



Universidade Federal do Piauí  
Centro de Ciências da Natureza  
Programa de Pós-Graduação em Ciência da Computação

# **Localização e Identificação de Pontos de Referência Facial para Cirurgias Plásticas**

**Ricardo Teles Freitas**

**Número de Ordem PPGCC: M028**

**Teresina-PI, 26 de agosto de 2016**

Ricardo Teles Freitas

# **Localização e Identificação de Pontos de Referência Facial para Cirurgias Plásticas**

**Dissertação de Mestrado** apresentada ao Programa de Pós-Graduação em Ciência da Computação da UFPI (área de concentração: Sistemas de Computação), como parte dos requisitos necessários para a obtenção do Título de Mestre em Ciência da Computação.

Universidade Federal do Piauí – UFPI

Centro de Ciências da Natureza

Programa de Pós-Graduação em Ciência da Computação

Orientador: Kelson Rômulo Teixeira Aires

Teresina-PI

26 de agosto de 2016

---

Ricardo Teles Freitas

Localização e Identificação de Pontos de Referência Facial para Cirurgias Plásticas/ Ricardo Teles Freitas. – Teresina-PI, 26 de agosto de 2016-  
102 p. : il. (algumas color.) ; 30 cm.

Orientador: Kelson Rômulo Teixeira Aires

Dissertação (Mestrado) – Universidade Federal do Piauí – UFPI  
Centro de Ciências da Natureza

Programa de Pós-Graduação em Ciência da Computação, 26 de agosto de 2016.

1. Pontos de Referência. 2. Cirurgia Plástica. I. Kelson Rômulo Teixeira Aires. II. Universidade Federal do Piauí. III. Centro de Ciências da Natureza. IV. Localização e Identificação de Pontos de Referência Facial para Cirurgias Plásticas

CDU 02:141:005.7

---

## Localização e Identificação de Pontos de Referência Facial para Cirurgias Plásticas

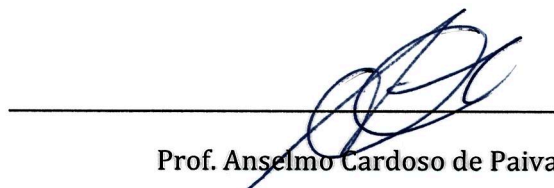
**RICARDO TELES FREITAS**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências da Natureza da Universidade Federal do Piauí, como parte integrante dos requisitos necessários para obtenção do grau de Mestre em Ciência da Computação.

Aprovado por:



Prof. Kelson Rômulo Teixeira Aires  
(Presidente da Banca Examinadora)



Prof. Anselmo Cardoso de Paiva  
(Examinador Externo)



Prof. Victor Eulálio Sousa Campêlo  
(Examinador Externo)



Prof. Rodrigo de Melo Souza Vêras  
(Examinador Interno)



Prof. Vinícius Ponte Machado  
(Examinador Interno)

Teresina, 26 de agosto de 2016

*Aos meus pais Marco Aurélio Carvalho de Freitas e Cleonice de Castro Teles,  
por sempre estarem comigo em todos os momentos.*

# Agradecimentos

Agradeço aos meus pais, Marco Aurélio Carvalho de Freitas e Cleonice de Castro Teles, por me apoiarem em todos os momentos da minha vida.

Ao meu irmão Marco Aurélio Teles Freitas, por sempre estar comigo e pelo apoio no trabalho. À minha companheira Marcelle Mendel Ribeiro Queiroz pelo apoio e incentivo.

Agradeço ao meu orientador, Kelson Rômulo Teixeira Aires, por todos os conselhos, pela paciência e ajuda nesse período.

Aos meus amigos Humberto, Maylson, Perichdey, Roniel pelo apoio incondicional nos momentos de necessidade. Aos colegas do corpo discente do Mestrado e aos colegas participantes deste projeto, Hugo e Orlando.

Aos professores da Universidade Federal do Piauí, em especial ao professor Victor Eulálio Sousa Campelo, participante deste projeto.

Agradeço à FAPEPI pelo apoio financeiro para a realização deste trabalho de pesquisa.

*“Reality is that which,  
when you stop believing in it,  
does not go away.”  
(Philip K. Dick)*

# Resumo

A beleza é uma qualidade que gera discussão e intriga o ser humano desde tempos remotos. Filósofos da Grécia Antiga como Platão e Aristóteles já discutiam sobre os aspectos objetivos e subjetivos dessa característica. Já na Renascença, o ideal de beleza humano foi revisto por vários artistas e cientistas daquele tempo. No século XX, os avanços da medicina proporcionaram uma abordagem ainda mais científica acerca da estética humana. Essa qualidade moveu a ciência através de um caminho que permite ao homem tanto corrigir anomalias congênitas como reparar desconformidades com os padrões estéticos de uma determinada sociedade ou grupo biológico e cultural. Para tanto, é indispensável que uma cirurgia corretiva seja precedida de uma rigorosa análise pré-operatória baseada em informações sobre os aspectos físicos do paciente. Nos procedimentos faciais, tais dados podem ser representados por pontos de referência, que são explorados para medir desvios estéticos. Este trabalho tem como objetivo propor uma metodologia que possibilite localizar e identificar os pontos de referência necessários para a realização de avaliações estéticas da face. A intenção é de auxiliar na definição de procedimentos cirúrgicos adequados às características de um paciente com base em cálculos de métricas de avaliação estética. A necessidade de separar a análise facial em duas diferentes perspectivas (perfil e frontal) gera abordagens distintas para a solução do problema. Para localizar os pontos do perfil foi elaborado um método inovador a partir de um conjunto de 30 imagens. Para a perspectiva frontal, foram realizadas modificações no método *Active Shape Models* (ASM) de modo a adaptá-lo ao conjunto de pontos frontais necessários para análises estéticas. As metodologias desenvolvidas permitem a localização e identificação de 35 pontos faciais nas perspectivas de perfil e frontal e foram testadas em mais de 600 imagens. O método de localização de pontos do perfil apresentou resultados que podem ser comparados a marcações de especialistas. O método de localização de pontos frontais foi capaz de melhorar a precisão do ASM clássico em até 44%.

**Palavras-chaves:** Pontos de Referência. Análise Facial. Cirurgia Plástica.



# Abstract

Beauty is a quality that arises discussion and intrigues the human being since old times. Ancient philosophers such as Plato and Aristotle already discussed on objective and subjective aspects of this feature. During the Renaissance, the beauty ideal was revised by many artists and scientists of that time. In the XX century, advances in medicine provided an approach even more critical over human aesthetics. This quality has guided the science through a path that allows man to both correct congenital anomalies and repair nonconformities with the aesthetic standards of a society. For such purpose, it is indispensable a corrective surgery to be preceded of a rigorous pre-operative analysis based on information about the physical aspects of the patient. In facial procedures, such data can be represented by landmarks, that are used to measure aesthetic deviation. This work aims to propose a methodology that allows the automatic location and identification of landmarks needed to evaluate a face aesthetically. It intends to help defining surgical procedures according to the patient's features based on aesthetic evaluation metrics. The need to divide the facial analysis into two different perspectives (profile and frontal) generates distinctive approaches to solve the problem. To locate profile landmarks a novel method was elaborated based on a set of 30 images. For the frontal perspective, the *Active Shape Models* (ASM) method was modified in order to locate the frontal landmarks needed for aesthetic analysis. The methodologies developed allow to locate and identify 35 facial landmarks in both profile and frontal perspectives and were tested in over 600 images. The method for locating profile landmarks presented results comparable to the ground truth. The method for locating frontal landmarks was able to improve classic ASM precision rate in 44%.

**Keywords:** Landmarks. Facial Analysis. Plastic Surgery.

# Lista de ilustrações

Figura 1	– <i>Doríforo</i> de Policleto, resultado de um tratado sobre proporções ideais (POLICLETO, 1797 (Excavada)). . . . .	3
Figura 2	– À esquerda, um recorte de <i>O Homem Vitruviano</i> de Leonardo Da Vinci. À direita, uma representação do tamanho da cabeça em relação ao corpo. . . . .	4
Figura 3	– Exemplo de estudo de proporções faciais realizado por Albrecht Durer. . . . .	4
Figura 4	– Exemplos de proporções ideais observados na literatura médica (PRENDERGAST, 2012). (a) Exemplos de proporções faciais verticais. (b) Exemplos de proporções faciais horizontais. . . . .	6
Figura 5	– Padrões de variação da forma da mão direita modelados pelo <i>Active Shape Models</i> (COOTES et al., 1995). . . . .	10
Figura 6	– Localização de pontos faciais (DU et al., 2008). (a) Resultado do ASM tradicional. (b) Resultado do SVMBASM. . . . .	12
Figura 7	– Localização de estruturas cerebrais subcorticais (CERROLAZA; VILLANUEVA; CABEZA, 2012). (a) Marcação manual. (b) Resultado do ASM tradicional. (c) Resultado do AHSM. . . . .	13
Figura 8	– Exemplo de aplicação do AAM. Mais à esquerda, o modelo inicial é aplicado em imagem de baixa resolução. Ao longo das iterações (esquerda para direita), a instância resultante das aplicações de resolução mais baixa serve de entrada para versões de maior resolução da imagem original. As duas últimas colunas mostram a comparação da imagem original (mais à direita) com a iteração final do método (à esquerda) (COOTES; EDWARDS; TAYLOR, 2001). . . . .	14
Figura 9	– Conjunto de pontos faciais explorados pela estratégia <i>Landmark Model Matching</i> . É possível observar que a localização de alguns pontos é auto-explicativa. Entretanto muitos outros não são definidos com precisão, um exemplo disso é o ponto 44 (SENARATNE; HALGAMUGE, 2006b). . . . .	15
Figura 10	– Os 29 pontos apresentados por Shi, Samal e Marx (2006). (a) A localização dos pontos faciais. (b) Distribuição dos pontos em um conjunto de imagens. . . . .	16
Figura 11	– Algumas etapas de processamento exploradas por Sohail e Bhattacharya (2008). (a) Fluxo para localizar pontos do olho. (b) Fluxo para localizar fossas nasais. . . . .	17
Figura 12	– Conjunto de pontos faciais explorados por Bottino e Cumani (2008). . . . .	18

Figura 13 – Imagem com faces em diversas posições obtida em ambiente não-controlado. O algoritmo detecta a face (caixa maior ao redor da face), localiza os pontos (caixas menores dentro da face) e estima a pose (rótulo de inclinação acima da face) (ZHU; RAMANAN, 2012). . . . .	20
Figura 14 – Imagens do nariz em duas visões. (a) Visão do perfil para cálculo da razão de Baum. (b) Visão frontal. . . . .	23
Figura 15 – Pontos faciais frontais. (a) Conjunto de pontos definidos por Prendergast (2012). (b) Conjunto de pontos definidos por Zimble e Ham (2010). . . . .	23
Figura 16 – Localização dos pontos frontais definidos na Tabela 2. . . . .	24
Figura 17 – Pontos do perfil facial. (a) Conjunto de pontos definidos por Prendergast (2012). (b) Conjunto de pontos definidos por Zimble e Ham (2010). . . . .	25
Figura 18 – Localização dos pontos do perfil facial definidos na Tabela 3. . . . .	26
Figura 19 – Representação de uma imagem digital com resolução $6 \times 6$ em níveis de cinza. . . . .	28
Figura 20 – Formas de representação da imagem da Figura 19. (a) Representação numérica em 8 bits. (b) Representação em matriz com os valores normalizados. . . . .	28
Figura 21 – Aplicação da filtragem mediana em imagem ruidosa. (a) Imagem capturada em condições adversas de iluminação. (b) Imagem após sucessivas aplicações da mediana com janela $3 \times 3$ . . . . .	29
Figura 22 – Convolução de uma imagem com <i>kernel</i> da média de tamanho $5 \times 5$ . (a) Imagem original. (b) Imagem com valores complementares. (c) Imagem final. . . . .	31
Figura 23 – Aplicação da filtragem Gaussiana. (a) Imagem ruidosa. (b) $\sigma = 2, 0$ . (c) $\sigma = 5, 0$ . (d) $\sigma = 15, 0$ . . . . .	32
Figura 24 – Aplicação do filtro Sobel para detecção de bordas. (a) Imagem ruidosa. (b) Magnitude do gradiente da imagem ruidosa. . . . .	33
Figura 25 – Aplicação do método de Canny para a detecção de bordas. (a) Imagem ruidosa. (b) Suavização. (c) Extração do gradiente. (d) Supressão dos não-máximos. (e) Dupla limiarização. (f) Rastreamento de bordas. . . . .	35
Figura 26 – Resultado da aplicação da análise de componentes principais em conjunto de dados com duas variáveis. (a) Distribuição de vetores bidimensionais. (b) Eixos ortogonais de maior variância com média $\vec{\mu}$ . (c) Amostra $\vec{x}$ reconstruída com $\vec{b}$ . . . . .	38
Figura 27 – Rasterização de figura geométrica. (a) Triângulo formado por 3 segmentos de reta. (b) Resultado do algoritmo de Bresenham (1965) para rasterização do triângulo. . . . .	40
Figura 28 – Exemplo de coleta do perfil de textura do ponto $(x_2, y_2)$ em forma composta por 6 pontos. . . . .	43

Figura 29 – Exemplo de escolha de pontos para definir a forma da boca. Os pontos 2, 3, 4, 7, 9, 11, 13 e 14 são os pontos principais que definem a forma da boca. Os outros são pontos usados para preencher a forma e estão posicionados de modo equidistante aos pontos adjacentes no contorno.	46
Figura 30 – Exemplo de extração do perfil de textura. (a) Segmento determinado pelos pontos da forma. (b) Perfil de textura coletado. (c) Perfil de textura com diferenças de intensidade normalizadas.	47
Figura 31 – Alinhamento de um conjunto de bocas marcadas com os 16 pontos ilustrados na Figura 29. (a) Distribuição inicial dos pontos. (b) Distribuição após algumas iterações do algoritmo. (c) Distribuição final (PDM).	50
Figura 32 – Modos de variação do conjunto de bocas marcadas da Figura 31. (a) Média das formas variando ao longo de 3 componentes. (b) Modos de variação sobrepostos.	52
Figura 33 – Coleta de textura de um ponto do olho. (a) Reta perpendicular à forma passando pelo ponto $(x_2, y_2)$ . (b) Limites da amostra de textura.	54
Figura 34 – Perfis de textura ao redor de pontos da forma. (a) Média dos perfis para o ponto $(x_2, y_2)$ da Figura 33. (b) Faixas de valores que devem ser comparadas com a média dos perfis.	55
Figura 35 – Demonstração de iteração do Algoritmo 3. (a) Localização da forma antes da execução do algoritmo. (b) Forma $\vec{f}_s$ gerada pelo Algoritmo 4. (c) Forma $\vec{f}_c$ gerada pelo Algoritmo 5. (d) Forma $\vec{f}_c$ alinhada a $\vec{f}_s$ .	57
Figura 36 – Metodologia proposta.	59
Figura 37 – Esquema de funcionamento do LPP.	60
Figura 38 – Aplicações sucessivas da filtragem mediana. (a) Imagem ruidosa. (b) Filtragem mediana ( $\times 3$ ). (c) ( $\times 6$ ). (d) ( $\times 9$ ).	61
Figura 39 – Resultado esperado da etapa de detecção de bordas. (a) Imagem de entrada. (b) Imagem com bordas detectadas.	61
Figura 40 – Busca pelo perfil. (a) Localização do ponto com $y_{max}$ . (b) Definição do contorno.	62
Figura 41 – Vetorização do contorno. (a) Localização da ponta do nariz. (b) Vetorização da parte frontal da face em perfil.	63
Figura 42 – Adição de vetores. (a) Vetores $\vec{p_i p_j}$ e $\vec{p_j p_k}$ originais. (b) Vetor $\vec{p_i p_k}$ resultante da soma.	64
Figura 43 – Protuberâncias e concavidades do perfil facial. Os círculos em vermelho indicam as protuberâncias. Os azuis representam as concavidades.	67
Figura 44 – Esquema de funcionamento do LPF ASM. O fluxo $I$ do diagrama é percorrido pelo conjunto de treinamento. A imagem de teste percorre o fluxo $II$ .	68

Figura 45 – Conjunto de pontos explorados no trabalho. (a) 21 pontos frontais alvos do trabalho. (b) 145 pontos usados para melhorar os resultados de busca.	70
Figura 46 – <i>Software</i> desenvolvido para auxiliar a marcação manual das imagens.	71
Figura 47 – Aplicação da filtragem gaussiana com $\sigma = 10, 0$ . (a) Imagem de entrada. (b) Imagem filtrada.	71
Figura 48 – Pirâmide de imagens usada na execução do ASM. (a) Imagem original. (b) Imagem redimensionada a $\frac{1}{4}$ . (c) Imagem redimensionada a $\frac{1}{16}$ .	72
Figura 49 – Comparação dos perfis extraídos de duas formas diferentes. (a) Perfil com gradiente em uma direção. (b) Perfil com gradiente em várias direções.	73
Figura 50 – Demonstração gráfica do erro médio na localização dos pontos do perfil.	79
Figura 51 – Região abaixo do nariz. (a) Contorno original. (b) Contorno com limiar $\gamma = 0$ . (c) Contorno com $\gamma > 0$ . (d) Contorno com $\gamma \gg 0$ .	80
Figura 52 – Relação entre o limiar e o erro médio. Dados coletados de imagens com $480 \times 640$ <i>pixels</i> .	81
Figura 53 – Resultados do LPF ASM com diferentes valores para o desvio padrão da filtragem Gaussiana. (a) Erro médio do LPF ASM com diferentes valores para $\sigma$ . (b) Comparação do ASM com e sem filtragem ao longo de todo o conjunto de teste.	84
Figura 54 – Comparação entre o ASM clássico e o ASM implementado com Sobel.	84
Figura 55 – Resultados do LPF ASM com diferentes valores de tamanho para o perfil de textura. (a) Erro médio do LPF ASM com diferentes valores para $k$ . (b) Comparação do LPF ASM com 4 valores para $k$ ao longo de todo o conjunto de teste.	86
Figura 56 – Resultados do LPF ASM com diferentes valores para o fator de multiplicação das restrições. (a) Erro médio do LPF ASM com diferentes valores para as restrições. (b) Comparação do LPF ASM com 3 valores para as restrições ao longo de todo o conjunto de teste.	87
Figura 57 – Comparativo das variantes do ASM e LPF ASM ao longo de todo o conjunto de teste.	87
Figura 58 – Demonstração gráfica do erro médio na localização dos pontos frontais. (a) Resultado para o ASM. (b) Resultado para o LPF ASM.	91

# Lista de tabelas

Tabela 1 – Pontos de referência segundo Bottino e Cumani (2008). . . . .	18
Tabela 2 – Descrição dos pontos de referência frontais. . . . .	24
Tabela 3 – Descrição dos pontos de referência do perfil facial. . . . .	25
Tabela 4 – Sumário dos métodos ASM clássico e LPF ASM. . . . .	76
Tabela 5 – Medidas em imagens de perfil de resolução $2592 \times 3456$ <i>pixels</i> . . . . .	78
Tabela 6 – Medidas em imagens frontais de resolução $1200 \times 1600$ <i>pixels</i> . . . . .	82
Tabela 7 – Comparativo, ponto a ponto, entre o ASM clássico e o LPF ASM. . . . .	89
Tabela 8 – Comparativo, ponto a ponto, entre o ASM com as modificações externas propostas e o LPF ASM. . . . .	90

# Lista de abreviaturas e siglas

<b>AAM</b>	<i>Active Appearance Models</i>
<b>AHSM</b>	<i>Active Hierarchical Shape Models</i>
<b>ANd</b>	Asa Nasal Direita
<b>ANe</b>	Asa Nasal Esquerda
<b>ASd</b>	Arco da Sobrancelha Direita
<b>ASe</b>	Arco da Sobrancelha Esquerda
<b>ASM</b>	<i>Active Shape Models</i>
<b>CLd</b>	Canto Lateral Direito
<b>CLe</b>	Canto Lateral Esquerdo
<b>CMd</b>	Canto Medial Direito
<b>CMe</b>	Canto Medial Esquerdo
<b>Col</b>	Columela
<b>Crv</b>	Cervical
<b>Gl</b>	Glabela
<b>LI</b>	Lábio Inferior
<b>LLd</b>	Limbo Lateral Direito
<b>LLe</b>	Limbo Lateral Esquerdo
<b>LPF</b>	Localizador de Pontos Frontais
<b>LPP</b>	Localizador de Pontos do Perfil
<b>LS</b>	Lábio Superior
<b>Mt</b>	Mentoniano
<b>Ns</b>	Násio
<b>Od</b>	Orbitário Direito

<b>Oe</b>	Orbitário Esquerdo
<b>PCA</b>	Análise de Componentes Principais
<b>px</b>	<i>pixels</i>
<b>PDI</b>	Processamento Digital de Imagens
<b>PDM</b>	<i>Point Distribution Model</i>
<b>Pg</b>	Pogônio
<b>PN</b>	Ponta do Nariz
<b>Rn</b>	Rínio
<b>Sm</b>	Supramental
<b>Sn</b>	Subnasal
<b>St</b>	Stomion
<b>SVMBASM</b>	<i>Support Vector Machine Based Active Shape Models</i>
<b>Tr</b>	Tríquio



# Lista de símbolos

$\alpha$	Ângulo de inclinação
$\theta$	Ângulo de rotação
$\approx$	Aproximadamente igual
$\leftarrow$	Atribuição
$\lambda$	Autovalor
$\otimes$	Convolução
$\sigma$	Desvio padrão
$\gamma$	Limiar
$\mu$	Média
$\gg$	Muito maior que
$\forall$	Para todo
$\in$	Pertence
$\Sigma$	Somatório
$\Delta$	Variação
$\emptyset$	Vazio

# Lista de publicações

**FREITAS, R. T.**, AIRES, K. T., CAMPELO, V. S. Automatic Location of Facial Landmarks for Plastic Surgery Procedures. *IEEE International Conference on Systems, Man and Cybernetics (SMC '14)*, 2014, San Diego.

**FREITAS, R. T.**, AIRES, K. T., CAMPELO, V. S. Locating Facial Landmarks Towards Plastic Surgery. *28th SIBGRAPI Conference on Graphics, Patterns and Images*, 2015, Salvador.

CAMPELO, V. S., **FREITAS, R. T.**, LEITE FILHO, O. A., MELO, M. B., AIRES, K. T. Towards an Automatic Facial Landmark Locator Tool for Plastic Surgery (Poster). *8th World Congress of Facial Plastic Surgery*, 2016, Rio de Janeiro.

CAMPELO, V. S., **FREITAS, R. T.**, PINTO, L. G., LEITE FILHO, O. A., PIAUILINO NETO, H. S., AIRES, K. T. Mobile Support Application for Facial Plastic Surgery Planning (2º lugar na categoria Poster). *8th World Congress of Facial Plastic Surgery*, 2016, Rio de Janeiro.

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>1</b>
<b>1.1</b>	<b>Contextualização Histórica</b>	<b>1</b>
1.1.1	Estética: Objetividade x Subjetividade	2
1.1.2	Evolução da Objetividade Estética	2
1.1.3	A Estética na Medicina Moderna	5
1.1.4	Motivação	6
<b>1.2</b>	<b>Objetivos</b>	<b>7</b>
1.2.1	Objetivo Geral	7
1.2.2	Objetivos Específicos	7
<b>1.3</b>	<b>Justificativa</b>	<b>7</b>
1.3.1	Contribuições deste Trabalho	8
<b>1.4</b>	<b>Organização do Trabalho</b>	<b>8</b>
<b>2</b>	<b>TRABALHOS RELACIONADOS</b>	<b>9</b>
<b>2.1</b>	<b><i>Eigenfaces</i> para Reconhecimento Facial</b>	<b>9</b>
<b>2.2</b>	<b><i>Active Shape Models</i> e Variações</b>	<b>10</b>
<b>2.3</b>	<b><i>Active Appearance Models (AAM)</i></b>	<b>13</b>
<b>2.4</b>	<b>Pontos Faciais para Reconhecimento</b>	<b>14</b>
<b>2.5</b>	<b>Detecção de Pontos Faciais Baseados em Modelos Antropométricos</b>	<b>16</b>
<b>2.6</b>	<b>Identificação de Pontos Faciais em Perspectivas Diversas</b>	<b>17</b>
<b>2.7</b>	<b>Considerações Finais</b>	<b>21</b>
<b>3</b>	<b>MATERIAIS E MÉTODOS</b>	<b>22</b>
<b>3.1</b>	<b>Elementos da Análise Estética</b>	<b>22</b>
3.1.1	Pontos Faciais Frontais	23
3.1.2	Pontos Faciais de Perfil	25
<b>3.2</b>	<b>Processamento Digital de Imagens</b>	<b>26</b>
3.2.1	A Imagem Digital	27
3.2.2	Filtragem Mediana	29
3.2.3	Convolução	30
3.2.3.1	Filtragem Gaussiana	31
3.2.3.2	Filtro Sobel	32
3.2.4	Detecção de Bordas com Canny	33
<b>3.3</b>	<b>Análises Estatísticas</b>	<b>36</b>
3.3.1	Notação de Variáveis	36
3.3.2	Análise de Componentes Principais (PCA)	37

3.3.2.1	Autovalor e Autovetor . . . . .	37
3.3.2.2	Computando as Componentes Principais . . . . .	37
3.3.3	Aferição de Resultados . . . . .	39
<b>3.4</b>	<b>Rasterização para Coleta de Textura . . . . .</b>	<b>39</b>
<b>3.5</b>	<b>Considerações Finais . . . . .</b>	<b>41</b>
<b>4</b>	<b>ACTIVE SHAPE MODELS . . . . .</b>	<b>42</b>
<b>4.1</b>	<b>Definições Preliminares . . . . .</b>	<b>42</b>
4.1.1	Forma . . . . .	42
4.1.2	Perfil de Textura . . . . .	43
4.1.3	Resumo do ASM . . . . .	44
<b>4.2</b>	<b>Construção do Modelo . . . . .</b>	<b>44</b>
4.2.1	Marcação de Imagens . . . . .	45
4.2.2	Extração do Perfil dos Pontos . . . . .	46
4.2.3	Cálculo dos Pesos . . . . .	48
4.2.4	Alinhamento das Formas . . . . .	48
4.2.5	Captura dos Modos de Variação . . . . .	49
<b>4.3</b>	<b>Busca pela Forma . . . . .</b>	<b>52</b>
4.3.1	Buscar Nova Posição . . . . .	53
4.3.2	Ajustar Forma . . . . .	55
<b>4.4</b>	<b>Considerações Finais . . . . .</b>	<b>56</b>
<b>5</b>	<b>METODOLOGIA PROPOSTA . . . . .</b>	<b>58</b>
<b>5.1</b>	<b>Localizador de Pontos do Perfil (LPP) . . . . .</b>	<b>58</b>
5.1.1	Processamento da Imagem . . . . .	60
5.1.2	Vetorizar Contorno . . . . .	61
5.1.3	Reduzir Vetores . . . . .	63
5.1.4	Agrupar Vetores . . . . .	64
5.1.5	Localizar Extremos . . . . .	65
<b>5.2</b>	<b>Localizador de Pontos Frontais (LPF ASM) . . . . .</b>	<b>67</b>
5.2.1	Formas e Conjunto de Pontos . . . . .	68
5.2.2	Reduzir Ruído . . . . .	70
5.2.3	Múltiplas Dimensões . . . . .	72
5.2.4	Extrair Gradiente . . . . .	72
5.2.5	Construção do Modelo e Busca dos Pontos - ASM . . . . .	73
<b>5.3</b>	<b>Considerações Finais . . . . .</b>	<b>76</b>
<b>6</b>	<b>RESULTADOS E DISCUSSÃO . . . . .</b>	<b>77</b>
<b>6.1</b>	<b>Resultados do Localizador de Pontos do Perfil . . . . .</b>	<b>77</b>
6.1.1	Discussão das Medidas . . . . .	78

6.1.2	Discussão do Limiar . . . . .	79
6.1.3	Problemas com o Contorno . . . . .	81
<b>6.2</b>	<b>Resultados do Localizador de Pontos Frontais . . . . .</b>	<b>82</b>
6.2.1	Resultados de Modificações Externas ao ASM . . . . .	83
6.2.2	Resultados de Parâmetros Internos ao ASM . . . . .	85
6.2.3	LPF ASM Ideal . . . . .	86
<b>6.3</b>	<b>Considerações Finais . . . . .</b>	<b>91</b>
	 <b>REFERÊNCIAS . . . . .</b>	 <b>96</b>
	 <b>ANEXOS . . . . .</b>	 <b>100</b>
	 <b>ANEXO A – DEFINIÇÕES ESTATÍSTICAS . . . . .</b>	 <b>101</b>
A.1	Médias Aritmética e Ponderada . . . . .	101
A.2	Variância . . . . .	101
A.3	Covariância . . . . .	101
A.4	Distância de <i>Mahalanobis</i> . . . . .	102

# 1 Introdução

Os grandes avanços nas ciências da saúde, acelerados a partir da primeira metade do século XX, proporcionaram tratamento e cura para uma imensa variedade de disfunções e doenças. Não por acaso, a expectativa de vida das populações que têm acesso a tais avanços aumentou drasticamente, assim como a qualidade de vida de seus indivíduos.

Dirigindo-se nesse caminho, a medicina passou a possibilitar uma outra perspectiva para a saúde. Foram desenvolvidas técnicas cirúrgicas que garantem harmonia entre o organismo do paciente e a equipagem necessária para a realização de procedimentos invasivos. Com adventos dessa natureza, as ciências da saúde rumaram não somente para avanços que objetivam a cura de enfermidades como também para a possibilidade de se corrigir imperfeições de ordem estética.

Nas últimas décadas, esses avanços se tornaram mais evidentes à medida que incluíram uma parcela cada vez maior da população ([WIMALAWANSA; MCKNIGHT; BULLOCKS, 2009](#)). Procedimentos estéticos que antes se restringiam aos extratos financeiramente mais privilegiados das sociedades, passaram a se encontrar bem mais acessíveis às camadas populares. Além da redução nos custos das operações, o aumento na procura por procedimentos dessa natureza também foi motivado por avanços tecnológicos e por uma aceitação crescente do público alvo.

Decorrente desse aumento, os cirurgiões plásticos, profissionais especialistas da área, se encontram em uma posição de destaque, tendo conseqüentemente que realizar ajustes na maneira como abordam seus trabalhos. Já não é suficiente organizar e planejar as cirurgias nos moldes do passado. Os métodos originais foram superados a partir do momento em que a tecnologia passou a possibilitar novas maneiras de se resolver o problema.

Nosso trabalho busca contribuir para a forma como os cirurgiões plásticos devem trabalhar os planos operatórios das cirurgias, proporcionando uma maneira rápida e eficaz de localização e identificação de pontos de referência facial largamente utilizados como origem para a definição de procedimentos cirúrgicos.

## 1.1 Contextualização Histórica

A qualidade da beleza era tradicionalmente considerada uma virtude assim como a benevolência, a paciência e a justiça. Na Grécia Antiga, o adjetivo grego *kalon* remetia ao significado de belo, embora nem sempre representasse a mesma coisa. Para Platão e Aristóteles, *kalon* tinha um importante papel como aprovação ética. Em algumas ocasiões, o termo era sinônimo de nobreza e admiração. Contudo, dizer que a beleza é uma virtude,

em tempos atuais, seria uma colocação incorreta e até mesmo perversa (PAPPAS, 2015).

Em uma análise mais restrita do termo, é possível notar que o ideal de beleza humano é comumente encarado como um aspecto subjetivo e dependente da predileção daqueles que se empenham em qualificar a aparência do próximo. Tal subjetividade sofre influência direta do contexto no qual o indivíduo está inserido, sendo afetada pelos padrões estéticos da sua sociedade e pelo gosto particular de cada um. É raro existir unanimidade em relação à beleza de algum indivíduo.

A discussão acerca da estética de uma maneira geral não é recente. Há muito se discute esse tema, principalmente sob o ponto de vista filosófico. Mas ao longo dos séculos, essa discussão invadiu outras áreas do conhecimento e ganhou contornos mais objetivos.

### 1.1.1 Estética: Objetividade x Subjetividade

No campo filosófico, a discussão sobre o caráter da estética ser uma qualidade subjetiva ou objetiva sempre foi uma das maiores fontes de discórdia na literatura. Alguns pensadores afirmavam que a qualidade da beleza era dependente dos observadores, argumentando que onde uns percebiam deformidades, outros enxergavam beleza. Contudo, não haveria propósito em considerar a beleza como sendo completamente subjetiva, uma vez que o próprio sentido da palavra seria perdido (SARTWELL, 2014).

Apesar desse viés subjetivo, quantificar e qualificar a beleza humana em termos objetivos é um desafio tão antigo que suas origens podem ser comparadas ao próprio surgimento de ciências como a filosofia e a matemática. Um dos mais importantes filósofos do período clássico da Grécia Antiga, Platão começou a esboçar uma abordagem matemática para a beleza humana. Ainda que considerasse tal característica como sendo resultado de bom gosto e equilíbrio, fatores que remetem a um certo nível de subjetividade, Platão percebeu que as proporções físicas também desempenhavam papel importante para se concluir que um indivíduo é belo (ZIMBLER; HAM, 2010).

Um exemplo de que o conceito clássico de beleza está relacionado objetivamente com proporções e relações ideais é o tratado *Cânone* elaborado por Policleto. Neste trabalho, o escultor definiu proporções sobre o corpo humano e corroborou sua obra com a escultura *Doríforo*<sup>1</sup> (STEINER, 2001). É possível observar, por exemplo, que *Doríforo* possui altura equivalente a sete vezes e meia o tamanho de sua cabeça, proporção que foi revista durante a Renascença. A Figura 1 ilustra uma representação da escultura.

### 1.1.2 Evolução da Objetividade Estética

Aos poucos, o homem foi incorporando nas análises estéticas um rigor outrora pouco explorado. Na Renascença, os ideais de beleza foram revisitados pelos artistas e

<sup>1</sup> Frequentemente chamada de *Cânone* em homenagem ao tratado.



**Figura 1** – *Doríforo* de Policleto, resultado de um tratado sobre proporções ideais (POLICLETO, 1797 (Excavada)).

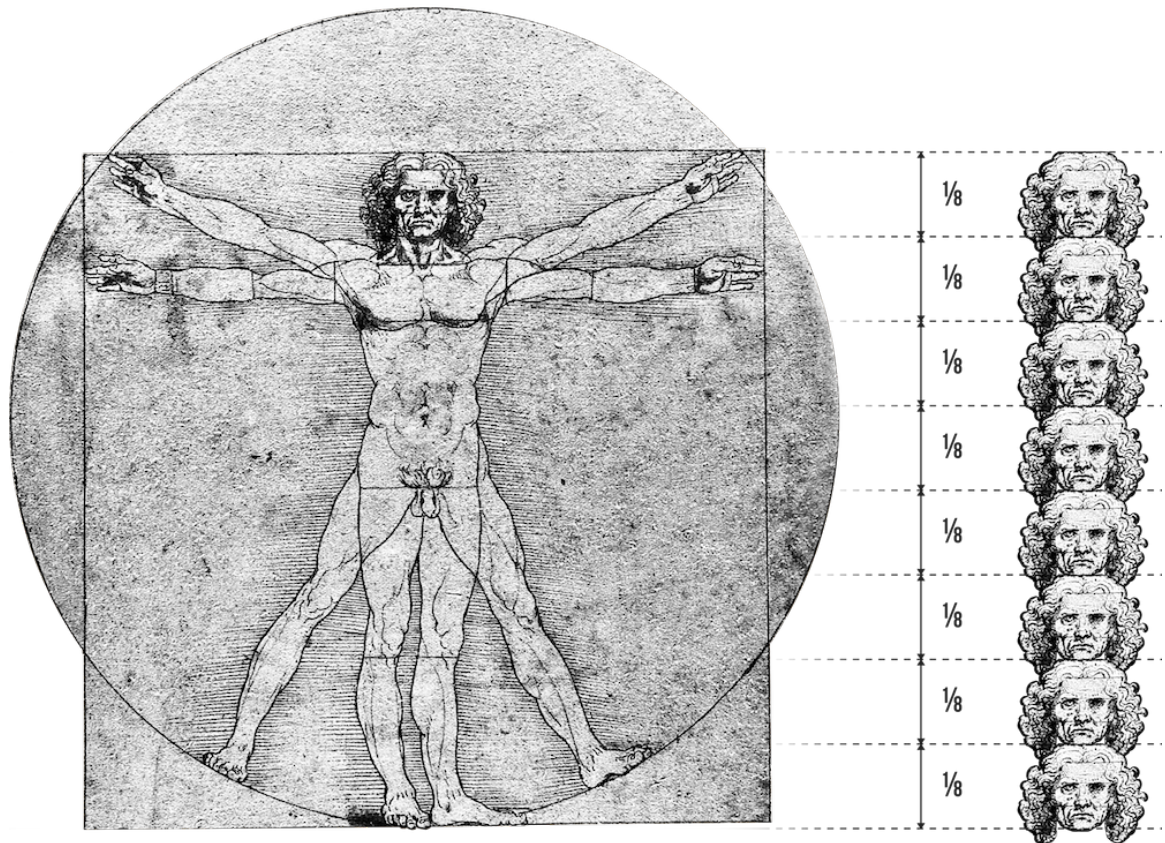
cientistas daquele tempo. Um dos mais aparentes exemplos disso foi *O Homem Vitruviano* de Leonardo Da Vinci (ZIMBLER; HAM, 2010). Observa-se que Leonardo desenhou um corpo cuja relação entre a altura ideal e o tamanho da cabeça seria igual a oito, em ligeira discordância com *Doríforo* de Policleto. A relação entre corpo e cabeça pode ser observada na Figura 2.

*O Homem Vitruviano* foi resultado de um estudo feito por Da Vinci sobre uma série de livros intitulada *De Architectura* do arquiteto romano Vitrúvio. Leonardo conseguiu desenhar aquilo que Vitrúvio havia preconizado sobre o corpo humano em sua obra. Uma demonstração de que havia esforço de outras áreas do conhecimento na busca pelo caráter objetivo da estética.

Albrecht Durer, pintor alemão da Renascença, foi outro artista inspirado a propor medidas e relações do corpo humano que denotassem perfeição estética. Assim como Policleto, Durer elaborou uma espécie de tratado<sup>2</sup> (dividido em quatro livros) sobre

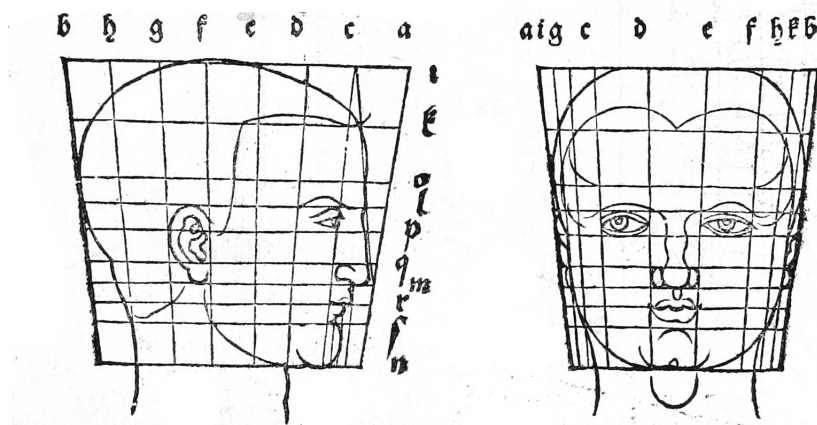
<sup>2</sup> *Vier Bücher von Menschlicher Proportion* - Quatro Livros sobre Proporções Humanas (tradução livre).





**Figura 2** – À esquerda, um recorte de *O Homem Vitruviano* de Leonardo Da Vinci. À direita, uma representação do tamanho da cabeça em relação ao corpo.

proporções humanas (ZIMBLER; HAM, 2010). A Figura 3 mostra um rascunho de Durer sobre seu estudo das proporções faciais.



**Figura 3** – Exemplo de estudo de proporções faciais realizado por Albrecht Durer.

Apesar de a preocupação com os padrões estéticos datar da Grécia Antiga e se estender pela Renascença, não houve grandes evoluções no rigor científico com o qual esse aspecto foi tratado. Os grandes avanços com relação à análise estética começaram, de fato,

na primeira metade do século XX. Tais avanços foram impulsionados pelo desenvolvimento da medicina, especialmente no campo das cirurgias plásticas. A face humana passou a ser vista como um objeto de estudo científico fundamental não somente para a realização de cirurgias corretivas, mas também, para o levantamento de dados necessários para a formulação de uma análise pré-operatória e a construção do modelo estético.

Esses avanços tomaram forma no momento em que a estética havia perdido força como objeto de investigação filosófica (SARTWELL, 2014). O homem pôde migrar de uma discussão teórica e conceitual para uma possibilidade de mudança prática, pois naquele momento, passou a ser viável atingir os ideais tão discutidos ao longo dos séculos.

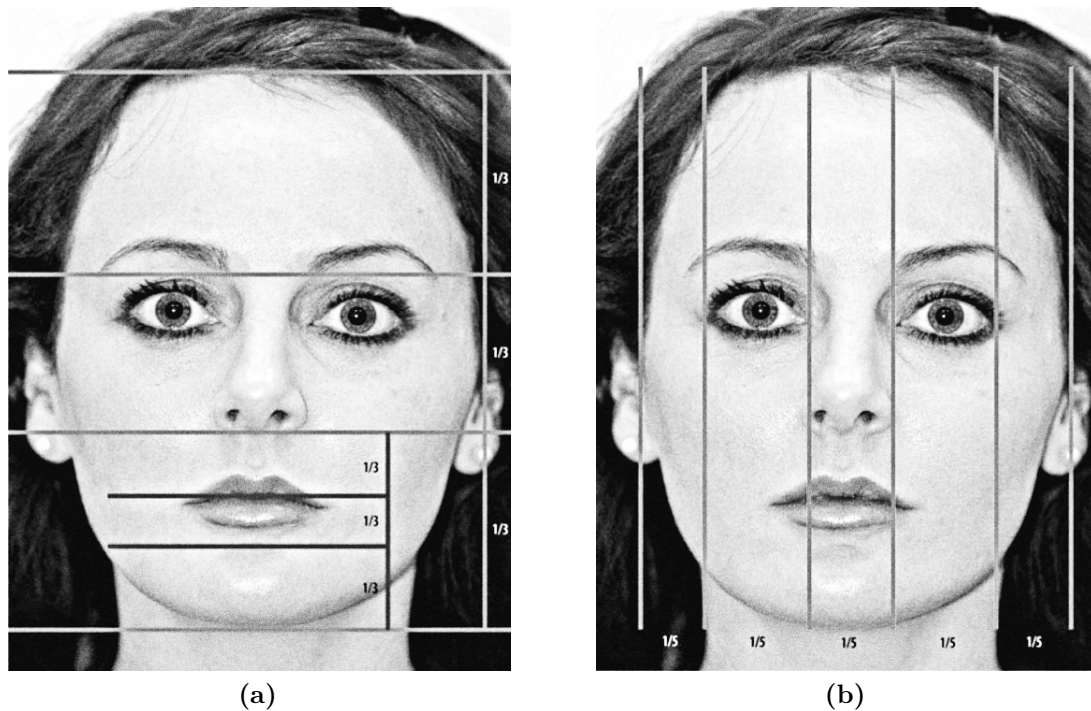
### 1.1.3 A Estética na Medicina Moderna

Já no século XX, o antropometrista Leslie Farkas realizou um estudo com duzentas mulheres para testar os tratados clássicos sobre proporções humanas, como os definidos por Policleto e Durer. Ele concluiu que algumas das medidas e relações eram apenas idealizações artísticas (ZIMBLER; HAM, 2010). Os avanços modernos deixaram de lado a motivação dos artistas e evidenciaram que cientificamente ainda havia bastante a se evoluir na definição da estética humana.

Desse modo, à medida que a estética passou a ser mais explorada nas ciências da saúde, os avanços se direcionaram para as implicações práticas que passaram a fazer parte do cotidiano médico. Com a possibilidade de se realizar procedimentos cirúrgicos de maneira cada vez mais segura, o estudo da estética avançou drasticamente durante o século XX e encontrou seu ápice nas cirurgias plásticas.

Para incorporar tal avanço em seus procedimentos cirúrgicos, os médicos precisaram formalizar uma vasta quantidade de informação acerca do rosto de seus pacientes. Uma das mais relevantes para qualquer análise facial é a definição dos pontos de referência. Esses pontos representam as partes mais importantes da face humana, que são responsáveis pelas principais características faciais (ponta do nariz, cantos da boca, cantos dos olhos, etc). Alterações cirúrgicas feitas nesses pontos indicam um maior impacto estético no paciente, por isso eles guardam informações vitais para tais análises.

Formalizadas as definições de cada ponto, os médicos precisaram determinar o significado que eles teriam em uma análise facial. Então eles conceberam métricas baseadas em estudos de diversos modelos estéticos, variados de acordo com a etnia do indivíduo, para, junto com a informação sobre os pontos, indicar que características se alinhavam com os padrões estéticos definidos. Um exemplo de métrica pode ser observado na Figura 4.



**Figura 4** – Exemplos de proporções ideais observados na literatura médica (PRENDERGAST, 2012). (a) Exemplos de proporções faciais verticais. (b) Exemplos de proporções faciais horizontais.

#### 1.1.4 Motivação

Como foi mencionado, os cirurgiões necessitam de uma série de informações sobre a face de um paciente no intuito de realizar uma análise pré-operatória para uma cirurgia plástica. A coleta dessas informações deve passar inicialmente pela localização e identificação de um conjunto de pontos faciais de referência. Considerando a literatura médica, a quantidade de pontos que devem ser buscados para uma análise sob essa perspectiva dificulta a coleta manual. Com mais de 30 pontos a serem localizados e identificados, os cirurgiões frequentemente abrem mão desta etapa da análise, o que representa um risco para o sucesso de um futuro procedimento e, sobretudo, para a saúde do paciente.

A ideia deste trabalho é reduzir a complexidade e esforço de que o médico deve dispor para completar essa etapa da análise. Isso deve ser feito através de um método automático de localização e identificação dos pontos necessários para se cumprir essa tarefa. A metodologia reduz o trabalho de análise do cirurgião, deixando para ele somente a preocupação de avaliar a posição dos pontos e tomar a decisão de que procedimento ele deve adotar, caso haja necessidade.

## 1.2 Objetivos

### 1.2.1 Objetivo Geral

O objetivo deste trabalho é desenvolver metodologias que permitam localizar e identificar pontos de referência faciais em imagens. Eles servirão de entrada para a realização dos cálculos de métricas necessários para definir procedimentos cirúrgicos adequados ao paciente. A estratégia aplicada dividiu o problema em duas partes de acordo com as necessidades da análise facial pré-operatória: busca de pontos em imagens faciais frontais e de perfil.

### 1.2.2 Objetivos Específicos

Os seguintes objetivos específicos foram propostos:

- estudo e experimentação de técnicas de Processamento de Imagens;
- desenvolvimento de um método para localizar e identificar pontos faciais em imagens de perfil;
- desenvolvimento de um método para localizar e identificar pontos faciais em imagens frontais;
- coleta e análise de resultados das metodologias propostas.

## 1.3 Justificativa

A localização e identificação de características faciais é um problema bastante recorrente na literatura e que avançou consideravelmente a partir da década de 1990. [Turk e Pentland \(1991\)](#) propuseram uma metodologia para reconhecer faces baseada nas componentes principais de um conjunto de imagens. [Cootes et al. \(1995\)](#) prestaram grande contribuição para a solução do problema quando criaram *Active Shape Models* (ASM). Um modelo estatístico que leva em consideração a forma dos elementos e pode ser explorado para localizar atributos faciais.

Outro trabalho do gênero, foi o de [Senaratne e Halgamuge \(2006b\)](#). Seu método chamado *Landmark Model Matching* visa reconhecimento facial. Para tanto, eles exploram a localização de pontos faciais.

O trabalho de [Milborrow \(2007\)](#) sugeriu modificações no ASM clássico para melhorar a busca por pontos faciais. Já [Bottino e Cumani \(2008\)](#) idealizaram um método para localização de pontos de referência em imagens de perfil, problema que nosso trabalho também ataca.

Apesar de os avanços serem inegáveis, a precisão na localização das principais características, em especial dos pontos de referência, ainda precisa de melhorias para ser aplicada ao problema proposto. Muitos trabalhos relacionados apresentam deficiências que podem ser melhoradas para resolver o problema da localização de pontos para auxílio em cirurgias plásticas. Uma análise detalhada dos trabalhos mais importantes para o problema se encontra no Capítulo 2.

### 1.3.1 Contribuições deste Trabalho

A principal contribuição deste trabalho é propor um método computacional capaz de localizar e identificar pontos de referência faciais tipicamente abordados em cirurgias plásticas. O método facilita a etapa de análise pré-operatória e remove parte da complexidade inerente a essa fase, agilizando e simplificando o trabalho do cirurgião.

Parte da proposta tem como base métodos já consagrados, como ASM. Porém, ela traz algumas inovações no sentido de melhorar metodologias existentes, resumidamente listadas abaixo:

- aprimoramentos no algoritmo ASM para localizar pontos em imagens faciais frontais;
- método de localização e identificação de pontos do perfil facial.

As inovações estão descritas em detalhes no Capítulo 5.

## 1.4 Organização do Trabalho

O Capítulo 2 discute detalhadamente os trabalhos que já foram propostos para resolver problemas semelhantes. O Capítulo 3 faz uma explanação sobre cada componente que é aproveitado para o desenvolvimento da metodologia proposta. O Capítulo 4 é dedicado a detalhar os principais pontos do método ASM, que é responsável por grande parte da nossa proposta. O Capítulo 5 descreve toda a composição da metodologia, como as ferramentas se relacionam para resolver o problema em questão, e que alterações precisaram ser realizadas nos métodos apresentados. O Capítulo 6 expõe os resultados dos testes realizados, justifica a escolha de alguns parâmetros adotados, e levanta discussões sobre as dificuldades encontradas. Finalmente, o último capítulo apresenta as conclusões sobre a proposta e perspectivas futuras para a solução do problema abordado.

## 2 Trabalhos Relacionados

O foco deste trabalho é a localização e identificação de um conjunto de pontos faciais de referência. Por ser um problema que atrai bastante interesse, a literatura está repleta de propostas que têm o mesmo objetivo, porém, com motivações distintas. Algumas fazem uso da localização desses pontos com o propósito de identificar um indivíduo, outras visam explorá-la para reconhecer expressões faciais, e existem, ainda, aquelas que objetivam auxiliar procedimentos médicos. De fato, muitos trabalhos abordam características faciais com as mais diversas finalidades e portanto foram alvos de nosso estudo.

Além disso, foram abordados trabalhos com objetivos diferentes dos elencados para esta dissertação. Outras propostas visam localizar e identificar padrões em imagens que se distinguem daqueles que foram explorados por nosso trabalho. Contudo, tais padrões, apesar de não serem especificamente características faciais, podem ser aproveitados de uma forma genérica para resolver o problema apresentado.

Este capítulo ressalta a importância do tema proposto destacando trabalhos que visam atingir objetivos semelhantes, aponta metodologias que podem ser exploradas para a solução do problema, e discute aspectos – positivos e negativos – dos trabalhos estudados que contribuem de alguma forma para nossa proposta.

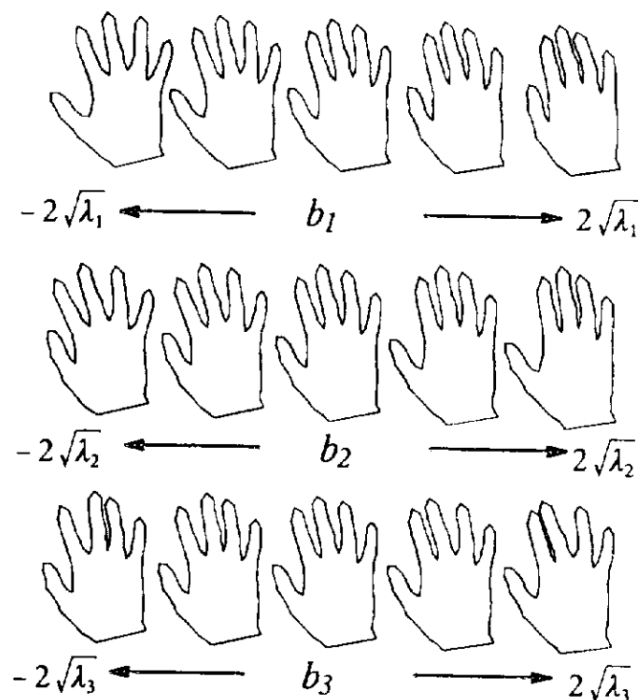
### 2.1 *Eigenfaces* para Reconhecimento Facial

O trabalho de Turk e Pentland (1991) apresentou os *eigenfaces*, que são os componentes principais de um conjunto de faces. Este método projeta imagens faciais em um espaço de atributos que abrange as variações mais significativas das faces usadas no processo. Os *eigenfaces* não necessariamente representam características pontuais da face como nariz, olhos ou boca. Na realidade, o método leva em conta os atributos faciais sob um ponto de vista mais amplo.

Apesar de ser capaz de extrair características faciais, o método de Turk e Pentland (1991) não garante que elas possam ser entendidas como pontos de referência. Assim, ele não pode ser aproveitado integralmente para suprir as necessidades da nossa proposta. Contudo, a maneira como os *eigenfaces* representam as componentes principais da face é bastante útil para nossa metodologia e é discutida no Capítulo 4.

## 2.2 Active Shape Models e Variações

Cootes et al. (1995) elaboraram um método que assimila os padrões de variação de uma determinada forma<sup>1</sup> e, com isso, constrói um modelo que a generaliza. O método, intitulado *Active Shape Models*, apresenta um modelo deformável que impõe restrições à forma durante sua busca. A Figura 5 ilustra como o modelo pode ser modificado até um limite que determina sua integridade.



**Figura 5** – Padrões de variação da forma da mão direita modelados pelo *Active Shape Models* (COOTES et al., 1995).

Resumidamente, o ASM possui uma etapa de construção do modelo de forma, que leva em consideração várias imagens com diversos pontos marcados, e uma etapa de busca em uma imagem desconhecida. Esta última consiste em, inicialmente, posicionar uma instância do modelo em alguma parte da imagem e aplicar as seguintes instruções alternadamente:

- encontrar uma nova posição para cada ponto da forma;
- aplicar restrições para as novas posições encontradas.

As duas instruções devem ser executadas iterativamente até que haja convergência na busca. Este método é um exemplo de proposta que não foi originalmente elaborada para

<sup>1</sup> Neste contexto, uma forma significa um conjunto de pontos em um espaço bidimensional.

a localização de pontos faciais, muito embora os próprios autores tenham, posteriormente, apontado para esta possibilidade em outro trabalho (COOTES; TAYLOR, 2004).

O *Active Shape Models* recebeu grande aceitação da comunidade e logo passou a pautar amplamente diversos trabalhos relacionados. Isso pode ser observado tanto na forma de novas aplicações propostas para a metodologia, quanto na forma de melhorias e otimizações do algoritmo original<sup>2</sup>.

Davatzikos, Tao e Shen (2003) observaram que a forma explorada pelo ASM pode ser decomposta em vários níveis de hierarquia. Com isso, eles propuseram uma variação no ASM tradicional, decompondo a forma em várias partes menores através da transformada de *Wavelet* (REISSELL, 1996). Eles demonstraram que esta alteração proporciona ao ASM um desempenho melhor quando o conjunto de treinamento é escasso. Para essa avaliação, a proposta foi testada na localização do contorno das mãos e na busca do corpo caloso do cérebro em imagens de ressonância magnética.

A dissertação de Milborrow (2007) traz um estudo sobre variações do ASM proposto por Cootes et al. (1995). A proposta dele é aplicar modificações no modelo original para obter uma boa taxa de acerto na identificação de pontos faciais em imagens frontais. Nosso trabalho também aproveita os conceitos propostos originalmente por Cootes et al. (1995) e considera alguns aspectos relevantes levantados por Milborrow (2007).

Um exemplo deles é a extensão do conjunto de pontos faciais. Milborrow (2007) apontou que o tamanho do conjunto é capaz de influenciar na precisão do algoritmo. Escolher uma quantidade maior de pontos é uma abordagem interessante, mesmo que haja algum subconjunto de pontos sobressalentes.

Outras sugestões não representam grandes mudanças para nosso trabalho. Por exemplo, aplicar duas instâncias do ASM em série pode ser vantajoso caso a estimativa de face inicial não seja muito boa. Contudo, como nossa abordagem considera faces obtidas em ambiente controlado e são adquiridas justamente para facilitar a solução do problema, essa melhoria não é muito interessante.

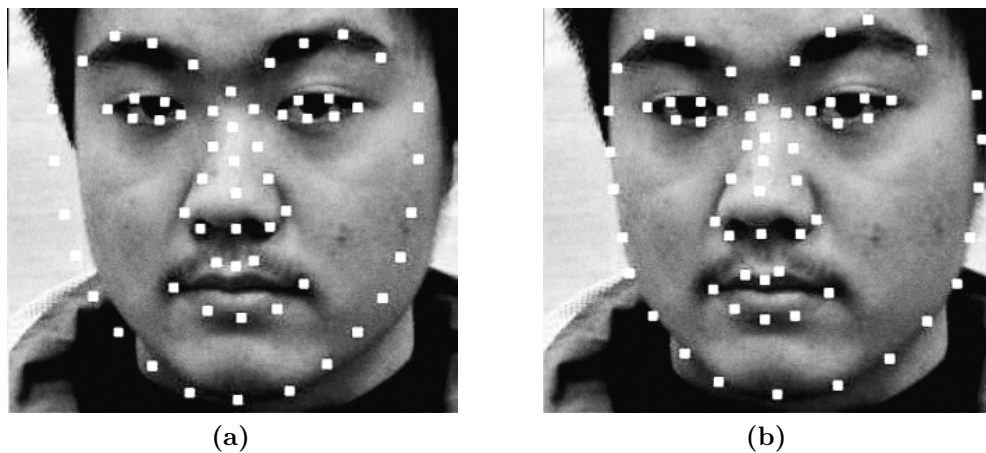
Du et al. (2008) também propuseram um trabalho que visa melhorar o ASM clássico. Uma das etapas de aplicação do algoritmo idealizado por Cootes et al. (1995) é a busca por novas localizações para os pontos. Nesta etapa, uma estimativa de forma é gerada com base no perfil de textura ao redor de cada ponto na imagem. A nova localização é escolhida a partir da comparação entre a média dos perfis do modelo e o perfil atual na imagem alvo através da minimização da distância *Mahalanobis*.

Entretanto, o método se mostra ineficaz quando o conjunto de imagens de treinamento não é grande o suficiente para garantir que os perfis obedeçam uma distribuição

<sup>2</sup> Ao longo desta dissertação, o ASM originalmente proposto por Cootes et al. (1995) é referido também por ASM clássico ou tradicional.



normal multivariada. [Du et al. \(2008\)](#) se propuseram a melhorar a busca pelos perfis que melhor se encaixam, transformando este em um problema de classificação. Para resolvê-lo, sugeriram a aplicação de *Support Vector Machine* ([VAPNIK, 2000](#)). A Figura 6 mostra os resultados das buscas com o ASM clássico e com o *Support Vector Machine Based Active Shape Models* (SVMBASM).



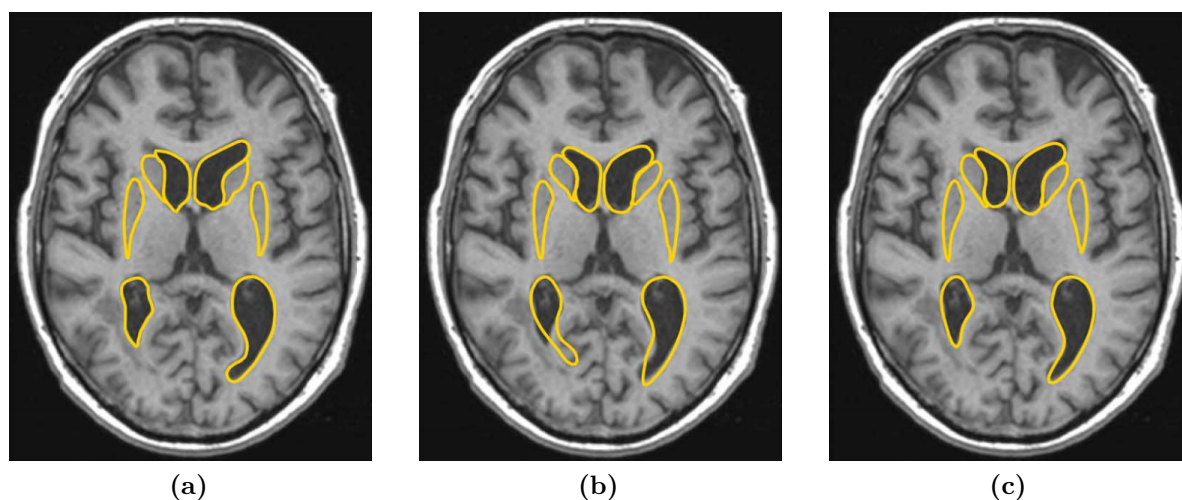
**Figura 6** – Localização de pontos faciais ([DU et al., 2008](#)). (a) Resultado do ASM tradicional. (b) Resultado do SVMBASM.

Assim como [Milborrow \(2007\)](#) e [Du et al. \(2008\)](#), [Seshadri e Savvides \(2009\)](#) apresentaram avanços para o ASM clássico no objetivo de localizar pontos faciais. Em seu trabalho eles adotaram uma das melhorias apontadas anteriormente para o ASM. [Milborrow \(2007\)](#) destacou que a quantidade de pontos pode servir para aperfeiçoar a precisão do algoritmo. Esse é um dos aspectos atacados por [Seshadri e Savvides \(2009\)](#). Eles citaram 79 pontos faciais para se levar em consideração na construção do modelo e na busca por atributos em imagens faciais.

Outra alteração importante visa atacar o problema da estimativa para a nova localização de um ponto. [Seshadri e Savvides \(2009\)](#) resolveram modificar a forma como a distância *Mahalanobis* é calculada, aumentando a taxa de acerto dos pontos a cada iteração.

O trabalho de [Cerrolaza, Villanueva e Cabeza \(2012\)](#) segue a linha apontada por [Davatzikos, Tao e Shen \(2003\)](#) de decompor a forma explorada pelo ASM através da transformada de *Wavelet*. A grande diferença é que eles estenderam a ideia para um ambiente com vários objetos, enquanto [Davatzikos, Tao e Shen \(2003\)](#) trabalharam com a restrição de decompor uma única forma. A Figura 7 ilustra um comparativo com a alteração proposta para o método, chamada *Active Hierarchical Shape Models* (AHSM).

Tais abordagens são bastante benéficas para nosso trabalho, pois demonstram a ideia de que uma forma explorada pelo ASM pode ser decomposta em diversos objetos. Apesar de nenhum dos dois trabalhos ter abordado características faciais, a própria



**Figura 7** – Localização de estruturas cerebrais subcorticais (CERROLAZA; VILLANUEVA; CABEZA, 2012). (a) Marcação manual. (b) Resultado do ASM tradicional. (c) Resultado do AHSM.

estrutura dos elementos da face indica que essas propostas podem ter êxito na solução desse problema.

### 2.3 *Active Appearance Models (AAM)*

Uma evolução do ASM foi proposta também por Cootes, Edwards e Taylor (2001). Além de usar informações sobre a forma de uma classe de objetos, *Active Appearance Models* também faz uso da textura global para buscar características de um objeto em imagens. Eles propuseram o uso do AAM para reconstrução facial com um modelo capaz de buscar simultaneamente forma e textura, diferente do ASM, que se restringe a coletar textura somente ao redor dos pontos que dão forma ao modelo.

Em relação ao ASM, o AAM é bem mais custoso do ponto de vista de construção do modelo. Por ser muito restrito quanto a textura, o ASM basicamente só faz uso da posição dos pontos e de algumas intensidades ao redor de cada um. Já o AAM precisa mapear toda a textura da região que compreende a face, o que aumenta bastante a complexidade desta solução. A Figura 8 mostra a aplicação do método para reconstruir faces em imagens.

Yan et al. (2003) apresentaram uma proposta similar ao AAM. O método intitulado *Texture Constrained Active Shape Models* explora as vantagens de se coletar amostras de intensidades ao redor dos pontos usados no ASM. Além disso, faz uso da textura global, como no AAM, para restringir a forma e prover um melhor critério de otimização para determiná-la.

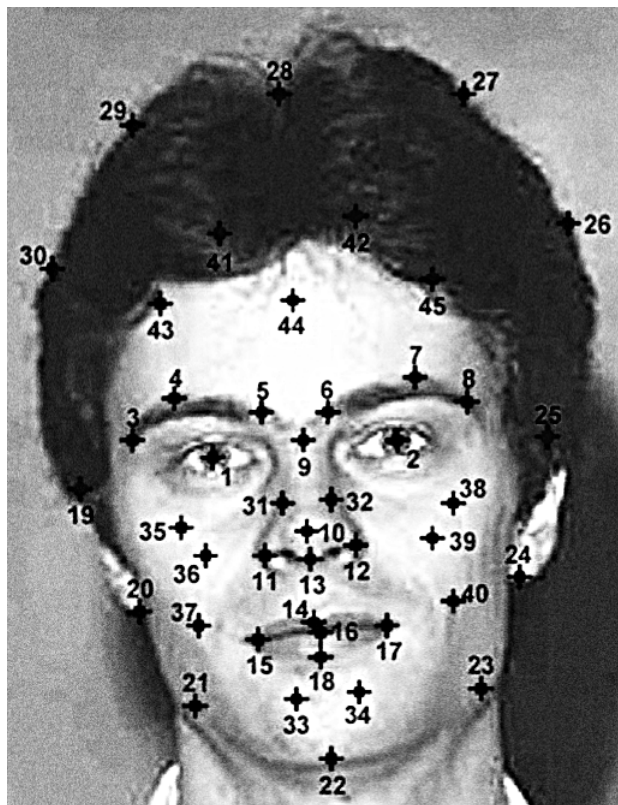


**Figura 8** – Exemplo de aplicação do AAM. Mais à esquerda, o modelo inicial é aplicado em imagem de baixa resolução. Ao longo das iterações (esquerda para direita), a instância resultante das aplicações de resolução mais baixa serve de entrada para versões de maior resolução da imagem original. As duas últimas colunas mostram a comparação da imagem original (mais à direita) com a iteração final do método (à esquerda) (COOTES; EDWARDS; TAYLOR, 2001).

## 2.4 Pontos Faciais para Reconhecimento

É bastante comum observar na literatura diferentes estratégias para a identificação de pessoas através de imagens faciais. Senaratne e Halgamuge (2006a) propuseram um método que faz uso de pontos faciais para resolver tal problema. A estratégia, chamada *Landmark Model Matching*, utiliza um modelo de distribuição de pontos criado a partir de algumas imagens de treinamento. Um modelo de pontos faciais, que representa a face de entrada, é deformado e encaixado na imagem. Posteriormente, eles propuseram uma evolução na fase de reconhecimento onde os pesos dados a cada ponto facial são calculados experimentalmente. No novo trabalho, eles fizeram o uso de otimização por enxame de partículas para calcular tais pesos e assim melhorar as taxas de reconhecimento (SENARATNE; HALGAMUGE, 2006b).

A grande relevância desse trabalho para esta investigação é a existência de uma etapa de localização de pontos faciais em imagens frontais. Eles fizeram uso de uma seleção bastante robusta de 45 pontos de referência para localização da face e identificação do indivíduo, ilustrados na Figura 9. No entanto, não foi feita nenhuma descrição formal dos pontos usados. A escolha dos mesmos foi feita de maneira intuitiva, considerando pontos que são largamente explorados (centro dos olhos, ponta do nariz, etc). Isso expõe uma deficiência que seria prejudicial na formulação de um plano operatório de cirurgia plástica, principal motivação do nosso trabalho. Em uma análise facial pré-operatória, cada ponto deve ser descrito com bastante cuidado para que não haja dúvidas com relação à sua correta localização.

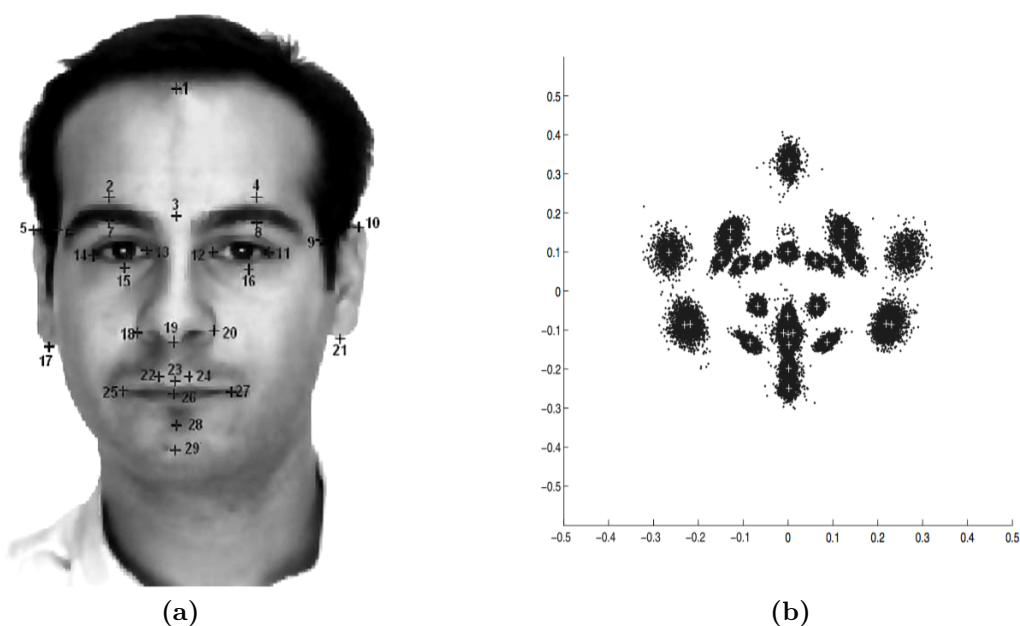


**Figura 9** – Conjunto de pontos faciais explorados pela estratégia *Landmark Model Matching*. É possível observar que a localização de alguns pontos é auto-explicativa. Entretanto muitos outros não são definidos com precisão, um exemplo disso é o ponto 44 (SENARATNE; HALGAMUGE, 2006b).

O trabalho de Shi, Samal e Marx (2006) também coloca em pauta os pontos faciais e discute a importância deles no papel da identificação. Nesse trabalho, foi introduzida uma preocupação com o tamanho crescente dos conjuntos de dados para reconhecimento facial. Eles propuseram um método baseado em pontos faciais e suas geometrias para reduzir o espaço de busca das faces. Além disso, eles sugeriram que seu método é eficiente para um reconhecimento facial direto.

Diferente do trabalho de Senaratne e Halgamuge (2006b), Shi, Samal e Marx (2006) foram mais cuidadosos com o conjunto de pontos faciais abordado. Apesar de utilizarem um conjunto mais modesto (29 pontos em vez de 45), eles apresentaram uma tabela com a descrição precisa de cada ponto. Muitos deles são, de fato, utilizados em análises pré-operatórias.

A Figura 10a indica o local de cada ponto em uma imagem facial. A Figura 10b mostra como os pontos coletados de 994 imagens rotuladas estão distribuídos. Essa distribuição é muito importante em vários trabalhos relacionados pois pode ser usada para construir modelos de distribuição de pontos. Esses modelos são explorados, por exemplo, no método ASM e consequentemente em nosso trabalho.



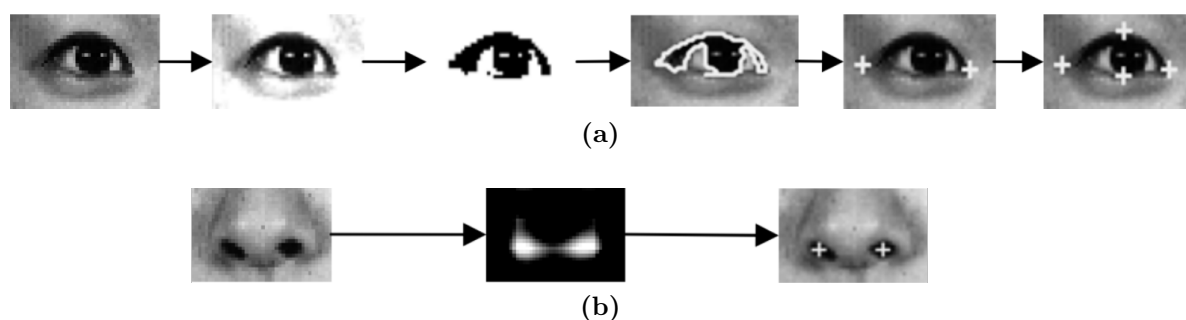
**Figura 10** – Os 29 pontos apresentados por [Shi, Samal e Marx \(2006\)](#). (a) A localização dos pontos faciais. (b) Distribuição dos pontos em um conjunto de imagens.

## 2.5 Detecção de Pontos Faciais Baseados em Modelos Antropométricos

Uma abordagem pouco usual na literatura é conduzir a busca dos pontos diretamente a partir de modelos antropométricos. Esses modelos são úteis para a localização dos pontos à medida que indicam as regiões mais prováveis onde eles podem ser encontrados. O trabalho de [Batista \(2007\)](#) visa obter informações sobre o ponto de atenção do indivíduo a partir da direção do seu olhar. O objetivo é atingido com a localização dos pontos faciais mais importantes feita a partir de um modelo antropométrico da face humana. Tal estratégia é baseada na probabilidade de um ponto se encontrar em determinada região da face de acordo com o modelo. A partir dessa análise estatística, algumas etapas de processamento digital de imagens (PDI) são responsáveis por alterar a imagem e facilitar a busca pelos pontos.

A proposta de [Sohail e Bhattacharya \(2008\)](#) também se apresenta bastante diferente da maioria dos trabalhos estudados. Seu objetivo é localizar os 18 pontos faciais mais relevantes. Ela também utiliza um modelo antropométrico para restringir o campo de busca por alguns pontos. Após encontradas as regiões de cada característica facial, ele propõe a aplicação de técnicas de PDI específicas para cada ponto. A [Figura 11](#) exemplifica algumas etapas de processamento exploradas por eles.

A grande vantagem tanto do método de [Batista \(2007\)](#) quanto de [Sohail e Bhattacharya \(2008\)](#) está na ausência de se realizar etapas de pré-processamento para a aplicação



**Figura 11** – Algumas etapas de processamento exploradas por [Sohail e Bhattacharya \(2008\)](#). (a) Fluxo para localizar pontos do olho. (b) Fluxo para localizar fossas nasais.

dos algoritmos. Contudo, para o nosso trabalho a grande contribuição se encontra na utilização de um modelo antropométrico, pois isso é bastante importante na redução do espaço de busca pelos pontos.

[Martinez et al. \(2013\)](#) propuseram um algoritmo para detectar pontos faciais em imagens frontais. O método combina uma abordagem baseada em regressão com um modelo probabilístico da face humana. Isso ajuda a restringir o campo de busca a regiões da face que são consistentes de um ponto de vista antropométrico. O trabalho foi testado em mais de 7500 imagens de diversas bases de dados bastante exploradas e apropriadas para o objetivo proposto.

Esse trabalho explora 20 pontos fiduciais<sup>3</sup> largamente utilizados em trabalhos do gênero. A avaliação dos resultados considerou uma série de diferentes cenários que envolveram imagens com muitas adversidades, incluindo variações étnicas e etárias, variações em luminosidade e até mesmo oclusão.

A proposta de avaliação da precisão do algoritmo foi bastante interessante. Eles fizeram uso de uma medida normalizada com relação à distância das pupilas de cada indivíduo. Tal medida mostra-se bastante adequada pois permite que o trabalho possa ser comparado a outros que utilizarem o mesmo conjunto de teste.

## 2.6 Identificação de Pontos Faciais em Perspectivas Diversas

Outro trabalho que faz uso de pontos faciais foi proposto por [Bottino e Cumani \(2008\)](#). Eles propuseram um método para localização e identificação de pontos do perfil facial visando análises para cirurgias plásticas. A proposta deles baseia-se em uma análise das curvaturas do perfil usando as derivadas de primeira e segunda ordem, avaliadas por uma convolução com *kernels* derivados de Gauss. O resultado final consiste nos pontos que representam extremos locais do perfil. Esta é uma abordagem interessante para este

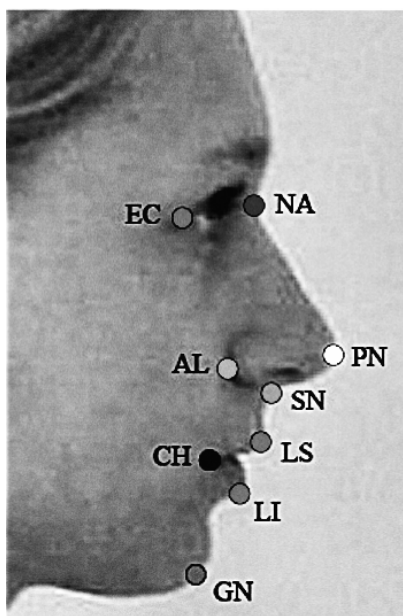
<sup>3</sup> Pontos de referência facial.

trabalho pois leva em conta a identificação de cada ponto já com o objetivo de realizar análises estéticas.

Assim como [Shi, Samal e Marx \(2006\)](#), eles tomaram cuidado em descrever, minuciosamente, cada ponto explorado. A Tabela 1 detalha cada ponto usado. A Figura 12 ilustra os pontos abordados. De todos os trabalhos investigados, é o que se apresenta mais similar ao nosso por ter motivação e objetivos parecidos ([BOTTINO et al., 2012](#)). Por isso, faz-se necessário para esta investigação colocar em pauta algumas ressalvas com relação aos objetivos e, também, aos resultados.

**Tabela 1** – Pontos de referência segundo [Bottino e Cumani \(2008\)](#).

Nome	Sigla	Descrição
<i>Pronasale</i>	PN	ponta do nariz
<i>Nasion</i>	NA	região profunda diretamente entre os olhos
<i>Subnasal</i>	SN	extremidade inferior da linha do nariz
<i>Labrale Superius</i>	LS	lábio superior
<i>Labrale Inferius</i>	LI	lábio inferior
<i>Cheilion</i>	CH	canto da boca
<i>Gnathion</i>	GN	ponto inferior da mandíbula
<i>Exocanthion</i>	EC	canto externo do olho
<i>Alare</i>	AL	ponto mais externo da asa nasal



**Figura 12** – Conjunto de pontos faciais explorados por [Bottino e Cumani \(2008\)](#).

Em primeiro lugar, tal proposta se resume a identificar pontos no perfil facial. Em oposição aos vários trabalhos estudados, eles deixaram de fora os pontos da imagem frontal. Fica evidente que o objetivo deles era dividir a análise dos pontos. Isso representa

uma deficiência, pois nossa proposta necessita de um conjunto de pontos diferente para a realização das análises pré-operatórias.

Os outros aspectos que merecem discussão dizem respeito aos testes e resultados. Em todo trabalho relacionado, é razoável que os autores realizem testes para mostrar, na prática, a eficácia das suas propostas. Não fugindo desta abordagem, [Bottino e Cumani \(2008\)](#) utilizaram um banco de imagens de perfil com cerca de 60 imagens de vários indivíduos. Em comparação com os outros trabalhos investigados, a proposta deles se mostrou a mais frágil nesse aspecto, já que outras propostas chegaram a avaliar os resultados com até 1000 imagens. Seria mais interessante a utilização de um banco de imagens mais robusto para aumentar o grau de confiança do algoritmo.

Deixando para trás a análise quantitativa e partindo para um ponto de vista qualitativo, o banco apresenta uma discrepância com a realidade estudada em nosso trabalho. Este aspecto não deve ser encarado como um ponto negativo do artigo de [Bottino e Cumani \(2008\)](#), mas sim como uma característica relativa a uma determinada região. As imagens por eles usadas nos testes são todas de indivíduos que pertencem a uma mesma localidade. Nos exemplos que foram expostos, não foi possível observar grandes diferenças étnicas.

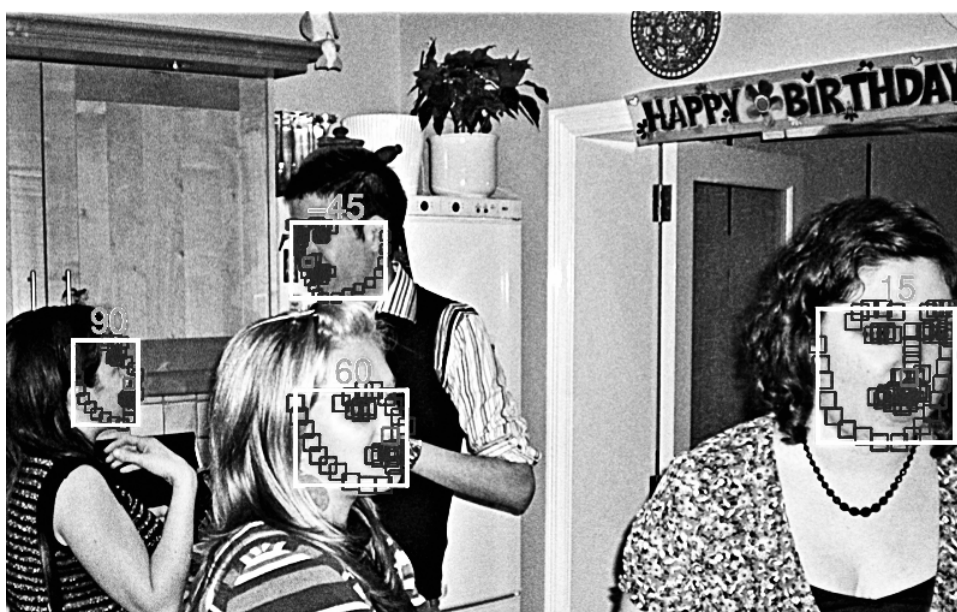
Um dos pontos importantes do nosso trabalho é abordar a diversidade étnica para refinar a análise estética, pois tais diferenças indicam padrões que devem ser respeitados. A partir do momento em que não há diversidade nas imagens testadas, é possível presumir que o algoritmo apresentará resultados diferentes em outros contextos. Além disso, os indivíduos do banco de imagens aparentam estar na mesma faixa etária, que é um outro aspecto importante a ser considerado nos testes.

O último ponto de reflexão com relação ao trabalho de [Bottino e Cumani \(2008\)](#) diz respeito à avaliação dos resultados. Em se tratando de uma imagem digital, a localização de um ponto pode até mesmo ser exata, contudo é muito provável que existam diferenças entre as marcações feitas por especialistas e pelo algoritmo. Por isso deve-se encontrar algum meio para calcular a precisão que o algoritmo atinge. Para isso, [Bottino e Cumani \(2008\)](#) fizeram uso da distância Euclidiana entre marcações manuais e automáticas. O grande problema é que eles não se preocuparam em relativizar as distâncias medidas, o que prejudica qualquer tentativa de comparar seus resultados com outros trabalhos, mesmo aqueles que usem a mesma base de imagens. Eles mostraram apenas medições com valores absolutos em *pixels* (px). Porém, tais números são altamente dependentes do tamanho das imagens faciais.

Outras estratégias de avaliação procuram relacionar o erro absoluto com o tamanho da face ou com a distância entre os olhos. As métricas de aferição do erro na localização dos pontos são discutidas no Capítulo 3.



Um dos trabalhos mais completos com relação à localização de pontos faciais é o de [Zhu e Ramanan \(2012\)](#). Esse trabalho traz uma solução unificada para três problemas comuns na biometria e identificação de indivíduos em um ambiente real: detecção facial, estimativa de pose facial e localização de pontos de referência. Um importante diferencial do trabalho de [Zhu e Ramanan \(2012\)](#) é a necessidade por realizar tarefas de identificação em ambientes não controlados, onde as condições para aquisição de imagens devem ser adversas. Assim, eles propuseram um método que unifica os três problemas, já que eles são comumente abordados de maneira separada. A Figura 13 mostra um exemplo de imagem usada para testar o algoritmo.



**Figura 13** – Imagem com faces em diversas posições obtida em ambiente não-controlado. O algoritmo detecta a face (caixa maior ao redor da face), localiza os pontos (caixas menores dentro da face) e estima a pose (rótulo de inclinação acima da face) ([ZHU; RAMANAN, 2012](#)).

Outro diferencial do método, é a utilização de árvores ao invés de grafos densos para representar os pontos faciais e suas conexões. Isto permite a modelagem de muitas poses com baixa complexidade. Os resultados foram comparados com sistemas acadêmicos e comerciais como *Google Picasa* e *face.com*. O algoritmo deles atingiu uma taxa de 100% de acerto na localização de pontos em uma face frontal, considerando um erro de até 5% em relação ao tamanho da imagem<sup>4</sup>.

A detecção facial realizada pelo algoritmo alcançou performance bastante próxima à das soluções comerciais. Com relação à outra tarefa, a estimativa de pose não é exatamente o foco do nosso trabalho mas é importante para ajudar na localização de pontos em imagens de perfil. Essa perspectiva também é explorada pelo trabalho de [Zhu e Ramanan \(2012\)](#) com bons resultados.

<sup>4</sup> O erro baseado no tamanho da imagem serve para normalizar a medida. Esse tamanho mencionado é uma média entre altura e largura da imagem usada. O resultado foi relativo à base de dados *MultiPIE*.

## 2.7 Considerações Finais

No que diz respeito aos objetivos desta dissertação, este capítulo relacionou basicamente dois tipos de trabalhos. Alguns são caracterizados por propostas que têm o mesmo objetivo de localizar pontos faciais automaticamente. Deles, foram destacados aspectos positivos para nosso trabalho e discutidos pontos negativos, que também contribuíram para o enriquecimento desta proposta. Os outros são compostos por metodologias que se propõem a localizar características diferentes do nosso objetivo, mas que podem de alguma forma ser aproveitadas e adaptadas para nossa solução.

Propostas como de [Cootes et al. \(1995\)](#) e de [Bottino e Cumani \(2008\)](#) foram largamente discutidas porque se mostraram bastante similares à nossa ideia. [Bottino e Cumani \(2008\)](#) expuseram motivação e objetivos muito parecidos com nosso trabalho, ao passo que [Cootes et al. \(1995\)](#) apresentaram uma estratégia bastante reconhecida e que norteou grande parte da nossa solução. Esses e outros trabalhos estudados indicam que, nesta área de pesquisa, existe bastante preocupação com o tema apresentado. Na literatura, surgem continuamente tanto novas aplicações para o problema proposto quanto novas soluções e formas de aprimoramento de metodologias existentes.

Diante desses trabalhos, é possível concluir que este tema é bastante relevante para a área, permanece amplamente explorado pela comunidade científica e ainda tem lacunas a serem preenchidas.

## 3 Materiais e Métodos

Para viabilizar a metodologia proposta neste trabalho, é necessário que se divida o problema em etapas. Desse modo, cada parte dela pode ser construída e testada separadamente, agilizando a resolução do problema e deixando-a mais robusta.

Este capítulo apresenta todo o arcabouço de técnicas utilizado em nosso trabalho e os métodos explorados para avaliar a qualidade da solução proposta. Primeiramente, é apresentada uma investigação preliminar que lança luz sobre os aspectos estéticos da face. Posteriormente, são discutidas técnicas de PDI, algumas notações relevantes e outros métodos explorados.

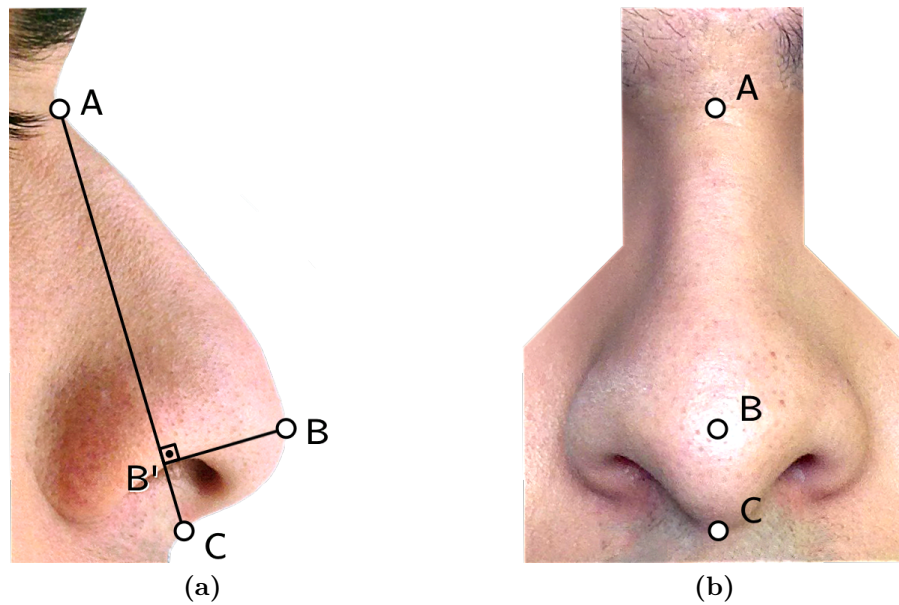
### 3.1 Elementos da Análise Estética

A análise estética da face é quem deve ditar que pontos faciais precisam ser aproveitados no nosso trabalho. Cada métrica de avaliação usada pelos cirurgiões leva em conta subconjuntos específicos de pontos faciais. Além disso, essa análise também impõe quais perspectivas ou visões devem ser levadas em consideração no momento de se avaliar esteticamente a face. É importante notar que a análise facial não deve ser conduzida somente a partir de uma visão frontal, já muito explorada em trabalhos relacionados.

Apesar de não ser um trabalho recente, *Proportions of the Aesthetic Face* define variadas métricas e proporções ideais para muitas características faciais (POWELL; HUMPHREYS, 1984). Um exemplo disso é a razão de Baum, que é uma das métricas mais citadas para avaliar-se a projeção da ponta do nariz. A Figura 14a mostra um segmento de reta  $\overline{AC}$  e outro segmento  $\overline{BB'}$ , sendo  $B'$  a projeção de  $B$  em  $\overline{AC}$ . Essa métrica é calculada através da divisão do comprimento de  $\overline{AC}$  pelo comprimento do segmento  $\overline{BB'}$ .

A razão de Baum é um bom exemplo da necessidade de separar a análise em diferentes visões. É possível observar que os pontos necessários para calculá-la são visíveis em uma imagem frontal. Contudo, é extremamente complicado o cálculo da projeção da ponta do nariz a partir de tal imagem. Nesse caso, o cálculo dessa métrica não é aplicável e deve ser feito pela visão do perfil facial. A Figura 14 expõe uma comparação entre duas visões no cálculo dessa métrica.

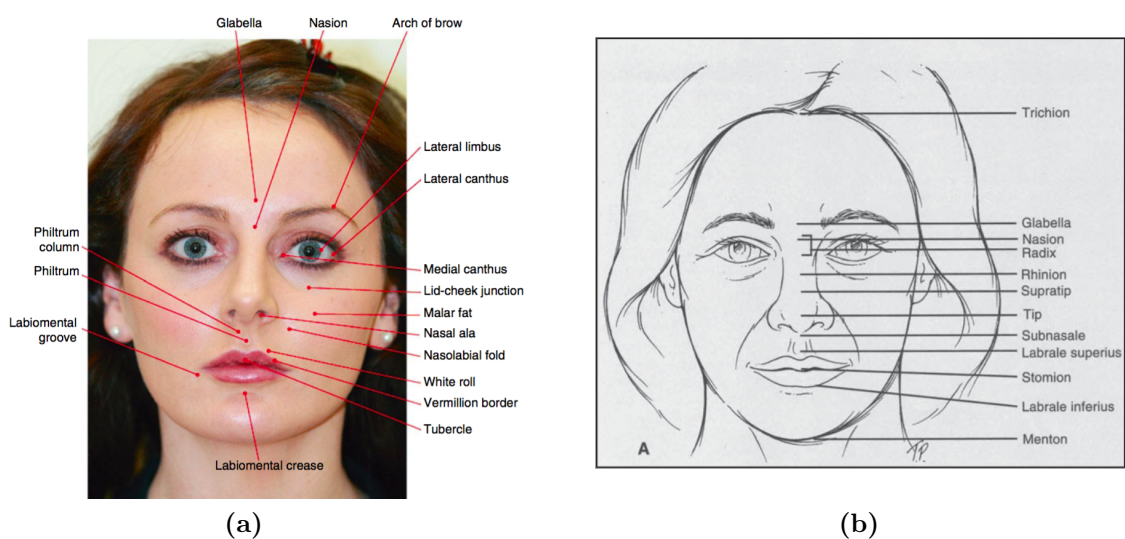
Nosso trabalho não tem como objetivo calcular essas métricas ou conduzir a avaliação estética, portanto esses aspectos não são levantados em detalhes. Contudo, as métricas são importantes pois justificam a escolha dos pontos de referência e das perspectivas de análise.



**Figura 14** – Imagens do nariz em duas visões. (a) Visão do perfil para cálculo da razão de Baum. (b) Visão frontal.

### 3.1.1 Pontos Faciais Frontais

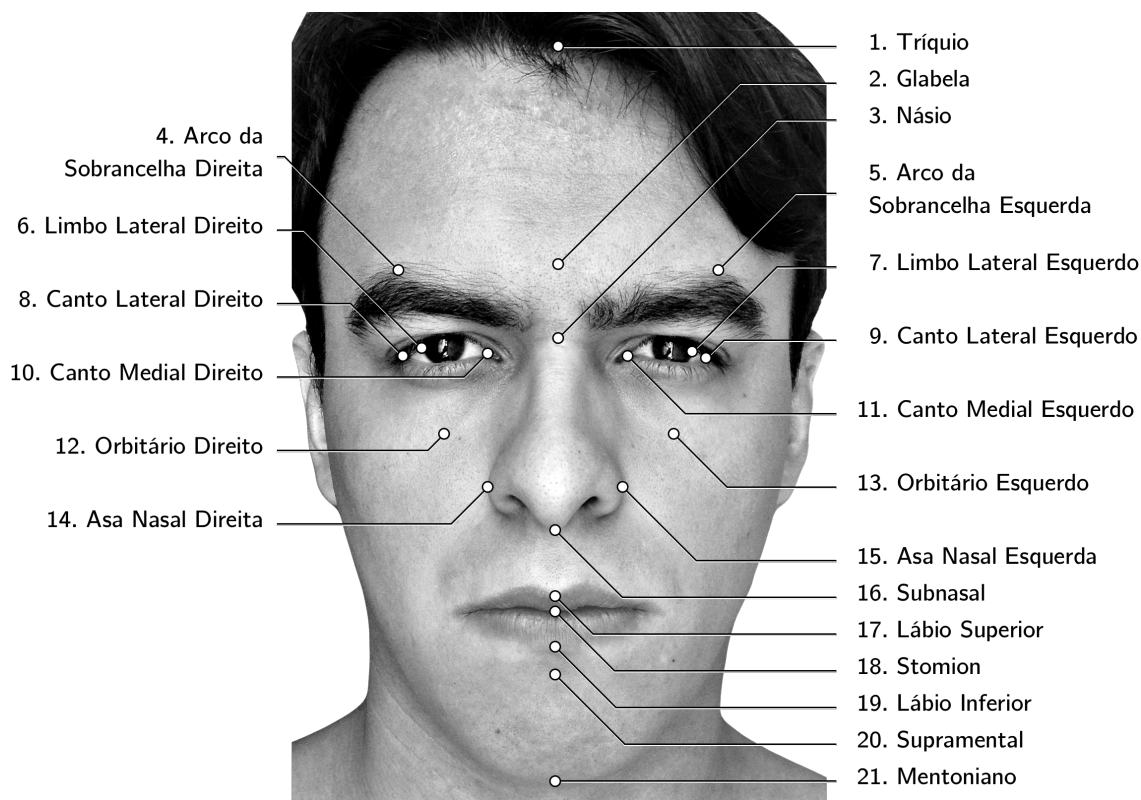
Zimblér e Ham (2010) propuseram uma análise estética da face salientando um conjunto de pontos necessários para realizá-la. Para complementar algumas definições, neste trabalho também foram utilizados os conceitos apresentados por Prendergast (2012). Assim como Zimblér e Ham (2010), ele levanta uma discussão sobre análise estética da face e complementa o conjunto de pontos necessários para tanto. A Figura 15 mostra os pontos abordados pelos dois trabalhos. A Tabela 2 mostra a descrição de todos os pontos de referência frontais explorados neste trabalho e ilustrados na Figura 16.



**Figura 15** – Pontos faciais frontais. (a) Conjunto de pontos definidos por Prendergast (2012). (b) Conjunto de pontos definidos por Zimblér e Ham (2010).

**Tabela 2** – Descrição dos pontos de referência frontais.

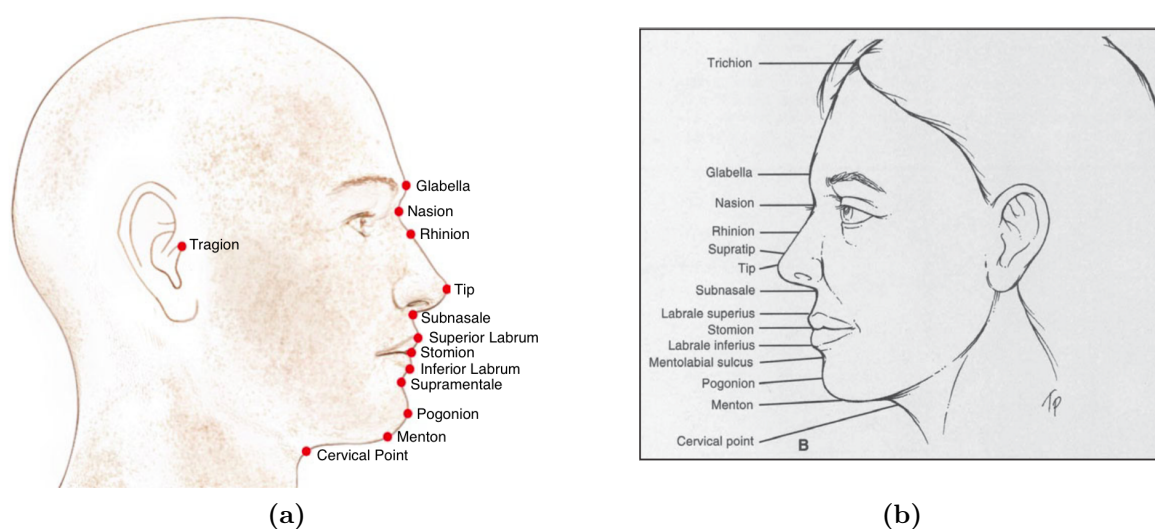
Nº	Sigla	Descrição
1	Tr	linha do couro cabeludo ao centro
2	Gl	parte mais proeminente, ao centro, entre as sobrancelhas
3	Ns	raiz do nariz ao centro
4	ASd	ponto mais superior da sobrancelha direita
5	ASe	ponto mais superior da sobrancelha esquerda
6	LLd	ponto lateral da íris direita mais afastado do nariz
7	LLe	ponto lateral da íris esquerda mais afastado do nariz
8	CLd	ponto lateral do olho direito mais afastado do nariz
9	CLe	ponto lateral do olho esquerdo mais afastado do nariz
10	CMd	ponto do olho direito mais próximo do nariz
11	CMe	ponto do olho esquerdo mais próximo do nariz
12	Od	ponto mais inferior da órbita direita
13	Oe	ponto mais inferior da órbita esquerda
14	ANd	ponto mais lateral da asa do nariz à direita
15	ANe	ponto mais lateral da asa do nariz à esquerda
16	Sn	junção da columela com a pele acima do lábio superior
17	LS	junção das partes vermelha e cutânea do lábio
18	St	ponto onde os lábios se encontram ao centro
19	LI	ponto, ao centro, do lábio de baixo na borda de vermilion
20	Sm	dobra labiamental entre o lábio inferior e o queixo
21	Mt	a parte mais inferior do queixo



**Figura 16** – Localização dos pontos frontais definidos na Tabela 2.

### 3.1.2 Pontos Faciais de Perfil

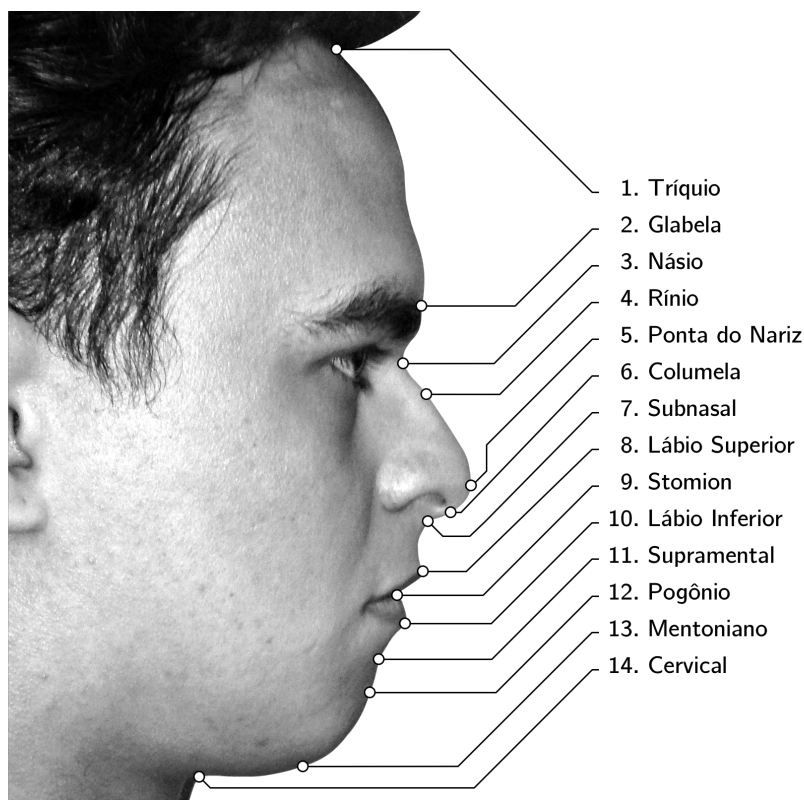
Como as métricas de avaliação estética da visão frontal não podem ser aplicadas para a avaliação do perfil, é necessário buscar outro conjunto que complemente a análise. [Prendergast \(2012\)](#) também definiu um conjunto de pontos do perfil facial para análise estética. As duas perspectivas apresentam pontos em comum, cada uma à sua própria ótica, como exemplificado na Figura 14. [Zimbler e Ham \(2010\)](#) também propuseram um conjunto de pontos para o perfil facial. Ambos os conjuntos estão ilustrado na Figura 17. A Tabela 3 mostra a descrição de todos os pontos de referência do perfil facial explorados neste trabalho e ilustrados na Figura 18.



**Figura 17** – Pontos do perfil facial. (a) Conjunto de pontos definidos por [Prendergast \(2012\)](#). (b) Conjunto de pontos definidos por [Zimbler e Ham \(2010\)](#).

**Tabela 3** – Descrição dos pontos de referência do perfil facial.

Nº	Sigla	Descrição
1	Tr	linha do couro cabeludo ao centro
2	Gl	parte mais proeminente, ao centro, entre as sobrancelhas
3	Ns	raiz do nariz ao centro
4	Rn	junção ósseo cartilaginosa no dorso nasal
5	PN	parte mais anterior do nariz
6	Col	conecta o ápice do nariz ao filtro labial
7	Sn	junção da columela com a pele acima do lábio superior
8	LS	junção das partes vermelha e cutânea do lábio
9	St	ponto onde os lábios se encontram ao centro
10	LI	ponto, ao centro, do lábio de baixo na borda de vermilion
11	Sm	dobra labiamental entre o lábio inferior e o queixo
12	Pg	a parte mais anterior do queixo
13	Mt	a parte mais inferior do queixo
14	Crv	a parte mais interna entre a área submental e o pescoço



**Figura 18** – Localização dos pontos do perfil facial definidos na Tabela 3.

Os pontos apontados são considerados anatômicos, pois foram escolhidos por um especialista com base em descrições antropométricas. Em alguns cenários discutidos nesta dissertação houve a necessidade de se incrementar o conjunto de pontos apresentado nesta seção. Esses novos pontos são derivados matematicamente, através de algum conceito formal, ou são apenas pontos marcados para preencher o contorno da forma (DRYDEN; MARDIA, 1998). Contudo, independente do cenário em questão, os pontos anatômicos sempre devem compor as análises.

## 3.2 Processamento Digital de Imagens

Tendo em vista o caráter multidisciplinar do trabalho, é preciso destacar, além dos aspectos da análise estética, os conceitos que devem ser aproveitados para a metodologia propriamente dita. Nosso trabalho busca resolver o problema apresentado através da análise de imagens. A metodologia deve fazer uso de uma imagem facial e determinar onde se encontram, dentro daquela imagem, os pontos de referência levantados na Seção 3.1.

A maioria das técnicas de análise de imagens e visão computacional são compostas, inicialmente, por uma etapa de processamento digital de imagens. As técnicas de PDI permitem alterar uma imagem para que ela se torne mais agradável ao olho humano, ou modificá-la de maneira a facilitar uma extração automática de características, que é a intenção do nosso trabalho.

Nesta seção são apresentados vários elementos necessários para a preparação da imagem que deve ser analisada. São definidos alguns conceitos preliminares, que são exaustivamente abordados nesta dissertação. Além disso, ao longo da seção são apresentadas técnicas consolidadas na área que são indispensáveis à solução do problema.

### 3.2.1 A Imagem Digital

Uma definição bastante recorrente trata a imagem como sendo uma função bi-dimensional  $f(x, y)$ , onde  $x$  e  $y$  são as coordenadas de um plano e a amplitude de  $f$  no par ordenado  $(x, y)$  é chamada de intensidade ou nível de cinza da imagem naquele ponto. Se considerarmos que  $x$ ,  $y$  e os valores da amplitude de  $f$  pertencem a domínios discretos e finitos, podemos dizer que  $f(x, y)$  se trata de uma imagem digital. A partir dessa definição, pode-se concluir que a imagem digital é formada por uma quantidade finita de elementos, cada um com uma posição. Eles são amplamente conhecidos como *pixels* (GONZALEZ; WOODS, 2011).

A Equação 3.1 mostra como essa função pode ser expressa na forma de uma matriz. Ela representa uma imagem onde existem  $M$  linhas e  $N$  colunas e, portanto, pode ser mapeada em até  $M \times N$  valores de amplitude.

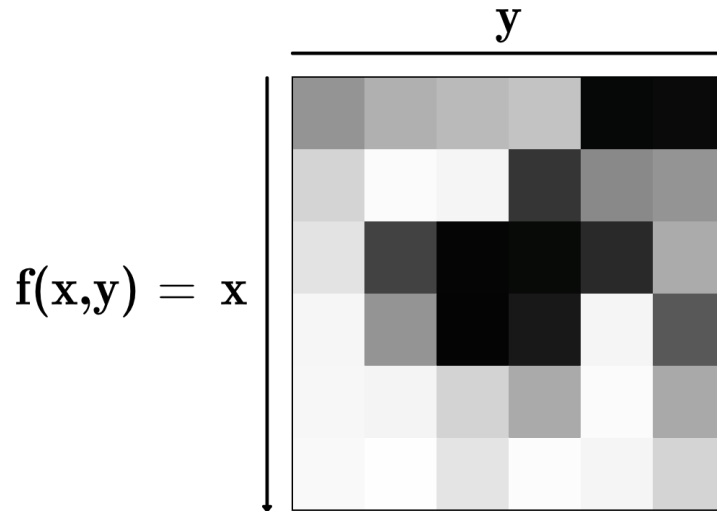
$$f(x, y) = \begin{bmatrix} f(0, 0) & f(0, 1) & \dots & f(0, N - 1) \\ f(1, 0) & f(1, 1) & \dots & f(1, N - 1) \\ \vdots & \vdots & \ddots & \vdots \\ f(M - 1, 0) & f(M - 1, 1) & \dots & f(M - 1, N - 1) \end{bmatrix} \quad (3.1)$$

Além da forma apresentada na Equação 3.1, existem outras representações bastante exploradas. Uma das mais relevantes para nosso trabalho é a representação em forma de vetor. Ela consiste em dispor toda a imagem digital em um vetor coluna. Nessa representação, as linhas da imagem são organizadas em colunas e dispostas uma abaixo da outra, formando um vetor. A Equação 3.2 exhibe uma forma equivalente da função  $f(x, y)$  expressa em vetor coluna de tamanho  $MN \times 1$ .

$$f(x, y) = [f(0, 0) \quad \dots \quad f(0, N - 1) \quad \dots \quad f(M - 1, 0) \quad \dots \quad f(M - 1, N - 1)]^T \quad (3.2)$$

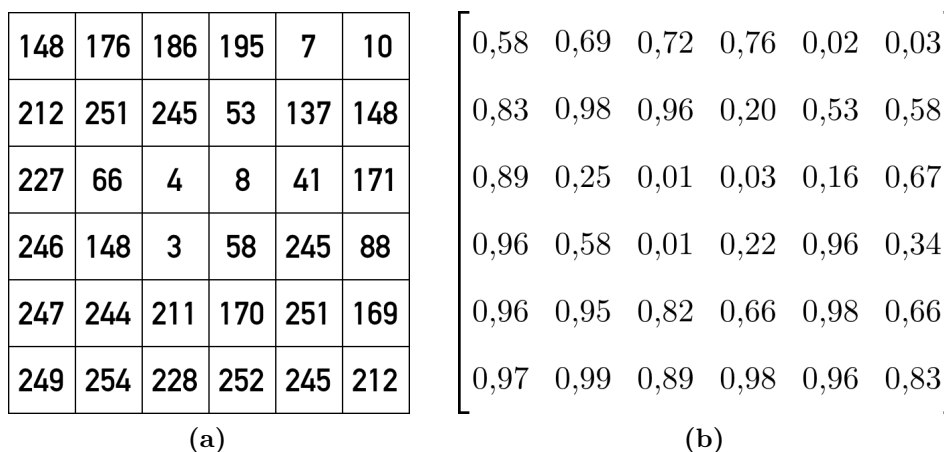
A Figura 19 mostra uma representação visual de uma imagem digital de dimensão  $6 \times 6$  *pixels*. Isso significa que a função pode ser mapeada em 36 valores de amplitude. As setas indicam o sentido de crescimento dos valores  $x$  e  $y$  pela imagem. Observa-se então, que a origem da representação visual da imagem digital é localizada, por convenção, no canto esquerdo superior. Tal convenção condiz com a representação em forma de matriz, cujos crescimentos dos índices respeitam a mesma ordem.





**Figura 19** – Representação de uma imagem digital com resolução  $6 \times 6$  em níveis de cinza.

Imagens de 8 *bits* em níveis de cinza são imagens digitais onde o valor da amplitude é expresso em 8 *bits*. Assim, o domínio das intensidades que um *pixel* pode assumir fica restrito a 256 valores, variando de 0 até 255. Além da representação visual da Figura 19, uma imagem de 8 *bits* em níveis de cinza pode ser expressa de outras maneiras relevantes. A Figura 20a mostra uma representação numérica das intensidades da imagem anterior. Observe que os *pixels* mais claros da imagem apresentam valores numéricos próximos a 255 - valor máximo de intensidade. Por outro lado, os *pixels* mais escuros têm valores próximos a 0. A Figura 20b exibe uma forma de representação onde as intensidades estão normalizadas (intervalo  $[0,1]$ ), outra abordagem bem utilizada em trabalhos do gênero.



**Figura 20** – Formas de representação da imagem da Figura 19. (a) Representação numérica em 8 *bits*. (b) Representação em matriz com os valores normalizados.

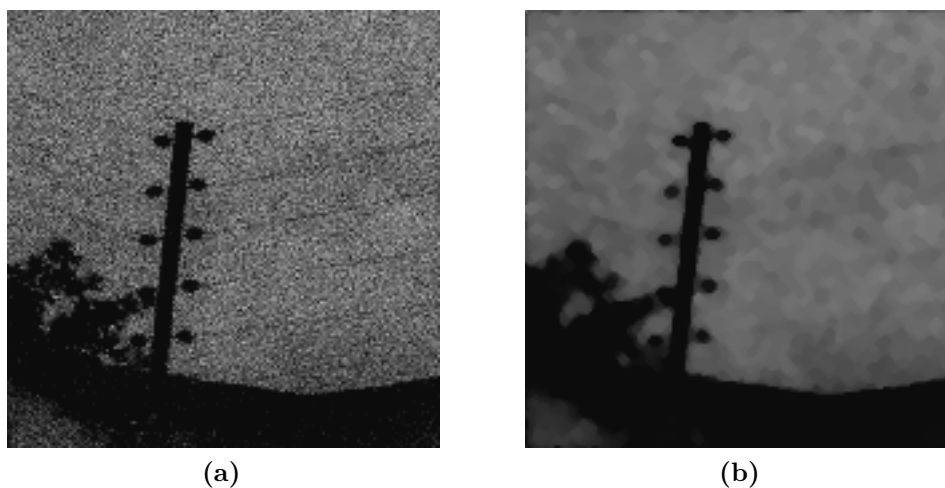
Neste trabalho, a imagem pode assumir as diferentes formas mencionadas nesta seção. Dependendo da circunstância, é mais vantajoso optar por uma representação em detrimento de outras.

### 3.2.2 Filtragem Mediana

As imagens são frequentemente alvos de um problema durante a fase de captura ou aquisição. Em alguns casos, o sinal que dá origem à imagem digital é acometido por perturbações aleatórias. Isso provoca alterações no brilho original da cena capturada e o resultado é uma imagem que apresenta divergências com o objeto real. A esse tipo de variação na captura dá-se o nome de ruído. Seu surgimento pode acontecer, também, durante a transmissão do sinal da imagem de um meio para outro.

Esses sinais indesejados atrapalham as etapas mais agudas de qualquer sistema de análise de imagem, pois dificultam a extração de atributos. Se este problema não for tratado, os ruídos acabam sendo interpretados como características próprias da imagem e causam prejuízo no entendimento do objeto observado.

Dada a natureza aleatória desses ruídos, a filtragem mediana se mostra uma boa opção para tratar esse problema. Ela é uma técnica de processamento de sinais não linear que consiste na aplicação de uma janela deslizante sobre a imagem. Esta janela possui uma quantidade ímpar de elementos que são coletados da região na imagem onde ela se encontra. O *pixel* central da janela deve ser substituído pela mediana daqueles que estiverem dentro da janela. Geralmente, o tamanho desta estrutura é de  $3 \times 3$  *pixels*, mas pode-se usar outras variações dependendo da quantidade de ruído (PRATT, 2007). A Figura 21 mostra como a filtragem mediana pode ser útil no tratamento do ruído.



**Figura 21** – Aplicação da filtragem mediana em imagem ruidosa. (a) Imagem capturada em condições adversas de iluminação. (b) Imagem após sucessivas aplicações da mediana com janela  $3 \times 3$ .

Pratt (1975) foi um dos primeiros a propor o uso do filtro da mediana em imagens. Além disso, trabalhos como o de Huang, Yang e Tang (1979) teorizaram melhorias para a filtragem mediana como proposta para redução de ruído. Em trabalho mais recente, Arias-Castro e Donoho (2009) trouxeram uma reflexão muito interessante para o tópico.

Eles discutiram sobre a capacidade que o filtro da mediana tem de preservar as bordas<sup>1</sup> presentes em uma imagem. No seu estudo, eles colocaram alguns cenários onde ela não era capaz de preservar bordas melhor que uma filtragem linear.

Ainda assim, segundo eles, o filtro da mediana é uma boa escolha para eliminação de ruídos sem que as bordas sejam prejudicadas. Isso é especialmente importante para metodologias dependentes de bordas, onde as imagens precisam ser suavizadas.

### 3.2.3 Convolução

Além da mediana, outros tipos de filtragem são utilizados neste trabalho. Para compreender o funcionamento deles é necessário ter em mente o conceito matemático da convolução em domínio discreto.

A convolução discreta é uma operação linear realizada entre uma função  $g$  e uma função  $f$  e tem como resultado  $h$ . A função gerada da aplicação pode ser encarada como uma versão modificada de uma das funções usadas no processo. No âmbito das imagens digitais - que também são funções - a convolução é feita a partir de um *kernel*, também chamado de matriz de convolução. A Equação 3.4 mostra como é calculado o elemento  $(x, y)$  de uma função  $h$  gerada a partir da convolução de um *kernel*  $g$  de tamanho  $N \times N$ , onde  $N$  deve ser ímpar, na função  $f$ .

$$m = \left\lfloor \frac{N}{2} \right\rfloor \quad P = \begin{cases} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} g(i, j) & , \text{ se } \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} g(i, j) \neq 0 \\ 1 & , \text{ caso contrário} \end{cases} \quad (3.3)$$

$$h(x, y) = \frac{\sum_{i=-m}^m \sum_{j=-m}^m g(m-i, m-j) \cdot f(x+i, y+j)}{P} \quad (3.4)$$

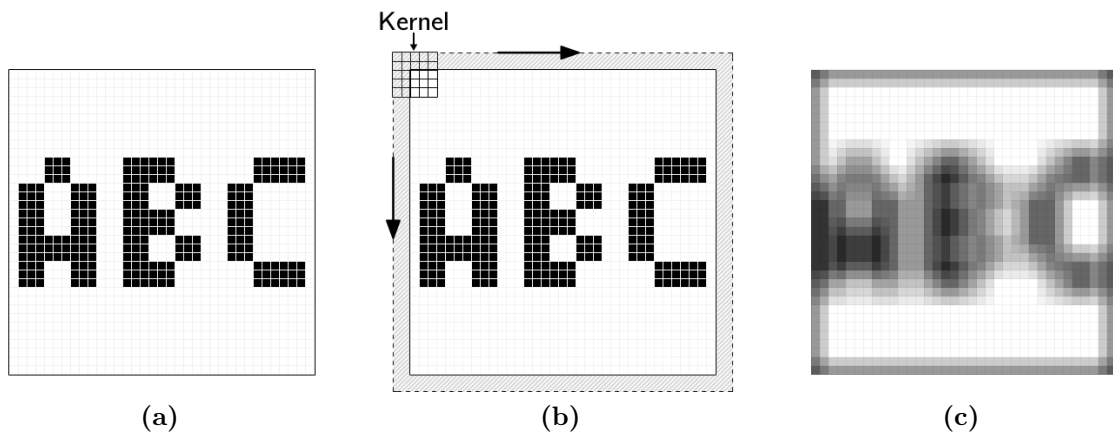
A convolução pode ser usada para filtrar uma imagem (função  $f$ ) a partir de simples modificações no *kernel* (função  $g$ ). Quando aplicada em todos os *pixels* da imagem, denota-se a convolução da seguinte maneira:  $h = g \otimes f$ . A Equação 3.5 mostra um exemplo de *kernel*  $5 \times 5$  para realizar a filtragem média.

$$k_{\text{média}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \div 25 \quad (3.5)$$

<sup>1</sup> As áreas de alto contraste em imagens podem ser interpretadas como bordas ou contornos de objetos.

Pratt (2007) cita um problema bastante comum quando trata desta operação. O cálculo da convolução de um *kernel* com os elementos de borda da imagem (limites da imagem) não pode ser executado exatamente como exposto na Equação 3.4, pois a função  $f$  não está definida para todos os valores explorados neste cálculo. Observe que para  $h(0,0)$ , o valor  $f(-1,-1)$  deve fazer parte do cálculo, porém, tal valor não é definido como parte da imagem.

Para superar esse problema, Pratt (2007) enumera várias soluções, dentre as quais, adicionar à função  $f$  valores que complementem a definição necessária para o cálculo. Em nosso trabalho, foram adicionados *pixels* com valor nulo para esse complemento. Ao final da convolução, os *pixels* extra são removidos do resultado. A Figura 22 ilustra a filtragem média realizada com auxílio desse processo.



**Figura 22** – Convolução de uma imagem com *kernel* da média de tamanho  $5 \times 5$ . (a) Imagem original. (b) Imagem com valores complementares. (c) Imagem final.

### 3.2.3.1 Filtragem Gaussiana

A filtragem Gaussiana é feita através da convolução da imagem com um *kernel* obtido a partir da amostragem de uma função Gaussiana. Ela é uma técnica muito usada para borrar a imagem, e conseqüentemente reduzir tanto ruído quanto pequenos detalhes. O resultado de sua aplicação é uma versão mais homogênea e suave da imagem, o que beneficia algumas estratégias de extração de características.

Em duas dimensões, a função Gaussiana pode ser representada através da Equação 3.6, a seguir:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \quad (3.6)$$

onde  $\sigma$  é uma constante que indica o desvio padrão da distribuição Gaussiana e o par  $(x, y)$  indica a posição do valor amostrado em relação à origem.

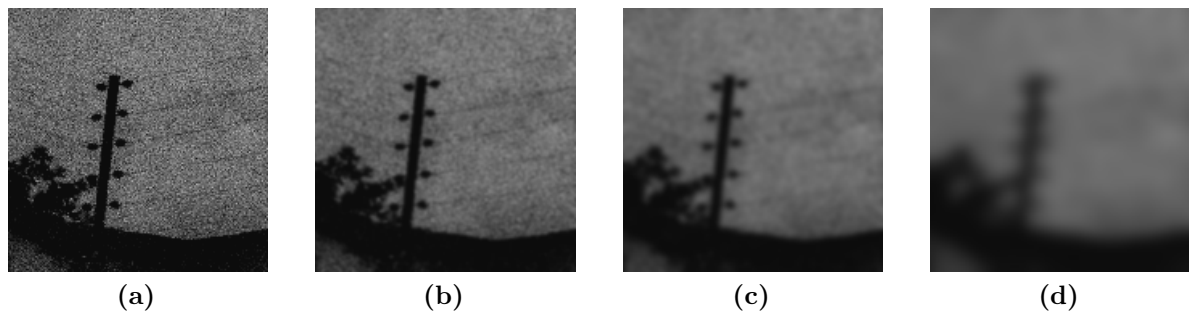
Para a construção do *kernel*, é necessário capturar apenas alguns valores da função, por isso deve-se fazer uma amostragem dela. Posteriormente, aplica-se a convolução da

imagem com o *kernel* gerado. O resultado é uma imagem suavizada de acordo com o valor  $\sigma$  da função. Valores mais elevados deixam a imagem mais borrada, enquanto que valores menores provocam menores alterações.

O *kernel* da Equação 3.7 é resultado de uma amostragem feita da função Gaussiana apresentada na Equação 3.6 com  $\sigma = 5,0$ .

$$k_{gauss} = \begin{bmatrix} 37 & 39 & 40 & 39 & 37 \\ 39 & 41 & 42 & 41 & 39 \\ 40 & 42 & 43 & 42 & 40 \\ 39 & 41 & 42 & 41 & 39 \\ 37 & 39 & 40 & 39 & 37 \end{bmatrix} \div 992 \quad (3.7)$$

A Figura 23 mostra a aplicação do *kernel* Gaussiano construído com diferentes valores de  $\sigma$  em uma imagem com forte presença de ruído.



**Figura 23** – Aplicação da filtragem Gaussiana. (a) Imagem ruidosa. (b)  $\sigma = 2,0$ . (c)  $\sigma = 5,0$ . (d)  $\sigma = 15,0$ .

### 3.2.3.2 Filtro Sobel

O filtro Sobel calcula a mudança de intensidade de um *pixel* para outro em toda a extensão da imagem. Essa variação tem direção e magnitude e é conhecida como gradiente. A filtragem é feita através da convolução de 2 *kernels* de tamanho  $3 \times 3$  em uma imagem  $f$  (Equação 3.8).

Cada *kernel* é responsável por extrair a variação de intensidade em uma direção diferente. Após a convolução deles na imagem, têm-se duas resultantes, e elas são combinadas para se obter as informações de direção e magnitude do gradiente (SOBEL, 1990).

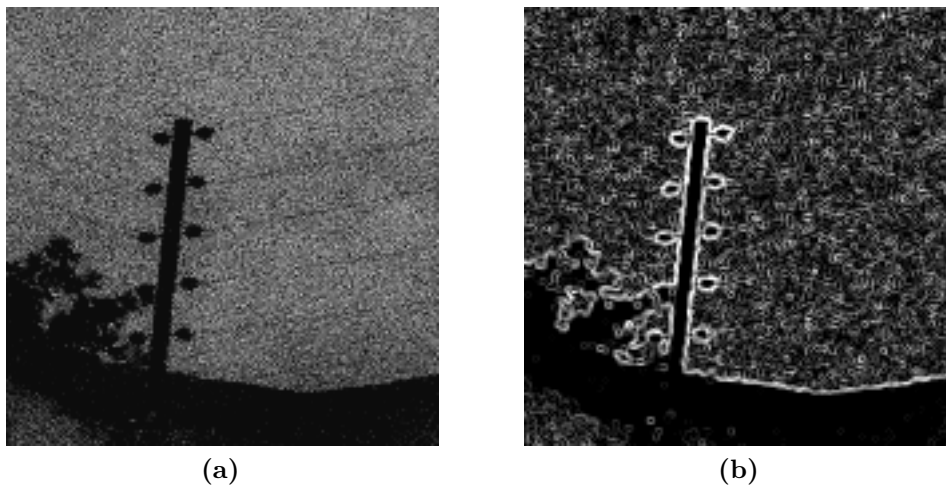
Esta operação é de grande utilidade quando o objeto do problema envolve as bordas dos elementos presentes na imagem. Elas podem ser interpretadas como regiões onde há grande variação de intensidade. Uma vez que o operador Sobel é capaz de quantificar essas variações, ele pode ser usado na detecção dessas bordas.

A Equação 3.9 mostra como são calculados tanto a magnitude  $M$  do gradiente quanto sua direção  $D$  após as operações de convolução da imagem com os *kernels* apresentados na Equação 3.8.

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \otimes f \quad G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \otimes f \quad (3.8)$$

$$M = \sqrt{G_x^2 + G_y^2} \quad D = \arctan \frac{G_y}{G_x} \quad (3.9)$$

A Figura 24 busca ilustrar o resultado da aplicação do operador Sobel. Na Figura 24b, é possível observar que as regiões mais escuras indicam que não há grandes mudanças na intensidade local. Já os *pixels* mais claros apontam maiores valores de magnitude do gradiente.



**Figura 24** – Aplicação do filtro Sobel para detecção de bordas. (a) Imagem ruidosa. (b) Magnitude do gradiente da imagem ruidosa.

Apesar de indicar as possíveis regiões com presença de bordas, o filtro Sobel não define o que de fato é borda e o que é somente sinal indesejado. Como mencionado anteriormente, os ruídos são incorporados na imagem e acabam dificultando a correta interpretação da cena. Contudo, o resultado da aplicação desse operador é bastante útil para que métodos mais sofisticados possam definir melhor as bordas da imagem.

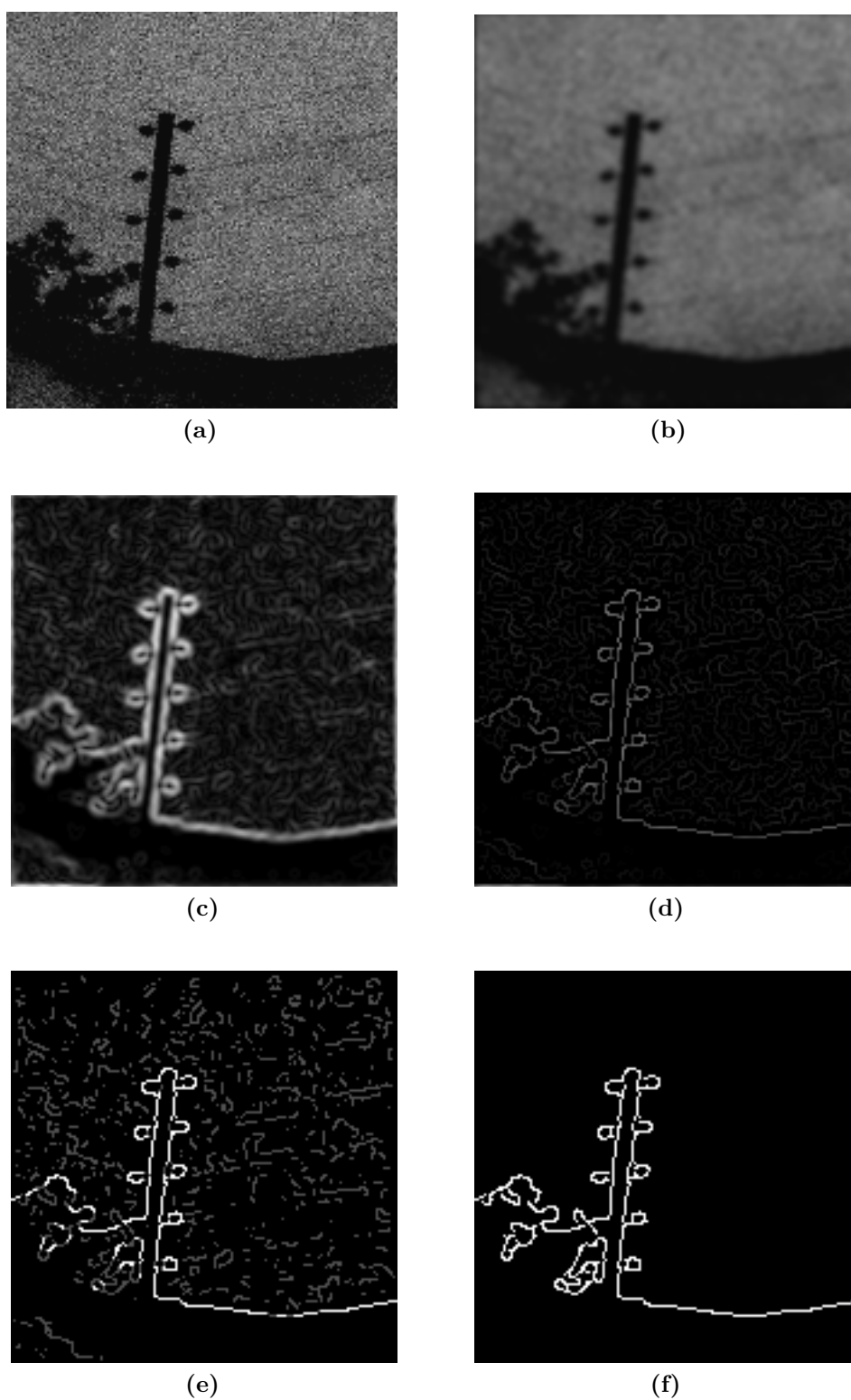
### 3.2.4 Detecção de Bordas com Canny

O método de John F. Canny para detecção de bordas é mais complexo que a convolução de filtro Sobel. Ele é dividido em 5 etapas e resulta em uma imagem binária<sup>2</sup>

<sup>2</sup> Imagem que contém apenas dois níveis de intensidade.

que contém apenas as bordas detectadas (CANNY, 1986). A seguir estão detalhadas as etapas que constituem o método de Canny.

1. **Suavização** - A primeira etapa do algoritmo de Canny (1986) consiste em realizar uma suavização da imagem. Para tal, a filtragem Gaussiana com *kernel*  $5 \times 5$  e  $\sigma = 1,5$  é frequentemente apontada como a mais adequada para o processo. A necessidade dessa etapa consiste em eliminar qualquer ruído que possa prejudicar a detecção de bordas. Os parâmetros apontados, em geral, funcionam bem, mas podem ser modificados dependendo do tipo de aplicação.
2. **Extração do Gradiente** - A partir de uma imagem suavizada e com pouco ruído, é necessário extrair tanto magnitude quanto direção do gradiente. Para tanto, utiliza-se operadores que sejam capazes de cumprir a tarefa como o filtro Sobel (1990) ou Prewitt (1970). As informações sobre o gradiente são vitais para a etapa seguinte, conhecida como supressão dos não-máximos.
3. **Supressão dos Não-Máximos** - A etapa de supressão dos não-máximos visa eliminar grande parte das diferenças de intensidade capturadas pelo operador, deixando somente a região que pode ser entendida como centro da borda. Nessa etapa, são utilizados tanto direção como magnitude do gradiente. Os *pixels* considerados não-máximos são suprimidos (i.e. sua intensidade fica igual a zero) e já não são considerados *pixels* de borda.
4. **Dupla Limiarização** - Esta etapa objetiva classificar os *pixels* em 3 grupos de acordo com a intensidade. O grupo 1 corresponde aos *pixels* de borda forte e é composto por todos aqueles que têm intensidade acima de um valor pré-definido chamado limiar superior. O grupo 2 compreende *pixels* cuja intensidade varia entre o limiar superior e o limiar inferior (também pré-definido). Estes são considerados de borda fraca. Do grupo 3 fazem parte *pixels* que estão abaixo do limiar inferior e estes são suprimidos. Os limiares mencionados são valores de intensidade determinados empiricamente.
5. **Rastreamento de Bordas** - A última fase é uma análise dos *pixels* de borda fraca. Aqueles que forem adjacentes, mesmo que indiretamente, a *pixels* de borda forte, devem ser movidos para o grupo 1. Os outros são suprimidos da imagem final. Ao término do processo, restam apenas *pixels* de borda forte e *pixels* suprimidos compondo uma imagem binária com as bordas detectadas. A Figura 25 mostra a aplicação do método em todas as suas etapas.



**Figura 25** – Aplicação do método de Canny para a detecção de bordas. (a) Imagem ruidosa. (b) Suavização. (c) Extração do gradiente. (d) Supressão dos não-máximos. (e) Dupla limiarização. (f) Rastreamento de bordas.



### 3.3 Análises Estatísticas

A metodologia leva em consideração alguns conceitos estatísticos já consagrados na literatura. Além de serem primordiais para a nossa proposta, alguns deles são explorados, também, nas etapas de avaliação da solução. Esta seção define algumas convenções e técnicas estatísticas essenciais exploradas em nosso trabalho. Outras definições relevantes estão descritas no Anexo A.

#### 3.3.1 Notação de Variáveis

São abordados na literatura conceitos que, frequentemente, misturam a ideia de variável com vetor de variáveis. Como as variáveis, em nosso trabalho, são majoritariamente multidimensionais, é bastante comum a utilização da terminologia variável para expressar, na realidade, o que se entende por vetor de variáveis. Ademais, outro conceito que pode causar ambiguidade é o de observação de variável, também referido neste trabalho por amostra de variável.

Para expressar a  $j$ -ésima variável ou coordenada é utilizada uma letra maiúscula com o indicador  $j$ . A  $i$ -ésima observação ou amostra da  $j$ -ésima variável denota-se pela respectiva letra minúscula com os indicadores da amostra e da variável, enquanto que um vetor da  $i$ -ésima observação de todas as variáveis se denota por  $\vec{v}_i$  (caso se refira ao vetor de todas as variáveis, a letra deve ser maiúscula)<sup>3</sup>.

A Equação 3.10 mostra como deve ser entendida essa notação. A matriz apresenta todo o conjunto de valores observados. Cada coluna indica todos os valores observados para uma variável  $V_j$ . Ao passo que cada linha mostra um vetor de valores  $\vec{v}_i$  observado para todas as variáveis.

$$\begin{aligned} \vec{V} &= \begin{matrix} V_1 & V_2 & \dots & V_m \end{matrix} \\ \vec{v}_1 &= \begin{pmatrix} v_{11} & v_{12} & \dots & v_{1m} \end{pmatrix} \\ \vec{v}_2 &= \begin{pmatrix} v_{21} & v_{22} & \dots & v_{2m} \end{pmatrix} \\ \vdots & \\ \vec{v}_n &= \begin{pmatrix} v_{n1} & v_{n2} & \dots & v_{nm} \end{pmatrix} \end{aligned} \quad (3.10)$$

Outras notações, para finalidades mais específicas, são definidas ao longo desta seção.

<sup>3</sup> Neste contexto, a dimensão do vetor (i.e. a quantidade de variáveis que compõem o vetor) é expressa pela letra **m**, enquanto que a quantidade de observações ou amostras de um conjunto é denotada sempre por **n**.

### 3.3.2 Análise de Componentes Principais (PCA)

Este é um dos métodos estatísticos mais simples para análise de dados. O PCA é uma operação que transforma um conjunto de observações de variáveis para um novo sistema de coordenadas. Neste novo sistema, a projeção dos dados apresenta a maior variância observada sob a primeira coordenada, a segunda maior variância fica ao longo da segunda coordenada e assim sucessivamente. A aplicação do PCA depende, além dos conceitos que já foram apresentados nesta seção e no Anexo A, das noções de autovalor e autovetor.

#### 3.3.2.1 Autovalor e Autovetor

Na álgebra linear, o autovalor é um escalar, geralmente denotado por  $\lambda$ , de uma matriz quadrada  $A$  se existir um vetor  $\vec{v} \neq 0$  tal que:

$$A\vec{v} = \lambda\vec{v}. \quad (3.11)$$

Nesse caso,  $\vec{v}$  é chamado de autovetor relativo ao autovalor  $\lambda$ . Tal valor é capaz de transformar a matriz  $A$  em uma matriz singular (que não pode ser invertida) quando sua diagonal principal é subtraída de  $\lambda$ . Uma consequência importante da existência dos autovalores e autovetores é que a matriz  $A$  representa uma transformação incapaz de alterar a direção do vetor  $\vec{v}$ . Ela pode apenas mudar sua magnitude. Esses conceitos são essenciais para a aplicação do PCA.

#### 3.3.2.2 Computando as Componentes Principais

A análise das componentes principais necessária para nosso trabalho pode ser realizada em cinco passos (COOTES; TAYLOR, 2004):

1. Calcula-se a média  $\vec{\mu}$  das formas do conjunto de dados, como exposto na Seção A.1;
2. Calcula-se a matriz de covariância  $S$  das formas, como na Seção A.3;
3. Calcula-se os autovetores  $\vec{p}_i$  e autovalores correspondentes  $\lambda_i$  da matriz  $S$ , ordenados tal que  $\lambda_i \geq \lambda_{i+1}$ ;
4. Calcula-se a variância total  $\sigma_T^2 = \sum \lambda_i$ ;
5. Escolhe-se os  $t$  maiores autovalores tal que:

$$\sum_{i=1}^t \lambda_i \geq f_v \sigma_T^2, \quad (3.12)$$

onde  $f_v$  é uma proporção da variância total que se deseja expressar.

Os autovetores que expressam as componentes principais do conjunto observado são o resultado final dessa análise, tal como seus respectivos autovalores. Para facilitar os cálculos posteriores, os autovetores são organizados em uma matriz  $P$  tal que:

$$P = (\vec{p}_1 | \vec{p}_2 | \dots | \vec{p}_t). \quad (3.13)$$

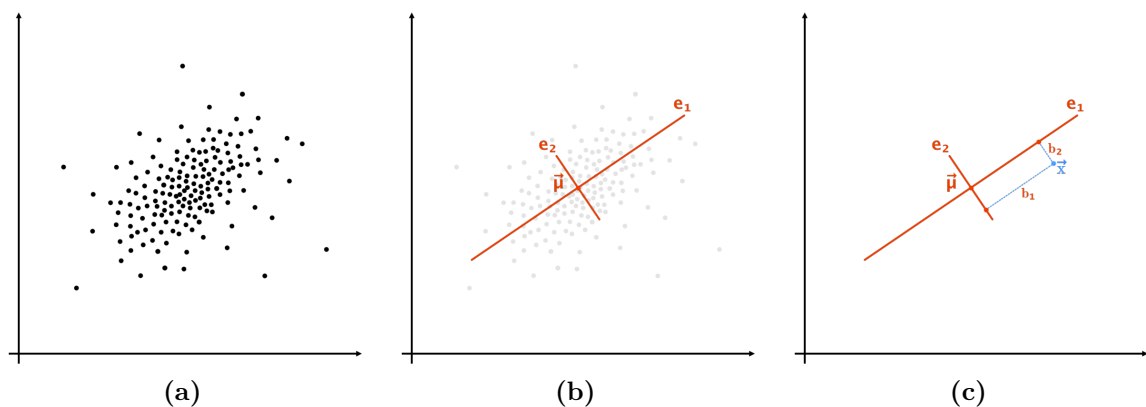
Um dado qualquer pode ser aproximado do conjunto usado na análise com base na Equação 3.14,

$$\vec{x} \approx \vec{\mu} + P\vec{b}, \quad (3.14)$$

onde  $\vec{b}$  é um vetor de parâmetros que indica a distância do vetor que se deseja aproximar para cada componente analisada e pode ser calculado através da Equação 3.15,

$$\vec{b} = P^T(\vec{x} - \vec{\mu}). \quad (3.15)$$

A Figura 26 ilustra a aplicação do PCA em um conjunto de dados com duas variáveis. No gráfico da Figura 26a, estão dispostas várias observações bidimensionais. Em 26b estão ilustradas a média  $\vec{\mu}$  dessas amostras e dois eixos ortogonais  $e_1$  e  $e_2$  que representam, respectivamente, a primeira e segunda componentes em variância dos dados observados. A Figura 26c mostra o papel do parâmetro  $\vec{b}$  na reconstrução de um dado do conjunto. Os valores de  $\vec{b}$  representam as coordenadas do conjunto (amplitude em cada eixo) após a aplicação do PCA.



**Figura 26** – Resultado da aplicação da análise de componentes principais em conjunto de dados com duas variáveis. (a) Distribuição de vetores bidimensionais. (b) Eixos ortogonais de maior variância com média  $\vec{\mu}$ . (c) Amostra  $\vec{x}$  reconstruída com  $\vec{b}$ .

### 3.3.3 Aferição de Resultados

A medição do tipo de resultados da aplicação da metodologia proposta conta com alguns precedentes na literatura. [Belhumeur et al. \(2011\)](#) faz uma comparação, através da distância Euclidiana, entre as marcações automáticas de sua proposta com uma média de três rotulações manuais para indicar que o método pode ser mais preciso até mesmo que a marcação de um especialista<sup>4</sup>. Além disso, seu trabalho relativiza o erro absoluto entre pontos, tendo como base a distância entre as pupilas (tal proposta também foi abordada por [Milborrow \(2007\)](#), [Valstar et al. \(2010\)](#) e [Martinez et al. \(2013\)](#)).

Contudo, como explanado por [Zhu e Ramanan \(2012\)](#), nem sempre é possível medir a distância entre os olhos, visto que a perspectiva avaliada pode não permitir tal medição (a perspectiva do perfil é um exemplo disso). Como alternativa, eles propuseram o uso da média entre a altura e largura da face para normalizar a distância absoluta.

Em nosso trabalho, foi utilizada a distância entre os olhos para relativizar as marcações frontais. No caso da perspectiva do perfil, optou-se pela abordagem de [Zhu e Ramanan \(2012\)](#). Neste trabalho, a diferença entre uma marcação manual e uma marcação apontada pela metodologia é referida, também, como erro e sempre é calculada a partir da distância Euclidiana. A Equação 3.16 mostra o cálculo do erro de um ponto marcado manualmente  $(x_1, y_1)$  para outro marcado de forma automática  $(x_2, y_2)$ .

$$D_E = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (3.16)$$

## 3.4 Rasterização para Coleta de Textura

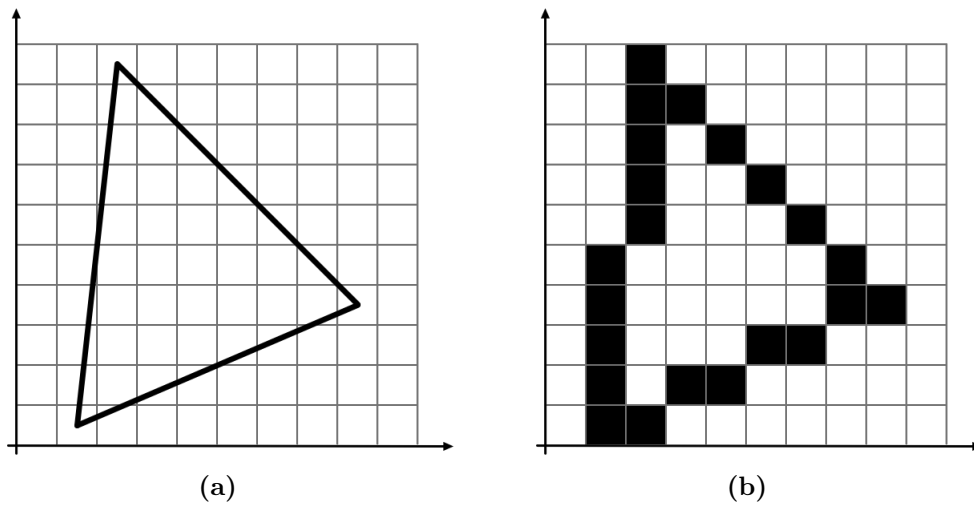
Em linhas gerais, um perfil de textura é um subconjunto de amplitudes de uma função  $f(x, y)$  que representa uma imagem digital. Em algumas ocasiões, este perfil deve corresponder a um segmento de reta cuja coleta deve ser realizada a partir de uma função do tipo  $f(x) = mx + b$  que determina uma reta. Contudo, a função deve ser aproximada para um domínio discreto de modo a possibilitar a coleta do perfil.

Uma alternativa para realizar essa aproximação é o uso de um conhecido algoritmo de rasterização<sup>5</sup>. O algoritmo de [Bresenham \(1965\)](#) descreve como uma reta pode ser aproximada e representada em uma imagem digital.

Foi feita uma pequena alteração no algoritmo original para que, ao invés de marcar os *pixels* da reta, ele retorne diretamente o perfil de textura desejado. A Figura 27 exemplifica a atuação do algoritmo em uma forma triangular.

<sup>4</sup> Não raro, há divergências entre marcações dos próprios especialistas.

<sup>5</sup> Processo que consiste em escolher dentre os *pixels* de uma imagem aqueles que melhor aproximam a representação de um gráfico vetorial.



**Figura 27** – Rasterização de figura geométrica. (a) Triângulo formado por 3 segmentos de reta. (b) Resultado do algoritmo de [Bresenham \(1965\)](#) para rasterização do triângulo.

A modificação realizada no método original está expressa no Algoritmo 1. É importante frisar que esta definição funciona apenas para pontos pertencentes ao octante 0 do plano cartesiano. Contudo, a generalização pode ser feita a partir de transformações nas coordenadas do plano. O algoritmo tem como entradas a imagem  $f(x, y)$ , e os pontos  $(x_1, y_1)$   $(x_2, y_2)$  que definem o segmento de reta. O resultado de sua aplicação é um vetor com as intensidades da imagem coletadas ao longo do segmento.

---

**Algoritmo 1:** Algoritmo para coletar textura ao longo de segmento de reta em uma imagem.

---

**Entrada:**  $f(x, y)$ ,  $(x_1, y_1)$ ,  $(x_2, y_2)$   
**Saída:**  $\vec{tx}$

```

1 início
2    $\Delta x \leftarrow x_2 - x_1$ 
3    $\Delta y \leftarrow y_2 - y_1$ 
4    $\Delta D \leftarrow \Delta y - \Delta x$ 
5    $y \leftarrow y_1$ 
6   para cada  $x \in [x_1, x_2 - 1]$  faça
7      $(a_x, b_x) \leftarrow f(x, y)$ 
8     se  $\Delta D \geq 0$  então
9        $y \leftarrow y + 1$ 
10       $\Delta D \leftarrow \Delta D - \Delta x$ 
11    fim
12     $\Delta D \leftarrow \Delta D + \Delta y$ 
13  fim
14   $\vec{tx} \leftarrow [(a_{x_1}, b_{x_1}) \dots (a_{x_2-1}, b_{x_2-1})]$ 
15 fim
16 retorna  $\vec{tx}$ 

```

---

## 3.5 Considerações Finais

Este capítulo expôs os materiais e métodos que foram explorados desde o projeto da metodologia, passando por seu desenvolvimento e culminando com a avaliação dos resultados obtidos com a proposta. Nele, foram mostrados o conjunto de pontos de referência necessários para a realização de análises estéticas, expostos alguns métodos e técnicas necessários para a concretização da etapa de localização e identificação automática dos pontos de referência, e detalhados aspectos relativos à avaliação dos resultados.

O próximo capítulo complementa o arcabouço de técnicas necessárias para a solução do problema. Nele, o método *Active Shape Models* elaborado por (COOTES et al., 1995) é explicado em detalhes.

## 4 Active Shape Models

Um problema bastante recorrente no âmbito da análise de imagens e visão computacional é localizar, em uma cena, objetos rígidos e com forma bem definida. Algumas abordagens buscam resolver esse problema através da localização das bordas dos objetos. [Kass, Witkin e Terzopoulos \(1988\)](#) propuseram um método para detecção da forma a partir do contorno dos objetos. Esta abordagem baseia-se em um modelo deformável que procura se encaixar nas bordas.

Entretanto, como apontado por [Cootes et al. \(1995\)](#), existe uma deficiência neste tipo de metodologia. O problema relatado reside na tentativa de priorizar o encaixe do modelo em detrimento da forma correspondente. Isso resulta, frequentemente, na localização de um objeto que não corresponde à forma proposta, pois o modelo não respeita restrições pré-definidas da forma.

Foi então que [Cootes et al. \(1995\)](#) propuseram o *Active Shape Models*, que são capazes de preservar a forma proposta para se localizar um determinado objeto. A grande vantagem em relação a outras metodologias é que a aplicação do método só permite que o modelo assuma modos de variação compatíveis com os observados em um conjunto de treinamento. Assim, o método preserva a forma geral do modelo enquanto busca o encaixe na imagem.

Este capítulo explica o funcionamento do ASM tradicional proposto por [Cootes et al. \(1995\)](#) desde a construção do modelo até a fase de busca pela forma. A proposta deles é de grande importância na solução do problema de localização de pontos frontais. Nesta dissertação, algumas modificações foram realizadas na tentativa de melhorar a precisão do método para o problema proposto. Essas modificações estão detalhadas no Capítulo 5.

### 4.1 Definições Preliminares

Para compreender o algoritmo ASM, alguns conceitos preliminares precisam ser apresentados. Esta seção é essencial para o entendimento do algoritmo pois define não somente conceitos relevantes como, também, mostra a forma como eles são representados.

#### 4.1.1 Forma

No âmbito do ASM, a forma deve ser encarada como uma sequência de pontos bidimensionais. Tais pontos devem fazer parte do contorno da forma e serem escolhidos com base em características dependentes da aplicação em questão. [Cootes et al. \(1995\)](#) relata que eles podem ser definidos, também, através da interpolação de pontos previamente

marcados. A forma pode ser denotada como na Equação 4.1, onde  $(x_{ij}, y_{ij})$  representa o  $j$ -ésimo ponto de referência da  $i$ -ésima forma observada em um determinado conjunto:

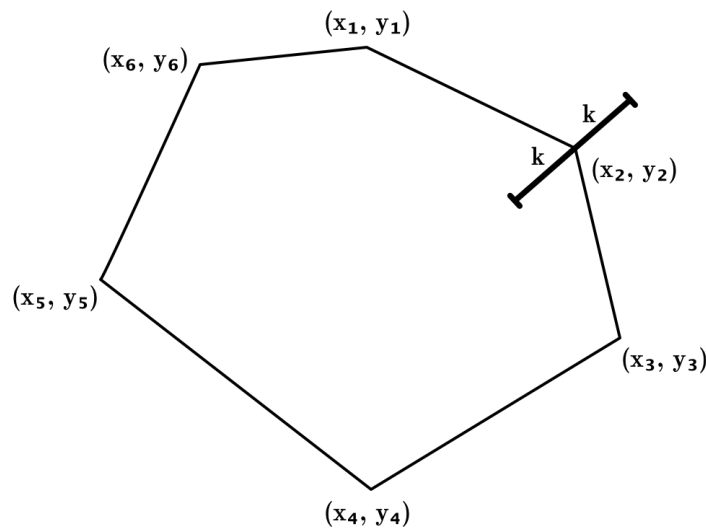
$$\vec{f}_i = [(x_{i1}, y_{i1}) \quad (x_{i2}, y_{i2}) \dots (x_{im}, y_{im})]. \quad (4.1)$$

Para simplificar os cálculos realizados com a forma, é mais comum expressá-la de acordo com a Equação 4.2,

$$\vec{f}_i = [x_{i1} \quad y_{i1} \quad x_{i2} \quad y_{i2} \quad \dots \quad x_{im} \quad y_{im}]^T. \quad (4.2)$$

### 4.1.2 Perfil de Textura

Como destacado na Seção 3.4, existe a necessidade de se identificar um perfil de textura para cada ponto de referência. Ele deve ser unidimensional, normal à forma e ter um tamanho  $2k + 1$ , onde  $k \in \mathbb{N}^+$  representa o tamanho da amostra tirada a partir do ponto estendendo-se para os lados interno e externo da forma. A Figura 28 ilustra como deve ser coletado o perfil de um ponto.



**Figura 28** – Exemplo de coleta do perfil de textura do ponto  $(x_2, y_2)$  em forma composta por 6 pontos.

Para expressar o perfil de textura do  $j$ -ésimo ponto da  $i$ -ésima forma de um conjunto utiliza-se a notação da Equação 4.3:

$$\vec{tx}_{ij} = [tx_{ij1} \quad tx_{ij2} \quad \dots \quad tx_{ij(2k+1)}]^T. \quad (4.3)$$



Uma vez que o perfil coincide com um subconjunto da imagem  $f(x, y)$ , ele também pode ser expresso como na Equação 4.4,

$$\vec{t}x_{ij} = [f(x_{ij1}, y_{ij1}) \quad f(x_{ij2}, y_{ij2}) \quad \dots \quad f(x_{ij(2k+1)}, y_{ij(2k+1)})]^T. \quad (4.4)$$

### 4.1.3 Resumo do ASM

Em linhas gerais, *Active Shape Models* é um método capaz de localizar uma forma rígida em uma imagem através de seu contorno. Como apontado na Seção 2.2, ele pode ser dividido em duas etapas: construção do modelo e busca pela forma.

Dado um conjunto de formas  $\vec{F}$ , anotadas manualmente por um especialista, e os perfis de textura  $\vec{T}x$  dos pontos que as compõem, a construção do modelo da forma consiste na elaboração de um modelo de distribuição de pontos (PDM<sup>1</sup>), que corresponde ao conjunto alinhado, e no cálculo das médias  $\vec{t}x_{ij}$  dos perfis de textura. A análise de componentes principais deve ser aplicada ao PDM para extração dos modos de variação, resultando em uma matriz de autovetores  $P_M$  e um vetor de autovalores  $\vec{\lambda}$  usados para reconstruir uma determinada forma.

Dados o modelo da forma e uma imagem  $f(x, y)$  onde deseja-se localizar a forma do modelo, a busca pode ser dividida em duas etapas:

1. gerar uma nova localização para a forma através da comparação entre os perfis de textura extraídos da imagem  $f(x, y)$  com as médias  $\vec{t}x_{ij}$  dos perfis de textura do modelo;
2. ajustar a nova localização da forma de acordo com as restrições  $\vec{\lambda}$  do modelo.

Estes passos são aplicados alternadamente até que haja convergência na busca. A construção do modelo e a busca pela forma estão detalhadas nas Seções 4.2 e 4.3.

## 4.2 Construção do Modelo

O objetivo do método é localizar objetos cujas formas sejam coerentes com o conjunto usado no treinamento do modelo. Em outras palavras, o algoritmo busca generalizar o objeto através de sua forma.

O primeiro passo para possibilitar essa generalização é a construção do modelo de distribuição dos pontos da forma. Essa estratégia de construção do modelo descreve a geometria da forma visada e captura os modos de variação que ela pode assumir dentro

<sup>1</sup> *Point Distribution Model* proposto por Cootes et al. (1995).

do conjunto de treinamento. Neste caso, faz-se o uso da análise de componentes principais (Seção 3.3.2).

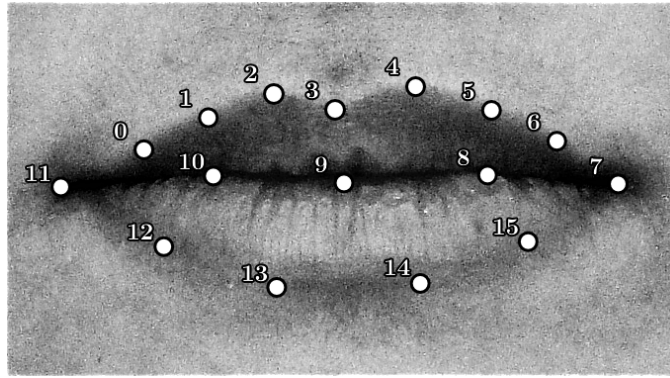
Para a construção do modelo, existe a necessidade que um especialista realize a marcação manual dos pontos que compõem a forma do objeto alvo. Mas antes de se iniciar o processo de rotulação, é importante que os pontos a serem marcados estejam bem definidos. Esse cuidado preserva a consistência do modelo uma vez que cada ponto deve apresentar um significado único na forma. Marcações imprecisas oriundas da má definição dos pontos ou de descuido do especialista costumam introduzir aspectos não-lineares no modelo. O resultado disso é que ele passa a descrever formas incompatíveis com os dados observados no conjunto de treinamento (BOWDEN; MITCHELL; SARHADI, 2000). As etapas de construção do modelo brevemente listadas a seguir serão discutidas em detalhes posteriormente:

1. **Marcação de imagens** - Nesta etapa, o conjunto de treinamento deve ter os pontos corretamente localizados por um especialista;
2. **Extração do perfil dos pontos** - Com as imagens marcadas, é necessário identificar o perfil de textura de cada ponto, que deve ser usado como critério de busca;
3. **Cálculo dos pesos** - Com relação à forma, cada ponto possui um peso específico. Eles devem ser calculados com base na variação de posição dentro da base de imagens;
4. **Alinhamento do conjunto** - Antes da captura dos modos de variação, é fundamental que as formas marcadas estejam alinhadas. Esta etapa deve levar em consideração o peso dos pontos na construção da forma média;
5. **Captura dos modos de variação** - Nesta etapa, deve-se usar o conjunto alinhado para identificar as possíveis variações que a forma pode sofrer.

#### 4.2.1 Marcação de Imagens

A primeira preocupação para a construção do modelo é a escolha dos pontos que fazem parte da forma. Uma vez que o método busca formas rígidas, a escolha dos pontos deve ser realizada levando-se em conta o contraste deixado pelo objeto. Logo, é necessário definir pontos que se encontram no contorno dos elementos. A Figura 29 ilustra um exemplo de marcação de pontos para a boca.

Como apontado na Seção 3.1, existem pontos que necessariamente devem compor o modelo. No caso ilustrado na Figura 29 alguns deles não são anatômicos mas ajudam na construção. Para o especialista, é importante observar que os pontos devem ser marcados de acordo com a descrição em todas as imagens. Por exemplo, na Figura 29 o ponto 9 é descrito anatomicamente como sendo o ponto onde os lábios se encontram no meio da face.



**Figura 29** – Exemplo de escolha de pontos para definir a forma da boca. Os pontos 2, 3, 4, 7, 9, 11, 13 e 14 são os pontos principais que definem a forma da boca. Os outros são pontos usados para preencher a forma e estão posicionados de modo equidistante aos pontos adjacentes no contorno.

Entretanto, os pontos 0, 1, 5 e 6 não possuem uma descrição formal que pode ser transferida para outro exemplo sem que haja prejuízo na marcação. Na realidade, a localização deles é dependente de outros pontos e da textura ao redor (valores dos *pixels* ao redor). É possível observar na imagem, que eles se encontram no contorno da boca e estão separados dos pontos adjacentes na borda por distâncias aproximadamente iguais, assim eles são definidos.

Definidos os pontos que irão compor a forma alvo do modelo, um especialista deve marcá-los manualmente em todas as imagens do conjunto de treinamento. O resultado desta etapa são  $n$  formas  $\vec{f}_i$  definidas manualmente que serão utilizadas nas fases mais agudas da construção.

#### 4.2.2 Extração do Perfil dos Pontos

Esta etapa de construção do modelo não diz respeito à forma do conjunto de pontos, mas sim à textura ao redor deles. A forma que o modelo assume é capturada a partir da posição  $(x, y)$  de cada ponto na imagem digital. Contudo, para que o modelo possa se ajustar a um novo objeto, é preciso extrair mais informação da imagem do que somente a localização de cada ponto. Neste caso, é necessário identificar um subconjunto de amplitudes da imagem  $f(x, y)$ . Este subconjunto pode ser entendido como uma faixa de *pixels* que passa pelo ponto. Ela é usada para determinar qual deve ser a nova posição daquele ponto a cada iteração do algoritmo de busca.

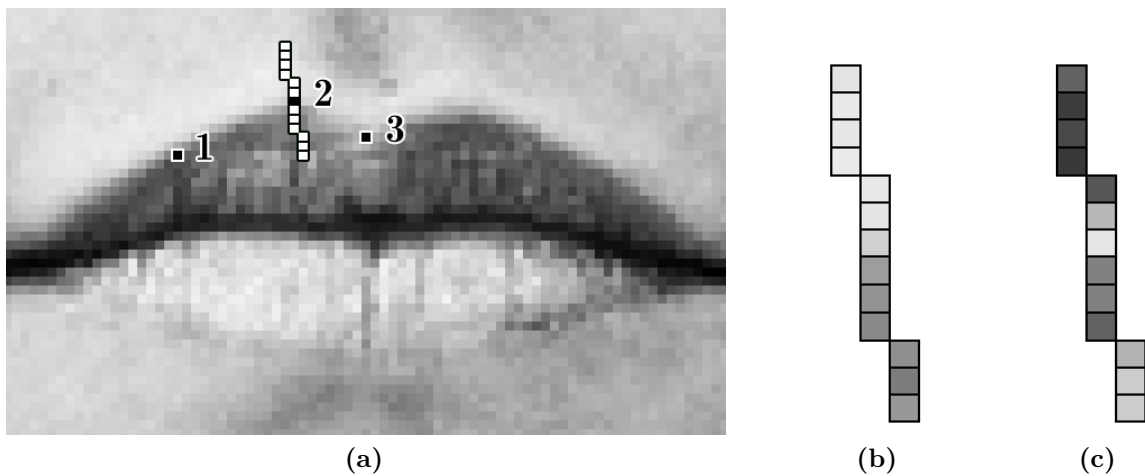
A proposta de [Cootes et al. \(1995\)](#) indica que o perfil de textura do ponto deve ser coletado como apontado na Seção 4.1.2. A Figura 30 exemplifica a extração do perfil do ponto 2. A imagem em questão tem dimensão  $250 \times 75$  *pixels*, e o perfil retirado possui tamanho  $2k + 1 = 13$  *pixels*.

A transformação dos valores contínuos do segmento de reta em valores discretos do

perfil pode ser realizada como descrito na Seção 3.4. Após a extração, é recomendado que se faça uma normalização das diferenças de intensidade entre os *pixels* do perfil para amenizar discrepâncias de iluminação comuns nas imagens da base. Ela é realizada dividindo-se esses valores  $tx_{ijl}$  do perfil pela soma  $N_{tx_{ij}}$  de todas as diferenças de intensidades. As Equações 4.5 e 4.6 mostram como o cálculo pode ser realizado:

$$N_{tx_{ij}} = \sum_{l=1}^{2k+1} tx_{ijl}, \quad (4.5)$$

$$N(\vec{tx}_{ij}) = \frac{[tx_{ij1} \quad tx_{ij2} \quad \dots \quad tx_{ij(2k+1)}]^T}{N_{tx}}. \quad (4.6)$$



**Figura 30** – Exemplo de extração do perfil de textura. (a) Segmento determinado pelos pontos da forma. (b) Perfil de textura coletado. (c) Perfil de textura com diferenças de intensidade normalizadas.

Esse processo deve ser feito para todos os pontos de todas as formas do conjunto de treinamento. Ao final, deve-se calcular a média dos perfis de cada ponto através de uma média aritmética simples com seus respectivos perfis extraídos de cada imagem e normalizados. A Equação 4.7 mostra como é calculada a média dos perfis de textura  $\vec{\mu}_{txj}$  do  $j$ -ésimo ponto em um conjunto com  $n$  amostras.

$$\vec{\mu}_{txj} = \frac{1}{n} \sum_{i=1}^n N(\vec{tx}_{ij}) \quad (4.7)$$

Muitos trabalhos relacionados procuram melhorar o desempenho do ASM tradicional neste aspecto. Milborrow (2007) propôs escolher perfis bidimensionais para alguns pontos e Du et al. (2008) optou por transformar a busca de novas posições para o ponto em um problema de classificação. Em nosso trabalho, uma outra alternativa foi elaborada para a busca, e a extração do perfil introduz uma alteração no ASM tradicional. Esses detalhes são discutidos no Capítulo 5.

### 4.2.3 Cálculo dos Pesos

A construção do modelo deve levar em consideração que cada ponto possui um peso específico. A Equação 4.8 mostra como se calcula a variância  $\sigma_{R(k,l)}^2$  em um conjunto com  $n$  amostras. Nela,  $R(k, l)_i$  representa a distância euclidiana do ponto  $k$  para o  $l$  observada na amostra  $i$  e  $\mu_{R(k,l)}$  é a distância média de  $k$  para  $l$  consideradas todas as formas. A Equação 4.9 mostra como se calcula o peso  $w_k$  do  $k$ -ésimo ponto do modelo.

$$\sigma_{R(k,l)}^2 = \frac{1}{n} \sum_{i=1}^n (R(k, l)_i - \mu_{R(k,l)})^2 \quad (4.8)$$

$$w_k = \left( \sum_{l=1}^m \sigma_{R(k,l)}^2 \right)^{-1} \quad (4.9)$$

Da equação, depreende-se que o peso é inversamente proporcional à variância das distâncias relativas a um ponto, ou seja, quanto menor a variação de um ponto dentro do conjunto, maior seu peso, utilizado na etapa seguinte: alinhamento das formas.

### 4.2.4 Alinhamento das Formas

Com os pesos definidos, é iniciado o alinhamento das  $n$  formas  $\vec{f}_i$  marcadas anteriormente. Esse processo tem como objetivo posicioná-las de maneira a minimizar as distâncias entre pontos equivalentes nas amostras ponderados por seus respectivos pesos. As formas devem sempre ser alinhadas duas a duas, por isso, primeiramente é necessário ter noção do alinhamento de duas formas quaisquer do conjunto.

Esse processo consiste em determinar as transformações que devem ser aplicadas em uma forma  $\vec{f}_b$  para que ela se posicione o mais próximo de uma outra  $\vec{f}_a$ . As transformações incluem escala  $s$ , rotação  $\theta$  e translação  $\vec{t}$ .

A Equação 4.10 mostra como o ponto  $(x'_{ij}, y'_{ij})$  é calculado partir da escala  $s$  e rotação  $\theta$  do ponto  $(x_{ij}, y_{ij})$ . Uma notação para expressar as transformações de escala e rotação aplicadas a todos os pontos de uma forma  $\vec{f}_i$  é expressa na Equação 4.11. Para não distorcer a forma, todos os pontos devem sofrer a mesma translação em cada coordenada como detalhado na Equação 4.12. Os pesos de cada ponto devem ser organizados em um matriz como exposto na Equação 4.13. A minimização de  $E_b$ , na Equação 4.14, realiza o alinhamento pretendido.

$$\begin{bmatrix} x'_{ij} \\ y'_{ij} \end{bmatrix} = M(s, \theta) \begin{bmatrix} x_{ij} \\ y_{ij} \end{bmatrix} = \begin{bmatrix} (s \cos \theta)x_{ij} - (s \sin \theta)y_{ij} \\ (s \sin \theta)x_{ij} + (s \cos \theta)y_{ij} \end{bmatrix} \quad (4.10)$$

$$M(s, \theta)[\vec{f}_i] = [x'_{i1} \quad y'_{i1} \quad x'_{i2} \quad y'_{i2} \quad \dots \quad x'_{im} \quad y'_{im}]^T \quad (4.11)$$

$$\vec{t} = [t_x \ t_y \ t_x \ t_y \ \dots \ t_x \ t_y]^T \quad (4.12)$$

$$W = \begin{bmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_m \end{bmatrix} \quad (4.13)$$

$$E_b = (\vec{f}_a - M(s_b, \theta_b)[\vec{f}_b] - \vec{t}_b)^T W (\vec{f}_a - M(s_b, \theta_b)[\vec{f}_b] - \vec{t}_b) \quad (4.14)$$

Aplicando o método dos mínimos quadrados para minimizar  $E_b$ , os valores de  $s$ ,  $\theta$  e  $\vec{t}$  podem ser encontrados a partir da solução de um sistema de equações lineares.

O alinhamento do conjunto requer que todas as formas sejam alinhadas duas a duas com base na Equação 4.14. Primeiramente, alinha-se todo o conjunto com uma forma  $\vec{f}_j$  selecionada arbitrariamente. Depois calcula-se a média de todas as formas através da Equação 4.15. Alinha-se a média com a forma  $\vec{f}_j$  para garantir convergência (caso contrário, a média iria encolher infinitamente). Esse processo se repete até que o conjunto não sofra grandes alterações. O Algoritmo 2 mostra como deve ser feito o processo. Ele tem como entrada um conjunto de formas desalinhadas  $C_{formas}$ , e tem como saídas a média das formas  $\vec{\mu}_f$  e o mesmo conjunto com as formas alinhadas. O conjunto alinhado, conhecido como modelo de distribuição de pontos, está ilustrado na Figura 31.

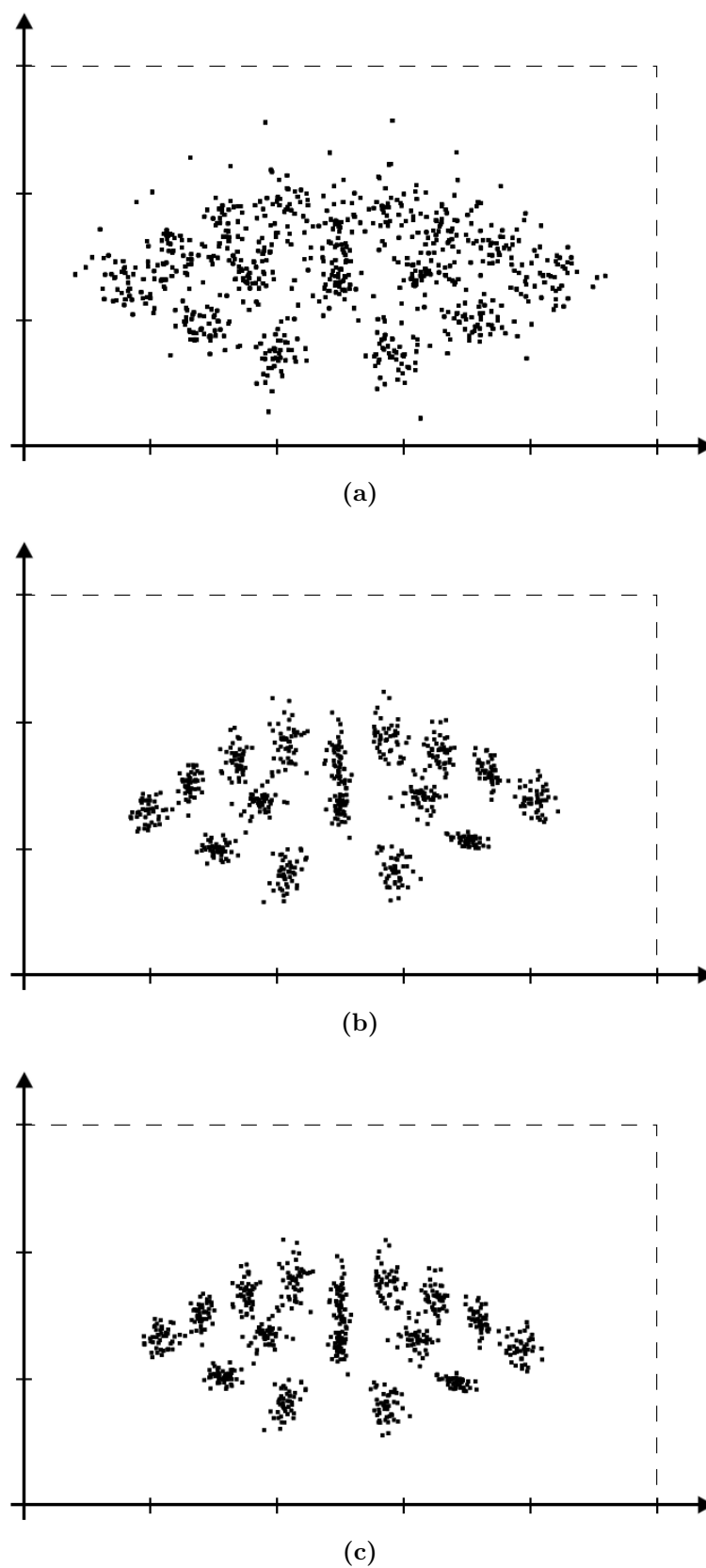
$$\vec{\mu}_f = \frac{1}{n} \sum_{i=1}^n \vec{f}_i \quad (4.15)$$

O resultado das formas alinhadas é especialmente importante para a última etapa de construção do modelo: a determinação de todos os modos de variação que a média das formas pode sofrer.

### 4.2.5 Captura dos Modos de Variação

Esta etapa representa um aspecto chave do ASM. Diferente de alguns outros modelos deformáveis, este método tem a capacidade de preservar a média das formas observadas no conjunto de treinamento. Ao identificar as possíveis variações que o modelo sofre, é possível adaptar uma forma modificada para que ela seja compatível com o conjunto observado.

Para possibilitar a captura dos modos de variação da forma, é preciso que seja feita uma análise estatística do modelo de distribuição de pontos. Cootes et al. (1995) determina que seja feita uma análise das componentes principais do PDM.



**Figura 31** – Alinhamento de um conjunto de bocas marcadas com os 16 pontos ilustrados na Figura 29. (a) Distribuição inicial dos pontos. (b) Distribuição após algumas iterações do algoritmo. (c) Distribuição final (PDM).

---

**Algoritmo 2:** Algoritmo de alinhamento de formas do conjunto de treinamento.

---

**Entrada:**  $C_{formas}$   
**Saída:**  $\vec{\mu}_f, C_{formas}$

```

1 início
2   para cada  $\vec{f}_i \in C_{formas}$  faça
3     | Alinhar  $\vec{f}_i$  com  $\vec{f}_0$ 
4   fim
5   repita
6     | Calcular  $\vec{\mu}_f$ 
7     | Alinhar  $\vec{\mu}_f$  com  $\vec{f}_0$ 
8     | para cada  $\vec{f}_i \in C_{formas}$  faça
9       | Alinhar  $\vec{f}_i$  com  $\vec{\mu}_f$ 
10    fim
11  até haver convergência;
12 fim
13 retorna  $\vec{\mu}_f, C_{formas}$ 

```

---

Como exposto na Seção 3.3.2, o PCA realiza uma transformação que leva um conjunto de observações para um novo sistema de coordenadas. As novas coordenadas indicam onde se encontram as maiores variâncias no conjunto.

O resultado da aplicação do PCA, tendo o PDM como entrada, é uma matriz de  $t$  autovetores  $P_M$  que expressam as componentes principais do modelo e seus respectivos autovalores  $\vec{\lambda}$ . Através dessa matriz, da média das formas e dos autovalores é possível reconstruir qualquer forma de maneira que ela seja compatível com o conjunto de treinamento.

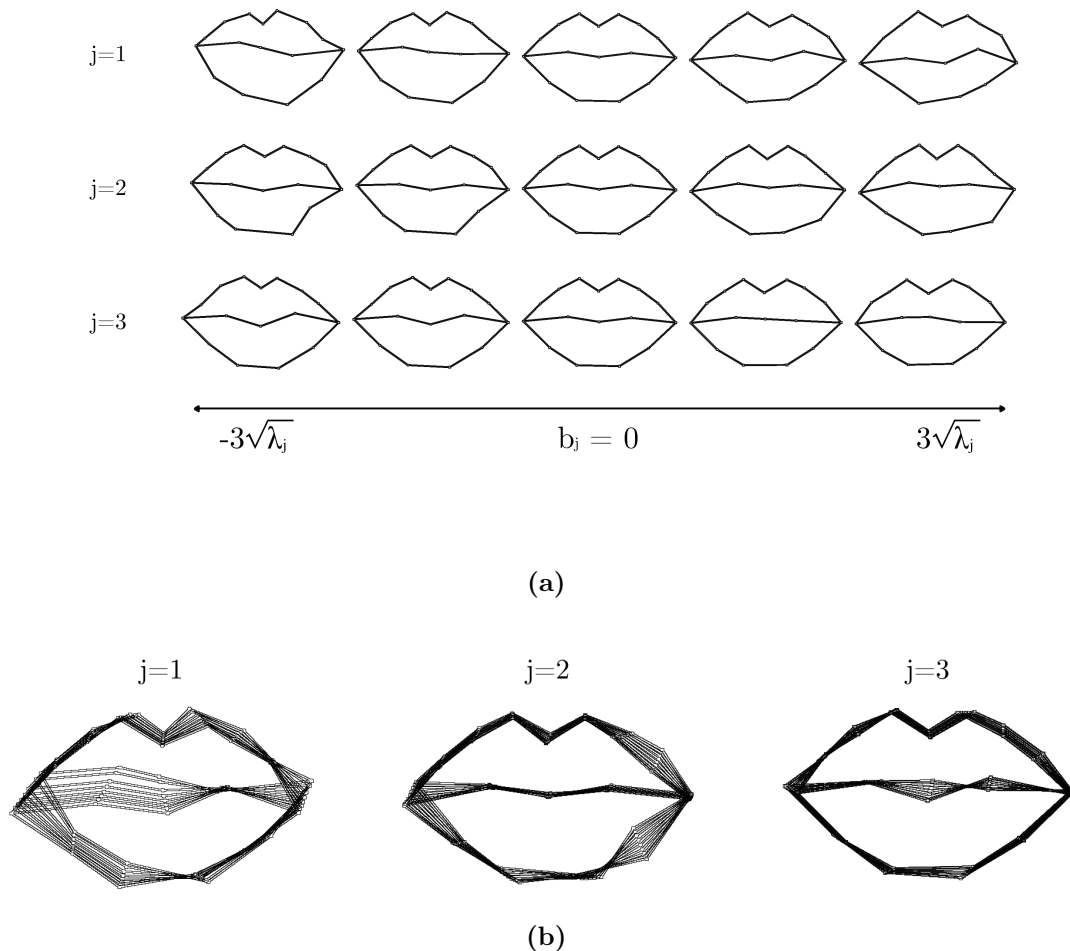
Para isso, é preciso descobrir um vetor de parâmetros  $\vec{b}$ . Como mencionado na Seção 3.3.2, cada elemento de  $\vec{b}$  representa a variação em uma coordenada do conjunto de observações transformado pelo PCA. Contudo, não se pode permitir que os valores de  $\vec{b}$  excedam um determinado limite, caso contrário a forma reconstruída não será compatível com o conjunto observado. Segundo Cootes et al. (1995) tal limite deve ser determinado como mostra a Equação 4.16,

$$|b_j| \leq 3\sqrt{\lambda_j}, \quad j = 1, 2, \dots, t. \quad (4.16)$$

A Figura 32 mostra os tipos de alterações que a média das formas pode sofrer mantendo-se coerente com o conjunto de treinamento.

Todos esses elementos são necessários para a outra etapa do método que é a busca, em uma nova imagem, pela forma observada durante o treinamento. A seção seguinte faz uso dos elementos produzidos na etapa de construção do algoritmo e de outros métodos, relatados no Capítulo 3, para conduzir a busca.





**Figura 32** – Modos de variação do conjunto de bocas marcadas da Figura 31. (a) Média das formas variando ao longo de 3 componentes. (b) Modos de variação sobrepostos.

### 4.3 Busca pela Forma

Nesta etapa, é realizada a aplicação da busca pela forma em uma imagem desconhecida – não utilizada no treino do modelo. Ela consiste, resumidamente, de duas operações aplicadas alternadamente até que não se observe modificações significativas na posição da forma. O Algoritmo 3 mostra como essas etapas se alternam para resolver o problema. Ele tem como entradas a média das formas  $\vec{\mu}_f$ , as médias dos perfis de textura  $\vec{\mu}_{tx}$ , a matriz de autovetores  $P_M$ , os autovalores correspondentes  $\vec{\lambda}$ , a imagem  $f(x, y)$  e a matriz de pesos  $W$ . Seu resultado é a forma localizada  $\vec{f}_s$ .

Nesta seção, são detalhados os métodos de busca por uma nova posição para os pontos (*BuscarNovaPosicao*) e de ajuste da forma gerada (*AjustarForma*).

**Algoritmo 3:** Etapa de busca do ASM.

---

**Entrada:**  $\vec{\mu}_f, \vec{\mu}_{tx}, P_M, \vec{\lambda}, f(x, y), W$   
**Saída:**  $\vec{f}_s$

- 1 **início**
- 2      $\vec{f}_c \leftarrow \vec{\mu}_f$
- 3     **repita**
- 4          $\vec{f}_s \leftarrow \emptyset$
- 5         **para** cada  $(x_{cj}, y_{cj}) \in \vec{f}_c$  **faça**
- 6              $(x_{cj}, y_{cj}) \leftarrow \text{BuscarNovaPosicao}(x_{cj}, y_{cj}, f(x, y), \vec{\mu}_{txj})$
- 7             Adicione  $(x_{cj}, y_{cj})$  em  $\vec{f}_s$
- 8         **fim**
- 9          $\vec{f}_s \leftarrow \text{AjustarForma}(\vec{f}_s, \vec{\mu}_f, P_M, \vec{\lambda}, W)$
- 10          $\vec{f}_c \leftarrow \vec{f}_s$
- 11     **até** haver convergência;
- 12 **fim**
- 13 **retorna**  $\vec{f}_s$

---

### 4.3.1 Buscar Nova Posição

A nova posição de cada ponto é encontrada com base na sua atual posição, na imagem alvo da busca e na média dos perfis de textura. O primeiro passo deve ser a coleta de vários perfis de textura próximos ao ponto atual. Para isso, deve ser retirada uma amostra da imagem que corresponde à reta normal à forma que passa pelo ponto. O tamanho da amostra deve ter  $2h + 1$ , com  $h > k$ , e deve gerar  $2(h - k) + 1$  possíveis novas posições para o ponto. De cada posição possível deve ser retirada uma amostra de tamanho  $2k + 1$  para ser comparada com a média  $\vec{\mu}_{tx}$ . A comparação é feita a partir da primeira derivada normalizada de cada amostra<sup>2</sup>.

O processo está descrito no Algoritmo 4 que tem como entradas as coordenadas do ponto  $x_{cj}$  e  $y_{cj}$ , o ângulo de inclinação  $\alpha$  da reta normal à forma passando pelo ponto  $(x_{cj}, y_{cj})$ , a imagem  $f(x, y)$  e a média dos perfis de textura  $\vec{\mu}_{txj}$  do  $j$ -ésimo ponto. O resultado do algoritmo é o par ordenado  $(cx, cy)$  indicando a nova posição do ponto.

$\text{Bresenham}(f, p_a, p_b)$  é a mesma rotina definida no Algoritmo 1.  $\text{Textura}(\vec{tx}, 2n+1)$  determina um conjunto composto por  $2(m - n) + 1$  perfis de textura extraídos de  $\vec{tx}$  com tamanho  $m > n$ .  $\text{Mahalanobis}(\vec{\mu}, \vec{v})$  está definida no Anexo A.

Após a aplicação do algoritmo,  $2(h - k) + 1$  vetores de textura terão sido comparados com a textura média daquele ponto. O centro  $(cx, cy)$ <sup>3</sup> daquele vetor que for mais semelhante à média, segundo a distância de *Mahalanobis*, será escolhido como nova posição

<sup>2</sup> A amostra de textura é transformada para reduzir os efeitos negativos das diferenças de iluminação nas imagens da base. Esse processo é feito substituindo a textura  $f(x_i, y_i)$  do vetor  $\vec{tx}_i$  pela diferença  $f(x_i, y_i) - f(x_{i+1}, y_{i+1})$ .

<sup>3</sup> O par  $(cx_i, cy_i)$  indica as coordenadas, na imagem  $f(x, y)$ , do centro do vetor  $\vec{tx}_i$ .

---

**Algoritmo 4:** Buscar nova posição do ponto.
 

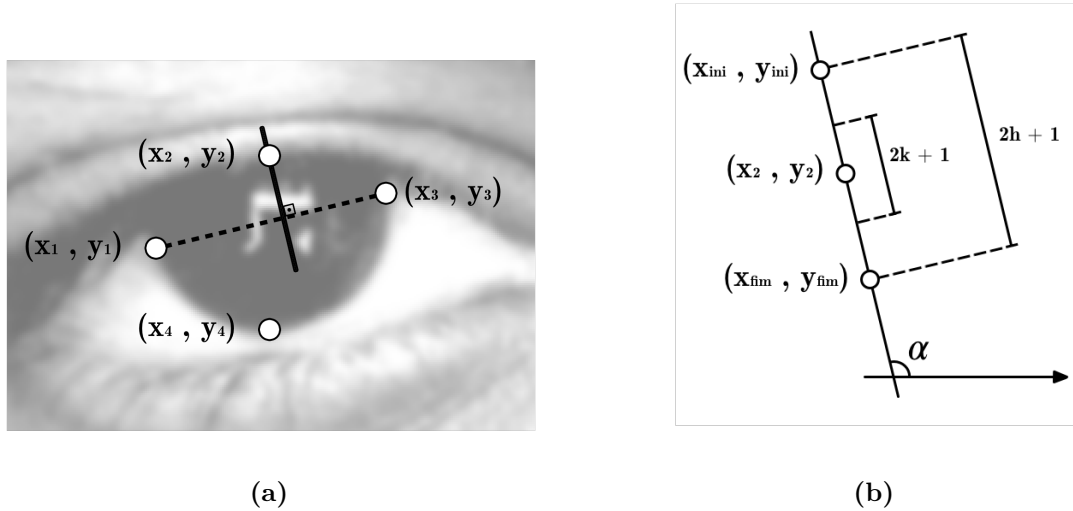
---

**Entrada:**  $x_{cj}, y_{cj}, \alpha, f(x, y), \vec{\mu}_{txj}$   
**Saída:**  $(cx, cy)$

- 1 **início**
- 2    $\overline{Dm} \leftarrow \emptyset$
- 3    $(x_{ini}, y_{ini}) \leftarrow (x_{cj} + h \cos \alpha, y_{cj} + h \sin \alpha)$
- 4    $(x_{fim}, y_{fim}) \leftarrow (x_{cj} - h \cos \alpha, y_{cj} - h \sin \alpha)$
- 5    $\vec{tx} \leftarrow \text{Bresenham}(f(x, y), (x_{ini}, y_{ini}), (x_{fim}, y_{fim}))$
- 6   **para** cada  $\vec{tx}_i \in \text{Textura}(\vec{tx}, 2k + 1)$  **faça**
- 7      $Dm_i \leftarrow \text{Mahalanobis}(\vec{\mu}_{txj}, \vec{tx}_i)$
- 8     Adicionar  $(Dm_i, (cx_i, cy_i))$  em  $\overline{Dm}$
- 9   **fim**
- 10 **fim**
- 11 **retorna**  $(cx_i, cy_i) \in \overline{Dm} | Dm_i \leq Dm_j, \forall j \in (1, 2, \dots, 2(h - k) + 1)$

---

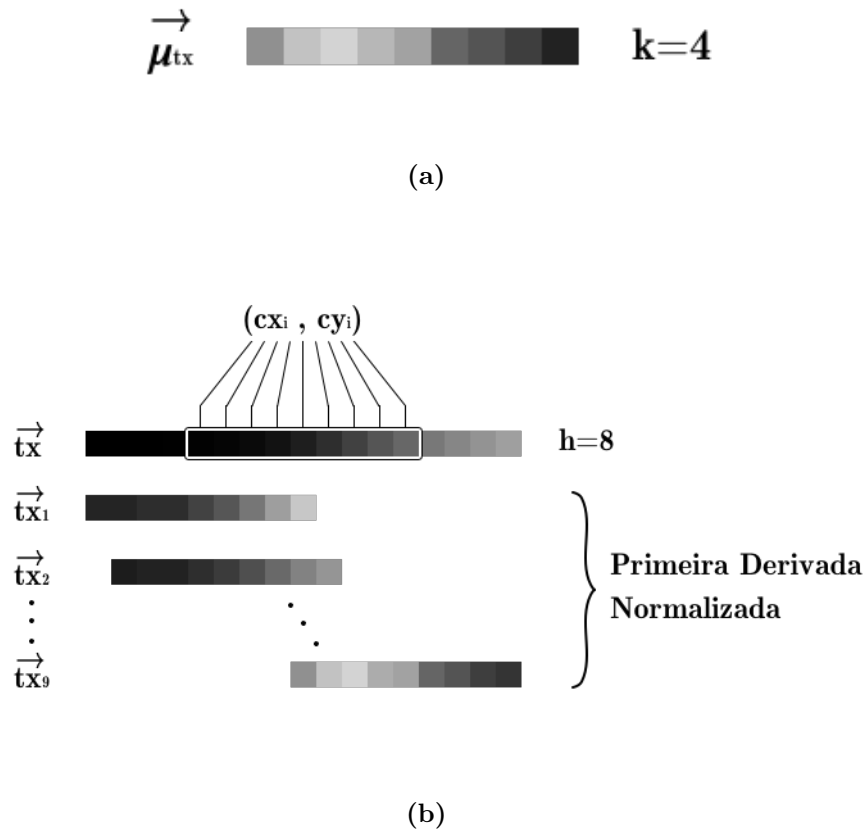
para o ponto avaliado. A Figura 33 traz um exemplo do cálculo da reta normal à forma e ilustra como os limites  $(x_{ini}, y_{ini})$  e  $(x_{fim}, y_{fim})$  estão relacionados com o ponto.



**Figura 33** – Coleta de textura de um ponto do olho. (a) Reta perpendicular à forma passando pelo ponto  $(x_2, y_2)$ . (b) Limites da amostra de textura.

A Figura 34 ilustra uma amostra de tamanho  $2h + 1$ , retirada do exemplo exposto na Figura 33, e indica a posição de cada centro  $(cx_i, cy_i)$ . Antes de efetuar a comparação, é calculada a primeira derivada de cada amostra e os valores resultantes são normalizados. Na Figura 34a, está ilustrado o perfil de textura médio para o ponto  $(x_2, y_2)$ . A Figura 34b aponta como são compostos os  $2(h - k) + 1$  perfis de texturas que serão usados na comparação com a média dos perfis.

O perfil que estiver mais próximo da média, ou seja, que apresentar menor valor  $Dm_i$ , terá seu centro  $(cx_i, cy_i)$  escolhido como nova localização do ponto. No caso ilustrado



**Figura 34** – Perfis de textura ao redor de pontos da forma. (a) Média dos perfis para o ponto  $(x_2, y_2)$  da Figura 33. (b) Faixas de valores que devem ser comparadas com a média dos perfis.

na Figura 34, o perfil de textura mais próximo da média foi o  $\vec{tx}_9$ .

### 4.3.2 Ajustar Forma

Eventualmente, durante a fase de busca, é gerada uma forma não compatível com o conjunto utilizado na criação do modelo. Para identificar esse problema, é necessário recorrer à restrição definida na Equação 4.16. O cálculo de  $b_j$  pode ser feito como na Equação 3.15. Caso algum  $b_j$  ultrapasse a sua respectiva restrição, é necessário limitá-lo ao valor  $\lambda_j$  e recalculá-la a forma correspondente através da Equação 3.14.

Após ajustada a forma, é necessário realizar um alinhamento para aproximar a forma corrigida daquela que foi previamente gerada. Isso é feito de acordo com a Equação 4.14. O Algoritmo 5 descreve formalmente como é feito o ajuste. As suas entradas são a forma  $\vec{f}_s$  gerada pelo Algoritmo 4, a média das formas  $\vec{\mu}_f$ , a matriz de autovetores  $P_M$ , os autovalores correspondentes  $\vec{\lambda}$  e a matriz de pesos  $W$ . O resultado é a forma  $\vec{f}_c$  com os pontos localizados.

A execução desse algoritmo garante que a nova forma gerada faça parte do conjunto

**Algoritmo 5:** Ajuste de forma.

---

**Entrada:**  $\vec{f}_s, \vec{\mu}_f, P_M, \vec{\lambda}, W$   
**Saída:**  $\vec{f}_c$

- 1 **início**
- 2      $\vec{b} \leftarrow P_M^T(\vec{f}_s - \vec{\mu}_f)$
- 3     **para** cada  $b_j \in \vec{b}$  onde  $|b_j| > 3\sqrt{\lambda_j}$  **faça**
- 4         **se**  $b_j \geq 0$  **então**
- 5              $b_j \leftarrow 3\sqrt{\lambda_j}$
- 6         **senão**
- 7              $b_j \leftarrow -3\sqrt{\lambda_j}$
- 8         **fim**
- 9     **fim**
- 10     $\vec{f}_c \leftarrow \vec{\mu}_f + P_M \vec{b}$
- 11    Alinhar  $\vec{f}_c$  com  $\vec{f}_s$  pesos  $W$
- 12 **fim**
- 13 **retorna**  $\vec{f}_c$

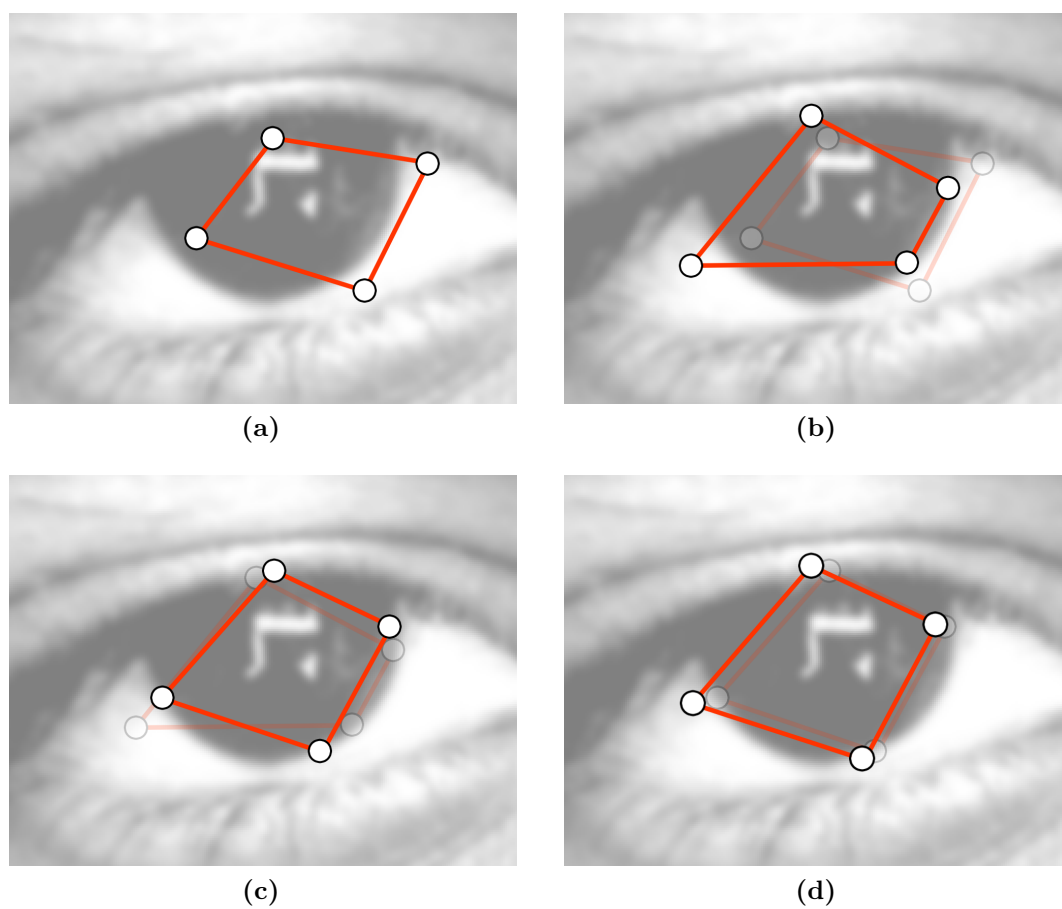
---

de treinamento, segundo a restrição apresentada na Equação 4.16. A Figura 35 exemplifica o ajuste de uma forma, composta por quatro pontos, gerada para o olho. É possível observar que, após a execução do algoritmo, a localização dos pontos pode ser deslocada da melhor posição como resultado de um efeito colateral do ajuste.

## 4.4 Considerações Finais

Este capítulo mostrou como o método ASM executa a busca por uma forma em uma imagem. Foram detalhados aspectos que se iniciam na construção do modelo (treinamento do algoritmo), passando pelas etapas de rotulação do conjunto de treinamento até a captura dos modos de variação. Ao final, foi mostrado como ele executa, efetivamente, a busca pela forma em uma nova imagem dividida em duas etapas: geração de novas posições e ajuste de forma.

O algoritmo ASM é essencial para a resolução de parte do problema apresentado. Nossa proposta realiza algumas alterações no método tradicional. O intuito é de resolver o problema da localização de alguns pontos faciais sob a ótica das cirurgias plásticas. O Capítulo 5 introduz o sistema proposto e expõe os detalhes dos ajustes, para este método, propostos em nosso trabalho, além de apresentar um método para localização de pontos do perfil.



**Figura 35** – Demonstração de iteração do Algoritmo 3. (a) Localização da forma antes da execução do algoritmo. (b) Forma  $\vec{f}_s$  gerada pelo Algoritmo 4. (c) Forma  $\vec{f}_c$  gerada pelo Algoritmo 5. (d) Forma  $\vec{f}_c$  alinhada a  $\vec{f}_s$ .

## 5 Metodologia Proposta

Este trabalho propõe uma metodologia de localização e identificação de pontos faciais em imagens frontais e de perfil. Ela deve ser explorada em avaliações estéticas destinadas a análises pré-operatórias de cirurgias plásticas. A metodologia não é responsável por nenhuma análise estética, mas deve ser aproveitada como ferramenta para facilitá-las.

O problema proposto pode ser separado em duas etapas:

- localização e identificação dos pontos de referência do perfil facial;
- localização e identificação dos pontos de referência faciais frontais.

Da forma como a metodologia está proposta, a solução do problema de localização dos pontos do perfil necessita apenas de uma imagem facial correspondente, uma vez que o método não realiza nenhuma etapa de pré-processamento. Para o caso dos pontos frontais, há uma necessidade de se fornecer ao método imagens previamente anotadas. Após treinado o modelo, basta inserir uma nova imagem facial frontal para dar início à busca pelos pontos. A Figura 36 mostra um esquema simplificado que ilustra a metodologia proposta.

Neste capítulo, inicialmente é tratada a etapa que diz respeito aos pontos do perfil facial. São discutidas algumas técnicas exploradas na resolução do problema relacionadas com os métodos apresentados no Capítulo 3. A solução proposta é um método *ad-hoc*<sup>1</sup> que divide as curvas do perfil em grupos para facilitar a análise de padrões.

Posteriormente, é apresentado o método que localiza pontos frontais, elaborado com base no ASM. Esta proposta realiza algumas mudanças no *Active Shape Models* para buscar, em uma imagem, os pontos necessários ao planejamento cirúrgico. Apesar de ter sido testado para o problema em questão, ele é um método que pode ser aproveitado para situações semelhantes.

### 5.1 Localizador de Pontos do Perfil (LPP)

Esta parte da metodologia realiza exclusivamente a localização de pontos do perfil facial. Na Seção 3.1.2 estão descritos os pontos de referência alvos abordados pelo LPP. A localização dos pontos do perfil deve ser iniciada com uma imagem facial compatível visto que não está prevista a detecção ou localização facial. Além disso, devem ser observadas três restrições importantes:

---

<sup>1</sup> Metodologia criada exclusivamente para este problema.

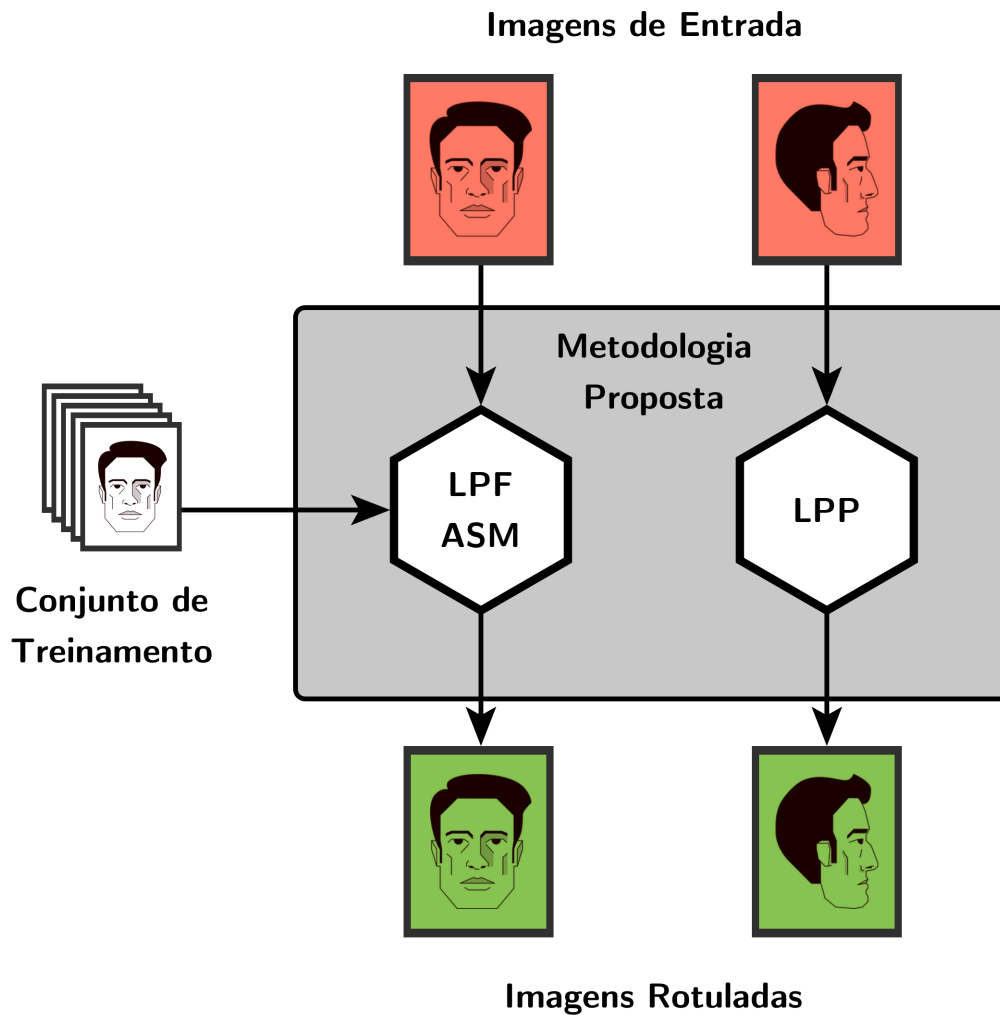


Figura 36 – Metodologia proposta.

- a imagem deve conter apenas uma face;
- o perfil deve estar na posição horizontal natural<sup>2</sup>;
- o fundo da imagem deve ser uniforme e contrastar com a face.

Essas restrições limitam bastante o escopo do problema e normalmente são observadas apenas em ambiente controlado. Existem métodos que tratam esse problema em situações mais espontâneas e que são apropriadas para todo tipo de imagem (ZHU; RAMANAN, 2012). Contudo, para nosso trabalho, um ambiente de aquisição controlado é mais conveniente, já que tal fator contribui para a precisão do método, e ele deve ser aplicado justamente neste tipo de situação (consultório médico, sala de cirurgia, etc).

<sup>2</sup> Alinhado com a superfície para não deformar a face.



A Figura 37 ilustra as etapas de funcionamento do LPP desde a entrada da imagem até o resultado final.

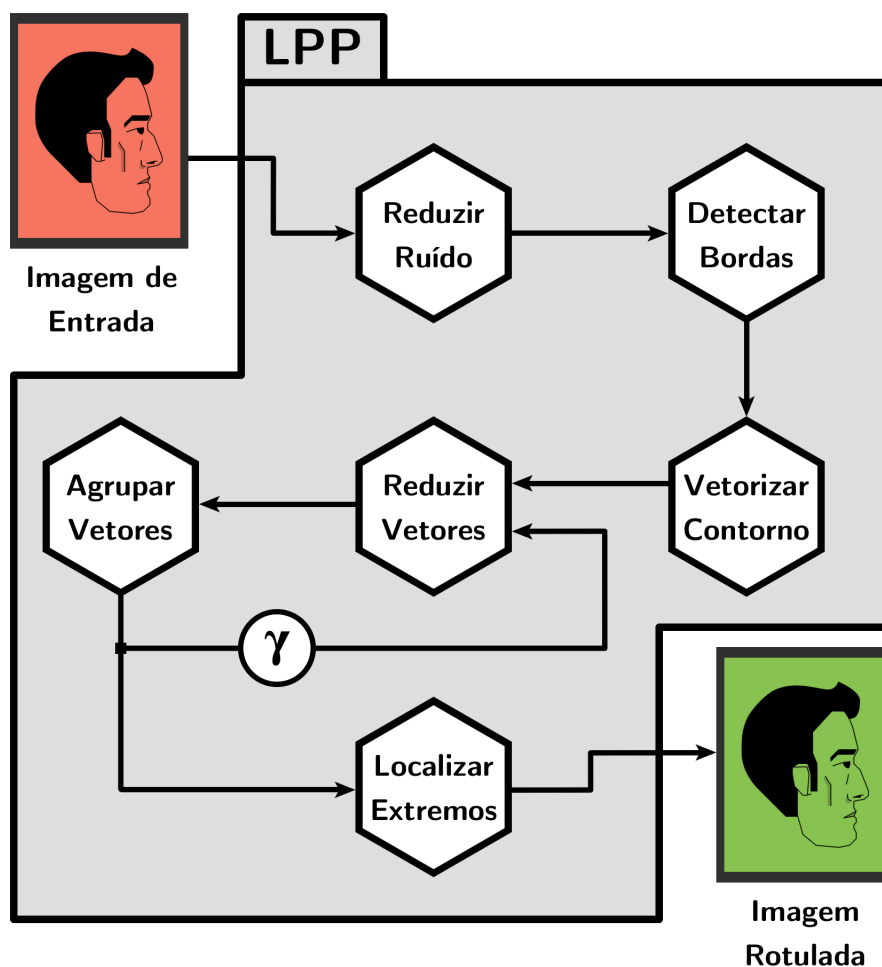


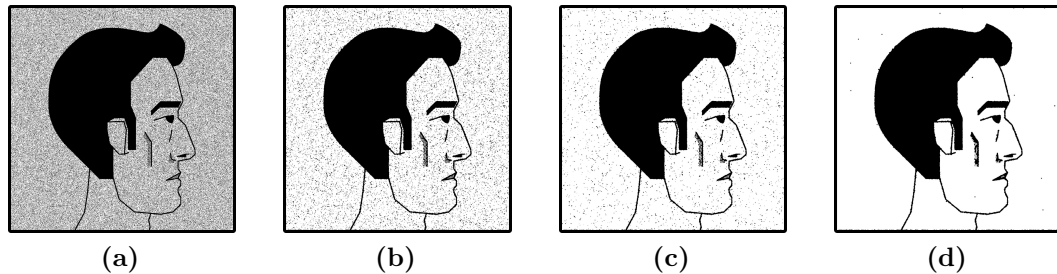
Figura 37 – Esquema de funcionamento do LPP.

Resumidamente, este método de busca é baseado em um processo de simplificação de uma representação vetorial do contorno do perfil. Primeiramente deve ser determinada a borda do perfil facial. Depois, ela segue por um processo que a transforma em vetores. A partir daí, eles passam por uma rotina iterativa onde alguns vetores são somados, descartados e agrupados para determinar as regiões onde cada ponto se encontra.

### 5.1.1 Processamento da Imagem

As duas primeiras etapas do LPP são relativas às técnicas de PDI apresentadas no Capítulo 3. Como mencionado anteriormente, a filtragem mediana é uma importante ferramenta que ajuda na redução de ruído da imagem ao passo que mantém uma característica essencial para a solução deste problema: as regiões de alto contraste. Isso significa que a identificação de contornos ou bordas, em qualquer parte da imagem, não sofre prejuízos significativos. Para esta metodologia, é importante tanto preservar os contornos na imagem quanto remover vestígios de ruído.

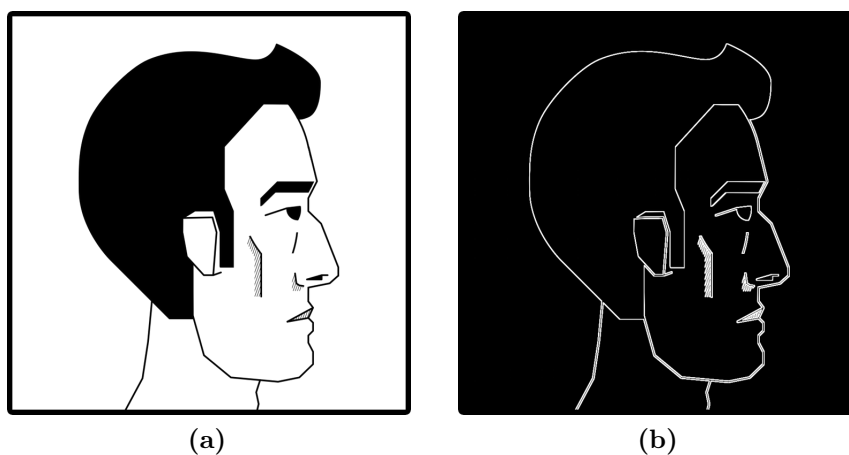
A primeira etapa é responsável por este trabalho, onde a imagem é filtrada, sucessivamente, para que sejam reduzidos ruídos que possam prejudicar as etapas posteriores. A Figura 38 ilustra esse processo em algumas etapas. Uma imagem ruidosa (Figura 38a) é filtrada diversas vezes através desse processo (Figuras 38b, 38c e 38d).



**Figura 38** – Aplicações sucessivas da filtragem mediana. (a) Imagem ruidosa. (b) Filtragem mediana ( $\times 3$ ). (c) ( $\times 6$ ). (d) ( $\times 9$ ).

A etapa seguinte realiza a extração das bordas presentes na imagem. Para isto, é utilizado o método de Canny apresentado na Seção 3.2.4. O resultado desta etapa é uma imagem binária  $I_b$  com as bordas detectadas.

Concluídas as etapas de processamento, o resultado é uma imagem onde sobram apenas as bordas da face e, principalmente, o contorno do perfil. Como todos os pontos alvo desta parte da metodologia se encontram no contorno, essas duas etapas são fundamentais para a solução do problema. A Figura 39 ilustra o resultado desta etapa.



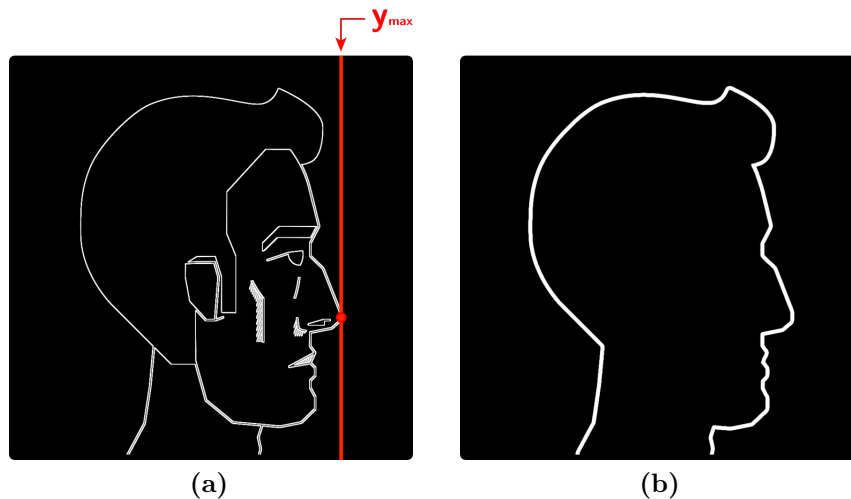
**Figura 39** – Resultado esperado da etapa de detecção de bordas. (a) Imagem de entrada. (b) Imagem com bordas detectadas.

### 5.1.2 Vetorizar Contorno

A estratégia do método proposto envolve um processo de simplificação de uma representação vetorial do perfil. O contorno deve ser transformado em vetores que substituem os *pixels* na representação do perfil facial.

Para transformar o contorno em uma lista de vetores, inicialmente é preciso determinar as bordas do perfil facial. A busca pelo contorno que representa o perfil do indivíduo é baseada na noção de vizinhança de *pixels* descrita em Gonzalez e Woods (2011) e em teoria de grafos. Nesta etapa, a busca é semelhante a uma busca em profundidade realizada em um grafo, supondo que cada  $I_b(x, y) > 0$  da imagem binária representa um vértice e as ligações entre eles e seus vizinhos  $I_b(vx, vy) > 0$  indicam arestas.

Para determinar os *pixels* que fazem parte do contorno, basta realizar a busca em profundidade a partir de um ponto que pertence ao contorno. O ponto de localização imediata para realizar tal busca é a ponta do nariz, pois é o ponto mais proeminente horizontalmente do perfil facial. Para encontrar a ponta do nariz na imagem  $I_b$  é necessário localizar o pixel  $I_b(x, y) > 0$  com  $y$  máximo ou mínimo<sup>3</sup>. O processo está esquematizado na Figura 40.



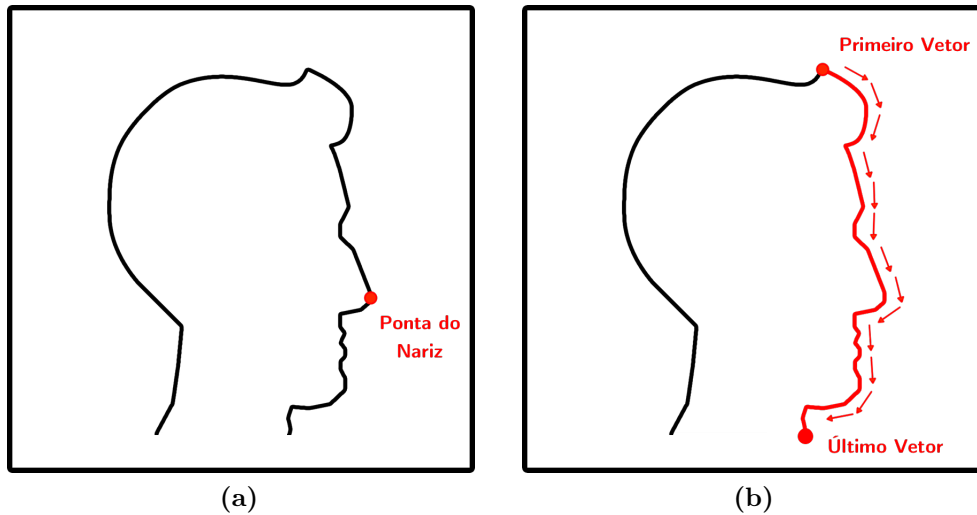
**Figura 40** – Busca pelo perfil. (a) Localização do ponto com  $y_{max}$ . (b) Definição do contorno.

Com a ponta do nariz determinada, realiza-se a busca em profundidade pelo contorno. As coordenadas de todos os *pixels* encontrados pertencem a um conjunto chamado  $C_p = \{p_1, p_2, p_3, \dots, p_n\}$ . Uma vez que  $C_p$  contém todos os *pixels* do contorno do perfil, é necessário apenas descobrir o subconjunto de  $C_p$  que corresponde às coordenadas dos pontos de referência.

A última parte deste processo consiste em dispor os elementos do conjunto  $C_p$  em forma de vetores. Dentro do conjunto não há noção de ordem, porém, os pares ordenados de  $C_p$  estão conectados pela sua vizinhança. Essa informação é usada para construir uma cadeia de vetores. Um vetor  $\vec{v}_x$  é construído a partir de duas coordenadas:  $p_i$ , indicando o início do vetor; e  $p_j$ , indicando o fim. Necessariamente,  $p_i$  e  $p_j$  devem ser vizinhos na imagem  $I_b$ .

<sup>3</sup> Para fotos de perfil onde a face está voltada para a direita da imagem, deve-se buscar o  $y$  máximo. Caso contrário, deve-se procurar pelo  $y$  mínimo.

Após descobertos todos os vetores  $v_x$  formados a partir do conjunto  $C_p$ , basta ordená-los em uma lista  $L_v$  de modo que eles descrevam o perfil facial. Neste caso, o vetor que tiver a coordenada  $x$  de menor valor<sup>4</sup> deve iniciar a lista, assim como o último vetor é aquele que tiver maior coordenada  $x$ . A Figura 41 mostra a vetorização do contorno.



**Figura 41** – Vetorização do contorno. (a) Localização da ponta do nariz. (b) Vetorização da parte frontal da face em perfil.

### 5.1.3 Reduzir Vetores

Para encontrar o subconjunto de  $C_p$  que corresponde à solução do nosso problema, deve-se eliminar elementos dele que não se enquadram nas características dos pontos de referência. Como esses pontos representam regiões extremas no contorno<sup>5</sup>, deve-se remover do conjunto aqueles que não fazem parte dessas áreas.

Para isso, a lista de vetores  $L_v$  que constitui o perfil deve sofrer uma modificação. Os vetores que constituem regiões planas devem ser somados, enquanto que aqueles que fazem parte dos extremos faciais devem ser preservados. A soma de vetores consecutivos de  $L_v$  nesse processo é análogo à determinação de um subconjunto de  $C_p$ . Quando dois vetores consecutivos são somados, a coordenada comum aos dois desaparece.

Contudo, caso não haja controle durante esta operação, poderá restar somente um elemento em  $L_v$ . Para restringir esse processo, deve ser usado um valor limiar  $\gamma$  responsável por impedir que a soma ocorra.

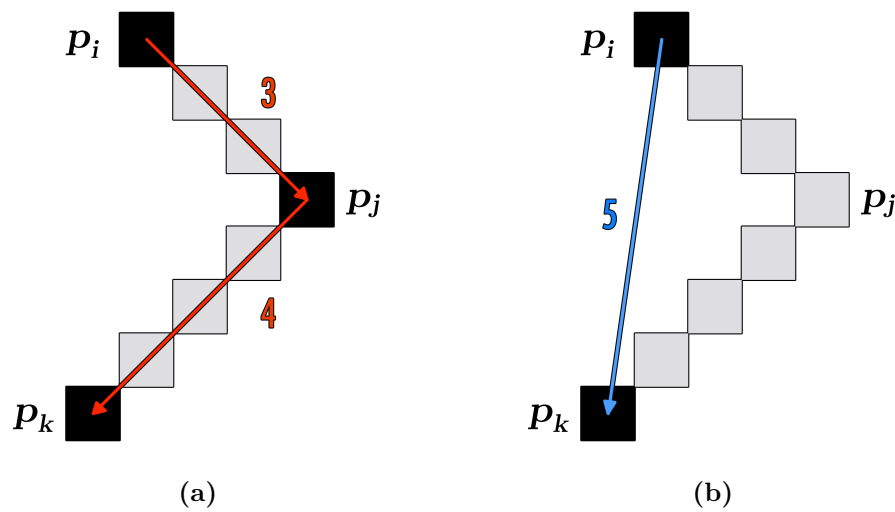
Nesse processo, sempre que dois vetores são somados, alguma informação é descartada. O controle da soma é feito com base nesse descarte, que deve ser acumulado toda vez que isso ocorrer. Suponha que:  $\overrightarrow{p_i p_j}, \overrightarrow{p_j p_k} \in L_v \mid \|\overrightarrow{p_i p_j}\| = 3, \|\overrightarrow{p_j p_k}\| = 4$  sejam somados de acordo com a Figura 42.

<sup>4</sup> Este vetor indica a região mais elevada da cabeça.

<sup>5</sup> Protuberâncias ou concavidades no perfil.

O vetor  $\overrightarrow{p_i p_k}$  resultante tem módulo menor que a soma dos módulos de  $\overrightarrow{p_i p_j}$  e  $\overrightarrow{p_j p_k}$ . Neste caso, diz-se que a soma foi responsável por uma redução  $R_j = 2$  no contorno avaliado. A informação perdida é mensurada dessa maneira e é regulada pelo valor limiar  $\gamma$ . Caso a redução seja maior que o limiar, os dois vetores são preservados na lista  $L_v$ . Caso contrário, eles são substituídos pelo vetor resultante.

Além disso, os vetores envolvidos na soma podem ser resultantes de somas anteriores. Logo, deve-se adicionar a esse cálculo os valores descartados  $D_i$  (do vetor  $\overrightarrow{p_i p_j}$ ) e  $D_j$  (do vetor  $\overrightarrow{p_j p_k}$ ) que foram acumulados pelos vetores envolvidos. No caso ilustrado na Figura 42a, os valores  $D_i$  e  $D_j$  são nulos. Na Figura 42b,  $D_i = 2$ .



**Figura 42** – Adição de vetores. (a) Vetores  $\overrightarrow{p_i p_j}$  e  $\overrightarrow{p_j p_k}$  originais. (b) Vetor  $\overrightarrow{p_i p_k}$  resultante da soma.

O processo de redução da lista de vetores  $L_v$  está detalhado no Algoritmo 6. Ele tem como parâmetros de entrada a lista de vetores  $L_v$  a ser reduzida e o valor do limiar  $\gamma$  que controla o processo. Sua saída é uma versão reduzida da lista  $L_v$ .

Inicialmente, os valores de  $\gamma$  e todos de os elementos da lista  $D$  são nulos. Contudo, como ilustrado na Figura 37, o fluxo da metodologia permite que a etapa de redução de vetores seja executada repetidas vezes, com novos valores  $\gamma$  definidos pela etapa de agrupamento de vetores.

### 5.1.4 Agrupar Vetores

Este processo recebe uma lista  $L_v$  e agrupa os vetores com respeito ao ângulo que eles formam entre si. O agrupamento é realizado como descrito no Algoritmo 7. Ele recebe uma lista de vetores  $L_v$  para agrupar e retorna uma lista de grupos de vetores  $L_G$ . A rotina  $Angulo(\vec{x}, \vec{y})$  determina o ângulo entre dois vetores  $\vec{x}$  e  $\vec{y}$ . Os grupos  $G_n$  e  $G_c$  guardam,

**Algoritmo 6:** Algoritmo de redução de vetores.

---

**Entrada:**  $L_v, \gamma$   
**Saída:**  $L_v$

```

1 início
2    $D \leftarrow \{0, 0, \dots, 0\}$ 
3   para cada  $\overrightarrow{p_i p_j}, \overrightarrow{p_j p_k} \in L_v$  faça
4      $R_j \leftarrow \|\overrightarrow{p_i p_j}\| + \|\overrightarrow{p_j p_k}\| - \|\overrightarrow{p_i p_k}\|$ 
5     se  $R_j + D_i + D_j < \gamma$  então
6       Remover  $\overrightarrow{p_i p_j}$  e  $\overrightarrow{p_j p_k}$  de  $L_v$ 
7       Adicionar  $\overrightarrow{p_i p_k}$  em  $L_v$ 
8        $D_i \leftarrow R_j + D_i + D_j$ 
9     fim
10  fim
11 fim
12 retorna  $L_v$ 

```

---

respectivamente, vetores que formam ângulos não-côncavos e vetores que formam ângulos côncavos.

Para finalizar, o processo define se deve ser ou não aplicada novamente a etapa de redução de vetores. Ele cumpre essa tarefa comparando o número de grupos  $S(L_G)$  criados com a quantidade  $S_E$  de protuberâncias e concavidades presentes no perfil facial médio<sup>6</sup>. Com  $S(L_G) > S_E$  significa que o contorno não foi reduzido o suficiente para a localização dos pontos, assim, o valor  $\gamma$  precisa ser aumentado. Caso  $S(L_G) < S_E$ , o contorno foi simplificado demais e o valor  $\gamma$  precisa ser reduzido. Se  $S(L_G) = S_E$ , o processo de agrupamento é finalizado e é feita uma análise individual em cada grupo, para determinar a localização de cada ponto.

A Figura 43 ilustra um exemplar médio do perfil com seus extremos marcados por círculos. Neste caso, existem 12 extremos faciais ( $S_E = 12$ ). Em algumas situações, o valor  $\gamma$  não é preciso o suficiente para fazer  $S(L_G) = S_E$ , então, a comparação pode ser flexibilizada para que  $S(L_G)$  se adeque a  $S_E$ .

### 5.1.5 Localizar Extremos

Para finalizar o LPP, os grupos formados devem ser avaliados individualmente junto com as definições de cada ponto de referência. Uma vez que a definição de vários pontos coincide com os extremos locais do contorno (protuberâncias e concavidades do perfil), a maioria deles é identificada através dos vetores de maior e menor ângulos dentro dos grupos.

Para facilitar a localização dos pontos, os vetores devem ser organizados de maneira não-crescente no interior de cada grupo com respeito ao seus ângulos. Denomina-se o

<sup>6</sup> O valor  $S_E$  foi determinado empiricamente através da análise de 30 imagens de perfil.

**Algoritmo 7:** Algoritmo de agrupamento de vetores.

---

**Entrada:**  $L_v$   
**Saída:**  $L_G$

```

1 início
2    $L_G \leftarrow \emptyset$ 
3    $G_n \leftarrow \emptyset$ 
4    $G_c \leftarrow \emptyset$ 
5   se  $Angulo(\overrightarrow{p_0p_1}, \overrightarrow{p_1p_2}) \leq 180^\circ$  então
6     | Adicionar  $\overrightarrow{p_0p_1}$  em  $G_n$ 
7   senão
8     | Adicionar  $\overrightarrow{p_0p_1}$  em  $G_c$ 
9   fim
10  para cada  $\overrightarrow{p_i p_j}, \overrightarrow{p_j p_k} \in L_v$  faça
11    se  $\overrightarrow{p_i p_j} \in G_n$  então
12      se  $Angulo(\overrightarrow{p_i p_j}, \overrightarrow{p_j p_k}) \leq 180^\circ$  então
13        | Adicionar  $\overrightarrow{p_j p_k}$  em  $G_n$ 
14      senão
15        | Adicionar  $\overrightarrow{p_j p_k}$  em  $G_c$ 
16        | Adicionar  $G_n$  em  $L_g$ 
17        |  $G_n \leftarrow \emptyset$ 
18      fim
19    senão
20      se  $Angulo(\overrightarrow{p_i p_j}, \overrightarrow{p_j p_k}) > 180^\circ$  então
21        | Adicionar  $\overrightarrow{p_j p_k}$  em  $G_c$ 
22      senão
23        | Adicionar  $\overrightarrow{p_j p_k}$  em  $G_n$ 
24        | Adicionar  $G_c$  em  $L_g$ 
25        |  $G_c \leftarrow \emptyset$ 
26      fim
27    fim
28  fim
29 fim
30 retorna  $L_G$ 

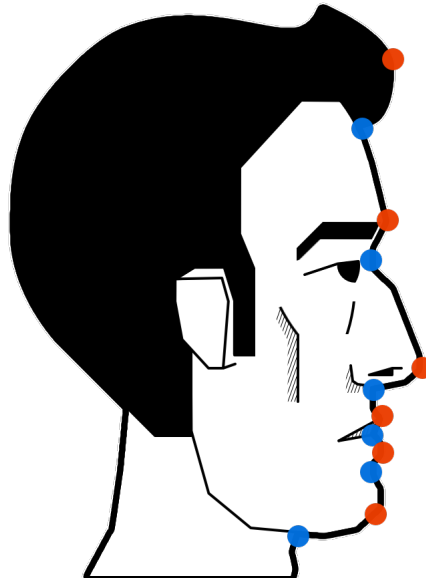
```

---

ângulo formado por  $\overrightarrow{p_i p_j}$  e  $\overrightarrow{p_j p_k}$  de  $\hat{p}_j$ . O valor considerado para ordenar um grupo de vetores é o módulo da diferença entre  $\hat{p}_j$  e o ângulo raso  $180^\circ$ . Quanto maior o valor de  $|\hat{p}_j - 180|$ , mais brusca é a protuberância ou concavidade. Os maiores valores de cada grupo indicam os pontos de referência alvos do nosso trabalho.

Existe apenas uma ressalva para essa abordagem. Os pontos rínio, columela, pogônio e mentoniano pertencem a grupos onde se encontram mais de um ponto alvo. Para cada um deles, existe uma maneira diferente de localização dentro dos grupos ordenados.

- **Rínio** : No grupo de pontos onde se encontra o nariz, o rínio é representado pelo vetor acima da ponta do nariz que tem maior valor  $|\hat{p}_i - 180|$ .



**Figura 43** – Protuberâncias e concavidades do perfil facial. Os círculos em vermelho indicam as protuberâncias. Os azuis representam as concavidades.

- **Columela** : No grupo de pontos onde se encontra o nariz, a columela é representada pelo vetor abaixo da ponta do nariz que tem maior valor  $|\hat{p}_i - 180|$ .
- **Pogônio** : No grupo de pontos do queixo, o pogônio é representado pelo vetor com maior coordenada  $x$ .
- **Mentoniano** : No grupo de pontos do queixo, o mento é representado pelo vetor com maior coordenada  $y$ .

De acordo com as definições apresentadas nesta seção, o método descrito é capaz de localizar os pontos de referência do perfil indicados no Capítulo 3.

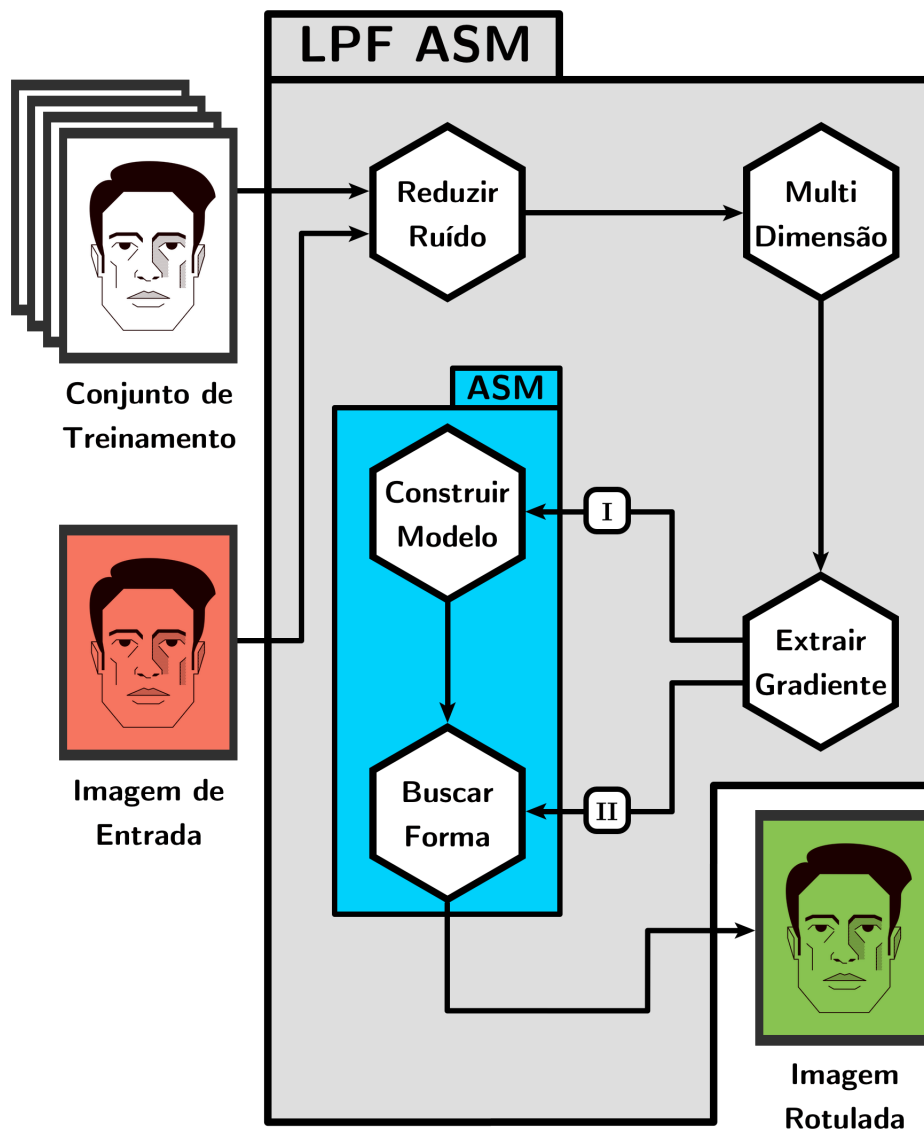
## 5.2 Localizador de Pontos Frontais (LPF ASM)

A outra parte que compõe a metodologia faz a localização dos pontos frontais. Ela tem como base o algoritmo *Active Shape Models* descrito no Capítulo 4. A diferença mais significativa do LPF para o LPP é a necessidade de uma etapa de treinamento do algoritmo antes de executar a localização propriamente dita.

A Figura 44 mostra o funcionamento da metodologia proposta para a localização de pontos frontais. Ela deve contar com dois tipos de entradas: as imagens do conjunto de treinamento e as imagens de entrada que serão alvos da busca. Cada parte do LPF gera um elemento aproveitado em outra. As etapas de redução de ruído, multi-dimensão e extração de gradiente geram imagens. A construção do modelo no ASM gera a média das formas, as médias dos perfis de textura e os modos de variação como detalhados no Capítulo



4. A etapa de busca gera a localização dos pontos. Algumas das etapas ilustradas não pertencem ao ASM original mas foram melhorias observadas em trabalhos relacionados.



**Figura 44** – Esquema de funcionamento do LPF ASM. O fluxo I do diagrama é percorrido pelo conjunto de treinamento. A imagem de teste percorre o fluxo II.

Esta seção destaca alguns assuntos preliminares empregados no método, explica como funciona o fluxo do LPF e detalha as mudanças aplicadas no algoritmo ASM original para melhorar a busca pelos pontos frontais abordados no Capítulo 3.

### 5.2.1 Formas e Conjunto de Pontos

Antes de descrever o LPF ASM é necessário ressaltar como as formas devem ser determinadas no conjunto de treinamento e na etapa de busca.

Originalmente, o ASM foi desenvolvido para localizar objetos sólidos em imagens.

A descrição de [Cootes et al. \(1995\)](#) aponta três grupos de pontos que definem o contorno desses objetos e devem ser escolhidos para realizar as etapas de marcação, treinamento e busca. São eles, em ordem decrescente de importância:

1. pontos dependentes da aplicação com significado particular;
2. pontos independentes da aplicação, normalmente extremos locais do contorno;
3. pontos que devem ser interpolados a partir dos dois primeiros grupos, normalmente apresentando distâncias iguais dos pontos adjacentes.

Contudo, é possível encontrar trabalhos na literatura que relacionam pontos que não pertencem a forma alguma. [Milborrow \(2007\)](#) cita, por exemplo, a ponta do nariz em uma imagem frontal como sendo alvo da busca pelo ASM. Em uma imagem frontal, a ponta do nariz, apesar de se encaixar no primeiro grupo, não faz parte do contorno de nenhum objeto.

Como apontado na Seção 3.1, existem 21 pontos de referência alvos explorados nessa parte do trabalho. Muitos desses pontos pertencem a formas triviais na face (limbo lateral, por exemplo, pertence à íris), entretanto existem alguns que fogem a essa regra. Esse é o caso do orbitário, que dificilmente pode ser localizado em uma imagem comum<sup>7</sup>.

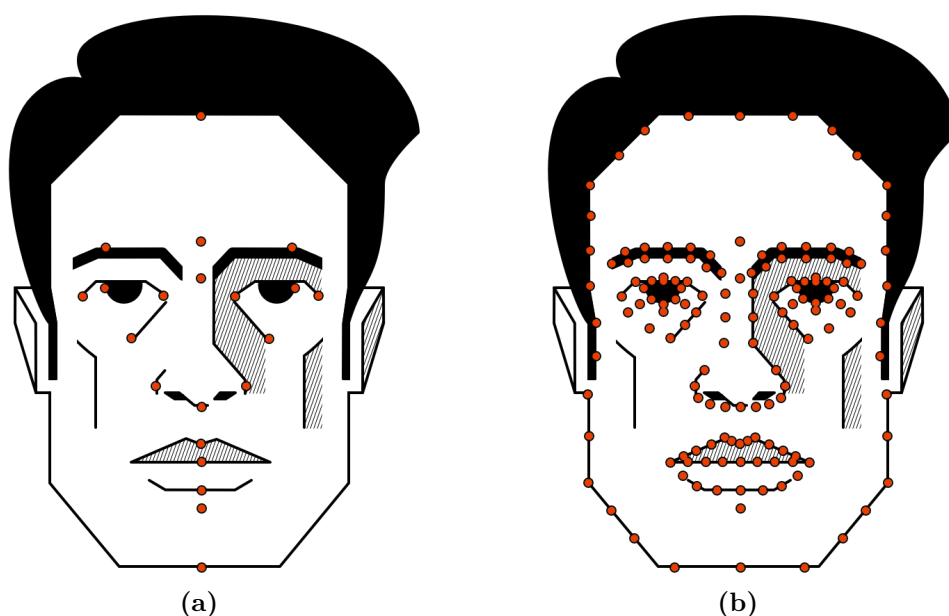
Além disso, devido a pequena quantidade de pontos descritos na Seção 3.1, mesmo aqueles que estão presentes em alguma forma trivial da face podem não ser facilmente observados em sua forma correspondente. Isso acontece porque os pontos apresentados são insuficientes para que as formas sejam corretamente observadas.

A solução deste problema consiste em adicionar pontos para compor as formas que contêm os alvos do trabalho. Eles devem ser adicionados como descrito por [Cootes et al. \(1995\)](#) e [Dryden e Mardia \(1998\)](#), ou seja, devem servir para preencher as formas buscadas de modo que os pontos fiquem equidistantes ao longo do contorno. Nesta proposta são usados, no total, 145 pontos que compreendem todas as formas necessárias para a localizar os pontos de referência. A Figura 45b ilustra tais pontos.

Os pontos escolhidos podem ser ainda mais numerosos. Já foi observado, em trabalhos relacionados, que existe uma correlação entre a quantidade de pontos que compõem o modelo e o nível de precisão do ASM. [Milborrow \(2007\)](#) estabeleceu esta relação ao aplicar o ASM para localização de pontos faciais. Segundo ele, quanto maior a quantidade de pontos, mais precisa é a localização.

Em nosso trabalho, o único empecilho para a escolha de um conjunto ainda mais amplo é o custo da marcação dos pontos. A rotulação manual de uma imagem facial leva

<sup>7</sup> O orbitário é o ponto mais inferior de uma das órbitas e geralmente só pode ser localizado com precisão em imagens de raio-x.



**Figura 45** – Conjunto de pontos explorados no trabalho. (a) 21 pontos frontais alvos do trabalho. (b) 145 pontos usados para melhorar os resultados de busca.

cerca de 9 minutos com o auxílio de um *software* (Figura 46), especialmente desenvolvido para auxiliar as marcações dos conjuntos de treinamento e teste. Supondo que um conjunto de 400 imagens deva ser rotulado, o tempo total necessário para marcar todas as imagens seria de 60 horas ininterruptas.

### 5.2.2 Reduzir Ruído

A primeira etapa de processamento do LPF tanto para o conjunto de treinamento quanto para as imagens de teste é a redução de ruído. Em nosso método, a filtragem Gaussiana foi escolhida para realizar essa etapa. Esse processo tem como função transformar a imagem de maneira que as variações de intensidade fiquem mais suaves. O resultado é uma imagem embaçada, com bordas pouco definidas e com pouco ruído. Esse processo é de extrema importância na comparação dos perfis de textura apresentados nas Seções 4.2.2 e 4.3.1.

Para o LPF ASM é preciso definir um parâmetro essencial da função que indica o nível de suavidade da imagem resultante. Como visto na Equação 3.6, a retirada de uma amostra da função exige a definição de um valor para  $\sigma$ . No caso do LPF, considerando o tamanho das imagens usadas<sup>8</sup>, o valor mais adequado foi  $\sigma = 10,0$ . Uma discussão sobre a escolha desse parâmetro se encontra no Capítulo 6.

Com esse valor, é coletada uma amostra da função em duas dimensões. Os valores observados são usados para compor um *kernel*  $5 \times 5$  que é operado (convolução) na imagem

<sup>8</sup> Todas as imagens usadas têm dimensão  $1600 \times 1200$  *pixels*.

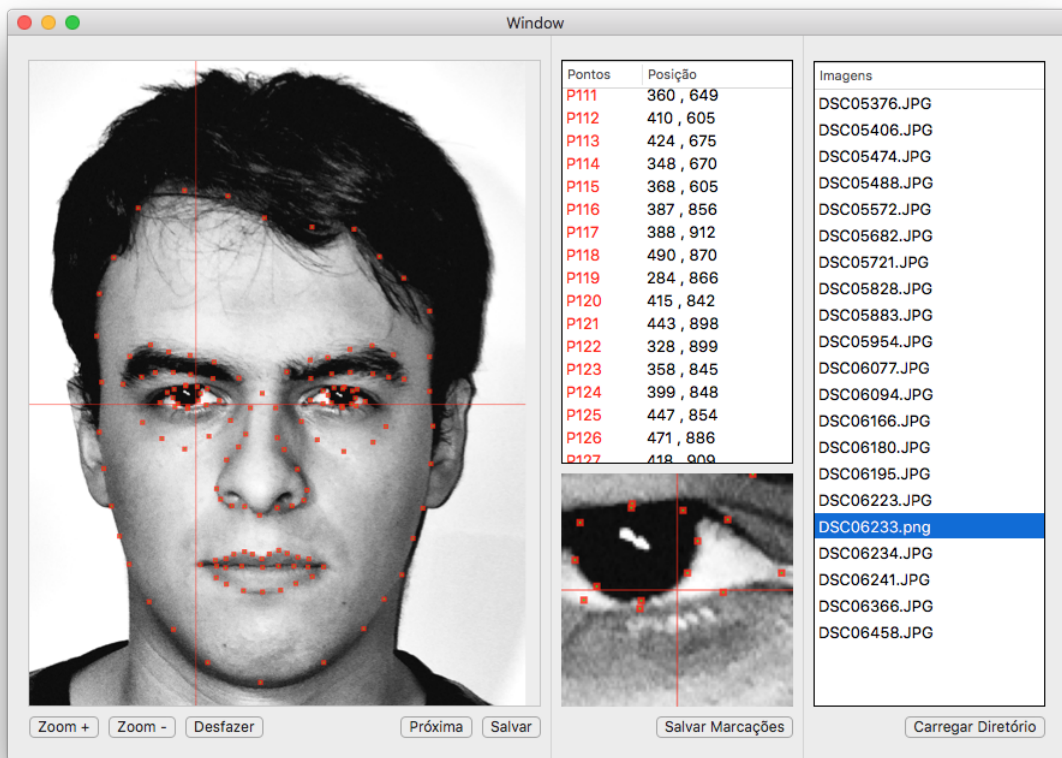


Figura 46 – Software desenvolvido para auxiliar a marcação manual das imagens.

para realizar a filtragem. A Figura 47 ilustra o processo de aplicação da filtragem dentro do LPF.

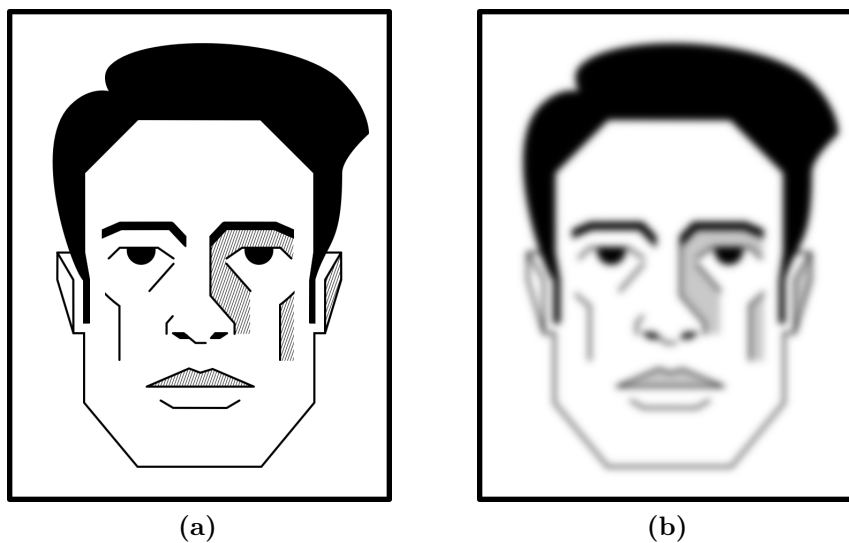
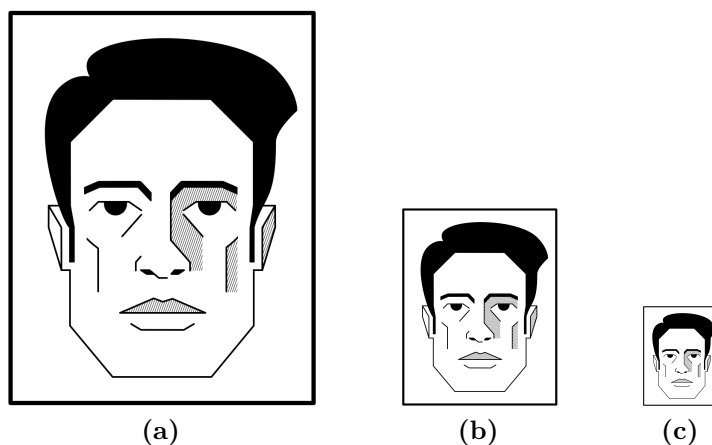


Figura 47 – Aplicação da filtragem gaussiana com  $\sigma = 10,0$ . (a) Imagem de entrada. (b) Imagem filtrada.

### 5.2.3 Múltiplas Dimensões

A aplicação do ASM em versões de uma imagem com diferentes dimensões é uma estratégia que já foi abordada por [Cootes e Taylor \(2004\)](#) e [Milborrow \(2007\)](#). O processo consiste em diminuir as dimensões da imagem de entrada com diferentes fatores de redução. As imagens geradas são exploradas pelo ASM em ordem crescente de resolução.

Nesta metodologia, as imagens são redimensionadas com o uso da interpolação bilinear definida em [Gonzalez e Woods \(2011\)](#). São usadas no algoritmo além da imagem original, uma versão com  $\frac{1}{4}$  do tamanho original e outra com  $\frac{1}{16}$ . A Figura 48 mostra um exemplo onde uma imagem facial é redimensionada dessa maneira. Durante a etapa de busca, a imagem de menor resolução serve como entrada com uma localização padrão inicial para a forma.



**Figura 48** – Pirâmide de imagens usada na execução do ASM. (a) Imagem original. (b) Imagem redimensionada a  $\frac{1}{4}$ . (c) Imagem redimensionada a  $\frac{1}{16}$ .

A imagem original junto com as duas geradas nesta etapa servem como entrada para a etapa seguinte do LPF.

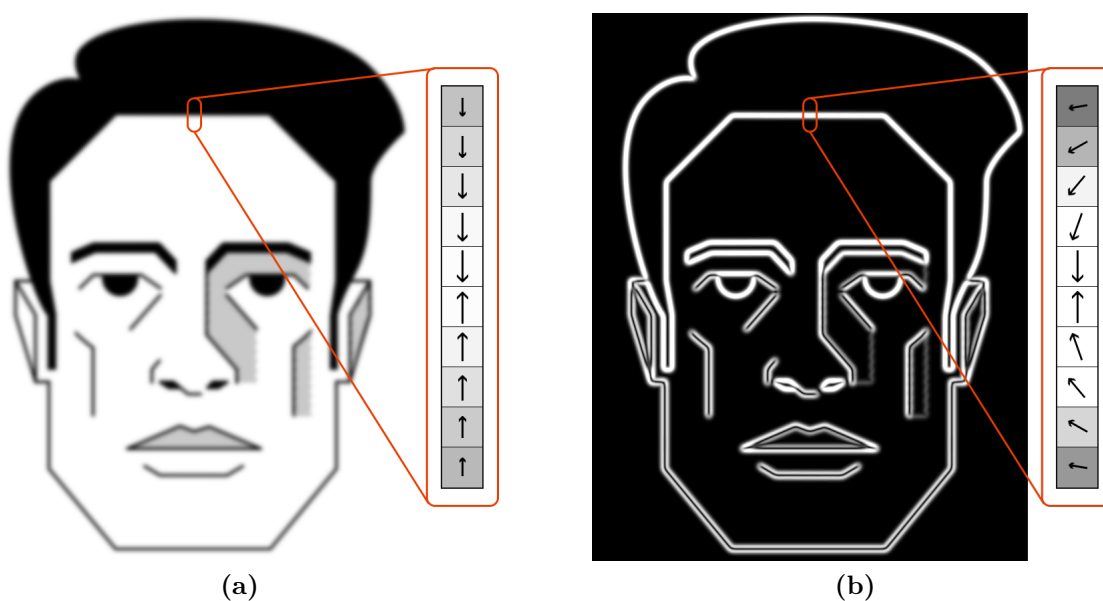
### 5.2.4 Extrair Gradiente

Esta etapa apresenta uma diferença com relação ao ASM original durante as fases de construção da média dos perfis de textura e de localização de novas posições detalhadas nas Seções 4.2.2 e 4.3.1.

De acordo com o ASM tradicional, a média dos perfis de textura é calculada a partir de vários perfis observados durante treinamento. Os valores avaliados devem ser as diferenças de intensidade entre os elementos daquele perfil. Para cada ponto, é calculada a média desses valores. Isso é feito para reduzir as diferenças de iluminação nas imagens avaliadas dentro do conjunto de treinamento. Contudo, apenas os valores ao longo da reta normal à forma passando pelo ponto são avaliados durante a construção do perfil médio.

A diferença da nossa abordagem para a tradicional é a utilização do operador Sobel para extrair as mudanças de intensidade na imagem. Enquanto o método tradicional observa o gradiente apenas na direção da reta normal, nossa proposta é de avaliar o gradiente em várias direções.

Como mencionado na Seção 3.2.3.2, a aplicação do filtro Sobel é realizada com auxílio de dois *kernels*: um que detecta a mudança de intensidade horizontal na imagem e outro que detecta a mudança vertical (Equação 3.8). Uma vez que os *kernels* detectam variações de intensidades ortogonais, é possível realizar uma composição (após a convolução dos *kernels* na imagem) para descobrir a direção do gradiente (Equação 3.9). Como consequência deste processo, o gradiente ao longo do perfil de textura pode se apresentar em várias direções, enquanto que no ASM original, só é possível observar uma direção para todo o gradiente. A Figura 49 ilustra os resultados dessas duas diferentes formas para calcular o perfil de textura com base nas diferenças de intensidade da imagem.



**Figura 49** – Comparação dos perfis extraídos de duas formas diferentes. (a) Perfil com gradiente em uma direção. (b) Perfil com gradiente em várias direções.

Dessa maneira, a média do perfil é construída com um refinamento maior de informações sobre as mudanças de intensidade. Durante a fase de busca pela forma essa alteração também é explorada.

### 5.2.5 Construção do Modelo e Busca dos Pontos - ASM

A construção do modelo usado no algoritmo segue como detalhado no Capítulo 4, com a diferença que os perfis de textura devem ser calculados a partir da imagem gerada pela convolução com o operador Sobel. Como são utilizados três tipos de imagens (original,  $\frac{1}{4}$  e  $\frac{1}{16}$  do tamanho original) um modelo deve ser construído para cada uma das resoluções.

Os três modelos gerados na etapa de construção são utilizados para realizar a busca das formas que contêm os pontos de referência. Diferente da construção, esta fase apresenta algumas mudanças em relação ao ASM tradicional.

Relembrando o Capítulo 4, a busca do ASM pela forma é dividida em duas etapas: busca por novas posições e ajuste de forma (Algoritmo 3). A primeira leva em consideração os perfis de textura para a localização de novas posições para os pontos. A segunda é responsável por manter a forma consistente com os elementos observados no conjunto do treinamento.

De acordo com a descrição original do ASM, as novas posições são definidas a partir da comparação, através da distância de *Mahalanobis*, entre a média dos perfis e os perfis observados na imagem de entrada. Contudo, a forma pode não ser coerente com o conjunto de treinamento. Nesse caso deve-se realizar seu ajuste e realinhá-la com a forma gerada na etapa anterior.

A maneira como a forma é realinhada pode levá-la a uma posição pior que a observada na iteração anterior do algoritmo. Isso pode acontecer quando os pontos que obtiveram melhores acertos na comparação dos perfis forem deslocados em virtude do alinhamento. Tal alinhamento é feito com uma matriz de pesos  $W$  que não leva em consideração a textura dos perfis. Nessas situações, tais pontos são “punidos” e acabam sendo afastados de posições mais próximas do alvo correto. A matriz é calculada como mostra a Equação 4.13.

Nosso método modifica o alinhamento criando uma nova matriz de pesos  $W'$  de modo que ela reflita o acerto nos perfis de textura, em vez de ser uma matriz identidade ou uma que seja composta pelo desvio padrão de cada ponto com relação à forma. Nesse caso, sempre que um perfil de textura avaliado se aproximar da média, o ponto correspondente tenderá a permanecer próximo do local, mesmo após o alinhamento. A matriz  $W'$  é calculada de acordo com a Equação 5.2. Nela, a função *Mahalanobis* (Equação A.6) calcula a referida distância entre a média dos perfis de textura  $\vec{\mu}_{txk}$  do ponto  $k$  para o perfil de textura  $\vec{txk}_i$  escolhido na etapa de busca por novas posições.

$$w'_k = Mahalanobis(\vec{\mu}_{txk}, \vec{txk}_i)^{-1} \quad (5.1)$$

$$W' = \begin{bmatrix} w'_1 & 0 & \dots & 0 \\ 0 & w'_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w'_m \end{bmatrix} \quad (5.2)$$

Além disso, o LPF ASM aplica uma outra alteração para aproveitar as melhores posições da etapa de busca. Após o alinhamento na fase de ajuste, pode acontecer de a

forma adquirir uma localização melhor, sob o ponto de vista da textura, com o alinhamento tradicional. Portanto, nesses casos deve-se optar pelo algoritmo clássico em detrimento de nossa proposta. Caso isso não seja observado, escolhe-se a localização que leva em conta a textura no momento do alinhamento.

Essa tomada de decisão é responsabilidade de uma função objetivo  $F(w)$  que leva em consideração a soma de todos os acertos dos perfis de textura. Caso o valor da função com o nosso método seja menor, escolhe-se a nossa proposta de alinhamento. Caso contrário, opta-se pela proposta tradicional. A função  $F(w)$  é calculada de acordo com a Equação 5.3,

$$F(w) = \sum_{k=1}^m w_k. \quad (5.3)$$

A modificação na etapa de busca do ASM está detalhada no Algoritmo 8. As linhas em vermelho indicam as mudanças realizadas no algoritmo original (Algoritmo 3). É possível observar que a forma  $f_c$  resultante de uma iteração é sempre a melhor opção, com base na função objetivo, entre nossa modificação e a proposta original.

---

**Algoritmo 8:** Etapa de busca do LPF ASM.
 

---

**Entrada:**  $\vec{\mu}_f, \vec{\mu}_{tx}, P_M, \vec{\lambda}, f(x, y), W$   
**Saída:**  $\vec{f}_c$

- 1 **início**
- 2      $\vec{f}_c \leftarrow \vec{\mu}_f$
- 3     **repita**
- 4          $\vec{f}_s \leftarrow \emptyset$
- 5         **para** cada  $(x_{cj}, y_{cj}) \in \vec{f}_c$  **faça**
- 6              $(x_{cj}, y_{cj}) \leftarrow \text{BuscarNovaPosicao}(x_{cj}, y_{cj}, f(x, y), \vec{\mu}_{txj})$
- 7             Adicione  $(x_{cj}, y_{cj})$  em  $\vec{f}_s$
- 8         **fim**
- 9         Calcular  $W'$
- 10          $\vec{f}'_s \leftarrow \text{AjustarForma}(\vec{f}_s, \vec{\mu}_f, P_M, \vec{\lambda}, W')$
- 11          $\vec{f}_s \leftarrow \text{AjustarForma}(\vec{f}_s, \vec{\mu}_f, P_M, \vec{\lambda}, W)$
- 12         Calcular  $F(w')$  e  $F(w)$
- 13         **se**  $F(w') > F(w)$  **então**
- 14              $\vec{f}_c \leftarrow \vec{f}'_s$
- 15         **senão**
- 16              $\vec{f}_c \leftarrow \vec{f}_s$
- 17         **fim**
- 18     **até** haver convergência;
- 19 **fim**
- 20 **retorna**  $\vec{f}_c$

---



A Tabela 4 sumariza as principais características do ASM clássico, assim como os pontos mais importantes da nossa proposta de metodologia.

### 5.3 Considerações Finais

Em ambientes controlados, os cirurgiões plásticos precisam realizar manualmente todas essas etapas descritas neste capítulo para chegar a uma definição dos procedimentos cirúrgicos que devem ser adotados. Marcar dezenas de pontos faciais para calcular e avaliar mais algumas dezenas de métricas é um trabalho bastante desgastante que acaba afastando os médicos de uma análise pré-operatória bem fundamentada. O resultado de tal descuido pode ser uma alteração cirúrgica indesejada e irreversível para o paciente.

A proposta desta metodologia tem o intuito de reduzir a complexidade de uma análise facial para cirurgia plástica. Ela deve ser capaz de facilitar e dar mais segurança ao cirurgião durante o planejamento pré-operatório. O próximo capítulo apresenta resultados das etapas de localização e identificação dos pontos de referência e levanta algumas discussões sobre as dificuldades encontradas.

**Tabela 4** – Sumário dos métodos ASM clássico e LPF ASM.

	<b>ASM clássico</b>	<b>LPF ASM</b>
<b>Fase PDI</b>	—	Filtragem Gaussiana Sobel
<b>Treinamento</b>	Perfil unidirecional Alinhamento por forma PCA	Perfil com Sobel Alinhamento por forma PCA
<b>Busca</b>	<i>Mahalanobis</i> por textura Alinhamento por forma	<i>Mahalanobis</i> por textura Alinhamento por forma ou textura

## 6 Resultados e Discussão

Os resultados apresentados neste capítulo dizem respeito à implementação das duas partes que compõem a metodologia proposta: o localizador de pontos do perfil (LPP) e o localizador de pontos frontais baseado no ASM (LPF ASM).

A parte que compreende o LPP foi avaliada com base em uma implementação feita na linguagem *Objective-C* (INC; DAVIDSON, 2002). Durante o desenvolvimento do método foram utilizadas 30 imagens de perfil. Os testes foram realizados em 410 diferentes imagens de perfis faciais. Os resultados foram comparados com marcações manuais feitas por dois estudantes de medicina em nível de graduação. As imagens usadas para a coleta destes resultados têm dimensão  $2536 \times 3456$  *pixels*.

Com relação aos pontos frontais, o LPF ASM foi implementado na linguagem *Swift* (INC, 2014). Foram realizados vários cenários de testes com diferentes parâmetros e comparações com o ASM clássico. Durante o desenvolvimento, foram utilizadas 400 imagens que foram divididas em dois conjuntos de mesmo tamanho: treinamento e teste. As imagens exploradas possuem dimensão  $1200 \times 1600$  *pixels*. Os resultados foram comparados com marcações manuais realizadas por um mestrando com auxílio de um software desenvolvido especialmente para este trabalho.

A base de imagens utilizada foi construída especificamente para atender às necessidades do problema proposto, ou seja, elas foram adquiridas em ambiente controlado. Ela é composta por indivíduos de todas as faixas etárias e de várias etnias, portanto é bastante diversificada.

### 6.1 Resultados do Localizador de Pontos do Perfil

A geometria Euclidiana foi escolhida para medir a distância entre as marcações manuais e automáticas. Antes de compilar-se os resultados, a análise da distância bruta deve ser colocada em perspectiva. A utilização da distância real (medida em *pixels*) entre as marcações, como métrica de avaliação, mostra-se um engano pois várias imagens diferem em tamanho. Isso implica que as diferenças absolutas podem ser mais ou menos significantes, dependendo do tamanho da imagem.

A abordagem de avaliação adotada neste trabalho leva em consideração não apenas as diferenças absolutas entre as marcações, mas, também, as suas relações com o tamanho da imagem.

Como explicado por Zhu e Ramanan (2012), o erro deve ser calculado em função da média entre altura e largura da imagem. A Tabela 5 relaciona os resultados obtidos

com o LPP.

**Tabela 5** – Medidas em imagens de perfil de resolução  $2592 \times 3456$  pixels.

Sigla	Erro Médio (px)	Desvio Padrão	Erro Relativo
Tr	216,71	133,43	7,14 %
Gl	66,26	32,07	2,19 %
Ns	67,12	41,43	2,21 %
Rn	81,40	49,61	2,69 %
PN	8,93	4,30	0,29 %
Col	45,57	27,36	1,50 %
Sn	23,14	11,93	0,76 %
LS	36,97	18,18	1,22 %
St	22,60	15,37	0,74 %
LI	29,80	19,27	0,98 %
Sm	30,30	16,36	1,00 %
Pg	38,90	23,55	1,28 %
Mt	80,06	42,54	2,64 %
Crv	206,70	119,67	6,83 %

### 6.1.1 Discussão das Medidas

Os pontos onde a curvatura local no perfil é mais óbvia apresentaram os melhores resultados. A ponta do nariz (PN), o subnasal (Sn) e o stomion (St) são, na maioria dos casos, extremos locais muito aparentes no contorno. Assim, as localizações automáticas desses pontos apresentam os melhores resultados. Considerando a proposta de avaliação de Bottino e Cumani (2008), que considera um erro de aproximadamente 1% da largura da imagem como aceitável<sup>1</sup>, a localização média desses pontos pode ser comparada à marcação manual. Os lábios superior (LS) e inferior (LI), o supramental (Sm) e o násio (Ns) foram localizados com razoável precisão.

O tríquio (Tr) e o cervical (Crv) são os pontos mais desafiadores. Os resultados das localizações desses pontos se apresentaram inconsistentes e dificilmente seriam aproveitados em uma análise real. Contudo, as imagens de teste apresentaram alguns desafios a mais para esses pontos, o que adicionou dificuldades que não haviam sido enfrentadas, tornando a localização mais difícil até mesmo para os especialistas.

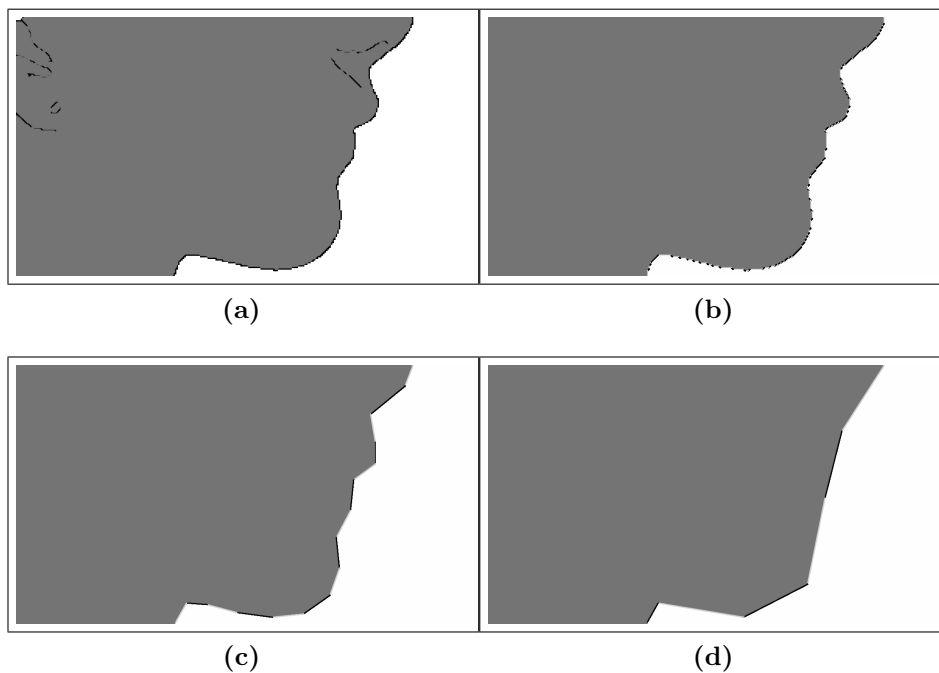
De acordo com a definição apresentada no Capítulo 3, o tríquio se encontra na linha do couro cabeludo. Em alguns casos, este ponto está escondido pelo próprio cabelo do indivíduo ou, em situações mais complicadas, não há couro cabeludo aparente (geralmente por conta da calvície). Nesses casos, a localização do ponto é extremamente difícil.

<sup>1</sup> Erro de 5 pixels em imagens de  $480 \times 640$  pixels.



e o contorno torna-se mais simples. Este processo preserva vetores consecutivos cujas direções são mais divergentes. Então, o valor do limiar deve ser cuidadosamente escolhido para que o contorno  $C_p$  seja reduzido apenas a seus *pixels* mais importantes.

A Figura 51 mostra diferentes valores de limiar sendo aplicados ao mesmo contorno. Na Figura 51b, o limiar é igual a zero, nenhuma característica é perdida. Na Figura 51c, o limiar ajuda a manter as características mais importantes para o propósito deste trabalho, portanto, é considerado um valor próximo ao ideal. Na Figura 51d o limiar é muito alto, resultando em um contorno muito mais simples.

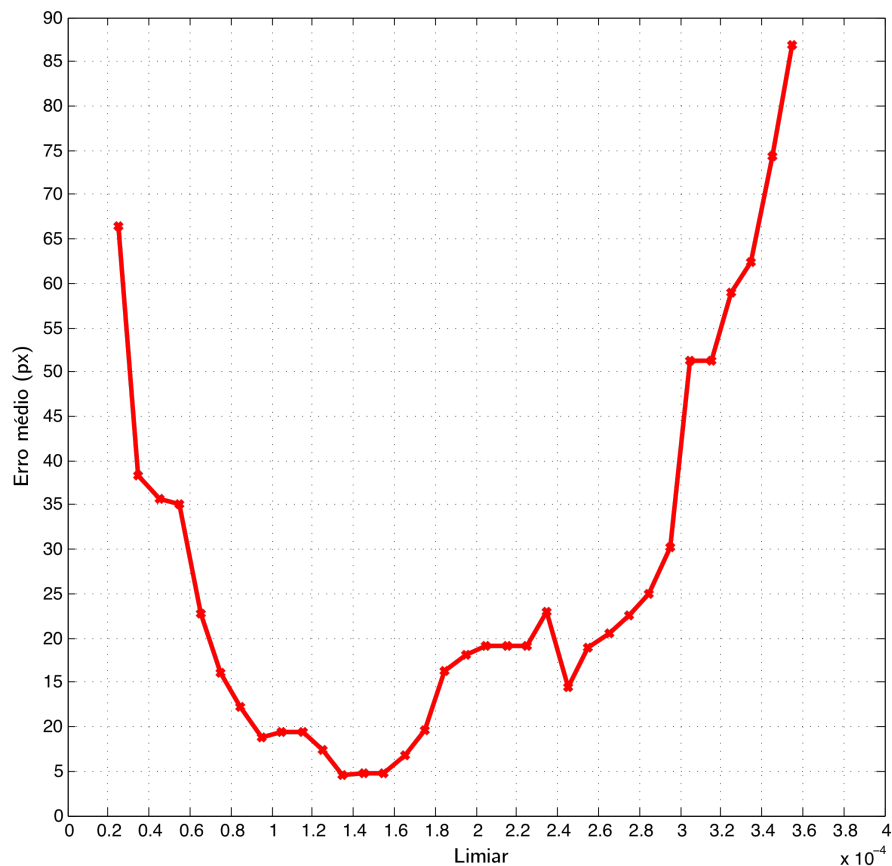


**Figura 51** – Região abaixo do nariz. (a) Contorno original. (b) Contorno com limiar  $\gamma = 0$ . (c) Contorno com  $\gamma > 0$ . (d) Contorno com  $\gamma \gg 0$ .

A Figura 52 exprime a relação entre o valor do limiar responsável por determinar a simplificação do contorno e o erro médio calculado com os 10 pontos que apresentaram melhor resultado. O eixo horizontal expressa o valor limiar em função do tamanho da imagem. Pode-se observar que valores muito elevados ou muito reduzidos, apresentam os piores resultados. Esses dados foram coletados em versões reduzidas das imagens originais.

A primeira proposta para esta parte da metodologia levava em consideração um limiar estático que deveria ser determinado a partir de uma análise empírica do conjunto usado para desenvolvimento do método. O gráfico da Figura 52 indica como essa análise poderia ser realizada.

Contudo, devido a grande variedade de perfis em um conjunto bastante rico etnicamente, adotar um valor estático para esse parâmetro mostrou-se uma fragilidade no algoritmo. Para superar esse problema, foi desenvolvida a versão atual do método que



**Figura 52** – Relação entre o limiar e o erro médio. Dados coletados de imagens com  $480 \times 640$  pixels.

dinamiza o cálculo do limiar (Seções 5.1.3 e 5.1.4). Assim, o valor se ajusta de acordo com a quantidade de extremos faciais.

### 6.1.3 Problemas com o Contorno

A proposta do LPP é um método altamente dependente do processo de detecção de bordas. Portanto, qualquer problema relacionado à extração de contorno tem grande influência nos resultados. Um fundo escuro, por exemplo, pode facilmente se confundir com a sobrancelha do paciente ou com seu cabelo, tornando a tarefa de localizar a glabella ou o tríquio muito difícil. A busca em grafo adotada para extrair o perfil exige um contorno sem falhas. Isso abre muitas possibilidades para erros e resultados ruins.

Alguns trabalhos podem ser explorados para resolver o problema da extração de contorno. *Active Contour Models* (KASS; WITKIN; TERZOPOULOS, 1988) é um exemplo de técnica que pode ser usada em trabalhos futuros para tornar o algoritmo mais robusto e menos dependente de uma detecção de bordas perfeita.

## 6.2 Resultados do Localizador de Pontos Frontais

Assim como no LPP, os resultados do método que efetua a localização de pontos frontais foram relativizados em função de alguma medida conhecida. Na Seção 3.3.3 é apontado que o erro deve ser calculado em função da distância entre os olhos. Para a visão frontal, o cálculo pode ser facilmente realizado, uma vez que os dois olhos estão sempre visíveis nas imagens do conjunto de testes.

Considerando as combinações ideais de parâmetros para a solução proposta, é possível chegar aos resultados apresentados na Tabela 6. As imagens usadas são versões redimensionadas do conjunto original, que é semelhante ao usado para avaliar o LPP. A decisão de usar versões reduzidas foi motivada pelo custo computacional da realização dos cenários de teste elencados nesta seção. A utilização das imagens originais tem imenso impacto tanto no custo de tempo quanto no custo de memória dos testes realizados.

**Tabela 6** – Medidas em imagens frontais de resolução  $1200 \times 1600$  pixels.

<b>Sigla</b>	<b>Erro Médio (px)</b>	<b>Desvio Padrão</b>	<b>Erro Relativo</b>
Tr	38,44	24,84	8,08%
Gl	18,41	9,75	3,87%
Ns	16,40	10,61	3,44%
ASd	21,90	12,58	4,59%
ASe	19,72	11,00	4,14%
LLd	13,11	7,52	2,76%
LLe	15,09	12,45	3,17%
CLd	14,12	7,99	2,96%
CLe	14,90	8,34	3,12%
CMd	13,56	9,17	2,86%
CMe	12,77	8,46	2,67%
Od	18,43	9,67	3,85%
Oe	17,79	11,98	3,70%
ANd	17,32	9,83	3,66%
ANe	15,69	9,06	3,28%
Sn	16,71	10,54	3,52%
LS	16,95	9,86	3,55%
St	15,28	8,49	3,21%
LI	18,93	10,46	3,95%
Sm	23,30	13,23	4,88%
Mt	22,70	12,76	4,76%

A Tabela 6 mostra que o LPF é mais regular que o LPP na localização dos pontos, uma vez que os resultados não apresentam variância tão grande. Contudo, os piores resultados de ambas as soluções estão fortemente relacionados com os pontos extremos

verticais. No caso do LPF, isso ocorre por conta da variância desses pontos após a etapa de alinhamento do conjunto. As subseções seguintes detalham melhor este problema.

### 6.2.1 Resultados de Modificações Externas ao ASM

O localizador de pontos frontais levou em conta modificações externas ao algoritmo ASM. Elas foram responsáveis por alterar as imagens de treinamento e teste, antes do processamento feito pelo método, com o objetivo de atacar fraquezas da proposta original. Algumas dessas propostas já foram abordadas em trabalhos relacionados e nosso trabalho apenas corrobora os resultados já obtidos. Além disso, nosso trabalho também contribui com uma nova abordagem para uma das etapas cruciais do algoritmo.

A análise multi-resolução proposta no Capítulo 5 é uma adaptação comumente feita em aplicações do ASM ((COOTES; TAYLOR, 2004), (MILBORROW, 2007)). A metodologia proposta explora a configuração mais sugerida na literatura para apurar a localização dos pontos. Com relação à aplicação da filtragem Gaussiana, ela tem a função de suavizar a região de borda das formas e aproximar os perfis de textura de uma distribuição normal.

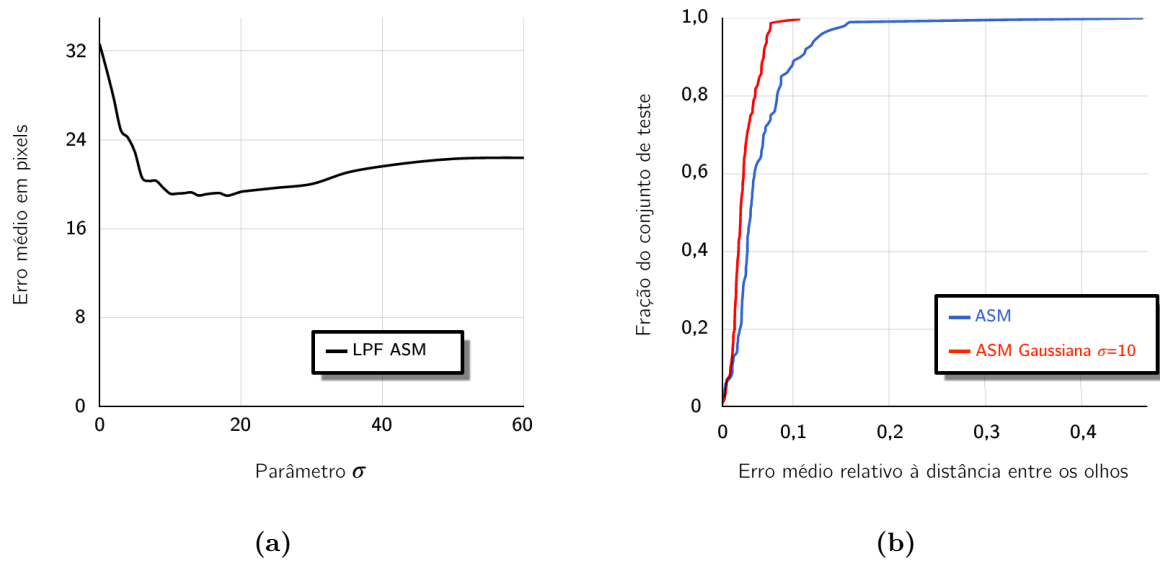
O conjunto de treinamento foi explorado para determinar um valor  $\sigma = 10$  ideal para suavizar a imagem. Além disso, foi realizada uma série de experimentos com o conjunto de teste para avaliar o comportamento do LPF ASM com diferentes valores para esse parâmetro. O gráfico da Figura 53a mostra que baixos valores para  $\sigma$  não são capazes de suavizar os perfis de textura a ponto de gerar resultados satisfatórios para as localizações dos pontos. À medida que o valor é aumentado, as localizações passam a ser mais precisas até ficarem estáveis.

A Figura 53b estabelece um comparativo entre o ASM clássico e uma versão implementada com a filtragem Gaussiana com o melhor valor observado para  $\sigma$ . O gráfico cumulativo mostra no eixo horizontal o erro de localização em função da distância entre os olhos do indivíduo. Para o eixo vertical, os valores indicam a proporção do conjunto de teste avaliado e variam de 0 (nenhuma amostra avaliada) a 1 (todas as amostras avaliadas).

O erro médio apresentado pelo ASM clássico ficou em torno de 6,84% da distância entre os olhos. A versão que utiliza a filtragem Gaussiana apresentou erro médio igual a 4,77% da mesma distância. A melhora foi de 30,3%, ou seja, em um caso que o ASM clássico erra a marcação por 10 *pixels*, a versão com a filtragem erra por 6,9 *pixels* em média.

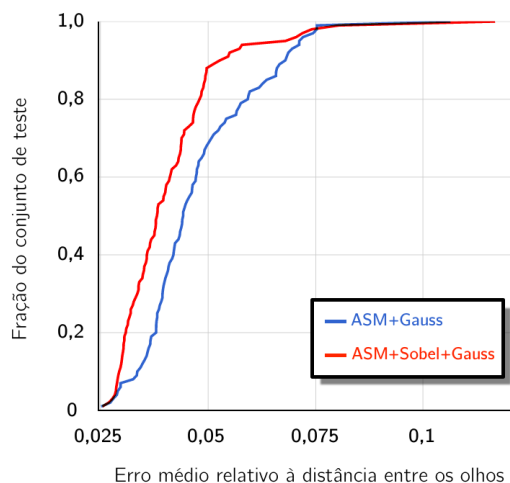
Outra alteração externa ao algoritmo ASM proposta em nossa metodologia é a utilização do operador Sobel para extrair o gradiente dos perfis. No ASM clássico são utilizadas as diferenças de intensidade ao longo do perfil de textura. Nossa proposta é aplicar o operador Sobel para refinar o gradiente ao longo do perfil. O gráfico da Figura 54





**Figura 53** – Resultados do LPF ASM com diferentes valores para o desvio padrão da filtragem Gaussiana. (a) Erro médio do LPF ASM com diferentes valores para  $\sigma$ . (b) Comparação do ASM com e sem filtragem ao longo de todo o conjunto de teste.

compara o ASM implementado com a filtragem Gaussiana com uma versão utilizando as mesmas imagens operadas com o filtro Sobel.



**Figura 54** – Comparação entre o ASM clássico e o ASM implementado com Sobel.

Neste caso, o erro médio de localização do ASM implementado com o filtro Sobel é de 4,12% da distância entre os olhos. Comparada ao ASM implementado com a filtragem Gaussiana, esta alteração foi capaz de melhorar o resultado de localização média em 13,5%.

Os resultados dos testes com esses parâmetros externos ao ASM evidenciam a possibilidade de se obter melhoras no algoritmo a partir da textura, modificando a maneira como são determinadas novas localizações para os pontos da forma. Além disso, é necessário determinar valores ideais para os parâmetros internos do algoritmo que são responsáveis por mudanças não somente da textura, mas também da forma.

## 6.2.2 Resultados de Parâmetros Internos ao ASM

Para aproveitar todo o potencial do ASM na metodologia proposta foi preciso realizar testes com diferentes combinações para os valores dos parâmetros internos do método. Assim, foi possível determinar qual a configuração ideal para a aplicação do método tendo em vista o conjunto de imagens usado.

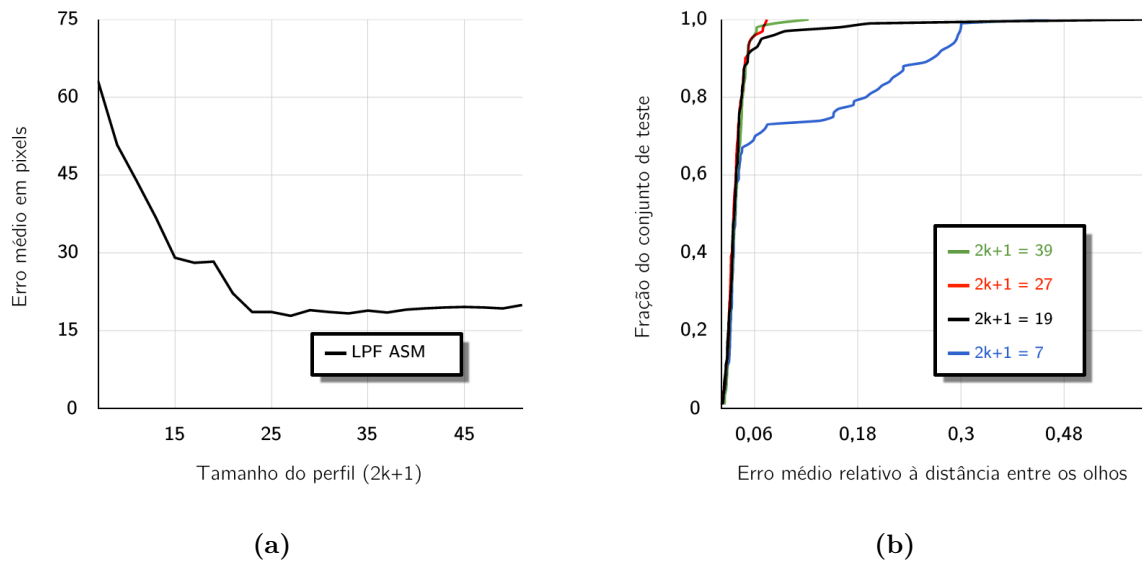
Ainda com relação à textura, existem algumas considerações importantes a serem relacionadas. Em nossa metodologia, é utilizado um perfil unidimensional de textura a partir da imagem operada com o filtro Sobel. Algumas propostas estudadas sugerem que o perfil de textura seja bidimensional para capturar mais informação sobre as bordas da forma. Milborrow (2007) explica que esta proposta não funciona com todos os pontos, então sugere que apenas alguns usem esse tipo de perfil. Como essa abordagem não apresentou resultados consistentes para toda a forma, nossa metodologia fez uso apenas de perfis unidimensionais originalmente propostos por Cootes et al. (1995).

Foram realizados testes para se avaliar uma característica essencial dos perfis unidimensionais: o tamanho do perfil. A Figura 55a mostra como o tamanho do perfil pode influenciar na precisão do método. A precisão do algoritmo aumenta à medida que aumentamos o tamanho do perfil de textura coletado. A partir de um certo valor, o aumento passa a não exercer grandes melhoras na solução do problema.

A Figura 55b estabelece uma comparação do LPF ASM com diversos valores para o tamanho do perfil. É possível observar a partir dos dois gráficos, que a precisão aumenta progressivamente até  $2k + 1 = 19$ , depois desse valor, o erro se mantém estável.

Existe outro aspecto interno ao ASM que deve ser destacado. Ao contrário dos parâmetros anteriores, que diziam respeito à textura do pontos, este aspecto influencia na maneira como a forma é ajustada. Como mostrado na Seção 4.3.2, a forma gerada na etapa de busca por novas posições deve respeitar certos limites para garantir que ela seja compatível com o conjunto de treinamento. Os valores apontados como ideais para definir essas restrições foram apresentados na Equação 4.16.

Contudo, esses valores não apresentaram bons resultados quando observados no conjunto de treinamento. No conjunto de teste, a avaliação do algoritmo com o vetor  $\vec{b}$  seguindo essas restrições também não apresentou resultados satisfatórios. A Figura 56a mostra como o erro de localização evolui à medida que “relaxamos” o valor das restrições.



**Figura 55** – Resultados do LPF ASM com diferentes valores de tamanho para o perfil de textura. (a) Erro médio do LPF ASM com diferentes valores para  $k$ . (b) Comparação do LPF ASM com 4 valores para  $k$  ao longo de todo o conjunto de teste.

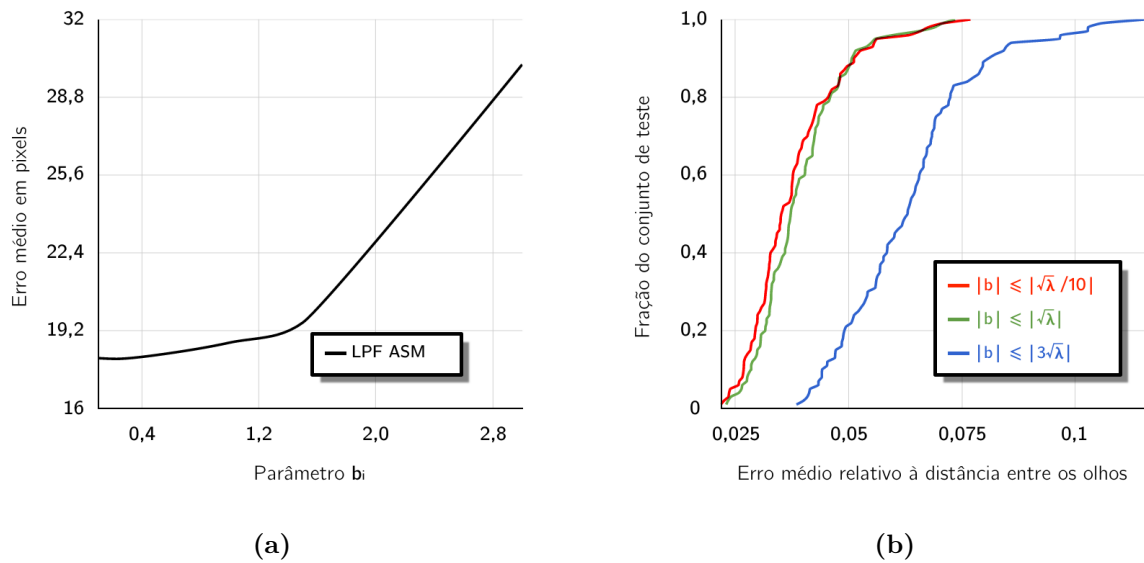
Os valores ideais para o fator de multiplicação das restrições variam de 0, 1 a 1, 3, a partir daí, a precisão decresce progressivamente.

A Figura 56b mostra uma comparação do algoritmo com dois valores satisfatórios para o fator de multiplicação das restrições e o valor recomendado por Cootes et al. (1995). Em todo o conjunto de testes a utilização das restrições originais se mostrou ineficiente.

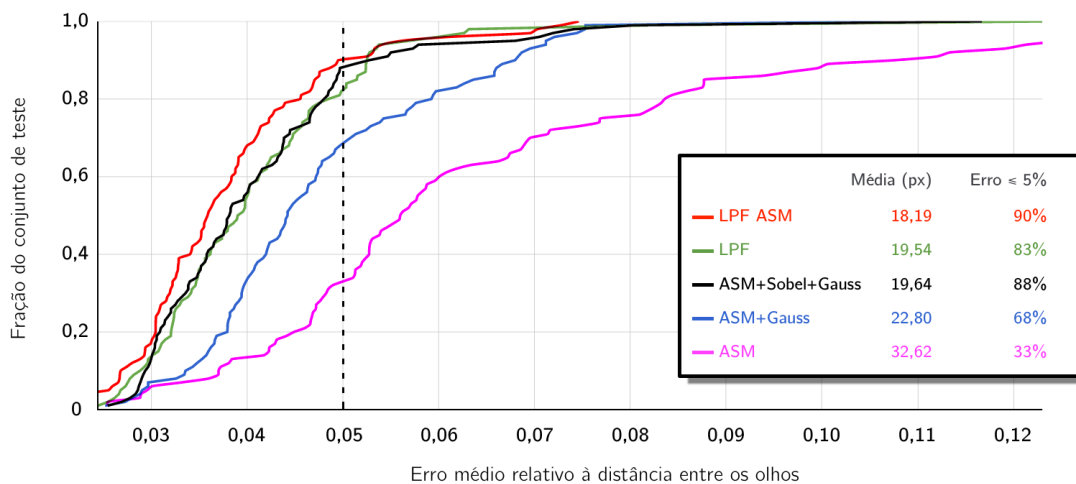
É importante notar que essas restrições são fortemente influenciadas pelo conjunto de treinamento utilizado para construir o modelo do algoritmo. A construção de um modelo cujas formas observadas apresentam pequena variação, tende a apresentar melhores resultados quando combinada com restrições mais “relaxadas”, mesmo considerando os valores de  $\vec{\lambda}$ . Para o caso de o conjunto de treinamento apresentar grande variação nas formas (fato bastante comum em formas faciais) as restrições não podem ser tão brandas, pois o modelo já apresenta bastante riqueza de variações devido às formas observadas.

### 6.2.3 LPF ASM Ideal

Após analisadas alterações externas e parâmetros internos do método proposto em combinação com o ASM, foi possível determinar as melhores configurações para a solução do problema. O gráfico exposto na Figura 57 relaciona os principais resultados obtidos das alterações propostas, assim como os resultados para o algoritmo ASM clássico.



**Figura 56** – Resultados do LPF ASM com diferentes valores para o fator de multiplicação das restrições. (a) Erro médio do LPF ASM com diferentes valores para as restrições. (b) Comparação do LPF ASM com 3 valores para as restrições ao longo de todo o conjunto de teste.



**Figura 57** – Comparativo das variantes do ASM e LPF ASM ao longo de todo o conjunto de teste.

As curvas do gráfico indicam a relação entre a média do erro<sup>2</sup> na localização dos 21 pontos de referência e o percentual de amostras avaliadas.

Os trabalhos de Bottino e Cumani (2008) e Zhu e Ramanan (2012) propuseram a observação dos resultados sob uma ótica diferente. Eles procuraram classificar os seus

<sup>2</sup> Como exposto no gráfico, o erro médio de uma amostra é calculado em função da distância entre olhos do indivíduo presente na imagem.

resultados de localização de pontos em dois conjuntos. Um conjunto de resultados satisfatórios onde o erro se encontra dentro de um determinado limite e outro onde o erro ultrapassa o limite estipulado. Bottino e Cumani (2008) definiram um limite de maneira absoluta, onde localizações com erro menor ou igual a 5 *pixels* são consideradas aceitáveis. Zhu e Ramanan (2012) estipularam um erro relativo à média entre a altura e a largura facial. Para eles, caso o erro se encontre em até 5% desse valor, o resultado é considerado satisfatório.

No gráfico da Figura 57, considerando um erro de até 5% da distância entre os olhos, nossa proposta obteve um acerto de 90% na localização média dos 21 pontos. Isso significa que a localização dos 21 pontos em 180 das 200 imagens usadas para teste apresentou um erro inferior ou igual ao citado. No caso do ASM clássico, esse valor foi de apenas 33%.

Verificando as modificações externas ao algoritmo, também é possível observar uma expressiva melhora na localização dos pontos. Para o ASM usado com uma imagem previamente filtrada com um *kernel* Gaussiano, observou-se que o acerto ficou em torno de dois terços do conjunto. A simples adição do operador Sobel para extrair o gradiente dos perfis de textura elevou substancialmente a taxa de acerto do método para 88%.

O LPF baseado no ASM, porém sem aproveitar os resultados gerados pela estratégia de busca tradicional, apresentou melhora discreta com relação ao ASM acrescido das modificações externas (Sobel e filtragem Gaussiana) apresentadas neste capítulo. Contudo, a versão completa de nossa proposta que aproveita os melhores resultados gerados na busca com base na função objetivo apresentada ao final do Capítulo 5, mostrou-se consistentemente superior a todas as versões modificadas do algoritmo analisadas em nosso trabalho.

A análise dos resultados ponto a ponto, exibidos na Tabela 7, mostra que o erro médio de localização apresenta uma diminuição de 44% na comparação do ASM clássico com a nossa proposta de aprimoramento do mesmo. Além disso, é possível observar que ambas as propostas apresentaram maior dificuldade na localização dos pontos extremos verticalmente: tríquio, os arcos das duas sobrancelhas, supramental e mentoniano.

Uma explicação imediata para estes resultados guarda relação com a variação da forma facial verticalmente. Esta característica pode ser observada após o alinhamento do conjunto de treinamento, onde os pontos apresentam maior variação no eixo vertical. Isso indica que os pontos extremos verticalmente têm maior chance de serem deslocados após as etapas de ajuste da forma.

A comparação da proposta de adição de algumas modificações externas ao ASM com a nossa versão final também aponta um redução no erro médio dos pontos, o que indica que as alterações internas propostas para o ASM também apresentaram bons resultados. A nossa proposta foi superior na localização média de quase todos os pontos, como exposto

**Tabela 7** – Comparativo, ponto a ponto, entre o ASM clássico e o LPF ASM.

Sigla	Erro ASM		Erro LPF ASM		Evolução
	<i>pixels</i>	Relativo	<i>pixels</i>	Relativo	
Tr	58,42	12,28%	38,44	8,08%	-34,20%
Gl	32,57	6,85%	18,41	3,87%	-43,48%
Ns	28,28	5,93%	16,40	3,44%	-42,01%
ASd	38,35	8,04%	21,90	4,59%	-42,89%
ASe	37,40	7,85%	19,72	4,14%	-47,27%
LLd	30,81	6,49%	13,11	2,76%	-57,45%
LLe	30,82	6,47%	15,09	3,17%	-51,04%
CLd	32,64	6,84%	14,12	2,96%	-56,74%
CLe	31,78	6,65%	14,90	3,12%	-53,12%
CMd	28,31	5,97%	13,56	2,86%	-52,10%
CMe	28,09	5,87%	12,77	2,67%	-54,54%
Od	31,83	6,65%	18,43	3,85%	-42,10%
Oe	28,84	6,00%	17,79	3,70%	-38,31%
ANd	27,23	5,75%	17,32	3,66%	-36,39%
ANe	25,89	5,41%	15,69	3,28%	-39,40%
Sn	24,91	5,25%	16,71	3,52%	-32,92%
LS	27,32	5,72%	16,95	3,55%	-37,96%
St	27,31	5,74%	15,28	3,21%	-44,05%
LI	31,87	6,65%	18,93	3,95%	-40,60%
Sm	38,71	8,11%	23,30	4,88%	-39,81%
Mt	43,75	9,17%	22,70	4,76%	-48,11%
<b>Média</b>	<b>32,63</b>	<b>6,84%</b>	<b>18,17</b>	<b>3,81%</b>	<b>-44,50%</b>

na Tabela 8, e a evolução média na precisão ficou em torno de 7%.

Assim como na comparação anterior, os pontos extremos verticalmente apresentaram os maiores desafios. Por outro lado, os dois métodos exibiram os melhores resultados nas localizações dos pontos centrais que, observados no conjunto de formas alinhadas, se encontram mais próximos.

O problema da grande variância para alguns pontos (extremos verticais), que resulta em respectivas localizações precárias, é oriundo de um conjunto de treinamento bastante diverso. A base de imagens utilizada para construção e avaliação do modelo é composta por indivíduos de várias etnias, de todas as faixas etárias e de ambos os gêneros. Além disso, a forma modelada leva em conta todos os 145 pontos destacados no Capítulo 5. Isso tem impacto direto no alinhamento das formas, pois alguns pontos fatalmente são sacrificados durante essa etapa. Eles acabam apresentando grande variância dentro do conjunto alinhado.

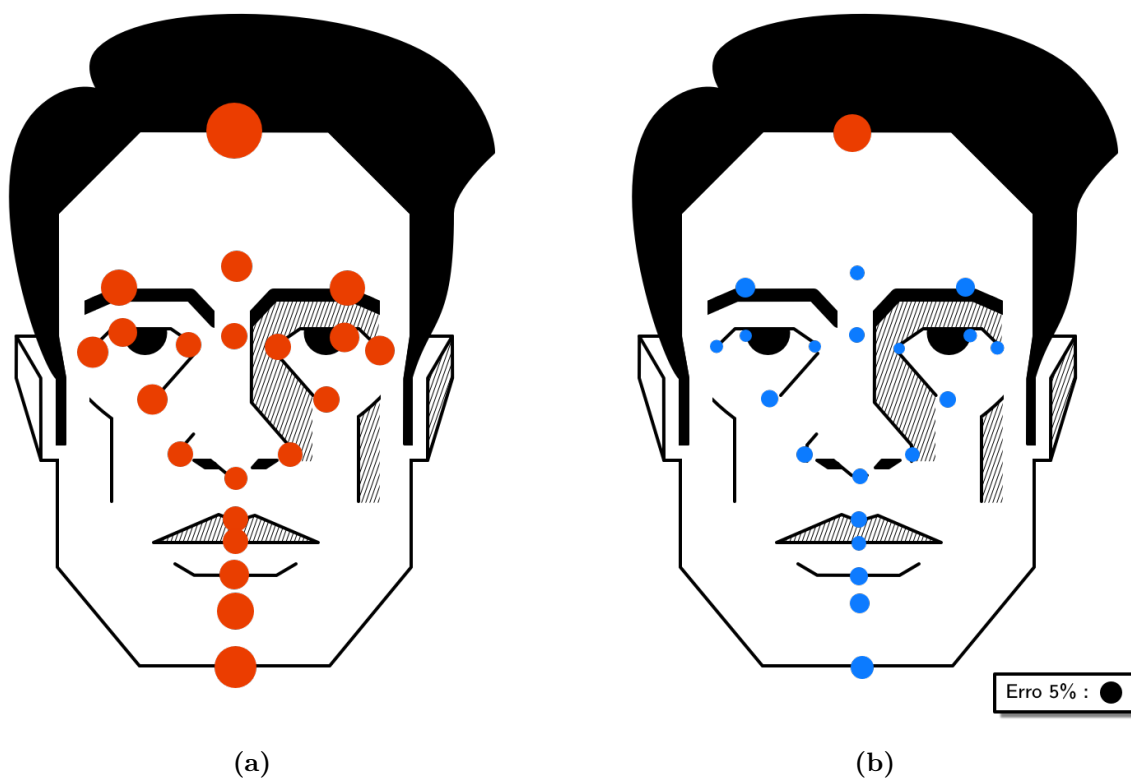
Uma alternativa interessante para mitigar este efeito seria a divisão da forma em

**Tabela 8** – Comparativo, ponto a ponto, entre o ASM com as modificações externas propostas e o LPF ASM.

Sigla	Erro ASM Ext.		Erro LPF ASM		Evolução
	<i>pixels</i>	Relativo	<i>pixels</i>	Relativo	
Tr	43,13	9,07%	38,44	8,08%	-10,87%
Gl	16,90	3,55%	18,41	3,87%	8,93%
Ns	17,57	3,69%	16,40	3,44%	-6,66%
ASd	22,20	4,65%	21,90	4,59%	-1,35%
ASe	21,86	4,59%	19,72	4,14%	-9,79%
LLd	14,15	2,98%	13,11	2,76%	-7,35%
LLe	16,55	3,48%	15,09	3,17%	-8,82%
CLd	14,29	3,00%	14,12	2,96%	-1,19%
CLe	16,99	3,56%	14,90	3,12%	-12,30%
CMd	13,95	2,94%	13,56	2,86%	-2,80%
CMe	13,94	2,91%	12,77	2,67%	-8,39%
Od	19,34	4,04%	18,43	3,85%	-4,71%
Oe	20,47	4,26%	17,79	3,70%	-13,09%
ANd	17,94	3,79%	17,32	3,66%	-3,46%
ANe	18,04	3,77%	15,69	3,28%	-13,03%
Sn	18,85	3,97%	16,71	3,52%	-11,35%
LS	18,06	3,78%	16,95	3,55%	-6,15%
St	17,29	3,63%	15,28	3,21%	-11,63%
LI	20,24	4,22%	18,93	3,95%	-6,47%
Sm	25,64	5,37%	23,30	4,88%	-9,13%
Mt	25,02	5,25%	22,70	4,76%	-9,27%
<b>Média</b>	<b>19,64</b>	<b>4,12%</b>	<b>18,17</b>	<b>3,81%</b>	<b>-7,09%</b>

várias partes. Cada pedaço do modelo seria capaz de representar um único elemento facial ao invés de ter apenas um modelo para descrever toda a face. Dessa maneira, a variância dos pontos extremos poderia ser reduzida, uma vez que os elementos faciais isoladamente não apresentam grande variância no conjunto de treinamento. Este caminho já foi apontado por Davatzikos, Tao e Shen (2003) e Cerrolaza, Villanueva e Cabeza (2012) para localizar formas em imagens de ressonância magnética com auxílio do ASM.

A Figura 58 mostra o erro médio na localização dos pontos em uma ilustração de face frontal. Nas imagens, os centros dos círculos apontam as localizações corretas de cada ponto de referência. O raio de cada um indica o valor do erro de localização médio do respectivo ponto. Os círculos vermelhos indicam pontos cujos erros médios ultrapassam o valor de 5% da distância entre os olhos. Os azuis se encontram dentro deste limite.



**Figura 58** – Demonstração gráfica do erro médio na localização dos pontos frontais. (a) Resultado para o ASM. (b) Resultado para o LPF ASM.

### 6.3 Considerações Finais

O método de localização de pontos do perfil apresentou bons resultados na localização da maioria dos pontos. Contudo, os dois pontos extremos verticalmente, tríquio e cervical, tiveram péssimos resultados. Essa parte da metodologia ainda precisa ser aprimorada de maneira que seu uso esteja adequado às necessidades de uma análise pré-operatória.

O método que aproveita os conceitos do ASM para resolver o problema da localização de pontos frontais apresentou resultados mais regulares. A nossa proposta de alteração do ASM foi bem sucedida na intenção de melhorar o algoritmo original aplicado ao problema exposto. Todas as alterações realizadas apresentaram alguma contribuição na localização dos pontos. Entretanto, a análise dos resultados mostrou que ainda existe espaço para a aplicação de melhorias.



## Conclusões e Trabalhos Futuros

A localização e identificação automáticas de pontos de referência da face são tarefas com inúmeras possibilidades de aplicação. Tais soluções são largamente associadas com a Biometria (i.e. identificação facial ou de expressões), com a computação gráfica (construção de modelos faciais), etc. Contudo, essas soluções ainda não são exploradas na área médica com a mesma intensidade.

As análises pré-operatórias, essenciais na preparação das cirurgias plásticas, exigem um rigor, já na escolha dos pontos, que não é observado em outros tipos de aplicação. Muitos trabalhos que propõem o uso de pontos faciais com outras motivações pecam por ignorar características importantes e por isso não são adequados para resolver o problema proposto neste trabalho.

Um exemplo disso é a abordagem do perfil facial. São poucos os trabalhos que se preocupam em explorar tal perspectiva. Contudo, não se pode negligenciar o contorno do perfil em uma análise pré-operatória, pois o cálculo de muitas métricas utilizadas na avaliação estética depende de uma localização precisa dos pontos nessa perspectiva.

Uma importante diferença entre trabalhos motivados por outros tipos de aplicação e o nosso é o conjunto de pontos faciais que foi abordado. Tais aspectos foram tratados com extremo cuidado pois são eles que guiam toda a análise estética. Nenhum ponto facial foi escolhido sem que haja um objetivo claro para seu uso. Além disso, este trabalho introduziu conceitos que possibilitaram a localização e identificação de novos pontos faciais.

### Conclusões

O LPP proposto é uma alternativa simples e eficaz para a localização de pontos faciais do perfil. Ele introduziu uma nova proposta para localizar pontos como o rínio e a columela. O testes indicam que o algoritmo pode ser comparado, para alguns pontos, com uma localização manual feita por especialista. Por outro lado, as localizações de tríquio e cervical se provaram difíceis e apresentaram resultados ruins.

Uma das dificuldades mais marcantes está relacionada justamente com a diversidade étnica dos indivíduos. As definições dos pontos faciais encontradas na literatura funcionam como um guia para cirurgias plásticas. Entretanto, dadas as diferenças das faces humanas, tais definições nem sempre são aplicáveis. Em alguns casos, as características eram bem óbvias, mas em outros elas eram extremamente sutis. Superar este obstáculo foi um dos maiores desafios deste trabalho.

Com relação a nossa proposta de modificação do ASM, ela se mostrou efetivamente

mais robusta que o método original. Ainda assim, de acordo com o que foi observado, algumas melhorias devem ser propostas visando resolver as deficiências apontadas no Capítulo 6. Contudo, sua aplicação já pode ser aproveitada para resolver o problema proposto.

## Limitações

Este trabalho propôs uma metodologia para localizar pontos faciais em imagens para o planejamento de cirurgias plásticas. Apesar disso, ele tem restrições com relação ao que pode ser feito dentro da etapa de planejamento cirúrgico. A seguir estão resumidas as limitações do trabalho:

1. a proposta não realiza o cálculo de métricas de avaliação estética ou análise pré-operatória;
2. a localização dos pontos é restrita a imagens faciais frontais e de perfil em posição natural. Faces rotacionadas fogem ao escopo da proposta;
3. as imagens devem ser adquiridas em ambiente controlado e devem ser neutras em relação a expressões faciais;
4. a metodologia não é capaz de localizar ou reconhecer a face. A localização é restrita ao conjunto de pontos de referência;
5. a metodologia não é capaz de determinar a posição facial.

## Publicações

A metodologia responsável pela localização de pontos do perfil (LPP) desenvolvida inicialmente como exposto na Seção 6.1.2 apresentou resultados bastante promissores e foi publicada na trilha principal da *IEEE International Conference on Systems, Man, and Cybernetics* de 2014 em San Diego (FREITAS; AIRES; CAMPELO, 2014). A versão atual do algoritmo, com os resultados apresentados neste capítulo, foi publicada na *28th SIBGRAPI Conference on Graphics, Patterns and Images* em 2015, Salvador, também na trilha principal (FREITAS; AIRES; CAMPELO, 2015).

O localizador de ponto frontais baseado no ASM foi apresentado na categoria poster do *8th World Congress of Facial Plastic Surgery* realizado em 2016 no Rio de Janeiro. Um outro trabalho, fruto da metodologia proposta foi apresentado neste mesmo evento na forma de poster e foi escolhido como o 2º melhor trabalho da categoria. Este outro, trata-se de uma implementação do LPP em dispositivos móveis.

## Trabalhos Futuros

Um dos maiores problemas na etapa de localização de pontos do perfil considera a extração do contorno. Uma vez que o método para a localização desses pontos é baseado em detecção de bordas, podem surgir alguns problemas com o fundo da imagem. Dependendo de onde ela é adquirida, o fundo pode se misturar com o perfil, causando terríveis efeitos nos resultados. Trabalhos futuros devem tornar a extração de contorno mais robusta a essas variações. Trabalhos como *Active Contour Models* podem ser usados para resolver problemas similares.

A análise dos resultados do localizador de pontos frontais baseado no ASM apresentou um problema imediato. O algoritmo encontrou bastante dificuldade para localizar os pontos extremos verticalmente. Uma proposta de solução para isso consiste em dividir a forma facial nos elementos que compõem a face. Cada elemento deve corresponder a uma forma (olhos, boca, nariz, etc). A ideia é que esta estratégia reduza a variância observada no conjunto de treino. Uma outra alteração para trabalhos futuros é a divisão do conjunto de treino de acordo com a etnia. Isso pode ser capaz de criar modelos faciais específicos, que podem ajudar na etapa de busca.

A intenção deste trabalho foi propor uma metodologia que facilite uma das etapas de planejamento de cirurgias plásticas faciais. Contudo, é importante lembrar que a localização e identificação dos pontos de referência é apenas a primeira etapa de um processo de análise estética da face. Após concluída esta fase, existem alguns caminhos a serem trilhados no objetivo de remover toda a complexidade de um planejamento para cirurgias faciais:

- **Avaliação estética** - A etapa seguinte a um processo de marcação manual de pontos para planejamento cirúrgico é o cálculo de métricas para avaliação estética. Uma proposta futura deve relacionar as principais métricas e aplicá-las de acordo com os pontos localizados e identificados;
- **Definição de procedimentos** - Uma outra possibilidade é a definição automática de procedimentos adequados a um paciente. Neste caso, poderia ser desenvolvida uma metodologia para adiantar o trabalho do cirurgião e oferecer possibilidades de intervenções cirúrgicas de acordo com a localização dos pontos e as avaliações das métricas existentes;
- **Sistema inteligente** - Tendo em vista uma interação entre o usuário e o computador, o médico pode sugerir correções de todas as ordens, seja na localização dos pontos, seja na definição de procedimentos. Uma outra proposta para trabalhos futuros seria de um sistema capaz de observar suas falhas e buscar corrigir os erros em iterações posteriores;

- **Outras perspectivas de análise** - Neste trabalho foram exploradas as duas perspectivas principais de análise estética. Contudo, a literatura médica elenca algumas outras perspectivas para avaliação estética da face. Trabalhos futuros podem abordar a face por uma perspectiva diagonal, por exemplo, já que algumas métricas são específicas para este ponto de vista.

# Referências

ARIAS-CASTRO, E.; DONOHO, D. L. Does median filtering truly preserve edges better than linear filtering? *The Annals of Statistics*, The Institute of Mathematical Statistics, v. 37, n. 3, p. 1172–1206, jun. 2009. Citado na página 29.

BATISTA, J. P. Locating facial features using an anthropometric face model for determining the gaze of faces in image sequences. In: KAMEL, M.; CAMPILHO, A. (Ed.). *4th International Conference Image Analysis and Recognition (ICIAR '07)*. [S.l.]: Springer Berlin Heidelberg, 2007. p. 839–853. ISBN 978-3-540-74260-9. Citado na página 16.

BELHUMEUR, P. N.; JACOBS, D. W.; KRIEGMAN, D. J.; KUMAR, N. Localizing parts of faces using a consensus of exemplars. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '11)*. [S.l.: s.n.], 2011. p. 545–552. ISSN 1063-6919. Citado na página 39.

BOTTINO, A.; CUMANI, S. Robust identification of face landmarks in profile images. In: *Proceedings of the 12th WSEAS international conference on Computers*. [S.l.: s.n.], 2008. p. 107–114. Citado 11 vezes nas páginas 9, 13, 7, 17, 18, 19, 21, 78, 79, 87 e 88.

BOTTINO, A.; SIMONE, M. D.; LAURENTINI, A.; SFORZA, C. A new 3-d tool for planning plastic surgery. *IEEE Transactions on Biomedical Engineering*, v. 59, n. 12, p. 3439–3449, dez. 2012. ISSN 0018-9294. Citado na página 18.

BOWDEN, R.; MITCHELL, T.; SARHADI, M. Non-linear statistical models for the 3d reconstruction of human pose and motion from monocular image sequences. *Image and Vision Computing*, v. 18, n. 9, p. 729 – 737, 2000. ISSN 0262-8856. Citado na página 45.

BRESENHAM, J. Algorithm for computer control of a digital plotter. *IBM Systems Journal*, v. 4, n. 1, p. 25–30, 1965. ISSN 0018-8670. Citado 3 vezes nas páginas 10, 39 e 40.

CANNY, J. A computational approach to edge detection. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*. [S.l.: s.n.], 1986. p. 679–698. Citado na página 34.

CERROLAZA, J. J.; VILLANUEVA, A.; CABEZA, R. Hierarchical statistical shape models of multiobject anatomical structures: Application to brain mri. *IEEE Transactions on Medical Imaging*, v. 31, n. 3, p. 713–724, mar. 2012. ISSN 0278-0062. Citado 4 vezes nas páginas 9, 12, 13 e 90.

COOTES, T.; TAYLOR, C.; COOPER, D.; GRAHAM, J. Active shape models-their training and application. *Computer Vision and Image Understanding*, v. 61, n. 1, p. 38 – 59, 1995. ISSN 1077-3142. Citado 14 vezes nas páginas 9, 7, 10, 11, 21, 41, 42, 44, 46, 49, 51, 69, 85 e 86.

COOTES, T. F.; EDWARDS, G. J.; TAYLOR, C. J. Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE Computer Society, Los Alamitos, CA, USA, v. 23, n. 6, p. 681–685, 2001. ISSN 0162-8828. Citado 3 vezes nas páginas 9, 13 e 14.

- COOTES, T. F.; TAYLOR, C. *Statistical Models of Appearance for Computer Vision*. [S.l.], 2004. Citado 4 vezes nas páginas 11, 37, 72 e 83.
- DAVATZIKOS, C.; TAO, X.; SHEN, D. Hierarchical active shape models, using the wavelet transform. *IEEE Transactions on Medical Imaging*, v. 22, n. 3, p. 414–423, mar. 2003. ISSN 0278-0062. Citado 3 vezes nas páginas 11, 12 e 90.
- DRYDEN, I.; MARDIA, K. *Statistical Shape Analysis*. [S.l.]: Wiley, 1998. (Wiley Series in Probability & Statistics). ISBN 9780471958161. Citado 2 vezes nas páginas 26 e 69.
- DU, C.; WU, Q.; YANG, J.; WU, Z. Svm based asm for facial landmarks location. In: *8th IEEE International Conference on Computer and Information Technology (CIT '08)*. [S.l.: s.n.], 2008. p. 321–326. Citado 4 vezes nas páginas 9, 11, 12 e 47.
- FREITAS, R. T.; AIRES, K. R. T.; CAMPELO, V. E. S. Locating facial landmarks towards plastic surgery. In: *28th SIBGRAPI Conference on Graphics, Patterns and Images*. [S.l.: s.n.], 2015. p. 219–225. ISSN 1530-1834. Citado na página 93.
- FREITAS, R. T.; AIRES, K. T.; CAMPELO, V. S. Automatic location of facial landmarks for plastic surgery procedures. In: *IEEE International Conference on Systems, Man and Cybernetics (SMC '14)*. [S.l.: s.n.], 2014. p. 1444–1449. Citado na página 93.
- GONZALEZ, R.; WOODS, R. *Digital Image Processing*. [S.l.]: Pearson Education, 2011. ISBN 9780133002324. Citado 3 vezes nas páginas 27, 62 e 72.
- HUANG, T.; YANG, G.; TANG, G. A fast two-dimensional median filtering algorithm. *IEEE Transactions on Acoustics, Speech and Signal Processing*, v. 27, n. 1, p. 13–18, Feb 1979. ISSN 0096-3518. Citado na página 29.
- INC, A. C. *The Swift Programming Language (Swift 2.2)*. [S.l.], 2014. Citado na página 77.
- INC, A. C.; DAVIDSON, J. D. *Learning Cocoa with Objective-C*. [S.l.]: O'Reilly Media, 2002. ISBN 9781449399696. Citado na página 77.
- KASS, M.; WITKIN, A.; TERZOPOULOS, D. Snakes: Active contour models. *International journal of computer vision*, Springer, v. 1, n. 4, p. 321–331, 1988. Citado 2 vezes nas páginas 42 e 81.
- MARTINEZ, B.; VALSTAR, M.; BINEFA, X.; PANTIC, M. Local evidence aggregation for regression-based facial point detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 35, n. 5, p. 1149–1163, mai. 2013. ISSN 0162-8828. Citado 2 vezes nas páginas 17 e 39.
- MILBORROW, S. *Locating Facial Features with Active Shape Models*. Dissertação (Mestrado) — Faculty of Engineering, University of Cape Town, nov. 2007. Citado 9 vezes nas páginas 7, 11, 12, 39, 47, 69, 72, 83 e 85.
- PAPPAS, N. Plato's aesthetics. In: ZALTA, E. N. (Ed.). *The Stanford Encyclopedia of Philosophy*. Summer 2015. [S.l.: s.n.], 2015. Citado na página 2.
- POLICLETO. *Estátua de Doriforo*. 1797 (Excavada). Estátua de jovem guerreiro. Original perdida. Cópia em mármore feita durante o império Romano. Museu Arqueológico Nacional de Nápoles. 2,12 m de alt. Citado 2 vezes nas páginas 9 e 3.

- POWELL, N.; HUMPHREYS, B. *Proportions of the Aesthetic Face*. [S.l.]: Thieme Medical Pub, 1984. (The American Academy of Facial Plastic and Reconstructive Surgery). ISBN 9780865771178. Citado na página 22.
- PRATT, W. *Digital Image Processing: PIKS Scientific Inside*. [S.l.]: Wiley, 2007. ISBN 9780470097441. Citado 2 vezes nas páginas 29 e 31.
- PRATT, W. K. Semiannual technical report. In: \_\_\_\_\_. [S.l.]: University of Southern California, 1975. cap. Median Filtering, p. 116 – 123. Citado na página 29.
- PRENDERGAST, P. M. Advanced surgical facial rejuvenation. In: \_\_\_\_\_. 1. ed. [S.l.]: Springer-Verlag Berlin Heidelberg, 2012. cap. Facial Proportions, p. 15–22. Citado 5 vezes nas páginas 9, 10, 6, 23 e 25.
- PREWITT, J. M. S. Object enhancement and extraction. In: *Picture Processing and Psychopictorics*. [S.l.]: Academic Press, 1970. Citado na página 34.
- REISSELL, L. M. Wavelet multiresolution representation of curves and surfaces. *Graphical Models and Image Processing*, v. 58, n. 3, p. 198 – 217, 1996. ISSN 1077-3169. Citado na página 11.
- SARTWELL, C. Beauty. In: ZALTA, E. N. (Ed.). *The Stanford Encyclopedia of Philosophy*. Spring 2014. [S.l.: s.n.], 2014. Citado 2 vezes nas páginas 2 e 5.
- SENARATNE, R.; HALGAMUGE, S. Optimised landmark model matching for face recognition. In: *7th International Conference on Automatic Face and Gesture Recognition (FGR '06)*. [S.l.: s.n.], 2006. p. 6 pp.–125. Citado na página 14.
- SENARATNE, R. S.; HALGAMUGE, S. K. Optimal weighting of landmarks for face recognition. *Journal of Multimedia*, v. 1, n. 3, p. 31–41, jun. 2006. Citado 4 vezes nas páginas 9, 7, 14 e 15.
- SESHADRI, K.; SAVVIDES, M. Robust modified active shape model for automatic facial landmark annotation of frontal faces. In: *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems (BTAS '09)*. [S.l.: s.n.], 2009. p. 1–8. Citado na página 12.
- SHI, J.; SAMAL, A.; MARX, D. How effective are landmarks and their geometry for face recognition? *Computer Vision and Image Understanding*, v. 102, n. 2, p. 117–133, mai. 2006. Citado 4 vezes nas páginas 9, 15, 16 e 18.
- SOBEL, I. An isotropic 3x3 image gradient operator. In: *Machine Vision for Three-dimensional Scenes*. [S.l.]: Academic Press, 1990. Citado 2 vezes nas páginas 32 e 34.
- SOHAIL, A.; BHATTACHARYA, P. Detection of facial feature points using anthropometric face model. In: DAMIANI, E.; YÉTONGNON, K.; SCHELKENS, P.; DIPANDA, A.; LEGRAND, L.; CHBEIR, R. (Ed.). *Signal Processing for Image Enhancement and Multimedia Processing*. [S.l.]: Springer US, 2008, (Multimedia Systems and Applications Series, v. 31). p. 189–200. ISBN 978-0-387-72499-7. Citado 3 vezes nas páginas 9, 16 e 17.

- STEINER, D. *Images in Mind: Statues in Archaic and Classical Greek Literature and Thought*. [S.l.]: Princeton University Press, 2001. (Princeton paperbacks). ISBN 9780691094885. Citado na página 2.
- TURK, M.; PENTLAND, A. Face recognition using eigenfaces. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '91)*. [S.l.]: IEEE, 1991. p. 586–591. Citado 2 vezes nas páginas 7 e 9.
- VALSTAR, M. F.; MARTINEZ, B.; BINEFA, X.; PANTIC, M. Facial point detection using boosted regression and graph models. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '10)*. [S.l.: s.n.], 2010. p. 2729–2736. Citado na página 39.
- VAPNIK, V. N. *The Nature of Statistical Learning Theory*. [S.l.]: Springer New York, 2000. Citado na página 12.
- WIMALAWANSA, S.; MCKNIGHT, A.; BULLOCKS, J. M. Socioeconomic impact of ethnic cosmetic surgery: Trends and potential financial impact the african american, asian american, latin american, and middle eastern communities have on cosmetic surgery. *Seminars in Plastic Surgery*, v. 23, p. 159–162, 2009. Citado na página 1.
- YAN, S.; LIU, C.; LI, S. Z.; ZHANG, H.; SHUM, H.-Y.; CHENG, Q. Face alignment using texture-constrained active shape models. *Image and Vision Computing*, v. 21, n. 1, p. 69 – 75, 2003. ISSN 0262-8856. Citado na página 13.
- ZHU, X.; RAMANAN, D. Face detection, pose estimation, and landmark localization in the wild. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*. [S.l.]: IEEE, 2012. p. 2879 – 2886. Citado 7 vezes nas páginas 10, 20, 39, 59, 77, 87 e 88.
- ZIMBLER, M. S.; HAM, J. Cummings otolaryngology - head and neck surgery. In: \_\_\_\_\_. 5. ed. [S.l.]: Elsevier Health Sciences, 2010. cap. Aesthetic Facial Analysis. Citado 7 vezes nas páginas 10, 2, 3, 4, 5, 23 e 25.



# Anexos

# ANEXO A – Definições Estatísticas

## A.1 Médias Aritmética e Ponderada

Em uma distribuição de variáveis, o valor da média indica onde se concentram os dados que a compõem. A média é essencial para vários cálculos estatísticos detalhados neste anexo.

Suponha que o conjunto de vetores  $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$  esteja definido em um espaço  $m$ -dimensional. A média aritmética do conjunto é definida pela soma dos vetores dividida por  $n$ , como na Equação A.1,

$$\vec{\mu} = \frac{1}{n} \sum_{i=1}^n \vec{v}_i, \quad (\text{A.1})$$

e é denotada por  $\vec{\mu} = [\mu_1 \ \mu_2 \ \dots \ \mu_m]^T$ , onde  $\mu_j$  é o valor da média da  $j$ -ésima coordenada.

A média ponderada permite que cada elemento do cálculo seja apresentado com um peso específico  $w_i$ . Ela pode ser calculada como mostra a Equação A.2,

$$\vec{\mu} = \frac{\sum_{i=1}^n \vec{v}_i w_i}{\sum_{i=1}^n w_i}. \quad (\text{A.2})$$

## A.2 Variância

A variância de uma variável é uma medida de dispersão que indica a distância que seus valores se encontram da média do conjunto observado. Para avaliar essa dispersão em uma população de  $n$  vetores de variáveis  $\vec{v}_i$ , deve-se aplicar o cálculo da Equação A.3,

$$\vec{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (\vec{v}_i - \vec{\mu})^2. \quad (\text{A.3})$$

## A.3 Covariância

A covariância é um valor que mede o grau de interdependência entre duas variáveis aleatórias. Suponha que se queira determinar a covariância entre duas variáveis  $V_l$  e  $V_k$ , a

partir de um conjunto com  $n$  amostras. O cálculo da relação de dependência entre elas pode ser realizado com base na Equação A.4,

$$\text{cov}(V_l, V_k) = \frac{1}{n} \sum_{i=1}^n (v_{il} - \mu_l)(v_{ik} - \mu_k), \quad (\text{A.4})$$

onde  $v_{il}$  é o valor da  $i$ -ésima amostra na  $l$ -ésima coordenada, assim como  $v_{ik}$  é o valor da  $i$ -ésima amostra na  $k$ -ésima coordenada.

Em alguns casos pode ser interessante computar todas as covariâncias em uma única estrutura. Isso pode ser feito através do cálculo da matriz de covariância, que sumariza as covariâncias de todas as variáveis, assim como as variâncias de cada uma. A Equação A.5 mostra como é efetuado o cálculo dessa matriz.

$$S = \frac{1}{n-1} \sum_{i=1}^n (\vec{v}_i - \vec{\mu})(\vec{v}_i - \vec{\mu})^T \quad (\text{A.5})$$

Nessa estrutura, a diagonal principal é composta pela variância de cada variável, enquanto que os outros elementos representam a covariância entre as respectivas variáveis.

## A.4 Distância de *Mahalanobis*

Esta medida expressa a distância de uma observação  $\vec{v}_i = [v_{i1} \ v_{i2} \ \dots \ v_{im}]^T$ , onde  $m$  representa a dimensão do vetor, para um conjunto de observações com média  $\vec{\mu} = [\mu_1 \ \mu_2 \ \dots \ \mu_m]^T$  e matriz de covariância  $S$ . A equação a seguir mostra como se calcula a medida:

$$D_M(\vec{v}_i) = \text{Mahalanobis}(\vec{\mu}, \vec{v}_i) = \sqrt{(\vec{v}_i - \vec{\mu})^T S^{-1} (\vec{v}_i - \vec{\mu})}. \quad (\text{A.6})$$