

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Eduardo Vieira e Sousa

**Reconhecimento facial aplicado a imagens  
infravermelhas: um estudo comparativo  
utilizando abordagens holísticas**

**Uberlândia, Brasil**

**2018**

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Eduardo Vieira e Sousa

**Reconhecimento facial aplicado a imagens  
infravermelhas: um estudo comparativo utilizando  
abordagens holísticas**

Trabalho de conclusão de curso apresentado  
à Faculdade de Computação da Universidade  
Federal de Uberlândia, Minas Gerais, como  
requisito exigido parcial à obtenção do grau  
de Bacharel em Ciência da Computação.

Orientador: Henrique Coelho Fernandes

Universidade Federal de Uberlândia – UFU

Faculdade de Computação

Bacharelado em Ciência da Computação

Uberlândia, Brasil

2018

Eduardo Vieira e Sousa

# **Reconhecimento facial aplicado a imagens infravermelhas: um estudo comparativo utilizando abordagens holísticas**

Trabalho de conclusão de curso apresentado à Faculdade de Computação da Universidade Federal de Uberlândia, Minas Gerais, como requisito exigido parcial à obtenção do grau de Bacharel em Ciência da Computação.

Trabalho aprovado. Uberlândia, Brasil, 19 de dezembro de 2018:

---

**Henrique Coelho Fernandes**  
Orientador

---

**Maria Adriana Vidigal de Lima**

---

**Rita Maria da Silva Julia**

Uberlândia, Brasil  
2018

*Dedico aos meus pais, Regina e Fernando, aos meu irmãos, Henrique e Fernanda, e à minha  
namorada, Livia.*

# Agradecimentos

Primeiramente, agradeço ao meu orientador, Henrique Coelho Fernandes, que desempenhou um papel fundamental não apenas no desenvolvimento deste trabalho, mas também nas minhas decisões em relação à carreira acadêmica e profissional. Foi por meio dele que pude conhecer a área que acabaria se tornando o caminho pelo qual eu decidi seguir durante os próximos anos da minha vida. Além disso, sou muito grato por todo o auxílio que recebi durante esse período em que trabalhamos juntos, fazendo com que ele se tornasse para mim não apenas um excelente orientador, mas também um verdadeiro amigo.

Gostaria de agradecer à minha família, por sempre me apoiar em todas as minhas decisões. Deixo meu obrigado aos meus pais, Regina e Fernando, pela paciência, confiança e suporte, sem os quais eu não teria conseguido chegar até aqui. Sou grato ao meu irmão, Henrique, pelo apoio, inspiração e ajuda durante toda a minha trajetória no curso.

Agradeço à minha irmã, Fernanda, que apesar de não estar mais entre nós, foi a maior responsável por moldar a pessoa na qual eu me tornei. Serei eternamente grato à ela por transformar minha maneira de enxergar o mundo e de enfrentar a vida. Também não posso deixar de agradecer à minha namorada, Lívia, que tem sido tão importante durante esses anos. Obrigado por todo apoio e carinho, que serviram de combustível para que eu pudesse alcançar meus objetivos.

Por fim, agradeço aos excelentes professores que me ajudaram durante o decorrer do curso, indo muito além de seus deveres como educadores, e abrindo portas para o futuro.

*"My destiny is my own!"*  
*(Illidan Stormrage)*

# Resumo

Os sistemas de reconhecimento facial modernos possuem um desempenho extraordinário quando aplicados a imagens em alta resolução e bem iluminadas. No entanto, essas não são as condições encontradas em um cenário real, onde aplicações desse tipo se fazem necessárias. Deste modo, as taxas de reconhecimento tendem a se degradar a medida que esses sistemas são expostos a imagens com variações na posição, iluminação e expressões faciais, situações comuns com as quais uma aplicação de reconhecimento facial deve lidar, especialmente em casos onde não se dispõe de um ambiente controlado. Por esses motivos, a utilização de imagens infravermelhas tem se tornado uma das abordagens mais promissoras, já que elas são mais robustas aos fatores que afetam a performance dos sistemas de detecção e reconhecimento facial. Esse trabalho apresenta o desenvolvimento de uma aplicação capaz de realizar o reconhecimento utilizando imagens de faces no espectro visível e infravermelho, e posteriormente, o sistema foi empregado na elaboração de um estudo comparativo entre ambos os resultados. Para isso, foi utilizada uma base de dados contendo conjuntos de imagens de 30 indivíduos nos dois espectros. Cada conjunto possui variações de iluminação, expressão e na posição da face, o que tornou possível mensurar o impacto dessas alterações nas taxas de reconhecimento, e verificar robustez proporcionada por cada tipo de imagem frente às diferentes condições. Os métodos utilizados para o reconhecimento foram o Eigenfaces e o Fisherfaces, duas abordagens holísticas bastante conhecidas e estudadas na literatura. Na etapa de classificação, essas técnicas foram combinadas ao algoritmo dos  $k$  vizinhos mais próximos, e os testes foram executados variando a métrica de similaridade entre: Distância Euclidiana, Manhattan, Chebyshev e Mahalanobis. Além disso, foram testadas variações nos valores de  $k$  e no número de componentes principais utilizadas para a redução de dimensionalidade e projeção das imagens. Os resultados comprovaram a hipótese de que as imagens infravermelhas são mais robustas à variação de iluminação, preservando as taxas de reconhecimento mesmo nas condições mais adversas, ao contrário das imagens do espectro visível. Quanto às variações de posição e expressões faciais, os testes foram inconclusivos devido à pequena quantidade de imagens disponíveis. O estudo indica que a área do reconhecimento facial ainda carece de bases de dados complexas para que testes mais precisos possam ser realizados. Apesar disso, os resultados sugerem que a fusão de imagens em um sistema multispectral parece ser a melhor alternativa, aproveitando assim as vantagens de cada tipo de imagem.

**Palavras-chave:** Reconhecimento Facial, Infravermelho, Eigenfaces, Fisherfaces, K-NN.

# Lista de ilustrações

|                                                                                                                                                                                                                                                     |    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 1 – Exemplo de variações na iluminação e como as imagens infravermelhas são afetadas - extraída de (LI et al., 2007). . . . .                                                                                                                | 16 |
| Figura 2 – Bandas do espectro eletromagnético - extraída de (RONAN, 2013). . . . .                                                                                                                                                                  | 20 |
| Figura 3 – Imagem de uma face em diferentes bandas do espectro infravermelho - extraída de (AKHLOUFI; BENDADA; BATSALE, 2008). . . . .                                                                                                              | 22 |
| Figura 4 – Espectro da emissão de calor do corpo humano - extraída de (GHIASS et al., 2014). . . . .                                                                                                                                                | 22 |
| Figura 5 – Fase de registro. . . . .                                                                                                                                                                                                                | 25 |
| Figura 6 – Fase de reconhecimento. . . . .                                                                                                                                                                                                          | 26 |
| Figura 7 – Taxas de erro em um sistema biométrico: (a) <i>Threshold</i> e (b) Aplicações - extraída de (JAIN; ROSS; PRABHAKAR, 2004). . . . .                                                                                                       | 27 |
| Figura 8 – Exemplo de classificação usando o k-NN. Para $k = 3$ o objeto é classificado como triângulo vermelho, para $k = 5$ a classe atribuída é quadrado azul - extraída de (ANAJ, 2007). . . . .                                                | 30 |
| Figura 9 – Sistema de cores aditivas RGB - extraída de (QUARK67, 2006). . . . .                                                                                                                                                                     | 33 |
| Figura 10 – Representação de <i>features</i> globais e locais - extraída de (HASSABALLAH; ALI; ALSHAZLY, 2016). . . . .                                                                                                                             | 35 |
| Figura 11 – Exemplos dos tipos de <i>features</i> locais encontradas na visão computacional - extraída de (JOHANSSON, 2015). . . . .                                                                                                                | 36 |
| Figura 12 – Arquitetura de um sistema de reconhecimento facial genérico - extraída de (ZHAO et al., 2003). . . . .                                                                                                                                  | 37 |
| Figura 13 – Conjunto de dados e suas componentes principais em verde e rosa - extraída de (SPRUYT, 2014). . . . .                                                                                                                                   | 42 |
| Figura 14 – Reconstrução de uma face utilizando o banco (CAMBRIDGE, 2002). . . . .                                                                                                                                                                  | 43 |
| Figura 15 – Face média do banco (CAMBRIDGE, 2002). . . . .                                                                                                                                                                                          | 44 |
| Figura 16 – Conjunto das 20 primeiras <i>eigenfaces</i> geradas com o banco (CAMBRIDGE, 2002). . . . .                                                                                                                                              | 46 |
| Figura 17 – Exemplos de projeções utilizando o Eigenfaces sobre a base de dados (CMU-PIE, 2008): na primeira linha, as imagens originais de uma face e um objeto não-face, e na segunda, as imagens resultantes - extraída de (TURK, 2008). . . . . | 47 |
| Figura 18 – LD1 representa o eixo que provavelmente seria escolhido pelo LDA, aumentando a separabilidade entre as classes, ao contrário de LD2 - extraída de (HO, 2017). . . . .                                                                   | 48 |
| Figura 19 – Conjunto da 20 primeiras <i>fisherfaces</i> gerados com o banco (CAMBRIDGE, 2002). . . . .                                                                                                                                              | 51 |

|                                                                                                                                        |    |
|----------------------------------------------------------------------------------------------------------------------------------------|----|
| Figura 20 – Exemplo de formas elementares - extraída de (PROKOSKI; RIEDEL; COFFIN, 1992). . . . .                                      | 53 |
| Figura 21 – Arquitetura do sistema de reconhecimento facial implementado. . . . .                                                      | 58 |
| Figura 22 – Iluminação: 2on, Lon, Ron, Off e Dark respectivamente - extraída de (UTK-IRIS, 2007). . . . .                              | 60 |
| Figura 23 – Expressões: Surpresa, alegria e raiva respectivamente - extraída de (UTK-IRIS, 2007). . . . .                              | 61 |
| Figura 24 – Face média gerada para o conjunto de treinamento de cada espectro. . . .                                                   | 63 |
| Figura 25 – Conjunto das 20 primeiras <i>eigenfaces</i> geradas para as imagens visíveis. . .                                          | 64 |
| Figura 26 – Conjunto das 20 primeiras <i>eigenfaces</i> geradas para as imagens infravermelhas. . . . .                                | 64 |
| Figura 27 – Distâncias das faces do conjunto de treinamento em relação à face teste para as imagens do espectro visível. . . . .       | 65 |
| Figura 28 – Distâncias das faces do conjunto de treinamento em relação à face teste para as imagens do espectro infravermelho. . . . . | 65 |
| Figura 29 – Resultado para as imagens do espectro visível. . . . .                                                                     | 66 |
| Figura 30 – Resultado para as imagens do espectro infravermelho. . . . .                                                               | 66 |
| Figura 31 – Conjunto das 20 primeiras <i>fisherfaces</i> geradas para as imagens visíveis. . .                                         | 68 |
| Figura 32 – Conjunto das 20 primeiras <i>fisherfaces</i> geradas para as imagens infravermelhas. . . . .                               | 68 |
| Figura 33 – Distâncias das faces do conjunto de treinamento em relação à face teste para as imagens do espectro visível. . . . .       | 69 |
| Figura 34 – Distâncias das faces do conjunto de treinamento em relação à face teste para as imagens do espectro infravermelho. . . . . | 69 |
| Figura 35 – Resultado para as imagens do espectro visível. . . . .                                                                     | 70 |
| Figura 36 – Resultado para as imagens do espectro infravermelho. . . . .                                                               | 70 |
| Figura 37 – Representação da metodologia de testes. . . . .                                                                            | 71 |
| Figura 38 – Taxas de reconhecimento: Eigenfaces e Distância Euclidiana. . . . .                                                        | 73 |
| Figura 39 – Taxas de reconhecimento: Fisherfaces e Distância Euclidiana. . . . .                                                       | 74 |
| Figura 40 – Taxas de reconhecimento: Eigenfaces e Distância de Manhattan. . . . .                                                      | 74 |
| Figura 41 – Taxas de reconhecimento: Fisherfaces e Distância de Manhattan. . . . .                                                     | 75 |
| Figura 42 – Taxas de reconhecimento: Eigenfaces e Distância de Chebyshev. . . . .                                                      | 75 |
| Figura 43 – Taxas de reconhecimento: Fisherfaces e Distância de Chebyshev. . . . .                                                     | 76 |
| Figura 44 – Taxas de reconhecimento: Eigenfaces e Distância de Mahalanobis. . . . .                                                    | 76 |
| Figura 45 – Taxas de reconhecimento: Fisherfaces e Distância de Mahalanobis. . . . .                                                   | 77 |

# Lista de tabelas

|                                                                                                                                                                                           |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Tabela 1 – Sub-bandas do espectro IR, e seus respectivos intervalos e aplicações - adaptada de (CASTANEDO, 2005). . . . .                                                                 | 21 |
| Tabela 2 – Comparação entre várias tecnologias biométricas: alta, média e baixa es-<br>tão denotadas por A, M e B respectivamente - adaptada de (JAIN; ROSS;<br>PRABHAKAR, 2004). . . . . | 24 |
| Tabela 3 – Exemplos de aplicações do RP - adaptada de (JAIN; DUIN; MAO, 2000). . .                                                                                                        | 28 |
| Tabela 4 – Modelos de RP - adaptada de (JAIN; DUIN; MAO, 2000). . . . .                                                                                                                   | 29 |
| Tabela 5 – Categorização dos métodos de detecção facial - adaptada de (YANG; KRI-<br>EGMAN; AHUJA, 2002). . . . .                                                                         | 39 |
| Tabela 6 – Comparação de sensibilidade a variações - adaptada de (TAN et al., 2006). .                                                                                                    | 40 |
| Tabela 7 – Categorização dos principais métodos de reconhecimento facial conheci-<br>dos - adaptada de (ZHAO et al., 2003). . . . .                                                       | 41 |
| Tabela 8 – Metodologia dos testes aplicados - adaptada de (WILDER et al., 1997). . . .                                                                                                    | 54 |
| Tabela 9 – Resultados - adaptada de (WILDER et al., 1997). . . . .                                                                                                                        | 54 |
| Tabela 10 – Taxas de acerto: média/mínima - adaptada de (SOCOLINSKY et al., 2001). .                                                                                                      | 55 |
| Tabela 11 – Taxas de acerto: média/mínima/máxima - adaptada de (SOCOLINSKY et<br>al., 2001). . . . .                                                                                      | 55 |
| Tabela 12 – Taxa de reconhecimento média para a variação no número de componentes.                                                                                                        | 72 |
| Tabela 13 – Taxa de reconhecimento média para a variação do valor de k. . . . .                                                                                                           | 72 |

# Lista de abreviaturas e siglas

|        |                                                       |
|--------|-------------------------------------------------------|
| 2on    | Two on                                                |
| 3D     | 3 Dimensões                                           |
| ASM    | Active Shape Model                                    |
| CNN    | Convolutional Neural Network                          |
| EM     | Electromagnetic                                       |
| FMR    | False Match Rate                                      |
| FNMR   | False Nonmatch Rate                                   |
| ICA    | Independent Component Analysis                        |
| IR     | Infrared                                              |
| IRIS   | Imaging, Robotics, and Intelligent Systems Laboratory |
| k-NN   | k-Nearest Neighbors                                   |
| LDA    | Linear Discriminant Analysis                          |
| Lon    | Left on                                               |
| LWIR   | Long Wavelength Infrared                              |
| MATLAB | Matrix Laboratory                                     |
| MLP    | Multilayer Perceptron                                 |
| MPF    | Matching Pursuit Filters                              |
| MWIR   | Medium Wavelength Infrared                            |
| NIR    | Near Infrared                                         |
| NN     | Neural Network                                        |
| PCA    | Principal Component Analysis                          |
| PR     | Pattern Recognition                                   |
| RGB    | Red Green Blue                                        |

|       |                               |
|-------|-------------------------------|
| Ron   | Right on                      |
| SGLD  | Space Gray-Level Dependence   |
| SVM   | Support Vector Machine        |
| SWIR  | Short Wavelength Infrared     |
| UV    | Ultraviolet                   |
| VLWIR | Very Long Wavelength Infrared |

# Lista de símbolos

|               |                  |
|---------------|------------------|
| $\epsilon$    | Pertence         |
| $\Gamma$      | Gamma            |
| $\gamma$      | Gamma minúsculo  |
| $\Lambda$     | Lambda           |
| $\lambda$     | Lambda minúsculo |
| $\mu$         | Mi minúsculo     |
| $\mu\text{m}$ | Micrômetro       |
| $\text{mm}$   | Milímetro        |
| $m$           | Metro            |
| $\omega$      | Omega minúsculo  |
| $\Phi$        | Phi              |
| $\Psi$        | Psi              |

# Sumário

|            |                                         |           |
|------------|-----------------------------------------|-----------|
| <b>1</b>   | <b>INTRODUÇÃO</b>                       | <b>15</b> |
| <b>1.1</b> | <b>Objetivos</b>                        | <b>17</b> |
| <b>1.2</b> | <b>Método</b>                           | <b>17</b> |
| <b>1.3</b> | <b>Resultados</b>                       | <b>18</b> |
| <b>1.4</b> | <b>Organização do Trabalho</b>          | <b>18</b> |
| <b>2</b>   | <b>FUNDAMENTAÇÃO TEÓRICA</b>            | <b>20</b> |
| <b>2.1</b> | <b>Espectro eletromagnético</b>         | <b>20</b> |
| 2.1.1      | Espectro Infravermelho                  | 21        |
| <b>2.2</b> | <b>Biometria</b>                        | <b>23</b> |
| 2.2.1      | Sistemas Biométricos                    | 24        |
| <b>2.3</b> | <b>Reconhecimento de Padrões</b>        | <b>27</b> |
| 2.3.1      | Classificação                           | 28        |
| 2.3.2      | Classificadores                         | 28        |
| 2.3.3      | k-Nearest Neighbors                     | 30        |
| 2.3.4      | Métricas de similaridade                | 31        |
| <b>2.4</b> | <b>Imagem Digital</b>                   | <b>32</b> |
| 2.4.1      | Sistema RGB                             | 33        |
| <b>2.5</b> | <b>Processamento Digital de Imagens</b> | <b>33</b> |
| <b>2.6</b> | <b>Visão computacional</b>              | <b>34</b> |
| 2.6.1      | Features                                | 34        |
| <b>2.7</b> | <b>Reconhecimento Facial</b>            | <b>36</b> |
| 2.7.1      | Detecção                                | 37        |
| 2.7.2      | Extração de Features e Reconhecimento   | 39        |
| 2.7.3      | PCA                                     | 41        |
| 2.7.3.1    | Eigenfaces                              | 42        |
| 2.7.3.1.1  | Treino                                  | 43        |
| 2.7.3.1.2  | Teste                                   | 47        |
| 2.7.4      | LDA                                     | 48        |
| 2.7.4.1    | Fisherfaces                             | 49        |
| <b>3</b>   | <b>TRABALHOS RELACIONADOS</b>           | <b>52</b> |
| <b>4</b>   | <b>MATERIAIS E MÉTODOS</b>              | <b>57</b> |
| <b>4.1</b> | <b>Desenvolvimento</b>                  | <b>57</b> |
| <b>4.2</b> | <b>Arquitetura do Sistema</b>           | <b>58</b> |

|            |                                   |           |
|------------|-----------------------------------|-----------|
| <b>4.3</b> | <b>Banco de Imagens</b> . . . . . | <b>60</b> |
| <b>4.4</b> | <b>Processos</b> . . . . .        | <b>62</b> |
| 4.4.1      | Eigenfaces . . . . .              | 62        |
| 4.4.1.1    | Execução . . . . .                | 63        |
| 4.4.2      | Fisherfaces . . . . .             | 67        |
| 4.4.2.1    | Execução . . . . .                | 67        |
| <b>4.5</b> | <b>Experimentos</b> . . . . .     | <b>70</b> |
| 4.5.1      | Metodologia . . . . .             | 71        |
| <b>5</b>   | <b>RESULTADOS</b> . . . . .       | <b>72</b> |
| <b>6</b>   | <b>CONCLUSÃO</b> . . . . .        | <b>78</b> |
| 6.1        | Trabalhos Futuros . . . . .       | 79        |
|            | <b>REFERÊNCIAS</b> . . . . .      | <b>80</b> |

# 1 Introdução

Nas últimas três décadas, uma das áreas que mais vem chamando a atenção, tanto da indústria como da comunidade acadêmica, é a do reconhecimento facial automatizado (OLSZEWSKA, 2016). Isso se deve à sua gama de aplicações comerciais, especialmente nos campos da segurança, ciência forense e mídias sociais. As vantagens de se utilizar a face na biometria são muitas, em particular para a área da segurança, que passa a dispor de um método de identificação à distância não invasivo, ou seja, sem a necessidade de colaboração do indivíduo. Existe também a possibilidade de aplicar sistemas como esses para a análise de expressões faciais espontâneas, o que se torna uma grande vantagem, já que seriam capazes de detectar microexpressões que geralmente passam despercebidas ao olho humano (BARTLETT et al., 2006).

Para realizar o reconhecimento facial, as máquinas necessitam do uso de algoritmos complexos e um banco de imagens compatível, o que torna essa tarefa um enorme desafio. Os problemas que a tecnologia enfrenta são muitos, surgindo desde o momento da aquisição da imagem, como por exemplo o nível de luminosidade, a posição do rosto, ou até mesmo expressões faciais. Os sistemas de reconhecimento também podem ser facilmente enganados pelo uso de acessórios como óculos ou perucas (SHARIF et al., 2016). Além disso, a falta de um banco de imagens adequado pode degradar as taxas de reconhecimento e inviabilizar todo o processo. Deste modo, é necessário que o banco esteja sempre atualizado, pois até mesmo mudanças no rosto causadas pelo envelhecimento podem afetar os resultados.

Além disso, pesquisadores compararam o impacto de variações nos sistemas de reconhecimento facial com outros tipos de biometria, e chegaram à conclusão de que eles são mais sensíveis a mudanças do que as demais tecnologias (LANITIS, 2009). Por isso, é imprescindível que os algoritmos utilizados no reconhecimento sejam robustos a essas adversidades, com objetivo de preservar o desempenho e a segurança nesses tipos de sistemas.

A utilização de imagens infravermelhas é apresentada como uma das alternativas mais promissoras para os problemas do reconhecimento facial, isso se deve ao fato de que a radiação nessa região do espectro eletromagnético é emitida ao invés de ser refletida, tornando o método independente de iluminação (GHIASS et al., 2013). Rostos capturados dessa maneira, mostram-se quase que inalterados frente às variações na iluminação, ou até mesmo em sua ausência (GHIASS et al., 2014), como mostra a figura (1). Isso torna possível a aplicação de técnicas de reconhecimento facial em ambientes externos com ou sem iluminação controlada, algo que seria inviável utilizando o espectro visível, pois até mesmo pequenas mudanças no ângulo da fonte de luz podem projetar sombras, diminuindo a performance do sistema.



Figura 1 – Exemplo de variações na iluminação e como as imagens infravermelhas são afetadas - extraída de (LI et al., 2007).

Ao comparar os efeitos de mudanças de posição e expressões faciais, é possível perceber que as imagens térmicas sofrem menores alterações (GHIASS et al., 2013), além de ajudarem a separar a face do resto do ambiente, facilitando a localização de rostos na imagem e contribuindo para etapa de detecção facial. Outra grande vantagem do infravermelho é a utilização do sistema de vasos sanguíneos da face como traço biométrico, a ideia foi apresentada por Prokoski, Riedel e Coffin (1992), em um dos primeiros trabalhos publicados que abordavam o reconhecimento facial utilizando imagens térmicas. Isso é possível pois o sangue que passa por eles é um pouco mais quente do que os tecidos ao seu redor, fazendo com que possam ser extraídos de uma imagem térmica. Esta técnica se tornou uma das mais revolucionárias, pois esses vasos possuem um padrão único para cada indivíduo, permanecendo inalterados pelo envelhecimento (PERSSON; BUSCHMANN, 2011), e por se localizarem debaixo da pele, sofrem menos com distorções causadas por fatores externos.

Além de todas as razões citadas anteriormente, é importante ressaltar que os preços das câmeras térmicas vêm caindo abruptamente desde os últimos 15 anos (KOCHANKE, 2015), de modo que já existem no mercado celulares com câmeras térmicas e dispositivos que podem ser acoplados a um *smartphone* comum, permitindo a captura de imagens infravermelhas. Juntos, esses fatores constroem um cenário favorável à pesquisa e ao desenvolvimento de tecnologias que utilizam o espectro infravermelho, tornando as imagens térmicas uma alternativa viável e promissora para os problemas do reconhecimento facial.

Este trabalho baseia-se na hipótese de que as imagens infravermelhas são mais robustas a variações de iluminação, posição e expressões faciais do que suas contrapartes do espectro visível. Ou seja, os métodos de reconhecimento facial utilizando o espectro infravermelho possuem melhor desempenho, em relação às situações e problemas mencionados, quando comparados às técnicas aplicadas ao espectro visível. Deste modo, para a verificação da hipótese, será implementado um sistema de reconhecimento facial baseado em

imagens de ambos os espectros, que será testado utilizando um banco de imagens dentre os vários disponíveis na literatura. Posteriormente, os resultados serão comparados e assim será possível verificar se os métodos de reconhecimento facial com imagens infravermelhas possuem características necessárias para justificar sua aplicação na solução dos problemas do reconhecimento facial explicitados.

## 1.1 Objetivos

O objetivo geral deste trabalho de conclusão de curso é demonstrar que o uso de imagens infravermelhas no reconhecimento facial é mais robusto, em cenários onde há variação de iluminação, posição e expressões faciais, do que o uso de imagens do espectro visível. Desta forma, os objetivos específicos são:

- Implementar técnicas eficientes para extrair *features*<sup>1</sup> de faces no espectro visível e infravermelho.
- Desenvolver um sistema para realizar o reconhecimento facial com base nas características extraídas.
- Comparar o desempenho dos métodos implementados nas imagens de ambos os espectros.
- Investigar os efeitos das variações de luminosidade, posição e expressão facial sobre as taxas de reconhecimento.
- Delimitar os cenários onde cada tipo de imagem é mais eficiente.

## 1.2 Método

As imagens utilizadas para o desenvolvimento e teste do classificador serão extraídas de bases de dados disponíveis na literatura, com preferência para as que possuem amostras multi-espectrais, ou seja, imagens de ambos os espectros: visível e infravermelho. Além disso, é necessário que o banco de imagens também contenha variações de iluminação, para que a comparação possa ser devidamente executada. Após a escolha da base de dados, inicia-se o processo de reconhecimento para as imagens selecionadas, sendo este composto por três etapas.

A primeira consiste no pré-processamento da imagem, em seguida são realizadas a extração e seleção de *features*, e por último é feita a classificação da imagem, comparando as características obtidas na segunda etapa com as presentes nas demais imagens do banco.

---

<sup>1</sup> Características que descrevem uma imagem.

Os processos de extração e seleção de *features* utilizam algoritmos que reduzem a dimensionalidade da imagem, escolhendo apenas os conjuntos de características que a definam da melhor maneira possível. Já na classificação, a imagem de teste, ou seja, a face que se deseja identificar será comparada utilizando esse conjunto de atributos selecionados no treinamento. Para isso, será utilizado algoritmo dos  $k$  vizinhos mais próximos ( $k$ -NN) combinado a diferentes métricas de similaridade, como por exemplo as distâncias: Euclidiana, Manhattan e Chebyshev.

Após a etapa de classificação, os resultados são conferidos comparando-se a classe atribuída à imagem com a identidade real do indivíduo. Desta maneira, após a repetição do processo para outras imagens do banco, e a variação dos parâmetros possíveis nos algoritmos utilizados, uma taxa de acerto poderá ser obtida. Finalmente, os resultados para ambos os espectros serão confrontados.

### 1.3 Resultados

Os seguintes resultados são esperados para este trabalho:

- Identificar os cenários ideais para a utilização das imagens de cada espectro.
- Determinar como cada tipo de imagem é afetado frente às variações na aquisição.
- Despertar um maior interesse em pesquisa e desenvolvimento de tecnologias que utilizam imagens térmicas para o reconhecimento de faces.
- Motivar a aplicação de imagens infravermelhas para o reconhecimento facial automatizado.

### 1.4 Organização do Trabalho

Para uma melhor compreensão e separação do conteúdo, os próximos capítulos deste trabalho estão organizados da seguinte maneira:

- Capítulo 2: Contém as informações necessárias para a plena compreensão das atividades desenvolvidas neste trabalho.
- Capítulo 3: Apresenta uma pesquisa sobre trabalhos relacionados e uma breve análise de seus resultados.
- Capítulo 4: Neste capítulo estão expostas as técnicas e métodos aplicados, assim como as etapas envolvidas nos experimentos propostos.

- Capítulo 5: Expõe os resultados e as análises dos experimentos conduzidos.
- Capítulo 6: Por fim, neste capítulo são apresentadas as considerações finais sobre o desenvolvimento do trabalho e os resultados obtidos, além de propor temas e atividades para trabalhos futuros

## 2 Fundamentação Teórica

Neste capítulo, são apresentadas as definições e técnicas relacionadas ao reconhecimento facial e a termografia infravermelha. Deste modo, também se faz necessária a compreensão de conceitos básicos associados ao processamento de imagens, reconhecimento de padrões e visão computacional.

### 2.1 Espectro eletromagnético

Toda a matéria emite radiação eletromagnética, independentemente do seu estado ou composição (CASTANEDO, 2005). O espectro eletromagnético (EM) nada mais é do que o intervalo que abrange todas as frequências de radiação eletromagnética, classificando-as pelo seu comprimento de onda, que pode variar de quilômetros de distância até uma pequena fração do tamanho do núcleo de um átomo. As ondas EM são divididas em 7 tipos: ondas de rádio, micro-ondas, infravermelho, luz visível, ultravioleta, raios- $x$  e raios- $\gamma$ , como demonstrado na figura (2).

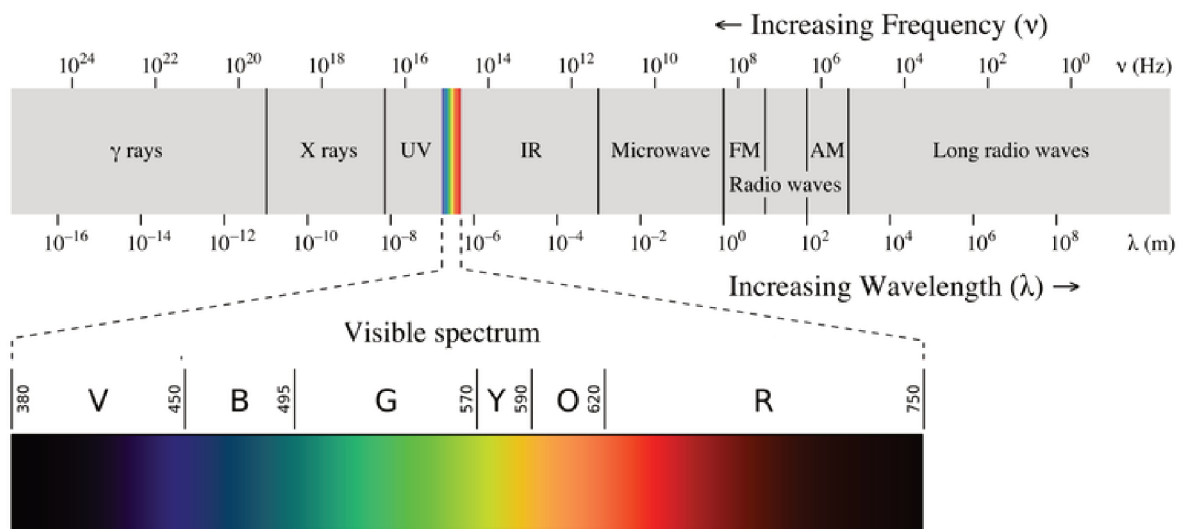


Figura 2 – Bandas do espectro eletromagnético - extraída de (RONAN, 2013).

As ondas em cada uma dessas bandas possuem diferentes características, como por exemplo a maneira como são produzidas, como interagem com a matéria e suas aplicações práticas. Dentre os tipos de radiação, as formas mais energéticas possuem o menor comprimento de onda, variando de  $10^{-14}m$  até  $10^{-8}m$ . Nessa região do espectro se encontram os raios- $\gamma$  e os raios- $x$ , que possuem importantíssimas aplicações na medicina e na segurança.

Do outro lado do espectro, com comprimentos variando de  $1\text{ mm}$  até alguns metros, se localizam as micro-ondas e as ondas de rádio, sendo que a segunda possui várias aplicações nas telecomunicações devido a sua capacidade de penetrar na atmosfera terrestre. Além disso, outros tipos de ondas com as mais variadas utilidades estão localizadas entre esses dois extremos, essa região é conhecida como o espectro térmico.

### 2.1.1 Espectro Infravermelho

Qualquer objeto com temperatura acima do zero absoluto emite radiação na forma de calor, que é trocado com o meio (CASTANEDO, 2005). Esse tipo é conhecido como radiação térmica, estando localizada no intervalo de  $0.1\mu\text{m}$  até  $1000\mu\text{m}$  do espectro EM. A região térmica pode ser dividida em 3 bandas: ultravioleta (UV), luz visível e infravermelha (IR).

A radiação IR, também conhecida por luz infravermelha, foi descoberta pelo astrônomo William Herschel em 1800. Sua emissão e absorção ocorrem devido a mudanças nos movimentos vibracionais e rotacionais das moléculas que compõe o objeto. Ela está localizada no espectro térmico, em uma região que vai de  $0.74\mu\text{m}$  até  $1000\mu\text{m}$ , e pode ser dividida em cinco sub-bandas, como mostra a tabela (1).

Tabela 1 – Sub-bandas do espectro IR, e seus respectivos intervalos e aplicações - adaptada de (CASTANEDO, 2005).

| Sub-banda | Intervalo [ $\mu\text{m}$ ] | Aplicações                 |
|-----------|-----------------------------|----------------------------|
| NIR       | 0,74 – 1                    | Telecomunicações           |
| SWIR      | 1 – 3                       | Sensoriamento remoto       |
| MWIR      | 3 – 5                       | Pesquisa científica        |
| LWIR      | 8 – 14                      | Inspeção industrial        |
| VLWIR     | 14 – 1000                   | Espectrometria, astronomia |

A partir de  $0,74\mu\text{m}$  até  $1\mu\text{m}$  está localizada a área conhecida como *near infrared* (NIR), a região *short wavelength infrared* (SWIR) fica entre  $1\mu\text{m}$  a  $3\mu\text{m}$ , já a *medium wavelength infrared* (MWIR) começa em  $3\mu\text{m}$  e termina em  $5\mu\text{m}$ , entre  $8\mu\text{m}$  e  $14\mu\text{m}$  fica a região chamada de *long wavelength infrared* (LWIR), e por último a *very long wavelength infrared* (VLWIR), que vai de  $14\mu\text{m}$  a  $1000\mu\text{m}$ .

Dentro do contexto do reconhecimento facial, é possível obter imagens térmicas em todas as regiões do espectro IR (AKHLOUFI; BENDADA; BATSALE, 2008), como mostra a figura (3). No entanto, algumas dessas sub-bandas são mais importantes, devido à forma como o corpo humano emite calor (GHIASS et al., 2014), que está demonstrada no gráfico (4).

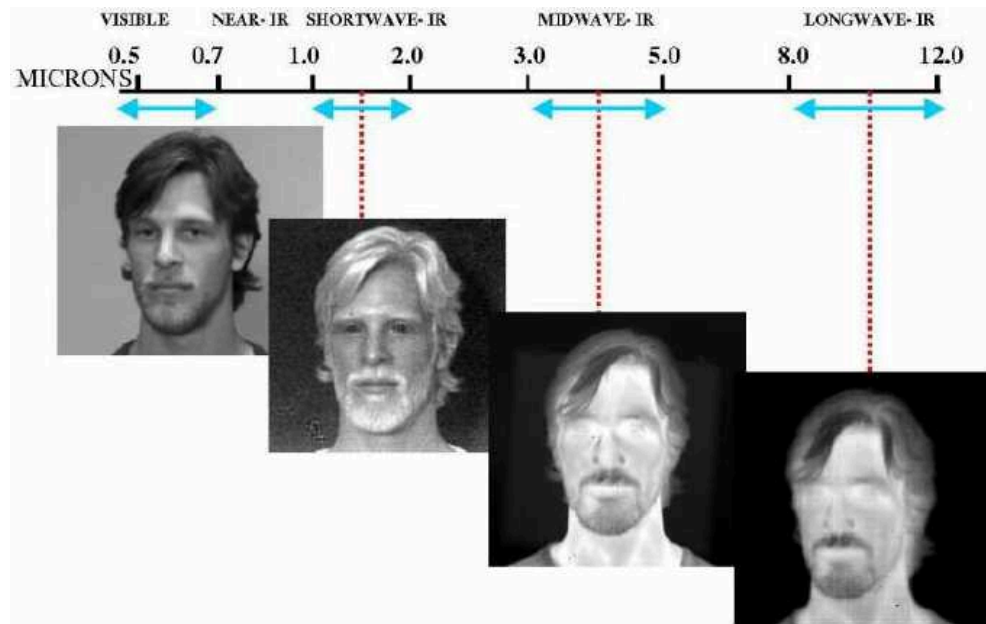


Figura 3 – Imagem de uma face em diferentes bandas do espectro infravermelho - extraída de (AKHLOUFI; BENDADA; BATSALE, 2008).

A maior parte da energia térmica do corpo humano é emitida nas sub-bandas LWIR e MWIR (GHIASS et al., 2014), logo ambas podem ser utilizadas de forma passiva no reconhecimento, ou seja, sem a necessidade de uma fonte de luz externa.

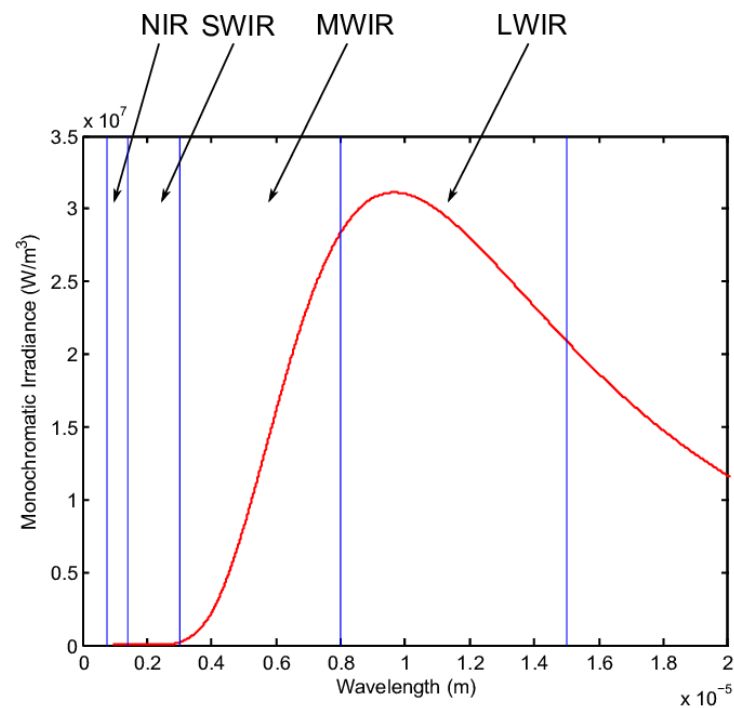


Figura 4 – Espectro da emissão de calor do corpo humano - extraída de (GHIASS et al., 2014).

## 2.2 Biometria

O termo biometria é derivado das palavras gregas *bios* (vida) e *metrikos* (medir). A área consiste no estudo de atributos dos seres humanos, que geralmente são aplicados na identificação de indivíduos (DELAC; GRGIC, 2004). As características utilizadas na biometria podem ser divididas em duas categorias (SRIVASTAVA, 2013), que são definidas por sua procedência, sendo elas:

**Fisiológicas:** Estão relacionadas às características físicas do corpo humano e por isso são difíceis de adulterar, além de pouco se modificarem durante a vida do indivíduo. Esse tipo inclui traços como: íris, impressão digital, rosto e DNA.

**Comportamentais:** São características dependentes da psicologia humana, podendo variar de acordo com o estado mental do indivíduo e com o ambiente. Alguns exemplos são a voz, a assinatura e até mesmo o ritmo de digitação pode ser utilizado para identificar uma pessoa.

Qualquer traço fisiológico ou comportamental pode servir como biometria, desde que possua os seguintes atributos (SRIVASTAVA, 2013):

1. **Universalidade:** A característica deve estar presente em todos os indivíduos que se deseja comparar.
2. **Distinção:** Nenhum par de indivíduos deve possuir o mesmo traço, ou seja, ele deve identificar cada um de maneira única.
3. **Permanência:** Deve conservar suas características distintivas durante o decorrer do tempo.
4. **Acessibilidade:** O traço biométrico deve ser acessível para a coleta e verificação.

Além das quatro características básicas citadas, também existem fatores adicionais que devem ser levados em consideração no momento da aplicação prática em um sistema biométrico:

- **Performance:** Diz respeito à eficiência do traço biométrico em termos de robustez, acurácia, velocidade e custo computacional para a aplicação.
- **Aceitabilidade:** Define quão invasiva é a extração e verificação biométrica.
- **Alterabilidade:** A característica deve ser robusta contra fraudes e adulterações.

Desta maneira, é possível classificar as abordagens biométricas modernas de acordo com suas características (JAIN; ROSS; PRABHAKAR, 2004), como mostra a tabela (2).

Tabela 2 – Comparação entre várias tecnologias biométricas: alta, média e baixa estão denotadas por A, M e B respectivamente - adaptada de (JAIN; ROSS; PRABHAKAR, 2004).

| Identificador Biométrico | Universalidade | Distinção | Permanência | Acessibilidade | Performance | Aceitabilidade | Alterabilidade |
|--------------------------|----------------|-----------|-------------|----------------|-------------|----------------|----------------|
| 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 facial        | 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              |
| Íris                     | A              | A         | A           | M              | A           | B              | B              |
| Ritmo de digitação       | B              | B         | B           | M              | B           | M              | M              |
| Cheiro                   | A              | A         | A           | B              | B           | M              | B              |
| Impressão da mão         | M              | A         | A           | M              | A           | M              | M              |
| Retina                   | A              | A         | M           | B              | A           | B              | B              |
| Assinatura               | B              | B         | B           | A              | B           | A              | A              |
| Voz                      | M              | B         | B           | M              | B           | A              | A              |

### 2.2.1 Sistemas Biométricos

Um sistema biométrico automático consiste de um reconhecedor de padrões que identifica uma pessoa por meio de um conjunto de atributos derivados de suas características físicas (DELAC; GRGIC, 2004). Geralmente existem duas abordagens para esses tipos de sistemas:

**Unimodal:** Utiliza apenas um traço biométrico para a comparação, o que deixa o sistema vulnerável às deficiências do atributo escolhido.

**Multimodal:** Faz a combinação de múltiplas características biométricas de um mesmo indivíduo para gerar um padrão mais complexo e robusto, aumentando assim a segurança e contornando qualquer problema inerente a um tipo específico de biometria.

Além disso, um sistema biométrico frequentemente opera em dois modos (CHELLAPPA; SINHA; PHILLIPS, 2010):

1. Verificação: Confere se a identidade do indivíduo corresponde à apresentada (1:1).
2. Identificação: Busca uma identificação baseando-se nos dados de todos as pessoas armazenadas pela aplicação (1:n).

De maneira geral, a arquitetura dos sistemas de reconhecimento biométrico pode ser dividida nos seguintes módulos (JAIN; ROSS; PRABHAKAR, 2004):

**Sensor:** Hardware que colhe os dados biométricos a serem analisados, como por exemplo uma câmera de segurança que registra imagens da face de um indivíduo para realizar o reconhecimento facial.

**Extrator de características:** Responsável por processar os dados colhidos pelo sensor e extrair deles as características discriminatórias, também conhecidas como *features*. Cada sistema biométrico possui algoritmos que extraem características diferentes de acordo com a conveniência e abordagem necessária.

**Classificador:** Nele são comparados os conjuntos de características geradas pelo módulo extrator. A maneira como essas características serão analisadas vai depender do algoritmo e do tipo de aprendizado utilizado pelo sistema. A saída desse módulo é uma decisão, ou classificação referente aos dados iniciais de entrada.

**Banco de dados:** Módulo onde fica armazenado o conjunto de *features* para cada um dos indivíduos já analisados pelo sistema, assim como outras informações relevantes que possam ser utilizadas na identificação.

Por fim, o sistema biométrico funciona em duas fases (JAIN; ROSS; PRABHAKAR, 2004):

- Registro (*Enrollment*): Etapa onde uma nova amostra é cadastrada no banco de dados, ela consiste na coleta dos dados biométricos pelo módulo de sensor, em seguida as características são geradas pelo extrator de *features*, que por sua vez são armazenadas pelo módulo de banco de dados de acordo com a figura (5).

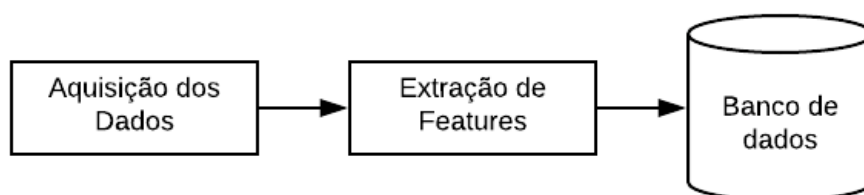


Figura 5 – Fase de registro.

- **Reconhecimento:** Consiste em identificar os dados biométricos de entrada comparando-os com as informações provenientes do módulo do banco de dados. Novamente as informações são colhidas e as características extraídas, porém nessa etapa elas são verificadas pelo módulo comparador e uma decisão é tomada, ou seja, a amostra é classificada como mostra a figura (6).

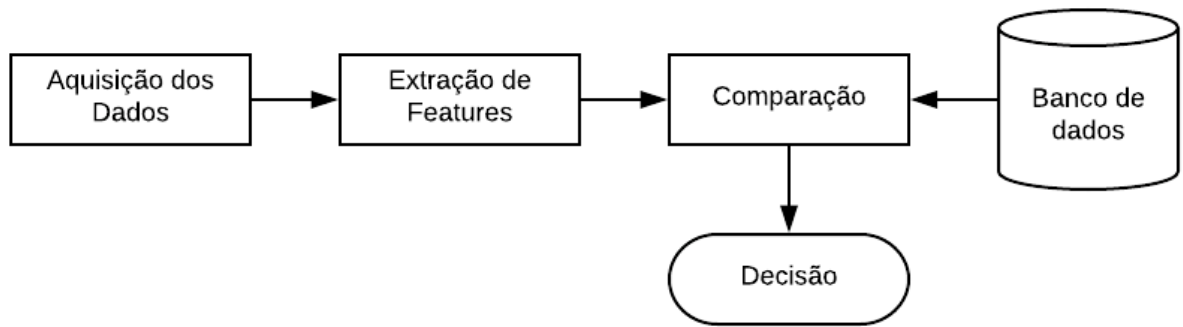


Figura 6 – Fase de reconhecimento.

Fatores como o posicionamento do módulo de sensor, interferências ambientais e má interação com o usuário podem causar ruídos e deformações nas amostras colhidas (DE-LAC; GRGIC, 2004). Isso faz com que o mesmo traço biométrico de um indivíduo possa gerar amostras diferentes, de maneira que elas não possam ser comparadas rigorosamente, o que pode causar falhas no processo de identificação. Existem 2 tipos de erros em um sistema de reconhecimento (JAIN; ROSS; PRABHAKAR, 2004):

**Falso positivo (*False match*):** Ocorre quando duas pessoas diferentes são classificadas como sendo uma mesma pessoa.

**Falso negativo (*False nonmatch*):** Acontece quando amostras de uma mesma pessoa são classificadas como de pessoas diferentes.

Para a redução da taxa de erro de um sistema biométrico, pode ser adotado um valor de decisão conhecido como *threshold*. Ele serve como parâmetro para uma pontuação (*score*), que ao atingir um valor alto o suficiente, permite a classificação da amostra. Esse valor pode ser alterado a fim de encontrar um ponto ótimo que proporciona uma menor taxa de erro, como mostra o gráfico (7).

Quanto às aplicações dos sistemas biométricos, segundo Jain, Ross e Prabhakar (2004), taxas de falsos positivos menores são mais úteis na área forense, enquanto baixas proporções de falso negativo beneficiam mais as aplicações na segurança.

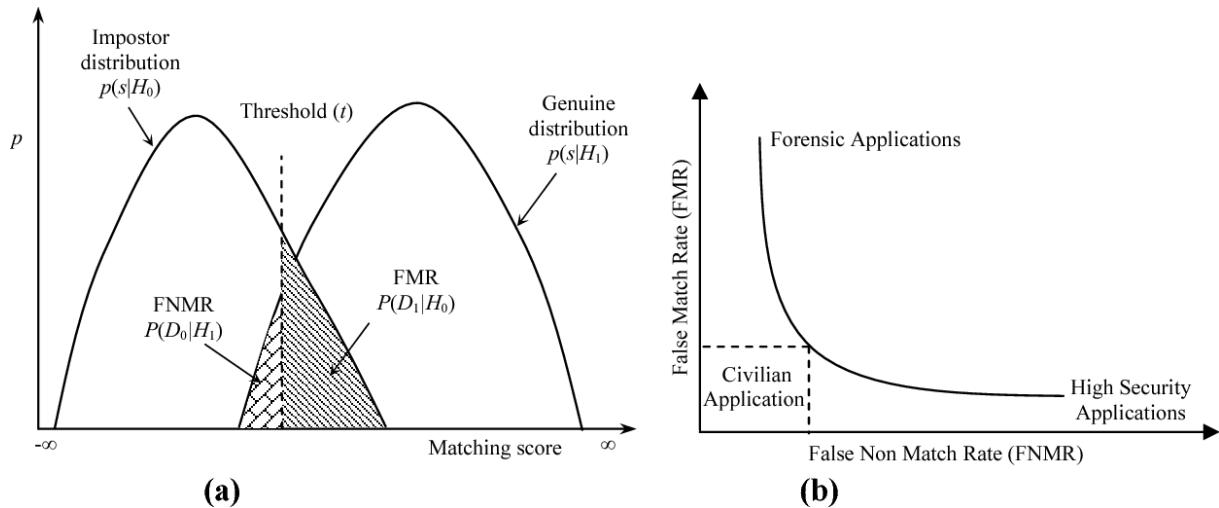


Figura 7 – Taxas de erro em um sistema biométrico: (a) *Threshold* e (b) Aplicações - extraída de (JAIN; ROSS; PRABHAKAR, 2004).

## 2.3 Reconhecimento de Padrões

A execução automática de tarefas como reconhecimento, descrição, classificação e agrupamento são problemas importantes nas mais variadas áreas do conhecimento (JAIN; DUIN; MAO, 2000), como por exemplo: biologia, psicologia, medicina, marketing, visão computacional.

De acordo com Watanabe (1985), um padrão pode ser interpretado como o oposto do caos, uma entidade vagamente definida a qual pode ser atribuída um nome. Desta maneira, o reconhecimento de padrões pode ser definido com um campo de estudo que tem por objetivo a classificação de objetos (padrões) em categorias ou classes (THEODORIDIS; KOUTROUMBAS, 2008). Isso é feito através da observação e extração de informações dos dados brutos, que servem de base para tomada de decisões (DUDA; HART; STORK, 2000), como por exemplo a classificação.

Existem diversos tipos de padrões que podem ser observados e classificados, a tabela (3) fornece alguns exemplos de aplicações, assim como seus padrões de entrada e classes aplicáveis.

Tabela 3 – Exemplos de aplicações do RP - adaptada de (JAIN; DUIN; MAO, 2000).

| Domínio do Problema             | Aplicação                       | Padrão de Entrada             | Classes                                         |
|---------------------------------|---------------------------------|-------------------------------|-------------------------------------------------|
| Bioinformática                  | Análise de sequências           | DNA/Sequência de proteínas    | Tipos de genes                                  |
| Mineração de dados              | Procura por padrões             | Pontos no espaço              | Agrupamentos compactos e bem separados          |
| Classificação de documentos     | Buscas na internet              | Arquivo de texto              | Categorias semânticas (Ex: música, jogos, etc.) |
| Análise de imagens              | Máquina de leitura para cegos   | Arquivo de imagem             | Caracteres alfanuméricos, palavras              |
| Automação industrial            | Inspeção de circuitos impressos | Imagem de intensidade         | Produto com defeito, sem defeito                |
| Recuperação de dados multimídia | Buscas na internet              | Arquivo de vídeo              | Gêneros de vídeo (Ex: ação, romance, etc.)      |
| Reconhecimento biométrico       | Identificação pessoal           | Face, íris, impressão digital | Usuário autorizado, não autorizado              |
| Sensoriamento remoto            | Previsão de colheita            | Imagens multiespectrais       | Padrão de crescimento da colheita               |
| Reconhecimento de fala          | Suporte telefônico automático   | Arquivo de áudio              | Palavras pronunciadas                           |

### 2.3.1 Classificação

Um objeto ou padrão pode ser representado por um vetor de dados  $x$  de  $p$  dimensões tal que  $x = (x_1, x_2, \dots, x_p)$  onde  $x_i$  são medições de seus atributos, também conhecidas como *features*, que por sua vez são características extraídas de uma amostra de entrada, o que pode ser feito através de medição ou processamento.

Desta maneira, esses padrões podem ser associados à uma classe, que pode ser definida como um conjunto de objetos que possuem certos atributos em comum. Dado que existem  $C$  classes denotadas por  $\omega_1, \omega_2, \dots, \omega_C$ , e associada a cada padrão existe uma variável categórica  $z$  (WEBB, 2011), que denota a associação de classe, isso significa que se  $z = i$ , então o objeto pertence a  $\omega_i$ , com  $i \in 1, 2, \dots, C$ .

### 2.3.2 Classificadores

São as ferramentas que realizam a classificação de padrões ou objetos a partir de suas características. Seu objetivo é estabelecer uma função de separação entre as classes do

problema, desta maneira, ao receber um conjunto de características, a função retorna um rótulo para a amostra de entrada.

Para que o classificador possa reconhecer um padrão, é necessário que ele saiba associar um conjunto de atributos à uma classe. Esse processo é conhecido como aprendizado, e para que ele ocorra, geralmente os dados disponíveis são divididos em dois grupos:

- Treino: Conjunto utilizado para “aprender” os padrões.
- Teste: Dados usados na avaliação do aprendizado.

Deste modo, é comum que os classificadores sejam separados de acordo com a abordagem utilizada na etapa de aprendizado (JAIN; DUIN; MAO, 2000), da seguinte maneira:

**Supervisionado:** Consiste em apresentar ao sistema de reconhecimento os objetos e as classes correspondentes, para que ambos possam ser associados de uma maneira prévia.

**Não-Supervisionado:** Permite que o classificador defina de maneira autônoma as classes, baseando-se exclusivamente nas características presentes no conjunto de treino.

Os dados obtidos na fase de aquisição são utilizados para treinar o classificador, desta maneira, a taxa de acerto está diretamente relacionada à representatividade do conjunto de dados. Além disso, a construção do modelo deve ser cuidadosa, já que um ajuste muito preciso pode causar um problema conhecido como *overfitting* (DUDA; HART; STORK, 2000), que ocorre quando o classificador se especializa demais em um conjunto de treino específico, gerando um modelo que não necessariamente corresponde à realidade.

Segundo Jain, Duin e Mao (2000), as principais abordagens para o reconhecimento de padrões podem ser classificadas da seguintes maneiras, de acordo com a tabela (4):

Tabela 4 – Modelos de RP - adaptada de (JAIN; DUIN; MAO, 2000).

| Abordagem               | Representação                             | Função de Reconhecimento | Critério de Avaliação |
|-------------------------|-------------------------------------------|--------------------------|-----------------------|
| Template Matching       | Amostras, pixels, curvas                  | Correlação, distância    | Erro de classificação |
| Estatístico             | <i>Features</i>                           | Função discriminante     | Erro de classificação |
| Sintático ou Estrutural | Primitivas                                | Regras, gramática        | Erro de aceitação     |
| Redes Neurais           | Amostras, <i>pixels</i> , <i>features</i> | Função de ativação       | Erro quadrático médio |

### 2.3.3 k-Nearest Neighbors

O classificador dos  $k$  vizinhos mais próximos (k-NN) se encaixa em uma categoria de algoritmos conhecida como aprendizado preguiçoso (*lazy-learning*), que consiste em atrasar o processo de indução ou generalização até o momento em que a classificação é executada (KOTSIANTIS; ZAHARAKIS; PINTELAS, 2006).

Esse método se baseia no princípio de que objetos que possuem propriedades similares estarão localizados mais próximos dentro de um espaço multidimensional. Deste modo, para classificar uma amostra, basta verificar a classe dos vizinhos mais próximos. A proximidade entre os padrões é dada por uma métrica de similaridade, como por exemplo a distância euclidiana.

Dado um padrão desconhecido  $X$ , a classificação é feita da seguinte maneira:

1. Calcular a distância entre  $X$  e todos os outros objetos do conjunto de treinamento.
2. Verificar as classes dos  $k$  vizinhos mais próximos.
3. Selecionar entre as classes verificadas, a que possui maior frequência.
4. Se houver empate, decrementar o valor de  $k$  e voltar para o passo 2.
5. Rotular  $X$  com a classe selecionada.

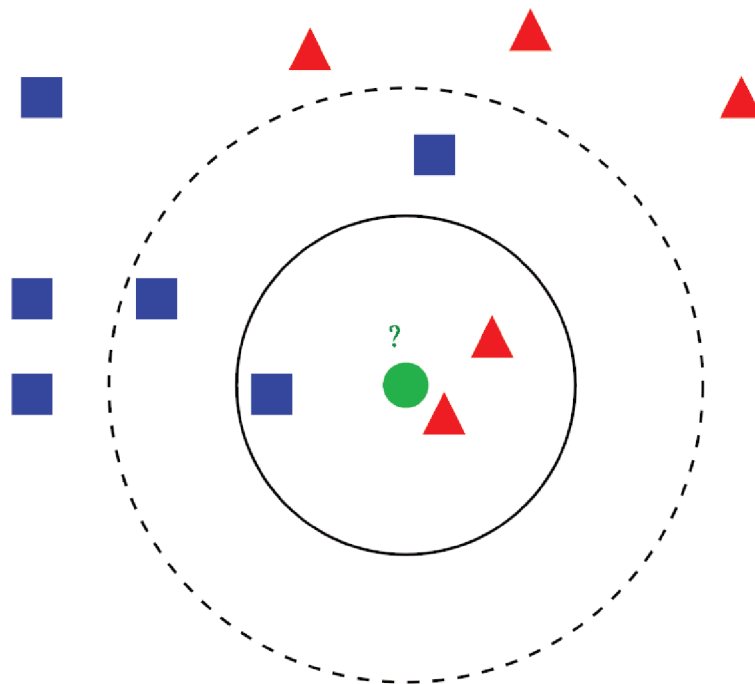


Figura 8 – Exemplo de classificação usando o k-NN. Para  $k = 3$  o objeto é classificado como triângulo vermelho, para  $k = 5$  a classe atribuída é quadrado azul - extraída de (ANAJ, 2007).

O k-NN tem sido amplamente utilizado de maneira bem-sucedida em várias aplicações (KOTSIANTIS; ZAHARAKIS; PINTELAS, 2006), como por exemplo algoritmos de reconhecimento facial como o Eigenfaces e o Fisherfaces, porém é importante ressaltar alguns de seus aspectos:

- Requer grandes espaços de armazenamento.
- É sensível à escolha da métrica de similaridade utilizada.
- Não existem princípios claros para escolha do valor de  $k$ , sendo necessária a aplicação de técnicas como por exemplo a validação cruzada.

### 2.3.4 Métricas de similaridade

Entende-se por distância, a medida de separação entre dois objetos. No reconhecimento de padrões, a distância indica similaridade ou afastamento entre dois atributos ou classes. Para que uma função  $d()$  seja considerada uma métrica, ela deve possuir as seguintes propriedades (DUDA; HART; STORK, 2000):

1. Não negatividade:  $d(a,b) \geq 0$
2. Reflexibilidade:  $d(a,b) = 0 \iff a = b$
3. Simetria:  $d(a,b) = d(b,a)$
4. Desigualdade triangular:  $d(a,b) + d(b,c) \geq d(a,c)$

Dados dois pontos  $n$ -dimensionais,  $A = (a_1, a_2, \dots, a_n)$  e  $B = (b_1, b_2, \dots, b_n)$ , as métricas mais conhecidas são:

**Distância Euclidiana:** A mais utilizada dentre as métricas de similaridade (EL-SHOROUK, 2011), pode ser calculada de acordo com a fórmula (2.1):

$$d(A,B) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2} \quad (2.1)$$

**Distância de Manhattan:** Também conhecida como *City-Block* ou distância retilínea (EL-SHOROUK, 2011), ela examina a diferença em valores absolutos da distância entre dois pontos da seguinte maneira (2.2):

$$d(A,B) = |a_1 - b_1| + |a_2 - b_2| + \dots + |a_n - b_n| \quad (2.2)$$

**Distância de Chebyshev:** A distância entre dois pontos é dada pela maior entre as diferenças de suas coordenadas, também é conhecida como distância máxima (EL-SHOROUK, 2011) ou *chessboard*. Ela pode ser obtida por (2.3):

$$d(A,B) = \max(|a_1 - b_1|, |a_2 - b_2|, \dots, |a_n - b_n|) \quad (2.3)$$

**Distância de Mahalanobis:** Baseada na correlação entre as variáveis e os padrões analisados, ela se difere da distância euclidiana por ser invariante à escala. Seu cálculo é dado de acordo com (2.4):

$$d(A,B) = \sqrt{(A - B)^T S^{-1} (A - B)} \quad (2.4)$$

onde  $S$  é a matriz de covariância.

## 2.4 Imagem Digital

Uma imagem pode ser definida como uma função bidimensional  $f(x,y)$ , onde  $x$  e  $y$  são coordenadas em um plano, e a amplitude de  $f$  em qualquer par coordenado  $(x,y)$  é chamada de intensidade ou nível de cinza da imagem no ponto. Quando os valores  $x,y$  e sua amplitude são finitos, ou seja, representados por valores discretos, chamamos então de uma imagem digital (GONZALEZ; WOODS, 2006). Essa função geralmente é denotada por uma matriz, onde cada um de seus elementos carrega um valor que indica a cor do respectivo ponto na imagem, que é conhecido como *pixel*.

A qualidade da imagem está diretamente relacionada à dois fatores (GONZALEZ; WOODS, 2006) presentes na etapa de aquisição:

**Amostragem:** Diz respeito à resolução da imagem, ou seja, quanto maior a quantidade de *pixels*, menor o tamanho dos mesmos e menores são as discontinuidades entre eles, tornando a imagem mais nítida.

**Quantização:** Está relacionada à quantidade de *bits* utilizados para a codificação de cada valor de intensidade. Uma maior quantidade de *bits* permite um maior alcance de cores, proporcionando uma imagem de melhor qualidade.

Durante a fase de aquisição das imagens, surgem as principais fontes de ruídos (GONZALEZ; WOODS, 2006), podendo causar sérios prejuízos para as aplicações que as utilizam. Para a correção desses problemas, são utilizadas técnicas de processamento de imagens.

### 2.4.1 Sistema RGB

O sistema de cores mais utilizado é o RGB, representado na figura (9). Trata-se de um esquema aditivo que consiste na mistura de três cores primárias: vermelho (*Red*), verde (*Green*) e azul (*Blue*) em diferentes proporções, denotadas por um valor de intensidade, à partir das quais é possível gerar todas as demais cores do espectro visível (SZELISKI, 2010). Nesse caso, o branco é produzido misturando-se as três cores primárias com suas intensidades máximas, e o preto é gerado quando todas possuem valor igual à zero.

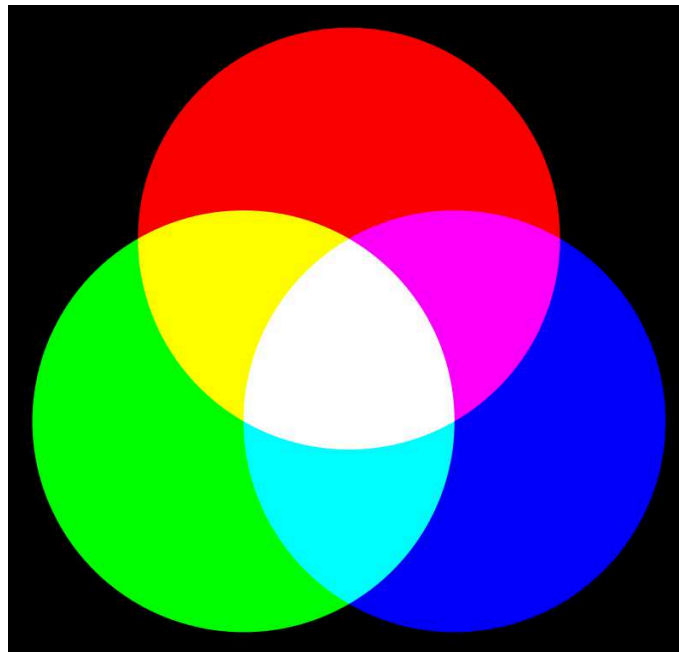


Figura 9 – Sistema de cores aditivas RGB - extraída de (QUARK67, 2006).

## 2.5 Processamento Digital de Imagens

Alguns autores definem o processamento de imagens como sendo uma área onde tanto os dados de entrada como os de saída devem ser uma imagem, porém não existe uma fronteira clara de onde termina o processamento de imagens, e onde começa a visão computacional (GONZALEZ; WOODS, 2006). Uma maneira comum de se abordar o processamento de imagens, é dividi-lo em três níveis de processos:

**Baixo nível:** Trata-se de operações primitivas, como redução de ruído e aumento de contraste.

**Nível médio:** Tarefas relacionadas ao conteúdo da imagem, como por exemplo a segmentação, que consiste em particionar uma imagem em regiões ou objetos.

**Alto nível:** Diz respeito aos processos que envolvem interpretação da imagem, muitas vezes de formas associadas à visão humana.

Desta maneira, ficou estabelecido que a área do processamento de imagens compreende os procedimentos de baixo nível (GONZALEZ; WOODS, 2006). Já a visão computacional está relacionada aos de alto nível, entretanto nos processos de nível médio é possível perceber uma sobreposição onde ambos os conceitos se misturam.

## 2.6 Visão computacional

O objetivo da visão computacional é a utilização de máquinas para emular a visão humana, incluindo a capacidade de aprender, fazer inferências e realizar ações baseadas em informações visuais (SHAPIRO; STOCKMAN, 2001). Deste modo, fica claro que a visão computacional é um campo interdisciplinar localizado entre processamento digital de imagens, inteligência artificial e reconhecimento de padrões.

Para que seja possível a tomada de decisão sobre objetos reais representados em imagens, se faz necessária a construção de um modelo, um descritor da imagem (SHAPIRO; STOCKMAN, 2001). Desta maneira, alguns especialistas argumentam que o problema fundamental da visão computacional é na verdade a construção e descrição de uma cena através de imagens.

Os recentes avanços no campo do aprendizado de máquina possibilitaram o desenvolvimento de uma série de tecnologias que viabilizaram a aplicação da visão computacional nas mais variadas áreas (SZELISKI, 2010), como por exemplo:

- Processamento de texto manuscrito
- Inspeção de materiais
- Construção automática de modelos de 3D
- Diagnósticos médicos
- Veículos autônomos

### 2.6.1 Features

Tarefas como a representação, classificação e reconhecimento de imagens são extremamente dependentes da extração de características robustas e representativas. Para conseguir estabelecer uma correspondência entre duas ou mais imagens, é necessário localizar atributos que as caracterizam de uma maneira única, essas informações são conhecidas na visão computacional como *features* (HASSABALLAH; ALI; ALSHAZLY, 2016). É importante também que imagens semelhantes produzam valores de *features* próximos, para que algoritmos que utilizam medidas de similaridade na classificação, como por exemplo a distância

euclediana, possam classificar as imagens da maneira mais precisa possível. As *features* podem ser separadas em dois tipos, representados na figura (10) e explicados a seguir:

**Globais:** Descrevem a imagem como um todo, ou seja, podem ser interpretadas como uma propriedade particular da imagem envolvendo todos os seus *pixels*. Desta maneira, esse tipo de abordagem acaba sendo sensível a alterações de iluminação e posição, além de não ser possível isolar regiões de interesse na imagem. Entre os exemplos de *features* globais estão os histogramas de cor, forma e textura.

**Locais:** Buscam representar a imagem através de suas regiões ou pontos de interesse (*key-points*). Cada *keypoint* contido na imagem é considerado uma *feature*.

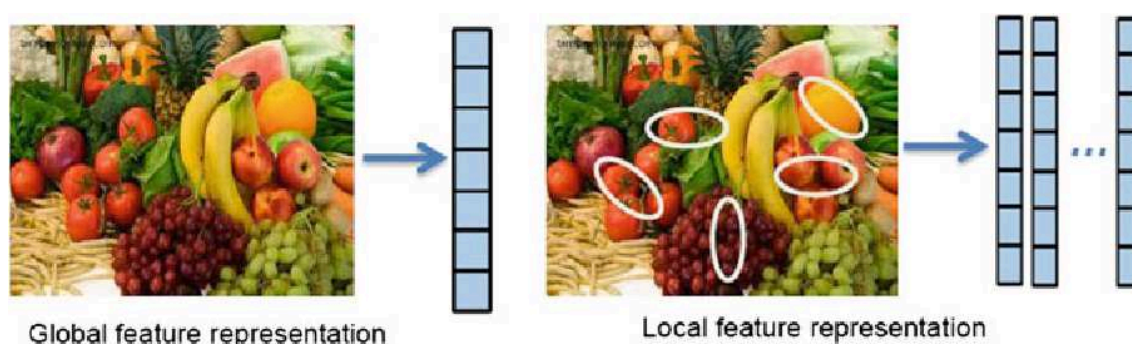


Figura 10 – Representação de *features* globais e locais - extraída de (HASSABALLAH; ALI; ALSHAZLY, 2016).

O tipo de *feature* a ser utilizado varia de acordo com a aplicação, cada um possui suas vantagens e desvantagens. Para conjuntos de dados muito extensos, as *features* globais geralmente são as preferidas, devido ao fato de serem mais compactas e rápidas de se computar, geralmente requisitando menos memória. No entanto, elas são mais vulneráveis a variações como rotação, oclusão e outros ruídos.

As *features* locais por sua vez possuem um desempenho muito superior ao de suas contrapartes (HASSABALLAH; ALI; ALSHAZLY, 2016), por esse motivo são mais utilizadas nas tarefas de detecção e reconhecimento facial. Apesar de suas vantagens, elas demandam uma quantidade significativa de memória devido a grande quantidade de atributos gerados para uma determinada imagem. Na maioria dos casos, esse problema é contornado com algoritmos de redução da dimensionalidade. As *features* locais mais frequentemente utilizadas são:

**Edges:** Podem ser visualizadas como mudanças bruscas na cor ou na intensidade dos pixels entre regiões vizinhas em uma certa direção .

**Corners:** Similares às *edges*, com a diferença de que a variação ocorre nas direções de duas linhas que se intersectam.

**Blobs:** São áreas que contém pontos com propriedades similares, e geralmente possuem uma forma circular.

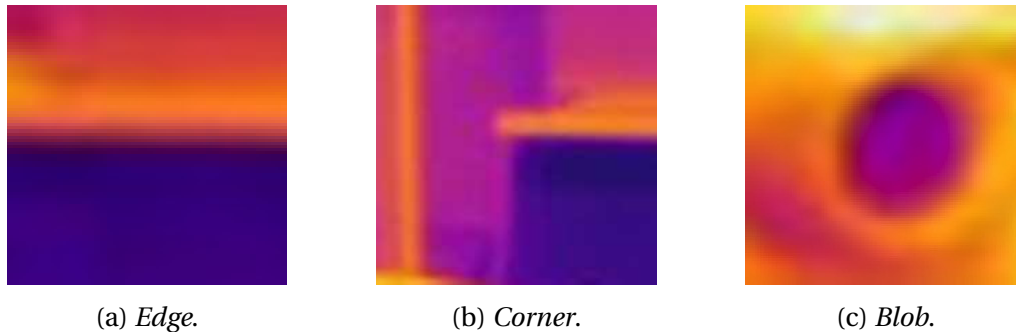


Figura 11 – Exemplos dos tipos de *features* locais encontradas na visão computacional - extraída de (JOHANSSON, 2015).

Uma boa *feature* deve possuir os seguintes atributos (JOHANSSON, 2015):

**Reprodução:** Um conjunto de valores similares de *features* deve ser extraído para imagens diferentes de uma mesma cena ou objeto.

**Distinção:** As características extraídas devem ser altamente distintivas, para que o processo de classificação seja mais preciso.

**Quantidade:** O número de valores de *features* gerados precisa ser grande o suficiente para possibilitar uma classificação eficiente, mas de maneira que não torne o processo computacionalmente dispendioso.

**Acurácia:** Deve ser possível decidir a localização de uma *feature*, assim como sua forma e tamanho.

**Eficiência:** Baixo custo computacional e de memória.

Após a etapa de extração, que obtém todos os atributos contidos em uma determinada imagem, ocorre um processo conhecido por seleção de *features*, no qual as características extraídas são selecionadas de acordo com sua capacidade de representação e distinção. Geralmente, ambos os processos são executados por detectores de *features*, que por sua vez combinam as características selecionadas, gerando um descritor para a imagem.

## 2.7 Reconhecimento Facial

A tarefa do reconhecimento facial pode ser trivial para um ser humano, mas do ponto de vista computacional ela se torna extremamente complexa. De uma maneira geral, ela

pode ser enxergada como um caso particular de reconhecimento de padrões, onde o objeto de entrada é uma imagem na qual será buscado um padrão facial. Na maioria dos casos, o reconhecimento é composto das seguintes etapas (ZHAO et al., 2003):

1. Detecção da face: Consiste em localizar e segmentar a face da imagem de entrada.
2. Extração de *features*: Etapa na qual são extraídas as características da face utilizadas na construção de um descritor. Os atributos escolhidos variam de acordo com o tipo de abordagem usada para o reconhecimento.
3. Reconhecimento: Momento final onde o conjunto de *features* da face de entrada é comparado com as informações registradas no banco de dados do sistema. Para isso, podem ser utilizados classificadores como redes neurais, SVMs, k-NN e outros.

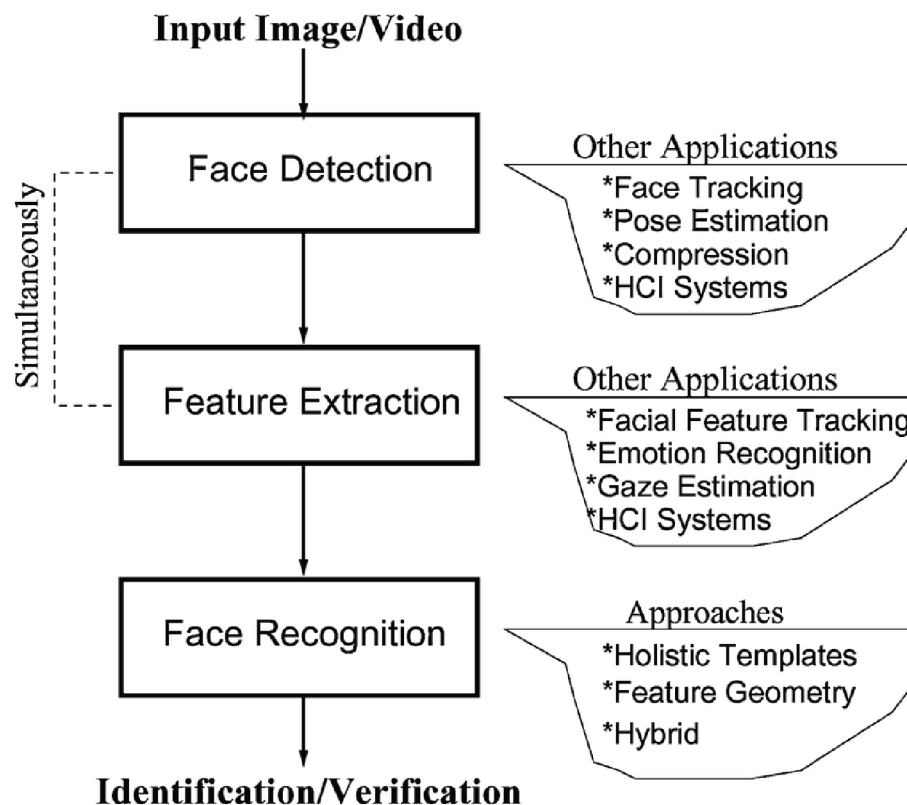


Figura 12 – Arquitetura de um sistema de reconhecimento facial genérico - extraída de (ZHAO et al., 2003).

### 2.7.1 Detecção

Detectar e rastrear faces em cenários compostos por um amontoado de outras informações é um estágio crucial na realização do reconhecimento facial automático. O processo geralmente começa com a busca de padrões faciais na imagem teste, como por exemplo olhos, boca e nariz. Após ser normalizada, a face é segmentada da imagem ou vídeo, já que

a maioria dos algoritmos de reconhecimento assumem que a região da face é conhecida (KONG et al., 2005).

A informações de cores da imagem são uma importante abordagem na detecção de faces, entretanto podem haver dificuldades em ambientes mal iluminados ou muito complexos. Além disso, as técnicas de detecção que utilizam os olhos como referência também são muito bem-sucedidas (KONG et al., 2005), apesar de serem afetadas por problemas como uso de óculos e más condições de iluminação. De uma maneira geral, as abordagens convencionais para a detecção facial podem ser divididas nas seguintes (YANG; KRIEGMAN; AHUJA, 2002):

**Knowledge-based:** Nesses tipos de métodos, o algoritmo é desenvolvido baseando-se em regras derivadas do conhecimento do pesquisador sobre a face humana. Geralmente, são denotados os relacionamentos entre as características faciais, como por exemplo o formato do rosto, e as distâncias entre olhos, boca e nariz.

**Invariant Features:** Consistem na localização de *features* robustas que se mantenham invariantes a mudanças na posição e iluminação. As características extraídas na maioria dos casos são: olhos, nariz, boca, sobrancelhas, linha do cabelo, e geralmente isso é feito utilizando detectores de bordas (*edges*).

**Template Matching:** As imagens de teste são comparadas a padrões faciais pré-definidos, que em sua maior parte são gerados com base em uma face frontal. A presença de um rosto na imagem é determinada através de valores de correlação entre ela e o *template* construído.

**Appearance-based:** Os padrões faciais são gerados por meio de um grupo de imagens de treinamento, capturando as variações presentes em todas as faces do conjunto. Posteriormente, eles são utilizados na comparação com a imagem de teste que se deseja classificar. Além disso, existe a possibilidade de que técnicas como essas sejam combinadas a métodos de redução de dimensionalidade, com o objetivo de minimizar o custo computacional.

A tabela (5) a seguir mostra exemplos de métodos de detecção desenvolvidos utilizando cada uma das abordagens mencionadas:

Tabela 5 – Categorização dos métodos de detecção facial - adaptada de (YANG; KRIEGMAN; AHUJA, 2002).

| Abordagem                    | Trabalhos Importantes                              |
|------------------------------|----------------------------------------------------|
| <i>Knowledge-based</i>       |                                                    |
|                              | Método de multi-resolução baseado em regras        |
| <i>Invariant features</i>    |                                                    |
| - Features faciais           | Agrupamento de arestas                             |
| - Textura                    | SGLD                                               |
| - Cor da pele                | Mistura Gaussiana                                  |
| - Múltiplas features         | Integração de cor, tamanho e forma                 |
| <i>Template matching</i>     |                                                    |
| - Templates pré-definidos    | Template de forma                                  |
| - Templates deformáveis      | Modelo de forma ativa (ASM)                        |
| <i>Appearance-based</i>      |                                                    |
| - Eigenfaces                 | Decomposição de autovetores e agrupamento          |
| - Distribution-based         | Distribuição gaussiana e MLP                       |
| - Redes Neurais              | Conjunto de redes neurais e esquemas de arbitragem |
| - Support Vector Machines    | SVM com kernel polinomial                          |
| - Classificadores Bayesianos | Estatísticas da aparência local e posição          |
| - Modelo de Markov           | Estatísticas de alta ordem com HMM                 |
| - Informação teórica         | Informação relativa de Kullback                    |

### 2.7.2 Extração de Features e Reconhecimento

O reconhecimento facial é uma tarefa extremamente desafiadora, e por esse motivo, a literatura é bastante diversificada quanto às abordagens relacionadas ao problema. Frequentemente, uma única aplicação utiliza múltiplos métodos motivados por diferentes princípios, isso faz com que esses sistemas sejam difíceis de se classificar baseando-se apenas nas técnicas de extração de *features* e identificação implementadas (ZHAO et al., 2003). Apesar disso, de uma maneira geral é possível definir as técnicas de reconhecimento em três categorias:

**Métodos Holísticos:** Desde 1990, esses tipos de técnicas foram a principal direção adotada nas abordagens relacionadas ao reconhecimento facial. Também conhecidos como métodos baseados em aparência (*appearance-based*), ou métodos globais, eles consistem em utilizar toda a informação contida na imagem sem que sejam extraídas *features* locais. Desta maneira, é possível preservar implicitamente todos os detalhes de textura e forma que contribuem bastante como características distintivas de uma face. No entanto, ao armazenar toda essa informação, é muito comum que métodos desse tipo sofram com a maldição da dimensionalidade<sup>1</sup> (*Curse of Dimensionality*) (TAN et al., 2006), algo que pode ser contornado projetando os dados sobre um subespaço ve-

<sup>1</sup> Termo utilizado para se referir a problemas que surgem durante a análise de dados em espaços com muitas dimensões.

torial (KONG et al., 2005), esta técnica é conhecida como redução de dimensionalidade. Dentre os métodos holísticos, os mais conhecidos (ZHAO et al., 2003) são:

- Eigenfaces: Baseado na análise de componentes principais (PCA).
- Fisherfaces: Utiliza uma combinação da análise de discriminante linear (LDA) e o PCA.
- Análise de componentes independentes (ICA).

**Métodos Locais:** Também conhecidos como métodos baseados em características (*feature-based*), estas abordagens consistem na extração de *features* relacionadas a pontos de referência (*keypoints*), como por exemplo ângulos e distâncias entre sobrancelhas, olhos, nariz, boca e queixo. Essas informações são aplicadas na construção de um descritor para a face, que é utilizado na comparação. Entre as vantagens dos métodos locais, está a baixa dimensionalidade do vetor de características extraído, o que possibilita um menor custo de memória em relação aos métodos holísticos (KONG et al., 2005). Além disso, também é possível a classificação baseando-se em uma característica específica da face, como por exemplo o tamanho do nariz, o que proporciona muita flexibilidade a esse tipo de abordagem (KONG et al., 2005). A performance desses métodos está intimamente ligada ao algoritmo utilizado na extração das *features*.

**Métodos Híbridos:** Assim como o sistema perceptivo humano, que usa todas as informações disponíveis para a avaliação, ou seja, tanto as características marcantes da face (locais) como todo o conjunto (globais). Existem técnicas que fazem o mesmo, tentando extrair o melhor de cada abordagem (ZHAO et al., 2003).

Os métodos híbridos combinam algoritmos globais e locais, de maneira que ambos possam oferecer ao sistema de reconhecimento informações complementares, incrementando a acurácia e a precisão da classificação. A tabela (6) apresenta a sensibilidade das principais abordagens, em relação às adversidades encontradas por um sistema de reconhecimento facial.

Tabela 6 – Comparação de sensibilidade a variações - adaptada de (TAN et al., 2006).

| Fontes de Variação | Locais              | Holísticos     |
|--------------------|---------------------|----------------|
| Pequenas variações | Baixa sensibilidade | Sensível       |
| Grandes variações  | Sensível            | Muito Sensível |
| Iluminação         | Muito Sensível      | Sensível       |
| Expressões         | Baixa sensibilidade | Sensível       |
| Posição            | Sensível            | Muito Sensível |
| Ruído              | Muito Sensível      | Sensível       |
| Oclusão            | Baixa sensibilidade | Muito Sensível |

A seguir, na tabela (7) estão apresentados os principais métodos de reconhecimento facial relacionados à cada tipo de abordagem:

Tabela 7 – Categorização dos principais métodos de reconhecimento facial conhecidos - adaptada de (ZHAO et al., 2003).

| Abordagem                                                                                                                             | Trabalhos Importantes                                                                                                                                                                                                                                         |
|---------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Holísticos                                                                                                                            |                                                                                                                                                                                                                                                               |
| * PCA:<br>- Eigenfaces<br>- Eigenfaces Probabilístico<br>- Fisherfaces<br>- SVM<br>- Algoritmos genéticos<br>- Feature lines<br>- ICA | Aplicação direta do PCA<br>Problema de duas classes com medidas de probabilidade<br>FLD sobre o autoespaço<br>Problema de duas classes baseado em SVM<br>Aprendizado GA aprimorado<br>Baseado na distância ponto-linha<br>Análise de features baseadas em ICA |
| * Outras representações:<br>- LDA/FLD<br>- PDBNN                                                                                      | LDA/FLD em uma imagem crua<br>NN baseada em decisões probabilísticas                                                                                                                                                                                          |
| Feature-based                                                                                                                         |                                                                                                                                                                                                                                                               |
| - Métodos de geometria pura<br>- Arquitetura de link dinâmico<br>- Modelo de Markov<br>- Redes convolucionais                         | Métodos antigos<br>Combinação de grafos<br>Método HMM<br>CNN baseada em aprendizado SOM                                                                                                                                                                       |
| Híbridos                                                                                                                              |                                                                                                                                                                                                                                                               |
| - Eigenfaces modular<br>- LFA híbrido<br>- Forma normalizada<br>- Component-based                                                     | Eigenfaces e eigenmodules<br>Método das features locais<br>Modelos de aparência flexível<br>Regiões e componentes da face                                                                                                                                     |

### 2.7.3 PCA

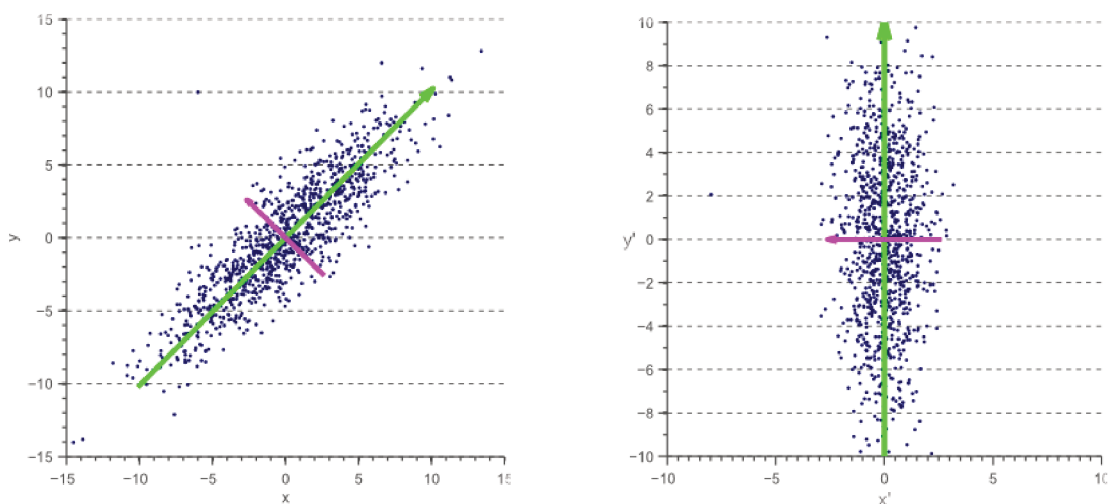
A Análise de Componentes Principais (*Principal Component Analysis* – PCA) é um método estatístico utilizado na redução da dimensionalidade de um conjunto de dados, o que possibilita sua aplicação para extração de *features*, compressão e visualização de dados (BISHOP, 2011). Para isso, são aplicadas transformações lineares sobre os padrões de entrada, descorrelacionando os dados, ou seja, removendo informações redundantes. Desta maneira, é obtido um conjunto menor de dados capazes de representar toda a variabilidade do conjunto original.

O PCA foi uma das primeiras técnicas desenvolvidas para a análise exploratória de dados e construção de modelos preditivos (ABDI; WILLIAMS, 2010). O método foi apresentado pelo britânico Karl Pearson em 1901, também sendo conhecido como Expansão Karhunen-Loeve. Segundo Bishop (2011), existem duas definições possíveis para o PCA que resultam em um mesmo algoritmo:

**Máxima Variância:** Uma projeção ortogonal dos dados sobre um espaço dimensional menor, conhecido como subespaço principal, de maneira que a variância dos dados projetados é maximizada.

**Erro Mínimo:** A projeção linear que minimiza o custo médio de projeção, definido pela média da distância entre os dados e suas projeções elevada ao quadrado.

Em outras palavras, a transformação linear gerada pelo PCA projeta os dados de um espaço  $n$ -dimensional para um espaço  $m$ -dimensional, tal que  $m < n$ . Esse novo sistema de coordenadas obtido é ortonormal, e captura a maior variância entre os dados (ABDI; WILLIAMS, 2010). Isso é feito calculando os autovalores (*eigenvalues*) e autovetores (*eigenvectors*) da matriz de covariância do conjunto de dados inicial. Os autovetores associados aos maiores autovalores são chamados de componentes principais, pois são responsáveis por representar a maior variabilidade dos dados. A figura (13) mostra um conjunto de dados de entrada e suas componentes principais, em seguida os dados descorrelacionados.



(a) Dados correlacionados antes do PCA.

(b) Dados descorrelacionados depois do PCA.

Figura 13 – Conjunto de dados e suas componentes principais em verde e rosa - extraída de (SPRUYT, 2014).

### 2.7.3.1 Eigenfaces

O Eigenfaces é uma das abordagens holísticas mais famosas no reconhecimento facial. A ideia teve suas raízes nos experimentos conduzidos por Kirby e Sirovich (1990), Sirovich e Kirby (1987), onde a análise de componentes principais (PCA) foi utilizada para representar de forma eficiente uma imagem. Deste modo, foi proposto que qualquer face poderia ser reconstruída por uma pequena coleção de pesos e um conjunto de faces padrão conhecidas por autofaces (*eigenfaces*), como mostra o exemplo da figura (14). Esses pesos podem ser obtidos projetando a imagem de uma face de teste sobre o espaço gerado pelas autofaces.

Posteriormente, a ideia foi aplicada com sucesso para a detecção e reconhecimento facial por [Turk e Pentland \(1991\)](#), e assim a técnica ficou conhecida como Eigenfaces.



Figura 14 – Reconstrução de uma face utilizando o banco ([CAMBRIDGE, 2002](#)).

O algoritmo proposto por [Turk e Pentland \(1991\)](#) é composto por uma etapa de treinamento, onde as *eigenfaces* são obtidas e as imagens do conjunto de treino são projetadas gerando um novo espaço de faces. Posteriormente, inicia-se a fase de teste, que consiste em projetar uma face de entrada sobre o espaço gerado, e comparar sua distância com as demais imagens do conjunto de treinamento, a fim de obter uma identificação. A seguir, ambos os procedimentos estão apresentados de uma maneira mais detalhada.

#### 2.7.3.1.1 Treino

Essa etapa consiste no cálculo do PCA para o conjunto de treinamento, obtendo assim as *eigenfaces* que serão utilizadas na projeção dos dados de treino e da face de teste. Todas as imagens do conjunto de treinamento devem possuir as mesmas dimensões e as faces devem estar alinhadas. Além disso, é ideal que o cenário e quaisquer outros objetos “não-face” sejam removidos da imagem. Deste modo, o treinamento pode ser iniciado, sendo composto das seguintes etapas:

1. Dado um conjunto de  $M$  imagens faciais na forma de matrizes de intensidade com dimensões  $h \times w$ , onde seus termos representam cada *pixel* da imagem, sendo  $h$  o número de linhas (altura) e  $w$  o número de colunas (largura). O primeiro passo consiste na geração dos vetores-imagem  $\Gamma_i$  de tamanho  $D = hw$ , que são obtidos concatenando cada uma das matrizes da seguinte maneira (2.5):

$$Imagem_i = Matriz_i = \begin{bmatrix} m_{11} & m_{12} & \dots & m_{1w} \\ m_{21} & m_{22} & \dots & m_{2w} \\ \vdots & \vdots & \ddots & \vdots \\ m_{h1} & m_{h2} & \dots & m_{hw} \end{bmatrix} \Leftrightarrow Vetor_i = \begin{bmatrix} m_{11} \\ m_{12} \\ \vdots \\ m_{1w} \\ m_{21} \\ m_{22} \\ \vdots \\ m_{2w} \\ \vdots \\ m_{h1} \\ m_{h2} \\ \vdots \\ m_{hw} \end{bmatrix} = \Gamma_i \quad (2.5)$$

obtendo assim, o conjunto de vetores-imagem (2.6)

$$\Gamma = [\Gamma_1, \Gamma_2, \dots, \Gamma_M] \quad (2.6)$$

2. O próximo passo consiste em remover a redundância do conjunto obtido, centralizando os dados, isso pode ser feito calculando a face média (2.7), representada na figura (15), e subtraindo-a das faces de treino.



Figura 15 – Face média do banco (CAMBRIDGE, 2002).

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \quad (2.7)$$

Cada imagem do conjunto de treino se distancia da face média de acordo com (2.8):

$$\Phi_i = \Gamma_i - \Psi \quad (2.8)$$

onde  $i = 1, 2, \dots, M$ . Desta maneira é obtido o vetor  $\Phi_i$  que contém as variações da face  $\Gamma_i$  em relação a face média  $\Psi$ , ou seja, cada vetor preserva apenas as características únicas de cada face, gerando assim o conjunto de faces de treinamento  $A$  (2.9) normalizado.

$$A = [\Phi_1, \Phi_2, \dots, \Phi_M] \quad (2.9)$$

3. Na etapa seguinte, é necessário calcular a matriz de covariância dada por (2.10):

$$C = AA^T \quad (2.10)$$

e encontrar seus autovalores  $\lambda_k$  contidos na diagonal principal e autovetores  $u_k$  associados, tais que (2.11):

$$(AA^T)u_k = \lambda_k u_k \quad (2.11)$$

Nesse momento surge um problema, para os casos onde  $M < D$ , o cálculo da matriz de covariância torna-se extremamente dispendioso. Mesmo se imagens pequenas, de  $100 \times 100$  *pixels*, forem utilizadas, será gerado um espaço de  $100^2 = 10000$  dimensões. Deste modo, a matriz de covariância calculada possuirá dimensões  $10000 \times 10000$ , contendo  $10^8$  elementos, o que torna os cálculos inviáveis. Para contornar esse problema, é possível encontrar os autovalores  $\mu_i$  e autovetores  $v_i$  de  $A^T A$ , como demonstrado em (2.12). Isso facilita bastante o processo, pois a matriz  $AA^T$  possui dimensões  $D \times D$ , já o cálculo de  $A^T A$  resulta em uma matriz de tamanho  $M \times M$ .

$$(A^T A)v_i = \mu_i v_i \quad (2.12)$$

Ao multiplicar ambos os lados de (2.12) por  $A$ , obtém-se (2.13):

$$(AA^T)Av_i = \mu_i (Av_i) \quad (2.13)$$

Assim, é possível perceber que  $Av_i = u_k$ , ou seja, os autovetores  $u_k$  de  $AA^T$  são combinação linear dos autovetores de  $A^T A$ , com a matriz normalizada  $A$ . Além disso, ambas

as matrizes possuem os mesmos autovalores associados, ou seja,  $\lambda_k = \mu_i$ . Desta maneira, obtém-se as autofaces dos dados originais como mostra a figura (16), gerando o conjunto  $p_i$  (2.14).



Figura 16 – Conjunto das 20 primeiras *eigenfaces* geradas com o banco (CAMBRIDGE, 2002).

$$p_i = \sum_{j=1}^M v_{ji} \Phi_j \quad (2.14)$$

4. Finalmente, basta selecionar as componentes principais, ou seja, os autovetores associados aos maiores autovalores, e utilizá-las para representar todo o conjunto de dados. Isso é feito projetando as faces de treino normalizadas sobre o novo espaço de faces, encontrando assim os pesos  $w_i$  necessários para representar cada uma delas da seguinte maneira (2.15):

$$w_i = p_k^T (\Gamma_k - \Psi) \quad (2.15)$$

para  $k = 1, 2, \dots, M'$ , onde  $M'$  é a quantidade de *eigenfaces* selecionadas.

Os pesos podem então ser utilizados para representar cada uma das imagens do conjunto de treino, ao invés de todos os seus *pixels*. Deste modo, deixando a variabilidade por conta das *eigenfaces*, que são capazes de representar todo o conjunto de dados.

## 2.7.3.1.2 Teste

Seja  $\Gamma_{teste}$  uma imagem de entrada contendo a face que se deseja classificar, e  $\Psi$  a face média do conjunto de treinamento, a fase de teste pode ser dividida nas seguintes etapas:

1. Subtrair a face média  $\Psi$  da face teste  $\Gamma_{teste}$ , obtendo a imagem normalizada  $\Phi_{teste}$  de acordo com (2.16):

$$\Phi_{teste} = \Gamma_{teste} - \Psi \quad (2.16)$$

2. Projetar a face teste sobre o subespaço das *eigenfaces*, encontrando desta maneira seu conjunto de pesos  $w_{teste}$ , como demonstrado em (2.17):

$$w_{teste} = p_k^T \Phi_{teste} \quad (2.17)$$

3. Comparar os pesos obtidos com as projeções geradas na fase de treinamento. Isso pode ser feito com a ajuda de um classificador, como por exemplo o k-NN ou uma rede neural.

Após a execução do Eigenfaces, um subespaço de faces é gerado. Deste modo, a detecção facial no algoritmo também pode ser encarada como a proximidade entre a imagem de teste e este espaço. Assim sendo, uma imagem qualquer pode ser projetada, mesmo que nenhuma face esteja contida nela, como mostra a figura (17).



Figura 17 – Exemplos de projeções utilizando o Eigenfaces sobre a base de dados (CMU-PIE, 2008): na primeira linha, as imagens originais de uma face e um objeto não-face, e na segunda, as imagens resultantes - extraída de (TURK, 2008).

Para evitar que uma imagem desse tipo seja classificada gerando um falso positivo, é possível estabelecer um valor de distância mínima conhecido como *threshold*. Caso a distância entre a imagem projetada e o conjunto de treino seja maior do que esse valor, a entrada é considerada como “não-face”. Isso ajuda a diminuir as taxas de erro do algoritmo, melhorando a performance do sistema de reconhecimento. O valor ideal para o *threshold* vai variar de acordo com o conjunto de treinamento disponível, podendo ser obtido através de métodos de validação, que por sua vez dependem do classificador utilizado.

#### 2.7.4 LDA

A Análise de Discriminante Linear (*Linear Discriminant Analysis* – LDA), também conhecida como Análise Discriminante de Fisher, é uma técnica estatística aplicada na redução da dimensionalidade de um conjunto de dados, assim como o PCA. Ela permite discriminar e classificar observações com base em um número de variáveis independentes, conhecidas como variáveis discriminantes (DUDA; HART; STORK, 2000). Isso é feito através do cálculo de uma função discriminante composta por índices e pesos específicos, porém diferentemente do PCA, que realiza uma projeção buscando preservar a variância dos dados, o LDA procura por eixos que maximizam a separabilidade entre as classes, como mostra a figura (18):

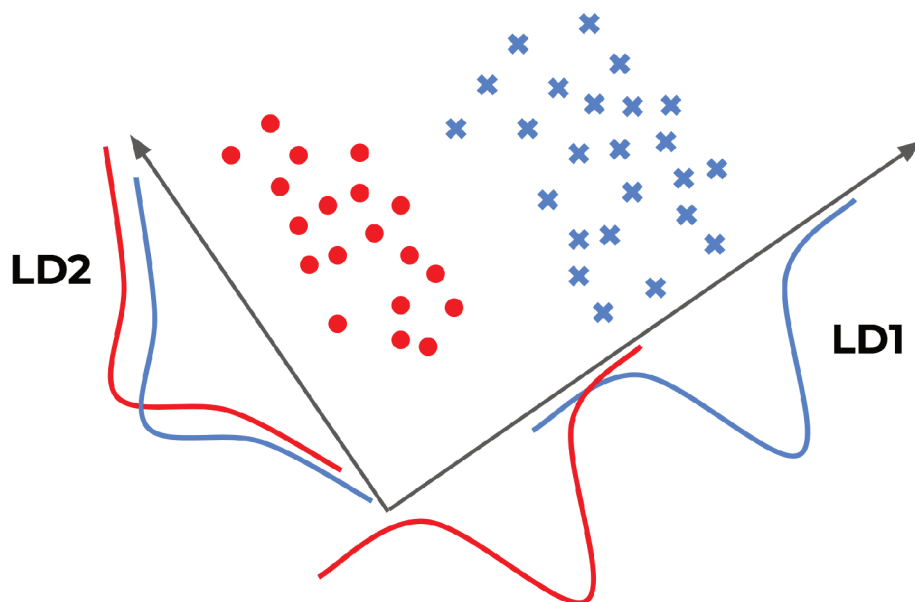


Figura 18 – LD1 representa o eixo que provavelmente seria escolhido pelo LDA, aumentando a separabilidade entre as classes, ao contrário de LD2 - extraída de (HO, 2017).

O principal objetivo do LDA é a projeção do conjunto de dados sobre um sistema de coordenadas de menor dimensionalidade, de modo que a dispersão (*scatter*) intra-classe é

minimizada e a dispersão entre-classes é maximizada. Dessa maneira, tornando separação entre as classes, através de uma medida de similaridade, mais eficiente.

#### 2.7.4.1 Fisherfaces

O Fisherfaces é um algoritmo de reconhecimento facial proposto por [Belhumeur, Hespanha e Kriegman \(1997\)](#), que originalmente pesquisavam a aplicação do LDA em sistemas de processamento linguísticos. O método combina o PCA com o discriminante linear de Fisher para produzir um subespaço vetorial de maneira bem similar ao Eigenfaces, porém com a separabilidade entre as classes aumentada, o que aprimora bastante as taxas de reconhecimento em relação aos métodos que utilizam apenas o PCA.

Da mesma maneira que o Eigenfaces, o método é dividido em etapas de teste e treino, que ocorrem utilizando a mesma lógica, com a diferença que o Fisherfaces adiciona alguns passos a mais no algoritmo, projetando os dados gerados pelo PCA novamente sobre um eixo que maximize a dispersão entre-classe, e minimize a dispersão intra-classe. Isso é feito da seguinte maneira:

1. Considerando uma imagem  $\Gamma_{ji}$  como a  $i$ -ésima imagem da uma classe de imagens  $C_j$ , a face média da classe é denotada por  $\Psi_j$ , da seguinte maneira (2.18):

$$\Psi_j = \frac{1}{c_j} \sum_{i=1}^{c_j} \Gamma_{ji} \quad (2.18)$$

onde  $c_j$  é a quantidade de faces que pertencem a cada uma das  $k$  classes, tal que:

$$C_1, C_2, \dots, C_k \quad (2.19)$$

onde

$$C_j = \Gamma_{j1}, \dots, \Gamma_{jc_j} \quad (2.20)$$

2. Calcula-se então a dispersão das faces médias pertencentes a cada classe em relação a face média geral  $\Psi$ , como mostra a fórmula (2.21):

$$\Psi = \frac{1}{\sum_{j=1}^k c_j} \sum_{j=1}^k \sum_{i=1}^{c_j} \Gamma_{ji} \quad (2.21)$$

3. Em seguida, obtém-se as matrizes de dispersão intra-classe  $S_w$  (2.22) e entre-classe  $S_b$  (2.23).

$$S_w = \sum_{j=1}^k \sum_{i=1}^{c_j} (\Gamma_{ji} - \Psi_j)(\Gamma_{ji} - \Psi_j)^T \quad (2.22)$$

$$S_b = \sum_{j=1}^k c_j (\Psi_j - \Psi)(\Psi_j - \Psi)^T \quad (2.23)$$

4. O critério de Fisher, denotado por  $P_{LDA}$ , pode então ser obtido por meio da fórmula (2.24), selecionando o valor que maximize a razão dos determinantes das matrizes  $S_b$  e  $S_w$ .

$$P_{LDA} = \max(P) \frac{|P^T S_b P|}{|P^T S_w P|} \quad (2.24)$$

Onde  $P = [P_1, P_2, \dots, P_M]$  é o conjunto de autovetores de  $S_b$  e  $S_w$  associados aos maiores autovalores  $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_M]$ .

5. O próximo passo consiste em obter a projeção dos dados com o PCA, da seguinte maneira (2.25):

$$P_{PCA} = \max(P) |P^T (S_w + S_b) P| \quad (2.25)$$

6. Por fim, após a utilização do PCA para redução da dimensionalidade no conjunto inicial de dados, é possível obter uma projeção em um novo espaço vetorial utilizando o LDA calculado, como mostra a fórmula (2.26):

$$P_{otima} = P_{LDA} P_{PCA} \quad (2.26)$$

Desta maneira, gerando as *fisherfaces* para o conjunto de dados, como exemplificado na figura (19):

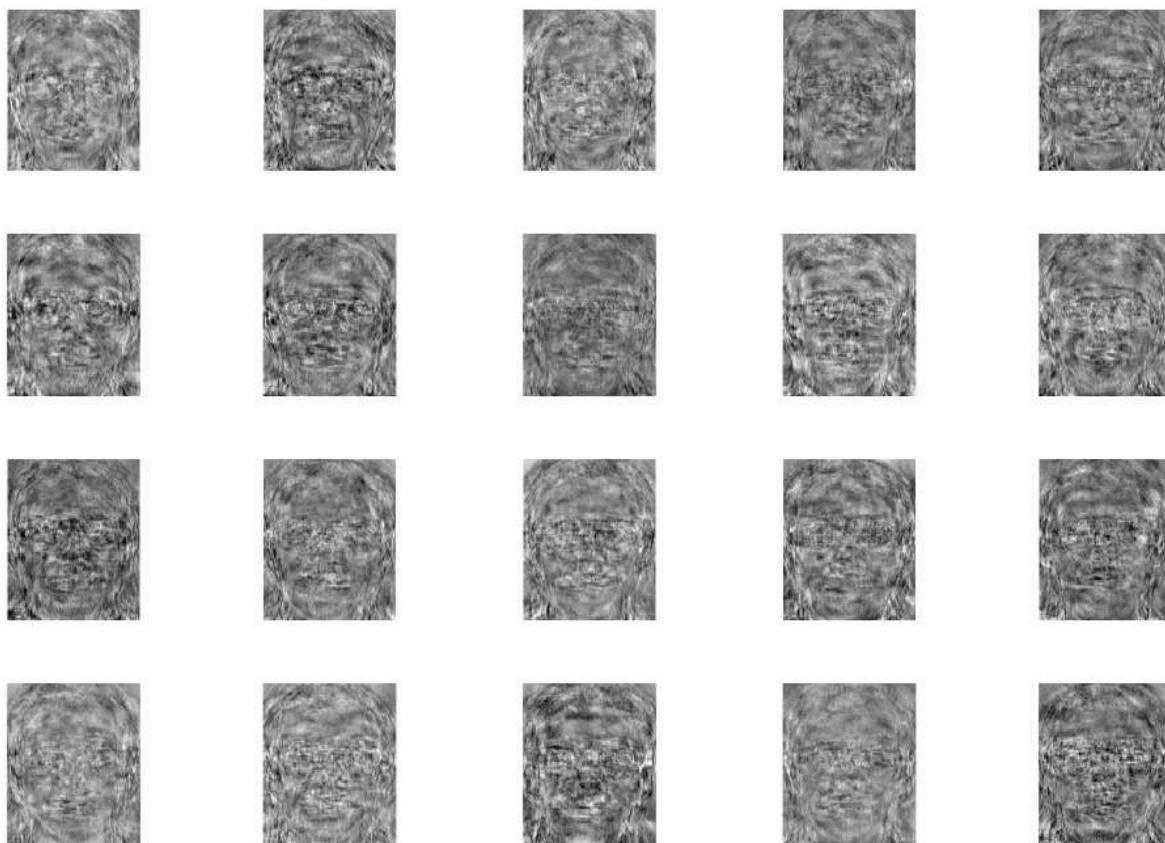


Figura 19 – Conjunto da 20 primeiras *fisherfaces* gerados com o banco ([CAMBRIDGE, 2002](#)).

### 3 Trabalhos Relacionados

O reconhecimento facial por meio de imagens do espectro visível é uma área madura e bem-sucedida. Para chegar à essa conclusão, basta observar o desempenho desses tipos de sistemas quando aplicados à imagens em alta resolução e com boas condições de iluminação. Além de disso, eles contam com uma grande variedade de abordagens tanto para a detecção quanto para a identificação facial, e dezenas de trabalhos utilizando essas técnicas de formas complementares já foram produzidos. No entanto, todos esses métodos possuem um problema em comum: a sensibilidade a variações de iluminação, que é uma característica inerente ao espectro visível, e altera de forma significativa as imagens a ponto de reduzir bastante a eficiência desses sistemas, dependendo das condições.

A utilização de imagens térmicas no reconhecimento surgiu como uma alternativa para esse problema, e desde então alguns trabalhos já foram realizados abordando o tema. Mesmo que as técnicas de reconhecimento facial aplicadas à imagens térmicas ainda estejam em seus estágios iniciais de desenvolvimento, elas já oferecem resultados promissores, tornando-se uma área de pesquisa bastante atrativa.

O trabalho realizado por [Prokoski, Riedel e Coffin \(1992\)](#) foi um dos pioneiros em relação ao uso de imagens infravermelhas como solução para os problemas do reconhecimento facial. Baseando-se no cenário tecnológico e no estado da arte da biometria facial da época, [Prokoski, Riedel e Coffin \(1992\)](#) elencaram os principais empecilhos para a aplicação comercial dos sistemas de reconhecimento facial baseados em imagens do espectro visível, e compararam como eles afetariam a termografia infravermelha. Além disso, foram sugeridas diversas vantagens que poderiam justificar a utilização das imagens térmicas, além da invariância à luminosidade, como por exemplo uma menor sensibilidade a mudanças de posição, expressões, disfarces e também vantagens menos imediatas, como a possibilidade de se utilizar o padrão de vasos sanguíneos faciais como traço biométrico. Desta maneira, para fundamentar tais afirmações, introduziram o conceito das formas elementares, exemplificadas na figura (20), que funcionariam como impressões digitais para uma imagem térmica, viabilizando sua aplicação. Após a apresentação de várias aplicações e métodos que poderiam ser utilizados com a termografia infravermelha, [Prokoski, Riedel e Coffin \(1992\)](#) também abordaram as desvantagens do uso desse tipo de imagem, como por exemplo vulnerabilidades às condições de saúde e emocionais dos indivíduos, além de alterações na temperatura ambiente.

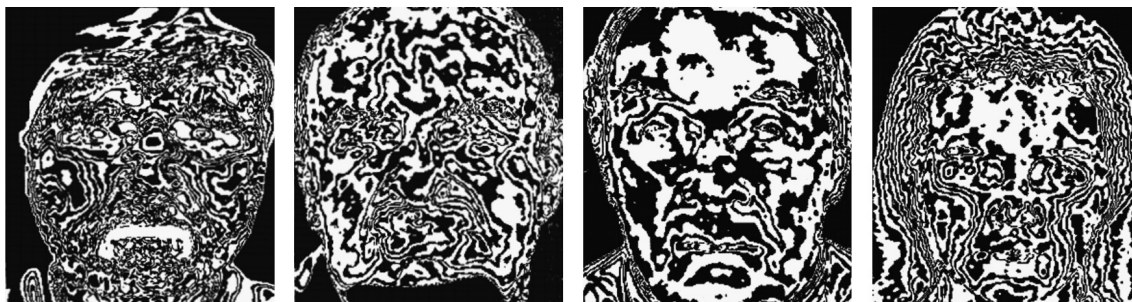


Figura 20 – Exemplo de formas elementares - extraída de (PROKOSKI; RIEDEL; COFFIN, 1992).

Uma das primeiras aplicações bem-sucedidas das imagens infravermelhas no reconhecimento facial foi reportada por Cutler (1996), que usou um método baseado inteiramente no Eigenfaces. Para isso, foi utilizada uma base de dados com 24 indivíduos, contendo 12 imagens para cada. As taxas de reconhecimento obtidas foram bastante satisfatórias, atingindo 96% de acerto em imagens frontais, e incríveis 100% em fotos de perfil. No entanto, as imagens utilizadas não possuíam quase nenhuma variação de posição e expressão, tornando os resultados demasiadamente otimistas e improváveis para uma aplicação prática do sistema. Mesmo assim, os resultados alcançados por Cutler (1996) demonstraram que as imagens térmicas não apenas eram aplicáveis ao reconhecimento facial, como também eram viáveis até mesmo utilizando um método que não foi desenvolvido para esse tipo de imagem.

Enquanto isso, Wilder et al. (1997) desenvolviam de maneira independente uma aplicação que também utilizava imagens infravermelhas no reconhecimento, além de elaborarem um conjunto de imagens próprias para seu treinamento e teste, de modo similar ao que ocorria com outros bancos de imagens existentes para o espectro visível. A grande diferença desse trabalho para os demais, é que ele se propunha a comparar objetivamente as taxas de reconhecimento entre as imagens de ambos os espectros, desta maneira não apenas testando a viabilidade das imagens térmicas, mas também medindo sua eficiência frente aos outros métodos disponíveis na época. Para isso, foram colhidas 8 imagens de cada indivíduo, 4 no espectro visível e 4 infravermelhas. Dentre as imagens obtidas para cada espectro, 2 delas eram frontais e as outras 2 imagens foram feitas em um ângulo de 22,5°, e por fim, para cada par de imagens, uma foi obtida em posição estática, e a outra foi registrada enquanto a pessoa falava. O objetivo das variações aplicadas era comparar as taxas de reconhecimento utilizando os dois tipos de imagens, em situações um pouco mais realistas. Para isso, o treino e o teste foram realizados em imagens sob diferentes condições, como apresentado na tabela (8) a seguir:

Tabela 8 – Metodologia dos testes aplicados - adaptada de (WILDER et al., 1997).

| Treino                          | Teste                            |
|---------------------------------|----------------------------------|
| Visível - Estática (VS5)        | Visível - Movimento (VT11)       |
| Visível - Movimento (VT5)       | Visível - Estática (VS11)        |
| Infravermelho - Estática (IRS5) | Infravermelho- Movimento (IRT11) |
| Infravermelho- Movimento (IRT5) | Infravermelho - Estática (IRS11) |

Os algoritmos utilizados no treinamento e no reconhecimento foram o Eigenfaces e o Matching Pursuit Filters (MPF) combinado a diferentes tipos de *features*, como por exemplo olhos, nariz, boca e sobrancelhas. Desta maneira, os resultados obtidos corroboram com os resultados alcançados por Cutler (1996), porém as taxas de reconhecimento obtidas com as imagens infravermelhas foram inferiores às do espectro visível, como mostra a tabela (9):

Tabela 9 – Resultados - adaptada de (WILDER et al., 1997).

| Treino > Teste | Eigenfaces | MPF |
|----------------|------------|-----|
| IRS5 > IRT11   | 85%        | 93% |
| IRT5 > IRS11   | 85%        | 95% |
| VS5 > VT11     | 86%        | 95% |
| VT5 > VS11     | 88%        | 97% |

Além disso, o trabalho também apontou uma grande vulnerabilidade das imagens térmicas: a utilização de óculos, que bloqueia a radiação infravermelha ocultando regiões discriminatórias importantes do rosto. No entanto, um fator muito importante não foi abordado nos experimentos: a variação de iluminação, que é a principal justificativa da utilização de imagens térmicas, e que poderia trazer resultados significativos e esclarecedores.

Posteriormente, uma série de artigos (SOCOLINSKY et al., 2001; Socolinsky; SALGIAN, 2002; Socolinsky; SALGIAN; NEUHEISEL, 2003) abordando o tema foram publicados por Socolinsky enquanto ele trabalhava na Corporação Equinox, que se tornaria uma das empresas mais importantes para o desenvolvimento de tecnologias relacionadas à radiação infravermelha. Em seus trabalhos, foram analisadas a viabilidade e a eficiência das imagens infravermelhas no reconhecimento facial, assim como a comparação do seu desempenho frente às imagens do espectro visível. Para isso, foi construída uma base de dados com 91 indivíduos, obtendo simultaneamente imagens do espectro visível e infravermelho. Além disso, também foi percebida a necessidade de se investigar o efeito das variações mais comuns sobre as taxas de reconhecimento, para isso as imagens foram classificadas nos seguintes tipos:

- Movimento
- Estática

- Iluminação
- Expressão
- Óculos

Esses tipos foram combinados de diversas maneiras, gerando um total de 24.500 imagens, além de um conjunto de 500 imagens classificadas como “aleatórias” pelos pesquisadores, que continham poses e iluminações arbitrárias.

Os algoritmos utilizados por [Socolinsky et al. \(2001\)](#) foram o ARENA e o Eigenfaces, e os resultados obtidos não apenas corroboraram com a narrativa de que as imagens térmicas seriam uma abordagem viável, como também corresponderam às expectativas dos pesquisadores da área no que diz respeito às taxas de reconhecimento, como apresentado na tabela (10).

Tabela 10 – Taxas de acerto: média/mínima - adaptada de ([SOCOLINSKY et al., 2001](#)).

| <b>Espectro</b> | <b>Arena</b> | <b>Eigenfaces</b> |
|-----------------|--------------|-------------------|
| Visível         | 0.860/0.765  | 0.639/0.319       |
| Infravermelho   | 0.990/0.980  | 0.933/0.868       |

Após obter resultados animadores em seu primeiro trabalho, [Socolinsky e Salgian \(2002\)](#) decidiram abordar outros tipos de algoritmos, comparando os performance entre os dois espectros da mesma maneira do trabalho anterior. As técnicas utilizadas foram algoritmos baseados no PCA, LDA, LFA, ICA e algumas de suas variações, obtendo os resultados da tabela (11).

Tabela 11 – Taxas de acerto: média/mínima/máxima - adaptada de ([SOCOLINSKY et al., 2001](#)).

| <b>Algoritmo</b>       | <b>Visível</b> | <b>Infravermelho</b> |
|------------------------|----------------|----------------------|
| <i>PCA</i>             | 0.73/0.36/0.97 | 0.95/0.89/1.00       |
| <i>LDA<sub>g</sub></i> | 0.93/0.85/0.99 | 0.97/0.92/1.00       |
| <i>LDA<sub>t</sub></i> | 0.92/0.76/1.00 | 0.98/0.94/1.00       |
| <i>LFA<sub>e</sub></i> | 0.82/0.62/0.97 | 0.93/0.84/0.99       |
| <i>LFA<sub>b</sub></i> | 0.85/0.63/0.98 | 0.93/0.83/0.99       |
| <i>ICA</i>             | 0.88/0.72/0.99 | 0.94/0.86/1.00       |

Dentre os trabalhos mais recentes, o de [Chen, Flynn e Bowyer \(2003\)](#), pela Universidade de Notre Dame, foi um dos que obteve maior destaque. A pesquisa se consistiu na coleta de pares de imagens no espectro visível e infravermelho de 101 indivíduos, e as variações presentes nas imagens foram de expressão, pose e iluminação, baseando-se nos trabalhos anteriormente citados.

O grande diferencial desse trabalho, está no intervalo de tempo entre a coleta das imagens, gerando um conjunto que aborda não apenas as diferenças diretas entre as imagens nos dois espectros, como também os efeitos de mudanças provocadas pelo tempo. As imagens foram colhidas durante um período de 10 semanas, e embora seja um intervalo relativamente curto, seus resultados foram extremamente relevantes para o estado da arte. Os algoritmos utilizados foram o Eigenfaces e o FaceIt, além da fusão dos dois tipos de imagens já propostas em alguns trabalhos anteriores.

Os resultados mostraram que as imagens do espectro infravermelho obtiveram taxas de reconhecimento inferiores às do espectro visível, porém a fusão dos dois tipos de imagens obteve o melhor resultado. Apesar disso, segundo as conclusões citadas no próprio artigo, é possível que as variações de luminosidade utilizadas para os procedimentos não fossem realmente significativas para que as imagens do espectro IR pudessem se destacar, gerando assim um resultado parecido com a comparação de imagens sem alterações de iluminação.

Dentre todos os trabalhos desenvolvidos sobre o tema, ainda não existe um consenso de que nível de variação na iluminação poderia proporcionar melhores resultados para as imagens infravermelhas, ou quais seriam as bandas do espectro infravermelho adequadas para o reconhecimento facial, assim como as taxas de acerto para cada uma delas. Apesar disso, em todos os trabalhos pesquisados fica clara a viabilidade do reconhecimento nesse tipo de imagem, e também a possibilidade de que a junção de ambos os espectros possa obter melhores resultados do que eles separadamente. Deste modo, esse trabalho se propõe a investigar mais a fundo os efeitos das variações de iluminação, posição e expressões faciais nas imagens do espectro visível e infravermelho, delimitando os melhores e piores cenários para cada tipo.

## 4 Materiais e Métodos

Nesse capítulo, são abordadas as ferramentas e as metodologias utilizadas no desenvolvimento deste trabalho. Além disso, é apresentada a estrutura do banco de imagens utilizado, assim como a arquitetura do sistema de reconhecimento facial implementado e os experimentos conduzidos para sua avaliação.

### 4.1 Desenvolvimento

O sistema de reconhecimento facial foi desenvolvido por meio do software MATLAB (Matrix Laboratory) ([MATHWORKS, 2018](#)), que foi escolhido devido à sua alta performance relacionada aos cálculos matriciais envolvidos na implementação dos algoritmos selecionados. Além disso, o programa também conta com um conjunto de ferramentas dedicadas ao processamento de imagens e visualização de dados, oferecendo assim uma série de vantagens na sua aplicação à visão computacional.

O primeiro algoritmo de reconhecimento implementado foi o Eigenfaces ([TURK; PENTLAND, 1991](#)), por ser o método mais conhecido e estudado na literatura, fornecendo assim uma grande quantidade de informações e resultados obtidos da sua aplicação nos mais diversos cenários. Esses dados servirão na comparação dos resultados obtidos por este trabalho, não apenas para a verificação da corretude das implementações, mas também para a validação dos resultados alcançados e dos novos conhecimentos produzidos.

O segundo método abordado foi o Fisherfaces ([BELHUMEUR; HESPANHA; KRIEGMAN, 1997](#)), sua escolha se deu por possuir um desempenho sensivelmente melhor quando comparado ao do Eigenfaces, algo que é bem documentado na literatura. Desta maneira, é possível verificar se os resultados obtidos pelo Eigenfaces se sustentam em um algoritmo derivado dele, e se também é possível que os benefícios das imagens infravermelhas escalem da mesma maneira que no primeiro método.

Por fim, as *features* extraídas pelas duas técnicas de reconhecimento serão comparadas utilizando o algoritmo dos  $k$  vizinhos mais próximos ( $k$ -NN), que foi escolhido não só por sua simplicidade e eficiência, mas também pela possibilidade de modificação de alguns parâmetros, como por exemplo os valores de  $k$  e a utilização de diversas medidas de similaridade, que podem ter sua eficiência comparada frente aos tipos de imagem utilizadas.

## 4.2 Arquitetura do Sistema

O sistema desenvolvido foi dividido em módulos, como mostra a figura (21), com o objetivo de facilitar sua implementação. Cada parte possui uma função específica, sendo através da interação entre elas que a tarefa do reconhecimento se torna possível.

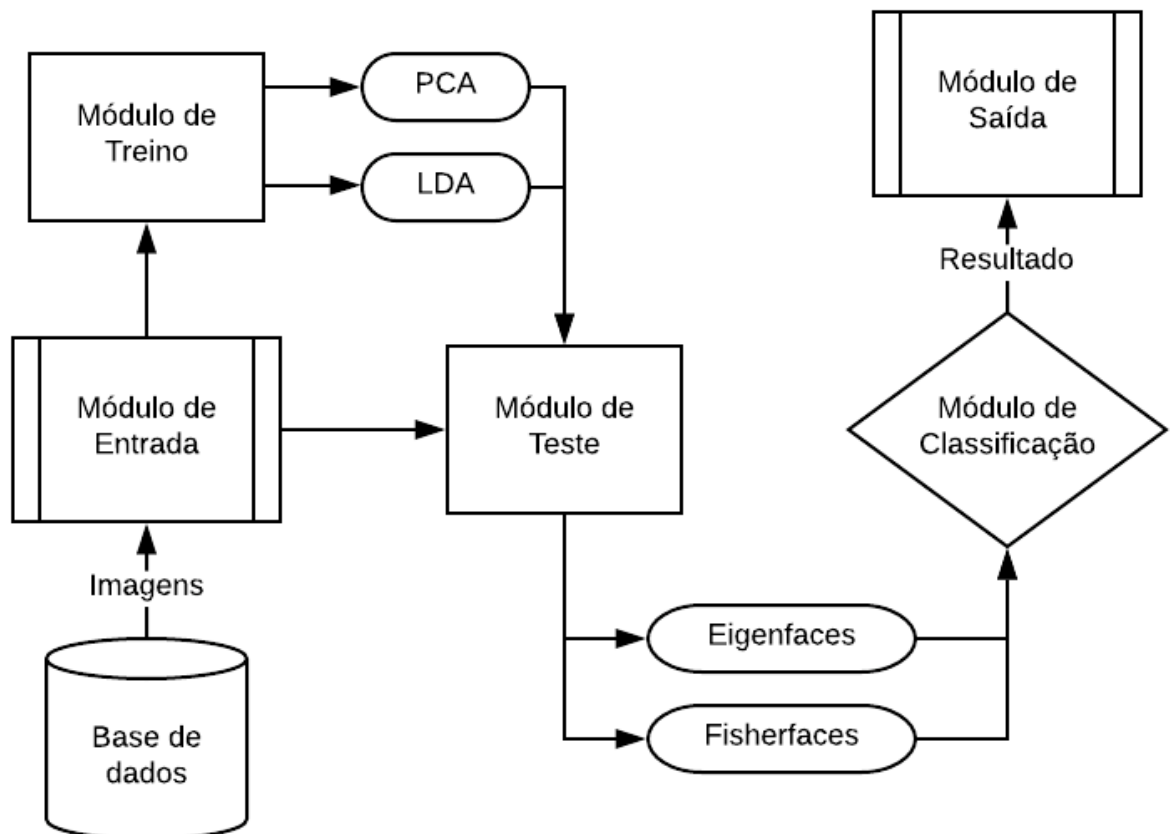


Figura 21 – Arquitetura do sistema de reconhecimento facial implementado.

Cada módulo executa suas tarefas da seguinte maneira:

**Entrada:** Recebe os caminhos para os diretórios onde estão localizadas as bases de dados que serão utilizadas nas etapas de treinamento e teste. As imagens pertencentes a um mesmo indivíduo devem estar separadas das demais em subdiretórios, de acordo com o seu tipo. Esse módulo pode ser adaptado para a estrutura de cada banco, dependendo da variedade de imagens e indivíduos presentes. Após a leitura dos dados de entrada, o módulo realiza a etapa de pré-processamento, convertendo cada imagem para uma escala de cinza, caso seja necessário, e construindo o vetor-imagem. Em seguida, é gerada uma matriz para cada um dos conjuntos de imagens selecionados, concatenando os seus vetores. Por fim, a matriz é alocada em um diretório específico pelo usuário, para evitar que o processo seja repetido em cada execução do

programa, podendo assim ser acessada livremente sem a necessidade de execução do módulo de entrada, graças ao processo de pré-computação executado.

**Treino:** Esse módulo é responsável pela extração das *features* dos conjuntos de treinamento selecionados. Para isso, de acordo com o algoritmo de reconhecimento que está sendo executado, são aplicados o PCA ou LDA sobre as matrizes geradas pelo módulo de entrada. O resultado é o conjunto das componentes principais, cuja quantidade pode ser selecionada por parâmetro. Elas serão utilizadas para representar os dados iniciais, que por sua vez são projetados sobre o novo espaço vetorial, gerando uma matriz de dimensionalidade reduzida, que é utilizada na classificação.

**Teste:** O vetor-imagem da face que se deseja reconhecer é recebido e projetado por esse módulo, juntamente com as matrizes das imagens de treinamento geradas pelo módulo de entrada. A projeção é feita multiplicando esses dados pelas componentes fornecidas pelo módulo treino, obtendo assim seus respectivos pesos. Desta maneira, um novo espaço de faces é gerado e enviado ao módulo de classificação.

**Classificação:** Responsável por receber as projeções dos dados de treino e teste, compará-las e classificar a imagem de entrada de acordo com os padrões de treinamento. Isso é feito através do k-NN, especificando a quantidade k de vizinhos desejada, assim como a métrica utilizada entre:

- Euclidiana
- Manhattan
- Chebyshev
- Mahalanobis

O algoritmo classifica a imagem de entrada de acordo com a sua distância em relação às outras faces localizadas no espaço gerado pelo módulo de teste.

**Saída:** Recebe e exibe o resultado da classificação, além de outras informações pertinentes ao usuário, sendo elas:

- Face média
- Eigenfaces
- Fisherfaces
- Reconstrução da face teste
- Gráficos das distâncias

### 4.3 Banco de Imagens

A base de dados escolhida para os testes foi o *IRIS Thermal/Visible Face Database* (UTK-IRIS, 2007), disponibilizado pelo *Imaging, Robotics, and Intelligent Systems Laboratory* (IRIS), que fica localizado na Universidade do Tennessee, em Knoxville. O banco é composto por 31 conjuntos de imagens de  $320 \times 240$  pixels, colhidas de 30 indivíduos, sendo que um deles possui imagens de alguns desses indivíduos utilizando disfarces e máscaras. Cada conjunto contém aproximadamente 11 pares de imagens registradas simultaneamente, uma para o espectro visível e outra no infravermelho, expostas a cada uma das variações abordadas pela base de dados, da seguinte maneira:

- Expressão: Surpresa, alegria e raiva
- Iluminação: Luz esquerda acesa (Lon), luz direita acesa (Ron), duas luzes acesas (2on), duas luzes apagadas (Off) e todas as luzes apagadas (Dark)

Deste modo, cada indivíduo possui uma quantidade entre 176 e 250 imagens, totalizando 4228 imagens registradas por uma câmera que rotacionava ao seu redor em um ângulo de  $180^\circ$ , gerando as mudanças na posição da face, enquanto as demais alterações eram aplicadas. As variações de iluminação presentes na base de dados foram geradas por duas lâmpadas, uma posicionada do lado direito e outra do lado esquerdo do indivíduo, além da luz de teto do cômodo onde foram conduzidas a sessões de fotos.

As figuras (22) e (23) exibem exemplos das variações de iluminação e expressões faciais contidas no banco.



Figura 22 – Iluminação: 2on, Lon, Ron, Off e Dark respectivamente - extraída de (UTK-IRIS, 2007).



Figura 23 – Expressões: Surpresa, alegria e raiva respectivamente - extraída de (UTK-IRIS, 2007).

Cada imagem presente na base de dados possui em seu nome um prefixo onde um V (*Visible*) indica que ela pertence ao espectro visível, e a letra L (*Long Wavelength Infrared*) indica que pertence ao espectro infravermelho. Em seguida, aparece uma série de números que identificam os pares de imagens, de maneira que cada contraparte possui a mesma sequência numérica, facilitando a identificação. Os diretórios do banco de imagens estão dispostos da seguinte maneira:

- Indivíduo 1
  - Expressão
    - \* Expressão 1
      - L-156.bmp
      - L-158.bmp
      - ...
      - V-156.bmp
      - V-158.bmp
      - ...
    - \* Expressão 2
    - \* Expressão 3
  - Iluminação
    - \* 2on
      - L-598.bmp
      - ...
    - \* Dark

- \* Lon
- \* Off
- \* Ron
- Indivíduo 2
- ...

A escolha do [UTK-IRIS \(2007\)](#) se deu pelo fato de que este é o único banco que possui pares de imagens do espectro visível e infravermelho, onde há variação de iluminação, posição e expressão simultaneamente. A maneira que as luzes foram posicionadas, não gera uma diferença notável nas sombras projetadas sobre o rosto, porém as imagens com as duas luzes apagadas e o quarto escuro oferecem variações de iluminação similares às encontradas em aplicações reais desses tipos sistemas. Desta maneira, as imagens podem ser úteis na comparação das taxas de reconhecimento em cenários com 3 tipos de iluminação: boa, média e ruim.

No entanto, os conjuntos de imagens contendo disfarces, como por exemplo óculos, máscaras e perucas, foram ignorados por possuírem poucos exemplares, além de uma pequena quantidade de indivíduos, não apresentando as condições necessárias para a produção de resultados relevantes. Dentro dos diretórios de alguns indivíduos, também existiam conjuntos de imagens nomeados "*glasson*" e "*glassoff*", indicando uma sessão extra de fotografias na qual a pessoa colocou ou removeu o par de óculos que portava. Essas imagens também não foram utilizadas, pois não abrangiam todas as variações de iluminação e expressões originais, o que foi um grande problema do banco ao qual os autores não se atentaram inicialmente. Além disso, pequenas adaptações, como por exemplo a renomeação de algumas imagens que estavam fora dos padrões, precisaram ser realizadas sobre os arquivos originais obtidos da página oficial do [UTK-IRIS \(2007\)](#).

## 4.4 Processos

Essa seção apresenta os procedimentos envolvidos nas etapas do reconhecimento. Além disso, para exemplificar cada processo, são fornecidas as imagens exibidas pelo módulo de saída. Desta maneira, ambos os métodos foram aplicados sobre uma mesma face de teste, a título de demonstração.

### 4.4.1 Eigenfaces

O Eigenfaces consiste dos seguintes passos:

- Carregar as matrizes de treino e teste geradas pelo módulo de leitura.

- Calcular a face média do conjunto de treinamento.
- Normalizar as faces de treino subtraindo a face média.
- Obter os autovalores e autovetores da matriz de covariância dos dados normalizados.
- Escolher os autovetores com maiores autovalores, conhecidos por componentes principais ou *eigenfaces*.
- Selecionar uma imagem de teste.
- Projetar a matriz de treino e a face teste gerando um novo espaço de faces.
- Comparar a distância entre a imagem teste e o conjunto de treinamento dentro desse subespaço, por meio de uma métrica de similaridade.
- Identificar a face de acordo com a classe das  $k$  faces mais próximas.

#### 4.4.1.1 Execução

Para exemplificar o processo, o Eigenfaces foi executado utilizando 2on como treinamento, e Dark como teste. Deste modo, tentando identificar uma imagem escura por meio de um conjunto de treinamento composto por imagens claras. Além disso, os parâmetros selecionados foram:  $k = 1$ , distância euclidiana e 50 componentes. As figuras foram geradas pelo módulo de saída, através da execução do algoritmo sobre a face de número 239 do conjunto de teste.

Primeiramente, é possível visualizar na figura (24) as faces médias geradas para cada tipo de imagem.

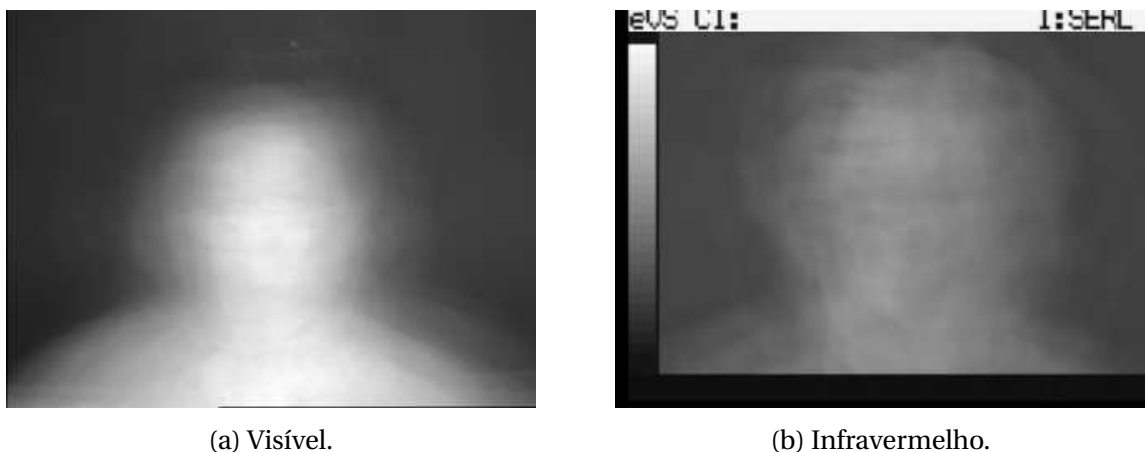


Figura 24 – Face média gerada para o conjunto de treinamento de cada espectro.

Em seguida, nas figuras (25) e (26) estão apresentadas as 20 primeiras *eigenfaces* geradas para cada conjunto de imagens.

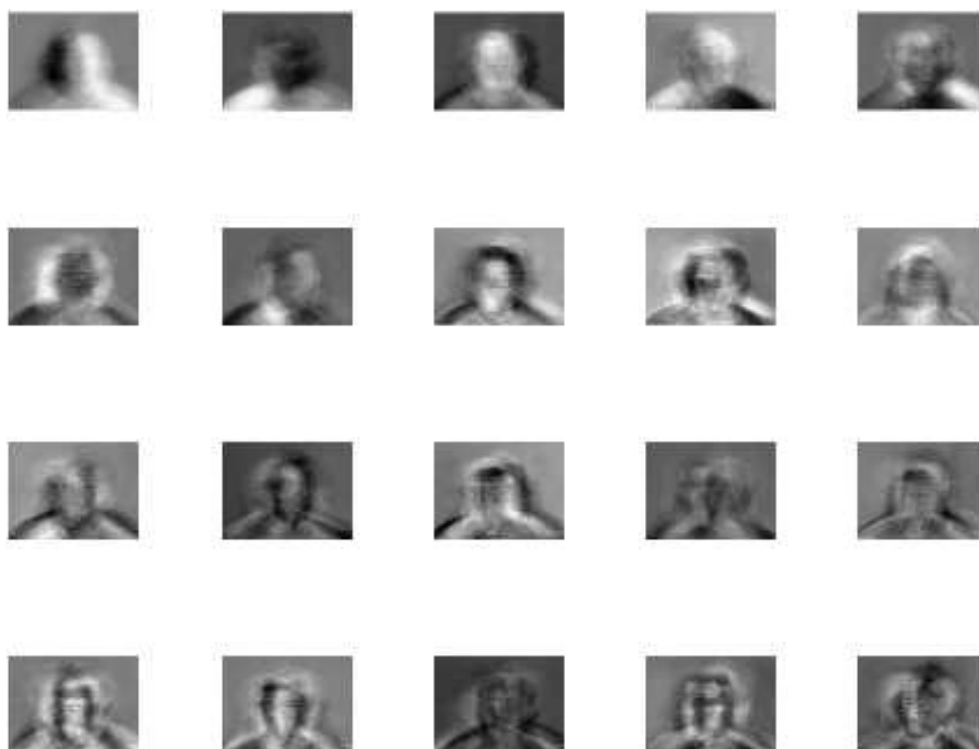


Figura 25 – Conjunto das 20 primeiras *eigenfaces* geradas para as imagens visíveis.



Figura 26 – Conjunto das 20 primeiras *eigenfaces* geradas para as imagens infravermelhas.

Os gráficos (27) e (28) representam as distâncias em relação à face de teste e as faces do conjunto de treino para ambos os tipos de imagem.

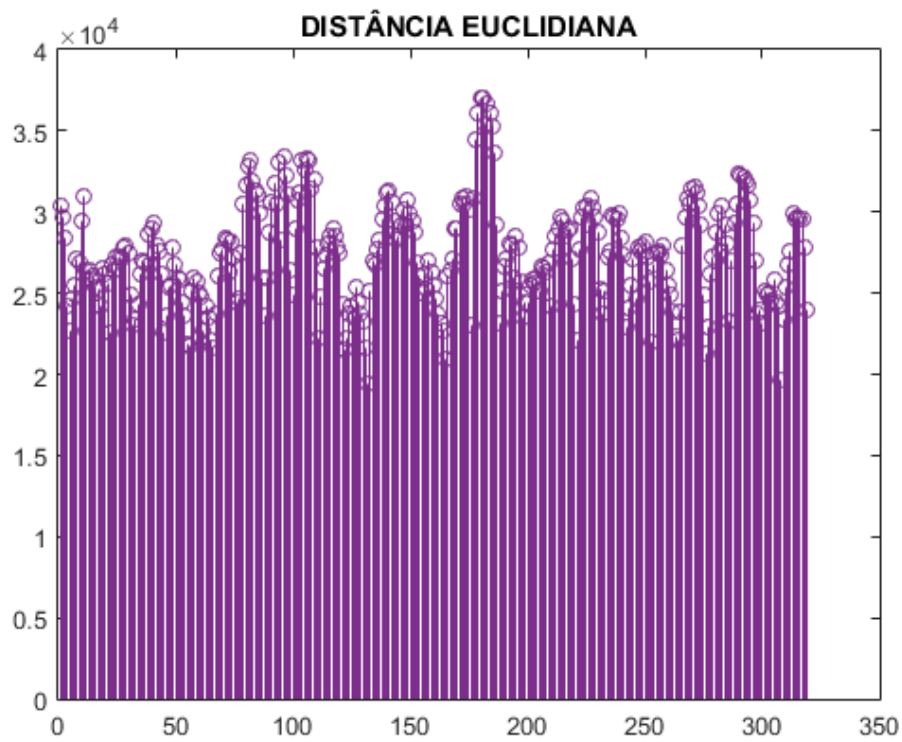


Figura 27 – Distâncias das faces do conjunto de treinamento em relação à face teste para as imagens do espectro visível.

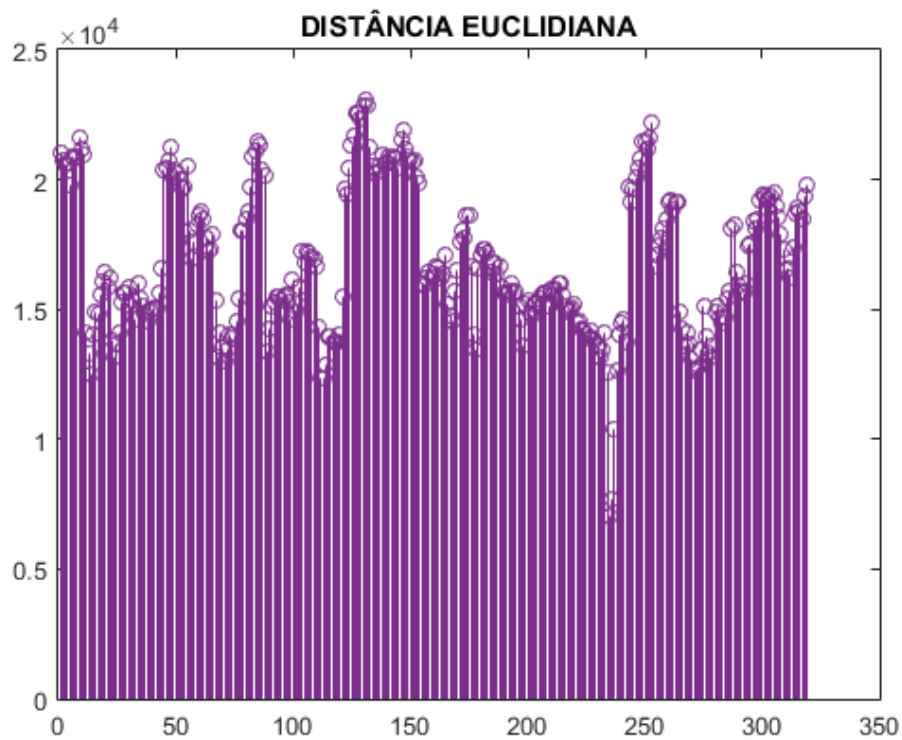


Figura 28 – Distâncias das faces do conjunto de treinamento em relação à face teste para as imagens do espectro infravermelho.

Observando os gráficos, é possível perceber que nos valores próximos ao número 239, que correspondem às imagens do mesmo indivíduo, estão muito mais próximos da imagem teste no espectro infravermelho, do que no espectro visível. Isso ocorre pois essas elas são muito mais robustas a variações de iluminação. Por fim, nas figuras (29) e (30), estão os resultados da classificação da face de teste para cada tipo de imagem.

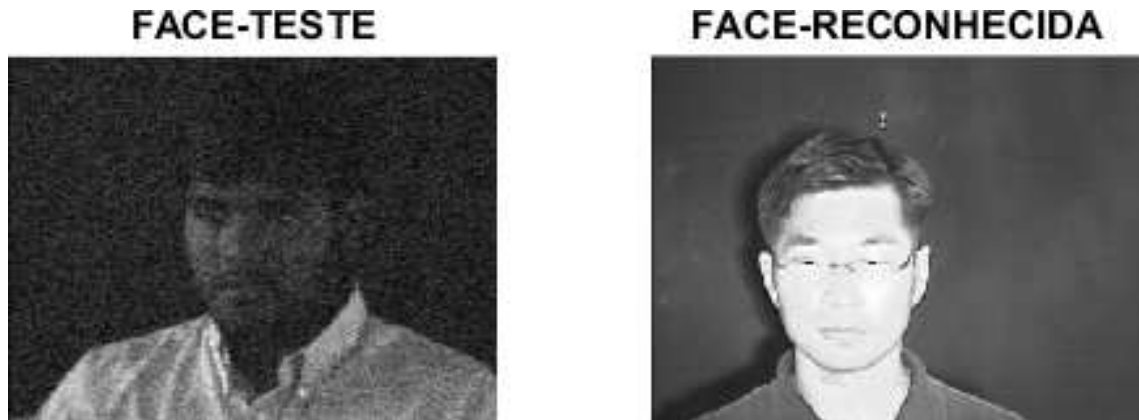


Figura 29 – Resultado para as imagens do espectro visível.



Figura 30 – Resultado para as imagens do espectro infravermelho.

É possível constatar que a imagem foi classificada corretamente nas imagens infravermelhas, porém o mesmo não ocorreu com as imagens do espectro visível.

### 4.4.2 Fisherfaces

Primeiramente, o Fisherfaces obtém o PCA da mesma maneira que no Eigenfaces, para só então calcular o LDA e aplicá-lo. Essa implementação é muito mais eficiente, pois aplicar o LDA diretamente sobre um conjunto de dados tão grande se torna impraticável, sendo necessária a redução de sua dimensionalidade. De maneira simplificada, o algoritmo consiste das seguintes etapas:

- Carregar as matrizes de treino e teste geradas pelo módulo de leitura.
- Executar o PCA sobre o conjunto de treinamento, obtendo assim as componentes principais.
- Projetar a matriz de treinamento original sobre o subespaço gerado, obtendo um novo conjunto de dados com dimensionalidade reduzida.
- Calcular a dispersão intra-classe e entre-classe da nova matriz de treinamento.
- Aplicar o LDA sobre o conjunto de dados, gerando as *fisherfaces*.
- Selecionar uma imagem de teste.
- Projetar novamente a matriz de treino e a face teste gerando um novo espaço de faces.
- Comparar a distância entre a imagem teste e o conjunto de treinamento dentro desse subespaço, por meio de uma métrica de similaridade.
- Identificar a face de acordo com a classe das  $k$  faces mais próximas.

#### 4.4.2.1 Execução

Assim como na demonstração anterior, o Fisherfaces foi executado utilizando 20n como treinamento e a face de número 239 do conjunto Dark como teste. Da mesma maneira, os parâmetros selecionados foram:  $k = 1$ , distância euclidiana e 50 componentes.

Os processos relacionados ao PCA já foram exemplificados, por isso serão mostradas apenas as etapas únicas do Fisherfaces. Deste modo, nas figuras (31) e (32) estão apresentadas as 20 primeiras *fisherfaces* geradas para cada conjunto de imagens.

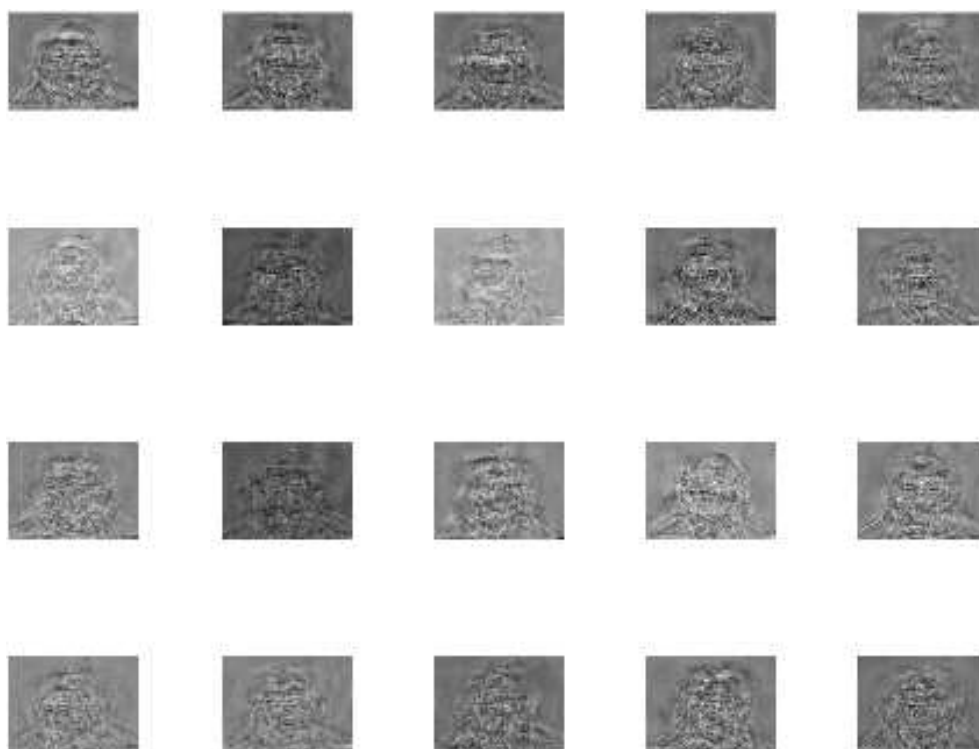


Figura 31 – Conjunto das 20 primeiras *fisherfaces* geradas para as imagens visíveis.

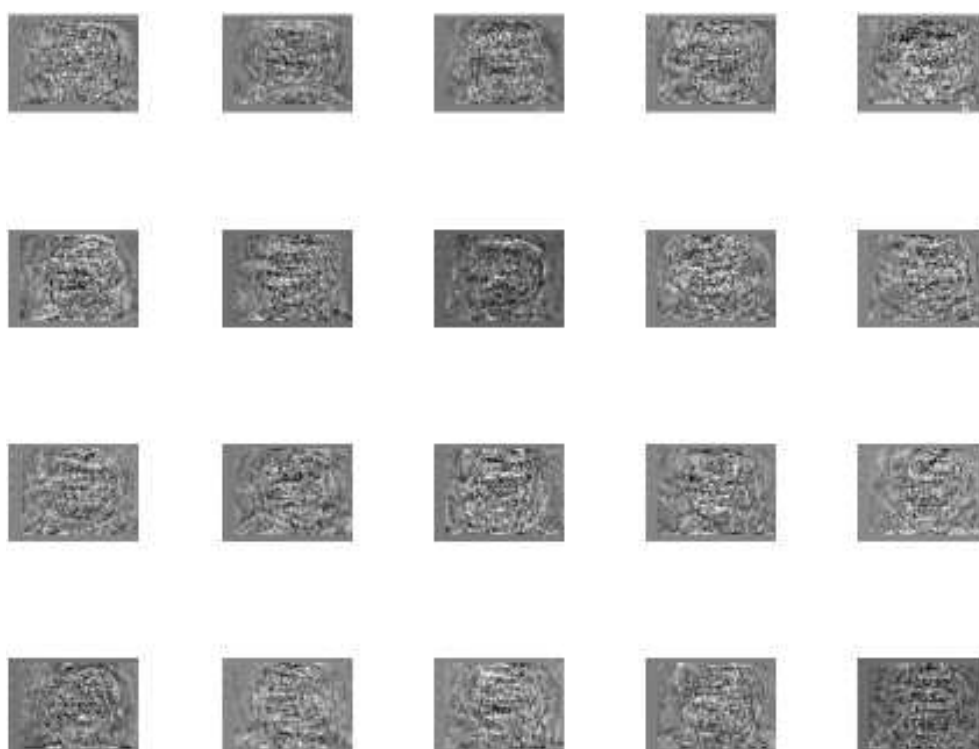


Figura 32 – Conjunto das 20 primeiras *fisherfaces* geradas para as imagens infravermelhas.

Os gráficos (33) e (34) representam as distâncias em relação à face de teste e as faces do conjunto de treino para ambos os tipos de imagem.

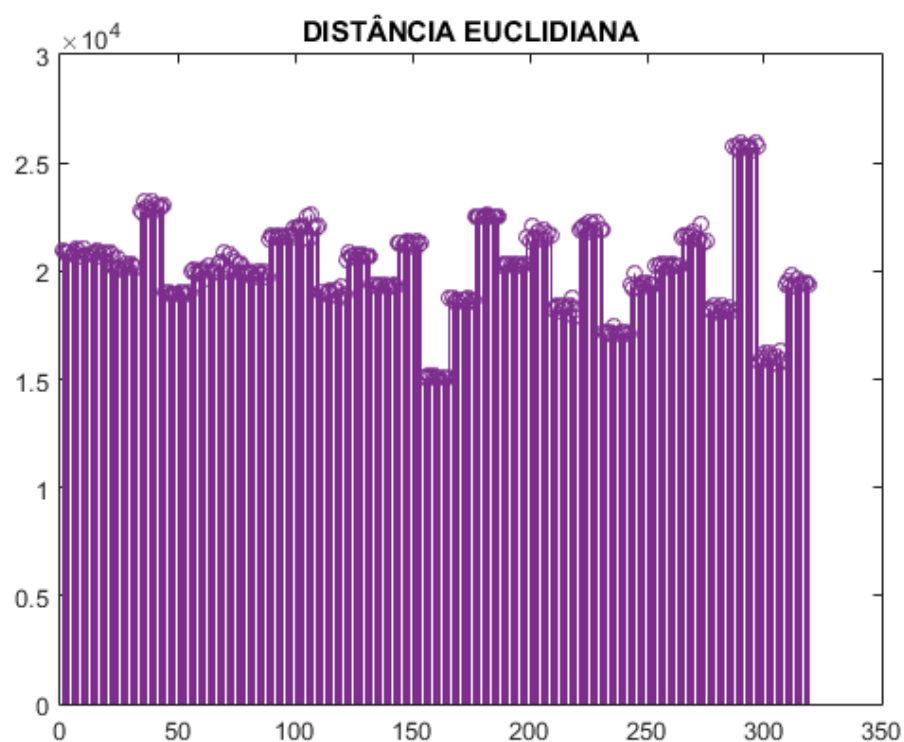


Figura 33 – Distâncias das faces do conjunto de treinamento em relação à face teste para as imagens do espectro visível.

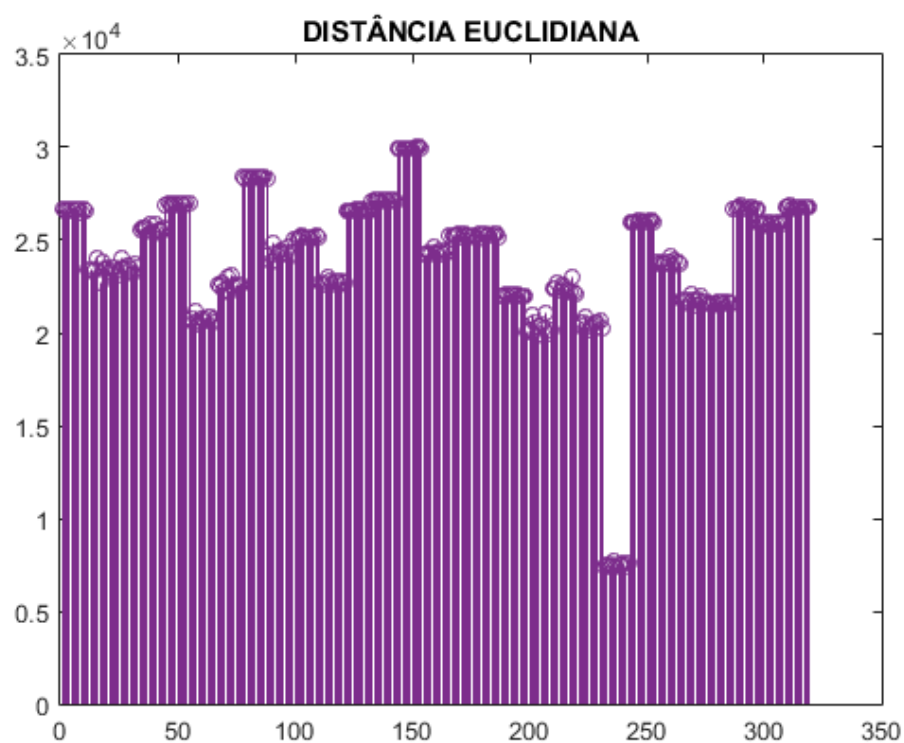


Figura 34 – Distâncias das faces do conjunto de treinamento em relação à face teste para as imagens do espectro infravermelho.

Diferentemente do PCA, fica é visível a forte separação entre as classes e a proximidade dos elementos dentro delas, deste modo o resultado é aprimorado substancialmente no Fisherfaces. Por fim, nas figuras (35) e (36) estão os resultados da classificação da face de teste para cada tipo de imagem.

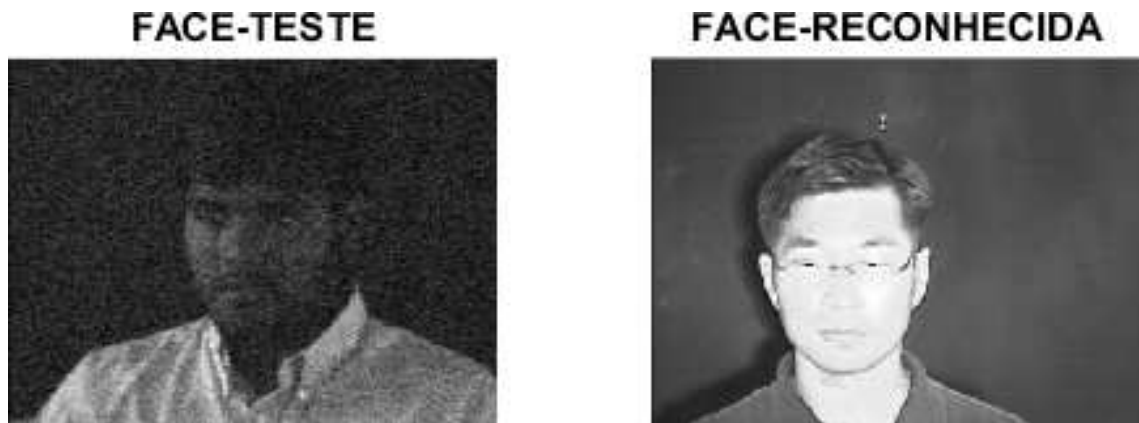


Figura 35 – Resultado para as imagens do espectro visível.



Figura 36 – Resultado para as imagens do espectro infravermelho.

Embora o Fisherfaces possua um desempenho superior ao do Eigenfaces, o algoritmo ainda não foi capaz de reconhecer a face em situações tão adversas, justificando o uso das imagens infravermelhas.

## 4.5 Experimentos

Para medir a degradação das taxas de reconhecimento proporcionada pela variação de iluminação, posição e expressões nas imagens do espectro visível, é preciso variar os conjuntos de teste e treino como propôs Wilder et al. (1997). Deste modo, é necessário treinar o sistema utilizando imagens bem iluminadas e estáticas, como geralmente são apresentadas aos algoritmos, e depois testar o reconhecimento em imagens com iluminação inadequada, posições e expressões variadas, representando situações bastante comuns para aplicações

de um sistema de reconhecimento automatizado. Com esse método, as imagens térmicas correspondentes também podem ser testadas e ter suas taxas de reconhecimento comparadas as do espectro visível. Em seguida basta repetir os testes variando os parâmetros dos classificadores com o objetivo encontrar a melhor configuração possível para cada combinação de imagens.

#### 4.5.1 Metodologia

As imagens escolhidas como padrão de iluminação foram aquelas que possuíam as duas luzes acesas (2on), por possuírem melhor iluminação e maior nitidez em relação às outras imagens da base de dados. Os próximos passos consistem em gerar o conjunto de treinamento a partir delas, e utilizar como teste cada uma das imagens pertencentes aos outros grupos, como mostra a figura (37).

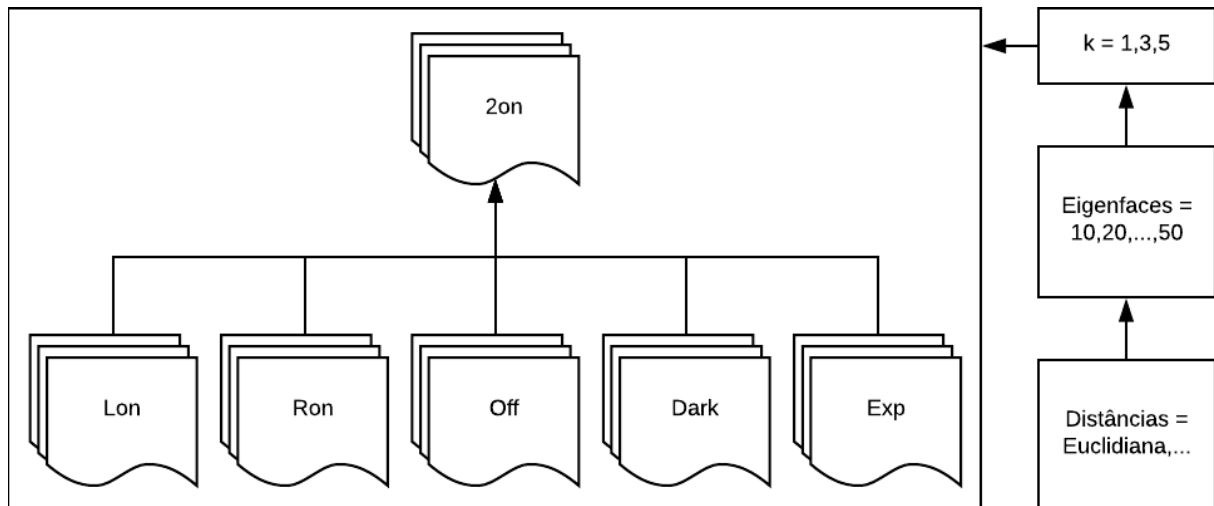


Figura 37 – Representação da metodologia de testes.

Esse processo é executado novamente para cada variação de parâmetros do PCA e do classificador, da seguinte maneira:

1.  $k = 1, 3, 5$
2. Componentes Principais = 10, 20, 30, 40, 50
3. Distâncias = Euclidiana, Manhattan, Chebyshev, Mahalanobis

Após a execução de cada uma dessas variações de parâmetros, os resultados são armazenados para uma análise posterior.

## 5 Resultados

Esse capítulo apresenta os resultados obtidos pela metodologia e testes propostos na seção anterior. Para uma melhor visualização e entendimento, eles foram compilados em tabelas e representados em gráficos.

Dentre as variações de parâmetros propostas, os valores que proporcionaram os melhores resultados foram: 40 componentes principais e  $k = 1$  para a maioria das distâncias, métodos de reconhecimento e tipos de imagem, como demonstrado nas tabelas (12) e (13):

Tabela 12 – Taxa de reconhecimento média para a variação no número de componentes.

| <b>Componentes</b> | <b>Eigenfaces</b> |        | <b>Fisherfaces</b> |        |
|--------------------|-------------------|--------|--------------------|--------|
|                    | VIS               | IR     | VIS                | IR     |
| 10                 | 0.4887            | 0.4995 | 0.6332             | 0.6236 |
| 20                 | 0.5727            | 0.6328 | 0.7011             | 0.6577 |
| 30                 | 0.5989            | 0.6468 | 0.7308             | 0.6793 |
| 40                 | 0.6067            | 0.6535 | 0.7308             | 0.6793 |
| 50                 | 0.6147            | 0.6517 | 0.7308             | 0.6793 |

Analisando os resultados, é possível perceber que o número de componentes necessárias para obter a taxa de reconhecimento máxima no Fisherfaces é bem menor para ambos os espectros quando comparado ao Eigenfaces, que por sua vez possui uma taxa de reconhecimento que converge por volta 40 componentes para as imagens térmicas, e 50 para as do espectro visível. As diferenças nas quantidades de componentes para ambos os espectros existem devido ao fato de que as imagens do espectro visível possuem uma maior variância, e portanto necessitam de mais informação para serem representadas. Já a diferença entre os algoritmos é causada por uma segunda redução de dimensionalidade executada pelo LDA, gerando um conjunto de *features* bem menor e mais preciso, como é possível observar nos resultados.

Tabela 13 – Taxa de reconhecimento média para a variação do valor de  $k$ .

| <b>k</b> | <b>Eigenfaces</b> |        | <b>Fisherfaces</b> |        |
|----------|-------------------|--------|--------------------|--------|
|          | VIS               | IR     | VIS                | IR     |
| 1        | 0.6408            | 0.6938 | 0.7302             | 0.6785 |
| 3        | 0.6209            | 0.6575 | 0.7310             | 0.6793 |
| 5        | 0.5823            | 0.6038 | 0.7313             | 0.6801 |

Para os valores de  $k$ , é possível constatar que a melhor opção para o Eigenfaces é o valor de  $k = 1$ , e para o Fisherfaces o aumento no  $k$  gera uma melhora quase que imperceptível. Isso ocorre porque o LDA separa as classes de uma maneira mais eficiente, fazendo com

que os membros da mesma classe fiquem mais próximos uns dos outros e ao mesmo tempo mais distantes das demais classes, de maneira que não seja necessário olhar muito além dos vizinhos mais imediatos. Já para o PCA, o foco é a redução da dimensionalidade, e muitas vezes isso resulta em uma projeção que não gera uma boa separação entre as classes.

Além disso, os resultados indicam uma taxa média de reconhecimento maior para as imagens infravermelhas quando o Eigenfaces é utilizado, já para o Fisherfaces, as imagens do espectro visível se saíram melhor. Isso ocorre porque no Eigenfaces, não existe uma grande vantagem das imagens visíveis em nenhum dos casos de teste, mas para o cenário onde o conjunto de entrada é composto pelas imagens escuras (Dark), as imagens térmicas obtiveram taxas de reconhecimento que chegam a ser 7 vezes maiores em algumas das distâncias, o que gera um desbalanceamento da média a favor delas. No entanto, quando o Fisherfaces é aplicado, a melhora na taxa de reconhecimento é muito maior nas imagens do espectro visível, o que faz com que a desvantagem seja superada.

Finalmente, resta analisar os resultados em relação às distâncias, que podem ser observados nos gráficos a seguir:

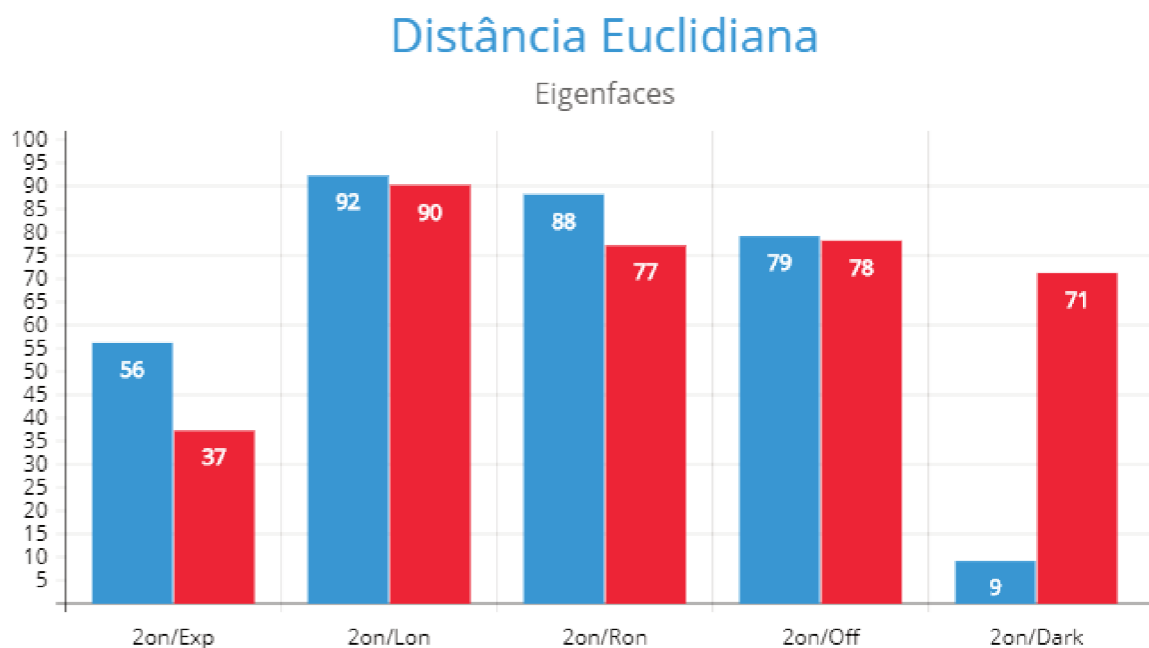


Figura 38 – Taxas de reconhecimento: Eigenfaces e Distância Euclidiana.

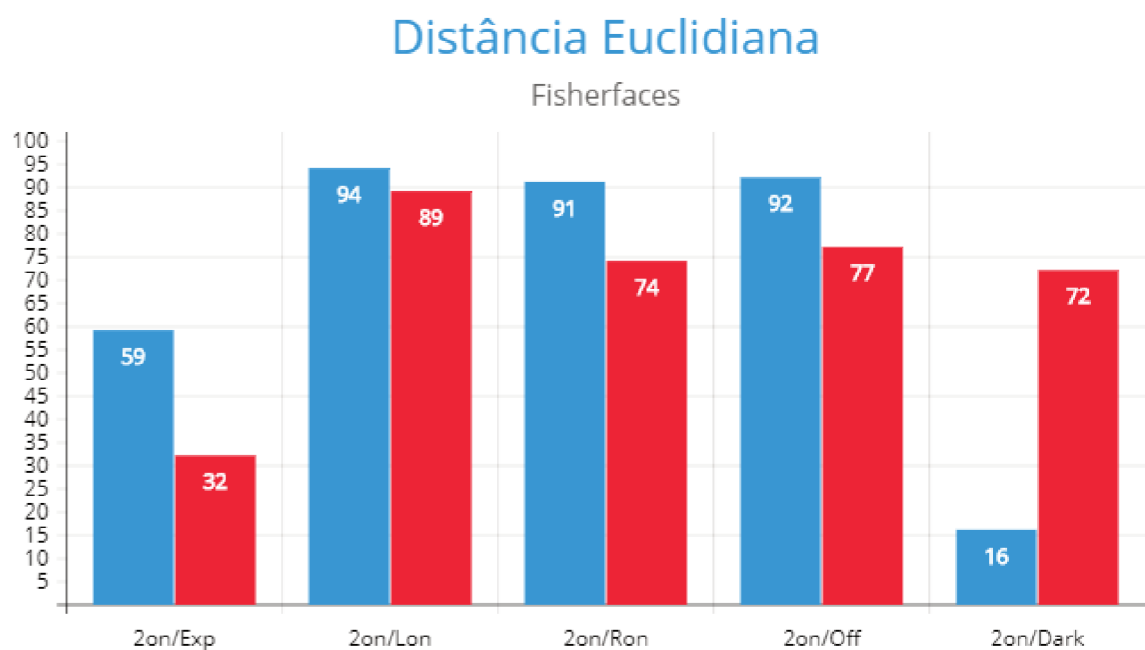


Figura 39 – Taxas de reconhecimento: Fisherfaces e Distância Euclidiana.

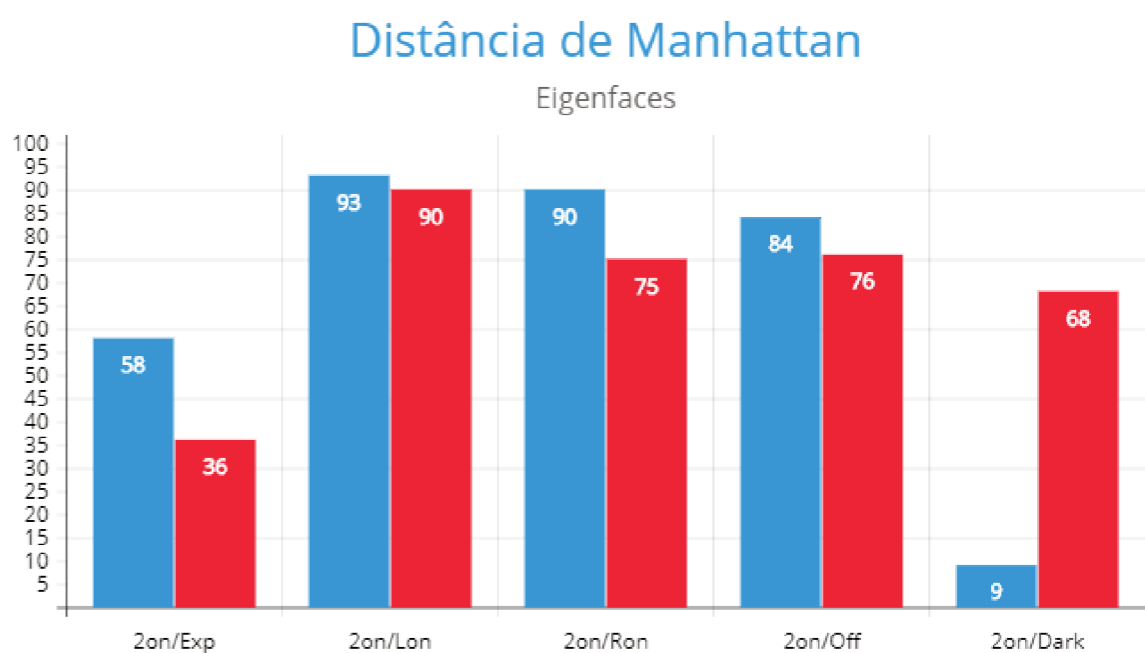


Figura 40 – Taxas de reconhecimento: Eigenfaces e Distância de Manhattan.

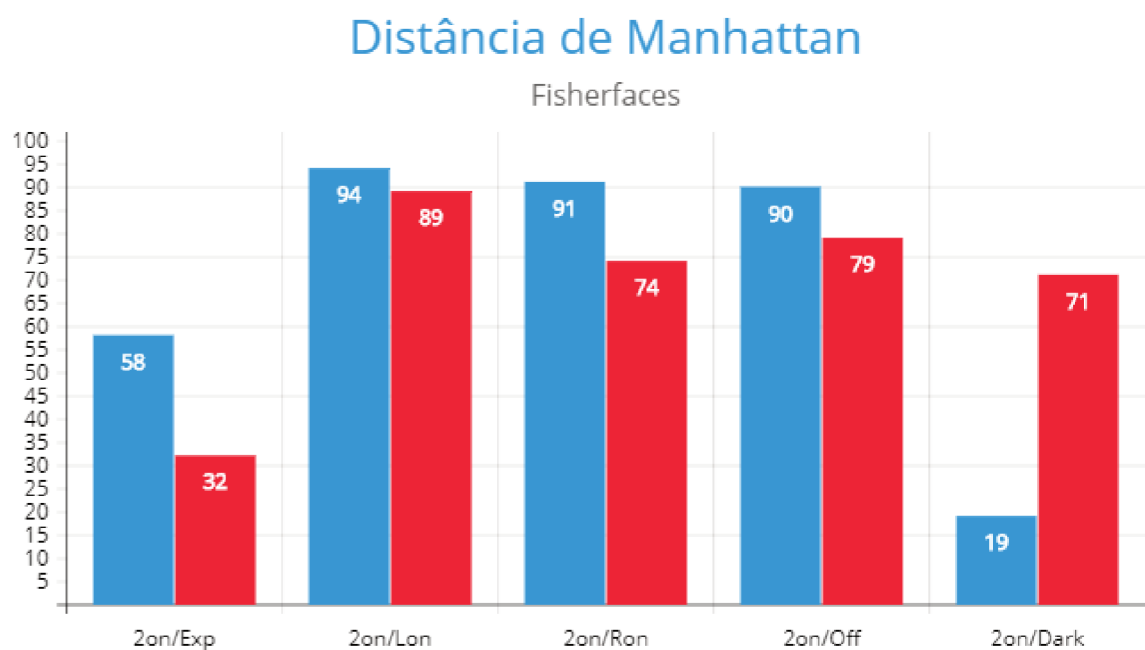


Figura 41 – Taxas de reconhecimento: Fisherfaces e Distância de Manhattan.

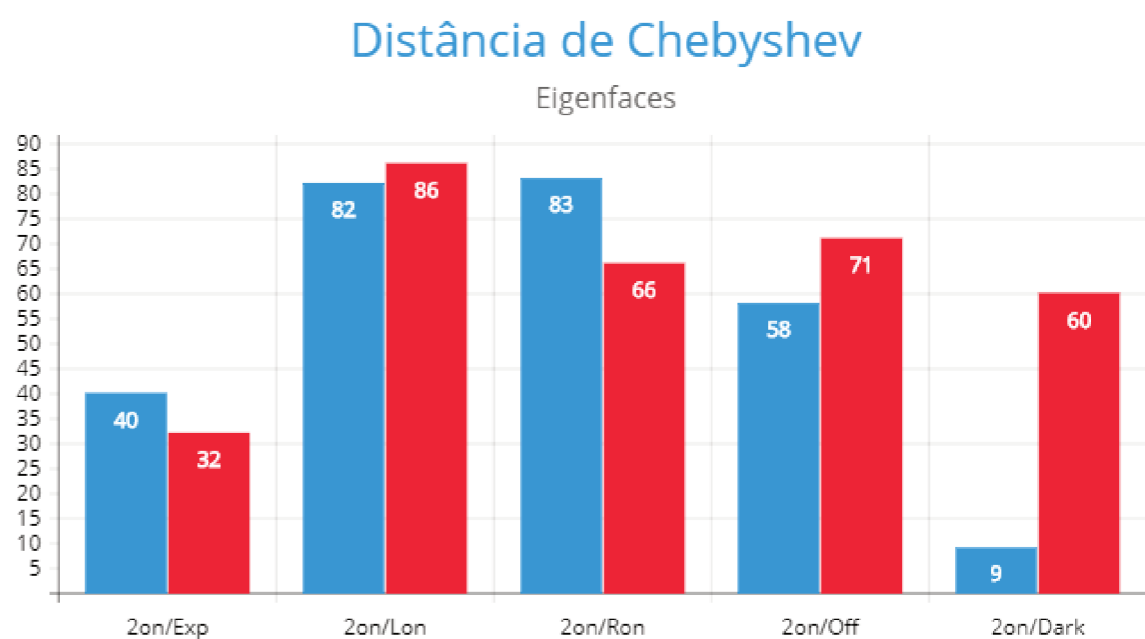


Figura 42 – Taxas de reconhecimento: Eigenfaces e Distância de Chebyshev.

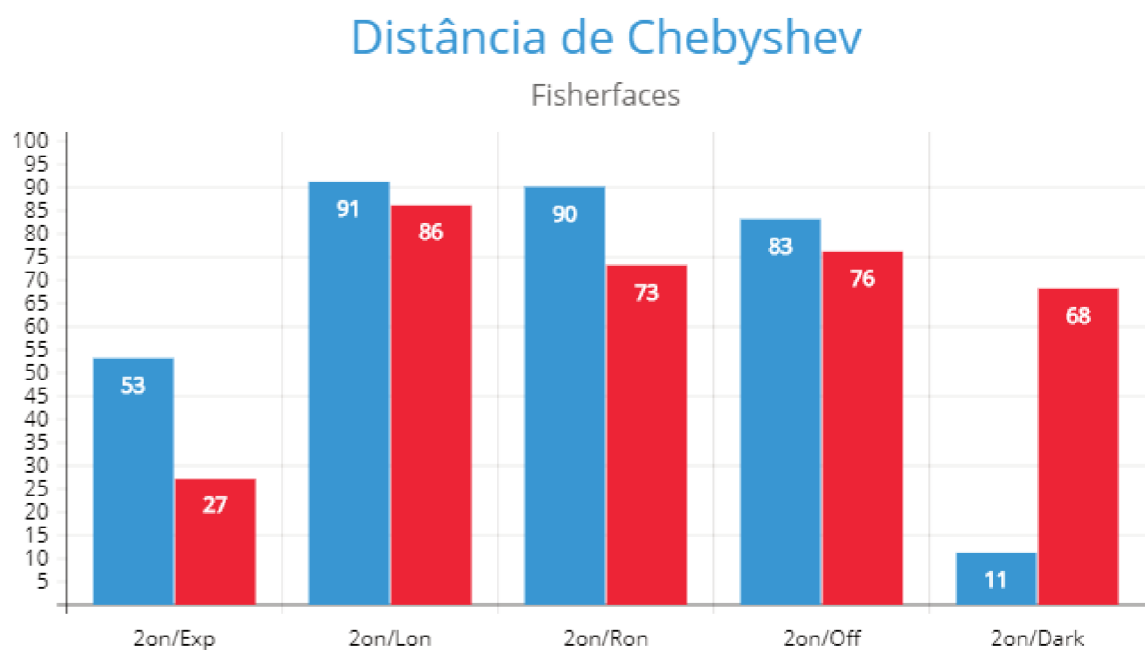


Figura 43 – Taxas de reconhecimento: Fisherfaces e Distância de Chebyshev.

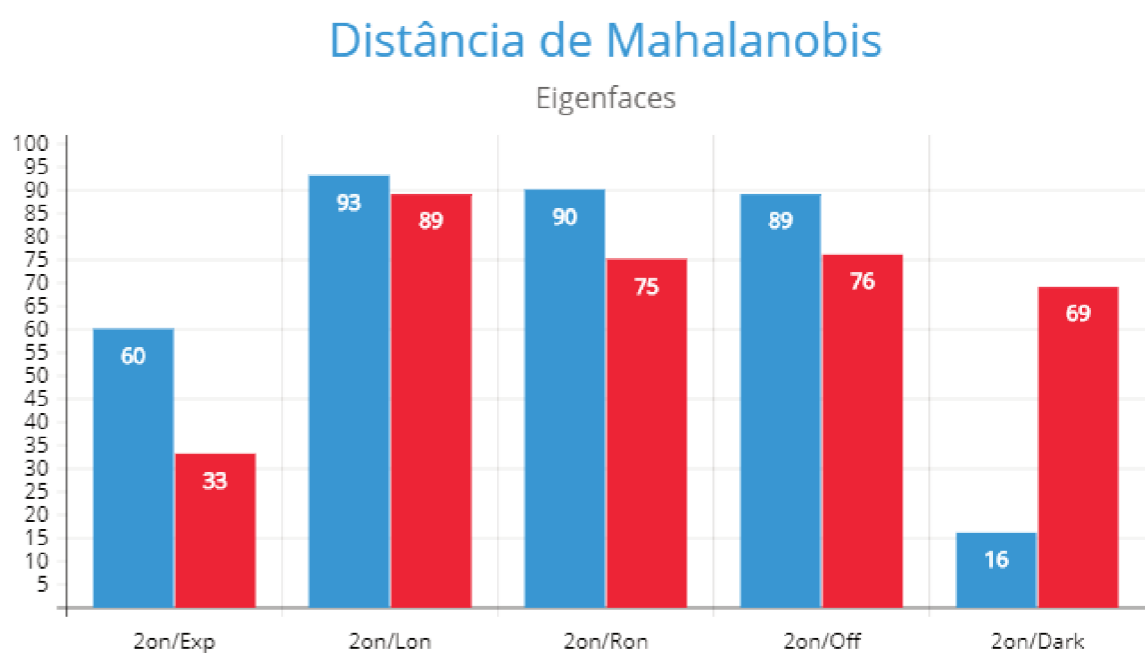


Figura 44 – Taxas de reconhecimento: Eigenfaces e Distância de Mahalanobis.

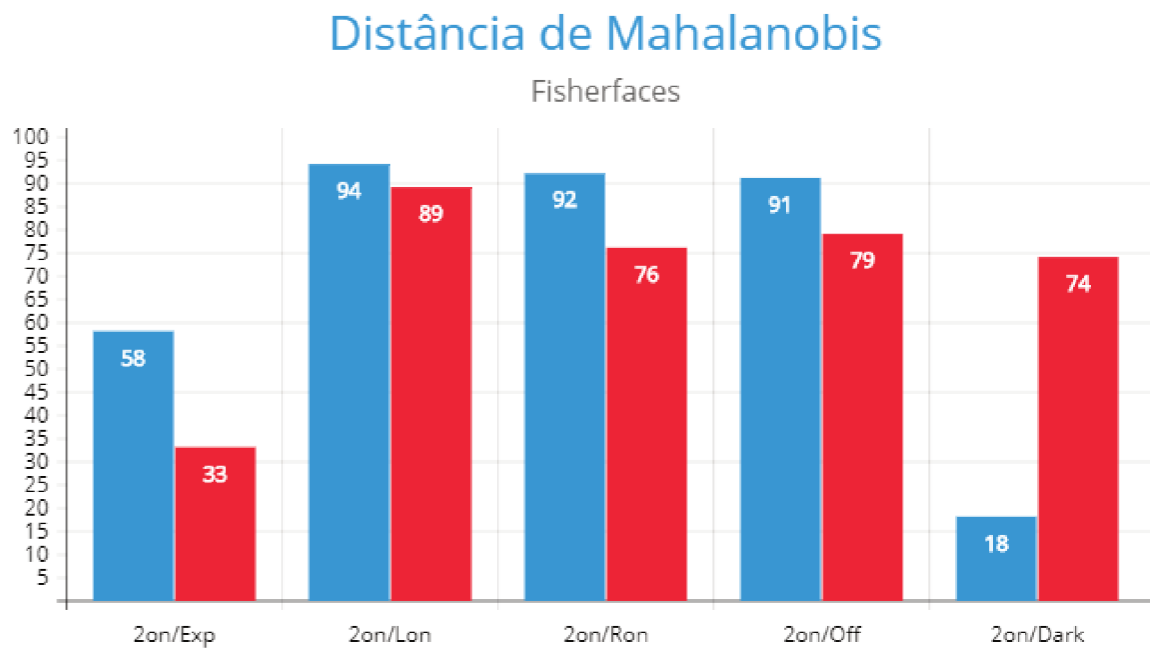


Figura 45 – Taxas de reconhecimento: Fisherfaces e Distância de Mahalanobis.

Os resultados indicam que não existe uma métrica que se destaca para todos os cenários. No entanto, é possível constatar no caso das imagens do espectro visível, tanto para o Eigenfaces quanto para o Fisherfaces, que a distância de Mahalanobis proporcionou a melhor performance, gerando uma média de acertos de 67,6% e 70,4% respectivamente. Já no caso das imagens térmicas, a distância euclidiana produziu uma melhor média de 70,6% para o Eigenfaces e o valor máximo obtido no Fisherfaces foi de 70,2%, com a distância de Mahalanobis.

## 6 Conclusão

Da mesma maneira que nos trabalhos de [Cutler \(1996\)](#), [Wilder et al. \(1997\)](#), [Socolinsky et al. \(2001\)](#), as imagens infravermelhas obtiveram uma menor taxa de reconhecimento frente as do espectro visível quando ambos os conjuntos, de treino e teste, possuíam pouca variação de luminosidade entre si, como foi o caso das imagens Lon e Ron. No entanto, a vantagem do espectro infravermelho se torna bastante evidente a medida que essa variação aumenta, como no caso das imagens do conjunto Dark, quando comparadas ao 2on, onde as imagens térmicas obtiveram uma taxa de reconhecimento até 7 vezes maior em alguns dos testes, confirmando a hipótese da sua superioridade nesse tipo de cenário.

Já em casos onde existe uma variação média de iluminação, como ocorre nas imagens do conjunto Off, as taxas de reconhecimento para ambos os espectros foram bem próximas, hora com vantagem para um ou para outro. Além disso, as imagens infravermelhas não obtiveram uma melhor taxa de reconhecimento nas variações de posição e expressões faciais como haviam sugerido [Prokoski, Riedel e Coffin \(1992\)](#), entretanto os resultados ainda são inconclusivos para as expressões, já que a quantidade presente no banco de imagens é muito pequena.

Além disso, nas aplicações reais dos sistemas de reconhecimento facial, as imagens infravermelhas podem ser bastante úteis no processo de detecção da face, o que pode interferir nos resultados finais em um processo completo de reconhecimento, porém esse fator não foi considerado neste trabalho.

Os resultados obtidos demonstram que as imagens infravermelhas possuem suas vantagens e desvantagens de acordo com cada cenário, e que ainda não está clara a fronteira entre as variações de iluminação onde cada tipo de imagem pode se tornar mais eficiente. Para determinar isso, seria necessária a utilização de imagens iluminadas com uma fonte de luz direcionada, o que causaria a projeção de sombras mais consistentes e que não podem ser encontradas nos bancos de imagens disponíveis na literatura.

Apesar das dificuldades relacionadas à falta de dados para realização dos testes, as taxas de reconhecimento obtidas sugerem que talvez as imagens do espectro visível ainda sejam uma boa opção para cenários onde há cooperação do indivíduo e iluminação controlada, mas para todos os outros casos a termografia infravermelha pode oferecer um melhor desempenho de maneira geral. Existem também boas razões para considerar a possibilidade de que a melhor saída para os problemas do reconhecimento facial seja a fusão de ambos os tipos de imagens, aproveitando o melhor de cada espectro e aprimorando as taxas de reconhecimento, mesmo nas condições mais adversas para um sistema de reconhecimento facial automatizado.

## 6.1 Trabalhos Futuros

Ainda é necessário investigar as taxas de reconhecimentos em imagens infravermelhas, pois as bases de dados disponíveis na literatura não possuem as variações necessárias para obtenção de resultados conclusivos. Desta maneira, um trabalho muito relevante é a criação de um banco de imagens multimodal e multi-iluminado, ou seja, contendo imagens do espectro visível e infravermelho, expostas a variações de iluminação. No entanto, essas alterações na iluminação precisam ser produzidas por meio de luzes direcionadas em diferentes ângulos, a fim de projetar sombras consistentes e que simulem de uma maneira mais próxima as sombras encontradas nas aplicações comerciais desses tipos de sistemas. A fusão de imagens de ambos os espectros também é um assunto bastante relevante a ser estudado, e pode oferecer alternativas promissoras para aplicações práticas, aproveitando o melhor de cada abordagem.

# Referências

- ABDI, H.; WILLIAMS, L. J. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, v. 2, n. 4, p. 433–459, 2010. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/wics.101>>. Citado 2 vezes nas páginas 41 e 42.
- AKHLOUFI, M.; BENDADA, A.; BATSALE, J.-C. State of the art in infrared face recognition. *Qirt Journal - QIRT J*, v. 5, p. 3–26, 06 2008. Citado 3 vezes nas páginas 7, 21 e 22.
- ANAJ, A. A. *Example of k-NN classification*. 2007. Disponível em: <[https://en.wikipedia.org/wiki/K-nearest\\_neighbors\\_algorithm#/media/File:KnnClassification.svg](https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm#/media/File:KnnClassification.svg)>. Citado 2 vezes nas páginas 7 e 30.
- BARTLETT, M. S. et al. Automatic recognition of facial actions in spontaneous expressions. v. 1, 09 2006. Citado na página 15.
- BELHUMEUR, P. N.; HESPANHA, J. P.; KRIEGMAN, D. J. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.*, v. 19, p. 711–720, 1997. Citado 2 vezes nas páginas 49 e 57.
- BISHOP, C. M. *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer-Verlag, 2011. Citado na página 41.
- CAMBRIDGE, A. L. *The Database of Faces*. 2002. Disponível em: <<https://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>>. Citado 5 vezes nas páginas 7, 43, 44, 46 e 51.
- CASTANEDO, C. I. Quantitative subsurface defect evaluation by pulsed phase thermography: depth retrieval with the phase [ressource électronique] /. 01 2005. Citado 3 vezes nas páginas 9, 20 e 21.
- CHELLAPPA, R.; SINHA, P.; PHILLIPS, P. J. Face recognition by computers and humans. *Computer*, v. 43, 2010. Citado na página 24.
- CHEN, X.; FLYNN, P. J.; BOWYER, K. Visible-light and infrared face recognition. *Computer Vision and Image Understanding - CVIU*, 01 2003. Citado na página 55.
- CMU-PIE. *The CMU Multi-PIE Face Database*. 2008. Disponível em: <<http://www.cs.cmu.edu/afs/cs/project/PIE/MultiPie/Multi-Pie/Home.html>>. Citado 2 vezes nas páginas 7 e 47.
- CUTLER, R. *Face Recognition Using Infrared Images and Eigenfaces*. College Park, MD, USA, 1996. Citado 3 vezes nas páginas 53, 54 e 78.
- DELAC, K.; GRGIC, M. A survey of biometric recognition methods. *Proceedings. Elmar-2004. 46th International Symposium on Electronics in Marine*, p. 184–193, 2004. Citado 3 vezes nas páginas 23, 24 e 26.
- DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern Classification (2Nd Edition)*. New York, NY, USA: Wiley-Interscience, 2000. Citado 4 vezes nas páginas 27, 29, 31 e 48.

- EL-SHOROUEK, H. R. Face recognition using principle component analysis with different distance classifiers. *International Journal of Computer Science and Network Security*, v. 11, 2011. Citado 2 vezes nas páginas 31 e 32.
- GHIASS, R. S. et al. Infrared face recognition: a comprehensive review of methodologies and databases. *Pattern Recognition*, v. 47, p. 2807–2824, 2014. Citado 4 vezes nas páginas 7, 15, 21 e 22.
- GHIASS, R. S. et al. Infrared face recognition: A literature review. p. 1–10, Aug 2013. ISSN 2161-4393. Citado 2 vezes nas páginas 15 e 16.
- GONZALEZ, R. C.; WOODS, R. E. *Digital Image Processing (3rd Edition)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006. ISBN 013168728X. Citado 3 vezes nas páginas 32, 33 e 34.
- HASSABALLAH, M.; ALI, A.; ALSHAZLY, H. *Image Features Detection, Description and Matching*. Switzerland: Springer International Publishing, 2016. v. 630. 11-45 p. ISBN 978-3-319-28852-9. Citado 3 vezes nas páginas 7, 34 e 35.
- HO, G. *Linear Discriminant Analysis for Starters*. 2017. Disponível em: <<https://eigenfoo.xyz/lda/>>. Citado 2 vezes nas páginas 7 e 48.
- JAIN, A. K.; DUIN, R. P. W.; MAO, J. Statistical pattern recognition: A review. *IEEE Trans. Pattern Anal. Mach. Intell.*, v. 22, p. 4–37, 2000. Citado 4 vezes nas páginas 9, 27, 28 e 29.
- JAIN, A. K.; ROSS, A.; PRABHAKAR, S. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, v. 14, p. 4–20, 2004. Citado 6 vezes nas páginas 7, 9, 24, 25, 26 e 27.
- JOHANSSON, J. *Interest Point Detectors and Descriptors for IR Images: An Evaluation of Common Detectors and Descriptors on IR images*. Tese (Doutorado) — School of Computer Science and Communication (CSC), 2015. Citado 2 vezes nas páginas 7 e 36.
- KIRBY, M.; SIROVICH, L. Application of the karhunen-loeve procedure for the characterization of human faces. *IEEE Trans. Pattern Anal. Mach. Intell.*, v. 12, p. 103–108, 1990. Citado na página 42.
- KOCHANNEK, E. *Are We in the Golden Age of Thermal Imaging?* 2015. Disponível em: <<https://www.irinfo.org/10-01-2015-kochanek/>>. Citado na página 16.
- KONG, S. G. et al. Recent advances in visual and infrared face recognition - a review. *Computer Vision and Image Understanding*, v. 97, p. 103–135, 2005. Citado 2 vezes nas páginas 38 e 40.
- KOTSIANTIS, S. B.; ZAHARAKIS, I. D.; PINTELAS, P. E. Machine learning: a review of classification and combining techniques. *Artificial Intelligence Review*, v. 26, p. 159–190, 2006. Citado 2 vezes nas páginas 30 e 31.
- LANITIS, A. Facial biometric templates and aging: Problems and challenges for artificial intelligence. 2009. Citado na página 15.
- LI, S. Z. et al. Illumination invariant face recognition using near-infrared images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 29, n. 4, p. 627–639, April 2007. ISSN 0162-8828. Citado 2 vezes nas páginas 7 e 16.

- MATHWORKS. *MATLAB*. 2018. Disponível em: <<https://www.mathworks.com/products/matlab.html>>. Citado na página 57.
- OLSZEWSKA, J. I. Automated face recognition: Challenges and solutions. In: RAMAKRISHNAN, S. (Ed.). *Pattern Recognition*. Rijeka: InTech, 2016. cap. 4. Disponível em: <<https://doi.org/10.5772/66013>>. Citado na página 15.
- PERSSON, A. B.; BUSCHMANN, I. Vascular growth in health and disease. v. 4, p. 14, 08 2011. Citado na página 16.
- PROKOSKI, F. J.; RIEDEL, R. B.; COFFIN, J. S. Identification of individuals by means of facial thermography. p. 120–125, Oct 1992. Citado 5 vezes nas páginas 8, 16, 52, 53 e 78.
- QUARK67. *Sistema de cores aditivas RGB*. 2006. Disponível em: <<https://pt.wikipedia.org/wiki/RGB#/media/File:Synthese2B.svg>>. Citado 2 vezes nas páginas 7 e 33.
- RONAN, G. P. *Electromagnetic spectrum with visible light highlighted*. 2013. Disponível em: <[https://en.wikipedia.org/wiki/Electromagnetic\\_radiation#/media/File:EM\\_spectrumrevised.png](https://en.wikipedia.org/wiki/Electromagnetic_radiation#/media/File:EM_spectrumrevised.png)>. Citado 2 vezes nas páginas 7 e 20.
- SHAPIRO, L. G.; STOCKMAN, G. C. *Computer Vision*. London, England: Pearson, 2001. Citado na página 34.
- SHARIF, M. et al. Accessorize to a crime: Real and stealthy attacks on state-of-the-art face recognition. In: *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. New York, NY, USA: ACM, 2016. (CCS '16), p. 1528–1540. ISBN 978-1-4503-4139-4. Disponível em: <<http://doi.acm.org/10.1145/2976749.2978392>>. Citado na página 15.
- SIROVICH, L.; KIRBY, M. Low-dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America. A, Optics and image science*, v. 4 3, p. 519–24, 1987. Citado na página 42.
- SOCOLINSKY, D. A.; SALGIAN, A. A comparative analysis of face recognition performance with visible and thermal infrared imagery. 2002. Citado 2 vezes nas páginas 54 e 55.
- SOCOLINSKY, D. A.; SALGIAN, A.; NEUHEISEL, J. D. Face recognition with visible and thermal infrared imagery. *Computer Vision and Image Understanding*, v. 91, p. 72–114, 2003. Citado na página 54.
- SOCOLINSKY, D. A. et al. Illumination invariant face recognition using thermal infrared imagery. 2001. Citado 4 vezes nas páginas 9, 54, 55 e 78.
- SPRUYT, V. *Feature extraction using PCA*. 2014. Disponível em: <<http://www.visiondummy.com/2014/05/feature-extraction-using-pca/>>. Citado 2 vezes nas páginas 7 e 42.
- SRIVASTAVA, H. A comparison based study on biometrics for human recognition. 2013. Citado na página 23.
- SZELISKI, R. *Computer Vision: Algorithms and Applications*. New York, NY, USA: Springer-Verlag, 2010. Citado 2 vezes nas páginas 33 e 34.
- TAN, X. et al. Face recognition from a single image per person: A survey. *Pattern Recognition*, v. 39, p. 1725–1745, 2006. Citado 3 vezes nas páginas 9, 39 e 40.

- THEODORIDIS, S.; KOUTROUMBAS, K. *Pattern Recognition, Fourth Edition*. 4th. ed. Orlando, FL, USA: Academic Press, Inc., 2008. ISBN 1597492728, 9781597492720. Citado na página 27.
- TURK, M. *Recons*. 2008. Disponível em: <<http://www.scholarpedia.org/article/File:Recons.jpg>>. Citado 2 vezes nas páginas 7 e 47.
- TURK, M. A.; PENTLAND, A. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, v. 3, p. 71–86, 1991. Citado 2 vezes nas páginas 43 e 57.
- UTK-IRIS. *Thermal/Visible Face Database*. 2007. Disponível em: <<http://vcip1-okstate.org/pbvs/bench/Data/02/download.html>>. Citado 4 vezes nas páginas 8, 60, 61 e 62.
- WATANABE, S. *Pattern Recognition: Human and Mechanical*. New York, NY, USA: John Wiley & Sons, Inc., 1985. ISBN 0-471-80815-6. Citado na página 27.
- WEBB, K. D. C. A. R. *Statistical Pattern Recognition*. New York, NY, USA: John Wiley & Sons, Inc., 2011. Citado na página 28.
- WILDER, J. et al. Comparison of view-based face recognition algorithms on visible and infrared imagery. *Proceedings of SPIE - The International Society for Optical Engineering*, 02 1997. Citado 5 vezes nas páginas 9, 53, 54, 70 e 78.
- YANG, M.-H.; KRIEGMAN, D. J.; AHUJA, N. Detecting faces in images: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.*, v. 24, p. 34–58, 2002. Citado 3 vezes nas páginas 9, 38 e 39.
- ZHAO, W. et al. Face recognition: A literature survey. *ACM Comput. Surv.*, v. 35, p. 399–458, 2003. Citado 6 vezes nas páginas 7, 9, 37, 39, 40 e 41.