

Sistemas de Identificação Pessoal Utilizando Técnicas de Reconhecimento e Verificação Facial Automáticas

Tese submetida à Faculdade de Engenharia Elétrica e de
Computação da Universidade Estadual de Campinas, Departamento
de Comunicações, como parte dos requisitos exigidos para a
obtenção do título de

Mestre em Engenharia Elétrica

Autor

Ernesto Luiz Andrade Neto

Engenheiro Eletricista pela UFSM em 1994

Orientador

Prof. Dr. Lee Luan Ling

PhD em Engenharia Elétrica Cornell University em 1992

Campinas, maio de 1997.

Este exemplar corresponde à redação final da tese	Ernesto Luiz Andrade
defendida por	Neto
Julgadora em	28 / 04 / 1997
	<i>[Assinatura]</i> Orientador

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

An25s Andrade Neto, Ernesto Luiz
Sistemas de identificação pessoal utilizando técnicas de reconhecimento e verificação facial automáticas / Ernesto Luiz Andrade Neto.--Campinas, SP: [s.n.], 1996.

Orientador: Lee Luan Ling.
Dissertação (mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Reconhecimento de padrões. 2. Processamento de imagens. 3. Análise de componentes principais. I. Lee, Luan Ling. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

**Para minha família, fonte de amor
compreensão e carinho em todos os
momentos.**

"Nature speaks to us patiently in many
tongues, most of which we understand
only imperfectly."

Dudley Herschbach

AGRADECIMENTOS

Aos meus pais, Ernesto e Ruth, e irmãos, Cristiano e Rubens, pela grande apoio que sempre me deram em todas as etapas de minha vida.

Ao meu orientador Lee Luan Ling pela colaboração, incentivo e oportunidade de desenvolver este trabalho sob sua orientação.

A Roberto de Alencar Lotufo pela apoio e interesse na realização deste trabalho.

A Alex Pentland pela colaboração e presteza em fornecer informações relevantes para o desenvolvimento do trabalho.

A Leonardo Mendes por sua colaboração na revisão do manuscrito.

Aos colegas do Departamento de Comunicações Natanael, Jackson, Miguel, Antônio Alberti, Marcelo Pellenz, Antônio Moraes, Sandro, Mylene, Marcelo, Paula, Fábio, Osmar, Marco, Luís Romulo, Maria Luiza, Gean, Albanita, Lin Voigt, Kin, Renato, Paulo, Alexandre Amorin, Rivelino pela convivência agradável e pelo apoio que sempre nos proporcionamos.

A todas os colegas e pessoas que colaboraram fornecendo pacientemente suas imagens para serem usadas neste trabalho.

SUMÁRIO

Lista de Tabelas.....	iv
Lista de Figuras.....	v
Resumo.....	ix
Abstract.....	x

1. INTRODUÇÃO.....1

1.1 PROBLEMAS BÁSICOS NO RECONHECIMENTO DE FACES	9
1.1.1 Detecção de Faces	12
1.1.2 Representação de Faces	12
1.1.3 Identificação de Faces	13
1.1.4 Análise de Expressão Facial	13
1.1.5 Classificação Baseada em Características Físicas.....	14
1.2 TRABALHO DESENVOLVIDO	14
1.3 ESTRUTURA DO TRABALHO.....	15

2. PROCESSAMENTO DE IMAGENS FACIAIS.....16

2.1 O RECONHECIMENTO FACIAL	16
2.2 DETECÇÃO DE FACES EM IMAGENS	18
2.2.1 Casamento de Padrões	19
2.2.1.1 Correlação de Padrões.....	19
2.2.1.2 Correlação de Padrões Normalizada	20
2.2.1.3 Padrões Deformáveis.....	21
2.2.1.4 Casamento de Cores	22
2.2.1.5 Casamento de Texturas	22
2.2.2 Redes Neurais	22
2.2.3 Eigentemplates.....	23
2.3 NORMALIZAÇÃO DAS IMAGENS FACIAIS.....	24
2.4 EXTRAÇÃO DE CARACTERÍSTICAS E RECONHECIMENTO FACIAL	26
2.4.1 Redes Neurais	27

2.4.1.1 Perceptrons Multi-Camada.....	27
2.4.1.2 Autoassociador Linear.....	28
2.4.1.3 Redes de Compressão de Imagem	29
2.4.1.4 Mapas Auto-Organizáveis de Kohonen.....	30
2.4.1.5 Casamentos de Ligações Dinâmico	30
2.4.1.6 Rede de Neural de Convolução	31
2.4.1.7 Árvores de Redes Neurais	32
2.4.2 Métodos Estatísticos	33
2.4.2.1 <i>Eigenfaces</i>	33
2.4.2.2 Análise Discriminante	35
2.4.2.3 HMM.....	36
2.4.3 Coeficientes de Autocorrelação	36
2.5 DISCUSSÃO.....	37
3. O MÉTODO DE <i>EIGENFACES</i>.....	39
3.1 INTRODUÇÃO	40
3.2 ORIGENS DO MÉTODO	41
3.3 OBTENÇÃO DO MODELO DE <i>EIGENFACES</i>	42
3.4 CÁLCULO DAS <i>EIGENFACES</i>	46
3.5 REPRESENTAÇÃO USANDO <i>EIGENFACES</i>	50
3.6 CLASSIFICAÇÃO COM <i>EIGENFACES</i>	51
3.7 DISCUSSÃO.....	54
4. SISTEMA DE RECONHECIMENTO FACIAL.....	55
4.1 BASE DE DADOS.....	55
4.1.1 Metodologia de Aquisição.....	56
4.1.2 Catalogação das Imagens.....	61
4.2 IMPLEMENTAÇÃO DO MÉTODO DE <i>EIGENFACES</i>	64
4.2.1 Normalização da Imagem Facial	66
4.2.2 Construção do Modelo de <i>Eigenfaces</i>	73
4.2.3 Extração de Características.....	78

4.2.4 Classificação.....	79
4.3 RECONSTRUÇÃO DE IMAGENS COM <i>EIGENFACES</i>	80
4.4 DISCUSSÃO.....	83
5. RESULTADOS DO MODELO DE <i>EIGENFACES</i>	84
5.1 SISTEMA DE RECONHECIMENTO FACIAL.....	84
5.1.1 Determinação do Centro de Classe	85
5.1.2 Determinação do Número de Características.....	87
5.1.3 Análise dos Erros de Classificação	97
5.1.4 Comparação com Outros Classificadores	102
5.2 VERIFICAÇÃO	104
5.2.1 Análise da Verificação.....	105
5.3 DISCUSSÃO.....	109
6. DISCUSSÕES E CONCLUSÕES	110
6.1 CONTRIBUIÇÕES.....	110
6.2 DISCUSSÕES	111
6.2.1 A Qualidade das Imagens da Base de Dados	111
6.2.2 O Processo de Normalização.....	112
6.2.3 Outros Resultados com o Método de <i>Eigenfaces</i>	113
6.3 CONCLUSÕES	114
6.4 SUGESTÕES PARA FUTUROS TRABALHOS	114
7. REFERÊNCIAS	116

Lista de Tabelas

Tabela 1-1: Comparação do desempenho de sistemas biométricos, [PENTLAND-a].	3
Tabela 4-1: Distribuição da ocorrência dos grupos de faces de indivíduos na base de dados.	57
Tabela 4-2: Imagens da base de dados segundo as características selecionadas. Uma mesma imagem pode apresentar mais de uma característica. O grupo modelo é composto por frontais, sem óculos e com expressão neutra que serão usadas na construção dos modelos de eigenfaces.	63
Tabela 4-3: Análise da variação da distância ao espaço de face do modelo de <i>eigenfaces</i> com diferentes tipos de face para as imagens da Figura 4-15.	81
Tabela 5-1: Valores do percentual de reconhecimento para o modelo de 98 <i>eigenfaces</i> sem pré-processamento, utilizando como centro de classe uma face arbitrariamente escolhida.	87
Tabela 5-2: Valores do percentual de reconhecimento para o modelo de 98 <i>eigenfaces</i> sem pré-processamento, utilizando como centro de classe a média das características. .	92
Tabela 5-3: Valores do percentual de reconhecimento para o modelo de 98 <i>eigenfaces</i> com pré-processamento, utilizando como centro de classe a média das características. .	93
Tabela 5-4: Valores do percentual de reconhecimento para o modelo de 178 <i>eigenfaces</i> sem pré-processamento, utilizando como centro de classe a média das características. .	95
Tabela 5-5: Valores do percentual de reconhecimento para o modelo de 178 <i>eigenfaces</i> com pré-processamento, utilizando como centro de classe a média das características.	97
Tabela 5-6: Número de erros associados a cada tipo de imagem.	98

Lista de Figuras

Figura 1-1: Componentes básicos de um sistema para reconhecimento facial automático.	4
Figura 1-2: Diagrama de blocos da operação de um sistema de reconhecimento facial.	5
Figura 1-3: Componentes básicos de um sistema para verificação facial automática.	7
Figura 1-4: Diagrama de blocos da operação de um sistema de verificação facial.	7
Figura 1-5: Diagrama de blocos de um sistema de reconhecimento de padrões clássico [SCHALLKOFF].	11
Figura 2-1: Exemplo de codificação da posição e área da face na imagem de entrada.	18
Figura 2-2: Detecção e extração das características faciais da imagem de entrada.	20
Figura 2-3: Esquema das etapas de um procedimento de normalização facial.	26
Figura 2-4: Exemplo de um modelo de eigenfaces para uma base de dados de 10 imagens frontais.	35
Figura 3-1: Representação de uma imagem na forma de vetor linha.	43
Figura 3-2: Exemplos de <i>eigenfaces</i> , que recebem este nome devido a sua semelhança com as faces humanas.	44
Figura 3-3: Imagem média de um conjunto de imagens faciais frontais e normalizadas.	47
Figura 4-1: Esquema dos equipamentos empregados para a aquisição das imagens da base de dados.	56
Figura 4-2: Base de dados de imagens faciais contendo 435 imagens de 102 indivíduos. Na figura percebe-se a variação das condições de aquisição para as diferentes imagens presentes na base de dados.	59
Figura 4-3: Base de dados do Olivetti Research Center para testes de algoritmos de reconhecimento de imagem. Esta base de dados foi montada por Ferdinando Samaria e utilizada em experimentos descritos em [SAMARIA] e [GOUDAIL].	60

Figura 4-4: Tela do aplicativo desenvolvido para o processo manual de normalização facial. As cruzes marcam os pontos das características que podem ser usadas na normalização.	62
Figura 4-5: Quadro do aplicativo de normalização onde são introduzidos os comentários sobre sexo, sorriso, barba, presença ou ausência de óculos, reflexão da luz na face, posição e iluminação.	63
Figura 4-6: Etapas da operação do sistema facial: 1) construção do modelo de <i>eigenfaces</i> ; 2) catalogação das imagens segundo o modelo; e 3) classificação das novas imagens apresentadas ao sistema.	65
Figura 4-7: Diagrama de blocos do processo de normalização da imagem facial.	69
Figura 4-8: Diagrama de blocos do pré-processamento seguido de normalização da imagem facial.	70
Figura 4-9: Imagens faciais normalizadas sem pré-processamento para construção do modelo facial e reconhecimento.	71
Figura 4-10: Imagens faciais pré-processadas e normalizadas para construção do modelo facial e reconhecimento.	72
Figura 4-11: Imagens selecionadas para a construção do modelo de 98 <i>eigenfaces</i>	73
Figura 4-12: Imagens selecionadas para obtenção do modelo após normalização sem pré-processamento.	74
Figura 4-13: Imagens selecionadas para a construção do modelo de 178 <i>eigenfaces</i>	75
Figura 4-14: Modelo de <i>eigenfaces</i> obtido das 99 imagens de treino selecionadas sem pré-processamento. As imagens estão ordenadas em ordem decrescente da proporção da variância normalizada associadas a suas <i>eigenfaces</i> . A imagem no canto superior esquerdo corresponde a <i>eigenface</i> com maior variância. No canto inferior direito está a imagem média das imagens de treino.	77

Figura 4-15: Reconstrução das imagens originais com diferentes quantidades de <i>eigenfaces</i> .	80
Figura 5-1: Variação da taxa de acerto de reconhecimento em função do número de características utilizadas para o modelo de 98 <i>eigenfaces</i> sem pré-processamento, com uma face arbitrária como centro de classe.	86
Figura 5-2: Distribuição da variância da base de dados associada a cada autovetor do modelo de <i>eigenfaces</i> . (a) modelo de 98 <i>eigenfaces</i> e (b) modelo de 178 <i>eigenfaces</i> sem pré-processamento. (c) modelo de 98 <i>eigenfaces</i> e (d) modelo de 178 <i>eigenfaces</i> com pré-processamento.	88
Figura 5-3: Variância acumulada nos autovetores do modelo de <i>eigenfaces</i> . (a) modelo de 98 <i>eigenfaces</i> e (b) modelo de 178 <i>eigenfaces</i> sem pré-processamento. (c) modelo de 98 <i>eigenfaces</i> e (d) modelo de 178 <i>eigenfaces</i> com pré-processamento.	89
Figura 5-4: Variação da taxa de acerto de reconhecimento em função do número de características utilizadas para o modelo de 98 <i>eigenfaces</i> sem pré-processamento.	92
Figura 5-5: Variação da taxa de acerto de reconhecimento em função do número de características utilizadas para o modelo de 98 <i>eigenfaces</i> com pré-processamento.	93
Figura 5-6: Variação da taxa de acerto de reconhecimento em função do número de características utilizadas para o modelo de 178 <i>eigenfaces</i> sem pré-processamento.	94
Figura 5-7: Variação da taxa de acerto de reconhecimento em função do número de características utilizadas para o modelo de 178 <i>eigenfaces</i> com pré-processamento.	96
Figura 5-8: Imagens que apresentaram erros de classificação para o modelo de 98 <i>eigenfaces</i> sem pré-processamento com sua taxa de reconhecimento máxima. As imagens destacadas são os erros das faces do grupo modelo.	100
Figura 5-9: Imagens que apresentaram erros de classificação para o modelo de 178 <i>eigenfaces</i> sem pré-processamento com sua taxa de reconhecimento máxima. As imagens destacadas são os erros das faces do grupo modelo.	101

Figura 5-10: Imagens que apresentaram erros de classificação para o modelo de 98 <i>eigenfaces</i> com pré-processamento para sua taxa de reconhecimento máxima. As imagens destacadas são os erros das faces do grupo modelo.	101
Figura 5-11: Imagens que apresentaram erros de classificação para o modelo de 178 <i>eigenfaces</i> sem pré-processamento com sua taxa de reconhecimento máxima. As imagens destacadas são os erros das faces do grupo modelo.	102
Figura 5-12: Curvas de percentual de acerto para treinamento e teste do MLP (65-80-102) em função do número de iterações do algoritmo de treinamento.	103
Figura 5-13: Curvas de erro absoluto da rede para treinamento e teste do MLP (65-80-102) em função do número de iterações do algoritmo de treinamento.	104
Figura 5-14: Curva de operação do receptor para em função do número de características utilizadas para o modelo de 178 <i>eigenfaces</i> com pré-processamento.	107
Figura 5-15: Curva da variação da taxa de erro de falsa aceitação em função do limiar de decisão.	108
Figura 5-16: Curva da variação da taxa de erro de falsa rejeição em função do limiar de decisão.	108

RESUMO

A presente tese apresenta o desenvolvimento de sistemas biométricos utilizando técnicas de reconhecimento de face. Sistemas biométricos faciais permitem reconhecer e identificar indivíduos usando apenas a informação extraída de suas imagens faciais. Os sistemas biométricos de reconhecimento de face necessitam ter um modelo para codificar e representar a informação da face para posterior reconhecimento. Escolhemos como modelo de representação da informação facial o método de eigenfaces [TURK]. Devido a capacidade do método de extrair e comprimir a informação da imagem facial ele compõe o núcleo para a extração de características de nosso sistema de reconhecimento. O processamento do método de eigenfaces em nosso sistema de reconhecimento facial foi implementado de modo semi-automático.

Os testes da arquitetura do sistema de reconhecimento facial com eigenfaces são realizados para as aplicações de reconhecimento e verificação facial. Na realização dos testes, adquirimos uma base de dados de imagens faciais composta de 435 imagens de 102 indivíduos. Não foram impostas restrições severas ao processo de aquisição da base de dados, resultando em grandes variações entre imagens do mesmo indivíduo. Avaliamos a tolerância do sistema as diversas variações presentes nas imagens da base de dados. Todas as rotinas necessárias às tarefas de processamento de imagens, reconhecimento e verificação de faces foram escritas em C++. Elas compõe uma biblioteca de rotinas que pode ser adaptada para outras tarefas como a segmentação automática de imagens faciais.

O sistema de reconhecimento de faces é testado na tarefa de classificar corretamente 435 imagens de 102 pessoas atribuindo a cada imagem sua classe correspondente. O sistema de reconhecimento de faces obteve uma taxa de acerto de 97.70% utilizando um classificador de distância mínima. O sistema de verificação facial de faces é testado com um classificador de distância mínima onde variamos o limiar de aceitação.

ABSTRACT

This thesis dissertation shows the development of biometrics systems based in face recognition technics. These biometric systems are able to recognize and identify people by extracting information from their facial images. Face recognition systems must have a model to encode and represent facial image information for recognition purposes. We choose the eigenface method [TURK], due to its remarkable capability to extract and compress facial image information, as the core for the feature extraction task. The eigenface method was implementd in a semi-automatic way.

The eigenface system framework is tested in two distinct tasks: face recognition and face verification. To perform the tests, we acquired a face image database composed of 435 views of 102 people. No hard constraints were imposed to the database acquisition process what means we obtained large variations between images of same people. We also evaluate the system's tolerance against such variations. All the routines needed to the task of image processing, face recognition and face verification were written in C++. They compose a library of routines that can be adapted for other task such as automatic segmentation of facial images.

The face recognition performance is evaluated against the full database, 435 images, trying to classify each image in its correspondent class. Each class represents one person present in the database. The face recognition system achieved a recognition rate of 97.70%. The face verification system was tested with the full database, 435 images, using a minimum distance classifier varying its acceptance threshold.

Capítulo 1

1. Introdução

A capacidade de reconhecer faces desempenha um papel importante em nossas vidas. Esta capacidade é uma habilidade que desenvolvemos muito cedo afetando diretamente a maneira como nos relacionamos na sociedade. Entre seus aspectos mais marcantes está sua robustez em termos de desempenho. Podemos identificar uma pessoa sob as mais diferentes condições de observação tais como em imagens distorcidas, com pouca luminosidade ou observando apenas partes da face. Nossa habilidade natural de reconhecimento é capaz de tratar estas variações presentes na face com facilidade. No entanto, não é uma tarefa trivial traduzir esta habilidade humana para um sistema automático, mesmo com o auxílio do computador. De fato a compreensão dos mecanismos que empregamos no reconhecimento de faces constitui um desafio para psicólogos e neurocientistas. A compreensão do processo de reconhecimento facial realizado pelo ser humano é muito importante para tentarmos simular o modo como

realizamos as tarefas de reconhecimento via um computador. Existe muita informação sobre o modo como o cérebro humano realiza os processos de reconhecimento. Mesmo assim, implementações de modelos de reconhecimento ainda encontram-se em fase inicial de desenvolvimento. Uma maneira prática de desenvolvermos sistemas que possuam a capacidade de reconhecer faces humanas é por meio do processamento de imagens faciais bidimensionais. Um sistema de processamento de imagens faciais eficaz deve ser capaz de reconhecer um indivíduo extraindo apenas a informação contida na imagem, do mesmo modo que fazemos ao olhar uma fotografia de alguém conhecido.

Os sistemas que desempenham as tarefas de reconhecimento facial são classificados como sistemas biométricos. Os sistemas biométricos são aqueles que identificam pessoas através de suas características físicas ou comportamentais. Entre as características utilizadas hoje em sistemas biométricos podemos citar: a mão, impressão digital, íris, voz, face, corpo e hábitos comportamentais (maneira de escrever ou teclar). Através dos sistemas biométricos, dispomos de uma maneira bastante prática de reconhecer um indivíduo. Aonde quer que ele vá estará sempre levando consigo suas características, dificultando fraudes.

O reconhecimento de face está entre as técnicas biométricas mais confiáveis e menos invasivas. Normalmente, o indivíduo deve apenas ficar diante de uma câmera para que sua imagem facial seja adquirida pelo sistema de reconhecimento. No caso de reconhecimento de mão ele deve colocar a mão, em um sensor de varredura. Isto para algumas pessoas pode ser algo pouco agradável. Para avaliarmos os diferentes tipos de sistemas biométricos em comparação com o sistema baseado em faces, podemos observar a Tabela 1-1 [PENTLAND-a]. Os dados tabela mostram os resultados dos testes de verificação com sistemas biométricos sob condições controladas desenvolvidos até o ano de 1987. Após esta data ocorreram vários avanços na utilização prática dos sistemas biométricos mas as relações de desempenho entre os vários sistemas mostrados

na Tabela 1-1 praticamente se mantêm. Através desta tabela podemos notar o bom desempenho do sistema de reconhecimento facial, que mostra-se particularmente adequado para aplicações como controle de acesso. Neste caso, o sistema facial apresenta uma das menores taxas de falsa aceitação entre os sistemas analisados. Note que taxa de falsa aceitação é o parâmetro de maior importância na avaliação da confiabilidade de um sistema de controle de acesso e segurança.

	Perfil da Mão	Voz (AT&T)	Voz (Voxtron)	Padrão da Retina	Impressão Digital	Face
Frequência de Tentativas de Verificação	1491	3206	3082	3384	3384	100
Frequência de Tentativas de Fraude	4055	3415	3795	5027	4849	742599
Tempo de Aquisição (s)	54	18	144	126	114	10
Tempo de Verificação (s)	4.4	8.8	10.1	7.5	9.8	1-10
Rejeições Falsas	0.9 %	12.5 %	17.0 %	10.8 %	9.1 %	1.5 %
Aceitações Falsas	0.4 %	0.1 %	0.6 %	0 %	9 %	0.01 %

Tabela 1-1: Comparação do desempenho de sistemas biométricos, [PENTLAND-a].

A técnica de análise de imagens faciais em sistemas biométricos é empregada para resolver os problemas de reconhecimento e verificação. Por ser a verificação um caso particular do reconhecimento serão sempre feitas referências a métodos ou modelos de reconhecimento facial. Assim, o termo verificação apenas será explicitado quando nos referirmos à utilização de técnicas específicas ao problema de verificação. A seguir apresentaremos as definições explícitas destes dois problemas, reconhecimento e verificação, para o caso de sistemas biométricos que empregam imagens faciais.

No reconhecimento facial, a identidade do indivíduo a ser testado não é declarada ao sistema. O processo de reconhecimento facial efetua apenas a comparação

das características da imagem facial, sem possuir a priori informação da identidade da pessoa presente na imagem. O reconhecimento deve determinar a quem pertence a imagem ou rejeitá-la caso não esteja presente na base de dados. O esboço de um sistema básico de reconhecimento facial pode ser visto na Figura 1-1. A informação para identificação do indivíduo é armazenada na base de dados. O computador processa a informação extraída da imagem adquirida para realizar o reconhecimento. Na Figura 1-2 temos o diagrama de blocos do processo de reconhecimento facial. A aquisição da imagem facial por uma câmera é o primeiro passo do processo. Em seguida são realizados os pré-processamentos na imagem adquirida necessários à extração das características da imagem facial pelo sistema. A extração de características varia com o tipo de implementação do sistema de reconhecimento facial. Baseado nas características extraídas, o sistema realiza o reconhecimento e toma uma decisão.

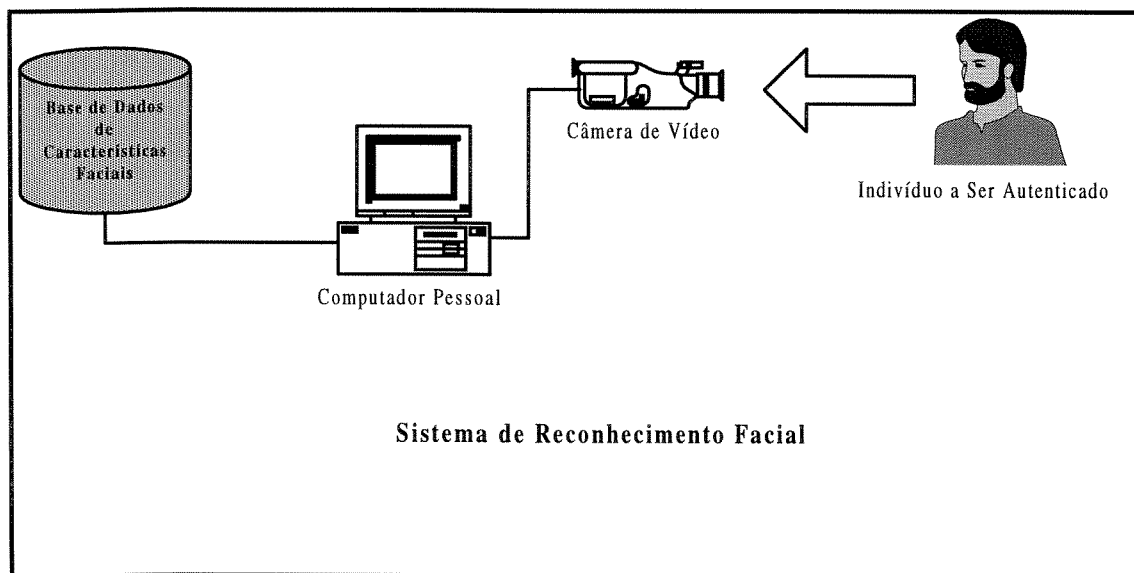


Figura 1-1: Componentes básicos de um sistema para reconhecimento facial automático.

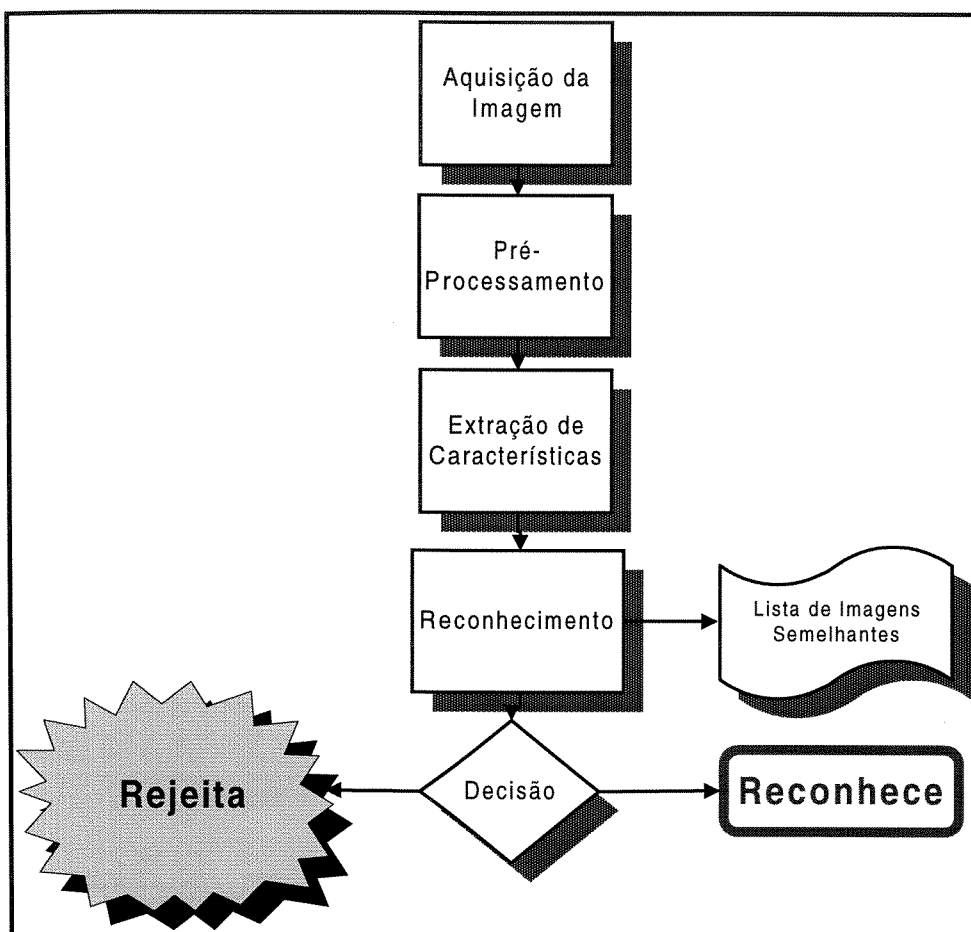


Figura 1-2: Diagrama de blocos da operação de um sistema de reconhecimento facial.

Na verificação facial, o sistema deve comparar as características extraídas da imagem facial com as características da face do indivíduo declaradas ao sistema. Normalmente, a pessoa declara sua identidade para ser autenticada através de um objeto tais como senha, nome, R.G. etc. A informação pessoal dada é então comparada à informação armazenada no banco de dados do sistema. Sua identidade é, então, validada ou não com base no resultado da comparação. Em geral, é possível estabelecer um limiar de aceitação. Este limiar expressa qual deve ser o valor mínimo de semelhança entre as características faciais extraídas para a identificação da pessoa com um certo grau de confiabilidade. Hoje é possível gravar todas as características que o indivíduo declara ao sistema para comparação com as extraídas da imagem facial em um cartão inteligente, evitando portanto um banco de dados centralizado. No caso do indivíduo declarar sua

identificação por meio de uma senha, as características faciais devem estar armazenadas em uma base de dados integrada ao sistema de verificação. Um esboço do sistema básico de verificação facial pode ser visto na Figura 1-3, onde são mostrados os componentes necessários ao processo de verificação: a câmera para a aquisição da imagem facial, um dispositivo para a entrada da identificação do indivíduo e o computador para processar os algoritmos necessários à verificação. Este esquema só difere do esquema para o processo de reconhecimento na Figura 1-1, pelo dispositivo onde indivíduo declara sua identificação ao sistema (entrada de informação via teclado ou leitora de cartões). No diagrama de blocos do processo de verificação na Figura 1-4, temos as etapas do processo de verificação de uma imagem facial. O processo de verificação começa com a aquisição da imagem facial e a identificação do indivíduo. Em seguida, realiza-se o pré-processamento e extração das características para o processo de verificação. Finalmente, é efetuada a decisão de aceitar ou rejeitar a identificação apresentada ao sistema. O diagrama de Figura 1-4 difere do diagrama da Figura 1-2 apenas no bloco que efetua a verificação. Este bloco toma a decisão de aceitação ou rejeição baseado na comparação da informação declarada ao sistema com a extraída da imagem facial. No caso do reconhecimento, o sistema apenas procura na base de dados a imagem que mais se assemelha à imagem de entrada.

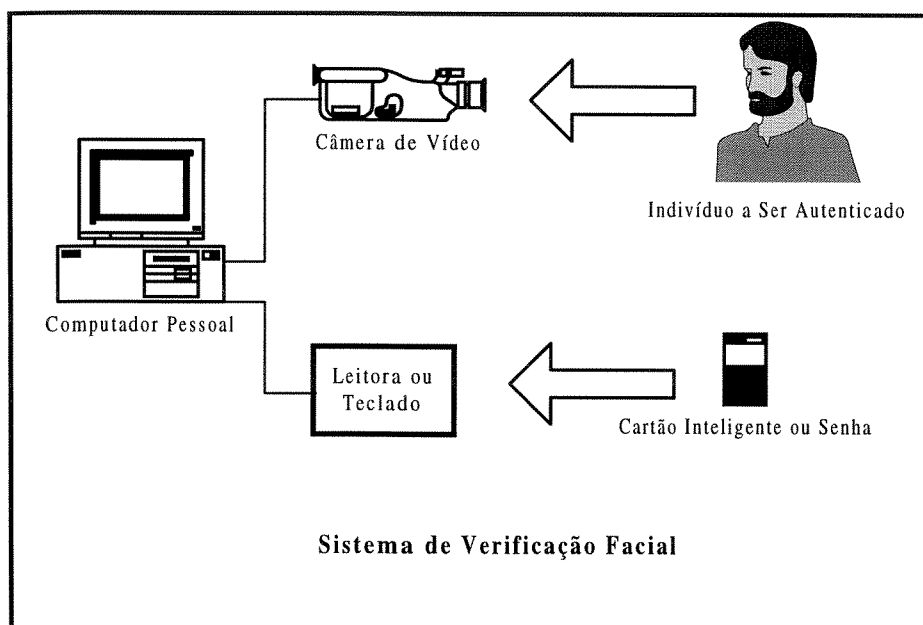


Figura 1-3: Componentes básicos de um sistema para verificação facial automática.

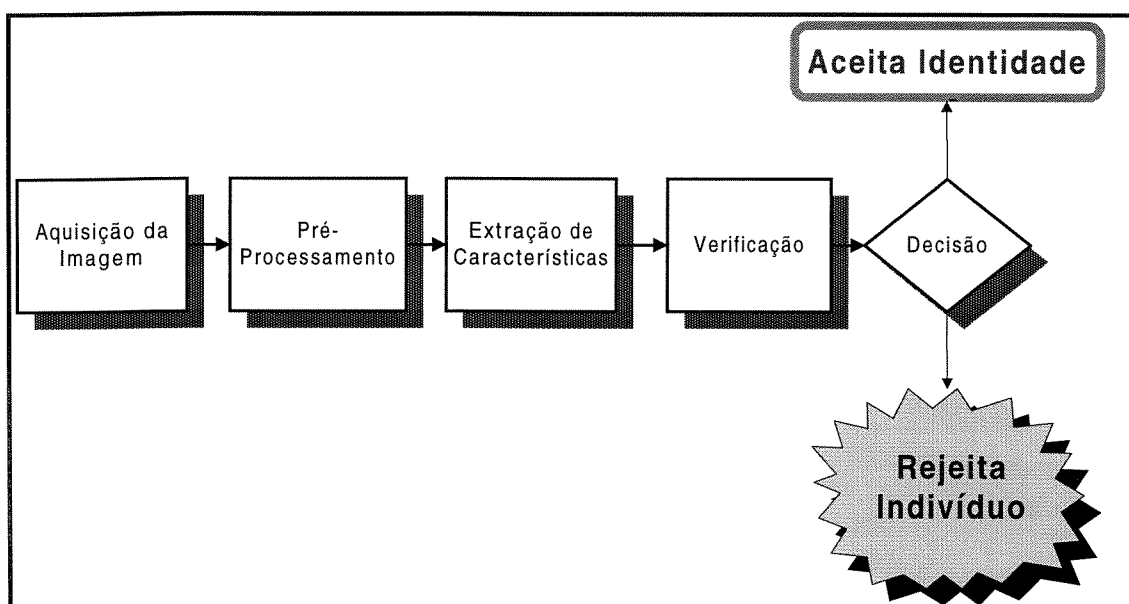


Figura 1-4: Diagrama de blocos da operação de um sistema de verificação facial.

Segundo [CHELLAPA], existem muitas aplicações práticas para um sistema que possua características de reconhecimento facial automático e robusto. Entre as principais aplicações previstas atualmente, podemos citar:

- tarefas gerais de segurança, como controle de acesso a prédios e instalações, verificando as faces captadas por câmeras de segurança;

- validação de transações bancárias e com cartões de crédito, verificando-se a imagem codificada no cartão magnético do cliente;
- tarefas policiais, implementação de banco de dados com imagens faciais de criminosos para reconhecimento através do uso de retrato falado eletrônico ou foto;
- identificação periódica dos usuários de redes de computadores, como forma de aumentar a segurança das redes corporativas, autenticando o uso dos números de identificação pessoal (*personal identification numbers* - PIN);

Além destas aplicações citadas por [CHELLAPA], o processamento de imagens faciais vem sendo empregado no desenvolvimento de interfaces homem-máquina, onde os sistemas tentam reconhecer e interpretar a face do usuário de um terminal de computador. Exemplos de tais sistemas podem ser vistos em [HASEGAWA] e [SAKO]. Esta é uma tendência que vem surgindo na indústria de computação como uma das formas de prover uma interação mais amigável na interface homem-máquina. Agora que sistemas de reconhecimento vocal se tornarão padrão em questão de poucos anos, os sistemas de reconhecimento facial estão entre as tecnologias que deverão receber uma maior atenção da indústria.

Na década de 90 tivemos uma renovação do interesse dos pesquisadores na área de reconhecimento facial. Este interesse foi motivado pelo surgimento de equipamentos capazes de realizar as tarefas de processamento de imagem em tempo real aliado com o grande potencial econômico das aplicações da tecnologia de reconhecimento facial. Um dos grandes interessados em sistemas biométricos são as instituições financeiras que perdem somas consideráveis anualmente devido ao uso fraudulento de cartões eletrônicos e cartões de crédito. Estas perdas devem-se à facilidade com que as identificações de cartões, senhas e assinaturas podem ser fraudadas. Assim estas instituições tem grande interesse em sistemas de identificação automática mais robustos que utilizem características biométricas [MILLER] como forma de validar a identificação

de um indivíduo, aumentando assim a confiabilidade do sistema.

Foi mostrado que o reconhecimento facial é uma técnica que possui um grande potencial em termos de aplicações práticas. É um método que tem boa aceitação por parte das pessoas, pois não exige muito esforço das mesmas na interação com o sistema. Todas estas razões motivaram este trabalho na direção de obtermos um sistema confiável para as tarefas de reconhecimento e verificação facial. Na próxima seção, é apresentada uma discussão geral dos aspectos do problema de reconhecimento de faces.

1.1 Problemas Básicos no Reconhecimento de Faces

O reconhecimento facial apresenta um problema realmente desafiador. A face é um objeto tridimensional e deformável, o que implica que ela não pode ser modelada de forma simples como objetos rígidos comuns tais como cones ou cubos em um sistema de visão robótica. Então, devemos nos restringir às variações mais comuns que uma face apresenta, construindo modelos faciais apenas para determinadas posições ou expressões da face. Estes modelos são a base de comparação do sistema de reconhecimento. A escolha de determinados modelos de face é uma maneira de reduzir a complexidade da tarefa de reconhecimento facial.

Um método prático de modelar a informação facial para um sistema de reconhecimento é trabalharmos com imagens bidimensionais da face. Em geral, as imagens faciais usadas para construir os modelos de reconhecimento podem ser bem diferentes das que serão processadas pelo sistema no processo de reconhecimento ou verificação. Em princípio o sistema de reconhecimento facial deve ser capaz de prover uma compensação para as variações presentes na imagem facial, permitindo um certo grau de liberdade nas condições de aquisição das imagens de entrada do sistema. Isto deve proporcionar ao usuário um maior conforto na utilização do sistema de reconhecimento facial. As principais variações na imagem facial que o sistema deverá

tratar são devidas a iluminação, posição da face, expressão facial, estado emocional, idade, maquiagem, presença ou ausência de óculos e barba. O tratamento das variações na imagem facial é uma tarefa bastante complexa. Mesmo seres humanos quando realizam a identificação de suspeitos em álbuns fotográficos cometem erros com imagens adquiridas em condições controladas. Somadas às dificuldades impostas pelas mudanças das condições de visualização, os métodos computacionais para reconhecimento de face têm que se confrontar com o fato de que as faces são muito semelhantes entre si. Esta é outra fonte de erros de classificação que dificultam ainda mais a tarefa de reconhecimento.

Os sistemas demonstrados na literatura, [CHELLAPPA], [VALENTIN], [SAMAL], [BEYMER], [SUNG] e [ZHANG], apresentam um bom grau de liberdade em relação a alguns parâmetros com os quais a imagem facial pode variar em relação ao modelo. Mas os sistemas demonstrados não tratam igualmente todas as variações que a face pode apresentar. As restrições com relação as variações tratadas se adequam à finalidade para a qual o sistema é destinado. Algo que todos têm em comum é a necessidade de uma grande quantidade de exemplos de faces, ou de não-faces, sob diferentes condições para caracterizar o modelo facial de forma a tratar com as variações presentes na face. É importante observar que a informação obtida com a variação dos parâmetros das imagens faciais é crítica para o desempenho de qualquer sistema. A qualidade da informação obtida afeta diretamente a performance do sistema de reconhecimento. A obtenção de um sistema de reconhecimento facial robusto em relação a variação de parâmetros presentes na imagem é um problema ainda em aberto.

No projeto e implementação de um sistema capaz de realizar a tarefa de reconhecer ou verificar faces humanas baseados em imagens faciais, há vários aspectos que devem ser considerados. Para identificar a face de uma pessoa, precisamos primeiramente localizar a face na imagem. Portanto, o primeiro passo dado pelo sistema

é detectar a presença de uma face e localizá-la. O próximo passo é extrair características da imagem facial. Por último, devemos identificar a face localizada com base nas informações extraídas da imagem. Existem mais duas características interessantes relacionadas ao contexto de discriminar o conteúdo de informação presente em uma face: a análise das expressões faciais e a classificação baseada em características físicas como sexo, idade e raça. Todos os problemas do reconhecimento de faces podem ser adequados à arquitetura clássica de um sistema de reconhecimento de padrões [SCHALLKOFF] mostrada na Figura 1-5. Nesta figura, observam-se três estágios: no estágio de aquisição, é efetuada a captura da imagem; o estágio de extração de características, onde a imagem adquirida é traduzida em termos da representação escolhida para caracterizar a informação da imagem facial; e por último, o estágio de classificação, onde as características faciais extraídas são comparadas e atribuídas a classes previamente definidas, que representam indivíduos presentes em uma base de dados.

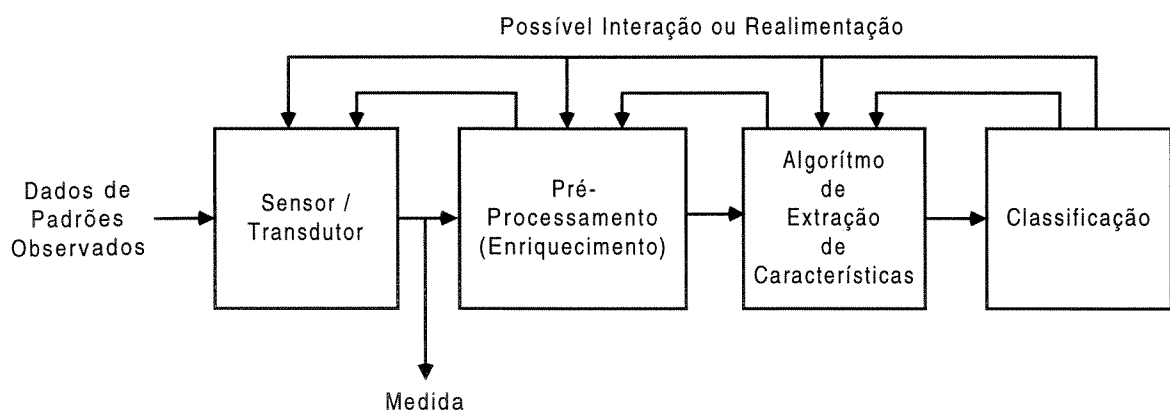


Figura 1-5: Diagrama de blocos de um sistema de reconhecimento de padrões clássico [SCHALLKOFF].

A seguir apresentamos maiores detalhes sobre a detecção, a representação e a identificação de faces, que formam os aspectos básicos do problema de reconhecimento de faces e sobre os problemas de análise de expressão e de características físicas.

1.1.1 Detecção de Faces

É a primeira tarefa a ser desempenhada por um sistema de reconhecimento ou verificação facial. Dada uma imagem de entrada devemos determinar se contém faces e quais suas posições. Este problema pode ser de fácil solução para o caso de imagens adquiridas em ambientes controlados, como faces em documentos de identificação como a carteira de identidade. Nestes casos as faces tem suas posições e tamanhos aproximados conhecidos ou facilmente determinados nas imagens. Para o caso de ambientes sem restrições nas condições de aquisição, a tarefa de determinar o tamanho e a posição das faces torna-se mais complexa. Neste caso temos um grande número de variações na imagem que devem ser tratadas no estágio de aquisição do sistema.

Esta área tem pouco destaque na literatura de reconhecimento facial em relação ao grande interesse dado aos modelos de representação da informação facial. Maiores detalhes sobre as técnicas utilizadas para a detecção de faces se encontram na revisão do Capítulo 2.

1.1.2 Representação de Faces

A representação das faces é uma das partes mais importantes em um sistema de reconhecimento de faces. A informação da face presente na imagem deve ser obtida em termos de medidas ou abstrações que possam ser processadas em um computador. A representação determina quais informações relevantes aos processos de reconhecimento o modelo deve extrair da imagem.

A busca de modelos de representação para a informação presente em imagens faciais é uma das áreas que recebe maior destaque na pesquisa de reconhecimento facial. Uma descrição dos diversos métodos empregados para representação encontra-se no Capítulo 2.

1.1.3 Identificação de Faces

Após a detecção e representação das faces, codificadas em termos das características do modelo de representação, o próximo passo do sistema é a identificação da face detectada na imagem. As faces conhecidas e suas características devem estar armazenadas em uma base de dados. O objetivo é casar a face na imagem com as faces presentes na base de dados. Podem ser usadas abordagens holísticas, aquelas que tratam com a imagem facial como um todo, ou baseadas em características locais da face como olhos, nariz e boca. Complicações adicionais ao processo de identificação podem advir de mudanças em uma face e em suas características com o passar dos anos. A escolha do método de comparação de características é baseada na natureza dos modelos de representação. Trataremos com maiores detalhes sobre a identificação de faces na descrição dos modelos faciais no Capítulo 2.

1.1.4 Análise de Expressão Facial

Este item, análise de expressão facial, e o próximo, classificação baseada em características físicas, não são necessariamente componentes de um sistema de reconhecimento de face mas são muito desejáveis para permitir uma interação mais amigável entre o sistema e seus usuários, além de adicionar uma maior capacidade de generalização ao sistema. A análise de expressões faciais, baseada apenas na informação contida na imagem, permite determinar se a pessoa está rindo, triste, surpresa ou zangada. O problema da análise de expressões é como modelar as emoções humanas e correlacioná-las com as características faciais e suas mudanças. A modelagem das emoções humanas não possui um tratamento exato pois mesmo seres humanos não são capazes de identificar corretamente algumas demonstrações emocionais. A análise de expressão facial não é coberto neste trabalho. Maiores informações podem ser encontradas em [ARAD], [ESSA], [MATSUNO], [YACOOB-a] e [YACOOB-b].

1.1.5 Classificação Baseada em Características Físicas

Uma outra tarefa de análise da informação facial que realizamos sem esforço é olharmos para uma face e determiná-la como masculina ou feminina, se pertence a uma pessoa jovem ou idosa e qual seu grupo étnico. A análise destas características é desejável em um sistema automático de processamento facial pois permite um maior grau de confiabilidade do sistema para discernir entre estas características e suas variações com o passar dos anos, aumentando sua robustez. Não abordamos neste trabalho a classificação baseada em características físicas, maiores informações podem ser encontradas em [ABDI-b], [ABDI-c], [VALENTIN-a] e [WISKOTT-b].

1.2 Trabalho Desenvolvido

Esta tese tem como objetivo estudar as técnicas de reconhecimento facial e sua aplicação a sistemas biométricos automáticos para identificação pessoal. Para implementar um sistema biométrico que pudesse ser empregado em reconhecimento e verificação pessoal, escolhemos o método de *eigenfaces* (veja Capítulo 3) para compor o núcleo do modelo de representação das características faciais do sistema.

O método foi avaliado em relação às variações que ele poderia admitir nas imagens de entrada com mudanças na posição, expressão, presença ou ausência de óculos ou barba, provendo um desempenho confiável para o reconhecimento das faces.

Na análise da aplicação do método à tarefa de verificação individual, observamos o desempenho do modelo em termos de suas taxas de falsa aceitação e falsa rejeição para ajustar o sistema de verificação à sua melhor performance.

Os procedimentos de extração de características, reconhecimento e verificação estão implementados de forma totalmente automática. As normalizações e as tarefas de processamento de imagem associadas à normalização também estão automatizadas. Apenas o estágio de segmentação para a normalização requer entrada de dados dos

usuários do sistema.

1.3 Estrutura do Trabalho

No Capítulo 2 realizamos uma revisão das técnicas utilizadas e desenvolvidas, até o presente, em termos de reconhecimento de faces e dos problemas relacionados, tais como a segmentação e a normalização da imagem facial.

Dentre os métodos descritos na literatura escolhemos o método de *eigenfaces*, descrito no Capítulo 3 para iniciarmos as análises dos sistema propostos e, posteriormente, testar as restrições dos vários métodos em relação a sua utilização em sistemas automáticos.

No Capítulo 4, estão descritas a construção da base de dados de imagens faciais e os resultados obtidos com o método de *eigenfaces* na representação da informação facial do conjunto de imagens.

No Capítulo 5, efetuamos uma análise dos resultados do método de *eigenfaces* nos problemas de reconhecimento e verificação facial.

No Capítulo 6, são apresentadas as discussões e conclusões finais sobre o trabalho desenvolvido e sugestões para futuros trabalhos na área de reconhecimento automático de faces.

Capítulo 2

2. Processamento de Imagens Faciais

Neste capítulo apresentamos uma revisão das várias técnicas presentes na literatura de reconhecimento de face. Entretanto, esta revisão não é exaustiva contando apenas com os trabalhos julgados mais relevantes na área, principalmente sobre os problemas de detecção de face em imagens e representação da informação facial. Uma fonte mais abrangente de referências pode ser encontrada nas revisões descritas em [SAMAL], [VALENTIN-a] e [CHELLAPPA].

2.1 O Reconhecimento Facial

O reconhecimento de faces por computadores, em fotografias ou seqüências de vídeo, está emergindo como uma área de pesquisa bastante ativa abrangendo várias disciplinas tais como: processamento de imagens, reconhecimento de padrões, visão computacional e redes neurais. O interesse no reconhecimento de face se deve ao fato das tecnologias de reconhecimento de face serem de fundamental importância em várias

aplicações práticas, tanto em campos comerciais como em campos criminais. Estas aplicações abrangem desde o casamento de fotografias com formato controlado, por exemplo passaportes, cartões de crédito e fotos de documentos de identidade, até o processamento em tempo real de imagens de circuitos fechados de vídeo para segurança. As diversas aplicações dos sistemas de reconhecimento de face impõem uma série de restrições para sua operação. Aplicações em segurança, por exemplo, devem contar com um grande grau de precisão na aquisição de imagens em condições não controladas e ser capazes de operar compensando as variações presentes nas imagens de entrada. Aplicações de procura por conteúdo presente na imagem em bases de dados de imagens faciais devem ser capazes de discernir entre uma grande quantidade de faces. Estas faces são sempre muito semelhantes entre si. Mesmo assim o sistema de pesquisa de imagens por conteúdo deve distinguir uma face particular das diversas outras presentes no banco de dados. A aplicação do reconhecimento facial realizado com fotografias ou seqüências de vídeo, apresenta os seguintes procedimentos básicos comuns a todos os sistemas de processamento de imagem facial:

- Pré-Processamento:
 - Segmentação da Imagem Facial, localizando e extraíndo a área facial relevante na imagem;
 - Normalização em relação às condições de iluminação e melhoramento da imagem adquirida;
- Extração de Características da Imagem Facial, segundo um modelo de representação das características consideradas discriminantes;

Estes procedimentos devem ser tratados pelo sistema facial para compensar as variações que as imagens apresentam. Existem várias abordagens propostas para a realização destes procedimentos. Nas seções seguintes faremos uma análise das técnicas descritas na literatura de processamento de imagens faciais abrangendo as áreas de

segmentação da imagem, normalização, modelagem e extração de características.

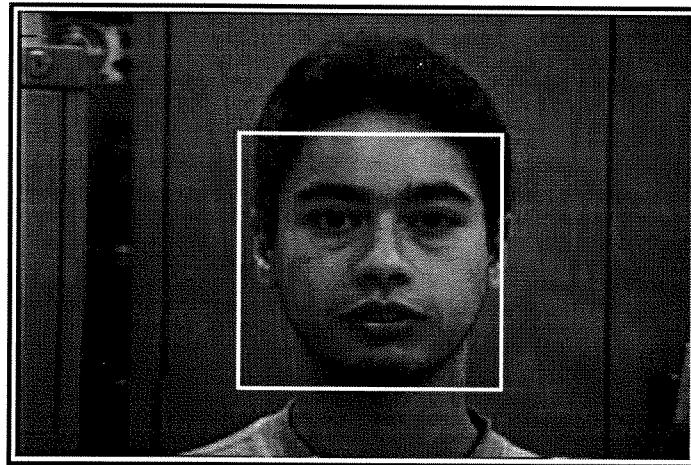


Figura 2-1: Exemplo de codificação da posição e área da face na imagem de entrada.

2.2 Detecção de Faces em Imagens

O problema de detecção de faces em imagens é definido na literatura de processamento de imagens como um problema de segmentação da imagem. Na fase de segmentação é que se separa a área de interesse para o processamento do restante da imagem. Na detecção de faces é determinado se existem faces e onde elas se localizam na imagem para então realizar a extração das faces detectadas. Se existem faces presentes na imagem os procedimentos de detecção e extração devem retornar respectivamente a posição de cada face e a área da imagem por estas ocupada. A informação de posição e área da face pode ser codificada por exemplo na forma dos vértices de um retângulo dentro do qual se encontra a face, conforme mostrado na Figura 2-1.

Os esforços iniciais na área de reconhecimento facial se dedicaram ao reconhecimento e verificação de faces em ambientes controlados (e.g. fotos de passaporte) [CHELLAPPA]. Entretanto, tornou-se necessário o desenvolvimento de técnicas que possam extrair uma face de uma foto ou sequência de imagens sob condições não supervisionadas, o que levou a abordagens que toleram grandes variações nas condições de aquisição para o procedimento de extração das imagens faciais como as

propostas em [SUNG] e [BUREL].

As soluções propostas para o problema de detecção de faces vão desde do clássico casamento de padrões, passando por padrões deformáveis, *snakelets*, redes neurais, cores características da face, texturas até *eigentemplates*. Entre todas as diferentes abordagens para o problema de segmentação da imagem facial temos compromissos em relação a capacidade de generalização, precisão e tempo de processamento. A seguir consideraremos as diferentes abordagens descritas na literatura para o problema de detecção de faces.

2.2.1 Casamento de Padrões

Várias técnicas empregadas na segmentação da imagem facial se enquadram no escopo do casamento de padrões. Os detalhes de implementação e de como se processa a informação podem variar, mas basicamente todas as técnicas visam a comparação da imagem com o padrão do objeto que se deseja extrair.

O padrão pode variar muito de acordo com a técnica empregada. Podemos ter como padrão do objeto procurado uma imagem, um modelo paramétrico ou outras características tais como cores ou texturas. A seguir, descrevemos as várias técnicas, dentro do escopo de casamento de padrões, descritas na literatura de processamento de imagens faciais. Estas técnicas são aplicadas ao problema de detecção da face e das características faciais, como olhos, nariz e boca, em imagens bidimensionais.

2.2.1.1 Correlação de Padrões

Esta é a técnica clássica de detecção de um objeto em uma imagem. Os padrões de correlação fixos desempenham a função de filtros casados. Eles calculam a diferença entre o padrão alvo pré-definido e as possíveis localizações deste padrão alvo na imagem, fornecendo como resultado uma medida do casamento do padrão alvo em cada ponto da imagem. Em seguida aplica-se um limiar de decisão na imagem resultante para obter as

possíveis áreas de casamento com o padrão alvo. Os resultados obtidos com esta técnica estão sujeitos a erros quando existem variações significativas em escala ou posição entre o padrão alvo e o objeto relacionado com o padrão que está sendo procurado dentro da imagem.

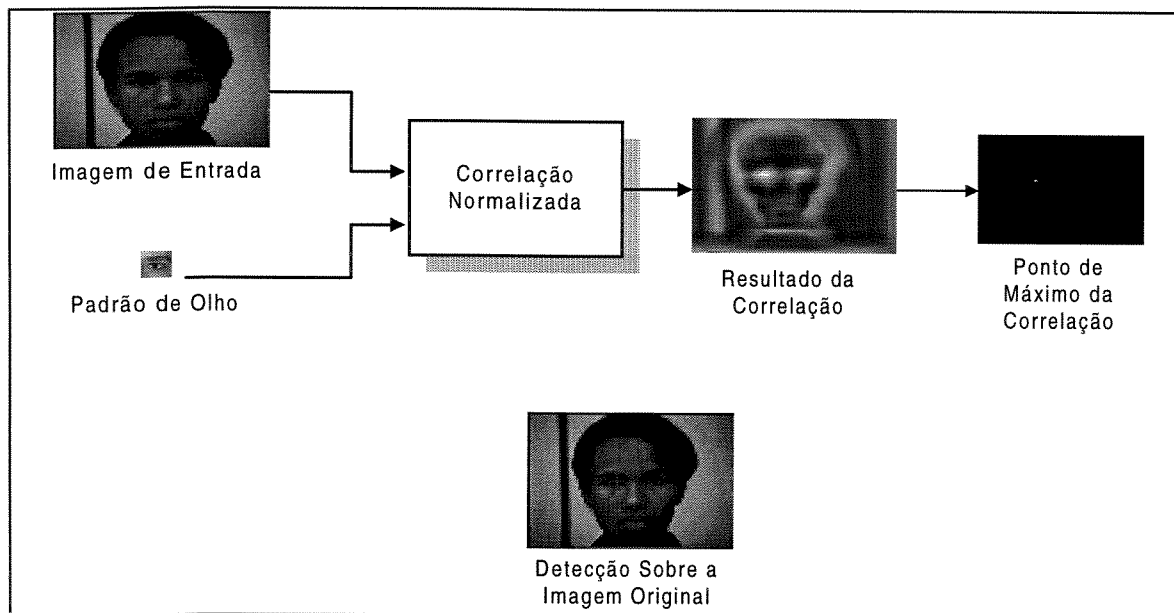


Figura 2-2: Detecção e extração das características faciais da imagem de entrada.

2.2.1.2 Correlação de Padrões Normalizada

A correlação de padrões normalizada, uma variante da técnica de correlação de padrões que permite um grau maior de deformação entre o padrão e a imagem a ser processada, é descrita em [BEYMER]. Nesta abordagem, tenta-se casar a energia do padrão com a energia das possíveis áreas da imagem que venham a conter este padrão de maneira que suas médias e variâncias se casem. Na Figura 2-2 temos um exemplo de teste da correlação normalizada. Neste caso, estamos tentando localizar o olho direito do indivíduo na imagem, dado um padrão de olho extraído de outro indivíduo. Vemos que o resultado da detecção do olho direito sobre a imagem original é satisfatório pois se aproxima da posição do olho direito na imagem original. A técnica de correlação

normalizada permite ainda a correlação dos padrões em multiescala. A correlação de padrões normalizada em multiescala consiste em correlacionar com a imagem uma série de padrões com diferentes dimensões. Devido ao fato da correlação ser normalizada podemos comparar os valores resultantes da correlação dos padrões de diferentes dimensões. A dimensão do padrão que possuir maior valor de correlação normalizada é atribuída como sendo a dimensão correspondente ao objeto detectado. Podemos então usar a correlação normalizada em multiescala para localizar um determinado objeto e determinar aproximadamente sua área na imagem.

2.2.1.3 Padrões Deformáveis

São similares em princípio às técnicas de correlação de padrões clássicas, exceto pelo fato de que os padrões possuem alguns componentes internos que são parametrizados. Em uma das abordagens dos padrões deformáveis, descrita em [XIE], temos curvas e superfícies parametrizadas para modelar os elementos não rígidos da face e das características faciais (olhos, nariz e boca). O processo de casamento tenta alinhar o padrão com uma ou mais versões da imagem de entrada pré-processadas, como os vales, picos e bordas da imagem original. A análise se torna um processo de minimização funcional, atraindo as curvas parametrizadas para as características correspondentes na imagem, enquanto penaliza as deformações acentuadas que o padrão sofre por se situar sobre uma área onde não se encaixa. O melhor encaixe se dá no mínimo desta função de deformação que serve como uma indicação da área a ser extraída.

Outra abordagem de padrões deformáveis utiliza *snakelets* [KWON] para realizar a extração dos contornos da face. Estes desempenham também uma minimização funcional tentando obter o melhor casamento dos elementos ativos do processo, *snakes*, com os contornos da face. No final do processo de minimização funcional da energia de deformação dos *snakelets* obtemos a região aproximada que descreve o contorno da face.

2.2.1.4 Casamento de Cores

Também dentro da abordagem de correlação pode-se tentar correlacionar o padrão das cores presentes na face humana com as cores da imagem de entrada. Exemplos da aplicação da informação de cor no casamento são descritos em [CHEN] e [HUNKE]. O padrão possui as cores características presentes na face humana, com a correlação deste padrão com o espaço de cores da imagem de entrada localizam-se as áreas correspondentes na imagem onde estas cores estão presentes. Esta é uma técnica rápida e utilizada para o rastreamento de faces em tempo real em ambientes parcialmente controlados proposta em [HUNKE].

2.2.1.5 Casamento de Texturas

A técnica casamento de texturas desenvolvida por Daí e Nakano em [DAI-a] usa como padrão uma matriz de níveis de cinza obtida com imagens de treino. Estes níveis representam as variações de textura que uma face pode possuir na imagem. Esta representação é bastante robusta obtendo boa performance sob condições variadas de aquisição da imagem. Os experimentos realizados por Daí e Nakano e [DAI-b] com esta técnica demonstram sua capacidade de localizar faces em ambientes complexos com uma grande precisão. Estes experimentos mostram que o modelo de texturas extraído das faces de treino é uma técnica que deve ser considerada no projeto de um sistema de detecção facial.

2.2.2 Redes Neurais

Os sistemas de detecção de face com redes neurais se baseiam no treinamento de perceptrons multi-camada (*multi layer perceptrons*) (MLP), através do algoritmo de *backpropagation* [HAYKIN]. Estes sistemas trabalham diretamente sobre a imagem de entrada em níveis de cinza, com uma janela de entrada para a rede que percorre a imagem detectando a presença ou ausência de faces dentro desta janela.

No treinamento dos MLPs são necessários vários exemplos para compensar as variações da imagem de entrada e permitir uma boa capacidade de generalização à rede. São fornecidas para a rede amostras de padrões de faces e de não-faces em diferentes escalas, como em [VAILLANT], ou a imagem é processada para uma escala padrão, como em [SUNG]. O sistema proposto em [BUREL], além de treinar uma os MLPs com uma grande variedade de dados e fazer as necessárias normalizações na imagem de entrada, utiliza uma rede de Kohonen para comprimir a informação facial extraída da imagem para treinamento dos MLPs.

Os métodos com redes neurais apresentam um bom desempenho, superior em muitos casos às técnicas de correlação. A única limitação dos métodos com redes neurais para a detecção de faces é a quantidade e a qualidade dos exemplos de faces a serem fornecidos à rede neural para obtermos uma bom desempenho. Se não possuímos uma quantidade suficiente de amostras do tipo de faces que queremos que a rede aprenda a identificar, não podemos garantir que a rede treinada com o algoritmo de *backpropagation* possua uma boa capacidade de generalização.

2.2.3 Eigentemplates

A técnica de detecção de faces com *eigentemplates* utiliza a abordagem descrita por Turk e Pentland em [TURK] e empregada por Moghaddam em [MOGHADDAM-d] para localizar as áreas de interesse na imagem. Esta técnica de detecção realiza a correlação dos padrões chamados de *eigentemplates* com a imagem. A detecção com *eigentemplates* não difere muito das outras abordagens de correlação em sua aplicação. Entretanto com o modelo de *eigentemplates* temos um banco de filtros estocásticos que são aplicados às áreas extraídas da imagem para detecção da face. As *eigentemplates* formam um banco de filtros estocásticos pois elas possuem a informação da distribuição da variação de intensidade dos pixels de um conjunto de imagens faciais de treino. Cada área

extraída é correlacionada com uma série de *eigentemplates*. Os resultados de cada correlação são adicionados para formar um mapa que possui o valor de correlação das *eigentemplates* em cada ponto da imagem. Então, depois de realizadas as correlações com todos os padrões, procura-se o ponto de mínimo deste mapa como sendo o ponto de casamento do padrão. Para uma descrição detalhada do modelo de *eigenfaces* em que se baseiam as *eigentemplates* veja o Capítulo 3.

2.3 Normalização das Imagens Faciais

As faces podem ser visualizadas como objetos tridimensionais, que ao serem processadas pelo sistema, são projetados em uma imagem bidimensional. Esta imagem pode não estar semelhante em termos de iluminação, translação, escala e rotação com a imagem do mesmo indivíduo armazenada na base de dados para efeito de comparação. Para compensar as variações devidas a translação, escala e rotação no plano da imagem e diminuir a variabilidade dos padrões sendo comparados, podemos normalizar a imagem facial detectada adotando uma distância interocular padrão e uma posição padrão para os olhos nas imagens processadas pelo sistema. Esta técnica está proposta em [PENTLAND-b] e [TURK]. Deste modo, após a detecção dos olhos na imagem facial, pode-se escalar, rotacionar e transladar a imagem de modo que ela se encontre na posição padrão e na escala padrão das imagens da base de dados, reduzindo a dimensionalidade do problema de reconhecimento. A normalização em relação à rotação compensa apenas as variações de posição no plano da imagem, as rotações da cabeça fora do plano da imagem não podem ser tratadas neste estágio do sistema pois a imagem em si não contém informação suficiente para resolver o problema. A tarefa de tentar compensar a rotação fora do plano da imagem deverá caber ao estágio de reconhecimento. Este estágio do sistema deverá possuir a informação do modelo facial que corresponde a uma dada rotação fora do plano da imagem.

A compensação da variação das condições de iluminação segundo [BRUNELLI] pode ser realizada utilizando-se um filtro de média na imagem. Efetuando a média na vizinhança de um pixel da imagem que seja grande o suficiente podemos diminuir os efeitos da variação da luminosidade. Além das filtragens pode-se normalizar a imagem para que possua média nula e variância unitária [VAILLANT] e [MOGHADDAM-c]. Para eliminar a contribuição de diferentes fundos da imagem no processo de reconhecimento aplica-se também uma máscara gaussiana centrada na face no processo de normalização [PENTLAND-c]. O esquema das etapas de normalização que combina as técnicas descritas anteriormente pode ser visto na Figura 2-3. Antes da normalização em si temos a detecção da face e de seus pontos característicos, no caso os olhos. Aplica-se a normalização geométrica (escala, translação e rotação) seguida da normalização em relação à iluminação. O último passo do processo é a aplicação da máscara gaussiana para ressaltar o centro da face e diminuir a contribuição do fundo da imagem. A máscara gaussiana enfatiza a região dos olhos que é uma das mais importantes para o reconhecimento.

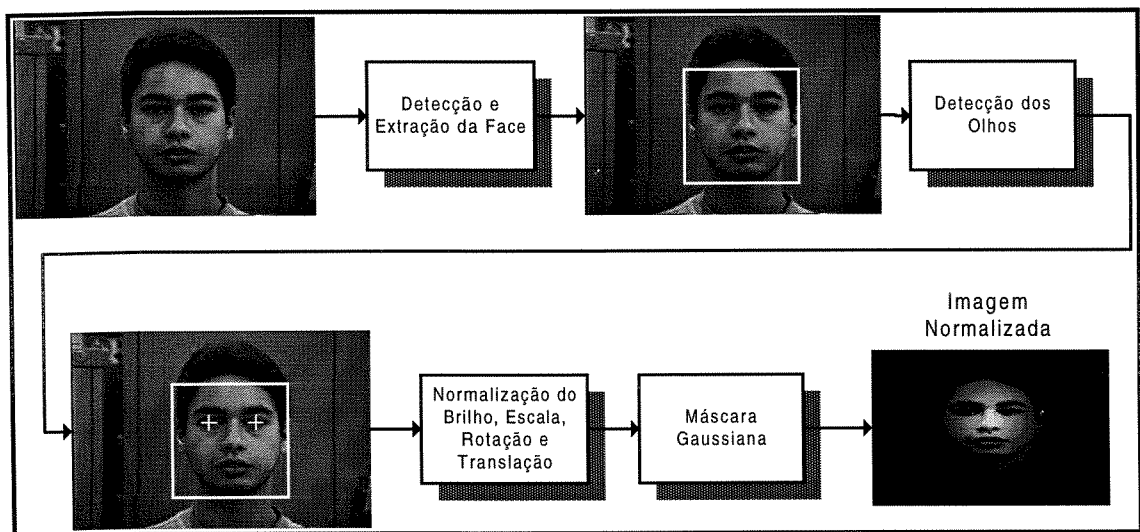


Figura 2-3: Esquema das etapas de um procedimento de normalização facial.

2.4 Extração de Características e Reconhecimento Facial

Existem várias tentativas de traduzir o modo como processamos e armazenamos as faces em implementações computacionais de modelos de reconhecimento facial. Qualquer sistema que consiga implementar esta analogia com a nossa capacidade natural possuirá um bom modelo de extração de características faciais [CHELLAPPA]. A modelagem pode ser obtida em termos de como nossa memória entende e armazena a face ([VALENTIN-a], [TURK] e [ZHANG]) ou em termos do modo que o sistema visual humano processa a face ([TACÁKS], [WISKOTT-a] e [WÜRTZ]). Atualmente estas duas abordagens de modelamento se encontram sob desenvolvimento por caminhos separados. No entanto, a tendência do desenvolvimento de modelos de reconhecimento facial é que haja uma interação destas abordagens para tentar obter uma maior robustez no processo de reconhecimento como um todo. Existem ainda codificações da imagem facial para extração de características que não possuem analogia com o modo como processamos a face. Exemplos disto são os coeficientes de autocorrelação descritos em [GOUDAIL] e os antigos trabalhos de medidas de distâncias sobre a face descritos em [SAMAL]. Entretanto os métodos que tentam aproximar o modo como processamos as imagens faciais tendem a ter um desempenho bem superior.

Como em todo o procedimento de reconhecimento de padrões, as várias abordagens de reconhecimento de face precisam aprender a caracterizar a face humana. Isto se realiza na etapa da construção do sistema, chamada de treinamento ou modelagem, onde são apresentados exemplos faciais ao modelo. Entre as abordagens empregadas para desempenhar a tarefa de reconhecimento de padrões faces temos: as que empregam redes neurais ([VALENTIN], [TACÁKS], [WISKOTT-a] e [WÜRTZ]); as que utilizam a análise de componentes principais das imagens faciais baseadas no modelo de eigenfaces ([TURK]) ou que realizam análise discriminante sobre este modelo ([SWETS]); as que utilizam cadeias de Markov, [SAMARIA-a] e [SAMARIA-b], dos

pixels da imagem; e as que empregam correlação de padrões para a codificação da imagem [GOUDAIL]. Estas são abordagens baseadas na análise da imagem como um todo. Estas abordagens diferem das primeiras aproximações do problema de reconhecimento facial que se baseavam em medidas realizadas sobre a imagem da face do indivíduo descritas na revisão de [SAMAL]. A seguir temos uma descrição dos vários modelos de extração de características presentes na literatura.

2.4.1 Redes Neurais

A área de Redes Neurais é uma das que mais tem atraído a atenção e os esforços dos pesquisadores envolvidos com o reconhecimento facial nos últimos anos. Isto se deve ao fato de que os modelos neurais traduzem em seus métodos e arquiteturas a maneira como o nosso cérebro processa e armazena os estímulos recebidos pelo sistema visual. As redes neurais realizam uma analogia computacional do funcionamento dos sistemas neurobiológicos reais. Um exemplo desta analogia pode ser visto em [WURTZ] em que a rede modela o processo de como nosso cérebro reconhece objetos. Nas próximas seções será apresentada uma revisão das técnicas presentes na literatura de redes neurais aplicadas ao processamento de imagens faciais. Uma revisão parcial das técnicas até 1994 pode ser encontrada em [VAILANT]. Detalhes sobre a implementação das estruturas de redes neurais clássicas encontram-se em [HAYKIN].

2.4.1.1 Perceptrons Multi-Camada

Os perceptrons multi-camadas treinados com o algoritmo de *backpropagation* são atualmente uma das arquiteturas básicas de Redes Neurais. Os MLPs podem ser utilizados separadamente, em grupos ou combinados a outras arquiteturas de rede. Aplicados isoladamente no processamento de imagens faciais MLPs são treinados com exemplos de faces dos diferentes indivíduos ou classes que deverão ser reconhecidos pelo sistema. Operam diretamente sobre a imagem em níveis de cinza ou sobre regiões

extraídas da mesma (somente face ou olhos, nariz e boca) como descrito em [TRATOR]. Os MLPs em conjunto com outras arquiteturas de rede são usados como classificadores para as características extraídas da imagem pelas redes às quais suas entradas estão conectadas. Um exemplo desta aplicação pode ser visto em [LAWRENCE] onde os MLPs são conectados à saída de uma rede neural convolucional. A aplicação dos MLPs diretamente sobre a imagem em níveis de cinza em geral não fornece bons resultados em termos de performance de reconhecimento [LAWRENCE], pois não apresentam invariância a translação ou deformações locais na imagem.

2.4.1.2 Autoassociador Linear

O autoassociador linear é uma memória treinada para reconstruir as imagens que lhe são apresentadas durante a fase de treino. Após o treinamento ele conterá as informações necessárias para reconstruir as imagens das faces treinadas. O autoassociador linear pode então ser usado como um extrator de características para as novas faces projetadas em sua entrada. A abordagem da autoassociação linear se assemelha conceitualmente ao método de eigenfaces, descrito no Capítulo 3. O ponto em comum é que ambas as abordagens realizam a análise de componentes principais sobre as imagens faciais de treino. Uma descrição mais detalhada desta semelhança é mostrada em [VALENTIN-c].

Em [VALENTIN-b] são estudadas as propriedades do autoassociador linear aplicado ao reconhecimento facial. É analisado como as propriedades do autoassociador linear se comportam com variações na posição e expressão das imagens faciais. Como autoassociador linear, a rede neural aprende a armazenar os padrões relevantes para a representação e reconstrução da face, extraídos diretamente da imagem em níveis de cinza. Esta é uma técnica holística, utiliza a informação global diretamente da imagem em níveis de cinza. O autoassociador linear possui uma analogia com modelos conhecidos

do cérebro humano que desempenham a função de memória (memórias de correlação).

Os autoassociadores lineares também foram investigados como meios de obter uma certa robustez no processo de reconhecimento, para o caso de não termos apenas vistas frontais de uma face. Assim em [VALENTIN-b] o autoassociador linear foi treinado para armazenar e reconhecer diferentes vistas da face (frontal, perfil, rotação de 45° , etc.), com isto adquirindo a capacidade de discernir entre as diferentes posições presentes nas faces observadas. Esta propriedade de determinação da posição da face pode assim ser utilizada para determinar em que posição está a face de entrada. Pode também aumentar a performance e a robustez de um sistema reconhecimento, realizando a procura apenas na base de dados que contém exemplos de faces na posição determinada pelo autoassociador linear.

2.4.1.3 Redes de Compressão de Imagem

Em [VALENTIN-a] uma rede neural MLP é treinada para comprimir as imagens que lhe são apresentadas através do algoritmo de *backpropagation*. Esta rede possui três camadas de entrada, escondida e de saída. As camadas de entrada e saída possuem as mesmas dimensões que são iguais às da imagem de treino. A camada escondida possui uma dimensão menor e por este fato é empregada para realizar compressão da imagem facial, pois a rede é treinada para reproduzir a imagem de entrada na saída passando pela redução dimensional da camada escondida. As características extraídas para o reconhecimento são as que a rede aprendeu a comprimir durante o treinamento na camada escondida. As características obtidas, resposta de saída dos neurônios da camada escondida, denominam-se holons. Os holons são usados para treinar um classificador baseado em um MLP tendo um neurônio de saída para cada indivíduo a ser classificado e um neurônio que corresponde a não existência de classe semelhante na base de dados.

2.4.1.4 Mapas Auto-Organizáveis de Kohonen

Os mapas auto-organizáveis de Kohonen são geralmente empregados como redutores dimensionais onde tenta-se extrair a informação significativa de um espaço de dados de entrada para uma dimensão menor, codificando desta forma os dados presentes na imagem numa forma mais compacta para o processo de reconhecimento. [TAKÁCS] utiliza, sobre a imagem facial, um mapa auto-organizável de Kohonen, tentando imitar o comportamento do olho humano aos estímulos visuais. O mapa de Kohonen simula os campos receptores e a inibição lateral dos neurônios do olho humano sobre a imagem. As características do "olho" artificial, obtidas com a convolução das entradas do mapa de Kohonen sobre a imagem, são empregadas no processo de reconhecimento.

2.4.1.5 Casamentos de Ligações Dinâmico

O casamento de ligações dinâmico é uma técnica de redes neurais que aproxima a maneira como nossa memória desempenha as tarefas de associação e recuperação da informação. No trabalho de [WISKOTT-a] o casamento de ligações dinâmico é empregado junto com outras características do sistema visual humano. Nesta abordagem a extração de características modela o olho como uma série de filtros de Gabor, utilizando *wavelets* de Gabor, para caracterizar a forma como processamos informações de objetos em diferentes escalas e orientações. Estas informações são extraídas sobre pontos característicos da face que são conectados para formar um grafo sobre a face, onde os ramos contêm a distância entre os pontos característicos e os nós as características extraídas com o banco de filtros de *wavelets* de Gabor. Quando uma nova imagem é apresentada ao sistema tenta-se casar o novo grafo obtido com o armazenado na base de dados através de uma rede neural que realiza o casamento dinâmico de ligações (*dynamic link matching - DLM*).

Em [WISKOT-c] são demonstrados os resultados do modelo com 300 imagens

de diferentes indivíduos comparadas com outras 300 imagens dos mesmos indivíduos. São relatadas taxas de acerto no reconhecimento de 97,3%. Este método é robusto em relação a deformações presentes na imagem sendo atualmente empregado em um sistema comercial para identificação pessoal descrito em [KONEN].

2.4.1.6 Rede de Neural de Convolução

A rede neural de convolução consiste tipicamente de um conjunto de camadas, cada uma com um ou mais planos receptores. Imagens aproximadamente centradas e normalizadas são inseridas na camada de entrada. Cada unidade (neurônio) em um plano recebe entradas de uma pequena região nos planos da camada anterior. Os pesos (conexões) que formam o campo receptivo para um plano receptor são forçados a serem iguais para todos os pontos neste plano receptor. Assim cada plano receptor pode ser considerado como um mapa de características que possui um detector de características fixo que é convolucionado com uma janela local que percorre os planos receptores da camada anterior. São empregados usualmente muitos planos receptores em cada camada, permitindo assim a detecção de múltiplas características. Estas camadas são denominadas de camadas convolucionais. Entre as camadas convolucionais, há geralmente uma camada que realiza uma sub-amostragem e retira a média local para alimentar então a próxima camada, diminuindo assim a dimensão desta. A rede inteira é treinada com o algoritmo de *backpropagation*.

O trabalho de [LAWRENCE] possui uma estrutura neural híbrida com mapas auto-organizáveis de Kohonen, para redução dimensional, e redes neurais de convolução, para a extração de características. As redes convolucionais permitem um certo grau de invariância com o deslocamento utilizando três conceitos: campos receptivos locais, pesos compartilhados e sub-amostragem espacial. Estas redes, quando testadas com as imagens faciais da base de dados do Olivetti Research Center (*ORL database*), obtiveram

o melhor índice de desempenho até o momento para esta base de dados (3.8 % de erro em 200 imagens de teste, com 200 imagens de treino para 40 indivíduos). Esta base de dados foi construída por Ferdinando Samaria, para ser utilizada em seus experimentos com HMM descritos em [SAMARIA-a] e [SAMARIA-b] e atualmente é de domínio público.

2.4.1.7 Árvores de Redes Neurais

Esta é uma arquitetura de redes neurais que vem sendo investigada recentemente em uma tentativa de tratar as variações que uma imagem facial pode apresentar. Zhang e Fulcher em [ZHANG] propõem uma nova arquitetura de rede neural, baseada em árvores de grupos de redes neurais com conexões adaptativas entre os grupos da árvore. Os grupos de redes são treinados para classificar, determinadas características, como presença ou ausência de barba e óculos. Durante o treinamento da árvore de grupos de redes neurais, novos grupos de redes são adicionados à árvore de modo a caracterizarem de forma correta a informação presente nas imagens de treino para cada indivíduo. No final do treinamento, a representação para cada indivíduo é armazenada em um nó da árvore de grupos de redes. O caminho até o nó do indivíduo passa pelas árvores que contêm toda a informação discriminatória, como os ramos da árvore que mostram se ele possui ou não possui barba e se ele possui ou não possui óculos. Por exemplo, se ele possui barba e óculos, ele estará localizado na árvore em um ramo diferente daquele dos indivíduos que não possuem esta característica. A procura pelo indivíduo vai descendo na estrutura da árvore até identificar o indivíduo chegando ao nó da árvore onde sua identificação foi armazenada durante o treinamento.

Os grupos de redes usam então a estratégia de dividir e conquistar para obter a classificação. Por exemplo, verificam primeiro as características gerais da imagem (barba ou óculos) e depois vão descendo nos ramos da árvore até atingirem o indivíduo correto.

Este modelo é empregado para reconhecimento facial invariante a translação e apropriado para tarefas como a verificação de segurança em aeroportos. Na segurança de aeroportos, apenas um pequeno número de faces precisa ser distinguida de um grande conjunto de pessoas que trafegam por ali diariamente. Dentro desta aplicação, foram realizados experimentos com uma base de dados de 780 imagens, 10 imagens diferentes de 78 pessoas, onde 20 imagens foram escolhidas para treino e 760 para teste. Em condições de laboratório, o modelo não apresentou erros em reconhecer 80 imagens (10 imagens diferentes de 8 pessoas) do total de 780 imagens faciais.

2.4.2 Métodos Estatísticos

Os métodos estatísticos baseiam-se na análise de imagens utilizando-se técnicas estatísticas clássicas como a análise de componentes principais e análise discriminante. Procuram encontrar uma representação para as características discriminantes presentes nas imagens. O método de HMM (*Hidden Markov Models* - HMM) está baseado nos modelos das cadeias escondidas de Markov aplicados às imagens faciais.

2.4.2.1 *Eigenfaces*

Esta é uma abordagem que trata de otimizar a representação da informação presente em uma imagem facial. Sobre uma base de dados de imagens faciais, é realizada a análise de componentes principais da matriz de correlação das imagens de treino. As componentes principais obtidas geram o modelo de *eigenfaces*, descrito primeiramente em [TURK]. As características obtidas na projeção possuem a forma de faces, como pode ser visto na Figura 2-4, daí o nome de *eigenfaces*. No reconhecimento facial, as características são extraídas projetando-se as imagens de entrada sobre as *eigenfaces* obtidas, gerando um coeficiente para cada *eigenface*. Novas imagens são comparadas medindo-se as distâncias que separam estes vetores de características. Este é atualmente um dos modelos mais aceitos e testados na literatura (a base de dados de teste possui

mais de 7000 imagens faciais), permitindo um bom reconhecimento e uma grande tolerância a variações que podem ocorrer na imagem a ser testada.

O modelo foi estendido para reconhecer diferentes expressões [ESSA] e faces em posições muito diferentes da frontal [MOGHADDAM-c]. Permite ainda implementar um método de detecção de faces e características faciais [MOGHADDAM-a]. Em [TURK] mostram-se os resultados do modelo com uma base de dados de 16 indivíduos com várias orientações da cabeça, escalas e condições de iluminação. As imagens parecem idênticas exceto por pequenas variações na expressão facial, detalhes da face e posição. Para variações em iluminação, orientação e escala o sistema obtêm 96%, 85% e 64% de classificações corretas, respectivamente. Em [PENTLAND-b] são apresentados bons resultados com uma grande base de dados: 95 % de acertos em reconhecer 200 pessoas de uma base de dados com 3000 indivíduos (com um total aproximado de 8000 imagens). Não é possível, entretanto, generalizarmos estes resultados, pois na base de dados, as imagens são muito similares e foram obtidas sob condições controladas. Em [MOGHADDAM-a] bons resultados são obtidos com a base de dados FERET do exército americano. Apenas um erro de classificação é cometido para classificar 150 vistas frontais contra 150 indivíduos presentes na base de dados. O sistema faz uso extensivo do pré-processamento da imagem para normalizá-la em relação a escala, iluminação, translação, contraste e rotação. O modelo de *eigenfaces* também foi utilizado para indexar uma base de dados de imagens faciais por conteúdo, descrito em [PENTLAND-b], onde, dada uma imagem de entrada, são retornadas as imagens que mais se assemelham a ela segundo o modelo.

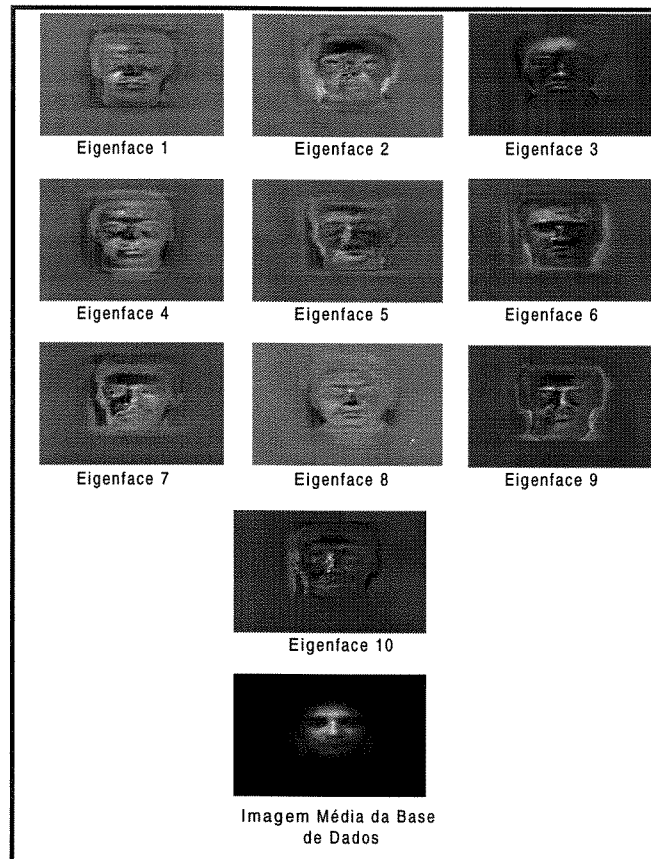


Figura 2-4: Exemplo de um modelo de eigenfaces para uma base de dados de 10 imagens frontais.

2.4.2.2 Análise Discriminante

O modelo de análise discriminante descrito em [SWETS] utiliza as teorias de projeção linear ótima para gerar uma divisão hierárquica do espaço definido pelas imagens de treino. Este espaço é gerado utilizando-se duas projeções: a projeção de Karhunen-Loève (análise de componentes principais) para produzir as chamadas características mais expressivas e uma subsequente análise discriminante desta projeção para obter as chamadas características mais discriminantes. O sistema de análise de características discriminantes implementa uma divisão do espaço de características geradas na análise discriminante. Esta divisão é feita com algoritmo de Voronoi e é usada para a classificação dos objetos, no caso imagens faciais, apresentados ao sistema. O

algoritmo de Voronoi divide o espaço de características entre as várias classes ou indivíduos que são apresentados como exemplos a serem reconhecidos pelo sistema.

No treinamento do modelo de análise discriminante foi utilizado um conjunto de treino de 1042 imagens de 384 indivíduos. Nos testes são usadas 246 imagens de 246 indivíduos. O sistema obtém uma taxa de acertos de reconhecimento de 95,5 % para os 246 indivíduos apresentados. Apesar de usar o método de *eigenfaces* como base da análise discriminante, o desempenho deste modelo não pode ser comparado ao método de *eigenfaces* pois não utilizam a mesma base de dados para testes.

2.4.2.3 HMM

Os modelos ocultos de Markov (*Hidden Markov Models* - HMM) foram utilizados com sucesso em áreas como a de reconhecimento de fala, onde uma palavra dita duas vezes pela mesma pessoa pode resultar em sinais substancialmente diferentes. Esta habilidade de reconhecer padrões que possuam grandes variações motivou o emprego destes modelos no reconhecimento de faces. Em [SAMARIA-b] encontramos a descrição da implementação do HMM, e os resultados obtidos para o reconhecimento de imagens faciais. Esta abordagem é baseada no casamento de padrões da imagem com uma cadeia de estados de um modelo estocástico duplamente conectado. Foram testados dois tipos de HMM o unidimensional e o bidimensional, obtendo respectivamente 90 % e 95 % de taxa de acerto para o reconhecimento. Empregaram-se 200 imagens para treino do modelo e 200 imagens para testes, com um total de 40 indivíduos (10 imagens por pessoa). As imagens utilizadas hoje compõem a chamada *ORL Database* disponível em domínio público.

2.4.3 Coeficientes de Autocorrelação

O trabalho de [GOUDAIL] investiga a utilização de 25 coeficientes de autocorrelação locais no reconhecimento de face. A forma dos núcleos de autocorrelação

é bastante simples. São matrizes de pixel $n \times n$ com um determinado padrão de pixels acesos ou apagados. As matrizes de pixels possuem várias escalas, iniciando com uma matriz de pixels de tamanho 3×3 . Os coeficientes de autocorrelação, que não exigem muito esforço computacional para serem determinados, possuem invariância ao deslocamento e são robustos em relação às mudanças na expressão facial.

Nos experimentos com este modelo utilizou-se uma base de dados de 11600 imagens, 100 imagens diferentes de 116 pessoas, divididas em dois conjuntos: um de 60 pessoas ditas conhecidas pelo sistema e o outro de 56 pessoas desconhecidas. Das 6000 imagens das pessoas conhecidas, metade foi usada para treinar o classificador que foi testado com as outras 5600 imagens. O sistema em multiescala obteve uma taxa de reconhecimento de 95 %, com uma taxa de falsa aceitação de 1,5 % de faces desconhecidas. Com rejeição das faces desconhecidas o desempenho de pico de reconhecimento do sistema foi de 99,9%. Dentre os sistemas descritos, este é o que apresenta maior número de exemplos para caracterizar uma única classe. Observamos também que as imagens usadas para treinar e testar o modelo não possuem grandes variações de posição e expressão facial.

2.5 Discussão

Observamos que há uma grande variedade de bons métodos que podem ser empregados para o reconhecimento de faces. Entretanto, a comparação direta dos vários métodos propostos em termos de taxas de classificações corretas não é adequada. As diferentes implementações em alguns casos diferem muito entre si, tanto conceitualmente, quanto na arquitetura do sistema. As condições de aquisição das bases de dados usadas para treinar e testar os sistemas é outro ponto que pode levar a resultados tendenciosos. Em alguns casos, elas possuem muitos exemplos de um mesmo indivíduo e poucos indivíduos a serem reconhecidos, em outros temos, um grande

número de indivíduos e pouca informação sobre cada um disponível. Ao menos o problema das bases de dados disponíveis tende a ser resolvido com o passar do tempo quando se estabelecerem bases de dados de referência internacionalmente aceitas e de uso público. Assim, as diferentes abordagens poderão ser realmente comparadas em termos de desempenho do conceito ou arquitetura implementada.

Neste trabalho, selecionamos o método de *eigenfaces* como modelo para a representação da informação facial. O método de *eigenfaces* é um dos mais amplamente testados e pode modelar uma grande gama de variações da imagem facial. Deste modo sua escolha como núcleo do sistema de reconhecimento provê um bom desempenho e a possibilidade de expansão das variações que podem ser tratadas com sucesso pelo sistema. A suas semelhanças conceituais com o autoassociador linear nos permitem também analisar de forma indireta o desempenho desta arquitetura de redes neurais. No Capítulo 3, descrevemos a concepção do método de *eigenfaces* para o processamento de imagens faciais.

Capítulo 3

3. O Método de *Eigenfaces*

O método de *eigenfaces* foi escolhido para compor o núcleo de nosso sistema de reconhecimento facial devido ao seu bom desempenho de reconhecimento. Os bons resultados obtidos com este método geraram uma série de trabalhos na literatura sobre o método em si e suas aplicações. Estes trabalhos estão descritos em [MOGHADDAM-a], [MOGHADDAM-b], [MOGHADDAM-c], [PENTLAND-a], [PENTLAND-b], [PENTLAND-c] e [TURK], nos quais podemos observar suas vantagens e limitações. Entre as principais limitações, está a necessidade de normalizarmos as imagens de entrada de modo a se assemelharem às imagens usadas para compor o modelo. Dentre as vantagens, temos um método rápido e prático de comparar faces humanas. Com o hardware adequado o método pode ser implementado em tempo real. O método de *eigenfaces* foi descrito pela primeira vez em [TURK], onde se estabeleceram as bases para a aplicação deste tipo de análise ao reconhecimento de faces humanas.

3.1 Introdução

A elaboração do método de *eigenfaces* foi motivada pela fisiologia e pela teoria de informação [TURK], visando também os requisitos práticos de processamento em tempo real. O método trata o problema de reconhecimento facial como um problema de reconhecimento de padrões bidimensionais, desta forma não exigindo a informação tridimensional da face. Em seqüências de vídeo ou fotografias, as faces estão quase sempre posicionadas frontalmente. O método de *eigenfaces* então se beneficia deste fato para descrever as faces com um pequeno conjunto de vistas características (frontal, perfil ou rotacionada) para reduzir a variabilidade do problema de reconhecimento facial. Assim, é necessário modelar apenas variações na posição, rotação ou expressão de faces frontais na imagem bidimensional.

O método de *eigenfaces* projeta, através do produto escalar das *eigenfaces* com as imagens, as imagens em um espaço de características que descreve as variações significativas das imagens faciais conhecidas. As características significativas são conhecidas como *eigenfaces* porque são constituídas dos autovetores (*eigenvectors*) do conjunto de faces. Note que estes autovetores possuem a forma de imagens de faces daí o nome *eigenfaces*. O resultado da operação de projeção do método caracteriza uma face individual. A face é deste modo representada por uma soma ponderada de coeficientes que medem a intensidade da correlação de cada imagem sobre uma dada *eigenface* do modelo. Para reconhecer uma face em particular, é necessário apenas comparar os pesos de uma nova imagem facial aos pesos de imagens faciais de indivíduos conhecidos. Uma vantagem em particular do método de *eigenfaces* é que ele provê a capacidade de aprender e mais tarde reconhecer novas faces sem ter que treinar o modelo novamente. A informação necessária para representar um determinado tipo de face fica armazenada nas *eigenfaces* do modelo. Assim pode ser usada para caracterizar qualquer face nova apresentada ao sistema, mesmo que esta face não esteja na composição do modelo.

Outras aplicações do método que não estão ligadas diretamente ao reconhecimento facial são a codificação, a compressão e a recuperação de imagens faciais.

3.2 Origens do Método

Muitos trabalhos anteriores em reconhecimento de face ignoravam a questão de quais aspectos dos estímulos faciais seriam importantes para a tarefa de reconhecimento [TURK]. O fato de determinar os aspectos mais significativos para o reconhecimento facial sugere o emprego de uma abordagem da teoria de informação para o problema de reconhecimento facial, em uma tentativa de determinar os componentes que mais contribuem para a caracterização da imagem facial. Assim, tornando possível codificar e decodificar imagens faciais baseado nas informações relevantes na imagem facial, extraídas através da análise de um conjunto de imagens faciais. Isto enfatizaria as características globais e locais mais significantes para a representação da face.

A implementação do método de *eigenfaces* foi motivada por técnicas desenvolvidas por Sirovich e Kirby, descritas em [KIRBY] e [SIROVICH] para a representação eficiente de imagens utilizando a análise de componentes principais. Nestes trabalhos, de um conjunto de imagens faciais calculou-se o melhor sistema de coordenadas (base de vetores ortonormais), no sentido do erro quadrático médio de reconstrução, para a compressão das imagens. Nesta base, cada eixo coordenado (vetor) representado por uma imagem que foi denominada de *eigenpicture*. Foi sugerido a princípio, que qualquer conjunto de faces poderia ser aproximado armazenando-se pequenos grupos de pesos para cada face e um conjunto de imagens padrão (*eigenpictures*). Os pesos que descrevem cada face são obtidos projetando-se a imagem facial em cada eixo de coordenadas de *eigenpicture*. Turk e Pentland em [TURK], propuseram que um grande número de imagens faciais podem ser reconstruídas em termos de somas ponderadas de um pequeno conjunto de características ou *eigenpictures*. Uma maneira eficiente para aprender e

reconhecer faces seria construir as *eigenpictures* e reconhecer faces em particular pela comparação dos pesos necessários à reconstrução da face com os pesos associados a indivíduos conhecidos.

Pelo procedimento acima descrito, os indivíduos seriam então caracterizados por um pequeno conjunto de características, ou pesos das *eigenpictures*, necessários para sua descrição e reconstrução. Após a descrição das imagens em termos das *eigenpictures* a reconstrução do método não é exata, mas é ótima no sentido do erro quadrático médio de reconstrução e fornece uma representação extremamente compacta quando comparada com as dimensões das imagens faciais originais. Deste modo Turk e Pentland propuseram a aplicação do modelo de *eigenfaces* para a solução do problema de extração e codificação das características das imagens faciais para o reconhecimento de faces.

3.3 Obtenção do Modelo de *Eigenfaces*

O método de *eigenfaces* deve extrair a informação relevante na imagem facial e codificar esta informação do modo mais eficiente possível. Esta eficiência é medida em termos do erro quadrático médio de reconstrução da imagem facial codificada em relação à imagem facial original. A informação da face codificada é comparada com uma base de dados de imagens codificadas da mesma maneira. Um modo de extrair a informação contida em conjunto de imagens faciais e usar esta informação para codificar outras imagens faciais é obtê-la capturando a variação entre as faces de um conjunto de imagens faciais. Isto é feito independente de qualquer julgamento ou atribuição de características a priori. A informação obtida da análise destas variações é usada então para codificar faces individuais.

No processamento de imagens faciais, tratamos cada imagem como vetor em um espaço de dimensão elevada, da ordem do número de pixels presentes na imagem. Esta

representação das imagens faciais na forma de um vetor linha pode ser vista na Figura 3-1.

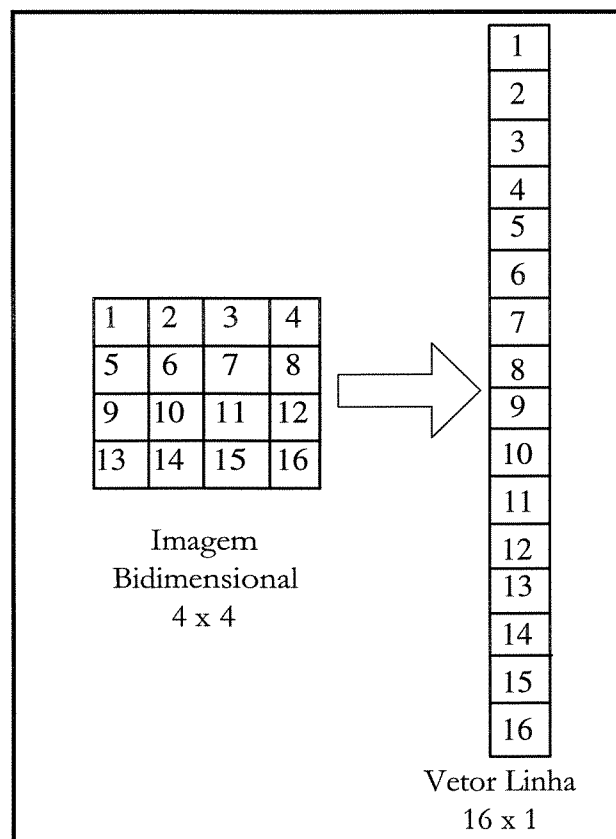


Figura 3-1: Representação de uma imagem na forma de vetor linha.

Na análise do conjunto de imagens faciais na forma de vetores linha obtemos as componentes principais da distribuição das faces. A análise de componentes principais fornece o conjunto de autovalores e autovetores da base de dados de imagens faciais. Os autovetores são ordenados, em ordem decrescente da proporção que retêm da variação entre as imagens faciais. O conjunto de autovetores pode ser imaginado como um conjunto de características que juntas representam as variações entre as faces. Os pontos das imagens faciais contribuem com maior ou menor intensidade na formação do autovetor, assim os autovetores possuem um aspecto semelhante a faces humanas, conforme mostrado na Figura 3-2 . Como foi dito anteriormente, devido a este fato Turk e Pentland em [TURK] denominaram estas características de *eigenfaces*. As *eigenfaces*

diferem de um padrão uniforme onde existem características diferentes entre as imagens do conjunto de treino. As *eigenfaces* realizam uma espécie de mapeamento das variações entre as faces. Na Figura 3-2 notamos a ausência de variação no fundo da imagem em contraste com as variações presentes na área da face já que o fundo é uniforme para imagens da base de dados.

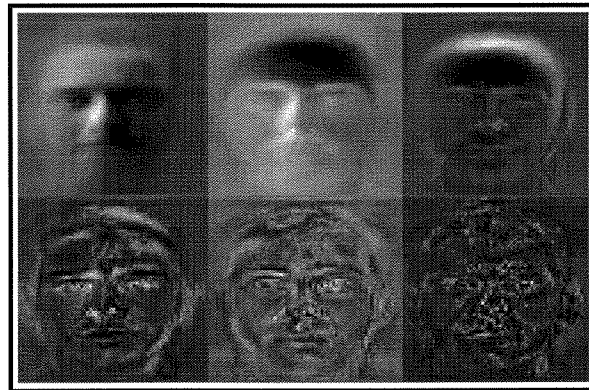


Figura 3-2: Exemplos de *eigenfaces*, que recebem este nome devido a sua semelhança com as faces humanas.

As faces individuais podem ser representadas como uma combinação linear das *eigenfaces*. As M *eigenfaces* obtidas do modelo formam um sub-espço M -dimensional chamado de espaço de face. O espaço de face contém a informação obtida das imagens do conjunto de treino. Podemos também aproximar cada face usando as M' "melhores" *eigenfaces*, aquelas com os maiores autovalores e que detêm a maior parte da variância dentro do conjunto de imagens. O procedimento para a obtenção do modelo de *eigenfaces* para o reconhecimento de faces envolve os seguintes passos:

1. Aquisição de um conjunto de imagens faciais para treinamento do modelo.
2. Cálculo das *eigenfaces* do conjunto de treino, mantendo apenas as M' *eigenfaces* que possuem os maiores autovalores.
3. Calcular a projeção de cada indivíduo conhecido no espaço M -dimensional de pesos projetando as faces sobre as M *eigenfaces*.

O conjunto de imagens escolhidas para compor o modelo determina a informação que vai ser reconhecida pelo modelo. Portanto, teoricamente, cada variação em expressão ou posição da face deveria ser modelada. Na prática, um determinado modelo pode suportar um certo grau de variação da imagem. Mas para um bom desempenho do método em relação às variações de posição, escala e expressão da face, é aconselhável modelar algumas variações da imagem facial. Esta composição de variações da imagem facial é usada em [MOGHADDAM-b] e [PENTLAND-c] para a construção de 5 diferentes modelos *eigenfaces* empregando faces frontais, em perfil e a 45° de rotação à esquerda e à direita em relação à câmera. Estes modelos *eigenfaces* são usados pelo sistema para determinar a posição da face antes de fazer o reconhecimento. Uma vez determinada a posição da face, a imagem de entrada tem suas características extraídas segundo os modelos de *eigenfaces* da posição determinada, então, a comparação de características é efetuada apenas entre as características das imagens com mesma posição da imagem de entrada. Isto aumenta o grau de tolerância em relação a rotação fora do plano da imagem, face não alinhada frontalmente com a câmera, do sistema proposto.

Após a construção do modelo de *eigenfaces*, podemos empregar a informação extraída pelo modelo para as tarefas de reconhecimento e verificação faciais. Com o modelo construído, os passos necessários para o reconhecimento de novas faces são:

1. Cálculo dos pesos da imagem realizando a projeção sobre as M' *eigenfaces*.
2. Determinar se a imagem é uma face, verificando sua distância do espaço de face.
3. Se for uma imagem facial, classificá-la como conhecida ou desconhecida.

O método de *eigenfaces* fornece um modo simples de análise através da comparação das distâncias entre vetores de pesos de imagens projetados no espaço de faces. No caso da imagem de entrada não ser uma imagem facial, sua distância ao espaço de face será elevada em comparação com as imagens faciais. Este fato ocorre porque o

modelo é treinado com imagens faciais para reconstruí-las com o menor erro quadrático médio possível. Assim, um objeto que não é uma face não é bem representado pelo modelo e apresenta uma grande distância ao espaço de faces do modelo. A análise destas distâncias pode ser empregada para estabelecer um limiar de diferenciação entre imagens faciais e não faciais.

3.4 Cálculo das *Eigenfaces*

Descrevemos a seguir os passos utilizados para a obtenção do modelo de *eigenfaces* a partir de um conjunto de imagens de treino. Para o treinamento, as imagens devem ser previamente normalizadas em relação à posição, escala e iluminação. A normalização é necessária para diminuir a variabilidade do problema. Seleccionamos então o tipo de imagem com uma determinada posição e expressão que queremos modelar. No cálculo do modelo, a imagem facial é descrita pela função imagem $I(x,y)$. Esta imagem é uma matriz bidimensional de níveis de cinza de dimensão $N \times N$. Definimos a imagem como uma matriz quadrada apenas para simplificar a demonstração, no entanto a imagem pode ter qualquer dimensão. A imagem pode ser considerada também com um vetor de dimensão N^2 . Assim uma imagem típica de 256×256 pixels pode ser também representada como um vetor de dimensão 65536. A coleção de imagens é mapeada para um conjunto de pontos neste espaço de dimensão elevada.

A idéia central do método empregado no cálculo das *eigenfaces*, a análise de componentes principais (ou expansão de Karhunen-Loeve), é encontrar os autovetores que melhor correspondem à distribuição das imagens faciais. Estes autovetores definem um espaço de imagens faciais, o chamado espaço de face. As imagens faciais são muito similares entre si, por exemplo, todas possuem olhos, boca e nariz e estão quase sempre em posição frontal. Deste modo não estarão dispostas aleatoriamente no espaço de face, concentrando-se em uma região que descreve do espaço de face que descreve as

semelhanças das imagens faciais. Podemos, então, posteriormente, descrever as imagens faciais em um sub-espço do espaço de face em função da similaridade das imagens faciais. Descreveremos a seguir os passos necessários para a obtenção do modelo.

Seja o conjunto de M imagens de treino:

$$S = \left\{ \Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_M \right\} \quad \text{Equação 3-1}$$

onde Γ_i , $i = 1 \dots M$, são as imagens faciais na forma de vetores linha de dimensão $N^2 \times 1$.

A média do conjunto de faces é definida por:

$$\Psi = \frac{1}{M} \sum_{n=1}^M \Gamma_n \quad \text{Equação 3-2}$$

onde M é o número de imagens no conjunto de treino.

Representamos as imagens normalizadas pela média do conjunto de treino, por:

$$\Phi_i = \Gamma_i - \Psi, \quad i = 1 \dots M \quad \text{Equação 3-3}$$

Um exemplo da face média para um dado conjunto de treino, composto de exemplos de imagens frontais normalizadas, é mostrado na Figura 3-3.

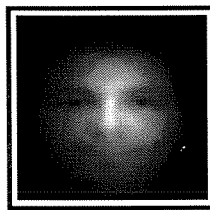


Figura 3-3: Imagem média de um conjunto de imagens faciais frontais e normalizadas.

Este conjunto de vetores de grandes dimensões, é então, submetido à análise de componentes principais a qual busca um conjunto de vetores ortonormais, $\mathbf{u}_n$, que descrevam a distribuição dos dados.

O k -ésimo vetor, $\mathbf{u}_k$, de dimensão $N^2 \times 1$, é definido de modo que,

$$\lambda_k = \frac{1}{M} \sum_{n=1}^M (\mathbf{u}_k^T \Phi_n)^2 \quad \text{Equação 3-4}$$

é máximo quando submetido a [JOHNSON],

$$\mathbf{u}_l^T \mathbf{u}_k = \delta_{lk} = \begin{cases} 1, & \text{se } l=k \\ 0, & \text{c.c.} \end{cases} \quad \text{Equação 3-5}$$

Os vetores $\mathbf{u}_k$ e os escalares λ_k são, respectivamente, os autovetores e autovalores da matriz de covariância:

$$\mathbf{C} = \frac{1}{M} \sum_{n=1}^M \Phi_n \Phi_n^T$$

$$\mathbf{C} = \mathbf{A} \mathbf{A}^T \quad \text{Equação 3-6}$$

onde a matriz $\mathbf{A}$ é dada por:

$$\mathbf{A} = \begin{bmatrix} \Phi_1 & \Phi_2 & \dots & \Phi_M \end{bmatrix} \quad \text{Equação 3-7}$$

Entretanto, a matriz $\mathbf{C}$, que representa de a matriz de covariância entre os N^2 pixels das imagem faciais, é de dimensão $N^2 \times N^2$ e determinar seus N^2 autovetores, um autovetor associado a variação de cada um dos N^2 pixels da imagem facial, e N^2 autovalores, não é uma tarefa prática em termos computacionais. Para imagens típicas da ordem de 256×256 pixels, por exemplo, os tempos envolvidos no cálculo podem inviabilizar a análise. Devemos então aplicar um método computacional viável para calcular os autovalores e autovetores. Podemos fazer a seguinte simplificação no método de cálculo descrito anteriormente. Se o número de dados no espaço de imagens, quantidade de imagens de treino, é menor que a dimensão do espaço, da ordem do número de pixels das imagens, ao invés de N^2 existirão apenas $M-1$ ($M < N^2$) autovetores

e autovalores significativos [TURK]. Os restantes estarão associados com autovalores nulos. Neste caso, podemos resolver o problema para os autovalores de dimensão N^2 , solucionando primeiramente os autovalores e autovetores de uma matriz $M \times M$ (onde M é o número de imagens de treino). Então, tomamos combinações lineares apropriadas das imagens normalizadas em relação à média do conjunto de treino, $\mathbf{A}$, com os autovetores da matriz $M \times M$, para obtermos os $M-1$ autovetores significativos de $\mathbf{C}$.

Para desenvolvermos estas simplificações no cálculo, considere os autovalores $\mathbf{v}_i$, com dimensão $M \times 1$, de $\mathbf{A}^T \mathbf{A}$, tais que:

$$\mathbf{A}^T \mathbf{A} \mathbf{v}_i = \mu_i \mathbf{v}_i \quad \text{Equação 3-8}$$

Pré-multiplicando ambos os lados por $\mathbf{A}$, nós temos:

$$\mathbf{A} \mathbf{A}^T \mathbf{A} \mathbf{v}_i = \mu_i \mathbf{A} \mathbf{v}_i \quad \text{Equação 3-9}$$

da qual vemos que $\mathbf{A} \mathbf{v}_i$ são os autovetores de $\mathbf{C} = \mathbf{A} \mathbf{A}^T$.

Seguindo a análise construímos a matriz de dimensão $M \times M$ de correlação das imagens de treino, dada por:

$$\mathbf{L} = \mathbf{A}^T \mathbf{A} \quad \text{Equação 3-10}$$

onde os elementos da matriz $\mathbf{L}$ também podem ser expressos como:

$$L_{mn} = \Phi_m^T \Phi_n \quad \text{Equação 3-11}$$

e calculamos os M autovetores de $\mathbf{L}$.

Para obtermos então os autovetores de $\mathbf{u}_l$, $l = 1..M$, de $\mathbf{C}$ baseados nos autovetores $\mathbf{v}_l$, $l = 1..M$, da matriz $\mathbf{L}$, temos que $\mathbf{A} \mathbf{v}_l$ são os autovetores $\mathbf{u}_l$, $l = 1..M$, da matriz $\mathbf{C} = \mathbf{A} \mathbf{A}^T$, logo podemos então expressar o produto $\mathbf{A} \mathbf{v}_l$ como:

$$\mathbf{u}_l = \sum_{k=1}^M \mathbf{v}_{lk} \Phi_k, l, 1..M \quad \text{Equação 3-12}$$

onde $\mathbf{v}_{lk}$ é o k-ésimo elemento do autovetor $\mathbf{v}_l$, $l = 1..M$, da matriz $\mathbf{L}$ e $\mathbf{u}_l$, $l = 1..M$ são então as M componentes principais da base de dados de treino, ou seja, suas *eigenfaces*. Com este método de análise, os cálculos dos autovetores são reduzidos da ordem de magnitude do número de pixels da imagem (N^2) para a ordem do número de imagens de treino (M). Na prática, teremos a dimensão do conjunto de treino sempre muito pequeno ($M \ll N^2$) e os cálculos se tornam de fácil manipulação.

3.5 Representação Usando *Eigenfaces*

Uma nova imagem facial, Γ , é codificada em termos das *eigenfaces* do modelo pela seguinte operação:

$$w_k = \mathbf{u}_k^T (\Gamma - \Psi) \quad k = 1..M' \quad \text{Equação 3-13}$$

onde,

- w_k é o peso calculado para a k-ésima *eigenface*;
- $\mathbf{u}_k$ é a k-ésima *eigenface*;
- Ψ é imagem média da base de dados de treino do modelo;
- M' o número de *eigenfaces* mais significativas, que possuem a maior variância da base de dados, escolhidas entre as M possíveis, sendo M o número de imagens de treino e do total de *eigenfaces*.

Os pesos formam o vetor,

$$\Omega_k^T = [w_1, w_2, \dots, w_{M'}] \quad \text{Equação 3-14}$$

o qual descreve a contribuição de cada *eigenface* na representação da nova imagem facial. Consideramos assim as *eigenfaces* com uma base para a descrição das imagens faciais

através da projeção da imagem facial sobre o espaço descrito pelas *eigenfaces*. Este vetor de pesos pode ser usado por algoritmos de reconhecimento de padrões para encontrar uma semelhança entre a nova imagem, Γ , e a classe pré-definida no conjunto de indivíduos da base de dados que mais se assemelha a ela. Este critério de semelhança varia com o tipo de método usado para a comparação. Podemos comparar por exemplo distâncias entre dois vetores de pesos ou utilizar um classificador de neural (codificador) treinado para classificar as imagens.

3.6 Classificação com *Eigenfaces*

O método mais simples e direto de avaliarmos a semelhança entre duas imagens faciais com o método de *eigenfaces* é compararmos as distâncias euclidianas (Equação 3-15) entre os vetores de pesos de *eigenfaces* das imagens. A classe da base de dados cujo vetor de pesos apresenta a menor distância em relação a nova imagem é atribuída como sendo a classe da nova imagem. A distância euclidiana a ser minimizada no processo de classificação é definida por,

$$\varepsilon_k = \left\| (\Omega - \Omega_k) \right\|^2 \quad \text{Equação 3-15}$$

onde,

- ε_k é a distância da nova face à k -ésima classe;
- Ω_k é o vetor de pesos que representa a k -ésima classe facial;
- Ω é o vetor de pesos da nova face a ser classificada.

O centro da k -ésima classe Ω_k pode ser obtido fazendo-se a média dos pesos obtidos para as imagens de um indivíduo tendo ao menos uma imagem de cada indivíduo para caracterizar a classe. Podemos também introduzir um limiar de aceitação na classificação. Uma face é classificada como pertencente a uma determinada classe k quando a distância mínima, ε_k , está abaixo de um limiar pré-estabelecido θ_ε . Caso

contrário a face é classificada como desconhecida e pode ser usada opcionalmente para criar uma nova classe de faces.

Podemos avaliar ainda a chamada distância do espaço de face dada por,

$$\varepsilon = \|\Phi - \Phi_f\|^2 \quad \text{Equação 3-16}$$

onde,

- Φ é a nova imagem normalizada em relação a média da base de dados de treino;
- Φ_f é dado por

$$\Phi_f = \sum_{i=1}^{M'} w_i \mathbf{u}_i \quad \text{Equação 3-17}$$

onde,

- M' é o número de *eigenfaces* escolhidas como significativas;
- w_i são os pesos da nova imagem sobre a i *eigenfaces*;
- $\mathbf{u}_i$ é a i -ésima *eigenface*.

A Equação 3-16 avalia a distância de dois vetores de imagens, normalizadas em relação ao espaço de face da base de dados. Os vetores que permitem avaliar esta distância são a imagem original e a imagem reconstruída em termos das *eigenfaces*. A distância do espaço de face, Equação 3-16, é uma medida da qualidade da reconstrução da imagem pelo modelo de *eigenfaces*. A distância do espaço de face pode indicar com alto grau de acerto quando um objeto é ou não uma imagem facial [TURK]. Em geral, imagens faciais normalmente irão se situar próximas ao espaço de face descrito pela *eigenfaces* do modelo e imagens que não contêm faces normalmente se situarão mais distantes. A análise de um limiar para a distância entre o objeto e o espaço de face permite rejeitar objetos que não sejam faces quando estes são apresentados ao modelo.

Observamos da Equação 3-17 que Φ_f é a reconstrução da imagem baseada nas *eigenfaces* sem estar adicionada a imagem média da base de dados. Uma reconstrução de imagem com o modelo de *eigenfaces* pode ser representada da seguinte forma:

$$\Phi_{fr} = \sum_{i=1}^{M''} w_i \mathbf{u}_i + \Psi \quad \text{Equação 3-18}$$

onde Ψ é a imagem média da base de dados. Ambas as reconstruções apresentam perdas em relação a imagem original. No entanto, se as imagens a serem reconstruídas se assemelharem às usadas na composição do modelo de *eigenfaces* a qualidade visual da reconstrução é satisfatória para representar a face original.

No processo de classificação com distância mínima existem quatro possibilidades para as distâncias do vetor de pesos da imagem em relação aos centros de classe e ao espaço de face, que é o espaço definido pelas *eigenfaces* do modelo facial. As combinações possíveis para as duas distâncias são as seguintes:

1. Próxima do espaço de face e próxima de uma classe facial;
2. Próxima do espaço de face mas distante das classes faciais;
3. Distante do espaço de face e próxima a uma classe facial;
4. Distante do espaço de face e distante das classes faciais.

No primeiro caso, o indivíduo é reconhecido e identificado. No segundo caso, um indivíduo desconhecido está sendo apresentado ao sistema. No terceiro e quarto casos temos uma imagem não facial apresentada ao sistema. O terceiro caso corresponde a uma aceitação falsa de uma imagem que não é uma face, neste caso ela pode ser rejeitada devido a sua grande distância do espaço de face. No quarto caso temos uma imagem não facial rejeitada sem maiores problemas pelo sistema.

3.7 Discussão

Apresentamos os passos necessários para a construção e utilização do modelo de *eigenfaces* como uma ferramenta para o reconhecimento e verificação facial. A análise de componentes principais fornece um modelo paramétrico das imagens que pode ser utilizado na compressão de imagens. Este modelo possui coeficientes que são estatisticamente descorrelacionados e ótimos no sentido do erro quadrático médio para reconstrução. As simplificações de cálculo permitem tornar tratáveis problemas que possuem um elevado número de imagens em alta resolução. A construção do modelo não afeta o tempo de processamento dos sistemas de reconhecimento e verificação. As únicas operações realizadas por estes sistemas são a projeção das novas imagens nas características obtidas pelo modelo e a classificação dos pesos obtidos da projeção.

Os detalhes da implementação dos sistemas de reconhecimento e verificação utilizando este método serão vistos no Capítulo 4.

Capítulo 4

4. Sistema de Reconhecimento Facial

Neste capítulo descrevemos a construção do sistema de reconhecimento facial. Mostramos a composição da base de dados utilizada para treinamento e teste. São comentados os algoritmos empregados para obtenção do modelo de *eigenfaces* com detalhes de sua implementação. Os resultados dos testes para estes sistemas podem ser observados no Capítulo 5.

4.1 Base de Dados

A qualidade da base de dados para treinamento e testes de um sistema de reconhecimento facial afeta diretamente o desempenho do sistema. Neste trabalho, o modelo de *eigenfaces* para o sistema de reconhecimento facial é construído apenas com imagens frontais das faces. Observamos as variações na iluminação, presença ou ausência de óculos, barba, expressões e rotação da face nas imagens. Estas variações devem

acrescentar uma maior probabilidade de erro de classificação ao modelo se comparadas aos resultados obtidos com bases de dados adquiridas sob condições controladas, descritas na literatura em [PENTLAND-a], [PENTLAND-b], [PENTLAND-c] e [TURK]. Nossa única restrição foi a utilização de um fundo uniforme para facilitar o processo de segmentação da imagem facial. Outras condições como iluminação e alinhamento não foram fortemente controladas. Poderemos posteriormente comparar estes resultados com bases de dados que também apresentam uma certo grau de variação, como por exemplo a base de dados do Olivetti Research Center [SAMARIA-b].

4.1.1 Metodologia de Aquisição

Montamos um conjunto para a aquisição das imagens composto por câmera de vídeo e microcomputador com placa digitalizadora. A placa digitalizadora adquire apenas fotografias das faces e não seqüências de imagens. Utilizamos um fundo uniforme para facilitar o processo de segmentação facial e normalização. O esquema deste conjunto pode ser visto na Figura 4-1.

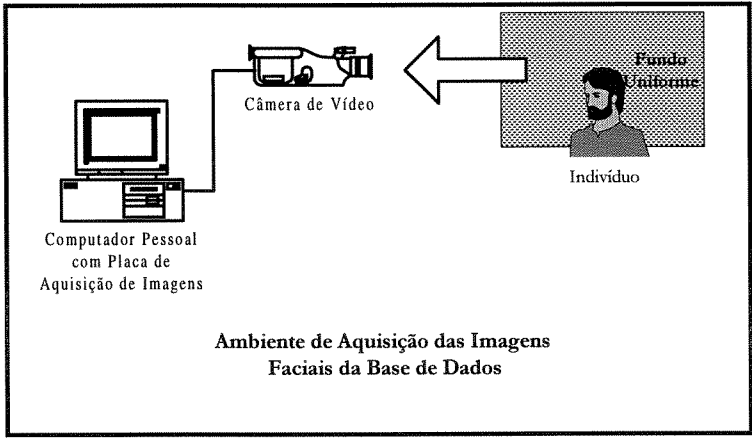


Figura 4-1: Esquema dos equipamentos empregados para a aquisição das imagens da base de dados.

Estas imagens foram adquiridas na resolução de 240 X 332 pixels em modo RGB, o que representa um compromisso entre qualidade da imagem e espaço de armazenamento, e convertidas para 8 níveis de cinza posteriormente, para simplificar o

processamento da informação contida na imagem. As imagens são armazenadas em formato de mapa de bits [LEVINE] para serem processadas pelo sistema.

Com o conjunto da Figura 4-1, adquirimos uma base de dados composta de 435 imagens de 102 indivíduos. Não temos um número igual de imagens para cada indivíduo. O número de imagens por indivíduos e quantidade de indivíduos com este número de imagens pode ser visto na Tabela 4-1.

Número de Imagens Por Indivíduo	Quantidade de Indivíduos com Este Número de Imagens	Total de Imagens no Grupo
3	37	111
4	1	4
5	64	320
Total de Imagens da base de dados		435

Tabela 4-1: Distribuição da ocorrência dos grupos de faces de indivíduos na base de dados.

Na maioria dos indivíduos, 428 pessoas, as imagens foram adquiridas na mesma ocasião. Somente alguns indivíduos, 7 pessoas, possuem imagens adquiridas em ocasiões diferentes. As imagens adquiridas em ocasiões diferentes apresentam, em alguns casos, um maior grau de variação. Em todas as imagens adquiridas não foram feitos esforços para garantir o correto alinhamento das faces frontais com a câmera nem a constância da expressão facial presente nos indivíduos. Não foram impostas restrições a presença de óculos ou barba. O controle da iluminação também não foi efetuado. No aspecto da iluminação, às vezes variávamos um pouco este parâmetro ligando ou desligando um conjunto de lâmpadas. Observamos isto na Figura 4-2 onde podemos ver a variação da iluminação em todas as imagens da base de dados.

A única restrição imposta no processo de aquisição foi a de que cada indivíduo deveria possuir duas imagens aproximadamente frontais, com expressão neutra e sem óculos, para compor o modelo de *eigenfaces*. Esta restrição não foi plenamente atendida e assim, das 204 imagens possíveis de 102 indivíduos com estas características, obtivemos

193, que são suficientes para a implementação e testes do modelo.

O exemplo da variabilidade da base de dados pode ser visto na Figura 4-2 onde estão presentes todas as 435 imagens dos 102 indivíduos. Podemos notar nesta figura que alguns grupos de indivíduos apresentam uma grande variabilidade em expressão, escala, posição e iluminação. As variações presentes na base de dados servem como um bom exemplo de uma situação real para um sistema de reconhecimento facial que opera sob condições onde não possuímos um controle rígido do processo de aquisição das imagens. Podemos desta forma avaliar a capacidade de tolerância do modelo de *eigenfaces* à estas variações.

Comparando nossa base de dados da Figura 4-2, com a base de dados da Figura 4-3 obtida por Ferdinando Samaria dos laboratórios do Olivetti Research Center e usada em seus testes de processamento de face, observamos que a base de dados de Samaria possui muito mais informação sobre as possíveis variações de cada classe de imagens do que nossa base de dados, ou seja, ela possui 10 imagens de cada um dos 40 indivíduos. Entretanto ela também trata com um número bem menor de classes que nossa base de dados. Esta diferença entre conjuntos de testes e funções em que são empregados os sistemas faciais dificulta muito a comparação do desempenho dos mesmos. Em nosso caso, estamos testando a capacidade do sistema reconhecer faces com poucos exemplos sobre cada indivíduo.

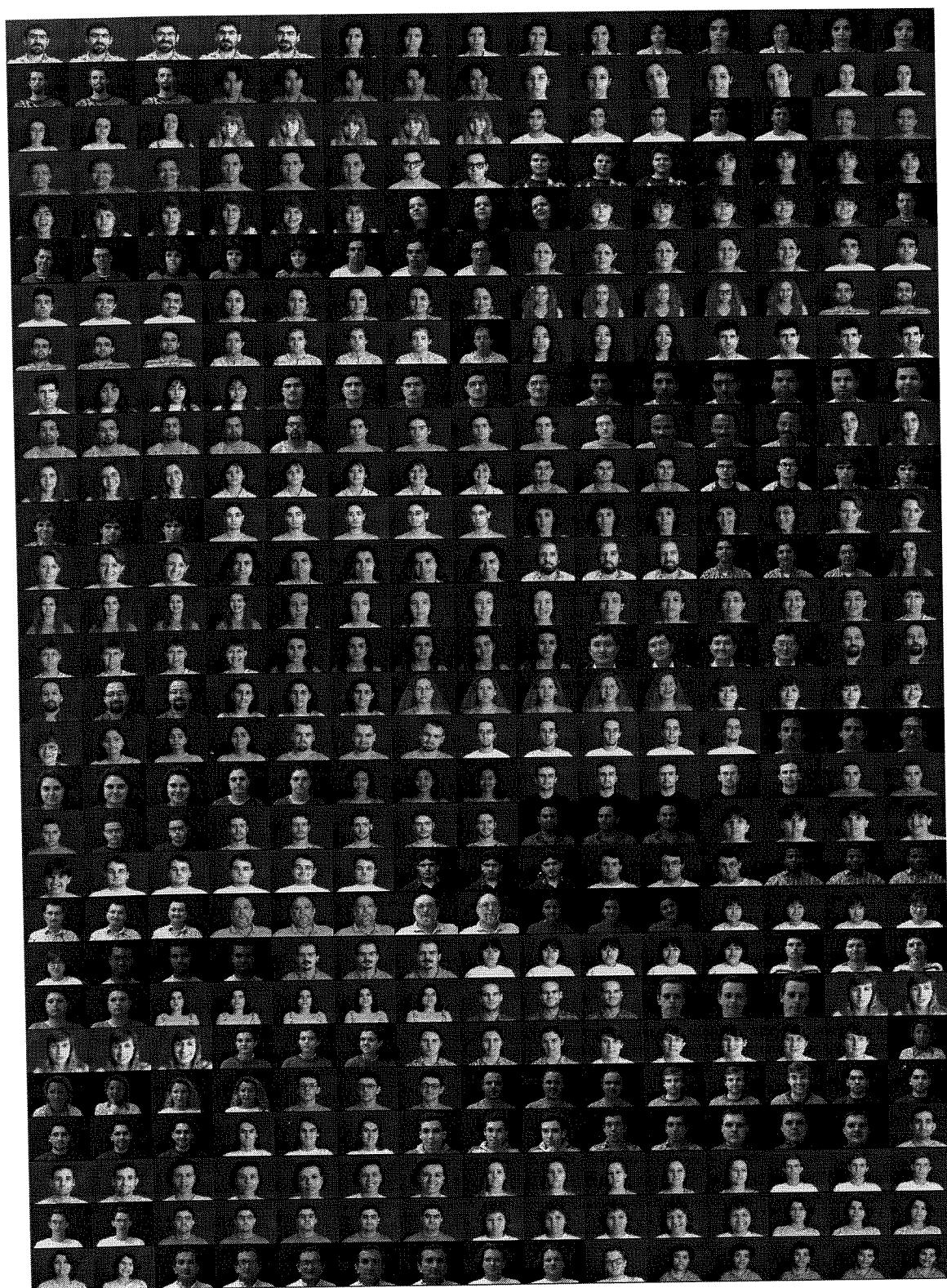


Figura 4-2: Base de dados de imagens faciais contendo 435 imagens de 102 indivíduos.

Na figura percebe-se a variação das condições de aquisição para as diferentes imagens presentes na base de dados.

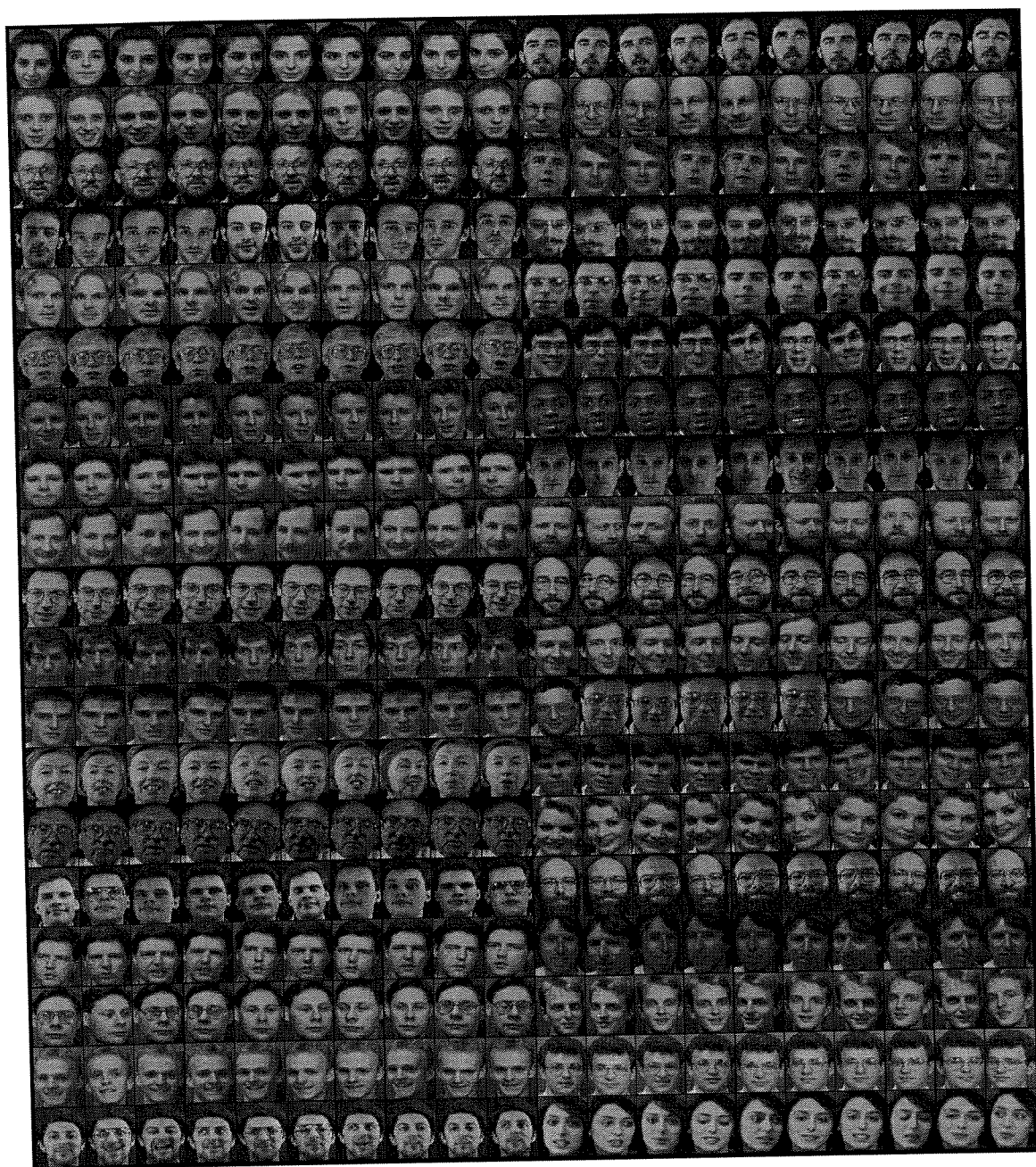


Figura 4-3: Base de dados do Olivetti Research Center para testes de algoritmos de reconhecimento de imagem. Esta base de dados foi montada por Ferdinando Samaria e utilizada em experimentos descritos em [SAMARIA] e [GOUDAIL].

4.1.2 Catalogação das Imagens

Como forma de obtermos a normalização necessária ao método de *eigenfaces* (em relação a escala, rotação e translação) e de colher dados sob as condições gerais da imagem (expressão, barba, óculos, rotação, etc.) foi desenvolvido um pequeno aplicativo em VISUAL BASIC™. O aplicativo de catalogação auxilia na tarefa marcar os pontos característicos necessários à normalização (olhos, nariz e boca) sobre a imagem com o mouse. Os resultados são então gravados em arquivos individuais para cada imagem da base de dados.

Na Figura 4-4, podemos ver um exemplo do processo de normalização facial onde manualmente marcamos sobre a imagem os pontos de interesse, olhos, nariz e boca. Na Figura 4-5 temos a tabela do aplicativo onde marcamos as condições em que determinada imagem se encontra em relação a presença ou ausência de óculos ou barba, condições de iluminação, sexo, posição e reflexões da luz sobre a face.

Dentre os itens catalogados, selecionamos os seguintes como de interesse no processo de comparação dos resultados do modelo de *eigenfaces*. Faces semelhantes a faces utilizadas para construir o modelo, ou seja, sem óculos, não rotacionadas e com expressão neutra, faces rotacionadas, faces com expressão (sorriso) rotacionadas ou não e faces com óculos. O resumo da distribuição da imagens da base de dados segundo estas categorias pode ser visto na Tabela 4-2. Estes dados serão usados posteriormente para compararmos o desempenho do sistema em relação a cada uma destas variações e avaliarmos quais são mais significativas.

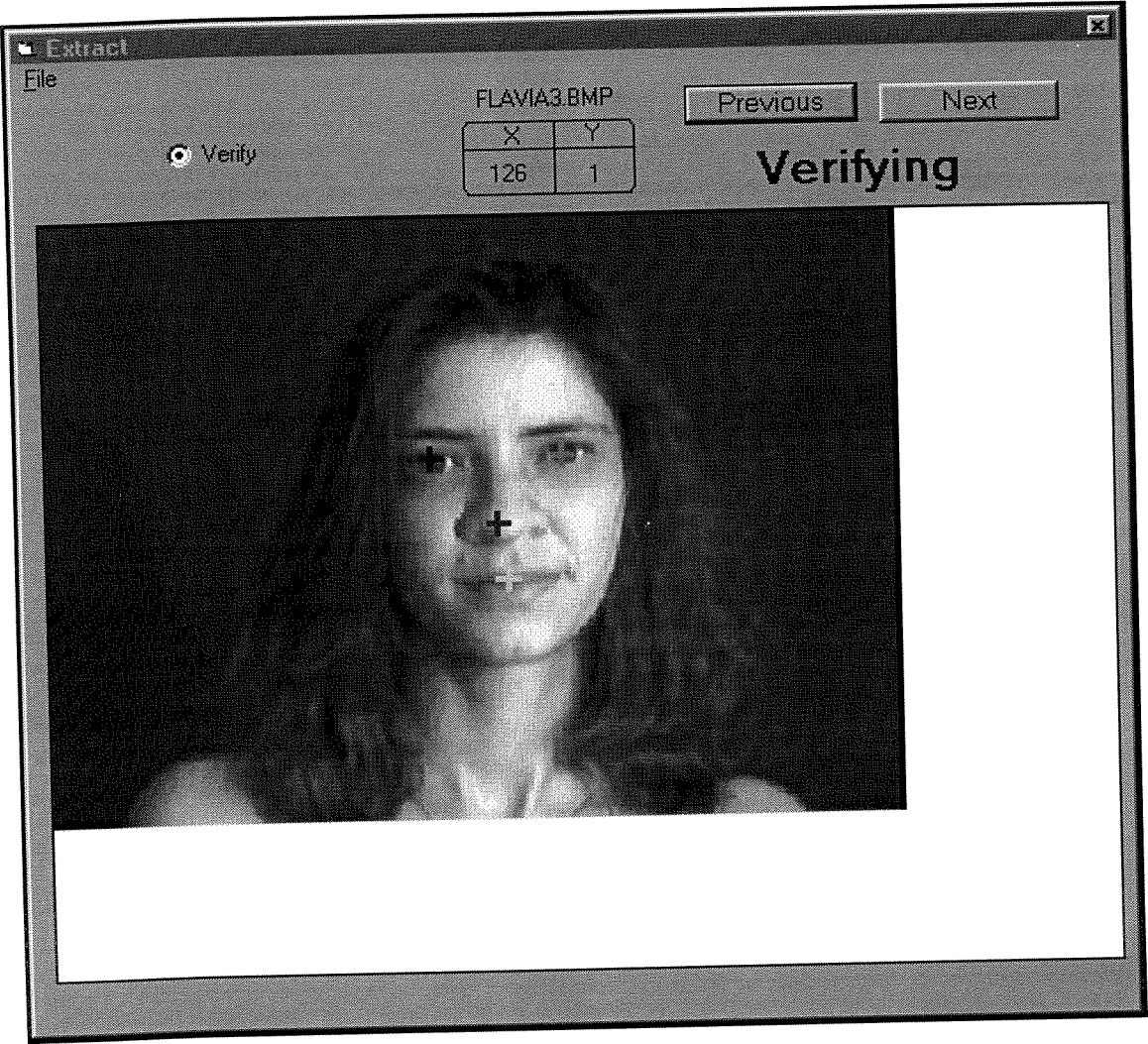


Figura 4-4: Tela do aplicativo desenvolvido para o processo manual de normalização facial. As cruces marcam os pontos das características que podem ser usadas na normalização.

Database Register

RightEye

148

96

LeftEye

199

95

Nose

174

122

Mouth

178

144

Sex

Male

Female

Glasses

Present

Absent

Smile

Normal

Smiling

Reflection

Present

None

Beard

Present

Absent

Aligment

Frontal

Rotated

Light

High

Normal

Low

Class :

flavia

OK

Redo

Figura 4-5: Quadro do aplicativo de normalização onde são introduzidos os comentários sobre sexo, sorriso, barba, presença ou ausência de óculos, reflexão da luz na face, posição e iluminação.

Categoria	Número de Imagens (Total 435)
Modelo	193
Frontal	319
Rotacionadas	116
Sorrindo	142
Óculos	62
Barba	141
Sorrindo e Frontal	95
Sorrindo e Rotacionada	47

Tabela 4-2: Imagens da base de dados segundo as características selecionadas. Uma mesma imagem pode apresentar mais de uma característica. O grupo modelo é composto por frontais, sem óculos e com expressão neutra que serão usadas na construção dos modelos de eigenfaces.

4.2 Implementação do Método de *Eigenfaces*

No desenvolvimento dos sistemas de reconhecimento e verificação facial foram implementadas todas as rotinas necessárias à construção do modelo de *eigenfaces* e ao processamento das imagens. A implementação das rotinas foi realizada em linguagem C++ [STROUSTRUP] e [ELLIS] utilizando os compiladores Borland 4.52 e gcc. Na Figura 4-6, temos os diferentes blocos que compõem os sistemas de reconhecimento e verificação. Em primeiro lugar, é construído o modelo de *eigenfaces* e as *eigenfaces* obtidas são gravadas em arquivos. Estas características serão usadas posteriormente para codificar as imagens da base de dados. O segundo passo é registrar as pessoas a serem reconhecidas pelo sistema. Atribuímos então um centro de classe para cada indivíduo. Este centro de classe é formado com as características de uma imagem exemplo ou com uma média aritmética das características de várias imagens do indivíduo. No terceiro passo, temos a operação dos sistemas de reconhecimento ou verificação onde temos a classificação das imagens com relação aos centros de classe pré-definidos com um classificador de distância mínima.

O tempo de processamento dos sistemas de reconhecimento e verificação facial não é afetado pelas etapas anteriores de cálculo das *eigenfaces* e registro dos indivíduos no sistema, as quais deixam seus resultados disponíveis em disco para processamento. O tempo de construção do modelo de *eigenfaces* é de aproximadamente 15 minutos em um PENTIUM 100 MHz e 32Mb de RAM para 99 imagens de treino e o tempo de registro das imagens da base de dados é de cerca de 18 minutos para todas as 435 imagens na mesma configuração, empregando as 98 características passíveis de serem obtidas com o modelo calculado. Os tempos de processamento descritos anteriormente já incluem a normalização das imagens faciais. Na operação dos sistemas de reconhecimento e verificação, os sistemas gastam mais tempo com a normalização do que com a extração

de características e a classificação. Os tempos de extração de características e classificação variam com o número de características extraídas e com o número de classes presentes na base de dados, respectivamente. Na configuração descrita anteriormente, tempos de extração de características e classificação são da ordem inferior de 0,5 segundos para cada imagem processada. Detalhes dos tempos de cada operação podem ser vistos nas seções correspondentes a seguir.

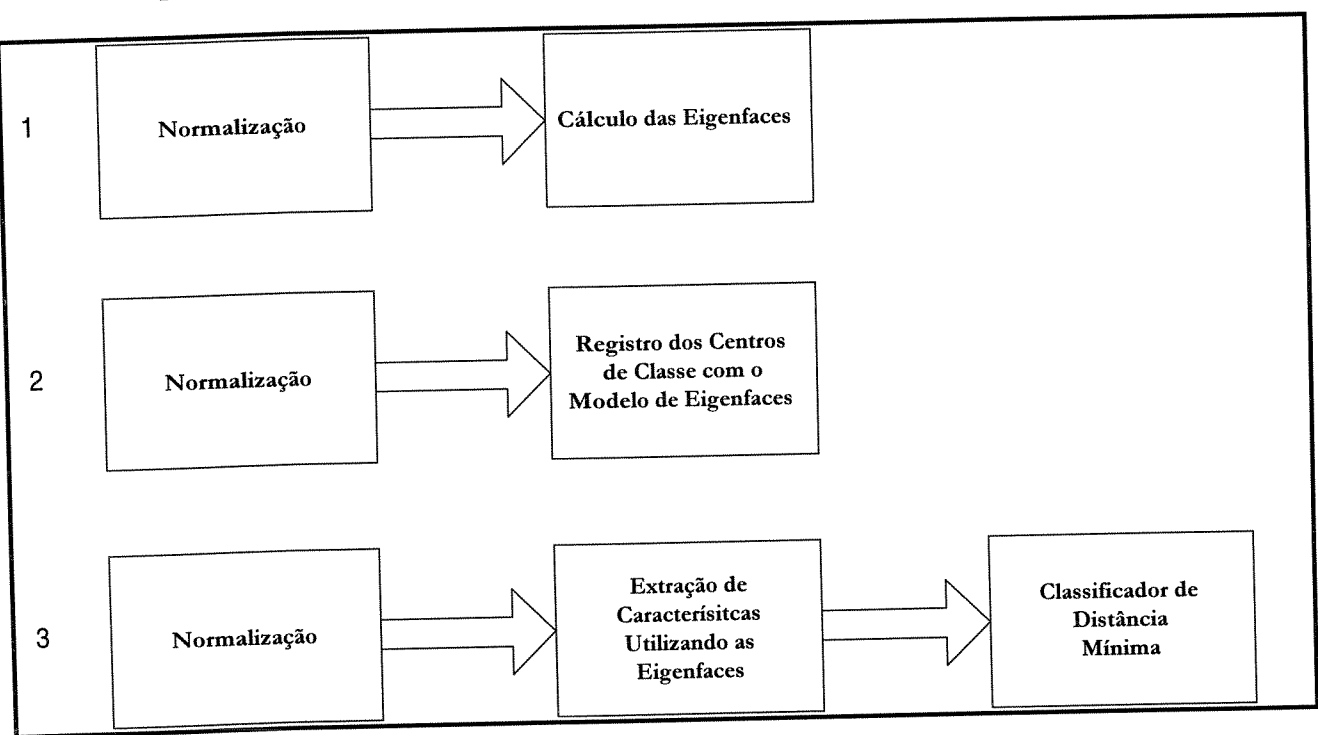


Figura 4-6: Etapas da operação do sistema facial: 1) construção do modelo de *eigenfaces*; 2) catalogação das imagens segundo o modelo; e 3) classificação das novas imagens apresentadas ao sistema.

As três etapas da Figura 4-6 foram implementadas em módulos independentes, para facilitar o tratamento dos dados. Os resultados intermediários são gravados em arquivos para posterior reutilização. Este procedimento permite que, se necessário, façamos a análise e classificação das características no sistema com outros aplicativos de análise de dados. Nas próximas seções, detalharemos os métodos implementados para a síntese de cada uma das partes do sistema que são empregadas nas diferentes etapas de

cálculo da *eigenfaces*, catalogação dos indivíduos no sistema e reconhecimento e verificação facial.

4.2.1 Normalização da Imagem Facial

A normalização das imagens faciais no modelo de *eigenfaces* é um passo importante. Os procedimentos de normalização da imagem facial foram divididos em dois grupos: pré-processamento da imagem e normalização da escala e posição com eliminação do fundo da imagem. O pré-processamento emprega um filtro de mediana sobre a imagem, para diminuir os ruídos presentes no processo de aquisição, seguido de obtenção do mapa de bordas da imagem. A função deste pré-processamento é buscar acentuar as características significativas da imagem facial, no caso, a informação das bordas da imagem.

As normalizações em relação a escala e posição, necessárias para operação do modelo de *eigenfaces*, foram efetuadas utilizando-se os dados das posições dos olhos nas imagens obtidos previamente com o aplicativo de catalogação manual. A distância média entre os olhos da base de dados foi determinada em 46 pixels, assim, esta distância foi assumida como a distância interocular padrão para normalização da imagem facial. Na Figura 4-7, podemos ver o esquema do processo de normalização facial. O primeiro passo é a detecção da posição dos olhos na imagem efetuado anteriormente com o aplicativo de catalogação manual. As imagens são então escaladas para a distância interocular padrão e transladadas de modo que a posição dos olhos, agora já normalizados em relação à escala, coincidam com a posição dos olhos do padrão de escala e posição previamente definido. Com as imagens na posição padrão, é efetuada a extração da área facial da imagem normalizada, para extrair apenas a região de interesse na imagem. Após, as imagens que contêm as áreas faciais extraídas sofrem uma equalização de histograma, para diminuir a variabilidade em relação a iluminação. A

seguir, é aplicada uma máscara gaussiana na imagem com variância de 30 pixels na direção x (horizontal) e 40 pixels na direção y (vertical) centrada 10 pixels abaixo do ponto médio entre os olhos. A máscara gaussiana ressalta as características da face mais significativas para o reconhecimento, desprezando áreas como o fundo da imagem e cabelos que degradariam a performance do sistema.

Finalizando o procedimento de normalização, as imagens de tamanho original 240 x 332 pixels que tiveram sua área facial normalizada extraída para uma imagem de 128 x 128 pixels, são redimensionadas por sub-amostragem para o tamanho de 64 x 64 pixels. Esta redução do tamanho da imagem é um compromisso entre qualidade do reconhecimento, o tempo de processamento e a capacidade de armazenamento do sistema de processamento de imagens faciais. Imagens de 64 x 64 podem ser processadas mais rapidamente pelo sistema e requerem menor espaço de armazenamento. Elas não garantem exatamente a mesma taxa de erros que imagens maiores mas estas variam apenas em décimos percentuais. Isto se deve ao fato de que em muitas das imagens a área facial em si está dentro deste limite de 64 x 64 pixels.

Na avaliação do desempenho dos sistemas de reconhecimento e verificação facial, apresentada no Capítulo 5, comparamos os desempenhos de ambos os sistemas utilizando imagens normalizadas com e sem pré-processamento na construção dos modelos de eigenfaces e extração de características. Na Figura 4-8, vemos o diagrama de blocos do processo de normalização com pré-processamento e comparando-o com a normalização facial sem pré-processamento da Figura 4-7, observamos que apenas os blocos de filtragem, com filtro de mediana, e mapa de bordas são adicionados ao processo de normalização. O mapa de bordas é obtido empregando-se os operadores de Sobel vertical e horizontal sobre a imagem e obtendo-se o módulo da soma destes dois operadores em cada pixel da imagem. Os valores do mapa de bordas que excederam a um determinado limiar de intensidade, no caso 50, tem sua intensidade mudada para 255

caso contrário seu valor de intensidade permanece o mesmo. Exemplos das imagens normalizadas sem pré-processamento podem ser vistos na Figura 4-9 e exemplos de imagens normalizadas com pré-processamento podem ser observadas na Figura 4-10 onde temos ressaltadas as bordas das imagens faciais. Podemos notar em ambas as figuras a ausência de normalizações em relação a rotação das faces na imagem que é um dos parâmetros avaliados na capacidade de tolerância do sistema de reconhecimento e verificação facial. Os teste preliminares com a normalização da rotação da face mostraram que ela não contribui significativamente para o aumento da taxa de reconhecimento do sistema na maioria dos casos, chegando até a degradar a performance em torno de 0,5% em alguns testes. O processo de normalização sem pré-processamento demora em torno de 2 segundos por imagem em um microcomputador PENTIUM de 100 MHz e 32 Mb de RAM e com pré-processamento, cerca de 6 segundos por imagem na mesma configuração.

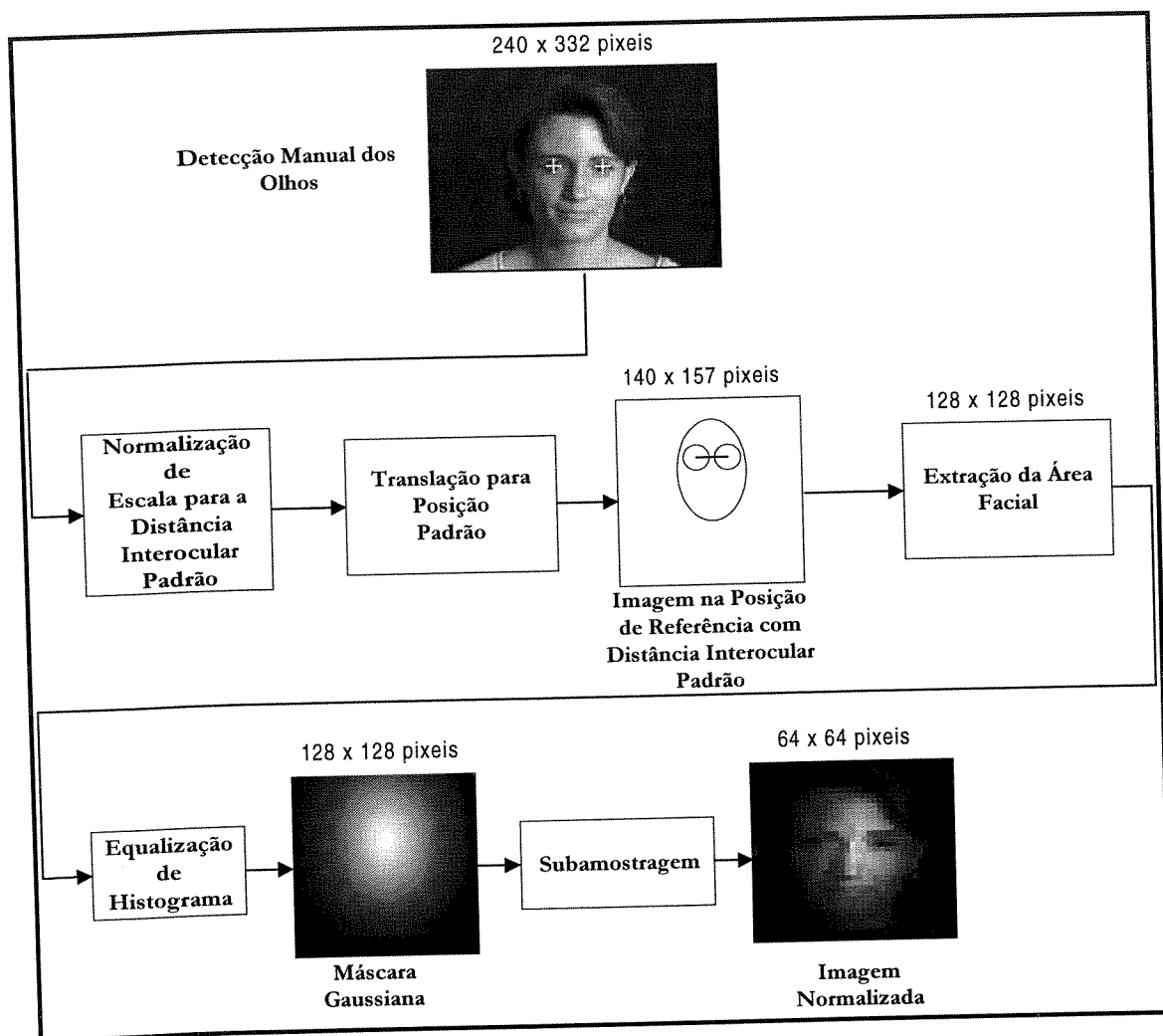


Figura 4-7: Diagrama de blocos do processo de normalização da imagem facial.

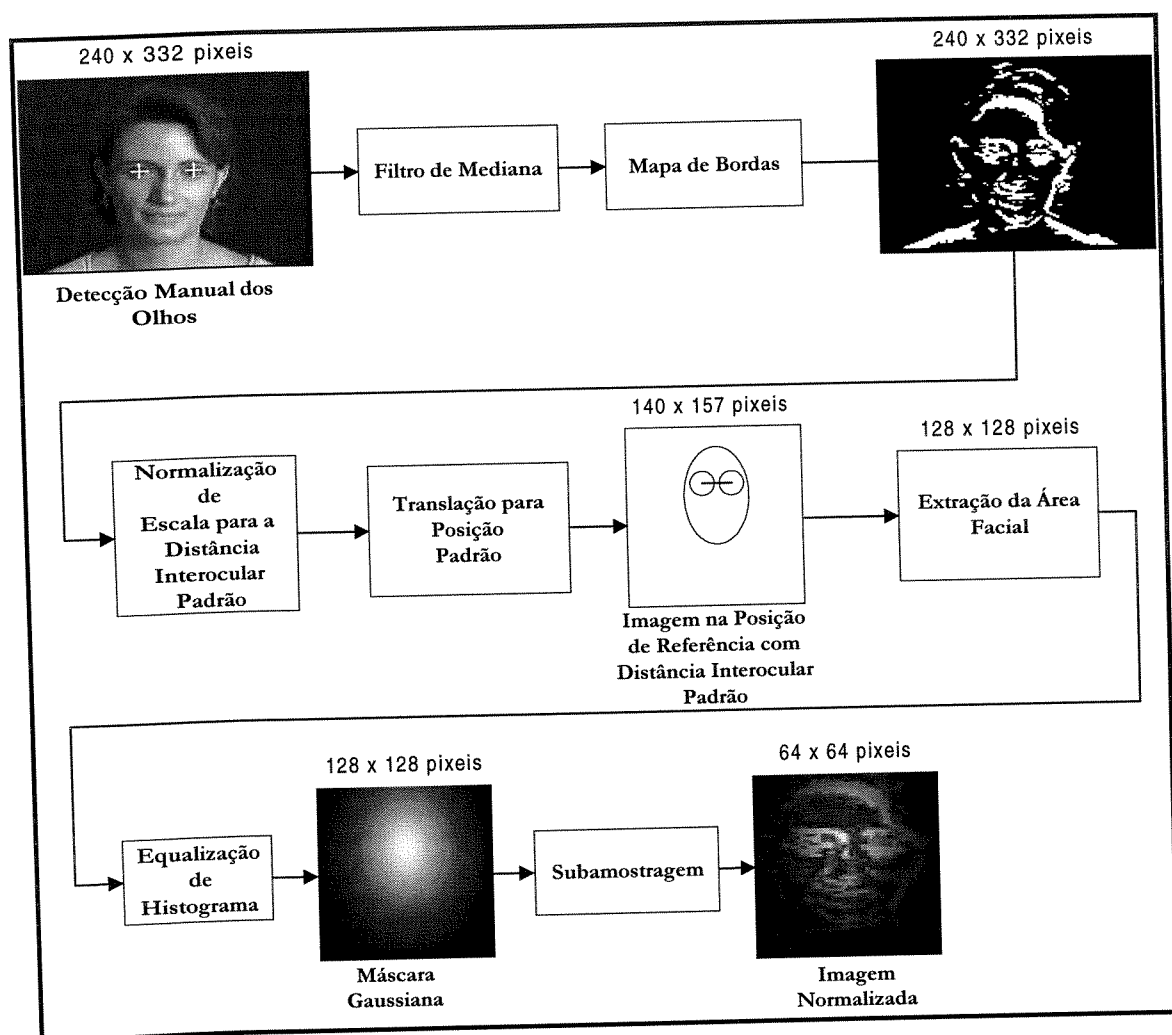


Figura 4-8: Diagrama de blocos do pré-processamento seguido de normalização da imagem facial.

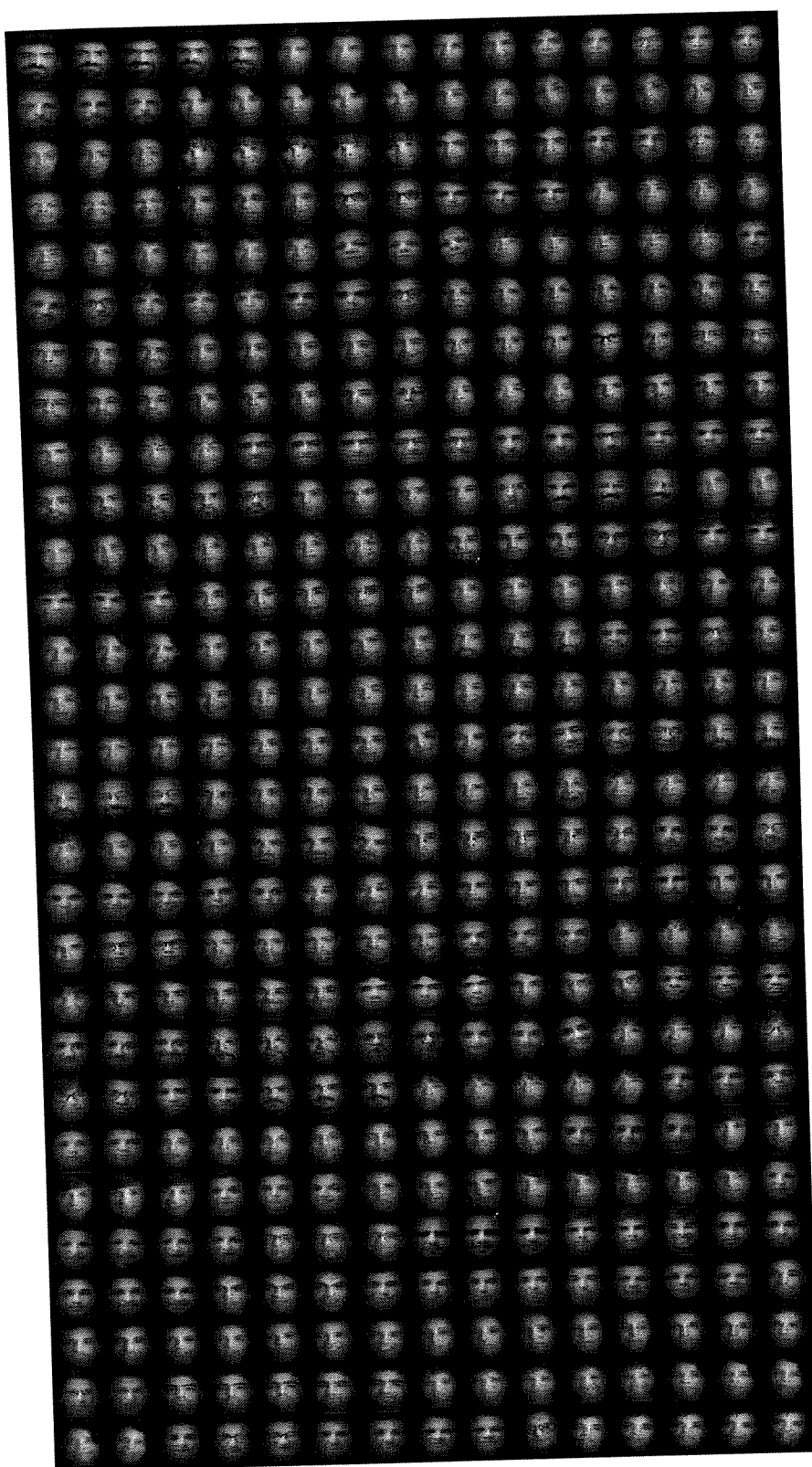


Figura 4-9: Imagens faciais normalizadas sem pré-processamento para construção do modelo facial e reconhecimento.

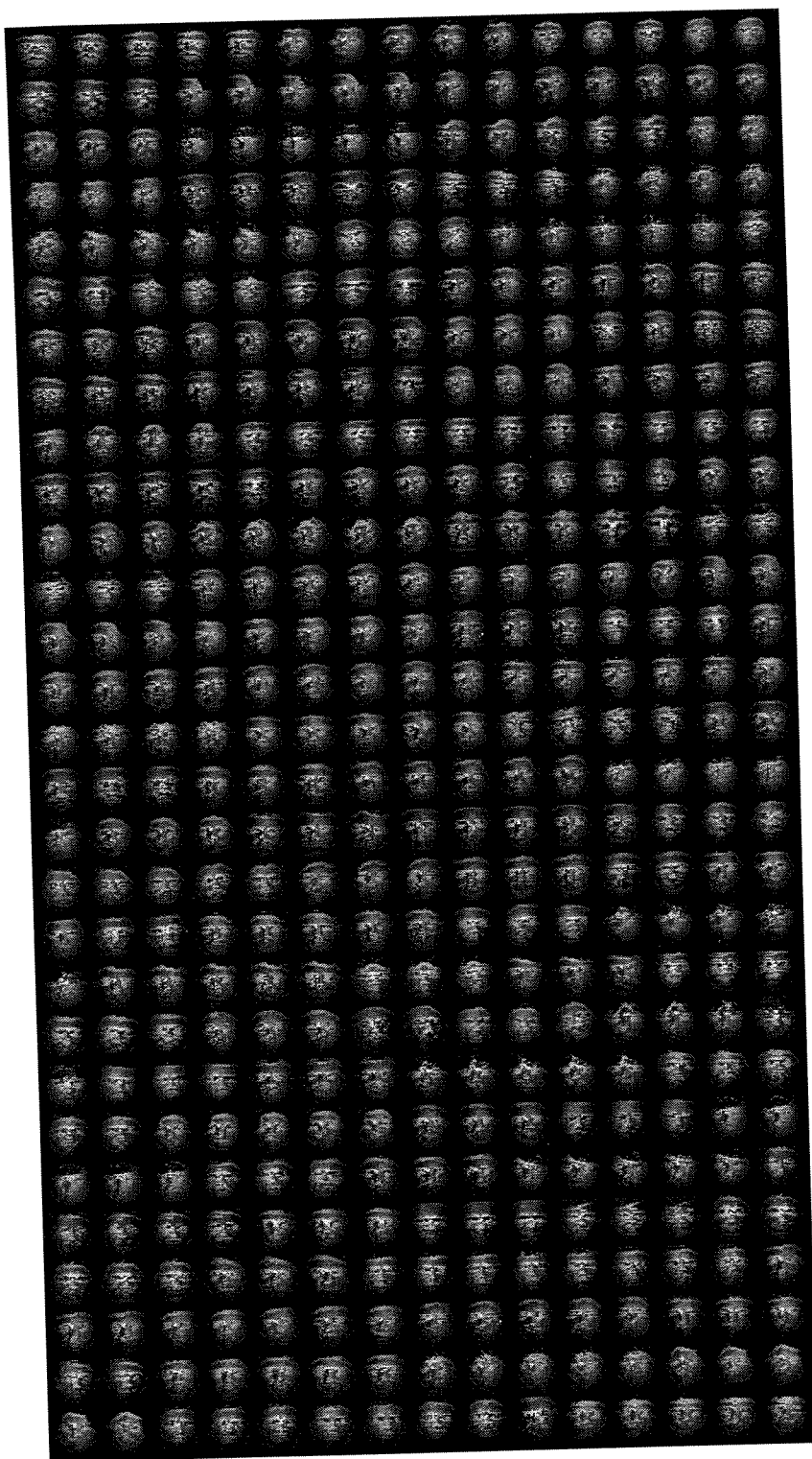


Figura 4-10: Imagens faciais pré-processadas e normalizadas para construção do modelo facial e reconhecimento.

4.2.2 Construção do Modelo de *Eigenfaces*

A construção do modelo de *eigenfaces* parte inicialmente da escolha de qual tipo de faces queremos modelar. Foram escolhidas faces que apresentassem as seguintes características: estivessem em posição frontal, não apresentassem óculos e possuísem uma expressão neutra. Foram selecionadas as 99 faces mostradas na Figura 4-11 e que foram posteriormente normalizadas, mostradas na Figura 4-12.

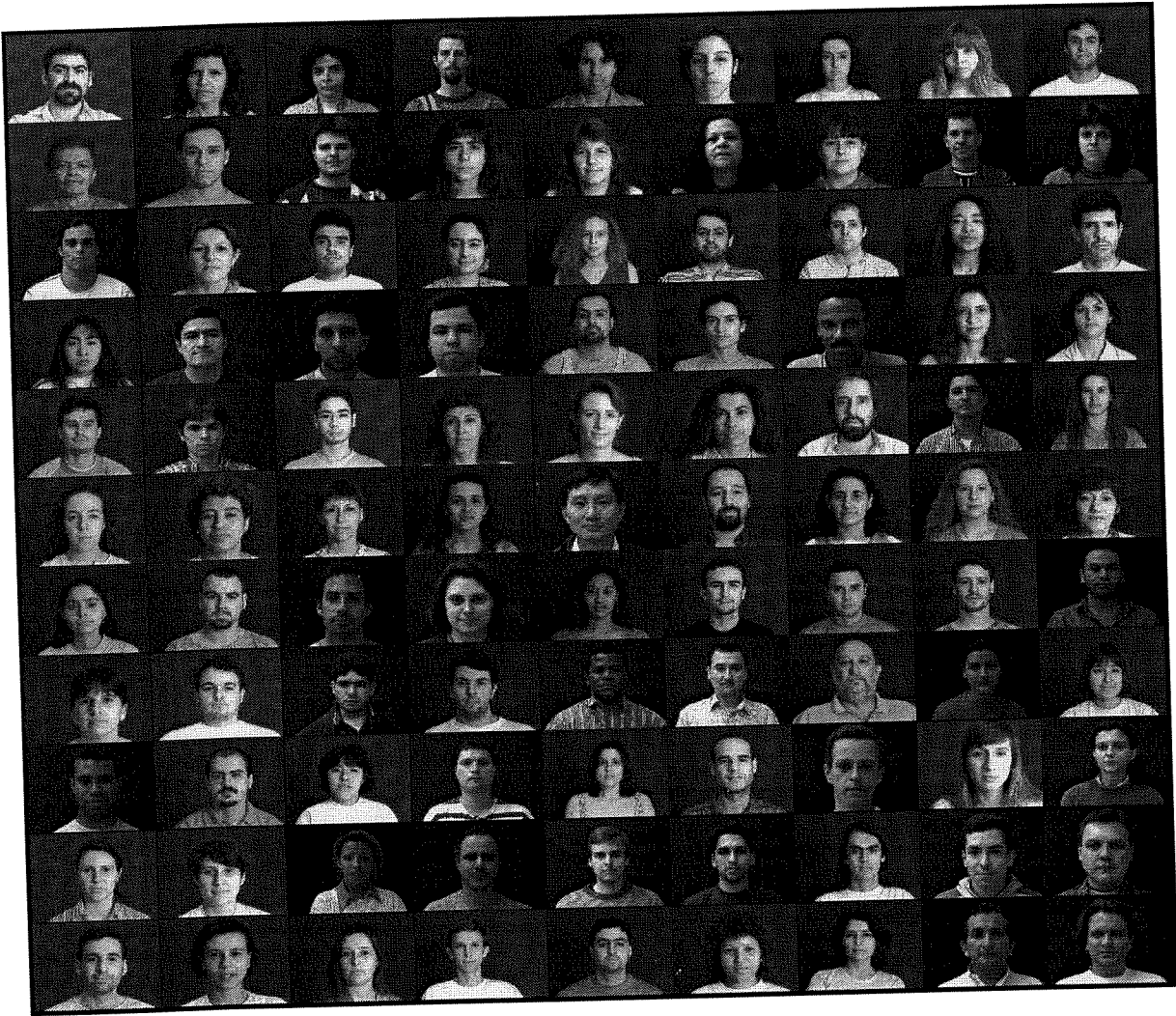


Figura 4-11: Imagens selecionadas para a construção do modelo de 98 *eigenfaces*.

Este tipo de face foi escolhido como modelo por se apresentar como o tipo mais semelhante às utilizadas em documentos de identificação como passaportes e carteiras de identidade. Podemos então avaliar a aplicação dos métodos de reconhecimento com

eigenfaces utilizando imagens provenientes destes tipos de documentação.

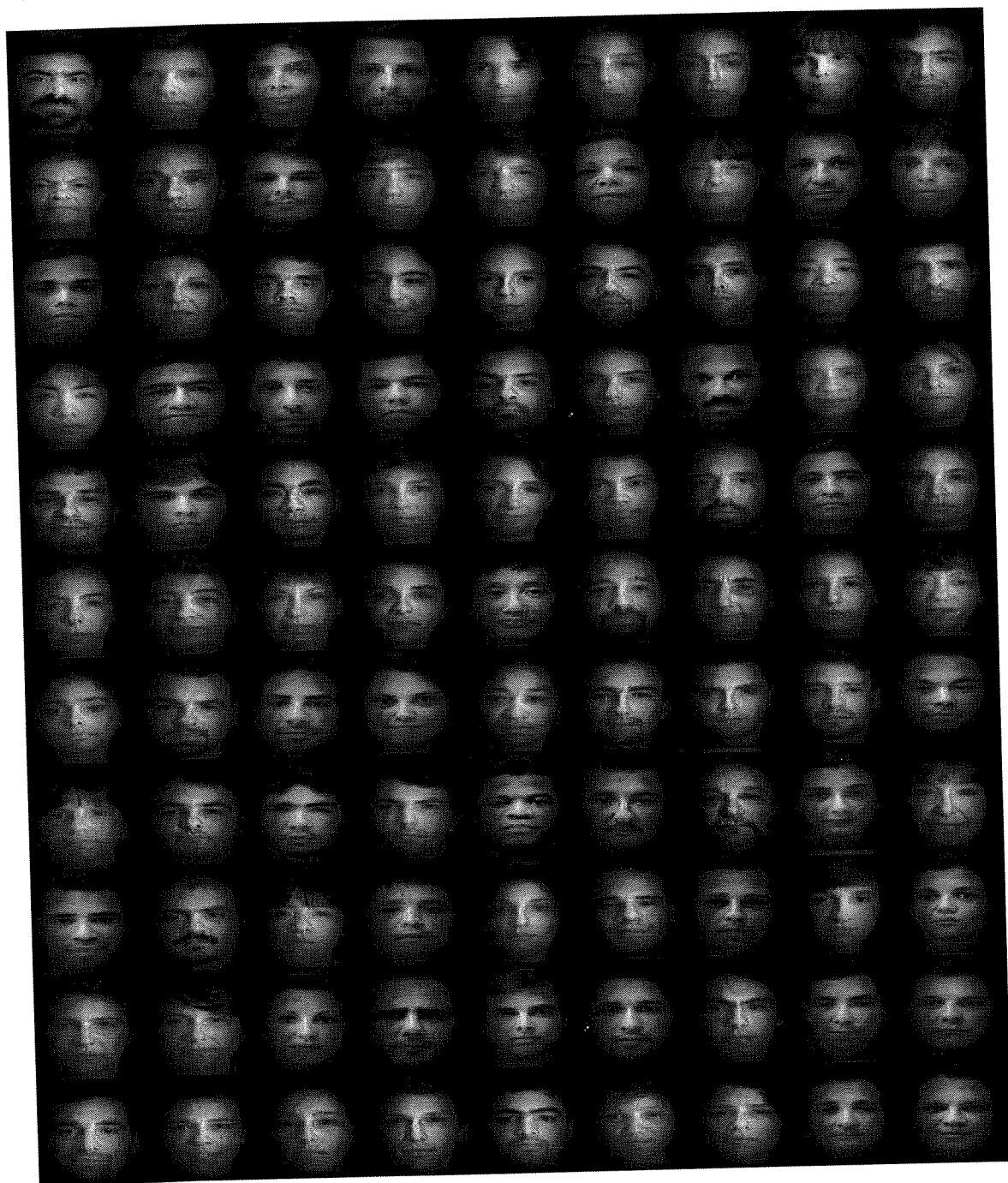


Figura 4-12: Imagens selecionadas para obtenção do modelo após normalização sem pré-processamento.

Na Figura 4-12 podemos observar que, mesmo após a normalização para posição e distância interocular padrão, existem variações em relação ao tamanho da área facial ocupada pela imagem. Isto é devido às grandes variações de escala das faces presentes na

base de dados adquirida. Estas variações devem estar presentes em situações de operação real para um sistema de reconhecimento, mas talvez em um grau bem menor do que o apresentado nesta base de dados. Um segundo conjunto de imagens treino, Figura 4-13, foi construído para fins de comparação com o conjunto de 99 imagens de treino.

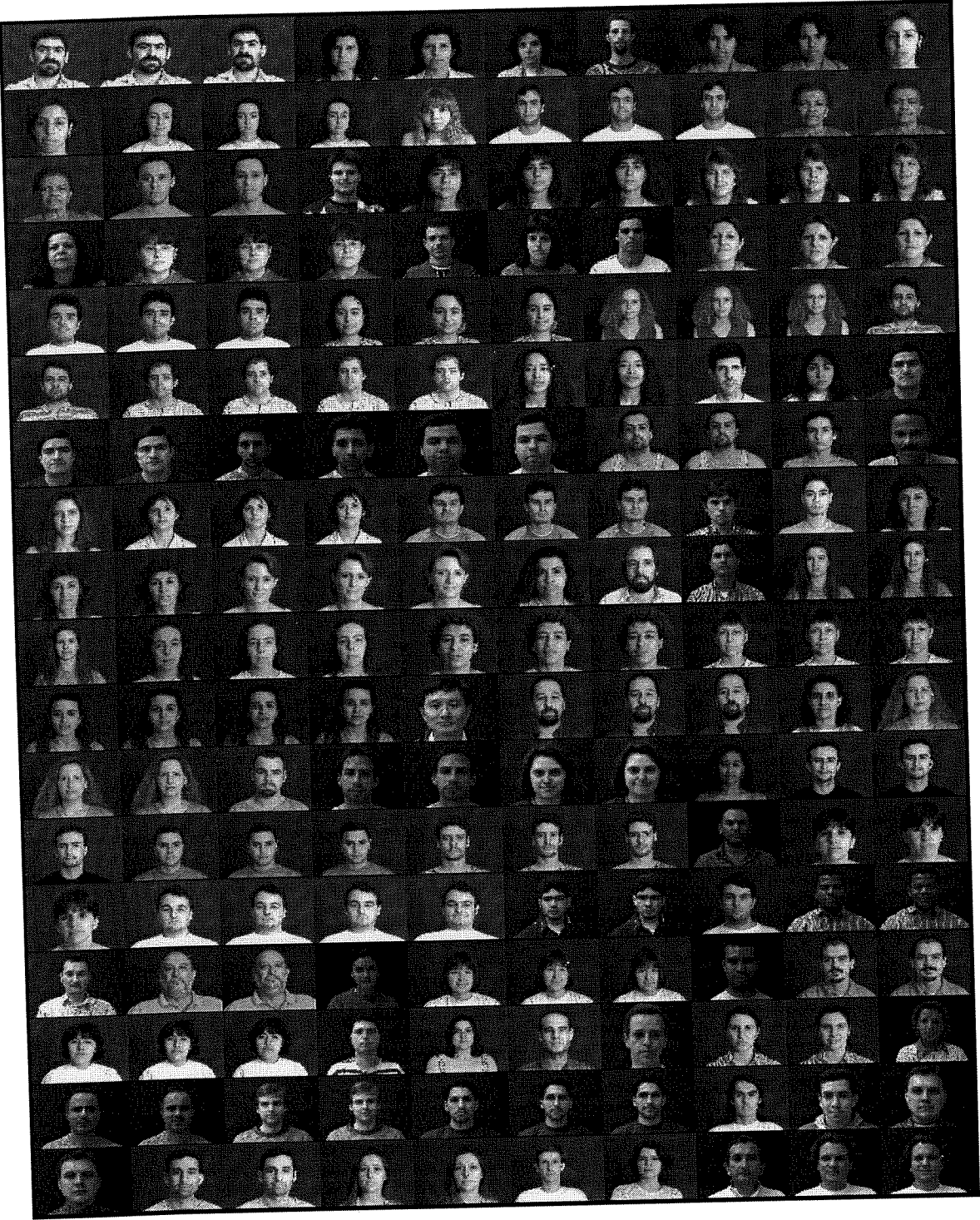


Figura 4-13: Imagens selecionadas para a construção do modelo de 178 *eigenfaces*.

Este conjunto é composto de 179 imagens exemplo com as mesmas características em relação a posição e expressão do conjunto de 99 imagens exemplo. Obtemos para este conjunto de imagens de treino 178 *eigenfaces* também construindo, como no modelo anterior, modelos separados para as imagens exemplo com e sem pré-processamento. Os detalhes das comparações entre os modelos de 98 e 178 *eigenfaces* são mostrados no Capítulo 5.

Para o cálculo do modelo de *eigenfaces*, foi implementado um algoritmo que realiza as operações necessárias para a obtenção das *eigenfaces* sobre as imagens já normalizadas. O algoritmo utilizado para o cálculo do modelo de *eigenfaces* é o seguinte:

1. Conversão das imagens de treino para a forma de vetores linha;
2. Construção da matriz de covariância dos vetores de imagens;
3. Utilização do método de Jacobi [PRESS] para o cálculo dos autovalores e autovetores da matriz de covariância;
4. Ordenação dos autovetores em ordem decrescente da variância do conjunto de treino associada a cada autovetor;
5. Soma da contribuição das imagens da base de dados sobre os autovetores da matriz de covariância das imagens obtendo as *eigenfaces* (Equação 3-12).

As rotinas necessárias para o cálculo do modelo foram implementadas de modo a gerar ao final do processamento o conjunto das *eigenfaces* e a imagem média da base de dados. Os dados gerados são armazenados para sua utilização pelo estágio de extração de características. Em nosso primeiro modelo, foram utilizadas 99 imagens de treino. Foram obtidos os modelos para as imagens de treino com e sem pré-processamento. Obtivemos então 98 *eigenfaces* para o modelo da informação facial. Elas representam o número máximo de características significativas que podem ser extraídas pelo sistema. Observamos na Figura 4-14 as imagens das *eigenfaces* ordenadas em ordem decrescente de seus autovalores, que correspondem à variância que cada autovetor detêm da base de

dados. No canto inferior direito da Figura 4-14, podemos observar também a imagem média da base de dados.

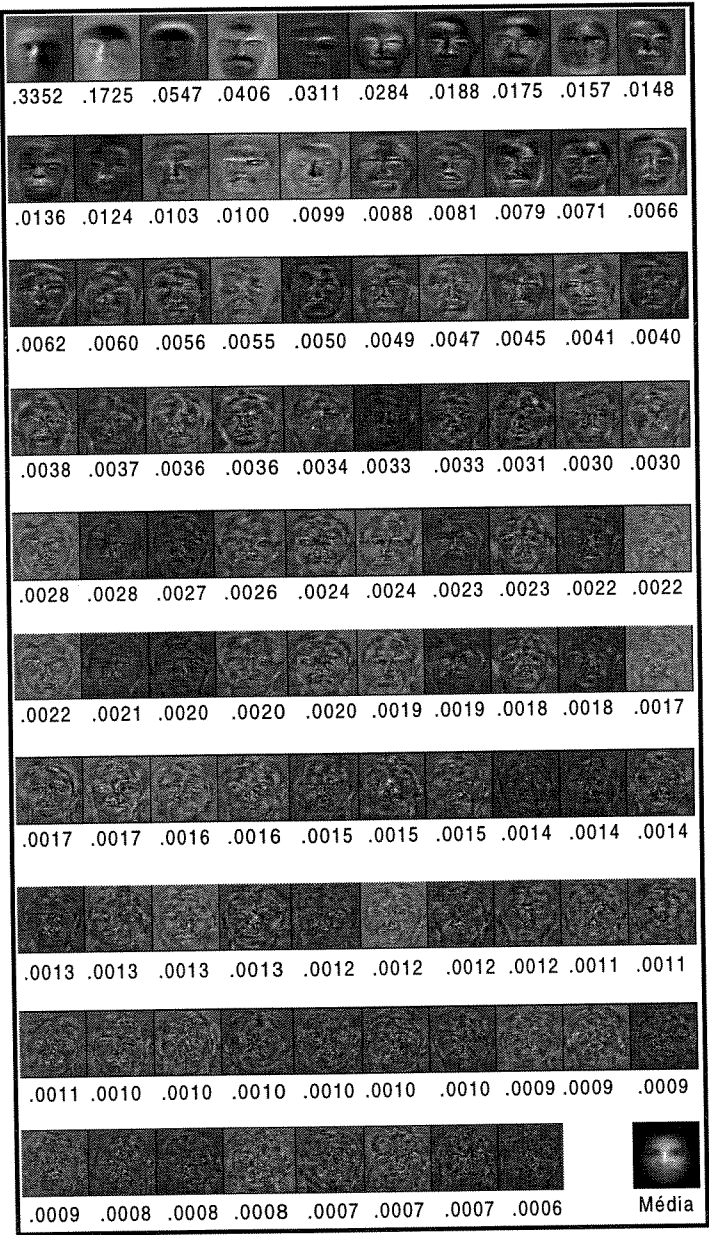


Figura 4-14: Modelo de *eigenfaces* obtido das 99 imagens de treino selecionadas sem pré-processamento. As imagens estão ordenadas em ordem decrescente da proporção da variância normalizada associadas a suas *eigenfaces*. A imagem no canto superior esquerdo corresponde a *eigenface* com maior variância. No canto inferior direito está a imagem média das imagens de treino.

As *eigenfaces* obtidas podem ser interpretadas da seguinte forma [VALENTIN-c]: as imagens com maiores autovalores detêm a maior parte da informação da base de dados. Estas *eigenfaces* representam as variações que são comuns à maior parte das imagens presentes na base de dados de treino. Estas componentes com maior variância guardam principalmente informações sobre a variação em relação a iluminação e sobre a forma da maioria das faces. Podemos observar isto olhando para a primeira linha de imagens da Figura 4-14. Nela vemos claramente imagens de faces que não possuem muitas variações. Estas *eigenfaces* não contêm, portanto, informações relevantes à diferenciação de faces individuais. Mas contêm informação importante para a determinação de um objeto como sendo uma face.

As componentes de ordem mais elevada, últimas linhas da Figura 4-14, são responsáveis, proporcionalmente, por uma parte menor da variância da base de dados. Por outro lado detêm as informações que diferenciam as faces entre si, o que se percebe na aparência ruidosa destas *eigenfaces* na Figura 4-14. Estas componentes são, deste modo, mais significativas para a diferenciação entre faces individuais que as componentes com maior variação. Atualmente, a obtenção do modelo de *eigenfaces* para as imagens normalizadas, com resolução de 64 x 64 pixels, é da ordem de 15 minutos, para as 98 *eigenfaces*, em um PENTIUM 100 MHz com 32Mb de RAM.

4.2.3 Extração de Características

A extração das características das imagens faciais é efetuada projetando-se as imagens faciais sobre todas as *eigenfaces* do modelo, segundo a Equação 3-13. A operação de projeção pode ser interpretada como o produto vetorial do vetor da imagem com o vetor da *eigenface*. Cada projeção sobre uma *eigenface* fornece um coeficiente. Este coeficiente caracteriza a imagem no espaço de face descrito pelo conjunto de *eigenfaces*. O conjunto de *eigenfaces* é escolhido para representar um determinado modelo facial. No

nosso caso, este modelo é frontal. Os resultados da projeção são armazenadas na forma de vetores com os valores ordenados segundo a mesma ordenação das *eigenfaces*, ou seja, ordem decrescente da variância proporcional da base de dados associada a cada *eigenface*. Os dados são armazenados em arquivos separados para cada imagem. Posteriormente serão utilizados pelo estágio de classificação. A extração de características para uma imagem facial demora aproximadamente 0.5 segundo para todas as 98 componentes em um microcomputador PENTIUM de 100 MHZ com 32 Mb de RAM.

4.2.4 Classificação

O módulo de classificação implementa um classificador de distância mínima. Este recebe como dados todos os centros das classes correspondentes aos indivíduos presentes na base de dados. Estes indivíduos são os que devem ser reconhecidos pelo sistema. Quando apresentamos uma nova imagem ao sistema, a distância euclidiana de seu vetor de características aos vetores que definem cada centro de classe é calculada e a classe com menor distância é considerada como a classe à qual a imagem de entrada pertence. Efetuamos o cálculo desta distância euclidiana no espaço descrito pelas *eigenfaces* da base de dados de treino. A ordem deste espaço é bem menor que a ordem do número de pixels presentes na imagem. Este fato reduz em muito a complexidade da tarefa de casamento e permite um aumento na velocidade de processamento.

No processo de reconhecimento não empregamos um limiar de aceitação. Portanto, para os testes do sistema de reconhecimento não existem critérios para rejeitar imagens. Este limiar de aceitação é implementado para a análise da performance do sistema de verificação. O processo de classificação de uma face contra 102 classes disponíveis na base de dados demora aproximadamente .5 segundo em um microcomputador PENTIUM de 100 MHz com 32 Mb de RAM.

4.3 Reconstrução de Imagens com *Eigenfaces*

A reconstrução de imagens com *eigenfaces* é um modo de avaliar o desempenho das características utilizadas para representar a informação facial no modelo de *eigenfaces*. Definimos a distância ao espaço de face descrito pelas *eigenfaces* selecionadas na Equação 3-16. Para ilustrarmos a qualidade de reconstrução das imagens segundo o modelo de *eigenfaces*, selecionamos diferentes tipos de imagens mostradas na Figura 4-15.

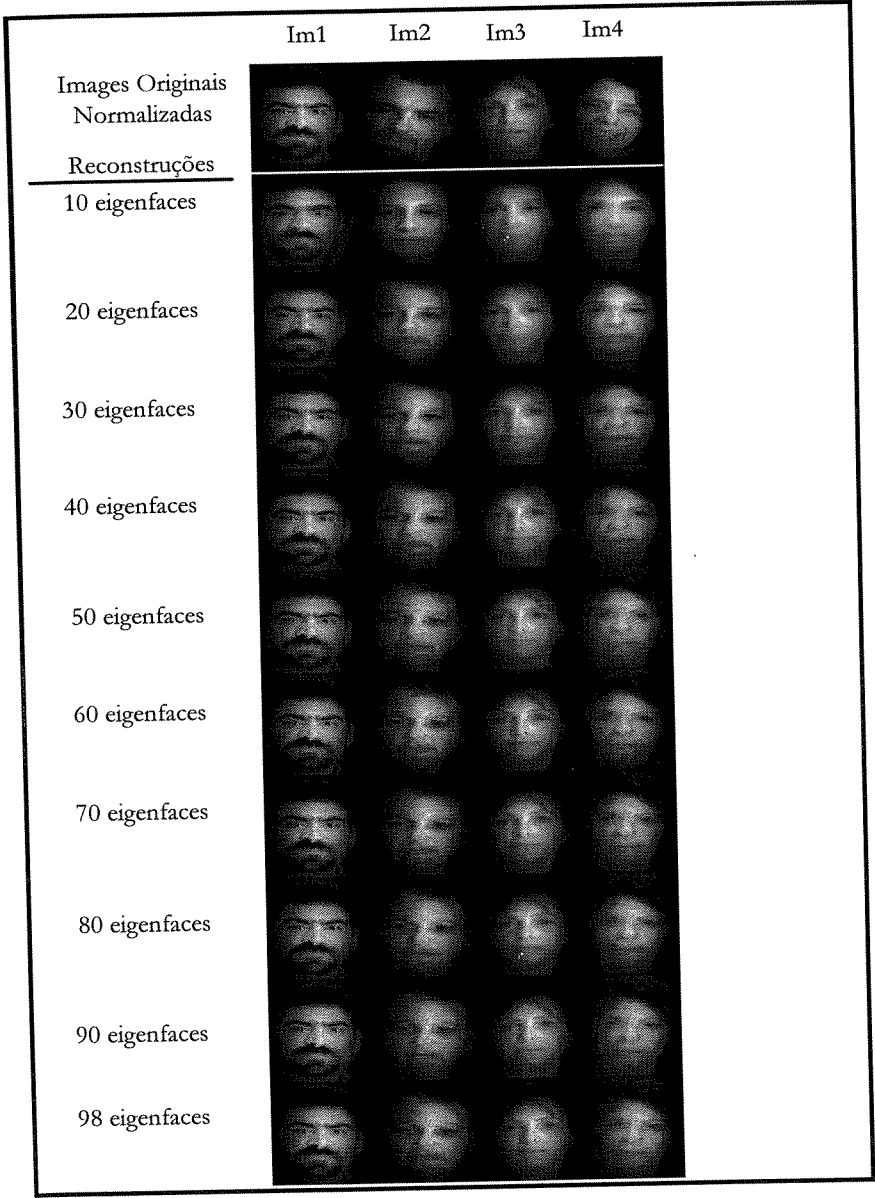


Figura 4-15: Reconstrução das imagens originais com diferentes quantidades de *eigenfaces*.

Distância do Espaço de Face x Número de <i>Eigenfaces</i>				
	Distância do Espaço de Face (normalizada pelo número de pixels da imagem)			
Número de <i>Eigenfaces</i>	Im1	Im2	Im3	Im4
10	0.1754170	0.283990	0.1453280	0.311085
20	0.0769745	0.249897	0.1373080	0.288446
30	0.0560184	0.238402	0.1203120	0.276339
40	0.0370397	0.232372	0.1071030	0.271768
50	0.0233832	0.229182	0.1018370	0.264228
60	0.0175802	0.224907	0.0961485	0.261601
70	0.0126939	0.224099	0.0934132	0.259495
80	0.0073542	0.221209	0.0888659	0.256424
90	0.0029621	0.218479	0.0862066	0.254472
99	0.0075630	0.217541	0.0845457	0.252380

Tabela 4-3: Análise da variação da distância ao espaço de face do modelo de *eigenfaces* com diferentes tipos de face para as imagens da Figura 4-15.

A imagem IM1 foi empregada na construção do modelo de *eigenfaces* e IM2 é uma imagem com rotação. A imagem IM3 é semelhante às empregadas no modelo e IM4 é uma imagem com rotação e sorriso. Obtemos a Tabela 4-3 comparando a distância do espaço de face, descrito pelas *eigenfaces* mostradas na Figura 4-14, para diferentes imagens selecionadas e varias quantidades de *eigenfaces*. As *eigenfaces* estão sempre ordenadas na ordem decrescente da proporção da variância associada com a base de dados que cada uma possui. Uma seleção de 10 *eigenfaces* representa estarmos trabalhando com as *eigenfaces* de número 1 a 10, uma seleção de 20 e temos as *eigenfaces* ordenadas de 1 a 20, e assim sucessivamente. Da análise da Tabela 4-3, observamos que IM1, que estava presente na construção do modelo de *eigenfaces*, apresenta as menores distâncias ao espaço de faces descrito por diferentes número de *eigenfaces* do modelo quando comparada com as outras imagens. IM3 é uma imagem semelhante às usadas para o treino do modelo frontal. Deste modo também apresenta uma pequena distância do espaço de faces descrito pelo modelo. A imagem IM2 é uma imagem rotacionada e IM4 é uma imagem com sorriso e rotação. Suas distâncias em relação ao espaço de faces do modelo, descrito por imagens

frontais e com expressão neutra são, portanto, bem maiores.

Na Figura 4-15, temos as imagens originais (IM1,IM2,IM3 e IM4) e a qualidade de sua reconstrução utilizando uma quantidade crescente de *eigenfaces*. Podemos observar que a qualidade visual da reconstrução melhora com o aumento no número de *eigenfaces*. A reconstrução da imagem é obtida somando-se a contribuição ponderada de cada *eigenface* que caracteriza a imagem e depois adiciona-se a imagem média da base de dados obtida na construção do modelo (Equação 3-18). As imagens com menor distância do modelo (IM1 e IM3) mostram uma boa qualidade reconstrução. As imagens com maior distância (IM2 e IM4) não apresentam uma boa reconstrução e estão bem distorcidas mesmo com o maior número de *eigenfaces* usados para a representação. IM2, em que o indivíduo se apresenta rotacionado, e IM4, em que o indivíduo está sorrindo e rotacionado, não são perfeitamente reconstruídos. O modelo facial não foi criado para representar características como sorriso e rotação. Esta deformação na reconstrução pode ser notada claramente nas distorções presentes nas imagens reconstruídas para IM2 e IM4 da Figura 4-15. Estes resultados confirmam os resultados obtidos em [PENTLAND-c]. Neste trabalho é sugerida a aplicação do modelo de *eigenfaces* com imagens de treino modelando posições específicas da face na imagem (frontal, inclinada e perfil), determinando primeiramente com qual destes modelos a imagem de entrada mais se aproximava, para posteriormente realizar reconhecimento segundo este modelo. Na Tabela 4-3 temos também uma confirmação dos resultados que medem a distância de imagens a modelos faciais obtidos com diferentes expressões (normal, riso, tristeza, surpresa, etc...). Observamos na Tabela 4-3, por exemplo, uma grande distância entre a imagem reconstruída com sorriso do modelo com expressão neutra. Os resultados da análise da distância a um determinado espaço de face são então uma boa medida para determinar que tipo de variação, em relação a posição ou expressão uma face apresenta, sem informações a priori sobre as características da imagem de entrada.

4.4 Discussão

Descrevemos a obtenção da base de dados e os blocos utilizados na construção dos sistemas de reconhecimento e verificação facial, salientando as características de reconstrução do modelo. O tempo total de processamento de uma imagem facial pelo sistema é de aproximadamente 3 segundos sem pré-processamento e 7 segundos com pré-processamento no microcomputador PENTIUM 100MHz. Mas isto não inclui os estágio de procura dos olhos na face que é efetuado manualmente por um aplicativo não integrado ao sistema. No próximo capítulo, faremos uma análise dos resultados obtidos com o modelo de *eigenfaces* descrito neste capítulo e com a base de dados adquirida.

Capítulo 5

5. Resultados do Modelo de *Eigenfaces*

Neste capítulo são apresentados os resultados de aplicação do modelo de *eigenfaces* descrito no Capítulo 4. O modelo obtido é empregado nas tarefas de reconhecimento e verificação facial. Realizamos uma comparação do desempenho dos sistemas para um número variado de características. São mostrados os fatores que mais contribuíram para os erros do sistema de reconhecimento.

5.1 Sistema de Reconhecimento Facial

O problema do reconhecimento consiste em, dada uma imagem facial, determinarmos dentre N classes faciais, a qual classe ela pertence, sem nenhuma informação a priori. Para avaliarmos o desempenho do sistema de reconhecimento facial com o modelo de *eigenfaces* utilizamos um variado número de *eigenfaces*. Na classificação,

comparamos a utilização de dois tipos de centros de classe. Um dos centros de classe é definido pelas características de uma face exemplo para cada indivíduo. O outro centro de classe é composto pela média aritmética de todas as características do indivíduo. Os testes de reconhecimento são realizados com imagens normalizadas com e sem pré-processamento. Estabelecemos nos testes quatro grupos de imagens. Estes grupos são definidos para efeitos de análise dos erros em relação às variações presentes em cada imagem. A distribuição do número de imagens em cada grupo pode ser vista na Tabela 4-

2. Os quatro grupos para efeitos de análise dos erros são:

1. geral: composto de todas as faces presentes na base de dados de treino;
2. modelo: composto de faces frontais, sem óculos e com expressão neutra.
3. frontal: composto de faces frontais que poderiam apresentar óculos e expressão;
4. rotacionada: composto de faces que apresentam uma rotação, dentro ou fora do plano da imagem, que podem estar com óculos e apresentar expressão.

Estes conjuntos permitem avaliar os tipos de variações que mais afetam o desempenho do modelo de reconhecimento.

5.1.1 Determinação do Centro de Classe

O centro de classe representa a característica que identifica um indivíduo que deve ser reconhecido. Ele é utilizado para efeito de comparação entre os indivíduos. Os dois tipo de centros de classe avaliados foram:

1. as características de uma face arbitrariamente escolhida como representativa de um indivíduo;
2. a média aritmética das características faciais de todas as faces presentes na base de dados para um indivíduo.

Podemos observar os resultados comparando a Figura 5-1 e a Tabela 5-1 com a

Figura 5-4 e a Tabela 5-2 para a face arbitrária e para a média aritmética com imagens normalizadas sem pré-processamento, respectivamente. Notamos que o procedimento que usa a face arbitrária apresenta um desempenho inferior ao da média aritmética como centro de classe para todos os grupos definidos. Nos testes da face arbitrária como centro de classe foram usadas 333 imagens de testes. As outras 102 imagens são excluídas dos testes por representarem os centros de classe. Devido ao seu melhor desempenho, foi utilizada como centro de classe, para a comparação de características nos sistemas de reconhecimento e verificação facial, a média aritmética das características faciais (projeções sobre as *eigenfaces* do modelo facial).

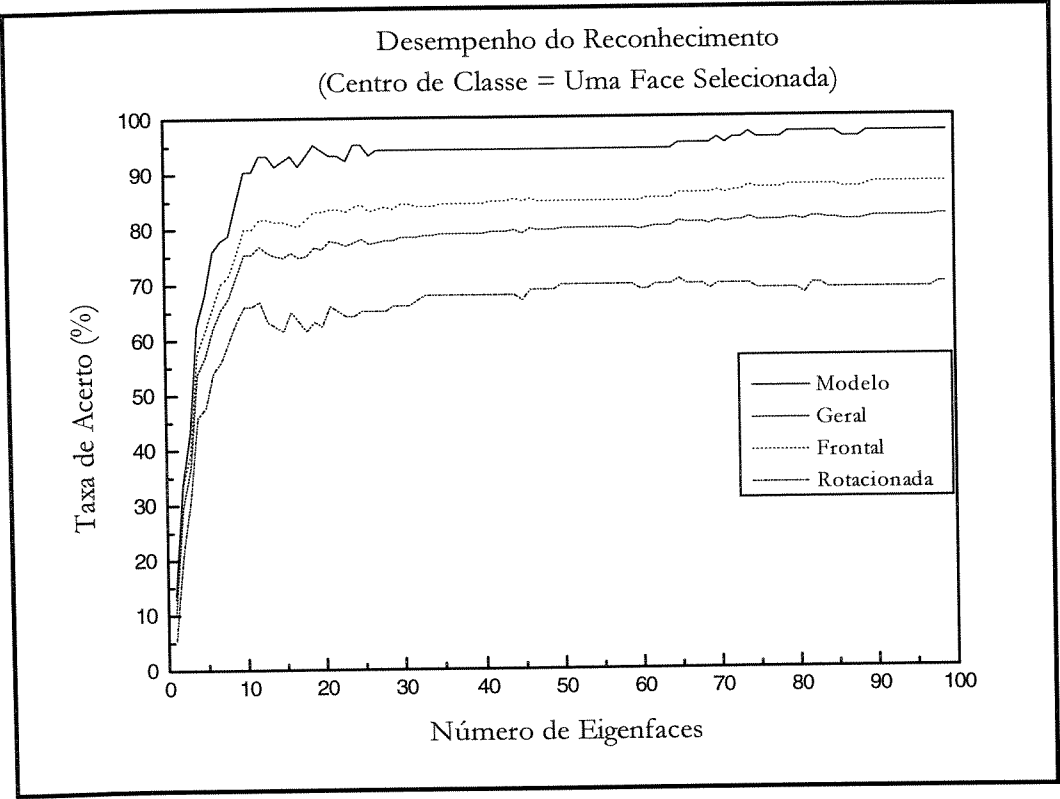


Figura 5-1: Variação da taxa de acerto de reconhecimento em função do número de características utilizadas para o modelo de 98 *eigenfaces* sem pré-processamento, com uma face arbitrária como centro de classe.

Número de <i>Eigenfaces</i>	Taxa de Acerto de Reconhecimento (%)			
	Geral	Modelo	Frontal	Rotacionada
1	10.5105	13.4615	12.9464	5.50459
10	75.3754	90.3846	79.9107	66.0550
20	76.2763	94.2308	83.0357	62.3853
30	78.3784	94.2308	84.3750	66.0550
40	78.9790	94.2308	84.3750	67.8899
50	79.8799	94.2308	84.8214	69.7248
60	79.5796	94.2308	84.8214	68.8073
70	81.0811	96.1538	86.6071	69.7248
80	81.3814	97.1154	87.5000	68.8073
90	81.6817	97.1154	87.9464	68.8073
98	81.9820	97.1154	87.9464	69.7248

Tabela 5-1: Valores do percentual de reconhecimento para o modelo de 98 *eigenfaces* sem pré-processamento, utilizando como centro de classe uma face arbitrariamente escolhida.

5.1.2 Determinação do Número de Características

Neste item, analisamos se o aumento do número de características usadas no sistema de reconhecimento melhora proporcionalmente a taxa de acertos. Para responder a esta questão, comparamos o modelo obtido anteriormente com 98 *eigenfaces* com outro modelo contendo 178 *eigenfaces*, onde ambos os modelos são analisados para os casos das imagens normalizadas com e sem pré-processamento. O modelo de 178 *eigenfaces* é obtido da mesma forma que o modelo anterior, sendo também composto de faces frontais, sem óculos e com expressão neutra, utilizando 179 exemplos dos 193 disponíveis. Primeiro, comparamos as curvas da Figura 5-2, observamos nas curvas que os modelos de 98 e 178 *eigenfaces* apresentam entre si uma distribuição semelhante na variância normalizada associada as componentes mais significativas (autovetores ou *eigenfaces* de ambos os modelos com maiores autovalores) para os casos com e sem pré-processamento. A proporção da variância normalizada de cada *eigenface* é definida como:

$$\sigma_i = \frac{\lambda_i}{\sum_{k=1}^{M-1} \lambda_k}$$

Equação 5-1

onde, λ_i é o autovalor do i -ésimo autovetor (*eigenface*) normalizado em relação à somatória de todos os autovalores dos autovetores significativos da base de dados de M imagens de treino [JOHNSON]. Entre os modelos de 98 e 178 *eigenfaces* sem pré-processamento, Figura 5-2 (a) e Figura 5-2 (b), e os modelos de 98 e 178 *eigenfaces* com pré-processamento, Figura 5-2 (c) e Figura 5-2 (d), constata-se que o pré-processamento acentua as diferenças entre as imagens de treino pois este realça as bordas das imagens faciais. Isto pode ser observado comparando-se a percentagem da variância normalizada das imagens de treino associadas aos autovetores (*eigenfaces*) de ambos os modelos construídos para estas imagens com e sem pré-processamento, mostrada na Figura 5-2.

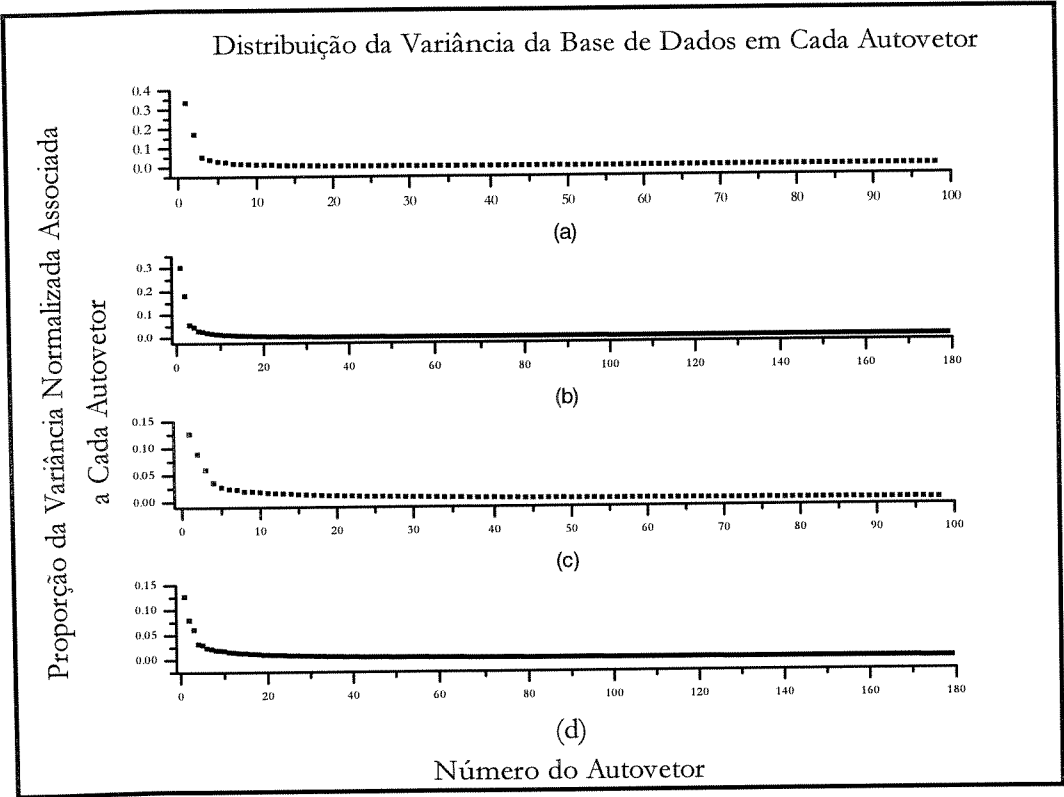


Figura 5-2: Distribuição da variância da base de dados associada a cada autovetor do modelo de *eigenfaces*. (a) modelo de 98 *eigenfaces* e (b) modelo de 178 *eigenfaces* sem pré-processamento. (c) modelo de 98 *eigenfaces* e (d) modelo de 178 *eigenfaces* com pré-processamento.

Os maiores valores de variância associados ao autovetor mais significativo são os dos modelos sem pré-processamento, em torno de 30% do total da variância normalizada do conjunto de imagens de treino, nos modelos sem pré-processamento este valor é da ordem de 10% da variância normalizada. Esta diferença implica em que as imagens pré-processadas possuem menor semelhança entre si, pois a percentagem de características comuns descrita por cada autovetor é menor nos modelos pré-processados, mostrando que o pré-processamento ressalta as diferenças entre as imagens de treino por salientar a informação de bordas e reduzir a contribuição da informação de textura das imagens faciais.

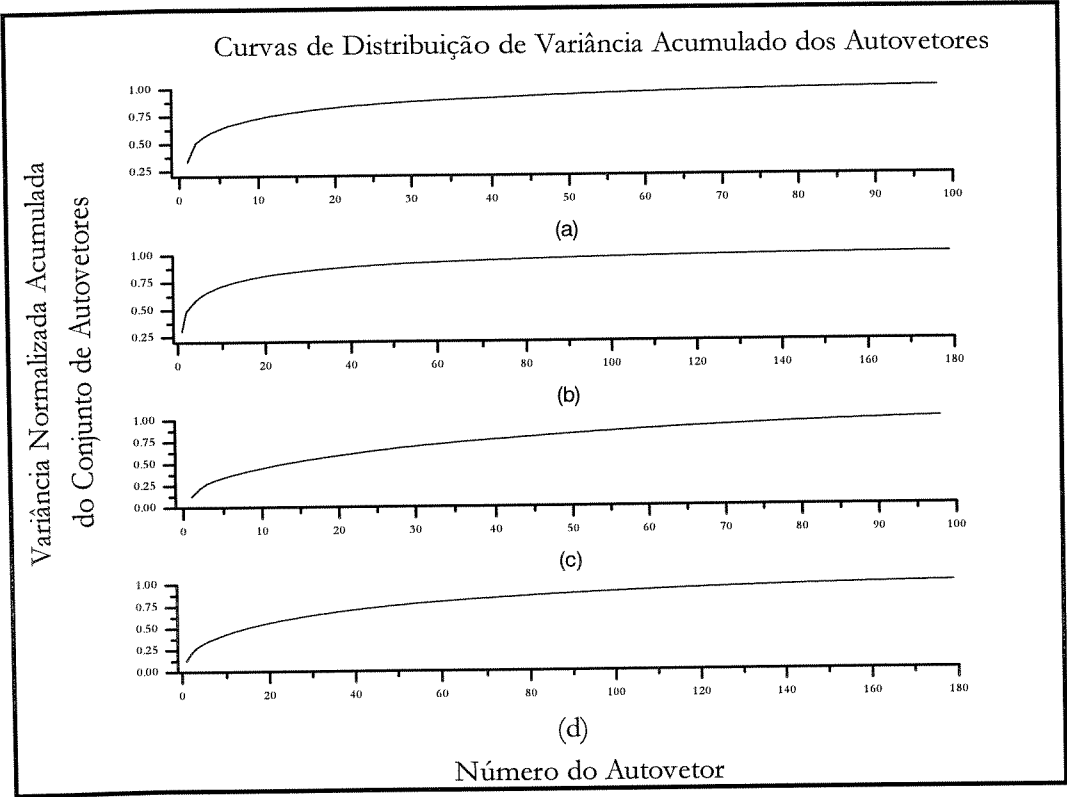


Figura 5-3: Variância acumulada nos autovetores do modelo de *eigenfaces*. (a) modelo de 98 *eigenfaces* e (b) modelo de 178 *eigenfaces* sem pré-processamento. (c) modelo de 98 *eigenfaces* e (d) modelo de 178 *eigenfaces* com pré-processamento.

Podemos também observar estas diferenças entre os modelos com e sem pré-processamento das imagens faciais na Figura 5-3, onde temos as curvas da distribuição acumulada da variância associada a cada autovetor (*eigenface*) do modelo. Nota-se que as imagens com pré-processamento, Figura 5-3 (c) e Figura 5-3 (d), possuem uma variância com distribuição mais uniforme entre os autovetores do modelo do que as imagens sem pré-processamento, Figura 5-3 (c) e Figura 5-3 (d). O fato do pré-processamento ressaltar as diferenças entre as imagens do modelo facial, mostrada na comparação das variâncias normalizadas obtidas com os modelos com e sem pré-processamento, colabora para melhorar o desempenho de reconhecimento reduzindo a similaridade dos padrões extraídos das imagens faciais.

Outro ponto a ser comparado entre os modelos de 98 e 178 *eigenfaces* são os desempenhos em relação à variação de taxa de reconhecimento com o número de características para cada modelo. Na comparação, utilizamos 435 imagens de teste para um total de 102 classes. Como centros de classe, foram adotadas as médias aritméticas de todas as características de cada indivíduo. A comparação dos modelos de 98 e 178 *eigenfaces* pode ser vista nas curvas das figuras 5-4 e 5-6 e nas tabelas 5-2 e 5-4, com as imagens sem pré-processamento, e nas curvas das figuras 5-5 e 5-7 e nas tabelas 5-3 e 5-5, com as imagens pré-processadas, respectivamente para o desempenho dos modelos com 98 e 178 *eigenfaces*. Percebe-se que não há diferenças significativas nestas desempenhos para os modelos de 98 e 178 *eigenfaces*, em ambos os casos com e sem pré-processamento o modelo de 178 *eigenfaces* apresenta um ganho na taxa de acerto de reconhecimento de menos de 1%. Entretanto, quando comparamos os modelos que utilizam imagens pré-processadas com aqueles que não as empregam, percebemos que o pré-processamento aumenta a taxa de acerto de reconhecimento dos modelos em torno de 3%, tornando o sistema mais preciso por salientar as diferenças entre as imagens faciais.

Na comparação dos resultados dos dois modelos para o grupo modelo, 193 exemplos, obtivemos as melhores taxas de reconhecimento; 98.44% para 98 *eigenfaces* e 98.96 % para 178 *eigenfaces*, com imagens sem pré-processamento, 98.96% para 98 *eigenfaces* e 98.96% para 178 *eigenfaces*, em imagens com pré-processamento. O grupo geral, 435 exemplos, apresentou um bom desempenho dada as grandes variações nele presentes; 94.48% de acertos para 98 *eigenfaces* e 94.71% para 178 *eigenfaces*, com imagens sem pré-processamento, 97.01% para 98 *eigenfaces* e 97.70% para 178 *eigenfaces*, em imagens com pré-processamento. No grupo frontal, 319 exemplos, tivemos também um bom desempenho, dado que ele possui variações em relação ao modelo utilizado para a representação como óculos e expressão; o desempenho foi de 95.61% de acertos para 98 *eigenfaces* e 96.23% para 178 *eigenfaces*, com imagens sem pré-processamento, 98.11% para 98 *eigenfaces* e 98.43% para 178 *eigenfaces*, em imagens com pré-processamento. O grupo que apresentou o pior desempenho é o das imagens que possuem rotação, 116 exemplos, pois dada a natureza de correlação do modelo de *eigenfaces*, não podemos representar adequadamente este tipo de faces com um modelo treinado com faces em posição frontal. Como consequência, teremos maior taxa de erro para as imagens com rotação. Para as imagens rotacionadas obtivemos 91.37% de acertos para 98 *eigenfaces* e 90.51% para 178 *eigenfaces*, com imagens sem pré-processamento, 93.96% para 98 *eigenfaces* e 95.69% para 178 *eigenfaces*, em imagens com pré-processamento. No entanto, não é necessário utilizarmos todas as *eigenfaces* dos modelos para termos um bom desempenho no reconhecimento, ligeiramente inferior ao máximo, como pode ser observado no comportamento de saturação das curvas nas figuras 5-4, 5-5, 5-6 4 e 5-7. Podemos então reduzir o número de características extraídas da imagem facial e ainda assim termos uma boa taxa de reconhecimento.

Nos testes de reconhecimento, todas as 102 classes (indivíduos) estavam registradas na base de dados como classes conhecidas. O problema de rejeitar uma face

como desconhecida da base de dados pelo sistema é analisado na verificação, onde determinam-se os limiares de aceitação dos indivíduos a serem autenticados.

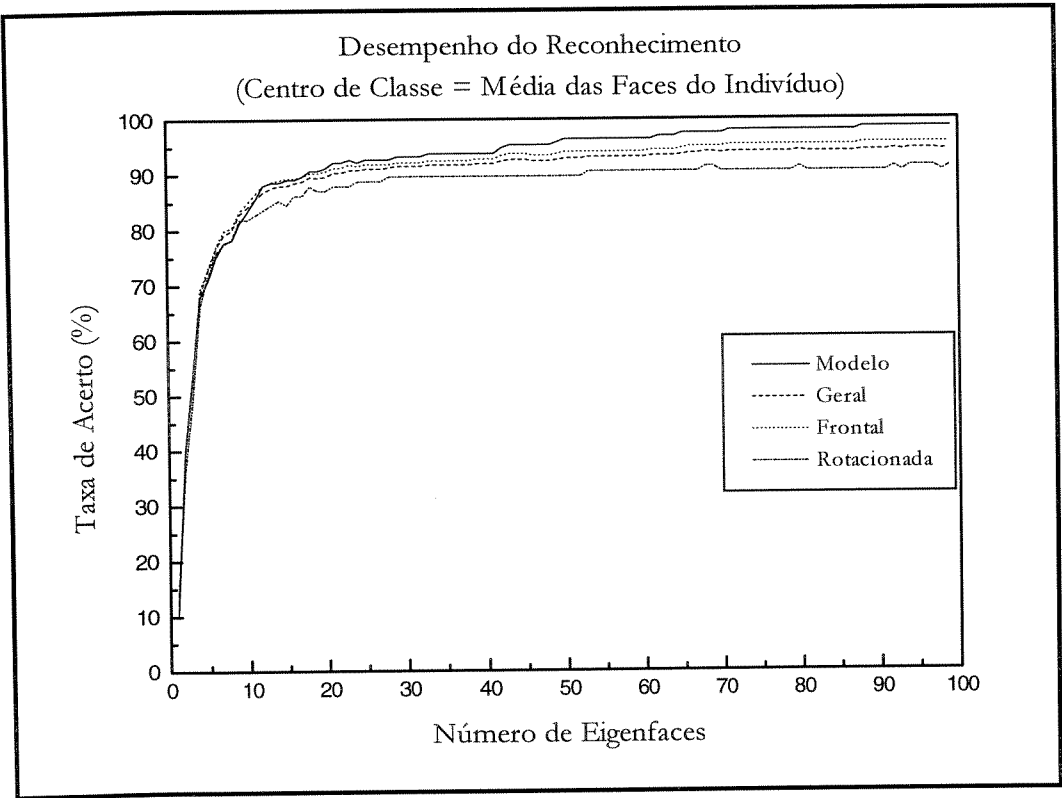


Figura 5-4: Variação da taxa de acerto de reconhecimento em função do número de características utilizadas para o modelo de 98 *eigenfaces* sem pré-processamento.

Número de Eigenfaces	Taxa de Acerto de Reconhecimento (%)			
	Geral	Modelo	Frontal	Rotacionada
1	12.6437	10.3627	12.2257	13.7931
10	84.1379	83.4197	84.953	81.8966
20	89.6552	91.1917	90.5956	87.069
30	91.4943	93.2642	92.163	89.6552
40	91.954	93.7824	92.79	89.6552
50	92.8736	96.3731	94.0439	89.6552
60	93.1034	96.3731	94.0439	90.5172
70	93.7931	97.4093	94.9843	90.5172
80	94.2529	97.9275	95.2978	91.3793
90	94.2529	98.4456	95.6113	90.5172
98	94.4828	98.4456	95.6113	91.3793

Tabela 5-2: Valores do percentual de reconhecimento para o modelo de 98 *eigenfaces* sem pré-processamento, utilizando como centro de classe a média das características.

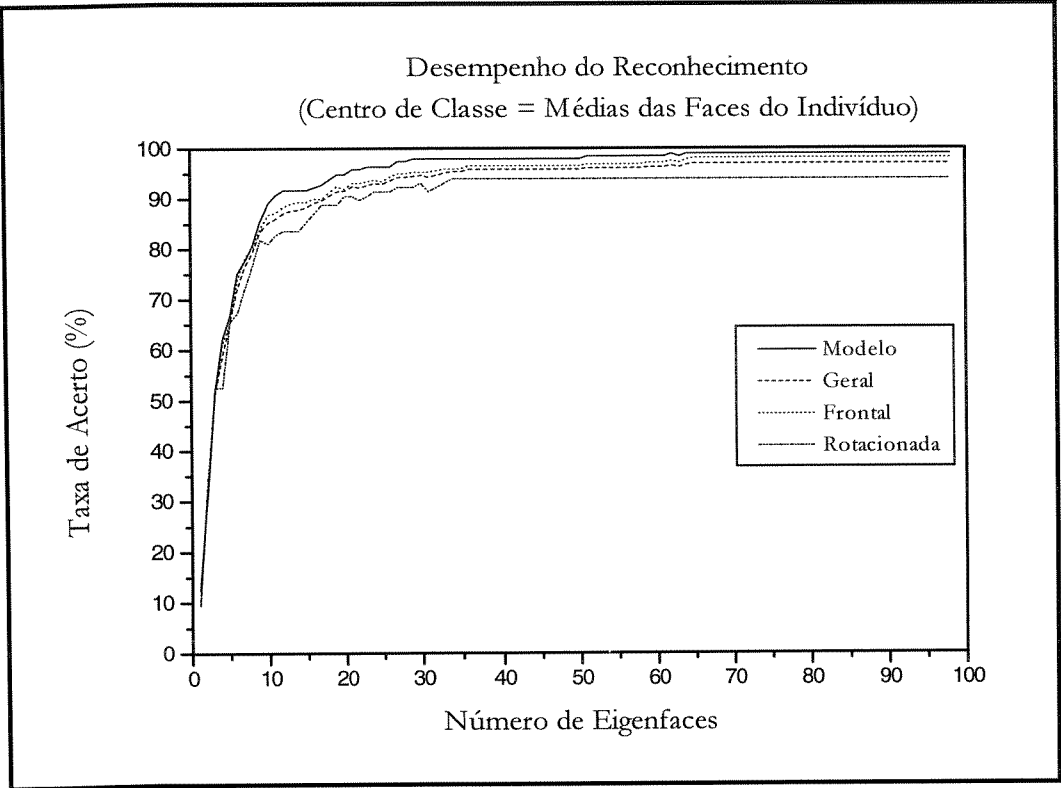


Figura 5-5: Variação da taxa de acerto de reconhecimento em função do número de características utilizadas para o modelo de 98 *eigenfaces* com pré-processamento.

Número de Eigenfaces	Taxa de Acerto de Reconhecimento (%)			
	Geral	Modelo	Frontal	Rotacionada
1	10.3448	12.4352	10.6583	9.48276
10	85.2874	89.1192	86.8339	81.0345
20	91.4943	94.8187	91.8495	90.5172
30	94.7126	97.9275	95.2978	93.1034
40	95.8621	97.9275	96.5517	93.9655
50	95.8621	97.9275	96.5517	93.9655
60	96.3218	98.4456	97.1787	93.9655
70	97.0115	98.9637	98.1191	93.9655
80	97.0115	98.9637	98.1191	93.9655
90	97.0115	98.9637	98.1191	93.9655
98	97.0115	98.9637	98.1191	93.9655

Tabela 5-3: Valores do percentual de reconhecimento para o modelo de 98 *eigenfaces* com pré-processamento, utilizando como centro de classe a média das características.

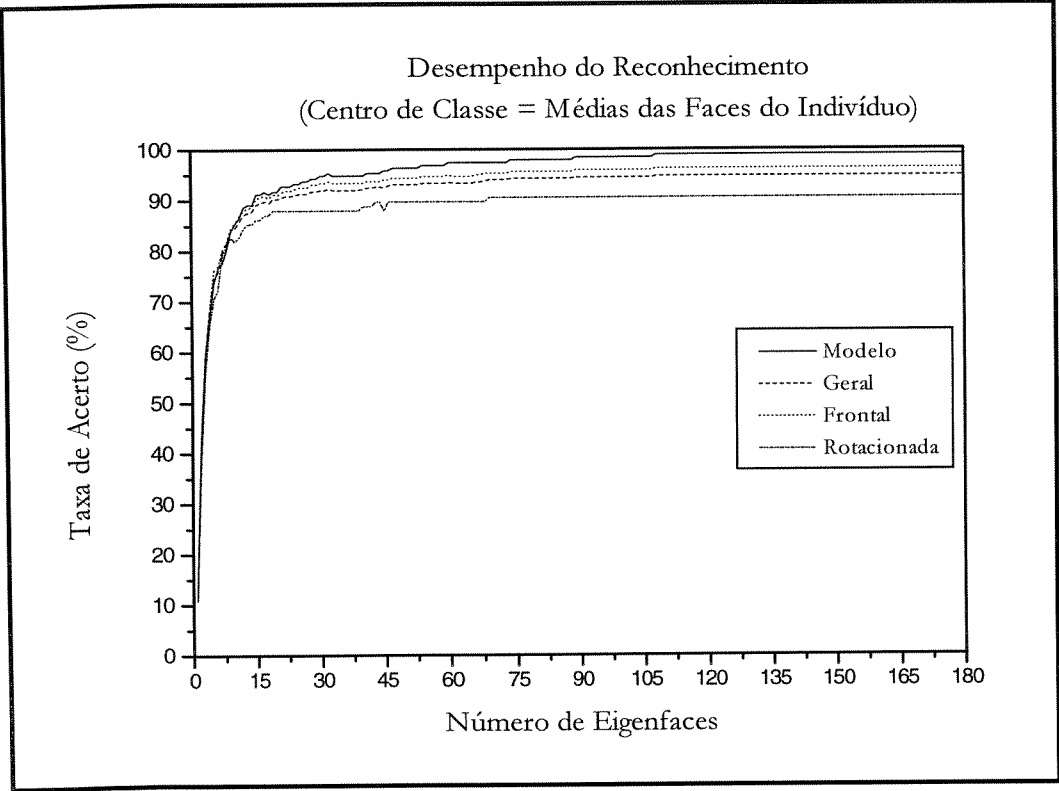


Figura 5-6: Variação do taxa de acerto de reconhecimento em função do número de características utilizadas para o modelo de 178 *eigenfaces* sem pré-processamento.

Número de <i>Eigenfaces</i>	Taxa de Acerto de Reconhecimento (%)			
	Geral	Modelo	Frontal	Rotacionada
1	13.1034	10.8808	13.7931	11.2069
10	84.5977	85.4922	85.5799	81.8966
20	90.1149	91.7098	90.9091	87.9310
30	91.954	94.8187	93.4169	87.9310
40	92.1839	94.8187	93.4169	88.7931
50	93.1034	96.3731	94.3574	89.6552
60	93.5632	97.4093	94.9843	89.6552
70	94.0230	97.4093	95.2978	90.5172
80	94.2529	97.9275	95.6113	90.5172
90	94.4828	98.4456	95.9248	90.5172
100	94.4828	98.4456	95.9248	90.5172
110	94.7126	98.9637	96.2382	90.5172
120	94.7126	98.9637	96.2382	90.5172
130	94.7126	98.9637	96.2382	90.5172
140	94.7126	98.9637	96.2382	90.5172
150	94.7126	98.9637	96.2382	90.5172
160	94.7126	98.9637	96.2382	90.5172
170	94.7126	98.9637	96.2382	90.5172
178	94.7126	98.9637	96.2382	90.5172

Tabela 5-4: Valores do percentual de reconhecimento para o modelo de 178 *eigenfaces* sem pré-processamento, utilizando como centro de classe a média das características.

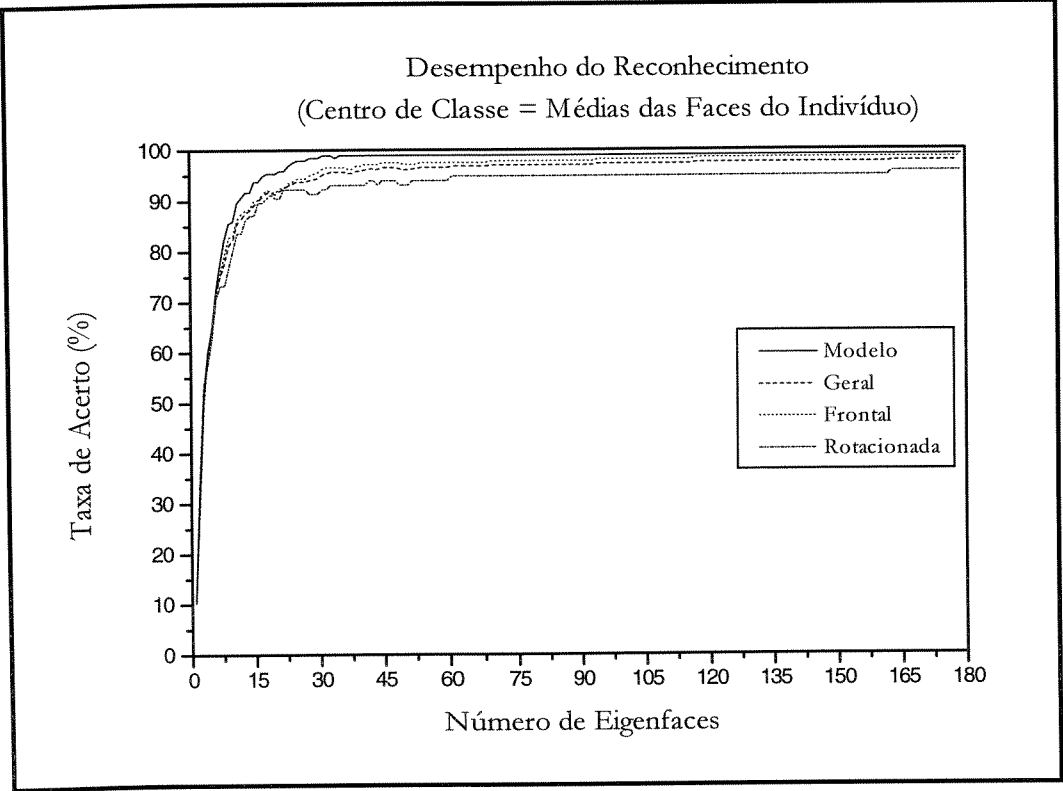


Figura 5-7: Variação do taxa de acerto de reconhecimento em função do número de características utilizadas para o modelo de 178 *eigenfaces* com pré-processamento.

Número de <i>Eigenfaces</i>	Taxa de Acerto de Reconhecimento (%)			
	Geral	Modelo	Frontal	Rotacionada
1	10.5747	10.8808	10.6583	10.3448
10	82.2989	86.0104	83.0721	80.1724
20	91.2644	95.3368	91.5361	90.5172
30	94.4828	98.4456	95.6113	91.3793
40	95.8621	98.9637	96.8652	93.1034
50	96.0920	98.9637	97.1787	93.1034
60	96.5517	98.9637	97.4922	93.9655
70	97.0115	98.9637	97.8056	94.8276
80	97.0115	98.9637	97.8056	94.8276
90	97.0115	98.9637	97.8056	94.8276
100	97.2414	98.9637	98.1191	94.8276
110	97.2414	98.9637	98.1191	94.8276
120	97.4713	98.9637	98.4326	94.8276
130	97.4713	98.9637	98.4326	94.8276
140	97.4713	98.9637	98.4326	94.8276
150	97.4713	98.9637	98.4326	94.8276
160	97.4713	98.9637	98.4326	94.8276
170	97.7011	98.9637	98.4326	95.6897
179	97.7011	98.9637	98.4326	95.6897

Tabela 5-5: Valores do percentual de reconhecimento para o modelo de 178 *eigenfaces* com pré-processamento, utilizando como centro de classe a média das características.

5.1.3 Análise dos Erros de Classificação

Na análise do tipo de característica presente na imagem, que acarreta erros no processo de classificação, empregando o classificador de distância mínima, utilizamos o número mínimo de características que provê a maior taxa de acerto no reconhecimento para um determinado modelo. Os valores do número mínimo de características empregadas para prover as maiores taxas de acerto são determinados observando os valores nas curvas de variação da taxa de erro em função do número de características empregadas (figuras 5-4, 5-5, 5-6 e 5-7). Este número de é 95 características (*eigenfaces*) para o modelo de 98 *eigenfaces* e 110 características para o modelo de 178 *eigenfaces*, com imagens sem pré-processamento, de 65 características para o modelo de 98 *eigenfaces* e de

174 características para o modelo de 178 *eigenfaces*, com imagens pré-processadas. Nas figuras 5-8 e 5-9 temos as imagens classificadas erroneamente nos modelos de 98 e 178 *eigenfaces*, respectivamente, com imagens sem pré-processamento, e nas figuras 5-10 e 5-11 temos os erros de classificação, com imagens pré-processadas para os modelos de 98 e 178 *eigenfaces*, respectivamente. O modelo de 98 *eigenfaces* cometeu 25 erros de classificação e o modelo de 178 *eigenfaces* 23 erros de classificação, sem pré-processamento, e o modelo de 98 *eigenfaces* cometeu 13 erros de classificação junto com 10 erros de *classificação* para o modelo de 178 *eigenfaces* com pré-processamento. Todos os modelos foram testados na classificação de 435 imagens em 102 classes. Para as imagens classificadas erroneamente nos modelos testados, temos a distribuição dos erros em relação ao tipo de imagem expressa na Tabela 5-2. Percebemos na tabela que a maior quantidade de erros é das imagens com óculos. Havia apenas 62 imagens com óculos na base de dados, os modelos sem pré-processamento erraram em 13 (20.97%) imagens e os modelos com pré-processamento erraram em 6 (9.77%) imagens. O baixo desempenho das imagens com óculos em relação a outras características ocorre porque a região dos olhos é uma das mais importantes para o processo de reconhecimento [CHELLAPA]. Qualquer reflexão presente nos óculos e a própria utilização dos mesmos altera muito a área dos olhos na imagem.

Grupo da imagem	Modelo			
	sem pré-processamento		pré-processado	
	98 <i>eigenfaces</i>	178 <i>eigenfaces</i>	98 <i>eigenfaces</i>	178 <i>eigenfaces</i>
modelo	3	2	2	2
óculos	13	13	8	6
óculos e rotação	1	1	-	-
expressão	1	1	-	-
expressão e rotação	6	4	2	2
rotação	1	2	1	-

Tabela 5-6: Número de erros associados a cada tipo de imagem.

Expressões e rotações foram tratadas satisfatoriamente pelo sistema. Apenas algumas imagens que apresentavam deformações muito acentuadas foram classificadas

erroneamente, o que demonstra a robustez do sistema às variações das imagens de entrada em relação às imagens empregadas na composição do modelo facial. Os erros do grupo de imagens modelo, composto das faces usadas para compor o modelo de *eigenfaces* e faces semelhantes a estas, foi bem baixo como esperado. Com um total de 192 exemplos houve apenas 3 erros para o modelo de 98 *eigenfaces* sem pré-processamento, 2 erros para o modelo de 98 *eigenfaces* com pré-processamento e 2 erros para os modelo de 178 *eigenfaces* com e sem pré-processamento. Os erros são devidos à posição relativa entre os centros de classe e as imagens de teste no espaço de face dos modelos. As imagens do grupo modelo classificadas incorretamente ficaram muito distantes do centro de classe correspondente. Isto é devido, em parte, à qualidade do processo de aquisição ter variado muito de imagem para imagem. No caso das duas imagens do grupo modelo que apresentaram erros em todos os modelos faciais tivemos uma acentuada variação de iluminação e também diferença de foco entre as imagens do indivíduo presentes na base de dados. Os modelos de 98 e 178 *eigenfaces* não apresentaram variações significativas para a taxa de reconhecimento e por isso escolhemos trabalhar com o modelo de 178 *eigenfaces*. Entretanto, escolhemos também o modelo de 178 *eigenfaces* por apresentar um erro de classificação ligeiramente menor que o modelo de 98 *eigenfaces*. Assim, o modelo de 178 *eigenfaces* será utilizado como núcleo de extração de características do sistema de reconhecimento facial e nos testes do sistema de verificação, apesar do mesmo elevar o custo computacional dos testes de verificação.

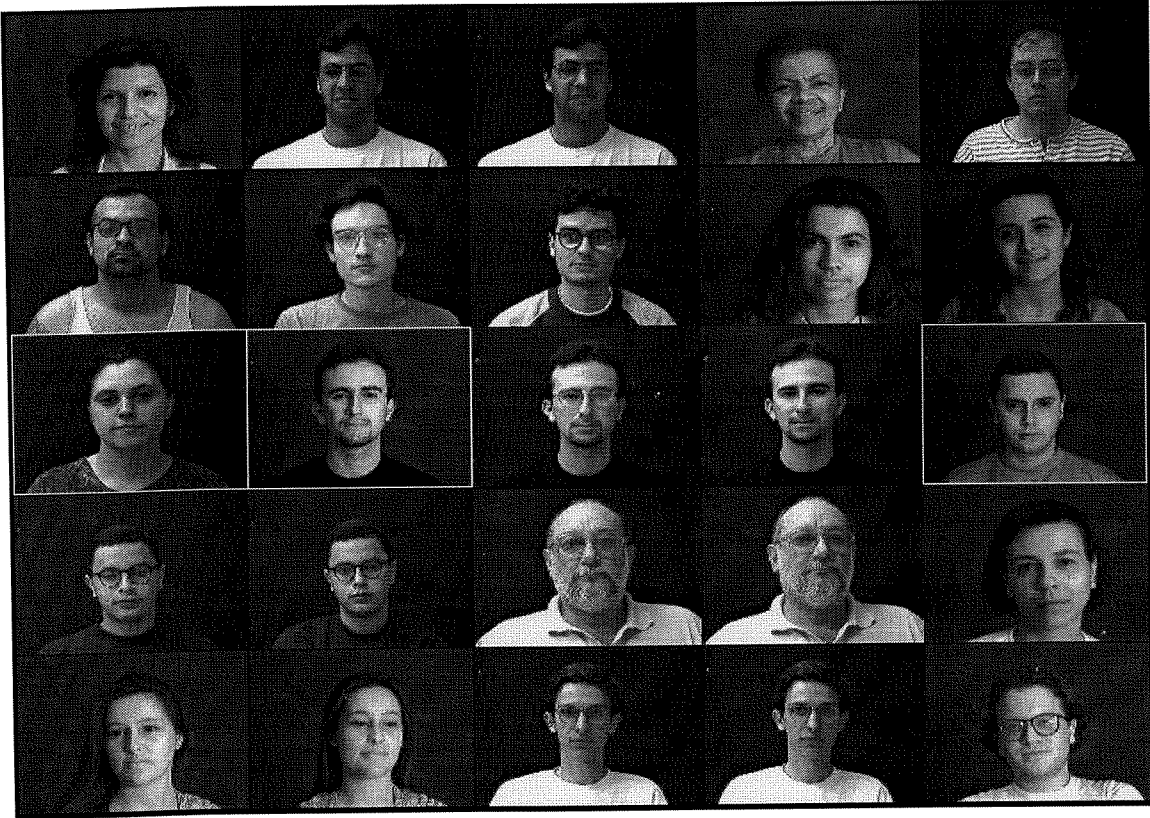


Figura 5-8: Imagens que apresentaram erros de classificação para o modelo de 98 *eigenfaces* sem pré-processamento com sua taxa de reconhecimento máxima. As imagens destacadas são os erros das faces do grupo modelo.

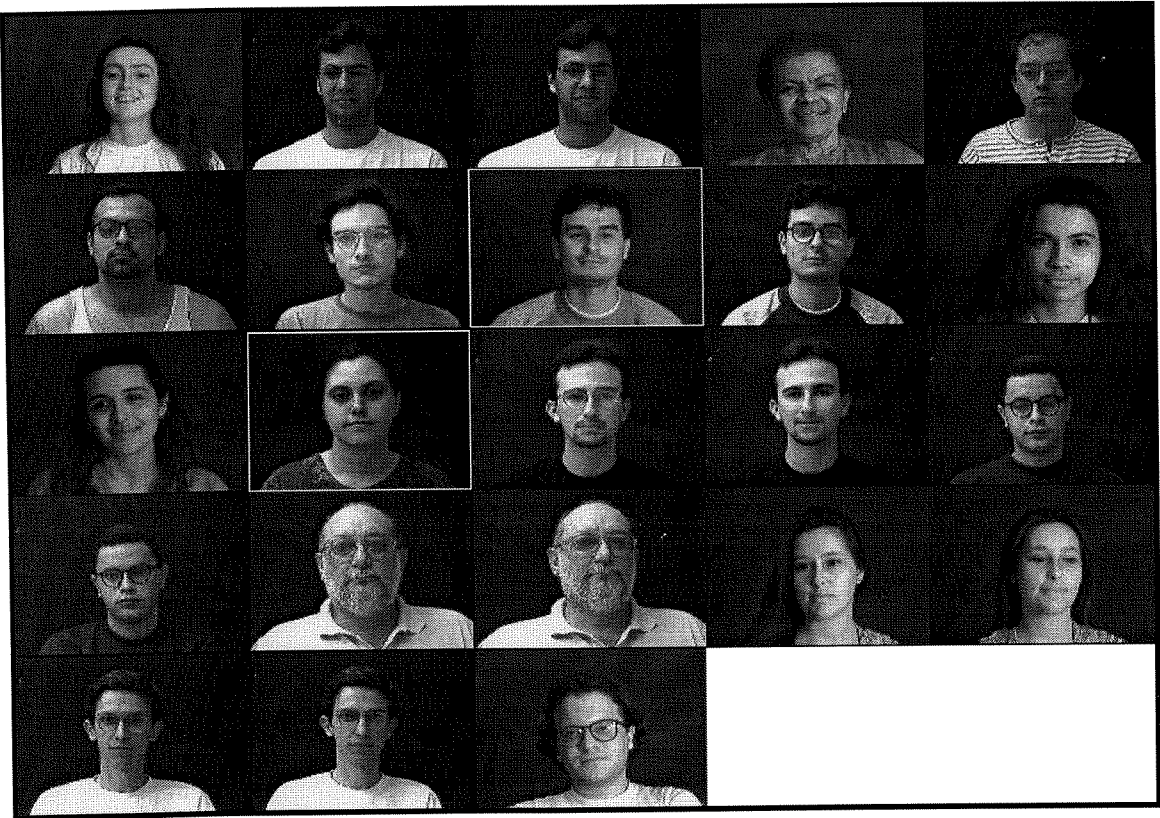


Figura 5-9: Imagens que apresentaram erros de classificação para o modelo de 178 *eigenfaces* sem pré-processamento com sua taxa de reconhecimento máxima. As imagens destacadas são os erros das faces do grupo modelo.



Figura 5-10: Imagens que apresentaram erros de classificação para o modelo de 98 *eigenfaces* com pré-processamento para sua taxa de reconhecimento máxima. As imagens destacadas são os erros das faces do grupo modelo.



Figura 5-11: Imagens que apresentaram erros de classificação para o modelo de 178 *eigenfaces* sem pré-processamento com sua taxa de reconhecimento máxima. As imagens destacadas são os erros das faces do grupo modelo.

5.1.4 Comparação com Outros Classificadores

Como forma de comparar o desempenho do classificador de distância mínima, que no melhor caso obteve um desempenho de 97,70% de acertos, com outros tipos de classificadores os dados do modelo de 98 *eigenfaces* foram empregados para treinar uma perceptron multi-camadas na tarefa de classificar corretamente as 435 imagens de 102 indivíduos. O treinamento do perceptron multi-camadas, foi realizado com os 102 centros de classe obtidos com a média aritmética das características de cada indivíduo, utilizando o algoritmo de QuickPropagation, o mesmo conseguiu obter 100 % de acertos no conjunto de treino. A arquitetura do perceptron multi-camadas empregado é de 65 neurônios na camada de entrada, 80 neurônios na camada escondida e 102 neurônios, um para cada classe, na camada de saída, todos os neurônios possuem função de ativação sigmóide. Os testes foram efetuados no conjunto de 435 imagens obtendo um taxa máxima de acerto de 94,71%, inferior portanto ao desempenho do classificador de distância mínima. As curvas de treinamento e teste em função do número de interações do algoritmo de treinamento podem ser observadas nas figuras 5-12 e 5-13 para a taxa de acerto e erro absoluto da rede respectivamente. O fato da rede neural não ter conseguido uma performance superior ao classificador de distância mínima pode ser explicado pela

natureza descorrelacionada e linearizada das características de componentes principais utilizadas na representação do modelo facial que são mais apropriadas para serem tratadas por um método linear de classificação.

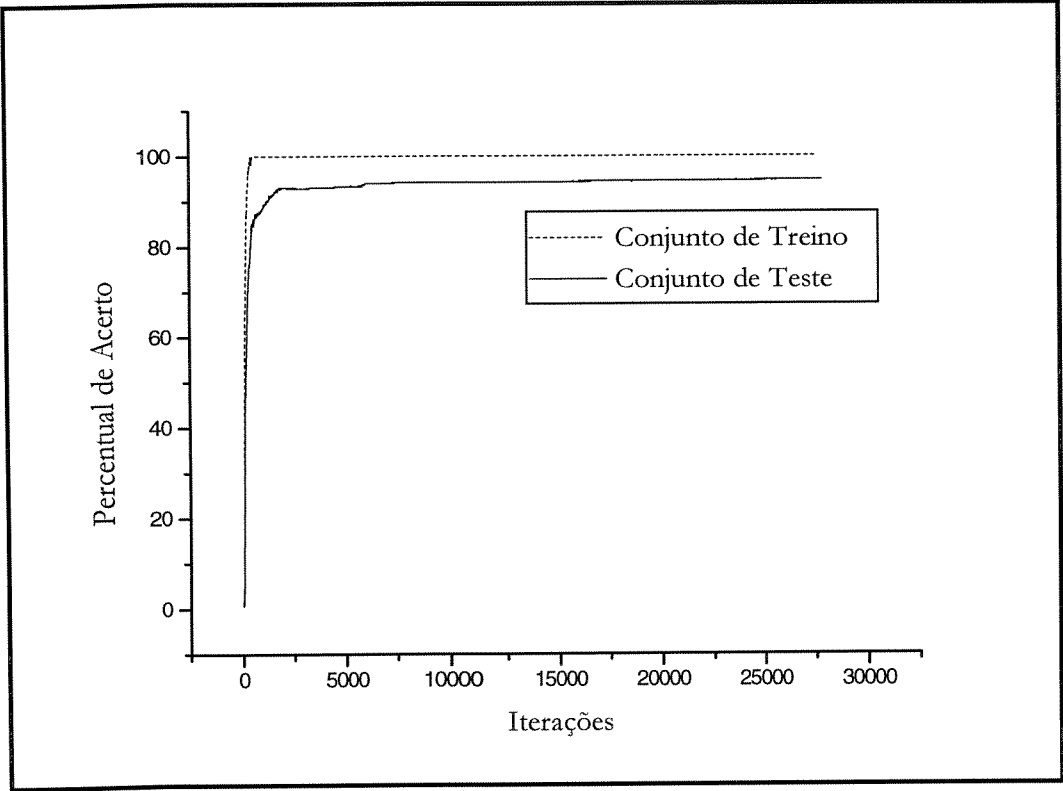


Figura 5-12: Curvas de percentual de acerto para treinamento e teste do MLP (65-80-102) em função do número de iterações do algoritmo de treinamento.

Foi efetuada também a análise discriminante das características das 435 imagens, utilizando-se o pacote de análise estatística SAS. A análise discriminante das características conseguiu separar corretamente 100 % das imagens nas suas 102 classes correspondentes usando as 98 características do modelo de 98 *eigenfaces*. O que comprova o grande grau de separabilidade linear entre as características obtidas com o modelo de *eigenfaces*.

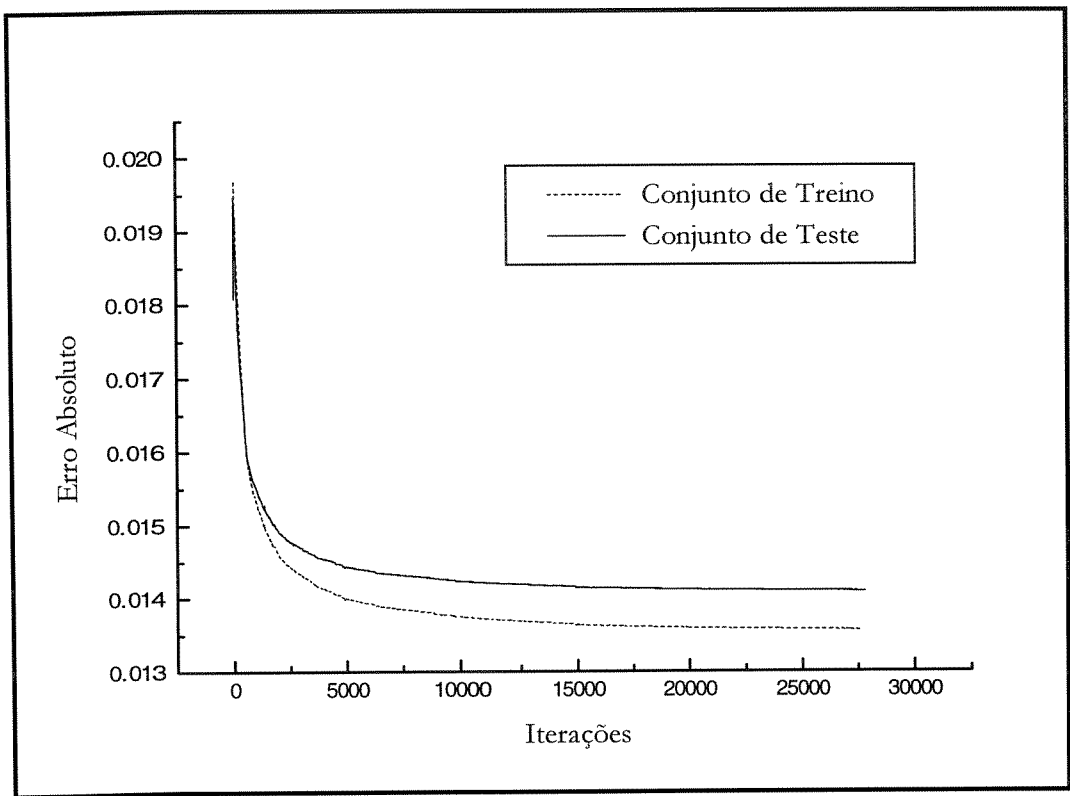


Figura 5-13: Curvas de erro absoluto da rede para treinamento e teste do MLP (65-80-102) em função do número de iterações do algoritmo de treinamento.

5.2 Verificação

O problema de verificação consiste em, dada uma imagem facial e declarada a classe a qual ela pertence, determinar se esta declaração é verdadeira. Em um sistema de verificação facial com aplicações tais como controle de acesso, procuramos uma otimização no desempenho para garantir a segurança do sistema. Assim, um sistema para controle de acesso deve rejeitar todos os indivíduos não autorizados. O sistema pode rejeitar indivíduos autorizados cujo padrão da imagem no instante da verificação varie muito em relação aos padrões registrados no sistema de verificação facial. Este fato faz com que o sistema de verificação tenha que readquirir a imagem do indivíduo até poder tomar uma decisão considerada confiável. Então, o objetivo da análise do sistema de

verificação é otimizar o ponto de operação do sistema, de modo que rejeite todos os indivíduos não autorizados, enquanto rejeitaria falsamente apenas um pequeno número de indivíduos autorizados mas com imagens muito diferentes dos padrões armazenados no sistema de verificação.

5.2.1 Análise da Verificação

Como forma de medirmos o desempenho do sistema de verificação facial, utilizamos as curvas de característica de operação do receptor (*receptor operation characteristic* - ROC [VALENTIN-b]) para um variado número de *eigenfaces*. As curvas de característica de operação do receptor fornecem o desempenho do sistema em relação às taxas de erro de falsa aceitação e falsa rejeição que o sistema provê para um determinado limiar de aceitação. O erro de falsa aceitação ocorre quando, dada uma classe a ser verificada, classificamos como pertencente a esta classe uma imagem que não é de um indivíduo desta classe. O erro de falsa rejeição ocorre quando classificamos como não pertencente a classe um indivíduo que realmente pertence a esta classe.

As curvas de operação do receptor foram obtidas variando-se o limiar de aceitação do classificador de distância mínima empregado no sistema. Os erros considerados foram os erros médios de verificação para as 102 classes de indivíduos presentes na base de dados. Assim, usamos todas as 435 imagens para teste tendo como centros de classe a serem comparados as médias aritméticas de todas as características de cada indivíduo presentes na base de dados.

A definição do erro médio de falsa aceitação (EMFA) empregado na análise é a seguinte:

$$EMFA = \frac{\sum_{i=1}^C EFA_i}{\sum_{i=1}^C EXF_i} \quad \text{Equação 5-2}$$

onde,

C é o número de classes presentes no processo;

EFA_i é o número de erros de falsa aceitação para a i -ésima classe;

EXF_i é o número de exemplos falsos apresentados à i -ésima classe.

A definição do erro médio de falsa rejeição (EMFR) é a seguinte:

$$EMFR = \frac{\sum_{i=1}^C EFR_i}{\sum_{i=1}^C EXV_i} \quad \text{Equação 5-3}$$

onde,

C é o número de classes presentes no processo;

EFR_i é o número de erros de falsa rejeição para a i -ésima classe;

EXV_i é o número de exemplos verdadeiros apresentados à i -ésima classe.

Na busca do ponto de operação ótimo do sistema de verificação observamos, na Figura 5-14, a variação das curvas de operação do receptor com o número de *eigenfaces*. Nos testes de verificação, é utilizado o modelo de 178 *eigenfaces* com pré-processamento pois este apresenta a melhor desempenho de reconhecimento. Como podemos observar, os pontos que mais se aproximam da taxa de erro médio de falsa aceitação nula e com um menor erro médio de falsa rejeição estão na curva de 178 *eigenfaces*. Este ponto de operação ótimo foi obtido para as 178 *eigenfaces* com um limiar de distância de 810 no classificador de distância mínima (Equação 3-15), provendo 100 % de aceitações corretas com a rejeição de 26.82% das amostras.

Nas curvas das figuras 5-15 e 5-16, observamos a variação dos erros de falsa aceitação e falsa rejeição, respectivamente, em função do limiar de decisão do

classificador de distância mínima (Equação 3-15). Os valores do limiar de decisão podem ser usados para estabelecermos as taxas desejadas dos dois tipos de erros no processo de verificação. Podem ainda ser utilizados no reconhecimento para atribuírmos uma determinada imagem como não estando presente, em nenhuma das classes da base de dados.

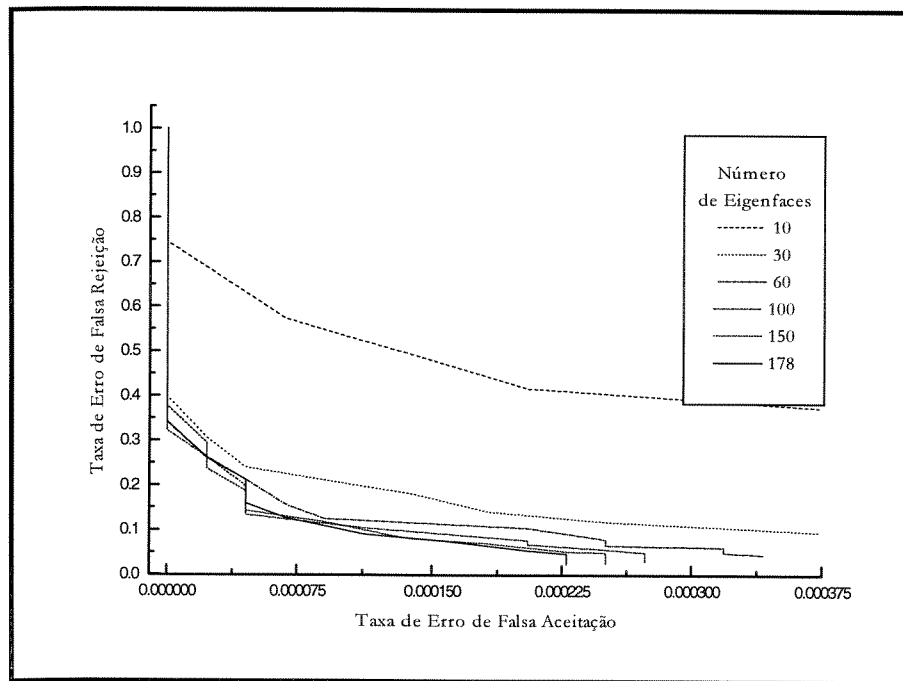


Figura 5-14: Curva de operação do receptor para em função do número de características utilizadas para o modelo de 178 *eigenfaces* com pré-processamento.

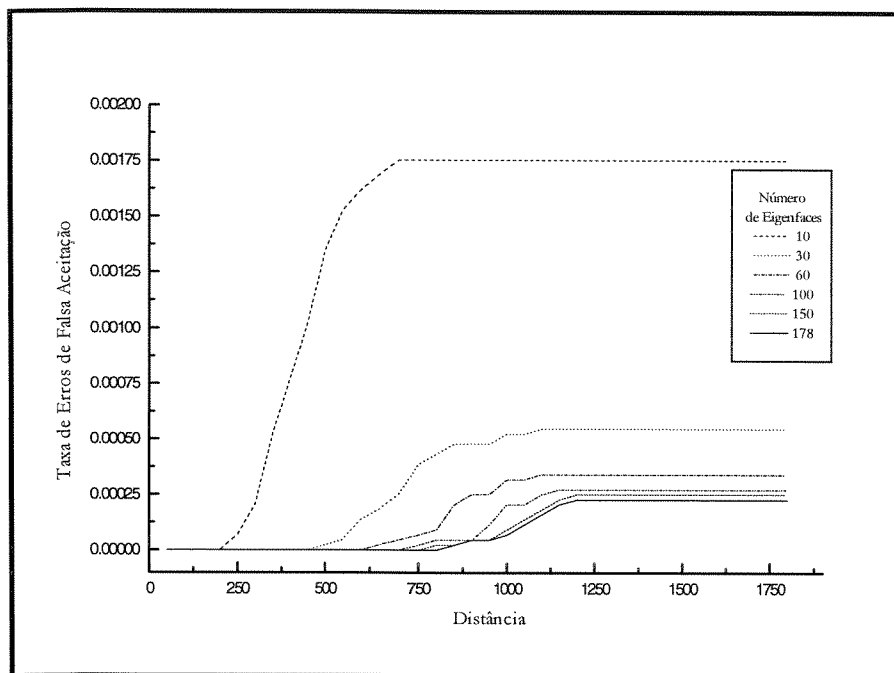


Figura 5-15: Curva da variação da taxa de erro de falsa aceitação em função do limiar de decisão.

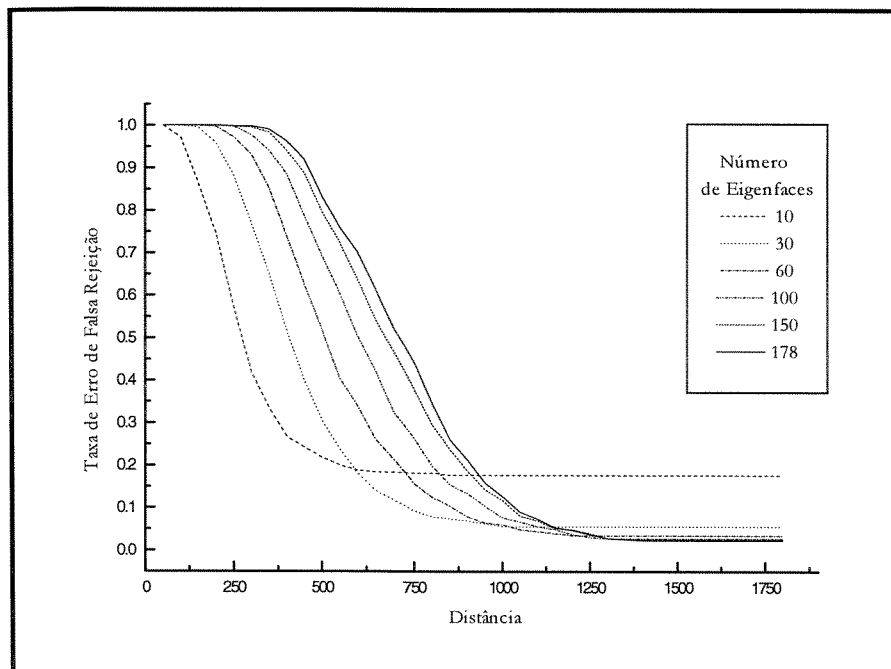


Figura 5-16: Curva da variação da taxa de erro de falsa rejeição em função do limiar de decisão.

5.3 Discussão

Neste capítulo, apresentamos os resultados do desempenho do sistema nas tarefas de reconhecimento e verificação facial, mostrando como os mesmos podem ser utilizados para termos um sistema de reconhecimento e verificação facial eficiente. No próximo capítulo, serão mostradas as comparações com os resultados apresentados por outros sistemas e a discussão sobre os resultados obtidos.

Capítulo 6

6. Discussões e Conclusões

Apresentamos neste capítulo as contribuições obtidas com o trabalho desenvolvido utilizando o método de *eigenfaces*, bem como as discussões e comparações com os outros resultados obtidos na literatura de processamento de imagens faciais. Na seqüência são apresentadas as conclusões finais e sugestões para aprimoramentos no sistema existente e utilização de outras técnicas de processamento de imagens faciais.

6.1 Contribuições

Este trabalho traduz a primeiro esforço na área de reconhecimento e verificação facial dentro do âmbito do Departamento de Comunicações da Faculdade de Engenharia Elétrica e Computação. Ele fornece resultados confiáveis que servem de base para a pesquisa em problemas relacionados ao reconhecimento e a verificação facial.

O sistema desenvolvido pode ser empregado como núcleo para um sistema totalmente automatizado de reconhecimento facial. Permite, sem muitas alterações, sua aplicação em indexação de uma base de dados por conteúdo presente na imagem, fazendo a procura de uma face a partir de uma face exemplo.

Na tarefa de verificação, este sistema pode ser empregado como núcleo de sistemas de controle de acesso, permitindo a verificação automática de indivíduos. O sistema de verificação provê confiabilidade suficiente para ser utilizado em aplicações que possuem fortes restrições de segurança. O método de *eigenfaces* implementado realiza uma boa compressão das características faciais que poderiam ser facilmente gravadas em um cartão magnético para a utilização em um sistema automático de verificação de cartões de crédito.

O trabalho desenvolvido contribui com novas técnicas de pré-processamento e normalização facial que ressaltam as diferenças entre as faces individuais quando comparado com as técnicas de normalização sem pré-processamento mostradas em [TURK], [PENTLAND-b] e [MOGGHADDAM-e].

6.2 Discussões

6.2.1 A Qualidade das Imagens da Base de Dados

As imagens presentes na base de dados foram adquiridas com o objetivo de termos uma grande variabilidade para testarmos a robustez do sistema. Devido inicialmente a problemas de armazenamento em disco, fomos obrigados a reduzir a resolução com que eram obtidas as imagens, de 480 x 640 pixels para 240 x 332 pixels. Isto provocou uma diminuição na qualidade das imagens adquiridas, pois reduz as definições de características importantes como os olhos. Acarreta então uma fonte de erros para o sistema que não dispõe da melhor qualidade de informação possível. Esta

resolução maior teria ajudado a compensar algumas falhas do processo de aquisição de imagens, principalmente naquelas imagens que foram adquiridas com uma pequena área facial. Dadas as limitações na qualidade de aquisição das imagens, o sistema de reconhecimento facial apresenta um bom desempenho, tratando de compensar as variações nas imagens faciais, apesar da baixa definição das características faciais presentes em algumas imagens.

6.2.2 O Processo de Normalização

Em nosso sistema de reconhecimento facial, para alinhar e posicionar as imagens faciais para comparação, realizamos normalizações de escala e translação. O processo de normalização não contou com normalizações em relação a rotação da face, que seria um parâmetro a ser avaliado na robustez do sistema. Nos resultados apresentados vemos que o sistema conseguiu tratar bem as variações de rotação em relação a face frontal sem a necessidade da normalização de rotação. As variações de rotação não eram tão acentuadas na maioria das imagens, deste modo, rotações no plano da imagem foram compensadas pelo sistema. Ocorreram apenas dois erros em relação a rotação para o modelo de 178 *eigenfaces* com pré-processamento.

A normalização com equalização de histograma, usada no sistema, não apresentou diferenças significativas em relação a imagens não normalizadas, mesmo assim foi mantida no sistema. Isto indica que o método de projeção nas *eigenfaces* do modelo apresentado possui certo grau de robustez em relação à iluminação, pois sendo um método que efetua correlação, a informação das bordas da imagem facial são as mais significativas para o desempenho do sistema.

Há apenas a necessidade de automatizar os procedimentos de localização da face e dos olhos na imagem de entrada. Isto vai permitir a normalização automática das imagens faciais que são analisadas pelo sistema de reconhecimento.

6.2.3 Outros Resultados com o Método de *Eigenfaces*

Devido ao fato de não estarmos trabalhando com a mesma base de dados com que os outros pesquisadores testaram suas implementações do método, não podemos comparar diretamente os resultados obtidos. O desempenho de nosso sistema no entanto se encontra dentro da mesma faixa de desempenho, entre 87% e 99% de acertos, apresentada para as outras implementações do método de *eigenfaces* [TURK] [PENTLAND-b] [MOGHADDAM-e]. Nosso desempenho também está dentro da desempenho das faixas de acerto, entre 90% e 99%, dos outros métodos de reconhecimento apresentados na literatura de processamento de imagens faciais, tais como, [GOUDAIL], [SAMARIA-a], [WISKOTT-a] e [ZHANG]. Temos ainda a restrição de estar trabalhando com imagens em resolução e qualidade inferior às dos sistemas descritos. Em nosso processo de aquisição de imagens, não foram impostas fortes restrições no procedimento de aquisição das imagens de teste, o que não é verdade para as outras abordagens do método de *eigenfaces* e de outros sistemas de reconhecimento.

A comparação dos nossos resultados de reconhecimento pode ser feita aproximadamente com um experimento semelhante descritos em [MOGHADDAM-a] que classifica 150 imagens faciais frontais de 150 indivíduos contra outras 150 imagens frontais dos mesmos indivíduos. Ele obtêm apenas um erro de classificação com um taxa de acerto 99.4%. Este experimento se assemelha ao que foi realizado para as imagens modelo, classificar 193 imagens de 102 indivíduos em 102 classes, que possui dois erros de classificação obtendo um desempenho de 98.96%. A pequena diferença entre os desempenhos pode ser atribuída à diferença na qualidade das imagens presentes em ambas as bases de dados.

6.3 Conclusões

O desempenho do sistema de reconhecimento foi de 97.70% de acertos para todas as imagens da base de dados e de 98.96% de acertos para as imagens da base de dados que se assemelham ao modelo usado para compor as características. O sistema de verificação apresentou um melhor desempenho empregando 178 características. Seu desempenho para aplicações como controle de acesso, quando testado com todas as imagens da base de dados, foi de 100% de acertos de verificação, rejeitando 26.82% das 435 imagens apresentadas ao sistema para verificação.

Os sistemas descritos neste trabalho apresentam um desempenho satisfatório para as tarefas de reconhecimento e verificação facial sob condições de aquisição de imagem sem um controle rígido. Eles oferecem uma certa robustez às condições de aquisição da imagem de entrada. Em caso de ambientes controlados, o desempenho dos sistemas seria ainda melhor com o aumento da qualidade das imagens adquiridas.

6.4 Sugestões para Futuros Trabalhos

Com o objetivo de obter uma maior robustez em relação a variações na posição da imagem de entrada em nosso sistema, pode-se implementar as adições ao sistema de *eigenfaces* descritas em [PENTLAND-c]. Nela treinamos diferentes modelos para cada posição que a face pode tomar na imagem de entrada.

Na composição da base de dados, podemos implementar um sistema com a aquisição de seqüências de vídeo, nas quais o indivíduo faria uma série de movimentos e expressões. Esta seqüência seria posteriormente amostrada e catalogada para prover uma grande gama de exemplos para a construção e testes de sistemas de reconhecimento. Estes exemplos são o fator limitante do desempenho dos sistemas de reconhecimento,

pois mesmo aqueles que possuem uma robustez a variações necessitam de bons exemplos durante sua construção ou treinamento.

Na mudança de arquitetura básica dos sistema, podemos observar e testar os métodos descritos em [GOUDAIL], [LAWRENCE] , [WISKOTT-a], [WISKOTT-c] e [ZHANG]. Estes métodos são mais recentes na literatura e apresentam novas características que trazem mais robustez ao processo de reconhecimento.

7. Referências

[ABDI-a] H. Abdi, **A generalized Approach for Connectionist Autoassociative Memories: Interpretation, Implications and Ilustration for Face Processing**, J. Demongeot (Ed.), *Artificail Intelligence and Cognitive Sciences, Manchester University Press*, Manchester, 1988.

[ABDI-b] H. Abdi, D. Valentin, A. J. O'Toole, **A generalized Autoassociative Model for Face Processing and Sex Categorization: From Principal Component Analysis to Multivariate Analysis**, *Optimality in Biological and Artificial Networks*, D. Levine (Ed.) em impressão, 1996.

[ABDI-c] H. Abdi, D. Valentin, T.B. Edelman, A. J. O'Toole, **More About the difference Between Men and Women : From Linear Neural Networks and The Principal Component Approach**, *Perception*, Vol. 24, pp. 539-562, 1996.

[ARAD] N. Arad, N. Dyn, D. Reisfield, **Image Warping by Radial Basis Functions: Aplications to Facial Expressions**, *CVGIP: Graphical Models and Image Processing*, Vol. 56, No 2, pp. 161-172, 1994.

[BEYMER] D. Beymer, T. Poggio, **Face Recognition from One Example View**, *Proceedings of the International Conference on Computer Vision* , Cambridge, MA, agosto 1995.

[BRUNELLI] R. Brunelli e T. Poggio, **Face Recognition: Features versus Templates**, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, Nº 10, pp 1042-1052, outubro, 1993.

[BUREL] G. Burel, D. Carel, **Detection and Localization of Faces on Digital Images**, *Pattern Recognition Letters*, Nº 15, pp 963-967, outubro, 1994.

[CYCHOSZ] L. W. Cychosz e J. Zhang, **A Face Recognition System Using Holistic Regional Feature Extraction**, *Proceedings of 13^o IEEE ICASSP*, Seattle, pp 3431-3434, 1995.

[CHELLAPPA] R. Chellappa, C. L. Wilson, S. Sirohey, **Human and Machine Recognition of Faces: A Survey**, *Proceedings of the IEEE*, Vol. 83, Nº 5, maio 1995.

[CHEN] Q. Chen, H. Wu, M. Yachida, **Face Detection by Fuzzy Pattern Matching**, *IEEE*, agosto, 1995.

[DAI-a] Y. Daí, Y. Nakano, **Face-Texture Model Based on SGLD and Its Application in Face Detection in a Color Scene**, *Pattern Recognition*, vol 29, No 6, pp 1007-1017, 1996.

[DAI-b] Y. Daí, Y. Nakano, **Extraction of Facial Images from a Complex Background Using SGLD Matrices**, *Proceedings of 12^o IEEE ICPAR*, Israel, pp 137-141, 1994.

[ELLIS] M. A. Ellis, B. Stroustrup, **C++ Manual de Referência Comentado**, Ed. Campus, 1993.

[ESSA] I. A. Essa e A. Pentland, **Facial Expression Recognition Using Dynamic Model and Model Energy**, *Proceedings of the International Conference on Computer Vision*, Cambridge, MA, maio, 1995.

[GOUDAIL] F. Goudail, E. Lange, T. Iwamoto, K. Kyume, N. Otsu, **Face Recognition System Using Local Autocorrelations and Multiscale Integration**, *IEEE Pattern Analysis and Machine Intelligence*, vol 18, No 10, pp 1024-1028, outubro 1996.

[HASEGAWA] O. Hasegawa, K. Yokosawa, M. Ishizuka, **Real-Time Parallel and Cooperative Recognition of Facial Images for an Interactive Visual Human Interface**, *Proceedings of 12^o IEEE ICPAR*, Israel, pp 384-387, 1994. .

[HAYKIN] S. Haykin, **Neural Networks - A Comprehensive Foundation**, *IEEE PRESS*, Macmillan, 1994.

[HUNKE] H. M. Hunke, **Locating and Tracking of Human Faces with Neural Networks**, *Carnegie Mellon University, School of Computer Science, Msc. Thesis, Technical Report CMU-CS-94-155*, 1994.

[ITRATOR] N. Itrator, D. Reisfield e Y. Yeshurun, **Face Recognition using a Hybrid Supervised/Unsupervised Neural Network**, *Proceedings of 12^o IEEE ICPAR*, Israel, pp 50-54, 1994.

[JAIN] A. K. Jain, Y. Zhong, S. Lakshmanan, **Object Matching Using Deformable Templates**, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, No. 3, março, 1996.

[JOHNSON] R. A. Johnson, D. W. Wichern, **Applied Multivariate Statistical Analysis**, *Prentice Hall*, 1992.

[KIRBY] M. Kirby, L. Sirovich, **Application of the Karhunen-Loève Procedure for the Characterization of Human Faces**, *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol 12, pp 103-108, janeiro, 1990.

[KWON] Y. H. Kwon e N. V. Lobo, **Face Detection Using Templates**, *Proceedings of 12^o IEEE ICPAR*, Israel, pp 764-767, 1994.

[KONEN] W. Konen, E. Schulze-Krüger, **ZN-Face: A System for Access Control Using Automated Face Recognition**, *International Workshop on Automatic Face and Gesture Recognition*, University of Zürich, 1995.

[LAWRENCE] S. Lawrence, C. L. Giles, A. C. Tsoi, A. D. Back, **Face Recognition: A Convolution Neural Network Approach**, aceito para a publicação em *IEEE Transactions on Neural Networks, Special Issue on Neural Networks and Pattern Recognition*.

[LEVINE] J. Levine, **Programming for Graphics Files in C and C++**, John Wiley & Sons, 1994.

[LINDLEY] C. A. Lindley, **Practical Image Processing in C**, John Wiley & Sons, 1992.

[MATSUNO] K. Matsuno e S. Tsuji, **Recognizing Facial Expressions in a Potential Field**, *Proceedings of 12th IEEE ICPAR*, Israel, pp 44-49, 1994.

[MILLER] B. Miller, **Vital Signs of Identity**, *IEEE Spectrum*, pp. 22-30, fevereiro, 1994.

[MOGHADDAM-a] B. Moghaddam, A. Pentland, **Probabilistic Visual Learning for Object Detection**, *International Conference on Computer Vision*, Cambridge, MA, junho, 1995.

[MOGHADDAM-b] B. Moghaddam, A. Pentland, **Face Recognition using View-Based and Modular Eigenspaces**, *M.I.T. Media Laboratory Perceptual Computing Section Technical Report N^o 301*.

[MOGHADDAM-c] B. Moghaddam, C. Nastar, A. Pentland, **Bayesian Face Recognition Using Deformable Intensity Surfaces**, *IEEE Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, junho, 1996.

[MOGHADDAM-d] B. Moghaddam, A. Pentland, **A Subspace Method for Maximum Likelihood Target Detection**, *IEEE International Conference on Image Processing*, Washington, DC, outubro, 1996.

[MOGHADDAM-e] B. Moghaddam, A. Pentland, **An Automatic System for Model-Based Coding of Faces**, *IEEE Data Compression Conference*, Snowbird, Utah, março, 1995.

[NASTAR] C. Nastar e A. Pentland, **Matching and Recognition Using Deformable Intensity Surfaces**, *IEEE International Symposium on Computer Vision*, Coral Gables, FL, maio, 1995.

[O'TOOLE] A. O'Toole, H. Abdi, K. H. Deffenbacher, D. Valentin, **Low Dimensional Representation of Faces in the Higher Dimensions of the Face Space**, *Journal of Optical Society of America*, Vol. 10, No 3, pp. 405-411, março, 1993.

[**PENTLAND-a**] A. Pentland, T. Starner, T. N. Etcoff, A. Masoiu, O. Oliyide, M. Turk, **Experiments with Eigenfaces**, *IJCAI*, Chambery, France, agosto, 1993.

[**PENTLAND-b**] A. Pentland, R.W. Picard, S. Sclaroff, **PHOTOBOOK: Content Based Manipulation of Image Databases**, *SPIE - Storage and Retrieval Image and Video Databases*, San Jose ,CA, 1994.

[**PENTLAND-c**] A. Pentland, M. Baback e T. Starner, **View-Based and Modular Eigenspaces for Face Recognition**, *IEEE Conference on Vison and Pattern Recognition*, Seattle, WA, julho, 1994.

[**POGGIO**] T. Poggio , D. Beymer, **Learning to See**, *IEEE Spectrum*, pp 60-69, maio, 1996.

[**PRESS**] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, **Numerical Recipes in C**, *Cambridge University Press*, 1991.

[**SAMAL**] A. Samal e P. A. Iyengar, **Automatic Recognition and Analysis of Human Faces and Facial Expressions: A Survey**, *Pattern Recognition*, Vol. 25, N^o 1, pp 65-77, 1992.

[**SAMARIA-a**] F.S. Samaria, **Face Recognition Using Hidden Markov Models**, *PhD Thesis, Trinity College, University of Cambridge*, Cambridge, 1994.

[**SAMARIA-b**] F.S. Samaria, **Parameterisation of a Stochastic Model for Human Face Identification**, *Proceedings of 2nd IEEE Workshop on Applications of Computer Vision*, Sarasota, Florida, 1994.

[**SAKO**] H. Sako, M. Whitehouse, A. Smith e A. Sutherland, **Real-Time Facial Feature Tracking Based on Matching Techniques and Its Applications**, *Proceedings of 12^o IEEE ICPAR*, Israel, pp 320-324, 1994.

[**SCHALLKOFF**] R. Schalkoff, **Pattern Recognition, Statistical, Structural and Neural Approaches**, John Willey & Sons, 1992.

[SIROVICH] L. Sirovich, M. Kirby, **Low-Dimensional Procedure for the Characterization of Human Faces**, *Journal of Optical Society of America*, Vol. 4, No. 3, pp 519-524, março, 1987.

[STROUSTRUP] B. Stroustrup, **The C++ Programming Language**, Addison Wesley, 1991.

[SUNG] K. Sung, T. Poggio, **Example-based Learning for View-based Human Face Detection**, *MIT - AI LAB, AI-Memo No. 1521*, dezembro, 1994.

[SWETS] D.L. Swets, J.Weng, **Using Discriminant Eigenfeatures for Image Retrieval**, *Technical Report Computer Science Department, Michingan State University*, 1996.

[TAKÁCS] B. Takács e H. Wechsler, **Locating Facial Features Using SOFM**, *Proceedings of 12th IEEE ICPAR*, Israel, pp 137-141, 1994.

[TURK] M. Turk, A. Pentland, **Eigenfaces for Recognition**, *Journal of Cognitive Neuroscience*, Vol. 3, No 1, pp. 71-86, 1991.

[VAILLANT] R.Vaillant, C. Monrocq e Y. Le Cun, **Original Approach for the localisation of objects in Images**, *IEE Proceedings on Vision, Image and Signal Processing*, ol 141, Nº 4, agosto, 1994.

[VALENTIN-a] D. Valentim, Hervé Abdi, A. J. O'Toole e G. W. Cottrell, **Connectionist Models of Face Processing: A Survey**, *Pattern Recognition*, Vol. 27, Nº 9, pp 1209-1230, 1994.

[VALENTIN-b] D. Valentin, H. Abdi, **Can a Linear Autoassociator Recognize Faces from New Orientations**, *Journal of the Optical Society of America*, Vol. 13, agosto, 1996.

[VALENTIN-c] D. Valentin, H. Abdi, A. J. O'Toole, **Categorization and Identification of Human Face Images by Neural Networks: A review of the Linear Autoassociative and Principal Component Approaches**, *Journal of Biological Systems*, Vol. 2, No. 3, 1994.

[XIE] X. Xie, R. Sudhakar e H. Zhuang, **On Improving Eye Feature Extraction Using Deformable Templates**, *Pattern Recognition*, Vol. 27, Nº 6, pp 791-799, 1994.

[WISKOTT-a] L. Wiskott, C. Von Der Malsburg, **Face Recognition by Dynamik Link Matching**, *Proceedings of the International Conference on Artificial Neural Networks - ICANN*, pp-347-352, Paris , 1995.

[WISKOTT-b] L. Wiskott, J. Felous, N. Krüger, C. Von Der Malsburg, **Face Recognition and Gender Determination**, *Proceedings of the International Workshop on Automatic Face and Gesture Recogniton*, Zürich , 1995.

[WISKOTT-c] L. Wiskott, **Labeled Graphs and Dynamic Link Matching for Face Recognition and Scene Analysis**, *Verlag Harri Deutsch, Thun - Frabkefurt am Main, Reihe Physik 53 (PhD thesis)*, 1995.

[WÜRTZ] R. P. Würtz, **Building Visaul Correspondence Maps - From Neuronal Dynamics to a face Recogniton System**, in *Roberto Moreno-Díaz and José Mira-Mira (eds.), Brain Process, Theories and Models*, pp. 420-429, MIT PRESS, novembro, 1995.

[YACOOB-a] Y. Yacoob e L. Davis, **Recognizing Facial Expressions by Spatio-Temporal Analysis**, *Proceedings of 12th IEEE ICPAR*, Israel, pp 747-749, 1994.

[YACOOB-b] Y. Yacoob e L. Davis, **Recognizing Human Facial Expressions from Long Image Sequences Using Optical Flow**, *IEEE Transactions on Pattern Analyis and Machine Intellige*, Vol. 18, No. 6, pp. 630-642, 1996.

[ZHANG] M. Zhang, J. Fulcher, **Face Recognition Using Artificial Network Group-Based Adaptive Tolerance (GAT) Tress**, *IEEE Transactions on Neural Networks*, Vol. 7, No. 3, pp. 555-567, maio, 1996.