



**SENAI CIMATEC**

**PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM  
COMPUTACIONAL E TECNOLOGIA INDUSTRIAL  
Mestrado em Modelagem Computacional e Tecnologia Industrial**

**Dissertação de mestrado**

**Reconhecimento de faces aplicado ao problema de  
pessoas desaparecidas - Estudo de caso do Eigenface**

Apresentada por: Marcos Batista Figueredo  
Orientador: Prof. Dr. Josemar Rodrigues de Souza

Março de 2011

Marcos Batista Figueredo

# Reconhecimento de faces aplicado ao problema de pessoas desaparecidas - Estudo de caso do Eigenface

Dissertação de mestrado apresentada ao Programa de Pós-graduação em Modelagem Computacional e Tecnologia Industrial, Curso de Mestrado em Modelagem Computacional e Tecnologia Industrial do SENAI CIMATEC, como requisito parcial para a obtenção do título de **Mestre em Modelagem Computacional e Tecnologia Industrial**.

Área de conhecimento: Interdisciplinar

Orientador: Prof. Dr. Josemar Rodrigues de Souza  
*SENAI CIMATEC*

Salvador  
SENAI CIMATEC  
2011

---

## Nota sobre o estilo do PPGMCTI

---

Esta dissertação de mestrado foi elaborada considerando as normas de estilo (i.e. estéticas e estruturais) propostas aprovadas pelo colegiado do Programa de Pós-graduação em Modelagem Computacional e Tecnologia Industrial e estão disponíveis em formato eletrônico (*download* na Página Web [http://ead.fieb.org.br/portal\\_faculdades/dissertacoes-e-teses-mcti.html](http://ead.fieb.org.br/portal_faculdades/dissertacoes-e-teses-mcti.html) ou solicitação via e-mail à secretaria do programa) e em formato impresso somente para consulta.

Ressalta-se que o formato proposto considera diversos itens das normas da Associação Brasileira de Normas Técnicas (ABNT), entretanto opta-se, em alguns aspectos, seguir um estilo próprio elaborado e amadurecido pelos professores do programa de pós-graduação supracitado.

# SENAI CIMATEC

Programa de Pós-graduação em Modelagem Computacional e Tecnologia Industrial

Mestrado em Modelagem Computacional e Tecnologia Industrial

A Banca Examinadora, constituída pelos professores abaixo listados, leram e recomendam a aprovação [com distinção] da Dissertação de mestrado, intitulada “Reconhecimento de faces aplicado ao problema de pessoas desaparecidas - Estudo de caso do Eigenface”, apresentada no dia (dia) de (mês) de (ano), como requisito parcial para a obtenção do título de **Mestre em Modelagem Computacional e Tecnologia Industrial**.

Orientador:

---

Prof. Dr. Prof. Dr. Josemar Rodrigues de Souza  
SENAI CIMATEC

Membro interno da Banca:

---

Prof. Dr. Gilney Figueira Zebende  
SENAI CIMATEC

Membro externo da Banca:

---

Prof. Dr. Rogério Junior Correia  
UNEB

Dedico este trabalho a todas as pessoas que contribuiriam direta ou indiretamente a ele, em especial: a minha amada esposa Soraia e amado filho Guilherme, que compreenderam e me deram sempre energia para construir este caminho; A minha família que sempre esteve e sempre estará ao meu lado, com especial carinho a Jussara e Jandiara, minhas queridas irmãs que sem o seu apoio não chegaria aqui Egberto, fonte de grande inspiração, Betinho, Flávia e Juliana meus queridos sobrinhos, a minha sogra Maria Rita e cunhada Suzana que pelas correções e ajuda nos momentos difíceis, a toda família Argoinense que estão presentes e sempre farão parte da minha vida. Não poderia de deixar de agradecer em especial aos meus avós Jandira e Arlindo, onde quer que estejam saibam que continuo amando vocês e com muitas saudades.

---

## Agradecimentos

---

Ao Prof. Dr. Josemar, mestre e amigo, pelos ensinamentos e orientações sempre objetivas e principalmente pela confiança e incentivo para realização desta pesquisa.

Aos Prof. Dr. Hernane Pereira, Dr. Marcelo Moret, Dr. Gilney, Dra. Lynn Alves por fundamentar os meus primeiros passos na área de pesquisa com muita paciência e dedicação, que possibilitou a busca por novos horizontes, meu sincero reconhecimento.

Aos amigos e do curso de Modelagem computacional pelo aprendizado diário, compartilhamento de idéias e bons momentos. A todos os colaboradores do sistema FIEB que ajudaram no meu percurso, em especial na figura da minha amiga Inês que sempre uma grande companheira.

Aos professores do Mestrado que estiveram sempre dispostos a colaborar, as secretárias pela atenção e ajuda burocrática.

Ao sistema FIEB que proporcionou o auxílio financeiro para a realização deste curso.

A minha família que sempre me apoiou de forma sábia, humilde e amorosa ensinando-me a vencer em cada momento de dificuldade, E neste momento especial, minha sincera gratidão pelo apoio incansável.

Salvador, Brasil  
11 de Março de 2011

Marcos Batista Figueredo

---

## Resumo

---

No Brasil, segundo dados do livro “Cadê Você”, cerca de 200 mil pessoas desaparecem todos os anos. Este problema é difícil de ser solucionado, pois demandam de órgãos governamentais investimentos em pessoal e recursos tecnológicos. Concomitantemente, sistemas biométricos melhoram seu desempenho a cada dia e alguns já estão em uso por diversas instituições e órgãos governamentais, estes sistemas permitem o reconhecimento/identificação de pessoas e alguns com margem de confiança acima de 95%. Neste cenário, propomos avaliar a eficiência de um método biométrico de cunho estatístico de reconhecimento de face aplicado a uma situação não controlada de detecção e reconhecimento de pessoas desaparecidas a partir de uma ou de um conjunto pequeno de imagens de entrada. Para realizar esta avaliação escolhe-se o modelo mais aplicado e consagrado da literatura atual, o Eigenface. Procurou-se estabelecer um método de pesquisa baseado em aspectos estatísticos, e como a maioria dos trabalhos nesta área se baseiam em condições ideais, procuramos reverter este panorama tentando encurtar o hiato entre os experimentos em condições ideais e em situações carregadas de ruído da realidade. Buscou-se minimizar os problemas de iluminação e pose com filtros especiais que tentam atenuar essas dificuldades e programamos o modelo em C# com auxílio da biblioteca OpenCV. Utilizou-se um banco de dados consagrado na literatura para testar o desempenho do programa e em seguida foi construído um banco de imagens próprio, constituído de 134 imagens, com pose frontal composto de homens e mulheres com tom de pele variado sem controle de iluminação. Como ponto relevante, mas não fundamental, analisamos também o comportamento do modelo em relação à detecção de idade. Utilizaram-se, como entrada de dados, imagens oriundas de fotografias digitais e vídeo câmera com hardware simples que constitui a maioria das entradas de dados em bancos de pessoas desaparecidas. O modelo ofereceu maior desempenho com imagens estáticas (entre 65% de reconhecimento positivo) do que com sequencias de vídeo (8% de reconhecimento positivo). Os testes revelaram que o modelo não funciona ou funciona de maneira precária (14% de reconhecimento positivo) com apenas uma imagem de treinamento por pessoa, sendo que o mínimo necessário de imagem de entrada estabelecido foi de mais de 15 outro fato que surgiu de maneira intrigante durante a investigação está relacionado a taxa de reconhecimento positivo em pessoas com tom de pele negra (4% de taxa positiva) e a taxa de reconhecimento da idade do indivíduo que se revelou muito destoante das pesquisas nesse campo (28% de taxa positiva). Estes dados permitiram concluir que o modelo falha no seu objetivo e necessita evoluir para ser considerado efetivo no reconhecimento preciso.

**Palavras-chave:** Eigenface, Pessoas Desaparecidas, PCA.

---

## Abstract

---

About two hundred people disappear every year in Brazil, according data from the book “Cadê você”. It is a difficult problem to solve as more investment in staff and technological resources from government is needed. At the same time, biological systems are getting better each day, which have already been used by the government and a range of institutions, allowing them to recognize or identify people within a margin of 95%. In this scenario, we propose to evaluate the efficiency of a biometric method of statistical nature of face recognition applied to a non controlled situation of detection and identification of missing people, from an image or a small set of entry images. In order to perform this evaluation, the Eigenface, the most used and consecrated current literature was used. A research method based on statistical aspects was established, and as the majority of those projects in this area is based in ideal conditions, we tried to reverse this panorama shortening the distance between those experiments in ideal conditions and the real ones full of noise. Problems of light and pose were decreased through special filters, which seemed to lower those difficulties and with the help of the Open CV library, the model C# was programmed. A database consecrated in literature was used to test the performance of the program and a data of images were built, consisting of 134 images of male and female frontal poses with different skin colors, however with no light control. As a relevant issue, but not fundamental, the behavior of the model in identifying age was analyzed. Images that came from digital photographs and video cameras with simple hardware were used as a data entry, which are the majority cases of missing people. The model offered the best performance with static images (65% of positive recognition), than with videos sequences (8% of positive recognition). Those tests showed that such models do not work or work very little (14% of positive recognition) with only one image of training per person, being the minimum necessary of entry image 15. Another intriguing fact that was raised during the investigation is related to the range of positive recognition in dark skin people (4% of positive recognition), and the range to recognize the age of a person, which researches were uncongenial in this field (28%positive). From this data, we come to the conclusion that such model fails in its objective and needs to be developed to be considered effective in its precise recognition.



---

# Sumário

---

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Definição do problema . . . . .	3
1.2	Objetivo . . . . .	4
1.2.1	Objetivo Geral . . . . .	5
1.2.2	Objetivos Específicos . . . . .	5
1.3	Importância da pesquisa . . . . .	5
1.4	Motivação . . . . .	6
1.5	Limites e limitações . . . . .	6
1.6	Aspectos metodológicos . . . . .	8
1.7	Organização da dissertação . . . . .	9
<b>2</b>	<b>Trabalhos Correlatos</b>	<b>11</b>
2.1	Introdução . . . . .	11
2.2	Estudo da Biometria . . . . .	11
2.2.1	Aplicações da Biometria . . . . .	18
2.3	Reconhecimento Facial . . . . .	19
2.3.1	Introdução . . . . .	19
2.3.2	Problemas e Limitações . . . . .	21
2.4	Técnicas de Reconhecimento Facial . . . . .	22
2.4.1	Técnicas baseadas na Transformada de Gabor . . . . .	23
2.4.2	Imagens de Perfil . . . . .	26
2.4.3	Técnicas baseadas em Imagens 3D . . . . .	27
2.4.3.1	Reconhecimento usando faixas de Imagens . . . . .	29
2.4.3.2	Reconhecimento baseado em características de profundi- dade e curvatura . . . . .	30
2.4.3.3	Reconhecimento facial usando Eigenface 3D . . . . .	31
2.4.4	Métodos Holísticos . . . . .	32
2.4.4.1	Análise discriminante Linear . . . . .	32
2.4.4.2	Análise Componentes Principais . . . . .	34
2.5	Considerações finais . . . . .	36
<b>3</b>	<b>Modelo Proposto</b>	<b>37</b>
3.1	Introdução . . . . .	37
3.2	Definição de Imagem Digital . . . . .	37
3.3	Definições aplicadas ao PCA . . . . .	39
3.3.1	Definições de Estatística Multivariada . . . . .	39
3.3.2	Definições de Álgebra Linear . . . . .	41
3.4	Eigenface . . . . .	42
3.4.1	Treinamento . . . . .	45
3.4.2	Reconhecimento . . . . .	46
3.5	Considerações Finais . . . . .	47

<b>4</b>	<b>Implementação</b>	<b>49</b>
4.1	Introdução . . . . .	49
4.2	Base de Dados . . . . .	49
4.3	Detecção de Face . . . . .	51
4.3.1	Características de Haar . . . . .	52
4.3.2	Imagem Integral . . . . .	53
4.3.3	Algoritmo de Aprendizagem . . . . .	55
4.3.3.1	Classificador Fraco . . . . .	55
4.3.3.2	AdaBoost . . . . .	56
4.3.4	Um classificador em cascata . . . . .	56
4.4	Pré-Processamento . . . . .	58
4.4.1	Introdução . . . . .	58
4.4.2	Correção de intensidade de luz . . . . .	58
4.4.3	Transformação Logarítmica . . . . .	59
4.4.4	Equalização de Histogramas . . . . .	60
4.4.5	Filtro Linear Espacial . . . . .	61
4.5	Testes e Resultados . . . . .	62
4.5.1	Experimento 1: Teste de tempo de processamento do treinamento . . . . .	65
4.5.2	Experimento 2: Teste de reconhecimento via imagem estática . . . . .	66
4.5.3	Experimento 3: Teste de reconhecimento via imagem dinâmica . . . . .	73
4.5.4	Análise do Reconhecimento levando em conta o tom de pele . . . . .	74
4.5.5	Estimação da idade . . . . .	74
4.5.6	Discussão . . . . .	75
4.5.6.1	Experimento 1: Teste de tempo de processamento do Trei- namento . . . . .	75
4.5.6.2	Experimento 2: Teste de reconhecimento via imagem estática . . . . .	76
4.5.6.3	Teste de reconhecimento via imagem dinâmica . . . . .	76
4.5.6.4	Reconhecimento - tom de pele . . . . .	78
4.5.6.5	Estimar da Idade . . . . .	78
4.6	Considerações Finais . . . . .	79
<b>5</b>	<b>Conclusão e Trabalhos Futuros</b>	<b>80</b>
5.1	Conclusões . . . . .	80
5.2	Contribuições e Trabalhos Futuros . . . . .	82
5.2.1	Trabalhos futuros . . . . .	83
	<b>Referências</b>	<b>84</b>

---

## Lista de Tabelas

---

1.1	Aplicações típicas para o reconhecimento de faces (CHELLAPPA; WILSON; SIROHEY, 1995) . . . . .	2
2.1	Tabela com as características gerais de um sistema biométrico,(JAIN; ROSS; PRABHAKAR, 2004) . . . . .	12
2.2	Tabela comparativa de métodos biométricos. A - Alta performance; M - performance Média; B - Baixa performance. (JAIN; ROSS; PRABHAKAR, 2004) . . . . .	18
2.3	Tabela dos mais conhecidos bancos de dados de faces 3D - (SAVRAN et al., 2008) . . . . .	28
2.4	Resultado comparativa entre Eigenfaces e HMM . . . . .	30
4.1	Tabela de resultados do reconhecimento utilizando a base ORL e a nossa implementação . . . . .	64
4.2	Tabela com resultados dos 10 primeiros dias . . . . .	69
4.3	Tabela com resultados do reconhecimento com imagens coletadas entre os dias 11-20 . . . . .	70
4.4	Dados dos últimos 10 dias de coleta de dados . . . . .	70
4.5	Dados condensados com resultados generalizado . . . . .	70
4.6	Reconhecimento utilizando uma única imagem. . . . .	73
4.7	Reconhecimento utilizando algum tipo de disfarce . . . . .	73
4.8	Dados de reconhecimento por tom de pele . . . . .	74
4.9	Dados do reconhecimento de idade utilizando imagens estáticas . . . . .	75
5.1	Comparativo entre a tabela de (JAIN; ROSS; PRABHAKAR, 2004) e nossa análise . . . . .	81

---

## Lista de Figuras

---

1.1	Modelo básico do Eigenfaces . . . . .	3
1.2	Fluxo básico para reconhecimento de faces . . . . .	4
1.3	Diferentes condições de iluminação . . . . .	7
1.4	Ilustração do funcionamento do sistema de reconhecimento de Face . . . . .	9
2.1	Figura mostra como o processo pode ser invasivo, além do indivíduo poder recusar-se a fazer o teste . . . . .	13
2.2	Identificação biométrica já utilizada em terminais de auto atendimento bancário. . . . .	14
2.3	Imagem mostra a utilização pela justiça eleitoral brasileira de um método biométrico . . . . .	15
2.4	Mostra um leitor biométrico da íris em versão comercial, também observamos que o usuário pode se recusar a utilizar o equipamento . . . . .	16
2.5	Processo de análise da retina além da necessidade de especialistas para coleta e análise do material . . . . .	16
2.6	Relógio de ponto com reconhecedor de voz, problemas com ronquidão e resfriados . . . . .	17
2.7	Resultado de uma imagem sendo aplicada a transformada de Fourier e depois sua inversa . . . . .	25
2.8	Imagens de face sendo aplicados jets em diferentes poses . . . . .	26
2.9	Alguns elementos utilizados no reconhecimento por perfil, 1 - Pontos extremos da face; 2 - Posição e Contorno da narina; 3- Centro da Boca; 4 - Cantos dos lábios, 5 - Posição das fossas nasais; 6 - Contorno do queixo; 7 - Saliência do pombo de adão; 8 -Características do pavilhão auricular (CAMPOS, 2001) . . . . .	27
2.10	Imagem inicial e extração 3D, (ETHEVE; RICCIO; DUGELAY, 2006) . . . . .	27
2.11	Modelo básico para reconhecimento 3D, (ETHEVE; RICCIO; DUGELAY, 2006) . . . . .	29
2.12	Handheld 3D scanner, de mão, utiliza 3 scanners simultâneos . . . . .	30
2.13	Imagem ilustrativa do banco de Dados RMA, (ETHEVE; RICCIO; DUGELAY, 2006) . . . . .	32
3.1	Imagem utilizando a convenção definida . . . . .	39
3.2	Espaço de Faces visualizado em 2D. Os eixos representam dois <i>Eigenfaces</i> . . . . .	43
3.3	Imagem média, obtida através de um conjunto de imagens . . . . .	44
3.4	Eigenfaces . . . . .	45
3.5	Modelo empregado no reconhecimento facial . . . . .	47
4.1	Algumas faces presentes no banco de faces ORL . . . . .	50
4.2	Algumas imagens da fonte de dados utilizada para o reconhecimento . . . . .	51
4.3	Imagens de um indivíduo da base IFDB, a idade varia de 2 a 45 anos . . . . .	51
4.4	Modelo básico de Viola e Jones . . . . .	52
4.5	Exemplo de Características de Haar utilizados no modelo . . . . .	53
4.6	Imagem Integral, Teoria e resultado . . . . .	54
4.7	Modelo do Classificador em Cascata . . . . .	57
4.8	Correção Gamma . . . . .	59

4.9	Ajuste da Imagem através da correção gama, imagem da correção com fatores modificados . . . . .	59
4.10	Transformação logarítmica aplicada a imagem original e com $c = 0.5$ . . . .	60
4.11	Equalização do histograma . . . . .	60
4.12	As imagens (a) e (b) representam a captura sem o filtro e as (af) e (bf) representam as mesmas após o uso do filtro . . . . .	62
4.13	Tela principal do modelo implementado . . . . .	63
4.14	Tempo de treinamento - Base ORL . . . . .	65
4.15	Tempo de treinamento - Nossa base . . . . .	66
4.16	Formas de iluminação. Partindo do canto superior esquerdo temos iluminação direta e unilateral, angular, axial ou perpendicular, axial difusa, frontal difusa, iluminação de fundo, iluminação difusa uniforme. Adaptado de (GONZALEZ; WOODS; EDDINS, 2004) . . . . .	67
4.17	Representa a captura da imagem, formação e discretização. Adaptado (GONZALEZ; WOODS; EDDINS, 2004) . . . . .	67
4.18	Imagem captada sem auxílio de iluminação artificial em ambiente interno .	68
4.19	Imagem com alta incidência de luz, captada às 12:00h com iluminação natural	69
4.20	Resultado representado graficamente . . . . .	71
4.21	Relação entre a quantidade de imagens de treinamento e a taxa de reconhecimento utilizado no modelo . . . . .	71
4.22	Relação entre a quantidade de imagens de treinamento e a taxa de reconhecimento utilizado o modelo e sem filtros . . . . .	72
4.23	O fundo verde representa a imagem original, as demais com fundo azul representam as imagens modificadas e inseridas para treinamento . . . . .	72
4.24	Resultados do reconhecimento obtido em câmera de vídeo . . . . .	73
4.25	Máquina de venda de cigarros com reconhecimento de idade . . . . .	79

---

## Lista de Siglas

---

PPGMCTI ..	Programa de Pós-graduação em Modelagem Computacional e Tecnologia Industrial
WWW .....	World Wide Web
PCA .....	Principal Component Analysis
PIN'S .....	Personal Identification Number
DNA .....	Ácido Desoxirribonucleico
LDA .....	Linear Discriminante Analysis
KLT .....	Karhunen-Loève Transformation
TF .....	Transformada de Fourier
FERET .....	The Facial Recognition Technology Database
FRVT .....	The Face Recognition Vendor Test
3DRMA ....	Database 3D of Royal Military Academy of Belgium
CASIA .....	Chinese Academy of Sciences Institute of Automation
FACS .....	Facial Action Coding System
FRGC .....	Face Recognition Grand Challenge
BU-3DFE ...	Binghamton University 3D Facial Expression
ND-2006 ....	Notre Dame 3D Database
GAVABDB .	Database - Grupo de Algoritmica para la Vision Artificial y la Biometria
FLDA .....	Discriminant Linear Fisher
Cov .....	Covariância
MATLAB ...	Matrix Laboratory
OpenCV ....	Open Source Computer Vision
IDE .....	Integrated Development Environment
SO .....	Sistema Operacional
ORL .....	Olivet Research Laboratory
IFDB .....	Iranian Face Database
RGB .....	Color System, Red, Green, Blue
AdaBoost ...	Adaptative Boosting
CDF .....	Comulative Distributions Functions
JPG .....	Joint Photographic Experts Group
ONG's .....	Organizações Não-Governamentais
AVI .....	Audio Video Interleave
HP .....	Hewlett-Packard
fps .....	cenos por segundo

## Introdução

---

A imagem da face humana é extremamente cativante para diversos segmentos tanto da ciência como das artes. Naturalmente intuitiva ela desempenha um papel importante na nossa interação social e podemos facilmente reconhecer, identificar, perceber traços etimológicos, estimar a idade, perceber mudanças de humor entre outras reações. Essas revelam-se, ao ser humano, como um algo extremamente trivial, porém no contexto computacional esta trivialidade dá lugar ao um problema extremamente complexo e instigante. No início da década de 90, foi proposto uma solução para este problema, dois pesquisadores Matthew Turk e Alex Pentland propuseram um modelo de reconhecimento chamado *Eigenface* (TURK; PENTLAND, 1991a), que utiliza a análise de componentes principais ou simplesmente PCA e pressupostos de álgebra linear.

Apesar de ser considerado na década de 80, inclusive, um problema insolúvel para sistemas computacionais (HUBEL, 1989),(MINSKY, 1986), os avanços na área de hardware e software associados a uma grande dose de investigação científica permitiu que a pesquisa de Matthew Turk e Alex Pentland (TURK; PENTLAND, 1990),(TURK, 1991),(TURK, 2005) pudesse compreender e caracterizar os limites, limitações, possibilidades e variações do problema. No entanto, uma compreensão completa de como os seres humanos e máquinas interagem com a face ainda permanece como um desafio (TURK, 2005).

Vários pesquisadores(ZHUJIE; YU, 1994),(CHELLAPPA; WILSON; SIROHEY, 1995), (LAKSHMANAN et al., 2001), (ZHAO et al., 2003), (CHICHIZOLA et al., 2005), (LIU; WANG; YAU, 2006),(TAN et al., 2006), discutiram este modelo que consiste basicamente em reconhecer uma imagem que contenha uma face humana e associa-la a outra que esteja no banco de dados do previamente treinado. Esta ideia parte da premissa que a face humana possui características únicas e mensuráveis, conforme vimos na tabela 2.1, e que de modo não-invasivo é possível identificar pessoas, através da face, eletronicamente com um certo grau de confiabilidade (PARENTE, 1993), (SANTOS, 2007).

Sistemas que executam este tipo de tarefa são classificados como sistemas biométricos que fazem parte do campo de pesquisa chamado de Biometria, que pode ser sintetizado como o uso sistematizado de características fisiológicas, ou comportamentais de seres humanos tais como impressão digital, altura, peso, íris, reconhecimento facial, entre outras (SANTOS, 2007).

Dentre os sistemas biométricos com viabilidade econômica e computacional o reconhecimento facial tem provado(JAIN; ROSS; PRABHAKAR, 2004) ser um dos mais proeminentes

meios de identificar pessoas, muito devido as mais diversas aplicações que podem ser imputadas as método, conforme vemos na tabela 1.1. A área de segurança tem impulsionado fortemente as pesquisas (JENKINS, 2008), permitindo a identificação de criminosos ou terroristas nos aeroportos, monitoramento de fronteiras, localização de pessoas desaparecidas, entre outras, com o máximo de precisão possível. Mas a implementação de tais sistemas deve levar em conta qual a tarefa que irá ser realizada (ALMEIDA, 2006), pois os problemas em cada situação são diferenciados.

Área	Aplicação Específica
Entretenimento	Vídeo Games Realidade Virtual Acesso a áreas de lazer
Identificação em Cartões	Validação de Softwares Questões de Imigração Identificação Nacional (ID) Passaportes Eleições
Informações de Segurança	Controle Conteúdo em TVs Acesso a Computadores Aplicações de segurança Codificações de Arquivos Segurança/Acesso na Internet Dados Médicos Terminais de comércio seguro
Leis	Localização e rastreamento de Suspeitos Controle de tráfego Reconhecimento pessoas desaparecidas

Tabela 1.1: Aplicações típicas para o reconhecimento de faces (CHELLAPPA; WILSON; SIROHEY, 1995)

Das muitas abordagens que são utilizadas para solucionar o problema de reconhecimento facial (CHELLAPPA; WILSON; SIROHEY, 1995), os métodos holísticos são considerados por Yang, Ahuja e Kriegman (2000), Tan et al. (2006) e Jenkins (2008) as técnicas mais bem sucedidas, ou seja, técnicas que buscam aproximar-se do método humano de reconhecimento, realizando o reconhecimento por comparação. Entre estas técnicas de reconhecimento facial, a mais utilizada na atualidade é o *Eigenface*, que utiliza uma combinação de álgebra linear e modelos estatísticos ao seu favor para gerar um espaço de faces, a fim de que futuramente se possa comparar com outras faces de forma holística.

O *Eigenface* é um algoritmo baseado na PCA e foi desenvolvido a partir dos trabalhos de Kirby e Sirovich (1990) e Sirovich e Kirby (1987) por Turk e Pentland (1991a) e consiste basicamente em compor todas as imagens de treinamento em uma única matriz. A partir dessa matriz de treinamento são encontrados os seus autovetores e autovalores para, em seguida, extrair os de maior relevância. Após a escolha destes autovetores eles são comparados aos autovetores de uma dada imagem e a partir de um limiar estabelecido pode-se



classificar essa face como positiva, falso-positivo e não-face, como fica claro na figura 1.1.

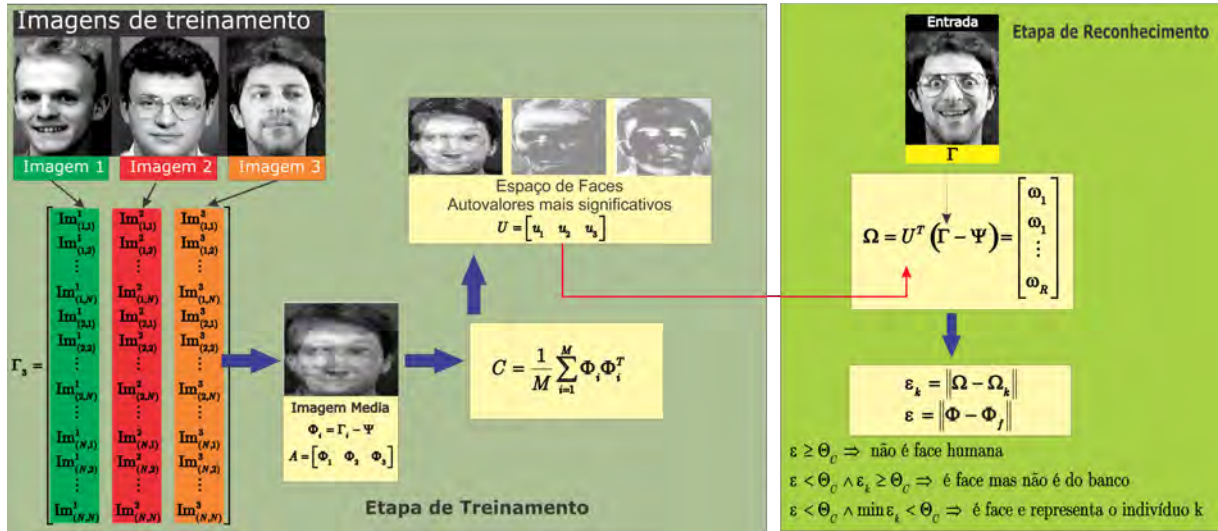


Figura 1.1: Modelo básico do Eigenfaces

Na realização deste trabalho foi construída uma base de dados para treinamento e reconhecimento que tenta simular uma situação real. A situação escolhida centra-se no problema de pessoas desaparecidas, pois apresenta uma série de situações peculiares como imagens em alto grau de degradação, variação de pose muito grande, iluminação não controlada, mudanças físicas decorridas pelo avançar da idade, ou mesmo por cirurgias plásticas e tom de pele. A soma de todas essas variáveis produziu uma série de dados que serão discutidos nas próximas sessões.

### 1.1 Definição do problema

De forma geral o problema pode ser formulado da seguinte maneira:

Dada uma ou um conjunto de imagens (fotografias em qualquer estado, que contenha a face frontal do indivíduo), como identificar e associar esta imagem a um quadro de vídeo ou outra imagem que contenha uma face?

Para realização desta tarefa teremos que, inicialmente, capturar a imagem da face do indivíduo pelo meio que ela for fornecida (estático ou dinâmico) logo em seguida transformá-la e duplicá-la para formar um conjunto de imagens. Aplicaremos em seguida, filtros de imagem que tornem o objeto capturado mais fácil de ser compreendido pelo modelo.

A esta etapa chamamos de pré-processamento, e estará presente tanto na fase de treina-

mento do algoritmo, como na fase de reconhecimento para que possamos minimizar os ruídos e deixar as imagens o máximo possível similares. Logo em seguida, por meio de uma entrada dinâmica ou estática, faremos a detecção da face do indivíduo comparando-a com a base de dados, gerando assim, uma massa de dados de objetos reconhecidos ou não.

A partir dos dados de reconhecimento, poderemos fazer uma análise, validando ou não o modelo para que ele venha apoiar ou não as autoridades para realizar a busca por pessoas desaparecidas. Para solucionar esta questão, utilizaremos um método estatístico considerado eficiente, por Zhuje e Yu (1994), Tan et al. (2006) e Khan e Alizai (2006), Zhao, Cai e Zhu (2009), chamado *Eigenfaces*, que podemos generalizar conforme a figura 1.2.



Figura 1.2: Fluxo básico para reconhecimento de faces

A percepção/reconhecimento da face é uma importante capacidade humana e extremamente comum a todas as pessoas, umas mais que outras. O grande desafio é construir sistemas computacionais que reproduzam essa capacidade e que busquem aperfeiçoar esta habilidade, desafio que tentaremos ajudar a solucionar ao longo deste trabalho. Com base nestes pressupostos podemos estabelecer logo a seguir os objetivos deste trabalho.

## 1.2 *Objetivo*

A técnica aplicada no modelo de reconhecimento que faremos uso classifica a face de maneira muito intuitiva, na qual toda a informação relevante da imagem facial é utilizada. Com isso a face deve ser extraída da imagem e tratada para obtermos um reconhecimento mais preciso possível.

Diferentemente de outras técnicas de reconhecimento facial (ZHOU; CHELLAPPA; ZHAO, 2006), este método utiliza uma quantidade maior de informação. Assim, levando-se em conta o grande número de dados que virá a ser gerado por um banco de dados, faz-se necessário fazer uma análise do desempenho do método quando aplicado a um grande número de informação.

### 1.2.1 *Objetivo Geral*

Avaliar a eficiência do método estatístico de reconhecimento de face, chamado *Eigen-face* aplicado a uma situação não controlada de detecção e reconhecimento de pessoas desaparecidas a partir de uma ou de um conjunto de imagens de entrada.

### 1.2.2 *Objetivos Específicos*

- Realizar em condições mínimas de iluminação a detecção de uma face;
- Estabelecer quantidades mínimas de imagens de entrada para o reconhecimento;
- Minimizar o problema da iluminação durante a aplicação do algoritmo com o uso de filtros de suavização;
- Aplicar e avaliar um modelo de identificação de idade;
- Avaliar o modelo quanto ao problema de reconhecimento de pessoas com diferentes etnias.

## 1.3 *Importância da pesquisa*

Com a criação do Cadastro Nacional de Crianças e Adolescentes Desaparecidos, através da Lei 12.127/09, diversos órgãos governamentais terão a sua disposição um conjunto de dados que facilitarão o processamento das informações. Porém, de nada adianta uma

enorme quantidade de dados se o contingente de policiais e investigadores que tratam do assunto é reduzido (OLIVEIRA, 2007). Além disso, entende-se que esse é um assunto que produz transformações significativas numa parcela considerável de famílias que vivem esse drama (OLIVEIRA, 2007). A tecnologia atual permite, a custos aceitáveis, a implementação de algoritmos de busca (ROSÁRIO, 2005) utilizando, por exemplo, câmeras de vigilância já instaladas nas principais capitais do país.

Complementando, temos atualmente, diversas tentativas de identificar pessoas através da sua face, com resultados promissores, porém estes testes foram realizados, em sua maioria, com imagens captadas em ambientes controlados, (CHICHIZOLA et al., 2005), (DELAC; GRGIC, 2004), (LIAU et al., 2008), em termos de iluminação e fundo. Como nossos dados se aproximam de uma realidade não controlada, ou sejam um ambiente completamente aleatório, poderemos avaliar o modelo contrapondo os dados.

## **1.4 Motivação**

Das muitas formas de reconhecimento biométrico já exploradas, o reconhecimento facial tem tido uma atenção significativa por parte da comunidade científica (ALMEIDA, 2006), (HJELMAS ERIK; LOW, 2001), (KHAN; ALIZAI, 2006), e apesar desse esforço vários problemas ainda precisam ser solucionados para que um sistema de reconhecimento obtenha índices de acerto acima de 95% (ZHAO et al., 2003). Além disso, muitos trabalhos exploram apenas a localização, sem preocupação com a identificação ou trabalham com bases de dados em condições ideais de iluminação e pose. Este cenário torna a pesquisa na área de grande valor científico.

O principal elemento motivador para atacar o problema envolve o potencial de aplicação dos algoritmos de reconhecimento, que podem ser utilizados para o monitoramento de multidões e no nosso caso específico no reconhecimento de pessoas desaparecidas.

A possibilidade de aplicação do *Eigenfaces* na solução do problema também foi motivado pela criação do Cadastro Nacional de Crianças e Adolescentes Desaparecidos que permite a integração das diversas corporações policiais do país. Isso permitirá que câmeras, espalhadas pela cidade em pontos chaves, possam identificar e informar as autoridades policiais se o indivíduo reconhecido pertence ou não a este banco de dados.

## 1.5 Limites e limitações

Mesmo sob condições ideais, os algoritmos e sistemas de reconhecimento estão longe de serem confiáveis para Jenkins (2008), no caso do *Eigenfaces* uma limitação que tem que ser administrada pelo sistema de reconhecimento são as diferentes condições de luz presente num ambiente não controlado.



Figura 1.3: Diferentes condições de iluminação

O efeito causado na figura 1.3, é um exemplo de situação em que ocorrem falhas no reconhecimento. Outrossim, destacamos:

1. Identificação de uma face única em uma multidão;
2. Baixa qualidade do hardware de captura de vídeo;
3. Problemas ligados ao algoritmo;
4. Construção do banco de dados de imagens (Fotos muito antigas ou com pouca iluminação);
5. Dificuldades com novos elementos acrescentados ao indivíduo e não ao banco de dados (barba, bigode, óculos, plásticas e etc);
6. Distância mínima para identificação e reconhecimento da face.

Os itens descritos descrevem limitações encontradas, basicamente, na relação entre hardware e software, mas outras limitações se interpõe a este trabalho de pesquisa, dos quais destacamos:

1. Construção do banco de dados para teste;
2. Número de pessoas envolvidas nos testes com o algoritmo;

3. Aquisição de hardware que compense problemas com iluminação.

Somando-se aos itens destacados, sabemos que como o reconhecimento de faces não tem retornos exatos, seria possível pensar na utilização do mesmo algoritmo (*Eigenfaces*) para tentar identificar outros aspectos do indivíduo como tonalidade de pele e idade. Tentamos avaliar o comportamento do modelo com relação nestes eventos no intuito de entender o funcionamento do espaço de face quando submetido a outras tarefas.

## **1.6 Aspectos metodológicos**

A face, dentre as características biométricas, é a mais usada para o reconhecimento ou identificação de indivíduos, devido ao seu processo de obtenção não intrusivo, isto é, não há necessidade da cooperação da pessoa para que sua face seja capturada (SILVA, 2006) como também não se configura invasão de privacidade, pois qualquer estabelecimento comercial ou órgão de polícia pode solicitar uma identificação do cidadão, para fins de concessão de crédito, acesso ao estabelecimento, etc. Isto permite que criminosos ou pessoas com problemas mentais não tenham a necessidade de se colocar em uma determinada posição para que se efetue o reconhecimento. Para isso, estabelecemos um método de trabalho baseado em aspectos estatísticos, aliado a isso trabalhos que se baseiam em faces humanas quase sempre se fundamentam suas investigações em condições ideais; poucos trabalhos exploram os algoritmos em situações reais. Este trabalho se propõe a tentar encurtar o hiato entre os diversos experimentos em condições ideais e as situações, carregadas de ruído, da realidade.

Para realizar esta tarefa fundamentaremos nosso trabalho numa abordagem experimental, na qual, sistematicamente provocaremos mudanças no ambiente a ser pesquisado e observaremos se cada intervenção produz ou não resultados esperados, sendo os experimentos e as observações tratados de maneira objetiva. Fundamentaremos esta pesquisa, na filosofia do empirismo (WAZLAWICK, 2008), trabalhando com variáveis discretas e dimensionando o espaço de pesquisa a um determinado meio.

Esta opção metodológica permitirá a realização de um certo número de experimentos controlados que, estatisticamente, comprovem a validade do método estudado.

De maneira geral temos as seguintes etapas:

1. Inicialmente, montamos um banco de imagens, com fotos que contenham a face frontal do indivíduo (voluntário) a ser localizado. Tomaremos dois grupos distintos sendo o primeiro compostos de apenas uma imagem, preferencialmente com fotos

antigas (até 3 anos anteriores à idade do indivíduo) com ruído ou processo de degradação em andamento. Segundo, tomaremos mais de uma imagem do mesmo indivíduo composta de diversas poses e em diferentes condições de luz.

2. A partir do software construído e do algoritmo implementado, serão feitas várias sessões para verificar a quantidade de pessoas identificadas (positivos) e não identificadas ou similares (falsos-positivos).
3. Nesta etapa faremos a análise dos dados coletados, estudaremos os resultados e validaremos ou não a utilização algoritmo estudado para o conjunto de teste.

A figura 1.4 retrata este processo metodológico:

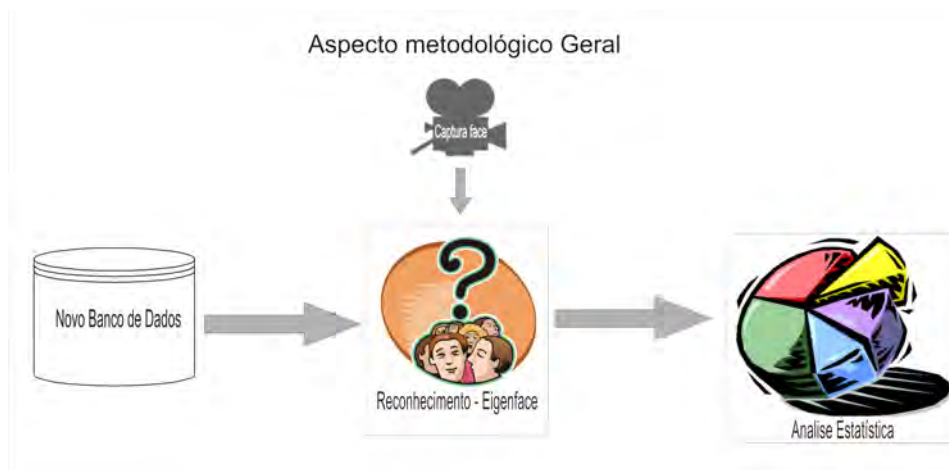


Figura 1.4: Ilustração do funcionamento do sistema de reconhecimento de Face

Separamos cada conjunto de dados (imagens dos indivíduos) em classes. Sendo que essas foram devidamente balanceadas, ou seja, cada classe contém um número equivalente de fotos e todas passarão por um processo de extração de características, melhoria nas condições de iluminação, equalização e dimensionalidade.

## 1.7 Organização da dissertação

Inicialmente, definimos nosso problema e traçamos os objetivos geral e específico. Mostramos a importância da pesquisa, centrada em dois pontos: a localização de pessoas desaparecidas e uma avaliação do método *Eigenface* aplicado a este problema. Buscamos estabelecer limites para o modelo e as limitações gerais da pesquisa além disso fomos norteados por uma metodologia consagrada na avaliação de um método estatístico de reconhecimento facial. Para cumprir nossos objetivos dividimos esta dissertação em 5 partes incluindo a que escrevemos inicialmente, a qual chamamos de introdução, da seguinte forma:

- O capítulo 2 apresenta uma explanação geral sobre os trabalhos correlatos do reconhecimento de face, focando nos principais avanços do *Eigenfaces* e nos problemas encontrados no reconhecimento em tempo real. Procurando ilustrar as principais técnicas de reconhecimento visando ressaltar a relação entre estas e o método adotado para dar sustentação ao nosso trabalho. Inicialmente fazemos um panorama sobre a área de maior abrangência do reconhecimento de padrões que está inserido o reconhecimento facial, a Biometria. Com base nestas informações restringimos nosso universo de trabalho ao reconhecimento facial para logo em seguida mostramos os principais avanços discutidos a luz da literatura mais recente possível.
- No capítulo 3 expusemos o modelo adotado. Esta exposição pautou-se em definições matemáticas de todos os elementos envolvidos no processo do modelo em específico as definições de algebra linear e do PCA. Fixada a base matemática, apresentamos, de maneira completa, todos os aspectos do reconhecimento facial através do modelo.
- O capítulo 4 refere-se à implementação do algoritmo apresentando objetivamente a sua estrutura e avaliações obtidas através dos testes efetuados. Buscamos descrever todos os passos que foram estruturados começando pela base de dados e a forma de detecção da face utilizado a etapa de pré-processamento. Estabelecidos estes parâmetros buscamos realizar os testes do modelo implementado e realizamos discussões sobre cada teste executado.
- Por fim no capítulo 5 finalizamos a dissertação com as conclusões gerais, perspectivas futuras e contribuições gerais do trabalho.



---

## Trabalhos Correlatos

---

### 2.1 Introdução

O reconhecimento de face pode parecer algo extremamente futurista, mas durante os últimos dez anos, inúmeros esforços tornaram possível que sistemas que propõem realizar esta tarefa possam se tornar uma realidade viável. Estes sistemas podem ser de valor inestimável e possuem várias aplicações, tais como, controle de acesso a áreas de segurança, defesa nacional e segurança internacional. Mas, simular a nossa capacidade de reconhecimento não é uma tarefa trivial, pois exige um grande esforço computacional e diversas limitações. Neste capítulo iremos discutir qual o estado da arte do reconhecimento facial tratando o assunto sob a ótica da biometria, passando pelos avanços mais recentes do reconhecimento facial e do algoritmo de reconhecimento mais utilizado pela comunidade científica o *Eigenfaces*. Portanto, temos como objetivo neste capítulo, fundamentar teoricamente a nossa opção metodológica para a escolha deste modelo.

### 2.2 Estudo da Biometria

A palavra biometria vem do grego *bios*(vida) e *metrikos*(medida), muito intuitiva aos seres humanos que a utilizam corriqueiramente ao identificar um amigo, reconhecer uma voz, entre outros aspectos. De maneira simples, podemos definir a biometria como o reconhecimento automático de uma pessoa com base em características físicas que a distingue das demais(WOODWARD et al., 2003), mas de maneira geral, a biometria é a ciência que utiliza qualquer medida mensurável do corpo humano para identificá-lo ao reconhecê-lo (DESCONHECIDO, 2004). Porém, a biometria vem ganhado um novo significado, isto porque do ponto de vista da segurança, biometria significa a verificação da identidade de uma pessoa através de uma característica única inerente a ela, sendo que esta característica pode ser fisiológica ou comportamental, desde que preencha os requisitos descritos na tabela 2.1.

Um sistema biométrico é essencialmente um sistema de reconhecimento de padrões que opera através de dados biométricos de uma pessoa, adquiridos e comparados com um gabarito previamente gravado. Dependendo do contexto ele pode operar como identificador ou verificador:

- No modo *Verificação*, o sistema valida a identidade de uma pessoa, comparando os dados biométricos capturados com seu próprio modelo biométrico armazenado

Universalidade	cada pessoa tem uma característica;
Distinção	quaisquer que sejam, duas pessoas devem ser suficientemente diferentes em termos de características;
Permanência	a característica deve ser invariante;
Mensuração	a característica pode ser medida quantitativamente;
Performance	que se refere a exatidão possível, velocidade e recursos necessários para atingir o objetivo, bem como fatores operacionais e ambientais;
Aceitabilidade	número de pessoas que estão dispostas a aceitar o uso de uma identificação biométrica;
Fraude	que reflete a facilidade com que o sistema pode ser enganado por métodos fraudulentos.

Tabela 2.1: Tabela com as características gerais de um sistema biométrico,(JAIN; ROSS; PRABHAKAR, 2004)

num banco de dados. Neste tipo de sistema, o indivíduo deseja ser reconhecido e reivindica sua identificação. O reconhecimento positivo, tem o objetivo de evitar que várias pessoas usem a mesma identidade (WAYMAN, 2001).

- No modo *Identificação*, o sistema reconhece o indivíduo procurando os modelos de todos os usuários no banco de dados, comparando os dados de entrada com vários modelos, para estabelecer a identidade do indivíduo sem que o esse solicite sua identificação. Este tipo de aplicação é crítica em aplicações de reconhecimento negativo, quando o sistema verifica se a pessoa é o que ela diz ser(WAYMAN, 2001).

A tecnologia biométrica provou ser eficiente na substituição de PIN'S e senhas, independentemente do mercado, tecnologia do cartão ou sistema e tem como principais vantagens a segurança, velocidade e aceitação do usuário (JAIN; ROSS; PRABHAKAR, 2004). Ao usar a biometria é possível reconhecer uma pessoa baseado no que você é, e não por aquilo que você possui, ou que você lembra. Além de aumentar a segurança, os sistemas biométricos também melhoram a conveniência do usuário.

Apesar das vantagens, a tecnologia tem problemas de confiabilidade principalmente pelo elevado número de falsos-positivos e rejeitados (DELAC; GRGIC, 2004). Diferente de estudos em condições ideais quando a taxa de acerto chega a 90%, em uma situação real, as condições deixam de ser suceptíveis e muitas pessoas deixam de utilizar o sistema por não confiar em sua eficiência(WAYMAN, 2001). Avanços recentes nas pesquisas (CHICHIZOLA et al., 2005),(DELAC; GRGIC, 2004),(EL-BAKRY; HAMADA, 2009),(LIU; WANG; YAU, 2006),(MONWAR et al., 2006), (PUYATI; WALAIRACHT, 2008), (WOODWARD et al., 2003), (ZHAO; ZHUANG; XU, 2008), (ZHAO et al., 2009), tem diminuído este hiato entre as situações reais e teóricas.

Existem vários métodos biométricos em uso (alguns comerciais e alguns ainda não). Fa-

remos um breve panorama das diferentes características biométricas iniciando com as tecnologias mais recentes até as mais antigas:

- **DNA** - O ácido desoxirribonucleico é uma molécula formada por duas cadeias na forma de uma dupla hélice, único em cada ser humano, exceto pelo fato que gêmeos univitelinos podem conter o mesmo DNA. Pode ser colhido através de uma gota de sangue, um fio de cabelo ou de osso, além da saliva conforme vemos na figura 2.1 . É portanto, bastante utilizado como método forense de reconhecimento de pessoas, muito preciso, mas possui três questões limitantes em aplicações:
  1. A contaminação e sensibilidade: é fácil roubar um pedaço de DNA de um sujeito.
  2. Problema no reconhecimento automático: requer análise através de recursos químicos por especialistas e torna-se bastante invasivo.
  3. Questões de privacidade: informações do DNA podem revelar susceptibilidade do indivíduo a determinadas doenças além de questões que envolvem discriminação.



Figura 2.1: Figura mostra como o processo pode ser invasivo, além do indivíduo poder recusar-se a fazer o teste

- **Varredura de veias da mão** - Esta tecnologia pode identificar uma pessoa a partir dos padrões dos vasos sanguíneos na parte posterior da mão. A tecnologia usa luz do infravermelha para detectar estes padrões. Mesmo entre gêmeos, este padrão é distinto, até mesmo entre a mão direita e esquerda de uma mesma pessoa. A tecnologia não é invasiva e funciona mesmo se a mão não estiver limpa. É comercialmente disponível como vemos na figura 2.2.
- **Pulso Arterial** - Esta técnica mede, através de sensores infravermelhos, o pulso sanguíneo em um dedo. Ainda experimental tem uma alta taxa de falsos positivos, tornando-se atualmente impraticável para a identificação pessoal



Figura 2.2: Identificação biométrica já utilizada em terminais de auto atendimento bancário.

- **Elementos de Pele** - A composição da pele é singular em cada pessoa. As camadas da pele diferem em espessura entre as camadas, ondulações, pigmentação, fibras de colágeno e leitos capilares. O funcionamento da técnica trabalha com um sensor de luz que ilumina um pequeno fragmento de pele com um feixe de infravermelho próximo e visível. A luz é medida com um espectroscópio depois de ser espalhado pela pele. As medições são analisadas, e um padrão óptico distinto é extraído.
- **ORELHA** - Sugere-se que a orelha tenha um uma estrutura singular em cada indivíduo, tanto na estrutura cartilaginosa do pavilhão interno como do externo. Combina-se daí a distância entre pontos mais proeminentes deste ouvido, além de contornos e tamanho do pavilhão. É pouco intrusivo, mas pode ser burlado com um molde ou plásticas que deformem o pavilhão. Embora promissor a biometria do ouvido ainda não desenvolveu sistemas comerciais.
- **Impressão digital** - Largamente utilizada por diversos governos como meio de identificação e comprovação da identidade do indivíduo há muito tempo e vem demonstrando ser eficaz no reconhecimento e identificação de pessoas (MAIO et al., 2002), como por exemplo nas eleições de 2010, figura 2.3. Analisa pequenos atributos da impressão conhecidos como minúcias (PRABHAKAR, 2001), que podem ser definidas como contornos das linhas papilares ou bifurcações (ramificações das linhas papilares). Também podem ser utilizados pequenos poros dos dedos que são posicionados de forma única em cada indivíduo. O seu baixo custo, torna-a acessível a um grande número de aplicações. Suas limitações são geradas por fatores de envelhecimento e rugosidade da pele, sujeira no leitor ou fatores ambientais que podem causar problemas no reconhecimento.
- **Reconhecimento de Marcha** - A técnica analisa o andar característico de uma pessoa. Captando, em uma sequência de imagens, a variação do movimento. No entanto, a marcha pode ser ocultada ou dissimulada, se o indivíduo, por exemplo,



Figura 2.3: Imagem mostra a utilização pela justiça eleitoral brasileira de um método biométrico

estiver vestindo roupas largas e soltas. Resultados preliminares confirmam o potencial do método, mas a sua aplicação ainda depende de maiores estudos que definam suas limitações e vantagens.

- **Termograma Facial** - Detecta padrões de calor gerados pela ramificação dos vasos sanguíneos e emitidos a partir da pele. Estes padrões chamados de termogramas, são distintos e capturados por uma câmera infravermelha. Sua grande vantagem é que ao contrário de sistemas que dependem de boas condições de iluminação os infravermelhos podem reconhecer com muito pouca luz. Embora os sistemas de identificação facial com termogramas tenham sido idealizados em 1997, sua pesquisa é difícil devido aos altos custos envolvidos no processo.
- **Íris** - A íris é a zona do globo ocular responsável pelo colorido do olho, tendo a função de separar a pupila da esclerótica, que constitui a parte branca do olho. O diâmetro médio da pupila é de 4,4 mm, e essa localiza-se no centro da íris. A íris trata-se de uma membrana circular de aproximadamente 12 mm de diâmetro. No contexto do globo, sua posição está entre a córnea e o cristalino, dilatando ou contraindo sua abertura central de acordo com a intensidade da luminosidade, controlando a passagem de luz. Atualmente, a precisão e a velocidade dos sistemas de identificação por íris tornam este método bastante promissor em aplicações comerciais de segurança(JAIN; ROSS; PRABHAKAR, 2004) como vemos na figura 2.4;
- **Padrão de Digitação** - Baseia-se na hipótese de que cada pessoa não digita em um teclado de forma característica. Este comportamento biométrico não é único em um indivíduo, mas oferece informações suficientes para que permita o reconhecimento.
- **Odor** - Sabe-se que cada ser humano exala um odor característico(JAIN; ROSS; PRABHAKAR, 2004), de sua composição química e isso poderia ser usado para distinguir o indivíduo. Por ser uma técnica recente(DELAC; GRGIC, 2004), não está



Figura 2.4: Mostra um leitor biométrico da íris em versão comercial, também observamos que o usuário pode se recusar a utilizar o equipamento

claro se a invariância do odor corporal(desodorantes, perfumes, cremes, etc) pode ser detectado ou não, tornando o reconhecimento extremamente difícil.

- **Palma da Mão** - A palma da mão possui padrões similares aos encontrados na impressão digital. Como a área da mão é maior que a área do dedo humano, espera-se que o reconhecimento torne-se mais apurado. Quando usada em um scanner de alta resolução, todas as características da palma da mão, como a geometria, cume e vale, linhas principais e rugas podem ser combinados e criar um sistema de verificação de alta precisão. Já utilizada por alguns bancos privados.
- **Varredura de Retina** - O reconhecimento de retina é uma tecnologia que tem a retina humana como meio de verificação e autenticação. Identifica o indivíduo, devido aos padrões de vasos sanguíneos existentes na retina obtidos através da reflexão da luz causada sobre eles. Considerado um meio extremamente seguro, já que não é fácil alterar ou reproduzir a vascularização da retina. Porém a aquisição da imagem é um processo voluntário o que afeta negativamente a utilização do método, além da necessidade de especialistas para fazer a identificação com vemos na figura 2.5.
- **Assinatura** - O modo como uma pessoa assina é uma característica singular. Analisar os dados de pressão, aceleração, velocidade e movimentos aéreos a partir da assinatura em um *tablet* e mediante a aplicação de fractais, comparar com os dados biométricos da assinatura cadastrada do usuário. É um padrão comportamental e portanto, muda ao longo da vida da pessoa devido a condições físicas e emocionais. Além disso é facilmente reproduzida por especialistas em falsificação.
- **Voz** - Consiste em identificar um indivíduo através da verificação de características comportamentais e fisiológicas como trato vocal, boca, fossas nasais e lábios que são usados na síntese do som. Sendo que estas características fisiológicas são invariantes



Figura 2.5: Processo de análise da retina além da necessidade de especialistas para coleta e análise do material

não ocorre o mesmo com as características comportamentais que fazem com que o discurso da pessoa mude ao longo das suas experiências. Muito utilizado em aplicativos de telefonia sofre com a degradação na qualidade do sinal devido ao microfone ou canal de comunicação e na marcação de entrada e saída do trabalho, figura 2.6.



Figura 2.6: Relógio de ponto com reconhecedor de voz, problemas com ronquidão e resfriados

- **FACE** - O reconhecimento de face é um método não-invasivo que utiliza informações relativas a face do indivíduo para realizar a identificação. Pode ser utilizado de duas formas:
  1. De forma local: análise de atributos faciais como olhos, sobrancelhas, nariz, lábios queixo e relações espaciais.



2. De forma global: a imagem do rosto é representada como uma combinação de uma série de faces e depois, comparada com uma face de entrada.

Sem muitas informações da face é extremamente difícil o reconhecimento de uma pessoa em uma multidão o que torna discutível o nível de confiança dos sistema(GOLFARELLI; MAIO; MALTONI, 1997). Um sistema de reconhecimento facial, na prática, deve *(i)* detectar a face, *(ii)* reconhecer a face.

A utilização de um método biométrico específico depende de vários fatores como logísticos, de confiabilidade, financeiros, mas de certo que isoladamente, nenhuma técnica é completa ou perfeita, mas aplicável a solução de um problema. Em evolução constante cada vez mais, estes sistemas estarão presentes no cotidiano das pessoas. A tabela 2.2 faz uma comparação entre os métodos e nos mostra como a face tem uma excelente aceitação, porém tem problemas com a sua performance e com a similaridade entre as pessoas.

Identificador Biométrico	Universalidade	Distinção	Permanência	Mensuração	Performance	Aceitabilidade	Fraude
DNA	A	A	A	B	A	B	B
Orelha	M	M	A	M	M	A	M
Face	A	B	M	A	B	A	A
Termograma da Face	A	A	B	A	M	A	B
Impressão digital	M	A	A	M	A	M	M
Marcha	M	B	B	A	B	A	M
Geometria da Mão	M	M	M	A	M	M	M
Veias da mão	M	M	M	M	M	M	B
Iris	A	A	A	M	A	B	B
Teclar	B	B	B	M	B	M	M
Odor	H	A	A	B	B	M	B
Impressão da mão	M	A	A	M	A	M	M
Retina	H	A	M	B	A	B	B
Assinatura	B	B	B	A	B	A	H
Voz	M	B	B	M	B	A	H

Tabela 2.2: Tabela comparativa de métodos biométricos. A - Alta performance; M - performance Média; B - Baixa performance. (JAIN; ROSS; PRABHAKAR, 2004)



### 2.2.1 Aplicações da Biometria

O universo das aplicações da biometria é vasto e versátil, podendo trazer ao usuário ganhos surpreendentes em termos de segurança, simplicidade e rapidez na validação de resultados. Contudo, as aplicações biométricas podem ser simplesmente categorizadas, sendo elas:

- Aplicações comerciais - *logon* de redes, segurança de dados eletrônicos, comércio eletrônico, acesso a internet, cartões de banco, ensino a distância entre outros.
- Aplicações Governamentais - Cartão de identidade, carteira nacional de trânsito, INSS, bolsa família, controle de passaportes, etc.
- Aplicações forenses - Identificação de cadáveres, investigação criminal, combate ao terrorismo, análise de paternidade e maternidade, localização de pessoas desaparecidas. (JAIN; ROSS; PRABHAKAR, 2004)

Sendo assim, nosso trabalho é focado em aplicações forenses. Visto que o estudo biométrico da face tem um bom rendimento de falsos-positivos (DELAC; GRGIC, 2004) como também percebendo a face como um componente chave na maneira com nos relacionamos (o que diminui o impacto ao usuário) uns aos outros, buscaremos concentrar esforços para minimizar o erros, tentando assim, tornar viável sua aplicabilidade no problema de pessoas desaparecidas. Embora o fato de reconhecer outro indivíduo pelas características faciais seja algo completamente intuitivo aos seres humanos, o reconhecimento biométrico de face não é um problema trivial, mas suas vantagens tornam-o um importante auxiliar de governos e sociedade civil organizada.

Sistemas que captam faces em áreas públicas são discretos, tem um custo baixo em relação a outras técnicas, além disso não oferecem contato físico do indivíduo com o equipamento. O funcionamento desses sistemas dependem de banco de dados para treinamento do algoritmo, com um número de imagens da face do indivíduo que pode variar de um para muitos, ou ainda retratos falados.

## 2.3 Reconhecimento Facial

### 2.3.1 Introdução

O reconhecimento facial é uma área da biometria de grande interesse da ciência da computação e as primeiras tentativas efetivas de resolver este problema aconteceram no início

dos anos 1970 e ganharam força na década seguinte (ZHAO et al., 2003), (DELAC; GRGIC, 2004).

Quando falamos em reconhecimento facial nos referimos a processos automatizados ou semi-automatizados que confrontam imagens faciais. A ideia mais utilizada e mais recomendada de realizar o reconhecimento facial não difere muito da capacidade humana (JENKINS, 2008), apesar dos humanos terem larga vantagem neste campo. Para reconhecer uma face de entrada num sistema pode-se mapear a geometria da face, suas proporções ou ainda decompô-la para extrair as principais características. Registrados estes delimitadores é possível estimar distâncias e iniciar as comparações com uma base de dados previamente montada. A tecnologia atual leva em conta medidas biométricas da face e enfrenta muitos desafios.

A chave para obter reconhecimento positivo, ou seja, sucesso no processo, está diretamente ligada à formação do banco de dados. Nesta etapa uma ou mais imagens da face são agrupadas e codificadas para o uso no modelo, esse conjunto forma o que chamamos de espaço de faces. Quão robusto for este espaço de face, mais chances teremos de sucesso, pois após a seleção de uma face alvo e estabelecer a similaridade com o banco de dados o algoritmo trabalha para buscar características entre postura da cabeça, expressão facial entre outras que definem a face. Como lidamos com *pixels*, lembramos o fato que na verdade analisamos a intensidade luminosa de cada um deles em relação aos que estão no banco de face (CHELLAPPA; WILSON; SIROHEY, 1995).

O reconhecimento de face integra-se a área de visão computacional que utilizam faces para tentar identificar uma pessoa ou verificar a identidade. Independentemente do método utilizado este processo pode ser definido em cinco etapas:

1. Aquisição da imagem: geralmente se utiliza uma fotografia ou uma imagem digitalizada obtida por uma câmera;
2. Localização de uma face nesta imagem: esta tarefa visa retirar da imagem partes que não interessam ao reconhecimento como o fundo e outros elementos, em geral utiliza-se o conceito de “parecido com”;
3. Análise da imagem: utiliza-se aqui diferentes técnicas para extrair as características de cada indivíduo. O método mais popular é o baseado na Análise de Componentes Principais (PCA) chamado de Eigenface(ZHOU; CHELLAPPA; ZHAO, 2006).
4. Comparação: comparar o modelo gerado com o banco de dados previamente estabelecido computando as pontuações e comparando-as para identificar qual a menor.
5. Escore: Determinado o escore compara-se com as regras estabelecidas e verifica se

a face encontrada é ou não do indivíduo. Podem ocorrer três casos: Positivo, falso-positivo e negativo.

Uma observação importante na aquisição da imagem é observá-la como um padrão visual em três dimensões que é transformado em duas dimensões. Esta análise permite que seja clara a ideia de que neste processo, certamente, existirão perdas e deformações.

### 2.3.2 Problemas e Limitações

Devido a natureza dinâmica do reconhecimento facial, podemos encontrar diversos problemas durante o processo de reconhecimento. Alguns desses problemas podem ser tratados e outros não, daí classificaremos o algoritmo de estudo como robusto ou fraco com base em seu desempenho de reconhecimento sobre qualquer circunstância. Segundo Atalay (1996), podemos classificar um sistema como robusto se o mesmo conseguir superar os seguintes problemas:

- Variância de Escala: a mesma face pode ser apresentada para o sistema de diferentes escalas, ou seja, a distância entre o rosto e a câmera influencia o reconhecimento.
- Variação de posição: A mesma face pode ser captada sob diferentes perspectivas e orientações, ou seja, de perfil, frontal, com rotações e translações.
- Variação de iluminação: uma mesma imagem da face captada sob diferentes condições de iluminação tais como posição e força da fonte de luz pode causar ambiguidades.
- Expressões emocionais e detalhes faciais: sorrir, fazer caretas, barbas e bigodes, usar óculos, entre outras características pode interferir na operação.
- Ruídos: ruídos gerados por granulações na imagem, defeitos de hardware, oclusão da câmera e objetos que venham interferir na precisão da imagem

Os problemas expostos acima não contemplam todas as limitações que vem a surgir no processo de reconhecimento. Pois outros entraves acontecem e somam-se ao que vimos ao longo do processo como falhas de hardware e software. Estes tópicos revelaram-se presentes nas maioria das pesquisas (BELHUMEUR; HESPANHA; KRIEGMAN, 1997),(YANG; AHUJA; KRIEGMAN, 2000),(LAKSHMANAN et al., 2001), (ALMEIDA, 2006),(TAN et al., 2006), (LIAU et al., 2008) e costumam se revelar o ponto fraco do modelo de reconhecimento facial.

Apesar de questões de performance atrapalharem o processo após o treinamento a velocidade do reconhecimento é extremamente rápida, pois, apenas calculamos uma distância

simples e fazemos comparações. Outro fato pitoresco ao reconhecimento facial é a não necessidade de registro, o que torna a técnica bastante eficaz, visto que uma pessoa não precisa se aproximar de algum aparelho ou ser perfurado de alguma forma.

Variações de técnicas antigas e novas técnicas aplicadas a este recurso biométrico tem sido propostas por vários pesquisadores e testados pela comunidade científica (CHELLAPPA; WILSON; SIROHEY, 1995), algumas com relativo sucesso científico e outros que despertam desconfiança ou que dependem de grandes avanços computacionais como o processamento quântico. Na próxima seção discutiremos alguns desses métodos e sua aplicabilidade.

## 2.4 Técnicas de Reconhecimento Facial

Segundo a *Biometric Data Interchange Formats*, pacote de requisitos de segurança que auxilia bancos, agências governamentais e outras organizações pelo mundo, (DESCONHECIDO, 2004), o reconhecimento de face vai se tornar parte integrante de nossas vidas, com o avanço dos métodos de identificação biométricos.

Dai, técnicas eficazes que minimizem as limitações e tornem o sistema robusto tem sido estudadas desde a década de 1970. Estes avanços são vistos como parte integrante de grandes tecnologias e da próxima geração de ambientes inteligentes (H.K.; SANKUR, 2005).

Atualmente os métodos de reconhecimento facial são divididos em dois tipos:

- Métodos baseados na aparência
- Métodos baseados em características

Os métodos baseados na aparência são mais populares e vem demonstrando grande sucesso, sobretudo por serem extremamente similares à forma com que os humanos fazem o reconhecimento de pessoas (JENKINS, 2008). São caracterizados desta forma por não utilizarem nenhum conhecimento *a priori* sobre o objeto ou características a serem detectadas. Surge neste modelo, as ideias de aprendizado e treinamento, visto que estas etapas são necessárias para a realização do reconhecimento, sem intervenção externa.

Nestes, uma imagem, de uma certa dimensão, é representada por um vetor no espaço de dimensão  $N$  para permitir o rápido reconhecimento. Como a dimensão pode, e é, um grande entrave para a técnica, foram desenvolvidas e/ou aplicadas técnicas de redução de dimensionalidade, como a Análise de Componentes Principais. Estas técnicas minimizam a questão da dimensionalidade mas se faz necessária uma reflexão sobre as ideias de (ELAD;

GOLDENBERG; KIMMEL, 2007), ao discutir que uma imagem pode ser comprimida, mas para ser identificada, em algum ponto ela terá que ser descompactada. Isso significa que sistemas de reconhecimento que lidam com grandes quantidades de imagem dependem de um hardware competente, que consiga, de maneira rápida, efetuar os cálculos e apresente seus resultados.

Os métodos baseados em características buscam utilizar características estruturais que existem mesmo quando a pose, o ponto de vista ou as condições de iluminação variam dentro do reconhecimento facial e sua importância não pode ser exagerada (CHELLAPPA; WILSON; SIROHEY, 1995). Estes métodos utilizam modelos deformáveis e análise matemática extensa para extrair informação das componentes básicas da face e, em seguida, converter essa informação num vetor característico.

Três tipos de métodos de extração de características podem ser destacados: (1) métodos genéricos baseados em bordas, linhas e curvas, (2) métodos apresentam-baseado em modelos que são usados para detectar características faciais tais como olhos, (3) estruturais métodos combinados que levem em consideração as restrições sobre as características geométricas (CHELLAPPA; WILSON; SIROHEY, 1995).

Estes métodos possuem um dificultador, quando as aparências dos indivíduos alterar-se significativamente, por exemplo, o indivíduo fechou os olhos, pôs óculos, abriu a boca. Comparado com o método anterior, nota-se que os métodos baseados na aparência são muito mais robustos em termos de manuseamento variações na intensidade da imagem, que das características (JENKINS, 2008).

O PCA é um método de autovetores também conhecido como transformada de Karhunen-Loève (KLT), projetado para modelar a variação linear de dados de alta dimensão. Seu objetivo é encontrar um conjunto de funções de base mutuamente ortogonais que capturam a direção da maior variação de dados (MINGOTI, 2005), estas bases são formadas por autovetores da matriz de co variância de um certo vetor. Por causa da elevada dimensão da imagem da face, apenas os autovetores principais são utilizados, em outras palavras a energia localizada no subespaço constituído pelo primeiros autovetores. Portanto, uma grande compressão pode ser atingida deixando de lado autovetores com autovalores que não representem a face X.

Durante a análise de diversos métodos de reconhecimento entendemos que:

1. Transformada de Gabor
2. Imagem de Perfil
3. imagem 3D

#### 4. Métodos Holísticos

Representam os métodos mais significativos (ZHAO; ZHUANG; XU, 2008), (FOWLER, 2009), faremos então uma breve descrição deles.

##### 2.4.1 Técnicas baseadas na Transformada de Gabor

Uma imagem se refere a uma função de intensidade de luz bi-dimensional  $f(x, y)$ , onde  $x$  e  $y$  são coordenadas e o valor de  $f$  representa o nível de cinza associado ao ponto. Portanto, uma imagem digital é uma imagem  $f(x, y)$  discretizada no espaço de intensidade de brilho e pode ser considerada uma matriz cujos elementos são chamados de *pixel*. Com esta abordagem, torna-se possível associar esta função a métodos que utilizem conceitos de função para analisar e extrair características dessa imagem. A Transformada de Fourier (TF) modela sinais ao longo do tempo, levando o domínio do tempo para o domínio da frequência e depois reconstruir esse sinal utilizando sua função inversa, ou seja, decompõe a imagem em dois componentes, seno e cosseno. Na imagem de domínio de Fourier, cada ponto representa uma frequência particular contida na imagem do domínio do tempo. A TF é uma ferramenta importante para o processamento de imagens e é utilizada em uma ampla gama de aplicações, tais como análise de imagens, filtragem, reconstrução e compressão (FERIS; KRUEGER; JR., 2004).

Nesse caso podemos considerar a transformada de Fourier uma função  $F : \mathbb{R} \rightarrow \mathbb{C}$  definida por:

$$F(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\omega) e^{j\omega t} dt \quad (2.1)$$

Definimos portanto, que a transformada de Fourier como sendo a função  $\hat{f}$  que associa a cada função absolutamente integrável  $f : \mathbb{R} \rightarrow \mathbb{R}$  a função  $F : \mathbb{R} \rightarrow \mathbb{C}$

Qualquer imagem pode ser representada por uma TF bi-dimensional ao qual pode ser considerada como uma imagem, com uma parte real e uma parte complexa. Seja então uma imagem bi-dimensional, podemos representar qualquer ponto na imagem por uma intensidade luminosa  $I(x, y)$  logo:

$$y(\omega_1, \omega_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} I(x, y) e^{-j(\omega_1 x + \omega_2 y)} dx dy \quad (2.2)$$

Conhecendo  $y(\omega_1, \omega_2)$  podemos recuperar a informação de intensidade através da transformada inversa de Fourier ou seja:

$$I(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y(\omega_1, \omega_2) e^{j(\omega_1 x + \omega_2 y)} d\omega_1 d\omega_2 \quad (2.3)$$

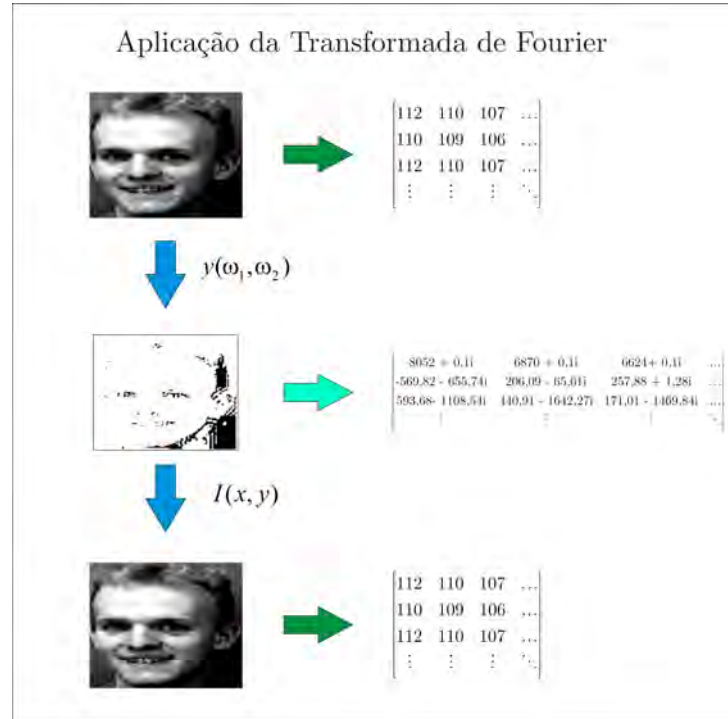


Figura 2.7: Resultado de uma imagem sendo aplicada a transformada de Fourier e depois sua inversa

Embora útil a TF não permite uma análise local em termos de frequência, consequentemente eventos podem ocorrer em intervalos de tempo distintos. Para sanar este problema há uma variante da transformada de Fourier denominada transformada janelada de Fourier, que permite analisar o sinal ao longo do tempo. Uma alternativa na análise de imagens e sinais é a chamada transformada de Wavelets, que permite variação no tamanho da janela juntamente com a frequência. Podemos defini-la como:

$$F(a, b) = \int f(t) \Psi_{(a,b)}(t) dt \quad (2.4)$$

onde  $a > b$  e  $b$  varia continuamente e  $\Psi_{(a,b)}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right)$  sendo chamada de transformada de Gabor e dada por:  $\psi(x, y) = e^{(-\frac{1}{2}(x^2+y^2) \sin(x))}$

Na utilização da transformada de Gabor as imagens são representadas com uma com-

binacão linear destas wavelets em que os parâmetros de cada uma das funções de Gabor são otimizados para refletir a estrutura local de uma imagem, a título de ilustração vemos estas transformações na figura 2.7. Esta abordagem trás inúmeras vantagens, Feris, Krueger e Jr. (2004) destaca as seguintes:

1. Pela sua própria natureza, as wavelets Gabor são invariantes em certos graus de deformações e alterações no brilho da imagem;
2. Filtros de Gabor são bons detectores de características e os parâmetros otimizados de cada wavelet de Gabor refletem a estrutura subjacente da imagem;
3. Os pesos das wavelet de Gabor são linearmente relacionados com as respostas do filtro Gabor;
4. A precisão da representação pode ser alterado a qualquer grau desejado desde uma representação grosseira a uma quase realista, simplesmente variando o número de wavelets utilizadas. Dependendo do poder computacional disponível e a precisão necessária ao experimento, do número de wavelets podem ser dinamicamente variada.

Outro método mais recente (WEEKS, 2007), é a utilização de *jets*, que é um vetor em que cada posição é determinada através de uma transformada de Gabor bidimensional com a janela Gaussiana (Modulada por uma exponencial complexa) em um determinado local da imagem. A figura 2.8 representa estes Jets.

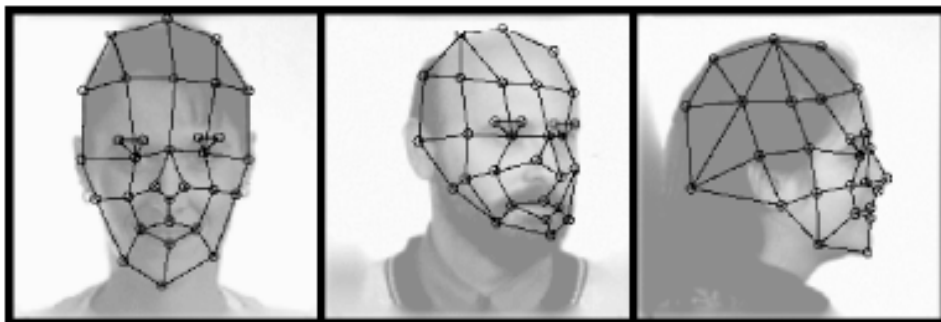


Figura 2.8: Imagens de face sendo aplicados jets em diferentes poses

### 2.4.2 Imagens de Perfil

O método baseia-se nas informações da face que estão em posição de perfil, figura 2.9 analisando o formato do rosto e o formato do cabelo. É formado um vetor característico a partir dessa imagem em escala cinza binarizada com um limiar automático ressaltando a região da face que é normalizada usando dilatação e erosão. Sendo bastante intuitiva



a vantagem desse tipo de reconhecimento está no uso de informações que não ficam disponíveis em imagens frontais bidimensionais, como tamanho do nariz e do papo, além do fato de não ser difícil extrair estes pontos quando o fundo é uniforme. Outra vantagem é a invariância à translação e rotação no plano, a desvantagem está nos problemas ocasionados com alterações devido a expressões faciais e a rotações em profundidade (CAMPOS, 2001).

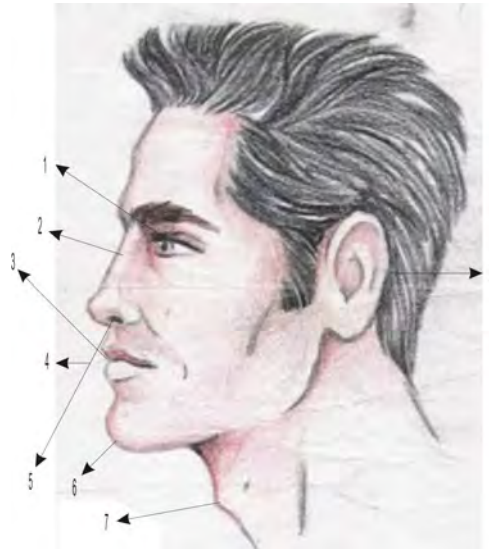


Figura 2.9: Alguns elementos utilizados no reconhecimento por perfil, 1 - Pontos extremos da face; 2 - Posição e Contorno da narina; 3- Centro da Boca; 4 - Cantos dos lábios, 5 - Posição das fossas nasais; 6 - Contorno do queixo; 7 - Saliência do pombo de adão; 8 -Características do pavilhão auricular (CAMPOS, 2001)

Alguns métodos automáticos de extração desses pontos baseiam-se em contornos. O uso de energia de curvatura é uma abordagem, por exemplo, baseada em perfil.

### 2.4.3 Técnicas baseadas em Imagens 3D

As avaliações, como o teste FERET (ZHAO et al., 2003), (ZHOU; CHELLAPPA; ZHAO, 2006), e FACE RECOGNITION VENDOR TEST 2006, sublinham que o atual estado da arte do reconhecimento facial 2D ainda não é suficiente para uso em aplicações biométricas. Embora a precisão dos algoritmos atuais mostre-se bastante elevada em condições restritas, suas performances não são as melhores quando estas condições são perturbadas. O reconhecimento em 3D pode ser um caminho para solucionar este problema, pois lida muito bem com problemas com pose, expressão facial, envelhecimento do rosto e iluminação, figura 2.10. Mas esta técnica foi deixada de lado pela comunidade devido ao custo e a expressão de sensores 3D.

Com o passar dos anos, estes sensores tem se tornado baratos, rápidos e precisos, permi-



Figura 2.10: Imagem inicial e extração 3D, (ETHEVE; RICCIO; DUGELAY, 2006)

tindo que este modelo de reconhecimento, figura 2.11, seja analisado com mais interesse pela comunidade acadêmica. Ao contrário do 2D o 3D não possui base de dados publicamente disponível e os que existem possuem baixa qualidade, na tabela 2.3 vemos uma relação destas, o que torna difícil o teste de novas metodologias, usando um exemplo o  $3D_{RMA}$ .

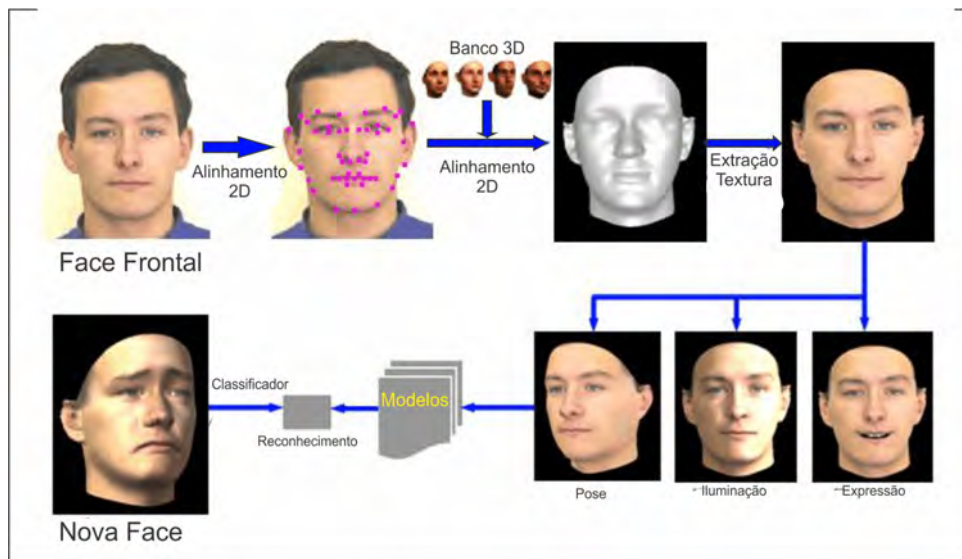


Figura 2.11: Modelo básico para reconhecimento 3D, (ETHEVE; RICCIO; DUGELAY, 2006)

No entanto, o progresso contínuo das tecnologias 3D estão influenciando cada vez mais a qualidade das técnicas de reconhecimento e novos algoritmos tem surgido (ETHEVE; RICCIO; DUGELAY, 2006), (ZHANG, 2000), (TAN et al., 2006) e tornando este modelo eficaz após uma triangulação eficaz. O presente estado da arte revela também, que a maioria dos trabalhos de investigação no domínio do reconhecimento facial e todos os principais sistemas de reconhecimento de face comercial usam imagem de intensidade da face. Este paradigma é conhecido como reconhecimento facial 2D, por outro lado o reconhecimento 3D leva em consideração a forma da cabeça, e em especial informações sobre profundidade,

Banco de Dados	Amostra	Amostras por Pose	Total	Expressões	Pose	Oclusão
FRGC v.2	466	1-22	4007	Raiva, felicidade, surpresa, desgosto, ofegante	NA	NA
BU-3DFE	100	25	2500	Raiva, Felicidade, Surpresa, Desgosto, Ofegante, Felicidade, medo ( em 4 níveis)	NA	NA
ND2006	888	1-63	13450	Felicidade, tristeza, desgosto, outros	NA	NA
York	350	15	5250	Felicidade, raiva, olhos fechados, sobrancelhas levantadas	Sem controle do movimento da cabeça	NA
CASIA	123	6	1845	Sorrindo, raiva, rindo, surpresa, olhos fechados	NA	NA
GavabDB	61	9	549	Sorrindo, testa acentuada, gestos aleatórios	esquerda, direita, abaixo e acima	NA
3DRMA	120	6	720	NA	ligeiramente a esquerda, direita, para baixo e para cima	NA
Bosphorus	81	31-53	3396	34 expressões diferentes	13 diferentes tipos de pose	4 tipos para cada elemento (cabeça, cabelo e óculos)

Tabela 2.3: Tabela dos mais conhecidos bancos de dados de faces 3D - (SAVRAN et al., 2008)

que é perdido na análise 2D. Em ambos os casos a relação é de um-para-muitos e pode ser multi-modal, ou seja, 2D e 3D. As técnicas mais utilizadas atualmente são:

- Reconhecimento usando faixas de Imagens;
- Reconhecimento baseado em características de profundidade e curvatura;
- Reconhecimento facial usando Eigenface 3D.

Vamos tecer algumas considerações sobre cada uma.

### 2.4.3.1 Reconhecimento usando faixas de Imagens

Uma das principais abordagens para o reconhecimento de faces 3D trabalhou no intervalo de dados obtidos diretamente através de scanners a laser, sendo as imagens adquiridas por meio de uma chamada abordagem de luz estruturada (ETHEVE; RICCIO; DUGELAY, 2006). A desvantagem dessa abordagem é a falta de dados devido a oclusão ou regiões mal-refletidas. Este problema pode ser evitado utilizando-se dois sensores, figura 2.12 ao invés de um e realizando dois passos:



Figura 2.12: Handheld 3D scanner, de mão, utiliza 3 scanners simultâneos

1. Calibração dos sensores (matriz de projeção, oclusão da câmera, etc)
2. Fundir as imagens, ou seja, para cada ponto de dados original em 3D as coordenadas na imagem mesclado são calculados.

O modelo utiliza o eigenfaces para redução da dimensionabilidade dos dados, treinamento e testes utilizando a distância euclidiana, pode também utilizar o Modelo de Markov

Processing	Eigenfaces	HMM
Sem Processamento	97,50%	90,83%
Suavização	98,33%	90,00%
Suavização	98,33%	76,67%
Rotação	100%	89,17%

Tabela 2.4: Resultado comparativa entre Eigenfaces e HMM

(HMM - Hidden Markov Models) (RABINER, 1989), (HUANG; JACK; ARIKI, 1990), (EPHRAIM; MERHAV, 2002) podemos analisar o desempenho na tabela 2.4

### 2.4.3.2 Reconhecimento baseado em características de profundidade e curvatura

Nesta abordagem, utiliza-se recursos que vão além dos usuais no Reconhecimento Facial (posição dos olhos, nariz boca, fronteiras) como limite da mandíbula, bochechas, testa. Então esta técnica baseia-se em um conjunto de dados mais amplo da face possuindo assim várias vantagens, em relação a imagens de intensidade em virtude das informações disponíveis. Ademais a caricatura, sob certos aspectos, é invariável (ETHEVE; RICCIO; DUGELAY, 2006). Definindo assim, um conjunto de funcionalidades de alto nível como:

- parte do nariz (násio)
- base do nariz(base do septo)
- cume do nariz
- cavidade do canto do olho (interior e exterior)
- centro do olho(globo ocular)
- limiar dos olhos
- fronteira do nariz
- partes opostas das bochechas

Sendo cada um dos aspectos acima descritos como um conjunto de relações de profundidade e os valores de curvatura, suas restrições são:

- sinal da gaussiana e curvatura média;
- medida absoluta da região sobre a superfície;

- distância entre os planos simétricos;
- proximidade do olho na superfície
- protusão da superfície circundada;
- configuração local da curvatura.

### 2.4.3.3 Reconhecimento facial usando Eigenface 3D

Aplicado ao banco de dados BD-RMA, figura 2.13 que é um banco disponível na web <[http://www.sic.rma.ac.be/~beumier/DB/3d\\_rma.html](http://www.sic.rma.ac.be/~beumier/DB/3d_rma.html)>, composto de 120 pessoas, sendo 14 mulheres e o restante homens, ambos de diferentes etnias, idade e em diferentes poses, no qual os modelos são representados por meio de nuvens de pontos espalhados. Portanto, centra-se numa malha de pontos por meio de algoritmos iterativos sendo analisado a ponta do nariz como valor mais proeminente, em seguida cria-se uma malha de pontos subdividindo e sintonizando passo a passo. A nuvem de pontos tem diferentes orientações e as malhas resultantes preservam as características da primeira. Ruídos podem resultar num modelo não-face. Após estas etapas utiliza-se o eigenface para treinamento e reconhecimento.



Figura 2.13: Imagem ilustrativa do banco de Dados RMA, (ETHEVE; RICCIO; DUGELAY, 2006)

### 2.4.4 Métodos Holísticos

Este tipo de método considera todos os *pixels* da imagem, ou de regiões características da face. Assim, a dimensionalidade dos dados é igual ao número de pixels da imagem considerada. Então uma imagem de  $128 \times 128$ , organizada para o treinamento, gera uma matriz do tipo  $16384 \times 16384$ , que torna o esforço computacional muito grande e

inviabilizaria o reconhecimento facial. Dai a razão para a utilização de métodos de redução de dimensionalidade que criam uma melhor precisão do classificador.

Sendo uma etapa importante no reconhecimento, a redução da dimensionalidade está conectada fortemente à capacidade dos algoritmos em utilizar espaços de características (dimensão dos dados) de tamanhos intratáveis para tamanhos manuseáveis.

Os métodos de redução de dimensionalidade estudam técnicas que permitem reduzir espaços de alta dimensão, para espaços de dimensão menores. Dentre as diversas técnicas, tem-se a Análise de Componentes Principais - PCA (JOLLIFFE, 2002),(HAIR et al., 2007), (MINGOTI, 2005) e Análise de Discriminante linear (ETEMAD; CHELLAPPA, 1997), (HILBE, 2009). O método de reconhecimento com PCA é o mais popular (DELAC; GRGIC, 2004),(EL-BAKRY; HAMADA, 2009),(FOWLER, 2009), sendo aplicado no pré-processamento das imagens para melhorar seu desempenho. Os métodos holísticos proporcionam resultados mais eficientes e sua principal vantagem é que pequenas variações nos dados não prejudicam demasiadamente o reconhecimento, enquanto sua maior desvantagem está nos problemas na variação de iluminação.

#### 2.4.4.1 *Análise discriminante Linear*

A análise de discriminante linear - LDA é uma técnica estatística para classificação de elementos em grupos previamente estabelecidos. O método é detalhadamente descrito em (ZHAO; CHELLAPPA; PHILLIPS, 1999), mas basicamente baseia-se em combinações lineares das variáveis independentes que dão uma espécie de score do objeto observado, obtendo-se em seguida uma probabilidade daquele objeto pertencer a um dos grupos. Permitindo assim, testar a integridade dos agrupamentos além de determinar a influência de variáveis individuais na formação dos grupos (KLECKA, 1980). A análise de discriminante linear também é conhecida como Discriminante linear de Fisher (FLDA) (BELHUMEUR; HESPANHA; KRIEGMAN, 1997), tornando-se uma alternativa ao uso do PCA.

Neste caso, o conjunto de imagens de treino da face é construído por várias classes. Assim cada classe representa uma identidade de uma pessoa e, com base num conjunto de classes previamente conhecidas, o problema do reconhecimento pode ser formulado de modo a determinar a que classe pertence uma determinada pessoa desconhecida.

Percebendo que a quantidade de imagens de treinamento é quase sempre menor que o total de pixels de cada imagem, a técnica LDA fica comprometida. Desse modo, é feita a projeção do conjunto de imagens de treinamento para um espaço de dimensão menor,  $N - c$  sendo  $N$  o número de imagens de treino e  $c$  o número de classes, usando o PCA.

A partir dessa redução de dimensionalidade, aplica-se o LDA. Os vetores próprios obtidos são usados para projetar as imagens de rosto num espaço de dimensão  $c - 1$ . O reconhecimento de uma imagem de teste se dá com a projeção no espaço dos vetores próprios criados e utilizando uma medida de semelhança, para efetuar a comparação com as outras imagens da face projetadas no mesmo espaço, utilizando um classificador específico ou a combinação de dois ou mais. Matematicamente, defini-se duas medidas:

1. Matriz de espalhamento intra-classes, dada por:

$$S_w = \sum_{j=1}^c \sum_{i=1}^{|T_j|} (x_i^j - \mu_j) \cdot (x_i^j - \mu_j)^t \quad (2.5)$$

em que  $x_i^j$  é o  $i$ -ésimo exemplo da classe  $j$ ,  $\mu_j$  é a média da classe  $j$ ,  $c$  é o número de classes, e  $|T_j|$  o número de exemplos na classe  $j$ ;

2. matriz de espalhamento inter-classes, dada por:

$$S_b = \sum_{j=1}^c (\mu_j - \mu) (\mu_j - \mu)^T \quad (2.6)$$

em que  $\mu$  representa a média de todas as classes.

A maximização da medida inter-classes e a minimização da intra-classe são obtidas a taxa de  $\frac{\det(S_b)}{\det(S_w)}$ . O espaço de projeção é então encontrado resolvendo a equação:

$$S_b W = \lambda S_w W \quad (2.7)$$

onde  $W$  é a matriz com autovetores generalizados associados com  $\lambda$ , que é a matriz diagonal com autovalores.

#### 2.4.4.2 Análise Componentes Principais

O PCA é uma das mais antigas e poderosas técnicas de estatística para a análise de dados (JOLLIFFE, 2002), que pode ser utilizada para redução do número de variáveis sendo portanto, um método de decorrelação de dados que possui algumas vantagens como a eliminação de informações redundantes de dados. Considerado um dos métodos da estatística multivariada mais simples ele foi introduzido por Karl Pearson em 1901 (PEARSON, 1901) e por (HOTELLING, 1933). Seu objetivo principal é o de explicar a estrutura de variância e co variância de um vetor aleatório, composto de  $p$ -variáveis aleatórias através da construção de combinações lineares das variáveis originais (MINGOTI, 2005). Neste



capítulo discutiremos este método, tendo como objetivo estabelecer uma fundamentação teórica para a *eigenfaces*. Por ser uma técnica muito utilizada no reconhecimento de face justifica-se aqui, uma abordagem matemática que torne evidente o potencial do método. Partindo dessa premissa, estabeleceremos o método baseado na literatura consagrada para em seguida aplicá-lo ao estudo de imagens.

O método pode ser descrito analisando que em um determinado dado se tivermos  $p$ -variáveis iniciais, é possível obtermos  $p$  componentes principais. Mas esta interpretação não nos ajudaria no estudo de dados de alta dimensionalidade, daí a idéia de reduzir esta dimensionalidade, eliminando partes repetidas ou que não acrescentem grande relevância no conjunto do dado, ou seja, a informação contida nas  $p$ -variáveis é transformada em  $k$  ( $k < p$ ) componentes principais não correlacionadas. Para obtermos estes componentes principais, com dimensionalidade menor, mas com grande relevância na informação, basicamente, realizaremos uma projeção dos dados de uma dimensão  $X$  para outra  $Y$ , onde na dimensão  $X$  as variáveis são correlacionadas, e na  $Y$  essas novas variáveis não são correlacionadas (ALMEIDA, 2006).

Seja, então,  $X = (X_1, X_2, \dots, X_p)'$  um vetor aleatório com vetor de médias  $\mu = (\mu_1, \mu_2, \dots, \mu_p)'$  e a matriz de co variâncias  $\Sigma_{pxp}$ . Sejam  $\lambda_1 \geq \lambda_2 \geq \dots \lambda_p$  os autovetores da matriz  $\Sigma_{pxp}$ , com os respectivos autovetores normalizados  $e_1, e_2, \dots, e_p$ , isto é, os autovetores  $e_i$  satisfazem as seguintes condições:

1.  $e_i' e_j = 0$  para todo  $i \neq j$ ;
2.  $e_i' e_i = 1$  para todo  $i = 1, 2, \dots, p$ ;
3.  $\sum_{pxp} e_i = \lambda_i e_i, I = 1, 2, \dots, p$

sendo o autovetor  $e_i$  denotado por  $e_i = (e_{i1}, e_{i2}, \dots, e_{ip})'$ . Considere o vetor aleatório  $Y = O'X$ , onde  $O_{pxp}$  é a matriz ortogonal de dimensão  $pxp$ , constituída dos autovetores normalizados da matriz  $\Sigma_{pxp}$ , isto é:

$$O_{pxp} = \begin{bmatrix} e_{11} & e_{21} & \dots & e_{p1} \\ e_{12} & e_{22} & \dots & e_{p2} \\ \vdots & \vdots & \vdots & \vdots \\ e_{1p} & e_{2p} & \dots & e_{pp} \end{bmatrix}$$

$$= \begin{bmatrix} e_1, e_2, \dots, e_p \end{bmatrix}$$

Percebemos que o vetor  $Y$  é composto de  $p$  combinações lineares das variáveis aleatórias

do vetor  $X$ , tem vetor de médias igual a  $O'\mu$  e matriz de co variâncias  $\Lambda_{p \times p}$ , matriz diagonal, cujos elementos são iguais a

$a_{ij} = \lambda_i, i = 1, 2, \dots, p$  isto é:

$$\Lambda_{p \times p} = \begin{bmatrix} \Lambda_1 & 0 & \dots & 0 \\ 0 & \Lambda_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \Lambda_p \end{bmatrix}$$

O vetor  $Y$  fica então constituído de variáveis não correlacionadas. Para (MINGOTI, 2005) surge nesse ponto a idéia de representar a estrutura da co variância do vetor  $X$ , para obter uma redução no espaço de variáveis tais que  $k < p$ , como combinações lineares em  $Y$ . Portanto, ao invés de se utilizar as  $p$  combinações utilizamos apenas as  $k$  combinações lineares principais.

**Definição 2.4.1.** *A  $j$ -ésima componente principal da matriz  $\Sigma_{p \times p}, j = 1, 2, \dots, p$  é definida como:*

$$Y_j = e_j X = e_{j1} X_1 + e_{j2} X_2 + \dots + e_{jp} X_p \quad (2.8)$$

A esperança e variância da componente  $Y_j$  são respectivamente, iguais a:

$$E[Y_j] = e_j \mu = e_{j1} \mu_1 + e_{j2} \mu_2 + \dots + e_{jp} \mu_p \quad (2.9)$$

$$Var[Y_j] = e_j \Sigma_{p \times p} e_j = \lambda_j \quad (2.10)$$

Sendo  $Cov[Y_j, Y_k] = 0, j \neq k$ . Cada autovalor  $\lambda_j$  representa a variância de uma componente principal  $Y_j$ . Como os autovalores estão ordenados de forma decrescente, a primeira componente é a de maior variabilidade e a  $p$ -ésima é a de menor (MINGOTI, 2005).

A partir dos pressupostos do PCA, foi proposto por Turk e Pentland em 1991 (TURK; PENTLAND, 1991b) motivado pelos trabalhos de Sirovich e Kirby (SIROVICH; KIRBY, 1987) o método conhecido com *Eigenfaces*. Nesse, cada imagem de face é inicialmente representada como um vetor unidimensional, no qual cada elemento dessa matriz representa um pixel. A idéia principal deste método consiste em reduzir a dimensão destes dados e, dessa forma executar o reconhecimento num espaço de dimensão menor. Neste método, a identificação de uma face pode ser realizada procurando na base de dados uma face com vetor de peso mais parecido ao da imagem de teste, ou seja, com a menor distância que pode ser medida usando a distância euclidiana ou outra qualquer.

## 2.5 *Considerações finais*

Neste capítulo mostramos que as pesquisas de reconhecimento facial estão se desenvolvendo de forma efervecente por diversos pesquisadores. Discutimos os principais avanços na área biométrica e pudemos visualizar o quão importantes são as questões levantadas como variância de escala, posição, iluminação, expressões faciais e ruídos independente do método são pontos cruciais para todos.

Várias frentes de trabalho convergiram para uma abordagem holística e continuam recebendo a atenção de várias pesquisas, isso se deve muito ao fato desse método se aproximar de como os seres humanos reconhecem e, para estas pesquisas esse caminho tornará esse modelo mais simples de ser interpretado pela máquina (JENKINS, 2008). Contrapondo a isso, é possível imaginar que a máquina necessite de um método não tão humano para efetuar algo aparentemente trivial aos humanos.

Dessa forma nossa atenção recai sobre este tipo de método e em particular ao *Eigenface*, que utiliza o PCA como redutor de dimensionalidade, pois é de simples implementação e funcionará como ponto de partida para outros trabalhos. Procuraremos nos capítulos seguintes estabelecer os elementos matemáticos e conceituais que permitem a implementação do modelo.

## Modelo Proposto

### 3.1 Introdução

Este capítulo apresenta o modelo que utilizamos para efetuar o reconhecimento facial, considerando uma base matemática, clara e rigorosa, centrada nas principais definições e propriedades. Estes princípios matemáticos são vitais para a compreensão dos algoritmos propostos e circulam basicamente pela estatística multivariada e álgebra Linear. O objetivo maior deste capítulo é a apresentação do modelo computacional desenvolvido por Turk e Pentland (1990) e aplicado neste trabalho. Muito conceituado na área de visão computacional, este modelo permitirá uma rápida implementação à avaliação dos seus resultados no sentido da velocidade e robustês. A modelagem do problema foi realizada e será apresentada na última sessão deste capítulo com base neste modelo. Partindo desta perspectiva, estabeleceremos, a partir da literatura clássica, os conceitos de imagem, autovetores e autovalores, matriz de co-variância que estão divididos em quatro blocos: no primeiro a definição de imagem, que mostramos como a representaremos matematicamente em seguida definições aplicadas ao PCA, a qual dividimos em dois sub tópicos (Definições de Estatística Multivariada e Definições de álgebra Linear), logo depois descrevemos o modelo do *Eigenface* e finalmente, apresentamos nosso modelo do problema. Fixaremos também a terminologia que permite a compreensão imediata dos conceitos aqui tratados e finalizamos o capítulo trazendo as considerações finais.

### 3.2 Definição de Imagem Digital

Antes de definirmos uma imagem digital, faremos uma pequena introdução a um conceito importante para o modelo, a definição de espaço vetorial.

**Definição 3.2.1.** *Seja  $V$  um conjunto não-vazio qualquer de objetos no qual estão definidas duas operações, a adição e a multiplicação por escalares. Por adição nós entenderemos uma regra que associa cada par de objetos  $u$  e  $v$  em  $V$  um objeto  $u + v$ , chamado a soma de  $u$  com  $v$ ; por multiplicação por escalar nós entendemos uma regra que associa a cada escalar  $k$  e cada objeto  $v$  em  $V$  um objeto  $kv$ , chamado o múltiplo de  $v$  por  $k$ . Se os seguintes axiomas são satisfeitos por todos os objetos  $u, v, w$  em  $V$  e quaisquer escalares  $k$  e  $l$ , então nós dizemos que  $V$  é um espaço vetorial e que os objetos de  $V$  são vetores.*

1. Se  $u$  e  $v$  são objetos de  $V$  então  $u + v$  é um objeto em  $V$

2.  $u + v = v + u$
3.  $u + (v + w) = (u + v) + w$
4. Existe um objeto  $0$  em  $V$ , chamado de vetor nulo de  $V$ , tal que  $0 + u = u + 0$  para cada  $u$  em  $V$
5. Para cada  $u$  em  $V$ , existe um objeto  $-u$ , chamado um negativo de  $u$ , tal que  $u + (-u) = (-u) + u = 0$ .
6. Se  $k$  é qualquer escalar e  $v$  é um objeto de  $V$ , então  $kv$  é um objeto em  $V$ .
7.  $l(u + v) = lu + lv$
8.  $(k + l)v = kv + lv$
9.  $k(lu) = kl(u)$
10.  $1u = u$

Esta definição não especifica nem a natureza dos vetores, nem suas operações. Qualquer tipo de objeto pode ser um vetor e as operações de adição e multiplicação por escalar podem não guardar semelhança, ou não ter relação alguma com as operações usuais em  $\mathbb{R}^n$ . A única exigência é que os dez axiomas de espaço vetorial sejam satisfeitos. Com isso, vamos agora integrar a ideia de espaço vetorial com o estudo de imagens.

Uma imagem digital pode ser definida como uma matriz quadrada ou não, em que cada elemento da matriz é denominado *pixel* do inglês *picture elements*. Cada *pixel* possui três coordenadas:  $x$  e  $y$  definem a sua posição na imagem e um valor  $i$  que representa o nível de cinza ou grau de brilho. Na maioria dos sistemas, um valor  $i$  igual a zero representa preto absoluto enquanto o valor mais alto representa o branco absoluto.

Mais formalmente, seja uma matriz  $M \times N$  cujos elementos pertencem a um espaço vetorial ou seja, uma aplicação  $i : R \rightarrow V$ , em que  $R$  é um subconjunto do plano e  $V$  é um espaço vetorial. Para cada *pixel*  $(x, y) \in R$  da imagem,  $i(x, y)$  contém informações da estrutura do *pixel* e é chamado de *valor do pixel*, figura 3.1.

Em geral:

- $R$  é um reticulado uniforme do plano;
- $V$  é um espaço de cor;
- $i$  associa cada *pixel* a sua intensidade de cor.

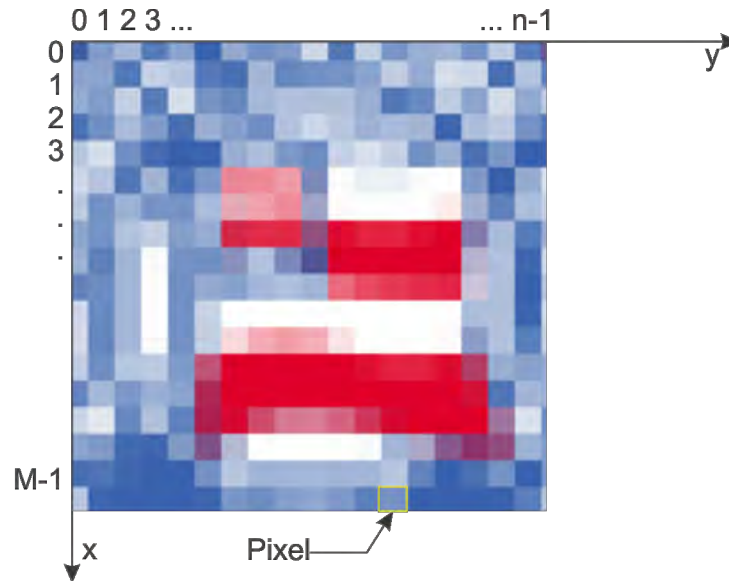


Figura 3.1: Imagem utilizando a convenção definida

A dimensão  $M \times N$  é denominada resolução geométrica da imagem. O número de *bits* utilizado para armazenar o valor do *pixel* é chamado de resolução de cor da imagem. Um espaço de cor unidimensional é também chamado de espaço monocromático. Uma imagem cujos *pixels* assumem valores em um espaço de cor monocromático é chamada imagem monocromática (GOMES; VELHO, 1990), (PARENTE, 1993), a cor associada ao *pixel* é dada por um número que representa a luminosidade de cada uma das cores primárias, vermelho, verde e azul (RGB, iniciais em inglês), representadas pelo pixel (TRIGO, 1998).

### 3.3 Definições aplicadas ao PCA

A seguir serão apresentadas algumas definições muito comuns no trato de dados multivariados e imagens.

#### 3.3.1 Definições de Estatística Multivariada

**Definição 3.3.1.** *Seja  $X$  um vetor aleatório contendo  $p$  componentes, onde cada componente é uma variável aleatória, isto é,  $X_i$  é uma variável aleatória,  $\forall i = 1, 2, \dots, p$ .*

Então,  $X$  é chamado de vetor aleatório e é denotado por:

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} \quad (3.1)$$

O vetor transposto do vetor aleatório  $X$  é denotado por  $X' = (X_1, X_2, \dots, X_p)$ .

**Definição 3.3.2** (Vetor de Médias). *Seja  $X$  um vetor aleatório. O vetor  $\mu = E(X)$  é chamado de vetor de médias do vetor  $X = (X_1, X_2, \dots, X_p)'$ , sendo*

$$\mu = E(X) = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_p) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} \quad (3.2)$$

onde  $\mu_i = E(X_i)$  denota a média aritmética, ou esperança da variável aleatória  $X_i$ ,  $i = 1, 2, \dots, p$ . A média  $\mu_i$  é uma das medidas mais utilizadas para sintetizar a informação de tendência central da distribuição de valores da variável  $X_i$  (MINGOTI, 2005).

**Definição 3.3.3** (Variância). *A variância do  $i$  – ésimo componente do vetor  $X$  é denotada por  $Var(X_i) = \sigma^2_i = \sigma_{ii}$ . O desvio padrão é denotado por  $\sigma_i$  ou  $\sqrt{\sigma_{ii}}$  e fornece a informação sobre a disposição dos valores da variável  $X_i$  em relação a  $\mu_i$ , isto é, indica se os valores de  $X_i$  estão próximos ou distantes da média  $\mu_i$ . Assim, valores grandes de  $\sigma_i$  indicam uma maior dispersão de valores em relação à média da distribuição.*

**Definição 3.3.4** (Covariância). *Seja  $(X, Y)$  uma variável aleatória bidimensional. A covariância entre os valores da  $i$  – ésima e  $j$  – ésima variáveis do vetor  $X$  é definida por:*

$$Cov(X_i, X_j) = \sigma_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)] \quad (3.3)$$

Quando  $i = j$ , a expressão (3.3) torna-se a variância da variável  $X_i$ ,  $i = 1, 2, \dots, p$ . Vê-se da definição que a covariância é a esperança do produto dos desvios dos valores de cada uma das duas variáveis em relação às suas médias (DANTAS, 2004). Para (MALHOTRA, 2004) a finalidade mais comum da covariância é remover variações estranhas da variável dependente, porque os efeitos dos fatores têm a maior importância. Outro ponto importante a se observar é que a  $Cov(X, Y)$  é o produto escalar dos vetores  $X - \mu_i$  e  $Y - \mu_i$ . Essa observação esclarece dois fatos: o primeiro o desvio-padrão, que é a distância de  $X$  até sua média, e da variância, que é o quadrado dessa distância (MYNBAEV, 2004). Quando se tem muitas variáveis, como é o caso de uma imagem, o procedimento mais

comum é guardar os valores de  $\mu_{ij}$  numa matriz chamada de matriz de co variâncias, como definimos a seguir:

**Definição 3.3.5.** *Matriz de co variância A matriz de variâncias e co variâncias do vetor aleatório  $X$  é denotada por:*

$$\text{Cov}(X) = V(X) = \text{Var}(X) = \Sigma_{p \times p} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{bmatrix} \quad (3.4)$$

**Propriedade 3.3.6.** *A matriz de co variâncias é uma matriz simétrica. De fato, basta observarmos que:*

$$\sigma_{ij}^2 = \int_x (x_i - \mu_i)(x_j - \mu_j)p(x)dx = \int_x (x_j - \mu_j)(x_i - \mu_i)p(x)dx = \sigma_{ji}^2 \quad (3.5)$$

**Propriedade 3.3.7.** *A matriz de co variâncias é positiva definida, ou seja  $a'\Sigma a \geq 0$  para todo vetor de constantes  $a_{p \times 1} \in \mathbb{R}^p$ . Esta condição implica que os autovetores da matriz  $\Sigma_{p \times p}$  denotados por  $\lambda_1, \lambda_2, \dots, \lambda_p$ , são não negativos, ou seja  $\lambda_i \geq 0, \forall i = 1, 2, \dots, p$*

### 3.3.2 Definições de Álgebra Linear

Definiremos aqui elementos importantes na construção e manipulação do espaço de face os chamados autovetores e autovalores

**Definição 3.3.8.** *Autovetores e Autovalores Dada uma matriz quadrada  $\mathbf{A}$  de ordem  $n$ , estaremos entendendo por autovalor e autovetor de  $\mathbf{A}$  os autovetores e autovalores de uma transformação linear  $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , associada à matriz  $\mathbf{A}$  em relação à base canônica, isto é,  $T_A(v) : Av$  na forma coluna. Assim, um autovalor  $\lambda \in \mathbb{R}$  de  $\mathbf{A}$ , e um autovetor  $v \in \mathbb{R}^n$ , são soluções da equação  $Av = \lambda v, v \neq 0$ .*

$\lambda$  é chamado de autovalor associado a matriz  $\mathbf{A}$  e  $v$  é chamado de autovetor.

Podemos encontrar os autovalores de uma matriz  $A$  de tamanho  $n \times n$  da seguinte forma:

$$Ax - \lambda x = 0 \quad (3.6)$$

$$x(A - \lambda I) = 0 \quad (3.7)$$

Para  $\lambda$  ser um autovalor, precisa haver uma solução não nula desta equação. Isto é  $x \neq 0$ ,



tais que  $x(A - \lambda I) = 0$ . Neste caso

$$\det(A - \lambda I) = 0 \quad (3.8)$$

Impondo esta condição determinamos primeiramente os autovalores de  $\lambda$  que satisfazem a equação e depois os autovetores a eles associados.

$$P(\lambda) = \det(A - \lambda I) = \begin{bmatrix} a_{11} - \lambda & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} - \lambda & \dots & a_{nn} - \lambda \end{bmatrix} \quad (3.9)$$

é um polinômio em  $\lambda$  de grau  $n$ .  $P(\lambda) = (a_{11} - \lambda) + \dots + (a_{nn} - \lambda)$  termos de grau  $< n$ , e os autovalores procurados são as raízes deste polinômio.  $P(\lambda)$  é chamado *polinômio característico* da matriz  $\mathbf{A}$ .

### 3.4 Eigenface

Basicamente a idéia do método visa representar um rosto como uma decomposição linear de imagens formando uma base (chamadas *eigenfaces*). O reconhecimento consistirá da inicialização dessa base durante a qual uma nova imagem é projetada sobre este espaço e através da distância euclidiana e de um parâmetro de erro definir se a face é conhecida, desconhecida ou apenas parecida com uma das faces da base (falso-positivo).

Suponha uma imagem em tons-de-cinza com dimensões horizontal e vertical  $n$  e  $m$ . Se imaginarmos que a escala varia de zero (preto) a 255 (branco), podemos mapear qualquer imagem com estas dimensões como um vetor de dimensão  $N = n.m$ , dentro do  $\mathbb{R}^n$ . As imagens que serão utilizadas são imagens de faces humanas, frontais com variação de iluminação e pose. Iremos aproveitar o fato de que existem semelhanças faciais entre indivíduos diferentes e, então capturando as principais feições desses, criaremos um espaço de dimensão menor. De fato, estamos interessados em um  $K$ -elipsóide ( $K \ll N$ ) centrado em algum ponto do  $\mathbb{R}^n$ . Este  $K$ -elipsóide é induzido por um subespaço afim, que iremos chamar de espaço das faces, conforme ilustramos na figura 3.2.

Sendo assim seja: primeiramente vamos obter um conjunto  $n$  com imagens em tom de cinza de faces  $I_1, I_2, \dots, I_n$ . Elas devem ser:

- Com olhos no mesmo nível e em escalas iguais

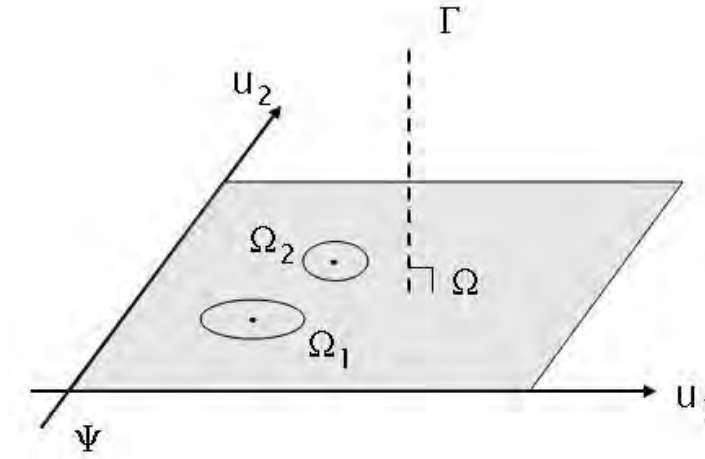


Figura 3.2: Espaço de Faces visualizado em 2D. Os eixos representam dois *Eigenfaces*

- Normalizados de modo que cada pixel tem um valor entre 0 e 255
- Do mesmo tamanho  $N \times N$ .

$$\{I_1, I_2, \dots, I_n\}, \text{ onde } I_k = \begin{bmatrix} p_{1,1}^k & p_{1,1}^k & \cdots & p_{1,N}^k \\ p_{1,1}^k & p_{1,1}^k & \vdots & p_{2,N}^k \\ \vdots & \vdots & \vdots & \vdots \\ p_{1,1}^k & p_{1,1}^k & \cdots & p_{N,N}^k \end{bmatrix} \text{ e } 0 \leq p_{i,j}^k \leq 255 \quad (3.10)$$

Assim que tivermos este conjunto, devemos mudar a representação da imagem da face  $I_k$  de uma matriz  $N \times N$  para uma matriz  $\Gamma_k, m \times n$ , em que  $n$  representa o número de imagens. Este o conjunto de faces representa um vetor de dimensão N. Ou seja:

$$\Gamma_k = \begin{bmatrix} p_{1,1}^k \\ p_{1,2}^k \\ \vdots \\ p_{1,N}^k \\ p_{2,1}^k \\ p_{2,2}^k \\ \vdots \\ p_{2,N}^k \\ \vdots \\ p_{N,1}^k \\ p_{N,2}^k \\ \vdots \\ p_{N,N}^k \end{bmatrix}, \text{ onde } k = 1, \dots, n \text{ e } p_{i,j}^k \in I_k \quad (3.11)$$

Neste momento formamos o conjunto  $S = \{\Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_n\}$  e após esta etapa vamos obter uma imagem média  $\Psi$ , que será dada por:

$$\Psi_{m \times n} = \frac{1}{n} \sum_{i=1}^n \Gamma_{i, m \times 1} \quad (3.12)$$

que resultará em algo parecido com a figura 3.3. A partir dessa imagem encontramos a diferença entre as imagens de entrada e a imagem média.



Figura 3.3: Imagem média, obtida através de um conjunto de imagens

$$\Phi_{i(m \times n)} = \Gamma_{i(m \times n)} - \Psi_{m \times n} \quad (3.13)$$

com

$$A_{m \times n} = \begin{bmatrix} \Phi_{1,1} & \cdots & \Phi_{1,n} \\ \vdots & \ddots & \vdots \\ \Phi_{M,1} & \cdots & \Phi_{m,n} \end{bmatrix} \quad (3.14)$$

Agora devemos tentar encontrar um conjunto de vetores ortonormais que melhor descrevem a distribuição dos nossos dados, ou seja, vamos extrair os *eigenfaces* de uma matriz  $C$  dada a seguir:

$$C_{M \times M} = A_{M \times n} A_{n \times M}^T \quad (3.15)$$

Daí encontraremos autovetores  $u_k$  *eigenvalues* e autovalores  $\lambda_k$  de  $C$ . Consideremos

$$L_{n \times m} = A_{m \times n}^T A_{M \times n} \quad (3.16)$$

Os primeiros  $n$  autovetores de  $C$  podem ser expressos como uma combinação linear entre os autovetores de  $L$  (denotados por  $V$ ) e as imagens contidas em  $A$ , conforme mostrado:

$$U_{M \times n} = A_{M \times n} \cdot V_{n \times n} \quad (3.17)$$



Figura 3.4: Eigenfaces

Criamos, portanto os eigenfaces, figura 3.4, armazenados na matriz  $U$ . Com base nestes vamos estabelecer duas etapas, treinamento e reconhecimento.

### 3.4.1 Treinamento

No estágio de treinamento a face de cada indivíduo  $\Gamma_k$  é projetada no espaço de imagens, gerando uma matriz  $\Omega$ :

$$\Omega_{n \times n} = U_{n \times M}^T (\Gamma_{M \times n} - \Psi_{M \times n}) \quad (3.18)$$

Com essa matriz realizaremos as comparações na hora de executarmos o reconhecimento. Para posterior classificação das imagens precisamos calcular um valor limiar  $\Theta_c$  para representar a maior distância aceitável entre as imagens de teste e as imagens de treinamento:

$$\Theta_c = \frac{1}{2} \max_{j,k}^n \{\Omega_j - \Omega_k\} \quad (3.19)$$

### 3.4.2 Reconhecimento

Consideremos uma imagem de prova  $I_{[m \times n]}$  no qual o algoritmo deve discernir se a imagem representa ou não uma face humana além de, se realmente representar, conseguir detectar se a face apresentada é de alguma pessoa cuja face está no banco de treinamento  $\Omega$  ou se é outra pessoa. Para iniciarmos o processo, projetamos  $I_{[m \times n]}$  no espaço de faces através da seguinte relação:

$$\Omega_{[n \times 1]} = U_{[n \times M]}^T (I_{[M \times 1]} - \Psi_{[M \times 1]}) \quad (3.20)$$

Reconstruímos a imagem de prova guardando-a em

$$I_{f,[m \times 1]} = U_{[m \times n]} \Omega_{n \times 1} + \Psi_{[m \times 1]} \quad (3.21)$$

Esta matriz será usada para que se possa realizar a distinção entre imagens de prova contendo faces humanas e imagens quaisquer. Para isso calculamos a distância (vetorial) entre a imagem de prova e a sua reconstrução:

$$\varepsilon = \|I - I_f\| \quad (3.22)$$

Determinamos também a distância entre a imagem de prova e todas as imagens do banco treinamento. Armazenamos esses resultados em

$$\varepsilon_k = \|\Omega - \Omega_k\|; k = 1, \dots, n \quad (3.23)$$

logo:

1.  $\varepsilon \geq \Theta_c \rightarrow$  a imagem não representa uma face humana;
2.  $\varepsilon < \Theta_c e \varepsilon_k \geq \Theta_c \rightarrow$  a imagem representa uma face humana, mas desconhecida;
3.  $\varepsilon < \Theta_c e \min \varepsilon_k < \Theta_c \rightarrow$  a imagem representa uma face do indivíduo contido no índice k do banco.

o que completa o modelo.

Nossa pesquisa implementa o modelo do *Eigenface* aplicando os filtros considerados mais eficientes na etapa de pré-processamento de tal forma que permita o reconhecimento facial com sucesso. Este modelo pode ser analisado conforme a figura 3.5

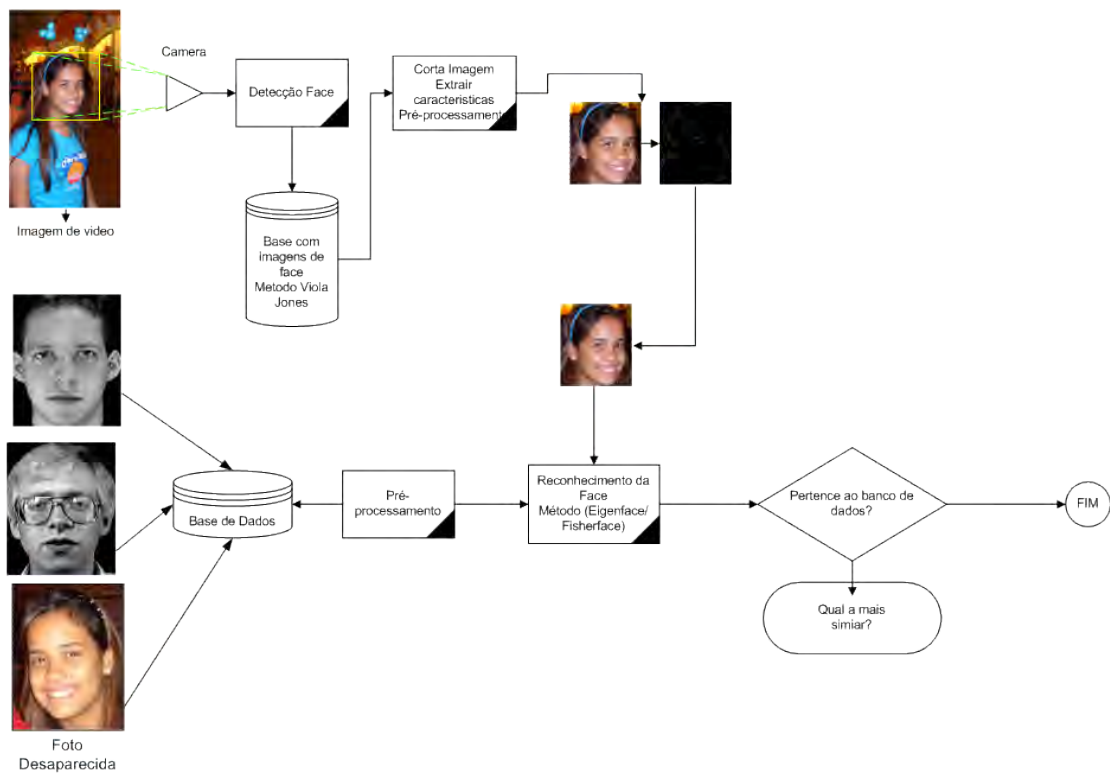


Figura 3.5: Modelo empregado no reconhecimento facial

### 3.5 Considerações Finais

O modelo constitui-se de duas fases: o treinamento e o reconhecimento. Na primeira os dados são gerados a partir de uma ou várias imagens de treinamento que são duplicadas e transformadas até compormos uma base com 20 faces, em escala cinza, dimensionada a  $92 \times 112$  com a face frontal do indivíduo. Criando assim, uma base ou uma memória de faces de 134 pessoas que desejamos localizar posteriormente.

Em seguida, utilizando uma câmera de vídeo ou uma imagem estática captamos uma imagem que contenha a face frontal de um indivíduo, normalizamos esta imagem às mesmas condições das imagens do banco de dados, ou seja, escala cinza, tamanho  $92 \times 112$ , etc e a projetamos no espaço de face. Nesse ínterim uma distância euclidiana é calculada e medimos, por meio de um limiar se a face está ou não presente neste banco de dados e se está a quem pertence.

Neste capítulo, portanto, apresentamos o modelo utilizado. Além disso, apresentamos todos os pressupostos matemáticos que achamos necessários ao entendimento do modelo

e, em seguida, foi discutida, de maneira simples e comentada o modelo de reconhecimento proposto por Turk e Pentland (1990). Assim, abrimos espaço para a implementação desse modelo que apresentaremos no próximo capítulo, bem como os testes realizados e seus resultados.

---

## Implementação

---

### 4.1 Introdução

O propósito deste capítulo é implementar, num ambiente computacional, o modelo para o reconhecimento facial com objetivo de analisar o comportamento desse em relação a sua performance e as taxas de reconhecimento positivo e falso positivo.

Utilizamos a linguagem de programação C#, para implementação do modelo em um computador pessoal com processador intel Core 2 Duo e 2MB de ram, utilizando o S.O. Windows 7 e alguns filtros foram implementados em MATLAB. A implementação utiliza uma biblioteca de domínio público chamada OpenCV (BRADSKI; KEAHLER, 2008) escrita em C e C++ que roda tanto no ambiente Windows como em Linux ou MacOSX que pode utilizar as IDE do Python, Ruby, Matlab, Visual Studio entre outras IDEs. Um dos objetivos principais da OpenCV e auxiliar a construção de aplicações de visão computacional com bastante sofisticação e rapidez. O OpenCV possui mais de 500 funções que abrangem diversas áreas de visão computacional entre elas destacamos as de reconhecimento de padrões que implementa inclusive o *Eigenfaces*.

### 4.2 Base de Dados

Para análise e calibração do software gerado a partir do modelo proposto, utilizamos inicialmente fontes de dados já consagrados na literatura. Neste caso foi utilizado o banco de faces produzido pela Olivetti a ORL. Esta base de dados é gratuita e contém 10 imagens diferentes de 40 indivíduos distintos perfazendo um total de 400 imagens com variação de pose e aspectos faciais em condições de iluminação diferentes mas padronizadas. Alguns indivíduos se apresentam com ou sem acessórios como por exemplo óculos. Todas as imagens possuem um fundo simples (preto ou cinza) com os indivíduos em posição frontal e com rotação de mais ou menos 20 graus em relação a posição central da imagem. Estas imagens estão em escala cinza (256 níveis) com uma resolução de 92x112 *pixels*. A figura 4.1 mostra uma parte dessa base.

Várias pesquisas utilizam esta base de dados com referência (SANTOS, 2005), (SILVA, 2006) para avaliar os seus modelos com 5 faces de treinamento e 5 faces de teste para cada pessoa. Em sua maioria este banco é aumentado utilizando imagens espelhadas, ou seja, um *flip* horizontal na imagem que faz a mesma girar em torno do eixo *y* gerando





Figura 4.1: Algumas faces presentes no banco de faces ORL

mais 400 imagens.

Porém, como nosso estudo centra as ações na localização de pessoas desaparecidas, procuramos criar uma base própria com algumas características que são intrínsecas ao problema, entre elas:

- Imagens com fundo complexos;
- Imagens com algum nível de degradação;
- Imagens em escala RGB;
- Imagens com resoluções diferentes;
- Imagens com defasagem de idade;

Muito dificilmente, encontraremos mais do que 7 imagens do mesmo indivíduo. Para nosso estudo foram adquiridas 2189 imagens diferentes de 134 indivíduos distintos que tentam simular as situações descritas inclusive de variação de pose e aspectos faciais em condições de iluminação diferentes, ou seja, não padronizadas e alguns indivíduos se apresentando com algum tipo de acessório, como óculos. As imagens possuem um fundo complexo, que dificulta a etapa de detecção, em escala RGB, os indivíduos apresentaram diferentes posições frontais com alguma rotação em relação a posição central da imagem. A resolução das imagens é diversa, mas após a detecção da face ela é normalizada para a condição da base ORL, isto é 92x112 *pixels* de dimensão, escala monocromática, etc. A figura 4.2 mostra uma parte dessa base.

Um dos problemas no reconhecimento de pessoas desaparecidas é a diferença de idade entre as fotos de treinamento e de entrada. Testamos o modelo utilizando uma base de dados com imagens de pessoas em diferentes fases da vida, o banco de dados de face iraniano (IFDB) (BASTANFARD et al., 2007) que contem imagens faciais em escala RGB de



Figura 4.2: Algumas imagens da fonte de dados utilizada para o reconhecimento



Figura 4.3: Imagens de um indivíduo da base IFDB, a idade varia de 2 a 45 anos

um grande número de iranianos, figura 4.3. Este banco de dados possui 616(487 homens e 129 mulheres) imagens sem restrições sobre roupas, acessórios, estilo de cabelo, barba ou bigode, perfazendo um total de 3600 imagens com resolução de  $480 \times 640$  *pixels*, essas imagens retratam pessoas entre 2 e 85 anos.

### 4.3 Detecção de Face

A detecção de face consiste em encontrar em imagens complexas ou não uma face humana. Para solucionar este problema alguns métodos propostos em (HJELMAS ERIK; LOW, 2001), tentaram solucioná-lo, mas sempre a um custo computacional muito alto. Este fato foi superado com o algoritmo proposto por Viola e Jones (VIOLA; JONES, 2004) que hoje, está muito perto de ser o padrão na detecção de face por resolver tarefas de detecção com baixo custo computacional. Este sucesso é atribuído principalmente à relativa simplicidade, a rapidez de execução e do desempenho notável do algoritmo (BRADSKI; KEAHLER, 2008). Basicamente o algoritmo combina quatro conceitos-chave:

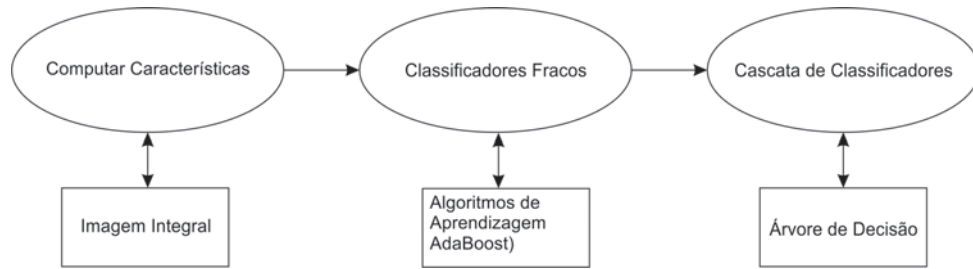


Figura 4.4: Modelo básico de Viola e Jones

- Características retangulares, chamadas de recursos de Haar;
- Imagem Integral
- Algoritmo de aprendizagem - AdaBoost
- Um classificador em cascata;

O modelo da figura 4.4, baseia-se no uso de características de Haar básicas, avaliadas rapidamente por meio de uma nova forma de representação da imagem chamada de imagem integral, para em seguida gerar um conjunto de características. Chama-se o algoritmo *boosting* para otimizar o número de características encontradas e finalmente cria-se uma estrutura de árvore dos classificadores fracos, para que esse gere inferências rápidas e robustas na construção de um classificador forte (VIOLA; JONES, 2004).

### 4.3.1 Características de Haar

As características de Haar codificam a existência de contrastes entre as regiões orientadas da imagem. Um conjunto desses recursos são utilizados para codificar os contrastes exibidos por um rosto humano e suas relações espaciais. Essas características são chamados de Haar, pois seu conceito é semelhante aos coeficientes da wavelet de Haar.

A presença de uma característica de Haar é determinada pela subtração do valor do pixel médio da região pelo valor médio do pixel da região clara. Se a diferença for acima de um limiar (definido durante a aprendizagem), dizemos que a característica está presente.

Viola e Jones (VIOLA; JONES, 2004), relatam o fato de que a escolha do uso de características, ao invés de modelos baseados em estatísticas de pontos da imagem (*pixel*), é importante devido aos benefícios do domínio de conhecimento *ad hoc*, sendo que podemos extrair conhecimentos encobertos nas imagens, dificilmente encontradas em um conjunto finito de treinamentos. No caso da detecção de face este fato muito importante é utilizado para a representação das informações faciais aproximadas e também, para as poucas informações relacionadas a planos de fundo de uma imagem teste.

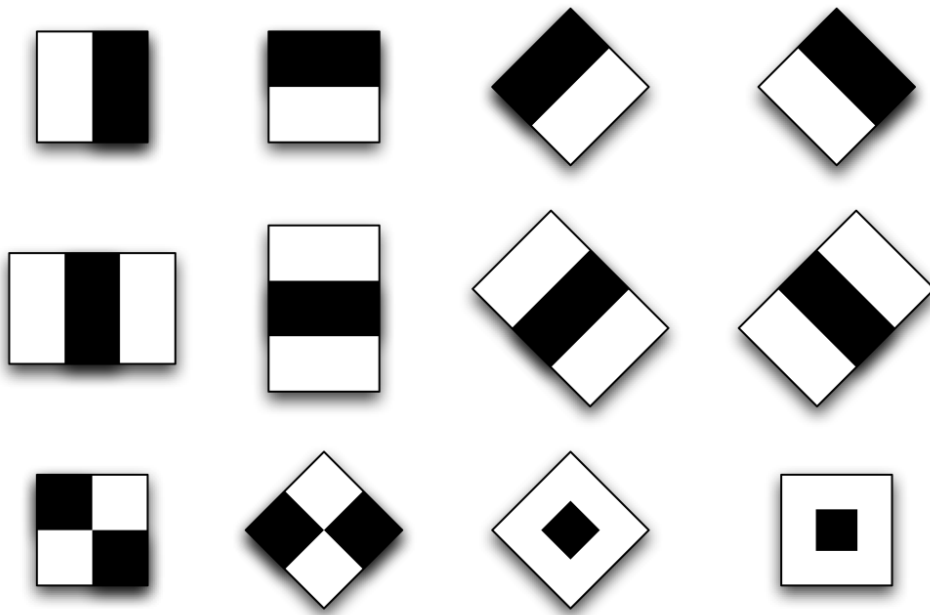


Figura 4.5: Exemplo de Características de Haar utilizados no modelo

De modo geral, portanto, as características nada mais são do que avaliações do conjunto das informações da intensidade de luz de um *pixel*. Esse processo consiste na soma da intensidade dos *pixels* de regiões brancas das características, subtraído da soma da intensidade do restante cinza da imagem. Os resultados são usados como valor da característica de uma determinada localização e podem ser combinados para a formação de hipóteses fracas nas imagens (HJELMAS ERIK; LOW, 2001).

Quatro características do tipo Haar foram escolhidas para uso conforme a figura 4.5, e para determinar a presença, ou ausência de centenas de características de Haar em cada posição de imagem e em varias escalas diferentes, Viola e Jones usou uma técnica chamada de imagem integral.

### 4.3.2 Imagem Integral

Um dos maiores problemas ao se trabalhar com imagens de fundos complexos é encontrar uma representação adequada para a imagem 4.6, pois as representações baseadas em pixel e bordas são inadequadas em algumas situações devido a fatores como variação de textura e alta quantidade de informações que não interessam ao estudo.

Em contrapartida, há uma maior facilidade em localizar áreas de uma face, ou partes da face (nariz, olhos e boca), em relação ao plano do fundo. Porem, é necessário uma nova representação da imagem, esta proposta foi descrita por Viola e Jones (VIOLA; JONES, 2004) e denominada de Imagem Integral.

As imagens integrais criam, a partir da imagem original, uma nova representação da imagem, simplesmente, somando-se os valores de cada *pixel* a esquerda e acima, inclusive. A idéia em usar essa representação é aumentar a velocidade na extração de características, pois qualquer retângulo de uma imagem pode ser calculado por meio da imagem integral. Sendo necessário apenas quatro índices da imagem integral para calcular qualquer retângulo e como consequência imediata, precisa-se de apenas uma passagem para obter dados desejados em sub regiões de uma imagem.

Formalmente uma imagem integral  $I$  é uma representação intermediária para uma imagem e contém a soma dos *pixels*, em escala cinza, da imagem  $N$  com altura  $y$  e largura  $x$ , ou seja:

$$I(x, y) = \sum_{x'=0}^x \sum_{y'=0}^y N(x', y') \quad (4.1)$$

A imagem integral é calculada de forma recursiva, através das fórmulas:  $I(x, y) = I(x, y - 1) + I(x - 1, y) + N(x, y) - I(x - 1, y - 1)$  com  $I(-1, y) = I(x, -1) = I(-1, -1) = 0$ , necessitando apenas uma varredura sobre os dados de entrada. Esta representação intermediária  $I(x, y)$  permite o cálculo de um valor recursivo do retângulo em  $(x, y)$  com altura e largura  $(h, w)$  utilizando quatro referências.

$$F(x, y, h, w) = I(x, y) + I(x + w, y + h) + I(x, y + h) - I(x + w, y) \quad (4.2)$$

O processo descrito na figura 4.6, corresponde ao processo de cálculo da imagem integral, note que cada pixel é a soma dos pixel anteriores (acima e à esquerda).

### 4.3.3 Algoritmo de Aprendizagem

Ao utilizar algoritmos de aprendizagem é possível fazer com que a implementação da detecção sempre aprimore a acurácia de uma classificação, preservando e melhorando tudo aquilo que já aprendeu, ou seja, prover a melhor estimativa possível para a classe de estudo.

Como uma imagem produz uma enorme quantidade de dados a serem traduzidos para obter o resultado desejado. Contudo, selecionar os dados que possuam maior relevância para esta inferência é um fator limitante para a performance do sistema como um todo. Para melhorar este quadro Viola e Jones (VIOLA; JONES, 2004) propõem combinar clas-

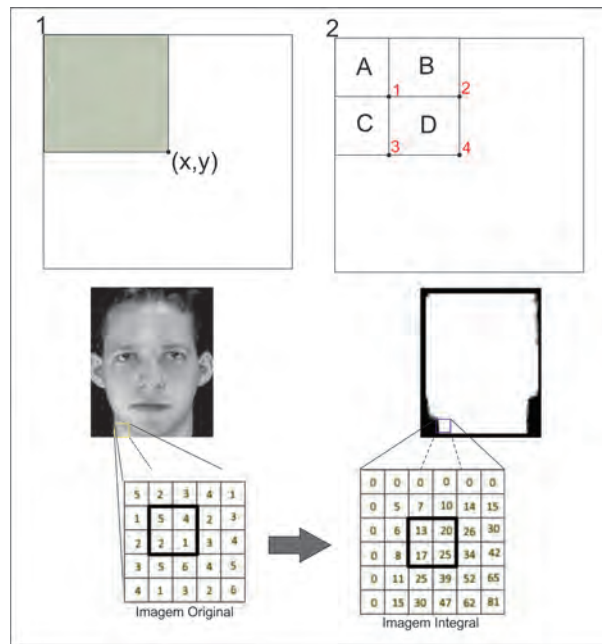


Figura 4.6: Imagem Integral, Teoria e resultado

sificadores fracos podendo eventualmente disponibilizar um classificador forte, em outras palavras o classificador forte produzido pelo algoritmo de aprendizagem é simplesmente um conjunto de características com pesos que classifica de forma precisa dois conjuntos de imagens pré-rotuladas as quais as características com pesos maiores são boas para a classificação de exemplos definidos como parte de um certo conjunto.

### 4.3.3.1 Classificador Fraco

Um classificador fraco  $h_j$  é uma simples estrutura contendo um vetor de característica  $f_j$ , um limiar  $\theta_j$  e uma paridade  $p_j$ . A idéia básica desse classificador é encontrar um limiar que melhor separe o valor de uma característica entre as imagens definidas como positivas, ou negativas.

$$\begin{cases} 1 & \text{se } p_j f_j(x) < p_j \theta_j \\ 0 & \text{Caso contrário} \end{cases} \quad (4.3)$$

Para que complete a tarefa de classificação é importante descobrir um limiar ótimo  $\theta_j$ . Este limiar não é definido de modo claro e é assumido apenas, que favorece o valor de menor erro de predição. Logo, temos como consequência vários modelos que procuram descobrir um bom limiar.

### 4.3.3.2 AdaBoost

O problema Boosting nasceu na área de computação conhecida com *machine learning* que informalmente pode ser exposto da seguinte forma: Suponha que existe um método de classificação que é ligeiramente melhor do que uma escolha aleatória, para qualquer distribuição em  $X$ . Chamado de *weak learner* ou classificador fraco. A existência de um classificador fraco implica na existência de um classificador forte, *strong learner*, com erro pequeno sobre todo o espaço  $X$ ? Em estatística consiste em perguntar se dado um método razoável de estimação, é possível obter um método próximo de ótimo.

Este problema foi resolvido por Schapire (SCHAPIRE, 1990), que apresentou um algoritmo que transforma um classificador fraco num classificador forte. A partir de então, foram desenvolvidos vários algoritmos dentro do contexto *boosting*, sendo um dos mais recentes e bem sucedido o algoritmo conhecido com AdaBoost que é oriundo do fato de que o *boosting* gera em cada passo uma distribuição sobre as observações da amostra, dando maior peso (maior probabilidade de estar na amostra perturbada) às observações classificadas incorretamente no passo anterior.

---

**Algoritmo 1:** Modelo básico do algoritmo AdaBoost

---

**Entrada:** Tome uma amostra:  $\mathbb{S} = \{(x_i, y_i), x_i \in X, y_i \in \{-1, 1\}\}$

$w_1 \leftarrow u$

**para**  $t = 1, 2, \dots, T$  **faça**

Tome  $(h_t : X \rightarrow S) \leftarrow WL(\mathbb{S}, w_t)$ ;

Encontre  $\alpha_t \in \mathbb{R}$

Atualize  $\forall 1 \leq i \leq m$ ,

$w_{t+1,i} \leftarrow w_{t,i} \cdot e^{(-\alpha_t y_i h_t(x_i)) / Z_t}$

**Saída:**  $H_T(x) = \sum_{t=1}^T \alpha_t h_t(x)$

---

Nesse sentido o AdaBoost está focado nas classificações ruins, ou então, nos dados de difícil classificação, e esta é a característica principal deste algoritmo: minimizar o erro sobre um conjunto de treinamento. Uma das vantagens do adaboost (NOCK; NIELSEN, 2006), (GAO; ZHOU, 2010), é a existência de outros parâmetros, além dos turnos  $T$ , para que melhore o aprendizado. O resultado, após sucessivas iterações do algoritmo, de um conjunto de hipóteses com pesos, em que aqueles que possuírem menores erros de classificação ganham maior importância e, é chamado de hipótese forte ou classificador forte.

### 4.3.4 Um classificador em cascata

Aumentar a velocidade de uma tarefa de classificação, em geral, resulta num aumento dos erros associados. Porém, para surtir este efeito teríamos que reduzir o número de avaliação dos classificadores fracos, o que resultaria numa perda na acurácia do sistema. Por isso

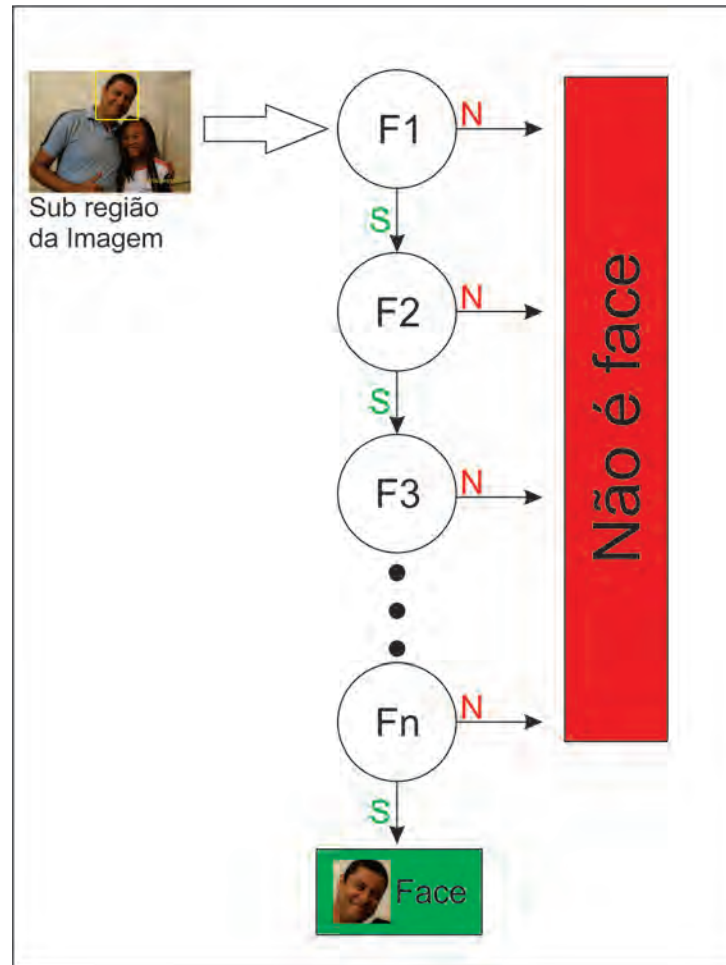


Figura 4.7: Modelo do Classificador em Cascata

Viola e Jones (VIOLA; JONES, 2004), propõem uma árvore degenerativa de decisão, *decision stump*, cuja estrutura contendo o encadeamento de classificadores do mais genérico ao mais específico, segundo o qual os primeiros níveis da cascata não são muito precisos, apesar de conseguir classificar uma grande quantidade de amostras com uma pequena quantidade de características.

O uso da cascata caracteriza-se pelo fato de que, em uma imagem durante uma tarefa de detecção, a maioria das sub janelas analisadas pelo classificador são rejeitadas. Por essa razão, uma generalização nos primeiros estágios deve ser suficientemente alta para que evite a passagem para estágios subsequentes, sub janelas classificadas como falso positivo (HORTON; CAMERON-JONES; WILLIAMS, 2007), conforme descrevemos na figura 4.7.



## 4.4 Pré-Processamento

### 4.4.1 Introdução

Entendemos por processamento digital de imagens como a manipulação numérica de imagens digitais por meio do uso de funções analisadores, que, em conjunto, formam os sistemas de tratamento de imagens.

Como já vimos, um imagem digital foi definida com uma matriz, em que cada elemento da matriz é chamada de *pixel*. Cada *pixel* possui três coordenadas: valores de X e Y que definem a sua posição na imagem e um valor de Z que representa o nível de cinza ou grau de brilho. O nível de cinza corresponde a uma unidade de informação em termos da quantidade de energia refletida e, ou, emitida pelo objeto imageado, em diferentes comprimentos de onda no espectro eletromagnético (ROSOT; DISPERATI; FILHO, 2001) Na maioria dos sistemas, um valor de  $Z = 0$  representa o preto absoluto e 255 o branco absoluto.

Considerando que a finalidade principal do processamento digital é fornecer ferramentas que facilitem a identificação e a extração de informações contidas nas imagens, para posterior interpretação e análise, ou seja, a rotina de pré-processamento consiste em uma série de operações que são aplicadas à imagem no nível de intensidade do *pixel*. Sendo que o objetivo desta rotina é a obtenção da mais regular das características da região facial dos indivíduos, eliminando possíveis distorções, redimensionamento e mudança do formato RGB para tons cinza. Atuando diretamente sobre os *pixels* da imagem este procedimento pode ser traduzido na equação (4.4).

$$g(x, y) = T[f(x, y)] \quad (4.4)$$

sendo  $f(x, y)$  a imagem de entrada,  $g(x, y)$  a imagem de saída(processada) que representa, na realidade, a intensidade de luz do *pixel* e  $T$  um operador associado a  $f$  definido previamente.

### 4.4.2 Correção de intensidade de luz

Esta correção mapeia os valores de  $f$  ajustando-os a uma nova função  $g$  que descreve uma relação ideal entre os valores de  $f$  e  $g$  Neste contexto o parâmetro gama especifica a curva criada pela intensidade de luz da função  $g$  em relação a  $f$ , basicamente as curvas descritas na figura 4.8.

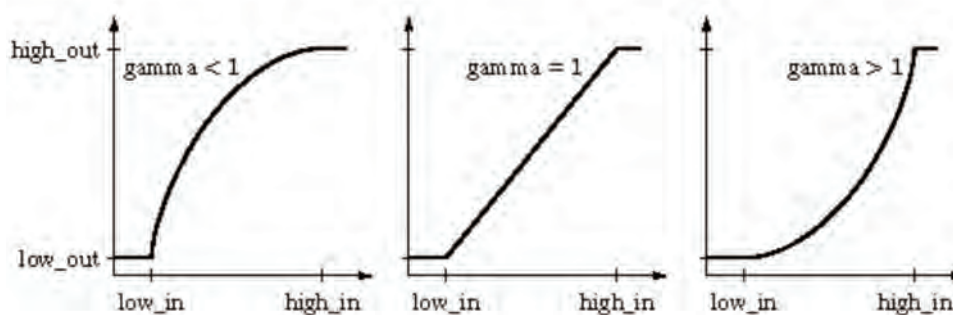


Figura 4.8: Correção Gamma

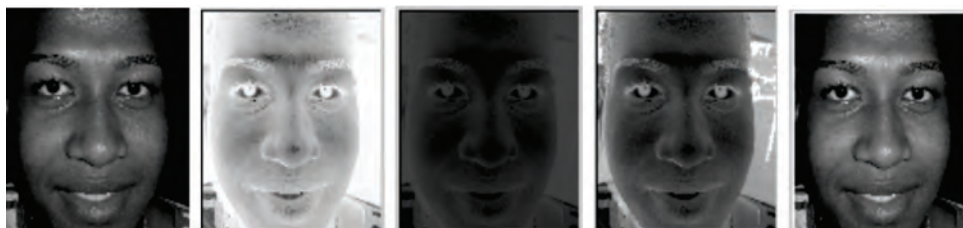


Figura 4.9: Ajuste da Imagem através da correção gama, imagem da correção com fatores modificados

A correção gama tem por objetivo aumentar a faixa dinâmica da imagem nas regiões mais escuras e diminuir nas mais iluminadas, podemos ver seus resultados na figura 4.9. Dados os níveis de cinza  $f(x, y)$  dos *pixels* de uma imagem de tamanho  $M \times N$  ( $0 \leq x \leq M; 0 \leq y \leq N$ ) (GONZALEZ; WOODS; EDDINS, 2004).

### 4.4.3 Transformação Logarítmica

A transformação logarítmica é similar a curva gama, onde o menor(Low) valor é 0 e o maior(high) é 1. formato da curva gama é variável, já o formato da função logarítmica é fixo. Um dos principais usos da transformação logarítmica é a compressão dinâmica. Por exemplo, para mostrar um espectro (Fourier) com valores que vão de 0 a 106 ou mais, em uma escala linear (8 bits), os valores mais altos prevalecem na figura, resultando em perda de detalhes dos valores de baixa intensidade. Através da transformação logarítmica, um valor da ordem de 106 é reduzido em aproximadamente 14 vezes (GONZALEZ; WOODS; EDDINS, 2004). A transformação logarítmica é dada da seguinte forma:

$$g = c * \log(1 + f) \quad (4.5)$$

Onde  $g$  é a imagem de saída,  $f$  a imagem de entrada e  $c$  uma variável de intensidade que é aplicada a cada um dos pontos da imagem (pontual em cada *pixel*). A transformação de cada *pixel* não está atrelada aos *pixels* da vizinhança. O resultado básico



Figura 4.10: Transformação logarítmica aplicada a imagem original e com  $c = 0.5$

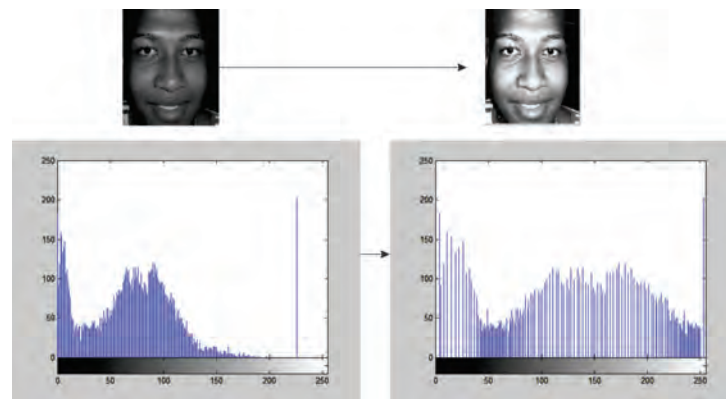


Figura 4.11: Equalização do histograma

desta transformação pode ser visto na figura 4.10.

#### 4.4.4 Equalização de Histogramas

O histograma de uma imagem é um gráfico que descreve o número de pontos por cada nível de cinza da imagem. Equalizar o histograma significa obter a máxima variância do histograma de uma imagem, obtendo assim uma imagem com o melhor contraste. O contraste é uma medida qualitativa que está relacionada com a distribuição dos tons de cinza em uma imagem.

Para tanto, utilizamos uma função auxiliar denominada função transformação. A forma mais usual de equalizar um histograma é utilizar a Função de Distribuição Acumulada (CDF) da distribuição de probabilidades, que pode ser expressa por (GONZALEZ; WOODS; EDDINS, 2004):

$$s_k = T(r_k) = \sum_{j=0}^k \frac{n_j}{n} = \sum_{j=0}^k P_r(r_j) \quad (4.6)$$

onde:

- $n$  é o número total de *pixels* da imagem,
- $0 \leq r_k \leq 1$
- $k = 0, 1, \dots, L - 1$  onde  $L$  é o número de níveis de cinza da imagem digitalizada
- $P_r(r_j)$  é a probabilidade do  $j$ -ésimo nível de cinza (MENDONÇA, 2008).

A equalização do histograma é um modelo global utilizado para visualizar mais claramente a imagem, porém. Produz um resultado não satisfatório em algumas regiões locais.

#### 4.4.5 Filtro Linear Espacial

Um filtro espacial para os quais a atribuição de cor cinza AT coordenadas  $(x, y)$  na transformação da imagem é feita por alguma média ponderada (combinação linear) de tons de cinza localizado a um determinado padrão espacial em torno de coordenadas  $(x, y)$  da imagem de domínio. Por exemplo, um filtro espacial linear, que enfatiza a altas frequências espaciais (high-pass), tende a aguçar, as bordas em uma imagem. Um filtro espacial linear, que enfatiza as frequências espaciais baixas (low-pass), tendem a borrar a imagem e reduzir o ruído. Em particular no nosso trabalho utilizamos o filtro Laplaciano que caracteriza uma imagem  $f(x, y)$ , denotado por  $\nabla^2 f(x, y)$  e definida como:

$$\nabla^2 f(x, y) = \frac{\partial^2 f(x, y)}{\partial x^2} \quad (4.7)$$

Comumente usamos aproximações digitais da segunda derivada, descritos como:

$$\frac{\partial^2 f}{\partial x^2} = f(x + 1, y) + f(x - 1, y) - 2f(x, y) \quad (4.8)$$

e

$$\frac{\partial^2 f}{\partial y^2} = f(x, y + 1) + f(x, y - 1) - 2f(x, y) \quad (4.9)$$

de modo que

$$\nabla^2 f = [f(x + 1, y) + f(x - 1, y) + f(x, y + 1) + f(x, y - 1)] - 4f(x, y) \quad (4.10)$$

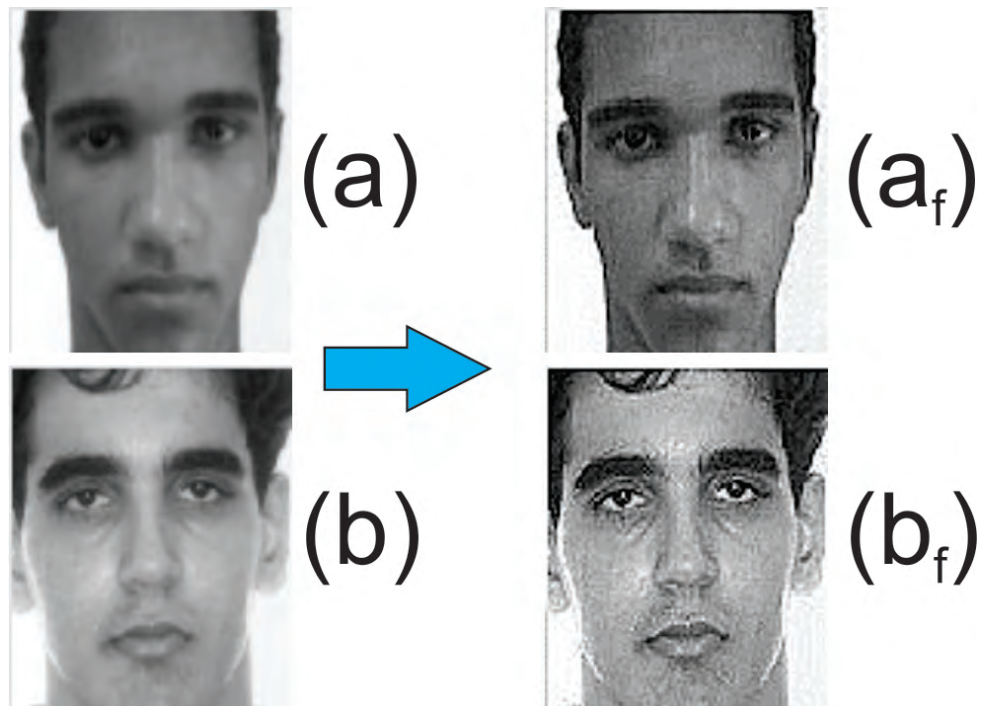


Figura 4.12: As imagens (a) e (b) representam a captura sem o filtro e as (a<sub>f</sub>) e (b<sub>f</sub>) representam as mesmas após o uso do filtro

Aplicando esta expressão e utilizando uma das máscaras (4.11) e (4.12):

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (4.11)$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (4.12)$$

Numa outra abordagem, podemos utilizar outra máscaras que leva em conta os elementos da diagonal, ilustramos estas aplicações na figura 4.12.

## 4.5 Testes e Resultados

Nesta seção, os experimentos realizados serão relatados, assim como seus resultados obtidos. Estes experimentos têm por finalidade demonstrar a viabilidade do modelo implementado e avaliar seu desempenho nas situações críticas para o reconhecimento de pessoas desaparecidas. Inicialmente realizamos os teste com uma base de dados (ORL),

que estabeleceu parâmetros mínimos de comparação tanto para treinamento como para reconhecimento. As estimativas de erro e acerto utilizando o banco ORL foram avaliadas de duas formas:

1. Todas as amostras foram utilizadas para o treinamento e teste;
2. As amostras foram separadas em conjuntos de 1, 2, 3, 4 e 5 classes.

Os testes ocorreram utilizando imagens estáticas (fotografias) e dinâmicas (vídeo) entre os dias 01 e 30 de outubro de 2010, com uma IDE desenvolvida pelo autor, figura 4.13. As imagens estáticas e dinâmicas foram captadas por uma câmera comum, Samsung L100 com resolução de 3 megapixels, que tenta reproduzir as condições em que as fotos de pessoas desaparecidas são disponibilizadas em diversos sites na web. Durante a realização dos testes os colaboradores da pesquisa estavam em poses comuns e em algumas imagens utilizando acessórios como óculos, chapéus, diferentes cortes de cabelo além de variarem bastante as feições faciais.

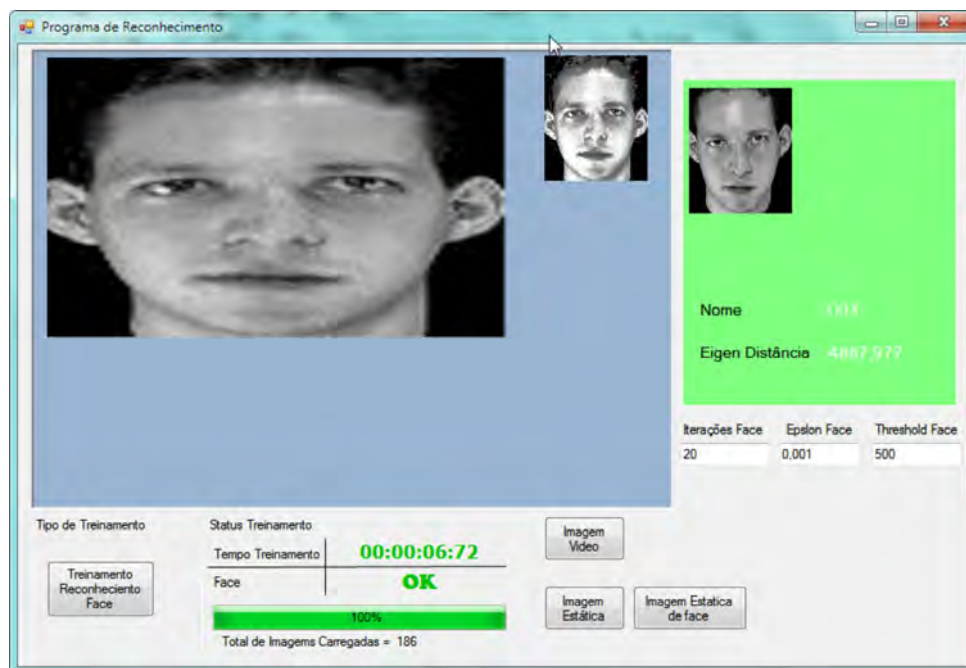


Figura 4.13: Tela principal do modelo implementado

Utilizamos nesta implementação e testes um laptop pessoal Itautec, com 2GB de memória RAM, HD de 80GB e processador Intel Core 2 Duo com sistema operacional Windows 7 e o Visual Studio 2010. Quanto a variação de expressões foi solicitado a cada indivíduo que fizesse expressões que chamamos de neutra, sorrindo, chateado, alegre, triste e tentando desfigurar a face (careta). Sendo capturada uma imagem em cada situação descrita.

As imagens obtidas são coloridas e possuem um tamanho de 640x486 *pixels* totalizando uma base de aproximadamente 20GB compostas por imagens no formato JPG. As ima-

gens utilizadas para reconhecimento foram captadas sem nenhuma preocupação com iluminação, pose ou alguma forma de melhorar a condição de captação ou representação. A partir dessas imagens extraiu-se a face de cada indivíduo projetá-la no espaço de faces para então compará-la e atribuir à face mais próxima.

Os testes foram motivados pelo fato de tais situações serem corriqueiras para pessoas e comuns aos bancos de dados de pessoas desaparecidas. Por exemplo, uma imagem adquirida por câmeras de segurança podem conter inúmeras situações que provocam ruídos que impedem o reconhecimento correto, nesse sentido realizamos experimentos nas seguintes condições:

- Experimento 1: Teste de tempo de processamento do treinamento;
- Experimento 2: Teste de reconhecimento via imagem estática;
- Experimento 3: Teste de reconhecimento via imagem dinâmica;
- Experimento 4: Análise do tom de pele;
- Experimento 5: Estimar a idade;

Para efeito de comparação realizamos inicialmente uma análise da implementação utilizando a base ORL e esta revelou os resultados impressos na tabela 4.1:

Total Indivíduos	Total de Imagens	Taxa de Reconhecimento		Acerto(%)	Erro(%)
		Positivo	Falso-positivo		
1	19	15	4	79%	21%
2	38	19	0	100%	0%
3	57	14	5	74%	26%
4	76	16	3	84%	16%
5	95	18	1	95%	5%
10	190	19	0	100%	0%

Tabela 4.1: Tabela de resultados do reconhecimento utilizando a base ORL e a nossa implementação

Os resultados obtidos pela implementação correspondem ao que foi verificado em (ATALAY, 1996), (LAKSHMANAN et al., 2001), (ZHUIE; YU, 1994), (PERLIBAKAS, 2008), que circulam em torno de 80% a 90% de resultados positivos chamando a atenção para o fato do treinamento com apenas uma imagem de entrada e a alta taxa de reconhecimento com o modelo. Chamamos a atenção para o fato da base ser constituída, exclusivamente, por pessoas de tom de pele branca o que nos impediu fazer comparações dessa base com a nossa em relação a este item. Este fato nos trouxe um parâmetro novo para estudo, seria o tom da pele um dificultador para o reconhecimento? Porque?

### 4.5.1 Experimento 1: Teste de tempo de processamento do treinamento

As imagens desse banco de dados passaram pelas etapas de pré-processamento descritas na sessão anterior como também foram testadas sem a utilização dos mesmos. Inicialmente, realizamos os testes de treinamento com relação ao tempo de treinamento. Este teste se fez necessário para analisar qual deverá ser o comportamento desse algoritmo em bancos com um grande número de pessoas incluídas. Não existem estimativas oficiais do total de pessoas desaparecidas no país mas algumas ONGs que disponibilizam sites para localização de pessoas desaparecidas no Brasil estimam que existem mais de 50 mil pessoas (OLIVEIRA, 2007).

Utilizando o banco ORL realizamos 10 testes de treinamento que resultaram na média descrita no gráfico 4.14, com isso podemos comparar com o treinamento da nossa base, grafico 4.15. Apenas modificamos a unidade de tempo, pois como o banco ORL é menor e possui uma capacidade de processamento muito mas rápida que o da nossa base.

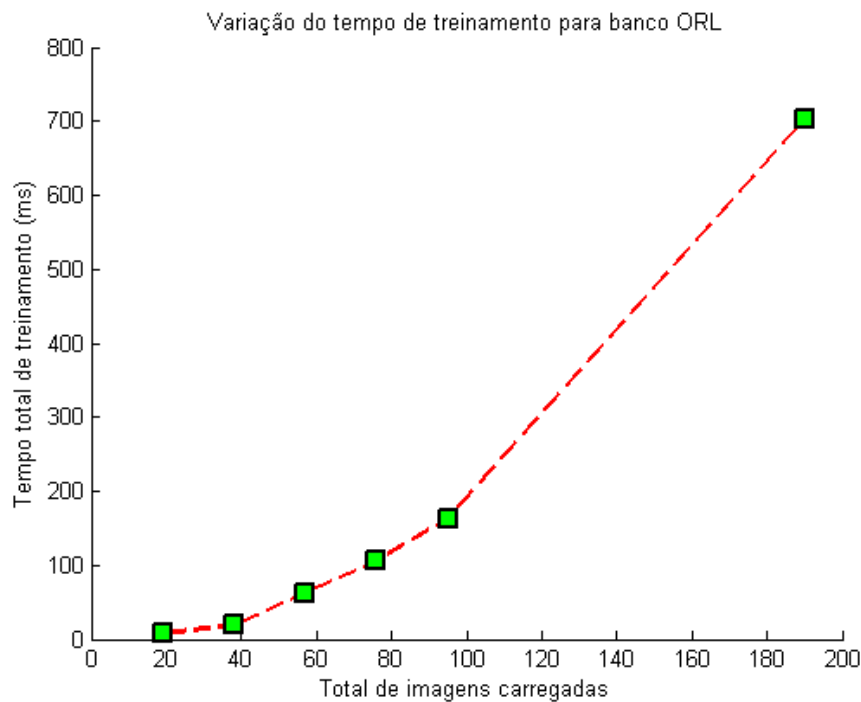


Figura 4.14: Tempo de treinamento - Base ORL

Os testes incluíram imagens com ou sem pré-processamento, na posição frontal com variação de iluminação e expressão, além de incluir elementos que não representam faces para efeito de análise do comportamento do algoritmo no caso de um inserção no espaço de face de algo que não é face.



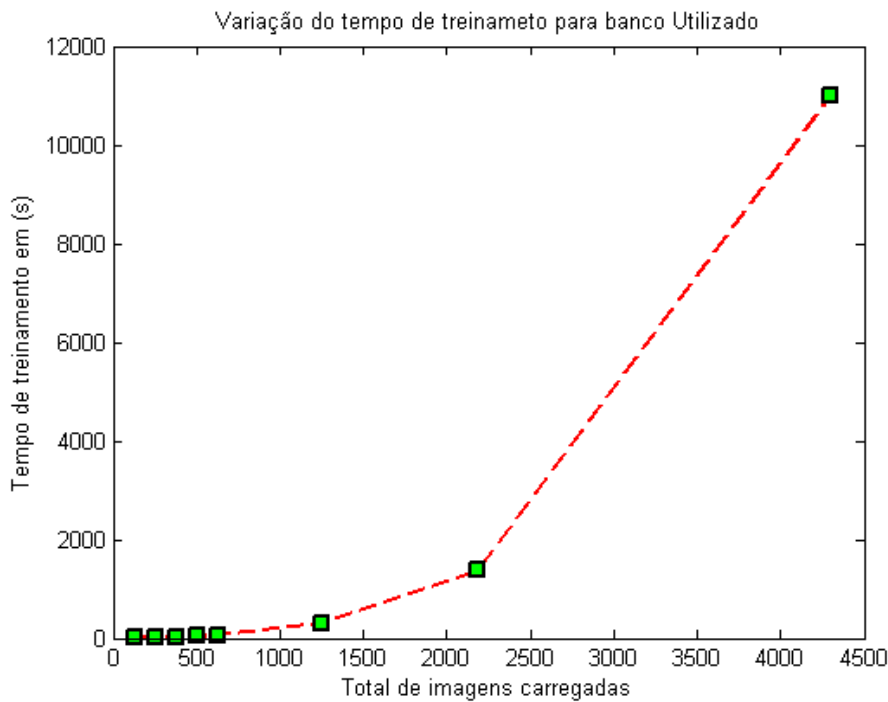


Figura 4.15: Tempo de treinamento - Nossa base

#### 4.5.2 Experimento 2: Teste de reconhecimento via imagem estática

Para a realização destes testes utilizamos imagens captadas entre às 07:00hs e às 18:30h durante 30 dias, as imagens não tem qualquer tipo de intervenção de processamento ou melhoria durante a captura. A câmera utilizada foi a mesma da que utilizamos para a construção do banco de dados e utilizada em modo automático de captura, que proporciona uma melhor captação dentro dos parâmetros da máquina.

Dessa forma, procuramos reproduzir a situação da qual algum órgão de localização de pessoas desaparecidas encontrará ao receber as imagens para localização. Todas as imagens utilizaram o formato JPG com 3MB de resolução e qualidade superfina (definição do fabricante). Ainda na fase de aquisição devemos considerar que não foi criada ou articulada nenhuma forma de compensação de iluminação, dependendo exclusivamente da luz natural do ambiente apesar de entender que a iluminação tem um peso significativo no processamento da imagem e para compensar essa falta, utilizamos os filtros descritos na fase de pré-processamento, que melhoram essa condição. As figuras 4.16 e 4.17 mostram o modelo básico de iluminação que nos deparamos e como a câmera reage a esta condição.

Com este esquema básico de captura foram inseridas para teste exatamente 978 imagens com múltiplas faces ou não, ou seja em várias ocasiões não foi capturada apenas a imagem de uma única pessoa mas de duas ou mais. Em geral as imagens mostram os individuo em posturas despojadas e que diferem algumas vezes significativamente da imagem de teste

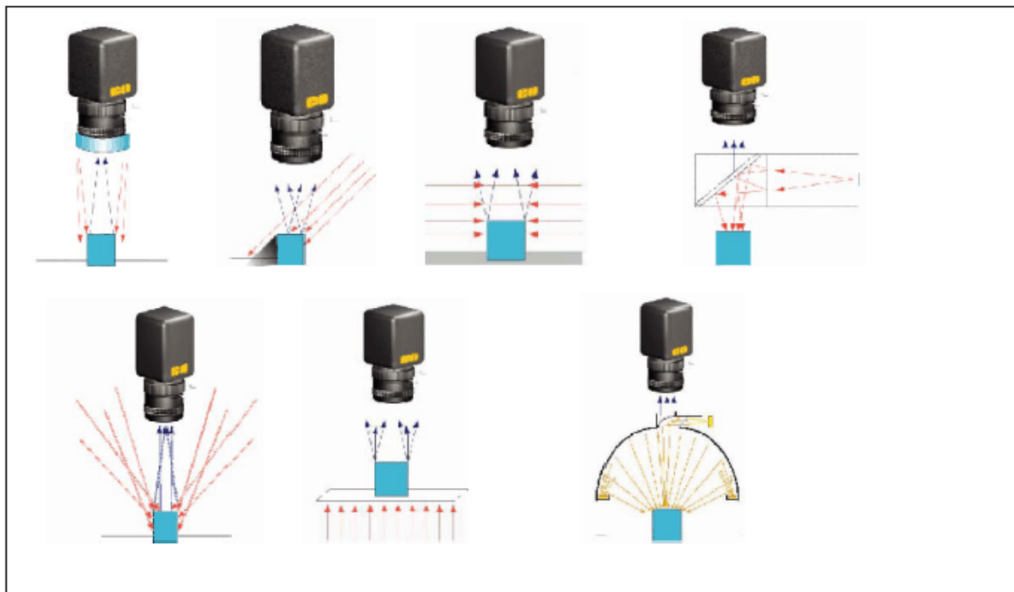


Figura 4.16: Formas de iluminação. Partindo do canto superior esquerdo temos iluminação direta e unilateral, angular, axial ou perpendicular, axial difusa, frontal difusa, iluminação de fundo, iluminação difusa uniforme. Adaptado de (GONZALEZ; WOODS; EDDINS, 2004)

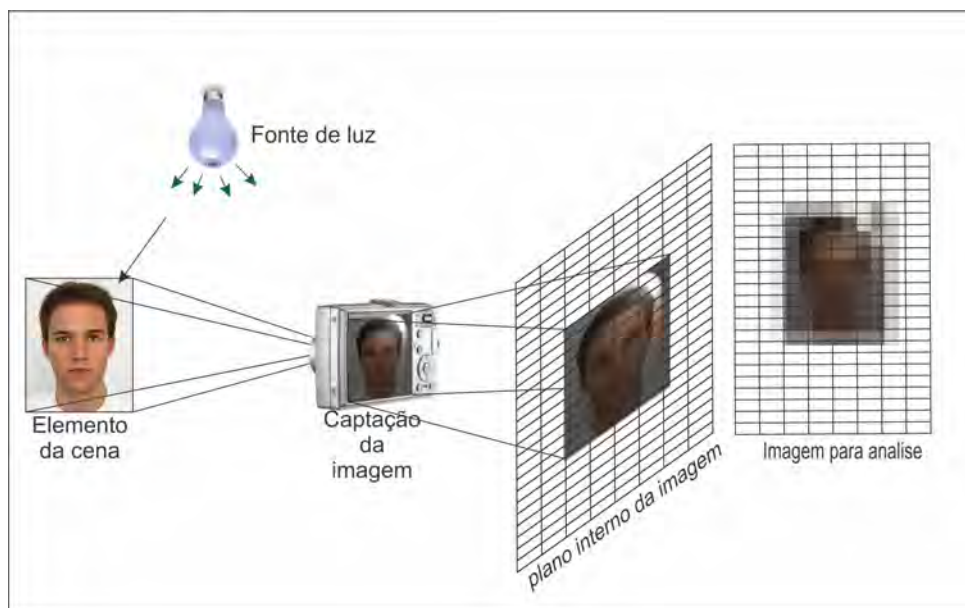


Figura 4.17: Representa a captura da imagem, formação e discretização. Adaptado (GONZALEZ; WOODS; EDDINS, 2004)

ou não. Em diversos momentos estavam com formas de cabelo, barba, bigode, chapéus, toucas, maquiagem excessiva ou não, pinturas, inclusive com fantasias com máscaras que impedem a visualização total da face.

No ambiente ao qual estas imagens foram captadas elementos de iluminação são naturais e artificiais, isto por que o prédio, onde foram coletadas as amostras, não possui somente iluminação natural. Em alguns momentos a iluminação artificial falhou, devido a quedas de energia, o que proporcionou uma coleta de imagens com iluminação unicamente natural. Esta captura representou um percentual muito pequeno do total de fotos ou seja 9 fotos que ficaram muito próximas da figura 4.18 e não houve reconhecimento positivo, mesmo utilizando filtros já propostos que compensam o baixo grau de luz.

Assim como tivemos imagens com baixa quantidade de luz, o oposto também se verificou 4.19 nesse caso também a amostra foi pequena, apenas 8 fotos. A dificuldade de coletar estes dados deveu-se a incompatibilidade de horários, durante a etapa de coleta, que impediu o registro da mesma pessoa varias vezes, para assim podermos analisar esta situação. Assim como com baixa iluminação o reconhecimento falhou o mesmo ocorre com alta incidência de luz.

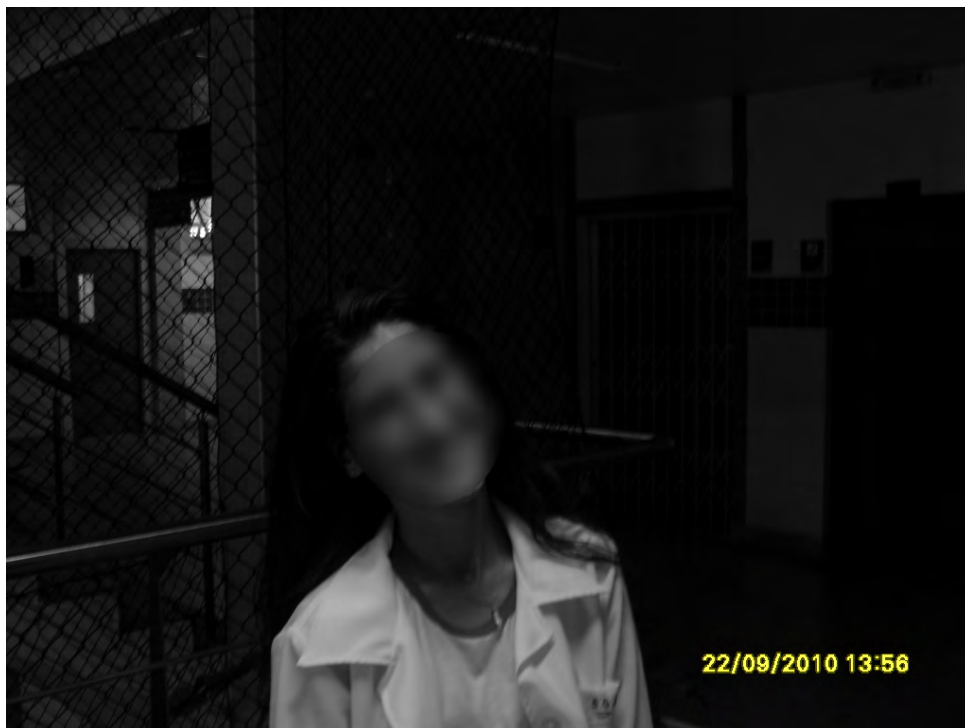


Figura 4.18: Imagem captada sem auxílio de iluminação artificial em ambiente interno

A idade média dos indivíduos que aceitaram ceder a imagem para este trabalho é 15 anos com variações de 2 anos para baixo e 2, 3 ou até 4 para cima. Apesar de conseguir 134 voluntários é importante ressaltar que a expectativa era da colaboração de 320 indivíduos, mas devido a conjuntura atual da sociedade muitos não aceitaram ceder a



Figura 4.19: Imagem com alta incidência de luz, captada às 12:00h com iluminação natural

imagem para o experimento. Com este cenário obtivemos os resultados discriminados nas tabelas 4.2,4.3,4.4 e 4.5 :

Dia de Teste	Testadas	Trei. p/pessoa	Reconhecimento		Acerto(%)	Erro(%)
			Positivo	Falso-positivo		
1	28	20	15	13	54%	46%
2	28	20	15	13	54%	46%
3	28	20	16	12	57%	43%
4	28	20	17	11	61%	39%
5	28	20	15	13	54%	46%
6	28	20	15	13	54%	46%
7	28	20	15	13	54%	46%
8	28	20	15	13	54%	46%
9	28	20	12	16	43%	57%
10	29	20	12	17	41%	59%
<b>Totais</b>	<b>281</b>		<b>147</b>	<b>134</b>	<b>52%</b>	<b>48%</b>

Tabela 4.2: Tabela com resultados dos 10 primeiros dias

Com estes dados coletados podemos realizar uma análise geral que resultou na tabela 4.5 e gráfico :

ou resumidamente:

Além disso, os comparamos como o modelo se comporta para uma base de dados com imagens variando entre 1 e 20 imagens de treinamento, que resultou no seguinte gráfico:

Dia de Teste	Testadas	Trei. p/pessoa	Reconhecimento		Acerto(%)	Erro(%)
			Positivo	Falso-positivo		
11	35	20	23	12	66%	34%
12	32	20	27	5	84%	16%
13	31	20	21	10	68%	32%
14	34	20	20	14	59%	41%
15	32	20	19	13	59%	41%
16	33	20	18	15	55%	45%
17	32	20	27	5	84%	16%
18	40	20	27	13	68%	33%
19	43	20	31	12	72%	28%
20	42	20	28	14	67%	33%
<b>Totais</b>	<b>354</b>		<b>241</b>	<b>113</b>	<b>68%</b>	<b>32%</b>

Tabela 4.3: Tabela com resultados do reconhecimento com imagens coletadas entre os dias 11-20

Dia Teste	Testadas	Trei. p/pessoa	Reconhecimento		Acerto(%)	Erro(%)
			Positivo	Falso-positivo		
21	34	20	22	12	65%	35%
22	33	20	19	14	58%	42%
23	35	20	23	12	66%	34%
24	37	20	26	11	70%	30%
25	27	20	18	9	67%	33%
26	33	20	22	11	67%	33%
27	37	20	24	13	65%	35%
28	33	20	21	12	64%	36%
29	32	20	29	3	91%	9%
30	42	20	34	8	81%	19%
<b>Totais</b>	<b>343</b>		<b>238</b>	<b>105</b>	<b>69%</b>	<b>31%</b>

Tabela 4.4: Dados dos últimos 10 dias de coleta de dados

Testadas	Trei. p/pessoa	Reconhecimento		Acerto(%)	Erro(%)
		Positivo	Falso-positivo		
281	20	147	134	52%	48%
354	20	241	113	68%	32%
343	20	238	105	69%	31%
<b>978</b>		<b>626</b>	<b>352</b>	<b>64%</b>	<b>36%</b>

Tabela 4.5: Dados condensados com resultados generalizado



Figura 4.20: Resultado representado graficamente

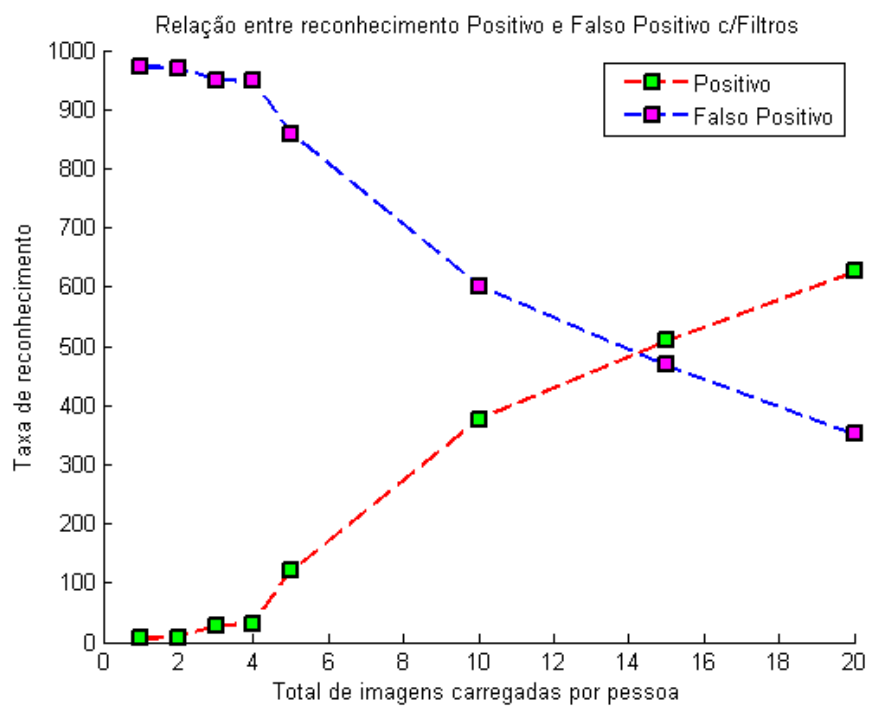


Figura 4.21: Relação entre a quantidade de imagens de treinamento e a taxa de reconhecimento utilizado no modelo

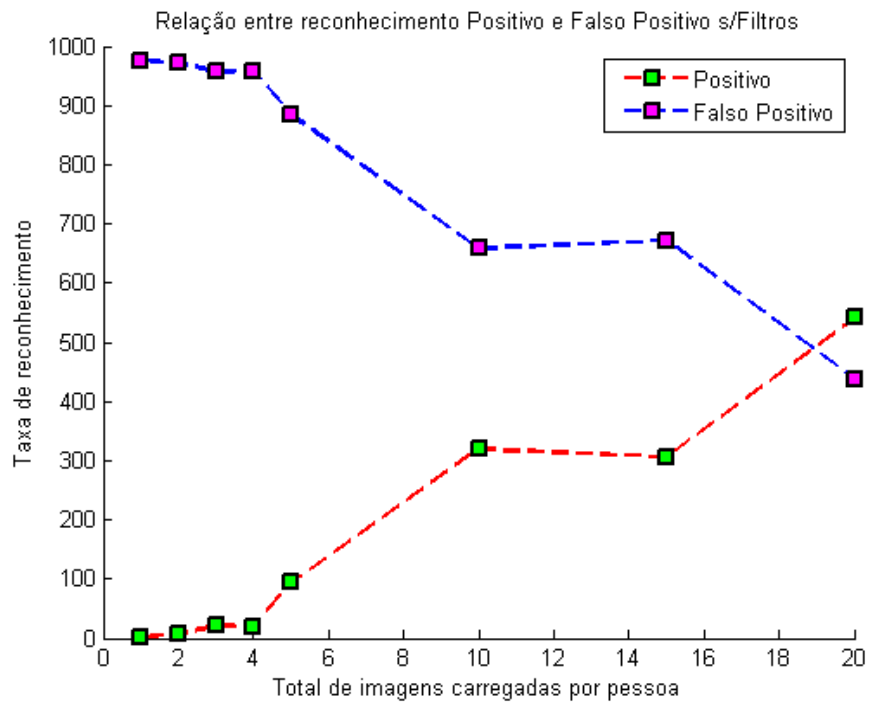


Figura 4.22: Relação entre a quantidade de imagens de treinamento e a taxa de reconhecimento utilizado o modelo e sem filtros

Os testes realizados com as imagens da nossa base com ou sem filtro, tentamos criar, a partir de uma imagem do indivíduo outras 10 a partir desta, para isso criamos a imagem utilizando o MATLAB e funções que geram imagens com filtros, correções e perturbações.



Figura 4.23: O fundo verde representa a imagem original, as demais com fundo azul representam as imagens modificadas e inseridas para treinamento

Somamos a estas 10 imagens outras 2551 que formaram um grupo de treinamento que levou 1:34:03h, com isso, percebemos que em relação ao treinamento realizado com imagens com apenas um tipo de filtro este grupo levou muito mais tempo que o anterior, além disso o resultado do reconhecimento não foi tão satisfatório quanto com imagens com diferença de pose.



Disfarce					
Treinamento	Analizadas	Positivo	Falso Positivo	Positivo(%)	Falso Positivo(%)
2561	294	41	253	14%	86%

Tabela 4.6: Reconhecimento utilizado uma única imagem.

Disfarce					
Disfarce	Total de Imagens	Positivo	Falso Positivo	Positivo(%)	Falso Positivo(%)
Óculos	22	2	20	9%	91%
Chapeu	22	2	20	9%	91%
Oculos+Chapeu	22	1	21	5%	95%
Distorção Face	62	1	61	2%	98%

Tabela 4.7: Reconhecimento utilizando algum tipo de disfarce

### 4.5.3 Experimento 3: Teste de reconhecimento via imagem dinâmica

Chamamos aqui de imagem dinâmica a todos os vídeos que foram captados, em formato AVI, utilizando o mesmo equipamento dos experimentos com fotos. Esta captura ocorreu sempre com iluminação natural somada a artificial, em um ambiente fechado, com 30fps e resolução de 1MB, com ajustes padronizados pelo fabricante do equipamento. Os vídeos tem em média 1 minuto de duração perfazendo um total 1800 *frames* que foram analisados pelo modelo. Todos os indivíduos presentes nas imagens estavam a uma distância média de no máximo 3 metros do equipamento em movimento ou não. A câmera estava sempre a mão e todos os vídeos totalizaram um tempo de 30 minutos. Os vídeos procuram captar as mesmas pessoas que compuseram a base de treinamento e em diferentes ângulos, poses e com acessórios. Com isso, obtivemos um resultado muito menor em termos de reconhecimento positivo.

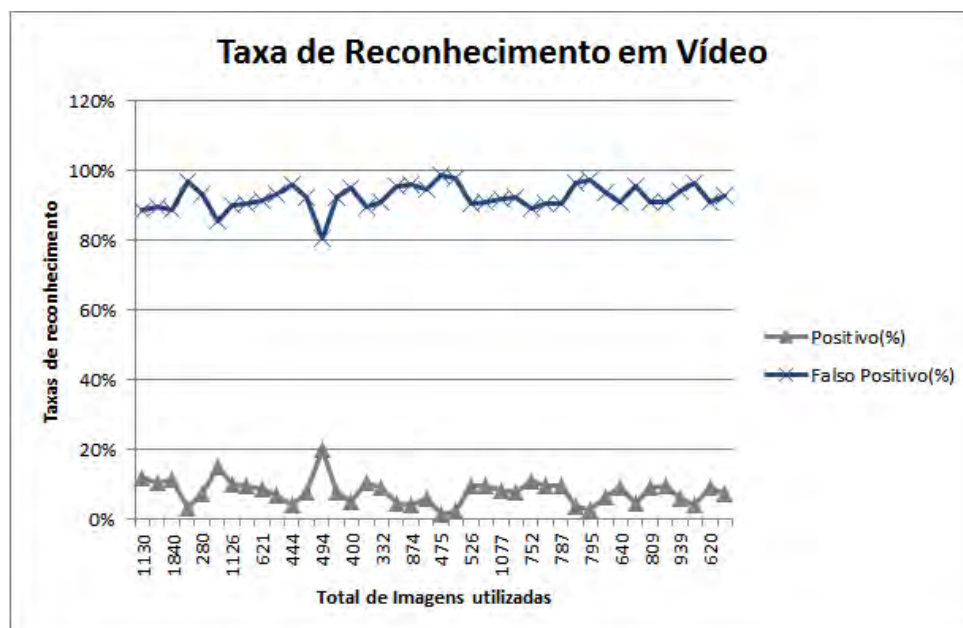


Figura 4.24: Resultados do reconhecimento obtido em câmera de vídeo



Durante o processo de detecção muitas imagens são descartadas devido a sua baixa qualidade para o reconhecimento, bem como imagens que não representam face ou imagens captadas por indivíduos em movimento, que em geral não são bem tratadas pelo modelo. Durante a captura notou-se que as imagens em escala RGB possuíam uma melhora no reconhecimento do que no formato *GRAYSCALE*.

#### 4.5.4 Análise do Reconhecimento levando em conta o tom de pele

Os testes realizados nessa sessão foram motivados pelo excessivo número de falsos positivos que notamos ao tentar reconhecer a face de pessoas com tom de pele negro. Trabalhos recentes afirmam que o reconhecimento facial independe do tom da pele (BERNSTEIN; YOUNG; HUGENBERG, 2007), (LEVIN, 2000), mas apesar destes trabalhos defenderem este ponto nosso trabalho identificou separou e classificou o reconhecimento por tom de pele e identificamos os seguintes resultados:

Tom pele branco (480 imagens)		Tom Pele Moreno (384 imagens)		Tom Pele Negra (114 imagens)	
Positivo	Falso Positivo	Positivo	Falso Positivo	Positivo	Falso Positivo
404	76	210	174	37	77
41%	8%	21%	18%	4%	8%

Tabela 4.8: Dados de reconhecimento por tom de pele

Os dados apresentam uma diferença muito acentuada entre o total de reconhecimento de faces com pele caucasiana e pele negra. Apesar desse tópico não ter figurado com objetivo específico do nosso trabalho ele chamou muito nossa atenção para o entendimento do porque desse tipo de problema. Algumas empresas como a Microsoft e a HP tem sido, a título de ilustração, alvo da imprensa escrita e via www que relatam que diversos dos seus clientes têm dificuldade em ser reconhecido por sistemas de identificação facial e o mesmo não ocorre com pessoas com tom de pele claro.

#### 4.5.5 Estimação da idade

A identificação da idade do indivíduo a partir de uma fotografia da face frontal é um problema que possui um grande relevância no que tange ao reconhecimento de pessoas desaparecidas. Várias pessoas permanecem desaparecidas durante anos o que causa modificações na rugosidade da pele ou ainda são raptadas ainda como bebês, ou na primeira infância e permanecem em cidades diferentes.

O *eigenface* foi testado para identificar não somente o reconhecimento facial, mas também a idade do indivíduo. Isso por que, como já mencionamos, a ideia básica do *eigenface* é a

busca por similaridade, então podemos configurar os pesos e seu limiar para retornar a face mais próxima do indivíduo em questão. Conhecendo a idade de cada pessoa incluída no treinamento é possível portanto, tentar identificar a idade dessa. Este teste foi realizado em conjunto com os demais e apresentou os seguintes resultados:

Idade						
Sessão	Imagens	Faces identificadas	Positivo	Falso Positivo	Acerto (%)	Erro (%)
1	359	343	79	264	23%	77%
2	160	320	80	240	25%	75%
3	160	300	80	220	27%	73%
4	116	70	21	49	30%	70%
5	145	81	27	54	33%	67%
6	149	115	31	84	27%	73%

Tabela 4.9: Dados do reconhecimento de idade utilizando imagens estáticas

Nesta etapa utilizamos o banco de face chamado de *IRABASE*, que dispõe de faces de pessoas que variam de 0 a 67 anos de idade.

#### 4.5.6 Discussão

Através dos resultados obtidos e apresentados anteriormente podemos inferir sobre a implementação e salientar algumas observações as seguintes:

##### 4.5.6.1 Experimento 1: Teste de tempo de processamento do Treinamento

Inicialmente, observamos que o modelo possui uma variação de desempenho alta quando o número de faces é aumentado significativamente. Revelando assim, que o treinamento tem um alto custo computacional, isto em parte devido a abordagem matricial o algoritmo leva muito tempo para processar a matriz que no nosso caso, utiliza o banco com mais de 4000 imagens chegamos a uma matriz de 10304x4000 para a partir daí calcularmos os autovalores e autovetores e criamos o espaço de faces. Ademais, cada *pixel* da imagem possui um valor entre 0 e 255 mas em se tratando de um banco de faces composto em sua maioria por pessoas afro-descendentes o valor dos *pixels* sempre se revelaram acima de 200 o que torna os cálculos ainda mais demorados. Realizamos 10 testes para cada faixa e resultou no gráfico da figura 4.15.

Quando aplicamos os tipos de pré-processamento propostos na sessão 4.4, verificou-se que a diferença entre imagens com e sem pré-processamento é muito pequena e não alterou de forma significativa da figura 4.15.

Outro ponto importante desta discussão é que, conforme visto na tabela 4.1 o reconhecimento melhora a partir de um número de imagens maior que 10 (LIZAMA; WALDOESTL; NICKOLAY, 1997), (ZHUIE; YU, 1994), (SILVA, 2006) mais precisamente o número ideal seriam 20 (ZHAO et al., 2003), o que tornaria uma aplicação, com um banco de dados real de pessoas desaparecidas, com alto custo de processamento, visto que existem mais de 50 mil pessoas desaparecidas.

#### *4.5.6.2 Experimento 2: Teste de reconhecimento via imagem estática*

Como mencionamos no capítulo 3 uma das dificuldades do modelo é sua sensibilidade a luminosidade. Esta deficiência ficou muito clara na tabelas 4.2 , 4.3, 4.4, que mostram que o reconhecimento sem o devido tratamento da luz incidente no indivíduo não permite que o modelo possa ser aplicado numa situação real.

Além disso, outros pontos chamaram a atenção durante o reconhecimento. Primeiramente, percebemos que o modelo procura as imagem com maior semelhança mas essa semelhança traz problemas no reconhecimento quando temos grupos de pessoas muito parecidas, fato que já havia sido notado por (ATALAY, 1996) que atribui a este problema que somente as informações mais relevantes da imagem facial são levantadas pelo modelo, porém para pessoas com alta similaridade, e a captação com pequenas nuances de luz faz com que o reconhecimento não seja efetuado com sucesso.

Comparativamente, os resultados obtidos em relação a base ORL e a base aplicada mostra um hiato muito grande entre os resultados obtidos. As imagens da base ORL foram trabalhadas para e normalizadas a uma condição de iluminação, que permite a este fator não ter influencia na sua comparação e no seu resultado, tomando-se mais eficaz. Porém, como nossa base não sofreu tal preparação o modelo provocou uma diferença de 36% na taxa de reconhecimento positivo.

Outro ponto importante do modelo implementado, é que o funcionamento mais eficiente do reconhecimento se dá ao momento em que utilizamos mais de 10 imagens de treinamento. Visto que no caso de pessoas desaparecidas essa situação é crítica, pois em geral não se tem muitas imagens da face frontal do indivíduo a ser localizado. Previmos tal situação e criamos, partir de uma única imagem, outras 9.

### 4.5.6.3 *Teste de reconhecimento via imagem dinâmica*

O processo de reconhecimento usando vídeos mostrou-se bastante instável e com dificuldade no reconhecimento, principalmente devido a angulação da câmera, que resulta numa pose, em geral, muito diferente da inserida no banco de treinamento. O modelo procura similaridades na imagem e com isso, pequenas nuances na mesma imagem produz um resultado positivo e falso-positivo com diferenças entre 4 a 5 frames. Ficou muito claro que para este tipo de entrada os resultados não podem ser em termos de precisão mas sim do mais próximo ou mais similar, nesta abordagem seria possível melhorar o desempenho do modelo, porém não testamos essa situação. Apesar de não constar esta análise em termos quantitativos, foi percebido um problema do modelo ao lidar com tons de pele não caucasianos.

Este problema será abordado no item a seguir mas também foi visto acontecer tanto na detecção como no reconhecimento, além disso todas as imagens (8%) que retornaram positivo são de pessoas caucasianas. Em nível de *pixel* isso quer dizer que Images frontais da face que vairam de 9 a 169, não possuem nem uma detecção nem um reconhecimento satisfatório.

Este problema parece mais complexo de ser resolvido do que uma imagem fixa (BAE; KIM, 2005), (OLIVER; PENTLAND; BEHRARD, 2000), principalmente devido ao fluxo de informações que chegam para ser processado. Dessa discussão, e comparação entre o total de reconhecimentos positivos entre as captação da imagem por foto ou vídeo inicialmente a última parece favorecer, de acordo com os dados apresentados, a localização de pessoas desaparecidas, até por que as uma câmera pode funcionar perfeitamente com uma máquina fotográfica.

O reconhecimento falha quando a imagem captada está a mais de 1,5m de distância da fonte de captação e quando o indivíduo está em movimentos bruscos tanto a detecção e o reconhecimento falham. O que não ocorre com frequência quando o indivíduo está próximo da câmera, além disso para o reconhecimento efetivo, observado neste trabalho, a imagem captada deve estar em termos de pose e iluminação muito próxima da imagem que foi armazenada na fase de treinamento.

Nas imagens de vídeo características extremas, disfarces e até mesmo fotografias de pessoas na cena são detectados e retornam falsos positivos, situação que permite que um indivíduo mal intencionado leve uma fotografia de uma pessoas desaparecida e engane o algoritmo.

Ser parecido, o que chamamos de similaridade, mostrou-se um problema durante a análise dos dados obtidos. Pessoas que são muito parecidas retornam autovalores muito próximos

e portanto, ligeiras modificações na pose confundem o modelo, também notamos que algumas imagens do banco de treinamento parecem concentrar mais energia e quase sempre são retornadas como falso positivo. Ao serem retiradas do banco de treino, outra imagem assume este papel e assim sucessivamente.

#### 4.5.6.4 *Reconhecimento - tom de pele*

Nossa ideia básica não era estudar a influencia do tom da pele no reconhecimento, mais devido a um aos resultados obtidos na tabela 4.8, ficou claro que existe alguma influencia do tom de pele no equilíbrio dos autovalores. Testamos o modelo, inclusive com apenas pessoas negras compondo a base e não obtivemos taxas de reconhecimento superiores, o que não ocorreu com pessoas de tom de pele caucasiano treinadas para reconhecer pessoas com o mesmo tom de pele.

Caso os indivíduos tenham um tom de pele negro e com algum tipo de disfarce a taxa de reconhecimento é quase zero. Percebemos que a iluminação perpendicular no indivíduo com esse tom de pele melhora significativamente o reconhecimento e alguns filtros também facilitam o trabalho. O reconhecimento não foi efetivo e teve taxas próximas de zero quando utilizamos a câmera no modo vídeo sendo que as taxas positivas só foram obtidas quando utilizamos imagens estáticas e com algum filtro.

#### 4.5.6.5 *Estimar da Idade*

Reconhecer a idade do indivíduo é uma área que, cada vez mais, faz parte do reconhecimento facial e de aplicações atuais, figura 4.25. Ter essa informação pode vir a ser crucial para diversos segmentos empresariais que podem focar sua publicidade naquela faixa etária.

Percebemos algumas dificuldades no reconhecimento, entre elas a face estar voltada completamente para a frente da câmera, pois ângulos laterais tornam o processo mais complicado. Aspectos relacionados a etnia, sexo ou emoções durante a captação também dificultam o processo. As respostas positivas foram calibradas em 2 anos para mais ou para menos.

O modelo compara as características faciais e traços marcantes são detectados com mais facilidade, por exemplo rugas ao redor do olhos, estrutura óssea, e pele flácida. Evidentemente, todo sistema biométrico está sujeito da a falhas mas o japão já possui uma máquina que vende cigarros e identifica menores de 20 anos.



Figura 4.25: Máquina de venda de cigarros com reconhecimento de idade

Como utilizamos uma base que possui imagens de iranianos e não de brasileiros, mas especificadamente baianos, o modelo teve uma certa dificuldade em ajustar o limiar, com alternativa sugerimos que sejam coletadas imagens da face de indivíduos baianos para que os testes tenham uma maior robustês.

## 4.6 Considerações Finais

Algumas dificuldades se impuseram ao modelo em sua implementação, iluminação, posição da face, alterações na composição facial, barba, cabelos e até mesmo o tom da pele. Apesar destes sobressaltos foi possível analisar o comportamento do modelo com relação a sua performance e resultados de reconhecimento. Centramos nosso esforços inicialmente na construção do banco de dados, que contou com a colaboração de indivíduos com idade variando entre 14 e 37 anos, com diferentes tons de pele e sem preocupação com a iluminação do local da captura. Para a tarefa de detecção da face na imagem escolhemos o padrão de detecção desenvolvido por Viola e Jones que revelou alguns problemas na detecção, como o resultado de face em imagens, que não representam face ou simplesmente, a não detecção da face num ambiente com inúmeras faces. Após a detecção da face aplicamos métodos de melhoramento da imagem que buscaram acentuar as taxas de reconhecimento. Os métodos visaram, de maneira geral, ressaltar a região da face e como os métodos trabalham diretamente sobre os *pixels* da imagem, equalizaram a intensidade de luz e quando possível distinguir a face dos ruídos provenientes da imagem. Procuramos detalhar os resultados obtidos em cada etapa, descrevendo-os por meio de tabelas e gráficos e percebemos principalmente que a força do modelo *eigenface*, não está

na exatidão, mas sim na similaridade. Todos os testes foram conduzidos buscando sempre analisar os resultados do ponto de vista do nosso recorte.

---

## Conclusão e Trabalhos Futuros

---

Neste trabalho tivemos como objetivo provar a eficiência de um método estatístico de reconhecimento de face aplicado a uma situação não controlada de detecção e reconhecimento de pessoas a partir de uma ou de um conjunto de imagens. Durante as próximas sessões mostraremos que este método se revelou frágil quando expostos a situações não controladas, apresenta lentidão em relação ao tempo de treinamento, alto número de falsos-positivos para imagens dinâmicas e razoável taxa de acerto para imagens estáticas. Também revelou dificuldades com relação a mudanças físicas (idade) e tom de pele.

### 5.1 Conclusões

Como os resultados obtidos nos testes realizados foram atingidos dentro do contexto no qual este modelo está inserido tivemos resultados que destoam das pesquisas analisadas durante o levantamento bibliográfico que fizemos, isso quer dizer que o modelo foi exposto a um ambiente carregado de elementos perturbadores.

Inicialmente, o custo computacional (tempo de processamento para o hardware utilizado) adjacente ao processo é alto, o fato de utilizar uma abordagem que analisa cada pixel das figuras, mesmo em conjunto causa uma redução da performance, conforme vimos na no gráfico 4.14 e 4.15, verificamos que o tempo total para a geração da matriz de treinamento e os autovalores revele-se o grande responsável por este alto custo

Outro ponto que detectamos, é que o número de faces mínimas que o banco de dados deve possuir para que ocorra reconhecimento positivo acima com a utilização do filtro linear espacial é da ordem de 15 imagens da face frontal e sem filtro acima de 18. O que se torna um impedimento para o caso de apenas uma única imagem de entrada. As taxas de falso positivo para uma única imagem e a o número ideal de imagens pode ser visto nas figuras 4.21 e 4.22 o que torna muito difícil o processamento de uma aplicação baseada neste modelo. Esta performance revelou-se portanto precária se pensarmos que existem mais de 50 mil pessoas desaparecidas no Brasil.

Os testes dos níveis de reconhecimento foram realizados com sucesso de 64%, para reconhecimento positivo em imagens estáticas e de apenas 8%, em imagens em vídeo o que revela que o algoritmo funciona melhor, dentro do nosso universo, em imagens estáticas. Isso se deu devido a vários fatores particulares ao nosso experimento, como a maior parte dos indivíduos da pesquisa estão na faixa etária de 12-19 anos de idade não se mostram



muito a vontade com filmagens, o que resultou em um grande número de imagens extremamente diferentes da encontradas no espaço de faces. Já com relação a imagens estáticas os jovens aceitaram mais tranquilamente este meio e as poses em geral estão em posição frontal.

Quanto a aceitabilidade do modelo, a literatura trás (CHELLAPPA; WILSON; SIROHEY, 1995), que o reconhecimento facial é extremamente não invasivo, porém para a realização dos testes muitos colaboradores não aceitaram muito bem o fato de serem fotografados ou filmados, mesmo em momentos em que sabiam que isto iria acontecer. Inicialmente imaginamos um banco de faces com cerca de 500 pessoas e só conseguimos a autorização para fotos e vídeos de apenas 134, ou seja em torno de 26% das pessoas aceitaram bem a proposta. Ainda, como vimos na tabela 4.7 o modelo é sensível a fraude, comparativamente a taxa que obtivemos de positivos sem disfarce a taxa obtida traduz uma certa fragilidade.

Fazendo um comparativo entre a tabela 2.2 e os dados que obtivemos podemos refazer esta tabela 5.1

Face	Universalidade	Distinção	Permanência	Mensuração	Performance	Aceitabilidade	Fraude
Literatura	A	B	M	A	B	A	A
Trabalho	M	B	B	M	B	M	B

Tabela 5.1: Comparativo entre a tabela de (JAIN; ROSS; PRABHAKAR, 2004) e nossa análise

Pessoas envelhecem e o estresse de viver longe dos familiares faz aparecer rugas e outras deformações na pele, além é claro de queda dos cabelos e cicatrizes de violência. Neste sentido fizemos uma análise se a mesma implementação seria capaz de analisar a idade do indivíduo. Nossos testes mostraram que o ambiente é promissor, mas necessita de ajustes devido ao seu grande numero de erros, tabela 4.9, estes ajustes de referem ao banco de face, utilizamos um banco gratuito disponível na *web*, porém cada sociedade possui características de envelhecimento e juventude muito diferentes, o que tornou um pouco complexo, por exemplo, o banco não possui pessoas negras.

Realizamos também, testes para verificar se o modelo é sensível ao tom de pele, imaginávamos que este dado seria irrelevante, mas não foi o que os testes mostraram. Existe uma certa sensibilidade do modelo a tons de pele escura, mesmo com a incidência de luz artificial. Mais estudos se fazem necessários, contudo os resultados preliminares retornaram reconhecimentos positivos na casa dos 4% o que distoa imensamente dos resultados obtidos ( em video não houve reconhecimento positivo em nenhuma situação para tons de pele escura). O problema pode ser relacionado a como o modelo enxerga o tom de pele

negro, ou seja, a quantidade de energia luminosa e provoque assim classificadores fracos. Tal suposição carece de maior análise e fica como trabalho futuro.

Com base em todas as percepções que tivemos dos resultados obtidos, e a luz da literatura que embasou este trabalho, concluímos que o modelo não é eficiente para as situações aos quais foi exposto, além disso, o reconhecimento não pode ser realizado com sucesso mínimo sem um número razoável de imagens de treinamento, como vimos na figura 4.21 acima de 10 imagens.

## 5.2 Contribuições e Trabalhos Futuros

Devido a complexidade do problema de reconhecimento de pessoas a partir de imagens ou sequências de vídeo perante a atual tecnologia disponível, vislumbramos que há muito trabalho a ser feito. Durante o desenvolvimento dessa dissertação passos foram dados no caminho desse da solução deste problema, mas restam várias tarefas a serem desenvolvidas futuramente.

Além dos resultados apresentados, este trabalho traz as seguintes contribuições:

- Descrição de algoritmos de amenização de iluminação e normalização de imagens encontrados na literatura, que se mostraram com razoável nível de dificuldade para implementação;
- mostramos que o método é sensível aos tons de peles mais claros que escuros;
- evidenciou-se que nem sempre as imagens que são melhores para os seres humanos reconhecerem são as melhores para o modelo;
- uma introdução teórica tanto ao *eigenface* como ao estudo do pré-processamento de imagens digitais.
- avaliação dos parâmetros e interação dos subespaços do PCA na obtenção das taxas de reconhecimento.
- Analisar se o problema da pele negra pode estar associado as aproximações lineares feitas e testar aproximações não lineares.
- Utilizar outras médias e pesos e comparar aos resultados obtidos.

### 5.2.1 *Trabalhos futuros*

A partir dos nossos objetivos vários outros estudos podem ser realizados visando sua complementação e o desenvolvimento de um modelo mais aplicado no contexto real. A seguir são apresentadas algumas sugestões de tópicos que podem ser abordados em estudos futuros:

- Desenvolver um trabalho cooperativo com as órgãos governamentais para aplicar o modelo em um banco de dados padronizado visando fazes buscas por similaridade.
- Integrar o modelo proposto com outros métodos biométricos para torna-lo mais eficaz e aumentar a taxa de reconhecimento positivo.
- Melhorar o tempo de processamento do treinamento utilizando outros meios matemáticos;
- Formalizar um modelo matemático que defina parâmetros de ótimo no reconhecimento facial.
- Estudar a relação entre o tom de pele negro ou mulato em relação ao reconhecimento.
- Aplicar e ajustar o *eigenface* ao problema de detecção de idade.
- Investigar formas de extrair informações obtidas exclusivamente do movimento de faces.
- Acentuar o reconhecimento em sequências de vídeo independente da iluminação.
- Construir um modelo que transforme a imagem em um sinal, no sentido de otimizar os calculos.

Um modelo de reconhecimento deve apresentar bons resultados isso quer dizer taxas de positivo acima de 95%, para que seja considerado eficiente e portanto garantir que seja utilizado, largamente com meio de solucionar problemas, como o nosso. Dai necessitamos de um estudo criterioso do espalhamento dos vetores de face no espaço altamente dimensional para ajudar então o modelo a ter taxas em imagens e vídeos sendo um caminho e através dos presupostos discutidos no capítulo 2, sessão 2.4.1.

---

## Referências Bibliográficas

---

- ALMEIDA, O. C. P. d. *Técnicas de processamento de imagens para localização e reconhecimento de faces*. Tese (Doutorado) — Mestrado em Ciência da Computação, ICMC-USP São Carlos, 2006.
- ATALAY, I. *FACE RECOGNITION USING EIGENFACES*. Dissertação (Mestrado) — Computer Engineer B.Sc., 1996.
- BAE, H.; KIM, S. Real-time face detection and recognition using hybrid-information extracted from face space and facial features. *Image and Vision Computing*, v. 23, p. 1181–1191, 2005.
- BASTANFARD, A. et al. Iranian face database with age, pose and expression . *International Conference on Machine Vision*, v. 1, p. 50–55, 2007.
- BELHUMEUR, P.; HESPANHA, J.; KRIEGMAN, D. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 19, p. 711–720, 1997.
- BERNSTEIN, M. J.; YOUNG, S. G.; HUGENBERG, K. The cross-category effect: Mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science*, v. 18, p. 706–712, 2007.
- BRADSKI, G.; KEAHLER, A. *OpenCV - Computer Vision with the OpenCV Library*. [S.l.]: O’Reilly Media Inc., 2008.
- CAMPOS, T. E. d. *Técnicas de Seleção de Características com Aplicações em Reconhecimento de Faces*. Dissertação (Mestrado) — Instituto de Matemática e Estatística da Universidade de São Paulo, 2001.
- CHELLAPPA, R.; WILSON, C. L.; SIROHEY, S. Human and machine recognition of faces: a survey. *IEEE Transactions on pattern Analysis and Machine Intelligence*, v. 83, n. 5, p. 705–741, maio 1995.
- CHICHIZOLA, F.; GIUSTI, L. D.; GIUSTI, A. D.; NAIIOUF, M. Face recognition: reduced image eigenfaces method. In: *Proc. 47th International Symposium ELMAR*. [S.l.: s.n.], 2005. p. 159–162.
- DANTAS, C. A. B. *PROBABILIDADE: Um Curso Introdutório*. 2. ed 1.reimpr. ed. São Paulo: [s.n.], 2004. (Acadêmica, 10).
- DELAC, K.; GRGIC, M. A survey of biometric recognition methods. *46th International Symposium Electronics in Marine*, v. 46, p. 16–18, 2004.

- DESCONHECIDO. *Biometric Data Interchange Formats - Part 5: Face Image Data*,. [S.l.], 2004.
- EL-BAKRY, H. M.; HAMADA, M. Fast principal component analysis for face detection using cross-correlation and image decomposition. In: *Proc. International Joint Conference on Neural Networks IJCNN 2009*. [S.l.: s.n.], 2009. p. 751–756.
- ELAD, M.; GOLDENBERG, R.; KIMMEL, R. Low bit-rate compression of facial images. *IEEE Trans. on Image Processing*, v. 16, n. 9, p. 2379–2383, 2007.
- EPHRAIM, Y.; MERHAV, N. Hidden markov processes. *IEEE Trans. Inform. Theory*, p. 1518–1569, 2002.
- ETEMAD, K.; CHELLAPPA, R. Discriminant analysis for recognition of human face images. *Lecture Notes in Computer Science*, v. 1206, p. 127–142, 1997.
- ETHEVE, R.; RICCIO, D.; DUGELAY, J.-L. State of the art in 3d face recognition. *Eurocom RR 06 160*, p. 14–31, 2006.
- FERIS, R. S.; KRUEGER, V.; JR., R. M. C. A wavelet subspace method for real-time face tracking. *science direct*, v. 10, p. 339–350, 2004.
- FOWLER, J. E. Compressive-projection principal component analysis. *IEEE Transactions on pattern Analysis and Machine Intelligence*, v. 18, n. 10, p. 2230–2242, out. 2009.
- GAO, W.; ZHOU, Z.-H. Approximation stability and boosting. *Algorithmic Learning Theory*, v. 21, p. 59–73, 2010.
- GOLFARELLI, M.; MAIO, D.; MALTONI, D. On the error-reject tradeoff in biometric verification systems. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, v. 19, n. 7, p. 786–796, Julho 1997.
- GOMES, J. M.; VELHO, L. C. *Conceitos Básicos de Computação Gráfica*. São Paulo: [s.n.], 1990. 311p p.
- GONZALEZ, R. C.; WOODS, R. E.; EDDINS, S. L. *Digital Image Processing using MATLAB*. New Jersey-USA: [s.n.], 2004.
- HAIR, J. F.; TATHAM, R. L.; ANDERSON, R. E.; BLACK, W. *Análise Multivariada de Dados*. [S.l.]: Bookman, 2007.
- HILBE, J. M. *Logistic Regression Models*. [S.l.]: Chapman & Hall/CRC, 2009.
- HJELMAS ERIK; LOW, B. K. Face detection: Asurvey. *Computer Vision and Image Undertanding*, v. 83', n. 83, p. 236–274, 2001.
- H.K., E.; SANKUR, B. Multiresolution face recognition, image and vision computing. *Issue 5*, v. 23, p. 469–477, 2005.

- HORTON, M.; CAMERON-JONES, M.; WILLIAMS, R. Multiple classifier object detection with confidence measures. *advances in artificial intelligence*, v. 20, p. 559–568, 2007.
- HOTELLING, H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, v. 24, p. 417–441, 498–520, 1933.
- HUANG, X.; JACK, M.; ARIKI, Y. *Hidden Markov Models for Speech Recognition*. [S.l.]: Edinburgh University Press, 1990.
- HUBEL, D. *Eye, Brain and Vision*. [S.l.]: W. H. Freeman and Company, 1989. (Scientific American Library, 22).
- JAIN, K. A.; ROSS, A.; PRABHAKAR, S. An introduction to biometric recognition. *Appeared in IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Image- and Video-Based Biometrics*, v. 14, n. 1, January 2004.
- JENKINS, R. Limitations in face identification: The evidence. *FEATURE OF THE WEEK*, p. 5–12, 2008.
- JOLLIFFE, I. *Principal component analysis*. [S.l.]: Springer verlag, 2002.
- KHAN, A. S.; ALIZAI, L. K. Introduction to face detection using eigenfaces. In: *Proc. International Conference on Emerging Technologies ICET '06*. [S.l.: s.n.], 2006. p. 128–132.
- KIRBY, M.; SIROVICH, L. Application of the karhunen-loeve procedure for the characterization of human faces. *IEEE Transactions on pattern Analysis and Machine Intelligence*, v. 12, p. 103–108, 1990.
- KLECKA, W. R. *Discriminant analysis*. 7-19. ed. [S.l.]: SAGE, 1980. 71 p.
- LAKSHMANAN, S.; WATTA, P.; HOU, Y. L.; GANDHI, N. Comparison between eigenfaces and fisherfaces for estimating driver pose. In: *Proc. IEEE Intelligent Transportation Systems*. [S.l.: s.n.], 2001. p. 889–894.
- LEVIN, D. T. Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition. *Journal of Experimental Psychology: General*, v. 129, p. 559–574, 2000.
- LIAU, H. F.; SENG, K. P.; ANG, L.-M.; CHIN, S. W. New parallel models for face recognition. *Recent Advances in Face Recognition*, v. 1, p. 15–26, 2008.
- LIU, N.; WANG, H.; YAU, W.-Y. Face recognition with weighted kernel principal component analysis. In: *Proc. 9th International Conference on Control, Automation, Robotics and Vision ICARCV '06*. [S.l.: s.n.], 2006. p. 1–5.

- LIZAMA, E.; WALDOESTL, D.; NICKOLAY, B. An eigenfaces-based automatic face recognition system. In: *Proc. IEEE International Conference on Systems, Man, and Cybernetics 'Computational Cybernetics and Simulation'*. [S.l.: s.n.], 1997. v. 1, p. 174–177.
- MAIO, D.; MALTONI, D.; CAPPELLI, R.; WAYMAN, J. L.; JAIN, A. K. Fvc2002: Fingerprint verification competition. *Proc. International Conference on Pattern Recognition (ICPR)*, p. 744–747, Agosto 2002. Quebec City, Canada.
- MALHOTRA, N. K. *PESQUISA DE MARKETING: Uma orientação aplicada*. 4ª edição. ed. São Paulo: [s.n.], 2004.
- MENDONÇA, M. M. *Comparação entre métodos de normalização da iluminação utilizados para melhorar a taxa de reconhecimento facial*. Dissertação (Mestrado) — Escola de Engenharia de São Carlos, 2008.
- MINGOTI, S. A. *ANALISE DE DADOS ATRAVÉS DE MÉTODOS DE ESTATÍSTICA MULTIVARIADA: Uma abordagem Aplicada*. Belo Horizonte: [s.n.], 2005. 297p p.
- MINSKY, M. *The Society of Mind*. [S.l.]: Simon and Schuster, 1986.
- MONWAR, M. M.; PAUL, P. P.; ISLAM, M. W.; REZAEI, S. A real-time face recognition approach from video sequence using skin color model and eigenface method. p. 2181–2185, maio 2006.
- MYNBAEV, K. T. *MANUAL DE ECONOMETRIA*. Rio de Janeiro: [s.n.], 2004. 348p p.
- NOCK, R.; NIELSEN, F. A real generalization of discrete adaboost. *17th European Conference on Artificial Intelligence*, v. 17, p. 509–515, 2006.
- OLIVEIRA, D. D. d. *Desaparecidos civis: conflitos familiares, institucionais e segurança públicas*. Tese (Doutorado) — Instituto de ciências sociais, Programa de pós graduação em sociologia, 2007.
- OLIVER, N.; PENTLAND, A.; BEHRARD, F. Lafter: a real-time face and lips tracker with facial expression recognition. *The Journal of the pattern recognition society*, v. 33, p. 1369–1382, 2000.
- PARENTE, a. *Imagem-maquina: a era das tecnologias do virtual*. Rio de janeiro: [s.n.], 1993. 304p p. (Coleção TRANS). Tradução de Rogério Luz et all.
- PEARSON, K. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, v. 2, p. 559–572, 1901.
- PERLIBAKAS, V. Recognition of expression variant faces using masked log-gabor features and principal component analysis. *IEEE Trans. on Image Processing*, v. 1, p. 06–27, 2008.

- PRABHAKAR, S. *Fingerprint classification and matching using a filterbank*. Tese (Doutorado) — Computer Science & Engineering, Michigan State University, East Lansing, MI, USA, 2001. Adviser-Jain, Anil K.
- PUYATI, W.; WALAIRACHT, A. Efficiency improvement for unconstrained face recognition by weighting probability values of modular pca and wavelet pca. In: *Proc. 10th International Conference on Advanced Communication Technology ICACT 2008*. [S.l.: s.n.], 2008. v. 2, p. 1449–1453.
- RABINER, L. R. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE 77*, p. 257–286, 1989.
- ROSOT, M. A. D.; DISPERATI, A. A.; FILHO, A. F. Medição da area de anéis de crescimento usando imagens digitais. *Revista Árvore/SIF*, v. 1, 2001.
- ROSÁRIO, J. M. *Princípios de Mecatronica*. [S.l.]: Prentice Hall, 2005.
- SANTOS, A. *Gerenciamento de identidades*. [S.l.]: Brasport, 2007. 26-30 p.
- SANTOS, A. R. d. *Identificação de Faces humanas através de PCA-LDA e Redes Neurais SOM*. Dissertação (Dissertação de Mestrado) — Escola de Engenharia de São Carlos da Universidade de São Paulo, 2005.
- SAVRAN, A. et al. Biometrics and identity management. In: SCHOUTEN, B.; JUUL, N. C.; DRYGAJLO, A.; TISTARELLI, M. (Ed.). Berlin, Heidelberg: Springer-Verlag, 2008. cap. Bosphorus Database for 3D Face Analysis, p. 47–56. ISBN 978-3-540-89990-7. Disponível em: <[http://cost2101.org/BIOID2008/papers/6\\_db\\_bioid08.pdf](http://cost2101.org/BIOID2008/papers/6_db_bioid08.pdf)>.
- SCHAPIRE, R. E. The strength of weak learnability. *Machine Learning*, v. 5, p. 197–227, 1990.
- SILVA, A. F. d. *Reconhecimento de Faces via PCA: Análise de Desempenho*. Dissertação (Dissertação) — Programa de Pós-Graduação da Faculdade de Computação da Universidade Federal de Uberlândia, 2006.
- SIROVICH, L.; KIRBY, M. Low-dimensional procedure for the characterization of human faces. *journal of the Optical Society of America*, v. 4(3), p. 519–524, 1987.
- TAN, X.; SONGCAN, C.; ZHI-HUA, Z.; FUYAN, Z. Face recognition from a single image per person: A survey. *PATTERN RECOGNITION, The Journal of the pattern recognition society*, v. 39, n. 39, p. 1725–1745, 2006.
- TRIGO, T. *EQUIPAMENTO FOTOGRÁFICO*. [S.l.]: EDITORA SENAC São Paulo, 1998.
- TURK, M. *Interactive-Time Vision: Face Recognition as a Visual Behavior*. Tese (Doutorado) — The Media Laboratory, Massachusetts Institute of Technology, 1991.



- TURK, M. Eigenfaces and beyond. *Face Processing: Advanced Modeling and Methods*, 2005.
- TURK, M.; PENTLAND, A. Face recognition without features. *Proc. IAPR Workshop on Machine Vision Applications*, p. 267–270, 1990.
- TURK, M.; PENTLAND, A. Eigenface for recognition. *J. Cognitive Neuroscience* 3(1), p. 71–96, 1991.
- TURK, M. A.; PENTLAND, A. P. Face recognition using eigenfaces. *Proc. CVPR '91. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, p. 586–591, jun. 3–6, 1991.
- VIOLA, P.; JONES, M. J. Robust real-time face detection. *International Journal of Computer Vision*, v. 2, n. 57, p. 137–154, 2004.
- WAYMAN, J. L. Fundamentals of biometric authentication technologies. *International Journal of Image and Graphics*, v. 1, n. 1, p. 93–113, 2001.
- WAZLAWICK, R. S. *Metodologia de Pesquisa para Ciência da Computação*. Rio de Janeiro: ELSEVIER, 2008.
- WEEKS, M. *Digital Signal Processing: Using MATLAB and Wavelets*. 1. ed. Hingham, Massachusetts: Infinity Science Press LLC, 2007.
- WOODWARD, J. D. J.; HORN, C.; GATUNE, J.; THOMAS, A. *Biometrics: A Look at Facial Recognition*. [S.l.]: RAND, 2003.
- YANG, M.-H.; AHUJA, N.; KRIEGMAN, D. Face recognition using kernel eigenfaces. In: *Proc. International Conference on Image Processing*. [S.l.: s.n.], 2000. v. 1, p. 37–40.
- ZHANG, D. *Automated biometrics: technologies and systems*. [S.l.]: Kluwer Academic Publishers, 2000.
- ZHAO, F.; CAI, C.; ZHU, J. Weed seeds recognition using color pca. In: *Proc. Second International Symposium on Knowledge Acquisition and Modeling KAM '09*. [S.l.: s.n.], 2009. v. 1, p. 375–378.
- ZHAO, L.; ZHUANG, G.; XU, X. Facial expression recognition based on pca and nmf. In: *Proc. 7th World Congress on Intelligent Control and Automation WCICA 2008*. [S.l.: s.n.], 2008. p. 6826–6829.
- ZHAO, W.; CHELLAPPA, R.; PHILLIPS, P.; A., R. Face recognition: A literature survey. *ACM Computing Surveys*, v. 35, n. 4, p. 399–458, Dezembro 2003.
- ZHAO, W.; CHELLAPPA, R.; PHILLIPS, P. J. Subspace linear discriminant analysis to face recognition. *Partially supported by the Office of Naval Research*, 1999.

ZHAO, Y.; SHEN, X.; GEORGANAS, N. D.; PETRIU, E. M. Part-based pca for facial feature extraction and classification. In: *Proc. IEEE International Workshop on Haptic Audio visual Environments and Games HAVE 2009*. [S.l.: s.n.], 2009. p. 99–104.

ZHOU, S. K.; CHELLAPPA, R.; ZHAO, W. *UNCONSTRAINED FACE RECOGNITION*. [S.l.]: Springer, 2006. 244p p. (International Series on Biometrics, XII).

ZHUJIE; YU, Y. L. Face recognition with eigenfaces. In: *Proc. IEEE International Conference on Industrial Technology*. [S.l.: s.n.], 1994. p. 434–438.

*Reconhecimento de faces aplicado ao problema de pessoas desaparecidas - Estudo de caso do Eigenface*

Marcos Batista Figueredo

Salvador, Março de 2011.