



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
Campus de São José do Rio Preto

Eduardo Machado Silva

*Padrões Mapeados Localmente em Multiescala
Aplicados ao Reconhecimento de Faces*

Dissertação de Mestrado
Pós-Graduação em Matemática

São José do Rio Preto

2018

Eduardo Machado Silva

*Padrões Mapeados Localmente em Multiescala
Aplicados ao Reconhecimento de Faces*

Dissertação apresentada para obtenção do título de Mestre em Matemática, área de Modelagem Matemática, junto ao Programa de Pós Graduação em Matemática do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

Financiadora: Capes.

Orientador: Prof. Dr. Maurílio Boaventura
Co-Orientador: Prof(a). Dr(a). Inês Aparecida Gasparotto Boaventura

São José do Rio Preto
2018

Silva, Eduardo Machado.

Padrões mapeados localmente em multiescala aplicados ao reconhecimento de faces / Eduardo Machado Silva. -- São José do Rio Preto, 2018

83 f. : il.

Orientador: Maurlio Boaventura

Coorientador: Inês Aparecida Gasparotto Boaventura

Dissertação (mestrado) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Instituto de Biociências, Letras e Ciências Exatas

1. Matemática. 2. Reconhecimento facial (Computação) 3. Reconhecimento de padrões. 4. Biometria. 5. Algoritmos genéticos. I. Universidade Estadual Paulista "Júlio de Mesquita Filho". Instituto de Biociências, Letras e Ciências Exatas. II. Título.

CDU – 518.72

Ficha catalográfica elaborada pela Biblioteca do IBILCE
UNESP - Câmpus de São José do Rio Preto

Eduardo Machado Silva

Padrões Mapeados Localmente em Multiescala Aplicados ao Reconhecimento de Faces

Dissertação apresentada para obtenção do título de Mestre em Matemática, área de Modelagem Matemática, junto ao Programa de Pós Graduação em Matemática do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

BANCA EXAMINADORA

Prof. Dr. Maurílio Boaventura
Orientador
UNESP - São José do Rio Preto

Prof. Dr. Silvio Alexandre de Araujo
UNESP - São José do Rio Preto

Prof. Dr. Aylton Pagamisse
UNESP - Presidente Prudente

São José do Rio Preto, 06 de abril de 2018

À minha avó, Afra *Dedico*.

Agradecimentos

Agradeço imensamente a minha família pelo apoio e confiança durante a minha trajetória para ingresso e conclusão do mestrado.

Aos amigos Eder, Luan e Luiz e a todos os outros que através de pequenas maneiras contribuíram para que eu realizasse esta pesquisa.

Um agradecimento especial a Cheienne por todo companheirismo durante todos os dias que dediquei a esta pesquisa.

Agradeço a meu orientador Prof. Maurílio Boaventura, pela confiança, pelo conhecimento transmitido, pela amizade, a paciência e as observações valiosas na elaboração desta dissertação.

Ao Rodrigo Contreras por ter contribuído de várias maneiras para este trabalho, desde dicas sobre implementações até os códigos raízes do Algoritmo Genético.

À CAPES, pelo suporte financeiro.

“Nada é perfeito, o mundo não é perfeito, mas ele está lá para nós, fazendo o melhor que pode. Isso é o que o torna tão maldito e belo.”

Edward Elric

Resumo

O Reconhecimento facial é uma das tecnologias biométricas mais utilizadas em sistemas automatizados que necessitam garantir a identidade de uma pessoa para acesso autorizado e monitoramento. A grande aceitação do uso da face tem várias vantagens sobre outras tecnologias biométricas: ela é natural, não exige equipamentos sofisticados, a aquisição de dados é baseada em abordagens não invasivas, e pode ser feito a distância, de maneira cooperativa ou não. Embora muitos estudos em reconhecimento facial tenham sido feitos, problemas com variação de iluminação, poses com oclusão facial, expressão facial e envelhecimento ainda são desafios, pois influenciam a performance dos sistemas de reconhecimento facial e motivam o desenvolvimento de novos sistemas de reconhecimento que lidam com esses problemas e sejam mais confiáveis. Este trabalho tem como objetivo avaliar a técnica de Padrões Localmente Mapeados em Multiescala (MSLMP) para o reconhecimento facial. Técnicas baseadas em algoritmos genéticos e processamento de imagens foram usadas para obter melhores resultados. Os resultados obtidos chegam a 100% de acurácia para alguns banco de dados. A base de dados MUCT é, em particular, bastante complexa, ela foi criada em 2010 com o objetivo de aumentar a quantidade de bancos de dados disponíveis com alta variação de iluminação, idade, posições e etnias, e por isso, é um banco de dados difícil quanto ao reconhecimento automático de faces. Uma nova técnica de processamento baseada na média dos níveis de cinza da base foi desenvolvida.

Palavras-chave: Reconhecimento Facial. Padrões Localmente Mapeados. Algoritmos Evolutivos. Reconhecimento de Padrões. Sistemas Biométricos.

Abstract

Facial recognition is one of the most used biometric technologies in automated systems which ensure a person's identity for authorized access and monitoring. The acceptance of face use has several advantages over other biometric technologies: it is natural, it does not require sophisticated equipment, data acquisition is based on non-invasive approaches, and can it be done remotely, cooperatively or not. Although many facial recognition studies have been done, problems with light variation, facial occlusion, position, expression, and aging are still challenges, because they influence the performance of facial recognition systems and motivate the development of more reliable recognition systems that deal with these problems. This work aim to evaluate the Multi-scale Local Mapped Pattern (MSLMP) technique for the facial recognition. Techniques based on genetic algorithms and image processing were applied to increase the performance of the method. The obtained results reach up to 100% of accuracy for some databases. A very difficult database to deal is the MUCT database which was created in 2010 with aim of providing images with high variation of lighting, age, positions and ethnicities in the facial biometry literature, which makes it a highly difficult base in relation to automated recognition. A new processing technique was developed based on the average gray levels of the images of the database.

Keywords: Facial Recognition. Local Mapped Pattern. Evolutionary Algorithms. Pattern Recognition. Biometric Systems.

Lista de Figuras

1.1	Exemplos de traços fisiológicos: 1, 2, 3, 5, 6, 9 e 12; e de traços comportamentais: 4, 7, 8 e 11. Fonte: (A. K. Jain, Flynn & Ross, 2008)	17
1.2	Situações Diárias em que a Biometria é Usada. Leitores de: a) iris; b) cartão e geometria da mão; c) digital; d) geometria da mão; e) digital; e f) face.	18
1.3	Exemplo da organização de dados em um banco de dados biométrico.	20
1.4	Processo de Identificação.	20
1.5	Processo de Verificação.	20
1.6	Pontos de ataque de um sistema biométrico. Fonte (R. Jain & Kant, 2015)	22
2.1	Ordem de processamento de um sistema para faces.	28
2.2	Exemplo de imagem pré-processada dividida em 64 regiões. Fonte:(Rahim, Hossain, Wahid & Azam, 2013)	32
2.3	Ilustração do Cálculo do LBP de uma Região. Fonte (Jia et al., 2014))	32
2.4	Exemplos das Imagens da Faces94.	35
2.5	Exemplos das Imagens da Faces95.	36
2.6	Exemplos das Imagens da Faces96.	37
2.7	Exemplos das Imagens da Grimace.	37
2.8	Exemplos das Imagens da JAFFE.	38
2.9	Exemplos das Imagens da Muct.	39
2.10	Posição das “Webcams”.	40
2.11	Imagens das cinco “Webcams” e das Iluminações q, r e s.	40
2.12	Imagens do ORL.	41
2.13	Exemplo das Expressões e Iluminação de uma Classe da YALE.	41
3.1	Região Circular para Três Diferentes Valores P , R . Fonte (Rahim, Hossain, Wahid & Azam, 2013))	43
3.2	Cálculo do $MSLBP2$ de uma matriz 9×9 . a) Matriz 9×9 , b) Média das regiões 3×3 da matriz original, c) Cálculo do $MSLBP2$. Fonte: (Contreras, 2015).	44
3.3	Distribuição ótima dos raios do padrão $MSLBP$ com respeito à redução de redundância da informação capturada pelo padrão. Fonte: (Jia et al., 2014).	45
3.4	Cálculo do mapeamento relativo ao pixel central. (a) Vizinhança 3×3 do pixel central $g(2, 2)$, (b) Função de mapeamento, (c) Matriz dos mapeamentos realizados na vizinhança do pixel central, (d) Matriz de pesos, (e) Mapeamento relativo à vizinhança de $g(2, 2)$. Fonte : (Contreras, Boaventura & Boaventura, 2015).	46
3.5	Gráfico da função sigmoide gerado pelo MATLAB.	47

3.6	Forma Geral de Um Algoritmo Genético. Fonte:(Miranda, s.d.)	50
3.7	Crossover de Ponto Único.	51
3.8	Crossover Uniforme.	51
3.9	Ilustração do método de seleção roleta-viciada. A) Indica o indivíduo menos apto da população. B) indica o indivíduo mais apto. C) É o ponto de seleção. Fonte: (“Roulette wheel selection”, s.d.)	53
3.10	Média dos níveis de cinza da MUCT.	58
3.11	Exemplos da Técnica Descritas com Imagens de 3 Iluminações distintas.	60
3.12	Média dos níveis de cinza da YALE em relação a quantidade de imagens.	61
3.13	3 imagens da Yale antes e depois da implementação da técnica descrita.	62
4.1	Processo de classificação de faces.	65
4.2	Cálculo da acurácia da Faces94 para uma escala.	66
4.3	Cálculo da acurácia da Faces95 para uma escala.	67
4.4	Cálculo da acurácia da Faces96 para uma escala.	68
4.5	Cálculo da acurácia da Grimace para uma escala.	69
4.6	Cálculo da acurácia da JAFFE para uma escala.	70
4.7	Cálculo da acurácia da MUCT para uma escala.	71
4.8	Cálculo da acurácia para uma escala.	73
4.9	Cálculo da acurácia da YALE para uma escala.	75

Lista de Tabelas

1.1	Indicativos Técnicos Quanto ao Uso de Faces como Biometria. Fonte: (Srivastava, 2013)	24
2.1	Algumas Aplicações do Reconhecimento Facial	27
2.2	Algumas Comparações entre Características Globais e Locais. Fonte (Meethongjan & Mohamad, 2007)	30
2.3	Versão Geral das Imagens ESSEX.	35
2.4	Visão Geral da Faces94.	35
2.5	Informações Sobre as Imagens da Faces95.	36
2.6	Informações Sobre as Imagens da Faces96.	37
2.7	Informações Sobre as Imagens da Grimace.	38
2.8	Distribuição das Iluminações nas Classes das Imagens.	39
3.1	Acurácias obtidas pelo conjunto de valores da população P	55
3.2	Conversão binária da população.	56
3.3	Processo de seleção de roleta-viciada.	56
3.4	Processo de recombinação.	56
3.5	Geração 1	56
3.6	Processo de Mutação	56
3.7	Conversão dos cromossomos para a base decimal.	57
3.8	População 0 para o caso 2.	57
4.1	Parâmetros Iniciais do AG utilizados neste Trabalho.	64
4.2	Taxa de acertos Essex 94 para três escalas.	67
4.3	Taxa de acerto Faces95 do <i>MSLMP</i> para duas escalas.	68
4.4	Taxa de acerto da Faces95, para duas escalas, utilizando o AG para as matrizes e $\lambda = 0,2849$	68
4.5	Taxa de acertos da Faces96 com cinco escalas.	68
4.6	Taxa de acerto da Faces96 para cinco escalas utilizando o AG.	68
4.7	Exemplo de formação dos grupos de uma classe da MUCT.	71
4.8	Taxa de acertos da MUCT para 8 escalas	71
4.9	Taxa de acertos da MUCT para 8 escalas com reajuste dos níveis de cinza.	72
4.10	Taxa de acertos para $\lambda = 0,2849$ e 8 escalas.	72
4.11	Taxa de acertos para $\lambda = 0,2849$ e otimizando a matriz do raio 29 com 8 escalas.	72
4.12	Taxa de acerto ORL.	73
4.13	Taxa de acerto ORL com o AG.	73
4.14	Taxa de acerto ORL com equalização das imagens.	74

4.15	Taxa de acerto ORL com equalização das imagens e AG.	74
4.16	Taxa de acerto YALE com 12 escalas.	74
4.17	Taxa de acerto YALE utilizando os raios 23 e 29.	75
4.18	Taxa de acerto YALE utilizando 5 amostras para extração em 12 escalas.	76
4.19	Taxa de acerto YALE através da alteração dos níveis de cinza.	76
4.20	Taxa de acerto YALE através da alteração dos níveis de cinza e exclusão dos padrões gerados pelo fundo das imagens.	76
4.21	Média da soma da frequência dos “uniform patterns” extraídos das imagens da ORL para diferentes valores de raio.	77

Símbolos

<i>BD</i>	Banco de Dados
<i>FAR</i>	False Accept Rate
<i>FRR</i>	False Reject Rate
<i>MSLMP</i>	Multi-Scale Local Mapped Pattern
<i>JAFFE</i>	Japanese Female Facial Expression
<i>PCA</i>	Principal Component Analysis
<i>LDA</i>	Linear Discriminant Analysis
<i>LBP</i>	Local Binary Pattern
<i>DNA</i>	Ácido Desoxiribonucleico
<i>AG</i>	Algoritmo Genético
<i>LMP</i>	Local Mapped Pattern
<i>MSLBP</i>	Multi-Scale Local Binary Pattern
<i>AE</i>	Algoritmos Evolutivos
<i>PE</i>	Programação Evolutiva

Sumário

1	Introdução	16
1.1	Biometria: Contexto Histórico	16
1.2	Sistemas Biométricos	18
1.2.1	Processo de Classificação	19
1.2.2	Performance	20
1.2.3	Sistemas Monomodais e Multimodais	21
1.2.4	Vulnerabilidade de Sistemas Biométricos	22
1.3	Motivação e Objetivos	23
1.4	Proposta do Trabalho	25
1.5	Organização do Trabalho	25
2	Reconhecimento de Faces	26
2.1	Introdução	26
2.2	Sistemas para Reconhecimento Facial	27
2.2.1	Módulos	28
2.2.2	Tipos de Algoritmos	28
2.3	Fatores que Afetam o Reconhecimento de Faces	29
2.4	Revisão da Literatura para Identificação de Faces	31
2.4.1	Técnicas Para o Reconhecimento Facial	33
2.5	Bancos de Dados de Imagens 2D	33
2.5.1	Banco de Dados ESSEX	34
2.5.2	Banco de Dados <i>Japanese Female Facial Expression (JAFFE)</i>	38
2.5.3	A Base de Dados MUCT	38
2.5.4	O Banco de Dados ORL	40
2.5.5	O Banco de Dados YALE	41
2.6	Considerações Finais	41
3	Técnicas Utilizadas	42
3.1	Introdução	42
3.2	Padrões Binários Locais em Multiescala	42
3.3	Padrões Mapeados Localmente	45
3.4	Padrões Mapeados Localmente em Multiescala	46
3.5	Extração de características	48
3.6	Algoritmo para Implementação do MSLMP	48
3.7	Algoritmos Evolutivos	49
3.7.1	Algoritmos Genéticos (AG's)	49

3.8	Implementação do Algoritmo Genético	52
3.9	Ajuste nos Níveis de Cinza	58
3.9.1	Correção de Brilho e Homogeneização	58
3.9.2	Ajuste na Intensidade dos Pixels	59
3.10	Considerações Finais	63
4	Metodologias e Resultados Experimentais	64
4.1	Considerações Iniciais	64
4.2	Metodologia Aplicada	64
4.2.1	Parâmetros do AG	64
4.2.2	Validação Cruzada	64
4.2.3	Processo de Classificação	65
4.2.4	Uso das Escalas	65
4.3	Resultados Experimentais	66
4.3.1	ESSEX	66
4.3.2	JAFFE	69
4.3.3	MUCT	70
4.3.4	ORL	72
4.3.5	YALE	74
4.4	Considerações Finais	76
5	Considerações Finais	78
5.1	Conclusões	78
5.2	Contribuições	78
5.3	Trabalhos Futuros	79
5.4	Artigos Aceitos e Publicados	79
	Referências	80

Introdução

1.1 Biometria: Contexto Histórico

Com os avanços de tecnologias e serviços dos dias atuais, atividades e transações humanas evoluíram muito, desta forma, métodos rápidos e confiáveis de identificação de pessoas se tornaram indispensáveis. Uma ampla variedade de sistemas requer segurança nos processos de identificação de pessoas tanto para confirmar a identidade do indivíduo quanto para tentar identificá-la. O propósito de tais sistemas é assegurar que a tarefa prestada seja realizado apenas por pessoas que possuam o devido acesso ao serviço. A falta de robustez nos métodos empregados para reconhecimento pode tornar o sistema vulnerável para impostores com propósitos ilícitos ou prejudicar o próprio ato de identificação, o que torna o método não confiável. Chama-se de **Reconhecimento Biométrico**, ou simplesmente **Biometria**, o reconhecimento automático de pessoas baseado em suas características fisiológicas ou comportamentais (A. K. Jain, 2005).

Características fisiológicas são aquelas que se referem a partes do corpo humano enquanto que as comportamentais são as que resultam de alguma ação humana. A Figura 1.1 exemplifica melhor algumas dessas características.

De acordo com (A. K. Jain, Flynn & Ross, 2008) as aplicações da biometria podem ser caracterizadas em três grupos principais:

1. aplicações comerciais, como o acesso a rede de computadores, segurança de dados eletrônicos, comércio online, acesso a internet, uso de cartão de crédito, controladores de acesso físico, registros de dados médicos, smartphones, etc;
2. aplicações governamentais, como documento de identidade, carteira de motorista, segurança social, controle de passaporte, controle de imigração em fronteiras, etc;
3. aplicações em ciências forenses, como identificação de corpos, investigação criminal, investigação de parentesco, etc.

Após definido o conceito de biometria surge a pergunta, o que é preciso para que uma característica biológica seja uma biometria? Para (A. K. Jain, Ross & Prabhakar, 2004) qualquer característica fisiológica ou comportamental se enquadra como uma biometria se possui os seguintes aspectos:

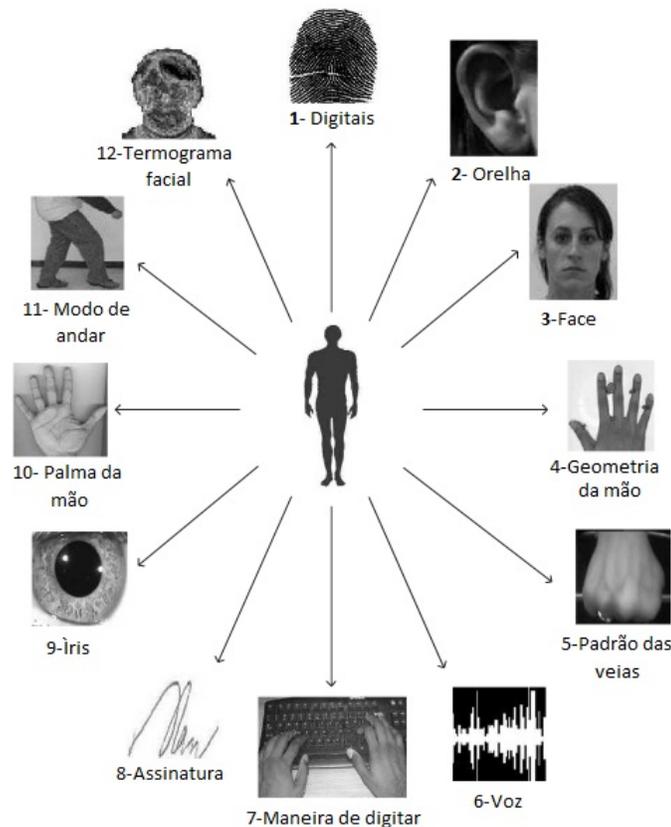


Figura 1.1: Exemplos de traços fisiológicos: 1, 2, 3, 5, 6, 9 e 12; e de traços comportamentais: 4, 7, 8 e 11. Fonte: (A. K. Jain, Flynn & Ross, 2008)

- **universalidade:** cada pessoa deve possuí-la;
- **distinção:** quaisquer duas pessoas devem possuir diferenças em termos da característica;
- **permanência:** a característica deve ser invariante em relação ao tempo.

No entanto, durante a aplicação real de um sistema biométrico existem alguns problemas a mais que não podem ser desconsiderados:

- **performance:** se refere a acurácia, ao tempo do reconhecimento e aos recursos necessários para seu funcionamento;
- **aceitabilidade:** indica a possibilidade de cada pessoa aceitar o uso de um particular sistema biométrico em seu dia-a-dia;
- **invasão:** é a facilidade que o sistema pode ser burlado através de métodos fraudulentos.

Não há uma biometria específica que cumpra todos os quesitos propostos anteriormente. A performance entre uma característica e sua aplicação depende de sua natureza e de suas requisições de aplicação.

Graças a biometria é possível confirmar a identidade de uma pessoa através do que ela é e não pelo que ela possui (um cartão de acesso) ou pelo que ela se lembra (uma senha). O uso da biometria torna muito prática várias situações, como mostrado na Figura 1.2: a)

o programa Schiphol Privium no aeroporto de Amsterdan que utiliza um scanner de íris para validar a identidade do viajante; b) o aeroporto Ben Gurion em Tel Aviv usa leitor de cartão acoplado com um sistema de geometria de mãos para acesso a quiosques com propósitos de segurança e controle de imigração; c) a urna eletrônica brasileira utiliza um sistema baseado em digitais para identificação do eleitor; d) agências bancárias utilizam um sensor para ler veias da mão do cliente; e) alguns smartphones vem com leitor de digitais acoplado para “destravar” o aparelho e possibilitar o acesso as suas funcionalidades como, por exemplo, gerenciador de contas online; f) na entrada no estádio nacional da China durante as olimpíadas de 2008 foi utilizado um sensor para reconhecimento de faces durante a cerimônia de abertura (“China’s Face Recognition System Serves Olympics”, 2008).



a)



b)



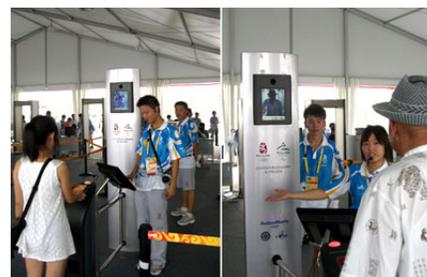
c)



d)



e)



f)

Figura 1.2: Situações Diárias em que a Biometria é Usada. Leitores de: a) íris; b) cartão e geometria da mão; c) digital; d) geometria da mão; e) digital; e f) face.

1.2 Sistemas Biométricos

Pode-se dizer que sistemas biométricos são mecanismos de captura e bloqueio (ou desbloqueio) a fim de controlar o acesso de pessoas a locais específicos. Para ter acesso a esses

esquemas de segurança, a pessoa deve providenciar a biometria específica do dispositivo para ser comparada com as demais no banco de dados do sistema. A relação entre a biometria e o sistema biométrico é também conhecida como o sistema de trava e chave (Le, 2011). Portanto, um sistema biométrico é um dispositivo de reconhecimento de padrões que precisa de uma característica física ou comportamental para avaliar a veracidade da identidade de um indivíduo, de forma geral, esse sistema possui quatro módulos principais:

1. **leitor**, é o responsável pela captura da biometria, em estado “bruto”, da pessoa. Para obter a face de uma pessoa, por exemplo, é necessário, em geral, um dispositivo que fotografe seu rosto quando posicionado em frente ao sensor. Ele é fundamental para o sistema biométrico, pois um leitor não muito preciso pode prejudicar a captura e afetar o processo de reconhecimento. Os leitores baseados em fotos podem ser prejudicados pela qualidade da câmera usada, pois em uma imagem de baixa qualidade a presença de ruído e até a iluminação do ambiente podem influenciar no processo de identificação (Levine & Yu, 2006). Devido a isso existe um módulo responsável por avaliar a qualidade dos dados capturados e aplicar técnicas de pré-processamento para melhorar a qualidade e, até mesmo, se necessário, solicitar ao indivíduo que repita o processo de leitura, como acontece com leitores de digitais, por exemplo. Para o caso de imagens, pode-se aplicar filtros para eliminação de ruídos e corrigir distorções com iluminação (Buciu & Gacsadi, 2016);
2. **extrator**, tem a função de extrair, dos dados da biometria adquirida pelo leitor, um conjunto de características discriminatórias que definam o indivíduo que utiliza o dispositivo. Na prática, esse módulo segue uma sequência de comandos que selecionam informações úteis da biometria. Se aumentada a qualidade do sensor durante a leitura e nos métodos utilizados para a extração, a validade da informação extraída se torna mais confiável, tornando o sistema mais robusto. Por exemplo, para o caso de reconhecimento de faces podem ser utilizados descritores de Padrões Locais Binários, tais como o Padr (LBP), (Ojala, Pietikäinen & Mäenpää, 2002), e suas extensões para gerar um vetor de características da imagem e armazená-la ou compará-la com os vetores inscritos no banco de dados;
3. **classificador** ou (“matching”), é o responsável por comparar a amostra obtida pelo extrator com as demais existentes no banco de dados, gerando o chamando “matching point”, que é um valor usado tanto para verificação quanto para identificação. Na prática o “matching point” é o valor obtido da amostra de teste através de uma função de similaridade, que é responsável por estabelecer o grau de semelhança entre dois vetores de amostras de características;
4. **banco de dados** (BD), é um depósito de informações biométricas. As características extraídas dos dados do leitor são armazenadas junto com informações que identificam o usuário a quem as características pertencem. A Figura 1.3 ilustra a organização dos dados em um banco de dados.

1.2.1 Processo de Classificação

O processo de classificação envolve, em geral, duas etapas: a de **verificação** (ou **autenticação**) e a de **identificação** (ou **reconhecimento**). *Verificação* é um processo individual que valida a identidade afirmada através de um cartão ou de documento de identificação, comparando

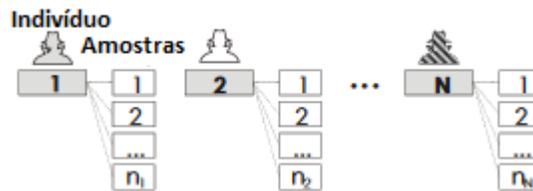


Figura 1.3: Exemplo da organização de dados em um banco de dados biométrico.

seus dados biométricos com os disponíveis no banco de dados referente à identidade que se deseja confirmar, esse processo possui como resposta apenas sim ou não, é um procedimento binário (Esta pessoa é realmente quem ela diz ser?). *Reconhecimento*, diferentemente da verificação, esse processo requer a comparação dos dados com todos os arquivos disponíveis no banco de dados do sistema a fim de identificar ou não a identidade da pessoa que o utiliza. As Figuras 1.4 e 1.5 representam esses esquemas, Fonte: (A. K. Jain et al., 2004).

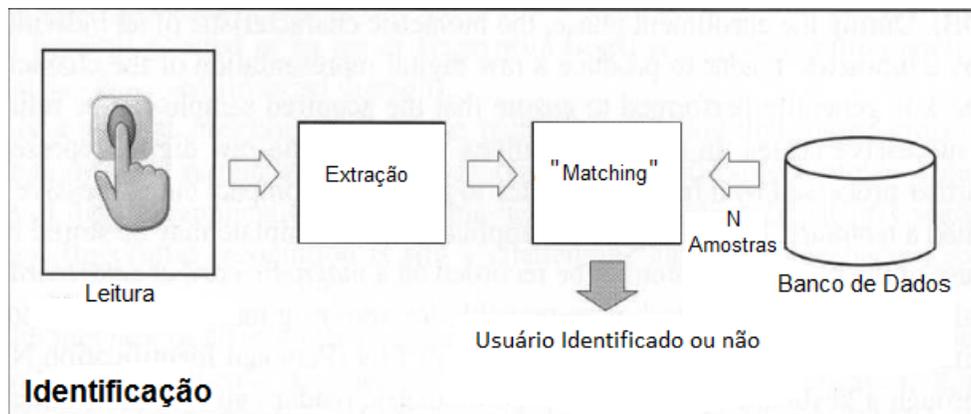


Figura 1.4: Processo de Identificação.

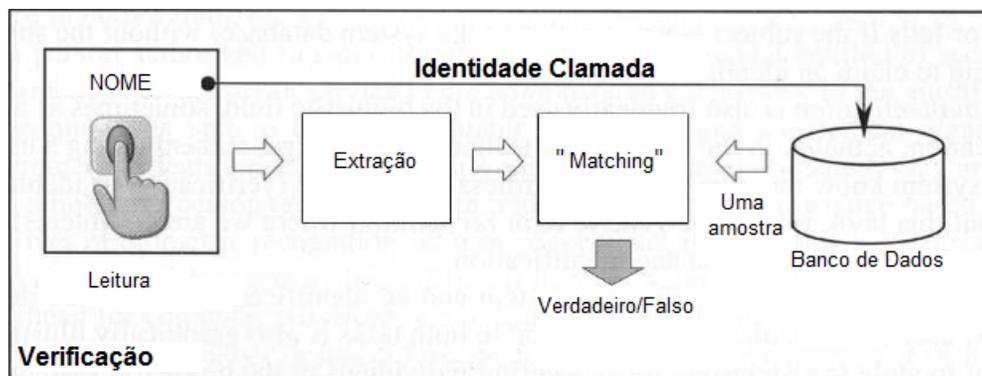


Figura 1.5: Processo de Verificação.

1.2.2 Performance

A performance de um sistema depende dos resultados obtidos, da forma de seleção dos dados, das tecnologias utilizadas e da taxa de acerto do método utilizado. Esse tipo de avaliação é, em geral, realizado em ambientes de pesquisas que avaliam o modelo utilizado para extração de características, juntamente com outros fatores do aparelho biométrico, como por exemplo,

a aceitação geral pelas pessoas quanto aos requisitos de funcionamento desse sistema (Philisp, Martin, C.L. Wilson & Przybocki, 2000).

A acurácia de medidas serve para avaliar a concordância entre os resultados obtidos de uma avaliação e os valores esperados ao se aplicar o sistema em um banco de dados. No caso da performance tecnológica existem índices de acurácia bastante usados e aceitos pela literatura de sistemas biométricos. Esses índices são caracterizados por dois erros estatísticos: a *taxa de falsos negativos*, (**FAR**) e *taxa de falsos positivos* (**FRR**). A (**FAR**) se refere aos casos em que o sistema aceita como verdadeira a identidade de uma pessoa que não é quem ela diz ser, e (**FRR**) é para casos em que o sistema valida como falsa a identidade de usuários legítimos.

FAR e **FRR** caminham lado a lado, se um sistema conceder acesso a todos os seus usuários têm-se **FRR** = 1 e **FAR** = 0, caso ele não garanta o acesso para nenhum **FAR** = 1 e **FRR** = 0. Em um sistema biométrico ideal ambas as taxas devem ser zero, porém isso está longe do que acontece na realidade.

1.2.3 Sistemas Monomodais e Multimodais

Os dados de uma pessoa extraídos de um banco de dados podem ser de uma única amostra retirada do leitor ou de duas ou mais amostras, retiradas a partir de um critério de aquisição específico, por exemplo, para sistemas de reconhecimento de face são retiradas fotos de diferentes posições do indivíduo (A. K. Jain et al., 2008). Sistemas biométricos que utilizam apenas uma amostra para gerar os dados característicos da biometria são chamados de *monomodais* e aqueles que usam $N, N > 1$, amostras são chamados de sistemas *multimodais*. Muitas pesquisas apostaram que sistemas multimodais conseguem melhor acurácia em relação aos monomodais (Marco, Massimo, Vincenzo, Daniele & Fabio, 2004).

Os sistemas monomodais em geral são afetados pelos seguintes problemas (A. K. Jain, 2005):

1. **ruído nos dados:** o ruído pode estar presente na amostra adquirida pelo sensor, devido a algum defeito de captura ou pela baixa qualidade do aparelho. A presença de ruído dificulta o processo de extração de características significativas da amostra, ocasionando a perda de acurácia no sistema;
2. **falta de individualidade:** amostras adquiridas de pessoas diferentes podem ser similares. De acordo com (Golfarelli, Maio & Malton, 1997), a aparência usual dos rostos durante a captura de imagens para o reconhecimento facial mostrou-se baixa em relação a sua capacidade discriminativa. Um exemplo clássico desse fato está nas características hereditárias presentes na população. Essa falta de individualidade biométrica pode aumentar a taxa de Falsos Negativos no sistema;
3. **falta de representação nos dados:** as características extraídas de uma pessoa durante o processo de identificação podem não ser as mesmas geradas durante o processo de cadastramento de identidade no banco de dados, esse fator é conhecido como variação intra-classe. Essa variação de dados pode vir de uma má interação do usuário com o sensor, como por exemplo, problemas de rotação e translação do indivíduo na hora da leitura, mudança de pose e expressão facial em frente à câmera, uso de diferentes sensores durante o cadastramento, mudanças nos ambientes de capturas e mudança de traços genéticos devido ao envelhecimento ou presença de cicatrizes. De maneira estrita, os dados extraídos pelo sistema devem ser relativamente invariantes a essas mudanças, no entanto, muitos dos procedimentos baseados em biometria hoje não cumprem esse requisito, devendo então ser compensados por algoritmos de extração mais complexos capazes de lidar com essas

variações. Uma grande quantidade de informação “intra-classe” aumenta a taxa de Falsos Negativos, o que é prejudicial ao sistema.

Comparado com sistemas unimodais, os multimodais superam algumas limitações e oferecem vantagens, tais como (Taouche, Berkane, Batouche & Talebahmed, 2014):

- melhor acurácia nos procedimentos de identificação e reconhecimento;
- menos passíveis a ataques impostores;
- lida melhor com o problema de não individualidade;
- insensitivo a ruídos presentes nos dados.

1.2.4 Vulnerabilidade de Sistemas Biométricos

Sistemas biométricos são vulneráveis a diversos tipos de ataques que comprometem a segurança oferecida por essas aplicações. Existem oito pontos em um sistema que são suscetíveis a ataques, eles se encontram representados na Figura 1.6.

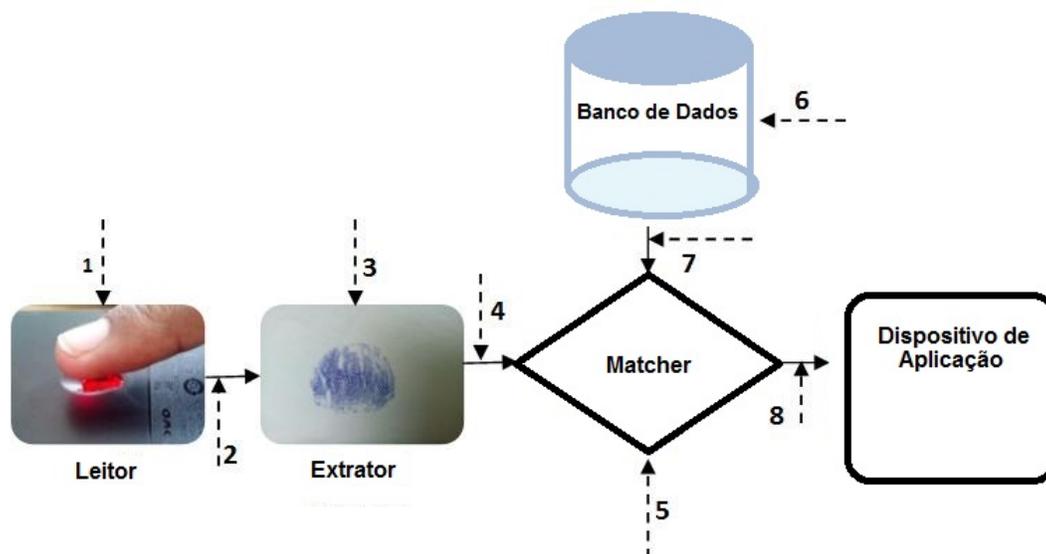


Figura 1.6: Pontos de ataque de um sistema biométrico. Fonte (R. Jain & Kant, 2015)

Os ataques a esses pontos podem ser divididos em dois tipos, os ataques diretos e os indiretos. **Ataques diretos** são aqueles que não necessitam de conhecimento específico sobre o funcionamento do sistema em questão. **Ataques indiretos**, ao contrário dos diretos, precisam de informações sobre o funcionamento interno do sistema para burlá-lo. A lista a seguir refere-se aos possíveis tipos de ataques em cada ponto mostrado na Figura 1.6, (R. Jain & Kant, 2015).

Ataques Diretos:

1. Os ataques a esse local são conhecidos como “Ataques ao sensor”. No qual, dados biométricos falsos, como dedos de silicone ou imagens de faces são apresentadas ao sensor pelo impostor a fim de burlar o sistema de reconhecimento, além disso, pode-se danificar o leitor, requisitando várias vezes acesso com informações falsas. Leitores, muitas vezes, são incapazes de diferenciar características reais das falsas e por isso podem ser burlados.

Ataques Indiretos:

2. Quando o leitor adquire dados biométricos, tais dados são enviados para o módulo de extração de característica para pré-processamento através de um canal de comunicação de dados. Esse canal fica entre o sensor e o módulo de extração. Durante o ataque ele é interceptado para “furtar” a informação da pessoa e é arquivado em algum lugar para que então possa ser usado novamente a vontade do impostor, a fim de burlar o sistema.
3. O impostor pode atacar o sistema de extração e fazer com que se obtenha valores escolhidos por ele, substituindo aqueles que seriam gerados pelos dados originais obtidos pelo sensor.
4. Esse ataque é similar ao de número 2, com a diferença de que o canal de interceptação para “furtar” as características fica entre o extrator e o classificador. Após obtidos esses dados, o farsante pode usá-los novamente no classificador.
5. Nesse ponto pode-se alterar o valor obtido pelo classificador durante o procedimento de autenticação, para tornar os dados inseridos no leitor parecidos com os dados originais da pessoa, a qual o impostor tenta se passar.
6. Nesse local é possível burlar a segurança do banco de dados para adicionar ou remover amostras. Essa não é uma tarefa fácil, pois os dados do banco de dados, em geral, são protegidos por mecanismos digitais como criptografia ou marca d’água. Para obter sucesso em ataques nesse local é exigido alto conhecimento do funcionamento interno do sistema.
7. Nesse ponto, pode-se interromper o canal de comunicação dos dados para “furtar”, substituir ou alterar as amostras biométricas disponíveis no banco de dados.
8. Pode-se substituir o resultado declarado pelo módulo de classificação. Nesse ponto o impostor pode interromper o canal de comunicação de dados e alterar o resultado final do sistema, aceitando ou rejeitando os dados inseridos no sensor.

De forma geral, um sistema biométrico deve ter uma boa acurácia de reconhecimento, ser rápido não deixando o indivíduo desconfortável com o tempo de espera para finalizar o processo, ser inofensivo ao usuário, ser aceito pelo público alvo pretendido, ser de fácil uso e ser suficientemente robusto contra métodos fraudulentos e ataques ao sistema.

1.3 Motivação e Objetivos

As pessoas tem usado características corporais e comportamentais como face, voz e a maneira de andar para reconhecer umas as outras há milhares de anos. De acordo com (A. K. Jain et al., 2004), Alphonse Bertillon, chefe da divisão de identificação criminal do departamento de Polícia de Paris, desenvolveu e colocou em prática a ideia de usar medidas físicas para identificar criminosos no século XIX. Quando sua ideia se tornou popular, no final do século, é que surgiram as primeiras descobertas do uso de impressões digitais como fator discriminativo de humanos. Após essa descoberta muitos departamentos de policia adotaram a ideia de primeiro arquivar (a princípio as técnicas de aquisição eram no papel) as digitais dos criminosos e guardá-las em um banco de dados (armário de arquivos) para que partes das digitais encontradas em cenas de crimes pudessem ser usadas para comparar com as digitais do banco de dados, e assim tentar identificar o criminoso.

Embora a biometria tenha surgido de sua extensiva aplicação na área criminal, hoje ela está sendo cada vez mais usada com propósitos de reconhecimento em diversas aplicações civis. Muitas das novas aplicações são atribuídas aos problemas de aumento de segurança, tanto para transações comerciais, como para proteção de informações virtuais. Os sistemas biométricos oferecem muitas vantagens comparados aos sistemas de reconhecimento tradicionais, porém, eles são vulneráveis a ataques que comprometem sua segurança, sendo esse um ponto altamente desfavorável (A. K. Jain, 2005). Sistemas Biométricos são mais convenientes que meios tradicionais, pois não necessitam memorização de senhas e são mais práticos e rápidos de serem utilizados.

Mesmo que vários fatores biométricos tenham sido propostos com o decorrer das décadas, as impressões digitais talvez sejam o exemplo de biometria mais utilizada até hoje. Como já é bem conhecido, pessoas diferentes possuem digitais diferentes e elas não se alteram com o passar do tempo, salvo menção ao casos de acidentes ou cirurgias. No entanto, é claro que a identificação através de digitais tem como desvantagem a necessidade de contato físico da pessoa. Não existem métodos que capturem digitais à distância ou pelo menos nenhum que pareça ser adequado para tal condição (Daugman, 1994).

Uma biometria alternativa e eficaz quanto as digitais é a face. Sistemas biométricos para reconhecimento de faces se tornaram bastante importantes e utilizados devidos ao rápido avanço nos aparelhos para captura de imagem (câmeras de segurança, câmeras de celulares,...), a grande quantidade de imagens de faces disponíveis na internet, ao aumento da demanda de sistemas de segurança (Li & Janil, 2011), à alta aceitabilidade entre as pessoas e, além disso, é natural e não invasivo, pois a imagem da face pode ser capturada a distância e sem nenhum contato físico do indivíduo. No trabalho de Srivastava, 2013, o autor realizou uma pesquisa comparativa entre vários tipos de biometrias tais como iris, digitais, faces, retina, orelha, voz e DNA, com relação a fatores técnicos, sociais e de implementação, os resultados para faces se encontram na Tabela 1.1, na qual, alto, médio e baixo estão substituídos por A, M e B, respectivamente.

Tabela 1.1: Indicativos Técnicos Quanto ao Uso de Faces como Biometria. Fonte: (Srivastava, 2013)

Unicidade	M	Circunvenção	A
Permanência	M	Privacidade	A
Universalidade	A	Segurança	M
Mensurabilidade	M	Custo	M
Compatibilidade	B	Introdução Social	2000
Colheitabilidade	A	Popularidade	A
Performance	B	Facilidade de Uso	A
Aceitabilidade	A	Higiene	B

Devido a todos esses aspectos positivos quanto a biometria facial, os investimentos nessa área vem aumentando cada vez mais. De acordo com a agência de notícia Bloomberg, a China e a Rússia estão investindo 460 milhões de dólares em tecnologias de reconhecimento facial (Che, 2017).

Apesar de existirem inúmeros sistemas comerciais de reconhecimento de face em uso, esse meio de identificação continua sendo um interessante tópico para pesquisas. Os sistemas atuais oferecem boa performance em relação à ambientes simples e controlados, porém deixam a desejar quando há variação de fatores, como pose, ângulo de vista, expressão facial e iluminação

(mudanças no brilho) (Rahim, Hossain, Wahid & Azam, 2013). Desta forma, desenvolver algoritmos robustos é um desafio ainda nos dias atuais. Esse desafio é a grande motivação para o desenvolvimento deste trabalho, cujo objetivo principal é o reconhecimento facial.

1.4 Proposta do Trabalho

Este trabalho tem três propósitos principais:

- avaliar o método de padrões mapeados localmente em multiescala (*MSLMP*), introduzido inicialmente em (Contreras, 2015) para a análise de impressões digitais fraudulentas, para o problema de reconhecimento de faces;
- aplicar técnicas de algoritmos genéticos para determinar melhores parâmetros exigidos pelo método *MSLMP*;
- desenvolver e aplicar novas técnicas de pré-processamento de imagens específicas para banco de dados que apresentam grandes variações de iluminação e poses nas imagens.

1.5 Organização do Trabalho

Além desse Capítulo em que foram apresentados conceitos gerais sobre sistemas biométricos, esta dissertação se divide em mais quatro capítulos.

No Capítulo 2 encontram-se apresentados os sistemas biométricos utilizando especificamente o reconhecimento facial. Além disso, são apresentados os bancos de dados utilizadas nesta pesquisa, ou seja os bancos de dados ESSEX, JAFFE, MUCT, ORL e YALE.

No Capítulo 3 são apresentadas as técnicas utilizadas para avaliar e otimizar a acurácia do método. São também descritos os métodos que deram origem ao *MSLMP* e a técnica de Algoritmos Genéticos, empregada neste trabalho para melhorar a performance do método. Também são apresentadas duas novas técnicas de pré-processamento baseadas nos níveis de cinza das imagens.

No Capítulo 4 é abordada a metodologia empregada neste trabalho e são apresentados os resultados obtidos utilizando as técnicas descritas no Capítulo 3.

No Capítulo 5 são dadas as conclusões sobre o trabalho, são apresentadas as possibilidades de trabalhos futuros e são listados os artigos publicados e aceitos para publicação.

Reconhecimento de Faces

Neste capítulo são tratados alguns aspectos sobre as características de uma face humana, modelos para sua caracterização em um sistema biométrico e alguns métodos relacionados ao reconhecimento facial presentes na literatura, além da descrição de algumas bases de dados utilizadas para mensurar a performance de modelos propostos ao reconhecimento facial.

2.1 Introdução

O rosto humano é repleto de informações sobre a identidade e o estado emocional de uma pessoa. Reconhecer um rosto é uma tarefa de rotina que as pessoas realizam sem esforços. Em geral, o nosso sistema de reconhecimento utiliza um amplo espectro de estímulos, obtidos por muitos, se não todos, os sentidos (visão, olfato, tato, audição e paladar). Esses estímulos são usados para armazenamento ou para recuperação de imagens de rostos já vistos anteriormente. A forma de captura das imagens, o ambiente interno ou externo, por exemplo, interferem em muito na tarefa de reconhecimento de um indivíduo. É impossível (com as tecnologias atuais) desenvolver um sistema que seja capaz de reproduzir, com a mesma precisão e habilidades humanas, o reconhecimento facial. No entanto, o cérebro humano possui uma limitação com relação a quantidade de pessoas que se pode lembrar para o reconhecimento. Um dos pontos chave de um sistema de computadores é sua capacidade de lidar com grandes quantidades de imagens.

De acordo com (Tolba, El-Baz & A.A., 2005), o primeiro método formal para classificação de faces foi proposto por (Galton, 1888). O autor propôs coletar perfis faciais como curvas e encontrar sua norma, para então classificar outras pessoas através de seus respectivos desvios. Essa classificação é multimodal, isto é, resultante de um vetor de medidas independentes que pode ser comparado com outros vetores de um mesmo banco de dados. A motivação para o avanço no progresso de reconhecimento facial foi o grande potencial em aplicações no mundo real como se vê na Tabela 2.1. Esse progresso é devido à combinação dos fatores: desenvolvimento ativo de algoritmos; disponibilidade de grandes banco de dados de faces; e métodos para avaliar a performance dos algoritmos desenvolvidos.

O primeiro sistema de reconhecimento automático facial para imagens 2D foi desenvolvido por Takeo Kanade em sua tese de doutorado de 1973, ele projetou e implementou um programa para reconhecimento de faces em um sistema de computadores feito para esse propósito. O algoritmo extrai 16 parâmetros faciais automaticamente, então, em sua tese, Kenade compara

o método de extração automática com a manual, mostrando uma pequena diferença nisso. Ele conseguiu uma taxa de identificação entre 45% – 75%. Logo após houve um período sem muitos resultados nessa área, até o trabalho de (Sirovich & Kirby, 1987), no qual foi proposto um sistema de reconhecimento baseado no redimensionamento da imagem através da técnica de Análise de Componentes Principais (PCA). Foi o trabalho pioneiro de (Turk & Pentland, 1991) sobre auto-face que consolidou as pesquisas em reconhecimento de faces. Outros marcos principais nessa área foram o método *Fisherface* (Etemad & Chellapa, 1996) que aplica Análise de Discriminantes Lineares (LDA), depois do PCA para obter melhor acurácia, e o uso de filtros locais como o Gabor (Lades et al., 1993), para melhorar as características faciais discriminativas (Li & Janil, 2011).

A seguir encontra-se a descrição de um sistema de reconhecimento de faces.

Tabela 2.1: Algumas Aplicações do Reconhecimento Facial

ÁREAS	APLICAÇÕES
Segurança de Informações	Acesso de Segurança (Sistemas Operacionais, banco de dados), Privacidade de Dados (registros médicos), Autenticação de Usuário (trocas, transações on-line).
Biometria	Identificação Pessoal (carteira de identidade, passaporte, título de eleitor, etc.), Sistemas Automáticos de Identificação (controle de fronteiras).
Aplicações Policiais	Vídeos de Segurança, Identificação de Suspeitos, Rastreamento de Suspeitos, Simuladores de Envelhecimento, Reconstrução Forense de Partes da Face.
Segurança Pessoal	Sistemas de Segurança Caseiros, Interpretador de Expressões (sistema de monitoramento de motoristas).
Entretenimento	Sistemas de Vídeo Games, Aplicações com Câmeras.

2.2 Sistemas para Reconhecimento Facial

Uma formulação geral do problema pode ser dada da seguinte forma: dada uma imagem ou um frame (quadros) de um vídeo, identificar ou verificar a identidade de uma ou mais pessoas na imagem, utilizando bancos de dados disponíveis. Caso acessíveis, informações como etnias, idade, gênero e expressão de rosto podem ser usadas para limitar a procura.

Um sistema biométrico para faces opera nos modos de verificação e identificação. A ideia geral desses modos já foi previamente discutida no Capítulo 1. A performance desse sistema está diretamente relacionada à variação de fatores como iluminação, pose, expressão, idade, cabelo, objetos no rosto (boné, máscaras, etc.) e movimento. Baseado nesses fatos, as aplicações de reconhecimento de faces são divididas em duas categorias, a cooperativa e a não-cooperativa.

O caso cooperativo é encontrado em aplicações que envolvem controle de acesso físico ou remoto, nos quais há a cooperação do usuário para utilizar o sistema, por exemplo, apresentar o rosto de maneira adequada para a captura da imagem (de frente com expressão neutra e de olhos abertos). Nos meios não cooperativos, que são típicos de aplicações de segurança, o

indivíduo não está ciente de estar sendo identificado. Em termos de distância entre o rosto e a câmera, a identificação em meios cooperativos não apresenta tantos desafios quanto a de não-cooperativos. Aplicações que misturam ambas abordagens podem também ser encontradas, em sistemas de segurança a distância, em que o usuário pretende ser cooperativo, porém, não conseguir mostrar seu rosto em um ângulo colaborativo à câmera é um exemplo disso. Isso gera grandes desafios para o reconhecimento na maioria dos casos, a ambientação é um dos fatores mais desafiadores nessas aplicações (Li & Janil, 2011).

2.2.1 Módulos

Os sistemas de reconhecimento facial podem ser compostos por quatro módulos, o de localização (detecção) da face, normalização, extração e o de classificação, Figura 2.1. Algoritmos que consistem de todos os módulos são chamados de automáticos, enquanto aqueles que possuem apenas os dois últimos são os parcialmente automáticos. Em algoritmos parcialmente automáticos são dadas a imagem e as coordenadas do centro dos olhos enquanto que para automáticos basta apenas a imagem da face (Tolba et al., 2005).

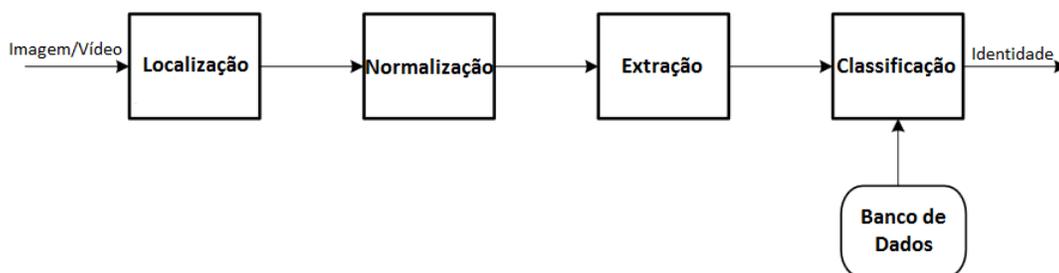


Figura 2.1: Ordem de processamento de um sistema para faces.

A localização facial divide a região do rosto em relação ao fundo da imagem. No caso de vídeo, o rosto detectado precisa ser localizado através de múltiplos frames do vídeo, utilizando um dispositivo próprio para isso. A detecção de faces pode não fornecer bons resultados do local e da escala do rosto na imagem, e por isso alguns módulos vêm acoplados com dispositivos para encontrar os “landmarkings” (pontos de referência) facilitando essas tarefas. Os pontos de referência de uma face são os olhos, nariz, boca e o contorno facial.

A normalização é realizada para ajustar o rosto geométrica e fotometricamente. Este módulo é necessário, pois, nos métodos espera-se robustez quanto a variação de iluminação e pose nas imagens. A normalização geométrica redimensiona a imagem para um tamanho padrão aplicado em todo o banco. A normalização fotométrica normaliza a face, baseando-se em propriedades como iluminação e escalas de cinza.

Os módulos de extração e classificação funcionam da mesma forma para qualquer sistema biométrico.

2.2.2 Tipos de Algoritmos

Os algoritmos para reconhecimento podem ser divididos em duas categorias: global e baseada em componentes locais. Em aproximações globais, também chamadas de *holísticas*, um único vetor de características que representa toda a face capturada pelo sensor é usada para a classificação, enquanto que sistemas de componentes locais detectam independentemente regiões específicas da imagem para a classificação. Muitos métodos foram propostos na literatura. Os descritos na introdução, PCA e LDA, são exemplos de métodos globais. Como métodos locais

pode-se citar o métodos de Padrões Binários Locais (*LBP*) e métodos que utilizam Redes Neurais.

Aproximações globais funcionam bem na classificação de imagens frontais da face, no entanto, elas não são tão robustas em relação à mudanças de pose, pois as características globais são sensíveis a translações e rotações da face.

Existem também métodos derivados da composição dessas duas aproximações, chamados de métodos híbridos. Essa ideia vem de como o sistema de visão humana percebe ambas as aproximações holística e a de componente local. O fator chave, que influencia a performance desse método híbrido, inclui em como determinar quais características devem ser combinadas e como se deve combiná-las, além da maneira de serem utilizados para que suas vantagens sejam consideradas e as desvantagens desconsideradas.

2.3 Fatores que Afetam o Reconhecimento de Faces

Mesmo com as vantagens mencionadas anteriormente, a biometria facial apresenta algumas limitações causadas pela estrutura das imagens de uma face humana. Alguns dos desafios são (Chihaoui, Elkefi, Bellil & Amar, 2016):

- **variações na iluminação**, mudanças causadas pela intensidade da luz nas imagens tem um grande impacto durante o processo de identificação. Como já foi dito, vários algoritmos possuem essa limitação e a essencialidade de novos métodos fizeram esse fato ser levado em conta;
- **expressões faciais**, variações faciais, tais como, expressões faciais, olhos e bocas fechadas, modificam a geometria e textura da face e, conseqüentemente, a acurácia de reconhecimento. Métodos locais, usando um histograma de características, vêm sendo usados com sucesso com o objetivo de superar problemas relacionados a expressões faciais;
- **idade**, a textura e o formato do rosto variam de acordo com a idade. O formato do crânio e da pele mudam da infância à adolescência, e isso representa um problema para o reconhecimento, pois as imagens usadas em passaportes e documentos de identificação não são frequentemente atualizadas;
- **mudança de resolução e escala de imagens**, mudança de escala é um desafio a esses modelos de reconhecimento. Em um sistema de monitoramento isso deveria funcionar bem em múltiplas escalas, pois os indivíduos podem estar em diferentes distâncias da câmera, por exemplo, uma pessoa localizada a dois metros da câmera tem cerca de dez vezes menos mudanças de escalas comparada com alguma à 20 metros. Alguns algoritmos para reconhecimento utilizam métodos de interpolação para redimensionar as imagens de acordo com alguma escala padrão;
- **mudança de pose**, variação de pose se refere principalmente a rotação em relação ao plano. É um desafio aos sistemas de reconhecimento devido a natureza das imagens 2D e 3D de faces. A diferença nas imagens causada pela troca de poses é, na maioria das vezes, maior que a diferença inter-classe das imagens. Em aplicações, como o controle de passaporte, as imagens são colocadas em um banco de dados. No entanto, em ambientes não controlados, como o monitoramento não invasivo, o indivíduo pode estar localizado em cima, em baixo, à esquerda ou à direita da câmera, causando uma rotação na imagem. Aproximações locais são mais robustas com variações de pose que as aproximações

holísticas, no entanto, essa tolerância de métodos locais é limitada a apenas pequenas rotações;

- **oclusão**, o uso de acessórios, tais como, óculos, cachecóis, chapéis, impedem parcialmente a visão de regiões da face. O próprio movimento da pessoa, como o movimento da mão, cria uma oclusão durante a extração de características, podendo perder ou substituir as informações a serem extraídas. Métodos locais também vêm se mostrando mais aptos para lidar com esse tipo de situação;
- **falsificação de imagens**, existem sistemas de reconhecimento que podem ser facilmente burlados por imagens de face apresentadas ao leitor. Por exemplo, dispositivos de desbloqueio de celulares baseados na identificação do rosto podem ser facilmente “enganados” por imagens disponíveis até mesmo nas redes sociais;
- **ruído**, em geral, o ruído é causado por limitações técnicas, “hardware” e “software”, do aparelho de captação de imagem;
- **borrão**, movimentos são a fonte principal dos borrões nas imagens. Esse efeito pode ser causado pelo movimento da pessoa durante a captura da imagem.

A Tabela 2.2 mostra a comparação de alguns fatores descritos anteriormente com relação aos métodos locais e globais.

Tabela 2.2: Algumas Comparações entre Características Globais e Locais. Fonte (Meethongjan & Mohamad, 2007)

Fatores de Variação	Características Locais	Características Globais
Poucas Variáveis	<i>Sem sensibilidade</i>	<i>Sensível</i>
Muitas Variáveis	<i>Sensível</i>	<i>Muito sensível</i>
Iluminação	<i>Muito Sensível</i>	<i>Sensível</i>
Expressões	<i>Sem sensibilidade</i>	<i>Sensível</i>
Pose	<i>Sensível</i>	<i>Muito sensível</i>
Ruído	<i>Muito sensível</i>	<i>Sensível</i>
Oclusão	<i>Sem sensibilidade</i>	<i>Muito sensível</i>

2.4 Revisão da Literatura para Identificação de Faces

Algumas das técnicas mais difundidas para o reconhecimento facial são versões baseadas ou estendidas da técnica de Análise de Componentes Locais (PCA) e da técnica de Padrões Binários Locais (LBP). Uma introdução sobre esses métodos encontra-se a seguir.

a) Análise de Componentes Locais (auto-faces)

O PCA é um método estatístico baseado na análise fatorial de dados. O propósito dessa técnica é o de diminuir a alta dimensão do espaço de dados (variáveis observadas) para a menor dimensão possível, sem que haja perda das informações contidas das amostras de dados (Kim, 2001).

Devido a sua grande aplicação em áreas de reconhecimento de imagens, essa técnica é bastante utilizada para o problema de reconhecimento facial. Em (Huang, Yuen, Lai & Li, 2004) é descrito a ideia de implementação do PCA para o problema de faces.

De forma geral, dada uma imagem I_n de dimensão $m \times n$, concatena-se as linhas da matriz de pixels em um único vetor de dimensão $d = m \times n$. Seja $X = \{X_n \in R^d | n = 1, \dots, N\}$ o conjunto de vetores das imagens de treino. Então X corresponde a um espaço de dimensão $d \times N$, em que N é a quantidade de imagens usadas para treinamento. O PCA tenta redimensionar o espaço X encontrando um espaço de dimensão menor que descreve o espaço original. Considere

$$E(X) = \frac{1}{N} \sum_{n=1}^N X_n$$

como sendo o vetor médio das imagens de treino. Após subtrair esse vetor de todos os elementos em X , encontra-se um conjunto de dados modificados, dado por:

$$\bar{X} = \{\bar{X}_n \in R^d | n = 1, \dots, N\}, \quad \bar{X} = X_n - E(X).$$

Para gerar esse espaço é necessário utilizar a matriz de covariância dos dados em \bar{X} . A matriz de covariância do conjunto \bar{X} é então, definida, por:

$$M = cov(\bar{X}) = E(\bar{X} \cdot \bar{X}),$$

em que M é uma matriz $d \times d$. Os autovalores da matriz M formam uma base ortonormal em R^d . No método PCA, o vetor de uma face y , relacionada com o conjunto X , é obtido através da projeção de y no subespaço gerado pelos k autovetores correspondentes aos maiores k autovalores da matriz de covariância em ordem decrescente, em que k é menor que N . Essa projeção resulta em um vetor contendo k coeficientes a_1, \dots, a_k , a face y é então representado pela combinação linear desses autovetores com pesos a_1, \dots, a_k .

b) Padrões Binários Locais (LBP)

Existem vários métodos para extrair as características discriminativas de uma imagem, um desse métodos é o *LBP*. O método foi introduzido em (Ojala, Pietikäinen & Hardwood, 1996). Com o *LBP* é possível diferenciar texturas em uma imagem digital. Isso é feito através da divisão da imagem em diversas regiões pequenas, das quais os fatores discriminantes são extraídos (Figura 2.2).

O operador *LBP* é aplicada em regiões 3×3 de uma imagem, sendo o valor do pixel localizado no centro dessa região, o ponto de referência. Se os pixels da vizinhança possuem níveis de cinza maiores que o central (ou o mesmo valor), então o valor 1 é designado para esses



Figura 2.2: Exemplo de imagem pré-processada dividida em 64 regiões. Fonte:(Rahim, Hossain, Wahid & Azam, 2013)

pixels, caso contrário é designado o valor 0. O padrão *LBP* do pixel central é então definido pela Equação (2.1):

$$LBP = \sum_{n=0}^7 s(g_n - g_c) \times 2^n, \quad (2.1)$$

na qual g_c é o valor do pixel central, g_n os da vizinhança e s é a função dada pela Equação (3.11a):

$$s(x) = \begin{cases} 1 & x \geq 0, \\ 0 & x < 0. \end{cases} \quad (2.2)$$

A Figura 2.3 ilustra o cálculo desses padrões para esse caso.

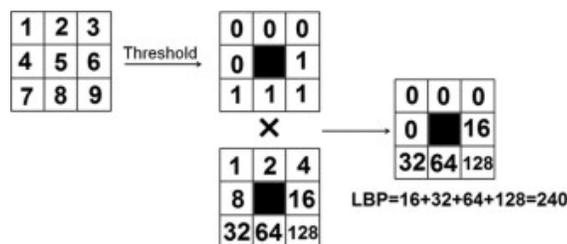


Figura 2.3: Ilustração do Cálculo do LBP de uma Região. Fonte (Jia et al., 2014))

Esses fatores consistem dos padrões que descrevem os arredores dos pixels. As características obtidas são concatenadas em um único histograma, o qual representa a imagem. As imagens podem, então, ser comparadas através da medida da similaridade (distância) entre seus respectivos histogramas.

Apesar de terem sido propostos para a análise de texturas, métodos baseados em *LBP* vem obtendo bons resultados no reconhecimento facial, tanto em termos de velocidade quanto de performance. Esse método se mostrou robusto ao tratar imagens com diferentes expressões faciais, diferentes condições de iluminações, rotação de imagem e envelhecimento (Rahim et al., 2013).

A seguir encontram-se algumas técnicas disponíveis na literatura de reconhecimento facial, baseadas nas técnicas PCA, LDA, LBP e Redes Neurais.

2.4.1 Técnicas Para o Reconhecimento Facial

Métodos de extração de característica, como o PCA e o LBP podem ser afetados pelas condições de iluminação das imagens, alterações de expressões faciais e outros fatores. Assim, antes de extrair as características pode-se pré-processar as imagens para aumentar a taxa de reconhecimento de faces (H. Zhang, 2011).

Em (Barnouti, Matti, Al-Dabbagh & Naser, 2016) é proposta uma técnica baseada na detecção de imagens da face através da Técnica Viola-Jones (Wang, 2014) para segmentação da face e, então, é aplicado o PCA para extração de características e na sequência a técnica de Análise de Discriminantes Lineares (LDA) (Chelali, Djeradi & Djeradi, 2009) é utilizada para redução na quantidade de características extraídas.

Como alternativa a técnica PCA, em (Jalled, 2017) é feito um comparativo do PCA com sua versão normalizada, chamada de N-PCA. O autor afirma que a versão normalizada obtém melhores resultados em relação ao PCA tradicional.

Na maioria dos casos de aplicações o LDA clássico gera amostras cuja dimensão é muito maior que a quantidade de amostras obtidas. Em (T. P. Zhang, Fang, Tang, Shang & B. Xu, 2010) é proposta a Análise de Discriminantes Exponenciais (EDA) para tratar esse problema. Em (Ran, Fang & Zhang, 2018) é proposta uma generalização da técnica EDA que se mostrou mais vantajosa.

Várias variações do *LBP* vem sendo propostas para o problema de reconhecimento facial. Em (Ferraz, Manzato & Gonzaga, 2017) é investigado o comportamento de uma dessas técnicas, o método de Padrões Médios Mapeados Localmente (M-LMP). O Padrão Local Binário Probabilístico (LBPP) proposto em (Dahmouni, Aharrane, Moutaoukil & Satori, 2016) usa o conceito de intervalo de confiança para avaliar os pixels.

Em (Liu, Fieguth, G., Pietikäinen & Hu, 2016) são apresentados descritores semelhantes ao *LBP* baseados no acúmulo local das diferenças dos pixels, tais como diferenças com relação a raios e ângulos, nas quais são decompostas em componentes complementares de sinais e magnitudes. Existem outras variações do *LBP* que vem sendo propostas com o objetivo de melhorar a capacidade discriminativa dos vetores de características, tais como o Padrão Binário Local de Blocos em Multiescala (MB-LBP) (Liao, Zhu, Lei, Zhang & Li, 2007; Schlett, Rathgeb & Busch, 2016).

Além de métodos baseados em PCA e LBP existem também métodos baseados em Redes Neurais. Em (Kasar, Bhattacharyya & Kim, 2016) é feita uma revisão de várias técnicas que utilizam Redes Neurais. No trabalho de (Peng, Wang, Long & Lu, 2015) é proposta uma extensão de métodos usuais de Redes Neurais, a abordagem proposta por esses autores apresenta bons resultados comparados à alguns métodos usuais.

Para análise de performance de métodos utilizados no reconhecimento facial existem algumas bases de dados disponíveis na literatura científica. Algumas delas são apresentadas na próxima seção.

2.5 Bancos de Dados de Imagens 2D

Muitos banco de dados (públicos ou privados) encontram-se disponíveis para propósitos de pesquisa na área de reconhecimento facial. Esses bancos de dados diferem entre si em diversos aspectos. Para (Huang et al., 2004) os aspectos mais interessantes são:

- número de imagens;

- número de indivíduos por classe. Sabendo que cada indivíduo está designado a uma única classe, e o número de imagens por classe representa o número de imagens representativas nela contida. Essas imagens são adquiridas sob diferentes condições de ambientes (orientação, expressão facial, etc.);
- tamanho das imagens;
- pose e orientação;
- mudança de iluminação;
- sexo das pessoas inscritas;
- presença de objetos (óculos, barba, etc.);
- presença de ruído nas imagens;
- presença de um plano de fundo uniforme;
- tempo decorrido entre uma foto e outra.

É recomendável que se escolha bancos de dados apropriados durante os testes de um algoritmo. Alguns bancos possuem um protocolo bem definido, permitindo uma comparação direta dos resultados. Além do mais, a escolha depende da característica a ser analisada: iluminação; envelhecimento; expressões faciais; e etc. A disponibilidade de uma alta quantidade de imagens por pessoa pode ser um fator decisivo para a análise da acurácia do método utilizado.

A seguir são apresentados alguns bancos de dados disponíveis na literatura e que foram utilizadas neste trabalho.

2.5.1 Banco de Dados ESSEX

Desenvolvido pelo Dr. Libor Spacek, professor da universidade de Essex, em 2008, esse banco de dados, disponível para acesso grátis em <http://cswww.essex.ac.uk/mv/allfaces/index.html>, tem como objetivo o encorajamento da pesquisa biométrica comparativa. Comparação de processos biométricos e publicações gerais de resultados são seus principais propósitos. As imagens estão contidas em quatro pastas, *faces94*, *faces95*, *faces96* e *grimace*. Com relação a dificuldade de reconhecimento automático, as pastas *faces96* e *grimace* são as mais difíceis por dois motivos: variação do fundo e das escalas e variação extrema de expressões. A Tabela 2.3 fornece uma visão geral desse banco de dados.

A seguir tem-se uma descrição sobre cada pasta do banco de dados ESSEX.

a) Faces94

Durante a aquisição das imagens desta pasta, os indivíduos se posicionaram a uma distância fixa da câmera e foi solicitado que eles falassem enquanto uma sequência de imagens fosse obtida, com o propósito de introduzir variação nas expressões faciais. A Figura 2.4 mostra algumas imagens desse banco de dados e a Tabela 2.4 mostra os dados das gerais desse bando de dados.

Tabela 2.3: Versão Geral das Imagens ESSEX.

Número Total de Indivíduos	395
Número Total de Imagens Por Indivíduos	20
Número Total de Imagens	7900
Gênero	Contém Imagens de Pessoas de Sexo Masculino e Feminino
Etnias	Contém Imagens de Pessoas de Várias Etnias
Idade de Alcance	18-20
Óculos	Sim
Barba	Sim
Formato da imagem	24bits, JPEG
Câmera Usada	S-VHS
Iluminação	Artificial, Mistura de Tungstênio e Fluorescente



Figura 2.4: Exemplos das Imagens da Faces94.

Tabela 2.4: Visão Geral da Faces94.

Número de Indivíduos	153
Resolução da Imagem	180 × 200
Diretórios	Contém Imagens de Homens(133) e Mulheres(20) Em Diferentes Diretórios
Plano de Fundo	É Uniforme e Verde
Movimentos com a Cabeça	Mínimos
Posição do Rosto na Imagem	Mínimas Mudanças
Variação de Iluminação	Nenhuma
Variação de Expressão	Mudanças Consideráveis
Comentário Adicional	Não Há Variação de Penteados

b) Faces95

Para as imagens contidas nesta pasta, foi obtida uma sequência de 20 imagens por classe, usando uma câmera fixa. Durante as fotos as pessoas deram um passo em direção a câmera, esse movimento foi usado para variar a escala do rosto do mesmo indivíduo. O tempo de obtenção entre uma imagem e outra foi de 0,5 segundos. A Figura 2.5 mostra algumas imagens desse banco de dados e a Tabela 2.5 contém informações sobre as imagens desse diretório.



Figura 2.5: Exemplos das Imagens da Faces95.

Tabela 2.5: Informações Sobre as Imagens da Faces95.

Número de Indivíduos	72
Resolução da Imagem	180 × 200
Sexo	Contém Imagens de Homens e de Mulheres
Plano de Fundo	Cortina Vermelha e Sombras Causadas pelo Movimento
Movimentos com a Cabeça	Mínimos
Posição do Rosto na Imagem	Algumas Translações
Variação de Iluminação	Quanto Mais Próximos Mais Há Iluminação
Variação de Expressão	Mudanças Consideráveis
Comentário Adicional	Não Há Variação de Penteados

c) Faces96

As imagens dessa pasta foram obtidas nas mesmas condições da pasta Faces95. A Figura 2.6 mostra algumas imagens desse banco de dados e a Tabela 2.6 contém informações sobre esse diretório.

d) Grimace

Uma sequência de 20 imagens por pessoa foi obtida usando uma câmera fixa, para as imagens contidas nesta pasta. Durante a captura das imagens os indivíduos moviam a cabeça a faziam



Figura 2.6: Exemplos das Imagens da Faces96.

Tabela 2.6: Informações Sobre as Imagens da Faces96.

Número de Indivíduos	152
Resolução da Imagem	196×196
Sexo	Contém Imagens de Homens e de Mulheres
Plano de Fundo	Cartazes Enfeitados e Sombras Causadas pelo Movimento
Movimentos com a Cabeça	Mínimos
Posição do Rosto na Imagem	Algumas Translações
Variação de Iluminação	Quanto Mais Próximos Mais Há Iluminação
Variação de Expressão	Mudanças Consideráveis
Comentário Adicional	Não Há Variação de Penteados

caretas, que ficaram mais extremas nas últimas poses. O tempo entre a captura de uma imagem e outra foi de 0,5 segundos. A Figura 2.7 mostra algumas imagens desse banco de dados e a Tabela 2.7 contém informações sobre as imagens desse diretório.



Figura 2.7: Exemplos das Imagens da Grimace.

Tabela 2.7: Informações Sobre as Imagens da Grimace.

Número de Indivíduos	18
Resolução da Imagem	180×200
Sexo	Contém Imagens de Homens e de Mulheres
Plano de Fundo	O Fundo é Uniforme
Movimentos com a Cabeça	Mínimos
Posição do Rosto na Imagem	Algumas Translações
Variação de Iluminação	Muito Pouca
Variação de Expressão	Alta
Comentário Adicional	Não Há Variação de Penteados

2.5.2 Banco de Dados *Japanese Female Facial Expression (JAFFE)*

As imagens do JAFFE foram coletadas por (Lyons, Akemastu & Gyoba, 1998). As fotografias foram obtidas de 10 modelos japonesas exibindo três ou quatro das seguintes expressões faciais: feliz, triste, surpresa, raiva, desgosto, medo e neutra, num total de 219 imagens de dimensão 256×256 . O cabelo das modelos estava preso, para que todas as regiões do rosto ficassem a mostra, sem perda de nenhuma característica da expressão. A iluminação do ambiente foi escolhida de maneira que as duas partes do rosto ficassem igualmente distribuídas quando fotografadas, foi também usado um material entre a câmera e o cenário de fundo para diminuir os reflexos do ambiente. As fotos foram tiradas em preto e branco e digitalizadas através de um “scanner”. A Figura 2.8 mostra algumas imagens do JAFFE.

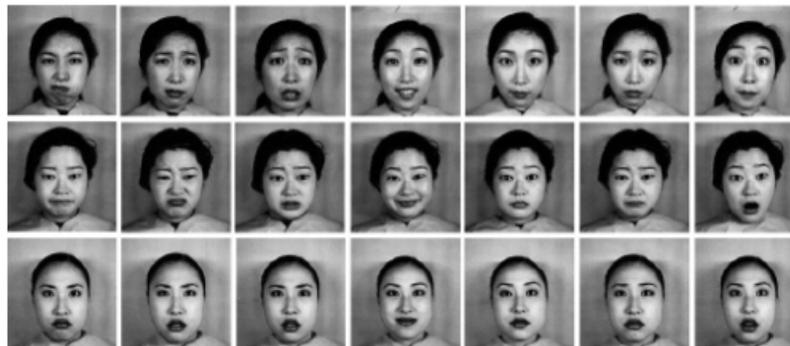


Figura 2.8: Exemplos das Imagens da JAFFE.

2.5.3 A Base de Dados MUCT

Esse banco de dados foi preparado por Stephen Milborrow, John Morkel e Fred Nicolls em dezembro de 2008 na Universidade de Cape Town, com descrição completa em (Milborrow, Morkel & Nicolls, 2010). A base de dados contém 3755 imagens de pessoas encontradas nos arredores da universidade. A população incluiu estudantes, pais de alunos que vieram prestigiar os filhos durante uma colação de grau, professores de ensino médio que estavam em um congresso e funcionários da universidade, como faxineiros e seguranças. Dentre todas as imagens, o

número de mulheres e homens fotografados é quase o mesmo, essa relação se manteve até mesmo com relação a suas idades e etnias.

Para convidar as pessoas para a foto os pesquisadores se aproximavam delas, explicavam a situação e ofereciam uma barra de chocolate como recompensa pela colaboração. As pessoas estavam com maquiagem, óculos e penteados. Durante as sessões não foi solicitado para elas esboçarem nenhuma expressão em particular. Na prática isso fez com que todos fossem fotografados com um expressão neutra ou sorrindo.



Figura 2.9: Exemplos das Imagens da Muct.

Cada indivíduo foi fotografado por cinco câmeras organizadas da maneira como está mostrado na Figura 2.10. A ideia era tirar as cinco fotos simultaneamente para deixar o banco de dados útil para avaliar a mesma pessoa em diferentes ângulos, mas devido ao atraso de cada “software” houve uma pequena diferença entre elas. As pessoas sentaram de frente para a câmera, mas devido a variação da altura entre as pessoas, ocorreram variações de orientação nas imagens.

Dez tipos de iluminações foram usadas e cada tipo foi representado, neste trabalho, por uma letra do alfabeto de q a z. Cada indivíduo foi fotografado utilizando 2 ou 3 tipos de iluminações. Como pessoas diferentes foram fotografadas, de modo geral, com diferentes tipos de iluminação, resultou em um banco de dados bastante diversificado, em questão de iluminação e ângulos de captura. A Figura 2.11 mostra a diferença entre as fotografias obtidas por cinco “webcams” e de três tipos de iluminação, q, r e s.

As imagens tem formato RGB de dimensão 640×480 . A Tabela 2.8 mostra a quantidade de imagens, a localização delas e o tipo de iluminação com a qual cada pessoa foi fotografada.

Tabela 2.8: Distribuição das Iluminações nas Classes das Imagens.

Iluminações	Posição das Imagens	N. Pessoas	N. Imagens
q r s	000 – 090	91	1365
t u v	200 – 307	108	1620
w x	400 – 451	52	520
y z	600 – 624	25	250
Total		276	3755

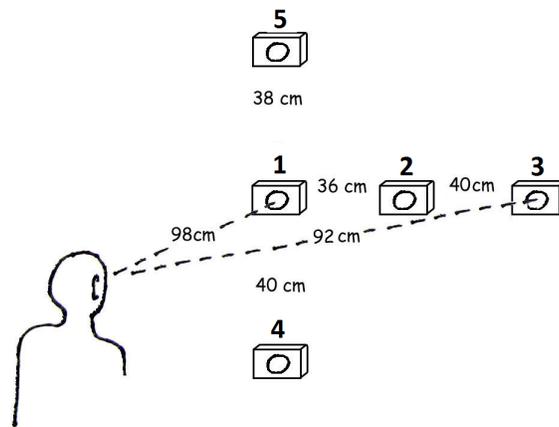


Figura 2.10: Posição das “Webcams”.



Figura 2.11: Imagens das cinco “Webcams” e das Iluminações q, r e s.

O acesso ao diretório das imagens está disponível em <http://www.milbo.org/muct/>. Nota-se, pela descrição da base, que ela apresenta alta variação de iluminação e poses, e como foi dito anteriormente, a maioria dos métodos para reconhecimento facial não são 100% robustos quanto a esses fatores e por isso esse banco de dados, em geral, apresenta grandes dificuldades para o processo de reconhecimento automático.

2.5.4 O Banco de Dados ORL

Esse banco de dados foi desenvolvido pelo Laboratório de Computadores da Universidade de Cambridge e contém um conjunto de imagens fotografadas entre 1992 e 1994.

Ele possui dez imagens de 40 pessoas, sendo assim composta por 400 imagens faciais. Para alguns indivíduos, as imagens foram obtidas em tempos diferentes, com variação de iluminação, expressões faciais (olhos/boca fechados, sorrindo ou não) e objetos no rosto (óculos). O fundo delas é uma região uniforme preta, e as pessoas ficaram com postura reta de frente para as câmeras, com tolerância para pequenos movimentos laterais. A Figura 2.12 mostra algumas das imagens desse banco de dados.



Figura 2.12: Imagens do ORL.

Os arquivos estão em formato PGM e são de dimensão 92×112 , com 256 escalas de cinza. Eles estão organizados em 40 diretórios (um para cada pessoa).

2.5.5 O Banco de Dados YALE

O banco de dados YALE, como o próprio nome já diz, foi desenvolvido pela universidade Yale. Ele contém 11 imagens de 15 pessoas (14 homens e uma mulher) totalizando 165 imagens em 256 escalas de cinza, de dimensões 320×243 .

Esse banco de dados é caracterizado pelas variações de iluminação e de expressão. Não há alteração de ângulo de captura. A Figura 2.13 exemplifica uma classe desse banco.

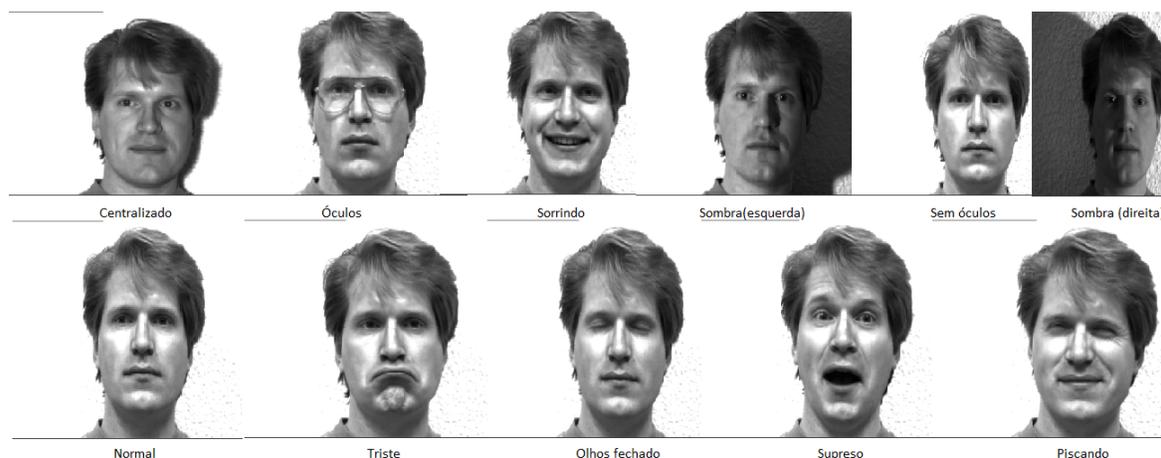


Figura 2.13: Exemplo das Expressões e Iluminação de uma Classe da YALE.

Essa base, em geral, apresenta dificuldades durante o processo de reconhecimento devido a grande variação de iluminação e à alta diversidade de expressões, como foi descrito anteriormente. O fundo branco está presente em quase todas as imagens dessa base.

2.6 Considerações Finais

Neste capítulo foram apresentados os sistemas biométricos baseados em reconhecimento facial. Além disso, foram apresentadas algumas técnicas presentes na literatura, dentre elas, a técnica LBP que é a técnica de características locais usada pra desenvolver o método MSLMP utilizado neste trabalho. Foram ainda mostrados alguns bancos de dados faciais disponíveis na literatura científica especializada.

Técnicas Utilizadas

3.1 Introdução

Este capítulo é dedicado às técnicas utilizadas neste trabalho, ou seja, são discutidos aqui o método dos Padrões Locais Mapeados em Multiescala (*MSLMP*), os algoritmos genéticos (AG's) e a forma de extração de características. São apresentadas também duas novas técnicas de pré-processamento de imagens baseadas nos níveis de cinza dos pixels, com o objetivo de melhorar a qualidade das imagens para posterior processamento.

O método *MSLMP* tem sua construção baseada no método dos Padrões Binários Locais em Multiescala (*MSLBP*), que é uma versão multiescalar dos métodos *LBP* e Padrões Localmente Mapeados (*LMP*), que por sua vez é também uma extensão do método *LBP*. Cada um desses métodos encontram-se descritos na sequência.

3.2 Padrões Binários Locais em Multiescala

O quadro das aproximações multiescalares, chamado de análise de multi resoluções, foi desenvolvido pelas áreas de Visão Computacional, Análise de Imagens e Processamento de Sinais com motivações complementares da física e da biologia. A ideia de ter uma representação multiescala de uma imagem veio da observação básica da vida de que objetos são compostos por estruturas diferentes em diferentes escalas (Chan, 2008).

Em (Mäenpää & Pietikäinen, 2003) é sugerido o uso de mais de uma extração do *LBP* em uma imagem de modo a variar o tamanho das regiões (R) ou a quantidade de amostras (P). Os autores justificaram isso pelo fato de uma das maiores limitações do *LBP* tradicional ser a pequena área tomada para os cálculos (região 3×3). De acordo com esses autores, isso não torna o operador robusto com relação a mudanças locais na textura da imagem causadas, por exemplo, pela rotação ou variação de iluminação.

Dessa forma, o método *LBP* foi estendido para uma versão multiescalar caracterizada pela mudança do tamanho do raio (R) em cada extração. Pode-se transformar a Equação (2.1) para essas mudanças através da Equação (3.1):

$$LBP_{P,R} = \sum_{n=0}^{P-1} s(g_n - g_c) \times 2^n, \quad (3.1)$$

na qual P é o número de amostras (pixels g_0, g_1, \dots, g_{P-1}) igualmente espaçados a uma distância R do pixel central g_c , contidos em sua vizinhança. A Figura 3.1 mostra algumas regiões através dessa variação de dados.

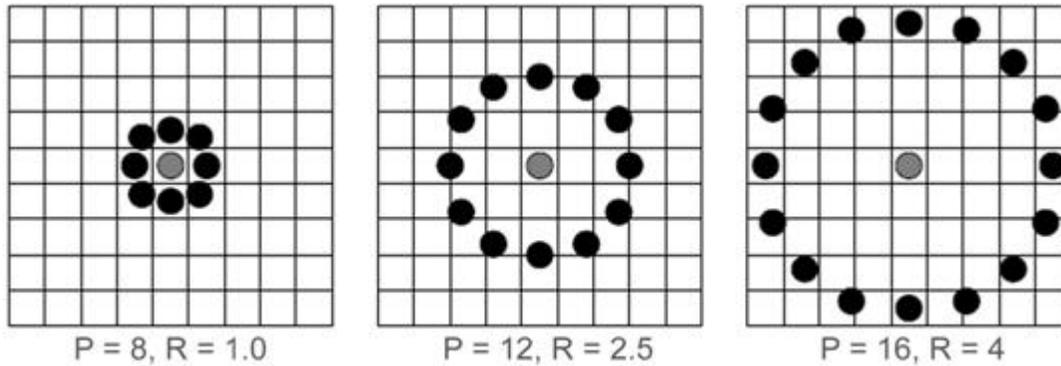


Figura 3.1: Região Circular para Três Diferentes Valores P, R . Fonte (Rahim, Hossain, Wahid & Azam, 2013))

Jia et al., 2014 apresentaram duas variações do método Padrões Binários Locais em Multiescala (*MSLBP*). O primeiro caso, chamado de *MSLBP1*, usa o *LBP* extraído-se 8 amostras de diversas regiões da imagem, após filtragem gaussiana, definidas por um raio R_n , em que n é a respectiva escala e é denotado por LBP_{8,R_n} . O segundo tipo, chamado de *MSLBP2* é uma maneira alternativa à primeira em que um conjunto de filtros de média é aplicado na região antes da extração do padrão $LBP_{8,1}$, como é mostrado na Figura 3.2.

O método aqui utilizado é inteiramente baseado no *MSLBP1*, em que em cada escala é aplicado um filtro Gaussiano de tamanho G_n , sendo a evolução das escalas dada pela troca dos valores dos raios. Essa mudança ocorre respeitando a distribuição ótima com relação à aplicação de filtros Gaussianos (Mäenpää & Pietikäinen, 2003). A Figura 3.3 exemplifica esse processo, são consideradas vizinhanças de um pixel quantizadas por oito amostras, as quais formam oito setores que distam R_n do pixel central. De acordo com (Mäenpää & Pietikäinen, 2003), o objetivo do uso de filtros Gaussianos de passa-baixa é para que em cada amostra na vizinhança se possa coletar uma quantidade de informação maior que aquela com um único pixel. Os círculos representam a área para extração do padrão *LBP* ao redor de cada amostra após filtragem. O raio externo da “área útil”, mostrada na figura 3.3, r_n , é usado para calcular o raio R_n para a extração do LBP_{P_n,R_n} e é dado por

$$r_n = r_{n-1} \cdot \left(\frac{2}{1 - \sin(\pi/P_n)} - 1 \right), n \in \{2, \dots, N\}, \quad (3.2)$$

na qual N é o número de escalas e P_n a amostragem de cada escala, que no caso do *MSLBP* foi considerado $P_n = 8 \forall n \in \mathbb{N}$.

Os raios r_n estabelecem a distância entre o pixel e a borda de uma vizinhança, assim r_1 é definido como 1,5, que é a menor distância entre o pixel e a borda de uma região 3×3 . Os raios das operações são escolhidos de forma que as “áreas úteis”, da Figura 3.3, estejam em contato. Consequentemente, define-se o raio do operador LBP_{8,R_n} , na escala n ($n \geq 2$), como o ponto médio entre r_n e r_{n-1} , assim:

$$R_n = \frac{r_n + r_{n-1}}{2}, n \in \{2, \dots, N\}. \quad (3.3)$$

134	45	98	225	15	85	43	111	229
234	76	85	63	68	74	64	40	95
4	48	178	38	220	32	36	24	39
65	79	3	92	227	117	90	100	92
34	46	111	86	30	226	25	5	101
244	53	38	126	60	33	119	98	94
189	28	239	52	6	50	45	32	106
245	60	117	92	110	223	32	12	61
89	39	141	63	44	63	78	40	94

(a)



100.2222	91.1111	75.6667
74.7778	110.7778	80.4444
127.4444	78.1111	55.5556

(b)



$$MSLBP2 = \frac{0.2^0 + 0.2^1 + 0.2^2 + 0.2^3 + 0.2^4 + 0.2^5 + 0.2^6 + 0.2^7 + 0.2^8}{64}$$

Figura 3.2: Cálculo do *MSLBP2* de uma matriz 9×9 . a) Matriz 9×9 , b) Média das regiões 3×3 da matriz original, c) Cálculo do *MSLBP2*. Fonte: (Contreras, 2015).

Analogamente ao caso dos r_n , os raios R_n dos operadores representam a distância entre o pixel central e os pixels de sua vizinhança, portanto R_1 é definido como 1 por ser a menor distância entre o pixel central e os de sua vizinhança 3×3 . Os raios r_n são usados para calcular o tamanho da janela dos filtros Gaussianos, G_n , na escala n como mostra a Equação (3.4):

$$G_n = 2 \cdot \text{round} \left(\frac{r_n - r_{n-1}}{2} \right) + 1, \quad (3.4)$$

na qual a função $\text{round} : \mathbb{R} \rightarrow \mathbb{N}^+$ associa a um valor o menor inteiro não negativo maior que ele. Em cada nível, o valor δ_n necessário para aplicar o filtro Gaussiano é dado por:

$$\delta_n = \frac{G_n}{\sqrt{-2 \ln(1-p)}}, \quad (3.5)$$

em que p foi estabelecido como 0,95.

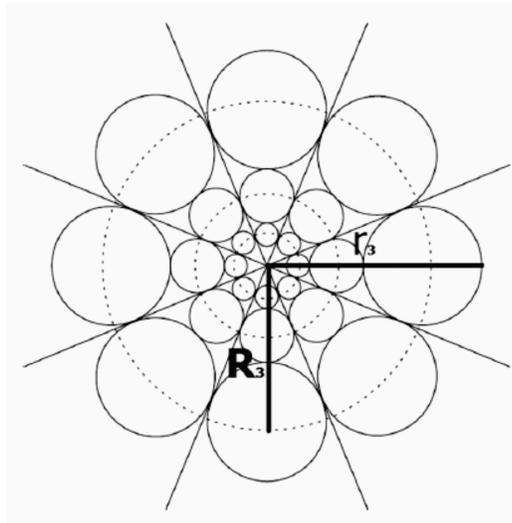


Figura 3.3: Distribuição ótica dos raios do padrão MSLBP com respeito à redução de redundância da informação capturada pelo padrão. Fonte: (Jia et al., 2014).

3.3 Padrões Mapeados Localmente

O *LMP*, que é uma extensão do $LBP_{8,1}$, assume que a distribuição dos níveis de cinza na vizinhança de um pixel é um padrão local e pode ser representada pela diferença dos níveis de cinza da vizinhança do pixel central (Ferraz, Junior & Gonzaga, 2014). Dessa forma, cada pixel $g(i, j)$ é mapeado em uma região $W \times W$ em torno deste ponto (Figura 3.4a), através da Equação (3.6) :

$$h(g(i, j)) = \frac{\sum_{k=1}^W \sum_{l=1}^W (f_{g(i,j)}(g(k, l) \cdot P(k, l)))}{\sum_{k=1}^W \sum_{l=1}^W P(k, l)}, \quad (3.6)$$

na qual $f_{g(i,j)}$ é a função de mapeamento (Figura 3.4b), responsável pela suavização dos níveis de cinza da vizinhança que mapeará os $W^2 - 1$ vizinhos do pixel central $g(i, j)$ (Figura 3.4c), e P é uma matriz de pesos $W \times W$ pré-definida (Figura 3.4d). Este mapeamento define, então, o micro padrão *LMP* da região $W \times W$ (Figura 3.4e).

Através da Equação (3.6) pode-se deduzir outros descritores de padrões presentes na literatura. Por exemplo, o LBP_{P_n, R_n} é obtido ao usar a função de mapeamento como sendo a função degrau de Heavisidade (H), dada pela Equação (3.7):

$$H[A(k, l) - g(i, j)] = \begin{cases} 1 & A(k, l) - g(i, j) > 0, \\ 0 & A(k, l) - g(i, j) \leq 0. \end{cases} \quad (3.7)$$

Nas condições anteriores, ao levar em conta o $LBP_{8,1}$, a matriz de pesos é dada por:

$$P = \begin{bmatrix} 1 & 2 & 4 \\ 8 & 0 & 16 \\ 32 & 64 & 128 \end{bmatrix}.$$

Para que os valores do padrão estejam em $[0, 255]$ é necessário normalizar o resultado obtido.

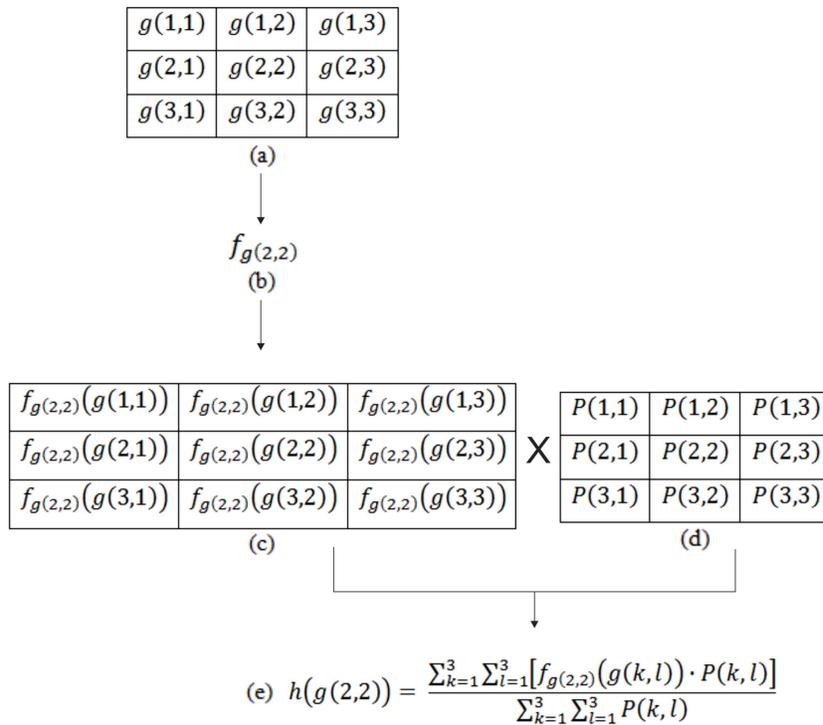


Figura 3.4: Cálculo do mapeamento relativo ao pixel central. (a) Vizinhança 3×3 do pixel central $g(2,2)$, (b) Função de mapeamento, (c) Matriz dos mapeamentos realizados na vizinhança do pixel central, (d) Matriz de pesos, (e) Mapeamento relativo à vizinhança de $g(2,2)$. Fonte : (Contreras, Boaventura & Boaventura, 2015).

Tal normalização é feita através da Equação (3.8):

$$N_{LMP} = \text{round} \left(\underbrace{h(g(i,j))}_{\in[0,1]} \cdot \underbrace{\sum_{k=1}^3 \sum_{l=1}^3 P(k,l)}_{255} \right). \quad (3.8)$$

3.4 Padrões Mapeados Localmente em Multiescala

A presença de ruídos na imagem prejudica a extração do padrão LBP, devido a isso, em (Contreras, Boaventura & Boaventura, 2015) foi proposta uma extensão multiescalar do método *LMP*, chamada Padrões Locais Mapeados em Multiescala (*MSLMP*), e que tem como propósito atenuar a ação ruidosa presente nas imagens, através do uso de aplicações que suavizam as altas frequências presentes nas vizinhanças de um pixel. Tais vizinhanças são representadas por raios que aumentam de acordo com as escalas desejadas, caracterizando, assim o método como multiescalar. A evolução de cada escala se faz de acordo com o aumento do tamanho da dimensão das vizinhanças do pixel $g(i,j)$. Portanto, as dimensões W_n (Equação (3.9)), da matriz de pesos P_n (Equação (3.10)) aumentam de acordo com a escala.

$$W_n = 2 \cdot \text{round}(R_n) + 1, \quad (3.9)$$

$$P_n = \begin{bmatrix} 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & 1 \\ 0 & 0 & & 0 & 0 & 0 & & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & & 0 & 0 & 0 & & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & & 0 & 0 & 0 & & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & & 0 & 0 & 0 & & 0 & 0 \\ 1 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & 1 \end{bmatrix}. \quad (3.10)$$

Este método utiliza a função Sigmoide, dada pela Equação (3.11), como aplicação de mapeamento, pois a suavidade da função Sigmoide transmite uma abordagem mais realística ao padrão *LMP*, o que vai atenuar a perda de informação sofrida pela presença de ruído na imagem.

$$f_{g(i,j)}(A(k,l)) = \frac{1}{1 + e^{-\frac{A(k,l) - g(i,j)}{\lambda}}}, \lambda \in (0, 1]. \quad (3.11)$$

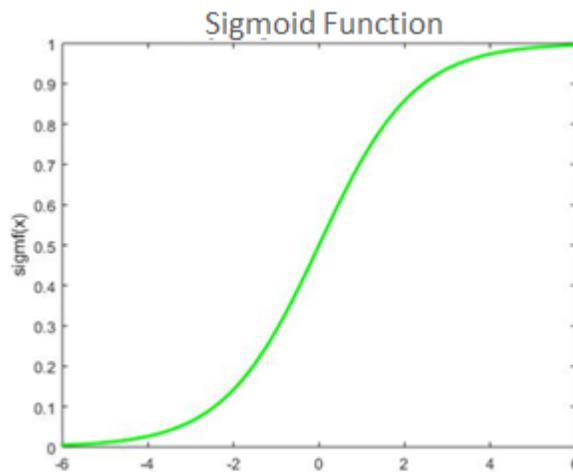


Figura 3.5: Gráfico da função sigmoide gerado pelo MATLAB.

Como $f_{g(i,j)}(g(i,j)) = 0,5$ seja qual for o valor de $g(i,j)$, o valor central de P foi considerado igual a 0, evitando esse cálculo e melhorando a performance do método. Nota-se que as matrizes P_n foram construídas de forma a considerar apenas oito amostras de cada região do pixel para o cálculo do *MSLMP*. Portanto, o padrão *MSLMP* de um pixel $g(i,j)$ na n -ésima escala é obtido pela Equação (3.12):

$$MSLMP(g(i,j)) = \frac{\sum_{(k,l) \in I_n - \{(i,j)\}} (f_{g(i,j)}(g(k,l)))}{8}, \quad (3.12)$$

na qual, $I_n = \{(k,l) : k = i - C, i - 1, i + C - 1, i + C \text{ e } l = j - (C - 1), j + 1, j - 1, j + C - 1\}$, em que $C = \text{round}(R_n)$.

Como os valores gerados pela função Sigmoide estão entre 0 e 1, o padrão *MSLMP* está contido no intervalo (0, 1). Portanto, para obter valores em [0, 255] é usada a versão normalizada

dos valores obtidos pela Equação (3.12), assim, o padrão final extraído de cada pixel $g(i, j)$ é dado pela Equação (3.13):

$$N_{MSLMP}(g(i, j)) = \text{round}(255 \cdot MSLMP(g(i, j))). \quad (3.13)$$

3.5 Extração de características

Pela Equação (3.13) nota-se que os padrões normalizados MSLMP são valores no intervalo $[0, 255]$. Em Ojala et al., 2002 são apresentados alguns valores dentro desse intervalo denominados “uniform patterns”. Esses valores representam uma classe que é responsável, na maioria das vezes, por mais de 90% das características mais representativas da imagem. De maneira geral, o “uniform pattern” é um valor $U \in [0, 255] \cup \mathbb{Z}$, que em sua representação binária $(b_1 b_2 b_3 b_4 b_5 b_6 b_7 b_8)$, com $b_i \in \{0, 1\}$, respeita a desigualdade (3.14):

$$|b_1 - b_8| + \sum_{i=2}^8 |b_i - b_{i-1}| \leq 2. \quad (3.14)$$

De acordo com (Jia et al., 2014), dadas K amostras da região de um pixel, existem $(K - 1) \cdot K + 2$ “uniform patterns”. Para $K = 8$, no método em questão, tem-se 58 valores entre 0 e 255 que satisfazem a relação (3.14).

Esses “Uniform Patterns” são então utilizados como as características principais de todo o processo. A seguir encontra-se descrito o algoritmo completo para a implementação do método *MSLMP*.

3.6 Algoritmo para Implementação do MSLMP

Com base nos conceitos apresentados nas seções anteriores (Contreras, 2015) definiu o seguinte algoritmo para a extração dos padrões *MSLMP* de uma imagem:

1. para cada pixel $g_m(i, j)$ da imagem g_m , calcula-se o padrão $N_{LMP}(g(i, j))$ descrito na Equação (3.13);
2. constrói-se um histograma intermediário formado pelos valores calculados em 1;
3. tomam-se as taxas de ocorrência dos 58 “uniform patterns” presentes no histograma construído em 2, seguidas da soma das taxas restantes, para formar um vetor x_m^n de 59 coordenadas;
4. atualiza-se o valor de P_n conforme (3.9) e repete-se os 3 primeiros passos para $n = 1, 2, \dots, N$, em que N é o número de escalas;
5. constrói-se o vetor $v_m = (x_m^1, \dots, x_m^N)$ que é o vetor de características da imagem g_m ;
6. repete-se os 5 primeiros passos com todas as imagens da base de dados construindo a matriz M , cujas colunas são formadas pelos vetores v_m .

3.7 Algoritmos Evolutivos

Muitas invenções foram resultados de aplicações biológicas ou de princípios naturais no estudo de sistemas humanos, como por exemplo, o radar foi baseado em uma característica dos morcegos e submarinos foram baseados nos peixes. A evolução natural das espécies pode ser interpretada como um processo de aprendizado sobre como se adaptar a um ambiente e sobreviver às suas alterações (Yu & Gen, 2010). Com a ideia de imitar esse comportamento de “sobrevivência do mais apto” foram criados os Algoritmos Evolutivos (**AEs**).

Os Algoritmos Evolutivos são métodos de otimização ou aprendizado com capacidade para evoluir. Eles tem três características principais:

1. **Baseados em População.** AEs possuem um grupo de soluções, que são chamados de “população”, para aumentar a quantidade de possíveis soluções do problema para ter mais chances de chegar na solução ótima. A população é um princípio básico do processo de evolução;
2. **Orientado por “Fitness”.** Toda solução em uma população é chamada de “indivíduo”. Todo indivíduo tem sua própria representação por genes, chamada de “código”, e seu valor de performance é chamado de “fitness” (aptidão). Os AEs escolhem os indivíduos “mais aptos” sendo esse o fundamento do processo de otimização e convergência desses algoritmos;
3. **Dirigido por Variações.** Os indivíduos passam por operações variadas a fim de imitar as mudanças genéticas, mutações e “crossovers”, o que é fundamental para procurar um espaço de solução ótima.

Desde 1960 diversos algoritmos baseados nesses três princípios foram propostos. Os algoritmos evolucionários básicos são: os algoritmos genéticos (**AGs**), programação evolucionária (**PE**) e programação genética (**PG**), todos eles são baseados em procuras aleatórias de espaços de soluções iniciais.

Neste trabalho foi utilizada apenas ideias de Algoritmos Genéticos. Na próxima seção é apresentada uma revisão sobre **AG's** e então mostrados os detalhes sobre sua implementação.

3.7.1 Algoritmos Genéticos (AG's)

De acordo com (Ahn, 2005), nas ultimas décadas, algoritmos genéticos foram aplicados com sucesso em problemas de engenharias e ciências. Isso seria consequência do grande progresso em sua teoria, design e desenvolvimento. O AG requer uma adaptação para cada aplicação prática em que for utilizado, e o uso de parâmetros não adequados afeta diretamente na performance do método utilizado, assim, é necessário um estudo prévio do caso antes de sua implementação.

Um AG simples possui a estrutura mostrada na Figura 3.6.

Os AG's são poderosos mecanismos de busca. Eles viajam pelo espaço de solução, buscando soluções ótimas. Com base no que foi descrito nesta seção, as etapas mostradas na Figura 3.6, de um modo genérico, são descritas da seguinte maneira:

- os AG's codificam as variáveis de decisão (parâmetros de entrada) do problema em strings de solução. Cada string, chamada de indivíduo (ou cromossomo), representa um candidato

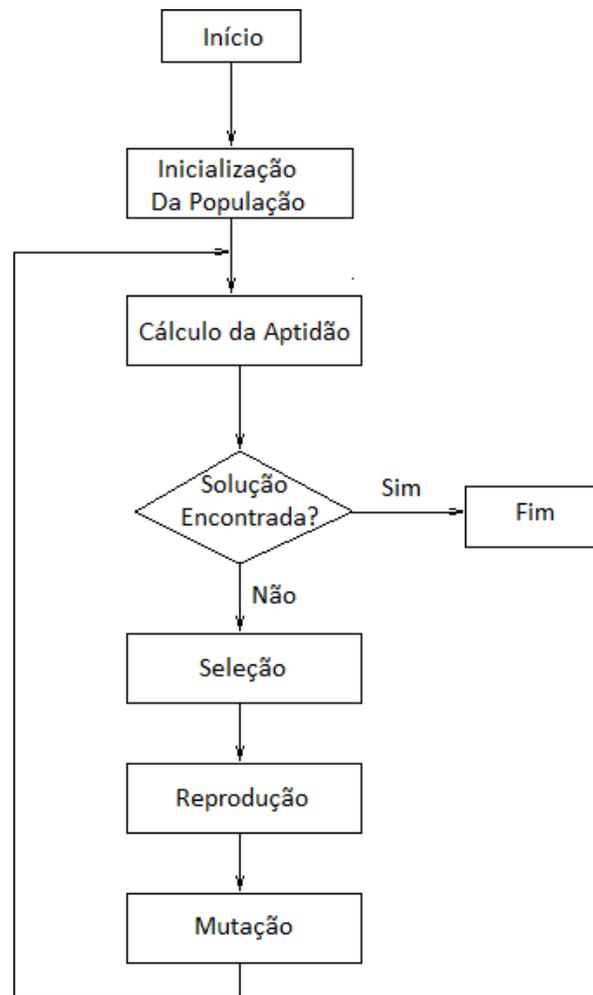


Figura 3.6: Forma Geral de Um Algoritmo Genético. Fonte:(Miranda, [s.d.](#))

a solução. Os caracteres das strings são chamados de genes. A posição e o valor na string de um gene são chamados de locus e alelo, respectivamente. Existem dois tipos de codificação, o genótipo e o fenótipo. O primeiro denota a codificação das variáveis e o segundo representa as próprias soluções;

- uma função de aptidão deve ser precisa para a diferenciação entre soluções boas e ruins. Essa função de aptidão pode ser representada em termos matemáticos ou como alguma simulação computacional, ou até em termos da avaliação subjetiva humana. A aptidão gera um sinal diferencial que guia o AG para a evolução do seu problema;
- a população inicial é gerada de forma aleatória ou com um conhecimento a priori do problema. Os indivíduos são avaliados de acordo com a qualidade de suas aptidões. Para que a prole seja gerada (novas soluções), operadores genéticos são aplicados à atual população.

Os operadores genéticos são:

- **seleção (reprodução):** seleciona-se os indivíduos com grandes aptidões como pais para a próxima geração. Em outras palavras, a seleção é usada para aumentar a qualidade

média da população, dando aos indivíduos superiores melhores chances de serem copiados para a próxima geração de soluções. Os dois tipos de seleções mais utilizados são: seleções baseadas em proporção e em ordem. A baseada em proporção escolhe os indivíduos baseados em seus valores de aptidão relativos comparadas as aptidões dos outros indivíduos; como exemplo, existem os métodos da roleta viciada, da seleção do restante estocástico (Booker, 1982) e da seleção estocástica universal (Baker, 1985). A seleção baseada em ordem escolhe os indivíduos segundo seus valores de “fitness”; seleção por torneio (Brindle, 1981) e seleção por truncamento (Muhlenbein & Schlierkamp-Voosen, 1993) são exemplos desse tipo;

- **crossover (recombinação):** o “crossover” troca e recombina soluções parciais de dois ou mais indivíduos de acordo com a possibilidade de “crossover”, p_c , para que haja diversidade na população. Esse operador explora as soluções atuais com o objetivo de melhorá-las. Dois dos mais populares são: o “crossover” de ponto único (“one-point crossover”) e o uniforme. Dados dois cromossomos, o “crossover” de **ponto único** aleatoriamente escolhe um alelo de algum deles (chamado ponto de crossover) e então troca todos os genes a partir desse ponto entre os dois indivíduos (Figura 3.7). O “crossover” uniforme troca genes entre dois indivíduos com probabilidade de 0,5, portanto, desse modo uma quantidade máxima de alelos é usada para troca (Figura 3.8). Em geral, os cromossomos do lado esquerdo são chamados de “pais” dos cromossomos do lado direito, que pertencem a próxima geração de soluções;

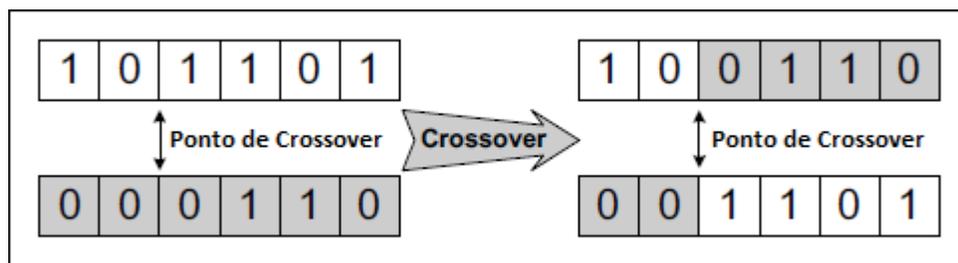


Figura 3.7: Crossover de Ponto Único.

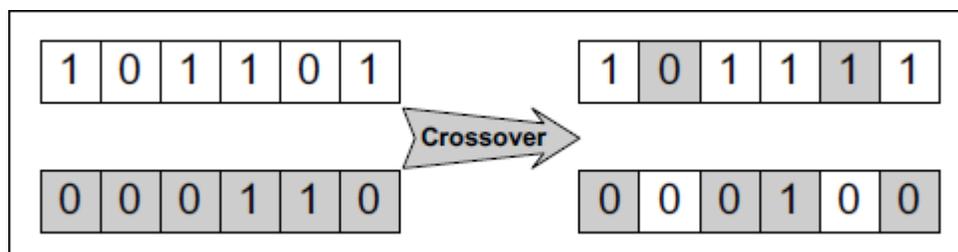


Figura 3.8: Crossover Uniforme.

- **mutação:** mutações agem alterando uma pequena porcentagem de genes na string dos indivíduos para que hajam mínimas alterações nas soluções recombinadas. Um operador clássico é a mutação por bits, na qual cada gene é complementado por uma porcentagem de mutação p_m . Por exemplo, para $p_m = 0,17$, um indivíduo $A = [1\ 1\ 1\ 1\ 1\ 1]$ se

tornará $A_1=[1 \ 1 \ 0 \ 1 \ 1 \ 1]$ quando o terceiro gene é escolhido (aleatoriamente) para mutação. Note que apenas um gene foi escolhido de modo aleatório por causa da taxa de mutação utilizada, de fato, como A possui 6 alelos, o método escolherá $\text{round}(p_m \cdot 6)$ locus aleatoriamente para sofrerem a mutação. Em geral, a probabilidade de mutação é tomada como sendo valores pequeno.

Quanto aos critérios de parada do método, eles podem ser relacionados a quantidade de gerações estipuladas ou com relação a diferença das soluções obtidas, de forma que o erro seja mínimo.

3.8 Implementação do Algoritmo Genético

Os operadores genéticos utilizados para seleção e “crossover”, neste trabalho, foram os métodos da Roleta Viciada e do Ponto Único, respectivamente. Considerando uma geração de indivíduos, o processo de seleção é feito da seguinte maneira:

1. toma-se S como sendo a soma de todos os valores de “fitness” da população em questão;
2. aleatoriamente, se escolhe um valor, R , entre 0 e S ;
3. somam-se os valores de “fitness”, do maior para o menor, até que essa soma seja maior que R . O indivíduo cujo valor “fitness” fez com que a soma ultrapasse R é, então, escolhido para ser um dos pais;
4. o processo continua até que os dois casais estejam determinados.

A Figura 3.9 representa o processo de seleção mencionado para uma população de 5 indivíduos. Nessa figura cada indivíduo está representado em um gráfico no qual as porcentagens de área apresentadas são a razão entre valor de “fitness” do indivíduo e S . Sem perda de generalidade e a fim de facilitar a explicação os valores de “fitness” dos indivíduos 1, 2, 3, 4 e 5 serão supostos 31, 5, 38, 12 e 15, respectivamente, assim o valor de S é 100. Os indivíduos 2 e 3 já foram selecionados formando o primeiro casal da geração, isto é, a roleta foi girada duas vezes e na primeira vez o valor R estava no intervalo $[95, 100]$ enquanto que na segunda estava no intervalo $[0, 38]$. Note que a quantidade de indivíduos é ímpar e portanto um cromossomo ficará sem par para o “crossover”. Para que a quantidade de indivíduos da próxima geração seja igual a da primeira pode-se repetir o indivíduo de maior “fitness”, assim, a próxima geração dessa população será composta pelo indivíduo 3 junto com os 4 filhos gerados pelos 2 casais selecionados.

O valor de aptidão (“fitness”) foi usado como a acurácia de reconhecimento dos padrões *MSLMP*.

O AG foi utilizado, neste trabalho, para:

1. encontrar o valor de λ , presente função Sigmoide (Equação (3.11)), visto que esse parâmetro exerce grande influência na acurácia do método *MSLMP*;
2. encontrar a matriz de pesos, P , que também exerce influência na performance do método.

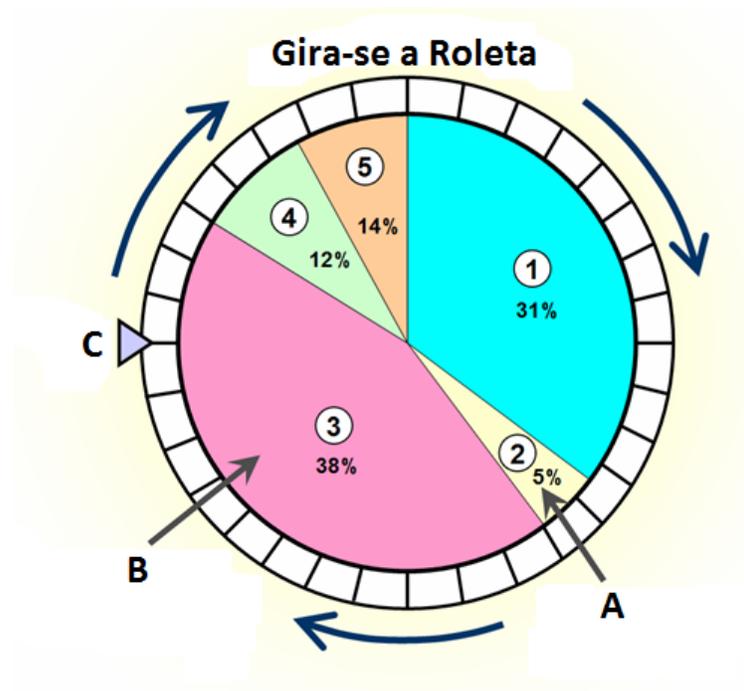


Figura 3.9: Ilustração do método de seleção roleta-viciada. A) Indica o indivíduo menos apto da população. B) indica o indivíduo mais apto. C) É o ponto de seleção. Fonte: (“Roulette wheel selection”, [s.d.](#))

Como já descrito anteriormente, o método *MSLMP* trabalha com várias escalas. Para cada escala foi ajustada uma matriz de pesos específica.

Como já foi dito anteriormente, os algoritmos genéticos foram utilizados neste trabalho na escolha do parâmetro λ que aparece na função sigmoide (3.11) e também na escolha das posições dos pesos que aparecem na matriz de pesos (3.10). O parâmetro λ é bastante sensível e pequenas variações podem levar a resultados insatisfatórios, portanto um ajuste adequado se faz necessário. A posição dos pesos, na matriz de pesos, também exerce grande influência na performance do método, assim, encontrar posições adequadas também é desejado.

A determinação dos parâmetros mencionados anteriormente dependem da base de dados considerada, além disso devem ser considerados os parâmetros:

- n : a escala escolhida para aplicação do AG;
- p : quantidade de indivíduos da população;
- p_m : taxa de mutação;
- M : a quantidade máxima de gerações;
- Te : as imagens da base de dados destinadas a teste;
- Tr : as imagens da base de dados destinadas a treino.

A função de aptidão será denotada por F . A implementação de ambos os casos é discutida a seguir.

a) Caso 1: determinação do parâmetro λ

Neste caso, o parâmetro da função de aptidão F é um escalar λ no intervalo $(0, 1]$, isto é, os valores $\lambda \in (0, 1]$ são os indivíduos da população. Como foi mencionado anteriormente, para implementar o AG em alguma aplicação prática deve-se fazer uma adaptação baseada na análise do contexto dessa aplicação. Considerando os parâmetros citados anteriormente, o último passo para a adaptação do AG para esse caso é selecionar a matriz de pesos usada para a extração dos padrões *MSLMP* das imagens. Neste trabalho, a matriz padrão de pesos utilizada para a determinação do parâmetro λ foi a matriz dada por (3.15).

A matriz (3.15) pode ser interpretada da seguinte maneira: considerando o valor $PJ(6, 6)$ como centro (PJ é 11×11), que representa o pixel da imagem a ser extraída o micro padrão, tem-se que a região 3×3 em torno desse ponto denota a P_1 (matriz de pesos para a escala um). A região 5×5 , tomando a região 3×3 como sendo a matriz nula, representa a P_2 . A vizinhança 7×7 , escolhida da mesma forma que a 5×5 , denota a P_3 e a vizinhança 11×11 denota a P_5 .

$$PJ = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}. \quad (3.15)$$

Essas matrizes foram obtidas de (Bravo, 2016). Nesse trabalho, foram exibidas 5 variações da matriz apresentada em 3.10. Um dos objetivos da autora foi o de identificar qual, dentre os 6 tipos, obtém a melhor acurácia na identificação de texturas utilizando o *MSLMP*. Em sua conclusão, ela apresentou a matriz 3.15 como sendo a que obteve melhores resultados. Devido a isso, essa matriz foi utilizada nesta pesquisa.

Voltando a atenção novamente para a obtenção do parâmetro λ , de acordo com a Figura 3.6, os procedimentos seguintes são:

1. gerar aleatoriamente p valores no intervalo $(0, 1]$, para compor a primeira geração de indivíduos;
2. extrair os padrões *LMP* dos conjuntos Te e Tc para cada indivíduo λ na população, e aplicar o processo de classificação para obter a acurácia de reconhecimento. A acurácia, $F(\lambda)$, é o valor de “fitness” do indivíduo λ ;
3. converter os indivíduos da população para a base binária. A quantidade de alelos de cada indivíduo deve ser a mesma para todos. A escolha dessa quantidade fica a critério do programador. Neste trabalho foram utilizados 60 alelos;
4. se alguma acurácia atingir 100%, o procedimento é finalizado, caso contrário, o processo de seleção prossegue escolhendo-se c casais, em que $c = \text{round}((p - 1)/2)$, caso p seja ímpar ou $c = p/2$ caso p seja par;

5. cada casal gera dois novos indivíduos para a próxima geração. Nota-se que, a seleção gera $p - 1$ indivíduos, caso p seja ímpar. Para que a quantidade de indivíduos da próxima geração continue sendo p repete-se o cromossomo de maior fitness na nova geração;
6. a nova população gerada pelo “crossover” passa pelo processo de mutação;
7. converte-se novamente os cromossomos da população para a base decimal, e com exceção da primeira etapa, repete-se todas as outras até que o número máximo de gerações seja atingido;
8. o valor de λ é escolhido como sendo o que obteve a melhor acurácia.

Exemplo:

A exemplificação passo a passo do algoritmo descrito anteriormente para a primeira geração de parâmetros λ encontra-se descrita a seguir. Foram utilizados os seguintes parâmetros: $n = 2$, $p = 3$ e $p_m = 0,05$, aplicados ao banco de dados ORL com 320 imagens para teste e 80 para treino. Nesse caso, foi considerada a matriz de pesos de dimensão 5 (Equação (3.2)), dada por:

$$P = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

Segue então, o passo a passo.

1. considera-se a população inicial gerada aleatoriamente $P_0 = [0, 12500, 50000, 1523]$.
2. aplica-se o algoritmo descrito na Subseção 3.6 p vezes, na qual cada extração será realizada para um λ diferente da população, obtendo assim duas matrizes M_e e M_t referentes aos padrões extraídos do grupos de treino e teste, respectivamente. Os valores obtidos pelo processo de classificação estão apresentados na Tabela 3.1.

Tabela 3.1: Acurácias obtidas pelo conjunto de valores da população P .

λ	Acurácia	Aptidão
0,1250	23%	$F(0, 1250) = 0, 23$
0,5000	16%	$F(0, 5000) = 0, 16$
0,1523	30%	$F(0, 1523) = 0, 30$

3. converte-se os valores de λ para binário. Esses valores encontram-se mostrados na Tabela 3.2.
4. como a acurácia não atingiu 100% para nenhum dos valores considerados, é aplicado o método de Seleção da Roleta-Viciada para gerar os pais para a próxima geração. A Tabela 3.3 exemplifica o processo descrito na Figura 3.9, tem-se $S = 0,69$ e $c = 1$. Para todos os cálculos do algoritmo foi considerada apenas a parte a esquerda da vírgula dos valores binários.

Tabela 3.2: Conversão binária da população.

λ	Binário
0,125	0,00100000
0,50	0,10000000
0,1523	0,00100111

Tabela 3.3: Processo de seleção de roleta-viciada.

Iteração	Valor Gerado	Soma Final	Indivíduo Selecionado
1	0,29	0,30	$\lambda = 0,1523$
2	0,62	0,69	$\lambda = 0,5000$

5. o processo de recombinação é então realizado. Note que cada cromossomo da população possui 8 alelos e, portanto, como mencionado, o “crossover” de ponto único escolherá um desses alelos para fazer a recombinação. A Tabela 3.4 ilustra esse procedimento:

Tabela 3.4: Processo de recombinação.

Pais	Alelo Selecionado	Filhos
0,00100111	Quinto	0,00100000
0,10000000		0,10000111

Observe que para que a nova população mantenha a mesma quantidade de indivíduos, deve-se repetir o elemento de maior acurácia da geração da população anterior. A nova geração é mostrada na Tabela 3.5.

Tabela 3.5: Geração 1

Geração 1
0,00100111 ← Indivíduo da geração 0
0,00100000 ← Filho
0,10000111 ← Filho

6. a próxima etapa é a mutação. Como a taxa de mutação igual a 0,05, a quantidade de alelos nos cromossomos a sofrerem mutação é de 1, pois $\text{round}(p_m \cdot 8) = 1$. A Tabela 3.6 mostra os detalhes desse processo.

Tabela 3.6: Processo de Mutação

Indivíduo	Alelo Selecionado	Novo Indivíduo
0,00100111	Primeiro	0,10100111
0,00100000	Quarto	0,00110000
0,10000111	Sétimo	0,10000101

7. os cromossomos da população 1 são convertidos para a base real e então os processos de 2 à 7 são repetidos até que a quantidade máxima de gerações seja atingida. A conversão,

Tabela 3.7: Conversão dos cromossomos para a base decimal.

Cromossomo	lambda
0,10100111	0,6523
0,00110000	0,1875
0,10000101	0,5195

dos dados binários que aparecem na Tabela 3.6, para a base decimal está apresentada na Tabela 3.7.

8. o valor de λ é escolhido como sendo o que obteve a melhor acurácia.

b) Caso 2: determinação da matriz de pesos P

Para esse caso, a matriz de peso P é o parâmetro da função de aptidão F , isto é, a população é composta por matrizes de pesos. Diferente do caso 1 em que tinha-se a matriz de pesos fixa e o valor de λ variável, tem-se agora o contrário e como já mencionado o valor de λ utilizado sendo aquele obtido pelo caso 1.

Para iniciar esse processo, a escala utilizada é definida como sendo a mesma da primeira aplicação, ou seja, apenas depois usam-se outras escalas para gerar a quantidade de matrizes que se deseja utilizar para compor o processo multiescalar.

O conveniente desse caso é que os elementos das matrizes de peso são apenas 0 e 1 e assim, não há necessidade da etapa de conversão binária, bastando transformar a matriz P em um vetor linha (cromossomo) concatenando-se as linhas das matrizes.

Assim, com exceção da primeira etapa, a implementação do AG segue idêntica ao Caso 1.

Exemplo:

A exemplificação desse caso se dá nas mesmas condições do exemplo para o caso 1, utilizando-se o λ obtido no caso 1. Neste exemplo é utilizada uma matriz 5×5 .

1. população 0: geram-se 3 vetores de dimensão 1×25 cujas entradas são valores binários. Foram gerados os vetores M_i mostrados na Tabela 3.8.

Tabela 3.8: População 0 para o caso 2.

Vetores	
M_1	00010 00101 01010 10100 00001
M_2	10100 00101 00010 10100 00100
M_3	10111 00001 00010 10100 00000

2. a cada 5 alelos dos cromossomos M_i é feita uma “quebra” na linha de modo que uma nova matriz 5×5 seja criada. Essas matrizes são as matrizes de pesos usadas para a extração dos padrões $MSLMP$ das imagens de treino e teste. As matrizes usadas para essa população encontram-se a seguir.

$$M_1 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, M_2 = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \text{ e } M_3 = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

As demais etapas são inteiramente análogas as do exemplo para o caso 1.

3.9 Ajuste nos Níveis de Cinza

Como já foi mencionado, duas técnicas de pré-processamento para ajuste nos níveis de cinza das imagens são propostas neste trabalho, os detalhes e motivações sobre essas técnicas encontram-se a seguir.

3.9.1 Correção de Brilho e Homogeneização

A motivação para o desenvolvimento desse método está no estudo de imagens em bases de dados que apresentam grandes variações de iluminação e ângulos de captura, que são situações que ocorrem na prática, quando por exemplo, se deseja reconhecer uma determinada pessoa em meio a tantas outras em ambientes internos ou externos, sem que haja a colaboração da pessoa para isso. Essa é uma aplicação típica da área criminal, na qual se tenta reconhecer criminosos em situações de ataque, com imagens capturadas em ambientes públicos, por exemplo. A base de Dados MUCT (Milborrow/University of Cape Town) é uma base bastante adequada para avaliar modelos propostos para essa finalidade. Assim, essa foi a principal base de dados utilizada neste estudo, para esse propósito.

Devido a alta variação de iluminação e poses, a média dos níveis de cinza das imagens da base MUCT variam bastante, como mostra a Figura 3.10, que mostra a quantidade de imagens distribuídas em função das médias dos níveis de cinza.

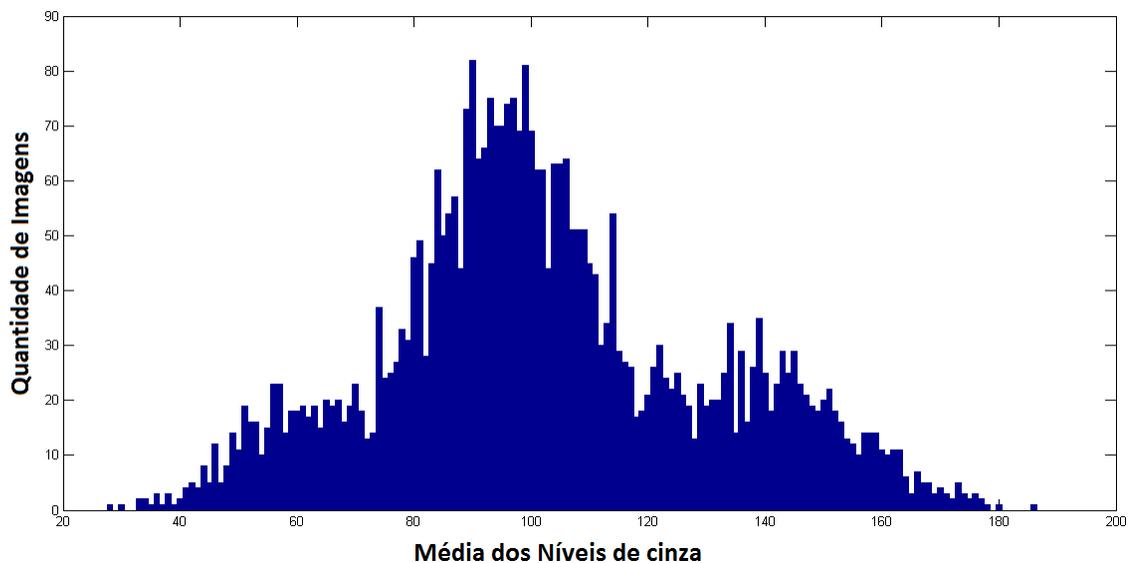


Figura 3.10: Média dos níveis de cinza da MUCT.

Os valores das médias encontram-se distribuídos entre $M_{min} = 28,46$ e $M_{max} = 185,66$. A média dos valores da Figura, isto é, a média das médias dos níveis de cinza das imagens é $M_m = 102,44$.

A ideia foi encontrar um valor padrão, P , $P \in [M_{min}, M_{max}]$, de forma que P dependa diretamente das médias das imagens contidas na base de dados, e, então, tentar fazer com que todas as médias das imagens estejam próximas a esse valor. Após determinar o valor

P , a aproximação é feita através de um parâmetro θ_I , com $\theta_I \in (-1, 1)$, que é unicamente determinado por P e M_I , em que M_I é média dos níveis de cinza da imagem I , a qual se deseja ajustar, assim cada imagem possui um determinado parâmetro de ajuste. Para isso, considerou-se P , como a média dos valores da Figura 3.10, ou seja, $P = M_m$, e o valor θ_I calculado por meio da Equação (3.16):

$$\theta_I = \frac{P - M_I}{M_{max} - M_{min}}. \quad (3.16)$$

Após a obtenção do parâmetro θ_I , cada imagem I , tem seus pixels $I(i, j)$, alterados pelo valor $(1 + \theta_I)I(i, j)$. A correção é, então, determinado da seguinte maneira:

- se $\theta_I > 0$, significa que a média da imagem I é menor que a média geral, assim, para que haja uma aproximação das médias é necessário somar a cada pixel $I(i, j)$ um valor positivo, ou seja o valor $\theta_I \cdot I(i, j)$;
- se $\theta_I < 0$, a média da imagem I é maior que a média geral e, dessa forma, para que haja uma aproximação das médias é necessário subtrair um valor de cada pixel $I(i, j)$, ou seja o valor $\theta_I \cdot I(i, j)$.

As Figuras de 3.11 exemplificam os efeitos da técnica descrita nas imagens da MUCT para diferentes iluminações de acordo com a Figura 2.11.

Em termos de visualização das imagens, o reajuste pode ser interpretado da seguinte maneira:

- Com relação ao valor de referência, P , se a imagem for “clara”, então a técnica irá “escurece-lá”;
- caso seja “escura”, então o ajuste irá “clareá-la”;
- caso seja uma imagem que não necessite de ajustes, isto é, a média dos seus níveis de cinza estiver próxima a P , a imagem sofrerá alterações mínimas, nestes casos $\theta_I \approx 0$.

Na próxima seção é apresentada uma outra técnica de correção dos níveis de cinza, porém para bases de dados que contém imagens predominantemente muito claras.

3.9.2 Ajuste na Intensidade dos Pixels

Com a motivação para lidar com bancos de dados cuja intensidade dos pixels é alta foi desenvolvida uma outra técnica para ajuste nos níveis de cinza de uma imagem clara.

Uma base de dados propícia para testes e exemplificações da ideia desse método é a base YALE. A Figura 3.12 mostra as médias dos níveis de cinza das imagens da YALE. Apesar da base apresentar alta iluminação, as médias das imagens não variam tanto, pois cerca de 80% delas se encontram no intervalo $[177, 197]$, que é um intervalo bastante reduzido, caracterizando as imagens, de um modo geral, como muito claras.

Observando as imagens da Figura 2.13, percebe-se que a grande quantidade de pixels brancos no fundo está presente em praticamente todas as imagens da base. Devido a isso, a ideia inicial do ajuste está relacionada com quantidade de pixels de valor 255 (branco). Diferente da técnica anterior, o parâmetro θ_I dependerá unicamente da imagem I . Também foi considerada uma constante $\theta^0 \in (0, 1)$ escolhida adequadamente, que será usada em todas as imagens.



(a) Imagem de Iluminação s



(b) Imagem após o reajuste dos níveis de cinza



(c) Imagem de Iluminação q



(d) Imagem após o reajuste dos níveis de cinza



(e) Imagem de Iluminação r



(f) Imagem após o reajuste dos níveis de cinza

Figura 3.11: Exemplos da Técnica Descritas com Imagens de 3 Iluminações distintas.

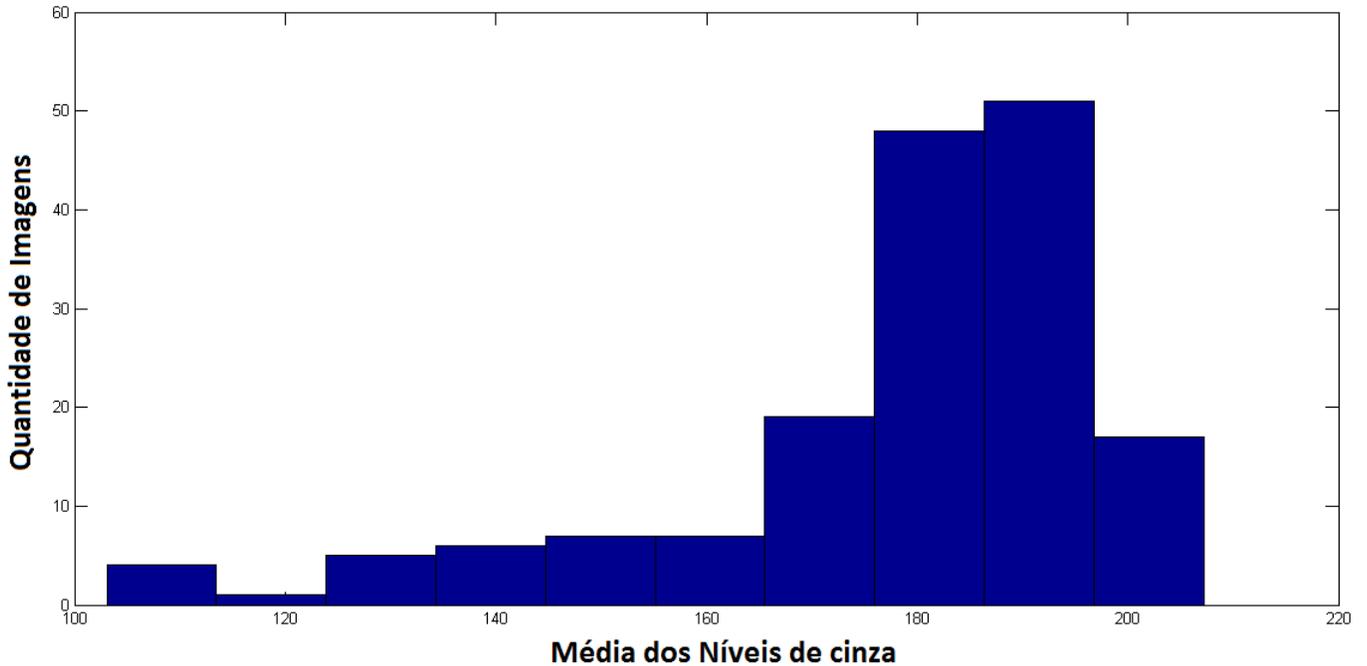


Figura 3.12: Média dos níveis de cinza da YALE em relação a quantidade de imagens.

Seja $P_I, P_I \in \mathbb{Z}$, a quantidade de pixels da imagem I , como em um banco de dados as imagens devem possuir as mesmas dimensões, caso contrário basta redimensioná-las para um valor padrão, tem-se que P_I é constante para toda imagem I da base de dados. Seja agora $P_I^b \in \mathbb{Z}$ a quantidade de pixels da imagem I cujo valor é maior ou igual 220 (pois, não só os pixels de extremo branco (255) contribuem para deixar a imagem clara), daí, define-se θ_I através da Equação (3.17):

$$\theta_I = \frac{P_I^b}{P_I}. \quad (3.17)$$

Claramente percebe-se que, $\theta_I \in (0, 1)$, para todas as imagens da base. Assim, dados θ_I e θ^0 o reajuste, inicial, das intensidades é feito da seguinte maneira:

- Se $\theta_I > \theta^0$ será subtraído o valor $\theta_I \cdot I(i, j)$, de cada pixel $I(i, j)$;
- Se $\theta_I < \theta^0$ será somado o valor $\theta_I \cdot I(i, j)$, de cada pixel $I(i, j)$.

As Figuras de 3.11 exemplificam os efeitos da técnica descrita nas imagens da YALE.



Imagem original



Imagem após o reajuste dos níveis de cinza



Imagem original



Imagem após o reajuste dos níveis de cinza



Imagem original



Imagem após o reajuste dos níveis de cinza

Figura 3.13: 3 imagens da Yale antes e depois da implementação da técnica descrita.

3.10 Considerações Finais

Neste capítulo foi introduzido o método *MSLMP*, utilizado como extrator de características para o reconhecimento de faces. Também foi trabalhado a teoria dos algoritmos genéticos na busca de parâmetros mais adequados, os quais foram utilizados para tentar melhorar a acurácia do método. Foram também apresentadas duas técnicas para ajuste de níveis de cinza de imagens. Uma delas dedicada ao ajuste de todas as imagens que compõem uma determinada base de dados e a outra para ajustar os níveis de cinza de imagens muito claras.

No Capítulo seguinte são apresentadas as metodologias e os resultados obtidos neste trabalho.

Metodologias e Resultados Experimentais

4.1 Considerações Iniciais

Neste capítulo são apresentados e discutidos os resultados da acurácia do Padrão Mapeado Localmente em Multiescalas para o reconhecimento de faces. O método foi avaliado nas base de dados apresentadas no Capítulo 2.

4.2 Metodologia Aplicada

A seguir são apresentadas as metodologias empregadas neste trabalho, utilizando as técnicas apresentadas no capítulo anterior.

4.2.1 Parâmetros do AG

As informações sobre os parâmetros de entrada do AG utilizados se encontram na Tabela 4.1.

Tabela 4.1: Parâmetros Iniciais do AG utilizados neste Trabalho.

População	30
Taxa de Mutação	0.05
Quantidade de Gerações	30

4.2.2 Validação Cruzada

Para cada base de dados utilizada foram separadas 20% das imagens para teste e 80% para o treinamento, em forma de rodízio, para que o processo de validação cruzada seja executado. A validação cruzada é um método estatístico utilizado para medir a capacidade de generalização de uma técnica. A ideia geral do método é dividir o conjunto de dados em grupos mutualmente exclusivos (grupos para validação) para a avaliação da performance do método.

Neste trabalho, cada banco de dados foi dividido em 5 grupos distintos para teste, com o objetivo de avaliar o *MSLMP* para o reconhecimento de todas as imagens da base. A escolha do grupo de imagens para validação cruzada foi diferente em cada base e por isso, os detalhes dessa escolha encontram-se apresentados juntos aos resultados experimentais de cada banco de

dados em suas respectivas seções. Em geral, todos os grupos foram escolhidos de forma que haja diversidade nas imagens em relação à pose, iluminação e expressões.

4.2.3 Processo de Classificação

O processo de classificação é feito calculando-se a distância Chi-quadrado (Pele & Werman, 2010) entre dois vetores de características $v_i, w_j \in \mathbb{R}^n$, através da Equação (4.1):

$$\chi^2(v_i, w_j) = \frac{1}{2} \sum_{k=1}^n \frac{(v_i^k - w_j^k)^2}{(v_i^k + w_j^k)}, \quad (4.1)$$

na qual, v_i representa o vetor de características da i -ésima amostra de teste a ser avaliada (teste) e w_j o vetor de característica da j -ésima amostra para comparação (treino). A figura (4.1) ilustra o processo de classificação.

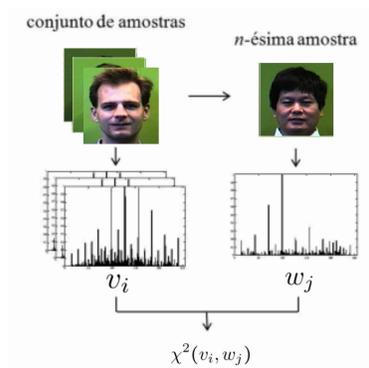


Figura 4.1: Processo de classificação de faces.

A acurácia A_i do grupo de validação i , $i \in \{1, 2, 3, 4, 5\}$, é então calculada pela Equação (4.2):

$$A_i = \frac{\text{Número de acertos}}{\text{Número total de imagens do grupo}}. \quad (4.2)$$

A taxa de acerto final é calculada pela média aritmética entre todos os grupos considerados neste caso.

4.2.4 Uso das Escalas

Como já foi ressaltado anteriormente o método *MSLMP* é multiescalar, no qual cada escala é caracterizada por um raio, e esse raio determina o tamanho da região das imagens de onde se extrai o padrão *LMP*. Normalmente, a primeira escala determina uma região de dimensão 3×3 e para maiores escalas, por exemplo, a oitava, tem-se uma região de 35×35 , de acordo com a Equação (3.9). Porém, neste trabalho notou-se durante testes que, ao aplicar o método com uma única escala a acurácia variava conforme o tamanho da região utilizada para a extração. A partir daí a multiescala foi usada apenas com as regiões que obtiveram as melhores acurácias e a conclusão foi que houve uma melhora geral de performance do método.

4.3 Resultados Experimentais

Os resultados obtidos em cada base se encontram a seguir, e estão separados por subseções, em que cada uma refere-se a um único banco de dados.

4.3.1 ESSEX

As informações sobre essa base de dados estão na subseção 2.5.1. Devido aos resultados serem satisfatórios nesse BD, o uso do AG não foi necessário. O procedimento de validação cruzada e acurácias encontram-se apresentados e discutidos a seguir.

a) Validação Cruzada

Como cada classe possui 20 imagens, cada grupo de validação contém 16 imagens para treino e 4 para teste. As imagens de teste foram escolhidas da forma que a partir da primeira, a cada 5 imagens da classe, uma é escolhida para compor esse conjunto.

b) Resultados Faces94

A Figura 4.2 mostra a relação de acurácias para uma única escala para vários raios. A Tabela 4.2 apresenta os resultados obtidos utilizando três escalas do método utilizando os raios 13, 17 e 19.

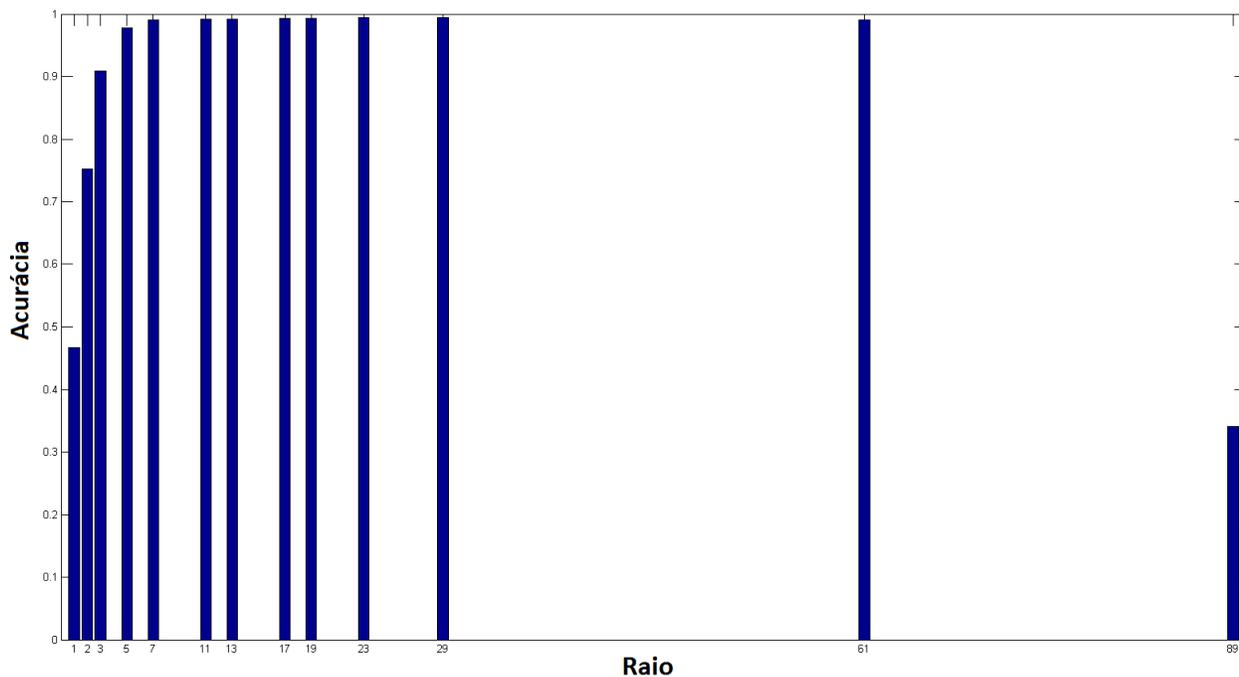


Figura 4.2: Cálculo da acurácia da Faces94 para uma escala.

c) Resultados Faces95

A Figura 4.3 mostra a relação de acurácias para uma única escala.

Tabela 4.2: Taxa de acertos Essex 94 para três escalas.

Grupo	Acertos
Grupo_a	98,37%
Grupo_b	99,51%
Grupo_c	99,67%
Grupo_d	99,51%
Grupo_e	99,51%
Média	99,31%

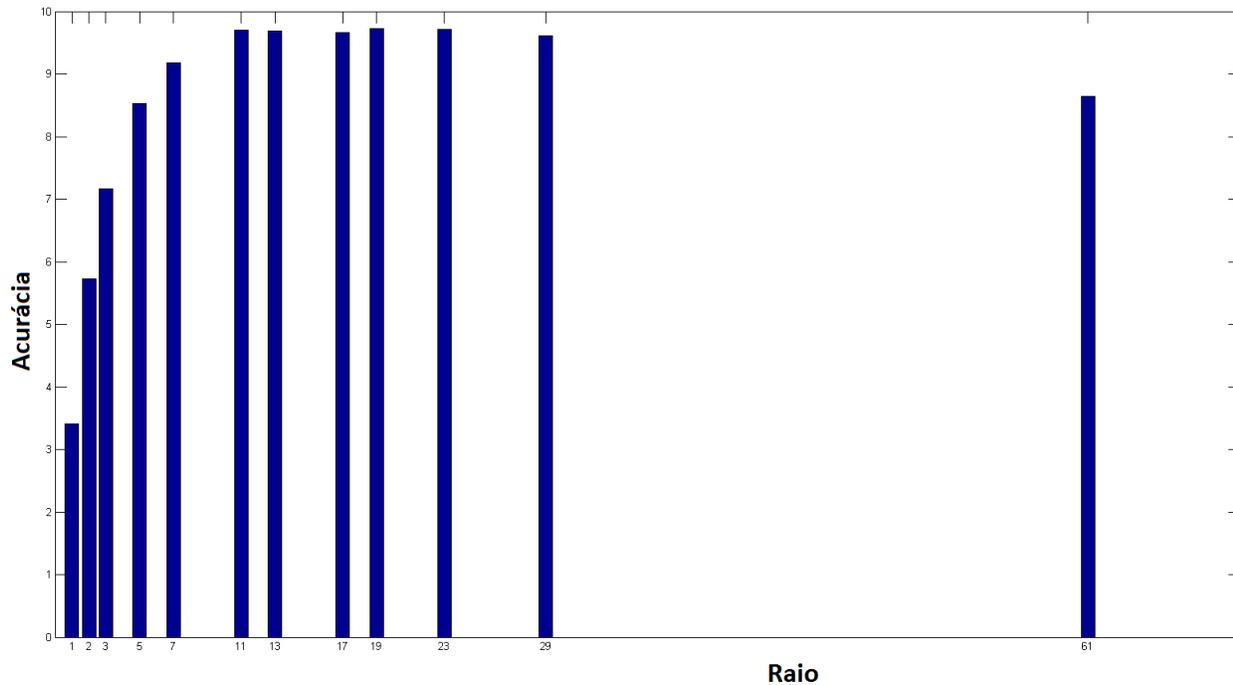


Figura 4.3: Cálculo da acurácia da Faces95 para uma escala.

Nesta base foi utilizado o algoritmo genético para melhorar os resultados. O raio escolhido para otimização da matriz de pesos e do λ , foi o 23, pois com a utilização desse raio obteve-se a maior acurácia dentre os raios mostrados na Figura 4.3. O valor λ obtido foi 0,2849. As Tabelas 4.3 e 4.4 mostram os resultados obtidos pelo método para duas escalas, considerando os raios 19 e 23, sem a utilização do AG e com a utilização, respectivamente.

d) Resultados Faces96

A Figura 4.4 mostra a relação de acurácias para uma única escala.

A utilização do AG para otimizar o λ não alterou o valor padrão, 0,25, porém, houve melhoras quando ele foi utilizado para otimizar as matrizes de pesos. A Tabela 4.5 mostra a acurácia do *MSLMP* para cinco escalas do método, utilizando os raios 13, 17, 19, 23 e 29, e a Tabela 4.6 mostra a acurácia para os mesmos raios utilizando o AG.

Tabela 4.3: Taxa de acerto Faces95 do *MSLMP* para duas escalas.

Grupo	Acertos
Grupo_a	100,00%
Grupo_b	98,61%
Grupo_c	97,92%
Grupo_d	96,68%
Grupo_e	97,92%
Média	98,23%

Tabela 4.4: Taxa de acerto da Faces95, para duas escalas, utilizando o AG para as matrizes e $\lambda = 0,2849$.

Grupo	Acertos
Grupo_a	100,00%
Grupo_b	99,65%
Grupo_c	98,26%
Grupo_d	97,57%
Grupo_e	98,61%
Média	98,82%

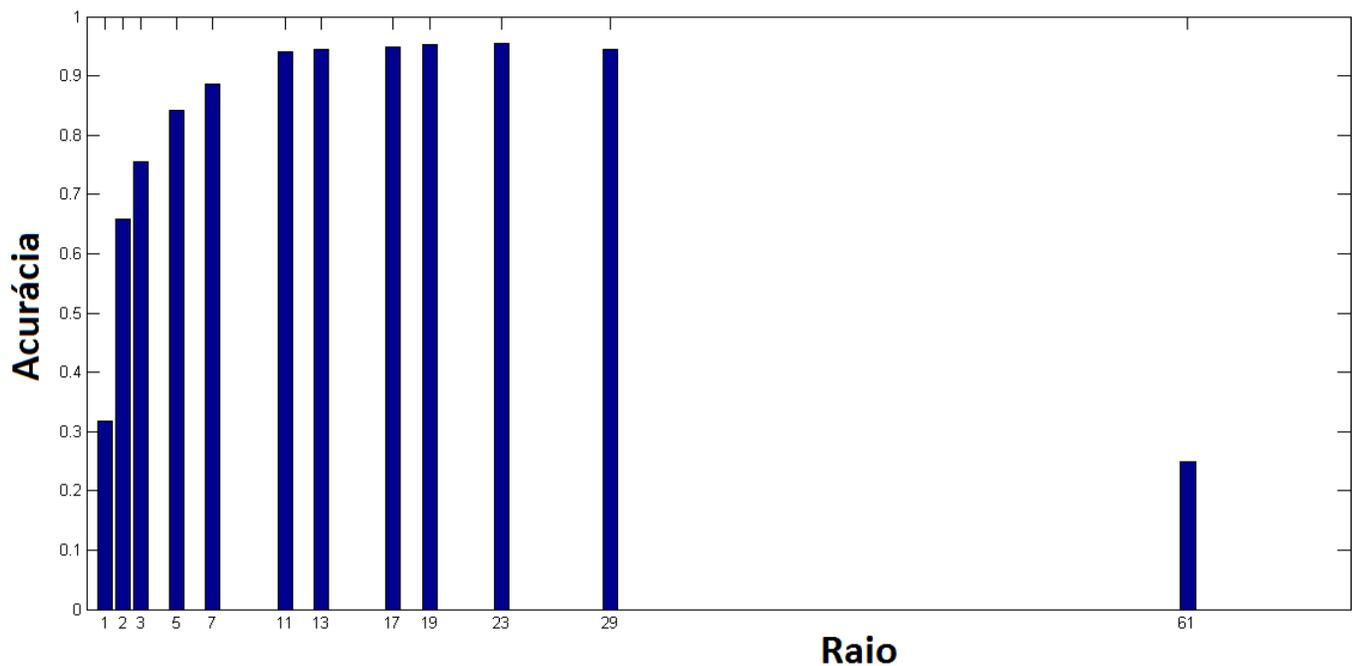


Figura 4.4: Cálculo da acurácia da Faces96 para uma escala.

Tabela 4.5: Taxa de acertos da Faces96 com cinco escalas.

Grupo	Acertos
Grupo_a	99,15%
Grupo_b	98,64%
Grupo_c	98,47%
Grupo_d	97,96%
Grupo_e	98,64%
Média	98,57%

Tabela 4.6: Taxa de acerto da Faces96 para cinco escalas utilizando o AG.

Grupo	Acertos
Grupo_a	99,49%
Grupo_b	99,32%
Grupo_c	98,81%
Grupo_d	99,15%
Grupo_e	99,66%
Média	99,29%

e) Resultados Grimace

A Figura 4.5 mostra a relação de acurácias para uma única escala. Nota-se que a acurácia foi de 100% em 5 raios do método, não sendo necessário o uso de múltiplas escalas e nem o uso de técnicas de otimização dos parâmetros envolvidos no método.

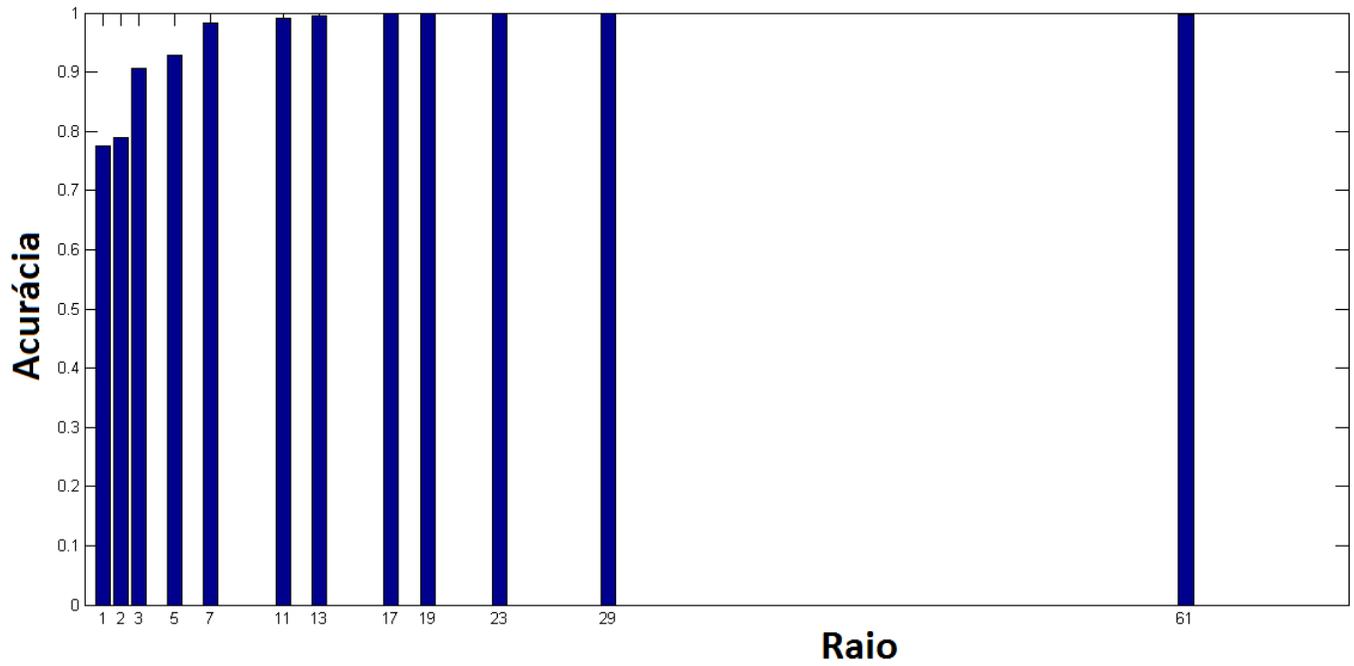


Figura 4.5: Cálculo da acurácia da Grimace para uma escala.

Em (Barnouti et al., 2016), os autores relatam 100% de acurácia nas bases Grimace e Faces94, porém das 20 imagens de cada classe os autores escolheram apenas quatro para realizar seus experimentos, não especificando a forma da escolha dessas imagens, e não relatando o uso do processo de validação cruzada.

4.3.2 JAFFE

As informações sobre essa base de dados estão na subseção 2.5.2. Devido aos resultados serem satisfatórios nesse BD não foi necessário o uso do AG.

a) Validação Cruzada

Neste BD tem-se um total de 219 imagens de 10 indivíduos, cada classe possui entre 20 e 22 imagens e para propósitos de simetria para a validação cruzada foram consideradas 20 imagens para todas as classes. As imagens excedentes foram escolhidas de expressões que já se repetiam no diretório. Portanto, o grupo de treino contém 16 faces e o de teste 4.

b) Resultados

A Figura 4.6 mostra os resultados para uma única escala. Nota-se que a acurácia foi de 100% em 3 raios do método, não fazendo uso sequer das multiescalas e portanto, não havendo necessidade do uso de técnicas de otimização dos parâmetros.

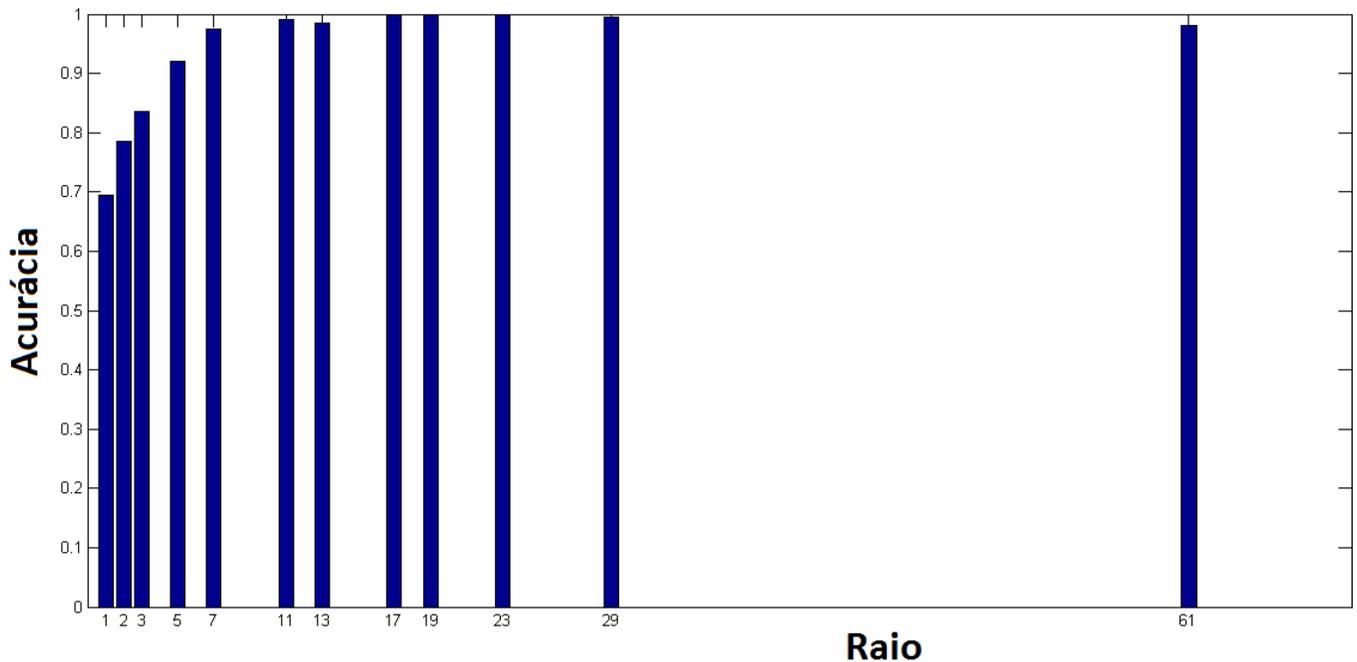


Figura 4.6: Cálculo da acurácia da JAFFE para uma escala.

4.3.3 MUCT

As informações sobre essa base de dados já foram mostradas na subseção 2.5.3.

a) Validação Cruzada

Neste BD algumas classes possuem 10 imagens e outras 15. Dentro de cada classe, as imagens foram agrupadas em conjuntos de 5 imagens, sendo cada uma fotografada por uma “webcam” diferente. Foram também utilizados 10 tipos distintos de iluminação. Para as classes de 15 imagens foram utilizados apenas 3 tipos distintos de iluminação e para classes com 10 imagens foram utilizados apenas dois.

Para formar os grupos de validação cruzada, neste trabalho, foi feita uma distribuição de forma que todos os grupos fossem contemplados com imagens das diferentes “webcams” e diferentes tipos de iluminação, distribuídos uniformemente. A Tabela 4.7 mostra os 5 grupos de teste/treino para uma única classe de 15 imagens, na qual, q , r e s representam os tipos de iluminações utilizadas e 1, 2, 3, 4 e 5 referem-se às câmeras, ou seja, $q3$ refere-se a iluminação q fotografada pela câmera 3.

b) Resultados

A Figura 4.7 mostra a relação de acurácias para uma única escala. A melhor acurácia da MUCT foi obtida utilizando 8 escalas do método a partir do raio $R_7 = 13$, cujo os resultados

Tabela 4.7: Exemplo de formação dos grupos de uma classe da MUCT.

Grupo a	Grupo b	Grupo c	Grupo d	Grupo e
$q1$	$q2$	$q3$	$q4$	$q5$
$r2$	$r3$	$r4$	$r5$	$r1$
$s3$	$s4$	$s5$	$s1$	$s2$

são mostrados na Tabela 4.8.

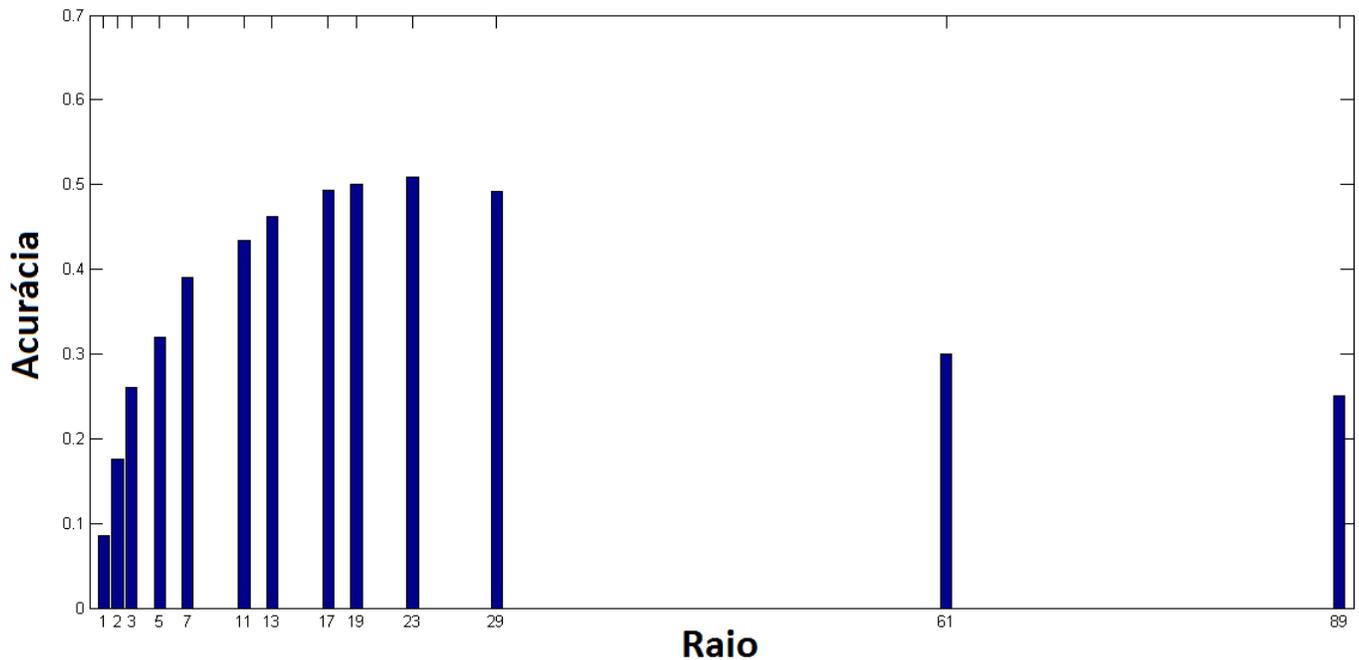


Figura 4.7: Cálculo da acurácia da MUCT para uma escala.

Tabela 4.8: Taxa de acertos da MUCT para 8 escalas

Grupo	Acertos
Grupo_a	92,41%
Grupo_b	83,49%
Grupo_c	83,09%
Grupo_d	84,02%
Grupo_e	85,75%
Média	85,75%

A Tabela 4.9 mostra os resultados após a aplicação da técnica de reajuste de níveis de cinza descrita na Seção 3.9.1.

O algoritmo genético foi utilizado nesta base para otimização do parâmetro λ e também para uma matriz de pesos. Foram utilizadas 8 escalas e dentre estas apenas uma escala, a de

Tabela 4.9: Taxa de acertos da MUCT para 8 escalas com reajuste dos níveis de cinza.

Grupo	Acertos
Grupo_a	94,44%
Grupo_b	92,68%
Grupo_c	84,29%
Grupo_d	87,08%
Grupo_e	89,88%
Média	89,68%

raio 29, foi escolhida para otimizar a matriz de pesos, sendo as demais escolhidas de forma usual. Os resultados obtidos encontram-se nas Tabelas 4.10 e 4.11.

Tabela 4.10: Taxa de acertos para $\lambda = 0, 2849$ e 8 escalas.

Grupo	Acertos
Grupo_a	94,41%
Grupo_b	95,47%
Grupo_c	89,08%
Grupo_d	90,68%
Grupo_e	91,88%
Média	92,30%

Tabela 4.11: Taxa de acertos para $\lambda = 0, 2849$ e otimizando a matriz do raio 29 com 8 escalas.

Grupo	Acertos
Grupo_a	94,94%
Grupo_b	95,87%
Grupo_c	89,88%
Grupo_d	92,14%
Grupo_e	92,41%
Média	93,00%

Em (Barnouti et al., 2016), os autores relatam uma taxa de acerto de 87,50%, porém foram usadas apenas 7 imagens para treino e 1 para teste não havendo menção quanto a escolha dessas imagens. O método foi baseado na extração de características através do *PCA*, na segmentação foi utilizado o método *Viola – Jones*, além de outras técnicas de pre-processamento que usam aplicações de filtros.

4.3.4 ORL

As informações sobre essa base de dados estão na Subseção 2.5.4. A acurácia deste banco de dados foi otimizada através do uso do algoritmo genético.

a) Validação Cruzada

Todas as classes contém 10 imagens e, portanto, os grupos de treino e teste possuem oito e duas imagens, respectivamente.

b) Resultados

A Figura 4.8 mostra a relação de acurácias nessa base em escalas únicas. A acurácia para três escalas utilizando os raios, $R = \{13, 17, 19\}$, são mostrados na Tabela 4.12. Durante os testes realizados notou-se que a acurácia permaneceu a mesma quando foram utilizados sete ou 12 escalas.

Para melhorar os resultados apresentados na Tabela 4.12 foi utilizado o algoritmo genético. Primeiramente foi escolhido o raio 17 para ajuste do parâmetro λ e da matriz de pesos, devido

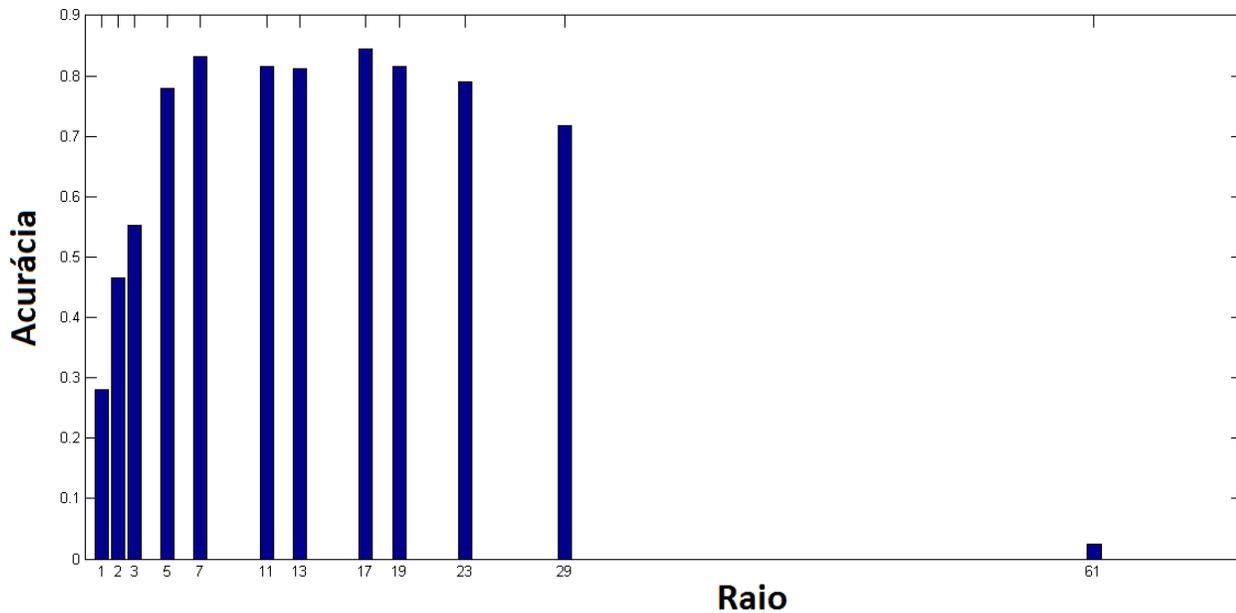


Figura 4.8: Cálculo da acurácia para uma escala.

Tabela 4.12: Taxa de acerto ORL.

Grupo	Acertos
Grupo_a	95,00%
Grupo_b	93,75%
Grupo_c	95,00%
Grupo_d	95,00%
Grupo_e	93,75%
Média	94,50%

ao fato de ter com esse raio conseguido a maior acurácia. O valor encontrado para o parâmetro λ foi $\lambda = 0,92118$. Esse valor foi usado para encontrar as matrizes referentes aos raios 13 e 19. Os resultados utilizando as 3 novas matrizes se encontram na Tabela 4.13.

Tabela 4.13: Taxa de acerto ORL com o AG.

Grupo	Acertos
Grupo_a	96,25%
Grupo_b	97,50%
Grupo_c	97,50%
Grupo_d	97,50%
Grupo_e	100,00%
Média	97,75%

Pelos testes realizados, notou-se que um ajuste nos brilhos, através do procedimento de equalização das imagens, melhorou a acurácia do método. Esses resultados são mostrados na

Tabela 4.14. O mesmo procedimento descrito no parágrafo anterior, para ajuste do parâmetro λ e da matriz de pesos, foi aplicado após a equalização das imagens. Foi obtido $\lambda = 0,16665$ e os resultados da classificação se encontram na Tabela 4.15.

Tabela 4.14: Taxa de acerto ORL com equalização das imagens.

Grupo	Acertos
Grupo_a	97,50%
Grupo_b	97,50%
Grupo_c	96,25%
Grupo_d	98,75%
Grupo_e	97,50%
Média	97,50%

Tabela 4.15: Taxa de acerto ORL com equalização das imagens e AG.

Grupo	Acertos
Grupo_a	100,00%
Grupo_b	96,25%
Grupo_c	98,75%
Grupo_d	98,75%
Grupo_e	97,50%
Média	98,25%

Em (Barnouti, 2016) e (Ferraz et al., 2017) é relatada uma taxa de 97,5% de acurácia com a base ORL.

4.3.5 YALE

As informações sobre essa base de dados encontram-se na subseção 2.5.5.

a) Validação Cruzada

Nesta BD tem-se um total de 165 imagens sendo 11 de cada pessoa, totalizando 15 classes. Devido a alta variação de iluminação e objetos de uma imagem para outra, não foram excluídas nenhuma figura durante a validação, porém, os grupos de validações foram composto com 9 imagens para treino e duas para teste.

b) Resultados

Como um primeiro resultado, a Figura 4.9 e a Tabela 4.16 mostram a acurácia do método para uma e para 12 escalas, respectivamente.

Tabela 4.16: Taxa de acerto YALE com 12 escalas.

Grupo	Acertos
Grupo_a	76,67%
Grupo_b	100,00%
Grupo_c	100,00%
Grupo_d	56,67%
Grupo_e	100,00%
Média	86,67%

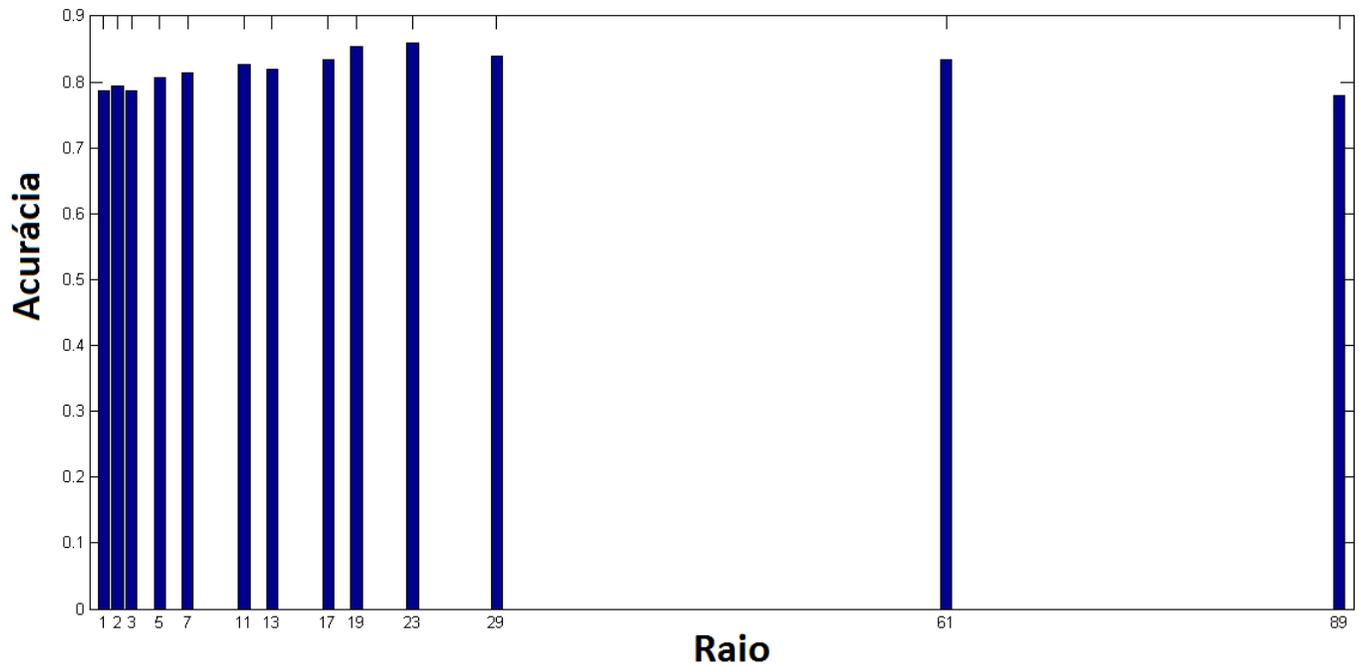


Figura 4.9: Cálculo da acurácia da YALE para uma escala.

A acurácia para duas escalas utilizando os raios , $R = \{23, 29\}$, são mostrados na Tabela 4.17. Durante os testes realizados notou-se que a acurácia permaneceu a mesma quando foram utilizados 12 escalas.

Tabela 4.17: Taxa de acerto YALE utilizando os raios 23 e 29.

Grupo	Acertos
Grupo_a	76,67%
Grupo_b	100,00%
Grupo_c	100,00%
Grupo_d	63,33%
Grupo_e	93,33%
Média	86,67%

Nesta base, melhores resultados foram obtidos utilizando 5 amostras na matriz de pesos. Essas 5 amostras foram distribuídas uniformemente com relação ao pixel central da região. A matriz M (Equação 4.3) exemplifica o caso para $R = 3$, que gera uma matriz 7×7 .

$$M = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}. \quad (4.3)$$

A Tabela 4.18 mostra a acurácia para o reconhecimento considerando 5 amostras para extração e 12 escalas do método.

Tabela 4.18: Taxa de acerto YALE utilizando 5 amostras para extração em 12 escalas.

Grupo	Acertos
Grupo_a	76,67%
Grupo_b	100,00%
Grupo_c	100,00%
Grupo_d	63,33%
Grupo_e	93,33%
Média	86,67%

Neste BD foi realizado o ajuste dos níveis de cinza proposto na Subseção 3.9.2. No entanto, apesar de ele ter melhorado a acurácia do método, a visualização foi prejudicada pelo escurecimento do fundo branco em alguns casos, como mostra a Figura 3.11. Ao analisar o vetor de características dessas imagens, foi notado que, o código *LMP* obtido desse fundo é o valor que aparece com maior frequência nas imagens. Uma melhora nos resultados foi observada quando esse padrão foi desconsiderado durante as extrações de características.

A combinação desse ajuste e da exclusão do padrão gerado pelo fundo durante os cálculos, obteve a maior acurácia da YALE. As Tabelas 4.19 e 4.20 apresentam os resultados com as técnicas descritas anteriormente.

Tabela 4.19: Taxa de acerto YALE através da alteração dos níveis de cinza.

Grupo	Acertos
Grupo_a	80,00%
Grupo_b	100,00%
Grupo_c	96,67%
Grupo_d	73,33%
Grupo_e	96,67%
Média	89,33%

Tabela 4.20: Taxa de acerto YALE através da alteração dos níveis de cinza e exclusão dos padrões gerados pelo fundo das imagens.

Grupo	Acertos
Grupo_a	83,33%
Grupo_b	100,00%
Grupo_c	96,67%
Grupo_d	73,33%
Grupo_e	100,00%
Média	90,66%

4.4 Considerações Finais

O algoritmo genético também foi aplicado na YALE, porém, a capacidade de generalização encontrada pelas matrizes ótima foi baixa, isto é, como mencionado na seção de metodologias,

a matriz não melhora todas os grupos de validação da base. Outro fator que influenciou nos teste do AG pra a MUCT foi o tempo de evolução de cada geração, devido a quantidade e a alta dimensão das imagens da base, exigindo um grande esforço computacional.

O motivo pelo qual o método funciona melhor para raios maiores que 1 (até um determinado raio, pois a partir deste os valores começam a decair) é devido ao baixo poder discriminativo de padrões extraídos de regiões pouco abrangentes, como é o caso de uma região 3×3 . A Tabela 4.21 mostra a soma da frequência dos “uniform patterns” das imagens da base ORL, para os raios 1, 2, 5, 17, 23 e 29.

Tabela 4.21: Média da soma da frequência dos “uniform patterns” extraídos das imagens da ORL para diferentes valores de raio.

Raio	Média
1	0,0032
2	0,0036
5	0,0041
17	0,0061
23	0,0065
29	0,0069

Nota-se que, quanto maior o raio, maior será a frequência das principais características das imagens, porém, seguindo lógica as acurácias nos raios 23 e 29 seriam maiores que a do raio 17, mas não é o que acontece (Figura 4.8), apesar dos raios maiores extraírem mais “uniform patterns” eles possuem menos regiões de extração quando comparados à raios menores. Essa quantidade menor de regiões para extrair características pode diminuir a acurácia do método. Por exemplo, através de uma álgebra simples pode-se calcular quantas regiões de dimensão $n \times n$ cabem em uma imagem de dimensão $X \times Y$, para o caso da ORL tem-se $X = 112$ e $Y = 92$, e assim existem 4.524 e 3.036 regiões para extração de características, utilizando os raios 17 e 23, respectivamente. A diferença de quase 1.500 regiões teve um impacto negativo na taxa de reconhecimento, visto que a acurácia do método ficou prejudicada com a utilização do raio 23, como pode ser observado na Figura 4.12.

Considerações Finais

Neste Capítulo são apresentadas as conclusões, as principais contribuições deste trabalho, as possibilidades de trabalhos futuros e a relação de artigos aceitos e publicados.

5.1 Conclusões

Nesta pesquisa foram apresentados e discutidos resultados sobre a acurácia do Padrão Mapeado Localmente em Multiescala (*MSLMP*) quanto ao reconhecimento de faces. As bases de dados utilizadas para a avaliação de performance do método foram a ESSEX, JAFFE, MUCT, ORL e YALE.

A fim de melhorar alguns resultados foi utilizado o algoritmo genético para seleção de parâmetros. Além disso, foram propostas duas técnicas de reajuste de níveis de cinza. Para avaliar a acurácia do método foi utilizado o procedimento de validação cruzada.

O *MSLMP* é robusto para classificação de imagens em ambientes controlados, além disso, obtém melhores resultados quando usadas regiões maiores da imagem para extração de característica devido à baixa variação intra-classe de características extraídas de regiões de raio menores que 5. Esse método não se demonstrou tão robusto quanto à alta variação de iluminação, no entanto, houve uma melhora na acurácia nos bancos com esse perfil quando técnicas de pré-processamento foram aplicadas, concluindo-se então que o método pode atingir boa acurácia caso o BD seja submetido a sofisticadas técnicas de processamento de imagens.

Em geral, o método mostrou grande potencial para o reconhecimento facial tornando-se competitivo ou mesmo superior a métodos disponíveis na literatura.

5.2 Contribuições

As contribuições deste trabalho foram apresentar uma discussão quanto à utilização do *MSLMP* para o reconhecimento de faces e usar técnicas de otimização na escolha de parâmetros para melhorar seus resultados. Essas contribuições podem ser divididas nas seguintes partes:

- análise do desempenho do método *MSLMP* quanto à classificação de imagens digitais de faces;
- apresentação de uma técnica de pré-processamento baseada na média dos níveis de cinza das imagens do banco de dados que apresentam alta variação de iluminação e poses;

- apresentação de uma técnica de pré-processamento baseada na quantidade de pixels de alta intensidade das imagens;
- análise da utilização do Algoritmo Genético para obter melhores parâmetros do método a fim de aumentar a taxa de reconhecimento.

5.3 **Trabalhos Futuros**

De imediato, sugere-se as seguintes frentes de trabalho:

- (Cai, Xiao, Wang & Zhang, 2011) apresenta uma técnica baseada na teoria de aprendizado de métrica, o intuito do método é de encontrar uma matriz que maximize a distância entre dois vetores de características de classes diferentes e minimize os de classes iguais. Pode-se aplicar este algoritmo usando como dados as matrizes de treino dos padrões *LMP*. Isto seria interessante, pois esse método foi desenvolvido exclusivamente para trabalhar com processos de classificação que se baseiam em histogramas;
- pode-se também explorar a técnica de ajuste dos níveis de cinza proposta para a base de dados YALE. A técnica apresenta melhoras significativas e, talvez, se melhor explorada pode trazer ganhos não só para a base de dados YALE, mas também para outras bases com características similares;
- a base de dados Face Recognition Technology (FERET) é uma renomada base presente na literatura dos métodos de reconhecimento facial. A FERET é financiada pelo Departamento de Defesa dos Estados Unidos da América. O objetivo desta base é o de avaliar novos algoritmos de reconhecimento automático de faces que possam ser usados em aplicações de segurança, inteligência e criminais. Pode-se considerar avaliar a performance do *MSLMP* nesta base e aplicar técnicas de otimização para melhorar a acurácia dos resultados, caso seja preciso. Para acesso a esse BD o pesquisador deve criar um usuário no site NIST, concordar com os termos de uso e submeter um pequeno resumo da sua proposta de trabalho com a base, o link para acesso está disponível em (Phillips, 2011);
- aplicação do método para reconhecimento de íris e de faces fraudulentas, por exemplo.

5.4 **Artigos Aceitos e Publicados**

- SILVA, E. M.; BOAVENTURA, M.; BOAVENTURA, I. A. G., *Face Recognition Based on a Multi-Scale Local Mapped Pattern Method*. In: XIII Workshop de Visão Computacional, 2017, Natal-RN. Anais do XIII Workshop de Visão Computacional. Natal-RN: EDUFRN, 2017. p. 82-87.
- SILVA, E. M.; BOAVENTURA, M.; BOAVENTURA, I. A. G., *Correção do brilho e homogeneização de imagens aplicados ao reconhecimento facial*. V Encontro Regional de Matemática Aplicada e Computacional 2018. Bauro-SP. (**Aceito**)
- SILVA, E. M.; BOAVENTURA, M.; BOAVENTURA, I. A. G.; CONTRERAS, R. C., *Face Recognition using Local Mapped Patterns and Genetic Algorithms*. International Conference on Pattern Recognition and Artificial Intelligence. Union-NJ. (**Aceito**)

Referências

- Ahn, C. K. (2005). *Advances in Evolutionary Algorithms Theory, Design and Practice* (Springer, Ed.).
- Baker, J. E. (1985). Adaptive selection methods for genetic algorithms. *International Conference on Genetic Algorithms and Their Applications*. 101–111.
- Barnouti, N. H. (2016). Improve Face Recognition Rate Using Different Image Pre-processing Techniques. *2016 American Journal of Engeneering Research(AJER)*, 5(4), 43–53.
- Barnouti, N. H., Matti, W. E., Al-Dabbagh, S. S. M. & Naser, M. A. S. (2016). Face Detection and Recognition Using Viola-Jones with PCA-LDA and Square Euclidean Distance. *Internacional Journal of Advanced Computer Science and Applications*, 7(5), 371–377.
- Booker, L. B. (1982). *Intelligent behavior as an adaptation to the task environment* (tese de doutorado, Universidade de Michigan).
- Bravo, M. J. A. (2016). *Análise do Descritor de Padrões Mapeados Localmente em Multiescala para classificação de textura em imagens digitais* (diss. de mestrado, Universidade Estadual Paulista “Júlio de Mesquita Filho”, São José do Rio Preto).
- Brindle, A. (1981). *Genetic algorithms for function optimization* (tese de doutorado, Universidade de Alberta).
- Buciu, I. & Gacsadi, A. (2016). Biometric Systems and Technologies: A Survey. *INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL*, 11(3), 315–330.
- Cai, X., Xiao, B., Wang, C. & Zhang, R. (2011). Quadratic-chi similarity metric learning for histogram feature. *Pattern Recognition (ACPR), 2011 First Asian Conference on*.
- Chan, C. H. (2008). *Multi-scale Local Binary Pattern Histogram for Face Recognition* (tese de doutorado, Universidade de Surrey).
- Che, L. Y. (2017). China, Russia Put Millions in This Startup to Recognize Your Face. Recuperado de <https://www.bloomberg.com/news/articles/2017-10-31/chinese-facial-recognition-startup-is-said-to-raise-460-million>
- Chelali, F. Z., Djeradi, A. & Djeradi, R. (2009). Linear discriminant analysis for face recognition. *Multimedia Computing and Systems*.
- Chihaoui, M., Elkefi, A., Bellil, W. & Amar, C. B. (2016). A Survey of 2D Face Recognition Techniques. *Computers*.
- China’s Face Recognition System Serves Olympics. (2008). Recuperado de http://english.china.com/zh_cn/Olympic/spotlight/11068320/20080813/15025578.html
- Contreras, R. C. (2015). *Detecção de impressões digitais fraudulentas utilizando padrões mapeados localmente em multiescala* (diss. de mestrado, Universidade Estadual Paulista “Júlio de Mesquita Filho”, São José do Rio Preto).

- Contreras, R. C., Boaventura, I. A. G. & Boaventura, M. (2015). Multi-scale Local Mapped Pattern for Image Texture Analysis. *XI Workshop de Visão Computacional, 2015, São Carlos, v. 11*.
- Dahmouni, A., Aharrane, N., Moutaoukil, K. E. & Satori, K. (2016). Face recognition using Local Binary Probabilistic Pattern (LBPP) and 2D-DCT frequency decomposition. *Comput. Graph. Imaging Vis. New Tech. Trends, CGiV 2016*, 73–77.
- Daugman, J. (1994). Biometric personal identification system based on iris analysis. Recuperado de <https://www.google.com/patents/US5291560>
- Etemad, K. & Chellapa, R. (1996). Face recognition using discriminant eigenvectors. *Proceedings of the International Conference on Acoustic, Speech and Signal Processing*.
- Ferraz, C. T., Junior, O. P. & Gonzaga, A. (2014). Feature description based on Mean Local Mapped Pattern. *Workshop de Visão Computacional - WVC, v. X*, 32–37.
- Ferraz, C. T., Manzato, M. G. & Gonzaga, A. (2017). Face Classification using a New Local Texture Descriptor. *23th Brazilian Symposium on Multimedia and the Web*, 237–240.
- Galton, F. (1888). Personal identification and description. *In Nature*, 173–177.
- Golfarelli, M., Maio, D. & Malton, D. (1997). On the error-reject trade-off in biometric verification systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7), 786–796.
- Huang, J., Yuen, P. C., Lai, J. H. & Li, C. (2004). Face Recognition Using Local and Global Features. *EURASIP Journal on Applied Signal Processing*, 4, 530–541.
- Jain, A. K. (2005). Biometric Recognition: How Do I Know Who You Are? *13 International Conference Image Analysis and Processing Proceedings*.
- Jain, A. K., Flynn, P. & Ross, A. (2008). *Handbook of Biometrics* (S.-V. US, Ed.). Springer US.
- Jain, A. K., Ross, A. & Prabhakar, S. (2004). An Introduction to Biometric Recognition. *IEEE Transactions on Circuits and Systems and Videos Technology*, 14(1), 4–20.
- Jain, R. & Kant, C. (2015). Attacks on Biometric Systems: An Overview. *International Journal of Advances in Scientific Research*, 1(7), 283–288.
- Jalled, F. (2017). Face Recognition Machine Vision System Using Eigenfaces. Recuperado de <https://dblp.org/rec/bib/journals/corr/Jalled17>
- Jia, X., Yang, X., Cao, K., Zang, Y., Zhang, N., Dai, R. & Z., X. (2014). Multi-scale local binary patterns with filters for spoof fingerprint detection. *Information Science, v. 268*, 91–102.
- Kasar, M. M., Bhattacharyya, D. & Kim, T. H. (2016). Face Recognition Using Neural Network: A Review. *International Journal of Security and Its Applications*, 10(3), 81–100.
- Kim, K. (2001). Face Recognition using Principle Component Analysis. Recuperado de <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.383.6655&rep=rep1&type=pdf>
- Lades, M., Vorbruggen, J., Buhmann, J., Lange, J., Von Der Malsburg, C., Wurtz, R. & Konen, W. (1993). Distortion invariant object recognition in the dynamic link architecture. *IEEE Trans Comput*, 42, 300–311.
- Le, C. (2011). A Survey of Biometrics Security Systems. Recuperado de <http://www.cse.wustl.edu/~jain/cse571-11/ftp/biomet/index.html>
- Levine, M. D. & Yu, Y. (2006). Face recognition subject to variations in facial expression, illumination and pose using correlation filters. *Computer Vision and Image Understanding*, 104(1), 1–15.
- Li, S. Z. & Janil, A. K. (Eds.). (2011). *Handbook of Face Recognition* (2ª ed.). Springer.

- Liao, S., Zhu, X., Lei, Z., Zhang, H. & Li, S. Z. (2007). Learning Multi-scale Block Local Binary Pattern for face recognition. *Adv. Biometrics*, 828–837.
- Liu, L., Fieguth, P., G., Z., Pietikäinen, M. & Hu, D. (2016). Extended Local binary Pattern for face recognition. *Inf. Sci (Ny)*, 56–72.
- Lyons, M. J., Akemastu, M., S. Kamachi & Gyoba, J. (1998). Coding Facial Expressions with Gabor Wavelets. *3rd IEEE International Conference on Automatic Face and Gesture Recognition*, 200–205.
- Mäenpää, T. & Pietikäinen, M. (2003). Multi-scale binary patterns for texture analysis. *Machine Vision Group*, 885–892.
- Marco, G., Massimo, L., Vincenzo, P., Daniele, S. & Fabio, S. (2004). Accuracy and performance of biometric systems. *Instrumentation and Measurement Technology Conference*.
- Meethongjan, K. & Mohamad, D. (2007). A Summary of literature review : Face Recognition. *Postgraduate Annual Research Seminar*.
- Milborrow, S., Morkel, J. & Nicolls, F. (2010). The MUCT Landmarked Face Database. *Pattern Recognition Association of South Africa*. Recuperado de <http://www.milbo.org/muct>
- Miranda, M. N. (s.d.). Algoritmos Genéticos: Fundamentos e Aplicações. Recuperado de <http://www.nce.ufrj.br/GINAPE/VIDA/alggenet.htm>
- Muhlenbein, H. & Schlierkamp-Voosen, D. (1993). Predictive models for the breeder genetic algorithm: I. Continuous parameter optimization. *Evolutionary Computation*.
- Ojala, T., Pietikäinen, M. & Harwood, D. (1996). A comparative study of texture measures with classification based on features distributions. *Pattern Recognition*, v. 29, 51–59.
- Ojala, T., Pietikäinen, M. & Mäenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *In: IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 24, n. 7, 971–987.
- Pele, O. & Werman, M. (2010). The Quadratic-chi Histogram Distance Family. *Proceedings of the 11th European Conference on Computer Vision*, 749–762.
- Peng, Y., Wang, S., Long, X. & Lu, B. (2015). Discriminative graph regularized extreme learning machine and its application to face recognition. *Neurocomputing*149, 340–353.
- Philips, P. J., Martin, A., C.L. Wilson, L. & Przybocki, M. (2000). Introduction to Evaluating Biometric Systems. *National Institute of Standards and Technology*, 56–63.
- Phillips, J. a. (2011). Face Recognition Technology (FERET). Recuperado de <https://www.nist.gov/programs-projects/face-recognition-technology-feret>
- Rahim, A., Hossain, N., Wahid, T. & Azam, S. (2013). Face Recognition using Local Binary Patterns (LBP). *Global Journal of Computer Science and Technology Graphics & Vision*, 13(4).
- Ran, R., Fang, X., B. Wu & Zhang, S. (2018). A Simple and Effective Generalization of Exponential Matrix Discriminant Analysis and Its Application to Face Recognition. *IEICE TRANS. INF. & SYST. E101-D(1)*, 265–268.
- Roulette wheel selection. (s.d.). Recuperado de <http://www.edc.ncl.ac.uk/highlight/rhjanuary2007g02.php>
- Schlett, T., Rathgeb, C. & Busch, C. (2016). A Binarization Scheme for Face Recognition based on Multi-scale Block Local Binary Pattern. doi:<https://doi.org/10.1109/BIOSIG.2016.7736901>
- Sirovich, L. & Kirby, M. (1987). Low-dimensional procedure for the characterization of human faces. *Journal of the optical society of America*, 4(3), 519–524.
- Srivastava, H. (2013). A Comparison Based Study on Biometrics for Human Recognition. *IOSR Journal of Computer Engineering*, 15(1), 22–29.

- Taouche, C., Berkane, M., Batouche, M. C. & Talebahmed, A. (2014). Multimodal Biometric Systems. *Multimedia Computing and Systems (ICMCS)*.
- Tolba, A. B., El-Baz, A. & A.A., E.-H. (2005). Face Recognition: A Literature Review. *International Journal of Signal Processing*, 2(2), 88–103.
- Turk, M. & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1), 71–86.
- Wang, Y. (2014). An Analysis of the Viola-Jones Face Detection Algorithm. *Image Processing On Line*.
- Yu, X. & Gen, M. (2010). *Introduction to Evolutionary Algorithms* (Springer, Ed.). Springer London Dordrecht Heidelberg New York.
- Zhang, H. (2011). Image Preprocessing Methods in Face Recognition. Em *Simpósio Photonics and Optoelectronics (SOPO)*.
- Zhang, T. P., Fang, B., Tang, Y., Shang, Z. & B. Xu, B. (2010). Generalized discriminant analysis: A matrix exponential approach. *IEEE Trans. Syst. Man, Cybern. B, Cybern*, 40(1), 186–197.

TERMO DE REPRODUÇÃO XEROGRÁFICA

Autorizo a reprodução xerográfica do presente Trabalho de Conclusão, na íntegra ou em partes, para fins de pesquisa.

São José do Rio Preto, ____/____/____

Assinatura do autor