizar o nariz com precisão, porém, na região dos olhos, que possui muito pouca iluminação, o algoritmo não é capaz de localizar os olhos.

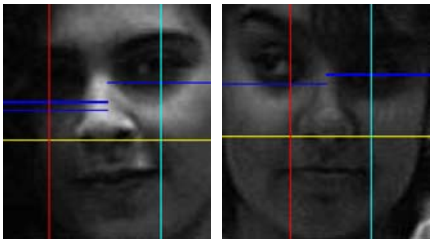


Figura 75: Problema na localização de componentes faciais em imagens de baixa luminosidade.

6.2.3.1.6 Assinatura horizontal na região dos olhos em imagens faciais com óculos

A Figura 76 apresenta um problema de iluminação no momento da aquisição da imagem facial. Observa-se que as imagens que possuem óculos, dependendo de como a imagem é adquirida, a iluminação do ambiente pode atrapalhar, causando um brilho nas lentes, impedindo que o algoritmo seja capaz de localizar os olhos.

Figura 76: Problema na localização dos olhos em imagens faciais usando óculos.

6.2.3.1.7 Assinatura horizontal para localização do nariz

A Figura 77 apresenta dois problemas parecidos. No primeiro, a Figura 77 (a) mostra que devido ao nariz da face ser muito longo, o algoritmo da assinatura horizontal (linha amarela) não consegue localizar o nariz. No segundo, ao contrário do primeiro problema, a face da Figura 77 (b) apresenta um nariz muito pequeno, induzindo o algoritmo a localizar a boca.

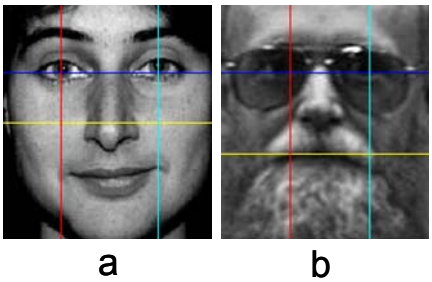


Figura 77: Problema na localização do nariz em imagens faciais.

6.2.4 Reconhecimento de faces usando rede neural MLP

Como já mencionado na Seção 5.3.4, para o reconhecimento das faces, foi utilizada a rede neural *perceptron* multicamadas (MLP), aplicando o algoritmo de treinamento *backpropagation*. A topologia e os parâmetros usados durante a fase de treinamento da rede foram os seguintes:

- **Topologia:** 30 neurônios na primeira camada oculta, seguida por 60 neurônios na segunda camada oculta e 20 neurônios na terceira camada oculta. Foi definido apenas 01 neurônio na camada de saída, definindo se a face pertence ou não ao banco de dados.
- **Matriz de pesos:** os pesos foram reinicializados randomicamente.
- **Função de ativação:** função sigmóide para ativar as saídas dos neurônios nas camadas ocultas e o neurônio na saída da rede.
- **Taxa de aprendizagem:** 0.01.
- **Termo de momento:** 0.09.
- **Número de épocas de treinamento:** 1.100 épocas.
- **Erro Quadrático Médio (MSE):** 10^{-3}.

Para se chegar à topologia utilizada, foram realizados testes cuidadosos que evidenciaram ser essa a melhor topologia para a função de ativação utilizada, isto é, que conduzia aos melhores resultados.

Os dados do conjunto de treinamento e teste foram normalizados devido à função de ativação utilizada (os dados devem estar entre 0 e 1).

Na etapa do treinamento da rede, como padrões de entrada foram aplicadas as características faciais extraídas anteriormente pelo módulo de extração de características dos três bancos de dados (PICS, ESSEX e AT&T), totalizando em aproximadamente 100 vetores de características faciais. Estas características foram divididas em dois conjuntos, o conjunto de treinamento, contendo 80% dos vetores de características faciais, e o conjunto de testes, contendo os vetores restantes. A rede MLP foi treinada utilizando o conjunto de treinamento e validada com o conjunto de teste.

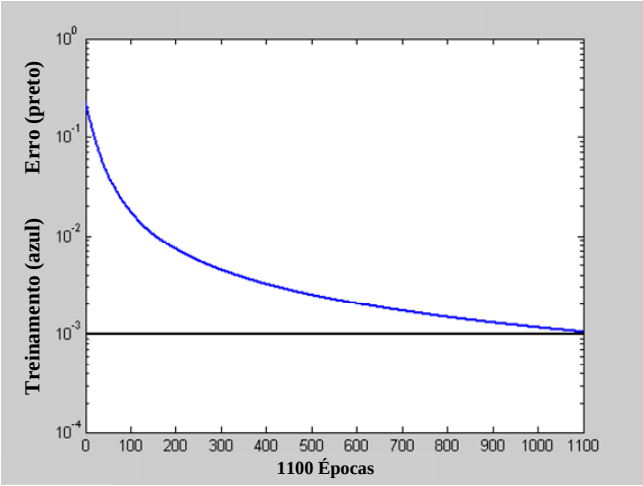


Figura 78: Curva de convergência da rede neural MLP para os dados de treinamento.

A Figura 78 mostra o poder de generalização da rede neural MLP, aplicando o conjunto de treinamento das características faciais. Observa-se que a curva de convergência aproximou-se do erro quadrático médio estabelecido de maneira apropriada.

Tabela 12: Resultado do reconhecimento de faces usando redes neurais MLP.

RNA	Características Faciais PICS / ESSEX / AT&T	
	Acerto	Erro
MLP *Backpropagation*	97%	3%

A Tabela 12 mostra a porcentagem de acerto aplicando a rede neural MLP para o reconhecimento facial. Apesar dos resultados satisfatórios obtidos para o reconhecimento facial constatou-se que essa rede não consegue separar adequadamente os *clusters* de características faciais, e portanto não é a rede mais eficiente para esse fim. Observou-se que as imagens com características faciais próximas que não pertenciam ao conjunto de treinamento foram classificadas como faces pertencentes ao banco de dados, o que resultou no erro de 3%.

Para melhor separar os vetores de características, foi aplicado um fator de multiplicação 7 e, posteriormente todas as características foram divididas pelo maior valor do conjunto de treinamento.

7 Conclusões e Sugestões para trabalhos futuros

Nesse capítulo são apresentadas as conclusões e são sugeridas melhorias no algoritmo, para trabalhos futuros.

7.1 Conclusões

Nesse trabalho foi implementado um algoritmo para reconhecimento de faces humanas, baseado em redes neurais. O algoritmo é constituído por quatro módulos: pré-processamento, segmentação, extração de características e o módulo de reconhecimento que usa a rede neural MLP.

Como imagens de entrada foram utilizados três bancos de dados distintos, totalizando aproximadamente 178 imagens faciais.

Considerando os resultados e discussões apresentadas no Capítulo 6, conclui-se que:

a) Para as etapas de pré-processamento, segmentação e extração de características foram obtidos resultados satisfatórios para os bancos PICS e AT&T, pois o algoritmo de assinatura horizontal e vertical conseguiu localizar de maneira satisfatória os componentes faciais (olhos e nariz). No entanto para o banco de dados ESSEX, a taxa de erro foi alta devido a problemas de resolução e iluminação das imagens.

b) O módulo de extração de características é simples, eficiente e de baixo custo computacional, para a detecção dos componentes faciais, apresentando resultados satisfatórios para imagens faciais frontais, com óculos e sem expressão facial e inclusive para faces com inclinação de até 25 graus.

As principais situações onde o algoritmo não consegue encontrar os componentes faciais de forma eficiente são as seguintes: olhos parcialmente abertos, brilho na lente dos óculos durante a aquisição da imagem facial, expressão facial sorrindo e baixa luminosidade.

c) Aplicando as características faciais extraídas como entrada para a rede neural, a taxa de acerto mostrou ser satisfatória, pois 97% das imagens foram reconhecidas corretamente.

Apesar dos resultados satisfatórios obtidos para os bancos de dados mencionados, o algoritmo ainda não é adequado para fins comerciais, pois imagens com características faciais próximas, não pertencentes ao conjunto de treinamento podem ser classificadas como pertencentes ao banco de dados.

Verificou-se posteriormente que o desempenho da rede MLP pode ser melhorado, utilizando-se como função de ativação a tangente hiperbólica, o que permite reduzir o número de camadas para duas camadas, com 30 neurônios na primeira camada oculta e 60 neurônios na segunda camada oculta.

7.2 Contribuições

Pelos resultados apresentados pode-se concluir que esta dissertação trouxe novas contribuições para a área de visão computacional, instigando maiores explorações no desenvolvimento específico do tema.

As contribuições desse trabalho são as seguintes:

- O estado da arte sobre extração de características fornece grandes contribuições, com trabalhos recentes sobre os métodos utilizados para extrair características faciais. A etapa de extração de características desenvolvida nesse trabalho, já foi

utilizada em outra pesquisa, sendo aplicado para a localização de falanges e metacarpos, com o propósito de extrair características da mão de crianças para estimar a idade óssea (QUEIROZ et. al, 2005).

- Nesse trabalho comprovou-se que o algoritmo de assinatura horizontal e vertical é simples e extremamente eficiente, desde que tenha sido aplicado um pré-processamento adequado sobre as imagens de entrada.
- Esse trabalho mostrou que apesar da rede neural MLP ser uma excelente rede utilizada para classificação, não é adequada para diferenciar faces com características próximas.

7.3 Propostas para trabalhos futuros

As propostas para melhorias nesse trabalho são as seguintes:

- Desenvolvimento de um algoritmo de pré-processamento capaz de padronizar as imagens no que se refere a tamanho, iluminação, inclinação e rotação.
- Utilização de uma rede neural RBF, LVQ ou Kohonen, que permita separar os *clusters*, de modo que as faces com características próximas sejam diferenciadas pelo sistema de reconhecimento.

Referências Bibliográficas

AHLBERG, J. (1999). "Facial Feature Extraction using Eigenspaces and Deformable Graphs", *International Workshop on Synthetic-Natural Hybrid Coding and Three Dimensional Imaging*, pp.8-11.

AITKENHEAD, M. J.; MCDONALD, A. J. S. (2003). "A neural network face recognition system", *Engineering Applications of Artificial Intelligence / Pergamon*, pp.167-176.

ALLINSON, N. M.; ELLIS, A. W. (1992). "Face recognition: combining cognitive psychology and image engineering", *Eletronics & Communication Engineering Journal*, vol.4, no.5, pp.291-300.

AT&T *AT&T Cambridge University Computer Laboratory*, http://www.uk.research.att.com/facedatabase.html ou http://www.cl.cam.ac.uk/Research/DTG/attarchive/facedatabase.html.

BANARSE, D. S.; FRANCE, I.; DULLER, A. W. G. (1999). "Analysis and application of a self-organising image recognition neural network", *Advances in Engineering Software / Elsevier Science*, vol.31, pp.937-944.

BARRETT, W. A. (1997). "A survey of face recognition algorithms and testing results", *Signals, Systems & Computers*, vol.1, pp.301-305.

BARTLETT, M. S.; LADES, H. M.; SEJNOWSKI, T. (1998). "Independent component representation for face recognition", *In Proceedings, SPIE Symposium on Eletronic Imaging: Science and Technology*, pp.528-539.

BASKAN, S.; BULUT, M. M.; ATALY, V. (2002). "Projection based method for segmentation of human face and its evaluation", *Pattern Recognition Letters / Elsevier Science*, vol.23, pp.1623-1629.

BECKER, S. (1991). "Unsupervised learning procedures for neural networks", *International Journal of Neural System*, vol.2, pp.17-33.

BELHUMEUR, P. N.; HESPANHA, J. P.; KRIEGMAN, D. J. (1997). “Eigenfaces vs. fisherfaces: Recognition using class specific linear projection”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.19, no.7, pp.711-720.

BELHUMEUR, P. N.; KRIEGMAN, D. J. The yale face database; avaliado no http://cvc.yale.edu/projects/yalefaces/yalefaces.html

BENNETT, K.; WU, D.; AUSLENDER, L. (1998). *On support vector decision trees for database marketing. Research Report*, Rensselaer Polytechnic Institute, Troy, NY,

BIEDERMAN, I. (1987). “Recognition by components: A theory of human image undestanding”, *Psych. Rev. 94*, pp.115-147.

BIEDERMAN, I.; KALOCSALK, P. (1998), *Neural and psychophysical analysis of object and face recognition*, In Face Recognition: From Theory to Applications, H. Wechsler, P. J. Phillips, V. Bruce, F. F. Soulie, and T. S. Huang, Eds. Springer-Verlag, Berlin, Germany.

BIENENSTOCK, E. L.; COOPER, L. N.; MUNRO, P. W. (1982). J. Neurosci, vol.2, pp.32-48.

BLEDSOE, W. W. (1996). *Man-machine facial recognition*, Technical Report Rep. PRI:22, Panoramic Researsh Inc,

BOSER, B. E.; GUYON, L. M.; VAPNIK, V. N. (1992). “A training algorithm for optimal margin classifiers”, *Proceedings of the 5th Annual ACM Workshop on Computation Learning Theory*, pp.144-152.

BROWN, M. P. S.; GRUNDY, W. N.; LIN, D.; CRISTIANINI, N.; SUGNET, C. W.; FUREY, T. S.; ARES, M. J.; HAUSSLER, S. (2000). ″Knowledge-based analysis of microarray gene expression data by using support vector machines″, *Proceedings of the National Academy of Sciences of the USA*, vol.97, pp.262-267.

BRUCE, V.; BURTON, M.; DENCH, N. (1994). “What's distinctive about a distinctive face?”, *Quart. J. Exp. Psych.*, vol.47A, pp.119-141.

BRUCE, V.; HANCOCK, P. J. B.; BURTON, A. M. (1998), *Human face perception and identification*, In Face Recognition: From Theory to Applications, H. Wechsler, P. J. Philips, V. Bruce, F. F. Soulie, and T. S. Huang, Eds. Springer-Verlag, Berlin, Germany.

BRUNELLI, R.; POGGIO, T. (1992). ″Face recognition through geometrical features″, *Proceedings 2nd European Conference Computer Vision*, pp.792-800.

BRUNELLI, R.; POGGIO, T. (1993). “Face recognition: Features versus templates”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.15, no.10, pp.1042-1052.

BURGES, C. J. C. (1998). “A tutorial on support vector machines for pattern recognition”, *Data Mining and Knowledge Discovery*, vol.2, no.2, pp.121-167.

CASTRO, L. N. D. (2003). *Fundamentals of Neurocomputing*, Faculdade de Engenharia Elétrica e de Computação da Universidade de Campinas (UNICAMP), http://www.dca.fee.unicamp.br/~lnunes.

CELLI, P. L. F. (1999). *Um Sistema de Apoio à Identificação de Suspeitos com Reconhecimento Automático de Faces*, Instituto de Ciências Matemática e Computação, Universidade de São Paulo.

CHELLAPPA, R.; ETEMAD, K. (1996). “Face recognition using discriminant eigenvectors”, *IEEE*, pp.2148-2151.

CHELLAPPA, R.; WILSON, C. L.; SIROHEY, S. (1995). “Human and Machine Recognition of Faces: A Survey”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.83, no.5,

CHEN, J.-L.; CHANG, J.-Y. (2000). “Fuzzy perceptron neural networks for classifiers with numerical data and linguistic rules as inputs”, *IEEE Transactions Fuzzy System*, vol.8, pp.730-745.

CHIANG, C. C.; TAI, W. K.; YANG, M. T.; HUANG, Y. T.; HUANG, C. J. (2003). “A novel method for detecting lips, eyes and faces in real time”, *Real Time Imaging / Elsevier Science*, no.9, pp.277-287.

CHOW, G.; LIN, X. (1993). “Towards a system for automatic facial feature detection”, *Pattern Recognition*, vol.26, pp.1739-1755.

DAUGMAN, J. (1993). “High confidence visual recognition of persons by a test of statistical independence”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.15, no.11, pp.1148-1161.

DIAMANTARAS, K. I.; KUNG, S. Y. (1996), *Principal Component Neural Networks: Theory and Applications*, New York, Wiley.

DUMAIS, S.; PLATT, J.; HECKERMAN, D.; SAHAMI, M. (1998). ″Inductive learning algorithms and representations for text categorization″, *In 7th International Conference on Information and Knowledge Management*, pp.148-155.

ELLIS, H. D. (1986). “Introduction to aspects of face processing: Ten questions in need of answers”, *Aspects of Face Processing*, pp.3-13.

ESSEX *Computer Science at the University of Essex*, http://cswww.essex.ac.uk/mv/allfaces/.

FDSW *Face Detector Software Web*, http://aias.csd.uch.gr:8999/cff/Upload.html.

FERNÁNDEZ, R.; VIENNET, E. (1999). ″Face Identification using Support Vector Machines″, *Proc. Symposium on Artificial Neural Networks*, pp.195-200.

FISHER, R. A. (1938). “The Statistical Utilization of Multiple Measurements”, *Annals of Eugenics*, vol.8, pp.376-386.

FRANCO, L.; TREVES, A. (2001). ″A neural network face expression recognition system using an unsupervised local processing″, *Proceedings of the Second International Symposium on Image and Signal Processing and Analysis (ISPA'01)*, pp.628-632.

GASPAR, T. L.; PAIVA, M. S. V. (2005). "Localização de Características Faciais Usando Projeções Horizontais e Verticais", *I Workshop de Visão Computacional (WVC'2005)*, pp.135-138.

GAUTHIER, I.; LOGOTHETIS, N. K. (2000). "Is face recognition so unique after All?", *Jounal Cognitive Neuropsych*, vol.17, pp.125-142.

GONZALEZ, R. C.; WOODS, R. E. (1992), *Digital Image Processing*, Reading, Ed. Addison-Wesley.

GROB, M. (1994), *Visual computing: the integration of computer graphics, visual perception and imaging*, Springer-Verlag, Berlin.

GU, H.; SU, G.; DU, C. (2003). "Feature Points Extraction from Faces", *Image and Vision Computing, New Zealand*, pp.154-158.

GUO, G.; LI, S. Z.; CHAN, K. (2000). "Face recognition by support vector machines", *Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, pp.196-201.

HASTIE, T.; STUETZLE, W. (1989). "Principal curves", *Jounal of the American Statistical Association*, vol.84, pp.502-516.

HAYKIN, S. (2001), *Redes Neurais: princípios e prática*, Editora Bookman.

HEISELE, B.; HO, P.; WU, J.; POGGIO, T. (2003). "Face recognition: component-based versus global approaches", *Computer Vision and Image Understanding / Academic Press*, pp.6-21.

HILL, H.; SCHYNS, P. G.; AKAMATSU, S. (1997). "Information and viewpoint dependence in face recognition", *Cognition*, vol.62, pp.201-222.

HJELMAS, E.; WROLDSEN, J. (1997). "Face recognition with multi-layer perceptrons", *Proceedings of Norwegian Image Processing and Pattern Recognition Conference*,

HOPFIELD, J. J. (1982). "Neural Networks and Physical Systems with Emergent Collective Computational Abilities", *Proceedings of National Academy of Sciences*, vol.79, pp.2554-2558.

HORIKAWA, S.-I.; FURUHASHI, T.; UCHIKAWA, Y. (1992). "On Fuzzy Modeling Using Fuzzy Neural Netwoks with the Back-Propagation Algorithm", *IEEE Transactions on Neural Networks*, vol.3, no.5, pp.801-806.

HOTTELING, H. (1933). "Analysis of a complex of statistical variables into principal components", *Journal of Educational Psychology*, vol.24, pp.417-441, 498-520.

HUANG, J.; WECHSLER, H. (1999). "Eye detection using optimal wavelet packets and radial basis functions (RBFs)", *International Journal of Pattern Recognition and Artificial Intelligence*, vol.13, pp.1009-1026.

INTRATOR, N. (1993). "Combining exploratory projection pursuit and projection pursuit regression with applications to neural networks", *Neural Computation*, vol.5, no.3, pp.443-455.

INTRATOR, N.; REISFELD, D.; YESHURUN, Y. (1996). "Face recognition using a hybrid supervised/unsupervised neural network", *Pattern Recognition Letters / Elsevier Science*, pp.67-76.

JOACHIMS, T. (1998). "Text categorization with support vector machines", *In Proceedings of European Conference on Machine Learning (ECML)*, pp.137-142.

KAMPMANN, M.; ZHANG, L. (1998). "Estimation of eye, eyebrow and nose features in videophone sequences", *Int. Workshop on Very Low Bitrate Video Coding*, pp.101-104.

KARHUNEN, J.; JOUTSENSALO, J. (1995). "Generalizations of principal component analysis, optimization problems, and neural networks", *Neural Networks*, vol.8, pp.549-562.

KARHUNEN, K. (1947). “Uber lineare methoden in der Wahrscheinlichkeitsrechnung”, *Annales Academiae Scientiarum Fennicae, Series AI: Mathematica-Physica*, vol.37, no.3-79,

KAWAGUCHI, T.; HIDAKA, D.; RIZON, M. (2000). “Robust Extraction of Eyes From Face”, *Proceedings 15th International Conference on Pattern Recognition*, vol.1, pp.1109-1114.

KIRBY, M.; SIROVICH, L. (1990). “Application of the karhunen-loève procedure for the characterization of human faces”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.12, no.1, pp.103-108.

KOHONEN, T. (1982). ″Clustering Taxonomy and Topological Maps of Patterns″, *International Conference on Pattern Recognition*, pp.114-128.

KOHONEN, T. (1990). ″Self-organizing maps″, *Proceedings of the IEEE*, vol.9, pp.1464-1480.

KRUSE, R.; NAUCK, D. (1995). ″Learning Methods for Fuzzy Systems″, *3rd German GI-Workshop "Neuro-Fuzzy-Systems"*, pp.15-17.

KUMAR, V. P.; POGGIO, T. (2000). ″Learning-Based Approach to Real Time Tracking and Analysis of Faces″, *Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000*,

KUNG, S. Y.; TAUR, J. S. (1995). “Decision-based neural networks with signal/image classification applications”, *IEEE Transactions Neural Networks*, vol.6, pp.170-181.

LAM, K.-M.; YAN, H. (1998). “An Analytic-to-Holistic Approach for Face Recognition Based on a Single Frontal View”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.20, no.7, pp.673-686.

LAWRENCE, S.; GILES, C. L.; TSOI, A. C.; BACK, A. D. (1996). *Face recognition: A hybrid neural network approach*, Technical report, Institute for Advanced Computer Studies, University of Maryland,

LAWRENCE, S.; GILES, C. L.; TSOI, A. C.; BACK, A. D. (1997). “Face Recognition: A Convolutional Neural Network Approach”, *IEEE Transactions on Neural Networks*, vol.8, no.1, pp.98-112.

LEE, S. Y.; HAM, Y. K.; PARK, R. H. (1996). “Recognition of Human Front Faces Using Knowledge-Based Feature Extraction and Neuro-Fuzzy Algorithm”, *Pattern Recognition*, vol.29, no.11, pp.1863-1876.

LIN, C. H.; WU, J. L. (1999). “Automatic Facial Feature Extraction by Genetic Algorithms”, *IEEE Transactions on Image Processing*, vol.8, no.6, pp.834-845.

LIN, S. H.; KUNG, S. Y. (1997). “Face Recognition/Detection by Probabilistic Decision-Based Neural Network”, *IEEE Transactions on Neural Networks*, vol.8, no.1, pp.114-132.

LOÈVE, M. (1963), *Probability Theory*, 3rd edition, New York: Van Nostrand.

LUO, M.; YUWAPAT (2003). *Eigenfaces for faces recognition - final project report*, Technical report, Departament of Eletrical and Computer Engineering - CS,

M2VTS Multi Modal Verification for Tele-services and Security applications, ACTS project No. 102,

MARR, D. (1982), *Vision*, W. H. Freeman and Company, New York.

MARTINEZ, M.; BENAVENTE, R. (1998). *The AR Face Database. CVC Technical Report*,

MOGHADDAM, B.; PENTLAND, A. P. (1997). “Probabilistic visual learning for object representation”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.19, no.7, pp.696-710.

MOGHADDAM, B.; WAHID, W.; PENTLAND, A. (1998). “Beyond eigenfaces: Probabilistic mathing for face recognition”, *In Third IEEE International Conference on Automatic Face and Gesgure Recognition, Nara, Japan. IEEE Computer Society Press*,

NAUCK, D.; KLAWONN, F.; KRUSE, R. (1992). “Fuzzy Sets, Fuzzy Controllers, and Neural Networks”, *Scientific Journal of the Humboldt-University of Berlin, Series Medicine 41*, vol.4, pp.99-120.

NAUCK, D.; KRUSE, R. (1997a). “A neuro-fuzzy method to learn fuzzy classification rules from data”, *Fuzzy sets and systems / Elsevier Science*, vol.89, pp.277-288.

NAUCK, D.; KRUSE, R. (1997b). ″What are Neuro-Fuzzy Classifiers?″ *7th International Fuzzy Systems Association World Congress IFSA'97*, vol.IV, pp.228-233.

NIELSEN, R. H. (1991), *Neurocomputing*, Addison-Wesley. Massachusetts, EUA.

ORL *ORL Database (Oliver Research Laboratory)*, http://www.orl.co.uk/facedatabase.html.

OSUMA, E.; FREUND, R.; GIROSI, F. (1997). “Training Support Vector Machines: an Application to Face Detection”, *Conference on Computer Vision and Pattern Recognition (CVPR'97)*,

O'TOOLE, A. J.; CHENG, Y.; ROSS, B.; WILD, H. A.; PHILIPS, P. J. (2000). “Face Recognition Algorithms as Models of Human Face Processing”, *Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000*, pp.552-557.

PEARSON, K. (1901). *On lines and planes of closest fit to systems of point in space*, Philosophical Magazine, vol.2, pp.559-572.

PENTLAND, A. (2000). “Looking at people: Sensing for ubiquitous and weareble computing”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.22, no.1, pp.107-119.

PENTLAND, A.; MOGHADDAM, B.; STARNER, T. (1994). “View-based and modular eigenspaces for face recognition”, *In Proceedings, IEEE Conference on Computer Vision and Pattern Recognition*, pp.84-91.

PERELMUTER, G.; E., E. V. C.; VELLASCO, M.; PACHECO, M. A. (1995). "Reconhecimento de imagens bidimensionais utilizando redes neurais artificiais", *Anais do VIII SIBGRAPI*, pp.197-203.

PERKINS, D. (1975). "A definition of caricature and recognition", *Stud. Anthro. Vis. Commun.*, vol.2, pp.1-24.

PHILIPS, P. J.; RIZVI, S. A.; RAUSS, P. J. (2000). "The FERET Evaluation Methodology for Face-Recognition Algorithms", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol.22, no.10, pp.1090-1104.

PICS *Psychological Image Collection at Stirling (PICS)*, http://pics.psych.stir.ac.uk/cgi-bin/PICS/New/pics.cgi.

PONTIL, M.; VERRI, A. (1998). "Object recognition with support vector machines", *IEEE Transaction on PAMI*, vol.20, pp.637-646.

QUEIROZ, A. C.; GASPAR, T. L.; OLIVETE, C. J.; AFFONSO, F. J.; RODRIGUES, E. L. L. (2005). "Metodologia para Extração de Características Automáticas da Mão Usando Assinatura Vertical", *I Workshop de Visão Computacional (WVC'2005)*, Piracicaba, pp.60-63.

RANGANATH, S.; ARUN, K. (1997). "Face recognition transform features and neural networks", *Pattern Recognition / Elsevier Science*, vol.30, no.10, pp.1615-1622.

RITTER, H.; MARTINETZ, T.; SCHULTEN, K. (1992), *Neural Computation and Self-Organizing Maps: An Introduction*, Reading MA: Addison-Wesley.

ROWLEY, H. A.; BALUJA, S.; KANADE, T. (1998). "Neural Network-Based Face Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.20, no.1, pp.23-38.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. (1986). "Learning representations of back-propagation errors", *Nature (London)*, vol.323, pp.533-536.

SHEPHERD, J. W.; DAVIES, G. M.; ELLIS, H. D. (1981), *Studies of cue saliency*, In Perceiving and Remembering Faces, G. M. Davies, H. D. Ellis, and J. W. Shepherd, Eds. Academic Press, London, U.K.

SIROVICH, L.; KIRBY, M. (1987). "Low-dimensional procedure for the characterization of human faces", *in J. Opt. Soc. Am. A.*, vol.4, no.3, pp.519-524.

SUNG, K.; POGGIO, T. (1994). "Example-based Learning for View-based Human Face Detection", *A. I. Memo 1521, MIT A. I. Lab.*,

SUNG, K.-K.; POGGIO, T. (1998). "Example-Based Learning for View-Based Human Face Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.20, no.1, pp.39-51.

TALUKDER, A.; CASASENT, D. (2001). "Adaptive activation function neural net for face recognition", *IEEE International Joint Conference on Neural Networks*, pp.549-552.

TAMURA, S.; KAWAI, H.; MITSUMOTO, H. (1996). "Male/female identification from 8 x 6 very low resolution face images by neural network", *Pattern Recognition / Pergamon*, vol.29, no.2, pp.331-335.

TARR, M. J.; BURTHOFF, H. H. (1993). "Is human objetct recognition better described by geon structural descriptions or by multiple views- comment on biederman and gerhardstein", *J. Exp. Psych.: Hum. Percep. Perf.*, vol.21, pp.71-86.

TOLBA, A. S.; ASHRAF, M.; ABU-REZQ (2000). "Eyes-strip extraction for detection of a human face", *The Eleventh International Conference on Microeletronics*, pp.109-112.

TURK, M. A.; PENTLAND, A. P. (1991). "Face recognition using eigenfaces", *IEEE Computer Society Conference on*, pp.586-591.

VAPNIK, V. N. (1995), *The Nature of Statistical Learning Theory*, Springer.

VAPNIK, V. N. (1998), *Statistical Learning Theory*, John Wiley and Sons, New York.

VAPNIK, V. N. (1999). "An Overview of Statistical Learning Theory", *IEEE Transactions on Neural Networks*, vol.10, pp.988-999.

VIGLIAZZI, D. (2003), *Biometria - Medidas de Segurança*, Editora Visual Books.

WANG, Y.; RONG, G. (1999). "A self-organizing neural-network-based fuzzy system", *Fuzzy sets and systems / Elsevier*, vol.103, pp.1-11.

WENG, J. J.; SWETS, D. L. (1999), *Face Recognition*, In A. K. Jain, R. Bolle (editors), Biometrics: Personal Identification in Nerworked Society, Kluwer Academic Press, Hingham, Massachusetts.

WIBERG, N. (1996). *Codes and Decoding on General Graphs, Dissertation*, Dept. of Eletrical Engineering, Linköpings Universitet, Sweden,

WISKOTT, L.; FELLOUS, J.-M.; KRÜGER, N.; MALSBURG, C. V. D. (1997). "Face recognition by elastic bunch graph matching", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.19, no.7, pp.775-779.

YANG, M. H.; KRIEGMAN, D. J.; AHUJA, N. (2002). "Detecting Faces in Images: A Survey", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.24, no.1,

YIN, L.; BASU, A. (1999). "Integrating active face tracking with model based coding", *Pattern Recognition Lett.*, vol.26, pp.651-657.

YOON, K. S.; HAN, Y. K.; PARK, R. H. (1998). "Hybrid Approaches to frontal view face recognition using the hidden Markov Model and Neural Network", *Pattern Recognition / Elsevier Science*, vol.31, no.3, pp.283-293.

ZHANG, D.; BAI, X.-L.; CAI, K.-Y. (2004). "Extended neuro-fuzzy models of multilayer perceptrons", *Fuzzy sets and systems / Elsevier Science*, vol.142, pp.221-242.

ZHAO, W.; CHELLAPPA, R.; PHILIPS, P. J.; ROSENFELD, A. (2003). "Face Recognition: A Literature Survey", *ACM Computing Surveys*, vol.35, no.4, pp.399-458.

ZHAO, W. Y.; CHELLAPPA, R.; PHILIPS, P. J. (1999). *Subspace linear discriminant analysis for face recognition*, Center for Automation Research, University of Maryland, College Park, Technical Report CAR-TR-914

APÊNDICE A – Imagens do banco de dados PICS

A.1

A.2

A.3

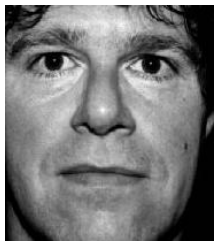
A.4

A.5

A.6

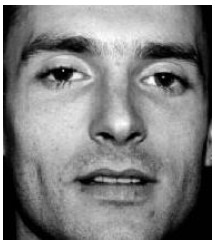
A.7

A.8


A.9

A.10

A.11


A.12

A.13

A.14

A.15

A.16

A.17

A.18

A.19

A.20

A.21

APÊNDICE B – Imagens do banco de dados ESSEX

B.1

B.2

B.3

B.4

B.5

B.6

B.7

B.8

B.9

B.10

B.11

B.12

B.13

B.14

B.15

B.16

B.17

B.18

B.19

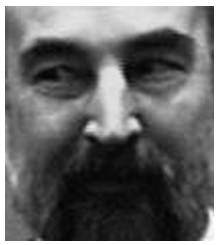
B.20

B.21

B.22

B.23

B.24


B.25

B.26

B.27

B.28

APÊNDICE C – Imagens do banco de dados AT&T

C.1

C.2

C.3

C.4

C.5

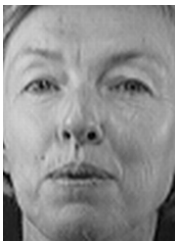
C.6

C.7

C.8

C.9

C.10


C.11

C.12

C.13

C.14

C.15

C.16

APÊNDICE D – Resultado do pré-processamento das imagens dos bancos de dados PICS, ESSEX e AT&T

D.1 - PICS

D.2 - PICS

D.3 - PICS

D.4 - PICS

D.5 - PICS

D.6 - PICS

D.7 - PICS

D.8 - PICS

D.9 - PICS

D.10 – PICS

D.11 - PICS

D.12 - PICS

D.13 - PICS

D.14 - PICS

D.15 – PICS

D.16 - PICS

D.17 - PICS

D.18 - PICS

D.19 - PICS

D.20 – PICS

D.21 - PICS

D.22 - ESSEX

D.23 - ESSEX

D.24 - ESSEX

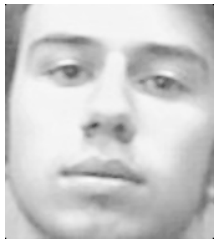
D.25 - ESSEX

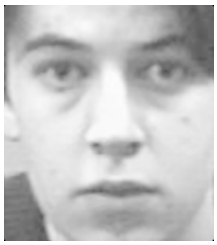
D.26 – ESSEX

D.27 – ESSEX

D.28 – ESSEX

D.29 – ESSEX


D.30 – ESSEX


D.31 – ESSEX

D.32 – ESSEX

D.33 – ESSEX

D.34 – ESSEX

D.35 – ESSEX

D.36 – ESSEX

D.37 – ESSEX

D.38 – ESSEX

D.39 – ESSEX

D.40 – ESSEX

D.41 – ESSEX

D.42 – ESSEX

D.43 – ESSEX

D.44 – ESSEX

D.45 – ESSEX

D.46 – ESSEX

D.47 – ESSEX

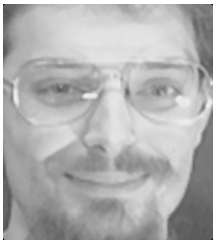
D.48 – ESSEX

D.49 – ESSEX

D.50 – AT&T

D.51 – AT&T

D.52 – AT&T


D.53 – AT&T

D.54 – AT&T

D.55 – AT&T

D.56 – AT&T

D.57 – AT&T

D.58 – AT&T

D.59 – AT&T

D.60 – AT&T

D.61 – AT&T

D.62 – AT&T

D.63 – AT&T

D.64 – AT&T

D.65 – AT&T

APÊNDICE E – Resultado da segmentação das imagens dos bancos de dados PICS, ESSEX e AT&T

E.1 - PICS

E.2 - PICS

E.3 - PICS

E.4 - PICS

E.5 - PICS

E.6 - PICS

E.7 - PICS

E.8 - PICS

E.9 - PICS

E.10 – PICS

E.11 - PICS

E.12 - PICS

E.13 - PICS

E.14 - PICS

E.15 - PICS

E.16 - PICS

E.17 - PICS

E.18 - PICS

E.19 - PICS

E.20 - PICS

E.21 - PICS

E.22 - ESSEX

E.23 - ESSEX

E.24 - ESSEX

E.25 - ESSEX

E.26 – ESSEX

E.27 – ESSEX

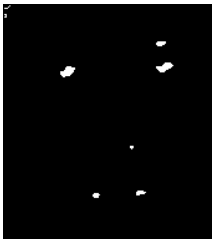
E.28 – ESSEX

E.29 – ESSEX

E.30 – ESSEX

E.31 – ESSEX

E.32 – ESSEX


E.33 – ESSEX

E.34 – ESSEX

E.35 – ESSEX

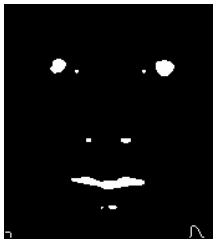
E.36 – ESSEX

E.37 – ESSEX

E.38 – ESSEX

E.39 – ESSEX

E.40 – ESSEX


E.41 – ESSEX

E.42 – ESSEX

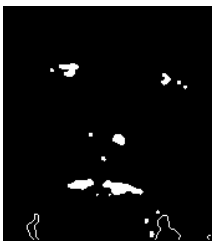
E.43 – ESSEX

E.44 – ESSEX

E.45 – ESSEX

E.46 – ESSEX

E.47 – ESSEX


E.48 – ESSEX

E.49 – ESSEX

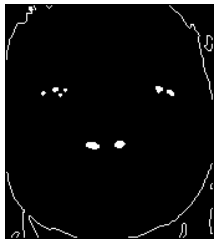
E.50 – AT&T

E.51 – AT&T

E.52 – AT&T

E.53 – AT&T

E.54 – AT&T


E.55 – AT&T

E.56 – AT&T

E.57 – AT&T

E.58 – AT&T

E.59 – AT&T

E.60 – AT&T

E.61 – AT&T

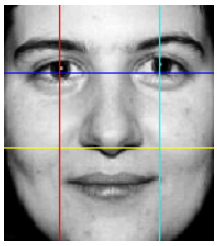
E.62 – AT&T

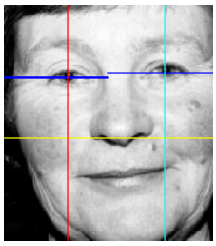
E.63 – AT&T

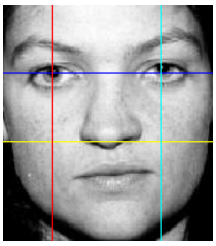
E.64 – AT&T

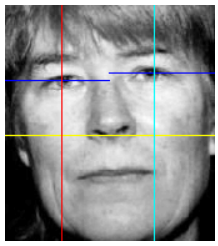
E.65 – AT&T

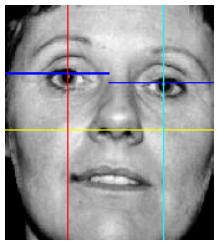
APÊNDICE F – Resultado das assinaturas horizontais e verticais dos bancos de dados PICS, ESSEX e AT&T

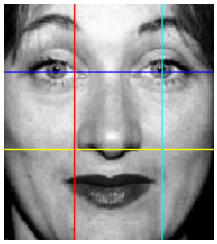

F.1 - PICS

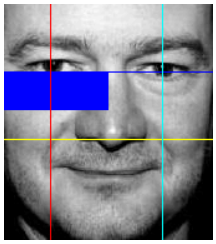

F.2 - PICS

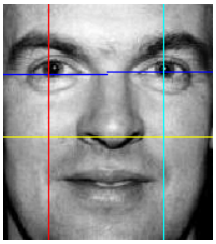

F.3 - PICS

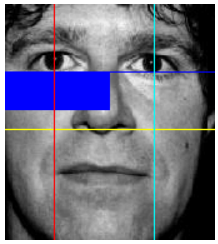

F.4 - PICS

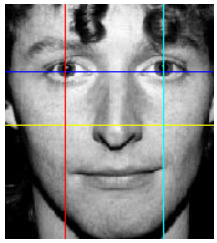

F.5 – PICS

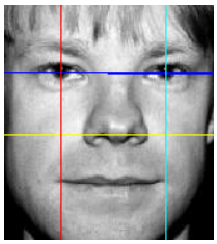

F.6 - PICS

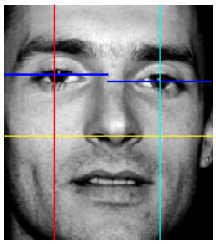

F.7 - PICS

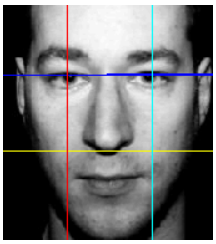

F.8 - PICS

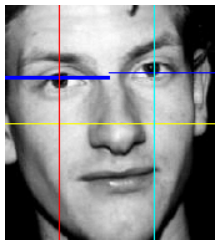

F.9 - PICS

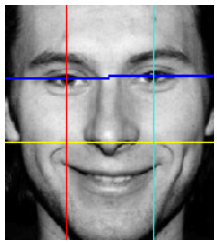

F.10 – PICS

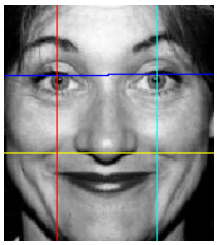

F.11 - PICS

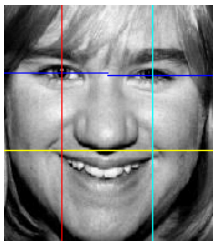

F.12 - PICS

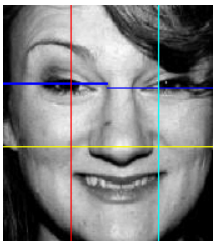

F.13 - PICS

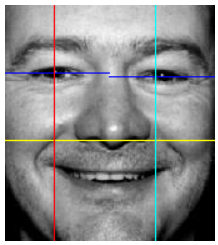

F.14 - PICS

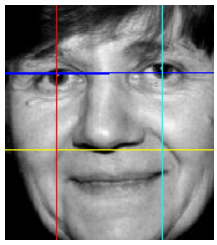

F.15 - PICS

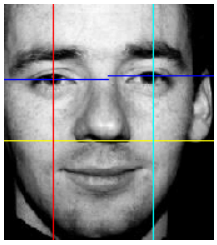

F.16 - PICS

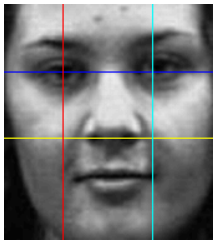

F.17 - PICS


F.18 - PICS

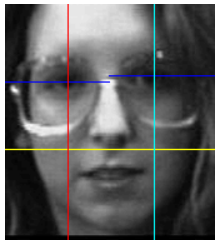

F.19 - PICS

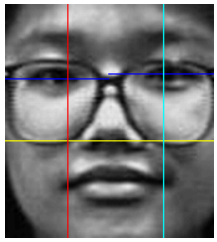

F.20 - PICS


F.21 - PICS

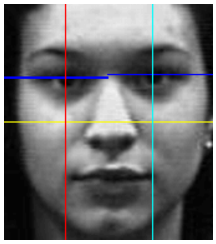

F.22 - ESSEX

F.23 - ESSEX

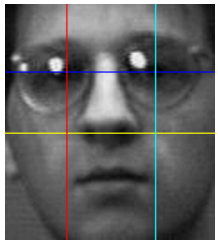

F.24 - ESSEX

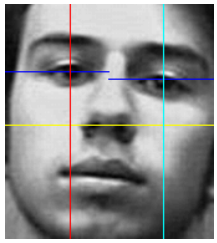

F.25 - ESSEX

F.26 – ESSEX

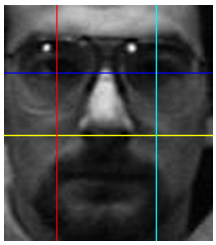

F.27 – ESSEX

F.28 – ESSEX

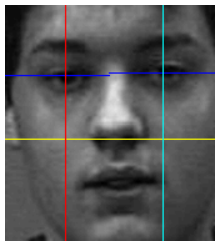

F.29 – ESSEX

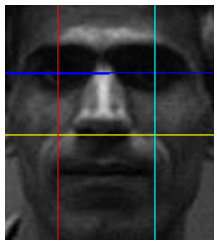

F.30 – ESSEX

F.31 – ESSEX


F.32 – ESSEX

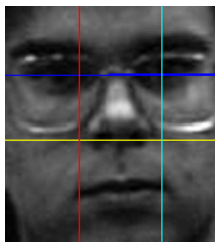
F.33 – ESSEX

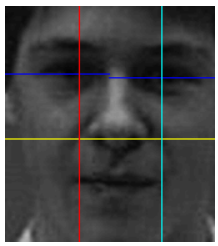

F.34 – ESSEX


F.35 – ESSEX

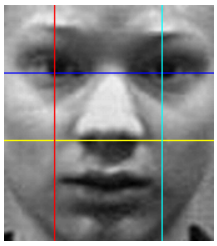
F.36 – ESSEX

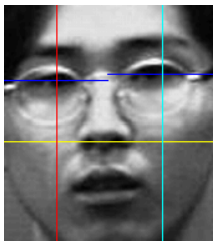
F.37 – ESSEX

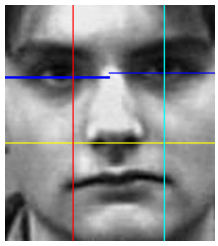

F.38 – ESSEX

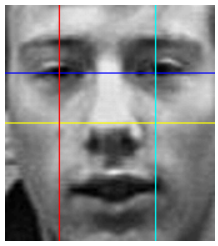

F.39 – ESSEX

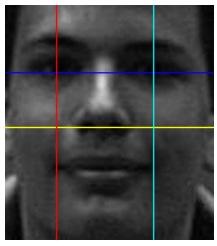
F.40 – ESSEX

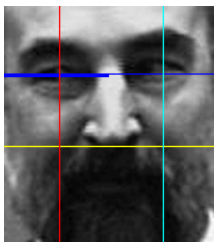

F.41 – ESSEX

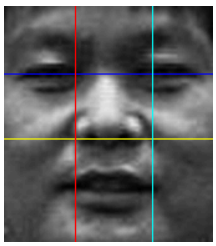

F.42 – ESSEX

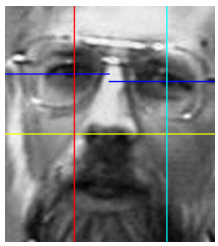

F.43 – ESSEX


F.44 – ESSEX


F.45 – ESSEX


F.46 – ESSEX

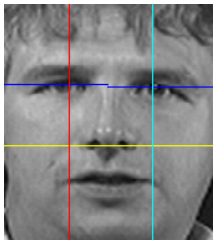

F.47 – ESSEX

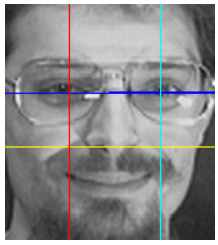

F.48 – ESSEX

F.49 – ESSEX

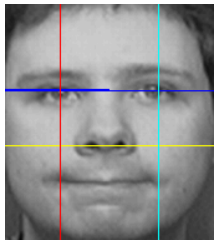
F.50 – AT&T

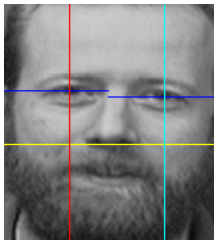
F.51 – AT&T


F.52 – AT&T


F.53 – AT&T

F.54 – AT&T


F.55 – AT&T


F.56 – AT&T

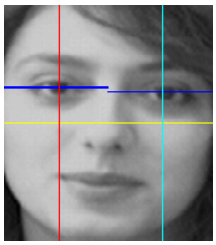
F.57 – AT&T

F.58 – AT&T

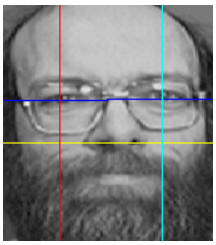
F.59 – AT&T

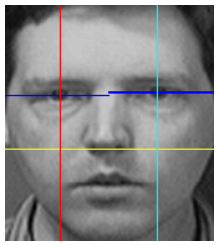
F.60 – AT&T

F.61 – AT&T


F.62 – AT&T

F.63 – AT&T


F.64 – AT&T


F.65 – AT&T