



**JAIME GOMES DA SILVA NETO**

**MÉTODO DE RECONHECIMENTO DE FACE  
BASEADO EM ESTATÍSTICAS DE ORDEM  
SUPERIOR**

**LAVRAS-MG  
2016**



**JAIME GOMES DA SILVA NETO**

**MÉTODO DE RECONHECIMENTO DE FACE BASEADO EM  
ESTATÍSTICAS DE ORDEM SUPERIOR**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Engenharia de Sistemas e Automação, área de concentração em Engenharia de Sistemas e Automação, para a obtenção do título de Mestre.

**Prof. Dr. Danton Diego Ferreira  
Orientador**

**LAVRAS-MG  
2016**

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca  
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Silva Neto, Jaime Gomes da.

Método de reconhecimento de face baseado em estatísticas de ordem superior / Jaime Gomes da Silva Neto. – Lavras : UFLA, 2016.

72 p. : il.

Dissertação (mestrado acadêmico)–Universidade Federal de Lavras, 2016.

Orientador(a): Danton Diego Ferreira.

Bibliografia.

1. Reconhecimento Facial. 2. Estatística de Ordem Superior. 3. Cumulantes. 4. Discriminante Linear de Fisher. 5. Classificador Bayesiano. I. Universidade Federal de Lavras. II. Título.

**JAIME GOMES DA SILVA NETO**

**MÉTODO DE RECONHECIMENTO DE FACE BASEADO EM  
ESTATÍSTICAS DE ORDEM SUPERIOR**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Engenharia de Sistemas e Automação, área de concentração em Engenharia de Sistemas e Automação, para a obtenção do título de Mestre.

APROVADA EM 18 DE ABRIL DE 2016.

Dr. Erivelton Geraldo Nepomuceno	UFSJ
Dr. Giovani Bernardes Vítor	UTC
Dr. Belisário Nina Huallpa	UFLA
Dr. Fábio Domingues de Jesus	UFLA

**Prof. Dr. Danton Diego Ferreira  
Orientador**

**LAVRAS/MG  
2016**

*Aos meus pais Cássio e Euriz, pois sem eles, nada seria possível e não estaríamos aqui reunidos desfrutando, juntos, destes momentos que nos são tão importantes.*

*Aos meus irmãos Pedro, Júnior e Ana pelo esforço, dedicação e compreensão, em todos os momentos desta e de outras caminhadas.*

*Em especial, a minha esposa Aline pelo apoio, amor e carinho a mim dedicados.*

*Aos meus filhos Heitor e Maria Luísa pela compreensão pelos momentos ausentes.*

*DEDICO*

## **AGRADECIMENTOS**

Tudo o que aprendemos a fazer, somente aprendemos quando realmente fazemos. Embora uma dissertação seja pela sua finalidade acadêmica um trabalho individual, há contribuições de natureza diversa que não podem nem devem deixar de ser realçadas. Por essa razão, é difícil criar uma lista contendo todas as pessoas que me apoiaram e contribuíram direta ou indiretamente para o desenvolvimento desta dissertação. No momento eu gostaria de ressaltar alguns colaboradores importantes e me desculpar por todos os nomes que eu omiti neste momento, mas que estão presentes de alguma forma nas entrelinhas desse projeto de mestrado.

Início meus agradecimentos a Deus, pela oportunidade de fazer o mestrado e pela força e sabedoria que me auxiliaram durante os meus estudos.

Agradeço ao meu orientador, Prof. Dr. Danton Diego Ferreira, que foi um verdadeiro mestre nesta minha caminhada, professor de grande saber, pelo constante incentivo sempre indicando à direção a ser tomada nos momentos de maior dificuldade, apoiando as minhas decisões nos momentos de afloração intelectual, e por despertar em mim o interesse pela área de Reconhecimento de Padrões. Agradeço, principalmente, pela confiança depositada no meu trabalho de mestrado.

Agradeço aos meus familiares – pai, mãe e irmãos, por me apoiarem e incentivarem diariamente. A minha esposa e filhos por me oferecer equilíbrio e entender a minha ausência nos momentos especiais de nossas vidas.

Agradeço aos colegas da Polícia Militar de Minas Gerais por creditar em mim a confiança para a realização desta pesquisa.

**MUITO OBRIGADO!**

## RESUMO

A identificação de uma pessoa pela sua face é um dos métodos não intrusivos mais efetivos em biometria. Por outro lado, é um dos maiores desafios para os pesquisadores na área, visto que implica em fatores que contemplam os aspectos de iluminação do ambiente, pose do indivíduo, qualidade da imagem, oclusão, disfarces, dentre outros. Os sistemas de reconhecimento facial apresentam vasta aplicabilidade, com destaque para os sistemas de segurança, realizando importante tarefa na sociedade. É chave de acesso para sistemas e locais como, por exemplo, computadores pessoais, smartphones, acesso a ambientes específicos do sistema bancário, dentre outros sistemas vinculados à interface homem-máquina. A presente dissertação de mestrado apresenta contribuições em dois aspectos: (i) explora as estatísticas de ordem superior para construção de assinaturas compactas de faces; (ii) considera um cenário cujo objetivo é detectar e identificar criminosos automaticamente com reconhecimento de face para auxiliar a Polícia Militar. O algoritmo proposto nesta dissertação para reconhecimento de faces foi desenvolvido em três etapas. Na primeira etapa foi realizada a extração de características utilizando Estatística de Ordem Superior (cumulantes de 2ª, 3ª e 4ª ordens). Na etapa seguinte, selecionaram-se os parâmetros por meio do discriminante de Fisher e análise de redundância com correlação linear. Na última fase, foi feita a classificação, utilizando o classificador de Bayes. Para verificar o desempenho do algoritmo de reconhecimento de faces proposto neste trabalho foram realizados testes utilizando o banco de imagens bem conhecido na área de processamento de imagens: ORL. Os resultados são motivadores, com taxas de detecção e classificação superiores a 70%, e evidenciam o potencial das estatísticas de ordem superior na construção de assinaturas de faces.

**Palavras-Chave:** Reconhecimento Facial. Estatística de Ordem Superior. Cumulantes. Discriminante Linear de Fisher. Classificador Bayesiano.

## ABSTRACT

Face recognition is one of the most effective non-intrusive methods in biometrics. On the other hand, it is a major challenge for researchers in the area, as it involves factors that include the ambient lighting aspects, the individual pose, image quality, occlusion, disguises, among others. Face recognition systems have broad applicability, especially the security systems, performing important task in society. It is key access to systems and locations, for example, personal computers, smartphones, access to specific rooms of the banking system, among other systems linked to human-machine interface. This master's thesis presents contributions in two aspects: (i) it explores the hither-order statistics to build compact signature of faces; (ii) it considers a scenario whose the goal is to detect and identify criminals automatically with face recognition to assist the military police. The algorithm proposed in this thesis for face recognition was developed in three stages. In the first stage the feature extraction using higher-order statistics (second-, third- and fourth-order cumulants) is performed. The next step comprises the feature selection, through the Fisher's discriminant, and redundancy analysis with linear correlation. In the last stage, the classification using the Bayes classifier is performed. To check the performance of face recognition algorithm proposed in this work it was carried out tests using the database ORL, which is a well-known dataset in the image processing area. Promising results were achieved in which detection and classification rates over 70% were reached, which shows the potential of higher-order statistics on building compact feature vector signatures of faces.

**Keywords:** Facial Recognition. Higher-Order Statistics. Cumulants. Fisher's Discriminant Ratio. Bayesian Classifier.



## LISTA DE FIGURAS

Figura 2.1 - Exemplo da conversão de uma face de dimensão 112x92 em um vetor coluna de 10.304 posições.....	25
Figura 2.2 - Arquitetura-I.....	28
Figura 2.3 - Arquitetura-I (figura acima): bases refletem características mais locais; Arquitetura-II (figura abaixo): bases refletem características mais globais.....	29
Figura 2.4 - Modelo não linear de um neurônio artificial.....	31
Figura 3.1 - Imagens do banco de dados de faces ORL.....	43
Figura 3.2 - Faces de criminosos sorteadas. ....	45
Figura 3.3 - Diagrama em blocos da implementação do sistema.....	47
Figura 3.4 - Diagrama em blocos do detalhamento de cada classificador.....	48
Figura 3.5 - Face representada pelo vetor.....	48
Figura 4.1 - Média dos Cumulantes de 2ª Ordem (linha contínua) $\pm$ desvio padrão (linha pontilhada): (a) indivíduo 1, (b) indivíduo 2, (c) indivíduo 3, (d) indivíduo 4, (e) indivíduo 5 .....	51
Figura 4.2 - Média dos Cumulantes de 3ª Ordem (linha contínua) $\pm$ desvio padrão (linha pontilhada): (a) indivíduo 1, (b) indivíduo 2, (c) indivíduo 3, (d) indivíduo 4, (e) indivíduo 5 .....	52
Figura 4.3 - Média dos Cumulantes de 4ª Ordem (linha contínua) $\pm$ desvio padrão (linha pontilhada): (a) indivíduo 1, (b) indivíduo 2, (c) indivíduo 3, (d) indivíduo 4, (e) indivíduo 5 .....	53
Figura 4.4 - Discriminante de Fisher para os parâmetros das faces dos indivíduos 1 (a), 2 (b), 3 (c), 4 (d) e 5 (e).....	56
Figura 4.5 - Espaço de parâmetros para o classificador 1.....	57
Figura 4.6 - Espaço de parâmetros para o classificador 2.....	58
Figura 4.7 - Espaço de parâmetros para o classificador 3.....	59

Figura 4.8 - Espaço de parâmetros para o classificador 4.....	60
Figura 4.9 - Espaço de parâmetros para o classificador 5.....	61

## **LISTA DE QUADROS**

Quadro 2.1 - Aplicações de reconhecimento de face.....	20
--	----

## LISTA DE TABELAS

Tabela 4.1 - Quantidade de Cumulantes de ordens 2, 3 e 4, por classificador ...	54
Tabela 4.2 - Resultado do treino para cada classificador.....	62
Tabela 4.3 - Resultado de teste para cada classificador.....	63
Tabela 4.4 - Resultado de classificação final.....	64
Tabela 4.5 – Resultado de detecção final.....	64

## LISTA DE SIGLAS

PMMG	Polícia Militar de Minas Gerais
EOS	Estatística de Ordem Superior
SO	Sistema Operacional
PCA	<i>Principal Component Analysis</i>
MIT	<i>Massachusetts Institute of Technology</i>
ICA	<i>Independent Component Analysis</i>
LDA	<i>Linear Discriminant Analysis</i>
RNA	<i>Redes Neurais Artificiais</i>
KNN	<i>K-Nearest Neighbour</i>
SVM	<i>Support Vector Machines</i>

## SUMÁRIO

<b>CAPÍTULO 1</b> .....	14
<b>1 INTRODUÇÃO</b> .....	14
1.1 Objetivos .....	18
1.2 Organização do Texto .....	18
<b>CAPÍTULO 2</b> .....	19
<b>2 REVISÃO DE LITERATURA</b> .....	19
2.1 Reconhecimento Facial: Contextualização e Estado da Arte.....	19
2.2 Métodos de Reconhecimento de Faces .....	23
2.3 Técnicas Utilizadas no Método Proposto.....	33
<b>CAPÍTULO 3</b> .....	42
<b>3 METODOLOGIA</b> .....	42
3.1 Plataforma de Desenvolvimento .....	42
3.2 Base de Dados .....	42
3.3 Método Proposto.....	46
<b>CAPÍTULO 4</b> .....	50
<b>4 RESULTADOS</b> .....	50
<b>CAPÍTULO 5</b> .....	65
<b>CONCLUSÃO</b> .....	65
<b>REFERÊNCIAS</b> .....	66

## **CAPÍTULO 1**

Este capítulo delimita as principais motivações da dissertação. Inicia-se pela definição do problema, erguido a partir de questões relacionadas ao reconhecimento automático da face. Após abordar alguns dos desafios sobre a técnica, realiza-se uma contextualização geral do problema com foco nas possíveis aplicações. Apresenta-se ainda nesta seção argumentos para vantagens do reconhecimento e sobre outros dados biométricos. Por fim, apresentam-se os objetivos e a organização geral do trabalho.

### **1 INTRODUÇÃO**

Nos últimos anos, o conhecimento científico produzido a respeito do reconhecimento de padrões tem contribuído para desenvolvimento e aperfeiçoamento de técnicas com diferentes finalidades e aplicações em diversas áreas como Astronomia, Engenharia, Medicina e Segurança. Uma importante subárea vinculada ao reconhecimento de padrões é a identificação facial, que utiliza técnicas avançadas de processamento de sinais e imagem através de métodos biométricos, estando presente em diversas aplicações (THEODORIDIS; KOUTROUMBAS, 2009).

O reconhecimento de faces por humanos se dá de forma natural e instantânea, quando realizada a captação da imagem pelos olhos, essa é automaticamente memorizada pelo indivíduo, ainda que o registo não seja resgatado a qualquer tempo, em função da dificuldade em lembrar-se de uma face familiar, ele estará disponível. Em outras palavras, se por um lado os seres humanos apresentam maior facilidade na detecção e identificação de padrões faciais, o mesmo não pode ser dito em relação ao resgate das informações da memória, que por diversos motivos podem ser falhos.

Esse procedimento de registro de faces humanas, aparentemente simples, do ponto de vista da interação cotidiana, torna-se significativamente complexo quando se busca construir um sistema computacional com a mesma habilidade. A complexidade ocorre em função da existência de uma diversidade de variações existentes em uma face. Além disso, sistemas computacionais são acompanhados de limitações, tornando o processo menos consistente quando comparado com a capacidade humana. Dentre as limitações, a capacidade de processamento do *hardware* e os ruídos presentes na imagem captada são significativos.

Além das diversas variações de uma face, o trabalho de reconhecimento facial implica em fatores que contemplam os aspectos de iluminação do ambiente, pose do indivíduo, qualidade da imagem, oclusão, disfarces, dentre outros.

Geralmente, a operação de reconhecimento de face é realizada através do recolhimento de amostras dos traços faciais. Em seguida, são utilizadas técnicas de visão computacional para extrair as características da face e transformá-las em dados. A última fase é representada pela comparação entre os elementos extraídos com imagens de faces armazenadas em um banco de dados, confirmando ou rejeitando o reconhecimento.

Os sistemas de reconhecimento facial apresentam vasta aplicabilidade, com destaque para os sistemas de segurança, realizando importante tarefa na sociedade. É chave de acesso para sistemas e locais como, por exemplo, computadores pessoais, *smartphones*, acesso a ambientes específicos do sistema bancário, dentre outros sistemas vinculados à interface homem-máquina.

Esses sistemas desempenham, portanto, um papel importante na segurança da informação, bens e pessoas. A vantagem do reconhecimento facial sobre outros métodos biométricos assenta-se que o reconhecimento é não intrusivo, isto é, pode ser feito sem que o indivíduo interaja de forma direta com o sistema.



O reconhecimento automático de faces ainda não alcançou um nível ótimo de maturidade. Diversos algoritmos de reconhecimento facial vêm sendo desenvolvidos, mas ainda não existe um que seja consideravelmente capaz de reconhecer o rosto de uma pessoa com a mesma eficiência que o ser humano, ou seja, que reconheça faces em qualquer ambiente, vista a qualquer ângulo, e não importando a expressão facial. Porém, em condições controladas, existem *softwares* que superam o desempenho humano, sendo até mesmo capazes de diferenciar gêmeos idênticos (PAONE et al., 2014).

No contexto de sistemas de segurança, a identificação de criminosos, usando reconhecimento facial, caracteriza-se como uma importante aplicação. Isso se justifica pela crescente sensação de insegurança, um dos principais fatores enfrentados pela sociedade atual, trazida por uma crescente cultura do medo. O crescimento da violência, representado por crimes patrimoniais diversos, homicídios, dentre outros, constitui motivação significativa para inserção de câmeras de vigilância, não apenas em ambientes privados, mas também em espaços de circulação pública.

O aumento dos índices de criminalidade no estado de Minas Gerais inspirou uma série de medidas por parte do Estado. Com a necessidade de alcançar uma maior efetividade das ações operacionais, uma das medidas encontradas com a finalidade de minimizar o crescimento da criminalidade foi a implantação de câmeras de vigilância. Iniciaram-se as instalações no ano de 2005 nos grandes centros e, atualmente, é crescente o número de câmeras espalhadas pelas cidades (CARVALHO; DINIZ, 2009).

Hoje o monitoramento dessas câmeras, denominadas “Olho Vivo”<sup>1</sup>, vem sendo realizado pela Polícia Militar de Minas Gerais (PMMG), a fim de identificar e capturar pessoas que estejam na prática de algum fato delituoso, procedimento

---

<sup>1</sup> O projeto Olho Vivo é formado por um conjunto de câmeras espalhadas por pontos específicos da cidade, permitindo um monitoramento da criminalidade.

que tem auxiliado na redução de alguns índices criminais, aumentando a sensação de segurança das pessoas. Contudo, os operadores desses equipamentos não conseguem identificar todas as pessoas que passam pelas câmeras e muito menos dizer se elas são foragidas da justiça ou se cometeram algum crime.

Motivado pelo contexto supracitado, esta dissertação usa-o como cenário para o desenvolvimento de um sistema de reconhecimento de faces e identificação de indivíduos com pendências judiciais, através de imagens das faces dos criminosos pré-cadastrados.

As técnicas de processamento de imagens e reconhecimento de padrões empregadas nos sistemas de reconhecimento de face são diversas e apresentam vantagens e desvantagens (AMARAL, 2011). Esta dissertação apresenta uma inovação no uso de estatísticas de ordem superior (EOS) como parâmetros a fim de gerar assinaturas compactas das faces a serem classificadas (NIKIAS; PETROPULU, 1993).

As principais motivações para o uso de EOS são (NIKIAS; PETROPULU, 1993):

- (i) são mais imunes a ruídos gaussianos do que as estatísticas de segunda ordem;
- (ii) são mais indicadas para representar processos não lineares;
- (iii) têm sido aplicadas em sistemas de reconhecimento de padrões em diferentes áreas, recentemente, com bastante sucesso (VIANELLO et al., 2009; MARQUES; FERREIRA; RIBEIRO, 2008; NAVES; BARBOSA; FERREIRA, 2014).

## 1.1 Objetivos

Desenvolver um método de reconhecimento de faces utilizando estatísticas de ordem superior, considerando o cenário de identificação de criminosos.

Os objetivos específicos foram:

- Adaptar o algoritmo de extração de EOS para imagens, visto que o mesmo foi originalmente desenvolvido para sinais;
- Fazer uma revisão bibliográfica no contexto de reconhecimento de face;
- Avaliar quais são as EOS mais relevantes para o reconhecimento de faces;
- Projetar um classificador para reconhecer as faces a partir das assinaturas formadas pelas EOS;

## 1.2 Organização do Texto

Esta dissertação está organizada em cinco capítulos, descritos a seguir. Neste primeiro capítulo apresentou-se as principais motivações da dissertação, uma contextualização geral do problema e os objetivos. O segundo capítulo apresenta uma revisão de literatura sobre o reconhecimento facial, com destaque para a contextualização e estado da arte, e apresenta as técnicas a serem utilizadas nesta dissertação. O terceiro capítulo descreve o método proposto para o reconhecimento de face no contexto de detecção e identificação de criminosos. No quarto capítulo, os resultados do método proposto são apresentados de uma forma qualitativa e quantitativa, sobre os módulos detector e identificador de faces. Por fim, o quinto capítulo traz a conclusão a respeito do trabalho realizado.

## CAPÍTULO 2

As inúmeras aplicações práticas do reconhecimento facial e seus desafios o tornaram uma área de pesquisa muito popular da visão computacional. Isto é evidenciado por um grande número de algoritmos de reconhecimento de face desenvolvidos ao longo das últimas três décadas e, nos últimos anos, o aparecimento de um número considerável de sistemas de reconhecimento de face comerciais. Este capítulo apresenta uma revisão de literatura sobre o reconhecimento facial, com destaque para a contextualização e estado da arte, e apresenta as técnicas a serem utilizadas nesta dissertação.

## 2 REVISÃO DE LITERATURA

### 2.1 Reconhecimento Facial: Contextualização e Estado da Arte

O reconhecimento facial é uma das técnicas biométricas mais pesquisadas na área computacional. É um desafio desenvolver um sistema computacional que tenha a mesma capacidade humana de reconhecer uma pessoa através da sua face. Embora os seres humanos sejam muito bons na identificação de pessoas conhecidas, não são muito hábeis quando têm que lidar com uma grande quantidade de rostos desconhecidos. Hoje existem algoritmos que conseguem superar essa limitação dos seres humanos, em ambientes controlados, conseguindo identificar gêmeos idênticos, com um alto grau de certeza.

Pesquisadores têm se interessado por trabalhos relacionados ao reconhecimento facial, prova disso foi uma simples procura eletrônica com a frase “*face recognition*” na Biblioteca Digital do IEEE (*Institute of Electrical and Electronics Engineers*) que apontou 22.957 resultados. Apenas no ano 2015,

foram encontrados 1.964 artigos científicos<sup>2</sup>. Já na biblioteca Google Acadêmico foram mostrados aproximadamente 346.000 resultados e, reduzindo a pesquisa para o ano de 2015, foram encontrados mais de 19.000 documentos. A mesma pesquisa, no Portal de Periódicos CAPES/MEC, apontou 9.091 periódicos, salientando o direcionamento de pesquisadores para trabalhos relacionados ao reconhecimento facial.

Quadro 2.1 - Aplicações de reconhecimento de face

<b>Áreas</b>	<b>Aplicações</b>
<b>Segurança da Informação</b>	Segurança de acesso (SO, bases de dados) Privacidade de dados (registros médicos) A autenticação do usuário ( <i>Internet Banking</i> )
<b>Gerenciamento de Acesso</b>	Autenticação de acesso seguro (instalações restritas) Sistemas baseados em permissão Log de acesso ou recursos de rastreamento
<b>Biometria</b>	Identificação de pessoa (passaportes, registros biométricos de eleitores, carteiras de motorista) Verificação de identidade automatizada
<b>Aplicação da Lei</b>	Vigilância por vídeo Identificação de suspeito Rastreamento de suspeito (investigação) Envelhecimento simulado Reconstrução Forense de rostos de restos mortais
<b>Segurança Pessoal</b>	Sistema de segurança por vídeo Interpretação de expressão (sistema de acompanhamento da fadiga do condutor)
<b>Entretenimento e Lazer</b>	Videogame Aplicações para câmeras fotográficas

Fonte: Adaptado de Marques (2010)

Segundo Marques (2010), existem muitas áreas diferentes da indústria interessadas em métodos de reconhecimento de padrões. Alguns exemplos incluem a vigilância de vídeo, interação homem-máquina, câmeras fotográficas, realidade virtual ou a aplicação da lei, conforme Quadro 2.1. Esse interesse

<sup>2</sup> Pesquisa feita em 21 de abril de 2016.

multidisciplinar alimenta a investigação e atrai o interesse de diversas disciplinas. Portanto, não é um problema restrito à pesquisa de visão por computador (ou visão de máquina). Reconhecimento facial é um assunto relevante no reconhecimento de padrões, redes neurais, computação gráfica, processamento de imagem e psicologia (ZHAO et al., 2003).

Um resumo da história do reconhecimento facial feito por Marques (2010), desde o início das pesquisas em 1950 até a década de 1990, revela aumento notável em publicações sobre o tema.

Os primeiros trabalhos sobre o assunto foram produzidos na década de 1950 em psicologia. Eles vieram ligados a outras questões como expressão da face, interpretação de emoção ou percepção de gestos.

A engenharia começou a mostrar interesse em reconhecimento facial na década de 1960. Uma das primeiras pesquisas sobre esse assunto foi realizada por Woodrow W. Bledsoe. Segundo Marques (2010), Woodrow W. Bledsoe identificou alguns dos principais problemas relacionados ao reconhecimento facial, que ainda na atualidade, representam dificuldades na materialização do método. Dentre as limitações, o autor identificou: variação na iluminação, rotação da cabeça, expressão facial e envelhecimento.

Outros pesquisadores como A. Jay Goldstein, Leon D. Harmon e Ann B. Lesk tentaram medir as características do rosto, como o tamanho da orelha ou a distância entre os olhos. Eles descreveram um vetor, contendo 21 características subjetivas, como tamanho da orelha, largura ou comprimentos do nariz e sobrelha, como base para reconhecer rostos usando técnicas de classificação padrão (MARQUES, 2010).

Em 1973, Fischler e Elschanger tentaram medir características semelhantes automaticamente (MARQUES, 2010). O algoritmo usado combina o local padrão e uma média global para localizar e medir características faciais. Ainda na década de 70, alguns pesquisadores tentaram definir um rosto como um

conjunto de parâmetros geométricos e, em seguida, executar um reconhecimento de padrões com base nesses parâmetros. Para Marques (2010), o primeiro pesquisador a desenvolver um sistema de reconhecimento facial totalmente automatizado foi Kenade em 1973. O algoritmo extraiu dezesseis parâmetros faciais automaticamente, com uma taxa de identificação correta compreendida no intervalo de 45-75%. Ele mostrou que os melhores resultados foram obtidos quando as características irrelevantes não foram utilizadas.

Já nos anos 80, seguiu-se uma diversidade na abordagem, a maioria delas com tendências anteriores. Alguns trabalhos tentaram melhorar os métodos usados para medir as características subjetivas. Por exemplo, Nixon (1985) apresentou uma medição geométrica para o espaçamento entre os olhos. Essa década também trouxe novas abordagens, como aquelas utilizadas por pesquisadores que construíram algoritmos de reconhecimento facial, utilizando redes neurais artificiais.

A primeira menção a *eigenfaces* em processamento de imagens foi feita por M. Kirby e L. Sirovich em 1986 (SIROVICH; KIRBY, 1987). Essa técnica se tornaria a abordagem dominante nos anos seguintes, utilizando a Análise de Componentes Principais (*Principal Component Analysis - PCA*). Seu objetivo era representar uma imagem numa dimensão inferior, sem perder muita informação, e depois reconstruí-la. O trabalho de Sirovich e Kirby (1987) passou a ser posteriormente a base para o desenvolvimento de novos algoritmos de reconhecimento de face.

Na década de 1990, houve um amplo estudo sobre o método *eigenface* mencionado como a base para o estado da arte e as primeiras aplicações industriais. Em 1992, Mathew Turk e Alex Pentland do MIT (*Massachusetts Institute of Technology*) apresentaram um trabalho que usou o *eigenface* como método de reconhecimento de faces. Seu algoritmo foi capaz de localizar, rastrear e classificar a face de uma pessoa (TURK; PENTLAND, 1991).

Marques (2010) apontam que desde a década de 1990, a área de reconhecimento facial tem recebido muita atenção, com um aumento significativo no número de publicações. Muitas abordagens foram feitas, o que levou a diferentes algoritmos, como os baseados em Análise de Componentes Principais, Análise de Componentes Independentes, Análise Discriminante Linear e seus derivados e Redes Neurais Artificiais - RNA.

As tecnologias que utilizam técnicas de reconhecimento de face também evoluíram ao longo dos anos. Atualmente, diversas empresas estão usando reconhecimento facial em seus produtos. Um exemplo são os videogames. Produtos como o *Project Natal* da *Microsoft* (DUDLEY, 2009) ou *Playstation Eye* da *Sony* (MCWHERTOR, 2009) usam o reconhecimento de face, permitindo uma nova forma de interagir com a máquina. A ideia de detectar pessoas e analisar o seu gesto também está sendo usada na indústria automotiva. Empresas como a *Toyota* estão desenvolvendo detectores de sono para aumentar a segurança (MASSY, 2008). Essas e também outras aplicações revelam o crescimento do interesse da comunidade científica no reconhecimento facial.

## **2.2 Métodos de Reconhecimento de Faces**

Os métodos aplicados ao reconhecimento de faces utilizam as informações globais das faces para executar o reconhecimento. Essas informações são representadas por um vetor de face em um sistema de coordenadas reduzidas, sendo o vetor derivado diretamente da informação dos pixels das imagens da face. Esse vetor no sistema de coordenadas reduzidas captura distintamente a variação entre as faces individuais, diferentes da média e, conseqüentemente, é utilizado para identificar excepcionalmente os indivíduos.

Para análises matemáticas, diversos métodos já foram propostos, os mais utilizados são Análise de Componentes Principais e Análise de Componentes



Independentes. Ao chegar a um determinado ponto do processo de reconhecimento da face, torna-se necessária a utilização de classificadores, para medirem a distância entre as imagens em um espaço N-dimensional. Os mais utilizados são Redes Neurais Artificiais, Análise Discriminante Linear, K-Vizinhos mais próximos, Classificador de Bayes, dentre outros.

### 2.2.1 Análise de Componentes Principais

Uma técnica estatística bastante utilizada, principalmente na área de reconhecimento de padrões, é a *Principal Component Analysis (PCA)*, em português Análise de Componentes Principais. Esse método foi um dos melhores e mais bem-sucedidos algoritmos para reconhecimento, utilizando auto faces (JOLLIFFE, 2002). Foi desenvolvido por M. Turk e A. P. Pentland (TURK; PENTLAND, 1991). Essa técnica matemática, em reconhecimento facial, visa eliminar a redundância de informação e, ainda assim, manter as principais características da face, através da seleção de características.

Para isso, faz-se necessário transformar a imagem em um padrão para efetivar o reconhecimento da face, ou seja, alterar o espaço da face em um espaço linear. Imagens faciais podem ser vistas como matrizes bidimensionais de dimensão  $A \times B$  características, onde  $A$  corresponde ao número de linhas da imagem e  $B$  é número de colunas da imagem, que deve ser transformado para o espaço linear, sendo então representado por um vetor coluna  $m$ , construído a partir da leitura coluna por coluna da imagem original.

Para exemplificar o processo de conversão de uma face em um vetor, a Figura 2.1 mostra uma matriz que representa uma imagem de dimensão 112x92, convertido em um vetor coluna de 10.304 posições, também mostrado nesta figura (MEDEIROS, 2012).

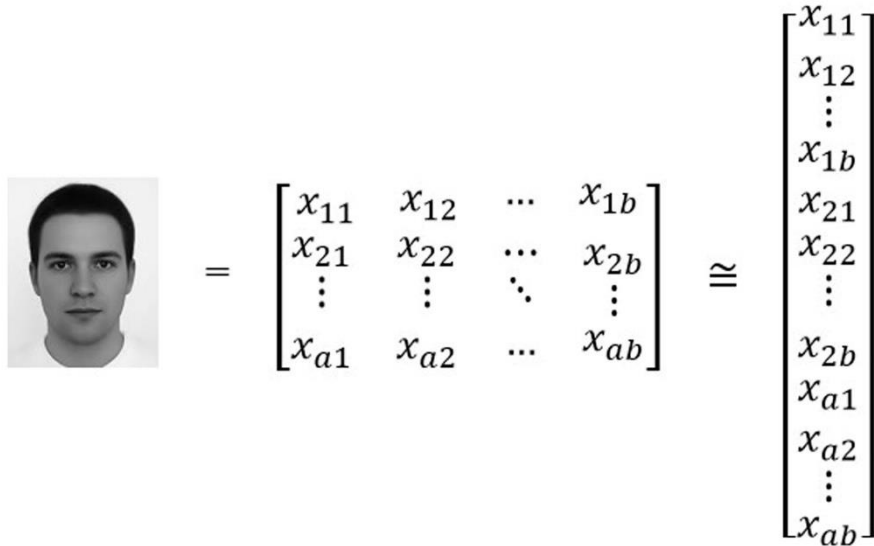


Figura 2.1 - Exemplo da conversão de uma face de dimensão 112x92 em um vetor coluna de 10.304 posições  
 Fonte: Adaptado de Medeiros (2012)

São propostos por Duan, Yan e Lin (2008) e Medeiros (2012) os passos do algoritmo para o reconhecimento facial utilizando o PCA:

O primeiro passo é converter todas as faces armazenadas no banco de faces e a face teste  $\mathbf{q}$  em vetores colunas, conforme a Figura 2.1. O resultado dessa conversão são vetores de  $n$  posições ( $n = AB$ , onde  $A$  e  $B$  são os números de linhas e de colunas das faces, respectivamente).

Em seguida é calculado o vetor médio  $\bar{X}_i$  para cada classe do conjunto de imagens  $X(i = 1, 2, \dots, L)$ , utilizando a Equação 2.1:

$$\bar{X}_i = \frac{1}{N} \sum_{k=1}^{N_i} X_k^{(i)} \quad (2.1)$$

onde:  $N = N_1 + N_2 + \dots + N_L$ .

O terceiro passo é determinar as matrizes de covariância  $\mathbf{S}_{x_i}$  de cada classe utilizada na Equação 2.2:

$$\mathbf{S}_{x_i} = \frac{1}{N-1} \sum_{k=1}^{N_i} (X_k^{(i)} - \bar{X}_i)(X_k^{(i)} - \bar{X}_i)^T \quad (2.2)$$

O próximo passo é calcular os  $m$  autovetores relacionados aos  $m$  maiores autovalores da matriz de covariância  $\mathbf{S}_{x_i}$ , obtidos na Equação 2.2. Esses autovetores relacionados a  $i$ -ésima classe são:  $\mathbf{u}_1^{(i)}, \mathbf{u}_2^{(i)}, \dots, \mathbf{u}_m^{(i)}$ , onde  $\mathbf{u}_1^{(i)}$  é o autovetor relacionado ao maior autovalor,  $\mathbf{u}_2^{(i)}$ , é o autovetor relacionado ao segundo maior autovalor e, assim por diante.

No passo cinco é gerada a matriz de transformação  $\mathbf{P}_{x_i}$ , definida como:  $\mathbf{P}_{x_i} = [\mathbf{u}_1^{(i)}, \mathbf{u}_2^{(i)}, \dots, \mathbf{u}_m^{(i)}]^T$ , onde  $\mathbf{P}_{x_i}^T \in R^{m \times n}$ .

O passo seguinte é projetar a imagem teste  $\mathbf{q}$  em cada autoespaço ( $i = 1, 2, \dots, L$ ), utilizando a Equação 2.3, onde  $\mathbf{w}_i$  indica a projeção da imagem teste na  $i$ -ésima classe.

$$\mathbf{w}_i = \mathbf{P}_{x_i}^T (\mathbf{q} - \bar{X}_i) \quad (2.3)$$

O penúltimo passo é determinar a imagem reconstruída  $\mathbf{q}_i$ , a partir do vetor projeção  $\mathbf{w}_i$ , utilizando a Equação 2.4, para ( $i = 1, 2, \dots, L$ ).

$$\mathbf{q}_i = \mathbf{P}_{x_i} \mathbf{w}_i \quad (2.4)$$

Por fim, calcula-se a distância Euclidiana relativa entre a imagem teste e todas as imagens reconstruídas ( $i = 1, 2, \dots, L$ ), utilizando a Equação 2.5.

$$d_i = \frac{\|\mathbf{q} - \mathbf{q}_i\|}{\|\mathbf{q}\|} \quad (2.5)$$

Como o índice  $i$  indica a  $i$ -ésima classe, a menor distância  $d_i$  calculada a partir da Equação 2.5 indica que a classe mais similar à imagem teste é a classe de índice  $i$  e a imagem é reconhecida como pertencente a essa classe (MEDEIROS, 2012).

### 2.2.2 Análise de Componentes Independentes

A Análise de Componentes Independentes (*Independent Component Analysis – ICA*) é uma técnica matemática que se resume em encontrar uma representação linear, em que os componentes são estatisticamente independentes. É também conhecida como “separação cega de sinais”, (HYVÄRINEN; KARHUNEN; OJA, 2001) e tem por finalidade separar sinais de diferentes fontes, todos misturados, em um conjunto de variáveis estatisticamente independentes.

O principal objetivo do *ICA* é decompor o sinal observado a partir das várias fontes em uma combinação linear dos sinais desconhecidos, que são independentes uns dos outros. Então, seja  $\mathbf{S}$  o vetor de sinais desconhecidos e o vetor das misturas dos sinais observados for representado por  $\mathbf{X}$ . Se o símbolo  $\mathbf{A}$  representa a matriz de mistura desconhecida, o modelo de mistura é escrito na Equação 2.6 (DRAPER et al., 2003):

$$\mathbf{X} = \mathbf{A}\mathbf{S} \quad (2.6)$$

Supõe-se que os sinais desconhecidos  $\mathbf{S}$  são independentes uns dos outros e que a matriz  $\mathbf{A}$  é invertível. Com base nestes dois pressupostos, o algoritmo *ICA* tenta encontrar  $\mathbf{A}$  (matriz maximizada) ou a matriz  $\mathbf{W}$  (matriz de separação), conforme a Equação 2.7

$$U = WX = WAS \quad (2.7)$$

De acordo com Cardoso (1997), a Equação 2.7 é uma estimativa dos sinais de fonte independentes.

No contexto do reconhecimento de face, o uso do ICA foi proposto, primeiramente por Bartlett (2001) e, posteriormente, por várias outras contribuições, sendo as mais citadas, Liu e Wechsler (1999); Baek et al. (2002); Draper et al. (2003). Em Bartlett (2001), duas arquiteturas diferentes são apresentadas para o reconhecimento de face. Na primeira, chamada de arquitetura-*I*, as imagens de faces são descritas como combinação linear de um conjunto desconhecido de componentes estatisticamente independentes e os componentes obtidos nessa arquitetura podem ser vistos como um conjunto de imagens de características faciais independentes. A Figura 2.2 exemplifica a arquitetura-*I*, onde a matriz de entrada,  $X$ , é organizado de modo que as linhas representem imagens diferentes e as colunas representem pixels diferentes. A projeção das imagens originais,  $X$ , no vetor de pesos,  $W$ , aprendido pelo *ICA*, produz as imagens base independentes,  $U$ , tal que  $U = WX$ .

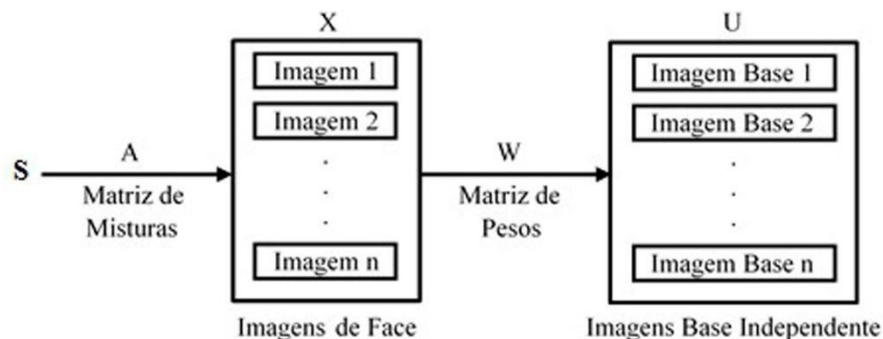


Figura 2.2 - Arquitetura-*I*  
Fonte: Adaptado de Bartlett (2001)

Na segunda, chamada de arquitetura-*II*, as representações dos coeficientes (pesos) são estatisticamente independentes e os componentes obtidos podem ser vistos como um conjunto de imagens com a aparência de uma face, assim como no PCA.

De acordo com os testes realizados em Draper et al. (2003), a arquitetura-*I* obteve melhores resultados do que a arquitetura-*II*. As bases derivadas da arquitetura-*I* refletem propriedades locais da face. Já na arquitetura-*II* apresentam propriedades mais globais da face, como é possível observar na Figura 2.3.

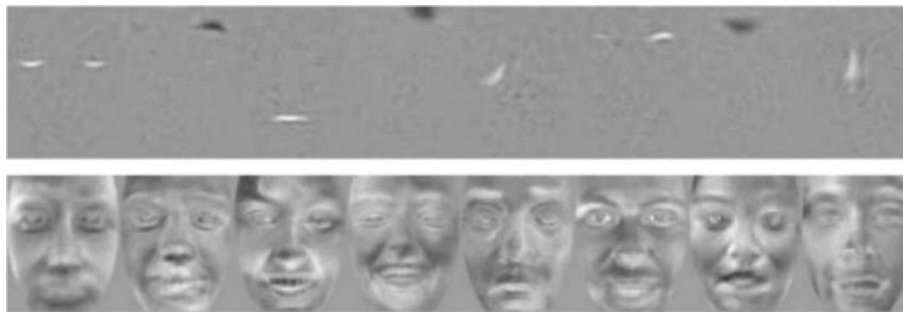


Figura 2.3 - Arquitetura-*I* (figura acima): bases refletem características mais locais; Arquitetura-*II* (figura abaixo): bases refletem características mais globais  
Fonte: imagem obtida de Draper et al. (2003)

### 2.2.3 Redes Neurais Artificiais

Uma Rede Neural Artificial (RNA) é uma técnica computacional que busca simular as conexões em rede estabelecidas pelo cérebro humano, permitindo autossuficiência na solução de situações problemas. Pode-se dizer ainda que são técnicas computacionais que buscam a superação autónoma de situações problemas que mesmo sem programação anterior são assimiladas e superadas.

Uma definição para RNA é apontado em:

sistemas inteligentes, cuja meta é interpretar e resolver problemas computacionais baseando seu funcionamento no cérebro humano. São compostas por unidades de processamento interconectadas, denominadas neurônios artificiais, que processam sinais para computar funções matemáticas (SILVA, 2016, p. 42).

Para Bordignon (2015, p. 20) RNA são:

sistemas não lineares baseados na funcionalidade dos neurônios biológicos, e são utilizados principalmente em problemas de classificação, reconhecimento de padrões e otimização na precisão de sistemas complexos.

A Figura 2.4 apresenta um modelo não linear de um neurônio artificial, que pode ser identificado por três elementos básicos (HAYKIN, 2001).

- Um conjunto de **sinapses** ou **elo de conexão**: cada uma consiste na entrada dos sinais por meio de um peso para o neurônio. Especificamente, um sinal  $x_j$ , na entrada da sinapse  $j$ , conectada ao neurônio  $k$ , é multiplicado pelo peso sináptico  $W_{kj}$ .
- Um **somador**  $\Sigma$ : para somar os sinais de entrada, ponderados pelas respectivas sinapses do neurônio; as operações descritas aqui constituem um combinador linear.
- Uma **função de ativação**  $\varphi(\cdot)$ : responsável por obter o resultado  $U_k$ , do somatório e delimitar o sinal de saída  $Y_k$  com um único valor, utilizado como entrada para o próximo neurônio ou como resposta da RNA.

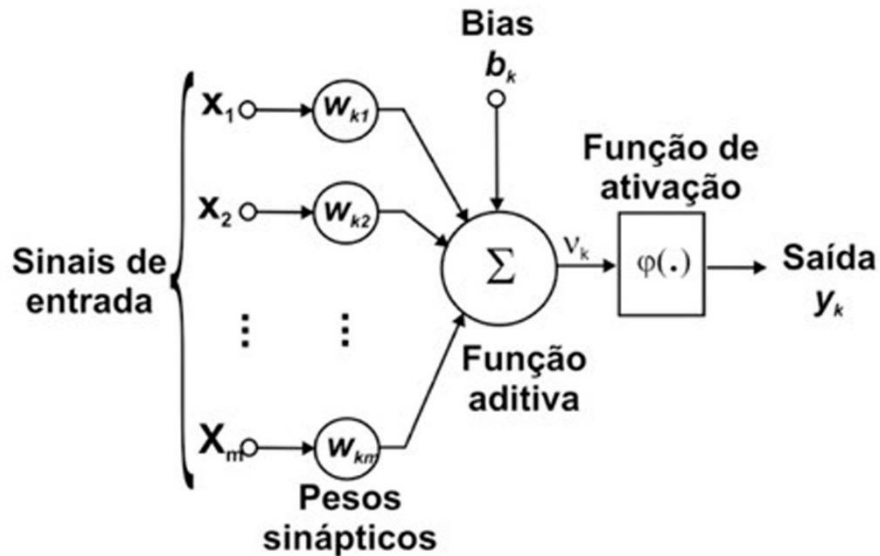


Figura 2.4 - Modelo não linear de um neurônio artificial  
 Fonte: Adaptado de Haykin (2001)

O objetivo das RNAs é solucionar problemas aprendendo e agindo diante de situações variadas, agindo através da experiência e da observação dos acontecimentos. McCulloch e Pitts (1943) foram os primeiros pesquisadores a apresentar um modelo matemático para uma rede neural.

Em reconhecimento facial, as RNAs vêm sendo aplicadas em diversos problemas, como por exemplo, classificação por sexo, reconhecimento de face, classificação de expressões faciais, dentre outros. Chellappa, Wilson e Sirohey (1995) foram uns dos primeiros a utilizar RNA para identificação de imagens.

Na detecção de faces, a utilização de redes neurais consiste em diferenciar imagens ou regiões da imagem que, possivelmente contenham uma face, daquelas imagens ou regiões que não contêm face, ou seja, a função da rede é apenas classificar as imagens em face e não face (VALENÇA, 2005).



Atualmente, existem numerosos estudos eficientes utilizando redes MLP's<sup>3</sup> para resolver problemas no processo de reconhecimento de faces (BOUGHRARA; CHTOUROU; AMAR, 2012; ARRUDA, 2013; BOUGHRARA et al., 2014; SILVA, 2016).

#### 2.2.4 Classificador K-Vizinhos mais Próximos

O classificador de  $k$ -Vizinhos Mais Próximos, do inglês *K-Nearest Neighbour (KNN)*, é um algoritmo que cria alcances de decisões complexas a partir de um conjunto de padrões de treinamento com classes conhecidas *a priori*. Dado um padrão desconhecido  $x$ , sua classificação ocorre da seguinte forma: calcula-se a distância entre  $x$  e todos os padrões de treinamento, verifica-se a quais classes pertencem os  $k$  padrões mais próximos de  $x$  e finalmente associa-se o padrão  $x$  à classe mais frequente entre esses  $k$  padrões. Como medida de distância utiliza-se a distância Euclidiana ou similar.

Matos (2008) aponta que uma das vantagens conhecidas do classificador KNN é que, quando o conjunto de treinamento é grande ou bem representativo, ele cria uma fronteira de decisão que se adapta à forma de distribuição dos dados de treinamento, possibilitando a obtenção de boas taxas de acertos.

Em reconhecimento facial, a abordagem de classificação KNN identifica as  $k$  faces mais próximas e classifica a face de teste, considerando a classe que for mais frequente entre as  $k$  faces de treinamento mais próximas.

Sejam  $\{y_{1j}, y_{2j}, \dots, y_{mj}\}$  os  $m$  coeficientes selecionados para representar a pessoa  $j$  e sejam  $\{w_{1jk}, w_{2jk}, \dots, w_{mjk}\}$  as amplitudes dos coeficientes de

---

<sup>3</sup> Do Inglês, *feed forward multilayer perceptron (MLP)*. Rede de múltiplas camadas, compostas por um conjunto de unidades sensoriais que formam a camada de entrada, uma ou mais camadas ocultas e uma camada de saída

treinamento da pessoa  $j$  na pose  $k$ , com  $w_{ijk}$  correspondendo ao coeficiente de mesma posição que  $y_{ij}$ .

Seja  $f$  a pessoa a ser classificada e sejam  $\{v_{1f}, v_{2f}, \dots, v_{mf}\}$  as amplitudes dos coeficientes Transformada Cosseno Discreta da pessoa  $f$ , com  $v_{if}$  correspondendo ao coeficiente de mesma posição que  $y_{ij}$ .

O classificador KNN classifica a face  $f$ , baseando-se nos seguintes passos (MATOS, 2008):

1. Calcula-se a distância entre a face  $f$  e a pessoa de treinamento  $j$  na pose  $k$ , com  $j = 1, 2, \dots, p$  e  $k = 1, 2, \dots, q$ , conforme Equação 2.8:

$$DKNN_{fjk} = \sum_{i=1}^m |w_{ijk} - v_{if}| \quad (2.8)$$

2. Identificam-se os  $k$  menores valores  $DKNN_{fjk}$  e define-se o  $i$ -ésimo vizinho mais próximo como a pessoa  $j$  associada ao  $i$ -menor valor de  $DKNN_{fjk}$ , com  $i = 1, 2, \dots, k$ ;
3. A classificação da pessoa  $f$  corresponde então à pessoa  $j$  mais frequente entre os  $k$ -vizinhos identificados;
4. Em caso de empate, foi definido nesse trabalho se adotar a abordagem KNN para desempate.

## 2.3 Técnicas Utilizadas no Método Proposto

### 2.3.1 Estatística de Ordem Superior

Conforme Mendel (1991); Ferreira (2010); Ferreira et al. (2011), as estatísticas de ordem superior (EOS) podem levar a resultados mais

representativos quando empregadas em processos não gaussianos e sistemas não lineares do que as estatísticas de segunda ordem. Sabe-se que um sinal gaussiano é completamente caracterizado pela sua média e variância. Consequentemente, as EOS de um sinal gaussiano são nulas ou possuem informações redundantes. Logo, as EOS são menos afetadas por ruídos gaussianos do que as estatísticas de segunda ordem.

Dado que muitas aplicações práticas possuem ruído aditivo gaussiano, as EOS têm sido utilizadas com bastante frequência, conforme é mostrado em (MENDEL, 1991; FERREIRA, 2010; FERREIRA et al., 2011).

Para os pesquisadores Mendel (1991); Ferreira (2010); Ferreira et al. (2011), as EOS podem ser definidas em termos de momentos e cumulantes. Momentos são mais indicados para sinais determinísticos, enquanto que cumulantes são mais aplicados para sinais aleatórios.

As expressões dos cumulantes de segunda, terceira e quarta ordens de um sinal aleatório  $x[n]$ , quando  $E\{x[n]\} = 0$ , são, respectivamente, a Equação (2.9)-(2.11):

$$C_{2,x}[i] = E\{x[n]x[n+i]\}, \quad (2.9)$$

$$C_{3,x}[i] = E\{x[n]x^2[n+i]\}, \quad (2.10)$$

$$C_{4,x}[i] = E\{x[n]x^3[n+i]\} - 3C_{2,x}[i]C_{2,x}[0], \quad (2.11)$$

em que  $i$  é o  $i$ -ésimo atraso, e  $x[n]$  é o  $N$ -ésimo elemento do vetor  $\mathbf{x}$ .

Para um vetor com comprimento finito ( $N$ ), as aproximações estocásticas oferecem as seguintes Equações:

$$\hat{C}_{2,x}[i] := \frac{2}{N} \sum_{n=0}^N x[n]x[n+i], \quad (2.12)$$

$$\hat{C}_{3,x}[i] := \frac{2}{N} \sum_{n=0}^N x[n]x^2[n+i], \quad (2.13)$$

$$\hat{C}_{4,x}[i] := \frac{2}{N} \sum_{n=0}^N x[n]x^3[n+i] - \frac{2}{N^2} \sum_{n=0}^N x[n]x[n+i] \sum_{n=0}^N x^2[n], \quad (2.14)$$

em que  $i = 0, 1, \dots, N/2 - 1$ .

Note que as Equações (2.12)-(2.14) não podem ser usadas se  $i > N/2 + 1$ , pois  $n + i$  será maior que  $N$ . Dessa forma, perdem-se informações no cálculo dos cumulantes. Em Ribeiro et al. (2007), é proposta uma abordagem alternativa, em que cada cumulante é calculado utilizando todas as  $N$  amostras de sinal disponível. Basicamente, forma-se um tipo de *Buffer* circular nos sinais, de tal forma que, se o valor de  $n + i$  for maior que  $N$  em  $k$  unidades, esse valor é substituído por  $k$ , tomando com isso as amostras antecessoras ao valor  $i$ , que anteriormente não eram utilizadas. Na verdade, é como se houvesse uma continuidade da última amostra para a primeira, fazendo com que o sinal seja virtualmente circular.

Matematicamente, a substituição de  $n + 1$ , quando  $n + 1 > N$ , pode ser dada pela Equação 2.15:

$$\text{mod}(n + i, N) = [n + i] - bN, \quad (2.15)$$

em que  $b$  é o número inteiro obtido quando se desconsideram as casas decimais da divisão de  $n + i$  por  $N$ , o que resulta no resto da divisão de  $n + i$  por  $N$ .

Assim, as Equações (2.12)-(2.14) podem ser estimadas, para  $N$  finito, por (RIBEIRO et al., 2007):

$$\hat{C}_{2,x}[i] := \frac{2}{N} \sum_{n=0}^N x[n]x[\text{mod}[n + i, N]], \quad (2.16)$$

$$\hat{C}_{3,x}[i] := \frac{2}{N} \sum_{n=0}^N x[n]x^2[\text{mod}[n + i, N]], \quad (2.17)$$

$$\begin{aligned} \hat{C}_{4,x}[i] := & \frac{2}{N} \sum_{n=0}^N x[n]x^3[\text{mod}[n + i, N]] \\ & - \frac{2}{N^2} \sum_{n=0}^N x[n]x[\text{mod}[n + i, N]] \sum_{n=0}^N x^2[n], \end{aligned} \quad (2.18)$$

em que  $\text{mod}[n + i, N]$  é o resto inteiro da divisão de  $n + i$  por  $N$ .

As aproximações apresentadas nas Equações (2.16)-(2.18) levam a uma boa simplificação para problemas em que é empregado um vetor de comprimento finito. Pode-se observar que para um sinal com  $N$  amostras, tem-se  $N$  cumulantes para cada ordem de EOS. Para efeitos de classificação e detecção, a combinação de poucos destes cumulantes é suficiente para a obtenção de um bom desempenho, conforme mostrado em (RIBEIRO et al., 2007; FERREIRA et al., 2011).

Para Mendel (1991); Ferreira (2010); Ferreira et al. (2011), o cumulante de segunda ordem é apenas a autocorrelação de  $x[n]$ . Se  $x[n]$  é um processo aleatório simetricamente distribuído, então o cumulante de terceira ordem é igual a zero. Nesse caso, o cumulante de quarta ordem seria o mais indicado.

Adicionalmente, em alguns processos os valores dos cumulantes de terceira ordem são muito pequenos enquanto que os valores dos cumulantes de quarta ordem são muito grandes. Nesse caso, o cumulante de quarta ordem também é o mais indicado. Para atraso (*lag*) igual a zero, ou seja, ( $i = 0$ ).

### 2.3.2 Análise de Discriminante Linear de Fisher

Uma das melhores abordagens, indicada na literatura, adotada por métodos de reconhecimento facial, sendo inclusive superior em alguns aspectos à abordagem PCA (NEFIAN, 1996; BELHUMEUR; HESPANHA; KRIEGMAN, 1997; RUIZ-DEL-SOLAR; NAVARRETE, 2005) é a *Linear Discriminant Analysis (LDA)*, em português Análise de Discriminante Linear, também conhecida como Fisherfaces. O LDA é uma técnica bastante simples para discriminação de dados multidimensionais. Conforme bem discutido em Duda, Hart e Stork (2000) e Ferreira (2010), o LDA tem duas principais e diferentes aplicações: pode ser usado como separador de classes e como ferramenta de seleção de parâmetros (FERREIRA, 2010).

Para entender ambas as aplicações do LDA, suponha um espaço de  $n$  eventos  $d$ -dimensionais,  $\mathbf{x}_1, \dots, \mathbf{x}_n$ , em que  $n_1$  eventos pertencem à classe 1 ( $C_1$ ) e  $n_2$  à classe 2 ( $C_2$ ). Se a média das amostras  $d$ -dimensionais da classe  $i$  é:

$$\boldsymbol{\mu}_i = \frac{1}{n_i} \sum_{\mathbf{x} \in C_i} \mathbf{x} \quad (2.19)$$

podem-se definir as matrizes *scatters*  $\mathbf{S}_i$ ,  $\mathbf{S}_\omega$  e  $\mathbf{S}_B$  por:

$$\mathbf{S}_i = \sum_{\mathbf{x} \in C_i} (\mathbf{x} - \boldsymbol{\mu}_i) (\mathbf{x} - \boldsymbol{\mu}_i)^T, \quad (2.20)$$

$$\mathbf{S}_\omega = \mathbf{S}_1 + \mathbf{S}_2 \quad (2.21)$$

e

$$\mathbf{S}_B = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \quad (2.22)$$

em que  $T$  significa transposto.

A matriz  $\mathbf{S}_\omega$  é conhecida por matriz *scatter* intraclases e é proporcional à matriz de covariância dos dados. É também simétrica, semidefinida positiva e, usualmente, não singular se  $n > d$ . A matriz  $\mathbf{S}_B$  é também simétrica e semidefinida positiva. É chamada de matriz *scatter* entreclases.

Em termos de  $\mathbf{S}_\omega$  e  $\mathbf{S}_B$ , a função custo para o critério LDA como discriminante é escrita como:

$$\mathbf{J}(\boldsymbol{\omega}) = \frac{\boldsymbol{\omega}^T \mathbf{S}_B \boldsymbol{\omega}}{\boldsymbol{\omega}^T \mathbf{S}_\omega \boldsymbol{\omega}} \quad (2.23)$$

O vetor  $\boldsymbol{\omega}$  que maximiza  $\mathbf{J}(\cdot)$  será:

$$\boldsymbol{\omega} = \mathbf{S}_\omega^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) \quad (2.24)$$

Assim, o vetor  $\boldsymbol{\omega}$  obtido representa a máxima razão entre a matriz *scatter* entreclases e a intraclases. Dessa forma, o processo de classificação é convertido de um espaço  $d$ -dimensional para um espaço unidimensional. Neste novo espaço, pode-se obter um limiar que fará a separação entre os dados projetados em  $\boldsymbol{\omega}$  das classes 1 e 2.

Diferentemente do critério LDA para classificação, a função custo do critério LDA como ferramenta de seleção de parâmetros é:

$$\mathbf{J}_C = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^2 \odot \frac{1}{\boldsymbol{\sigma}_1^2 - \boldsymbol{\sigma}_2^2} \quad (2.25)$$

em que  $\mathbf{J}_C = [J_1 \cdots J_{L_l}]^T$ ,  $L_l$  é o número total de parâmetros,  $\boldsymbol{\mu}_1$  e  $\boldsymbol{\mu}_2$ , e  $\boldsymbol{\sigma}_1^2$  e  $\boldsymbol{\sigma}_2^2$  são, respectivamente, os vetores de média e variância dos dados da classe 1 e da classe 2. O símbolo  $\odot$  refere-se ao produto de Hadamard, ou seja,  $r \odot s = [r_0 s_0 \cdots r_{L_l-1} s_{L_l-1}]^T$ .

Pode-se entender que os  $i$ -ésimos elementos do vetor de parâmetros, referentes aos maiores valores de  $J_C$ , representam os parâmetros que apresentam melhor separação entre classes. Essa técnica é bastante útil em processos de classificação mais complexos, uma vez que torna possível selecionar os melhores parâmetros (que mais discriminam as classes envolvidas) e, conseqüentemente, reduzir o número de parâmetros a serem utilizados na classificação. Essa redução de parâmetros é atraente em sistemas que necessitam de menor complexidade computacional, tais como sistemas que operam em tempo real. Ademais, o uso de um número reduzido de parâmetros pode levar ao projeto de classificadores mais simples, com melhor capacidade de generalização e melhor desempenho, já que os parâmetros responsáveis pela interseção entre classes são, teoricamente, descartados (FERREIRA, 2010).

A generalização da Equação (2.25) para o LDA para  $M$  classes é dada por:

$$\mathbf{J}_C = \sum_i^M \sum_{j \neq i}^M (\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)^2 \odot \frac{1}{\boldsymbol{\sigma}_i^2 + \boldsymbol{\sigma}_j^2}. \quad (2.26)$$

### 2.3.3 Classificador Bayesiano

Considere o problema de decisão entre duas hipóteses, uma hipótese  $\mathcal{H}_0$ , referente à classe 0, e outra hipótese  $\mathcal{H}_1$ , referente à classe 1. Diversas técnicas



lineares e não lineares podem ser aplicadas para determinar a superfície de separação entre as hipóteses  $\mathcal{H}_0$  e  $\mathcal{H}_1$ . O classificador de Bayes (DUDA, HART, STORK, 2000; THEODORIDIS, KOUTROUMBAS, 2009; FERREIRA, 2010) realiza a classificação através da seguinte inequação:

$$\frac{p(\mathbf{x}|\mathcal{H}_1)}{p(\mathbf{x}|\mathcal{H}_0)} \stackrel{?}{\leq} \frac{P(\mathcal{H}_1)}{P(\mathcal{H}_0)} \quad (2.27)$$

em que  $\mathbf{x}$  é um vetor constituído de elementos aleatórios,  $P(\mathcal{H}_i)$  representa a probabilidade *a priori* e  $p(\mathbf{x}|\mathcal{H}_i)$  representa a função densidade de probabilidade (pdf - *probability density function*) condicional de obter  $\mathbf{x}$ , dado que a classe é  $\mathcal{H}_i$ . Se a função densidade de probabilidade condicional for gaussiana, ela pode ser expressa como:

$$p(\mathbf{x}|\mathcal{H}_i) = \frac{1}{(2\pi)^{\frac{L}{2}} |\boldsymbol{\Sigma}_i|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i)}, \quad (2.28)$$

onde  $\boldsymbol{\mu}_i = E\{\mathbf{x}\}$  é o vetor valor médio da classe  $\mathcal{H}_i$ ,  $\boldsymbol{\Sigma}_i$  é a matriz de covariância da mesma classe definida por

$$\boldsymbol{\Sigma}_i = E\{(\mathbf{x} - \boldsymbol{\mu}_i)(\mathbf{x} - \boldsymbol{\mu}_i)^T\}, \quad (2.29)$$

e  $|\boldsymbol{\Sigma}_i|$  denota o determinante de  $\boldsymbol{\Sigma}_i$ .

Supondo classes equiprováveis, onde  $P(\mathcal{H}_0) = P(\mathcal{H}_1) = 1/2$  e as pdf's dadas por (2.20), o classificador dado por (2.19) assume a forma:

$$\frac{|\Sigma_0|^{\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_b)^T \Sigma_1^{-1}(\mathbf{x}-\boldsymbol{\mu}_1)}}{|\Sigma_1|^{\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_a)^T \Sigma_0^{-1}(\mathbf{x}-\boldsymbol{\mu}_0)}} \cong 1, \quad (2.30)$$

onde  $\boldsymbol{\mu}_0$  e  $\boldsymbol{\mu}_1$  são os vetores valor médio das classes  $\mathcal{H}_0$  e  $\mathcal{H}_1$  respectivamente, e  $\boldsymbol{\Sigma}_0$  e  $\boldsymbol{\Sigma}_1$  são as matrizes de covariância das respectivas classes.

Sendo assim, a Equação à esquerda de (2.22) é aplicada ao vetor de parâmetros extraído e, se o valor encontrado for maior ou igual a 1, o sinal é dado como pertencente à classe 0, caso contrário o mesmo é dado como pertencente à classe 1.

## CAPÍTULO 3

O presente capítulo descreve o método proposto para o reconhecimento de face no contexto de detecção e identificação de criminosos.

### 3 METODOLOGIA

#### 3.1 Plataforma de Desenvolvimento

A plataforma de desenvolvimento escolhida para elaborar este projeto foi o MatLab<sup>®</sup>. A principal razão que motivou tal escolha foi o fato do software em questão oferecer uma vasta biblioteca de funções matemáticas e algoritmos numéricos, além de um *toolbox* próprio para processamento de imagens. Outro ponto forte é a simplicidade da linguagem, que acompanha a álgebra aplicada a vetores e matrizes, facilitando a implementação de métodos de reconhecimento de padrões. Além disso, oferece boa portabilidade, sendo compatível com *Windows*, *Linux*, *Mac OS* e também com sistemas móveis como *Android IOS*.

#### 3.2 Base de Dados

Para conduzir os experimentos e testar o sistema de reconhecimento facial proposto foi utilizado o banco de faces ORL (*Olivetti Research Laboratories*), desenvolvido nos laboratórios da Olivettiem Cambridge, Inglaterra (AT&T LABORATORIES CAMBRIDGE, 1994). Essa base de dados contém um conjunto de imagens faciais, extraídas de 40 indivíduos distintos, e para cada indivíduo há 10 imagens.

A Figura 3.1 ilustra as imagens do banco ORL. Estas imagens foram tomadas em momentos diferentes, variando a iluminação, expressões faciais

(olhos abertos/fechados, sorrindo/não sorrindo) e detalhes faciais (com óculos/sem óculos). Todas as imagens foram adquiridas contra um fundo escuro homogêneo com as pessoas na posição vertical, frontal (com tolerância para algum movimento lateral).

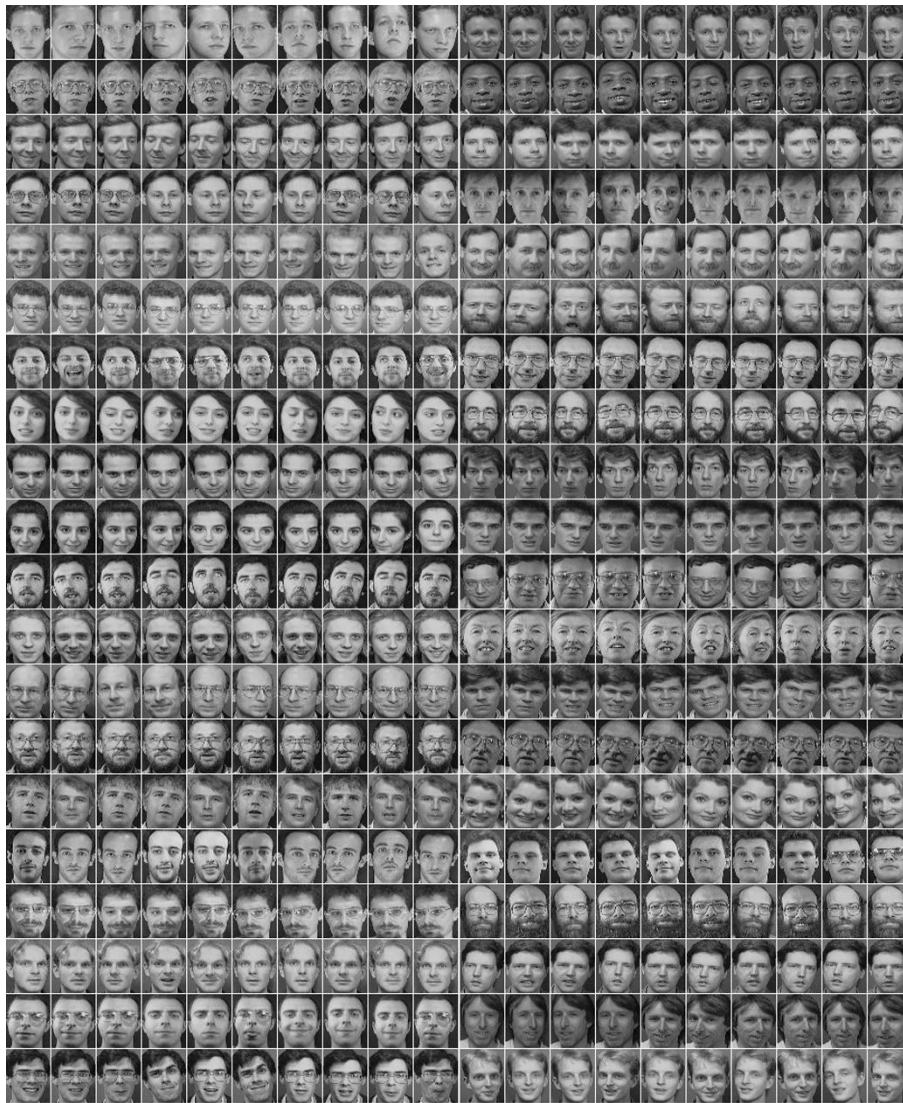


Figura 3.1 - Imagens do banco de dados de faces ORL  
Fonte: Adaptado de AT&T Laboratories Cambridge (1994)

As imagens estão em formato PGM e dimensão 92x112 pixels, com 256 níveis de cinza por pixel.

Esse banco de faces tem boa aceitação na literatura, tendo sido muito utilizado para validar métodos de reconhecimento de face (HAFED; LEVINE, 2001; BICEGO; CASTELLANI; MURINO, 2003; KUMAR; DEEPTI; PRABHAKAR, 2006; MATOS, 2008), e também para o treinamento de métodos de detecção (NEFIAN; HAYES, 2000; FLEURET; GEMAN, 2002; OMAIA, 2009).

Conforme discutido no Capítulo 1, foi considerado o cenário de identificação de criminosos por meio de reconhecimento de face através de imagens supostamente capturadas de câmeras de segurança pública. Sendo assim, pressupõe-se que as etapas de segmentação de imagens e detecção de faces já foram realizadas. Ademais, esse cenário leva à organização do banco de dados conforme descrito na próxima seção.

### **3.2.1 Organização dos Dados**

Para testar e avaliar o método proposto, dentro do contexto adotado de identificação de criminosos, o banco de imagens foi dividido em dois grupos, um grupo de faces que serão consideradas como faces de criminosos (grupo A) e o outro de faces de não criminosos (grupo B). Cumpre ressaltar que não há relação alguma entre as imagens do banco de dados utilizado e criminosos da vida real. Essa divisão de grupos e o uso do termo “criminoso” foram realizados apenas de maneira fictícia a fim de abordar o cenário proposto de identificação de criminosos por meio de reconhecimento de face.

O grupo A é composto por 15 indivíduos e o grupo B por 25. A divisão destes grupos foi feita de forma aleatória, seguindo uma distribuição uniforme no sorteio das faces.

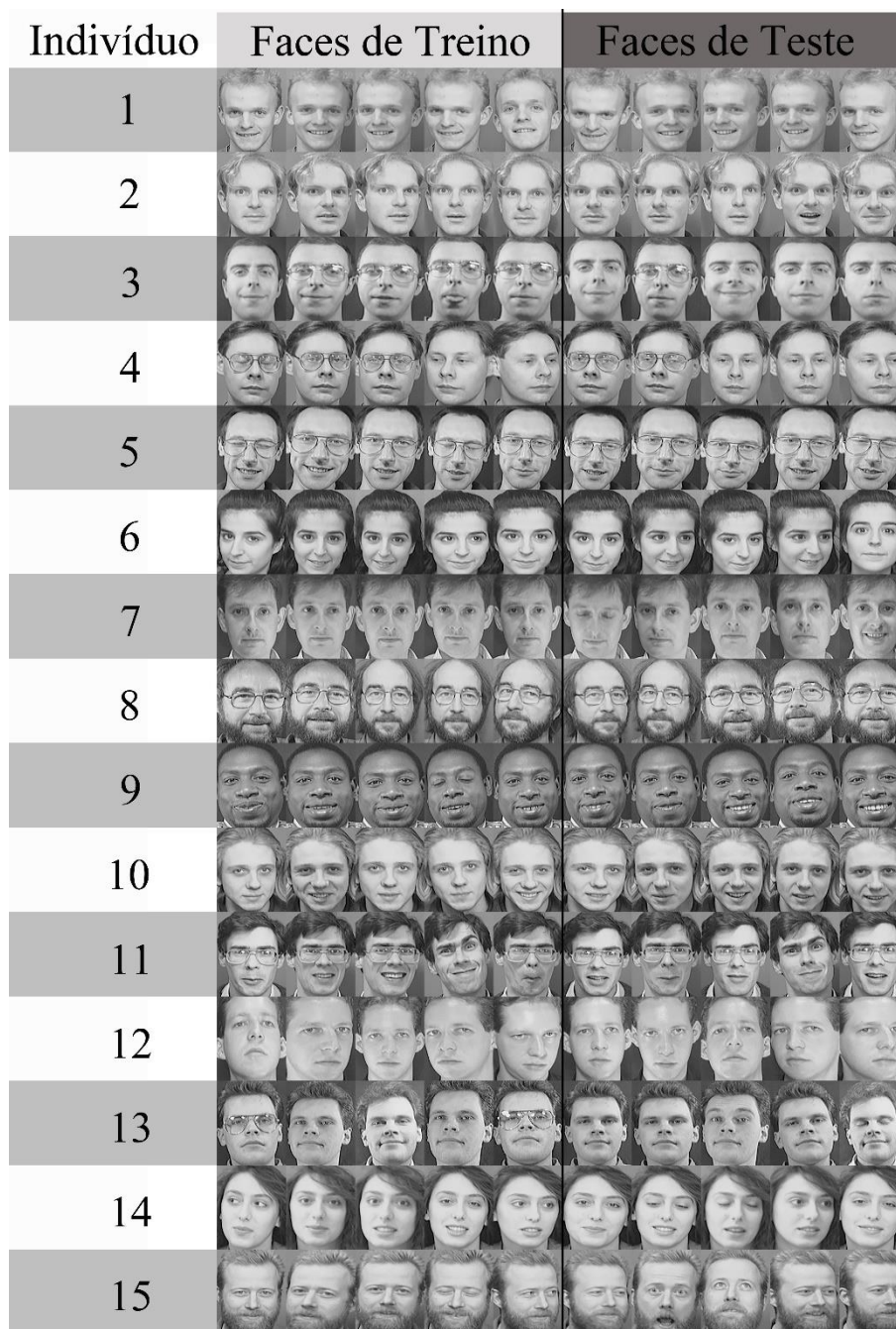


Figura 3.2 - Faces de criminosos sorteadas

O grupo A foi dividido em conjunto de treino e conjunto de teste. Como cada indivíduo é representado por dez faces, em poses diferentes (Figura 3.1), foram sorteadas aleatoriamente cinco faces para treino e cinco para teste. Portanto, o banco de dados de treinamento é composto por 75 (5x15) faces. As outras 75 faces serão usadas para teste. A Figura 3.2 ilustra as faces do grupo B, indivíduo a indivíduo, bem como as divisões entre dados de treino e de teste.

As faces do banco de dados do grupo B, que compreendem a um total de 250 faces, 25 indivíduos com 10 faces cada, serão usadas apenas na etapa de teste para avaliar o falso alarme do método proposto, já que os indivíduos que se deseja detectar e identificar são os do grupo A.

### **3.3 Método Proposto**

Os sistemas de reconhecimento de face que utilizam imagens que podem ou não conter mais de uma face requerem uma etapa de pré-processamento que deve realizar a segmentação das faces na imagem. Após essa segmentação, cada face presente na imagem é então recortada e enviada individualmente para o classificador, que finalmente fará a identificação da face.

Nesta dissertação, o foco será dado ao classificador e o estágio de segmentação das faces presentes na imagem, portanto, não será abordado.

O método de classificação proposto é representado, de forma simplificada, pelo diagrama em blocos da Figura 3.3. Será abordado o contexto de classificador especialista, em que um classificador será projetado para cada classe (indivíduos do grupo A).

No final, ou seja, após o processamento de uma face, uma lógica de decisão fará a combinação das saídas dos 15 classificadores (um para cada face do grupo A) e dará na saída a identificação da face, se a face processada pertencer

a algum dos 15 indivíduos do grupo A; ou apontará a face processada como não pertencente ao grupo A, não dando, nesse caso, a identificação da mesma.

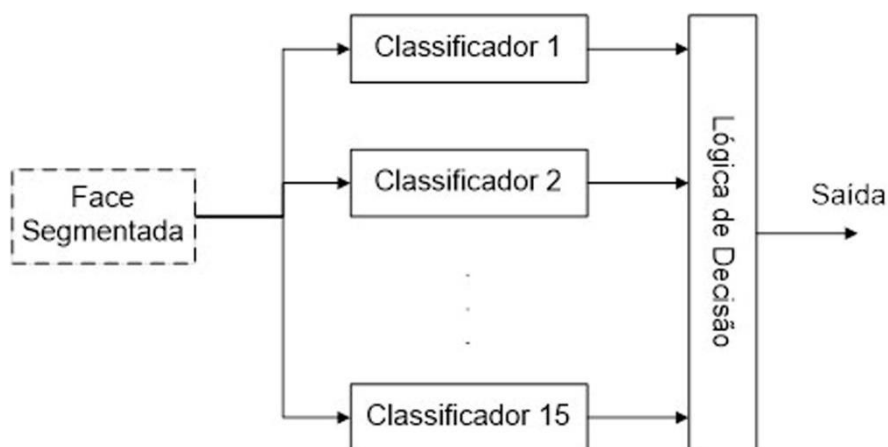


Figura 3.3 - Diagrama em blocos da implementação do sistema

A lógica de decisão implementada é uma porta lógica “OR”. Isso significa que se qualquer um dos classificadores classificar a face processada como um indivíduo pertencente ao grupo A, a saída da lógica de decisão indicará presença (detecção) de indivíduo pertencente ao grupo A. Todavia, para a identificação da face processada é feita a leitura das saídas de cada um dos classificadores. Se mais de um classificador classificar a face como pertencente ao grupo A, serão indicadas na saída tais classificações de forma que a identificação final possa ser feita visualmente por um especialista (policia).

Na saída do método proposto, duas informações são disponibilizadas:

- (i) se a face processada é pertencente ao banco de dados do grupo A ou não, e;
- (ii) se a face processada for pertencente ao grupo A, é apresentada a identificação da mesma (dentre os 15 indivíduos do grupo).



A informação (i) pode ser referida como detecção e a informação (ii) como classificação.

A implementação de cada um dos classificadores foi dividida em três etapas: extração de características, seleção de características (redução da dimensionalidade) e classificação, que está representada de forma simplificada, pelo diagrama em blocos da Figura 3.4.

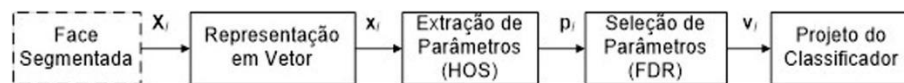


Figura 3.4 - Diagrama em blocos do detalhamento de cada classificador

O primeiro bloco da Figura 3.4 é a face detectada. Essa face, quando lida pelo *software* MatLab, é representada por uma matriz  $\mathbf{X}_i$ , na dimensão dos pixels (92x112). Essa matriz é então transformada em um vetor, seguindo a ordem de linha por linha, conforme mostrado na Figura 3.5.

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \dots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}}_{\mathbf{X}_i} \Rightarrow \underbrace{[a_{11} \ a_{12} \ \dots \ a_{1n} \ a_{21} \ a_{22} \ \dots \ a_{2n} \ \dots \ a_{n1} \ a_{n2} \ \dots \ a_{nm}]}_{\mathbf{x}_i}$$

Figura 3.5 - Face representada pelo vetor

O passo seguinte compreende a extração dos parâmetros utilizando Estatística de Ordem Superior. Extraíram-se os cumulantes de 2ª, 3ª e 4ª ordens, conforme descrito na seção 2.2.1.

A saída é o vetor de parâmetros  $\mathbf{p}_i$ , que é então enviado para o algoritmo de classificação. Como classificador foi implementado o classificador de Bayes, supondo distribuições gaussianas e probabilidades *a priori* estimadas

empiricamente (THEODORIDIS; KOUTROUMBAS, 2009), devido à sua simplicidade e obtenção de bons resultados.

Cumpramos ressaltar que a etapa de seleção de parâmetros e análise de redundância é realizada apenas na etapa de projeto do método. Na fase operacional, apenas os cumulantes presentes no vetor de parâmetros  $\mathbf{p}_i$ , que foram previamente selecionados, são calculados.

## CAPÍTULO 4

Este capítulo apresenta os resultados do método proposto de uma forma qualitativa e quantitativa, sobre os módulos detector e identificador de faces.

### 4 RESULTADOS

Após a extração dos cumulantes de 2<sup>a</sup>, 3<sup>a</sup> e 4<sup>a</sup> ordens, aplicou-se o Discriminante Linear de *Fisher* (Equação 2.25) para duas classes, ou seja, uma classe contra todas as outras 14 classes, para cada classe do grupo A.

Cada face é representada por um total de  $N = 10.304$  pixels. Portanto, ao extrair os cumulantes de 2<sup>a</sup>, 3<sup>a</sup> e 4<sup>a</sup> ordens obtém-se um total de  $3N$  pixels. No entanto, os cumulantes de ordem 2 são simétricos e, portanto, a segunda metade dos mesmos foi eliminada, o que leva a um total de  $5N/2$  (igual a 25.760) parâmetros (cumulantes de 2<sup>a</sup>, 3<sup>a</sup> e 4<sup>a</sup> ordens).

As Figuras 4.1, 4.2 e 4.3 apresentam as médias das assinaturas das faces do banco de dados de treino, para os cinco primeiros indivíduos da Figura 3.2, construídas pelos cumulantes de segunda, terceira e quarta ordens, respectivamente. O eixo das abscissas foi limitado de 0 a 1000 para não comprometer a visualização. Foram mostradas as médias dos cumulantes para apenas cinco das faces do banco de dados por simplificação.

Pode-se inferir que há pouca diferença entre as assinaturas das faces (diferentes poses) de um mesmo indivíduo, visto que o desvio padrão está próximo do valor da média obtido, exceto para as assinaturas construídas pelos cumulantes de terceira ordem do indivíduo 5 (Figura 4.2 (e)). As médias dos cumulantes de segunda e quarta ordens apresentam variações mais acentuadas de um indivíduo para o outro em relação à sua amplitude, enquanto que os de terceira ordem apresentam variações acentuadas na amplitude e forma das médias dos

cumulantes, o que os caracterizam como mais discriminativos no que tange à classificação.

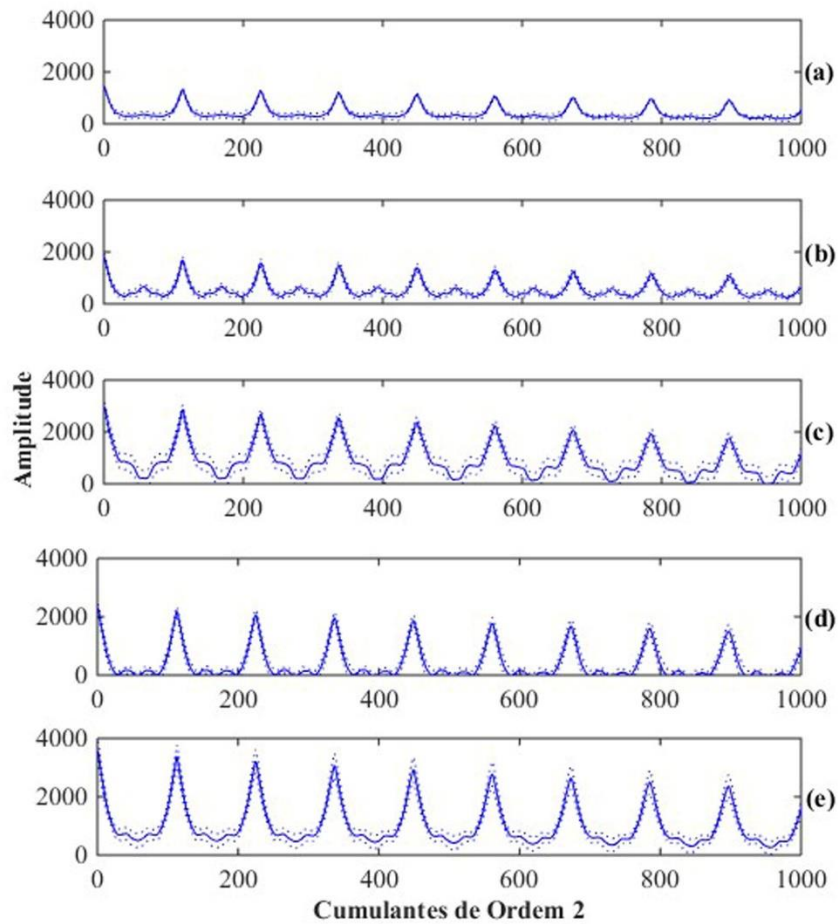


Figura 4.1 - Média dos Cumulantes de 2ª Ordem (linha contínua)  $\pm$  desvio padrão (linha pontilhada): (a) indivíduo 1, (b) indivíduo 2, (c) indivíduo 3, (d) indivíduo 4, (e) indivíduo 5

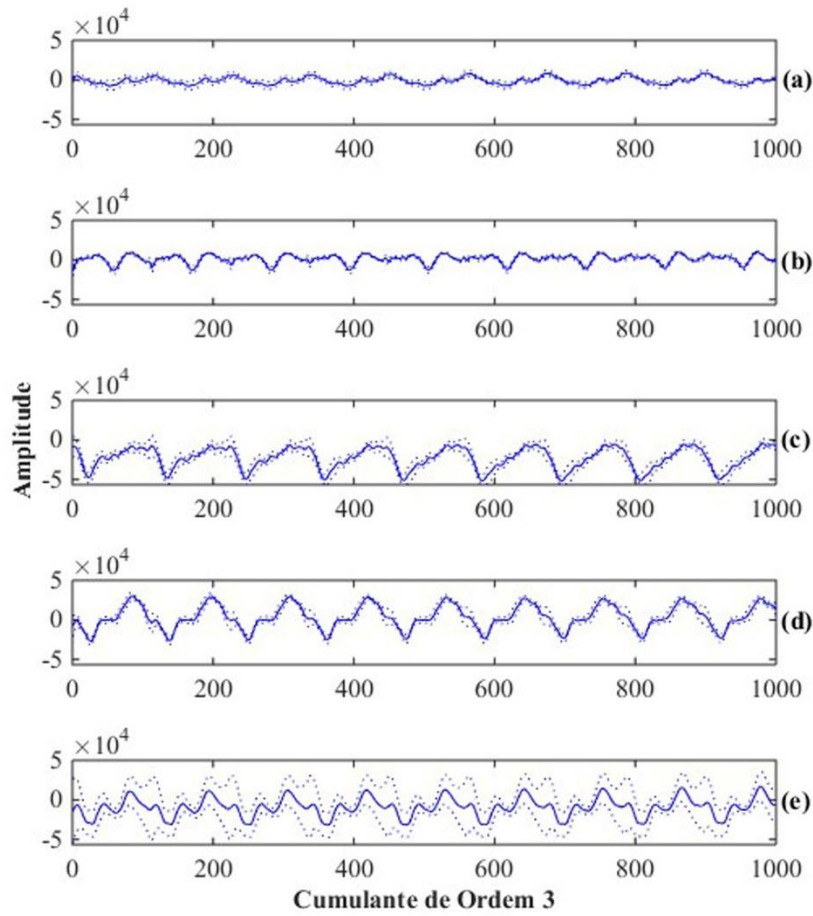


Figura 4.2 - Média dos Cumulantes de 3ª Ordem (linha contínua)  $\pm$  desvio padrão (linha pontilhada): (a) indivíduo 1, (b) indivíduo 2, (c) indivíduo 3, (d) indivíduo 4, (e) indivíduo 5

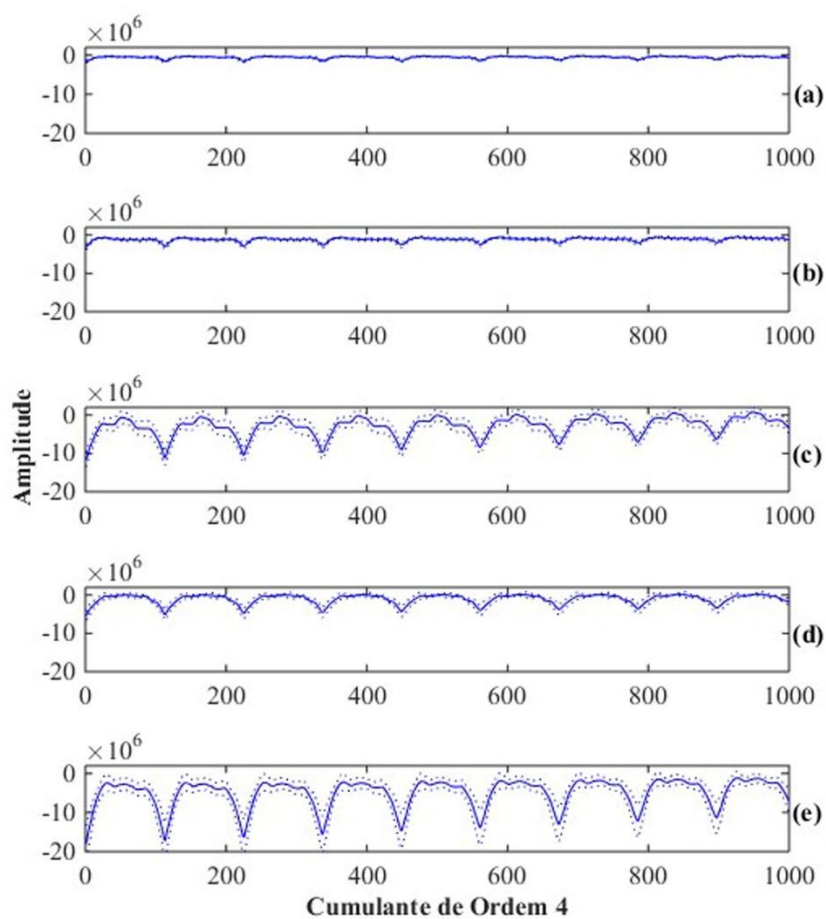


Figura 4.3 - Média dos Cumulantes de 4<sup>a</sup> Ordem (linha contínua)  $\pm$  desvio padrão (linha pontilhada): (a) indivíduo 1, (b) indivíduo 2, (c) indivíduo 3, (d) indivíduo 4, (e) indivíduo 5

Dentre os 25.760 cumulantes extraídos, foram selecionados os 1000 mais discriminantes de acordo com o critério FDR. Em seguida, foi aplicada a correlação linear entre os parâmetros para eliminar a redundância, em que os

cumulantes que apresentaram correlação com seus pares superior a 0,8 foram descartados.

A Tabela 4.1 mostra a quantidade de cumulantes de ordens 2, 3 e 4, selecionados para cada classificador, gerados após a aplicação do FDR (critério para duas classes) e eliminação da redundância usando o coeficiente de correlação. Cumpre ressaltar que a numeração dos classificadores mostrada na tabela segue a numeração dos indivíduos da Figura 3.2.

Tabela 4.1 - Quantidade de Cumulantes de ordens 2, 3 e 4, por classificador

Classificador	Cumulantes			Total
	2ª Ordem	3ª Ordem	4ª Ordem	
1	2	0	0	2
2	5	2	2	9
3	4	8	4	16
4	0	4	0	4
5	1	4	5	10
6	1	2	2	5
7	0	5	0	5
8	4	5	0	9
9	4	0	0	4
10	3	5	5	13
11	3	6	0	9
12	1	5	0	6
13	4	4	11	19
14	1	2	2	5
15	3	2	3	8
<b>Total</b>	<b>36</b>	<b>54</b>	<b>34</b>	<b>124</b>

Fonte: Próprio autor

Observa-se que do total de cumulantes utilizados, os de terceira ordem foram os mais presentes, 54 de 124. Esse resultado está de acordo com os resultados mostrados na Figura 4.3, que apresenta os cumulantes de terceira ordem como os mais representativos em comparação com os de segunda e quarta ordens.

Todavia, pode-se observar na Tabela 4.1 que para cada classificador há uma combinação de cumulantes diferente, o que indica que cada tipo de face é mais bem discriminada por um conjunto específico de cumulantes, que se diferenciam dos que representam as outras faces.

Observa-se também que a dimensão dos vetores de assinaturas de cada face variou de 2 a 19. Essa dimensão foi definida automaticamente ao aplicar o critério de Fisher e a correlação para eliminar redundância com corte em 0,8.

Importante ressaltar que, dada a dimensão original das faces,  $N = 10.304$  pixels, a redução de dimensão para 2, no melhor caso, e 19, no pior caso, é um ganho considerável no que tange a complexidade computacional do método.

A Figura 4.4 mostra o discriminante de *Fisher* para os cinco primeiros indivíduos (segundo a ordem listada na Figura 3.2), em que os primeiros 5.152 parâmetros representam os cumulantes de ordem 2, os próximos 10.304 os de ordem 3, e os últimos 10.304 parâmetros referem-se aos cumulantes de ordem 4. Observa-se que para cada classe há um cumulante que se destaca com amplitudes de  $J$  maiores, por exemplo, no *Fisher* obtido para os indivíduos 1 e 2, mostrado em (a) e (b), respectivamente, o cumulante que mais se destaca é o de 2ª ordem, enquanto que no indivíduo 4, em (d), os cumulantes de 3ª ordem apresentaram melhores resultados, e no indivíduo 5, em (e), são os cumulantes de 4ª ordem que apresentaram maior relevância.

Observando os resultados mostrados na Tabela 4.1, para os cinco primeiros indivíduos, pode-se concluir que os resultados mostrados na Figura 4.4 são concordantes, quando se verifica qual foi a ordem do cumulante dominante na representação do indivíduo. No entanto, deve-se ressaltar que há muita redundância entre os cumulantes extraídos das faces e que o discriminante de Fisher indica os mais discriminativos sem levar em consideração a redundância, ou seja, sem descartar os redundantes, e os resultados da Tabela 4.1 obtidos após a eliminação de redundância, utilizando a correlação linear.



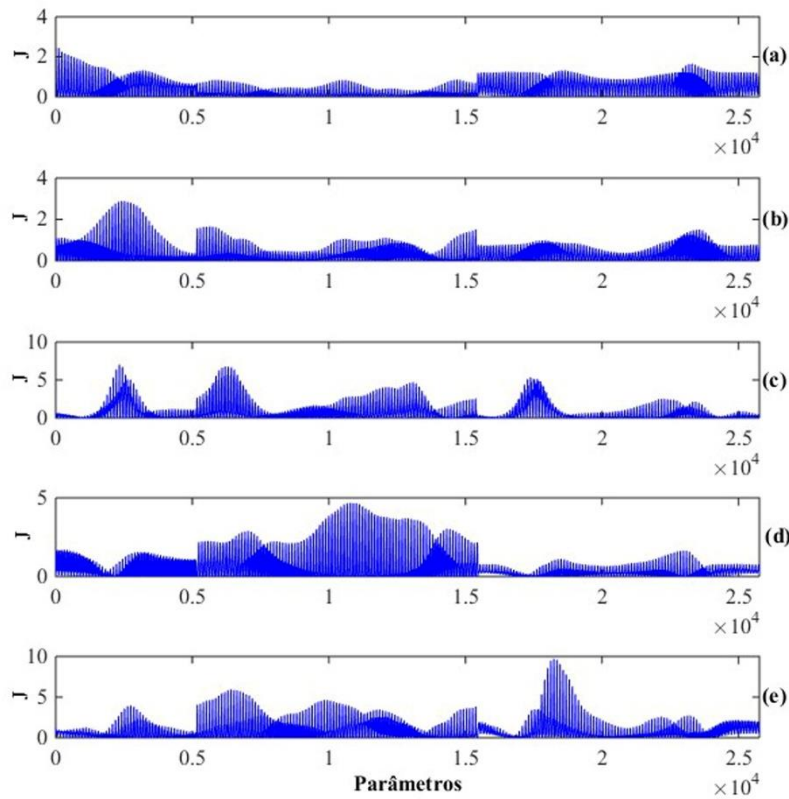


Figura 4.4 - Discriminante de Fisher para os parâmetros das faces dos indivíduos 1 (a), 2 (b), 3 (c), 4 (d) e 5 (e)

As Figuras 4.5–4.9 apresentam o espaço de parâmetros construído pelos dois cumulantes mais discriminativos para os classificadores 1-5, respectivamente, para os dados do conjunto de treino. Observa-se que não há uma separação linear entre as classes, o que indica que é um problema de classificação não simples. Por simplificação, foram mostrados os espaços de parâmetros apenas para os cinco primeiros classificadores, referentes às cinco primeiras classes (segundo a Figura 3.2). Contudo, dentre os quinze espaços de parâmetros, três

(referente aos indivíduos 9, 12 e 14, conforme numeração indicada na Figura 3.2) apresentaram separação linear com apenas dois parâmetros.

Interessante observar que não há coincidência entre os dois melhores parâmetros de cada classe, o que mostra não haver um padrão para a extração de cumulantes de faces para fins de classificação.

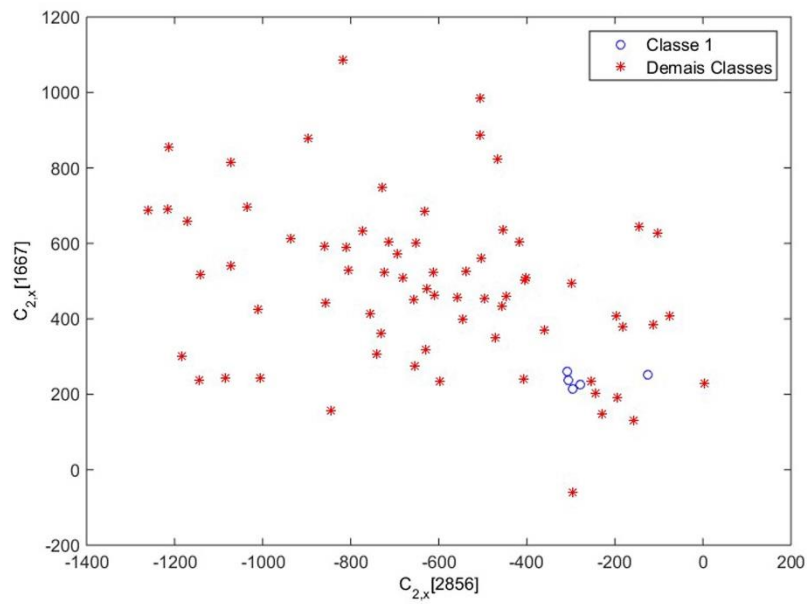


Figura 4.5 - Espaço de parâmetros para o classificador 1

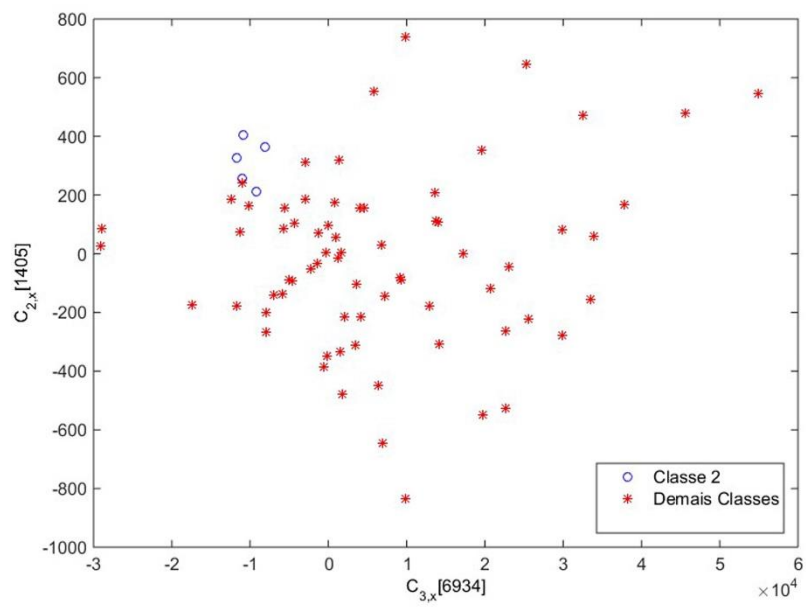


Figura 4.6 - Espaço de parâmetros para o classificador 2

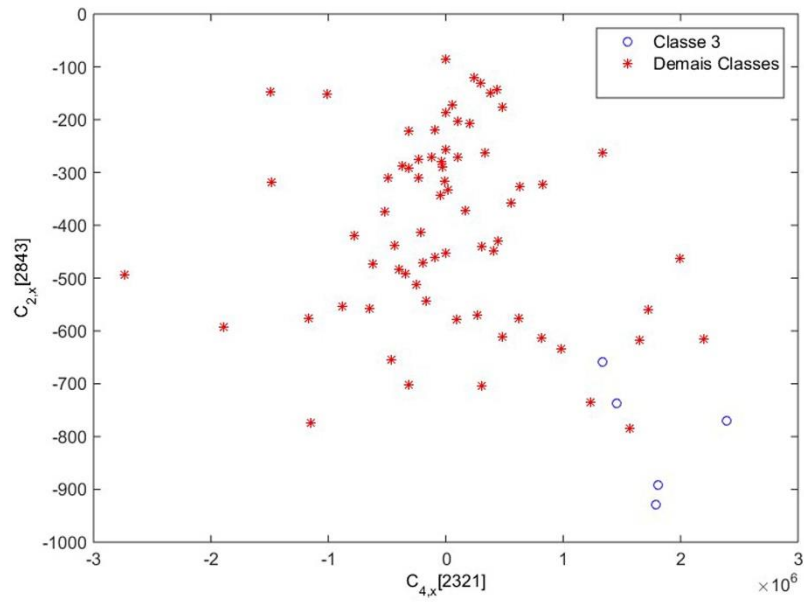


Figura 4.7 - Espaço de parâmetros para o classificador 3

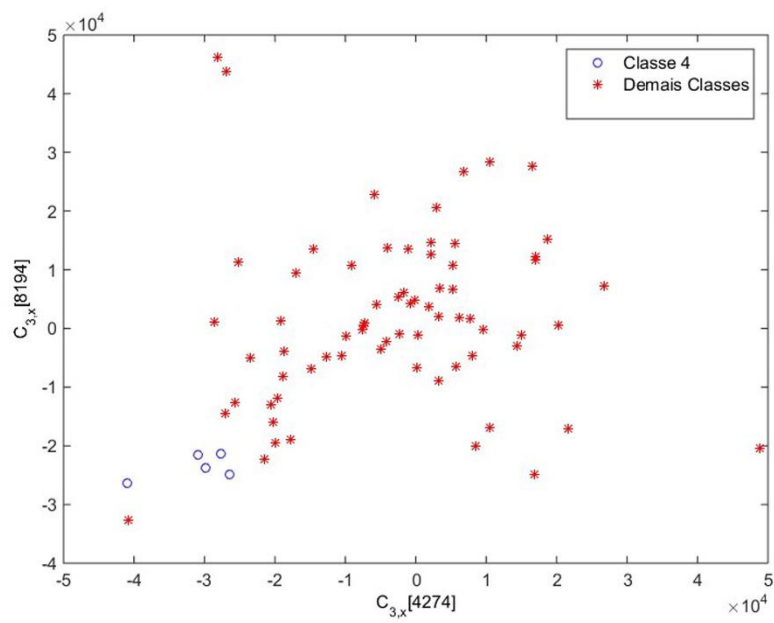


Figura 4.8 - Espaço de parâmetros para o classificador 4

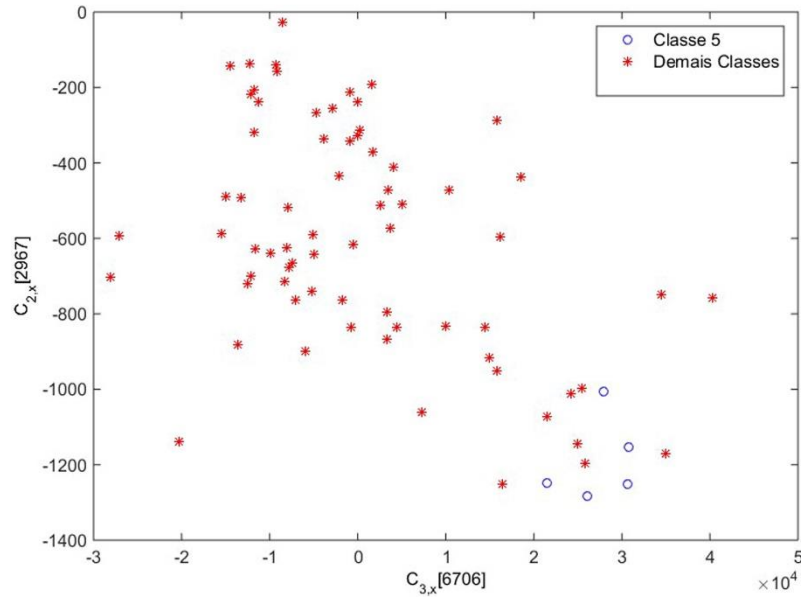


Figura 4.9 - Espaço de parâmetros para o classificador 5

A Tabela 4.2 mostra o desempenho de cada classificador para o conjunto de dados de treino. Como os classificadores foram projetados de forma independente um do outro, fazendo-se uma face contra todas as outras, os resultados foram mostrados em função dos acertos percentuais para o indivíduo  $i$  (faces do indivíduo que se deseja identificar) e para outros indivíduos (faces dos outros indivíduos).

Como os dados que compõem a classe dos outros indivíduos são mais heterogêneos, devido à diversidade entre os dados que são advindos de diferentes faces, o classificador apresentou maior dificuldade, no entanto, acertos superiores a 95,7 % foram alcançados.

Tabela 4.2 - Resultado do treino para cada classificador

<b>Classificador</b>	<b>Indivíduo (%)</b>	<b>Outros Indivíduos (%)</b>
<b>1</b>	100,0	95,7
<b>2</b>	100,0	98,6
<b>3</b>	100,0	100,0
<b>4</b>	100,0	100,0
<b>5</b>	100,0	100,0
<b>6</b>	100,0	100,0
<b>7</b>	100,0	98,6
<b>8</b>	100,0	100,0
<b>9</b>	100,0	100,0
<b>10</b>	100,0	100,0
<b>11</b>	100,0	98,6
<b>12</b>	100,0	98,6
<b>13</b>	100,0	100,0
<b>14</b>	100,0	100,0
<b>15</b>	100,0	100,0
<b>Total</b>	<b>100,0</b>	<b>99,3</b>

Fonte: Próprio autor

A Tabela 4.3 mostra o desempenho de cada classificador para o conjunto de dados de teste. O método apresentou melhor generalização para a classe de outros indivíduos, devido ao pouco número de faces de treino disponíveis para a classe de indivíduos, apenas 5 faces por indivíduo.

A Tabela 4.4 mostra a classificação final de acertos dos indivíduos, considerando o sistema completo representado pelo diagrama em blocos da Figura 3.3. Os resultados apontaram um total de acertos de 72%, o que é considerado um bom resultado, dado o pouco número de faces disponíveis para treinamento do método.

Tabela 4.3 - Resultado de teste para cada classificador

<b>Classificador</b>	<b>Indivíduo (%)</b>	<b>Outros Indivíduos (%)</b>
<b>1</b>	60,0	95,7
<b>2</b>	80,0	100,0
<b>3</b>	100,0	100,0
<b>4</b>	40,0	98,6
<b>5</b>	80,0	100,0
<b>6</b>	40,0	100,0
<b>7</b>	60,0	100,0
<b>8</b>	100,0	100,0
<b>9</b>	80,0	100,0
<b>10</b>	60,0	100,0
<b>11</b>	100,0	95,7
<b>12</b>	100,0	100,0
<b>13</b>	100,0	100,0
<b>14</b>	80,0	100,0
<b>15</b>	60,0	100,0
<b>Total</b>	<b>76,0</b>	<b>99,3</b>

Fonte: Próprio autor

A detecção final de indivíduos pertencentes à Classe A (considerada como a classe de criminosos) alcançada pelo método proposto é representada na Tabela 4.5. A fim de comparar o desempenho de diferentes classificadores, os classificadores KNN e SVM (THEODORIDIS; KOUTROUMBAS, 2009) foram aplicados ao problema abordado. O classificador de Bayes alcançou uma taxa de detecção intermediária seguida do menor falso alarme. Os classificadores KNN e SVM apresentaram taxas de falso alarme acima de 48% e, portanto, não são recomendados para essa aplicação.



Tabela 4.4 - Resultado de classificação final

<b>Classificador</b>	<b>Acertos (%)</b>
<b>1</b>	60,0
<b>2</b>	60,0
<b>3</b>	100,0
<b>4</b>	40,0
<b>5</b>	80,0
<b>6</b>	40,0
<b>7</b>	60,0
<b>8</b>	100,0
<b>9</b>	60,0
<b>10</b>	60,0
<b>11</b>	80,0
<b>12</b>	100,0
<b>13</b>	100,0
<b>14</b>	80,0
<b>15</b>	60,0
<b>Total</b>	<b>72,0</b>

Fonte: Próprio autor

Tabela 4.5 – Resultado de detecção final

<b>Método de Classificação</b>	<b>Deteção de Indivíduo da Classe A (%)</b>	<b>Falso Alarme (%)</b>
<b>Bayes</b>	81,3	25,2
<b>KNN (K = 3)</b>	92,0	52,4
<b>KNN (K = 1)</b>	88,0	54,0
<b>KNN (K = 5)</b>	89,3	54,0
<b>SVM (função kernel linear)</b>	92,0	72,8
<b>SVM (função kernel RBF)</b>	74,7	48,0

Fonte: Próprio autor

## **CAPÍTULO 5**

### **CONCLUSÃO**

A presente dissertação de mestrado propôs um método de reconhecimento de face usando estatísticas de ordem superior. Foram explorados os cumulantes de segunda, terceira e quarta ordens na construção de assinaturas das faces. Foi considerado um cenário em que o objetivo era identificar criminosos a fim de desenvolver um sistema automático para auxiliar o trabalho da PM. Dessa forma, o método proposto pode ser útil na triagem de faces presentes em imagens captadas por câmeras de segurança pública.

Os resultados são promissores, com taxas de detecção e classificação superiores a 70%. O trabalho apresentou uma inovação no uso de EOS para processamento de face e, com o objetivo de avaliar a capacidade das mesmas para reconhecimento de face, nenhum pré-processamento das imagens foi realizado. Espera-se alcançar desempenhos superiores com pré-processamentos adequados da área de processamento de imagens. Ademais, o ganho em redução de dimensão com o método proposto foi muito alto (de 10.304 parâmetros para 19, no máximo).

Em trabalhos futuros espera-se investigar o ganho em desempenho do método proposto com o acréscimo de ferramentas sofisticadas de processamento de imagens, bem como aplicar Algoritmos Genéticos na seleção de parâmetros, visto que esse método tem a capacidade de aperfeiçoar o desempenho dos classificadores.

## REFERÊNCIAS

AMARAL, Vagner do. **Extração e Comparação de Características Locais e Globais para o Reconhecimento Automático de Imagens de Faces**. 2011. 78 f. Dissertação (Mestrado) - Curso de Engenharia Elétrica, Centro Universitário da Fei, São Bernardo do Campo, 2011.

ARRUDA, Benedito Alencar de. **Comparação entre algoritmo genético, rede neural artificial e análise de componentes principais no reconhecimento de faces**. 2013. 237 f. Tese (Doutorado) - Curso de Ciências, Pós-graduação em Engenharia Elétrica, Universidade Federal de Uberlândia, Uberlândia, 2013.

AT&T LABORATORIES CAMBRIDGE (Cambridge). Cambridge University Computer Laboratory (Ed.). **The Digital Technology Group**. 1994. Disponível em: <<http://www.cl.cam.ac.uk/research/dtg/attarchive/facesataglace.html>>. Acesso em: 20 jan. 2014.

BAEK, Kyungim et al. PCA vs. ICA: A comparison on the FERET data set. In: INTERNATIONAL CONFERENCE ON COMPUTER VISION, 6., 2002, Durham. **Pattern Recognition and Image Processing in conjunction with the 6th JCIS**. Durham: [s.l.], 2002. p. 8 - 14.

BARTLETT, Marian Stewart. Independent Component Representations for Face Recognition. In: BARTLETT, Marian Stewart. **Face Image Analysis by Unsupervised Learning**. [s.l.]: Springer Us, 2001. Cap. 3. p. 39-67.

BELHUMEUR, Peter N.; HESPANHA, João P.; KRIEGMAN, David J. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. **IEEE Transactions On Pattern Analysis And Machine Intelligence**, Washington, v. 19, n. 7, p.711-720, jul. 1997. Institute of Electrical & Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/34.598228>.

BICEGO, M.; CASTELLANI, U.; MURINO, V.. Using hidden Markov models and wavelets for face recognition. **12th International Conference On Image Analysis And Processing, 2003.proceedings.**, [s.l.], p.52-56, 19 set. 2003.

Institute of Electrical & Electronics Engineers (IEEE).  
<http://dx.doi.org/10.1109/iciap.2003.1234024>.

BORDIGNON, Luiz Alberto. **Identificação de sementes utilizando visão computacional**. 2015. 75 f. Dissertação (Mestrado) - Curso de Sistemas Eletrônicos, Departamento de Engenharia Elétrica, Universidade Federal do Paraná, Curitiba, 2015.

BOUGHRARA, Hayet et al. Face recognition based on perceived facial images and multilayer perceptron neural network using constructive training algorithm. **IET Computer Vision**, [s.l.], v. 8, n. 6, p.729-739, 1 dez. 2014. Institution of Engineering and Technology (IET). <http://dx.doi.org/10.1049/iet-cvi.2013.0294>.

BOUGHRARA, Hayet; CHTOUROU, Mohamed; AMAR, Chokri Ben. MLP neural network based face recognition system using constructive training algorithm. **2012 International Conference On Multimedia Computing And Systems**, Tangier, p.233-238, 12 maio 2012. Institute of Electrical & Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/icmcs.2012.6320263>.

CARDOSO, Jean-françois. Infomax and maximum likelihood for source separation. **IEEE Letters On Signal Processing**, [s.l.], v. 4, n. 4, p.112-114, abr. 1997.

CARVALHO, Ederson de Assis; DINIZ, Alexandre Alves. **Projeto Olho Vivo: a íris dos olhos da segurança pública, uma análise geográfica**. 2008. 100 f. Dissertação (Mestrado) - Curso de Tratamento da Informação Espacial, Programa de Pós-graduação em Geografia, Pontifícia Universidade Católica de Minas Gerais, Belo Horizonte, 2008.

CHELLAPPA, Rama; WILSON, C.I.; A SIROHEY, Saad. Human and machine recognition of faces: a survey. **Proceedings Of The IEEE**, [s.l.], v. 83, n. 5, p.705-741, maio 1995. Institute of Electrical & Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/5.381842>.

DRAPER, Bruce A. et al. Recognizing faces with PCA and ICA. **Computer Vision And Image Understanding**, New York, v. 91, n. 1-2, p.115-137, jul. 2003. Elsevier BV. [http://dx.doi.org/10.1016/s1077-3142\(03\)00077-8](http://dx.doi.org/10.1016/s1077-3142(03)00077-8).

DUAN, Hong; YAN, Ruohe; LIN, Kunhui. Research on Face Recognition Based on PCA. **2008 International Seminar On Future Information Technology And Management Engineering**, Leicestershire, United Kingdom, p.29-32, 20 nov. 2008. Institute of Electrical & Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/fitme.2008.115>.

DUDA, Richard O.; HART, Peter E.; STORK, David G.. **Pattern Classification**. 2. ed. [s.l.]: Wiley-interscience, 2000. 680 p.

DUDLEY, Brier. **E3: New info on Microsoft's Natal - how it works, multiplayer and PC versions**. 2009. Disponível em: <http://www.seattletimes.com/html/technologybrierdudleysblog/>. Acesso em: 15 maio 2015.

FERREIRA, Danton Diego et al. Exploiting Higher-Order Statistics Information for Power Quality Monitoring. In: EBERHARD, A. **Power Quality: Higher-order statistics**. New York: Intech, 2011. Cap. 17. p. 346-362.

FERREIRA, Danton Diego. **Análise de Distúrbios Elétricos em Sistemas de Potência**. 2010. 210 f. Tese (Doutorado) - Curso de Engenharia Elétrica, Pós-graduação em Engenharia Elétrica, Coppe, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.

FLEURET, F.; GEMAN, D. Fast Face Detection With Precise Pose Estimation. **Pattern Recognition, 2002. Proceedings. 16th International Conference on**. Washington, DC, USA: IEEE Computer Society. 2002. p. 235-238.

HAFED, Ziad M.; LEVINE, Martin D.. Face Recognition Using the Discrete Cosine Transform. **International Journal Of Computer Vision**, Hingham, v. 43, n. 3, p.167-188, ago. 2001. Springer Science + Business Media. <http://dx.doi.org/10.1023/a:1011183429707>.

HAYKIN, Simon. **Redes Neurais: Princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001. 908 p. Tradução de Paulo Martins Engel.

HYVÄRINEN, Aapo; KARHUNEN, Juha; OJA, Erkki. **Independent Component Analysis**. New York: A Wiley-interscience Publication, 2001. 504 p.

JOLLIFFE, I.T. **Principal Component Analysis**. 2. ed. New York: Springer, 2002. 488 p. (Springer Series in Statistics).

KUMAR, S. A. S.; DEEPTI, D. R.; PRABHAKAR, B. Face Recognition Using Pseudo-2D Ergodic HMM. **2006 IEEE International Conference On Acoustics Speed And Signal Processing Proceedings**, Toulouse, v. 2. 19 maio 2006. Institute of Electrical & Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/icassp.2006.1660356>.

LIU, Chengjun; WECHSLER, Harry. Comparative Assessment of Independent Component Analysis (ICA) for Face Recognition. In: INTERNATIONAL CONFERENCE ON AUDIO AND VIDEO BASED BIOMETRIC PERSON AUTHENTICATION, 2., 1999, Washington. **Appears in the Second International Conference on Audio- and Video-based Biometric Person Authentication**. Washington: AVBPA'99, 1999. p. 22 - 24.

MARQUES, Cristiano Augusto Gomes; FERREIRA, Danton Diego; RIBEIRO, Moises Vidal. Wavelet e EOS para Detecção de Distúrbios em Aplicações de Qualidade de Energia: Análise Comparativa. In: XVII CONGRESSO BRASILEIRO DE AUTOMÁTICA, XVII., 2008, Juiz de Fora. **Anais do XVII Congresso Brasileiro de Automática**. Juiz de Fora, 2008.

MARQUES, Ion. **Face Recognition Algorithms**. 2010. 78 f. Proyectos Fin de Carrera, Universidad del Pais Vasco - Euskal Herriko Unibersitatea, Bizkaia, 2010.

MASSY, Kevin. **Toyota develops eyelid-monitoring system**. 2008. Disponível em: <<http://www.cnet.com/roadshow/news/toyota-develops-eyelid-monitoring-system/>>. Acesso em: 15 maio 2015.

MATOS, Fabrícia Medeiros de S.. **Reconhecimento de faces utilizando a transformada cosseno discreta**. 2008. 94 f. Dissertação (Mestrado) - Curso de Informática, Departamento de Informática, Universidade Federal da Paraíba, João Pessoa, 2008.

MCCULLOCH, Warren S.; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. **The Bulletin Of Mathematical Biophysics**, Cambridge, v. 5, n. 4, p.115-133, dez. 1943. Springer Science + Business Media. <http://dx.doi.org/10.1007/bf02478259>.

MCWHERTOR, Michael. **Sony Spills More PS3 Motion Controller Details To Devs**. 2009. Disponível em: <<http://kotaku.com/5297265/sony-spills-more-ps3-motion-controller-details-to-devs>>. Acesso em: 2 maio 2015.

MEDEIROS, Luciano Xavier. **Reconhecimento Facial Utilizando Análise de Componentes Principais e Algoritmos Genéticos em Imagens Segmentadas**. 2012. 156 f. Tese (Doutorado) - Curso de Engenharia Elétrica, Pós-graduação em Engenharia Elétrica, Universidade Federal de Uberlândia, Uberlândia, 2012.

MENDEL, J. M. Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. **Proceedings Of The IEEE**, [s.l.], v. 79, n. 3, p.278-305, mar. 1991. Institute of Electrical & Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/5.75086>.

NAVES, Raphael; BARBOSA, Bruno Henrique Groenner; FERREIRA, Danton Diego. Classificação de Sons Pulmonares Utilizando Estatísticas de Ordem Superior e Árvores de Decisão. In: XX CONGRESSO BRASILEIRO DE AUTOMÁTICA, XX., 2014, Belo Horizonte. **Anais do XX Congresso Brasileiro de Automática**. Belo Horizonte, 2014. p. 1196 - 1202.

NEFIAN, A.v.; HAYES, M.h.. Maximum likelihood training of the embedded HMM for face detection and recognition. **Proceedings 2000 International Conference On Image Processing (cat. No.00ch37101)**, Vancouver, v. 1, p.33-36, 13 set. 2000. Institute of Electrical & Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/icip.2000.900885>.

NEFIAN, Ara V. **Statistical Approaches To Face Recognition**. PhD Thesis, Georgia Institute Of Technology, 1996.

NIKIAS, Chrysostomos L.; PETROPULU, Athina P.. **Higher-order Spectra Analysis: A Nonlinear Signal Processing Framework**. 6. ed. New Jersey: Ptr Prentice Hall, 1993. 537 p.

NIXON, Mark. Eye Spacing Measurement for Facial Recognition. **Applications Of Digital Image Processing VIII**, San Diego, p.279-283, 19 dez. 1985. SPIE-Intl Soc Optical Eng. <http://dx.doi.org/10.1117/12.966515>.

OMAIA, Derzu. **Um sistema para detecção e reconhecimento de face em vídeo utilizando a transformada cosseno discreta**. 2009. 92 f. Dissertação (Mestrado) - Curso de Ciência da Computação, Informática, Universidade Federal da Paraíba, João Pessoa, 2009.

PAONE, Jeffrey R. et al. Double Trouble: Differentiating Identical Twins by Face Recognition. **IEEE Trans.inform.forensic Secur.**, [s.l.], v. 9, n. 2, p.285-295, fev. 2014. Institute of Electrical & Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/tifs.2013.2296373>.

RIBEIRO, Moisés V et al. Detection of Disturbances in Voltage Signals for Power Quality Analysis Using HOS. **Eurasip Journal On Advances In Signal Processing**, New York, v. 2007, n. 1, p.177-189, 30 abr. 2007. Springer Science + Business Media. <http://dx.doi.org/10.1155/2007/59786>.

RODRIGO, Vianello et al. Classificação de Múltiplos Distúrbios de Qualidade de Energia Baseado em Transformada Wavelet, EOS e Redes Neurais. In: VIII CBQEE CONFERÊNCIA BRASILEIRA SOBRE QUALIDADE DE



ENERGIA ELÉTRICA, VIII., 2009, Blumenau. **Anais do VIII CBQEE Conferência Brasileira Sobre Qualidade de Energia Elétrica**. Blumenau, Sc: VIII CBQEE, 2009. v. 1, p. 212 - 219.

RUIZ-DEL-SOLAR, J.; NAVARRETE, P. Eigenspace-Based Face Recognition: A Comparative Study of Different Approaches. **IEEE Trans. Syst., Man, Cybern. C**, [s.l.], v. 35, n. 3, p.315-325, ago. 2005. Institute of Electrical & Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/tsmcc.2005.848201>.

SILVA, Alex Lima. **Redução de Características para Classificação de Imagens de Faces**. 2015. 106 f. Dissertação (Mestrado) - Curso de Ciência da Computação, Pró-Reitoria de Pesquisa e Pós-graduação, Universidade Federal Rural do Semiárido, Mossoró, 2016.

SIROVICH, L.; KIRBY, M.. Low-dimensional procedure for the characterization of human faces. **Journal Of The Optical Society Of America A**, Providence, Rhode Island, v. 4, n. 3, p.519-524, 1 mar. 1987. The Optical Society. <http://dx.doi.org/10.1364/josaa.4.000519>.

THEODORIDIS, Sergios; KOUTROUMBAS, Konstantinos. **Pattern Recognition**. 4. ed. San Diego: Academic Press, 2009. 984 p.

TURK, M.a.; PENTLAND, A.p.. Face recognition using eigenfaces. **Proceedings. 1991 IEEE Computer Society Conference On Computer Vision And Pattern Recognition**, Maui, Hi, p.586-591, 6 jun. 1991. Institute of Electrical & Electronics Engineers (IEEE). <http://dx.doi.org/10.1109/cvpr.1991.139758>.

VALENÇA, Mêuser. **Aplicando redes neurais: um guia completo**. Olinda, PE: Ed. do Autor, 2005. 264 p.

ZHAO, W. et al. Face recognition. **Csur**, [s.l.], v. 35, n. 4, p.399-458, 1 dez. 2003. Association for Computing Machinery (ACM). <http://dx.doi.org/10.1145/954339.954342>.