



UNIVERSIDADE FEDERAL DA PARAÍBA
Programa de Pós-Graduação em Informática

Rodrigo Loureiro Parente

Algoritmos para análise da qualidade de imagens faciais de acordo com requisitos do padrão ISO/ICAO

João Pessoa, PB
Agosto 2017

Algoritmos para análise da qualidade de imagens faciais de acordo com requisitos do padrão ISO/ICAO

Rodrigo Loureiro Parente

Dissertação apresentada à Universidade Federal da Paraíba para obtenção de Título de Mestre em Ciência da Computação, na área de Sinais, Sistemas Digitais e Gráficos.

Orientador: Leonardo Vidal Batista

João Pessoa, PB
Agosto 2017

Catálogo na publicação
Seção de Catalogação e Classificação

P228a Parente, Rodrigo Loureiro.

Algoritmos para análise da qualidade de imagens faciais de acordo com requisitos do padrão ISO/IEC / Rodrigo Loureiro Parente. - João Pessoa, 2017.

93 f. : il.

Orientação: Leonardo Vidal Batista.

Dissertação (Mestrado) - UFPB/CT/PPGI.

1. Ciência da Computação. 2. Qualidade da imagem facial - processamento. 3. ICAO - Organização internacional de aviação civil. 4. ISO - Organização internacional de padronização. 5. ISO/IEC 19794-5. I. Batista, Leonardo Vidal. II. Título.

UFPB/BC



UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA



Ata da Sessão Pública de Defesa de Dissertação de Mestrado de Rodrigo Loureiro Parente, candidato ao título de Mestre em Informática na Área de Sistemas de Computação, realizada em 24 de agosto de 2017.

1 Aos vinte e quatro dias do mês de agosto, do ano de dois mil e dezessete, às quatorze
2 horas, no Centro de Informática da Universidade Federal da Paraíba, em Mangabeira,
3 reuniram-se os membros da Banca Examinadora constituída para julgar o Trabalho Final do
4 Sr. Rodrigo Loureiro Parente, vinculado a esta Universidade sob a matrícula nº 2015103984,
5 candidato ao grau de Mestre em Informática, na área de "Sistemas de Computação", na linha
6 de pesquisa "Computação distribuída", do Programa de Pós-Graduação em Informática, da
7 Universidade Federal da Paraíba. A comissão examinadora foi composta pelos professores:
8 Leonardo Vidal Batista (PPGI-UFPB), Orientador e Presidente da Banca, Clairton de
9 Albuquerque Siebra (PPGI-UFPB), Examinador Interno, Thais Gaudêncio do Rego (PPGI-
10 UFPB), Examinador interno, e Carlos Danilo Miranda Regis (IFPB), Examinador externo à
11 instituição. Dando início aos trabalhos, o Presidente da Banca, cumprimentou os presentes,
12 comunicou aos mesmos a finalidade da reunião e passou a palavra ao candidato para que o
13 mesmo fizesse a exposição oral do trabalho de dissertação intitulado "Algoritmos para
14 análise da qualidade de imagens faciais de acordo com requisitos do padrão ISO/IEC 30108".
15 Concluída a exposição, o candidato foi arguido pela Banca Examinadora que emitiu o
16 seguinte parecer: "**aprovado**". Do ocorrido, eu, Ruy Alberto Pisani Altafim, Vice-Coordenador
17 do Programa de Pós-Graduação em Informática, lavrei a presente ata que vai assinada por
18 mim e pelos membros da banca examinadora. João Pessoa, 24 de agosto de 2017.


Prof. Dr. Ruy Alberto Pisani Altafim

Ruy Alberto Pisani Altafim
Vice-Coordenador do Programa de
Pós-Graduação em Informática
SIAPE 1971934

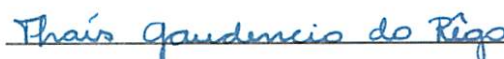
Prof. Dr. Leonardo Vidal Batista
Orientador (PPGI-UFPB)



Prof. Dr. Clairton de Albuquerque Siebra
Examinador interno (PPGI-UFPB)



Prof. Dr. Thais Gaudêncio do Rego
Examinador interno (PPGI-UFPB)



Prof. Dr. Carlos Danilo Miranda Regis
Examinador externo à instituição (IFPB)



A face humana apresenta
Muita variabilidade
Computador agora tenta
Adivinhar sua identidade
Mas é difícil acertar
Sem antes constatar
Que a foto tem qualidade.

Rodrigo Loureiro Parente

Agradecimentos

Agradeço primeiramente aos meus pais, qqMabel e Marcos Parente, por todo suporte, estímulo e amor. Agradeço loucamente à mãe natureza pela centelha da vida e pela oportunidade única de viver neste mundo mágico, tão degradado pela humanidade, suas crenças e sistemas ultrapassados. Agradeço à minha madrinha Tia Marialba, que incentivou na minha infância meu gosto por desenho e pintura, o que virou aptidão para arte digital, Photoshop etc. Tal aptidão virou profissão quando aprendi processamento de imagens digitais. Sinto gratidão especial ao meu grande parceiro de trabalho Iron Araújo, através de quem tive a oportunidade de ingressar no grupo de pesquisa do qual este trabalho é fruto. Com Iron, aprendi muito sobre análise de desempenho de software biométrico e também que não se pode dar cabimento a todo mundo. Através de sua indicação, teve início minha aventura no incrível mundo da ciência, até então desconhecido. Agradeço ao meu orientador, Prof. Leonardo Vidal Batista, pela inspiração e enorme quantidade de conhecimento passada a mim com toda a satisfação do mundo. Agradeço a Igor Andrezza, colega de trabalho e superior imediato em minhas funções no projeto BioPass ICAO da VSOFTE. Igor me orientou com o conhecimento de um sábio e paciência de motorista de ônibus em todo o meu trabalho no projeto de pesquisa. Igualmente agradeço a Erick Vagner, pelo acompanhamento e dicas pertinentes referentes à minha pesquisa no ICAO. Sou grato a Rajiv Albino, que me trouxe à luz todas as nuances do domínio das frequências. Agradeço à VSOFTE Tecnologia como um todo, pela boa acolhida, em especial ao chefe Pedro Alves por acreditar que incentivo à ciência é fundamental para o progresso de qualquer empresa ou sociedade. Finalmente agradeço a todos que, direta ou indiretamente, contribuíram para a realização deste trabalho: Adriano Marinho e meus bons amigos Victor Souza e Lígia Rosa etc.

Resumo

A face é uma das características físicas mais usadas para reconhecimento biométrico. A confirmação de identidade normalmente é feita por um ser humano através de documentos oficiais, tarefa que demanda tempo, recursos humanos e impede a agilidade de certos cenários como aeroportos etc. Assim, a ICAO, *International Civil Aviation Organization* (do inglês, Organização Internacional de Aviação Civil), iniciou um projeto para validação automática de identidade pessoal através de passaporte eletrônico. Entretanto, a precisão do reconhecimento biográfico de face depende diretamente da qualidade da amostra facial fornecida. Em seguida, a ISO, ou *International Standardization Organization* (do inglês, Organização Internacional de Padronização) desenvolveu um padrão internacional de qualidade para imagem facial a ser usado em passaportes eletrônicos. Este trabalho visa apresentar algoritmos para avaliação da qualidade de imagens faciais de acordo com os quatro requisitos do padrão ISO/ICAO. A verificação do nível de embaçamento é feita explorando a informação das altas frequências, calculando a variância dos pixels após filtragem passa-altas de Sobel. O algoritmo proposto obteve taxa de erro de 6,2%. O método de análise da presença de cabelo em frente aos olhos é avaliado através de detecção de bordas verticais na região ocular para detectar fios de cabelo na região dos olhos. O algoritmo proposto obteve taxa de erro de 11,9%. A análise da presença de chapéu é realizada através de segmentação de pele na região da testa. O algoritmo proposto obteve taxa de erro de 7,2%. Finalmente, o requisito da análise olhos encobertos por armação de óculos é averiguado através da detecção de bordas horizontais e detecção de linhas na região dos olhos. O método para detecção do encobrimento dos olhos por armação de óculos busca detectar segmentos de reta horizontais, causados pela armação, na região dos olhos. O algoritmo proposto obteve taxa de erro de 11,6%.

Palavras-chave: Qualidade de imagem facial, ICAO, ISO/IEC 19794-5.

Abstract

The face has traditionally been used in identity documents for visual recognition of a person, being this way one of the most used physical characteristics for biometric recognition. The task of confirming identity with traditional documents is done visually by a human being. Such a task demands time, human resources and precludes the agility of critical scenarios such as airports etc. In this context, the International Civil Aviation Organization (ICAO) has initiated a project for automatic validation of personal identity through an electronic passport. However, the accuracy of facial recognition is directly dependent on the quality of the captured face samples. Later, the International Organization for Standardization (ISO) has developed an international quality standard for facial imaging to be used in electronic passports. This work aims to present algorithms for evaluating the quality of facial images according to the four ISO/ICAO standard requirements described ahead. The blurring level analysis is performed by exploring the high frequency information, calculating the variance of the pixels after Sobel high-pass filtering. This proposed method achieved an error rate of 6,2%. The requirement for analysis of the presence of hair in front of the eyes is evaluated by detecting borders in the eyes region, aiming to detect hair threads in front of the eyes. This proposed method achieved an error rate of 11,9%. The detection of the presence of a hat/cap is performed through skin segmentation in the forehead region. This proposed method achieved an error rate of 7,2%. Finally, the detection of eyeglass frame covering eyes is ascertained by detecting horizontal edges and detecting straight line segments in the eye region. This proposed method achieved an error rate of 11,6%.

Keywords: Facial image quality, ICAO, ISO/IEC 19794-5.

Lista de Figuras

2.1	Exemplo de imagem binária. A Figura 2.1a exibe a famosa lena em cores, 2.1b apresenta a lena em níveis de cinza. A Figura 2.1c exibe a binarização de 2.1b, onde os pixels brancos (primeiro plano) representam a região de 2.1b com nível de cinza maior que 150. Este processo é chamado binarização.	8
2.2	Ilustração do experimento de Newton no século XVII.	9
2.3	Comparação entre o espectro de luz visível e o espectro eletromagnético como um todo. (Fonte: Gonzalez e Woods (2008))	9
2.4	Cores primárias do sistema RGB e suas respectivas combinações.	10
2.5	Ilustração de uma máscara 3x3 centrada num pixel (x,y)	14
2.6	Conversão para o domínio das frequências: 2.6a exibe a imagem de entrada; 2.6b mostra o espectro de frequências; 2.6c ilustra a localização das baixas frequências perto do centro; 2.6d ilustra a localização das frequências altas.	15
2.7	Exemplos de máscaras: 2.7a e 2.7c exibem máscaras de suavização e realce, respectivamente; 2.7b e 2.7d exibem o resultado da filtragem da Imagem 2.6a com uso das máscaras 2.7a e 2.7c, respectivamente.	16

2.8	Representações da função Gaussiana: 2.8a apresenta a distribuição da função para uma variável com valor de $\sigma = 1$; 2.8b apresenta a distribuição para duas variáveis com valor de $\sigma = 1$; 2.8c apresenta uma discretização em forma de matriz de convolução com valor de $\sigma = 1,4$	17
2.9	Matrizes de convolução utilizadas pelo filtro de Sobel. 2.9a apresenta a matriz que calcula a componente horizontal G_x do gradiente e 2.9b, a componente vertical G_y	18
2.10	Exemplos da execução do filtro de Sobel. A Figura 2.10a exibe uma imagem com bordas bem definidas e 2.10b ilustra o resultado da execução do filtro de Sobel. A Figura 2.10c exibe uma imagem que apresenta bordas suavizadas e 2.10d mostra o resultado da execução do filtro.	18
2.11	Exemplo de execução do filtro de Canny. 2.11a apresenta a imagem de entrada e 2.11b, a imagem filtrada.	19
2.12	Exemplos das execuções dos filtros Log e LogAbout: 2.12a exibe a imagem de entrada; 2.12b exibe o resultado da execução do filtro Log; 2.12c exibe o resultado da execução do filtro LogAbout.	20
2.13	Filtros de frequências. 2.13a e 2.13c apresentam filtros passa-baixas. 2.13b e 2.13d apresentam filtros passa-altas. Os filtros 2.13a e 2.13b são chamados ideais e os 2.13c e 2.13d são gerados pela função Gaussiana.	22
2.14	Exemplo de filtragem no domínio da frequência: a Figura 2.14b exibe a imagem da Figura 2.6a filtrada com a máscara passa-baixas de 2.14a; a Figura 2.14d exibe o resultado da filtragem da mesma imagem pela máscara passa-altas de 2.14c.	22
2.15	Resultado da execução do detector de faces Viola-Jones numa imagem com diversas faces (Fonte: Viola e Jones (2004)).	24
2.16	Resultado da execução do detector de elementos faciais de Kazemi e Sullivan em faces diferentes (Fonte: Kazemi e Sullivan (2014)).	25

2.17	Exemplo da execução das operações de dilatação e erosão de pixels. 2.17a exibe uma imagem binária onde os pixels brancos correspondem ao primeiro plano. A imagem 2.17b mostra o resultado da dilatação, enquanto que a 2.17c mostra o resultado da operação de erosão.	27
3.1	Mapas gerados para representação das quatro principais regiões da imagem: mapa de pele (3.1a); mapa de roupa (3.1b); de plano de fundo (3.1c); mapa de cabelo (3.1d). (Fonte: Ferrara et al. (2012a))	29
3.2	Etapas da expansão e determinação das regiões principais a partir da segmentação de baixo nível (3.2a) até a separação das quatro regiões principais (3.2d). (Fonte: Ferrara et al. (2012a))	30
3.3	Processo de detecção de oclusão facial através de segmentação de pele. Processo exibido para imagens com oclusão da face (3.3a e 3.3b) e sem oclusão (3.3c e 3.3d). (Fonte: Storer et al. (2010))	33
4.1	Exemplo de gráfico das curvas de taxa de falsa aceitação (FAR) e taxa de falsa rejeição. No cruzamento das duas está destacado o ponto de taxa de erro igual (EER).	37
4.2	Amostras de imagens do banco de faces AR: 4.2a, iluminação à esquerda; 4.2b, face neutra; 4.2c, uso de óculos; 4.2d, iluminação à direita (Fonte: Martinez (1998)).	38
4.3	Amostras de imagens do banco de faces FRGC: 4.3a, face com sombras; 4.3b, face neutra com iluminação controlada; 4.3c, presença de outras pessoas; 4.3d, face com sorriso.	38

- 4.4 Amostras de imagens do banco de faces PUT. A Figura 4.4a mostra a face voltada para a esquerda. As várias imagens de um mesmo indivíduo apresentam gradual variação horizontal de posição da face até a mesma estar voltada para a direita na Figura 4.4b. A Figura 4.4c mostra a face voltada para cima. Analogamente, a posição da face varia verticalmente até a mesma estar voltada para baixo, como na Figura 4.4d. 39
- 4.5 Delimitação da região dos olhos em função da região facial delimitada pelo Viola-Jones. X_0 e Y_0 são respectivamente a largura e altura da região facial. A origem da região ocular encontra-se no ponto $(1,5X_0, 1/3Y_0)$ e a mesma possui largura $L_o = 0,7X_0$ e altura $A_o = 1/6Y_0$ 42
- 4.6 Fluxo de execução do algoritmo de análise da presença de cabelo em frente aos olhos. 43
- 4.7 Delimitação da região da testa do indivíduo em função da região facial delimitada pelo Viola-Jones. X_0 e Y_0 são, respectivamente, a largura e altura da região facial. A origem da região da testa encontra-se no ponto $(1/5X_0, 0)$ e a mesma possui largura $L_t = 3/5X_0$ e altura $A_t = 1/4Y_0$ 45
- 4.8 Execução do algoritmo para uma face sem chapéu: 4.8a apresenta a imagem de entrada com a delimitação da face pelo Viola-Jones; 4.8b apresenta a face suavizada com a delimitação da região da testa; 4.8c apresenta a banda H do sistema HSV; 4.8d apresenta a binarização da banda H. 46
- 4.9 Execução do algoritmo para uma face com chapéu: 4.9a apresenta a imagem de entrada com a delimitação da face do Viola-Jones; 4.9b apresenta a face suavizada com a delimitação da região da testa; 4.9c apresenta a banda H do sistema HSV; 4.9d apresenta a binarização da banda H. 47

4.10	Processo de localização estimada dos olhos (a). Cálculo da Região S com relação à região da face detectada (b).	48
4.11	Principais etapas da execução do detector para imagem com e sem a presença de óculos: Seleção da Região S (4.11a e 4.11d), isolamento da banda Cr (4.11b e 4.11e), Binarização ($CrEdges_{bin}$) (4.11c e 4.11f). . .	49
4.12	Etapas principais do processo de análise do encobrimento dos olhos por armação de óculos (imagem com ocorrência do requisito). A primeira linha da figura exhibe um recorte colorido dos olhos. A segunda linha exhibe o recorte de cada olho em níveis de cinza. A terceira linha exhibe, à esquerda, o resultado da binarização do filtro de Sobel com expansão de histograma. À direita, a terceira linha exhibe o resultado da aplicação do filtro de Canny. A última linha da figura exhibe, destacadas em verde, as retas detectadas pela transformada de Hough, assim como a região superior da íris demarcada com um quadrilátero vermelho.	51
4.13	Etapas principais do processo de análise do encobrimento dos olhos por armação de óculos (imagem sem ocorrência do requisito). A primeira linha da figura exhibe um recorte colorido dos olhos. A segunda linha exhibe o recorte de cada olho em níveis de cinza. A terceira linha exhibe, à esquerda, o resultado da binarização do filtro de Sobel com expansão de histograma. À direita, a terceira linha exhibe o resultado da aplicação do filtro de Canny. A última linha da figura exhibe, destacadas em verde, as retas detectadas pela transformada de Hough, assim como a região superior da íris demarcada com um quadrilátero vermelho.	52
5.1	Gráfico de FAR x FRR para o requisito de análise do nível de foco. . . .	54
5.2	Gráfico de FAR x FRR para o requisito de análise da presença de cabelo em frente aos olhos.	55

5.3	Gráfico de FAR x FRR para o requisito de análise da presença de chapéu.	56
5.4	Gráfico de FAR x FRR para o requisito de encobrimento dos olhos por armação de óculos.	56
5.5	Exemplo de face que apresenta desfoque e reflexo na armação dos óculos. A Figura 5.5a apresenta a face desfocada enquanto que a 5.5b apresenta a detecção de bordas na região dos olhos.	57
5.6	Exemplo de imagem que apresenta pixelização e que o algoritmo de análise de foco rejeitou como se estivesse embaçada. A Figura 5.6a apresenta um recorte da face pixelizada enquanto que a 5.6b apresenta a detecção de bordas com o filtro de Sobel na região dos olhos demarcada.	58
5.7	Exemplo de imagem que apresenta pixelização e que o algoritmo de análise de foco rejeitou como se estivesse embaçada. A Figura 5.7a apresenta um recorte da face pixelizada enquanto que a 5.7b apresenta a detecção de bordas com o filtro de Sobel na região dos olhos demarcada.	58
5.8	Exemplos de fotos com alterações de iluminação onde ocorreram erros de falsa aceitação para o requisito de cabelo em frente aos olhos.	60
5.9	Etapas da execução do algoritmo para a imagem apresentada na Figura 5.8a. Recorte da região dos olhos (5.9a). Detecção de bordas com Canny (5.9b e 5.9c). Dilatação e erosão de bordas (5.9d e 5.9e). Operação XOR bit-a-bit (5.9f).	60
5.10	Etapas da execução do algoritmo para a imagem apresentada na Figura 5.8b. Recorte da região dos olhos (5.10a). Detecção de bordas com Canny (5.10b e 5.10c). Dilatação e erosão de bordas (5.10d e 5.10e). Operação XOR bit-a-bit (5.10f).	60

5.11 Etapas da execução do algoritmo para um caso de falsa rejeição. Recorte da região dos olhos (5.11a). Detecção de bordas com Canny (5.11b e 5.11c). Dilatação e erosão de bordas (5.11d e 5.11e). Operação XOR bit-a-bit (5.11f). 61

5.12 Exemplo de erro de falsa aceitação na detecção de chapéu. A Figura 5.12a apresenta a face do indivíduo, a Figura 5.12b ilustra o recorte da região da testa e a Figura 5.12c apresenta o resultado da detecção de pele. 62

5.13 Exemplo de erro de falsa aceitação na detecção de chapéu. A Figura 5.13a apresenta a face do indivíduo, a Figura 5.13b ilustra o recorte da região da testa e a Figura 5.13c apresenta o resultado da detecção de pele. 62

5.14 Exemplo de erro de falsa aceitação na detecção de chapéu. A Figura 5.14a apresenta a face do indivíduo, a Figura 5.14b ilustra o recorte da região da testa e a Figura 5.14c apresenta o resultado da detecção de pele. 63

5.15 Exemplo de erro de falsa aceitação na detecção de chapéu. A Figura 5.15a apresenta a face do indivíduo, a Figura 5.15b ilustra o recorte da região da testa e a Figura 5.15c apresenta o resultado da detecção de pele. 63

5.16 Exemplo de erro de falsa aceitação na detecção de chapéu. A Figura 5.16a apresenta a face do indivíduo, a Figura 5.16b ilustra o recorte da região da testa e a Figura 5.16c apresenta o resultado da detecção de pele. 64

5.17 Caso de erro de falsa aceitação. As Figuras 5.17b e 5.17c apresentam a detecção de bordas com Sobel para os olho direito e esquerdo, respectivamente, enquanto que as figuras 5.17d e 5.17e fazem o mesmo para detecção com canny. 65

5.18 Caso de erro de falsa aceitação. As Figuras 5.18b e 5.18c apresentam a detecção de bordas com Sobel para os olho direito e esquerdo, respectivamente, enquanto que as figuras 5.18d e 5.18e fazem o mesmo para detecção com Canny. 65

- 5.19 Caso de erro de falsa rejeição. As Figuras 5.19b e 5.19c apresentam a detecção de bordas com Sobel para os olho direito e esquerdo, respectivamente, enquanto que as figuras 5.19d e 5.19e fazem o mesmo para detecção com canny. 66
- 5.20 Caso de erro de falsa rejeição. As Figuras 5.20b e 5.20c apresentam a detecção de bordas com Sobel para os olho direito e esquerdo, respectivamente, enquanto que as figuras 5.20d e 5.20e fazem o mesmo para detecção com canny. 66

Lista de Tabelas

1.1	Descrição dos testes de qualidade de imagem facial realizados pelo Biolab ICAO, de acordo com o padrão ISO/IEC 19794-5.	5
5.1	Valores de EER e taxas de Rejeição (Rej.) dos softwares avaliados . . .	54

Lista de Abreviaturas e Siglas

Símbolo	Descrição
ICAO	<i>International Civil Aviation Organization</i>
ISO	<i>International Standardization Organization</i>
IEC	<i>International Electrotechnical Commission</i>
TI	<i>Tecnologia da Informação</i>
PDI	Processamento Digital de Imagens
DFT	<i>Discrete Fourier Transform</i>
EER	<i>Equal Error Rate</i>
FAR	<i>False Accept Rate</i>
FRR	<i>False Reject Rate</i>
FRGC	<i>Face Recognition Grand Challenge</i>
LFW	<i>Labeled Faces in the Wild</i>
OpenCV	<i>Open Computer Vision Library</i>
SDK	<i>Software Development Kit</i>

Sumário

1	Introdução	1
1.1	Objetivos	3
1.1.1	Objetivos Específicos	4
1.2	Estrutura da Dissertação	4
2	Fundamentação Teórica	6
2.1	Imagens Digitais e Processamento de Imagens	6
2.1.1	Tipos de imagem digital	7
2.2	As cores e suas formas de representação	8
2.2.1	Sistema de cores RGB	10
2.2.2	Sistema de cores YCbCr	11
2.2.3	Sistema de cores HSL	11
2.2.4	Uso das cores para detecção de pele em imagem digital	12
2.3	Filtros de imagem	13
2.4	Domínios do espaço e das frequências	14
2.4.1	Filtragem espacial	15
2.4.1.1	Filtro de suavização Gaussiana	16
2.4.1.2	Filtro de detecção de bordas de Sobel	17
2.4.1.3	Filtro de detecção de bordas de Canny	18
2.4.1.4	Filtro de normalização de iluminação - LogAbout	19

2.4.2	Convolução e o Teorema da Convolução	20
2.4.3	Filtragem no domínio da frequência	21
2.5	Detecção facial através do método Viola-Jones	23
2.6	Localização de elementos faciais	24
2.7	A transformada de Hough para detecção de linhas	25
2.8	Operações morfológicas: Erosão e Dilatação	26
3	Trabalhos Relacionados	28
4	Metodologia	34
4.1	Materiais	34
4.1.1	OpenCV	34
4.1.2	A plataforma Biolab-ICAO	35
4.1.2.1	Protocolo de testes do Biolab ICAO	36
4.1.2.2	Banco de faces AR	37
4.1.2.3	Banco de faces FRGC	38
4.1.2.4	Banco de faces PUT	38
4.1.2.5	Vsoft - A empresa	39
4.2	Métodos Propostos	40
4.2.1	Análise do Nível de Foco	40
4.2.2	Análise da presença de cabelo em frente aos olhos	42
4.2.3	Análise da presença de chapéu	44
4.2.4	Análise do encobrimento dos olhos por armação de óculos	47
4.2.4.1	Determinação da presença de óculos na imagem	47
4.2.4.2	Verificação do encobrimento dos olhos por armação de óculos	50
5	Avaliação Experimental	53

5.1	Resultados	53
5.2	Análise dos erros dos métodos propostos	57
5.2.1	Erros na aferição do nível de foco	57
5.2.1.1	Falsas aceitações	57
5.2.1.2	Falsas rejeições	58
5.2.2	Erros na detecção de cabelo encobrindo os olhos	59
5.2.2.1	Falsas aceitações	59
5.2.2.2	Falsas rejeições	61
5.2.3	Erros na detecção de chapéu ou boné	61
5.2.3.1	Falsas aceitações	61
5.2.3.2	Falsas rejeições	62
5.2.4	Armação encobrindo olhos	64
5.2.4.1	Falsas aceitações	64
5.2.4.2	Falsas rejeições	65
6	Considerações Finais e Trabalhos Futuros	67
	Referências Bibliográficas	69

Capítulo 1

Introdução

Segundo Ferrara et al. (2012b), a face tem sido usada tradicionalmente em documentos de identidade para reconhecimento visual de uma pessoa, e por isso representa uma das características físicas mais usadas para reconhecimento biométrico. Em comparação com outros atributos físicos, a face apresenta algumas vantagens: a aquisição das características faciais via fotografia digital é não intrusiva, pode ser realizada à distância e não requer cooperação do indivíduo retratado; não necessita de hardware adicional, o qual pode ser fonte de contaminação interpessoal por germes, o que acontece com equipamento de escaneamento de impressão digital ou leitura de íris.

A tarefa de confirmação de identidade com documentos tradicionais é feita visualmente por um ser humano, processo que demanda tempo e impede a agilidade de processos em cenários críticos como aeroportos internacionais, onde tal tarefa é realizada milhares de vezes por dia. Documentos de identificação tradicionais estão sendo substituídos por documentos eletrônicos com as características biométricas eletronicamente embutidas. Isso permite que uma máquina cheque rapidamente a identidade de uma pessoa.

De acordo com Ferrara et al. (2008), para permitir a interoperabilidade entre sistemas criados por desenvolvedores diferentes a integração do reconhecimento biométrico

em identificação de larga escala (passaporte, vistos etc.), um formato padronizado para imagem facial se faz necessário. Nesse contexto, a Organização Internacional de Aviação Civil (do inglês *International Civil Aviation Organization* - ICAO) começou, em 1980, um projeto focado na padronização da identificação biométrica automática de pessoas através de máquinas (Group et al., 2004). Inicialmente, três atributos físicos foram escolhidos para possível aplicação nesse contexto: face, impressão digital e íris. A face foi escolhida como a característica física mais apropriada pela praticidade na emissão de passaporte. Entretanto, os sistemas atuais de verificação automática de face apresentam taxas de acerto menores que os de impressão digital ou íris, existindo a possibilidade de uma confirmação através destes últimos métodos citados.

Segundo Maltoni et al. (2009), avaliações recentes mostraram que os algoritmos de reconhecimento facial geram resultados satisfatórios apenas operando com imagens capturadas sob condições controladas de pose e iluminação. Consequentemente, são necessárias fotografias faciais digitais de alta qualidade e livres de defeitos para maximizar a acurácia do reconhecimento feito por um ser humano ou por uma máquina. A Organização Internacional de Padronização (do inglês *International Standard Organization* - ISO) definiu, em conjunto com a Comissão Internacional de Eletrotécnica (do inglês *International Electrotechnical Commission* - IEC), um padrão para fotografia facial a ser usado em passaportes eletrônicos. Este padrão é o ISO/IEC 19794-5 (ISO, 2005), que especifica um conjunto de características que devem ser obedecidas, sendo estas relacionadas à posição da face na imagem e à ausência de falhas fotográficas como embaçamento, olhos vermelhos etc. Tais falhas podem afetar o resultado do reconhecimento, seja este feito por um ser humano ou uma máquina.

O desenvolvimento de sistemas de TI (Tecnologia da Informação) para verificação automatizada do nível de qualidade de uma imagem facial, de acordo com as normas do padrão ISO/IEC 19794-5, deu origem a estudos recentes em processamento de imagens digitais. A plataforma BioLab-ICAO (Maltoni et al., 2009) foi desenvolvida

com o objetivo de analisar e comparar o desempenho de tais sistemas. A mesma está disponível à comunidade científica através de um *website*, o qual possibilita aferir e comparar os desempenhos de algoritmos voltados para a análise de qualidade de imagens faciais de acordo com o padrão ISO/IEC. O Biolab-ICAO realiza testes padronizados que analisam diversos aspectos de qualidade da imagem, os quais estão listados na Tabela 1.1.

Este trabalho visa descrever a pesquisa e desenvolvimento de algoritmos voltados à análise de qualidade de imagens faciais para alguns dos requisitos especificados pelo padrão ISO/IEC 19794-5. Os tópicos abordados no presente trabalho são quatro: embaçamento, presença de cabelo na frente dos olhos, armação de óculos cobrindo os olhos e a presença de boné/chapéu. O requisito de análise do nível de foco da captura fotográfica (Biolab ICAO 08) verifica se houve erro de ajuste de foco no momento da captura da fotografia. A presença de cabelo encobrindo os olhos (Biolab ICAO 15) não é aceita pois afeta negativamente a localização precisa dos olhos, a qual é uma etapa crucial da biometria facial. Pelo mesmo motivo, é necessário analisar a ocorrência de óculos encobrindo os olhos (Biolab ICAO 26). A análise da presença de boné/chapéu (Biolab ICAO 27) avalia a presença de acessórios de cabeça, que não são permitidos pelas normas da ICAO.

1.1 Objetivos

O objetivo deste trabalho é o desenvolvimento de algoritmos de análise facial para alguns requisitos estabelecidos pelo padrão ISO/IEC 19794-5 (ISO, 2005). Estes requisitos são: verificação do nível de foco (Biolab ICAO 08), detecção de cabelo sobre os olhos (Biolab ICAO 15), verificação da ocorrência de armação de óculos encobrindo os olhos (Biolab ICAO 26) e, por último, detecção de chapéu (Biolab ICAO 27).

1.1.1 Objetivos Específicos

Os objetivos específicos deste trabalho estão listados a seguir:

1. Desenvolver métodos para calcular requisitos do ICAO já citados;
2. Estabelecer um comparativo com o estado-da-arte para cada requisito analisado;
3. Fazer uma análise quantitativa dos resultados tendo como referência a plataforma Biolab-ICAO.

1.2 Estrutura da Dissertação

Este trabalho está organizado da maneira a seguir: no Capítulo 2 a fundamentação teórica dos métodos implementados é explanada; o Capítulo 3 apresenta os trabalhos relacionados; no Capítulo 4 são apresentados os materiais e métodos; o Capítulo 5 traz a descrição da metodologia de avaliação, assim como os resultados alcançados pelos métodos propostos e um comparativo dos mesmos com os métodos já existentes; no Capítulo 6 são apresentadas as conclusões do trabalho.

Tabela 1.1: Descrição dos testes de qualidade de imagem facial realizados pelo Biolab ICAO, de acordo com o padrão ISO/IEC 19794-5.

<i>N</i> ^o	Descrição do Teste
Teste de Precisão da Extração de Características Faciais	
1	Precisão na localização dos olhos
2	Precisão na localização da face
Testes Geométricos	
3	Distância entre os olhos (min. 90 pixels)
4	Posição vertical relativa
5	Posição horizontal relativa
6	Razão da largura da cabeça
7	Razão da altura da cabeça
Teste de Pose e Fotografia	
8	Embaçamento
9	Desvio de olhar
10	Manchas ou amassões
11	Tom de pele não natural
12	Imagem muito clara/escuro
13	Imagem desbotada
14	Imagem pixelizada
15	Cabelos encobrindo os olhos
16	Olhos fechados
17	Fundo uniforme
18	Inclinações da cabeça
19	Reflexo de <i>flash</i> na pele
20	Olhos vermelhos
21	Sombra atrás da cabeça
22	Sombras projetadas na face
23	Óculos escuros
24	Óculos com reflexos nas lentes
25	Óculos com armação muito espessa
26	Armação de óculos cobrindo os olhos
27	Boné/Chapéu
28	Véu cobrindo a face
29	Boca aberta
30	Presença de outras faces ou bonecos perto da face

Capítulo 2

Fundamentação Teórica

Este capítulo é dedicado à apresentação de conceitos teóricos fundamentais para o entendimento geral do trabalho.

2.1 Imagens Digitais e Processamento de Imagens

De acordo com Gonzalez e Woods (2008), uma imagem monocromática pode ser definida como uma função bidimensional, $f(x,y)$, onde x e y são coordenadas espaciais ($x,y \in \mathbb{N}^2$) e o valor de f dado para qualquer par de coordenadas (x,y) é chamado de intensidade do nível de cinza para aquele ponto. Quando os valores de x , y e dos níveis de cinza de f são finitos, então tal imagem é chamada de imagem digital. Um ponto (x,y) do espaço de coordenadas é chamado de pixel, um acrônimo da expressão inglesa *picture element* ou "elemento da imagem". O Processamento Digital de Imagens (PDI) é o campo da ciência da computação que visa processar e/ou extrair informação de imagens em sua forma digital.

De forma sucinta, pode-se dizer que processamento de imagens digitais corresponde a qualquer processamento de dados cuja entrada é uma imagem digital e cuja saída é outra imagem ou valor qualquer referente à mesma. Esta área difere do tratamento de

imagens, que se limita à manipulação de imagens digitais com intuito de aperfeiçoamento do aspecto visual ou procedimentos afins. PDI serve, por exemplo, de etapa inicial para processamentos de dados mais elaborados como reconhecimento de padrões, diagnóstico por imagem de Raios-X etc.

2.1.1 Tipos de imagem digital

Uma imagem $f(x,y)$ pode ser representada em sua forma digital de formas distintas, tais como imagens monocromáticas, binárias ou coloridas.

Uma imagem monocromática, ou imagem em níveis de cinza, popularmente chamada de imagem em preto e branco, é uma função $f(x,y)$ onde o valor de f corresponde à intensidade do brilho do pixel (x,y) . O valor de $f(x,y)$ pode variar dentro de um intervalo pré-definido. Imagens digitais monocromáticas são armazenadas em memória com espaço de 8 bits para cada pixel. Assim sendo, os valores de $f(x,y)$ variam no intervalo $[0-255]$, podendo apresentar 256 níveis diferentes de cinza.

Uma imagem binária é uma função $f(x,y)$ onde o f pode assumir apenas dois valores possíveis. Estas imagens são, por isso, chamadas de mapas, máscaras ou ainda de imagens *booleanas*, em referência à lógica booleana e seus estados de verdadeiro ou falso. Os pixels de uma imagem binária podem ser representados por 0 e 1 ou qualquer outro padrão. Em alguns casos são representadas por 1 e "não importa" (do inglês, *don't care's*), ou seja, qualquer outro valor. Máscaras binárias são muito usadas como imagens auxiliares que, por exemplo, demarcam regiões de uma outra imagem. A Figura 2.1 exemplifica esta atividade. Na Figura 2.1c, os pixels brancos, chamados pixels em primeiro plano, representam a região de 2.1b onde os níveis de cinza são maiores que 150, ou seja, os pixels mais claros com relação ao limiar estabelecido (150).

Uma imagem colorida é uma função $f(x,y)$ onde para cada pixel (x,y) o valor de f é dado por componentes. Tais componentes podem ter naturezas distintas, representando informações distintas. Isto é abordado em mais detalhes na Seção 2.2 a seguir.



Figura 2.1: Exemplo de imagem binária. A Figura 2.1a exibe a famosa lena em cores, 2.1b apresenta a lena em níveis de cinza. A Figura 2.1c exibe a binarização de 2.1b, onde os pixels brancos (primeiro plano) representam a região de 2.1b com nível de cinza maior que 150. Este processo é chamado binarização.

2.2 As cores e suas formas de representação

O uso de cores em processamento de imagens é motivado por dois fatores principais. Primeiro, a cor é um descritor poderoso que ainda facilita a identificação e extração de objetos em uma cena. Segundo, o ser humano pode distinguir entre milhares de tonalidades diferentes de cor, entretanto, pode distinguir apenas dúzias de níveis diferentes de cinza (Gonzalez e Woods, 2008).

Como ilustrado na Figura 2.2, em 1666, Sir Isaac Newton observou que quando um raio de luz solar atravessa um prisma de vidro, o feixe emergente de luz não é branco, mas é composto de um espectro contínuo de cores que variam do violeta até o vermelho. De fato, as cores observadas pelo sistema visual são determinadas pela natureza da luz refletida pelos objetos observados (Newton, 1730).

A luz visível (Figura 2.3) é composta por uma gama relativamente pequena de frequências do espectro eletromagnético. Um corpo que reflete luz composta por todas as frequências do espectro visível parece branco para o observador. Por outro lado, um corpo que reflete apenas uma gama limitada de frequências apresenta um tonalidade dada pela combinação de tais frequências de onda luminosa.



Figura 2.2: Ilustração do experimento de Newton no século XVII.

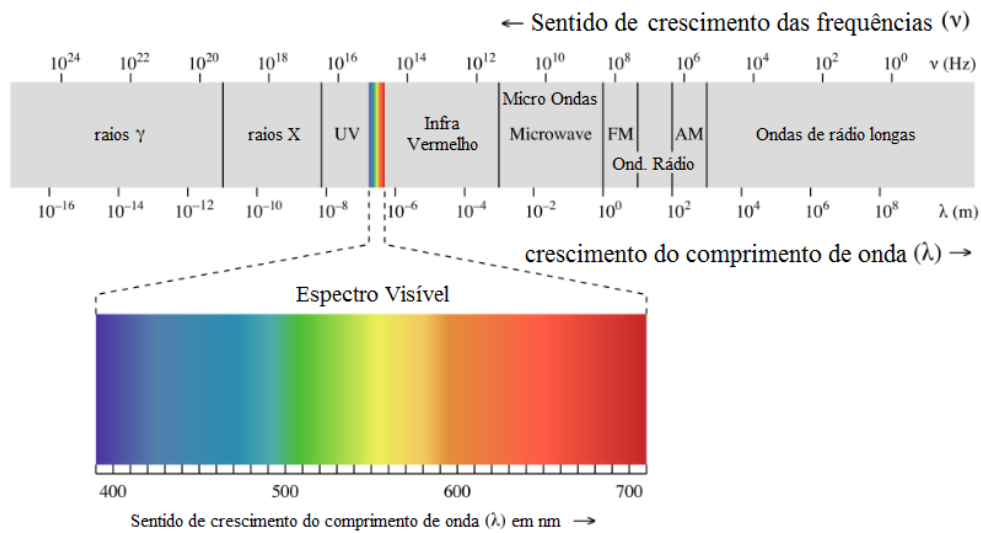


Figura 2.3: Comparação entre o espectro de luz visível e o espectro eletromagnético como um todo. (Fonte: Gonzalez e Woods (2008))

Os sistemas de cores, também chamados de *espaços de cores* ou *modelos de cores*, são uma forma de representar cores de maneira adequada para um sistema de computador. Os mesmos têm como propósito a padronização da descrição de cores e são, essencialmente, a especificação de um sistema de coordenadas (e um subespaço do mesmo) onde cada cor é representada por um único ponto deste sistema (Gonzalez e Woods, 2008). A maioria dos sistemas de cor utilizados atualmente são orientados a algum *hardware* (monitores, impressoras etc.) ou a aplicações cujo foco é a manipulação de

cores (computação gráfica). Alguns modelos que podem ser citados são o RGB (usado em telas e monitores de vídeo), o CMYK (dedicado a impressão), o YC_bC_r (voltado para transmissão de vídeo à distância) e o HSI (que visa a uma descrição mais intuitiva das cores). A seguir, alguns desses modelos são descritos em mais detalhes.

2.2.1 Sistema de cores RGB

O sistema de cores RGB é baseado na teoria tricromática da visão (Young, 1802), a qual remete ao trabalho no século XIX pelo inglês Thomas Young, um médico que se tornou notável pelas suas pesquisas em ótica. A teoria tricromática da visão estabelece que é possível a composição de uma grande variedade de cores a partir de três cores primárias: vermelho (R, *red*), verde (G, *green*) e azul (B, *blue*), como ilustrado na Figura 2.4.

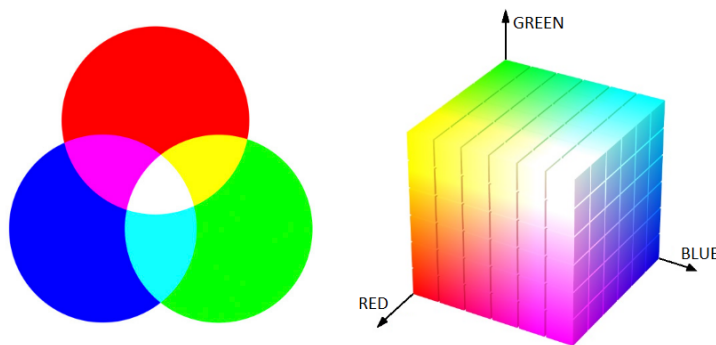


Figura 2.4: Cores primárias do sistema RGB e suas respectivas combinações.

De fato, o olho humano possui células sensíveis à intensidade da luz, chamadas de bastonetes, e outras sensíveis às cores, chamadas cones, os quais são sensíveis basicamente a estas três cores: vermelho, verde e azul. O sistema RGB é usado em dispositivos geradores de imagem como monitores de vídeo ou projetores, sendo o sistema de cores padrão para a maioria dos formatos de imagem digital existentes.

2.2.2 Sistema de cores YCbCr

O sistema de cores RGB apresenta muita redundância na forma de representar cores. Assim, em cenários onde a representação de cor precisa ser o mais otimizada possível, como é o caso das imagens de televisão transmitidas em tempo real a longas distâncias, uma alternativa ao RGB é necessária. O sistema YCbCr reduz a redundância presente nos canais do RGB e representa as cores em componentes estatisticamente independentes (ou o mais independentes possível). Este sistema representa as cores em termos de luminância (Y , luma) e crominância (C_b e C_r). O cálculo para conversão de cores do sistema RGB para YCbCr é feito de acordo com a Equação 2.1 (Vezhnevets et al., 2003).

$$\begin{cases} Y = 0,229 \cdot R + 0,587 \cdot G + 0,114 \cdot B \\ C_b = B - Y \\ C_r = R - Y \end{cases} \quad (2.1)$$

2.2.3 Sistema de cores HSL

Quando uma pessoa vê um objeto, a mesma não descreve tal imagem em função das componentes vermelho, verde e azul. Ela o faz descrevendo a cor, sua intensidade, e informando ainda se é claro ou escuro. O sistema HSL visa uma descrição mais intuitiva das cores e o faz através de três componentes. Estas são: *Hue* (do inglês, matiz); *Saturation* (do inglês, saturação); e *Lightness* (do inglês, luminosidade). A matiz indica a cor pura em questão (amarelo puro, vermelho puro etc.), a saturação indica a intensidade da coloração, ou seja, se a cor aparece vívida ou acinzentada. O brilho indica se a cor é clara ou escura. De acordo com Gonzalez e Woods (2008), o sistema RGB é ideal para a geração de cores, mas sua capacidade de descrição é limitada. Desta forma, o sistema HSL mostra-se ideal para o desenvolvimento de aplicações de processamento de imagens baseadas em descrição de cores que sejam

naturais e intuitivas para humanos, sendo estes os usuários finais de tais ferramentas.

A seção a seguir traz um exemplo clássico de como se pode explorar a informação contida nas cores para fins dos mais interessantes.

2.2.4 Uso das cores para detecção de pele em imagem digital

Segundo Kakumanu et al. (2007), detecção de pele tem um papel importante em uma vasta gama de aplicações no processamento digital de imagens, indo desde detecção facial, rastreamento facial, reconhecimento de gestos, recuperação de imagens baseada em conteúdo, até aplicações na interação humano-computador. Entretanto, detectar pele no espectro visível é uma tarefa desafiadora, uma vez que a cor da pele humana é suscetível a alguns fatores tais como iluminação, especificações da câmera, etnia do indivíduo, características pessoais como idade, marcas cutâneas etc. Vendo a tarefa de detecção de pele como um problema de classificação, esta resume-se a um problema de duas classes: pixels com cor de pele e pixels sem cor de pele. Os passos fundamentais da mesma são:

1. Representar os pixels da imagem em um sistema de cores adequado;
2. Criar uma modelagem da cor de pele dentro do espaço de cores. Buscar os pixels relacionados a cor de pele e identificar seus respectivos subespaços, ou seja, subconjunto dos vetores do espaço de cor que correspondem a cor de pele;
3. Classificar vetores (pixels) como pertencentes ou não ao subespaço modelado.

Muitos sistemas de cores foram propostos na literatura para detecção de pele, assim, o modelo escolhido também define quão efetivamente se pode modelar a distribuição da cor da pele. Explorar as cores é apenas uma forma de extrair informação de uma imagem digital, entre inúmeras outras, tais como detecção dos contornos ou análise de frequências. A filtragem de imagens é uma ferramenta importante para realização de

inúmeros procedimentos de processamento digital de imagens, ferramenta esta que é apresentada a seguir.

2.3 Filtros de imagem

Um filtro consiste no processamento dos valores dos pixels de uma imagem com intuito de modificar características da mesma. Tal processamento é descrito da seguinte maneira (Gonzalez e Woods, 2008):

$$g(x,y) = \mathcal{T}[f(x,y)] \quad (2.2)$$

Na Equação 2.2, $f(x,y)$ é a imagem de entrada, $g(x,y)$ é a imagem resultante do processamento e \mathcal{T} é um operador aplicado em f . O operador \mathcal{T} modifica o valor dos pixels, podendo atuar em cada ponto da imagem isoladamente (operador pontual) ou analisando uma dada vizinhança. Num operador pontual, o operador \mathcal{T} depende apenas do valor de $f(x,y)$ para cada ponto (x,y) . Esta modificação nos níveis de cinza é chamada de mapeamento. Quando a operação \mathcal{T} modifica os valores de $f(x,y)$ levando em consideração o valor de $f(x,y)$, assim como o valor dos pixels adjacentes, considerando a vizinhança, a operação é chamada de filtragem.

A análise da vizinhança de um pixel é feita através de uma sub-imagem, cujas dimensões delimitam os pixels adjacentes a um dado ponto (x,y) . A sub-imagem pode ter qualquer formato, mas em geral é quadricular, sendo chamada também de máscara ou *kernel*. A Figura 2.5 ilustra a análise de uma vizinhança de 3x3 pixels centrada em (x,y) .

O procedimento de filtragem é dependente da modelagem matemática utilizada para representação dos dados a serem exibidos. Há mais de uma forma de organizar ou representar os dados que compõem uma imagem digital. Estas formas são apresentadas e discutidas na próxima seção.

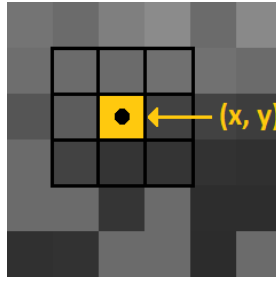


Figura 2.5: Ilustração de uma máscara 3x3 centrada num pixel (x,y) .

2.4 Domínios do espaço e das frequências

A forma tradicional de definir uma imagem digital como uma função bidimensional $f(x,y)$, onde x e y são coordenadas espaciais é chamada de representação no domínio do espaço cartesiano. Todavia, é possível representar uma imagem em um espaço vetorial diferente através de uma operação matemática chamada *transformada*. Operações como a transformada discreta de Fourier (Bracewell, 1965) (do inglês *Discrete Fourier Transform* - DFT) possuem ampla aplicação em processamento de imagem. A mesma transforma uma imagem $f(x,y)$ do domínio do espaço para uma versão no domínio das frequências, onde os valores dos pixels são dados em termos da frequência de sua ocorrência no sinal (imagem, função) original. As Equações 2.3 e 2.4 exibem, respectivamente, a transformada de Fourier $\mathfrak{F}\{f(x,y)\}$ e a transformada inversa $\mathfrak{F}^{-1}\{F(u,v)\}$. Esta última realiza a conversão de volta ao domínio do espaço.

$$\mathfrak{F}\{f(x,y)\} = F(u,v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) \exp[-j2\pi(ux + vy)] dx dy \quad (2.3)$$

$$\mathfrak{F}^{-1}\{F(u,v)\} = f(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u,v) \exp[j2\pi(ux + vy)] du dv \quad (2.4)$$

A Figura 2.6 mostra a conversão de uma imagem para o domínio das frequências. A Figura 2.6a mostra a imagem de entrada. A Figura 2.6b exibe o espectro de frequências

correspondente à imagem de entrada no domínio das frequências. As frequências neste espectro ficam distribuídas de tal forma que as baixas frequências se concentram perto do centro (Figura 2.6c) e as altas frequências ocupam a periferia do espectro (Figura 2.6d). As frequências baixas representam as regiões lisas da imagem de entrada, ou seja, com pequena variação do nível de cinza. As frequências altas representam as áreas detalhadas, regiões com grande contraste entre os pixels.

A representação de uma imagem no domínio da frequência é útil porque possibilita manipular a informação em domínio diferente, através de ferramentas algébricas intrínsecas ao mesmo. Por exemplo, eliminar frequências maiores que um dado limiar (como ilustrado na Figura 2.6c) equivale a uma operação chamada *filtragem passa-baixas*, o que elimina o detalhismo e suaviza os contornos da imagem no domínio do espaço. Analogamente, eliminar frequências menores que um limiar (Figura 2.6d) equivale à *filtragem passa-altas*, que elimina regiões espaciais com pequena variação na intensidade dos pixels, dando realce aos contornos no domínio do espaço.

2.4.1 Filtragem espacial

De acordo com o que foi introduzido na Seção 2.3, a filtragem de imagens no domínio do espaço se dá através do uso de máscaras. Inúmeras operações como suavização

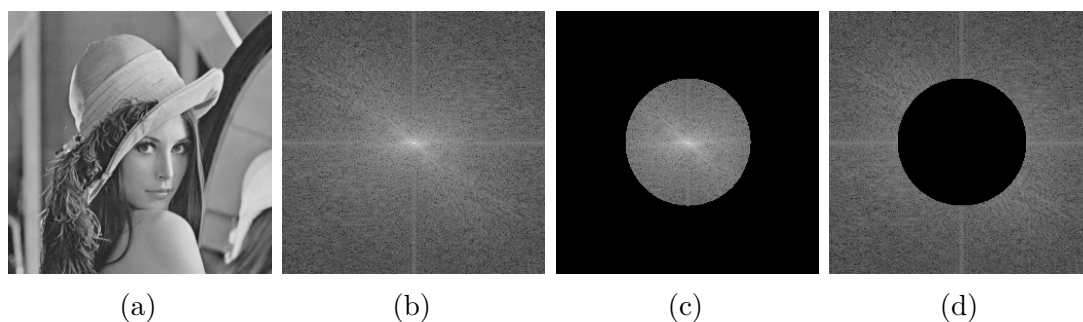


Figura 2.6: Conversão para o domínio das frequências: 2.6a exibe a imagem de entrada; 2.6b mostra o espectro de frequências; 2.6c ilustra a localização das baixas frequências perto do centro; 2.6d ilustra a localização das frequências altas.

ou realce de contornos podem ser realizadas com o uso de uma máscara, sendo esta específica para cada procedimento. A Figura 2.7 exhibe as máscaras usadas para realce e suavização, assim como os resultados do processo de filtragem com as mesmas.

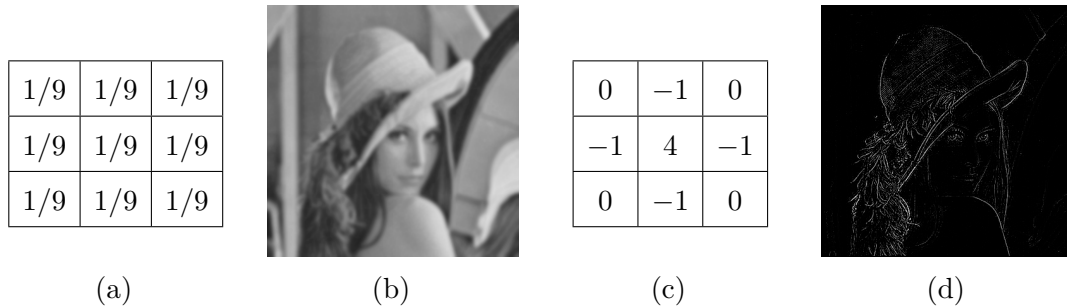


Figura 2.7: Exemplos de máscaras: 2.7a e 2.7c exibem máscaras de suavização e realce, respectivamente; 2.7b e 2.7d exibem o resultado da filtragem da Imagem 2.6a com uso das máscaras 2.7a e 2.7c, respectivamente.

Pode-se perceber que máscaras de suavização como a exibida na Figura 2.7a possuem somatório dos coeficientes igual a 1. Em contrapartida, máscaras de realce de bordas como 2.7c apresentam somatório igual a zero. Nas próximas seções são apresentados alguns métodos de filtragem no domínio do espaço.

2.4.1.1 Filtro de suavização Gaussiana

O filtro Gaussiano é um operador de convolução 2-D usado para suavizar imagens, removendo detalhes e ruído (Fisher et al., 1994). O filtro é baseado na função Gaussiana apresentada na Equação 2.5, onde σ representa o desvio padrão da distribuição. A Figura 2.8 apresenta a distribuição da função para uma e duas dimensões, assim como a representação discretizada numa matriz de convolução.

$$Gaussian(x,y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2.5)$$

A filtragem através da matriz de convolução gaussiana consiste em um filtro passa-baixas, tendo efeito de embaçamento e suavização.

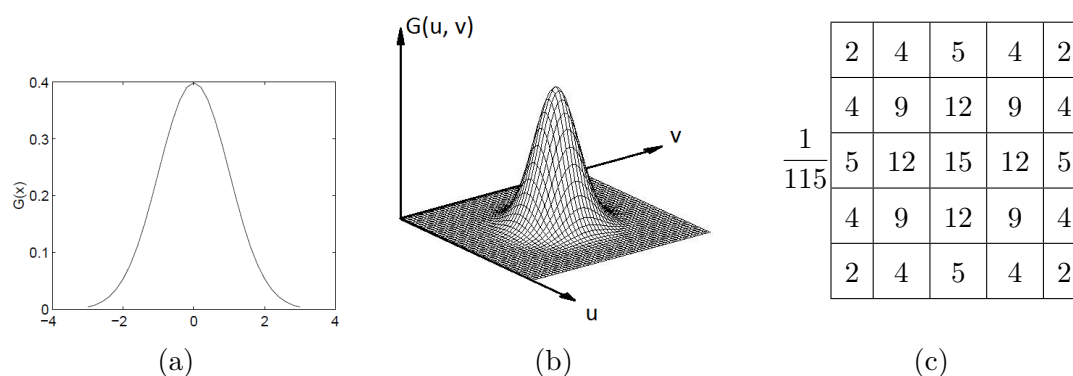


Figura 2.8: Representações da função Gaussiana: 2.8a apresenta a distribuição da função para uma variável com valor de $\sigma = 1$; 2.8b apresenta a distribuição para duas variáveis com valor de $\sigma = 1$; 2.8c apresenta uma discretização em forma de matriz de convolução com valor de $\sigma = 1,4$.

2.4.1.2 Filtro de detecção de bordas de Sobel

O operador de Sobel (Sobel, 2014) calcula o gradiente no espaço 2-D para cada ponto de uma imagem $f(x,y)$ (Fisher et al., 1994). O vetor gradiente é uma generalização do conceito de derivada e indica o sentido para o qual a função tem o crescimento mais acentuado. Calcular o gradiente resulta no realce das altas frequências espaciais que correspondem aos contornos da imagem.

O filtro funciona através da convolução de dois *kernels* que calculam as componentes do vetor gradiente nas direções horizontal (G_x) e vertical (G_y). As matrizes de convolução são apresentadas na Figura 2.9. Estas podem ser aplicadas individualmente com intuito de realçar apenas as bordas horizontais ou verticais da imagem. O gradiente resultante para cada pixel é dado por $|G| = \sqrt{G_x^2 + G_y^2}$.

O resultado da filtragem com o método de Sobel é uma imagem em níveis de cinza, onde os contornos da imagem original são destacados. O nível de cinza de um pixel da imagem filtrada equivale à intensidade do gradiente para tal ponto. A Figura 2.10 exibe o resultado da execução deste filtro para imagens com e sem foco. Percebe-se que nas Figuras 2.10b e 2.10d, de acordo com o nível de foco da imagem original, as bordas

-1	0	+1
-2	0	+2
-1	0	+1

(a)

+1	+2	+1
0	0	0
-1	-2	-1

(b)

Figura 2.9: Matrizes de convolução utilizadas pelo filtro de Sobel. 2.9a apresenta a matriz que calcula a componente horizontal G_x do gradiente e 2.9b, a componente vertical G_y .

estão realçadas de forma diferenciada, tanto em relação à espessura das mesmas como em relação ao nível de cinza.

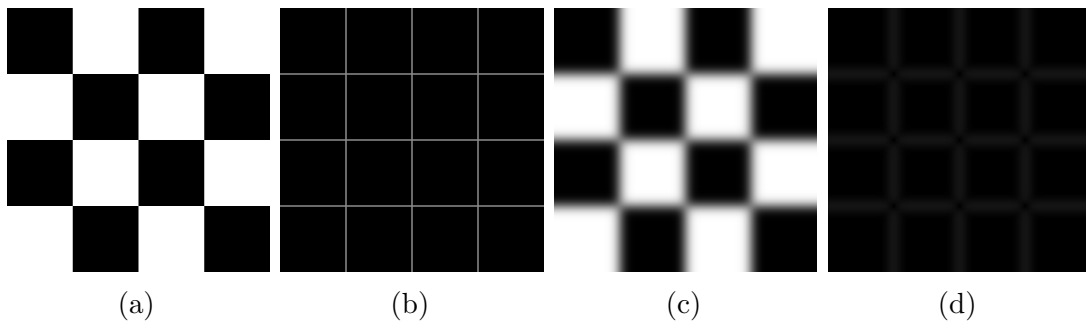


Figura 2.10: Exemplos da execução do filtro de Sobel. A Figura 2.10a exibe uma imagem com bordas bem definidas e 2.10b ilustra o resultado da execução do filtro de Sobel. A Figura 2.10c exibe uma imagem que apresenta bordas suavizadas e 2.10d mostra o resultado da execução do filtro.

2.4.1.3 Filtro de detecção de bordas de Canny

O filtro de Canny (Canny, 1986) foi desenvolvido para ser um detector de bordas otimizado (Fisher et al., 1994). O método faz uso de dois parâmetros (limiar mínimo e máximo) empiricamente definidos e usados para definir as bordas de forma personalizada de acordo com a aplicação. O algoritmo apresenta etapas bem definidas onde, primeiramente, a imagem de entrada é suavizada. Em seguida, um filtro de primeira derivada (e.g. Sobel) é aplicado com o intuito de realçar as bordas. Então, para cada

pixel, o valor do gradiente é comparado com dois parâmetros limiares. Se o gradiente é maior que o limiar máximo, então o pixel é demarcado como borda. Se o valor do gradiente está compreendido entre os limiares mínimo e máximo, será demarcado como borda se existir em sua vizinhança um pixel demarcado como tal. Se a intensidade do gradiente é menor que o limiar mínimo, determina-se que o pixel em questão não representa borda. A Figura 2.11 apresenta um exemplo da execução do filtro de Canny.



Figura 2.11: Exemplo de execução do filtro de Canny. 2.11a apresenta a imagem de entrada e 2.11b, a imagem filtrada.

2.4.1.4 Filtro de normalização de iluminação - LogAbout

A performance de métodos de reconhecimento de faces sofre impacto direto das condições da imagem facial. Um fator que tem impacto muito negativo é a presença de sombras sobre o rosto. Com o intuito de diminuir a influência da iluminação em imagens faciais, foram desenvolvidos métodos de normalização de iluminação que visam corrigir o efeito negativo das sombras ao mesmo tempo em que preservam as características relevantes da face. Um desses métodos é o LogAbout (Liu et al., 2002). O mesmo consiste em um passa-altas seguido do filtro de Log, o qual é baseado na função logarítmica que está descrita na Equação 2.6.

$$\text{Log}(x,y) = a_1 + \frac{\ln(f(x,y) + 1)}{b_1 \cdot \ln(c_1)} \quad (2.6)$$

As variáveis a_1 , b_1 e c_1 na Equação 2.6 são parâmetros que alteram a forma da curva gerada pela função logarítmica. O filtro de Log tem um resultado de alta qualidade na normalização da iluminação de uma imagem, entretanto, tem como ponto fraco o fato de não preservar os contornos da imagem. O filtro LogAbout faz uso de um passa-altas que precede o filtro de Log. Isto é feito com objetivo de realizar uma normalização eficiente da iluminação com preservação dos detalhes, superando desta forma a limitação do filtro do Log. A diferença de execuções destes filtros está ilustrada na Figura 2.12.



Figura 2.12: Exemplos das execuções dos filtros Log e LogAbout: 2.12a exibe a imagem de entrada; 2.12b exibe o resultado da execução do filtro Log; 2.12c exibe o resultado da execução do filtro LogAbout.

A seguir é apresentado um teorema, que correlaciona o procedimento de filtragem no domínio do espaço e o do domínio das frequências.

2.4.2 Convolução e o Teorema da Convolução

Convolução (simbolizada por $*$) é uma operação matricial espacial que consiste em filtrar uma imagem com uma máscara que é espelhada horizontal e verticalmente. O procedimento de convolução tem grande importância para o entendimento da filtragem no domínio da frequência, uma vez que esta estabelece uma ligação básica entre o domínio espacial e o da frequência. O fundamento teórico da filtragem no domínio da frequência é o Teorema da Convolução. De acordo com Gonzalez e Woods (2008), tal teorema afirma que se uma função $f(x,y)$ tem como transformada de Fourier $F(u,v)$

e uma máscara $g(x,y)$ tem como transformada de Fourier $G(u,v)$, então as seguintes relações são válidas:

$$f(x,y) * g(x,y) \iff F(u,v)G(u,v) \quad (2.7)$$

$$F(u,v) * G(u,v) \iff f(x,y)g(x,y) \quad (2.8)$$

Estas relações afirmam que a operação de convolução no domínio espacial é equivalente à operação de multiplicação no domínio da frequência, assim como a convolução no domínio da frequência é equivalente à operação de multiplicação no domínio do espaço.

2.4.3 Filtragem no domínio da frequência

No domínio da frequência o procedimento de filtragem consiste em realizar alterações no espectro de frequências. Tais alterações são feitas através da multiplicação das frequências por um valor estabelecido por uma função de filtro. Isto é fundamentado no teorema da convolução, o qual estabelece que: sendo uma imagem $f(x,y)$, que tem como transformada no domínio da frequência $F(u,v)$, filtrada espacialmente pela convolução de uma máscara $g(x,y)$, cuja transformada é $G(u,v)$, então a convolução $f(x,y) * g(x,y)$ no domínio das frequências equivale a $F(u,v) \cdot G(u,v)$.

Levando em consideração a distribuição de frequências no espectro, um filtro de suavização (passa-baixas) usa uma máscara que acentua as frequências próximas do centro do espectro. A Figura 2.13 apresenta representações espaciais de filtros de domínio da frequência para melhor visualização. Exemplos de máscaras passa-baixas estão ilustradas nas Figuras 2.13a e 2.13c e de máscaras passa-altas nas Figuras 2.13b e 2.13d.

Os filtros das Figuras 2.13a e 2.13b são chamados filtros ideais (Gonzalez e Woods,

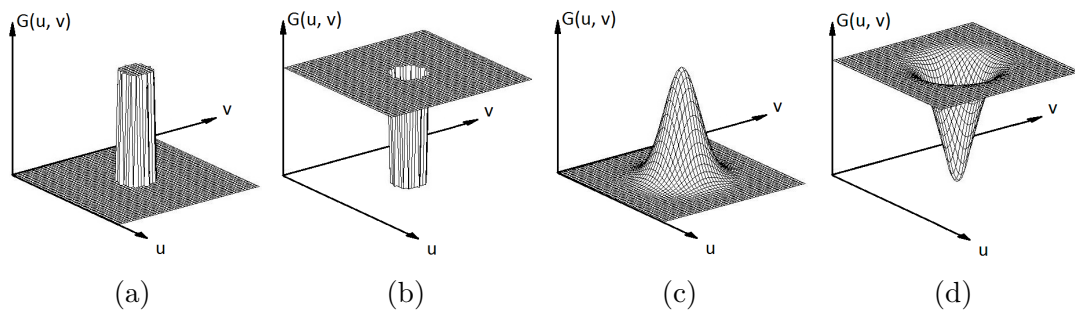


Figura 2.13: Filtros de frequências. 2.13a e 2.13c apresentam filtros passa-baixas. 2.13b e 2.13d apresentam filtros passa-altas. Os filtros 2.13a e 2.13b são chamados ideais e os 2.13c e 2.13d são gerados pela função Gaussiana.

2008) e realizam um corte bem definido das frequências de acordo com uma circunferência cujo centro coincide com o do espectro. O raio da mesma é o parâmetro de corte. Os filtros das Figuras 2.13c e 2.13d são gerados pela função Gaussiana. Tais máscaras apresentam uma transição menos abrupta na faixa de frequências a serem filtradas.

A Figura 2.14 ilustra o resultado da aplicação dos filtros ideais exibidos nas Figuras 2.6c e 2.6d à imagem da Figura 2.6a. Perceba que a Figura 2.14b, resultado da filtragem passa-baixas com a máscara da Figura 2.14a, apresenta bordas suavizadas, enquanto que a imagem da Figura 2.14d, resultado da filtragem passa-altas com a máscara da Figura 2.14c, apresenta o realce dos contornos da imagem original, presentes nas frequências mais altas.

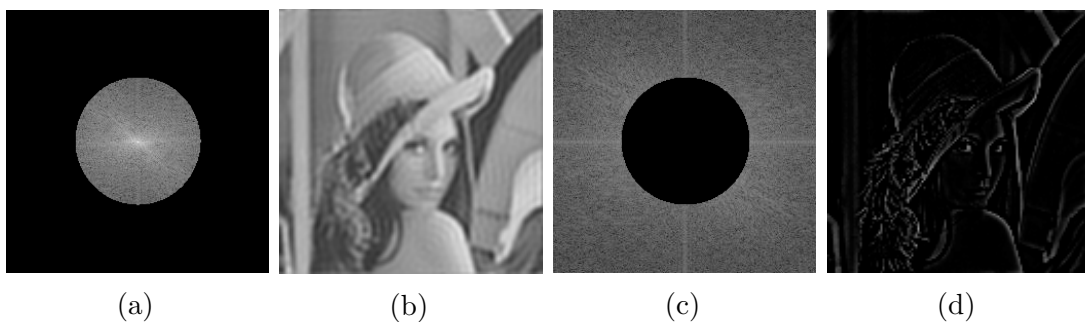


Figura 2.14: Exemplo de filtragem no domínio da frequência: a Figura 2.14b exibe a imagem da Figura 2.6a filtrada com a máscara passa-baixas de 2.14a; a Figura 2.14d exibe o resultado da filtragem da mesma imagem pela máscara passa-altas de 2.14c.

Analisar imagens faciais no contexto do padrão ICAO requer o uso de ferramentas clássicas de PDI, tais como filtragem, análise de frequências, assim como requer instrumentos bem mais elaborados. O reconhecimento de padrões em imagens digitais é uma área importante para a visão computacional. Tarefas como identificar e localizar o padrão de um rosto em uma imagem são primordiais em aplicações de PDI voltadas à análise facial. Assim sendo, são apresentados, a seguir, métodos de reconhecimento de padrões que foram úteis no desenvolvimento deste trabalho.

2.5 Detecção facial através do método Viola-Jones

O detector facial de Viola-Jones (Viola e Jones, 2004) é capaz de processar uma imagem em um espaço de tempo muito curto, atingindo, todavia, altas taxas de detecção correta. O método foi inicialmente proposto para o domínio de detecção facial em imagens, mas se trata essencialmente de um *framework* de detecção de objetos que combina três etapas principais: representação da imagem através de descritores de características Haar (Papageorgiou et al., 1998), que permitem um processamento super-rápido da imagem como um todo; classificação das características através de um classificador eficiente, construído através do algoritmo AdaBoost (Freund et al., 1999); combinação de classificadores em cascata, que permite a eliminação rápida do fundo da imagem, gastando processamento no que é mais provável que seja o objeto de busca. Como todo esse processo envolve aprendizado de máquina, o método pode ser treinado para detecção de quaisquer objetos numa imagem digital. Ao final do processo de classificação, o Viola-Jones indica através de quadriláteros as faces (ou objetos) encontradas. A Figura 2.15 ilustra um resultado da execução do algoritmo.



Figura 2.15: Resultado da execução do detector de faces Viola-Jones numa imagem com diversas faces (Fonte: Viola e Jones (2004)).

2.6 Localização de elementos faciais

No presente trabalho é usado um método de localização de elementos faciais proposto por Kazemi e Sullivan (2014). Trata-se de um *framework* baseado em aprendizado de máquina através de um *ensemble* de árvores de regressão, usado para estimar a posição dos elementos faciais a partir de um pequeno subconjunto de intensidades de pixels. Esse método é capaz de localizar inúmeras faces com precisão em um tempo de execução muito pequeno, resultando ainda na predição de alta qualidade dos elementos faciais. Como resultado da execução deste algoritmo, tem-se uma lista de pontos, os quais estão associados aos elementos faciais como nariz, boca, olhos etc. Tal resultado pode ser visualizado na Figura 2.16.

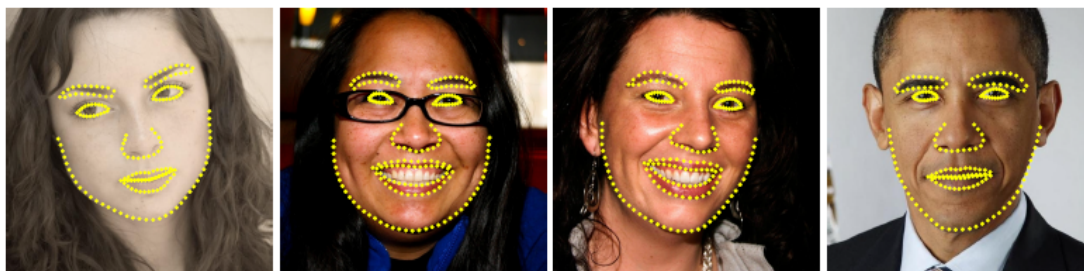


Figura 2.16: Resultado da execução do detector de elementos faciais de Kazemi e Sullivan em faces diferentes (Fonte: Kazemi e Sullivan (2014)).

2.7 A transformada de Hough para detecção de linhas

A transformada de Hough é um método bastante conhecido de reconhecimento de padrões, sendo computacionalmente simples, o que possivelmente explique tamanha popularidade. Inicialmente desenvolvido para detecção de linhas, o método foi expandido e aplicado na detecção de outras formas geométricas, tais como círculos e elipses. Tal adaptação é possível porque se trata de uma transformada paramétrica, ou seja, o método explora os parâmetros da equação que define a forma geométrica a ser detectada. No caso das retas a equação cartesiana é dada por $y = a_2x + b_2$, onde x e y são variáveis, a_2 e b_2 são os parâmetros. As variáveis representam coordenadas dos pontos no plano cartesiano, obviamente, enquanto os parâmetros descrevem as características da reta, ou seja, a inclinação (parâmetro a_2) e o ponto onde a mesma toca o eixo das ordenadas, ou zero da função (parâmetro b_2).

Pode-se, então, criar um espaço vetorial bidimensional representado em função dos parâmetros a_2 e b_2 , o chamado espaço paramétrico. Em tal espaço, uma reta é representada por apenas um ponto, o qual é dado por valores de a_2 e b_2 , respectivamente equivalentes à inclinação da reta e ao zero da função. A transformada de Hough explora o espaço paramétrico com objetivo de detectar retas. Em geral, uma etapa prévia de realce de bordas é realizada. Em seguida, o algoritmo busca nestas bordas realçadas um conjunto de pixels adjacentes, entre os quais segmentos de reta conectantes

sejam representados pelo mesmo ponto do espaço paramétrico. Se um ponto aparece repetidamente no espaço paramétrico, um segmento de reta está, assim, detectado. Duda e Hart (1972) apresentam a transformada de Hough para detecção de linhas através da representação polar da equação paramétrica de uma reta.

2.8 Operações morfológicas: Erosão e Dilatação

Operações morfológicas são realizadas geralmente sobre imagens binárias com o intuito de modificar a forma do conjunto de pixels em primeiro plano com relação ao fundo. Entre os usos comuns de tais operações podem-se citar detecção de bordas, remoção de ruído, aprimoramento e segmentação de imagem. As operações morfológicas mais comuns são dilatação e erosão.

A operação de dilatação tem como efeito básico o alargamento da fronteira do conjunto dos pixels em primeiro plano (*foreground*), tipicamente pixels brancos. A região dos pixels em primeiro plano cresce tendo como efeito a expansão das periferias e diminuição de possíveis buracos contidos nesta região. Já a operação de erosão tem como efeito básico a diminuição da região dos pixels em primeiro plano, de forma que a periferia da região encolhe e possíveis buracos na região aumentam de tamanho. A Figura 2.17 apresenta exemplos destas operações.

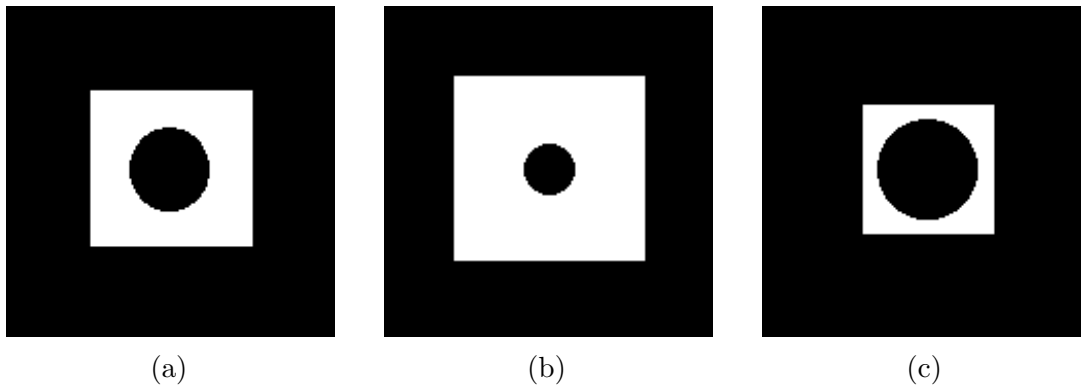


Figura 2.17: Exemplo da execução das operações de dilatação e erosão de pixels. 2.17a exibe uma imagem binária onde os pixels brancos correspondem ao primeiro plano. A imagem 2.17b mostra o resultado da dilatação, enquanto que a 2.17c mostra o resultado da operação de erosão.

Capítulo 3

Trabalhos Relacionados

Esta seção apresenta uma revisão literária relacionada ao padrão ICAO, assim como aos métodos abordados no presente trabalho, os quais foram citados previamente.

O trabalho de Ferrara et al. (2012b) apresenta a plataforma BioLab-ICAO, um sistema *on-line* para avaliação e comparação de algoritmos para análise de qualidade de imagem facial. Tal ferramenta foi disponibilizada à comunidade científica com objetivo de encorajar as pesquisas relacionadas ao tema. Esta consiste numa grande coleção de amostras faciais rotuladas, um protocolo de teste bem definido e alguns resultados de algoritmos, que servem como base de comparação.

O conjunto de imagens faciais do BioLab-ICAO é composto por 5.588 fotografias faciais. Destas, um subconjunto de 3.967 imagens é oriundo de diferentes bancos de imagens faciais existentes, outras 804 imagens foram artificialmente geradas, visando simular requisitos de qualidade como pixelização ou embaçamento e, por fim, 817 novas imagens foram capturadas. O protocolo de testes da ferramenta exige que os algoritmos analisem uma imagem facial e retorne, para cada requisito do ICAO, um valor no intervalo $[0,100]$, onde zero representa um nível mínimo de obediência ao requisito e cem, um nível máximo. A plataforma fornece ainda resultados de algoritmos próprios, apresentados no trabalho, para comparativo. Além dos algoritmos próprios, disponíveis

para todos os requisitos do ICAO, o trabalho ainda apresenta resultados para algoritmos de terceiros, os quais não englobam todos os requisitos do padrão.

Em Ferrara et al. (2012a) é proposta uma nova abordagem para segmentação de imagem facial. O método visa identificar as regiões principais da imagem e classificá-las em quatro categorias: plano de fundo, face, cabelo e roupa. A abordagem é baseada em sistema de múltiplos classificadores, os quais exploram informação das cores e texturas. Levando em conta um conhecimento previamente extraído de uma referência, o sistema tenta classificar as diferentes regiões da imagem. Este método consiste basicamente em segmentação da imagem facial, seguida da classificação das regiões segmentadas. Esta classificação das regiões é feita levando em consideração máscaras, ou mapas, que descrevem as quatro regiões principais: pele, roupa, cabelo e plano de fundo. Tais máscaras foram criadas previamente, a partir de uma coleção de imagens com as quatro regiões principais demarcadas, sendo cada máscara uma média das regiões demarcadas no banco de imagens faciais. Tais máscaras são apresentadas na Figura 3.1.

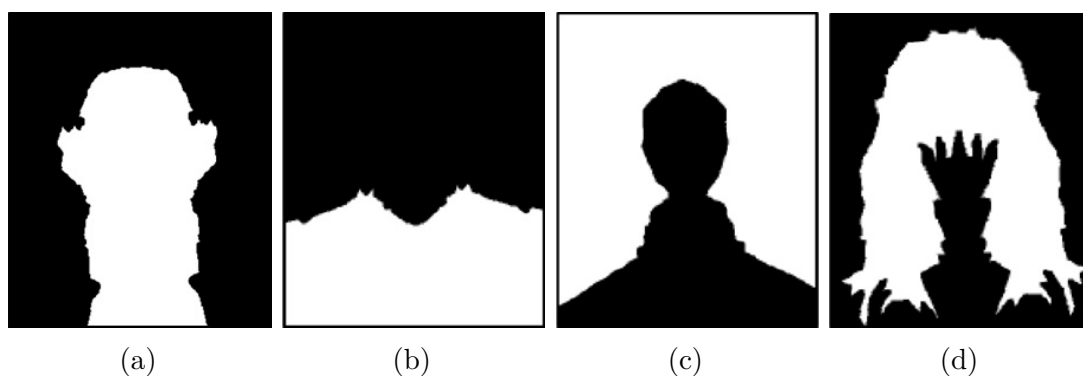


Figura 3.1: Mapas gerados para representação das quatro principais regiões da imagem: mapa de pele (3.1a); mapa de roupa (3.1b); de plano de fundo (3.1c); mapa de cabelo (3.1d). (Fonte: Ferrara et al. (2012a))

Primeiramente, a imagem passa por segmentação de baixo (do inglês, *low level segmentation*) nível ou, no inglês, *low-level segmentation*, processo que divide a imagem em sub-regiões que contêm pixels semelhantes. Em seguida, com base nos quatro ma-

pas de pele, roupa e outros, associados às regiões segmentadas, informação das cores, texturas, é feito um processo de classificação que resulta na segmentação das quatro regiões principais das imagens candidatas, exibido na Figura 3.2. Os mapas indicam a região onde provavelmente acontecem as regiões principais e o processo de classificação expande e adapta cada região à cor e à textura da imagem candidata. Estas regiões são tomadas como referência para análise de todos os requisitos estabelecidos pelo padrão ICAO de qualidade.

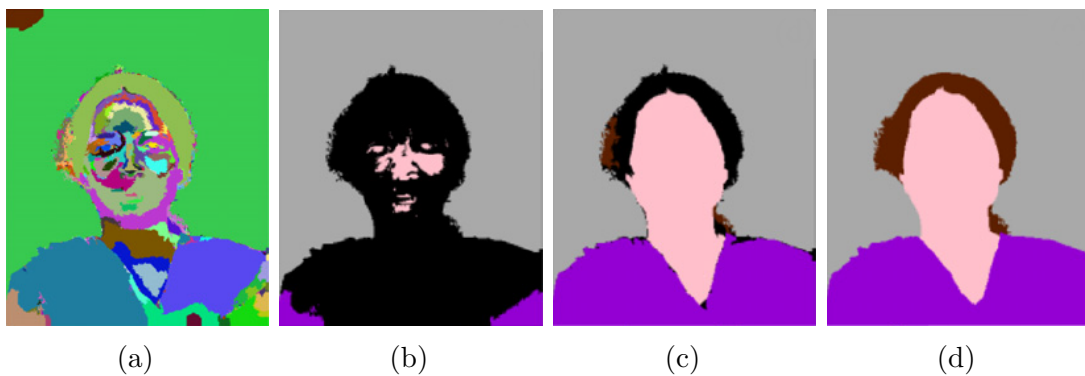


Figura 3.2: Etapas da expansão e determinação das regiões principais a partir da segmentação de baixo nível (3.2a) até a separação das quatro regiões principais (3.2d). (Fonte: Ferrara et al. (2012a))

Entre estes requisitos de qualidade há o requisito de análise de foco de uma imagem, propriedade relacionada à forma de captura da mesma. Verificar o nível de foco é uma tarefa comumente realizada por equipamentos de aquisição de imagens com ajuste automático de foco: máquinas fotográficas digitais, microscópios óticos etc. Pech-Pacheco et al. (2000) apresentam um algoritmo inovador para o ajuste automático de foco no cenário de microscopia. O trabalho compara o desempenho de métodos propostos para ajuste automático de foco em imagens de diatomáceas (algas unicelulares) oriundas de microscopia ótica. Analisar ou contar diatomáceas requer um especialista, sendo ainda uma tarefa demorada, tediosa e cuja automatização é de suma importância. Uma imagem de diatomáceas bem focada apresenta bordas muito bem definidas, ou seja,

com alto contraste na intensidade dos pixels dos contornos das células. Por outro lado, uma imagem fora de foco apresenta bordas difusas e sem definição, ou seja, com menor variação na intensidade dos pixels dos contornos.

Partindo dessa premissa, o trabalho de Pacheco faz uso de filtros de detecção de bordas para aferir o nível de definição dos contornos da imagem. Em seguida, a medição do nível de definição das bordas é feita através do cálculo da variância estatística da intensidade dos pixels da imagem filtrada. Os filtros utilizados são o operador de Sobel (Sobel, 2014), o operador Laplaciano (Wang, 2007) e Tenengrad (Krotkov, 1988). Resultados experimentais mostraram uma significativa melhora de desempenho comparando a abordagem proposta, baseada no cálculo da variância dos pixels após a detecção de bordas, com métodos presentes na literatura.

Para o requisito do ICAO que analisa se a armação de óculos aparece encobrendo os olhos é necessária uma prévia detecção de óculos na imagem facial. Jiang et al. (2000) apresentam um método clássico de detecção de óculos. A principal contribuição deste trabalho é determinar a região da face mais adequada para realizar a busca por óculos. Isto foi realizado através do Critério da Informação de Fisher (Frieden, 2004). Este método, desenvolvido pelo estatístico inglês Sir Ronald A. Fisher, indica a quantidade de informação contida numa variável aleatória X com relação a um dado parâmetro Y . Através desta ferramenta matemática constatou-se que a região localizada acima do nariz e entre os olhos é a que contém maior quantidade de informação referente a óculos. Isto se dá porque o apoio nasal dos óculos em geral encontra-se na região citada, apesar de óculos serem um acessório com grande variação de aparência. O trabalho de Jiang ainda apresenta um método para detecção baseado em realce de bordas pelo filtro de Sobel seguido pelo cálculo de três métricas definidas, as quais são apresentadas nas Equações 3.1, 3.2 e 3.3.

$$M1 = \frac{\sum_{p_{ij} \in F} edgeness_{ij}}{F} \quad (3.1)$$

$$M2 = \frac{\left(\sum_{p_{ij} \in F} edgeness_{ij} \geq T \right) \text{então } 1 \text{ caso contrário } 0}{F} \quad (3.2)$$

$$M3 = \frac{\left(\sum_{p_{ij} \in F} edgeness_{ij} \geq T \right) \text{então } edgeness_{ij} \text{ caso contrário } 0}{F} \quad (3.3)$$

Onde:

- T é o limiar de Otsu;
- $edgeness_{ij}$ é a intensidade da borda no pixel P_{ij} ;
- F é o número total de pixels da imagem.

O recente trabalho de Fernández et al. (2015) apresenta uma abordagem inovadora para o problema de detecção de óculos, usando representação de características através de LBP, ou Local Binary Patterns (do inglês, Padrões Binários Locais) e alinhamento robusto dos mesmos. Inicialmente, a imagem é pré-processada e a região dos olhos é delimitada. Em seguida, os descritores LBP para a região ocular são construídos. A classificação dos descritores é feita através de uma máquina de vetor de suporte (do inglês *Support Vector Machine* - SVM). Com objetivo de testar o método proposto em ambiente realístico, o algoritmo foi testado no banco de faces LFW (sigla do inglês *Labeled Faces in the Wild*, ou Faces Rotuladas em Ambiente Diverso em tradução livre, Huang et al. (2007)), cujo propósito é a análise de imagens faciais em ambiente não controlado. Este banco possui 13.233 imagens de pessoas famosas (ou nem tanto) adquiridas principalmente em páginas de notícias da internet. As faces variam bastante quanto a iluminação, pose, vestimentas etc. Este método atingiu uma taxa de detecção correta de 98,65% para o LFW. Como não há outro método da literatura relacionada a detecção de óculos avaliado com o banco LFW, o algoritmo também foi avaliado com

o banco FERET (Phillips et al., 1998) com objetivo de mostrar que o mesmo supera o estado-da-arte. A taxa de detecção correta para o FERET, com validação cruzada de 5 etapas, foi de 99,89%.

Para a avaliação da presença de acessório de cabeça tais como chapéus ou bonés, o trabalho de Storer et al. (2010) propõe uma abordagem baseada em detecção de pele na identificação de obstruções faciais em imagens do banco de faces AR (Martinez, 1998). Para o caso da obstrução causada por chapéu, a análise de pele é realizada na região da testa do indivíduo retratado. Uma forma clássica de modelar e detectar pele consiste em explorar a informação de cor disponível, principalmente fazendo uso de outros espaços de cores ao invés do tradicional modelo RGB. O trabalho de Storer utiliza a banda H do espaço de cores HSV (Joblove e Greenberg, 1978). Tal modelo representa as cores em três canais: (H) *Hue*, do inglês, a matiz, cor em questão; (S) *Saturation*, do inglês saturação, ou vivacidade da coloração; e (V) *Value*, do inglês valor, que é o brilho do pixel em questão. Após a realização de testes de detecção de pele em diversos espaços de cor, constatou-se que o HSV é o mais adequado para detecção de obstrução facial para tal banco de imagens. A Figura 3.3 apresenta o uso do canal H do HSV para segmentação de pele.

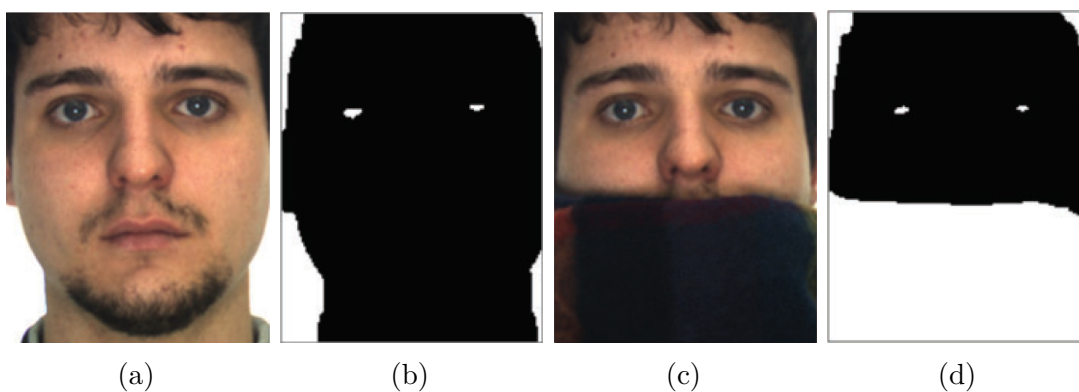


Figura 3.3: Processo de detecção de oclusão facial através de segmentação de pele. Processo exibido para imagens com oclusão da face (3.3a e 3.3b) e sem oclusão (3.3c e 3.3d). (Fonte: Storer et al. (2010))

Capítulo 4

Metodologia

Este capítulo apresenta os materiais utilizados para o desenvolvimento e avaliação dos métodos propostos, assim como uma descrição detalhada dos mesmos.

4.1 Materiais

Nesta seção são apresentados os materiais utilizados para implementação dos métodos propostos para avaliação dos requisitos do padrão ISO/IEC 19794-5 já citados. A Subseção 4.1.1 apresenta a biblioteca de software para visão computacional OpenCV. A Subseção 4.1.2 apresenta a plataforma de testes Biolab ICAO, através da qual foram realizados todos os testes de desempenho, e comparativo com outros métodos submetidos na mesma. São apresentados também os bancos de imagens faciais dos quais foram coletadas as amostras usadas pelo Biolab ICAO.

4.1.1 OpenCV

A OpenCV (do inglês *Open Source Computer Vision*) (OpenCV, 2008) é uma biblioteca de visão computacional disponível nas linguagens C/C++ e Python. Foi originalmente desenvolvida pela Intel Co. e atualmente é mantida pela Itseez, empresa que atua no

ramo de visão computacional. Como o foco da biblioteca é a operação em tempo real, a estrutura *Mat* (que armazena uma imagem digital) é extremamente otimizada, resultando em execuções rápidas e computacionalmente eficientes, com gerenciamento de memória encapsulado e transparente ao desenvolvedor. O módulo *ImageProcessing* inclui funcionalidades de PDI como filtros lineares e não lineares, transformações geométricas como redimensionamento, conversões de espaços de cor, entre muitas outras.

4.1.2 A plataforma Biolab-ICAO

O Biolab-ICAO (Maltoni et al., 2009) é uma plataforma desenvolvida para avaliação e comparação do desempenho de software voltado especificamente para análise de qualidade de imagem facial de acordo com o padrão ISO/IEC 19794-5. Tal plataforma está disponível para a comunidade científica através de um *site* na rede mundial de computadores. Este consiste em um sistema *on-line* onde se pode submeter arquivos executáveis, os quais são referentes aos métodos a serem testados, e obter, em seguida, os resultados experimentais.

O Biolab ICAO conta com um protocolo de testes bem definido, além de um conjunto de imagens faciais cuidadosamente construído, pois, em trabalhos anteriores (Ferrara et al., 2008), o conjunto de imagens faciais usado nos testes não incluía imagens fora de conformidade com alguns requisitos do padrão. Entre estes, pode-se citar a presença de véu sobre a face, uso de chapéu e outros. Tal conjunto de imagens foi significativamente estendido, passando a conter imagens que obedecem (e que não obedecem) a cada um dos requisitos listados na Tabela 1.1. Este novo conjunto conta com 7.740 faces coletadas de diferentes fontes (Maltoni et al., 2009). Nas próximas seções, o protocolo de testes e os bancos de imagens faciais são apresentados em mais detalhes.

4.1.2.1 Protocolo de testes do Biolab ICAO

Para a avaliação da conformidade de uma imagem facial com o padrão de qualidade, a plataforma BioLab-ICAO requer que os algoritmos a ela submetidos retornem um valor dentro do intervalo $[0,100]$. Zero representa o nível mínimo de conformidade com o requisito e 100 o nível máximo. Se o valor retornado na análise de uma imagem, para um dado requisito do ICAO, é maior ou igual a um limiar previamente definido para o mesmo então tal imagem é aceita, caso contrário, não aceita. Como a conformidade de uma amostra é um problema de classificação com duas classes possíveis (obedece ou não obedece um dado requisito), dois tipos de erros podem ocorrer neste contexto (Ross et al., 2006):

- Falsa Aceitação: declarar que uma amostra está em conformidade com um requisito quando a mesma de fato não está;
- Falsa Rejeição: declarar que uma amostra não está em conformidade quando a mesma está.

Realizar uma variação do limiar em um intervalo pré-determinado permite calcular a taxa de falsa aceitação (do inglês *False Acceptance Rate* - FAR) e a taxa de falsa rejeição (*False Rejection Rate* - FRR) para cada limiar e gerar as curvas de FAR e FRR. O ponto de intersecção destas curvas é chamado de ponto de ERR (do inglês *Equal Error Rate*, ou taxas de erro iguais). Este ponto indica o limiar para o qual a taxa de falsa aceitação é proporcionalmente igual à de falsa rejeição. A Figura 4.1 apresenta um exemplo de gráfico de curvas de FAR e FRR. Quando uma amostra de face é de difícil avaliação, por exemplo, analisar se os olhos estão fechados numa face que apresenta óculos escuros, a plataforma permite que os algoritmos rejeitem a imagem. Entretanto, com o intuito de não estimular a rejeição de amostras visando o ganho de performance, cada rejeição é considerada uma análise errada do requisito. A plataforma informa a taxa de rejeição junto com as taxas EER.

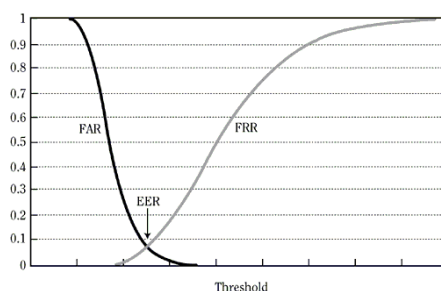


Figura 4.1: Exemplo de gráfico das curvas de taxa de falsa aceitação (FAR) e taxa de falsa rejeição. No cruzamento das duas está destacado o ponto de taxa de erro igual (EER).

A plataforma dispõe de dois conjuntos de imagens gabaritadas, um aberto para treinamento local e outro fechado para testes no sistema *on-line*. O conjunto aberto está disponível para testes locais e possível treinamento. Já o conjunto de teste é fechado e desconhecido por parte de quem submete algoritmos à plataforma. Este conjunto fechado de imagens possui subconjuntos (também desconhecidos) usados para a avaliação dos requisitos, de forma que tais subconjuntos apresentam imagens que obedecem e que não obedecem aos mesmos. A seguir, são apresentadas as características dos bancos de faces dos quais foram retiradas as imagens utilizadas pelo BioLab-ICAO, com exceção de uma minoria obtida pelos criadores ou gerada artificialmente (imagens pixelizadas, por exemplo).

4.1.2.2 Banco de faces AR

As imagens do banco AR (Martinez, 1998) apresentam dimensões de 768x576 pixels e as faces apresentam variações quanto a sorriso, boca aberta, uso de óculos escuros, cachecol, entre outras. Além disso, ocorre uma variação proposital e controlada da iluminação da face, variando entre frontal e lateral. O fundo apresenta pouca variação e é geralmente branco. Algumas amostras de imagens deste banco estão disponíveis na Figura 4.2.



Figura 4.2: Amostras de imagens do banco de faces AR: 4.2a, iluminação à esquerda; 4.2b, face neutra; 4.2c, uso de óculos; 4.2d, iluminação à direita (Fonte: Martinez (1998)).

4.1.2.3 Banco de faces FRGC

As imagens do banco FRGC (do inglês *Face Recognition Grand Challenge*, ou Grande Desafio de Reconhecimento Facial) apresentam formatos de 2272x1704 pixels e 1704x2272 pixels. Este banco de faces visa fornecer amostras de imagens faciais capturadas em ambiente mais realista, apresentando condições mais variadas quanto a fonte de iluminação, localização da face na imagem, proporção entre o tamanho da imagem e o tamanho da face, além de variações de *background* etc. Algumas amostras de imagens deste banco estão disponíveis na Figura 4.3.



Figura 4.3: Amostras de imagens do banco de faces FRGC: 4.3a, face com sombras; 4.3b, face neutra com iluminação controlada; 4.3c, presença de outras pessoas; 4.3d, face com sorriso.

4.1.2.4 Banco de faces PUT

No banco de faces PUT as imagens foram capturadas com iluminação parcialmente controlada e fundo uniforme. Apresentam dimensões de 2048x1536 pixels e a principal

variação de uma imagem para outra é a posição da face. Esta é variada horizontalmente, da esquerda para a direita, e verticalmente, de cima para baixo. Algumas amostras de imagens podem ser visualizadas na Figura 4.4.



Figura 4.4: Amostras de imagens do banco de faces PUT. A Figura 4.4a mostra a face voltada para a esquerda. As várias imagens de um mesmo indivíduo apresentam gradual variação horizontal de posição da face até a mesma estar voltada para a direita na Figura 4.4b. A Figura 4.4c mostra a face voltada para cima. Analogamente, a posição da face varia verticalmente até a mesma estar voltada para baixo, como na Figura 4.4d.

4.1.2.5 Vsoft - A empresa

A Vsoft Tecnologia é uma empresa que atua no segmento de software, como fornecedora independente especializada em Identificação Biométrica. Sediada em João Pessoa, Paraíba, a empresa foi fundada em 2000 e se especializou ao longo dos anos no desenvolvimento de software voltado ao processamento digital de imagens. A Biometria tem sido seu foco principal, com a pesquisa e desenvolvimento de algoritmos de reconhecimento de faces e de impressões digitais em parceria com a Universidade Federal da Paraíba.

Como a avaliação da qualidade de uma imagem facial é uma importante etapa prévia do reconhecimento biométrico de face, a empresa iniciou a pesquisa e o desenvolvimento de algoritmos para avaliação de qualidade facial de acordo com cada requisito do padrão ICAO. Esta ferramenta será um componente da solução de reconhecimento facial que está em desenvolvimento na empresa. Quatro dos métodos desenvolvidos para avaliação do ICAO são apresentados a seguir.

4.2 Métodos Propostos

Nesta seção são apresentados os métodos propostos para análise da qualidade de imagem facial de acordo com o padrão ISO/IEC 19794-5. São aqui analisados quatro requisitos: análise de embaçamento, uso de chapéu, análise do encobrimento dos olhos causado por armação de óculos e por fios de cabelo. A Subseção 4.2.1 descreve o método proposto para análise de embaçamento na imagem, a Subseção 4.2.2 apresenta de detecção de cabelo em frente aos olhos, a Subseção 4.2.3 descreve o método proposto para análise da presença de chapéu na imagem facial. Por fim, a Subseção 4.2.4 apresenta o método de verificação da ocorrência de encobrimento dos olhos causado por armação de óculos.

4.2.1 Análise do Nível de Foco

As normas ISO/IEC não permitem fotos desajuste acentuado de foco devido ao impacto claramente negativo que isto tem na tarefa de reconhecimento da face retratada, uma vez que desfoque é uma forma de difusão da informação capturada pela matriz da imagem. Verificar a qualidade de uma imagem facial quanto ao nível de foco consiste em determinar se a fotografia foi capturada com ajuste adequado de foco, ou seja, com boa definição dos contornos, o que implica numa representação apropriada da amostra biométrica, ou seja, a face. Uma fotografia sem foco apresenta uma informação difusa, dados embaralhados, o que não pode ser definido como uma boa coleta de amostra biométrica.

A abordagem proposta por Pech-Pacheco et al. (2000), inicialmente aplicada ao ajuste automático de foco em microscopia ótica, mostrou-se bastante eficaz para determinar o nível de foco de uma imagem facial. O método se baseia no fato de que o detalhismo apresentado por uma imagem digital se manifesta nas altas frequências. Assim sendo, analisar as frequências altas presentes numa imagem é uma forma direta de aferir o nível de definição da mesma. Uma das formas propostas por Pech-Pacheco

para tal consiste em ressaltar as altas frequências com uso do filtro espacial passa-altas de Sobel (Sobel, 2014). Posteriormente, a variância dos valores dos pixels na imagem é calculada e tomada como métrica do nível de foco.

Para aplicar tal metodologia ao contexto de imagens faciais, decidiu-se analisar o foco da imagem numa área restrita aos olhos do indivíduo. Isto é feito porque a face apresenta algumas regiões de pele com aspecto uniforme. É o caso da região da testa em geral e das bochechas onde não há presença de barba, espinhas etc. Regiões uniformes apresentam poucas frequências altas, então se escolheu a região dos olhos para análise deste requisito pois esta apresenta maior nível de detalhismo se comparada com a testa, bochechas, região nasal etc. A região ocular apresenta contornos nas pálpebras, cílios, pelos nas sobrancelhas e outros. Estes elementos detalhados se manifestam nas altas frequências que são afetadas pela ocorrência de desfoque.

O método proposto começa com a detecção e localização facial. Isto é feito através do detector facial de Viola-Jones (Viola e Jones, 2004), o qual retorna uma região quadrangular contendo a face. Uma vez que as dimensões desta região quadrada são proporcionais às dimensões da face, é possível identificar empiricamente uma sub-região retangular que delimite os olhos. Sejam X_0 e Y_0 , respectivamente, a largura e altura da região facial delimitada pelo Viola-Jones, a região ocular **Reg_{olhos}** corresponde ao retângulo com largura $0,7X_0$ e altura $1/6Y_0$, localizado como indicado na Figura 4.5.

Feita a delimitação de **Reg_{olhos}**, esta região é recortada da imagem original e convertida para escala de cinza (imagem em preto e branco). Em seguida, é aplicado o filtro de Sobel com intuito de realçar os contornos da imagem. Feito isto, o cálculo da variância é realizado e seu valor é tomado como métrica do nível de foco da imagem. Uma imagem bem focada apresentará muitas altas frequências e, conseqüentemente, uma variância maior nos níveis de cinza da imagem resultante do filtro de Sobel. Uma imagem com mau ajuste de foco, ou seja, com embaçamento acentuado, apresentará menos frequências altas detectadas pelo filtro de Sobel e, conseqüentemente, uma variância

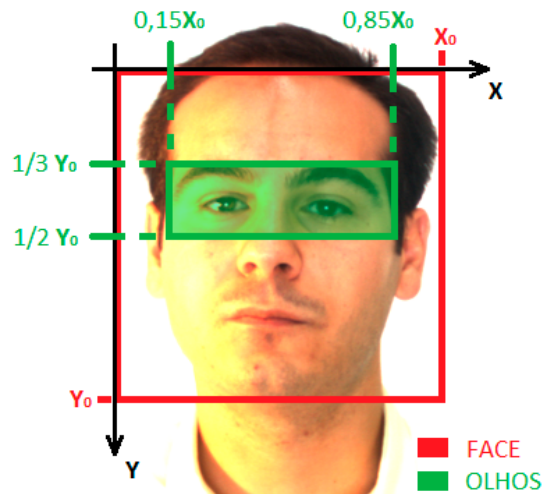


Figura 4.5: Delimitação da região dos olhos em função da região facial delimitada pelo Viola-Jones. X_0 e Y_0 são respectivamente a largura e altura da região facial. A origem da região ocular encontra-se no ponto $(1,5X_0, 1/3Y_0)$ e a mesma possui largura $L_o = 0,7X_0$ e altura $A_o = 1/6Y_0$.

menor.

4.2.2 Análise da presença de cabelo em frente aos olhos

Outro requisito de qualidade para imagem facial estabelece que não deve haver cabelos em frente aos olhos. Para avaliação deste requisito, um algoritmo baseado na simetria da face humana foi desenvolvido. Considerando que a face humana apresenta simetria bilateral, quando um recorte do olho direito é invertido horizontalmente e sobreposto pelo olho esquerdo a diferença entre as imagens deve ser pequena. Quando ocorre a presença de cabelo em frente aos olhos a diferença entre as mesmas deve ser maior do que quando não há presença de cabelo. Os passos principais do algoritmo estão ilustrados na Figura 4.6. A figura apresenta a execução do algoritmo para uma face sem cabelo em frente aos olhos na coluna da esquerda e uma face com cabelo em frente aos olhos na coluna da direita.

O primeiro passo do algoritmo é encontrar uma estimativa do centro dos olhos através do método de Kazemi e Sullivan (2014). Conhecendo a localização estimada dos

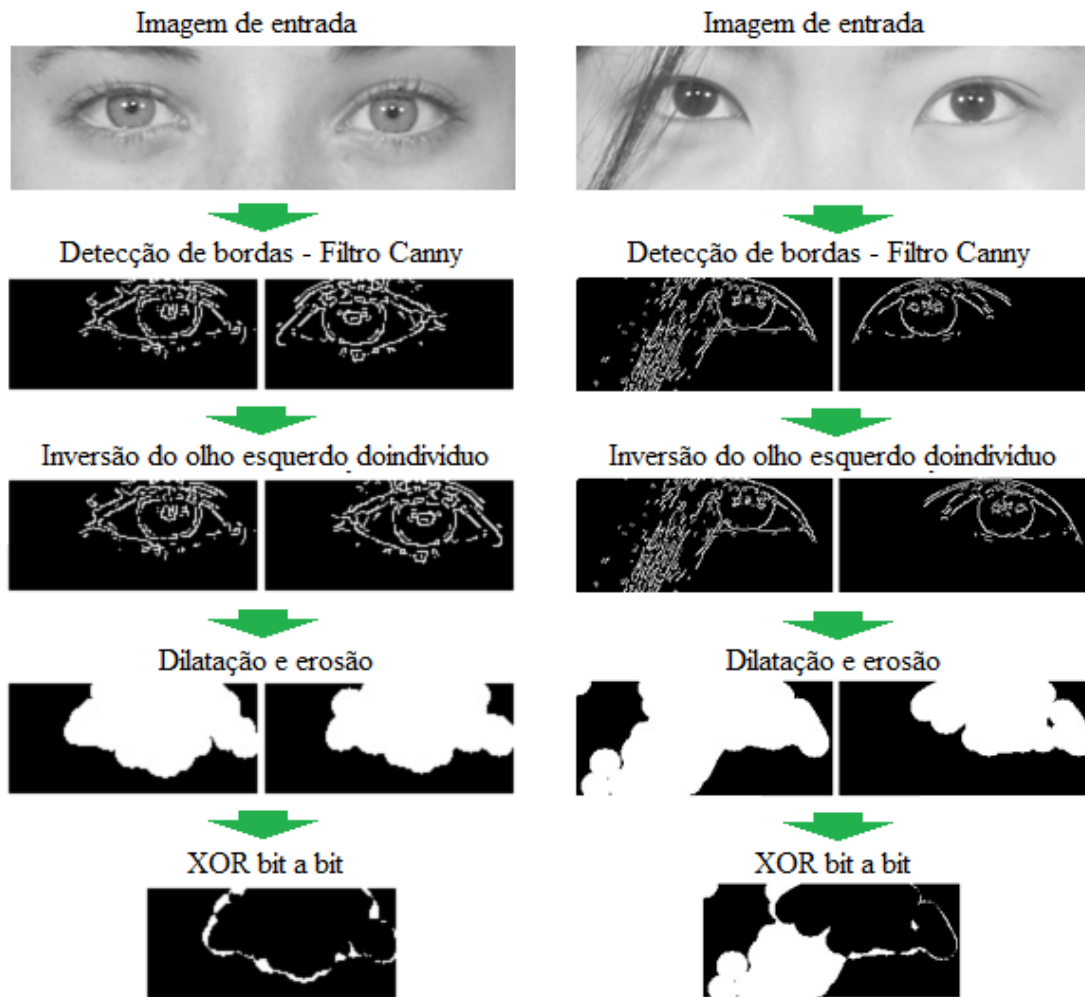


Figura 4.6: Fluxo de execução do algoritmo de análise da presença de cabelo em frente aos olhos.

olhos, encontra-se o ângulo entre a linha que conecta os olhos e a horizontal. A rotação do rosto é então compensada de acordo com este ângulo, de forma que a linha dos olhos fique alinhada com a horizontal. Em seguida, uma região de interesse é determinada em torno de cada olho do indivíduo. Estas regiões têm suas dimensões definidas de acordo com a distância entre os olhos (D_{olhos}), de forma que a altura vale $D_{\text{olhos}}/3$ e a largura vale $5D_{\text{olhos}}/8$.

Após a delimitação das regiões de interesse, a iluminação da face é normalizada

através do filtro LogAbout (Liu et al., 2002), com objetivo de diminuir a influência da iluminação e sombras projetadas sobre o rosto. No próximo passo, uma operação de suavização com filtro da média 3x3 é realizada para reduzir o ruído resultante do LogAbout. Então, o filtro de Canny (Canny, 1986) é executado com objetivo de detectar os contornos dos olhos, assim como a possível presença de cabelo. A imagem do olho esquerdo do indivíduo é invertida horizontalmente, de forma que se possa sobrepor o olho direito e explorar a simetria do rosto. Visando a união de linhas segmentadas e redução de ruído, uma sequência de operações de dilatação e erosão é realizada.

As imagens dos olhos direito e esquerdo resultantes deste processamento têm suas diferenças ressaltadas através de uma operação lógica XOR bit-a-bit, que realça os pixels não coincidentes nas duas imagens. A imagem resultante da operação XOR (I_{XOR}) é usada no cálculo de um valor (V_{cab}) que indicará a presença de cabelo na imagem. Este valor corresponde à porcentagem de pixels brancos (pixels não coincidentes) em I_{XOR} e seu cálculo está descrito na Equação 4.1.

$$V_{cab} = \frac{\sum_{j=1}^h \sum_{i=1}^w I_{XOR}(i,j)}{a_3 \cdot l} \quad (4.1)$$

$$I_{XOR}(i,j) = \begin{cases} 1, & \text{se o pixel for branco} \\ 0, & \text{caso contrário} \end{cases} \quad (4.2)$$

Dado que a_3 e l são a altura e largura de I_{XOR} .

4.2.3 Análise da presença de chapéu

O padrão ICAO não permite o uso de acessório de cabeça. Assim, para analisar a presença de chapéu, boné, em imagem facial visando atender aos requisitos de qualidade foi utilizada uma metodologia baseada no trabalho de Storer et al. (2010), que foca na detecção de obstruções faciais em imagem digital. Como o uso de acessórios na cabeça

pode obstruir a parte superior da face, este método visa detectar pele na região da testa da pessoa retratada.

De acordo com o trabalho de Storer, diversos experimentos com diferentes espaços de cor (RGB, YCbCr etc.) mostraram que o canal H do sistema de cores HSV é o mais adequado para detecção de obstruções em imagens faciais. O método binariza a banda H e realiza pós-processamento morfológico para remoção de regiões pequenas isoladas.

O método desenvolvido para análise da presença de chapéu no presente trabalho realiza primeiramente a delimitação da região da testa (\mathbf{Reg}_{testa}) com base na região facial delimitada pelo detector de faces de Viola-Jones, a qual é proporcional ao tamanho do rosto na imagem. Sendo X_0 e Y_0 respectivamente a largura e a altura da região facial delimitada, a região da testa empiricamente determinada corresponde a uma região retangular cujas largura e altura são, respectivamente, dadas por $\frac{3}{5}X_0$ e $\frac{1}{4}Y_0$, localizada conforme ilustrado na Figura 4.7.

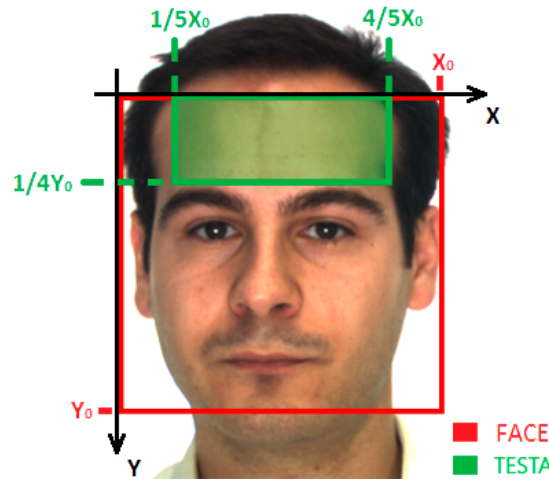


Figura 4.7: Delimitação da região da testa do indivíduo em função da região facial delimitada pelo Viola-Jones. X_0 e Y_0 são, respectivamente, a largura e a altura da região facial. A origem da região da testa encontra-se no ponto $(\frac{1}{5}X_0, 0)$ e a mesma possui largura $L_t = \frac{3}{5}X_0$ e altura $A_t = \frac{1}{4}Y_0$.

Após a localização da região da testa, uma suavização com filtro Gaussiano (com máscara de 5×5 e desvio padrão $\sigma = 1$) é realizada. A imagem suavizada é então

convertida para o espaço de cores HSV e a banda H é isolada. A imagem isolada da banda H é binarizada e a imagem binária (H_{bin}) é usada no cálculo de um valor (V_{chap}) que indicará a presença de obstrução na região da testa. Este valor corresponde à porcentagem de pixels correspondentes a pele em $\text{Reg}_{\text{testa}}$. Este cálculo está descrito formalmente na Equação 4.3, onde a_3 e l correspondem respectivamente às altura e largura de H_{bin} .

$$V_{\text{chap}} = \frac{\sum_{j=1}^h \sum_{i=1}^w H_{\text{bin}}(i,j)}{a_3 \cdot l} \quad (4.3)$$

As Figuras 4.8 e 4.9 apresentam a execução do algoritmo para faces, respectivamente, sem e com a presença de chapéu. Perceba a diferença na região da testa demarcada nas imagens 4.8d e 4.9d, as quais apresentam o resultado da binarização da banda H do HSV. Em 4.8d, imagem sem chapéu, a região da testa é uniforme e não apresenta pixels em primeiro plano (*foreground*). Já na imagem 4.9d, onde há presença de chapéu, a região da testa demarcada tem maioria de pixels não relacionados a pele.

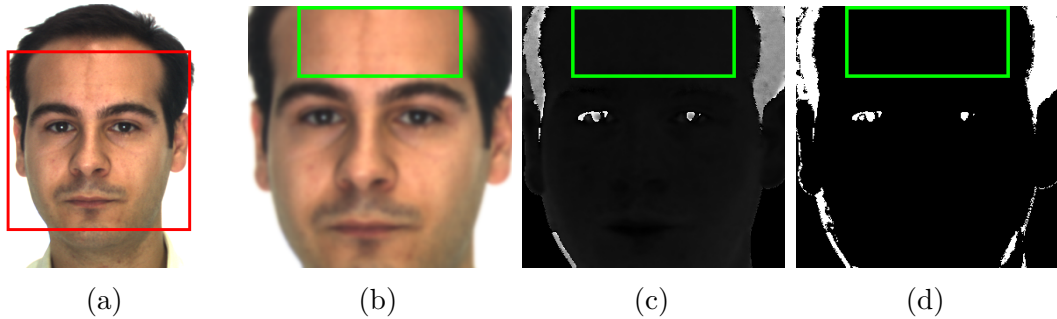


Figura 4.8: Execução do algoritmo para uma face sem chapéu: 4.8a apresenta a imagem de entrada com a delimitação da face pelo Viola-Jones; 4.8b apresenta a face suavizada com a delimitação da região da testa; 4.8c apresenta a banda H do sistema HSV; 4.8d apresenta a binarização da banda H.

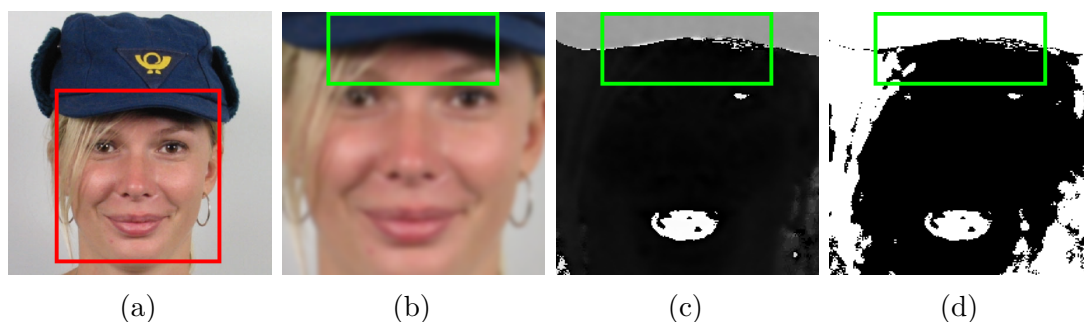


Figura 4.9: Execução do algoritmo para uma face com chapéu: 4.9a apresenta a imagem de entrada com a delimitação da face do Viola-Jones; 4.9b apresenta a face suavizada com a delimitação da região da testa; 4.9c apresenta a banda H do sistema HSV; 4.9d apresenta a binarização da banda H.

4.2.4 Análise do encobrimento dos olhos por armação de óculos

Esta seção apresenta o método proposto para a análise da presença de óculos encobrindo os olhos, o que não é permitido pelo padrão. Para averiguar tal requisito, é necessário determinar previamente a presença de óculos na imagem facial fornecida. Uma vez determinada a presença de óculos, é verificado se os mesmos encobrem os olhos.

4.2.4.1 Determinação da presença de óculos na imagem

Jiang et al. (2000) mostram, através do critério da informação de Fisher, que a região facial compreendida entre as sobrancelhas e a parte superior do nariz é a mais indicada para análise da presença de óculos em imagens faciais frontais. A localização desta região é calculada em função da posição dos olhos e distância entre eles. Assim sendo, a primeira etapa do algoritmo de detecção de óculos consiste em localizar os olhos.

O processo de localização dos olhos é feito em função do quadrilátero que delimita a região facial, fornecido pelo método Viola-Jones (Viola e Jones, 2004). A região dos olhos é calculada da forma ilustrada na Figura 4.5 e, em seguida, é dividida em duas partes iguais, de forma que cada uma contenha um dos olhos do indivíduo. Os pontos centrais dos quadriláteros resultantes são tidos como uma estimativa da localização dos

olhos na imagem. Este processo está ilustrado na Figura 4.10a.

Dados os pontos estimados de localização de cada olho, calcula-se o ponto médio dos mesmos, que aqui é denominado ponto P (Figura 4.10b). A região de interesse para detecção de óculos, denominada Região S, é calculada em função do ponto P e da distância entre os olhos (D_{olhos}). Como ilustrado na Figura 4.10b, o quadrilátero correspondente à região S tem como origem o ponto localizado a $0,125D_{\text{olhos}}$ à esquerda do Ponto P e $0,9D_{\text{olhos}}$ acima do mesmo. As dimensões de largura e altura da Região S são, respectivamente, D_{olhos} e $0,25D_{\text{olhos}}$, de acordo com o sistema de coordenadas cartesianas representado na figura.

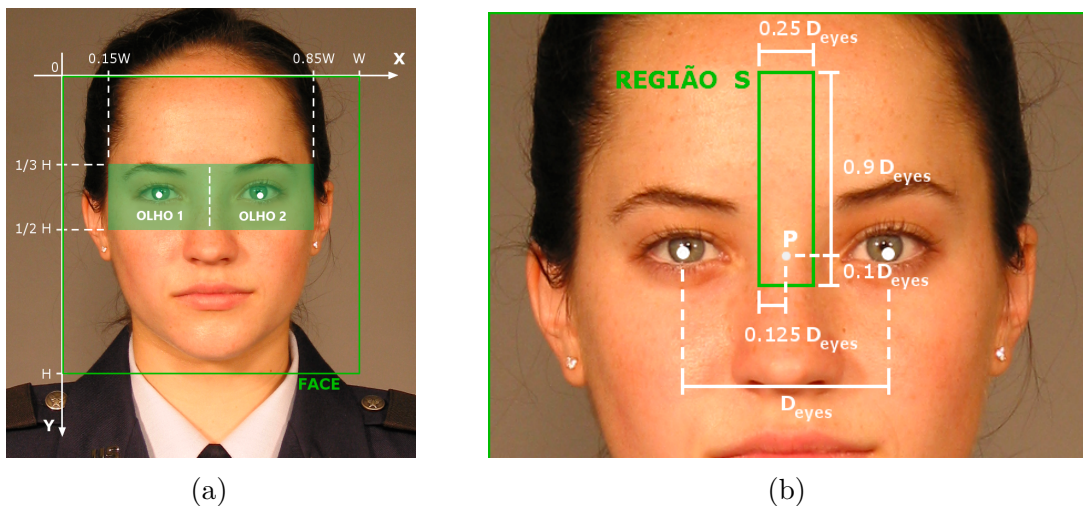


Figura 4.10: Processo de localização estimada dos olhos (a). Cálculo da Região S com relação à região da face detectada (b).

Estando definida a Região S, esta é convertida para o espaço de cores YCbCr e a banda Cr é isolada, pois o método proposto parte da premissa de que a banda Cr do sistema YCbCr, muito usada para detecção de pele (Kakumanu et al., 2007), destaca os contornos da junção das lentes dos óculos em contraste com a cor da pele e ainda é menos afetada pela variação de iluminação encontrada nas faces. Espera-se, assim, que a junção das lentes esteja contida na Região S e seja realçada após detecção de bordas horizontais, levando em conta que a mesma estende-se em geral na direção horizontal.

O filtro de Sobel para detecção de bordas horizontais é aplicado na banda Cr da Região S. Em seguida, a imagem resultante da detecção de bordas horizontais ($CrEdges$) é binarizada ($CrEdges_{bin}$). Busca-se, então, a linha da imagem $CrEdges_{bin}$ que contém maior número de pixels ($N_{máx}$) que correspondem a bordas. A probabilidade (P_O) de ocorrência de óculos na imagem de entrada é então calculada como a proporção entre $N_{máx}$ e a largura total da Região S (S_L) como na Equação 4.4.

$$P_O = \frac{4N_{max}}{D_{olhos}} \quad (4.4)$$

Quanto maior o valor de P_O , mais provável será a ocorrência de óculos na imagem de entrada. Dessa maneira, se P_O for maior que um limiar pré-definido, a imagem será marcada como possuindo óculos, caso contrário será classificada como não possuindo óculos. Na Figura 4.11 é possível visualizar o resultado de algumas etapas da execução do algoritmo. A Figura 4.11 ilustra, para faces com e sem óculos respectivamente, a delimitação da Região S, a banda Cr isolada e a binarização das bordas horizontais da banda Cr, $CrEdges_{bin}$.

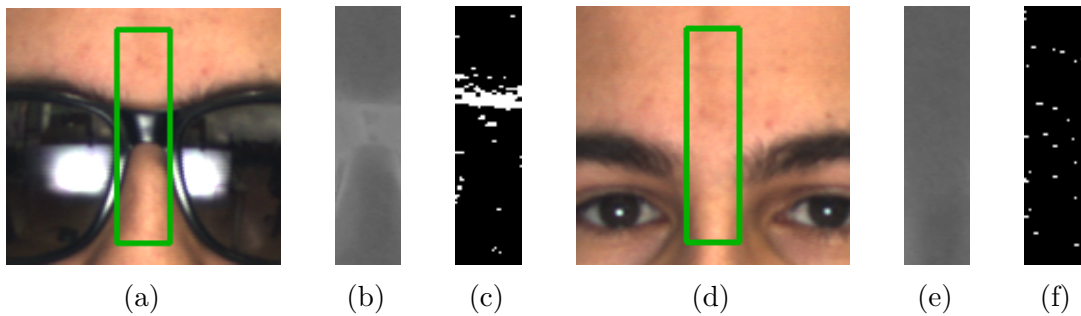


Figura 4.11: Principais etapas da execução do detector para imagem com e sem a presença de óculos: Seleção da Região S (4.11a e 4.11d), isolamento da banda Cr (4.11b e 4.11e), Binarização ($CrEdges_{bin}$) (4.11c e 4.11f).

4.2.4.2 Verificação do encobrimento dos olhos por armação de óculos

O algoritmo proposto para a verificação do encobrimento dos olhos causado por armação de óculos também explora a informação das bordas. Partindo da premissa de que a moldura dos óculos apresenta bordas bem definidas, o algoritmo busca linhas contínuas na região dos olhos. Para tal, usa-se realçamento de bordas através dos métodos de Sobel e Canny, seguido de detecção de linhas com a transformada de Hough.

Primeiramente, os olhos são localizados através de um detector de elementos faciais (Kazemi e Sullivan, 2014). Em seguida, uma área bem abrangente da região ocular, de cada olho separadamente, é recortada da imagem da face e convertida para níveis de cinza. Então, é realizada uma etapa de realce de bordas através de dois métodos: Canny e Sobel. As imagens resultantes do filtro de Sobel passam por uma etapa de expansão de histograma e binarização. Ao final deste processo, as imagens de realce de bordas resultantes (tanto do método de Canny quanto de Sobel) são submetidas à detecção de retas com a transformada de Hough.

Após a detecção e análise das linhas, é feita a contagem do número de segmentos que se encontram dentro da região correspondente à metade superior da íris. As Figuras 4.12 e 4.13 ilustram todo este processo, para imagens de faces onde, respectivamente, ocorre e não ocorre encobrimento dos olhos causado por óculos. Na figura, pode-se ver, ao final do processo, as áreas destacadas em verde que correspondem às linhas detectadas. Já os quadriláteros vermelhos correspondem à metade superior da íris. O valor que indica o nível de encobrimento de olhos é dado pela maior porcentagem de retas que se apresentam sobre a íris em relação ao número total de retas detectadas, considerando as quatro imagens tidas ao final do processo.

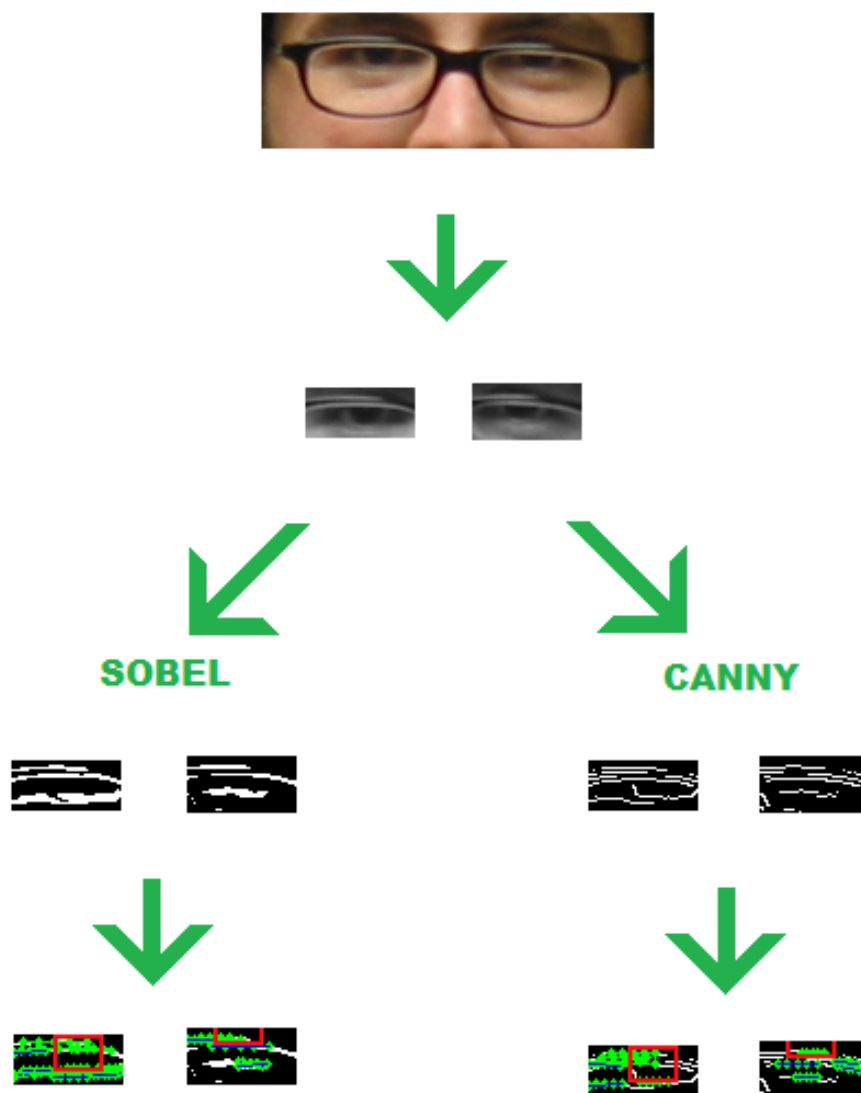


Figura 4.12: Etapas principais do processo de análise do encobrimento dos olhos por armação de óculos (imagem com ocorrência do requisito). A primeira linha da figura exibe um recorte colorido dos olhos. A segunda linha exibe o recorte de cada olho em níveis de cinza. A terceira linha exibe, à esquerda, o resultado da binarização do filtro de Sobel com expansão de histograma. À direita, a terceira linha exibe o resultado da aplicação do filtro de Canny. A última linha da figura exibe, destacadas em verde, as retas detectadas pela transformada de Hough, assim como a região superior da íris demarcada com um quadrilátero vermelho.

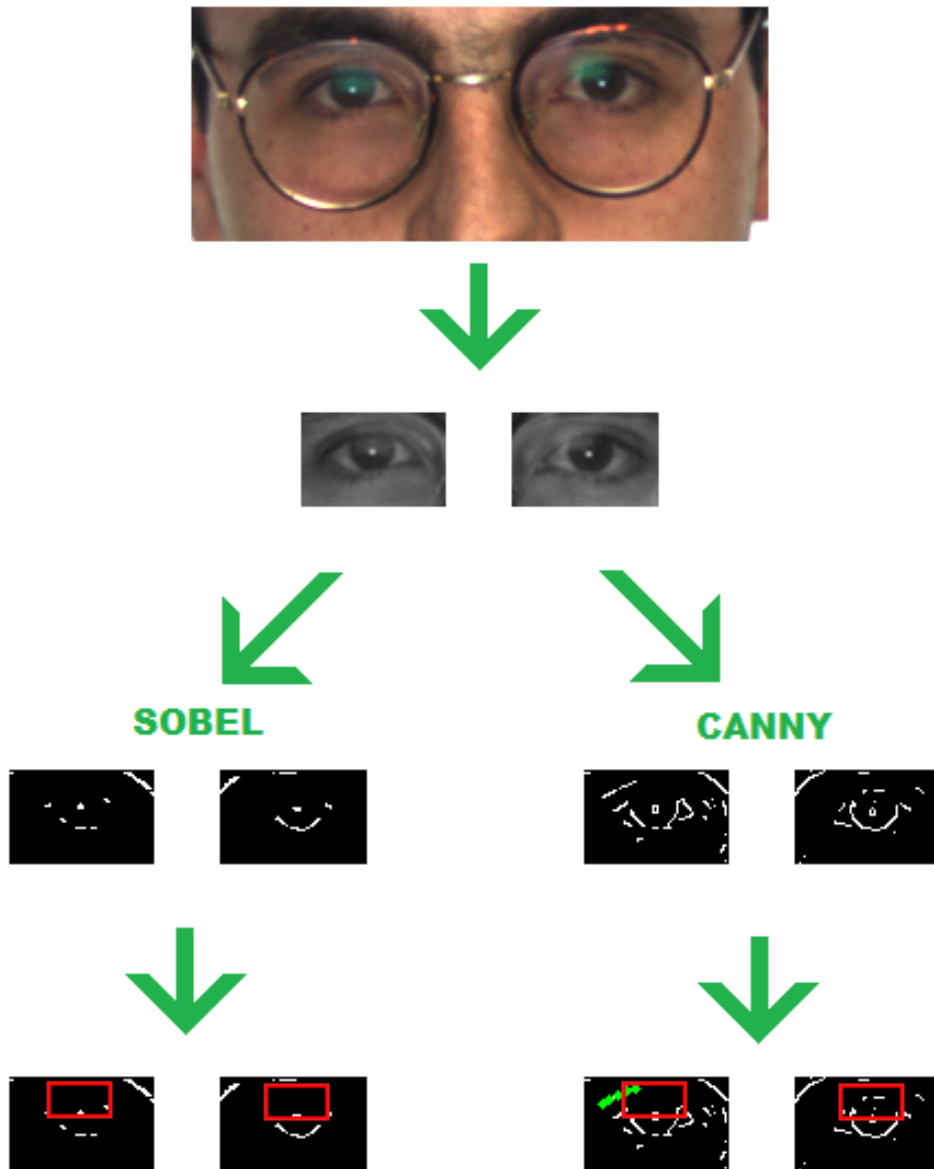


Figura 4.13: Etapas principais do processo de análise do encobrimento dos olhos por armação de óculos (imagem sem ocorrência do requisito). A primeira linha da figura exibe um recorte colorido dos olhos. A segunda linha exibe o recorte de cada olho em níveis de cinza. A terceira linha exibe, à esquerda, o resultado da binarização do filtro de Sobel com expansão de histograma. À direita, a terceira linha exibe o resultado da aplicação do filtro de Canny. A última linha da figura exibe, destacadas em verde, as retas detectadas pela transformada de Hough, assim como a região superior da íris demarcada com um quadrilátero vermelho.

Capítulo 5

Avaliação Experimental

Nesta seção são apresentados os resultados dos testes experimentais realizados através da plataforma Biolab ICAO.

5.1 Resultados

O desempenho dos algoritmos propostos no presente trabalho foi analisado através do *framework* Biolab-ICAO (Maltoni et al., 2009), que foi desenvolvido para prover à comunidade científica uma base de referência comum para a avaliação de métodos de validação de imagens faciais conforme o padrão ISO/ICAO. Os resultados experimentais dos métodos apresentados no Capítulo 4 são comparados com métodos propostos por cinco outros SDKs. O BiolabSDK, assim como SDK1 e SDK2 (que são comerciais e não tiveram seus nomes divulgados), foram avaliados na mesma plataforma e tiveram seus resultados publicados em Ferrara et al. (2012b). Os SDKs *BioTest* e *id3 ICAO* também foram avaliados, tendo seus resultados divulgados na plataforma (BioLab, 2006).

A Tabela 5.1 apresenta as taxas de erro (EER) alcançadas pelos métodos. A coluna "Rej." da tabela é referente à taxa de rejeição, ou seja, a porcentagem de imagens que não puderam ser avaliadas pelo SDK por quaisquer razões.

Tabela 5.1: Valores de EER e taxas de Rejeição (Rej.) dos softwares avaliados

Requisito	Foco		Cabelos		Chapéu		Armação	
	EER	Rej.	EER	Rej.	EER	Rej.	EER	Rej.
SDK1	26,0%	8,9%	50,0%	81,9%	-	-	50,6%	62,3%
SDK2	48,1%	0,6%	-	-	-	-	-	-
BioLabSDK	5,2%	0,0%	12,8%	0,0%	14,0%	0,0%	6,3%	0,0%
BioTest	30,5%	36,0%	12,4%	4,6%	16,5%	21,6%	4,0%	31,9%
id3 ICAO	1,7%	0,2%	-	-	6,8%	0,8%	6,6%	2,3%
Proposto	6,2%	0,8%	11,9%	4,2%	7,2%	2,3%	11,6%	1,2%

O método proposto para avaliação do nível de foco de uma imagem atingiu uma taxa de erro de 6,2% e rejeição de 0,8%. Este é um resultado compatível com os melhores resultados, alcançados pelo *BiolabSDK* (EER=5,2% e Rej.=0,0%) e *id3 ICAO* (EER=1,7% e Rej.=0,2%). A Figura 5.1 apresenta o gráfico das curvas de FAR x FRR para este algoritmo. No ponto onde as taxas de erro se equivalem o valor do limiar é baixo, o que indica que o método atribui valores altos para amostras com alto nível de conformidade com o requisito.

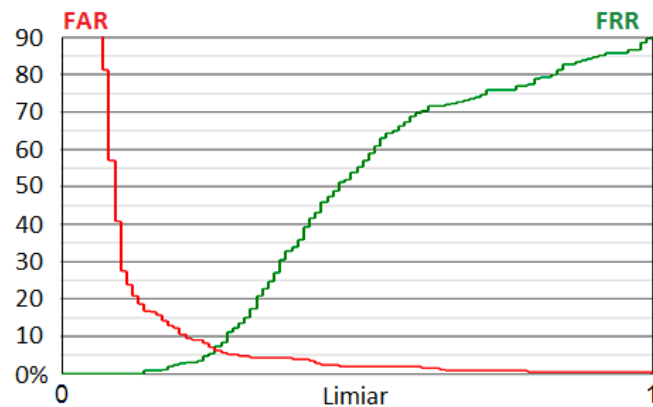


Figura 5.1: Gráfico de FAR x FRR para o requisito de análise do nível de foco.

O algoritmo proposto para avaliação do requisito de cabelos em frente aos olhos obteve uma taxa de erro de 11,9% e rejeição de 4,2%. Esta taxa de erro é a menor entre todos os outros métodos presentes na literatura. A Figura 5.2 apresenta o gráfico das

curvas de FAR x FRR para este algoritmo. Neste, o valor do limiar é alto no ponto onde as taxas de erro se equivalem. Isto acontece porque a análise deste requisito não é uma tarefa simples pois as classes não apresentam separabilidade alta. Assim, uma classificação adequada de uma amostra facial só é possível para valores altos do nível de conformidade do requisito.

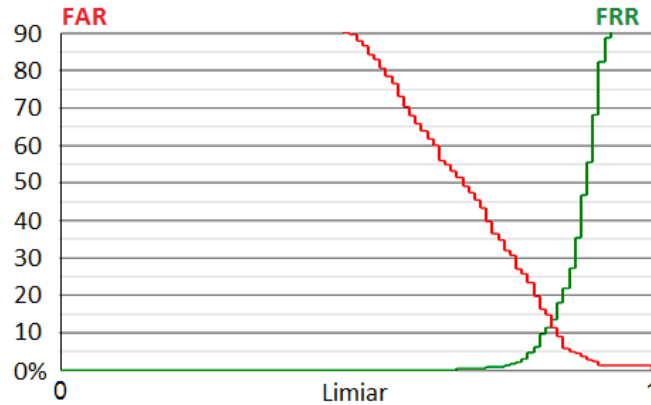


Figura 5.2: Gráfico de FAR x FRR para o requisito de análise da presença de cabelo em frente aos olhos.

O algoritmo proposto para avaliação da presença de chapéu atingiu uma taxa de erro de 7,2% e rejeição de 2,3%. Esta taxa de erro não é o melhor resultado obtido. Entretanto, o algoritmo proposto mostra-se competitivo, com um resultado compatível com o melhor, obtido pelo *id3 ICAO* (EER=6,8% e Rej.=0,8%). A Figura 5.3 apresenta o gráfico das curvas de FAR x FRR para este algoritmo. De forma análoga ao método de análise dos cabelos, uma classificação com mínima taxa de erro só é possível para valores altos do nível de conformidade, pois o limiar do ponto EER é muito próximo de 1. Isto mostra o quanto a avaliação deste requisito não é trivial. De toda forma, o resultado obtido pelo método proposto para avaliação deste requisito representa uma melhora significativa com relação aos resultados alcançados pelos demais SDKs. Estes são o BiolabSDK e BioTest, com taxas EER respectivamente de 14,0% e 16,5%.

O método de avaliação da ocorrência de encobrimento dos olhos por armação de

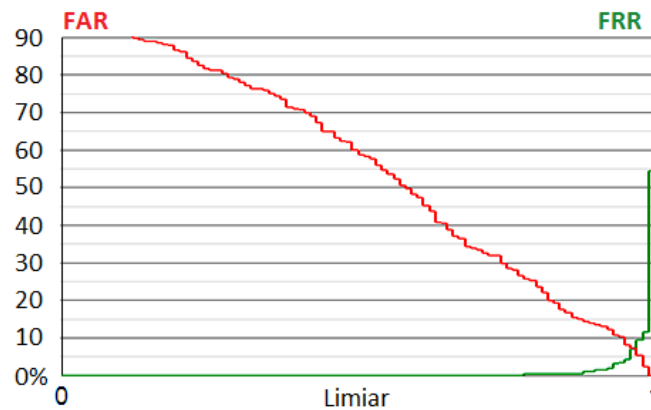


Figura 5.3: Gráfico de FAR x FRR para o requisito de análise da presença de chapéu.

óculos atingiu uma taxa de erro de 11,6% e rejeição de 1,2%. Este resultado fica um pouco atrás da concorrência, onde o melhor resultado foi conseguido pelo *BioTest* (EER=4,0% e Rej.=31,9%). A Figura 5.4 apresenta o gráfico das curvas de FAR x FRR para este método proposto. Novamente, uma classificação com mínima taxa de erro só é possível para valores altos do nível de conformidade. Assim como a análise da presença de cabelo em frente aos olhos, analisar se a armação dos óculos encobre os olhos trata-se de uma tarefa não trivial, uma vez que tal região apresenta bastantes detalhes.

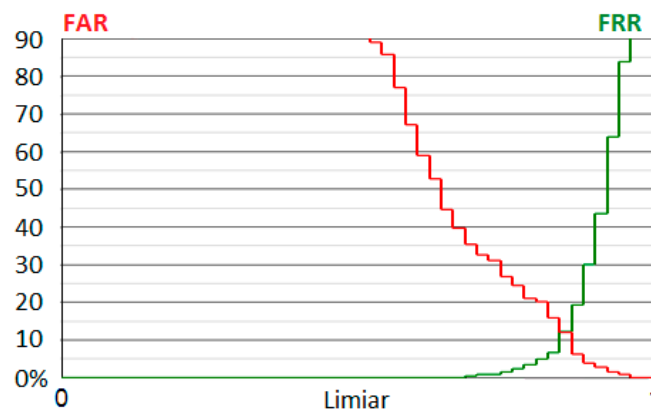


Figura 5.4: Gráfico de FAR x FRR para o requisito de encobrimento dos olhos por armação de óculos.

5.2 Análise dos erros dos métodos propostos

Os casos onde os métodos propostos cometem erros serão analisados nesta seção. Tal análise é feita para cada requisito abordado no presente trabalho, sendo dividida entre casos de falsa aceitação e falsa rejeição. A mesma foi realizada através de uma ferramenta de software própria, dedicada a testes de desempenho de algoritmos para análise facial de acordo com o ICAO, usando o conjunto de imagens de teste local disponibilizadas pela plataforma Biolab-ICAO.

5.2.1 Erros na aferição do nível de foco

Aqui são analisados os erros cometidos pelo algoritmo da análise de foco, tanto falsas aceitações quanto as falsas rejeições.

5.2.1.1 Falsas aceitações

O algoritmo proposto para análise de foco aceitou indevidamente apenas duas imagens nas quais a face apresenta desfoque e o indivíduo está usando óculos metálicos, os quais refletem a luz do flash da câmera fotográfica. Como o algoritmo avalia o nível de foco através de filtro de detecção de bordas, este detecta muito bem os contornos dos reflexos das armações, como apresentado na Figura 5.5.



Figura 5.5: Exemplo de face que apresenta desfoque e reflexo na armação dos óculos. A Figura 5.5a apresenta a face desfocada enquanto que a 5.5b apresenta a detecção de bordas na região dos olhos.

Este tipo de erro expõe um ponto fraco do algoritmo, uma vez que o fenômeno

de desfoque ocorrido é legítimo, originado por desfoque fotográfico e não por filtro de suavização. Os reflexos desfocados presentes na imagens apresentam forma redonda muito bem definida, que é detectada no processo de realce de bordas com algoritmo de Sobel.

5.2.1.2 Falsas rejeições

A maioria dos erros de falsa rejeição cometidos pelo algoritmo de análise de foco ocorreram em imagens pixelizadas, ou seja, imagens onde cada pixel lógico é representado por vários pixels físicos da imagem, como pode ser visualizado nas Figuras 5.6a e 5.7a.



Figura 5.6: Exemplo de imagem que apresenta pixelização e que o algoritmo de análise de foco rejeitou como se estivesse embaçada. A Figura 5.6a apresenta um recorte da face pixelizada enquanto que a 5.6b apresenta a detecção de bordas com o filtro de Sobel na região dos olhos demarcada.



Figura 5.7: Exemplo de imagem que apresenta pixelização e que o algoritmo de análise de foco rejeitou como se estivesse embaçada. A Figura 5.7a apresenta um recorte da face pixelizada enquanto que a 5.7b apresenta a detecção de bordas com o filtro de Sobel na região dos olhos demarcada.

É possível observar nas Figuras 5.6b e 5.7b que o algoritmo de Sobel detecta poucas bordas na região ocular. Isto acontece porque pixelização e desfoque são fenômenos

semelhantes, nos quais acontece uma ausência de altas frequências. Isso é devido ao fato de os blocos de pixelização não apresentarem transição de um pixel para outro, assim como o fato da transição entre blocos ser suave. Desta forma as bordas detectadas são pontuais e ocorrem nos vértices dos blocos quadrados de pixelização. A ausência de altas frequências, que causa pequena detecção de bordas, leva o algoritmo a confundir as imagens pixelizadas com as embaçadas.

A composição entre o algoritmo de aferição de foco e um de pixelização, com fusão da informação a respeito de cada requisito, pode ser uma boa alternativa para fugir de erros deste tipo.

5.2.2 Erros na detecção de cabelo encobrindo os olhos

Aqui são analisados os erros de falsa aceitação e falsa rejeição cometidos pelo método de detecção de cabelo em frente aos olhos. Para este requisito os erros aconteceram principalmente em função do fato do método proposto explorar a simetria do rosto.

5.2.2.1 Falsas aceitações

O método de detecção de cabelo em frente aos olhos apresenta erros de classificação em fotografias com anormalidades na iluminação. A Figura 5.8 apresenta dois exemplos de imagens onde o método proposto erra. Na Figura 5.8a ocorre excesso de iluminação em um dos lados da face, o que termina deixando imperceptível a presença de uma mecha de cabelos loiros que aparece encobrindo os olhos. Já na Figura 5.8b o contraste é baixo, o que tem impacto negativo na detecção de bordas, a qual não realça os contornos do cabelo aparente.

Em ambos os casos a detecção de bordas termina apresentando assimetria pequena entre os olhos, o que leva o método a aprovar tais amostras erroneamente. A execução do algoritmo para as faces das Figuras 5.8a e 5.8b é apresentada em detalhes nas Figuras 5.9 e 5.10, respectivamente.



Figura 5.8: Exemplos de fotos com alterações de iluminação onde ocorreram erros de falsa aceitação para o requisito de cabelo em frente aos olhos.

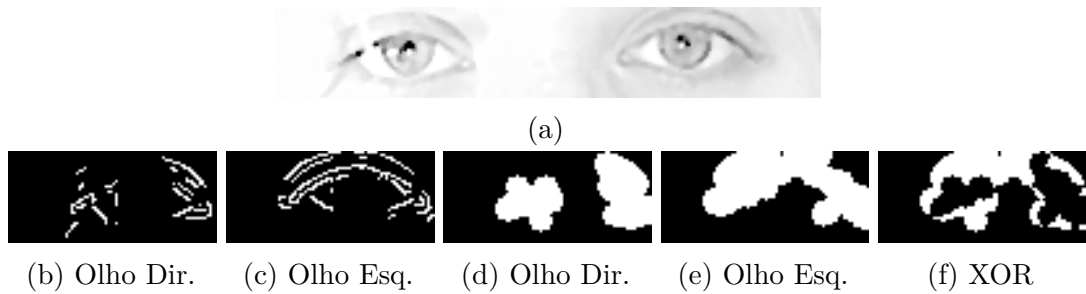


Figura 5.9: Etapas da execução do algoritmo para a imagem apresentada na Figura 5.8a. Recorte da região dos olhos (5.9a). Detecção de bordas com Canny (5.9b e 5.9c). Dilatação e erosão de bordas (5.9d e 5.9e). Operação XOR bit-a-bit (5.9f).

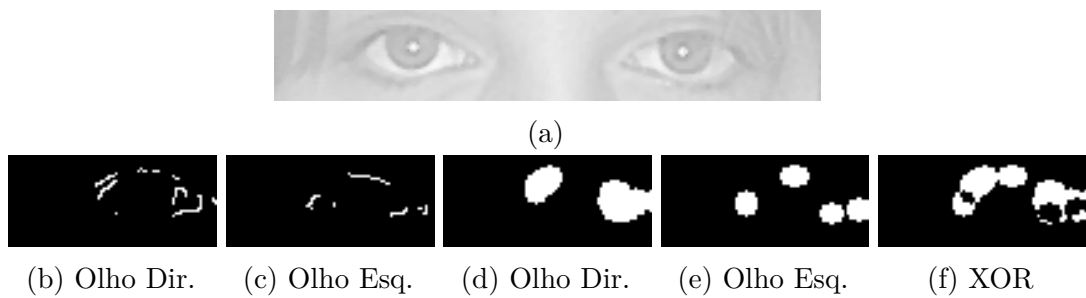


Figura 5.10: Etapas da execução do algoritmo para a imagem apresentada na Figura 5.8b. Recorte da região dos olhos (5.10a). Detecção de bordas com Canny (5.10b e 5.10c). Dilatação e erosão de bordas (5.10d e 5.10e). Operação XOR bit-a-bit (5.10f).

5.2.2.2 Falsas rejeições

Quando não há cabelo em frente aos olhos, detalhes quaisquer da região dos olhos tais como rugas, sinais na pele, levam a um resultado assimétrico na detecção das bordas. O algoritmo erra, então, considerando tal assimetria como encobrimento dos olhos por cabelo. A Figura 5.11 apresenta um destes casos, onde uma marca de nascença, próxima ao olho direito do indivíduo, provoca uma assimetria na etapa de detecção de bordas. Tal assimetria é amplificada na etapa de dilatação de bordas e detectada pela operação XOR, que destaca as diferenças entre olho direito e esquerdo.

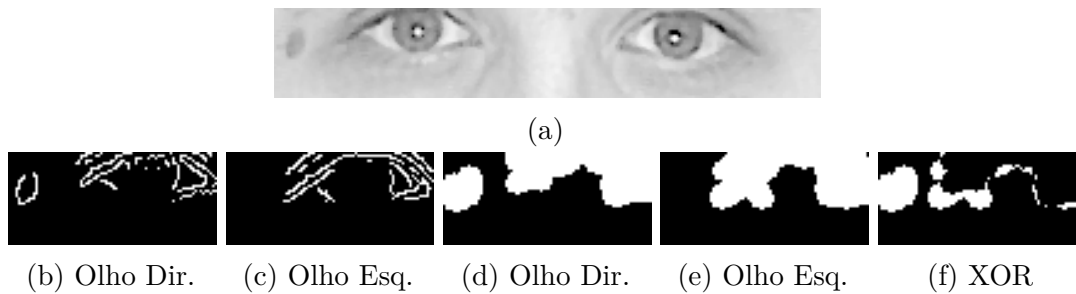


Figura 5.11: Etapas da execução do algoritmo para um caso de falsa rejeição. Recorte da região dos olhos (5.11a). Detecção de bordas com Canny (5.11b e 5.11c). Dilatação e erosão de bordas (5.11d e 5.11e). Operação XOR bit-a-bit (5.11f).

5.2.3 Erros na detecção de chapéu ou boné

Nesta seção são analisados os erros de falsa aceitação e falsa rejeição cometidos pelo algoritmo de detecção de chapéu ou boné.

5.2.3.1 Falsas aceitações

O algoritmo proposto para detecção de chapéu ou boné aceitou indevidamente algumas imagens nas quais o indivíduo usa o acessório de cabeça encobrindo uma fração pequena da região da testa. Ou seja, o método que é baseado em detecção de pele encontra muita pele quando o chapéu não encobre muito a região analisada, ou seja, a testa. As

Figuras 5.12 e 5.13 apresentam dois exemplos deste tipo de erro.

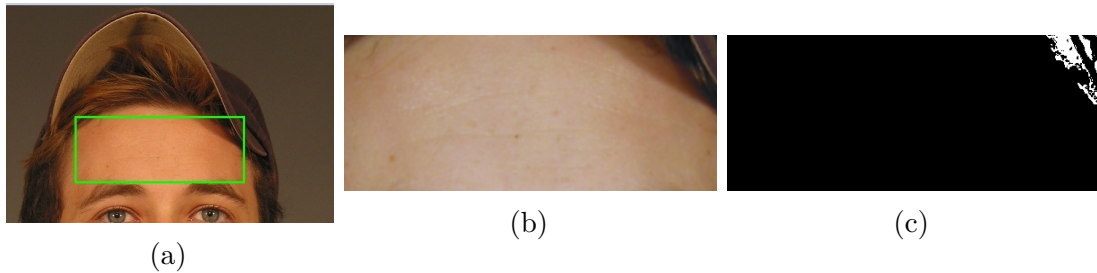


Figura 5.12: Exemplo de erro de falsa aceitação na detecção de chapéu. A Figura 5.12a apresenta a face do indivíduo, a Figura 5.12b ilustra o recorte da região da testa e a Figura 5.12c apresenta o resultado da detecção de pele.

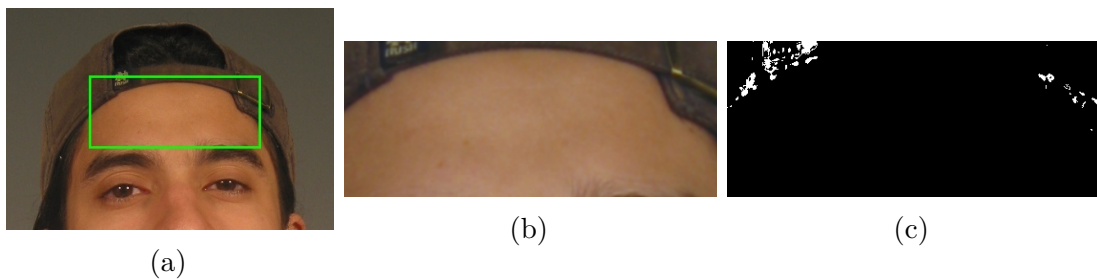


Figura 5.13: Exemplo de erro de falsa aceitação na detecção de chapéu. A Figura 5.13a apresenta a face do indivíduo, a Figura 5.13b ilustra o recorte da região da testa e a Figura 5.13c apresenta o resultado da detecção de pele.

Tais exemplos mostram que a abordagem baseada em detecção de pele na região da testa é vulnerável em casos como este, onde o gosto pessoal no uso de acessórios de cabeça dificulta a operação do método proposto. Assim, é possível que uma abordagem com aprendizado de máquina, com um classificador devidamente treinado para detectar chapéus, entre outros acessórios de cabeça, apresente um desempenho melhor, buscando o acessório inteiro, na imagem inteira, não apenas se limitando à região da testa.

5.2.3.2 Falsas rejeições

Os erros de falsas rejeições relacionadas à detecção de chapéu foram causados em sua maioria por obstrução da pele da testa causada por franja comprida, muitas vezes em

indivíduos do sexo feminino. Exemplos deste caso de erro podem ser visualizados nas Figuras 5.14, 5.15 e 5.16. Nestas, principalmente as da coluna (c), pode-se observar o quanto a obstrução da pele da testa leva o algoritmo à análise equivocada da presença de chapéu.

Novamente, técnicas de aprendizado de máquina, com uso de um classificador devidamente treinado para detecção de chapéu enquanto um objeto único, podem prover um desempenho mais satisfatório, uma vez que a análise do requisito não fica atrelada à presença de pele na região da testa.

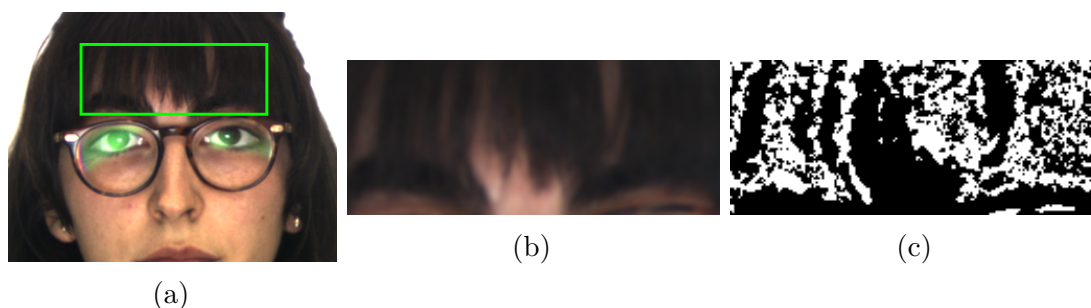


Figura 5.14: Exemplo de erro de falsa aceitação na detecção de chapéu. A Figura 5.14a apresenta a face do indivíduo, a Figura 5.14b ilustra o recorte da região da testa e a Figura 5.14c apresenta o resultado da detecção de pele.

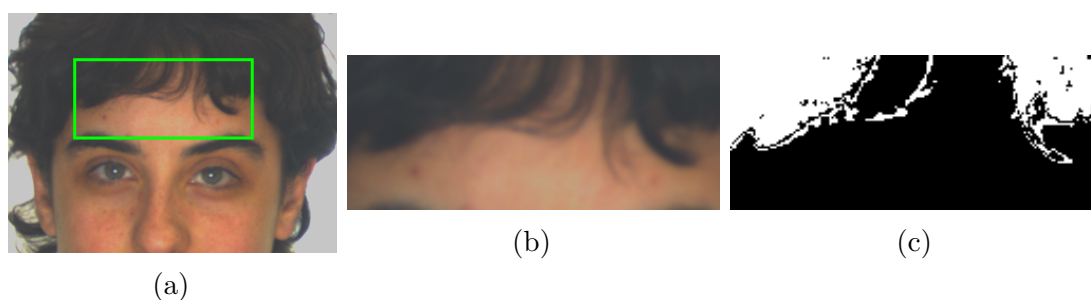


Figura 5.15: Exemplo de erro de falsa aceitação na detecção de chapéu. A Figura 5.15a apresenta a face do indivíduo, a Figura 5.15b ilustra o recorte da região da testa e a Figura 5.15c apresenta o resultado da detecção de pele.

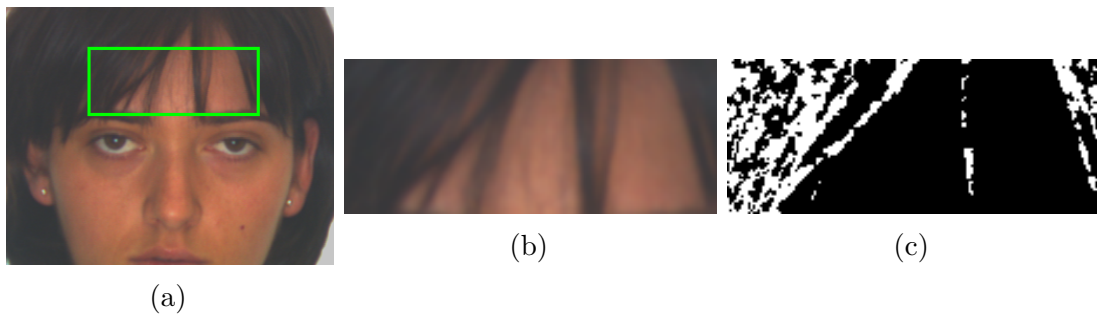


Figura 5.16: Exemplo de erro de falsa aceitação na detecção de chapéu. A Figura 5.16a apresenta a face do indivíduo, a Figura 5.16b ilustra o recorte da região da testa e a Figura 5.16c apresenta o resultado da detecção de pele.

5.2.4 Armação encobrindo olhos

Nesta seção são analisados os casos onde erra o algoritmo que analisa o encobrimento dos olhos causado por armação de óculos.

5.2.4.1 Falsas aceitações

Aconteceram dois erros de falsa aceitação. Ambos os casos envolvem fotografias onde a face aparece com desfoque acentuado, o que deixa claro que desajuste de foco tem impacto negativo na análise do requisito de encobrimento dos olhos por armação de óculos. As Figuras 5.17 e 5.18 apresentam estes dois casos de erro, exibindo os resultados da detecção de bordas, na região dos olhos, tanto através do método de Sobel como pelo de Canny. Pode-se observar que o foco ruim implica em poucas bordas detectadas.

Como o método busca bordas horizontais na região dos olhos, considerando que as mesmas são causadas por armação de óculos, a ocorrência de desfoque dificulta a detecção da armação nesta região. Nas imagens em questão a face aparece sombreada, o que é um agravante a mais para não detecção das armações. A sombra na região ocular contrasta pouco com as armações de cor escura, que claramente encobrem parte dos olhos. Entretanto, o desfoque pode ser considerado o principal causador dos erros, pois interfere negativamente na etapa de detecção de contornos.

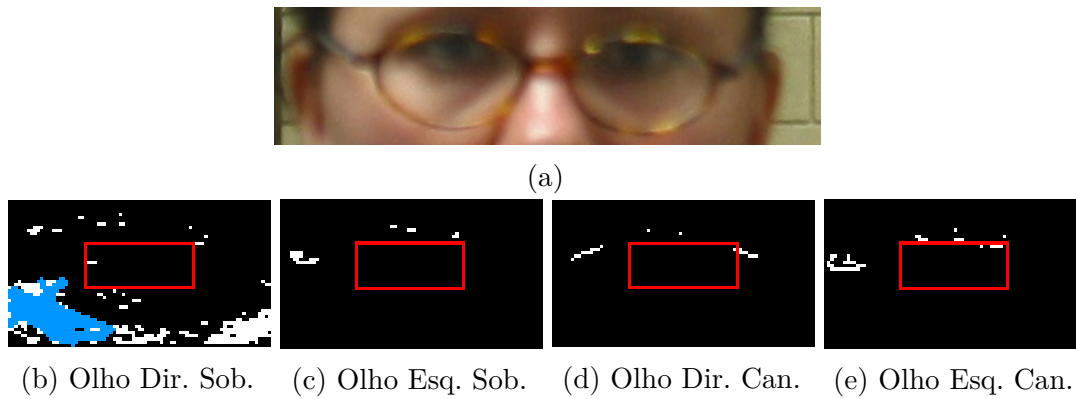


Figura 5.17: Caso de erro de falsa aceitação. As Figuras 5.17b e 5.17c apresentam a detecção de bordas com Sobel para os olho direito e esquerdo, respectivamente, enquanto que as figuras 5.17d e 5.17e fazem o mesmo para detecção com canny.

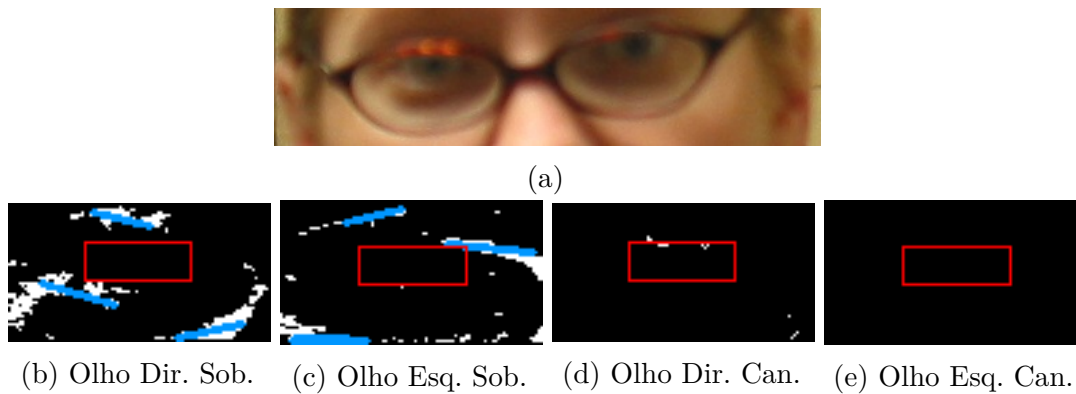


Figura 5.18: Caso de erro de falsa aceitação. As Figuras 5.18b e 5.18c apresentam a detecção de bordas com Sobel para os olho direito e esquerdo, respectivamente, enquanto que as figuras 5.18d e 5.18e fazem o mesmo para detecção com Canny.

5.2.4.2 Falsas rejeições

Entre os erros de falsa rejeição ocorreram muitos casos onde o método detecta retas nos olhos do indivíduo, interpretando-as como armação de óculos. O algoritmo espera detectar segmentos de reta na região dos olhos através da transformada de Hough. Entretanto, apesar da forma arredondada dos contornos dos olhos, há casos em que são detectadas retas nas pálpebras. As Figuras 5.19 e 5.20 apresentam casos onde é evidente que não há encobrimento dos olhos por armação. Pode-se observar nas imagens

de detecção de bordas que a transformada de Hough detecta retas nas pálpebras, destacadas em amarelo, interseccionando a região da íris.

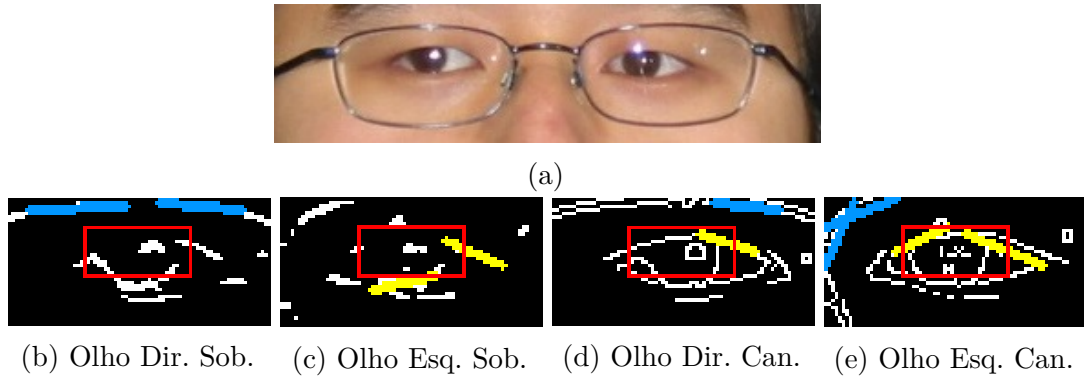


Figura 5.19: Caso de erro de falsa rejeição. As Figuras 5.19b e 5.19c apresentam a detecção de bordas com Sobel para os olho direito e esquerdo, respectivamente, enquanto que as figuras 5.19d e 5.19e fazem o mesmo para detecção com canny.

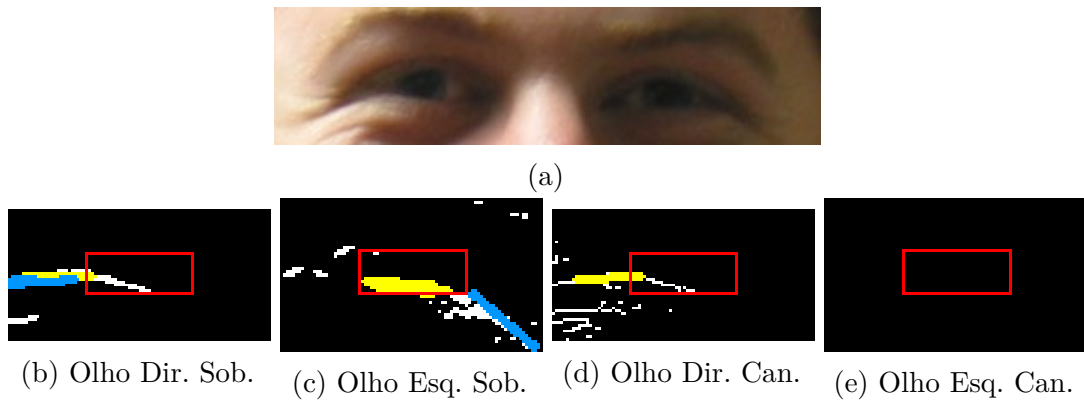


Figura 5.20: Caso de erro de falsa rejeição. As Figuras 5.20b e 5.20c apresentam a detecção de bordas com Sobel para os olho direito e esquerdo, respectivamente, enquanto que as figuras 5.20d e 5.20e fazem o mesmo para detecção com canny.

Capítulo 6

Considerações Finais e Trabalhos

Futuros

Este trabalho apresentou novos métodos para a avaliação da conformidade de imagens faciais com alguns dos requisitos de qualidade impostos pelo padrão internacional ISO/IEC 19794-5. Os métodos propostos foram avaliados através da plataforma FVC-OnGoing, mantido pelo laboratório BioLab e também conhecido como BioLab ICAO. Através dos dados apresentados no Capítulo 5, pode-se concluir que os métodos aqui propostos apresentam desempenho compatível com métodos do estado-da-arte, quando não superam os mesmos.

O método de avaliação do nível de foco tem um desempenho que concorre com o estado-da-arte, apresentando uma taxa EER de 6,2%. O de análise da presença de cabelo na frente dos olhos alcançou um resultado que supera todos os outros métodos presentes na literatura relacionada com uma EER de 11,9%. O método que analisa a presença de chapéu apresentou resultado muito próximo do estado-da-arte, uma EER de 7,2%. O algoritmo de detecção de armação de óculos em frente aos olhos apresentou taxa de erro um pouco superior à concorrência, EER de 11,6%, devido principalmente à grande quantidade de detalhes presentes na região ocular.

Como trabalho futuro pretende-se implementar soluções baseadas em aprendizado de máquina para os requisitos ICAO apresentados neste trabalho, assim como todos os outros. Aprendizagem de máquina tem se tornado uma ferramenta cada vez mais presente em aplicações de visão computacional e processamento de imagens, com resultados surpreendentemente satisfatórios. Será feito o treinamento de classificadores do tipo Máquina de Vetor de Suporte (do inglês, Support Vector Machine, ou SVM), o que depende de um grande conjunto de amostras faciais anotadas de acordo com os requisitos do ICAO. Tal conjunto de dados encontra-se em construção e será a base do prosseguimento das novas pesquisas do grupo de pesquisadores da VSOFTE, grupo do qual este trabalho é fruto.

Referências Bibliográficas

BioLab (2006). Fvc web site (<https://biolab.csr.unibo.it/fvcongoing>).

Bracewell, R. (1965). The fourier transform and iis applications. *New York*, 5.

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, pgs. 679–698.

Duda, R. O. e Hart, P. E. (1972). Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15.

Fernández, A., García, R., Usamentiaga, R., e Casado, R. (2015). Glasses detection on real images based on robust alignment. *Machine Vision and Applications*, 26(4):519–531.

Ferrara, M., Franco, A., e Maio, D. (2012a). A multi-classifier approach to face image segmentation for travel documents. *Expert Systems with Applications*, 39(9):8452–8466.

Ferrara, M., Franco, A., Maio, D., e Maltoni, D. (2012b). Face image conformance to iso/icao standards in machine readable travel documents. *Information Forensics and Security, IEEE Transactions on*, 7(4):1204–1213.

Ferrara, M., Franco, A., e Maltoni, D. (2008). Evaluating systems assessing face-image

- compliance with icao/iso standards. Em *Biometrics and Identity Management*, pgs. 191–199. Springer.
- Fisher, R., Perkins, S., Walker, A., e Wolfart, E. (1994). Hypermedia image processing reference. *Department of Artificial Intelligence, University of Edinburgh*.
- Freund, Y., Schapire, R., e Abe, N. (1999). A short introduction to boosting. *Journal-Japanese Society For Artificial Intelligence*, 14(771-780):1612.
- Frieden, B. R. (2004). *Science from Fisher information: a unification*. Cambridge University Press.
- Gonzalez, R. C. e Woods, R. E. (2008). Digital image processing. *Nueva Jersey*.
- Group, I. N. T. W. et al. (2004). Biometrics deployment of machine readable travel documents.
- Huang, G. B., Ramesh, M., Berg, T., e Learned-Miller, E. (2007). Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical report, Technical Report 07-49, University of Massachusetts, Amherst.
- ISO, I. S. O. (2005). *ISO/IEC 19794-5:2005, Information Technology - Biometric Data Interchange Formats - Part 5: Face Image Data/Amendment 1: Conditions for Taking Photographs or Face Image Data 2007*. Number pt. 5 in International Standard. ISO.
- Jiang, X., Binkert, M., Achermann, B., e Bunke, H. (2000). Towards detection of glasses in facial images. *Pattern Analysis & Applications*, 3(1):9–18.
- Joblove, G. H. e Greenberg, D. (1978). Color spaces for computer graphics. Em *ACM siggraph computer graphics*, volume 12, pgs. 20–25. ACM.
- Kakumanu, P., Makrogiannis, S., e Bourbakis, N. (2007). A survey of skin-color modeling and detection methods. *Pattern recognition*, 40(3):1106–1122.

- Kazemi, V. e Sullivan, J. (2014). One millisecond face alignment with an ensemble of regression trees. Em *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pgs. 1867–1874.
- Krotkov, E. (1988). Focusing. *International Journal of Computer Vision*, 1(3):223–237.
- Liu, H., Gao, W., Miao, J., e Li, J. (2002). A novel method to compensate variety of illumination in face detection. Em *JCIS*, pgs. 692–695. Citeseer.
- Maltoni, D., Franco, A., Ferrara, M., Maio, D., e Nardelli, A. (2009). Biolab-ica: A new benchmark to evaluate applications assessing face image compliance to iso/iec 19794-5 standard. Em *ICIP*, pgs. 41–44.
- Martinez, A. M. (1998). The ar face database. *CVC Technical Report*, 24.
- Newton, I. (1730). *Óptica*.
- OpenCV, L. (2008). Computer vision with the opencv library. *GaryBradski & Adrian Kaebler-O'Reilly*.
- Papageorgiou, C. P., Oren, M., e Poggio, T. (1998). A general framework for object detection. Em *Computer vision, 1998. sixth international conference on*, pgs. 555–562. IEEE.
- Pech-Pacheco, J. L., Cristóbal, G., Chamorro-Martinez, J., e Fernández-Valdivia, J. (2000). Diatom autofocusing in brightfield microscopy: a comparative study. Em *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 3, pgs. 314–317. IEEE.
- Phillips, P. J., Wechsler, H., Huang, J., e Rauss, P. J. (1998). The feret database and evaluation procedure for face-recognition algorithms. *Image and vision computing*, 16(5):295–306.

- Ross, A. A., Nandakumar, K., e Jain, A. (2006). *Handbook of multibiometrics*, volume 6. Springer Science & Business Media.
- Sobel, I. (2014). History and definition of the sobel operator. *Retrieved from the World Wide Web*.
- Storer, M., Urschler, M., e Bischof, H. (2010). Occlusion detection for icao compliant facial photographs. Em *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*, pgs. 122–129. IEEE.
- Vezhnevets, V., Sazonov, V., e Andreeva, A. (2003). A survey on pixel-based skin color detection techniques. Em *Proc. Graphicon*, volume 3, pgs. 85–92. Moscow, Russia.
- Viola, P. e Jones, M. J. (2004). Robust real-time face detection. *International journal of computer vision*, 57(2):137–154.
- Wang, X. (2007). Laplacian operator-based edge detectors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(5):886–890.
- Young, T. (1802). The bakerian lecture: On the theory of light and colours. *Philosophical transactions of the Royal Society of London*, 92:12–48.