



COPPE/UFRJ

DETECÇÃO DE PONTOS FIDUCIAIS FACIAIS USANDO FILTROS DE  
CORRELAÇÃO E CORRESPONDÊNCIAS ESPACIAIS

Axel Guimarães Hollanda

Dissertação de Mestrado apresentada ao  
Programa de Pós-graduação em Engenharia  
Elétrica, COPPE, da Universidade Federal  
do Rio de Janeiro, como parte dos requisitos  
necessários à obtenção do título de Mestre  
em Engenharia Elétrica.

Orientador: Eduardo Antônio Barros da  
Silva

Rio de Janeiro  
Março de 2011

DETECÇÃO DE PONTOS FIDUCIAIS FACIAIS USANDO FILTROS DE  
CORRELAÇÃO E CORRESPONDÊNCIAS ESPACIAIS

Axel Guimarães Hollanda

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO  
ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE  
ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE  
JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A  
OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA  
ELÉTRICA.

Examinada por:

---

Prof. Eduardo Atônio Barros da Silva, Ph.D.

---

Prof. José Ricardo de Almeida Torreão, Ph.D.

---

Prof. José Gabriel Rodriguez Carneiro Gomes, Ph.D.

RIO DE JANEIRO, RJ – BRASIL

MARÇO DE 2011

Hollanda, Axel Guimarães

Detecção de Pontos Fiduciais Faciais usando Filtros de Correlação e Correspondências Espaciais/Axel Guimarães Hollanda. – Rio de Janeiro: UFRJ/COPPE, 2011.

XVIII, 86 p.: il.; 29, 7cm.

Orientador: Eduardo Antônio Barros da Silva

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Elétrica, 2011.

Referências Bibliográficas: p. 81 – 85.

1. Reconhecimento de Padrões. 2. Detecção de pontos fiduciais. 3. Casamento de Grafos. 4. Biometria. I. Silva, Eduardo Antônio Barros da. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Elétrica. III. Título.

*Aos pais: Lu e Hollanda.*

# Agradecimentos

*No man is an island.*

– John Donne

Início meus agradecimentos com a certeza de que falharei em lembrar a todos. Meu muito obrigado àqueles que, de alguma forma, impactaram na concepção deste trabalho:

- A meus pais, Lu e Hollanda, pelo carinho, confiança, crédito, educação, valores, cuidado e amor.
- A meus irmãos, Breno e Eric, que contem sempre comigo.
- À minha namorada, Fernanda, que com paciência caminhou junto a mim nos últimos meses deste trabalho.
- Ao grande amigo Manuel Cardoso, por sua inspiração.
- Aos bons amigos.
- Ao corpo do Laboratório de Processamento de Sinais da COPPE/UFRJ, por criarem um ambiente mais que propício à realização da atividade da pesquisa.
- Ao professor Eduardo, pela orientação.

Este trabalho foi executado sob o suporte financeiro da Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM).

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

## DETECÇÃO DE PONTOS FIDUCIAIS FACIAIS USANDO FILTROS DE CORRELAÇÃO E CORRESPONDÊNCIAS ESPACIAIS

Axel Guimarães Hollanda

Março/2011

Orientador: Eduardo Antônio Barros da Silva

Programa: Engenharia Elétrica

Os sistemas de detecção de pontos fiduciais faciais via filtros de correlação conhecidos como *Detectores por Produto Interno* (DPIs) vêm sendo utilizados para a busca de padrões isolados em imagens de faces. Esses sistemas fazem uso de informações de alto nível como critério de decisão final para a classificação dos padrões em questão, como distribuições de probabilidade a priori.

Esta dissertação propõe o uso de correspondências espaciais entre os pontos fiduciais da face para a eleição conjunta desses padrões. Dois métodos foram desenvolvidos e aplicados sobre a combinação dos sistemas de detecção local propostos anteriormente. Isso resultou em um sistema complexo sob o paradigma de detecção global de pontos fiduciais faciais.

O primeiro método explora os padrões de simetria na distribuição espacial das saídas dos detectores locais, as quais apontam coordenadas no plano de imagens de faces. O segundo explora o casamento de grafos para a representação global dos pontos fiduciais.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

## FACIAL LANDMARKS DETECTION BY CORRELATION FILTERS AND SPATIAL SIMILARITIES

Axel Guimarães Hollanda

March/2011

Advisor: Eduardo Antônio Barros da Silva

Department: Electrical Engineering

Facial landmarks detection systems using correlation filters known as *Inner Product Detectors* (IPDs) have been used often for the search of isolated patterns over face images. These systems use high level information as decision criteria for pattern classification, such as prior probability distributions.

In this dissertation, we suggest the use of spatial similarities among the facial landmarks for the joint determination of these patterns. Two methods have been developed and applied over a combination of the early local detection systems. This resulted in a complex system under global landmark detection paradigm.

The first method explores the symmetry patterns among the several patterns output by the local detectors for different landmarks. The other one explores graph matching for global landmark representation.

# Sumário

<b>Lista de Figuras</b>	<b>x</b>
<b>Lista de Tabelas</b>	<b>xvi</b>
<b>Lista de Abreviaturas</b>	<b>xviii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Organização da Dissertação . . . . .	3
<b>2 Detecção de Pontos Fiduciais</b>	<b>4</b>
2.1 Técnicas . . . . .	4
2.2 O Detector por Produto Interno (DPI) . . . . .	5
2.2.1 DPI para uma única classe . . . . .	5
2.2.2 DPI generalizado . . . . .	6
2.2.3 DPI com Análise de Componentes Principais (PCA) . . . . .	7
2.2.4 Sistema de Detecção de Pontos Fiduciais Faciais via DPI com Análise de Componentes Principais . . . . .	9
2.3 Considerações Finais . . . . .	13
<b>3 Explorando as Simetrias da Face</b>	<b>14</b>
3.1 Método . . . . .	16
3.1.1 Eixos de Simetria . . . . .	18
3.1.2 Conversão de Base dos Elementos das Nuvens do DPI . . . . .	19
3.1.3 Assimetrias a partir das Marcações de Sabino e Araujo . . . . .	20
3.2 Treinamento do DPI-SIM . . . . .	21
3.3 Teste . . . . .	23
3.4 Análise dos Resultados . . . . .	26
3.4.1 Métrica . . . . .	26
3.4.2 Resultados para a base BioID . . . . .	26
3.4.3 Resultados para a base FERET . . . . .	34
3.4.4 Comparativo de Desempenho do DPI-SIM nas Bases FERET <i>versus</i> BioID . . . . .	42

3.5	Considerações Finais . . . . .	42
<b>4</b>	<b>Representação Global de Pontos Fiduciais Faciais por Grafos</b>	<b>44</b>
4.1	Grafos Elásticos: o <i>Elastic Bunch Graph Matching</i> (EBGM) e sua variante estatística . . . . .	44
4.1.1	EBGM Clássico . . . . .	45
4.1.2	EBGM com Máxima Verossimilhança . . . . .	48
4.2	DPI-GRAPH: o DPI com Casamento de Grafos . . . . .	50
4.2.1	Método . . . . .	51
4.2.2	Agregado de Grafos de Faces . . . . .	52
4.2.3	Treinamento do DPI-GRAPH . . . . .	53
4.2.4	Teste . . . . .	54
4.2.5	Análise dos Resultados . . . . .	58
4.2.6	Resultados para a base BioID . . . . .	58
4.2.7	Resultados para a base FERET . . . . .	66
4.2.8	Comparativo de Desempenho do DPI-GRAPH <i>versus</i> DPI-SIM . . . . .	74
4.3	Considerações Finais . . . . .	78
<b>5</b>	<b>Conclusões</b>	<b>79</b>
5.1	Trabalhos Futuros . . . . .	80
	<b>Referências Bibliográficas</b>	<b>81</b>
<b>A</b>	<b>Grafos: conceitos básicos</b>	<b>86</b>
A.1	Alguns Conceitos de Teoria dos Grafos . . . . .	86

# Lista de Figuras

1.1	Proporções da Face e Olhos, 1489, desenho à mão de Leonardo da Vinci ( <a href="http://drawingsofleonardo.org">http://drawingsofleonardo.org</a> , set/2010). . . . .	1
1.2	Marcações de Pontos Fiduciais Faciais por Sabino e Araujo na base BioID (a) e FERET (b) . . . . .	3
2.1	Exemplos de <i>eigenfaces</i> e <i>eigenpoints</i> . Em 2.1(b), componentes principais a partir de blocos centrados na pupila direita de imagens de faces. . . . .	8
2.2	Visão simplificada da seleção de blocos para classes de positivos e negativos para treinamento do sistema de detecção local de pontos fiduciais DPI-PCA. Em 2.2(a), o bloco centrado no PFF terá suas componentes principais extraídas para treinamento de um DPI $\mathbf{h}_{\phi}$ . Em 2.2(b), os blocos para a classe de negativos selecionados no interior da ROI desse ponto. Uma seleção pseudo-aleatória uniformemente distribuída de cinquenta desses blocos foi proposta em [1]. . . . .	11
2.3	Treinamento dos classificadores locais DPI $\mathbf{h}_{\phi_i}$ para um dos $i = 1, \dots, N$ <i>eigenpoints</i> de um ponto fiducial facial. O classificador AdaBoost também é uma saída desse procedimento, treinado a partir dos resultados do produto interno entre o DPI $\mathbf{h}_{\phi_i}$ e os blocos $B_{\mathbf{z}}$ centrados na coordenada $\mathbf{z}$ . . . . .	12
2.4	Diagrama de blocos do sistema de teste para detecção de um ponto fiducial facial pelo método DPI-PCA. . . . .	12
3.1	As Marcações de PFFs de Sabino e Araujo e as Nuvens de DPI. As marcações originais ( <i>ground truth</i> ) da imagem de face em prova em (a). Grandes reverses do DPI: a ordem e o espalhamento das nuvens em (b). . . . .	14

3.2	A reta de ajuste sobre as marcações originais ( <i>ground truth</i> ) da imagem de face em prova (a). Utilizando o ponto médio das marcações na região dos olhos, traça-se uma segunda reta formando o par de eixos de simetria para essa região (b). Ao se obter uma reta de ajuste sobre as nuvens de DPI para a região dos olhos, tomando as medianas vetoriais em cada nuvem como pontos de referência, é possível verificar padrões de inclinação e posicionamento próximos aos do <i>ground truth</i> (c). Com o par de eixos tomados agora a partir das nuvens de DPI a resposta é similar (d). . . . .	15
3.3	Maior estabilidade e confiabilidade dos eixos de simetria tomados a partir da região dos olhos. Em (a), os eixos de simetria tomados a partir das nuvens da região dos olhos da imagem de face em prova. Em (b), os eixos de simetria na região do nariz a partir das nuvens $\mathcal{N}(l)$ , $l = 6, \dots, 8$ . Em (c), os eixos para a região da boca, tomados a partir da nuvem de seus PFFs $\mathcal{L}(9)$ e $\mathcal{L}(10)$ . Por conter mais elementos, a região dos olhos permite a melhor observação para eixos de simetria nas várias poses das imagens de faces. É possível verificar que os eixos para as demais regiões se tornam mais verossímeis se ajustados pela reta de simetria bilateral tomada na região dos olhos, conforme (d). . . . .	17
3.4	Conversão de base dos elementos das nuvens $\mathcal{N}(l)$ . Em (a), para agilizar o processo de busca de pares de PFFs simétricos, uma mudança de base é executada previamente. Em (b), o processo de conversão se dá em 3 passos: 1) reflexão vertical dos eixos originais; 2) rotação proporcional a $\beta_x^{(o)}$ ; e 3) translação da origem $(0, 0)$ aos pontos médios dos PFFs da região onde se encontra o par de eixos de simetria. . . . .	19
3.5	Visão simplificada do produto do treinamento do método DPI-SIM. Em a) é localizado um ponto fiducial $\mathcal{L}(l)$ ; b) módulo $r_l$ ; c) ângulo $\varphi_l$ ; d) a reflexão em torno do eixo transversal $r_l \angle \varphi_l \rightarrow r_l \angle (\pi - \varphi_l)$ ; e) $r_l + e_l^{(min)} \angle (\pi - \varphi_l + \delta_l^{(max)})$ ; f) $r_l + e_l^{(min)} \angle (\pi - \varphi_l + \delta_l^{(min)})$ ; g) $r_l + e_l^{(max)} \angle (\pi - \varphi_l + \delta_l^{(max)})$ ; h) $r_l + e_l^{(max)} \angle (\pi - \varphi_l + \delta_l^{(min)})$ ; T descreve a área formada a partir dos limiares. . . . .	23
3.6	Diagrama de blocos do sistema de teste para detecção de pontos fiduciais faciais pelo método DPI-SIM. Como saída, os 11 PFFs da marcação de Sabino e Araujo definidos de forma global e os eixos de simetria posicionados sobre as regiões dos olhos, nariz e boca. . . . .	24
3.7	Resultados do DPI-SIM para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(0)$ e $\mathcal{L}(5)$ - cantos externos dos olhos. . . . .	28

3.8	Resultados do DPI-SIM para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(1)$ e $\mathcal{L}(4)$ - pupilas. . . . .	29
3.9	Resultados do DPI-SIM para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(2)$ e $\mathcal{L}(3)$ - cantos internos dos olhos. . . . .	30
3.10	Resultados do DPI-SIM para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para o pontol fiducial $\mathcal{L}(7)$ - centro do nariz. . . . .	31
3.11	Resultados do DPI-SIM para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(6)$ e $\mathcal{L}(8)$ - cantos do nariz. . . . .	32
3.12	Resultados do DPI-SIM para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(9)$ e $\mathcal{L}(10)$ - cantos da boca. . . . .	33
3.13	Resultados do DPI-SIM para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(0)$ e $\mathcal{L}(5)$ - cantos externos dos olhos. . . . .	36
3.14	Resultados do DPI-SIM para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(1)$ e $\mathcal{L}(4)$ - pupilas. . . . .	37
3.15	Resultados do DPI-SIM para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(2)$ e $\mathcal{L}(3)$ - cantos internos dos olhos. . . . .	38
3.16	Resultados do DPI-SIM para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para o pontol fiducial $\mathcal{L}(7)$ - centro do nariz. . . . .	39
3.17	Resultados do DPI-SIM para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(6)$ e $\mathcal{L}(8)$ - cantos do nariz. . . . .	40

3.18	Resultados do DPI-SIM para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(9)$ e $\mathcal{L}(10)$ - cantos da boca. . . . .	41
3.19	Exemplos de faces extraídas das bases de imagens FERET e BioID. Em (a), faces oriundas da base FERET. Em (b), faces da base BioID. . . . .	42
4.1	O <i>bunch graph</i> , estrutura de dados em formato de grafo rotulado. Seus vértices contêm pilhas de jatos extraídos de pontos fiduciais de faces humanas. Suas arestas são rotuladas com as distâncias médias entre esses pontos fiduciais. (Günther [2]). . . . .	47
4.2	Visão simplificada de um <i>agregado de grafos de faces</i> com $M$ elementos. Diferentemente dos <i>face bunch graphs</i> da Figura 4.1, empilhamos somente as informações qualitativas dos grafos em uma estrutura de dados. Além disso, não há mais rótulos nos vértices dos grafos com as informações quantitativas dos <i>landmarks</i> : os jatos. . . . .	52
4.3	Grafo definido a partir das Marcações de Sabino e Araujo. . . . .	53
4.4	Subgrafos induzidos de $\mathcal{G}$ da Figura 4.3. $G'$ registra os PFFs $\mathcal{L}(0)$ , $\mathcal{L}(1)$ e $\mathcal{L}(2)$ ; $G''$ , de $\mathcal{L}(6)$ a $\mathcal{L}(10)$ ; e $G'''$ , $\mathcal{L}(1)$ , $\mathcal{L}(2)$ , $\mathcal{L}(3)$ , $\mathcal{L}(4)$ e $\mathcal{L}(7)$ . . . . .	55
4.5	Diagrama de blocos do sistema de teste para detecção de pontos fiduciais faciais pelo método DPI-GRAPH. Diferentemente do método representado na Figura 2.4, aqui se buscam todos os 11 PFFs definidos de forma global. . . . .	56
4.6	Resultados do DPI-GRAPH para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(0)$ e $\mathcal{L}(5)$ - cantos externos dos olhos. . . . .	60
4.7	Resultados do DPI-GRAPH para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(1)$ e $\mathcal{L}(4)$ - pupilas. . . . .	61
4.8	Resultados do DPI-GRAPH para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(2)$ e $\mathcal{L}(3)$ - cantos internos dos olhos. . . . .	62
4.9	Resultados do DPI-GRAPH para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para o pontol fiducial $\mathcal{L}(7)$ - centro do nariz. . . . .	63

4.10	Resultados do DPI-GRAPH para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(6)$ e $\mathcal{L}(8)$ - cantos do nariz. . . . .	64
4.11	Resultados do DPI-GRAPH para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(9)$ e $\mathcal{L}(10)$ - cantos da boca. . . . .	65
4.12	Resultados do DPI-GRAPH para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(0)$ e $\mathcal{L}(5)$ - cantos externos dos olhos. . . . .	68
4.13	Resultados do DPI-GRAPH para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(1)$ e $\mathcal{L}(4)$ - pupilas. . . . .	69
4.14	Resultados do DPI-GRAPH para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(2)$ e $\mathcal{L}(3)$ - cantos internos dos olhos. . . . .	70
4.15	Resultados do DPI-GRAPH para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para o pontol fiducial $\mathcal{L}(7)$ - centro do nariz. . . . .	71
4.16	Resultados do DPI-GRAPH para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(6)$ e $\mathcal{L}(8)$ - cantos do nariz. . . . .	72
4.17	Resultados do DPI-GRAPH para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos $\mathcal{L}(9)$ e $\mathcal{L}(10)$ - cantos da boca. . . . .	73
4.18	Comparativos entre os resultados dos métodos DPI-GRAPH e DPI-SIM para as bases BioID e FERET para os pontos fiduciais $\mathcal{L}(0)$ e $\mathcal{L}(5)$ - cantos externos dos olhos. . . . .	74
4.19	Comparativos entre os resultados dos métodos DPI-GRAPH e DPI-SIM para as bases BioID e FERET para os pontos fiduciais $\mathcal{L}(1)$ e $\mathcal{L}(4)$ - pupilas. . . . .	75

4.20	Comparativos entre os resultados dos métodos DPI-GRAPH e DPI-SIM para as bases BioID e FERET para os pontos fiduciais $\mathcal{L}(2)$ e $\mathcal{L}(3)$ - cantos internos dos olhos. . . . .	76
4.21	Comparativos entre os resultados dos métodos DPI-GRAPH e DPI-SIM para as bases BioID e FERET para os pontos fiduciais $\mathcal{L}(6)$ e $\mathcal{L}(8)$ - cantos do nariz. . . . .	77
4.22	Comparativos entre os resultados dos métodos DPI-GRAPH e DPI-SIM para as bases BioID e FERET para o ponto fiducial $\mathcal{L}(7)$ - centro do nariz. . . . .	77
4.23	Comparativos entre os resultados dos métodos DPI-GRAPH e DPI-SIM para as bases BioID e FERET para os pontos fiduciais $\mathcal{L}(9)$ e $\mathcal{L}(10)$ - cantos da boca. . . . .	78

# Lista de Tabelas

2.1	Resumo de algumas técnicas de detecção automática de pontos fiduciais faciais. . . . .	4
-----	--	---

# Lista de Algoritmos

1	Treinamento do Método DPI-SIM. . . . .	22
2	Aplicação do Método DPI-SIM. . . . .	25
3	Treinamento do Método DPI-GRAPH. . . . .	54
4	Aplicação do Método DPI-GRAPH. . . . .	57

# Lista de Abreviaturas

AAM	<i>Active Appearance Model</i> , p. 5
ASM	<i>Active Shape Model</i> , p. 5
DLA	<i>Dynamic Link Architecture</i> , p. 45
DPI	Detector por Produto Interno, p. 2
EBGM	<i>Elastic Bunch Graph Matching</i> , p. 2
FBG	<i>Face Bunch Graph</i> , p. 47
ML	<i>Maximum Likelihood</i> , p. 48
PCA	<i>Principal Component Analysis</i> , p. 3
PPF	Ponto Fiducial Facial, p. 1
REP	Redução do Espaço de Busca, p. 10
ROI	<i>Region of Interest</i> , p. 10
RP	Reconhecimento de Padrões, p. 1
SVM	<i>Support Vector Machines</i> , p. 3

# Capítulo 1

## Introdução

Até hoje, um grande volume de pesquisa foi concebido focando na detecção ou estimação automática de pontos fiduciais em faces humanas, muito devido ao fato de serem de grande relevância em diferentes aplicações do campo de visão computacional acerca de faces como: detecção e reconhecimento [3], reconhecimento de expressões e emoções [4], estimação de pose [5] e rastreamento [6].

Por definição, um ponto fiducial facial (PFF) - em inglês, *facial landmark*, ou simplesmente *landmark* - é aquele localizado em uma posição específica da face humana que garanta sua existência na maioria das observações. Esses pontos normalmente marcam características salientes como: cantos e centros dos olhos, bordas da boca, centro do nariz e sobrancelhas, não se limitando a estes [7].



Figura 1.1: Proporções da Face e Olhos, 1489, desenho à mão de Leonardo da Vinci (<http://drawingsofleonardo.org, set/2010>).

Algumas tentativas de uma formulação matemática para o formato facial humano, o que está diretamente ligado ao intuito de localizar os PFFs, datam da Renascença, observáveis em estudos como os de Leonardo da Vinci (Fig. 1.1) para a beleza e envelhecimento. Há também registros bem mais remotos como na Grécia e Egito antigos. Regras de proporções rígidas como a *Razão Áurea* foram consideradas, embora essas medidas e resultados possam pouco valor além de significado histórico [8]. Foi de fato a partir dos anos 60s do século passado que se iniciou a

construção do entendimento da face que nós utilizamos ainda hoje em diversas disciplinas do saber: na psicologia experimental, como no trabalho de Ekman acerca das expressões [9]; na antropometria e anatomia com Farkas [10], na biométrica com Lu [11]; e na engenharia e computação, com Goldstein *et al* [12] destacando pela primeira vez a importância dos pontos fiduciais e características inerentes a estes para o reconhecimento de faces.

Dentre as numerosas técnicas de detecção de pontos fiduciais, destacamos duas: o casamento de grafos - *elastic bunch graph matching* (EBGM) e os detectores por produto interno (DPI). O EBGM é o método de detecção e identificação/verificação de faces original de Wiskott [3]. Neste, os *landmarks* são encontrados na fase de detecção da face na imagem em prova a partir do compromisso entre uma filtragem espacial por *wavelets* de motivação biológica [13] e o casamento de grafos elásticos com um conjunto de grafos a priori. O DPI é um método de reconhecimento de padrões (RP) baseado em filtragem de correlação. Esta técnica foi recentemente desenvolvida e apresentada em [14] e [1], e aplicada na detecção de PFFs. É nesses dois métodos que este trabalho se sustenta.

As saídas dos classificadores locais DPI apontam coordenadas no plano da imagem de uma face em prova referenciando candidatos a PFFs - os quais chamaremos a partir de agora de *nuvens do DPI* ou simplesmente *nuvens*. Os reverses desse método residem em algumas características dessas nuvens: sua ordem e espalhamento. Até agora, o DPI era utilizado para a busca isolada de PFFs em imagens de face. Esta dissertação propõe a busca global dos PFFs fazendo uso de suas posições relativas. Para a eleição dos candidatos mais prováveis, elencamos duas propostas. Primeiramente, exploramos a alternativa antropométrica para detectar os pontos fiduciais nas imagens em prova sobre regras de simetria facial. Aqui, a partir dos padrões de distribuição das nuvens, são traçados eixos de simetria longitudinais e transversais para realizar a busca dos melhores candidatos a PFFs entregues pelo DPI. A segunda proposta se dá através de um estimador de estrutura de grafos, concebido por Würtz [15] sobre o casamento de grafos de Wiskott, aplicando-o às saídas dos filtros de correlação. Nessa proposta, cada ponto fiducial passa a ser um vértice de um grafo que terá seus parâmetros estimados por uma função de máxima verossimilhança.

Para a avaliação dos métodos propostos, 11 pontos fiduciais faciais foram escolhidos - cantos e centro dos olhos e nariz, e laterais da boca - a partir das marcações de Sabino e Araujo [16] nas bases de imagens para testes de algoritmos de reconhecimentos de faces BioID [17] (Fig. 1.2(a)) e FERET [18] (Fig. 1.2(b)).

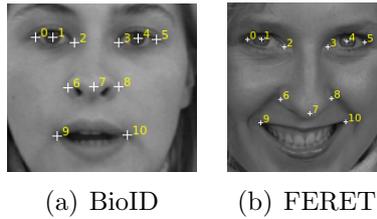


Figura 1.2: Marcações de Pontos Fiduciais Faciais por Sabino e Araujo na base BioID (a) e FERET (b)

## 1.1 Organização da Dissertação

O Capítulo 2 é composto pelo resumo de algumas técnicas para detecção de pontos fiduciais. Além disso, o método necessário para o entendimento dos capítulos subsequentes é detalhado: o DPI, desde sua forma primitiva até a mais depurada, alavancada pelo casamento com o método de Análise de Componentes Principais (PCA), o DPI-PCA. O Capítulo 3 descreve a primeira contribuição deste trabalho, definindo o método de detecção de pontos fiduciais faciais utilizando regras de simetria facial. Também são explicados os procedimentos experimentais, os bancos de imagens de faces utilizados, metodologia de treinamento e teste, e, finalmente, a análise dos resultados. Esta é feita em comparação com as taxas obtidas anteriormente pelo DPI-PCA [1] e classificadores baseados em Máquinas de Vetor de Suporte - do inglês, *Support Vector Machines* (SVM), resultados extraídos também de [1]. No Capítulo 4, a segunda proposta deste trabalho para detecção dos PFFs a partir das nuvens do DPI-PCA utilizando casamento de grafos é apresentada. Seus resultados são comparados com os do Capítulo 3. Além disso, o método em que se baseia essa segunda proposta é descrito: o EBGM, tanto em seu formato original quanto sua variante estatística.

E, finalmente, no Capítulo 5, estão as conclusões e trabalhos futuros.

# Capítulo 2

## Detecção de Pontos Fiduciais

Neste capítulo, listamos e resumimos algumas técnicas de sucesso em detecção automática de pontos fiduciais. Além disso, descrevemos o método DPI-PCA e o sistema de detecção local de pontos fiduciais faciais desenvolvidos em [1]. Eles serão usados mais adiante como blocos de construção para o desenvolvimento e validação dos métodos contidos neste trabalho.

### 2.1 Técnicas

Há duas vertentes principais que segmentam os métodos de detecção de pontos fiduciais: as técnicas de detecção local, e as de detecção global de PFFs. Alguns métodos de detecção automática de pontos fiduciais estão listados na Tabela 2.1.

As técnicas de detecção local de pontos fiduciais buscam os *landmarks* de forma individual, sem fazer uso de informação de outras regiões. Isto é, há detectores para a boca (cantos, lábios superior e inferior), olhos (cantos, pupila), narizes (cantos, ponta, base), etc. Entre outros métodos desta categoria [19], destacamos o DPI-PCA, sustentáculo deste trabalho.

Já as técnicas de detecção global usam, em sua maioria, modelos ou máscaras deformáveis [20, 21, 22] para representar de forma conjunta os PFFs, o que de fato é o necessário no uso prático em aplicações de visão computacional.

Tabela 2.1: Resumo de algumas técnicas de detecção automática de pontos fiduciais faciais.

Método	Descrição
<i>Active Shape Model</i> (ASM) [21]	Método de detecção de formas baseado em modelos deformáveis. Os detectores de PFFs desta linha buscam <i>landmarks</i> globalmente.

*Continua na próxima página*

Método	Descrição
<i>Active Appearance Model</i> (AAM) [22]	Realiza o trabalho do ASM e ainda agrega representação e casamento de textura na detecção dos padrões de interesse. Categoria global.
ASM baseado em SVM [20]	Combina SVM com ASM para a detecção de pontos fiduciais. Categoria global.
Detecção de PFFs por classificadores baseados em filtros de Gabor [19]	Utiliza cascatas de classificadores rápidos <i>AdaBoost</i> [23] para selecionar características extraídas por <i>wavelets</i> de Gabor. Categoria local.

É importante destacar que os métodos globais tendem a ser menos suscetíveis à variação de pose e iluminação do que os locais.

## 2.2 O Detector por Produto Interno (DPI)

O DPI é um método de classificação de padrões pertencente à categoria de detectores baseados em filtros de correlação. Esses filtros denominam a classe de métodos de RP que possui como principal característica o uso da correlação como métrica de similaridade no projeto de seus classificadores [24]. Os DPIs são ótimos sob o ponto de vista de minimização do erro médio quadrático de classificação. São considerados detectores rápidos por realizarem a classificação através de um simples produto interno.

O treinamento de um classificador DPI visa à construção de um vetor  $\mathbf{h}$  capaz de discriminar um padrão  $\mathbf{x}$  com a execução de um produto interno. Nesta seção, resumiremos a teoria por trás do DPI e descrevemos os passos para a aplicação desses filtros em blocos de imagens para detecção local de PFFs.

O conteúdo desta seção é um resumo do desenvolvimento da teoria e aplicação dos classificadores DPIs contidos na tese doutorado de Sabino, em [1].

### 2.2.1 DPI para uma única classe

Seja uma variável aleatória  $\mathcal{X}$  cujas realizações  $\mathbf{x}_j$  pertençam às classes  $\mathcal{A}_i$ ,  $i = 1, \dots, N$ , ou ainda à classe  $\mathcal{B}$ , que reúne demais classes de padrões que porventura possam existir. A probabilidade de  $\mathcal{X}$  pertencer à classe  $\mathcal{B}$  é dada por

$$p(\mathcal{B}) = p(\mathcal{X} \in \mathcal{B}),$$

e às classes  $\mathcal{A}_i$ :

$$p(\mathcal{A}_i) = p(\mathcal{X} \in \mathcal{A}_i) = p_i.$$

E  $\mathbf{h}_{\mathcal{A}_i}$  será de tal forma que, se  $\mathcal{X} \in \mathcal{A}_i$

$$\langle \mathbf{h}_{\mathcal{A}_i}, \mathcal{X} \rangle = 1,$$

caso contrário

$$\langle \mathbf{h}_{\mathcal{A}_i}, \mathcal{X} \rangle = 0.$$

De modo geral:

$$\langle \mathbf{h}_{\mathcal{A}_i}, \mathcal{X} \rangle = \mathbf{c},$$

com  $\mathbf{c}$  resultando após o produto interno:  $\mathbf{c} = 1 | \mathcal{X} \in \mathcal{A}_i$  ou  $\mathbf{c} = 0 | \mathcal{X} \notin \mathcal{A}_i$ .

Na busca de minimizar o erro médio quadrático de  $\mathbf{c}$ , escreve-se

$$e = \mathbf{h}_{\mathcal{A}_i}^T \mathcal{X} - \mathbf{c},$$

e fazendo

$$\frac{\partial E [||e||^2]}{\partial \mathbf{h}_{\mathcal{A}_i}} = 0$$

temos

$$\mathbf{h}_{\mathcal{A}_i} = \{E [\mathcal{X}\mathcal{X}^T]\}^{-1} E [\mathcal{X}\mathbf{c}]. \quad (2.1)$$

Conforme demonstrado em [1], tem-se:

$$\mathbf{h}_{\mathcal{A}_i} = \{p(\mathcal{B})\mathcal{R}_{\mathcal{B}} + \sum_{n=1}^N p_n \mathcal{R}_{\mathcal{A}_n}\}^{-1} p_i \boldsymbol{\mu}_{\mathcal{A}_i} \quad (2.2)$$

onde  $\mathcal{R}_{\mathcal{B}}$  é a matriz de autocovariância dos elementos da classe  $\mathcal{B}$ ;  $\mathcal{R}_{\mathcal{A}_n}$  é a matriz de autocovariância dos elementos das  $N$  classes  $\mathcal{A}_n$ ;  $\boldsymbol{\mu}_{\mathcal{A}_i}$ , a média dos elementos da classe  $\mathcal{A}_i$ ;  $p(\mathcal{B})$ ,  $p_n$  e  $p_i$ , as probabilidades de  $\mathcal{X}$  pertencer às classes  $\mathcal{B}$ ,  $\mathcal{A}_n$  e  $\mathcal{A}_i$  respectivamente. À essa equação, aplicam-se as seguintes restrições:

- A matriz  $E [\mathcal{X}\mathcal{X}^T]$  deve ser inversível;
- A dimensão dos vetores de entrada  $\mathbf{x}$  deve ser menor que a soma dos elementos das classes  $\mathcal{A}_i$  e  $\mathcal{B}$ .

## 2.2.2 DPI generalizado

A Equação 2.2 assume que as realizações da variável aleatória  $\mathcal{X}$  são reais e que somente os padrões pertencentes à cada uma das classes  $\mathcal{A}_i$  serão detectados de forma individual.

O DPI pode ser generalizado para permitir que múltiplas classes sejam detectadas a partir de um único filtro  $\mathbf{h}$  e ainda assumir como entrada variáveis aleatórias complexas. Para o primeiro caso, seja  $\mathcal{A}$  um subconjunto de classes com  $M$  das  $N$  classes de  $\mathcal{A}_i$  com  $M < N$ . Substituímos o produto  $p_i \boldsymbol{\mu}_{\mathcal{A}_i}$  por  $\sum_{m=1}^M p_m \boldsymbol{\mu}_{\mathcal{A}_m}$  na Equação 2.2, obtendo:

$$\mathbf{h}_{\mathcal{A}} = \{p(\mathcal{B})\mathcal{R}_{\mathcal{B}} + \sum_{n=1}^N p_n \mathcal{R}_{\mathcal{A}_n}\}^{-1} \sum_{m=1}^M p_m \boldsymbol{\mu}_{\mathcal{A}_m}. \quad (2.3)$$

Para o segundo caso, a Equação 2.3 não se modifica. A demonstração dessa solução também é encontrada em [1], notando que o primeiro termo da Equação 2.1 se torna  $E[\mathcal{X}\mathcal{X}^{*t}]$ .

### 2.2.3 DPI com Análise de Componentes Principais (PCA)

Com o intuito de aumentar a robustez dos detectores por produto interno, Sabino [1] propôs a construção desses classificadores para as chamadas *componentes principais* oriundas da transformada de Karhunen-Loève, também conhecida como Análise de Componentes Principais - *Principal Component Analysis* (PCA).

#### Transformada de Karhunen-Loève

A transformada de Karhunen-Loève, ou PCA, é um dos métodos mais populares para extração de características (dados de entrada para um sistema de classificação) e redução de dimensionalidade em reconhecimento de padrões [25].

Através da PCA, é possível representar de forma eficiente um vetor de características, reduzindo sua dimensionalidade ao projetá-los sobre *eixos principais*. Eles são conjuntos ortonormais de eixos que apontam na direção de máxima covariância dos dados. Assim, as características (entradas do classificador) tendem a ser mutuamente decorrelacionadas, no intuito de evitar ou minimizar redundância de informação. Além disso, minimiza o erro médio quadrático de projeção para um dado número de dimensões (eixos) [26].

A seguir, de forma sintética, descrevemos matematicamente a PCA por exemplos práticos com conceitos usados neste trabalho.

#### *Eigenfaces e Eigenpoints*

Suponha uma matriz  $\mathbf{X} = [\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_M]$ , onde cada coluna  $\mathbf{x}_i$  contém os pixels de uma imagem de face. Assuma, para efeitos de simplicidade, que esses dados tenham média zero (caso contrário, subtraia dos dados a média do seu conjunto). Para buscar um conjunto de  $M - 1$  vetores ortogonais  $\boldsymbol{\Phi}_i$  que melhor descreve a distribuição das

características no sentido dos mínimos quadrados, ache os autovetores da matriz de covariância

$$\Sigma_{\mathbf{x}} = \mathbf{X}\mathbf{X}^{*T}.$$

Essa solução é direta e dada pela Equação 2.4:

$$\Lambda = \Phi \Sigma_{\mathbf{x}} \Phi^{*T} \quad (2.4)$$

onde  $\Lambda$  é a matriz diagonal contendo os autovalores de  $\Sigma_{\mathbf{x}}$ .

De posse dos autovetores  $\Phi_i$ , é possível fazer uma seleção das “melhores” componentes através dos autovalores  $\lambda_i$ . Quanto maior for  $\lambda_i$ , mais da variância dos dados foi capturada pelo autovetor  $\Phi_i$ .

Desta forma, fica descrito o método de PCA, e os *eigenfaces*, que nada mais são do que as  $k$  componentes principais  $\Phi_i$ .

Migrar de *eigenfaces* para *eigenpoints* (Figuras 2.1(a) e 2.1(b)) é um procedimento direto, bastando modificar as colunas da matriz  $\mathbf{X}$  para conterem pixels de blocos centrados em pontos fiduciais ao invés de faces inteiras.



Figura 2.1: Exemplos de *eigenfaces* e *eigenpoints*. Em 2.1(b), componentes principais a partir de blocos centrados na pupila direita de imagens de faces.

## Formulação do DPI-PCA

De posse das ferramentas descritas a partir da Seção 2.2.1, é possível formular o DPI-PCA como segue.

Sejam  $\mathbf{U} = [\mathbf{u}_1 \mathbf{u}_2 \dots \mathbf{u}_M]$ ,  $\Sigma_{\mathbf{u}}$  e  $\Phi = [\Phi_1 \Phi_2 \dots \Phi_N]$  a matriz contendo os pixels normalizados de blocos de imagens de faces, a matriz de covariância desses dados e os *eigenpoints* respectivamente. Para projetar um vetor  $\mathbf{h}_{\Phi_i}$ , um DPI para discriminar as componentes  $\Phi_i$  e rejeitar as demais, são necessárias as seguintes suposições:

- A classe  $\mathcal{A}_1$  será composta pelo conjunto de componentes  $\Phi_i$  que serão discriminadas pelo DPI  $\mathbf{h}_{\Phi_i}$ ;
- A classe  $\mathcal{A}_2$  será composta por deslocamentos lineares de  $\Phi_i$  de  $n$  pixels horizontalmente e  $m$  pixels verticalmente na imagem de face em prova;
- A classe  $\mathcal{B}$  será composta pelas demais componentes  $\{\Phi_j\}_{j \neq i}$ .

Escrevendo a forma fechada para o DPI  $\mathbf{h}_{\Phi_i}$  pela Equação 2.2 teríamos:

$$\mathbf{h}_{\Phi_i} = \{p_1 \mathcal{R}_{\mathcal{A}_1} + p_2 \mathcal{R}_{\mathcal{A}_2} + p(\mathcal{B}) \mathcal{R}_{\mathcal{B}}\}^{-1} p_1 \boldsymbol{\mu}_{\mathcal{A}_i}. \quad (2.5)$$

De forma mais geral, essa Equação 2.5 pode ainda ser escrita em função das componentes, como:

$$\mathbf{h}_{\Phi_i} = \left\{ p_1 \Phi_i \Phi_i^{*T} + p_2 \frac{1}{(2L+1)^2 - 1} \sum_{\substack{n=-L \\ n \neq 0}}^L \sum_{\substack{m=-L \\ m \neq 0}}^L \Phi_i^{(n,m)} \Phi_i^{(n,m)*T} \right. \\ \left. + p(\mathcal{B}) \frac{1}{N-1} \sum_{\substack{j=1 \\ j \neq i}}^N \Phi_j \Phi_j^{*T} \right\}^{-1} p_1 \Phi_i, \quad (2.6)$$

onde  $L$  é o deslocamento linear máximo em  $m$  e  $n$ , com  $-L \leq m, n \leq L$  e  $m, n \neq 0$ . Os superescritos  $^{(n,m)}$  são usados como notação para representar esses deslocamentos.

## 2.2.4 Sistema de Detecção de Pontos Fiduciais Faciais via DPI com Análise de Componentes Principais

Para validar o método DPI, Sabino [1] e Araujo [14] propuseram um sistema de detecção de PFFs para imagens de faces em poses frontais e/ou pequenas rotações de perfil e laterais.

Para tanto, 11 pontos fiduciais faciais foram escolhidos e marcados manualmente sobre amostras de imagens de face provenientes das bases BioID [17] e FERET [18] (vide Figuras 1.2(a) e 1.2(b) no Capítulo 1). Cada uma dessas bases foi particionada em 7 *folds* para a validação cruzada do tipo *k-fold*. Nesse procedimento de avaliação,  $k - 1$  partições são usadas para treinamento, e uma para testes.

Nesta seção, descrevemos em resumo o sistema utilizado também como bloco de construção para validar os métodos desenvolvidos nesta dissertação. Maiores detalhes desse sistema podem ser encontrados na tese de doutorado de Sabino [1].

### Pré-processamento

O pré-processamento das imagens é um procedimento comum às operações de treino e teste do sistema. Nessa fase, dada uma imagem de prova que ao ser disponibilizada como entrada:

- Um detector de faces via algoritmo de Viola-Jones [27] realiza o enquadramento da face no plano da imagem.

- A imagem centralizada e recortada no passo anterior é reescalada para uma dimensão pré-definida: 220x220 pixels.
- É executada uma correção de iluminação na imagem de face através de correção do fator gama, filtragem por diferença de gaussianas e equalização de contraste [28].

## Redução do Espaço de Busca

A redução do espaço de busca (REP) define a região de interesse - *Region Of Interest* (ROI) - para o treinamento e aplicação dos detectores por produto interno. Essa região é definida segundo um modelo probabilístico gaussiano, através de uma seção transversal em uma função densidade de probabilidade  $N(\boldsymbol{\mu}_{\mathcal{L}(l)}, \boldsymbol{\Sigma}_{\mathcal{L}(l)})$ , onde  $\mathcal{L}(l)$  é a coordenada de um ponto fiducial  $l$  no plano de uma imagem de face.

A definição das ROIs para cada um dos onze PFFs definidos é feita da seguinte forma.

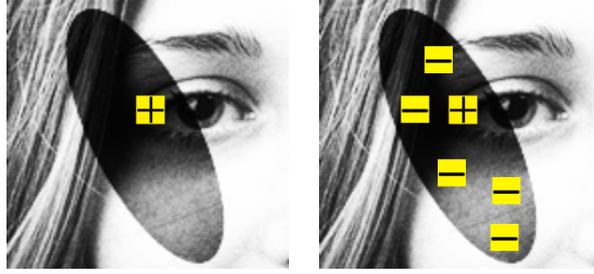
- Dado um conjunto de imagens de faces:
  1. Extraia as médias  $\boldsymbol{\mu}_{\mathcal{L}(l)}$  e matrizes de covariância  $\boldsymbol{\Sigma}_{\mathcal{L}(l)}$  dos conjuntos de PFFs  $\{\mathcal{L}(l)^{(p)}\}$ ,  $l = 0, \dots, 10$  pontos e  $p = 1, \dots, P$  imagens, manualmente marcados nas imagens de faces
  2. A distância entre o elemento  $\mathcal{L}(l)^{(p)}$  mais afastado da média  $\boldsymbol{\mu}_{\mathcal{L}(l)}$  acrescida de 5% define a ROI: uma região elíptica em torno da média da distribuição gaussiana de um ponto fiducial.
  3. A Equação 2.7 constrói a ROI para um ponto  $\mathcal{L}(l)$  dentro do plano da imagem:

$$1,05 \cdot \max_{\mathcal{L}(l)^{(p)}} \left\{ \|\mathcal{L}(l)^{(p)} - \boldsymbol{\mu}_{\mathcal{L}(l)}\|^2 \right\} \geq \left( \mathcal{L}(l)' - \boldsymbol{\mu}_{\mathcal{L}(l)} \right)^T \boldsymbol{\Sigma}_{\mathcal{L}(l)}^{-1} \left( \mathcal{L}(l)' - \boldsymbol{\mu}_{\mathcal{L}(l)} \right) \quad (2.7)$$

onde  $\mathcal{L}(l)'$  é um candidato a ponto fiducial válido por estar contido no interior da ROI.

## Classes

Para o treinamento dos DPIs para os *eigenpoints*, as classes de positivos e negativos são definidas conforme Seção 2.2.3. São consideradas três classes por DPI: uma classe de positivos,  $\mathcal{A}_1$ , formada pela componente a ser detectada a partir do bloco da imagem centrada sobre a marcação de um PFF  $\mathcal{L}(l)$ ; classes de negativos  $\mathcal{A}_2$  e  $\mathcal{B}$ , com  $\mathcal{A}_2$  sendo formada pelas componentes dos blocos na vizinhança de  $\mathcal{A}_1$ , e  $\mathcal{B}$ , pelas demais componentes. As Figuras 2.2(a) e 2.2(b) mostram de forma simplificada o processo de seleção das classes  $\mathcal{A}_1$  e  $\mathcal{A}_2$ .



(a) Bloco centrado sobre (b) Seleção de blocos da um PFF do canto do olho classe de negativos para esse *landmark*.

Figura 2.2: Visão simplificada da seleção de blocos para classes de positivos e negativos para treinamento do sistema de detecção local de pontos fiduciais DPI-PCA. Em 2.2(a), o bloco centrado no PFF terá suas componentes principais extraídas para treinamento de um DPI  $\mathbf{h}_{\Phi}$ . Em 2.2(b), os blocos para a classe de negativos selecionados no interior da ROI desse ponto. Uma seleção pseudo-aleatória uniformemente distribuída de cinquenta desses blocos foi proposta em [1].

## Treinamento

São dois os produtos do treinamento do sistema de detecção local de PFFs por DPI-PCA proposto em [1]: os vetores  $\mathbf{h}_{\Phi_i}$  e classificadores AdaBoost [23]. Estes últimos definem a superfície de decisão para as classes de positivos e negativos.

Para a obtenção desses classificadores, todos os passos descritos anteriormente nesta seção são executados previamente. Sejam eles:

- Pré-processamento: onde as faces nas imagens são enquadradas, recortadas, reescaladas e corrigidas em iluminação;
- Redução do Espaço de Busca: a ROI é definida na imagem a partir do modelo de distribuição gaussiana a priori para um PFF  $\mathcal{L}(l)$ ;
- Definição das classes de *Positivo* e *Negativos*.

Encerrados esses procedimentos iniciais, é iniciada a construção do classificador local  $\mathbf{h}_{\Phi_i}$  para a  $i$ -ésima componente principal,  $i = 1, \dots, N$ , do bloco centrado em um *landmark*  $\mathcal{L}(l)$  utilizando a Equação 2.6. De posse dos vetores  $\mathbf{h}_{\Phi_i}$ , é iniciado o treinamento do classificador AdaBoost que define a superfície separadora das classes de positivos e negativos. Isso é feito através do produto interno entre os DPIs  $\mathbf{h}_{\Phi_i}$  e os blocos  $B_{\mathbf{z}}$  de positivo e negativos, com  $\mathbf{z}$  sendo a coordenada referente ao centro do bloco.

O diagrama de blocos do procedimento de treinamento de um classificador local DPI-PCA pode ser visualizado na Figura 2.3. Para o treinamento do sistema desenvolvido em [1], o número de componentes principais na maioria dos pontos fiduciais

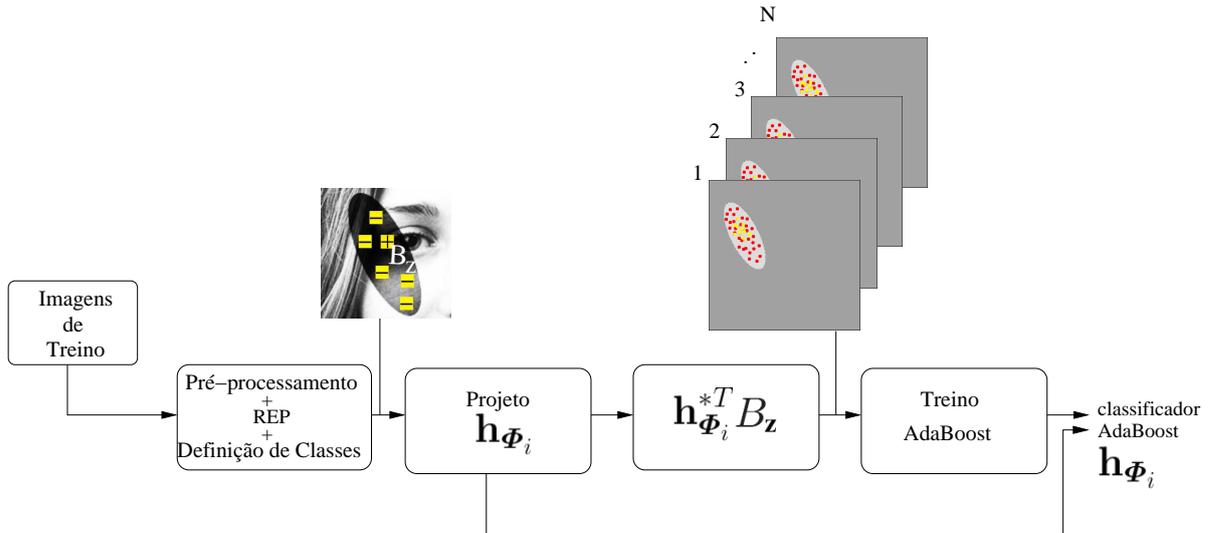


Figura 2.3: Treinamento dos classificadores locais DPI  $\mathbf{h}_{\phi_i}$  para um dos  $i = 1, \dots, N$  *eigenpoints* de um ponto fiducial facial. O classificador AdaBoost também é uma saída desse procedimento, treinado a partir dos resultados do produto interno entre o DPI  $\mathbf{h}_{\phi_i}$  e os blocos  $B_{\mathbf{z}}$  centrados na coordenada  $\mathbf{z}$ .

definidos na marcação de Sabino/Araujo foi de  $N = 63$ , com  $N = 53$  nos demais pontos.

### Teste

O diagrama de blocos que resume a fase de teste para os sistemas de detecção de PFFs por DPI-PCA pode ser visualizado na Figura 2.4.

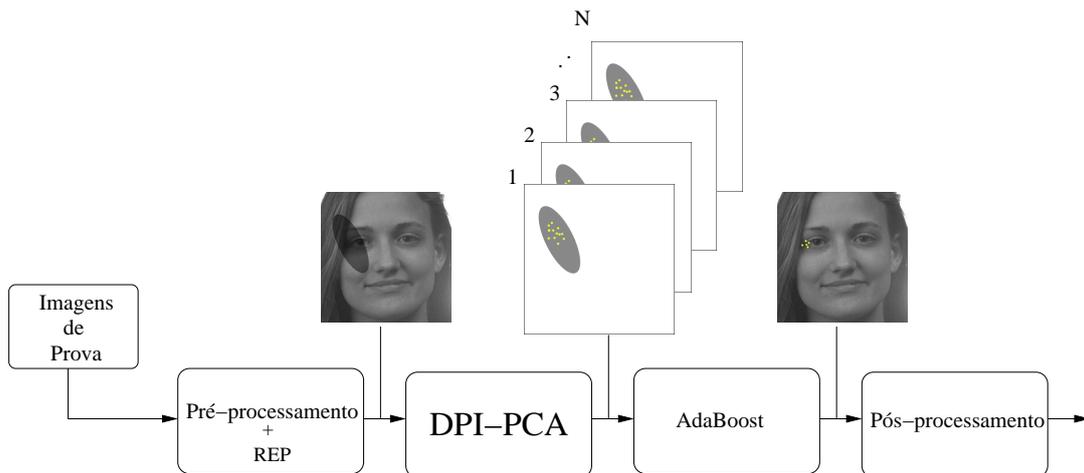


Figura 2.4: Diagrama de blocos do sistema de teste para detecção de um ponto fiducial facial pelo método DPI-PCA.

Nesta fase, uma imagem de prova é disponibilizada na entrada do sistema. Após

sofrer o pré-processamento, é feita a definição da ROI para aplicação do DPI  $\mathbf{h}_{\phi_i}$ . É possível perceber que o projeto desse sistema prevê a detecção de apenas um ponto fiducial por face. Para realizar a detecção local de todos os PFFs da marcação de Sabino/Araujo, 11 desses sistemas foram desenvolvidos.

Pela figura, nota-se que a saída anterior ao pós-processamento é uma nuvem de pontos, coordenadas no plano da imagem de prova reescalada. Para que dessa nuvem se escolha somente uma coordenada que será o candidato eleito a PFF, Sabino propôs algumas estratégias em [1], explicitadas a seguir.

### *Estratégias de Pós-processamento*

- Primeiramente, a estratégia de acrônimo *NA* em [1] não realiza nenhuma modificação na saída do bloco AdaBoost. Dessa forma, todos os elementos da nuvem são considerados saídas do método.
- A estratégia *ML* seleciona a coordenada mais provável, supondo o modelo gaussiano a priori como critério de decisão.
- A estratégia *GML* agrupa os elementos com distância menor que  $P$  pixels e seleciona os resultados dos agrupamentos conforme *ML*.
- A estratégia *A* dá como saída do método a média dos elementos da nuvem.

Para efeitos deste trabalho, a estratégia de pós-processamento utilizada foi a *NA*. Ou seja, todos os rótulos automáticos serão considerados como saída do sistema de detecção de PFFs por DPI-PCA para serem então atacados pelos métodos propostos nos capítulos subseqüentes.

## **2.3 Considerações Finais**

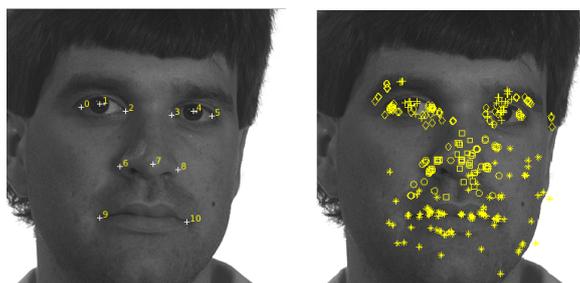
Neste capítulo, resumimos algumas técnicas de detecção de pontos fiduciais faciais, e descrevemos, de forma sucinta, o método de sumo interesse para este trabalho: o DPI-PCA, de Sabino [1]. O sistema de detecção local de PFFs descrito na Seção 2.2.4 se transforma em bloco de construção para validar os métodos desenvolvidos a seguir nos Capítulos 3 e 4.

No capítulo seguinte, será possível visualizar o produto dos classificadores DPI-PCA: coordenadas no plano das imagens de prova indicando a provável existência de um ponto fiducial. Além disso, iniciaremos a discussão sobre a busca de correspondências espaciais entre essas coordenadas para a localização global das características contidas nas marcações de Sabino/Araujo.

## Capítulo 3

# Explorando as Simetrias da Face

Conforme descrito na Seção 2.2.4, os classificadores DPI apontam através de suas saídas as coordenadas de uma imagem em prova onde há alta probabilidade de ocorrência de pontos fiduciais conforme a marcação de Sabino/Araujo. Algumas vezes, até dezenas de prováveis ocorrências são apontadas para um único ponto fiducial. Além disso, a distribuição desses pontos no plano da imagem pode ser demasiado espalhada, conforme Figura 3.1(b).



(a) Marcações dos PFFs da imagem de uma face. (b) Nuvens entregues pelo DPI em todos os 11 pontos.

Figura 3.1: As Marcações de PFFs de Sabino e Araujo e as Nuvens de DPI. As marcações originais (*ground truth*) da imagem de face em prova em (a). Grandes reverses do DPI: a ordem e o espalhamento das nuvens em (b).

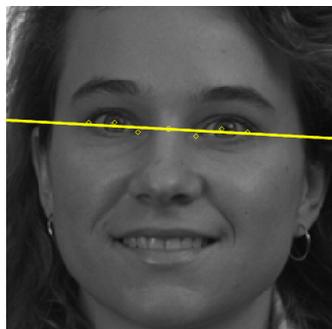
Tanto a partir das marcações de Sabino/Araujo quanto nas nuvens do DPI<sup>1</sup> é possível traçar eixos de simetria nas imagens de faces. Através desses eixos, criamos regras que permitem a eleição dos candidatos a PFFs a partir das nuvens, objetivo deste trabalho. Dessa forma, cada *landmark* eleito traz consigo o seu par simétrico, com exceção do ponto fiducial marcado no centro dos narizes das faces. Este ponto possui regras distintas dos demais, como se verá mais adiante.

As regras definidas neste método são baseadas em observação. Apesar da litera-

---

<sup>1</sup>Novamente, saídas do método DPI-PCA que apontam coordenadas no plano da imagem de prova.

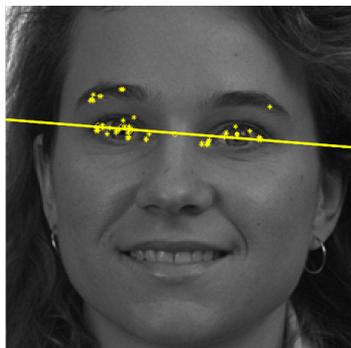
tura apontar meios de se definir eixos de simetria baseados em informação de alto nível como transformadas em [11] e [29], aqui, esses eixos são traçados através do método de mínimos quadrados [30], onde buscamos e definimos retas que melhor se ajustam a pontos (pixels onde se localizam PFFs ou nuvens) no plano da imagem. Isso torna o processo invariavelmente menos custoso computacionalmente, e mais adequado aos dados que nos são disponíveis. Primeiramente, buscamos a reta que se ajusta aos PFFs dos olhos, e, em seguida, traçamos uma segunda reta ortogonal à primeira no ponto médio daquela região. Isso é sempre feito obedecendo às informações das marcações ou nuvens, quando tratamos de treino ou teste respectivamente (vide Figuras 3.2(a), 3.2(b), 3.2(c) e 3.2(d)). Em seguida, aplicamos regras para a eleição dos candidatos a PFF em cada par de *landmarks* simétricos.



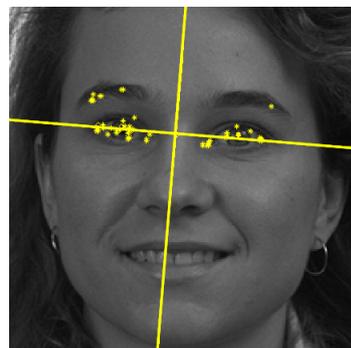
(a) Reta de ajuste aos PFFs da região dos olhos a partir da marcação de Sabino e Araujo.



(b) Eixos de simetria a partir dos PFFs da região dos olhos pela marcação de Sabino e Araujo.



(c) Reta de ajuste às nuvens de DPI da região dos olhos.



(d) Eixos de simetria a partir das nuvens de DPI da região dos olhos.

Figura 3.2: A reta de ajuste sobre as marcações originais (*ground truth*) da imagem de face em prova (a). Utilizando o ponto médio das marcações na região dos olhos, traça-se uma segunda reta formando o par de eixos de simetria para essa região (b). Ao se obter uma reta de ajuste sobre as nuvens de DPI para a região dos olhos, tomando as medianas vetoriais em cada nuvem como pontos de referência, é possível verificar padrões de inclinação e posicionamento próximos aos do *ground truth* (c). Com o par de eixos tomados agora a partir das nuvens de DPI a resposta é similar (d).

Neste capítulo, portanto, iniciaremos o estudo de relações espaciais entre os candidatos a PFFs das imagens de prova oriundos dos classificadores DPI para propor um método de eleição conjunta dessas características faciais: o DPI-SIM.

### 3.1 Método

Seja dada a saída do processamento de uma imagem de face em prova pelos sistemas de detecção de PFFs via DPI-PCA, os quais possuem as seguintes características (Figura 3.1(b)):

- Associada a cada PFF  $\mathcal{L}(l)$ ,  $l = 0, 1, \dots, 10$ , há uma nuvem de *landmarks* candidatos  $\mathcal{N}(l) = \{\nu_1(l), \nu_2(l), \dots, \nu_{n_l}(l)\}$ ,  $n_l = 1, \dots, N_l$ ;
- PFFs do olho direito:  $l = 0, 1, 2$ ;
- PFFs do olho esquerdo:  $l = 3, 4, 5$ ;
- PFFs do nariz:  $l = 6, 7, 8$ ;
- PFFs da boca:  $l = 9, 10$ .
- Pontos fiduciais simétricos entre si:
  - Cantos externos dos olhos:  $\mathcal{L}(0)$  e  $\mathcal{L}(5)$ ;
  - Pupilas:  $\mathcal{L}(1)$  e  $\mathcal{L}(4)$ ;
  - Cantos internos dos olhos:  $\mathcal{L}(2)$  e  $\mathcal{L}(3)$ ;
  - Cantos do nariz:  $\mathcal{L}(6)$  e  $\mathcal{L}(8)$ ;
  - Cantos da boca:  $\mathcal{L}(9)$  e  $\mathcal{L}(10)$ .

Buscamos eixos de simetria nas regiões dos olhos, nariz e boca que permitam a eleição dos PFFs mais simétricos entre si em relação à origem dos eixos. Haja vista que a região dos olhos possui mais pontos fiduciais a serem classificados que as demais, reutilizamos o coeficiente angular de sua reta de ajuste (Figura 3.2(c)) nas demais regiões. Isso porque observamos a tendência de maior estabilidade e confiabilidade na geração da reta para essa região (Figuras 3.3(a), 3.3(b) e 3.3(c)). Sendo assim, as novas retas da região do nariz e boca possuirão a mesma inclinação daquela da região dos olhos, conforme Figura 3.3(d).

Por observação, notamos que os candidatos mais prováveis a PFFs das nuvens  $\mathcal{N}(l)$  tendem a possuir mais pares simétricos. Com isso, realizamos uma busca para cada elemento  $\nu_i(l)$  com a finalidade de eliminar candidatos que não possuam pares dentro de uma pré-determinada tolerância. Tornando o método iterativo, é

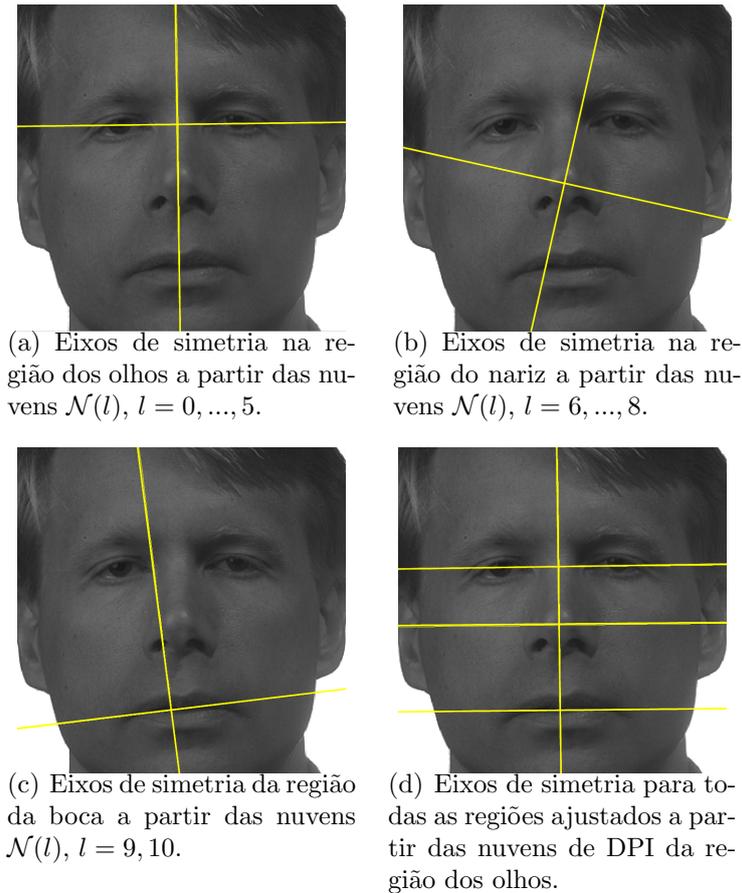


Figura 3.3: Maior estabilidade e confiabilidade dos eixos de simetria tomados a partir da região dos olhos. Em (a), os eixos de simetria tomados a partir das nuvens da região dos olhos da imagem de face em prova. Em (b), os eixos de simetria na região do nariz a partir das nuvens  $\mathcal{N}(l)$ ,  $l = 6, \dots, 8$ . Em (c), os eixos para a região da boca, tomados a partir da nuvem de seus PFFs  $\mathcal{L}(9)$  e  $\mathcal{L}(10)$ . Por conter mais elementos, a região dos olhos permite a melhor observação para eixos de simetria nas várias poses das imagens de faces. É possível verificar que os eixos para as demais regiões se tornam mais verossímeis se ajustados pela reta de simetria bilateral tomada na região dos olhos, conforme (d).

possível ajustar ainda mais os eixos de simetria, conduzindo as medianas vetoriais<sup>2</sup> das nuvens a uma posição estável até que os eixos não mais se modifiquem em posição e inclinação. É nesse momento em que as regras, definidas mais adiante, serão aplicadas a fim de eleger par a par os PFFs. Exceção feita ao ponto fiducial que referencia o centro do nariz,  $\mathcal{L}(7)$ , o qual não possui par simétrico.

<sup>2</sup>Coordenada no plano da imagem que minimiza o somatório das distâncias euclidianas desse ponto a todos os elementos de uma nuvem.

### 3.1.1 Eixos de Simetria

Dados os 6 PFFs  $\mathcal{L}(l)$  da região dos olhos,  $l = 0, \dots, 5$ , nas imagens de faces, buscamos os coeficientes angular  $\beta_x^{(o)}$  e linear  $\gamma_x^{(o)}$  de uma reta  $y = \beta_x^{(o)}x + \gamma_x^{(o)}$ <sup>3</sup> que melhor se ajusta às coordenadas  $(x, y)$  das medianas vetoriais extraídas de cada uma de suas nuvens  $\mathcal{N}(l)$  sob o ponto de vista dos mínimos quadrados (*linear least squares*).

Assim, sejam  $\{x_0, \dots, x_5\}$  e  $\{y_0, \dots, y_5\}$  os conjuntos de abscissas e ordenadas referentes às medianas vetoriais dos elementos das nuvens de  $\mathcal{L}(0)$  a  $\mathcal{L}(5)$ , e  $\beta_x^{(o)}$  e  $\gamma_x^{(o)}$ , os parâmetros da reta de ajuste a serem encontrados, fazemos:

$$\mathbf{A} = \begin{bmatrix} x_0 & 1 \\ \vdots & \vdots \\ x_5 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} y_0 \\ \vdots \\ y_5 \end{bmatrix} \text{ e } \mathbf{x} = \begin{bmatrix} \beta_x^{(o)} & \gamma_x^{(o)} \end{bmatrix}^T.$$

A solução de  $\mathbf{x}$  pelo método dos mínimos quadrados [30] é dada por:

$$\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}. \quad (3.1)$$

Esses dois parâmetros definem a reta longitudinal na região dos olhos, conforme Figura 3.2(c). A reta transversal que forma o primeiro par de eixos de simetria é registrada a partir da rotação e translação da primeira reta. De modo a fazer com que as retas fiquem ortogonais entre si e concorrentes no ponto médio dos PFFs na região em  $(x, y)$ . Assim, os coeficientes angular e linear do novo eixo,  $\beta_y^{(o)}$  e  $\gamma_y^{(o)}$ , são obtidos por:

$$\beta_y^{(o)} = \tan \left( \arctan(\beta_x^{(o)}) - \frac{\pi}{2} \right) \quad (3.2a)$$

$$\gamma_y^{(o)} = - \left( \frac{x_0 + x_1 + \dots + x_5}{6} \right) \beta_y^{(o)} + \frac{y_0 + y_1 + \dots + y_5}{6}. \quad (3.2b)$$

Com isso definimos o primeiro par de eixos de simetria para a região dos olhos. Uma ilustração desse processo pode ser visualizada ao início deste capítulo. As Figuras 3.2(a) e 3.2(b), pág. 15, mostram os eixos tomados a partir das marcações de Sabino e Araujo. E as Figuras 3.2(c) e 3.2(d), a partir das nuvens dos classificadores locais.

Ver-se-á, mais adiante, que o parâmetro  $\beta_x^{(o)}$  será transportado para as demais regiões para a formação das novas retas longitudinais que comporão os demais eixos de simetria.

---

<sup>3</sup>Os superescritos  $^{(o,n,b)}$  serão usados desta seção em diante referenciando a região dos olhos, nariz e boca respectivamente.

### 3.1.2 Conversão de Base dos Elementos das Nuvens do DPI

Para facilitar a busca de elementos simétricos em torno dos eixos definidos na Seção anterior, realizamos a mudança de base nos elementos das nuvens  $\mathcal{N}(l)$  de acordo com as suas regiões. Essa transformação envolve três passos simples: reflexão vertical dos eixos originais, rotação proporcional aos parâmetros  $\beta_x^{(o,n,b)}$  e translação da origem  $(0,0)$  aos pontos de concorrência dos eixos de simetria.

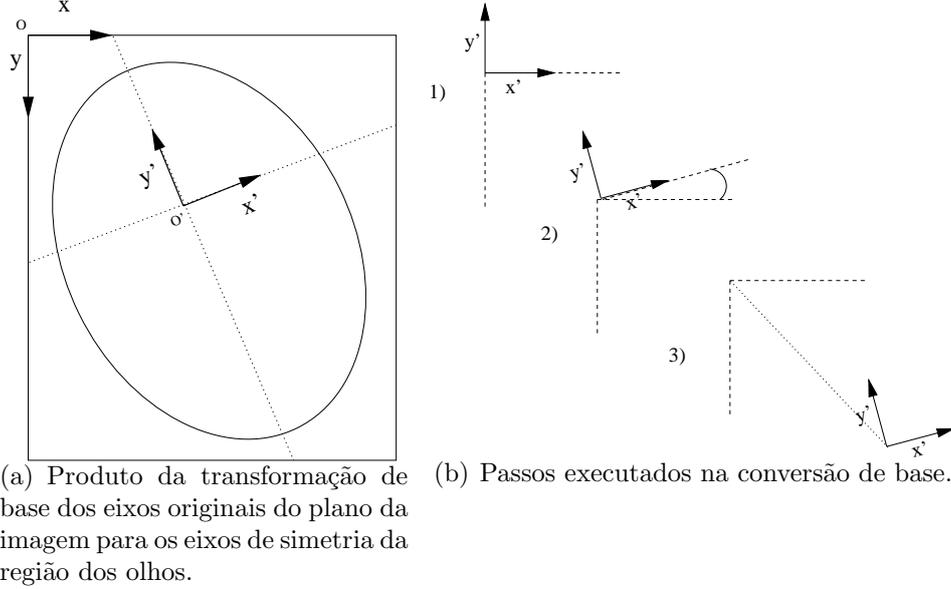


Figura 3.4: Conversão de base dos elementos das nuvens  $\mathcal{N}(l)$ . Em (a), para agilizar o processo de busca de pares de PFFs simétricos, uma mudança de base é executada previamente. Em (b), o processo de conversão se dá em 3 passos: 1) reflexão vertical dos eixos originais; 2) rotação proporcional a  $\beta_x^{(o)}$ ; e 3) translação da origem  $(0,0)$  aos pontos médios dos PFFs da região onde se encontra o par de eixos de simetria.

Assim, sejam  $\overline{\mathbf{Q}}$  a matriz de reflexão,  $\overline{\mathbf{P}}$  a matriz de rotação, faça:

$$\overline{\mathbf{Q}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

$$\overline{\mathbf{P}} = \begin{bmatrix} \cos\left(\arctan\left(\beta_x^{(o)}\right)\right) & \sin\left(\arctan\left(\beta_x^{(o)}\right)\right) \\ -\sin\left(\arctan\left(\beta_x^{(o)}\right)\right) & \cos\left(\arctan\left(\beta_x^{(o)}\right)\right) \end{bmatrix}.$$

A conversão de base, bem como seu processo inverso, é dada pelas equações abaixo:

$$\widetilde{\nu}_i(l) = \overline{\mathbf{Q}} \overline{\mathbf{P}} (\nu_i(l) - \mathbf{o}') \quad (3.4a)$$

$$\nu_i(l) = (\overline{\mathbf{Q}} \overline{\mathbf{P}})^{-1} \widetilde{\nu}_i(l) + \mathbf{o}' \quad (3.4b)$$

onde  $\mathbf{o}'$  é a origem dos eixos de simetria de uma dada região,  $\nu_i(l)$ , o  $i$ -ésimo elemento de uma nuvem  $\mathcal{N}(l)$  daquela região, e  $\widetilde{\nu_i(l)}$ , sua coordenada convertida para a nova base.

### 3.1.3 Assimetrias a partir das Marcações de Sabino e Araujo

Antes de avançarmos na estruturação do método, é preciso tomar conhecimento das assimetrias entre pares de PFFs oriundas da marcação manual nas bases de imagens de faces. Esse levantamento foi executado sobre dez (de onze<sup>4</sup>) pontos fiduciais da marcação de Sabino e Araujo (Figura 3.1(a)) em cada uma das imagens selecionadas dos bancos de faces a saber.

#### Bancos de Imagens de Faces

Duas foram as bases de imagens de faces utilizadas neste trabalho: FERET [18] e BioID [17].

Da base BioID, 503 imagens de faces frontais foram selecionadas. E da base FERET, 1988 imagens. Cada uma dessas bases foi particionada em 7 *folds* para a validação cruzada do tipo *k-fold*. Nesse procedimento de avaliação,  $k - 1$  partições são usadas para treinamento, e uma para testes.

Esse procedimento de levantamento de assimetrias pode ser encarado como o processo de treinamento para a aplicação do método sobre as saídas do processamento de imagens de faces pelo DPI-PCA. Buscamos, então, limiares de tolerância para as assimetrias entre pares de pontos fiduciais ditos simétricos entre si (conforme listagem do início da Seção 3.1, pág. 16).

Até aqui, a descrição do método levou em consideração as nuvens do DPI para a geração dos eixos de simetria e transformações de base. Para a análise de assimetrias, a fonte de dados será a base de marcações manuais de Sabino e Araujo. Isso apenas simplifica os procedimentos à medida em que não mais se torna necessária a extração de medianas vetoriais de nuvens  $\mathcal{N}(l)$ , mas sim a utilização direta dos pontos fiduciais  $\mathcal{L}(l)$  marcados.

#### Registrando as Assimetrias entre Pares Simétricos de PFFs

Dados: uma imagem de face devidamente marcada, um par de eixos de simetria em uma das regiões da face, as matrizes de conversão de base  $\overline{\mathbf{P}}$  e  $\overline{\mathbf{Q}}$  (Equações 3.4a e 3.4b), e um par simétrico de PFFs  $\{\mathcal{L}(l), \mathcal{L}(s)\}$ ,  $l \neq 7$ , daquela região, fazemos:

---

<sup>4</sup>Lembramos que o PFF  $\mathcal{L}(7)$ , centro do nariz, não possui par simétrico.

- Reescalonamento da imagem de face em análise para a dimensão 220x220 (conforme Seção 2.2.3, 8), bem como o transporte das coordenadas de  $\mathcal{L}(l)$  e  $\mathcal{L}(s)$  para essa dimensão.
- Conversão de  $\mathcal{L}(l) = (x_l, y_l)$  e  $\mathcal{L}(s) = (x_s, y_s)$  para a base dos eixos de simetria, resultando em:

$$\begin{aligned}\widetilde{\mathcal{L}}(l) &= (\widetilde{x}_l, \widetilde{y}_l) = \overline{\mathbf{Q}} \overline{\mathbf{P}} (\mathcal{L}(l) - \mathbf{o}'), \text{ e} \\ \widetilde{\mathcal{L}}(s) &= (\widetilde{x}_s, \widetilde{y}_s) = \overline{\mathbf{Q}} \overline{\mathbf{P}} (\mathcal{L}(s) - \mathbf{o}').\end{aligned}$$

- Conversão das coordenadas retangulares  $(x_l, y_l)$  e  $(x_s, y_s)$  em coordenadas polares:  $r_l \angle \varphi_l$  e  $r_s \angle \varphi_s$ .
- Reflexão de  $\widetilde{\mathcal{L}}(l)$  em torno do eixo de simetria  $y'$ :  $r_l \angle (\pi - \varphi_l)$ .
- Finalmente, os erros de assimetrias são registrados por:

$$e_l = r_l - r_s \tag{3.5a}$$

$$\delta_l = \varphi_s - (\pi - \varphi_l) \tag{3.5b}$$

onde  $e_l$  e  $\delta_l$  são os erros de assimetria em módulo e ângulo respectivamente.

## 3.2 Treinamento do DPI-SIM

As marcações de Sabino e Araujo sobre as bases de faces de imagens FERET e BioID contém as formações de interesse para o treinamento do método DPI-SIM. Nesta etapa do método, são levantados os limiares para os erros de assimetria descritos na seção anterior.

De forma simples, o treinamento do DPI-SIM se resume a registrar os erros admissíveis de assimetria em módulo  $\{e_l^{(min)}, e_l^{(max)}\}$  e ângulo  $\{\delta_l^{(min)}, \delta_l^{(max)}\}$ , conforme Equações 3.5a e 3.5b. Isso é feito tomando como referência um dos lados do eixo de simetria. Ou seja, os erros são extraídos a partir de cinco PFFs da marcação de Sabino e Araujo. Para efeitos deste trabalho, escolhemos o lado esquerdo da imagem, o qual apresenta o lado direito das faces dos bancos de imagens: PFFs  $\mathcal{L}(0)$ ,  $\mathcal{L}(1)$ ,  $\mathcal{L}(2)$ ,  $\mathcal{L}(6)$  e  $\mathcal{L}(9)$ .

O Algoritmo 1 mostra em detalhes os passos para obtenção dos limiares que são utilizados durante a aplicação do método DPI-SIM às nuvens de DPI. E a Figura 3.5 mostra de forma simplificada o produto deste procedimento.

**Entrada:** Marcações dos 11 pontos fiduciais faciais nas  $P$  imagens dos *folds* de treino e suas respectivas dimensões.

```

1 para cada imagem  $I^{(p)}$  com  $p = 1$  até  $P$  faça
2   Carregue a dimensão da imagem.
3   para cada  $l = 0$  até 10 faça
4     % Neste algoritmo, escolhemos o lado esquerdo das imagens (lado
5     direito das faces) como referência para levantamento de erros de
6     assimetria.
7     se  $l = 0, 1, 2, 6, 9$  então
8       Inicialize os registros de tolerâncias a erros de simetria para todos os
9       pontos fiduciais os quais possuam pares simétricos:  $e_l^{(min)}$ ,  $e_l^{(max)}$ ,  $\delta_l^{(min)}$  e
10       $\delta_l^{(max)}$ .
11     fim se
12     Traduza as coordenadas de  $\mathcal{L}(l)$  desta imagem para a dimensão 220x220.
13   fim para cada
14   Construa os eixos de simetria para a região dos olhos desta imagem, conforme
15   descrição na Seção 3.1.1, pág. 18, e Equações 3.1, 3.2a e 3.2b.
16   Construa os eixos de simetria para as demais regiões levando em consideração a
17   inclinação da reta longitudinal da região dos olhos.
18   De posse de  $\beta_x^{(o,n,b)}$  e  $\mathbf{o}'^{(o,n,b)}$  oriundos dos passos anteriores, preencha as matrizes
19   de conversão de base  $\overline{\mathbf{P}}^{(o,n,b)}$  e  $\overline{\mathbf{Q}}^{(o,n,b)}$  conforme Seção 3.1.2.
20   para cada  $l = 0$  até 2,  $l = 6$  e  $l = 9$  faça
21     Transforme  $\mathcal{L}(l)$  e  $\mathcal{L}(s)$  para a nova base dos eixos de simetria conforme
22     Equações 3.4a e 3.4b
23     Converta os pares  $(\tilde{x}_l, \tilde{y}_l)$  e  $(\tilde{x}_s, \tilde{y}_s)$  oriundos do passo anterior para
24     coordenadas polares:  $r_l \angle \varphi_l$  e  $r_s \angle \varphi_s$ .
25     Reflita  $r_l \angle \varphi_l$  em torno do eixo transversal  $y'$ :  $r_l \angle (\pi - \varphi_l)$ 
26     Apure o erro de assimetria em módulo e ângulo:  $e_l$  e  $\delta_l$  conforme Equações
27     3.5a e 3.5b.
28     se  $e_l < e_l^{(min)}$  então
29        $e_l^{(min)}$  recebe  $e_l$ 
30     fim se
31     se  $\delta_l < \delta_l^{(min)}$  então
32        $\delta_l^{(min)}$  recebe  $\delta_l$ 
33     fim se
34     se  $e_l > e_l^{(max)}$  então
35        $e_l^{(max)}$  recebe  $e_l$ 
36     fim se
37     se  $\delta_l > \delta_l^{(max)}$  então
38        $\delta_l^{(max)}$  recebe  $\delta_l$ 
39     fim se
40   fim para cada
41 fim para cada

```

**Saída:** Os limiares de erro de assimetria  $e_l^{(min)}$ ,  $e_l^{(max)}$ ,  $\delta_l^{(min)}$  e  $\delta_l^{(max)}$ , com  $l = 0, 1, 2, 6, 9$ .

**Algoritmo 1:** Treinamento do Método DPI-SIM.

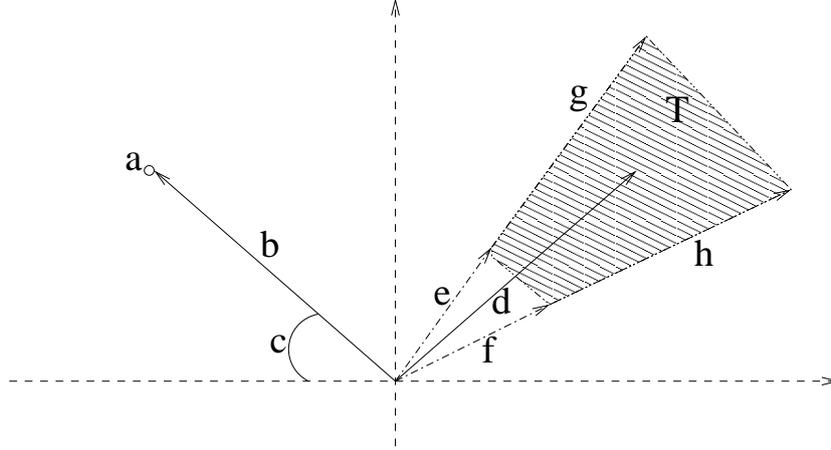


Figura 3.5: Visão simplificada do produto do treinamento do método DPI-SIM. Em a) é localizado um ponto fiducial  $\mathcal{L}(l)$ ; b) módulo  $r_l$ ; c) ângulo  $\varphi_l$ ; d) a reflexão em torno do eixo transversal  $r_l \angle \varphi_l \rightarrow r_l \angle (\pi - \varphi_l)$ ; e)  $r_l + e_l^{(min)} \angle (\pi - \varphi_l + \delta_l^{(max)})$ ; f)  $r_l + e_l^{(min)} \angle (\pi - \varphi_l + \delta_l^{(min)})$ ; g)  $r_l + e_l^{(max)} \angle (\pi - \varphi_l + \delta_l^{(max)})$ ; h)  $r_l + e_l^{(max)} \angle (\pi - \varphi_l + \delta_l^{(min)})$ ; T descreve a área formada a partir dos limiares.

### 3.3 Teste

Este método é aplicado sobre as saídas do processamento de uma imagem de face em prova pelos sistemas de detecção de pontos fiduciais faciais via DPI-PCA, apresentados na Seção 2.2.4. O procedimento de teste do método DPI-SIM pode ser resumido como um processo iterativo que visa a estabilização dos eixos de simetria seguida da aplicação de regras de decisão para a eleição de candidatos a PFFs nas imagens de face em prova. Conforme o treinamento deste método (Seção 3.2), adotamos o lado esquerdo da imagem - PFFs do lado direito das faces - como referência para os pares simétricos.

De posse de todas as nuvens  $\mathcal{N}(l)$  dos PFFs  $\mathcal{L}(l)$ ,  $l = 0, \dots, 10$ , começamos por construir os eixos de simetria para a região dos olhos. Nos casos em que  $\mathcal{N}(l) = \emptyset$ , criamos essa nuvem unitária com a coordenada  $\nu_1(l) = \mu_{\mathcal{L}(l)}$ , onde  $\mu_{\mathcal{L}(l)}$  é a média da distribuição normal definida na Seção 2.2.4, pág. 10. Então, de forma iterativa, passamos a eliminar os candidatos  $\nu_i(l)$ ,  $l = 0, 1, 2$  e  $i = 1, \dots, N_l$ , que não possuem pares simétricos  $\nu_j(s)$ ,  $s = 3, 4, 5$  e  $j = 1, \dots, N_s$ , dentro de uma *região de busca* formada pelos limiares de tolerância à assimetria descritos na Figura 3.5. Com isso, os eixos de simetria se estabilizam e não mais variam em inclinação e posição ( $\beta_x^{(o)}$ ,  $\gamma_x^{(o)}$ ,  $\beta_y^{(o)}$  e  $\gamma_y^{(o)}$ ). Neste momento, escolhemos os candidatos  $\nu_i(l)$  com menor distância ao eixo longitudinal  $x^{(o)}$ , e seus pares mais simétricos  $\nu_j(s)$ . Isto é, candidatos simétricos com menor distância à reflexão de  $\nu_i(l)$  em torno do eixo transversal  $y^{(o)}$ .

Da região dos olhos, passamos às regiões do nariz e boca. E, utilizando as mesmas

inclinações relativas a  $\beta_x^{(o)}$  e  $\beta_y^{(o)}$ , construímos seus eixos de simetria. As mesmas regras de mínimas distâncias aos eixos e pares mais simétricos são usadas para a eleição dos candidatos  $\nu_i(l)$ ,  $l = 6, 9$ , e  $\nu_j(s)$ ,  $s = 8, 10$ .

O centro do nariz recebeu regra especial e própria por não possuir par simétrico. Seu critério de decisão é o de mínima distância à origem dos eixos  $(x^{(n)}, y^{(n)})$ .

O Algoritmo 2 mostra em detalhes a aplicação do método DPI-SIM às nuvens de DPI. E a Figura 3.6 mostra o diagrama de blocos do sistema de detecção de pontos fiduciais faciais baseado em filtros de correlação DPI-PCA e correspondências espaciais por simetria facial.

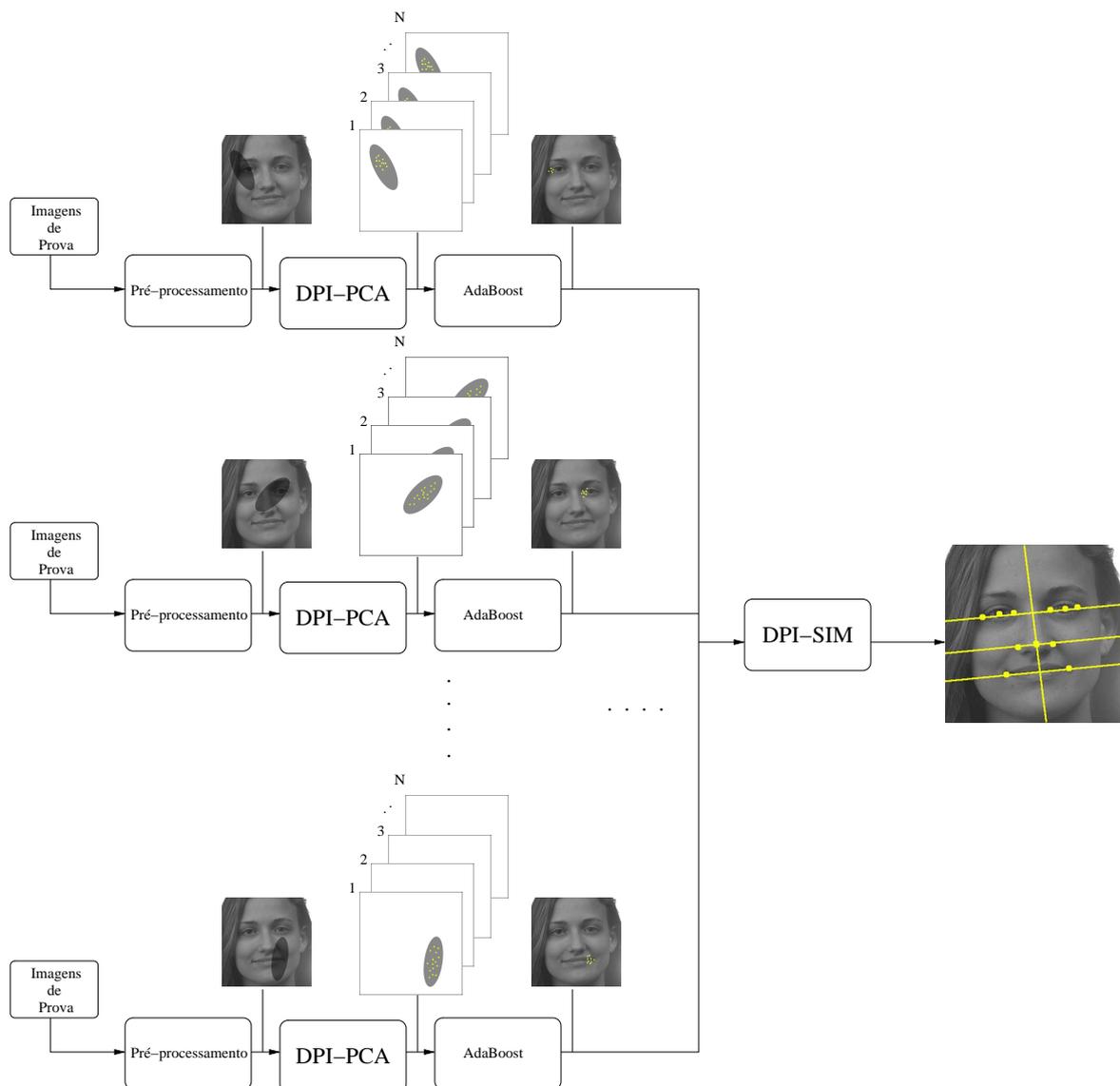


Figura 3.6: Diagrama de blocos do sistema de teste para detecção de pontos fiduciais faciais pelo método DPI-SIM. Como saída, os 11 PFFs da marcação de Sabino e Araujo definidos de forma global e os eixos de simetria posicionados sobre as regiões dos olhos, nariz e boca.

**Entrada:** Nuvens de DPI  $\mathcal{N}(l)$ ,  $l = 0, \dots, 10$ , dos pontos fiduciais faciais  $\mathcal{L}(l)$  em uma imagem de prova  $I$ . Erros de assimetria em módulo  $\{e_i^{(min)}, e_i^{(max)}\}$  e ângulo  $\{\delta_i^{(min)}, \delta_i^{(max)}\}$ .

```

1 Carregue a dimensão da imagem  $I$ .
  % Inicialização
2 para cada landmark  $\mathcal{L}(l)$ ,  $l = 0, \dots, 10$  faça
3   se  $\mathcal{N}(l) = \emptyset$  então
4      $\mathcal{N}(l) = \{\nu_1(l) = \mu_{\mathcal{L}(l)}\}$ , com  $\mu_{\mathcal{L}(l)}$  extraída do modelo gaussiano a priori.
5   fim se
6 fim para cada
  % Região dos olhos.
7 repita
8   Calcule as medianas vetoriais de  $\mathcal{N}(l)$ ,  $l = 0, 1, 2, 3, 4, 5$ , e seu ponto médio  $\mathbf{o}'^{(o)}$ .
9   Construa os eixos de simetria, conforme Seção 3.1.1, pág. 18, e Equações 3.1, 3.2a e 3.2b.
10  para cada  $l = 0$  até 2 faça
11    para cada elemento da nuvem  $\mathcal{N}(l) : \nu_i(l)$ ,  $i = 1, \dots, N_l$  faça
12      Construa a região de busca de elementos simétricos para o candidato  $\nu_i(l)$  e elimine os candidatos a  $\mathcal{L}(l)$  que não possuem pares simétricos nessa região.
13    fim para cada
14  fim para cada
15 até  $\beta_x^{(o)}, \gamma_x^{(o)}, \beta_y^{(o)}, \gamma_y^{(o)} = \text{constantes}$  ;
16 para cada  $l = 0$  até 2 faça
17   Escolha o candidato  $\nu_i(l)$  de menor distância ao eixo longitudinal  $x'^{(o)}$  e seu par mais simétrico  $\nu_i(s)$ .
18 fim para cada
  % Região do nariz.
19 Calcule as medianas vetoriais das nuvens  $\mathcal{N}(l)$ ,  $l = 6, 7, 8$ , e seu ponto médio  $\mathbf{o}'^{(n)}$ .
20 Construa os eixos de simetria utilizando a inclinação correspondente de  $\beta_x^{(o)}$  e a origem  $\mathbf{o}'^{(n)}$ .
21 repita
22   Construa a região de busca de elementos simétricos para o candidato  $\nu_i(6)$  e elimine os candidatos a  $\mathcal{L}(6)$  que não possuem pares simétricos nessa região.
23 até que não haja candidatos  $\nu_i(6)$ ,  $i = 1, \dots, N_6$  a serem eliminados ;
24 Escolha o candidato  $\nu_i(6)$  de menor distância ao eixo longitudinal  $x'^{(n)}$  e seu par mais simétrico  $\nu_i(8)$ .
25 Escolha o candidato  $\nu_i(7)$  de menor distância à origem  $\mathbf{o}'^{(n)}$ .
  % Região da boca.
26 Execute o procedimento de forma análoga ao da região do nariz, usando  $l = 9, 10$ .
27 Traduza os eixos de simetria e as coordenadas dos PFFs eleitos para a dimensão da imagem.
Saída: Os 11 candidatos eleitos dos pontos fiduciais  $\mathcal{L}(l)$  e os eixos de simetria para as regiões da boca, nariz e olhos.

```

**Algoritmo 2:** Aplicação do Método DPI-SIM.

## 3.4 Análise dos Resultados

### 3.4.1 Métrica

Assim como em Araujo [14] e Sabino [1], a métrica adotada para validação de um PFF eleito na aplicação do método descrito é a distância interocular [31]. Essa medida relativa de erro é dada pela Equação 3.6.

$$d_{pf} = \frac{\|\mathcal{L}(l) - \mathcal{L}(l)'\|}{\|\mathcal{L}(4) - \mathcal{L}(1)\|} \quad (3.6)$$

Essa métrica leva em consideração a distância relativa entre o ponto fiducial eleito  $\mathcal{L}(l)'$  contra a marcação *ground-truth*  $\mathcal{L}(l)$  pela distância entre os PFFs  $\mathcal{L}(4)$  e  $\mathcal{L}(1)$ , pupilas esquerda e direita respectivamente. A cada eleição, um PFF é considerado verdadeiro positivo caso  $d_{pf} < 0.1$ . Isto é,  $\mathcal{L}(l)'$  está a menos de 10% da distância interocular de sua marcação.

### 3.4.2 Resultados para a base BioID

Sintetizamos os desempenhos do DPI-SIM para a base de imagens de face BioID nos gráficos das Figuras 3.7, 3.8, 3.9, 3.10, 3.11 e 3.12.

Nos gráficos, as taxas médias de acerto cumulativo do DPI-SIM são comparadas com os melhores resultados do DPI-PCA de Sabino em [1], bem como as taxas obtidas pelo SVM com *kernel* polinomial (também oriundos de [1]). Os resultados são demonstrados por par simétrico, quando possível: cantos externos e internos dos olhos, pupilas, cantos do nariz e boca.

O ponto de corte verifica-se no erro relativo à distância interocular ( $d_{pf}$ ) igual a 10% (0,1). Após análise dos gráficos, pudemos estabelecer o que segue.

#### PFFs 0 e 5 - Cantos externos dos olhos

- As taxas médias de acerto para os cantos externos dos olhos utilizando o DPI-SIM ficaram em torno de 98% para  $\mathcal{L}(0)$  e 97% para  $\mathcal{L}(5)$ .
- O DPI-SIM mostra resultados pouco superiores ou muito próximos em taxa para ambos os PFFs relação aos resultados prévios do DPI-PCA.

#### PFFs 1 e 4 - Pupilas

- As taxas médias de acerto para as pupilas utilizando o DPI-SIM ficaram em torno de 98% para  $\mathcal{L}(1)$  (pupila direita) e 94% para  $\mathcal{L}(4)$  (pupila esquerda).
- O DPI-SIM se equivale aos resultados do DPI-PCA para a pupila direita e perde em desempenho na detecção da pupila esquerda.

### **PFFs 2 e 3 - Cantos internos dos olhos**

- As taxas médias de acerto para os cantos dos olhos utilizando o DPI-SIM ficaram em torno de 98% para  $\mathcal{L}(2)$  e 97% para  $\mathcal{L}(3)$ .
- Assim como nos cantos externos dos olhos, o DPI-SIM mostra resultados pouco superiores ou muito próximos em taxa para ambos os PFFs relação aos resultados prévios do DPI-PCA.

### **PFF 7 - Centro do Nariz**

- As taxas médias de acerto para o centro do nariz utilizando o DPI-SIM são em torno de 96,5%. Desempenho muito superior aos resultados dos antecessores.

### **PFFs 6 e 8 - Cantos do nariz**

- As taxas médias de acerto para os cantos do nariz utilizando o DPI-SIM ficaram em torno de 95% para o canto direito,  $\mathcal{L}(6)$ , e 92% para o esquerdo,  $\mathcal{L}(8)$ .
- Novamente, o DPI-SIM mostra resultados pouco superiores no PFF do lado direito da face, e inferiores em taxa para o PFF do lado esquerdo em relação aos resultados prévios do DPI-PCA.

### **PFFs 9 e 10 - Cantos da boca**

- As taxas médias de acerto para os cantos da boca utilizando o DPI-SIM ficaram em torno de 84% no canto direito da boca,  $\mathcal{L}(9)$ , e 78% no lado esquerdo da boca,  $\mathcal{L}(10)$ .
- O DPI-SIM possui taxas inferiores em relação aos resultados prévios do DPI-PCA. Especialmente pelo canto da boca do lado esquerdo da face.

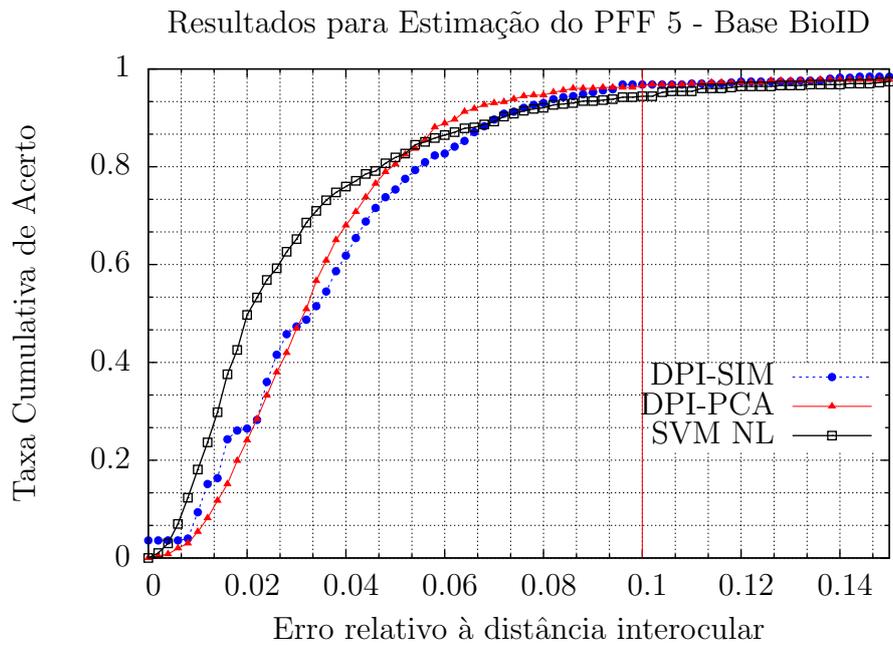
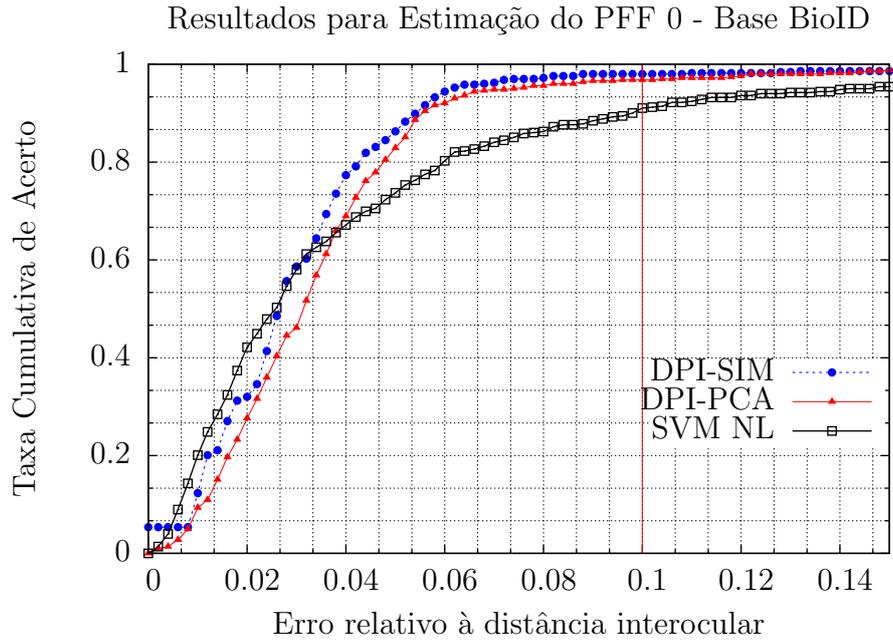


Figura 3.7: Resultados do DPI-SIM para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(0)$  e  $\mathcal{L}(5)$  - cantos externos dos olhos.

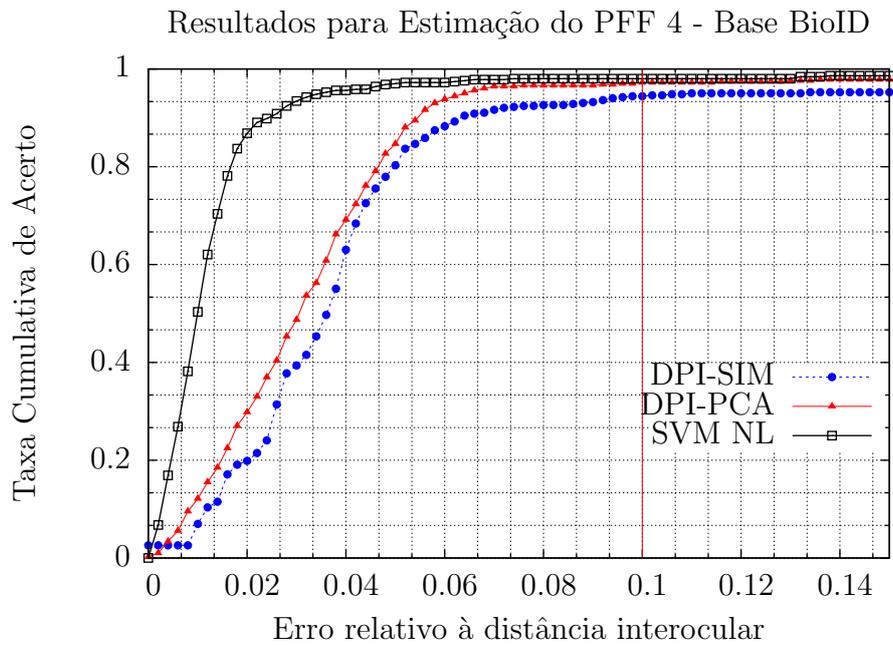
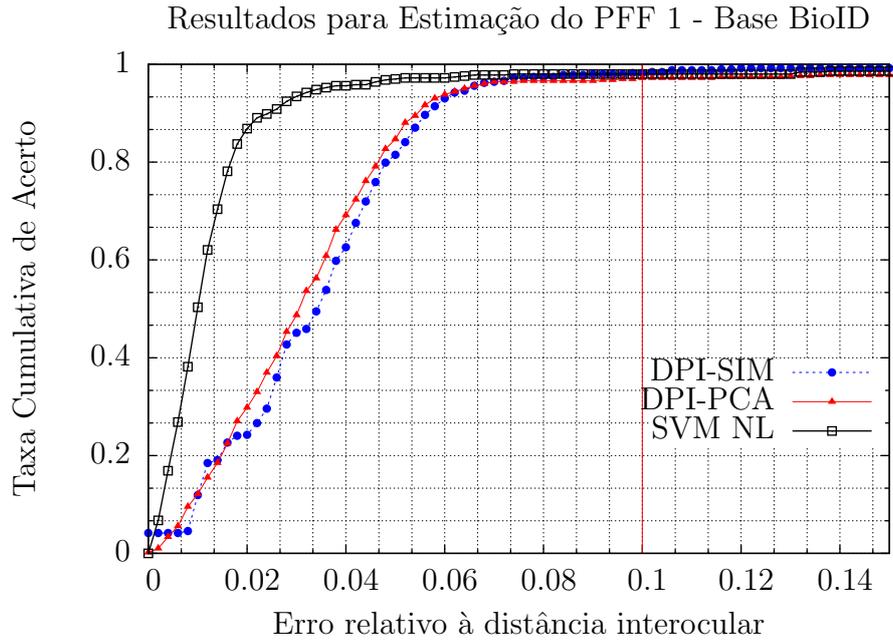


Figura 3.8: Resultados do DPI-SIM para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(1)$  e  $\mathcal{L}(4)$  - pupilas.

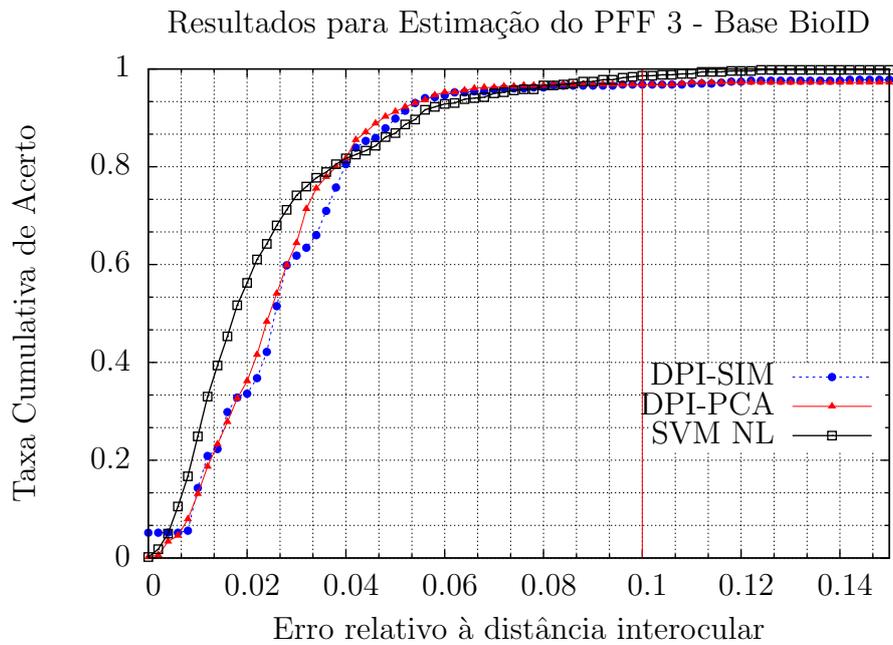
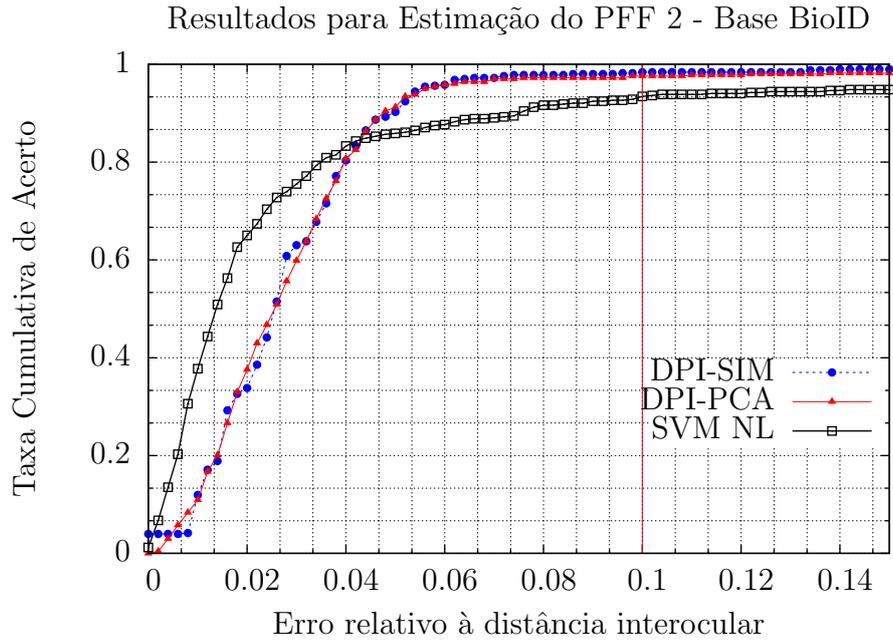


Figura 3.9: Resultados do DPI-SIM para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(2)$  e  $\mathcal{L}(3)$  - cantos internos dos olhos.

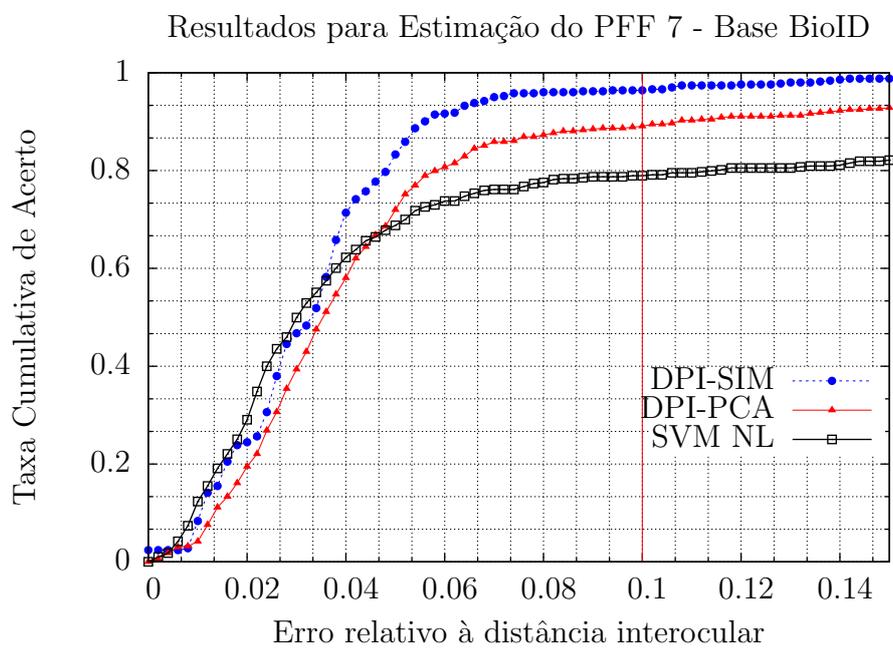


Figura 3.10: Resultados do DPI-SIM para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para o pontol fiducial  $\mathcal{L}(7)$  - centro do nariz.

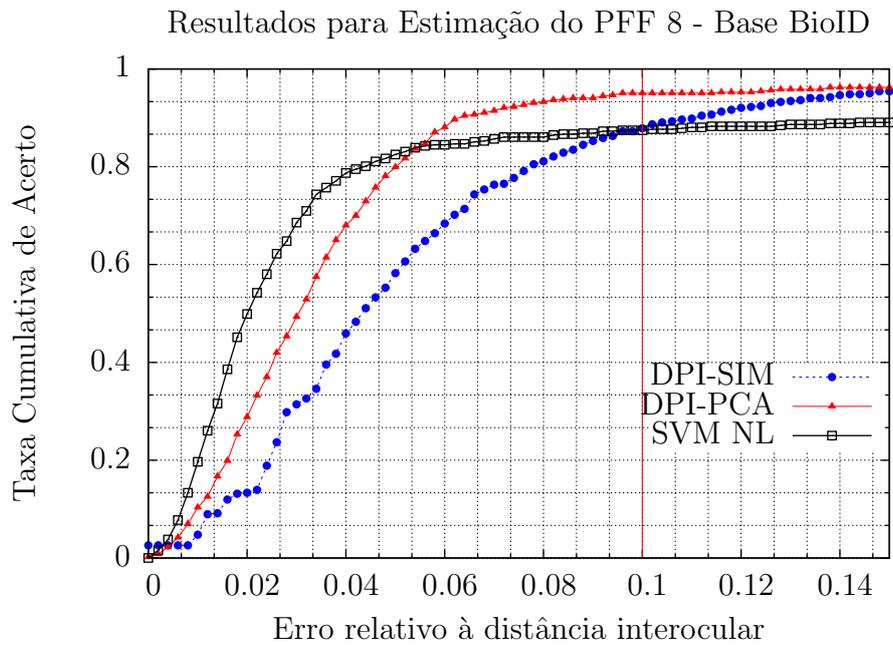
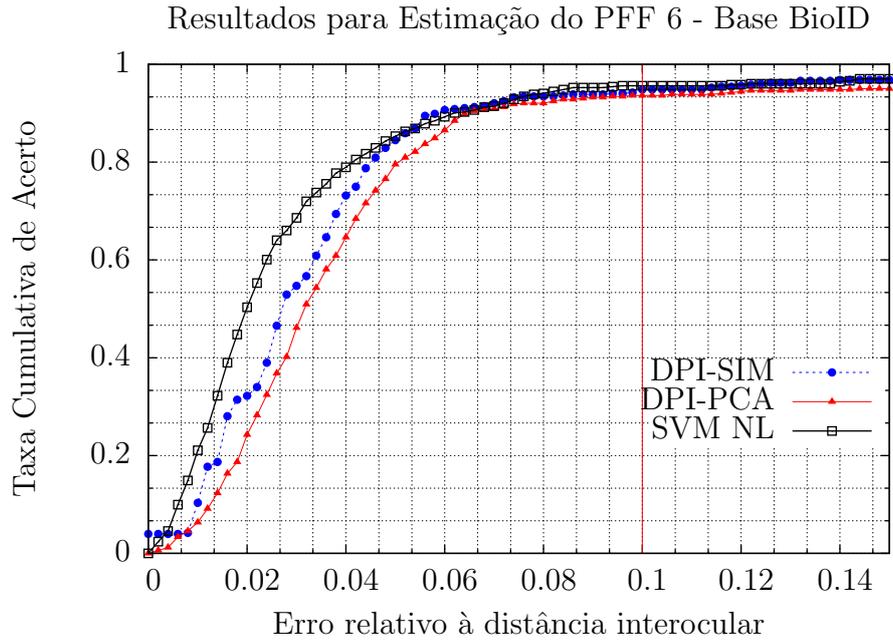


Figura 3.11: Resultados do DPI-SIM para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(6)$  e  $\mathcal{L}(8)$  - cantos do nariz.

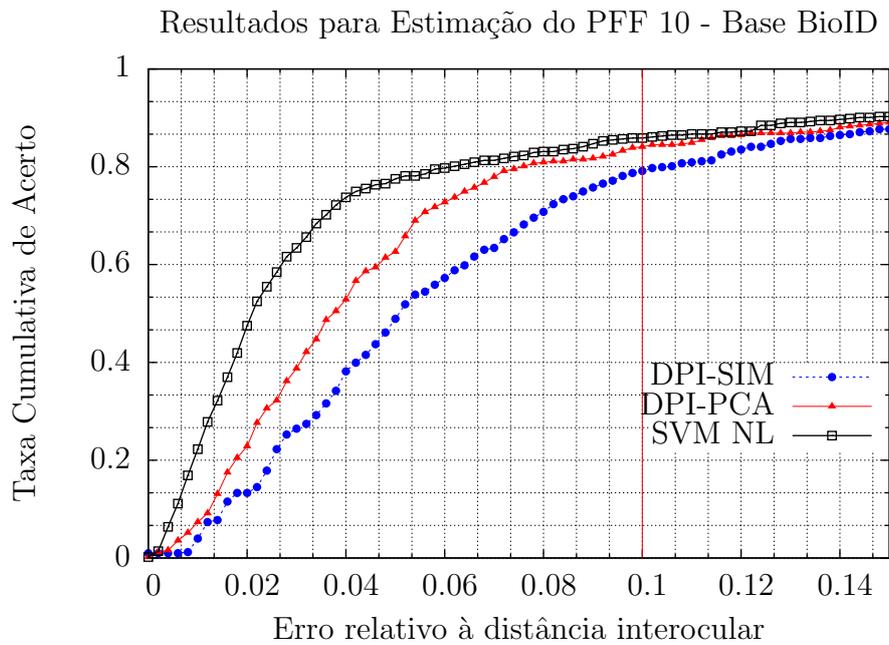
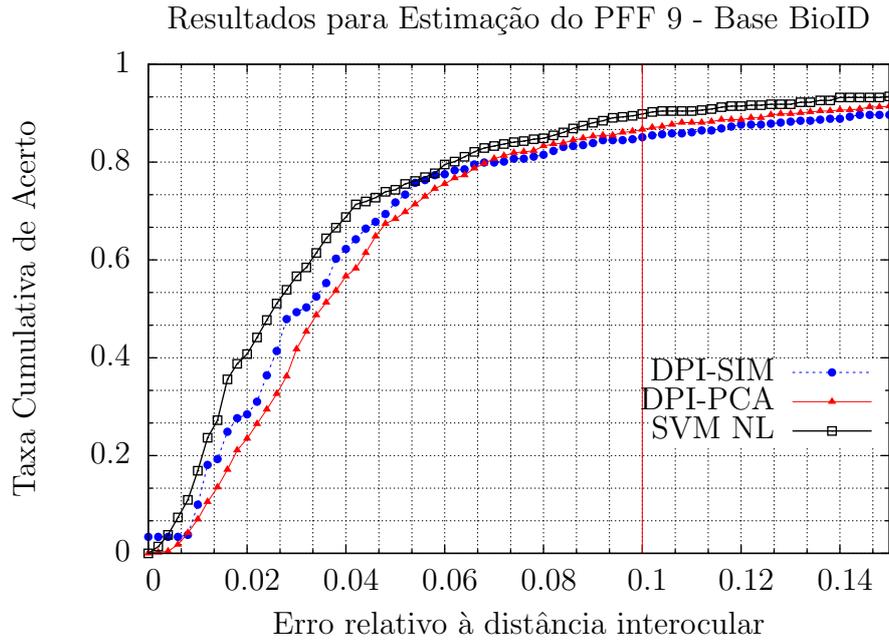


Figura 3.12: Resultados do DPI-SIM para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(9)$  e  $\mathcal{L}(10)$  - cantos da boca.

### 3.4.3 Resultados para a base FERET

Sintetizamos os desempenhos do DPI-SIM para a base de imagens de face FERET nos gráficos das Figuras 3.13, 3.14, 3.15, 3.16, 3.17 e 3.18.

Os mesmos comentários iniciais da Seção 3.4.2 valem aqui.

#### PFFs 0 e 5 - Cantos externos dos olhos

- As taxas médias de acerto para os cantos externos dos olhos utilizando o DPI-SIM ficaram em torno de 85% para  $\mathcal{L}(0)$  e 80% para  $\mathcal{L}(5)$ .
- O DPI-SIM se equivale aos resultados do DPI-PCA para o canto externo do olho direito e perde em desempenho na detecção do canto externo do olho esquerdo.

#### PFFs 1 e 4 - Pupilas

- As taxas médias de acerto para as pupilas utilizando o DPI-SIM ficaram em torno de 92% para  $\mathcal{L}(1)$  (pupila direita) e 84% para  $\mathcal{L}(4)$  (pupila esquerda).
- O DPI-SIM é superior ao DPI-PCA no sucesso em detecção da pupila direita, mas perde em desempenho na detecção da pupila esquerda.

#### PFFs 2 e 3 - Cantos internos dos olhos

- As taxas médias de acerto para os cantos internos dos olhos utilizando o DPI-SIM ficaram em torno de 89% para  $\mathcal{L}(2)$  e 92% para  $\mathcal{L}(3)$ .
- O DPI-SIM mostra resultados inferiores em taxa para ambos os PFFs em relação aos resultados prévios do DPI-PCA.
- Este par simétrico é o único dentre todos, em ambas bases de imagens de faces, onde as taxas de acerto para o PFF  $\mathcal{L}(s = 3)$  são superiores aos do *landmark*  $\mathcal{L}(l = 2)$ .

#### PFF 7 - Centro do Nariz

- As taxas médias de acerto para o centro do nariz utilizando o DPI-SIM são em torno de 91%. Desempenho superior aos resultados dos antecessores.

#### PFFs 6 e 8 - Cantos do nariz

- As taxas médias de acerto para os cantos do nariz utilizando o DPI-SIM ficaram em torno de 90% para o canto direito,  $\mathcal{L}(6)$ , e 86% para o esquerdo,  $\mathcal{L}(8)$ .

- Novamente, o DPI-SIM mostra resultados inferiores em taxa para ambos os PFFs em relação aos resultados prévios do DPI-PCA.

#### **PFFs 9 e 10 - Cantos da boca**

- As taxas médias de acerto para os cantos da boca utilizando o DPI-SIM ficaram em torno de 76% no canto direito da boca,  $\mathcal{L}(9)$ , e 63% no lado esquerdo da boca,  $\mathcal{L}(10)$ .
- O DPI-SIM possui taxas inferiores em relação aos resultados prévios do DPI-PCA. Especialmente pelo canto da boca do lado esquerdo da face, como nos resultados da base BioID.

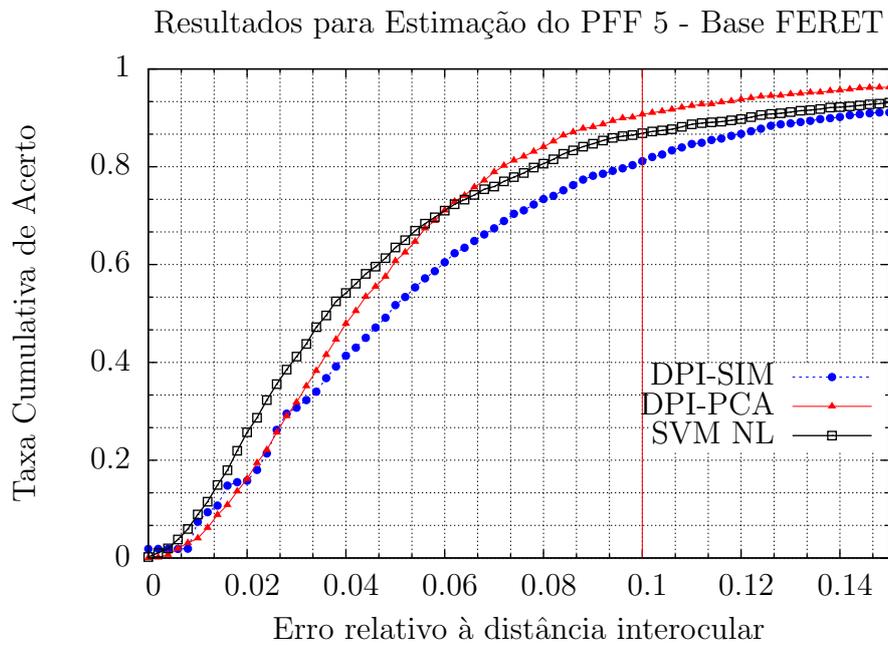
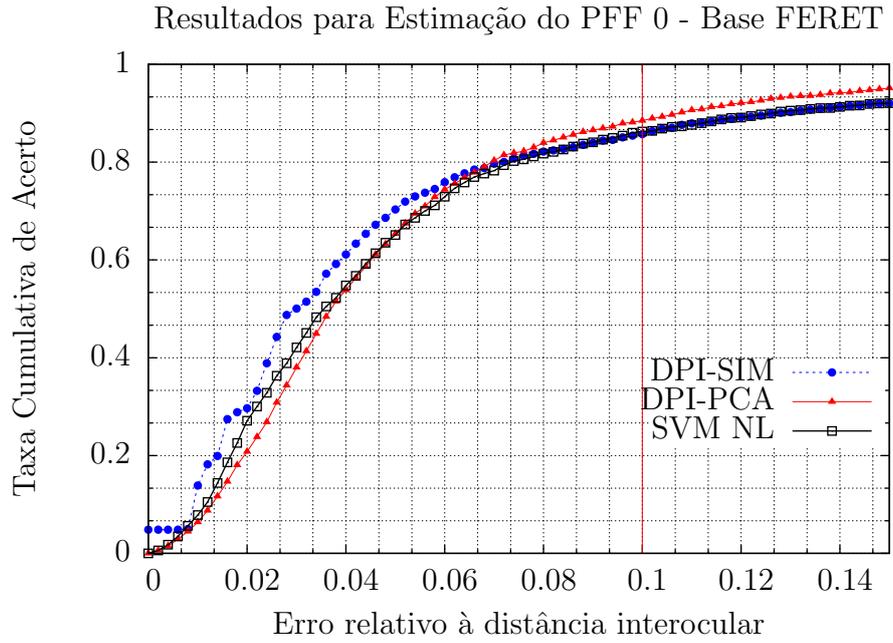


Figura 3.13: Resultados do DPI-SIM para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(0)$  e  $\mathcal{L}(5)$  - cantos externos dos olhos.

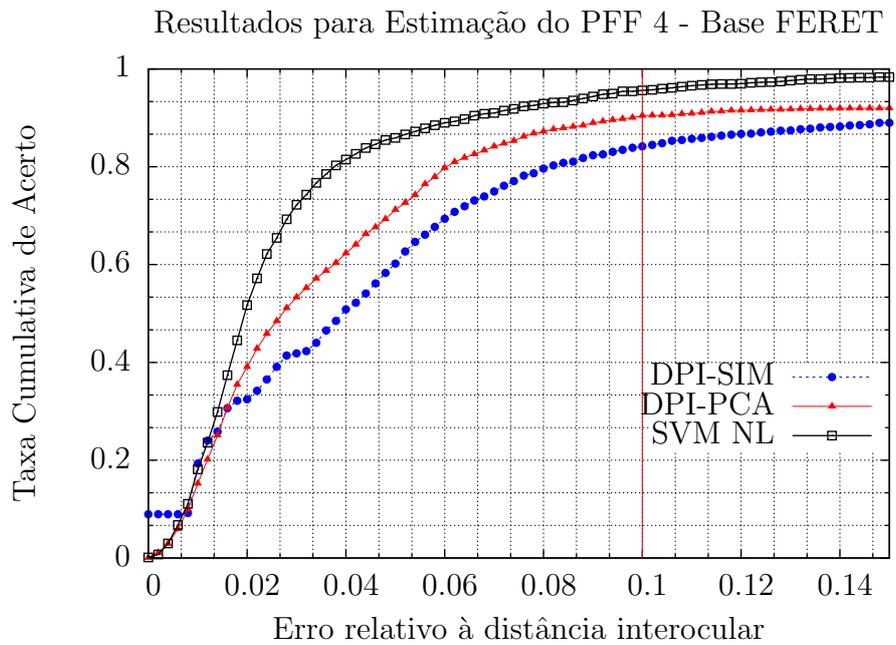
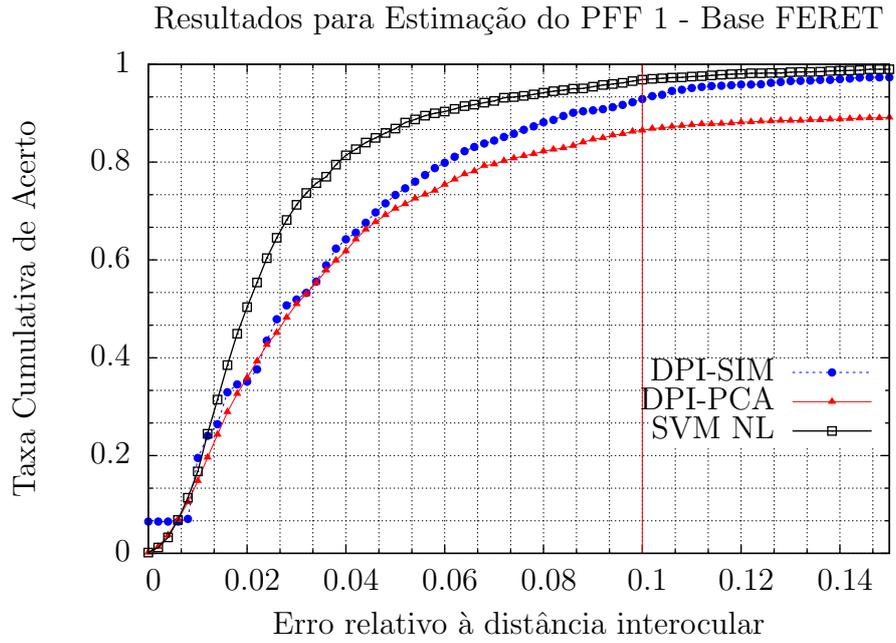


Figura 3.14: Resultados do DPI-SIM para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(1)$  e  $\mathcal{L}(4)$  - pupilas.

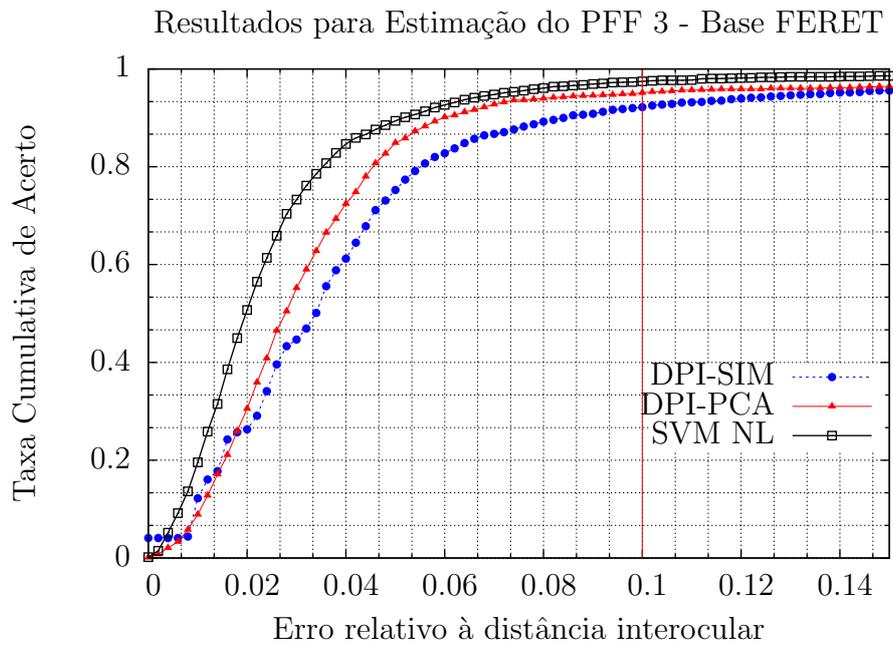
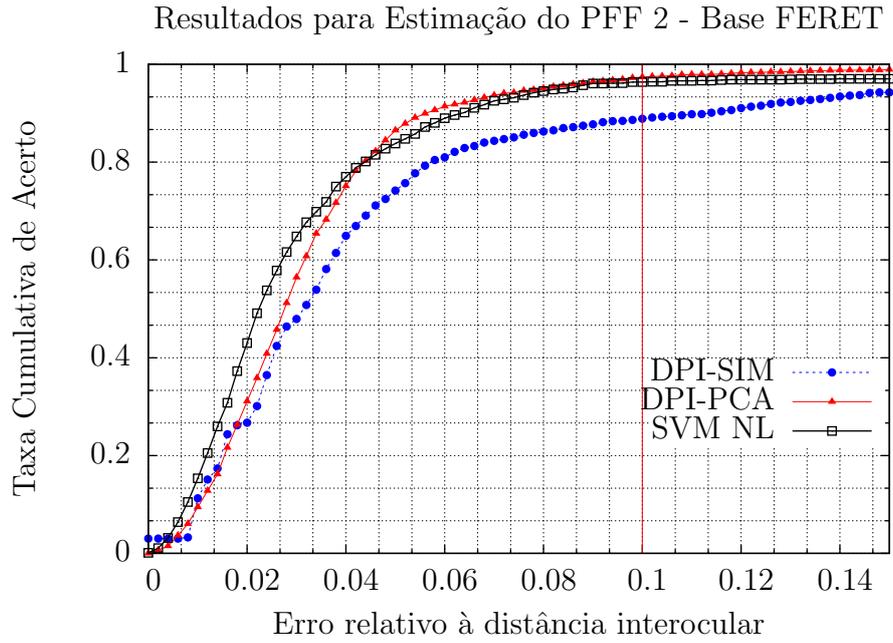


Figura 3.15: Resultados do DPI-SIM para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(2)$  e  $\mathcal{L}(3)$  - cantos internos dos olhos.

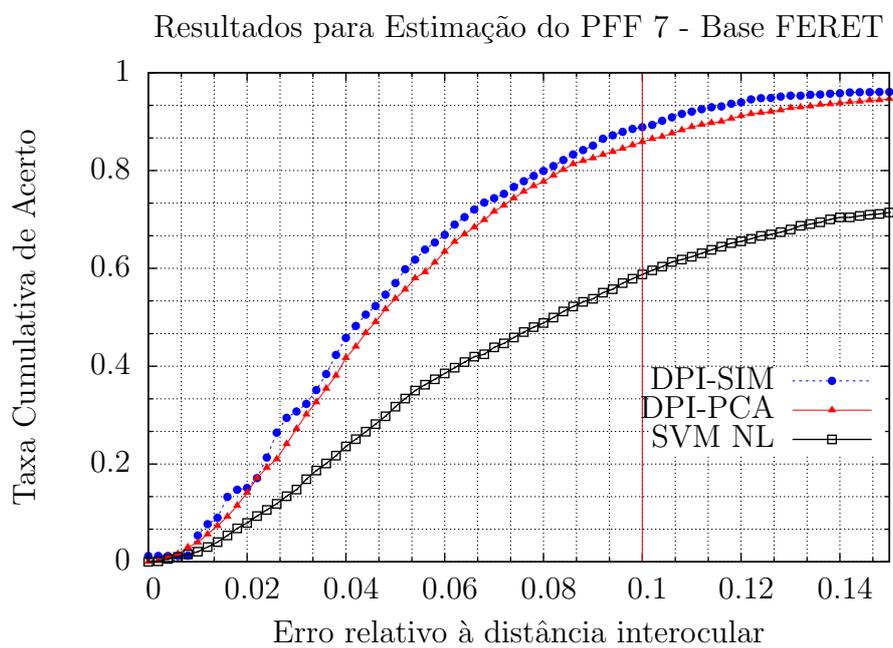


Figura 3.16: Resultados do DPI-SIM para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para o pontol fiducial  $\mathcal{L}(7)$  - centro do nariz.

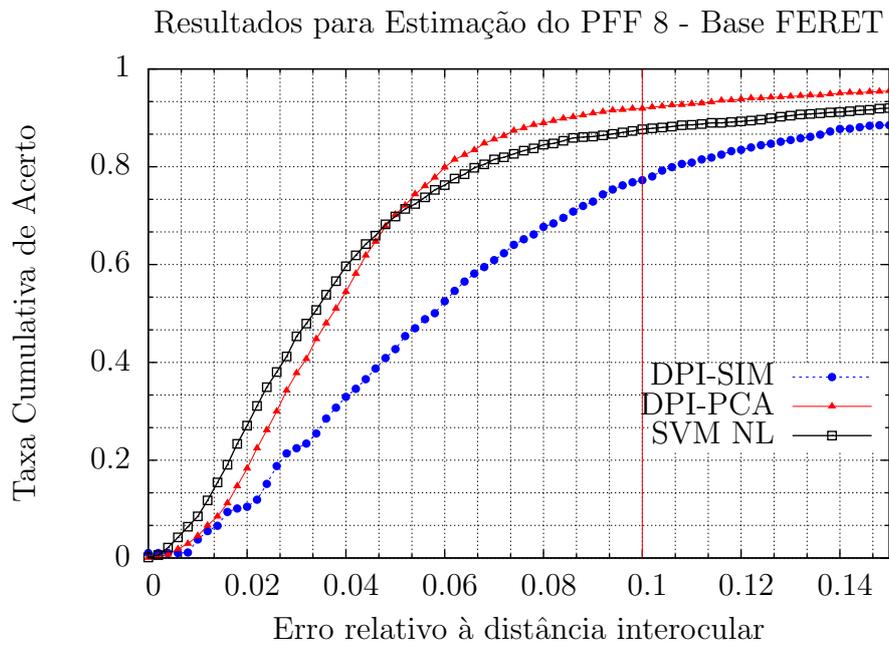
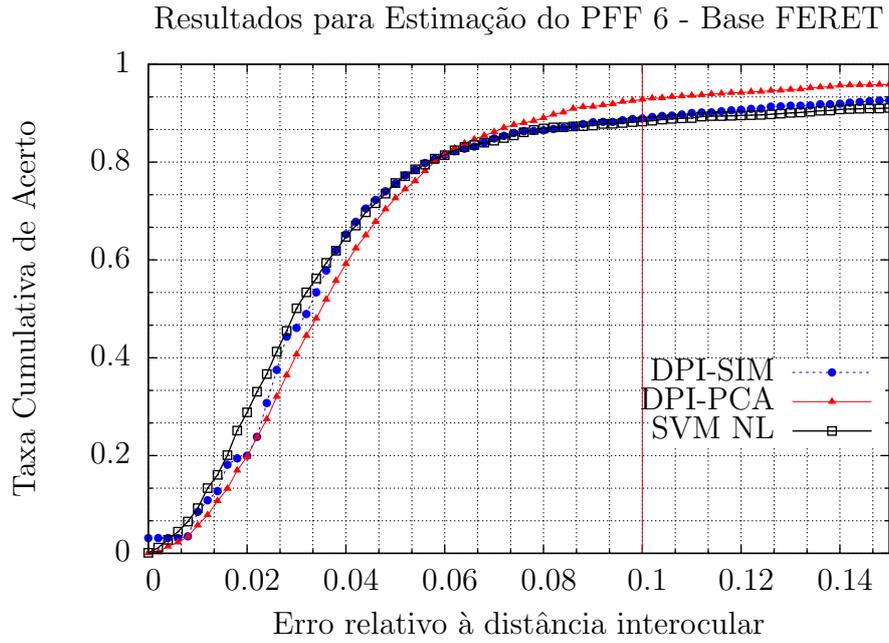


Figura 3.17: Resultados do DPI-SIM para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(6)$  e  $\mathcal{L}(8)$  - cantos do nariz.

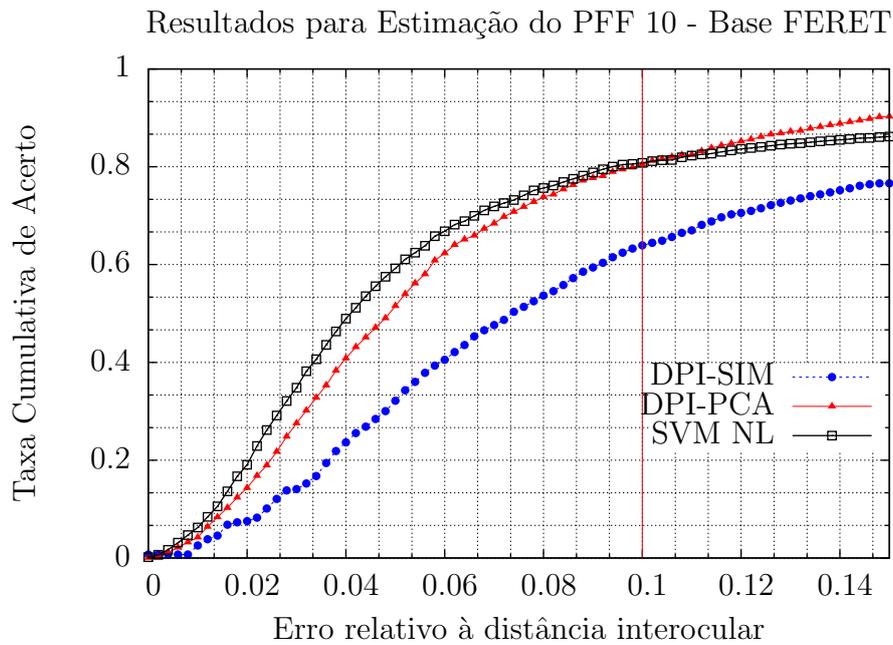
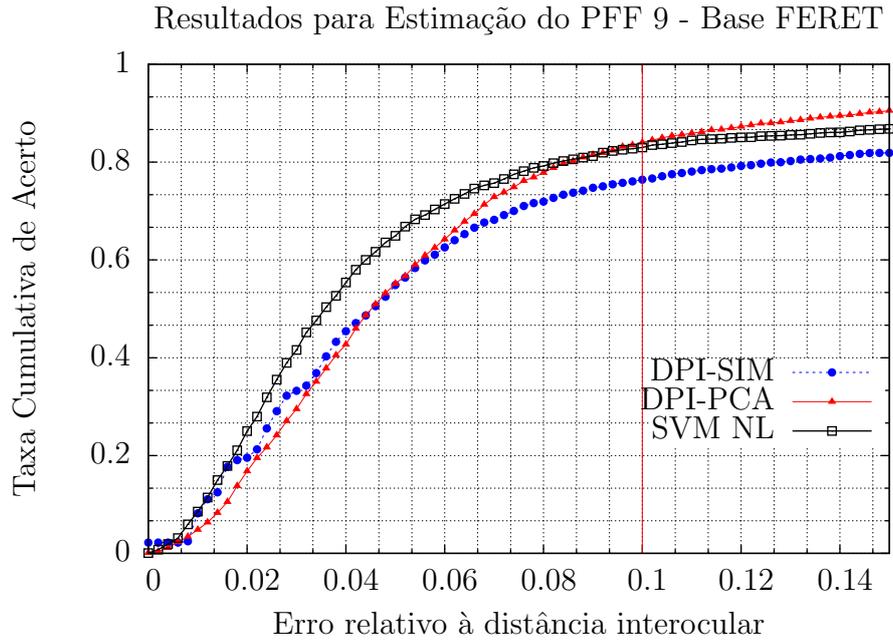


Figura 3.18: Resultados do DPI-SIM para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(9)$  e  $\mathcal{L}(10)$  - cantos da boca.

### 3.4.4 Comparativo de Desempenho do DPI-SIM nas Bases FERET *versus* BioID



(a) Imagens de faces extraídas da base FERET.



(b) Imagens de faces extraídas da base BioID.

Figura 3.19: Exemplos de faces extraídas das bases de imagens FERET e BioID. Em (a), faces oriundas da base FERET. Em (b), faces da base BioID.

Após análise dos gráficos nas Seções 3.4.2 e 3.4.3, cabem alguns comentários:

- O desempenho do DPI-SIM para a base BioID é superior se comparado com seus resultados para a base FERET.
- Isso também vale para os outros métodos comparados neste capítulo, DPI-PCA e SVM-NL, que atingiram marcas superiores na base BioID.
- Fica claro pelas Figuras 3.19(a) e 3.19(b) que há maior variabilidade em assimetria nas imagens de faces, seja por pose ou expressão, na base FERET. Isso implica direta e fortemente nos resultados obtidos pelos métodos na detecção dos pontos fiduciais das imagens de faces dessas bases.

## 3.5 Considerações Finais

Neste capítulo, demonstramos um método que explora correspondências espaciais da face através de eixos de simetria. Sendo estes definidos a partir das nuvens dos classificadores DPI para a eleição global de pontos fiduciais faciais. Para validar o procedimento, foram utilizadas imagens de faces oriundas de duas bases de dados conhecidas: BioID e FERET. A avaliação foi concebida através da metodologia de validação cruzada.

Embora os desempenhos do método DPI-SIM tenham sido em sua maior parte inferiores aos do DPI-PCA, ficou demonstrada a possibilidade de eleger de forma conjunta os PFFs levando em consideração somente relações espaciais. Note que isso é diferente do que foi feito em [1], onde informação de alto nível era usada para a

eleição local de um ponto fiducial (vide Seção 2.2.4, pág. 12): distribuição espacial de probabilidade a priori.

Atenção especial deve ser dada ao fato do único ponto fiducial sem par simétrico ter obtido taxas superiores, até de forma expressiva, sobre os métodos comparados (págs. 31 e 39). O centro do nariz, *landmark*  $\mathcal{L}(7)$ , se valeu da regra de menor distância à origem dos eixos de simetria para a sua região,  $O^{(n)}$ . O eixo de simetria bilateral definido pelo método DPI-SIM,  $y^{(o,n,b)}$ , tende a ocorrer, com frequência, de forma concorrente ao plano sagital que divide as faces. Assim, o PFF  $\mathcal{L}(7)$  leva vantagem por estar, na maioria das observações, sobre esse eixo. Além disso, embora tenhamos proposto um estudo de simetria, é importante destacar que as faces humanas não são perfeitamente simétricas, mas possuem, em média, algo muito aproximado de uma simetria bilateral [32].

Contudo, nos motivamos a buscar uma melhor forma de relacionar os pontos fiduciais da face através de correspondências espaciais. Uma alternativa conhecida e consolidada [33] de levar em consideração similaridades espaciais é dada através do uso de grafos, o que veremos no capítulo que segue.

## Capítulo 4

# Representação Global de Pontos Fiduciais Faciais por Grafos

Grafos são estruturas comprovadamente eficientes e poderosas para representação de objetos em visão computacional e reconhecimento de padrões [33]. Nessa linha, os vértices (também: nós ou pontos) de um grafo normalmente denotam um ponto (ou característica) de uma imagem, e suas arestas, as relações entre esses pontos. Os grafos podem ainda conter informação embutida (rótulos) em seus vértices e arestas, como se verá mais adiante.

Rigorosamente, o casamento de grafos é o processo pelo qual se busca uma correspondência entre os vértices e arestas de dois grafos. Com isso, restrições devem ser satisfeitas garantindo que as estruturas em um grafo sejam mapeadas em estruturas similares no outro grafo [33]. Por outro lado, em visão computacional esse casamento pode ser visto de forma simplificada, como a técnica pela qual dois ou mais grafos são comparados de modo a se obter o par de maior semelhança, seja qualitativamente (ou seja, em estrutura ou forma) ou quantitativamente (informações embutidas) [3].

Neste capítulo, apresentamos uma síntese dos métodos que motivaram o uso de grafos para a detecção global de pontos fiduciais a partir das nuvens de DPI: o EBGM [3] e sua variante estatística [15]. Finalmente, descrevemos o método proposto de casamento de grafos aplicado às nuvens de DPI: o DPI-GRAPH.

### 4.1 Grafos Elásticos: o *Elastic Bunch Graph Matching* (EBGM) e sua variante estatística

O EBGM [3] é um sistema de reconhecimento de faces com base na extração e comparação de características locais posicionadas sobre pontos fiduciais da face, que por sua vez é representada por um grafo.

Foi com Wiskott, em sua tese de doutorado [34], que grafos provaram ser ferramentas poderosas e praticáveis computacionalmente para o reconhecimento e identificação/verificação de faces. O método foi uma evolução natural do trabalho de Lades *et al* [35] que já se utilizava dessas estruturas em conjunto com filtros de Gabor, porém sob a perspectiva de redes neurais artificiais<sup>1</sup>. O sucesso do método de Wiskott e Lades para reconhecimento de faces é fortemente dependente do posicionamento dos vértices do grafo sobre pontos fiduciais. Recentemente, Würtz [15] propôs o uso de estimadores de aprendizado estatístico com o intuito de minimizar os erros desse posicionamento.

Neste breve resumo, enfatizaremos a descrição das ferramentas utilizadas durante a fase de posicionamento de *landmarks* em ambos métodos clássico e estatístico do EBGM. O detalhamento do processo no qual são utilizadas essas ferramentas, isto é, os algoritmos de posicionamento e reconhecimento, foge ao escopo deste trabalho.

#### 4.1.1 EBGM Clássico

Apesar do nome *Elastic Bunch Graph Matching*, o EBGM não denota um casamento de grafos no sentido matemático estrito [33]. Isso se deve à forma como são eleitos os grafos no processo de comparação.

O EBGM clássico se firma sobre três importantes ferramentas: os filtros espaciais de Gabor, o casamento de grafos e o *face bunch graph*.

#### Filtros de Gabor

Os filtros de Gabor chamaram atenção da comunidade de reconhecimento de padrões a partir do momento em que provaram obter boa aproximação para os perfis bidimensionais dos campos receptivos do neurocórtex dos mamíferos [13, 36, 37]. Esses filtros provêm descrições de um estímulo visual nos domínios espacial e da frequência. Isso é feito pela decomposição da informação em sinais elementares que possuem máxima localização simultânea no espaço e na frequência espacial.

A formulação original de Gabor foi desenvolvida no tempo [38]. Com os estudos de Daugman [36] e Marcelja [13], ganhou sua versão 2D. E no EBGM, é aplicada em *wavelets* de dimensões e rotações diversas para a extração de características em segmentos de imagens, conforme a Equação 4.1:

$$\mathcal{J}_j(\mathbf{x}) = \int I(\mathbf{x}')\psi_j(\mathbf{x} - \mathbf{x}')d\mathbf{x}' \quad (4.1)$$

---

<sup>1</sup>Como *Dynamic Link Architecture* (DLA), que foi proposta para servir como processos elementares de comparação e casamento de grafos rotulados em um sistema neuronal artificial. Com isso, os vértices dos grafos viram neurônios-modelo, onde cada um possui informações de característica agregadas como rótulos [34].

onde  $\mathcal{J}_j$  é um coeficiente complexo resultante da convolução de uma fatia da imagem  $I$  centrada em  $\mathbf{x}$  com uma função kernel de Gabor  $\psi_j$ . Através da aplicação dessa *wavelet* em diferentes escalas e rotações (Equação 4.2b abaixo), obtemos um conjunto de coeficientes complexos. A esse conjunto foi dado o nome de *jet* (jato). Os jatos são os vetores de características do EBGm os quais são agregados como *rótulos* nos vértices de grafos. E, portanto, definem as informações quantitativas dos grafos.

As funções kernel de Gabor  $\psi_j$  têm o formato de ondas planas encapsuladas por um sino gaussiano, e são definidas na Equação 4.2a:

$$\psi_j(\mathbf{x}) = \frac{\mathbf{k}_j^T \mathbf{k}_j}{\sigma^2} \exp\left(-\frac{\mathbf{k}_j^T \mathbf{k}_j \cdot \mathbf{x}^T \mathbf{x}}{2\sigma^2}\right) \left[ \exp(i \cdot \mathbf{k}_j^T \mathbf{k}_j) - \exp\left(-\frac{\sigma^2}{2}\right) \right] \quad (4.2a)$$

$$\mathbf{k}_j = \begin{pmatrix} k_{jx} \\ k_{jy} \end{pmatrix} = \begin{pmatrix} k_v \cos(\phi_\mu) \\ k_v \sin(\phi_\mu) \end{pmatrix} \quad (4.2b)$$

$$k_v = 2^{-\frac{v+2}{2}} \pi$$

$$\phi_\mu = \mu \frac{\pi}{8}$$

onde  $\mathbf{k}_j$  é o vetor onda;  $\sigma$ , o parâmetro que controla a largura da gaussiana em  $\sigma/k_v$ ;  $v$ , o índice que controla o número de frequências em  $k_v$ ;  $\mu$ , o índice do número de rotações em  $\phi_\mu$ ; o termo  $\exp\left(-\frac{\sigma^2}{2}\right)$  torna as funções livres da componente DC. Wiskott define um jato como o conjunto de 40 coeficientes complexos  $\mathcal{J} = \{\mathcal{J}_j\}$  obtidos em um ponto da imagem (um ponto fiducial). Esse número é oriundo de 5 frequências distintas, índice  $v = 0, 1, \dots, 4$ , e 8 rotações, índice  $\mu = 0, 1, \dots, 7$ , com o índice  $j$  variando conforme  $j = \mu + 8v$  na Equação 4.2b [3].

### **Face Bunch Graphs**

No EBGm, uma face humana é representada por um grafo rotulado  $\mathcal{G}$  chamado grafo facial. Este grafo possui  $L$  vértices e  $E$  arestas. Cada vértice é posicionado sobre um ponto fiducial facial definido. As arestas conectam os vértices adjacentes. À estrutura resultante do agrupamento de informações extraídas a partir desses grafos é dada o nome de *face bunch graph*.

Um *face bunch graph*  $\mathcal{B}$  (FBG) é a estrutura de dados que armazena simultaneamente  $m = 1, \dots, M$  grafos modelos  $\mathcal{G}^{\mathcal{B}^m}$  de mesma topologia, com vértices posicionados sobre PFFs de imagens de faces e conjuntos de *jets* extraídos nesses pontos. Os *bunch graphs* são rotulados em seus vértices por jatos  $\mathcal{J}(l)$ , os vetores de características<sup>2</sup>. Suas arestas<sup>3</sup> são rotuladas por vetores distância médios

<sup>2</sup>O índice  $l$  denota a qual ponto fiducial  $\mathcal{L}(l)$  o jato  $\mathcal{J}$  pertence.

<sup>3</sup>O subscrito  $e$  fará referência a arestas e adjacências de um *landmark* ao longo do texto.

$\Delta \mathbf{x}_e^{\mathcal{B}} = \sum_m \frac{\Delta \mathbf{x}_e^{\mathcal{B}_m}}{M}$  tomados a partir dos grafos modelos (vide Figura 4.1). Em outras palavras, seus vértices contêm pilhas de jatos extraídos a partir de pontos fiduciais de faces humanas, e suas arestas são rotuladas com as distâncias médias entre esses pontos. Os grafos modelos do *bunch* são criados durante o treinamento, com marcações manuais dos grafos sobre imagens de faces. O FBG é ainda incrementado automaticamente à medida em que novas faces vão sendo adicionadas ao sistema.

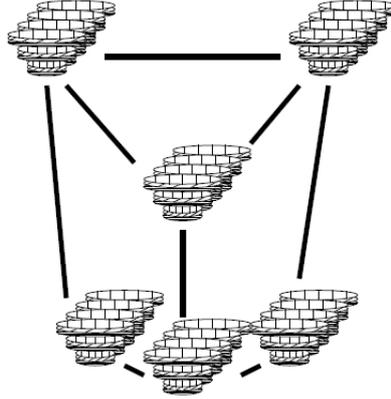


Figura 4.1: O *bunch graph*, estrutura de dados em formato de grafo rotulado. Seus vértices contêm pilhas de jatos extraídos de pontos fiduciais de faces humanas. Suas arestas são rotuladas com as distâncias médias entre esses pontos fiduciais. (Günther [2]).

### Casamento de Grafos

Tanto na fase de detecção quanto na fase de identificação/verificação, o casamento de grafos é determinado por uma função de similaridade. Essa função avalia simultaneamente a proximidade entre as informações quantitativas (jatos) e qualitativas (estrutura geométrica) do grafo na imagem de prova em relação ao FBG.

Devido à rotação de fase, jatos tomados em pontos da face distantes de poucos pixels possuem coeficientes muito diferentes, ainda que representem a mesma característica local [3]. Em Lades *et al* [35], a fase foi ignorada no processo de casamento. Wiskott [34] refinou o método utilizando a fase para estimar o erro de posicionamento de *landmarks*, melhorando seus resultados.

Seja  $\mathcal{J}_j = a_j \exp(i \cdot \phi_j)$ , a Equação 4.3 abaixo avalia a similaridade de jatos de forma sensível à mudança de fase.

$$S_\phi(\mathcal{J}, \mathcal{J}') = \frac{\sum_j a_j a'_j \cos(\phi_j - \phi'_j - \mathbf{d}^T \mathbf{k}_j)}{\sqrt{\sum_j a_j^2 \sum_j a_j'^2}} \quad (4.3)$$

Para a solução da Equação 4.3, o vetor erro de posicionamento  $\mathbf{d}$  deve ser esti-

mado por:

$$\mathbf{d}(\mathcal{J}, \mathcal{J}') = \begin{pmatrix} d_x \\ d_y \end{pmatrix} = \frac{1}{\Gamma_{xx}\Gamma_{yy} - \Gamma_{xy}\Gamma_{yx}} \begin{pmatrix} \Gamma_{yy} & -\Gamma_{yx} \\ -\Gamma_{xy} & \Gamma_{xx} \end{pmatrix} \begin{pmatrix} \Phi_x \\ \Phi_y \end{pmatrix}, \quad (4.4)$$

se  $\Gamma_{xx}\Gamma_{yy} - \Gamma_{xy}\Gamma_{yx} \neq 0$ , com

$$\begin{aligned} \Phi_x &= \sum_j a_j a'_j k_{jx} (\phi_j - \phi'_j), \text{ e} \\ \Gamma_{xy} &= \sum_j a_j a'_j k_{jx} k_{jy}, \end{aligned}$$

e de forma análoga, calcula-se  $\Phi_y$ ,  $\Gamma_{xx}$ ,  $\Gamma_{yy}$  e  $\Gamma_{yx}$ . A Equação 4.4 é resultado de um processo de otimização detalhado em [3].

De posse dessas ferramentas, é possível comparar de forma conjunta dois grafos quantitativa e qualitativamente, através da Equação 4.5 abaixo. Nesta função de similaridade, um grafo na imagem em prova  $\mathcal{G}^I$ , com vértices  $l = 1, \dots, L$  e arestas  $e = 1, \dots, E$ , é comparado com um *bunch graph*  $\mathcal{B}$  de grafos modelos  $m = 1, \dots, M$ :

$$S(\mathcal{G}^I, \mathcal{B}) = \frac{1}{L} \sum_{l=1}^L \max_m (S_\phi(\mathcal{J}^I(l), \mathcal{J}^{\mathcal{B}_m}(l))) - \frac{\lambda}{E} \sum_{e=1}^E \frac{\|\Delta \mathbf{x}_e^I - \Delta \mathbf{x}_e^{\mathcal{B}}\|^2}{\|\Delta \mathbf{x}_e^{\mathcal{B}}\|^2} \quad (4.5)$$

onde:  $S_\phi$  é definida na Equação 4.3,  $\Delta \mathbf{x}_e^I$  são os vetores distância entre vértices na imagem, e  $\Delta \mathbf{x}_e^{\mathcal{B}}$ , os vetores distância médios que rotulam as arestas do *bunch graph*. Essa equação, portanto, quantifica a correspondência entre um grafo e o FBG selecionando o conjunto de jatos do *bunch*  $\mathcal{J}^{\mathcal{B}}$  que maximizam a similaridade com os jatos  $\mathcal{J}^I$  da imagem em prova. Além disso, as arestas da imagem em prova são comparadas com as informações qualitativas do *bunch*, com  $\lambda$  adicionando ou retirando peso dessa comparação.

### 4.1.2 EBGM com Máxima Verossimilhança

Pela Equação 4.3, é possível verificar que as informações de posição de PFFs e de arestas dos grafos são ignoradas no momento de correção do erro de posicionamento de *landmarks*. Para agregar a comparação dessas informações de forma rápida, Würtz [15] propôs um estimador de máxima verossimilhança (*Maximum Likelihood* - ML [39]). Com isso, ficou possível combinar informações de mais alto nível acerca de vértices e arestas a jatos no processo de posicionamento.

Além disso, no teste de similaridade geral da Equação 4.5, as informações qualitativas são simplesmente as distâncias euclidianas entre vértices adjacentes  $\|\Delta \mathbf{x}_e\|$ . Isso ignora o fato de os *landmarks* possuírem tendências bem definidas com respeito aos seus posicionamentos no plano das imagens de faces.

Würtz baseou-se no desenvolvimento de Moghaddam *et al* [40] onde foi demonstrado o uso de estimadores ML aplicados a *eigenfaces* para detecção de faces. E utilizou as simplificações de Teixeira [41] sobre o trabalho anterior para chegar à forma fechada que descrevemos a seguir.

### Estimador da Estrutura do Grafo

O estimador da estrutura do grafo faz uso dos vértices  $\mathcal{L}$  e arestas  $\mathcal{E}$  de grafos para estimar a posição de um PFF em relação a seus adjacentes.

Da parcela da Equação 4.5 referente à estrutura do grafo, utilizam-se as componentes horizontal e vertical do vetor distância  $\Delta \mathbf{x}_e = \Delta(l, l_e) = \mathcal{L}(l) - \mathcal{L}(l_e)$ , isto é,  $\Delta_x(l, l_e)$  e  $\Delta_y(l, l_e)$ , sendo  $\mathcal{L}(l_e)$  a coordenada de um *landmark* adjacente a  $\mathcal{L}(l)$ .

Cada *landmark*  $\mathcal{L}(l)$  está conectado a *landmarks* vizinhos  $\mathcal{L}(l_e)$ . O número de conexões a partir de  $\mathcal{L}(l)$  define o grau do vértice  $d_G(\mathcal{L}(l))$  (detalhes no Apêndice A). As médias e variâncias  $\mu_x(l, l_e)$ ,  $\mu_y(l, l_e)$ ,  $\sigma_x(l, l_e)^2$  e  $\sigma_y(l, l_e)^2$  das variáveis  $\Delta_x(l, l_e)$  e  $\Delta_y(l, l_e)$  são tomadas a partir dos grafos modelos do conjunto de treinamento. Elas são extraídas para todos os vértices  $\mathcal{L}(l)$  e seus adjacentes  $\mathcal{L}(l_e)$  independentemente.

De posse dessas informações é possível calcular a similaridade da posição de um PFF com relação a suas adjacências a partir da Equação 4.6 abaixo.

$$S_{\Delta}(\mathcal{L}(l)) = -\frac{1}{2d_G(\mathcal{L}(l))} \sum_{l_e \text{ adjacente a } l} \left[ \frac{(\mathcal{L}_x(l) - \mathcal{L}_x(l_e) - \mu_x(l, l_e))^2}{\sigma_x(l, l_e)^2} + \frac{(\mathcal{L}_y(l) - \mathcal{L}_y(l_e) - \mu_y(l, l_e))^2}{\sigma_y(l, l_e)^2} \right] \quad (4.6)$$

Essa função ecoará seu valor máximo quando um candidato a  $\mathcal{L}(l)$  guardar maior proximidade em relação a suas adjacências de acordo com a tendência contida no conjunto de treinamento.

Este desenvolvimento é de fundamental importância para o método que descreveremos a seguir.

## 4.2 DPI-GRAPH: o DPI com Casamento de Grafos

Este método prevê a representação do conjunto dos onze PFFs da marcação de Sabino/Araujo através de um grafo. Assim como no DPI-SIM (Capítulo 3), demonstraremos a possibilidade de eleger candidatos a PFFs de forma global, objetivo deste trabalho, através de correspondências espaciais. Desta vez, escolheremos um dentre tantos grafos possíveis a partir do agrupamento de nuvens dos classificadores locais.

Para efeitos comparativos, e de forma simplificada, o EBGM inicia o posicionamento de *landmarks* ao sobrepor um grafo padrão, uma cópia do *grafo médio*  $\mathcal{B}$ , a uma imagem de face. O grafo derivado da imagem vai sofrendo modificações em sua forma através de um processo iterativo<sup>4</sup>, cujos detalhes fogem ao escopo desta dissertação, até que se encontre um grafo  $\mathcal{G}$  rotulado com maior similaridade em relação ao *bunch* [3].

No nosso caso, cada elemento de cada uma das nuvens dos classificadores locais se torna um potencial vértice de um grafo. O que, conhecendo os problemas de ordem e espalhamento das nuvens (Figura 3.1(b), pág. 14), remete, por vezes, a números muito grandes de possíveis grafos. Para se ter uma noção da dimensão possível desse problema, considere a simplificação abaixo.

Sejam dados: um grafo base  $\mathcal{G}_B = (\mathcal{L}, \mathcal{E})$  de 3 vértices  $\mathcal{L}(l)$ ,  $l = 1, 2, 3$ , e uma nuvem de prováveis vértices para todo  $\mathcal{L}(l)$ . Ache um grafo  $\mathcal{G}^*$  dentre todos os possíveis grafos  $\mathcal{G}_p$ ,  $p = 1, \dots, P$ , com a mesma ordem (número de vértices), graus (arestas a partir de um vértice) e adjacências de  $\mathcal{G}_B$ , que maximize a similaridade com aquele grafo. Seja ainda  $\mathcal{N}(l)$  o conjunto contendo a nuvem de prováveis vértices correspondentes de  $\mathcal{L}(l)$ , da forma:

$$\begin{aligned}\mathcal{N}(1) &= \{\nu_1(1); \nu_2(1); \nu_3(1); \nu_4(1)\}, \\ \mathcal{N}(2) &= \{\nu_1(2); \nu_2(2); \nu_3(2)\}, \\ \mathcal{N}(3) &= \{\nu_1(3); \nu_2(3)\},\end{aligned}$$

e seja  $V$  um subconjunto unitário de  $\mathcal{N}(l)$ , forme os  $P$  grafos  $\mathcal{G}_p = (V, E)$  possíveis e diferentes entre si,  $V = \{\nu_i(1), \nu_j(2), \nu_k(3)\}$ .

É possível perceber que  $P$  cresce independente do número de arestas que compõem o grafo em prova, e que é dado pela Equação 4.7:

$$P = \prod_{l=1}^3 |\mathcal{N}(l)|. \quad (4.7)$$

---

<sup>4</sup>Lembramos que a cada posição  $\mathbf{x}$  de um vértice, os jatos extraídos tendem a ser diferentes principalmente em fase [3].

Voltando ao problema original, suponha uma cardinalidade média pequena para as nuvens  $|\mathcal{N}| = 2$ . Haja vista que estamos tratando de 11 PFFs, a quantidade de possíveis grafos é então de  $2^{11}$ . Na prática, as ordens de possíveis formações de grafos chegam a  $10^{30}$  em algumas imagens de face em prova.

Para atacar esse problema notoriamente custoso do ponto de vista computacional, propomos uma adaptação do estimador da estrutura de grafos de Würtz [15] (Seção 4.1.2, pág. 49) aplicada às nuvens do DPI.

### 4.2.1 Método

Seja dada a saída do processamento de uma imagem de face em prova conforme Seção 3.1, pág. 16. Buscamos o conjunto de coordenadas das nuvens  $\mathcal{N}(l)$ ,  $\{\nu_i(l)\}$ , com  $l = 0, \dots, 10$  e  $i = 1, \dots, N_l$ , que sejam vértices de um grafo  $\mathcal{G}^*$  que possua máxima verossimilhança em relação a um *agregado de grafos de faces*<sup>5</sup>  $\mathcal{A} = \{\mathcal{G}^{\mathcal{A}_m}\}$  a priori, o qual definiremos mais adiante.

Como métrica de similaridade, usamos uma adaptação do estimador de Würtz dada pela Equação 4.8:

$$S_{\Delta}^{\mathcal{G}}(\mathcal{L}(l), \mathcal{A}) = -\frac{1}{2Md_{\mathcal{G}}(\mathcal{L}(l))} \sum_{m=1}^M \sum_{\substack{l_e \\ \text{adjacente a} \\ l}} \left[ \frac{(\mathcal{L}_x(l) - \mathcal{L}_x^{\mathcal{A}_m}(l_e) - \mu_x(l, l_e))^2}{\sigma_x(l, l_e)^2} + \frac{(\mathcal{L}_y(l) - \mathcal{L}_y^{\mathcal{A}_m}(l_e) - \mu_y(l, l_e))^2}{\sigma_y(l, l_e)^2} \right] \quad (4.8)$$

onde as adjacências do candidato  $\mathcal{L}(l)$  são oriundas de  $M$  grafos do agregado  $\mathcal{A}$ . A Equação 4.8 ecoará seu valor máximo quando um candidato a  $\mathcal{L}(l)$  guardar maior proximidade em relação a suas adjacências em  $M$  grafos do agregado  $\mathcal{A}$  de acordo com a tendência contida no conjunto de treinamento. Então, diferentemente do EBGm, não estamos comparando adjacências em um grafo elástico, mas em um agrupamento de grafos modelos de treinamento.

Dessa forma, cada ponto fiducial  $\mathcal{L}(l)$  terá seu candidato  $\nu^*(l)$  da nuvem  $\mathcal{N}(l)$  eleito se forem satisfeitas as condições:

- $\nu^*(l)$  é um vértice de um grafo  $\mathcal{G}$ ;
- $\nu^*(l)$  é adjacente a outros vértices  $\mathcal{L}(l_e)$ ,  $e = 1, \dots, d_{\mathcal{G}}(\mathcal{L}(l))$ ;

<sup>5</sup>Usaremos o superescrito  $\mathcal{A}$  para indicar que um grafo pertence a um agregado de grafos de faces, ou simplesmente *agregado*, nesta dissertação.

- $\nu^*(l)$  maximiza a correspondência de sua posição com relação a suas adjacências no agregado  $\mathcal{A}$  conforme Equação 4.9:

$$\nu^*(l) = \arg \max_{\nu_i(l) \in \mathcal{N}(l)} \{S_{\Delta}^{\mathcal{G}}(\nu_i(l), \mathcal{A})\} \quad (4.9)$$

onde  $S_{\Delta}^{\mathcal{G}}$  é definida na Equação 4.8.

Para isso, definimos as adjacências dos *landmarks* ao propor um grafo  $\mathcal{G}$  a partir das marcações de Sabino e Araujo, como veremos a seguir.

## 4.2.2 Agregado de Grafos de Faces

Antes de estruturarmos os procedimentos de treino e aplicação método, é preciso definir o objeto a ser comparado e manipulado ao longo do processo de treinamento e aplicação: o grafo  $\mathcal{G}$ .

É possível verificar na literatura em Wiskott [3] e Würtz [15], bem como na implementação disponível do EBGm de Bolme [42], que os grafos utilizados são, via de regra, planares. Isto é, não há pontos de concorrência entre arestas do grafo além de seus vértices. Partimos deste princípio ao propor um grafo a ser utilizado no estimador definido a seguir.

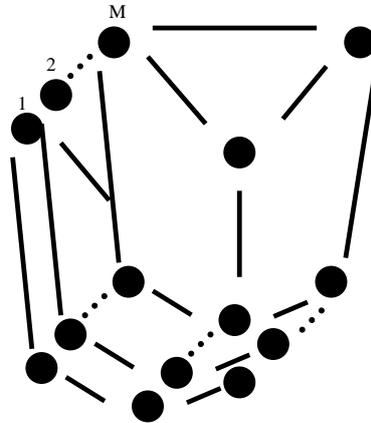


Figura 4.2: Visão simplificada de um *agregado de grafos de faces* com  $M$  elementos. Diferentemente dos *face bunch graphs* da Figura 4.1, empilhamos somente as informações qualitativas dos grafos em uma estrutura de dados. Além disso, não há mais rótulos nos vértices dos grafos com as informações quantitativas dos *landmarks*: os jatos.

Neste trabalho, convertemos as marcações de Sabino e Araujo dos 11 PFFs nas bases BioID e FERET em um *agregado de grafos de faces*  $\mathcal{A}$  (vide Figura 4.2). Esse *agregado* difere dos *face bunch graphs* propostos em [3] por armazenar somente as

informações qualitativas<sup>6</sup> dos grafos modelos. Com isso, não levamos em consideração os jatos do EBGM (Seção 4.1.1), que na proposta original rotulavam os vértices do *bunch*, e o grafo médio, que dava forma ao *bunch*.

Os 11 PFFs marcados nas imagens de faces se tornaram vértices. E entre estes, 16 arestas são conectadas de modo a minimizar a concentração de conexões em um único *landmark* (vide Figura 4.3). Esta topologia contribui na distribuição do tempo de processamento em cada PFF de forma mais uniforme no procedimento de avaliação (teste).

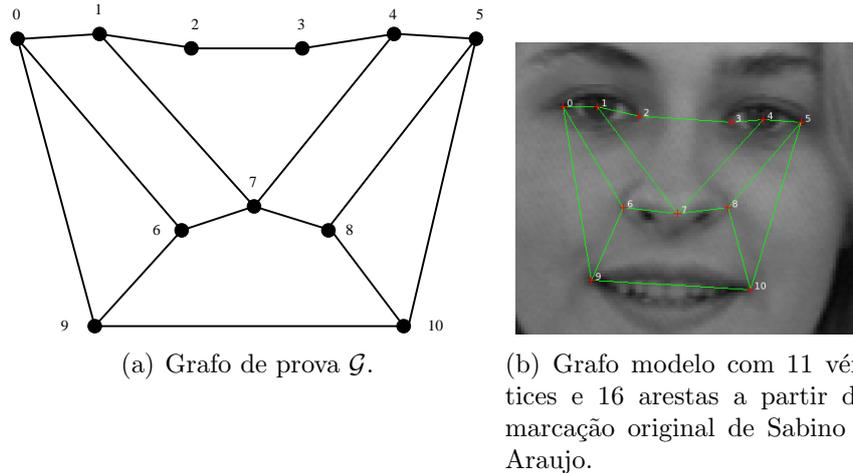


Figura 4.3: Grafo definido a partir das Marcações de Sabino e Araujo.

### 4.2.3 Treinamento do DPI-GRAPH

O agregado de grafos de faces construído a partir das marcações de Sabino/Araujo contém as informações de interesse para o treinamento do estimador da estrutura do grafo: as coordenadas dos PFFs  $\mathcal{L}(l)$ , agora vértices de um grafo, e as arestas  $\mathcal{E}$ , que apontam as adjacências no grafo.

Com isso, é possível calcular os vetores distância entre vértices adjacentes  $\Delta(l, l_e) = \mathcal{L}(l) - \mathcal{L}(l_e)$  nos grafos modelos do conjunto de treinamento. A partir das componentes horizontal e vertical desse vetor, são extraídas as médias e variâncias (Equações 4.10a, 4.10b, 4.10c e 4.10d abaixo) em  $x$  e  $y$ . Assim, obtemos informação de alto nível sobre as tendências e correspondências espaciais entre os

<sup>6</sup>Novamente, as informações qualitativas de grafos dizem respeito somente a sua estrutura e forma.

PFFs por:

$$\mu_x(l, l_e) = \frac{1}{M} \sum_{m=1}^M \Delta_x^{(m)}(l, l_e) \quad (4.10a)$$

$$\mu_y(l, l_e) = \frac{1}{M} \sum_{m=1}^M \Delta_y^{(m)}(l, l_e) \quad (4.10b)$$

$$\sigma_x^2(l, l_e) = \frac{1}{M-1} \sum_{m=1}^M (\Delta_x^{(m)}(l, l_e) - \mu_x(l, l_e))^2 \quad (4.10c)$$

$$\sigma_y^2(l, l_e) = \frac{1}{M-1} \sum_{m=1}^M (\Delta_y^{(m)}(l, l_e) - \mu_y(l, l_e))^2 \quad (4.10d)$$

com  $M$  sendo o número de grafos de um dado agregado  $\mathcal{A} = \{\mathcal{G}^{\mathcal{A}_m}\}$ ;  $\Delta_x(l, l_e)$ , a componente  $x$  do vetor distância entre os vértices adjacentes  $\mathcal{L}(l)$  e  $\mathcal{L}(l_e)$ ;  $\Delta_y(l, l_e)$ , a componente  $y$  desse vetor distância;  $\mu_x(l, l_e)$  e  $\mu_y(l, l_e)$ , as médias em  $x$  e  $y$ , respectivamente, dos vetores distância entre vértices adjacentes; e  $\sigma_x^2(l, l_e)$  e  $\sigma_y^2(l, l_e)$ , suas variâncias em  $x$  e  $y$  respectivamente.

De posse dessas informações, é possível resumir o procedimento de treinamento do estimador da estrutura do grafo no Algoritmo 3.

**Entrada:** Agregado  $\mathcal{A} = \{\mathcal{G}_k^{\mathcal{A}_m}\}_{k \neq i}$  contendo os 11 pontos fiduciais faciais nos  $m = 1, \dots, M$  grafos dos  $k - 1$  *folds* de treino dado um *fold* de teste  $i$ , e dimensões das  $P = M$  imagens associadas a cada grafo.

**1 para cada grafo  $\mathcal{G}_k^{\mathcal{A}_m}$  com  $m = 1$  até  $M$  faça**

**2** Calcule os vetores distância entre todos os vértices adjacentes:

$$\left( \Delta_x^{(m)}(l, l_e), \Delta_y^{(m)}(l, l_e) \right)$$

**3** Acumule as coordenadas em  $x$  e  $y$  de cada vértice para a construção de um grafo médio  $G_i^{\mathcal{A}}$ .

**4** Acumule a dimensão da imagem.

**5 fim para cada**

**6** Extraia as médias em  $x$  e  $y$  dos vetores distância entre vértices adjacentes de a partir das Equações 4.10a e 4.10b respectivamente.

**7** Extraia as variâncias em  $x$  e  $y$  dos vetores distância entre vértices adjacentes do grafo a partir das Equações 4.10c e 4.10d respectivamente.

**8** Salve o grafo médio  $G_i^{\mathcal{A}}$  a partir das coordenadas acumuladas em  $(x, y)$  e a dimensão média das  $P$  imagens.

**Saída:** Médias e variâncias em  $x$ ,  $\mu_x(l, l_e)$  e  $\sigma_x^2(l, l_e)$ , e  $y$ ,  $\mu_y(l, l_e)$  e  $\sigma_y^2(l, l_e)$ , o grafo médio  $G_i^{\mathcal{A}}$  para o *fold* de teste  $i$  e a dimensão média das imagens para esse *fold*.

**Algoritmo 3:** Treinamento do Método DPI-GRAPH.

#### 4.2.4 Teste

Este método é proposto sobre as saídas do processamento de uma imagem em prova pelos classificadores DPI-PCA de 11 PFFs já conhecidos. Conforme a Seção 2.2.4,

essa saída aponta coordenadas em um plano de dimensão 220x220 pixels. Com isso, é possível realizar a localização global dos pontos fiduciais levando em consideração suas correspondências na face através de um grafo.

Resumidamente, o procedimento de teste do método DPI-GRAPH é um processo iterativo que visa à construção de um grafo  $\mathcal{G}^*$  a partir das nuvens dos classificadores locais para a seleção conjunta dos candidatos a PFFs nas imagens de face em prova. Cada vértice do grafo  $\mathcal{G}^*$  é eleito observando o mínimo deslocamento em sua tendência relativa a suas adjacências. A quantificação dessas tendências de correspondências entre pontos fiduciais que são agora vértices de um grafo é executada quando do treinamento do método, conforme a seção anterior.

Iniciamos a aplicação do DPI-GRAPH limitando a quantidade de grafos modelos tomados do agregado. A escolha desse subconjunto de grafos é feita de forma pseudo-aleatória. Um número de grafos  $M = 60$  foi considerado um bom compromisso entre tempo de processamento e desempenho.

Para contornar as situações onde os classificadores locais não ecoam nenhuma coordenada da imagem de prova em suas saídas,  $\mathcal{N}(l) = \emptyset$ , selecionamos o vértice correspondente a  $l$  do grafo médio  $G^A$ . Além de escolher, de forma pseudo-aleatória, outras coordenadas em uma área limitada no interior da ROI (Seção 2.2.4, pág. 10) de  $\mathcal{L}(l)$  para o preenchimento daquela nuvem.

Embora tenhamos concebido esta proposta de busca global, o método é suficientemente flexível para registrar de 1 a  $n$  PFFs, com  $0 < n \leq 11$ . Para isso, os PFFs a serem detectados devem formar um subgrafo induzido (Figura 4.4) do grafo de prova  $\mathcal{G}$ . Assim, de posse de todas as nuvens  $\mathcal{N}(l)$  dos PFFs  $\mathcal{L}(l)$ ,  $l = 0, \dots, 10$ , começamos por buscar um grafo  $\mathcal{G}^*$ , ou um subgrafo induzido  $G^*$ , que contenha todos os seus vértices selecionados através do critério de máxima verossimilhança da Equação 4.9 (pág. 52).

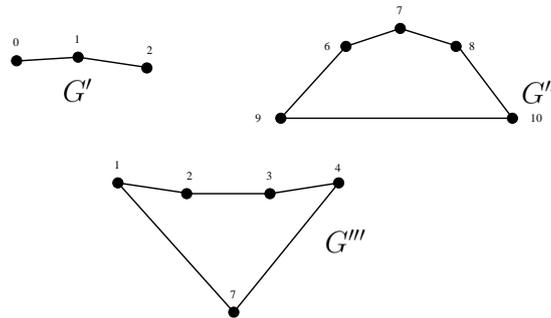


Figura 4.4: Subgrafos induzidos de  $\mathcal{G}$  da Figura 4.3.  $G'$  registra os PFFs  $\mathcal{L}(0)$ ,  $\mathcal{L}(1)$  e  $\mathcal{L}(2)$ ;  $G''$ , de  $\mathcal{L}(6)$  a  $\mathcal{L}(10)$ ; e  $G'''$ ,  $\mathcal{L}(1)$ ,  $\mathcal{L}(2)$ ,  $\mathcal{L}(3)$ ,  $\mathcal{L}(4)$  e  $\mathcal{L}(7)$ .

Inicialmente, cada candidato  $\nu_i(l)$ ,  $i = 1, \dots, N_l$ , de cada nuvem  $\mathcal{N}(l)$  é utilizado para o cálculo dos vetores distância  $\Delta^{(i,m)}(l, l_e) = \nu_i(l) - \mathcal{L}^{A_m}(l_e)$  entre suas

adjacências tomadas de  $m = 1, \dots, M$  grafos do agregado. A partir de uma nuvem inicial  $\mathcal{N}(j)$ , por exemplo  $\mathcal{N}(0)$ , cada candidato  $\nu_i(0)$  é submetido à avaliação de correspondência entre adjacências no agregado pela Equação 4.8, fazendo uso dos vetores distância pré-calculados. O candidato  $\mathcal{L}(0) = \nu^*(0)$  que obtiver o escore de máxima verossimilhança da Equação 4.9 é selecionado e adicionado ao grafo  $\mathcal{G}^*$ . Para as próximas iterações, se  $\mathcal{L}(0)$  for uma adjacência  $\mathcal{L}(l_e)$ , os vetores distância  $\Delta^{(i,m)}(l, 0)$  são atualizados para utilizarem a coordenada  $\nu^*(0)$  do grafo  $\mathcal{G}^*$  e não mais as coordenadas nesse vértice oriundas dos grafos do agregado. Isto fará com que, gradativamente, os vértices do grafo comecem a ser avaliados localmente, como na formulação original de Würtz (Equação 4.6, pág. 49). Isso é feito até que se tenha, finalmente, um grafo  $\mathcal{G}^*$  (ou um subgrafo  $G^*$ ) ótimo sob o ponto de vista de máxima verossimilhança de correspondência entre seus vértices.

Este procedimento de teste está detalhado no Algoritmo 4. O diagrama de blocos do sistema de detecção global de pontos fiduciais via DPI-PCA e casamento de grafos DPI-GRAPH pode ser visualizado na Figura 4.5.

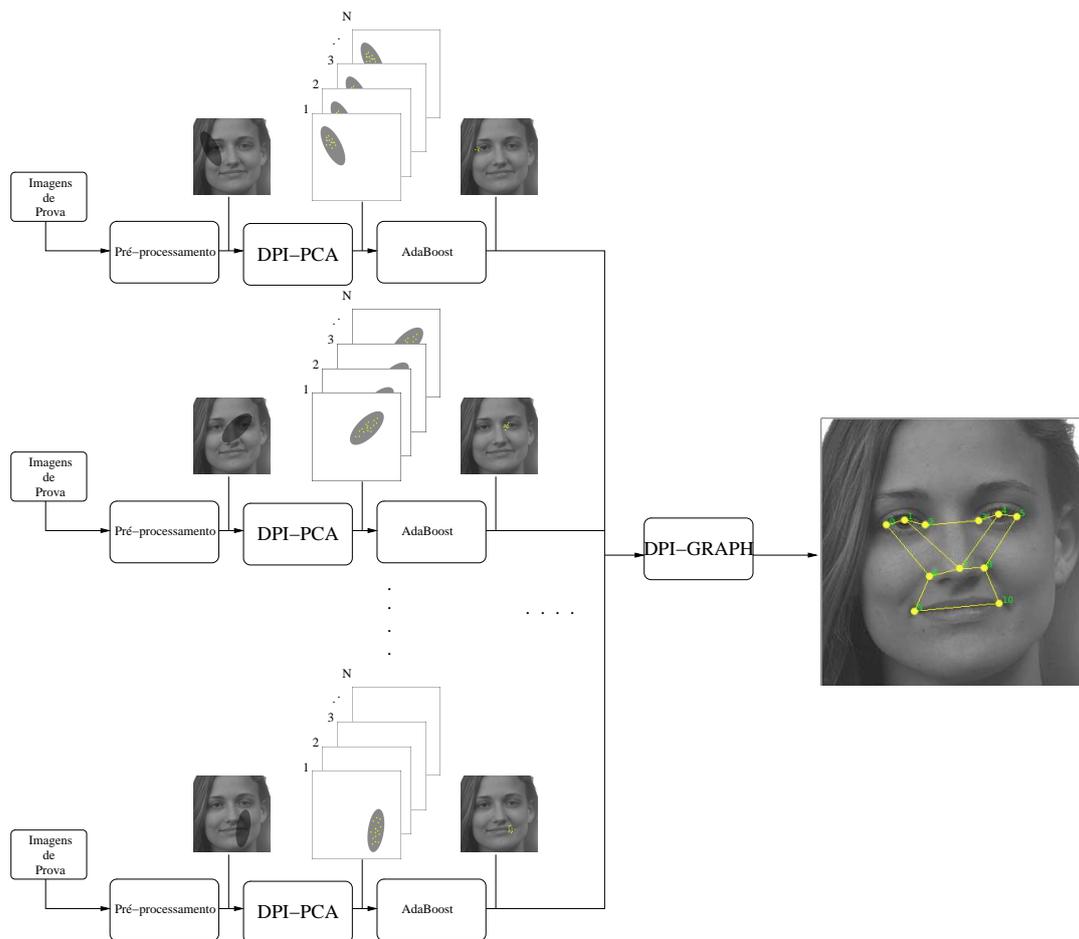


Figura 4.5: Diagrama de blocos do sistema de teste para detecção de pontos fiduciais faciais pelo método DPI-GRAPH. Diferentemente do método representado na Figura 2.4, aqui se buscam todos os 11 PFFs definidos de forma global.

**Entrada:** Grafo  $\mathcal{G}$  de prova; imagem de prova  $I_k^{(p)}$ , de  $p = 1, \dots, P^{(k)}$  imagens do fold  $k = 1, \dots, 7$ ; agregado  $\mathcal{A} = \{\mathcal{G}_k^{A_m}\}_{k \neq i, m = 1, \dots, M}$ , dado um *fold* de teste  $i$ ; grafo médio  $G_i^A$ ; e a dimensão média das imagens dos  $k - 1$  *folds* de treino.

- 1 Traduza as coordenadas das nuvens de DPI e de  $G_i^A$  para a dimensão da imagem de prova.
- 2 Eleja e copie  $N$  dos  $M$  grafos do agregado  $\mathcal{A}$  de forma pseudo-aleatória:  $\mathcal{A}' = \{\mathcal{G}^{A_n}\}$ .  
% Inicialização
- 3 **para cada** *landmark*  $\mathcal{L}(l)$ ,  $l = 0, \dots, 10$  **faça**
- 4     **se**  $\mathcal{L}(l)$  *for um vértice válido do grafo de prova*  $\mathcal{G}$  **então**
- 5         **se**  $\mathcal{N}(l) = \emptyset$  **então**
- 6             Construa uma nuvem utilizando como ponto de partida a posição de  $\mathcal{L}(l)$  extraída do grafo médio  $G_i^A$ .
- 7             Traduza essa coordenada do grafo médio  $G_i^A$  para a dimensão da imagem de prova.
- 8             Obtenha de forma pseudo-aleatória coordenadas extras em torno desse ponto, dentro de uma área limitada no interior da ROI de  $\mathcal{L}(l)$ , conforme Seção 2.2.4, pág. 10.
- 9         **fim se**
- 10        **para cada**  $\nu_j(l)$ ,  $j = 1, \dots, N_l$  **faça**
- 11            **para cada** grafo  $\mathcal{G}^{A_n}$ ,  $n = 1, \dots, N$  **do agregado**  $\mathcal{A}'$  **faça**
- 12                Extraia os vetores distância  $(\Delta_x^{(j,n)}(l, l_e), \Delta_y^{(j,n)}(l, l_e))$  de  $\mathcal{L}(l)$  com relação a suas adjacências  $\mathcal{L}^{A_n}(l_e)$ ,  $e = 1, \dots, d_G(\mathcal{L}(l))$ , no grafo  $\mathcal{G}^{A_n}$ .
- 13            **fim para cada**
- 14        **fim para cada**
- 15        **fim se**
- 16 **fim para cada**
- 17 **para cada** *landmark*  $\mathcal{L}(l)$ ,  $l = 0, \dots, 10$  **faça**
- 18     **se**  $\mathcal{L}(l)$  *for um vértice válido do grafo de prova*  $\mathcal{G}$  **então**
- 19        **para cada**  $\nu_j(l)$ ,  $j = 1, \dots, N_l$  **faça**
- 20            Avalie a função de similaridade  $S_{\Delta}^G(\nu_j(l), \mathcal{A}')$  conforme Equação 4.8.
- 21            Conserve o elemento de  $\mathcal{N}(l)$  de máxima verossimilhança com relação a  $\mathcal{A}'$  de acordo conforme Equação 4.9:  $\mathcal{L}(l) = \nu^*(l)$ .
- 22        **fim para cada**
- 23        **para cada** *landmark*  $\mathcal{L}(j)$ ,  $j = 0, \dots, 10$  e  $j \neq l$  **faça**
- 24            **se**  $\mathcal{L}(j)$  *for um vértice válido do grafo de prova*  $\mathcal{G}$  **então**
- 25                **se**  $\mathcal{L}(l)$  *é uma adjacência de*  $\mathcal{L}(j)$  **então**
- 26                    Atualize os vetores distância para cada nuvem de *landmark* onde  $\nu^*(l)$  é um vértice adjacente:  $(\Delta_x(j, l), \Delta_y(j, l))$ . Faça isso alterando na cópia dos  $N$  grafos em  $\mathcal{A}'$  os vértices  $\mathcal{L}^{A_n}(l)$  para  $\nu^*(l)$ .
- 27                **fim se**
- 28            **fim se**
- 29        **fim para cada**
- 30        **fim se**
- 31 **fim para cada**

**Saída:** Os 11 candidatos eleitos dos pontos fiduciais  $\mathcal{L}(l)$  e o grafo  $\mathcal{G}^*$  formado a partir deles.

**Algoritmo 4:** Aplicação do Método DPI-GRAPH.

## 4.2.5 Análise dos Resultados

Assim como na Seção 3.4, pág. 26, a métrica adotada para validação de um PFF eleito na aplicação do método DPI-GRAPH é a distância interocular.

## 4.2.6 Resultados para a base BioID

Sintetizamos os desempenhos do DPI-GRAPH para a base de imagens de face BioID nos gráficos das Figuras 4.6, 4.7, 4.8, 4.9, 4.10 e 4.11.

Os mesmos comentários iniciais da Seção 3.4.2 valem aqui. E lembramos que o ponto de corte verifica-se no erro relativo à distância interocular ( $d_{pf}$ ) igual a 10% (0,1). Após análise dos gráficos, pudemos estabelecer o que segue.

### PFFs 0 e 5 - Cantos externos dos olhos

- As taxas médias de acerto para os cantos dos olhos utilizando o DPI-GRAPH ficaram em torno de 98%.
- O DPI-GRAPH melhora as taxas em ambos os PFFs em 2% em relação aos resultados prévios do DPI-PCA.

### PFFs 1 e 4 - Pupilas

- As taxas médias de acerto para as pupilas utilizando o DPI-GRAPH ficaram em torno de 98,5%.
- Novamente, o DPI-GRAPH melhora as taxas em ambos os PFFs em 1,5% em relação aos resultados prévios do DPI-PCA.
- Anteriormente, o DPI-PCA levava desvantagem em competição com o SVM NL. Agora, o DPI-GRAPH é superior no ponto de corte ao SVM com *kernel* polinomial na pupila direita e equivalente àquele na pupila esquerda.

### PFFs 2 e 3 - Cantos internos dos olhos

- As taxas médias de acerto para os cantos internos dos olhos utilizando o DPI-GRAPH ficaram em torno de 99%.
- Novamente, o DPI-GRAPH melhora as taxas em ambos os PFFs em até 3% em relação aos resultados prévios do DPI-PCA.
- No canto interno do olho esquerdo, o erro é desprezível com o DPI-GRAPH.

### **PFF 7 - Centro do Nariz**

- As taxas médias de acerto para o centro do nariz utilizando o DPI-GRAPH são equivalentes aos resultados prévios com o DPI-PCA, em torno de 90%.

### **PFFs 6 e 8 - Cantos do nariz**

- As taxas médias de acerto para os cantos do nariz utilizando o DPI-GRAPH ficaram em torno de 94% para o canto direito e 97% para o esquerdo.
- Novamente, o DPI-GRAPH melhora as taxas em ambos os PFFs em relação aos resultados prévios do DPI-PCA, tendo uma melhora significativa no canto esquerdo do nariz, PFF 8.

### **PFFs 9 e 10 - Cantos da boca**

- As taxas médias de acerto para os cantos da boca utilizando o DPI-GRAPH ficaram em torno de 84%.
- O DPI-GRAPH possui taxas levemente inferiores (lado direito) e equivalentes (lado esquerdo) em relação aos resultados prévios do DPI-PCA. Ambos métodos obtiveram desempenhos inferiores em taxa comparados ao SVM NL.

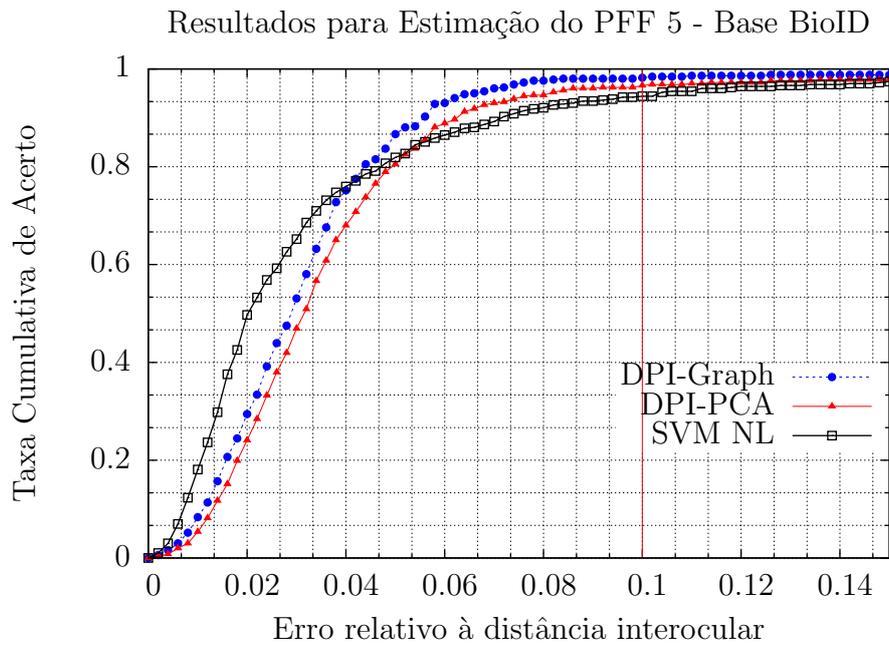
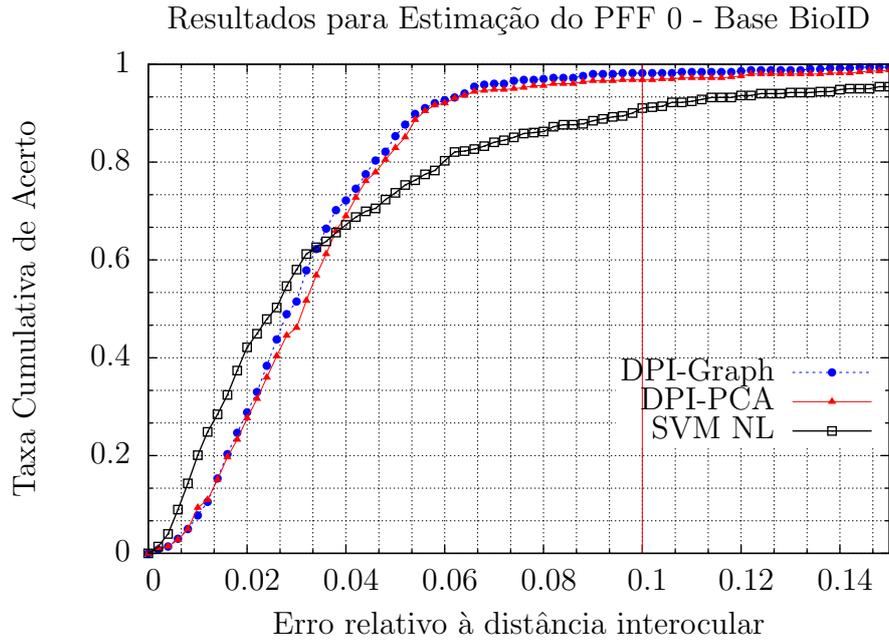


Figura 4.6: Resultados do DPI-GRAPH para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(0)$  e  $\mathcal{L}(5)$  - cantos externos dos olhos.

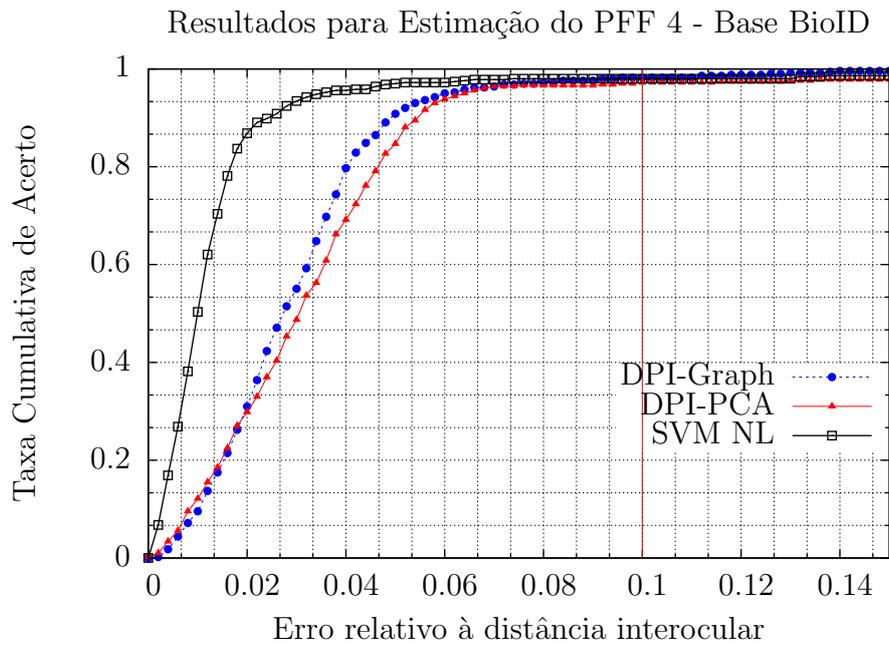
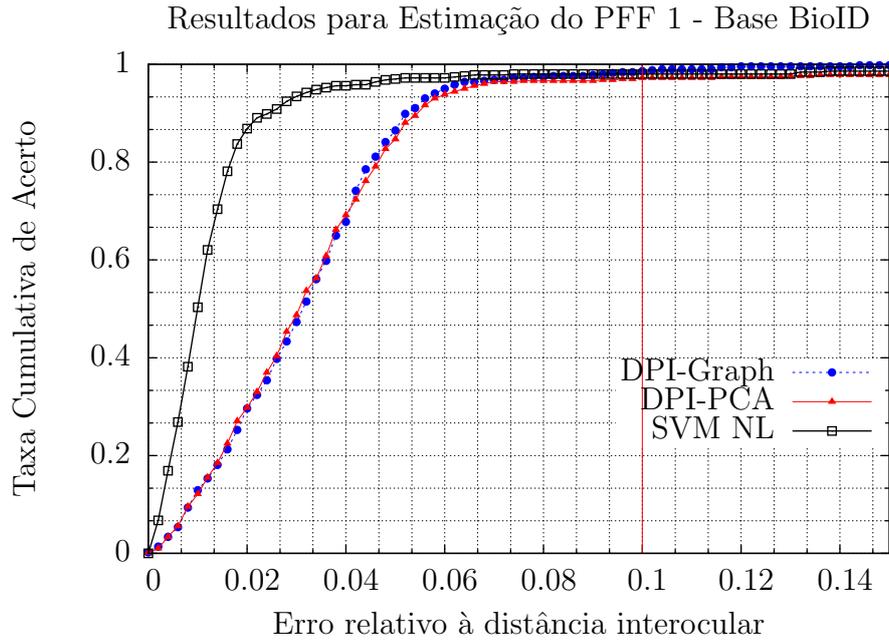


Figura 4.7: Resultados do DPI-GRAPH para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(1)$  e  $\mathcal{L}(4)$  - pupilas.

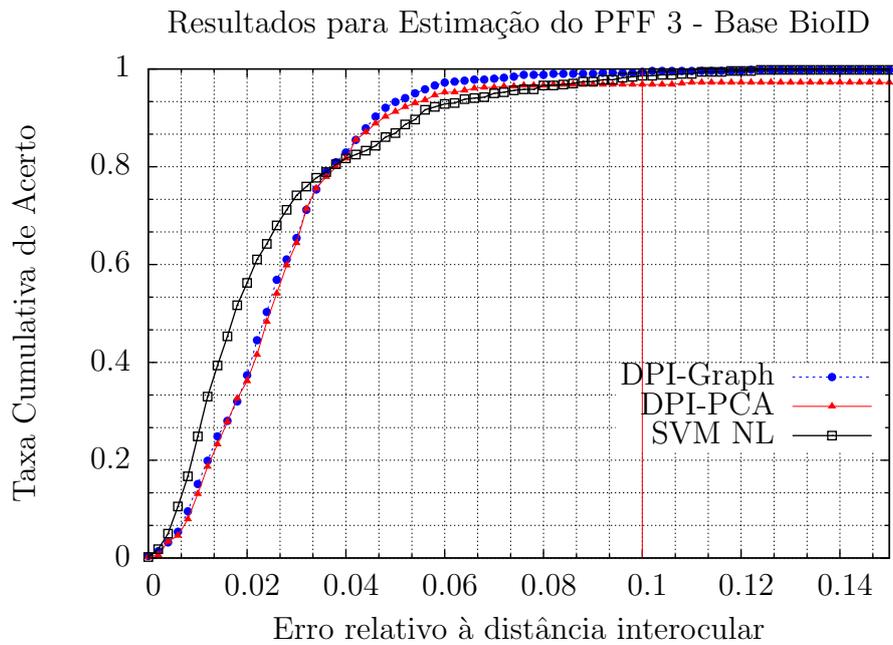
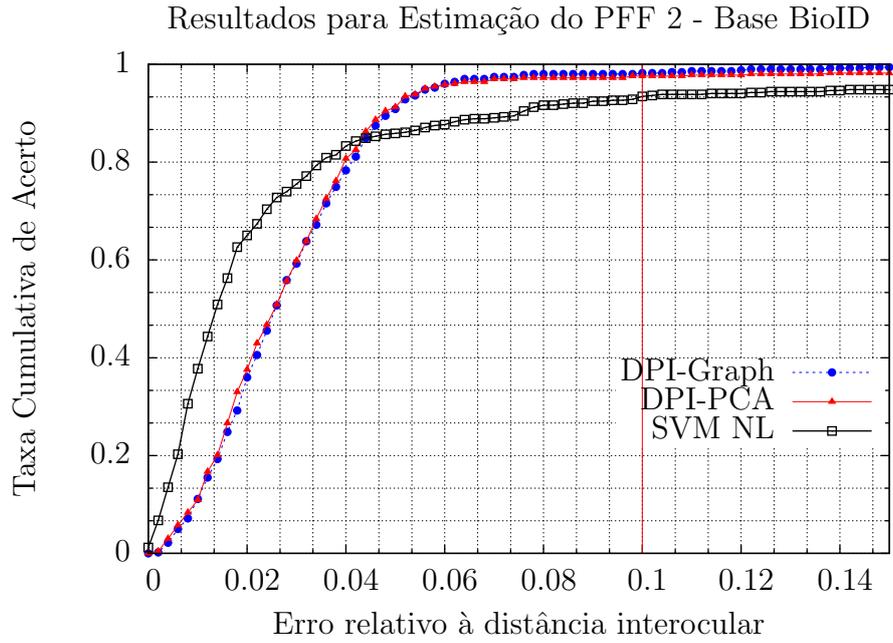


Figura 4.8: Resultados do DPI-GRAPH para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(2)$  e  $\mathcal{L}(3)$  - cantos internos dos olhos.

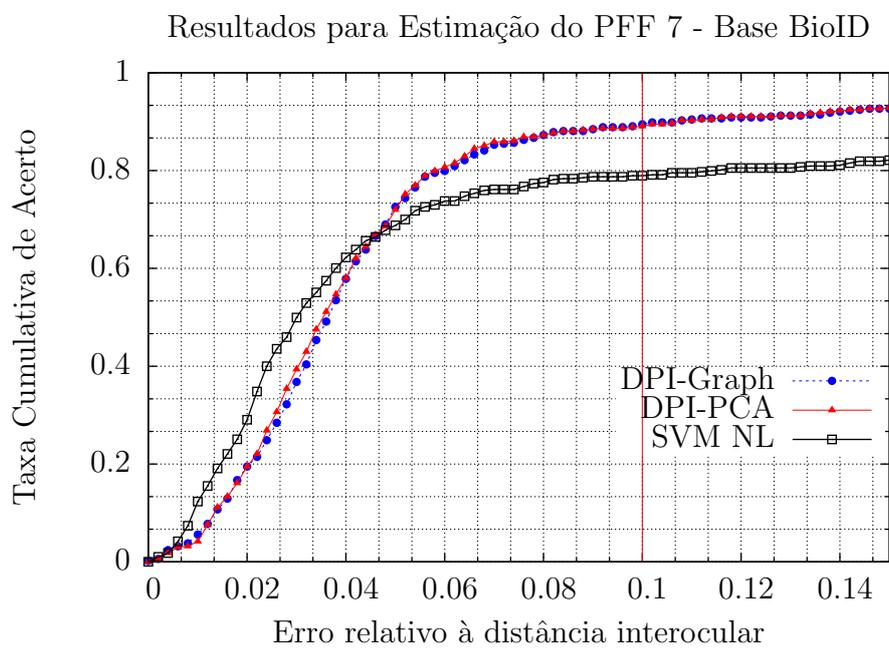


Figura 4.9: Resultados do DPI-GRAPH para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para o pontol fiducial  $\mathcal{L}(7)$  - centro do nariz.

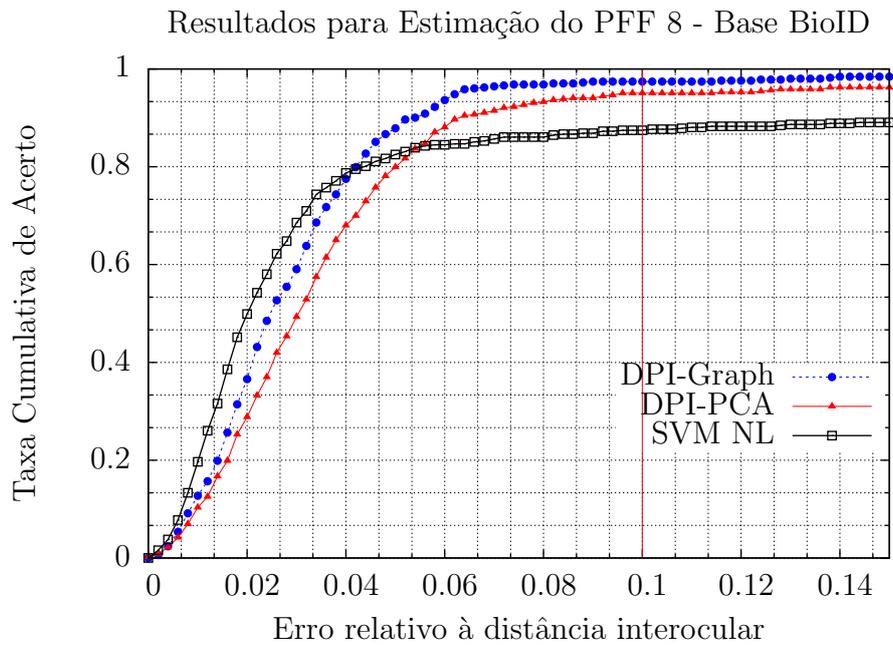
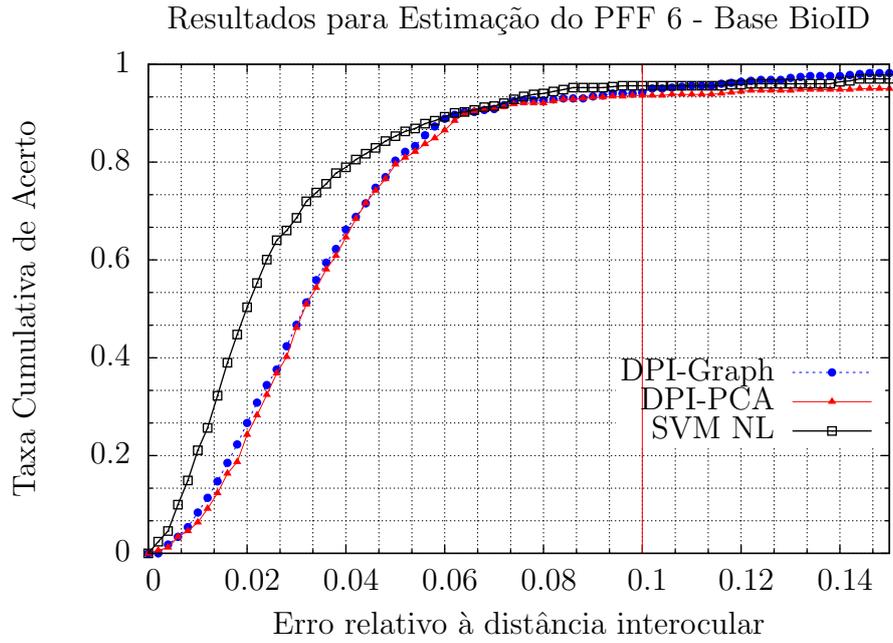


Figura 4.10: Resultados do DPI-GRAPH para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(6)$  e  $\mathcal{L}(8)$  - cantos do nariz.

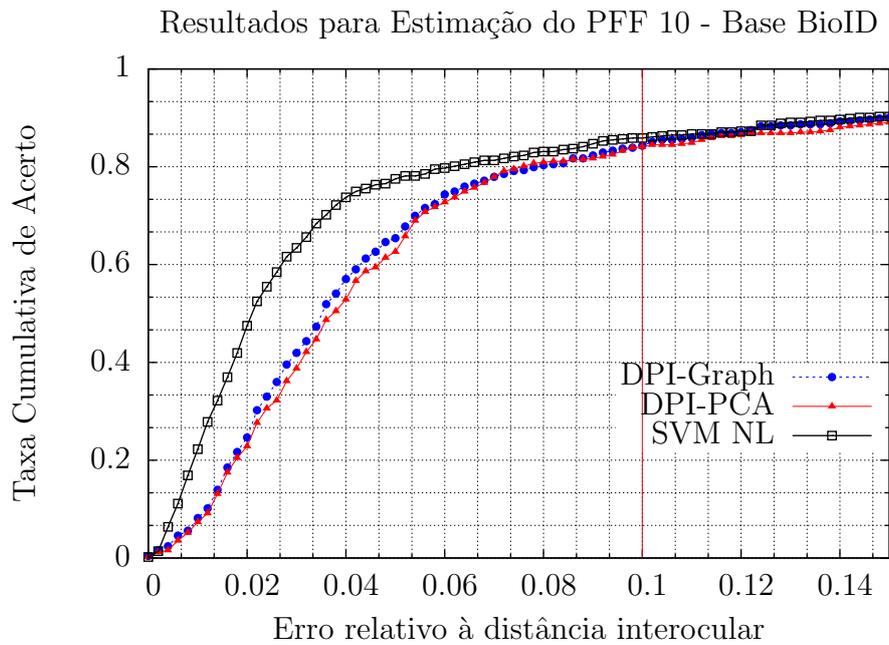
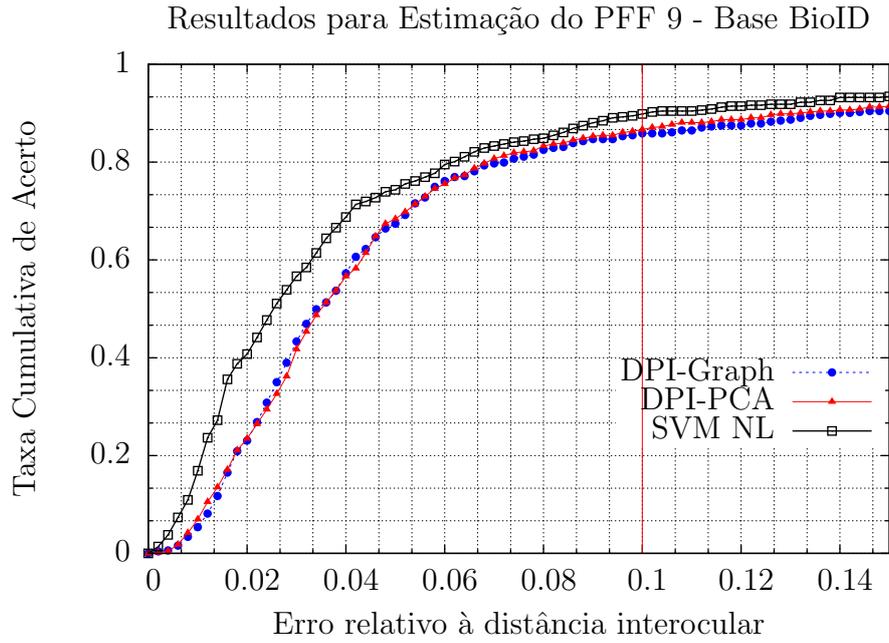


Figura 4.11: Resultados do DPI-GRAPH para a base BioID comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(9)$  e  $\mathcal{L}(10)$  - cantos da boca.

## 4.2.7 Resultados para a base FERET

Sintetizamos os desempenhos do DPI-GRAPH para a base de imagens de face FERET nos gráficos das Figuras 4.12, 4.13, 4.14, 4.15, 4.16 e 4.17.

Os mesmos comentários iniciais da Seção 3.4.2 valem aqui.

### PFFs 0 e 5 - Cantos externos dos olhos

- As taxas médias de acerto para os cantos dos olhos utilizando o DPI-GRAPH ficaram em torno de 90%.
- O DPI-GRAPH melhora as taxas em ambos os PFFs em até 2% em relação aos resultados prévios do DPI-PCA.

### PFFs 1 e 4 - Pupilas

- As taxas médias de acerto para as pupilas utilizando o DPI-GRAPH ficaram em torno de 93%.
- Novamente, o DPI-GRAPH melhora as taxas em ambos os PFFs em relação aos resultados prévios do DPI-PCA, com destaque ao ganho de aproximadamente 7% para a pupila direita.
- Apesar do SVM NL continuar obtendo melhores resultados, a melhora é significativa no ponto de corte.

### PFFs 2 e 3 - Cantos internos dos olhos

- As taxas médias de acerto para os cantos internos dos olhos utilizando o DPI-GRAPH ficaram em torno de 96% e 97% para os cantos direito e esquerdo respectivamente.
- O DPI-GRAPH possui leve desvantagem no canto interno direito, e desempenho levemente superior para o canto interno esquerdo dos olhos em relação ao DPI-PCA.
- No canto interno esquerdo, o desempenho do DPI-GRAPH é compatível com os melhores resultados do SVM NL.

### PFF 7 - Centro do Nariz

- As taxas médias de acerto para o canto do nariz utilizando o DPI-GRAPH são equivalentes aos resultados prévios com o DPI-PCA, em torno de 86%. Repetindo o que aconteceu na base BioID.

### **PFFs 6 e 8 - Cantos do nariz**

- As taxas médias de acerto para os cantos do nariz utilizando o DPI-GRAPH ficaram em torno de 92% e 93% para os cantos direito e esquerdo respectivamente.
- Assim como nos cantos internos dos olhos, o DPI-GRAPH possui leve desvantagem no canto interno direito, e desempenho levemente superior para o canto interno esquerdo do nariz em relação ao DPI-PCA.

### **PFFs 9 e 10 - Cantos da boca**

- As taxas médias de acerto para os cantos da boca utilizando o DPI-GRAPH ficaram em torno de 84%. Repetindo o desempenho para com a base BioID.
- O DPI-GRAPH possui taxas equivalentes pelo lado direito e superiores a 5% pelo lado esquerdo em relação aos resultados prévios do DPI-PCA.

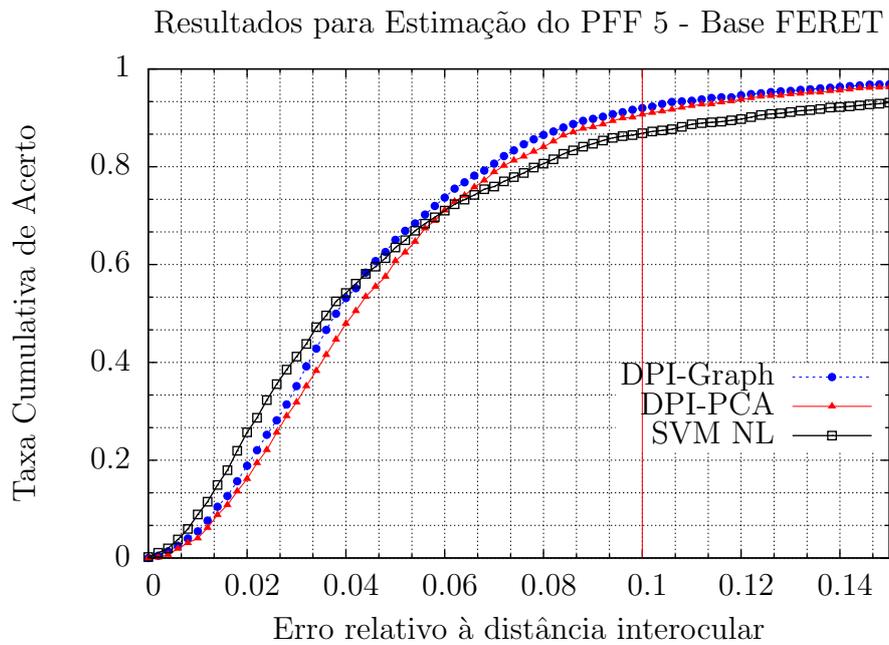
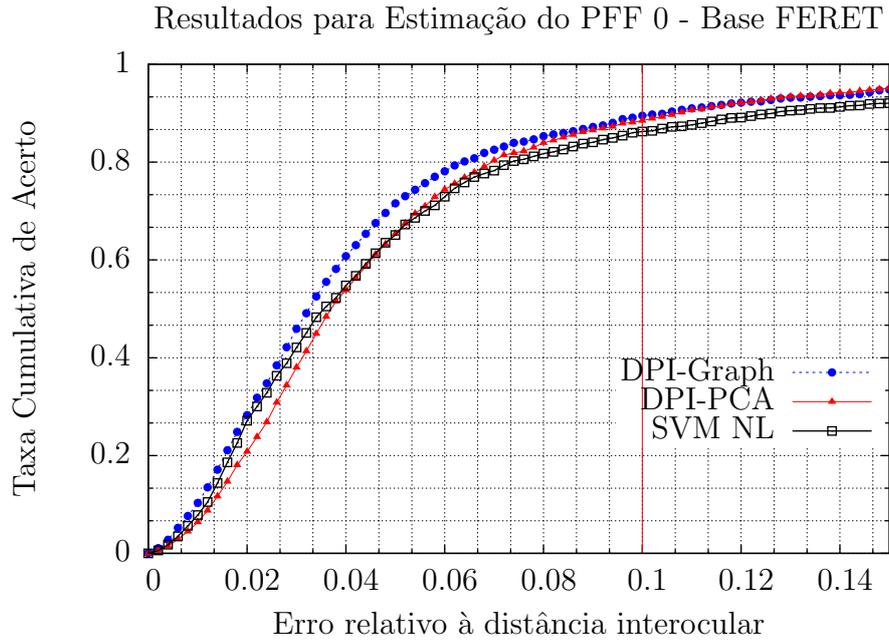


Figura 4.12: Resultados do DPI-GRAPH para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(0)$  e  $\mathcal{L}(5)$  - cantos externos dos olhos.

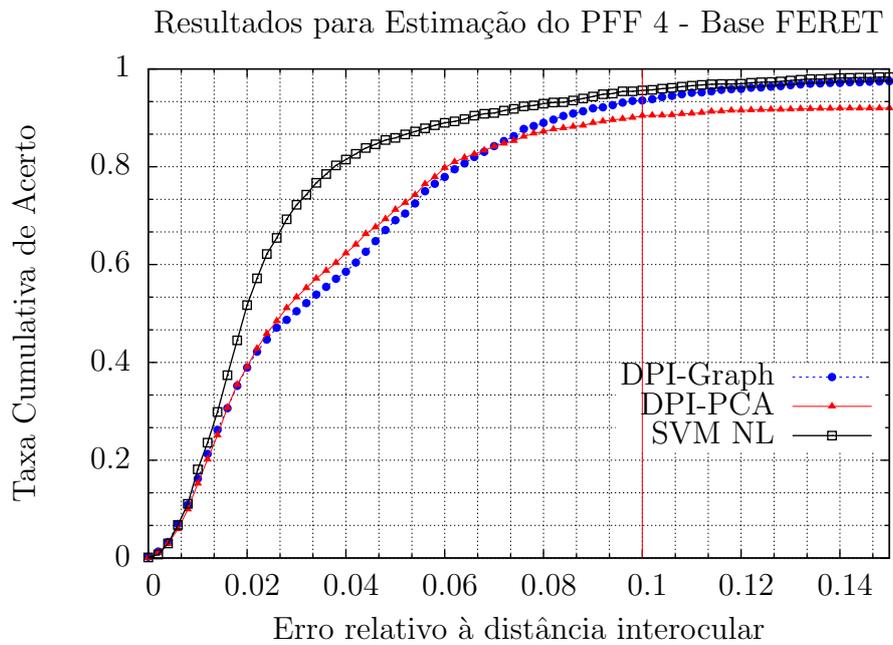
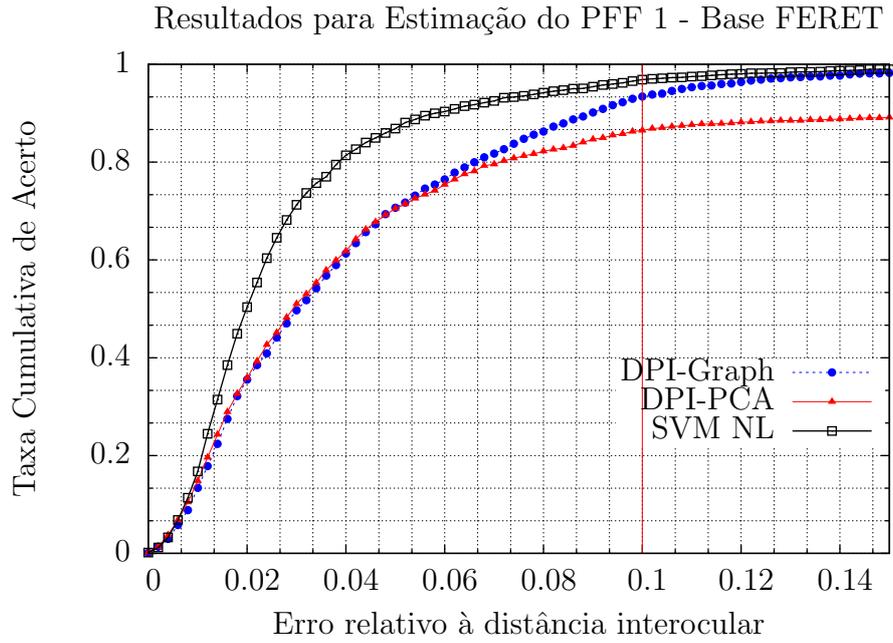


Figura 4.13: Resultados do DPI-GRAPH para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(1)$  e  $\mathcal{L}(4)$  - pupilas.

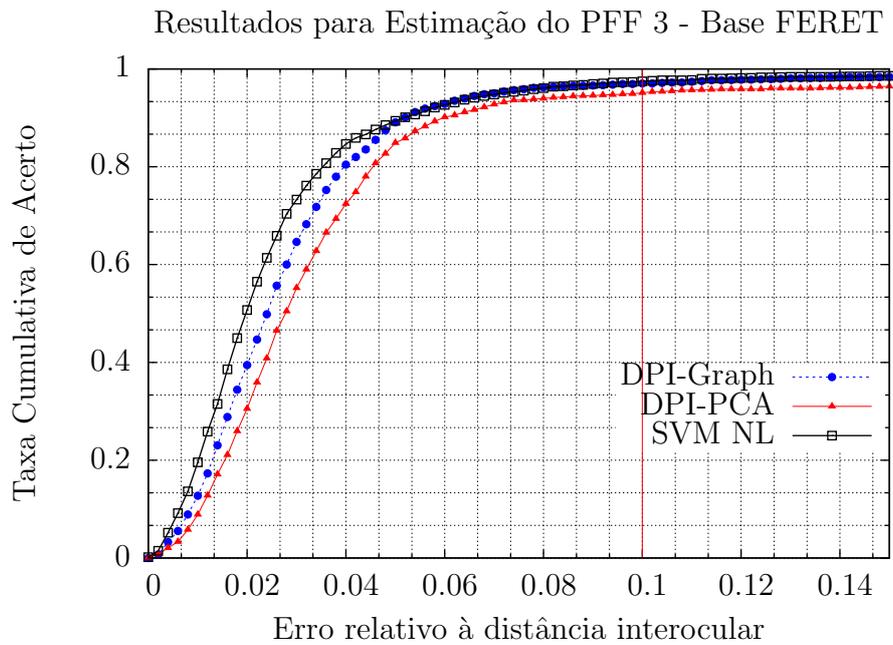
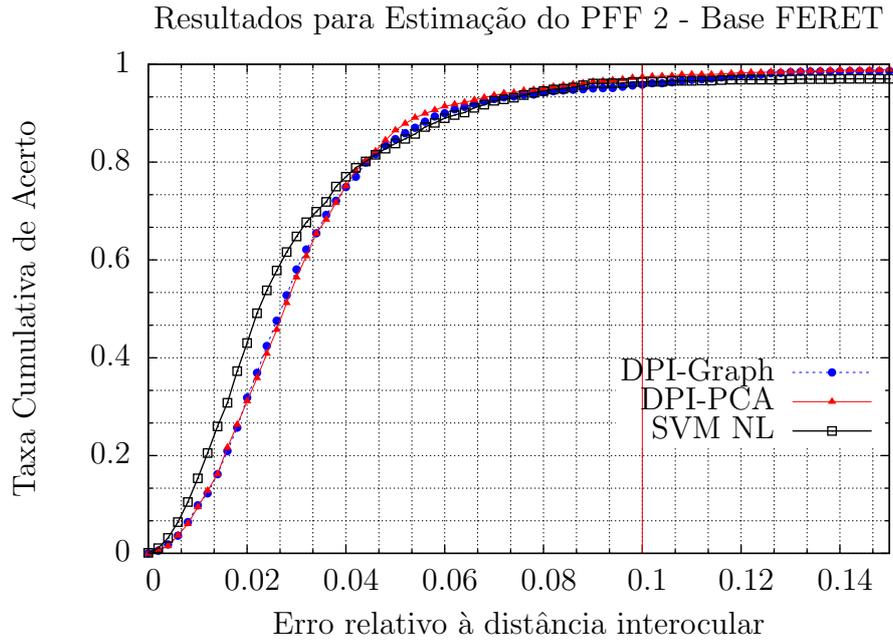


Figura 4.14: Resultados do DPI-GRAPH para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(2)$  e  $\mathcal{L}(3)$  - cantos internos dos olhos.

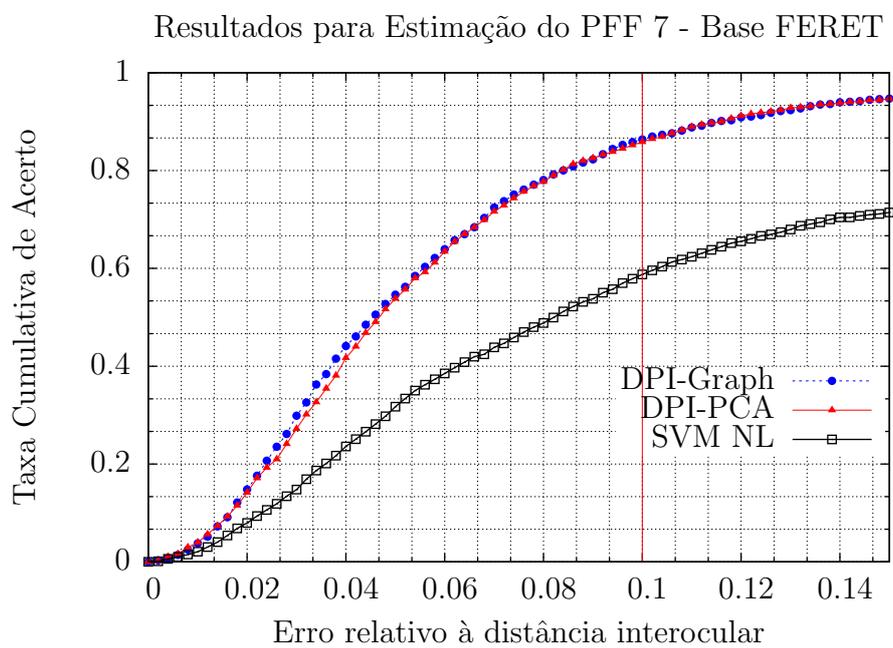


Figura 4.15: Resultados do DPI-GRAPH para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para o pontol fiducial  $\mathcal{L}(7)$  - centro do nariz.

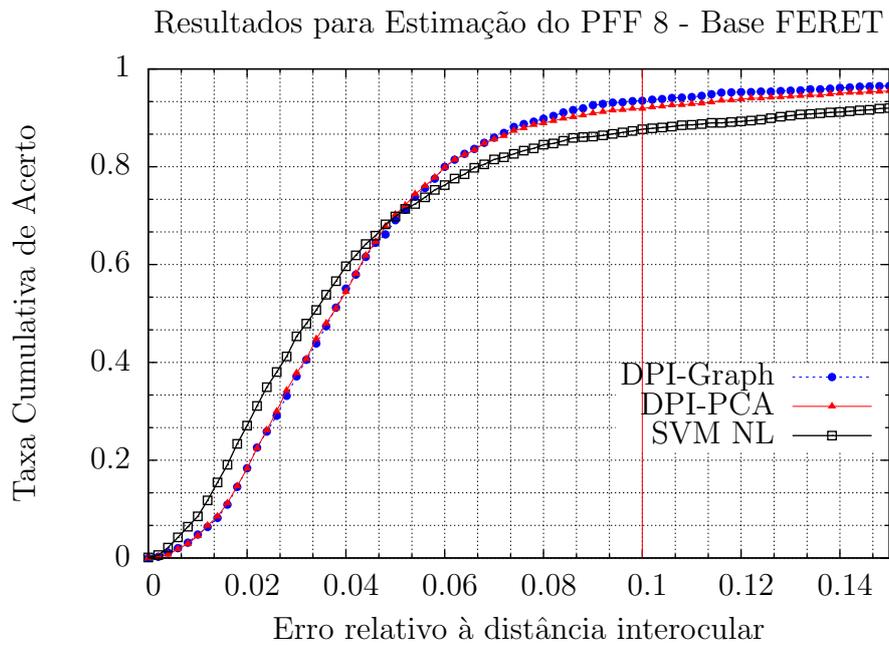
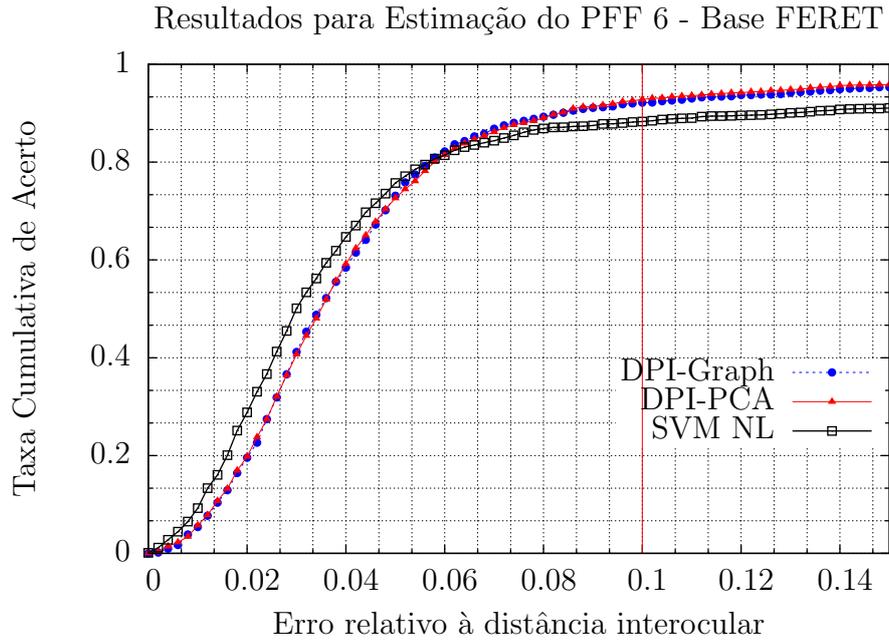


Figura 4.16: Resultados do DPI-GRAPH para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(6)$  e  $\mathcal{L}(8)$  - cantos do nariz.

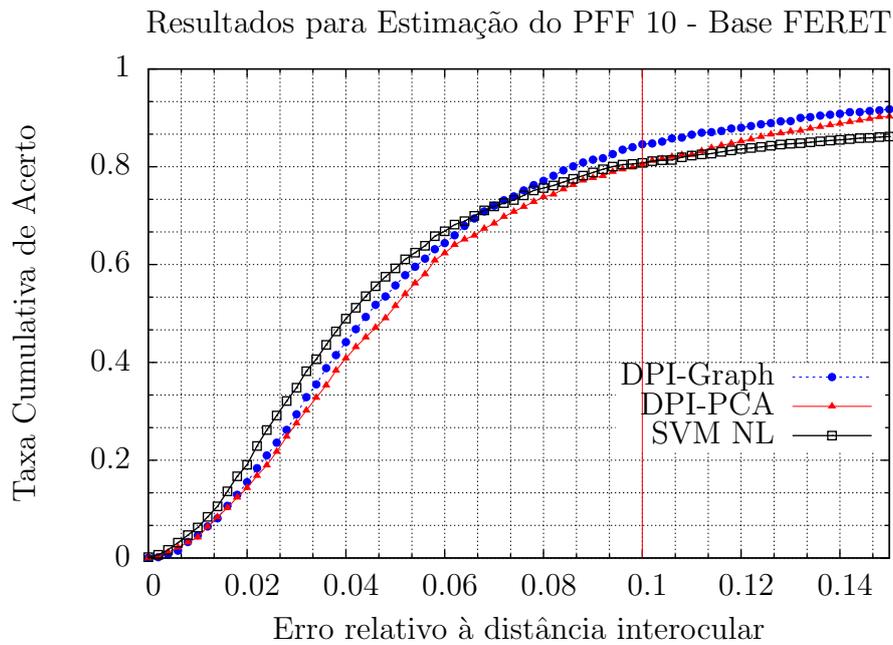
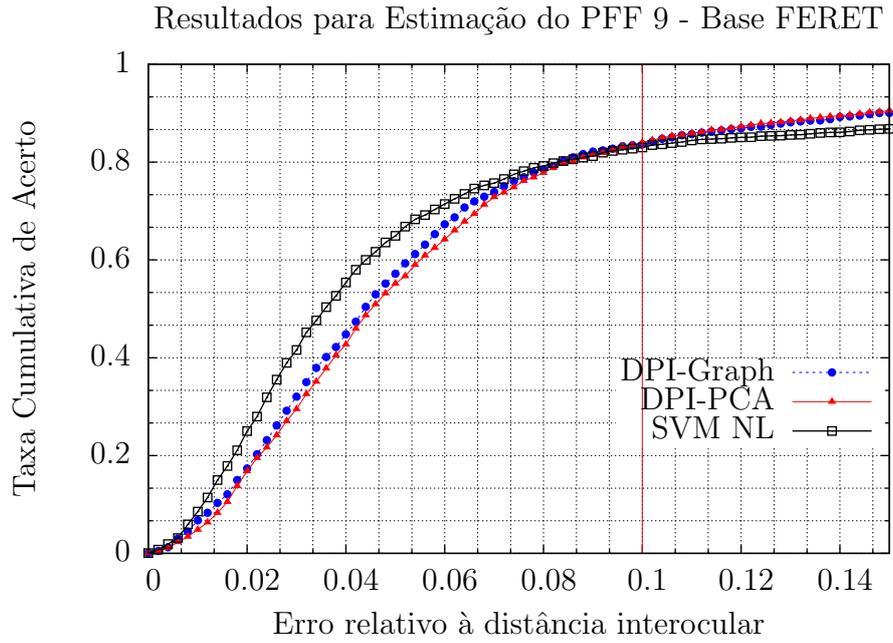


Figura 4.17: Resultados do DPI-GRAPH para a base FERET comparados com os melhores resultados de pós-processamento do DPI-PCA e SVM com kernel não-linear para os pontos fiduciais simétricos  $\mathcal{L}(9)$  e  $\mathcal{L}(10)$  - cantos da boca.

## 4.2.8 Comparativo de Desempenho do DPI-GRAPH *versus* DPI-SIM

Finalizamos a análise de resultados realizando a comparação de desempenhos entre os métodos desenvolvidos nesta dissertação: o DPI-GRAPH e o DPI-SIM. Nos gráficos das Figuras 4.18, 4.19, 4.20, 4.21, 4.22 e 4.23, as taxas médias de acerto cumulativo de ambos métodos são comparadas. Os resultados são demonstrados por par simétrico em ambas bases FERET e BioID, com exceção do PFF  $\mathcal{L}(7)$ , centro do nariz. Assim como nos estudos comparativos do Capítulo 3 e seções anteriores, o ponto de corte verifica-se no erro relativo à distância interocular de 10%.

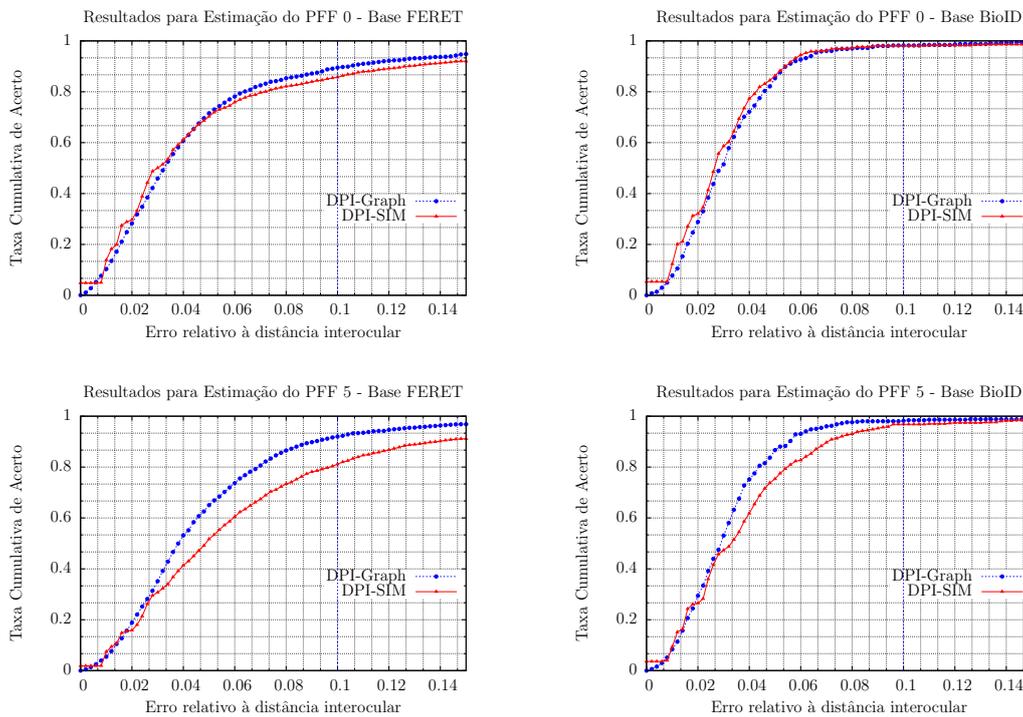


Figura 4.18: Comparativos entre os resultados dos métodos DPI-GRAPH e DPI-SIM para as bases BioID e FERET para os pontos fiduciais  $\mathcal{L}(0)$  e  $\mathcal{L}(5)$  - cantos externos dos olhos.

### PFFs 0 e 5 - Cantos externos dos olhos

- Para a base BioID, as taxas do DPI-GRAPH se mantêm estáveis em torno de 98%, enquanto o DPI-SIM, que atingiu o mesmo desempenho na detecção de  $\mathcal{L}(0)$ , tem sua taxa derrubada em torno de 1% para o PFF simétrico esquerdo,  $\mathcal{L}(5)$ .
- Para a base FERET, o DPI-GRAPH é superior em ambos *landmarks*  $\mathcal{L}(0)$  e  $\mathcal{L}(5)$  em relação ao DPI-SIM, que novamente mostra uma diferença negativa entre as taxas desses pontos simétricos.

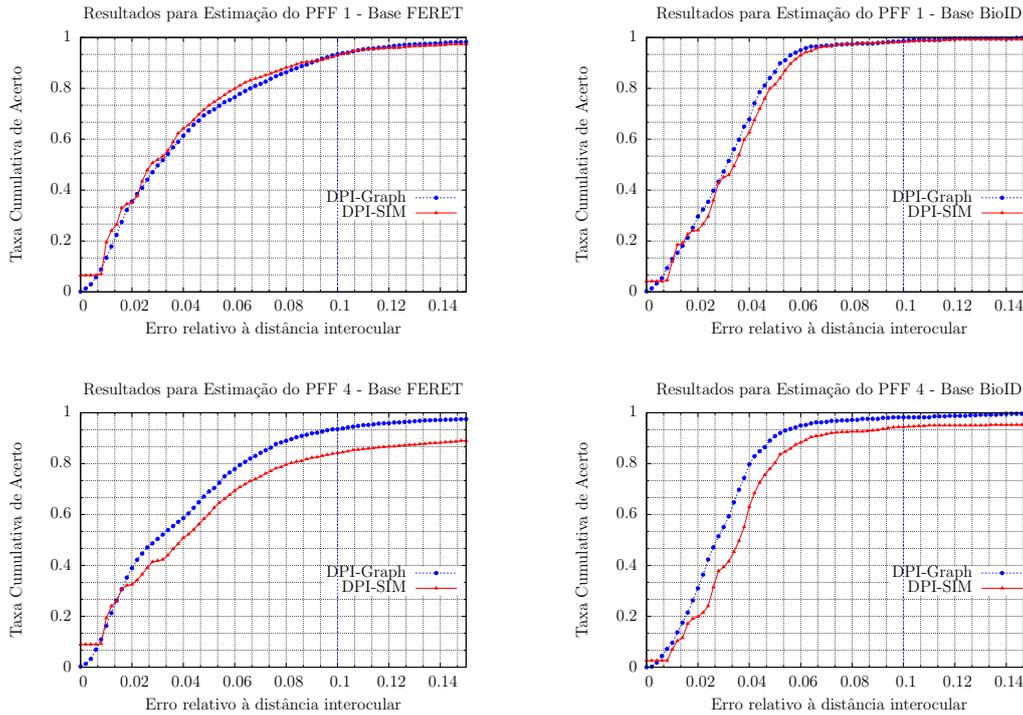


Figura 4.19: Comparativos entre os resultados dos métodos DPI-GRAPH e DPI-SIM para as bases BioID e FERET para os pontos fiduciais  $\mathcal{L}(1)$  e  $\mathcal{L}(4)$  - pupilas.

### PFFs 1 e 4 - Pupilas

- Novamente, em ambas bases FERET e BioID, o DPI-GRAPH é no mínimo equivalente ao DPI-SIM.
- É possível estabelecer que o DPI-GRAPH mantém um compromisso de desempenhos mais próximos entre os pares simétricos, essa tendência é observada até o final desta análise.
- O DPI-SIM possui a tendência de ecoar desempenhos de detecção inferiores no ponto simétrico esquerdo, o que também se observará ao longo desta seção.

### PFFs 2 e 3 - Cantos internos dos olhos

- Este par simétrico mostra um resultado interessante: é aquele onde o DPI-SIM mantém mais estáveis os seus desempenhos em taxa de detecção média. Isso se deve ao fato de serem  $\mathcal{L}(2)$  e  $\mathcal{L}(3)$  os PFFs simétricos mais próximos entre si, e portanto possuindo menores desvios em termos de assimetria.

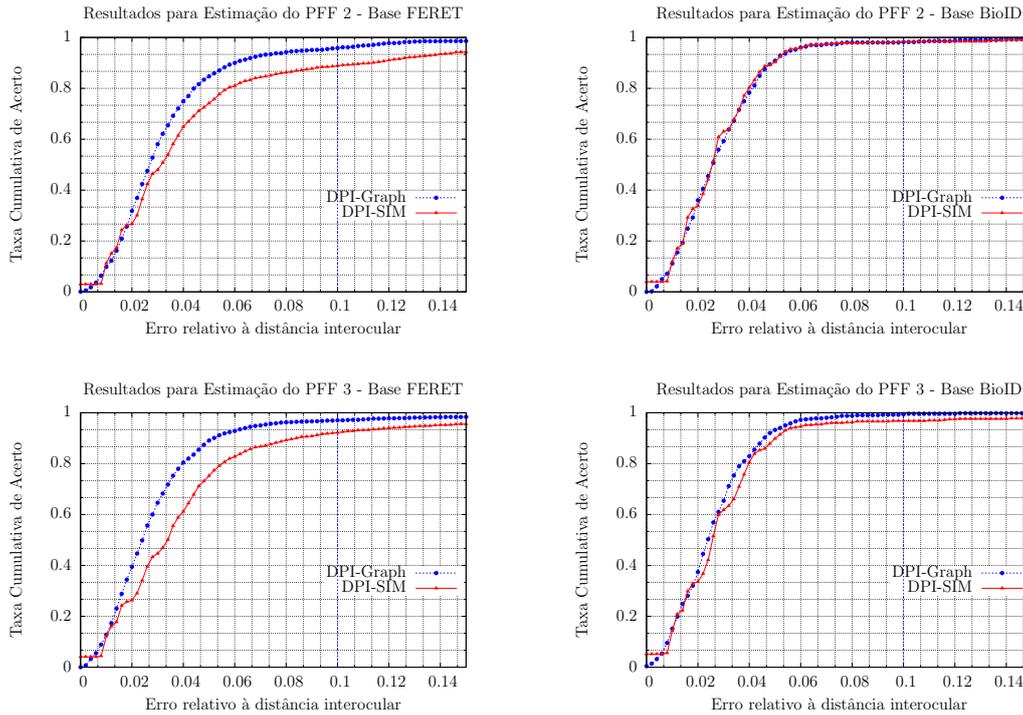


Figura 4.20: Comparativos entre os resultados dos métodos DPI-GRAPH e DPI-SIM para as bases BioID e FERET para os pontos fiduciais  $\mathcal{L}(2)$  e  $\mathcal{L}(3)$  - cantos internos dos olhos.

### PFFs 6 e 8 - Cantos do nariz

- Outro par que mantém as tendências anteriores, com o DPI-SIM possuindo desempenho superior próximo de 1% em relação ao DPI-GRAPH no ponto fiducial  $\mathcal{L}(6)$  na base BioID.

### PFF 7 - Centro do nariz

- Assim como nas análises do Capítulo 3 (págs. 31 e 39), e comentários da Seção 3.5, pág. 42, o DPI-SIM obteve melhor desempenho em taxa de detecção média do *landmark*  $\mathcal{L}(7)$  em ambas bases. Novamente, esse resultado é dado pelo sucesso daquele método em obter um bom eixo de simetria bilateral a partir das nuvens dos classificadores locais.

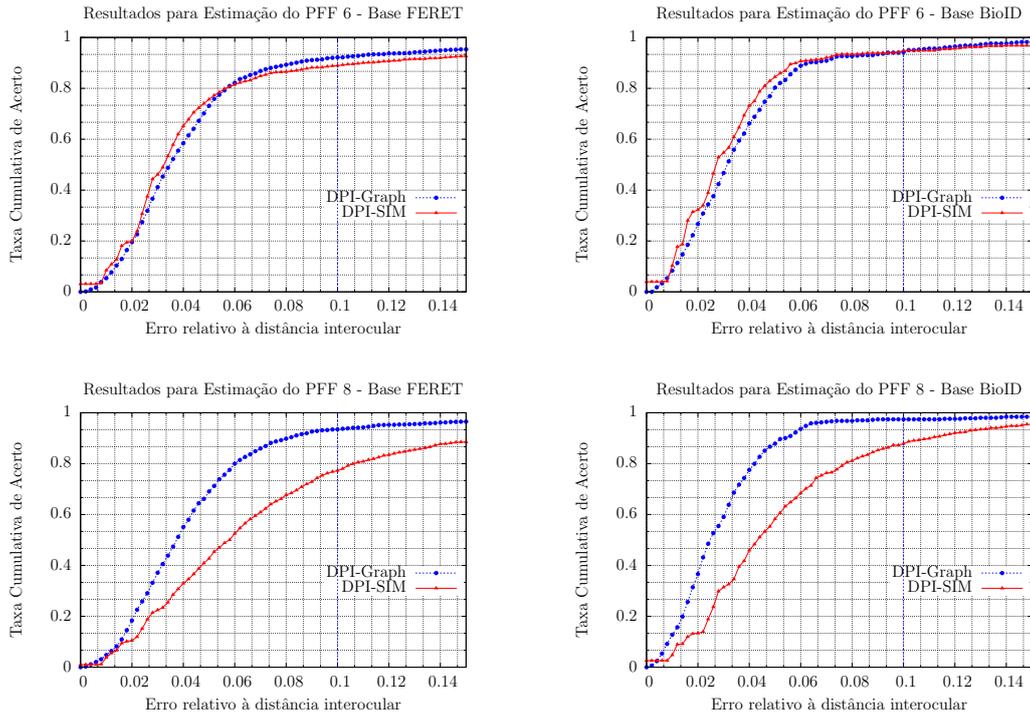


Figura 4.21: Comparativos entre os resultados dos métodos DPI-GRAPH e DPI-SIM para as bases BioID e FERET para os pontos fiduciais  $\mathcal{L}(6)$  e  $\mathcal{L}(8)$  - cantos do nariz.

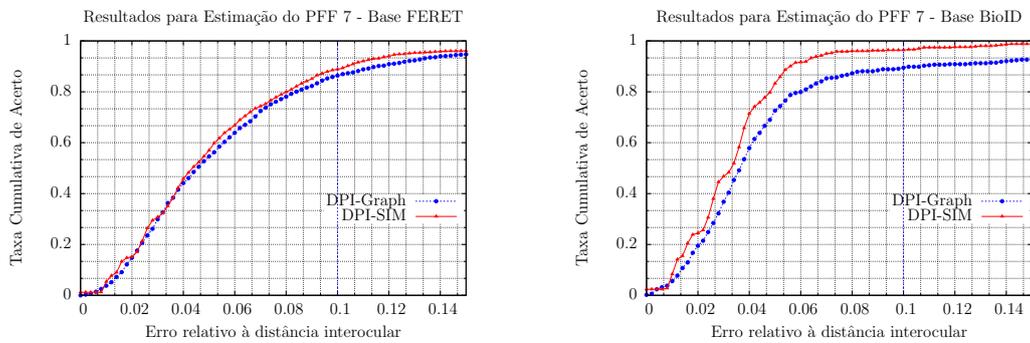


Figura 4.22: Comparativos entre os resultados dos métodos DPI-GRAPH e DPI-SIM para as bases BioID e FERET para o ponto fiducial  $\mathcal{L}(7)$  - centro do nariz.

### PFFs 9 e 10 - Cantos da boca

- Novamente, é um par que mantém as tendências anteriores, com maiores discrepâncias de desempenho para o método DPI-SIM, principalmente na base FERET.
- Com isso, confirmamos um melhor desempenho médio de detecção para o método desenvolvido neste Capítulo: o DPI-GRAPH.

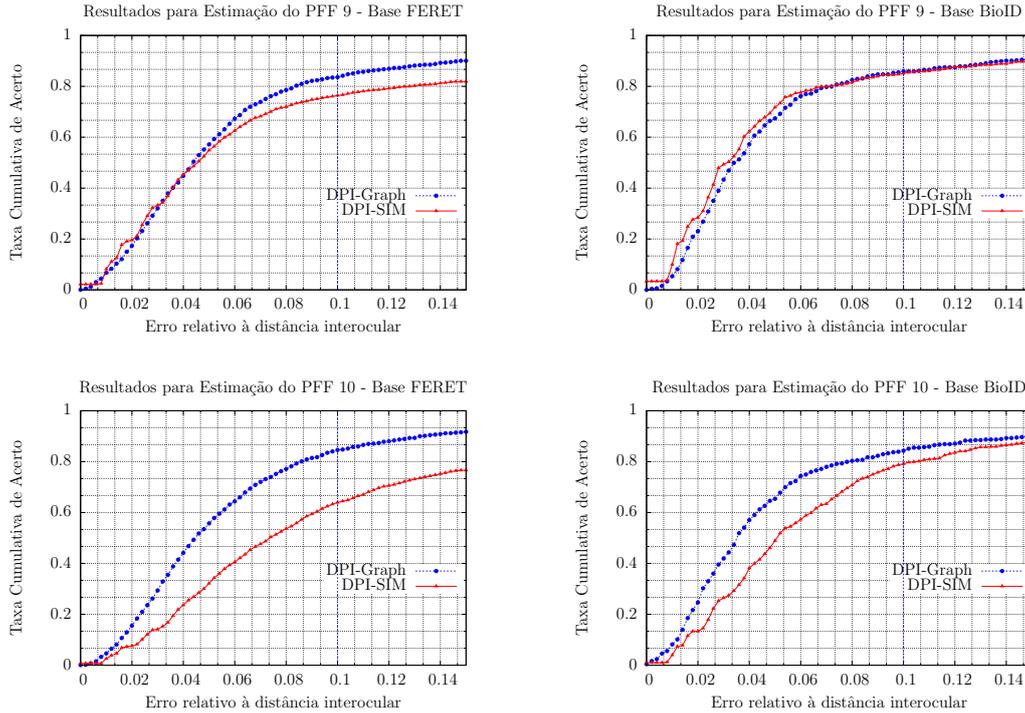


Figura 4.23: Comparativos entre os resultados dos métodos DPI-GRAPH e DPI-SIM para as bases BioID e FERET para os pontos fiduciais  $\mathcal{L}(9)$  e  $\mathcal{L}(10)$  - cantos da boca.

### 4.3 Considerações Finais

Neste capítulo, um método de casamento estatístico de grafos foi aplicado às saídas de classificadores DPI para detecção global de pontos fiduciais faciais humanos. Assim como no Capítulo 3, foram utilizadas imagens de faces oriundas das bases de dados BioID e FERET. A avaliação foi concebida através da metodologia de validação cruzada. Além disso, resumimos o método que motivou o uso de casamento de grafos nas nuvens do DPI: o EBGm e sua variante estatística.

O método desenvolvido aqui obteve ganhos por vezes significativos sobre os resultados anteriores do DPI em sua melhor forma, o DPI-PCA. Em comparação com o SVM-NL, o DPI-GRAPH equivalou e, em alguns casos, superou resultados para os pontos em que o primeiro método possuía desempenhos melhores que os do DPI.

Além disso, o DPI-GRAPH foi superior em praticamente todos os PFFs em relação ao primeiro método proposto neste trabalho: o DPI-SIM. A exceção se deu no *landmark*  $\mathcal{L}(7)$ , centro do nariz, por razões explicitadas ao final daquele capítulo.

# Capítulo 5

## Conclusões

Nesta dissertação, descrevemos dois métodos desenvolvidos para a detecção global de pontos fiduciais faciais: o DPI-SIM e o DPI-GRAPH. A partir da combinação de classificadores DPI locais, ambos métodos exploram os padrões de distribuição espacial das nuvens desses detectores. Isso foi feito com um só objetivo: eleger um conjunto de *landmarks* que representam faces humanas.

O primeiro método, DPI-SIM (Capítulo 3), explora as simetrias da face humana. Construindo eixos de simetria a partir das nuvens dos classificadores locais nas regiões dos olhos, nariz e boca, realiza a seleção de pares de PFFs simétricos entre si. Essa seleção é concebida a partir da aplicação de regras de cunho heurístico. Apesar de não obter os melhores desempenhos para a maioria dos PFFs da marcação escolhida neste trabalho, é bem sucedido na construção do eixo de simetria bilateral. Isso se confirma com os resultados expressivos de detecção no ponto fiducial  $\mathcal{L}(7)$  (págs. 31, 39 e 74), centro do nariz, o qual se localiza sobre o eixo de simetria bilateral na maioria das observações.

O segundo, o DPI-GRAPH (Capítulo 4), realiza o casamento estatístico de grafos. Esse método transforma os elementos das nuvens dos classificadores locais em vértices de um grafo planar. As bases de marcações de PFFs foram convertidas em grupos de grafos modelos, um *bunch graph* simplificado do método EBGM [3, 15], para extração de estatísticas. Com isso, tornou-se possível um casamento de grafos que representam as faces humanas por estimação de máxima verossimilhança. O DPI-GRAPH obteve desempenhos, em sua maioria, superiores aos métodos antecessores para os pontos fiduciais da marcação utilizada. Além disso, confirmou a possibilidade de fazer uso de correspondências espaciais para a detecção global de pontos fiduciais a partir de classificadores DPI locais.

## 5.1 Trabalhos Futuros

A partir dos métodos desenvolvidos e apresentados nesta dissertação, abrem-se novas oportunidades para a continuidade desta linha de pesquisa, como:

- Formulação de novas heurísticas para a eleição de pontos fiduciais utilizando eixos de simetria facial, de modo a promover melhores desempenhos de detecção;
- Retreino dos classificadores locais com *Regiões de Interesse* (ROI, Seção 2.2.4, pág. 10) construídas a partir do levantamento de assimetrias de faces (Seção 3.1.3, pág. 20) do método DPI-SIM. A suposição nesse caso seria que o espalhamento das nuvens possa se tornar mais controlável, além de diminuir o tempo de processamento dessas regiões;
- Uso dos eixos de simetria do método DPI-SIM para rastreamento de faces em vídeo. Esses eixos seriam usados como atratores para estabilizar os pontos fiduciais faciais em detecção. Para isso, os eixos obtidos no quadro  $N - 1$  seriam utilizados como informação a priori no processamento do quadro  $N$ .
- Uso dos grafos do método DPI-GRAPH para rastreamento de faces em vídeo. De forma análoga ao item anterior, os grafos obtidos em um dado quadro seriam usados como informação a priori para o processamento do quadro subsequente;
- Proposição de novas topologias de grafos para o método DPI-GRAPH, utilizando bases de marcações com mais pontos de controle.

# Referências Bibliográficas

- [1] SILVA JR., W. S. *Reconhecimento de Padrões Utilizando Filtros de Correlação com Análise de Componentes Principais*. Tese de D.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2010.
- [2] GÜNTER, M. “Sehen in Mensch und Maschine - Erweiterung des Elastic Bunch Graph Matching”. Selected Topics in Neurocomputing: Vision, Lectures at Ruhr-Universität, Bochum, 2009. Disponível em: <[http://www.neuroinformatik.ruhr-uni-bochum.de/ALL/lectures/detailed/simum\\_docs/07\\_ml\\_detection.pdf](http://www.neuroinformatik.ruhr-uni-bochum.de/ALL/lectures/detailed/simum_docs/07_ml_detection.pdf)>.
- [3] WISKOTT, L., FELLOUS, J.-M., KRÜGER, N., et al. “Face recognition by elastic bunch graph matching”, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 19, n. 7, pp. 775 –779, jul 1997. ISSN: 0162-8828. doi: 10.1109/34.598235.
- [4] PILOWSKY, I., KATSIKITIS, M. “The classification of facial emotions: a computer-based taxonomic approach”, *Journal of Affective Disorders*, v. 30, n. 1, pp. 61 – 71, 1994. ISSN: 0165-0327. doi: DOI:10.1016/0165-0327(94)90152-X. Disponível em: <<http://www.sciencedirect.com/science/article/B6T2X-45Y6PS0-R/2/4811663b376218d2b8c5efe32a38fde2>>.
- [5] KRÜGER, N., PÖTZSCH, M., V.D. MALSBURG, C. “Determination of Face Position and Pose With a Learned Representation Based on Labeled Graphs”, *Internal Report, Institut für Neuroinformatik Ruhr-Universität, Bochum*, 1996.
- [6] MAURER, T., VON DER MALSBURG, C. “Tracking and learning graphs and pose on image sequences of faces”. pp. 176 –181, oct. 1996. doi: 10.1109/AFGR.1996.557261.
- [7] KATSIKITIS, M. *The human face: measurement and meaning*. Kluwer Academic Publishers, 2003. ISBN: 9781402071676.

- [8] PECK, S., PECK, L. “Selected aspects of the art and science of facial esthetics”, *Seminars in Orthodontics*, v. 1, n. 2, pp. 105 – 126, 1995. ISSN: 1073-8746. doi: DOI:10.1016/S1073-8746(95)80097-2. Disponível em: <<http://www.sciencedirect.com/science/article/B75KK-4F4HDJJ-14/2/7df10ec9db7ac121f12db59307e42715>>. The Relevance of Facial Esthetics to Orthodontics.
- [9] EKMAN, P., FRIESEN, W. V. *Unmasking the Face*. Malor Books, 2003. ISBN: 1883536367.
- [10] FARKAS, L. *Anthropometry of the head and face in medicine*. Elsevier, 1981. ISBN: 97804444005571. Disponível em: <<http://books.google.com/books?id=BKBpAAAAMAAJ>>.
- [11] LU, K. H. “Harmonic Analysis of the Human Face”, *Biometrics*, v. 21, n. 2, pp. pp. 491–505, 1965. ISSN: 0006341X. Disponível em: <<http://www.jstor.org/stable/2528106>>.
- [12] GOLDSTEIN, A., HARMON, L., LESK, A. “Identification of human faces”, *Proceedings of the IEEE*, v. 59, n. 5, pp. 748 – 760, maio 1971. ISSN: 0018-9219. doi: 10.1109/PROC.1971.8254.
- [13] MARCELJA, S. “Mathematical description of the responses of simple cortical cells.” *J. Opt. Soc. Am.*, v. 70, n. 11, pp. 1297–1300, November 1980. ISSN: 0030-3941. Disponível em: <<http://view.ncbi.nlm.nih.gov/pubmed/7463179>>.
- [14] ARAUJO, G. M. *Algoritmo para Reconhecimento de Características Faciais Baseado em Filtros de Correlação*. Tese de M.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2010.
- [15] WÜRTZ, R. P., GÜNTHER, M. “Face Detection and Recognition Using Maximum Likelihood Classifiers on Gabor Graphs”, *International Journal on Pattern Recognition and Artificial Intelligence*, v. 23, n. 3, pp. 433–461, 2009.
- [16] SABINO, W., ARAUJO, G. “Base de Marcações de Pontos Fiduciais de Sabino e Araujo nas Bases de Imagens de Faces BioID e FERET.” . <http://www.lps.ufrj.br/gmatos>, 2009. [Último acesso em Fevereiro de 2010].
- [17] AG, B. “BIOID Database”. , 2010. Disponível em: <<http://www.bioid.com>>.

- [18] OF STANDARDS, N. I., TECHNOLOGY. “The facial recognition technology (FERET) database”. , 2010. Disponível em: <<http://www.itl.nist.gov/iad/humanid/feret/>>.
- [19] VUKADINOVIC, D., PANTIC, M. “Fully automatic facial feature point detection using Gabor feature based boosted classifiers”. In: *Systems, Man and Cybernetics, 2005 IEEE International Conference on*, v. 2, pp. 1692 – 1698 Vol. 2, 2005. doi: 10.1109/ICSMC.2005.1571392.
- [20] DU, C., WU, Q., YANG, J., et al. “SVM based ASM for facial landmarks location”. pp. 321 –326, july 2008.
- [21] COOTES, T. F., TAYLOR, C. J., COOPER, D. H., et al. “Active shape models - Their training and application”, *Comput. Vis. Image Underst.*, v. 61, pp. 38–59, January 1995. ISSN: 1077-3142. doi: 10.1006/cviu.1995.1004. Disponível em: <<http://portal.acm.org/citation.cfm?id=206543.206547>>.
- [22] COOTES, T. F., EDWARDS, G. J., TAYLOR, C. J. “Active appearance models”, *Proceedings of the European Conference on Computer Vision*, v. 2, pp. 484–498, 1998.
- [23] FREUND, Y., SCHAPIRE, R. E. “A decision-theoretic generalization of on-line learning and an application to boosting”. In: *Proceedings of the Second European Conference on Computational Learning Theory*, pp. 23–37, London, UK, 1995. Springer-Verlag. ISBN: 3-540-59119-2. Disponível em: <<http://portal.acm.org/citation.cfm?id=646943.712093>>.
- [24] XIE, C., SAVVIDES, M., KUMAR, B. V. K. V. “Redundant Class-Dependence Feature Analysis Based on Correlation Filters Using FRGC2.0 Data”. In: *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, v. 3, pp. 153–158 (6), San Diego, California, USA, 20-25 de Junho 2005.
- [25] THEODORIDIS, S., KOUTROUMBAS, K. *Pattern Recognition, Fourth Edition*. Academic Press, 2008. ISBN: 1597492728, 9781597492720.
- [26] ZHAO, W., CHELLAPPA, R. *Face Processing: Advanced Modeling and Methods*. Orlando, FL, USA, Academic Press, Inc., 2005. ISBN: 0120884526.
- [27] VIOLA, P., JONES, M. J. “Robust Real-time Face Detection”, *International Journal of Computer Vision*, v. 57, n. 2, pp. 137–154, 2004.

- [28] TAN, X., TRIGGS, B. “Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions”, *Image Processing, IEEE Transactions on*, v. 19, n. 6, pp. 1635–1650, 2010. ISSN: 1057-7149. doi: 10.1109/TIP.2010.2042645.
- [29] CORNELIUS, H., LOY, G. “Efficient Symmetry Detection Using Local Affine Frames”. . Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.92.6614>>.
- [30] STRANG, G. *Linear Algebra and Its Applications*. Brooks Cole, February 1988. ISBN: 0155510053.
- [31] JESORSKY, O., KIRCHBERG, K. J., FRISCHHOLZ, R. “Robust face detection using the Hausdorff distance”. In: *AVBPA '01: Proceedings of the Third International Conference on Audio- and Video-Based Biometric Person Authentication*, pp. 90–95, London, UK, 2001. Springer-Verlag.
- [32] MITRA, S., LIU, Y. “Local facial asymmetry for expression classification”. In: *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, 2004.
- [33] CONTE, D., FOGGIA, P., SANSONE, C., et al. “Thirty Years Of Graph Matching In Pattern Recognition”, *International Journal of Pattern Recognition and Artificial Intelligence*, 2004. doi: 10.1142/S0218001404003228. Disponível em: <<http://www.worldscinet.com/ijprai/18/1803/S0218001404003228.html>>.
- [34] WISKOTT, L. *Labeled Graphs and Dynamic Link Matching for Face Recognition and Scene Analysis*. PhD thesis, Ruhr-Universität, Bochum, 1995.
- [35] LADES, M., VORBRUGGEN, J., BUHMANN, J., et al. “Distortion invariant object recognition in the dynamic link architecture”, *Computers, IEEE Transactions on*, v. 42, n. 3, pp. 300–311, mar 1993. ISSN: 0018-9340. doi: 10.1109/12.210173.
- [36] DAUGMAN, J. G. “Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters”, *J. Opt. Soc. Am. A*, v. 2, n. 7, pp. 1160–1169, Jul 1985. doi: 10.1364/JOSAA.2.001160. Disponível em: <<http://josaa.osa.org/abstract.cfm?URI=josaa-2-7-1160>>.
- [37] JONES, J. P., PALMER, L. A. “An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex”, *J Neu-*

*rophysiol*, v. 58, n. 6, pp. 1233–1258, 1987. Disponível em: <<http://jn.physiology.org/cgi/content/abstract/58/6/1233>>.

- [38] GABOR, D. “Theory of communication. Part 1: The analysis of information”, *Electrical Engineers - Part III: Radio and Communication Engineering, Journal of the Institution of*, v. 93, n. 26, pp. 429 –441, 1946. doi: 10.1049/ji-3-2.1946.0074.
- [39] DUDA, R. O., HART, P. E., STORK, D. G. *Pattern Classification*. Wiley-Interscience Publication, 2000.
- [40] MOGHADDAM, B., PENTLAND, A. “Probabilistic visual learning for object representation”, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 19, n. 7, pp. 696 –710, jul. 1997. ISSN: 0162-8828. doi: 10.1109/34.598227.
- [41] TEIXEIRA, M. L. *The Bayesian Intrapersonal/Extrapersonal Classifier*. Masters thesis, Colorado State University, Fort Collins, Colorado, 2010.
- [42] BOLME, D. S. “Elastic Bunch Graph Matching”, *Colorado State University Masters Thesis*, 2003.
- [43] DIESTEL, R. *Graph Theory (Graduate Texts in Mathematics)*. Springer, August 2005.

# Apêndice A

## Grafos: conceitos básicos

Neste pequeno apêndice, apresentamos em resumo alguns conceitos que são utilizados ao longo desta dissertação onde grafos são ferramentas de representação de dados e objetos. Mais detalhes de Teoria de Grafos podem ser encontrados em [43].

### A.1 Alguns Conceitos de Teoria dos Grafos

- Um *grafo*  $\mathcal{G}$  é um par  $(\mathcal{L}, \mathcal{E})$ , onde  $\mathcal{E}$  é um conjunto de pares de  $\mathcal{L}$  com  $\{l_i, l_j\}_{i \neq j}$ . Os elementos do conjunto  $\mathcal{L}$  são chamados *vértices* do grafo  $\mathcal{G}$  (também *nós* ou *pontos*), e os elementos de  $\mathcal{E}$  de *arestas* (ou *linhas*).
- Se um par de vértices  $\{l_i, l_j\}_{i \neq j}$  é conectado por uma aresta  $e_m$  então  $l_i$  e  $l_j$  são ditos *vértices adjacentes*.
- Duas arestas  $e_m$  e  $e_n$  são ditas *adjacentes* se possuírem um vértice  $l_i$  em comum.
- A *ordem* de um grafo  $\mathcal{G}$ , escrita por  $|\mathcal{G}|$ , é dada pelo seu número de vértices; o seu número de arestas é escrito como  $||\mathcal{G}||$ .
- O *grau* de um vértice  $l_i$  é dado pelo número de arestas incidentes em  $l_i$ , e é escrito por  $d_{\mathcal{G}}(l_i)$ . O *grau máximo* de um grafo  $\mathcal{G}$  é escrito como  $\Delta(\mathcal{G}) = \max\{d_{\mathcal{G}}(l) : l \in \mathcal{L}\}$ .
- Se um grafo é estruturado de tal maneira que duas arestas jamais se encontram em um ponto que não um vértice, então esse grafo é dito *planar*.
- Sejam  $\mathcal{G} = (\mathcal{L}, \mathcal{E})$  e  $G = (L, E)$ , se  $L \subseteq \mathcal{L}$  e  $E \subseteq \mathcal{E}$  então  $G$  é um *subgrafo* de  $\mathcal{G}$ , e  $\mathcal{G}$  é o *supergrafo* de  $G$ .
- Se  $G$  é um subgrafo de  $\mathcal{G}$  e contém todas as suas arestas  $E \in \mathcal{E}$ , dizemos que  $G$  é um *subgrafo induzido* de  $\mathcal{G}$ .