

UNIVERSIDADE DA BEIRA INTERIOR

DEPARTAMENTO DE INFORMÁTICA



## **Reconhecimento Biométrico Periocular: Análise de Factores de Variabilidade**

MIGUEL AFONSO, N°23859  
LICENCIATURA ENG. INFORMÁTICA

ORIENTADOR DE PROJECTO: PROF. DOUTOR HUGO PROENÇA

Junho, 2011



# Agradecimentos

Em primeiro lugar não quero deixar passar a oportunidade de poder agradecer a todos aqueles que mais me apoiaram para poder concluir este projeto.

Gostaria de começar por agradecer ao meu orientador de projeto, o Professor Doutor Hugo Proença, pela enorme vontade e disponibilidade que teve para me ajudar, pela forma como facilmente me transmitiu tanto conhecimento importante para a realização deste projeto e acima de tudo por todos os conselhos que me deu.

Quero agradecer a todos os membros e colaboradores do grupo *Soft Computing and Image Analysis Laboratory (SOCIA Lab)*, por tudo o que me ensinaram e pela vontade com que todos me ajudaram, especialmente o Sílvio, o Gil e o Emanuel.

Por fim, mas essencial para mim, agradeço o apoio da minha família e da minha namorada, que nunca pararam de me apoiar e motivar para a conclusão deste projeto.

Os meus sinceros agradecimentos,

Miguel Afonso



# Conteúdo

Agradecimentos	i
Conteúdo	iii
Lista de Tabelas	v
Lista de Figuras	vii
Acrónimos	xi
Glossário	xiii
<b>1 Introdução</b>	<b>1</b>
1.1 Motivação . . . . .	2
1.2 Abordagem . . . . .	3
1.3 Estrutura do relatório . . . . .	5
<b>2 Reconhecimento Periocular</b>	<b>7</b>
2.1 Método base . . . . .	7
2.1.1 <i>Scale Invariant Feature Transform</i> . . . . .	8
2.1.2 <i>Local Binary Pattern</i> . . . . .	20
2.1.3 <i>Histogram of Oriented Gradients</i> . . . . .	22
2.2 Método proposto . . . . .	23
2.2.1 <i>Viola and Jones</i> . . . . .	23
2.2.2 Extração de características . . . . .	28
2.2.3 Comparação de imagens . . . . .	30
<b>3 Experiências e Resultados</b>	<b>31</b>
3.1 Conjunto de dados . . . . .	31

---

3.2	Medidas de desempenho . . . . .	35
3.3	Análise individual de cada método . . . . .	36
3.3.1	Resultados do <i>SIFT</i> . . . . .	36
3.3.2	Resultados do <i>HOG</i> . . . . .	37
3.3.3	Resultados do <i>LBP</i> . . . . .	38
3.3.4	Cálculo do vetor de pesos . . . . .	39
3.3.5	Resultados da fusão dos métodos . . . . .	40
3.4	Testes ao método proposto . . . . .	48
3.4.1	Análise ao conjunto <i>Frontal</i> . . . . .	49
3.4.2	Análise ao conjunto <i>Hair</i> . . . . .	54
3.4.3	Análise ao conjunto <i>Rotation</i> . . . . .	59
3.4.4	Análise ao conjunto <i>Glasses</i> . . . . .	64
3.4.5	Análise ao conjunto <i>Temporal</i> . . . . .	69
3.4.6	Análise ao conjunto <i>Nasal</i> . . . . .	74
3.5	Síntese dos resultados obtidos . . . . .	79
<b>4</b>	<b>Conclusão e trabalho futuro</b>	<b>83</b>
4.1	Conclusão . . . . .	83
4.2	Trabalho futuro . . . . .	84
<b>A</b>	<b>Ambiente de desenvolvimento</b>	<b>87</b>
A.1	Linguagens de desenvolvimento . . . . .	87
A.2	Bibliotecas utilizadas . . . . .	87
	<b>Bibliografia</b>	<b>89</b>

# Lista de Tabelas

3.1	Tabela de correlação entre os métodos . . . . .	42
3.2	Resultados dos diferentes métodos . . . . .	42
3.3	Resultados dos diferentes métodos no conjunto <i>Frontal</i> . . .	49
3.4	Resultados dos diferentes métodos no conjunto <i>Hair</i> . . . .	54
3.5	Resultados dos diferentes métodos no conjunto <i>Rotation</i> . . .	59
3.6	Resultados dos diferentes métodos no conjunto <i>Glasses</i> . . .	64
3.7	Resultados dos diferentes métodos no conjunto <i>Temporal</i> . .	69
3.8	Resultados dos diferentes métodos no conjunto <i>Nasal</i> . . . .	74
3.9	Resultados dos diferentes conjuntos . . . . .	79
3.10	Representação das matrizes de confusão dos conjuntos . . .	80



# Lista de Figuras

2.1	Representação da obtenção das DoG para as oitavas de uma imagem [9] . . . . .	10
2.2	Deteção de extremos no espaço-escala [9] . . . . .	11
2.3	Histograma de orientações dominantes dos gradientes [1] . . . . .	16
2.4	Construção do descritor para um ponto-chave de 2x2 com 48 elementos [9] . . . . .	17
2.5	Correspondência entre duas imagens . . . . .	18
2.6	Calcular o valor LBP de cada pixel [16] . . . . .	20
2.7	Divisão da face em descritores LBP locais [15] . . . . .	21
2.8	Gradientes de uma célula [9] . . . . .	22
2.9	Construção da imagem integral [17] . . . . .	24
2.10	Cálculo das somas [18] . . . . .	24
2.11	Diferentes tipos de características [18] . . . . .	25
2.12	Demonstração do funcionamento da cascata [6] . . . . .	27
2.13	Método de deteção: íris (vermelho), olho (verde) e região periocular (azul) . . . . .	28
2.14	Exemplo de pontos chave detetados . . . . .	29
2.15	Exemplo das regiões seleccionadas para o HOG e LBP . . . . .	29
3.1	Exemplo de imagens excluídas . . . . .	31
3.2	Exemplo de imagens do conjunto conjunto <i>ALL</i> . . . . .	32
3.3	Exemplo de imagens do conjunto <i>Frontal</i> . . . . .	32
3.4	Exemplo de imagens do conjunto <i>Hair</i> . . . . .	33
3.5	Exemplo de imagens do conjunto <i>Rotation</i> . . . . .	33
3.6	Exemplo de imagens do conjunto <i>Glasses</i> . . . . .	33
3.7	Exemplo de imagens do conjunto <i>Temporal</i> . . . . .	34
3.8	Exemplo de imagens do conjunto <i>Nasal</i> . . . . .	34

3.9	Curva ROC e distribuição de probabilidades por classe do SIFT . . . . .	36
3.10	Curva ROC e distribuição de probabilidades por classe do HOG . . . . .	37
3.11	Curva ROC e distribuição de probabilidades por classe do LBP . . . . .	38
3.12	Curva ROC e distribuição de probabilidades por classe da fusão . . . . .	40
3.13	Disposição das melhores e piores comparações de cada uma das classes . . . . .	43
3.14	Melhor resultado da classe <i>Intra</i> para o conjunto <i>All</i> . . . . .	44
3.15	Pior resultado da classe <i>Intra</i> para o conjunto <i>All</i> . . . . .	45
3.16	Melhor resultado da classe <i>Inter</i> para o conjunto <i>All</i> . . . . .	46
3.17	Pior resultado da classe <i>Inter</i> para o conjunto <i>All</i> . . . . .	47
3.18	Curva ROC e distribuição de probabilidades do conjunto <i>Frontal</i> . . . . .	49
3.19	Melhor resultado da classe <i>Intra</i> para o conjunto <i>Frontal</i> . . . . .	50
3.20	Pior resultado da classe <i>Intra</i> para o conjunto <i>Frontal</i> . . . . .	51
3.21	Melhor resultado da classe <i>Inter</i> para o conjunto <i>Frontal</i> . . . . .	52
3.22	Pior resultado da classe <i>Inter</i> para o conjunto <i>Frontal</i> . . . . .	53
3.23	Curva ROC e distribuição de probabilidades do conjunto <i>Hair</i> . . . . .	54
3.24	Melhor resultado da classe <i>Intra</i> para o conjunto <i>Hair</i> . . . . .	55
3.25	Pior resultado da classe <i>Intra</i> para o conjunto <i>Hair</i> . . . . .	56
3.26	Melhor resultado da classe <i>Inter</i> para o conjunto <i>Hair</i> . . . . .	57
3.27	Pior resultado da classe <i>Inter</i> para o conjunto <i>Hair</i> . . . . .	58
3.28	Curva ROC e distribuição de probabilidades do conjunto <i>Rotation</i> . . . . .	59
3.29	Melhor resultado da classe <i>Intra</i> para o conjunto <i>Rotation</i> . . . . .	60
3.30	Pior resultado da classe <i>Intra</i> para o conjunto <i>Rotation</i> . . . . .	61
3.31	Melhor resultado da classe <i>Inter</i> para o conjunto <i>Rotation</i> . . . . .	62
3.32	Pior resultado da classe <i>Inter</i> para o conjunto <i>Rotation</i> . . . . .	63
3.33	Curva ROC e distribuição de probabilidades do conjunto <i>Glasses</i> . . . . .	64
3.34	Melhor resultado da classe <i>Intra</i> para o conjunto <i>Glasses</i> . . . . .	65
3.35	Pior resultado da classe <i>Intra</i> para o conjunto <i>Glasses</i> . . . . .	66
3.36	Melhor resultado da classe <i>Inter</i> para o conjunto <i>Glasses</i> . . . . .	67
3.37	Pior resultado da classe <i>Inter</i> para o conjunto <i>Glasses</i> . . . . .	68

---

3.38	Curva ROC e distribuição de probabilidades do conjunto <i>Temporal</i> . . . . .	69
3.39	Melhor resultado da classe <i>Intra</i> para o conjunto <i>Temporal</i> . . .	70
3.40	Pior resultado da classe <i>Intra</i> para o conjunto <i>Temporal</i> . . . .	71
3.41	Melhor resultado da classe <i>Inter</i> para o conjunto <i>Temporal</i> . . .	72
3.42	Pior resultado da classe <i>Inter</i> para o conjunto <i>Temporal</i> . . . .	73
3.43	Curva ROC e distribuição de probabilidades do conjunto <i>Nasal</i> . . . . .	74
3.44	Melhor resultado da classe <i>Intra</i> para o conjunto <i>Nasal</i> . . . .	75
3.45	Pior resultado da classe <i>Intra</i> para o conjunto <i>Nasal</i> . . . . .	76
3.46	Melhor resultado da classe <i>Inter</i> para o conjunto <i>Nasal</i> . . . .	77
3.47	Pior resultado da classe <i>Inter</i> para o conjunto <i>Nasal</i> . . . . .	78
4.1	Perspetiva de trabalho futuro . . . . .	85



# Acrónimos

**SOCIA Lab** Soft Computing and Image Analysis Laboratory

**FNR** False Negative Rate

**FPR** False Positive Rate

**TNR** True Negative Rate

**TPR** True Positive Rate

**ROC** Receiver Operating CharacteristicROC

**SIFT** Scale Invariant Feature Transform

**DoG** Difference of Gaussian

**EER** Equal Error Rate

**AUC** Area Under Curve

**MatLab** Matrix Laboratory

**HOG** Histogram of Oriented Gradients

**LBP** Local Binary Pattern

**BBF** Best-Bin-First



# Glossário

**Equal Error Rate (EER)** - é a taxa de erro para a verificação do funcionamento de um sistema para se ajustar o limiar para a aceitação / rejeição da decisão, é adaptada de tal forma a que as taxas de falsos positivos e falsos negativos se tornam iguais.

**True Positive Rate (TPR)** - ou taxa de verdadeiros positivos é a taxa de comparações que foram identificadas corretamente como sendo comparações entre imagens de pessoas diferentes.

**False Positive Rate (FPR)** - ou taxa de falsos positivos é a taxa de comparações que foram identificadas erradamente como sendo comparações entre imagens da mesma pessoa.

**False Negative Rate (FNR)** - ou taxa de falsos negativos é a taxa de comparações que foram identificadas erradamente como sendo comparações entre imagens de pessoas diferentes.

**True Negative Rate (TNR)** - ou taxa de verdadeiros negativos é a taxa de comparações que foram identificadas corretamente como sendo comparações entre imagens da mesma pessoa.

**Receiver Operating Characteristic (ROC)** - ou simplesmente curva ROC é a representação gráfica da taxa de verdadeiros positivos com a taxa de falsos positivos.



# Capítulo 1

## Introdução

Este trabalho foi desenvolvido no âmbito do projecto de investigação PTDC/EIA/103945/2008, “NECOVID: Covert Negative Biometric Recognition”, financiado pela FCT/FEDER. O reconhecimento biométrico através da região periocular (região envolvente ao olho) baseia-se no princípio de que a “forma dos olhos” é das que mais informação discriminante entre indivíduos contém. Neste sentido, vários trabalhos foram propostos com vista ao reconhecimento biométrico através da região periocular.

Utilizámos na realização deste trabalho um dos métodos mais conhecidos de reconhecimento periocular [13], tendo como principal objetivo analisar as variações de desempenho do mesmo em função dos fatores de variabilidade clássicos dos dados, testando alterações nos parâmetros do método proposto, para a obtenção de melhores resultados.

Este trabalho foi concebido e desenvolvido principalmente no laboratório SOCIA Lab.

## 1.1 Motivação

Nos tempos em que vivemos, a humanidade é diariamente confrontada com vários desafios, para os quais anteriormente nunca existiram respostas eficazes. Foi necessário esperar muito tempo para concretizar dois grandes desafios: organização da sociedade e segurança.

Mas, pensando em segurança, hoje em dia um pouco por todo o lado, possuímos um vasto leque de métodos de autenticação que se desenvolveram ao sabor do requinte tecnológico que a todos diariamente surpreende. Para corroborar tal surpresa, podemos pensar no acesso a entradas dos edifícios, no controle de acesso a computadores, no acesso a caixas de multibanco, na investigação criminal, e muitas mais. Existe a crescente necessidade de identificação e autenticação que deve ser posta ao serviço de uma sociedade moderna e exigente.

A maioria dos métodos existentes para identificação e autenticação através das características biométricas de um indivíduo, tem a desvantagem do utilizador ter de cooperar com o sistema, ainda existindo pouca investigação por parte da comunidade científica para analisar características humanas em sistemas de identificação e autenticação que não exijam a cooperação por parte do utilizador.

As características biométricas consistem na medida de particularidades individuais, tais como: impressões digitais ou proporções entre características que podem identificar univocamente uma pessoa. Efetuando uma analogia, podemos dizer que a biometria é uma palavra-chave, isto é, não pode ser esquecida, perdida ou roubada, podendo autenticar ou identificar sem a cooperação por parte do utilizador.

A análise de características biométricas em Visão Computacional tem sido uma área em vasta expansão ao longo dos últimos anos, quer pela diversidade de aplicações, quer pela potenciação de recursos promovidos pelos inúmeros desenvolvimentos tecnológicos. A motivar estes trabalhos também se encontra a crescente necessidade de automatizar computacionalmente processos de identificação e autenticação, pelo que se tem investido cada vez mais no reconhecimento biométrico, utilizando técnicas de Visão Computacional.

Os principais métodos de análise de características biométricas utilizam processos automatizados de reconhecimento de padrões. Estes processos requerem eficiência, eficácia e agilidade crescentes, tornando nesse sentido as tarefas visuais automatizadas em excelentes soluções, especialmente pelo suporte conferido e pelo avanço da computação e dos sistemas automáticos.

Na simulação do comportamento visual humano, o reconhecimento através da região periocular numa pessoa constitui uma tarefa com crescente potencial, estando recentemente a ser visto como um meio de melhorar a robustez de sistemas biométricos do rosto ou da íris [22], quando o conteúdo destes inerentes nas imagens originais é pobre (por exemplo, devido à má qualidade da imagem) [21]. Também foi sugerido que as características da região periocular podem potencialmente ser utilizadas como um método de classificação biométrica fraco.

## 1.2 Abordagem

Para o tema proposto, utilizou-se a abordagem descrita por U. Park, A. Ross e A. K. Jain(2009) [13], extraindo informação da região periocular através da análise de pontos e texturas resultando num conjunto de características locais e globais.

Características locais são as que se focam em regiões da imagem, sendo que as utilizadas neste projeto começam por detetar um conjunto de pontos chave e codificam cada um destes pontos usando os valores dos pixels que o envolvem. Para efetuar a comparação entre duas imagens, o resultado consiste no total de pontos chave correspondentes entre essas duas imagens.

As características globais são extraídas da imagem inteira (ou de uma região de interesse) ou de um conjunto de regiões locais, sendo representadas por vetores de tamanho fixo e o processo de comparação simplesmente calcula as diferenças entre estes vetores de tamanho fixo, o que se torna muito eficiente temporalmente.

Nesta abordagem são implementados três métodos de extração de características produzindo o conjunto das mesmas que são usadas para efetuar a comparação entre diferentes imagens, sendo testado também a fusão destas características.

O primeiro método utilizado é um método de comparação local de extração de características invariantes, Scale Invariant Feature Transform (SIFT), descrito por David G. Lowe [9] [10]. Sendo muito utilizado devido à sua aproximação baseada em modelos, torna-se um modelo generalista, capitalizando numa estrutura específica e com a vantagem de ser robusto a ruído e a oclusão.

De seguida, utilizámos um método para extração de características globais, Local Binary Pattern (LBP), tendo sido descrito pela primeira vez por T. Ojala, M. Pietikäinen e D. Harwood [11] [12]. É um método muito eficiente que atribui um valor a cada pixel de uma imagem, através da limirização dos pixels vizinhos considerando o resultado como um número binário. Devido ao poder discriminativo que possui e a sua simplicidade computacional, este operador de texturas tem sido utilizado em diversas aplicações, principalmente como um classificador de características humanas [19].

Por fim, para extrair o mesmo tipo de características do segundo método, este terceiro método, Histogram of Oriented Gradients (HOG), é um descritor de características utilizado na Visão Computacional e processamento de imagens com o objetivo de deteção e reconhecimento de objetos [4]. Este método é processado numa grelha densa de células com a mesma área e usa uma sobreposição do contraste local normalizado para uma melhor precisão.

Para podermos utilizar esta abordagem precisamos de identificar a região periocular de cada imagem, algo que era efetuado manualmente na abordagem anterior. Automatizámos este processo utilizando o método descrito por Viola e Jones [18], devido à sua elevada fiabilidade, tanto em termos de tempo de processamento como de taxas de erro.

---

## 1.3 Estrutura do relatório

Este relatório encontra-se dividido em quatro capítulos.

**Capítulo 1 - Introdução** No primeiro capítulo optou-se por fazer uma introdução ao tema e descrever as motivações que originaram este projeto.

**Capítulo 2 - Reconhecimento Periocular** Descrição do método que foi usado na elaboração deste projeto e alterações efetuadas para a obtenção de melhores resultados.

**Capítulo 3 - Experiências e Resultados** Neste capítulo descreve-se as experiências realizadas, os conjuntos de dados e os resultados obtidos.

**Capítulo 4 - Conclusão e trabalho futuro** Como o próprio nome indica, pretende-se retirar algumas conclusões sobre a realização deste projeto bem como descrever um pouco o que poderá ser acrescentado a este trabalho na tentativa da obtenção de melhores resultados.



# Capítulo 2

## Reconhecimento Periocular

Antes de realizar algumas das etapas do trabalho, foi necessário um estudo aprofundado dos métodos utilizados para efetuar o reconhecimento através da região periocular, para uma melhor compreensão e implementação dos mesmos. Neste capítulo serão expostos alguns desses conceitos que facilitaram o desenvolvimento do trabalho.

### 2.1 Método base

Como método base utilizou-se a abordagem descrita por U. Park, A. Ross e A. K. Jain(2009) [13], extraindo informação da região periocular através da análise de pontos e texturas resultando num conjunto de características locais e globais.

Como foi referido anteriormente, usou-se o método SIFT para a extração de características locais e dois métodos para a extração de características globais, denominados por HOG e LBP.

### 2.1.1 *Scale Invariant Feature Transform*

Este método foi descrito por David G. Lowe [9] [10], sendo um método com uma aproximação baseada em modelos, criando uma estrutura específica e com a vantagem de ser robusta com ruído e oclusão.

Esta aproximação transforma uma imagem numa larga coleção de vetores de características, as quais apresentam invariância em relação à translação, escala, rotação de uma imagem, e parcialmente invariantes a mudanças de luminância, projeção 3D e enviesamento (transformada afim). As características são bem localizadas no domínio espacial e frequência, reduzindo a probabilidade de mau funcionamento por oclusão, bloqueio ou ruído, promovendo a sustentabilidade de reconhecimento de objetos ou cenas.

O custo de extração destas características é minimizado devido a uma filtragem em cascata, em que operações computacionalmente mais custosas são apenas aplicadas a localizações que passam num teste inicial. O algoritmo SIFT é executado através de cinco etapas principais: deteção de extremos, localização de pontos-chave, definição da orientação, descrição dos pontos-chave e comparação entre as imagens. Para melhor compreensão deste método explicamos de seguida cada uma das etapas.

#### **Deteção de Extremos**

A primeira etapa do SIFT consiste em procurar pontos que sejam invariantes a mudanças de escala da imagem, possibilitando a deteção de pontos independentemente da distância a que vemos a imagem. Tal objetivo é alcançado procurando características estáveis em diferentes escalas, utilizando uma função chamada de espaço de escala [20], que neste caso é a função Gaussiana.

Uma imagem  $I(x,y)$  passa a ser definida por uma  $L(x,y,\sigma)$ , no espaço-escala. Esta função é produzida pela convolução de uma função Gaussiana,  $G(x,y,\sigma)$ , com a imagem,  $I(x,y)$ :

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (2.1)$$

Sendo este filtro variável à escala através do parâmetro  $\sigma$ , na equação

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (2.2)$$

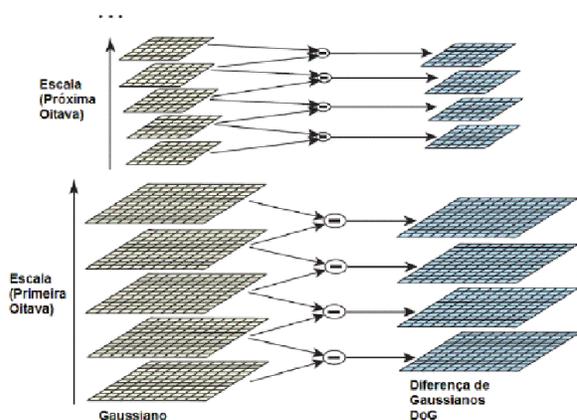
A eficiência na procura dos pontos-chave deve-se à utilização da função Difference of Gaussian (DoG) formada pela diferença de imagens filtradas em escalas próximas, separadas por uma constante de escala  $k$ . A função DoG é definida por

$$DoG = G(x, y, k\sigma) - G(x, y, \sigma) \quad (2.3)$$

O resultado de efetuar a convolução de uma imagem com o filtro DoG é dado por

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (2.4)$$

Esta função DoG suaviza as imagens, e pode ser calculada pela simples subtração de imagens desfocadas por um filtro Gaussiano em escalas  $\sigma$  e  $k\sigma$ . A utilização da função gaussiana tem o objetivo de obter amostras da imagem onde detalhes indesejados e ruídos são eliminados e características fortes realçadas. Variando  $\sigma$  é possível encontrar tais características em diferentes escalas.



**Figura 2.1:** Representação da obtenção das DoG para as oitavas de uma imagem [9]

Um modo eficiente para a criação das DoG é esquematizado na figura 2.1 seguindo-se uma breve descrição dos seus passos.

1. A imagem inicial é transformada com convoluções incrementais com filtros Gaussianos para produzir imagens separadas por um fator de escala  $k$  no espaço de escala, representados na coluna esquerda.
2. Lowe considera que é necessário fazer a convolução da imagem até  $2\sigma$  para ser possível a construção de descritores invariantes quanto à escala. Portanto, para se gerar em  $s$  intervalos, o fator de escala  $k$  é definido por  $k = 21/s$ , produzindo assim  $s + 3$  imagens na oitava de forma que a detecção de extremos cubra toda a oitava.
3. Imagens em escalas adjacentes são subtraídas para produzir as imagens resultantes da DoG mostradas à direita (na figura 2.1). Uma vez processada a oitava, é reduzida a resolução da imagem (*downsample*) tomando-se cada segundo pixel da imagem no centro da oitava, gerando-se uma nova oitava, e voltando-se ao passo número 1.
4. A partir de então, será feita a detecção de extremos em cada intervalo de cada oitava. Um extremo é definido como qualquer valor no DoG maior do que todos os seus vizinhos no espaço-escala.

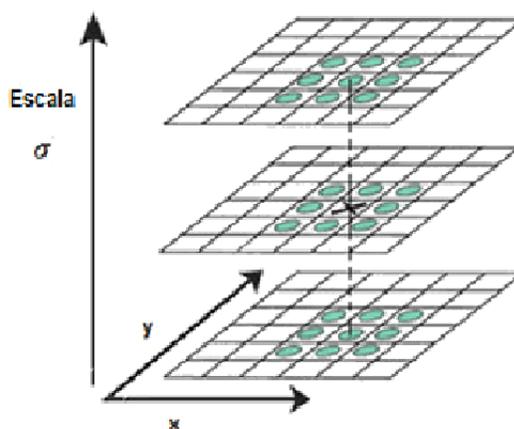


Figura 2.2: Detecção de extremos no espaço-escala [9]

Os extremos são dados por valores de máximo ou mínimo locais para cada  $D(x, y, \sigma)$ , que podem ser obtidos comparando-se a intensidade de cada ponto com as intensidades dos oito vizinhos na escala, com os nove pontos vizinhos na escala superior, e os nove vizinhos na escala inferior, representados na Figura 2.2. Na figura, o ponto marcado com "X" é comparado com seus vizinhos marcados como "O". As 3 imagens DoG apresentadas na figura correspondem à diferença entre imagens adjacentes da pirâmide gaussiana.

A próxima etapa é definir a localização dos pontos-chave e fazer o descarte de pontos instáveis.

### Localização de pontos chave

Todos os pontos detetados como extremos são candidatos a pontos-chave, os quais se deseja calcular a localização exata dos mesmos.

O método consiste em ajustar uma função quadrática 3D do ponto de amostragem local de modo a determinar uma localização interpolada do máximo. Isto é feito utilizando uma expansão de Taylor da DoG aplicada à imagem,  $D(x, y, \sigma)$ , deslocada de modo que a origem desta expansão esteja localizada no ponto de amostragem [3]:

$$D(\bar{x}) = D + \frac{\partial D^T}{\partial \bar{x}} \bar{x} + \frac{1}{2} \bar{x}^T \frac{\partial^2 D}{\partial \bar{x}^2} \bar{x} \dots \quad (2.5)$$

$$\bar{x} = (x, y, \sigma)^T \quad (2.6)$$

Onde o valor de  $D$ , a sua primeira e segunda derivada são calculados no ponto de amostragem e  $x$ , representa o deslocamento deste ponto. A localização em sub-pixéis do ponto de interesse é dada pelo extremo da função apresentada na equação 2.5. Esta localização,  $x$ , é determinada ao calcular-se a derivada de  $D(x)$  em relação a  $x$ , e igualando o resultado a zero:

$$\frac{\partial D}{\partial \bar{x}} + \frac{\partial^2 D}{\partial \bar{x}^2} \hat{x} = 0 \quad (2.7)$$

Obtemos então a posição do extremo, dada por:

$$\hat{x} = -\frac{\partial^2 D^{T-1}}{\partial \bar{x}^2} \frac{\partial^2 D}{\partial \bar{x}} \quad (2.8)$$

O valor da função no extremo,  $D(x)$ , é útil para a rejeição de extremos instáveis com baixo contraste, que seriam sensíveis a ruído. Substituindo-se a equação 2.8 na equação 2.5 obtém-se:

$$D(\hat{x}) = D + \frac{1}{2} \frac{\partial D^T}{\partial \bar{x}} \hat{x} \quad (2.9)$$

É aconselhável segundo Lowe que se rejeitem valores de  $|D(\hat{x})|$  inferiores a um determinado limiar. É aconselhado trabalhar com o valor 0.03 para esse limiar [3] (assumindo-se que os tons de cinza dos pixéis da imagem estejam normalizados em valores entre 0 e 1).

Além do procedimento apresentado para se descartar pontos, Lowe ainda aponta que a função DoG possui resposta "forte" ao longo de arestas, mesmo que a localização ao longo da borda seja mal determinada, ou seja, pontos em arestas poderiam ser escolhidos como pontos de interesse, o que não é desejável. Mas estes pontos podem ser detetados e eliminados, como discutidos a seguir.

A eliminação de pontos chave próximos de arestas é feita usando-se uma matriz Hessiana 2x2,  $H$ , computada na localização e escala dos pontos-chaves na função  $D$ .

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (2.10)$$

Onde  $D_{xx}$  é a derivada de  $D(x, y, \sigma)$  na localização e escala em relação a  $x$  e  $y$ ;  $D_{xx}$  é a segunda derivada em relação a  $x$ ; e  $D_{yy}$  é a segunda derivada em relação a  $y$ .

A Hessiana representa assim uma segunda derivada, permitindo medir as magnitudes das curvaturas de  $D$  a partir dos seus autovalores.

As derivadas são estimadas através das diferenças entre pontos vizinhos à localização e escala definida, e pode ser aproximada por

$$D_{xx} = D(x + 1, y, \sigma) - 2D(x, y, \sigma) + D(x - 1, y, \sigma) \quad (2.11)$$

$$D_{yy} = D(x, y + 1, \sigma) - 2D(x, y, \sigma) + D(x, y - 1, \sigma) \quad (2.12)$$

$$D_{xy} = \left( \begin{array}{l} D(x - 1, y + 1, \sigma) - D(x + 1, y + 1, \sigma) \\ +D(x + 1, y - 1, \sigma) - D(x - 1, y - 1, \sigma) \end{array} \right) / 4 \quad (2.13)$$

Determina-se  $\alpha$ , o valor com maior magnitude, e  $\beta$ , o de menor. Pode-se, então, calcular a soma dos autovalores pelo traço de H e o produto pelo determinante:

$$\text{Tr}(H) = D_{xx} + D_{yy} = \alpha + \beta \quad (2.14)$$

$$\text{Det}(H) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta \quad (2.15)$$

Para o caso em que o determinante for negativo, as curvaturas possuem sinais diferentes, e o ponto é descartado, não sendo considerado um extremo. Sendo  $r$  a razão entre o valor de maior magnitude e o de menor, de modo que  $\alpha = r\beta$ , então

$$\frac{\text{Tr}(H)^2}{\text{Det}(H)} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(r\beta + \beta)^2}{(r\beta)^2} = \frac{(r + 1)^2}{r} \quad (2.16)$$

A equação 2.16 depende apenas da razão entre os autovalores, sendo independente dos valores individuais. O valor de  $(r + 1)/r$  oferece uma medida de quanto os autovalores são distintos, ou seja, é mínimo quando são idênticos e cresce com respeito ao valor de  $r$ . Assim, eliminam-se pontos (indesejáveis) próximos a extremidades descartando-se pontos abaixo de determinado limiar ( $r$ ):

$$\frac{\text{Tr}(H)^2}{\text{Det}(H)} < \frac{(r + 1)^2}{r} \quad (2.17)$$

A equação 2.17 é altamente eficiente de ser computada. Lowe propõe o uso de  $r = 10$ , assim eliminam-se pontos chave que não são estáveis, apesar de estarem próximos de extremidades.

### Atribuição de orientações

Atribuindo uma orientação constante a cada ponto chave baseado em propriedades locais da imagem, o descritor do ponto chave pode ser representado relativamente à sua orientação e portanto apresentar invariância à rotação de imagem.

A escala do ponto chave é usada na seleção da imagem suavizada Gaussiana,  $L$ , com a escala mais próxima, para que todas as computações sejam realizadas de maneira que sejam invariantes à escala. Para cada imagem,  $L(x, y)$ , nesta escala, o gradiente de magnitude,  $m(x, y)$ , e a orientação,  $\theta(x, y)$ , são pré computados usando a diferença de pixels:

$$m(x, y) = \sqrt{\left( (L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2 \right)} \quad (2.18)$$

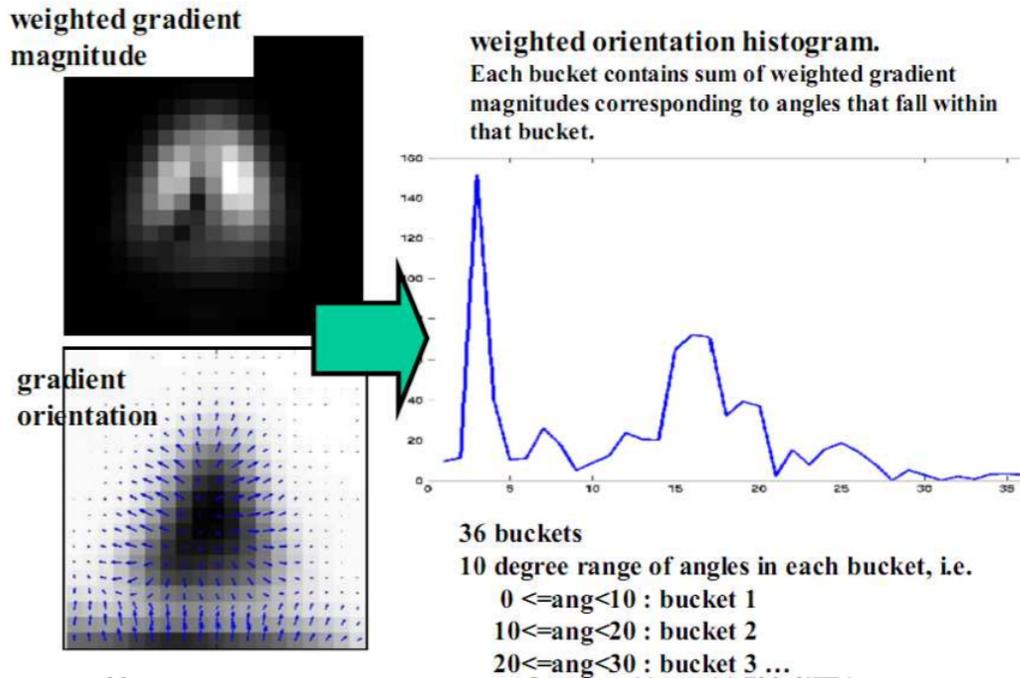
$$\theta(x, y) = \tan^{-1} \left( \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right) \quad (2.19)$$

Um histograma de orientações é formado a partir do gradiente de orientações de pontos amostrais dentro de uma região à volta do ponto chave. O histograma de orientações possui 36 porções, cobrindo 360 graus de orientações. Cada amostra adicionada ao histograma é pesada pela seu gradiente de magnitude e por janela Gaussiana pesada circular com  $\sigma$  de 1.5 vezes a escala do ponto chave. Picos no histograma de orientações correspondem a direções dominantes dos gradientes locais.

O pico mais alto no histograma é detetado, e todos os outros picos locais que atinjam 80% do pico mais alto também são usados para criar um ponto chave com essa orientação. Portanto, para localizações com múltiplos picos de magnitude semelhantes, irão haver múltiplos pontos chaves criadas com a mesma localização e escala, mas diferentes orientações. Apenas cerca de 15% dos pontos atribuídos contém múltiplas orientações, mas estes contribuem significativamente para a estabilidade do reconhecimento. Finalmente uma parábola é ajustada aos 3 valores do histograma mais próximos de cada pico, para interpolar a posição do pico para melhor precisão.

### Construção do descritor local

Nesta secção, será atribuído a cada ponto chave um descritor invariante a iluminação e ponto de vista 3D, tornando-os bem distinguíveis. É importante lembrar que os procedimentos a seguir serão feitos com os valores normalizados em relação à orientação e magnitude de gradiente definidos na secção anterior para cada ponto chave.

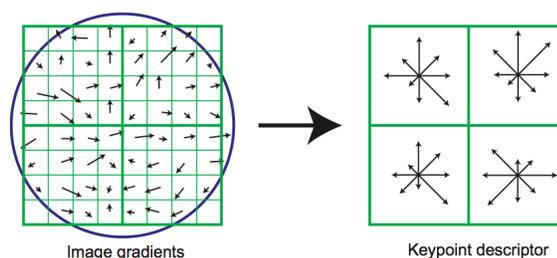


**Figura 2.3:** *Histograma de orientações dominantes dos gradientes [1]*

Para que os descritores tenham invariância à rotação, as orientações dos gradientes destes pontos são giradas de um ângulo correspondente à orientação do ponto-chave definida na secção anterior. O descriptor do ponto-chave é então criado computando-se as magnitudes e orientações dos gradientes que são amostradas ao redor da localização do ponto chave. Este procedimento está mostrado na Figura 2.4, onde os gradientes são representados pelas pequenas setas em cada amostra da localização. São definidas  $n * n$  regiões de amostragem com  $k * k$  pixels cada, ao redor da localização do ponto chave.

Uma função Gaussiana é utilizada para dar peso à magnitude do gradiente em cada ponto na vizinhança do ponto chave, com uma janela de suavização Gaussiana de escala  $\sigma$  igual à metade da largura da janela do descriptor. Esse Gaussiano evita mudanças súbitas do descriptor a pequenas mudanças na posição da janela, e também reduz a ênfase nos gradientes longe do centro do descriptor, que são mais afetados por erros.

Uma vez efetuada a suavização dos gradientes, o descritor consiste de um vetor contendo os valores do histograma. No exemplo da Figura 2.4, o histograma tem 8 valores de orientação, cada um criado ao longo de uma janela de apoio de 4x4 pixels. O vetor característico resultante tem 128 elementos com uma janela de apoio total de 16x16 pixels.



**Figura 2.4:** Construção do descritor para um ponto-chave de 2x2 com 48 elementos [9]

No entanto, duas imagens de um mesmo objeto podem possuir variações de luminosidade que modifiquem sensivelmente os descritores obtidos. Assim, para que o descritor tenha invariância à iluminação, este é normalizado.

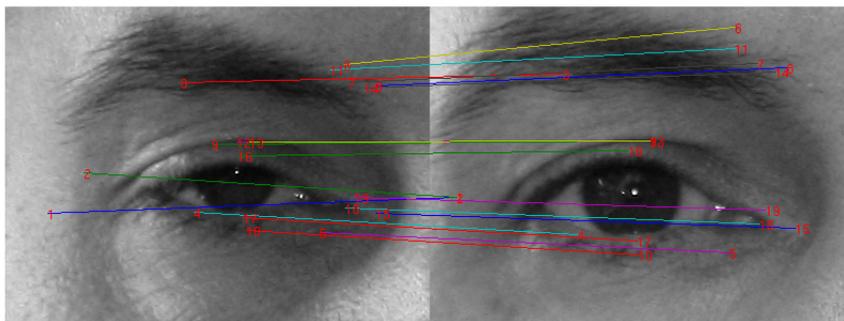
Os descritores são invariantes a mudanças homogêneas de brilho da imagem, uma vez que esta variação representa uma adição a todos os pixels da imagem de uma constante, e os descritores são calculados por diferenças de pixels. Quanto a mudanças homogêneas de contraste, representadas pela multiplicação de todos os pixels por uma constante, elas são corrigidas com a normalização dos descritores.

Variações não-lineares, causadas por saturação das câmaras ou por efeito de iluminação de superfícies tridimensionais em diferentes orientações, podem provocar elevada influência sobre as magnitudes dos descritores, mas com pouca influência na orientação. Reduz-se este efeito impondo um valor máximo às magnitudes. Após a normalização, todos os valores acima de um determinado limiar são ajustados para este limiar. Isto é feito para que direções com magnitude muito grande não dominem a representação do descritor. Lowe sugere usar um limiar de 0,2 [9], isto significa que a correspondência para as grandes magnitudes dos gradientes não é tão importante se comparada com a distribuição das orientações.

Para cada imagem, são construídos diversos descritores, cada um referente a um ponto chave. Tem-se como resultado, portanto, um conjunto de descritores robustos que podem ser usados para fazer a correspondência da imagem em outra imagem, como será detalhado na próxima secção.

### Comparação de imagens

Para comparar imagens diferentes é preciso extrair os pontos chave de cada uma delas, e procurar os pontos correspondentes, como exemplificado na Figura 2.5. A comparação de pontos é baseada nas semelhanças dos descritores correspondentes.



**Figura 2.5:** *Correspondência entre duas imagens*

Quando usamos o SIFT, os pontos de interesse são detetados pelo método e representados por descritores. Os descritores são vetores que podem ser comparados, por exemplo, utilizando-s a distância Euclidiana. Normalmente, os candidatos à melhor correspondência são pontos próximos, de maneira que o melhor candidato é o ponto que apresenta a menor distância Euclidiana.

Lowe utilizou uma modificação do algoritmo Árvore k-d chamado de método de Best-Bin-First (BBF) [2], que pode identificar os vizinhos mais próximos com elevada probabilidade, utilizando apenas uma quantidade limitada de esforço computacional.

O problema da correspondência, desta maneira, é reduzido à procura do vizinho mais próximo. No entanto, alguns pontos instáveis (*outliers*) são detetados ao longo do processo, levando a falsas correspondências.

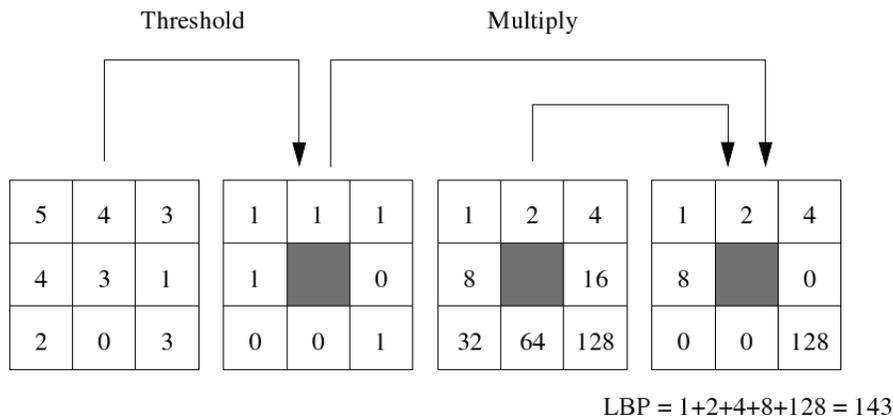
Para a eliminação desse problema, um método para comparar a menor distância com a segunda melhor distância é usado, selecionando somente correspondentes próximos por um limiar (*threshold*) [9].

Lowe rejeitou todas as correspondências em que a relação de distância é superior a 0.8, o que elimina 90% das falsas correlações, porém apenas descartando menos de 5% das correspondências corretas. Portanto, as correspondências são assim eficientemente refinadas, e os falsos pares são descartados.

### 2.1.2 Local Binary Pattern

Este método é utilizado para a extração de características globais tendo sido descrito pela primeira vez por T. Ojala, M. Pietikäinen e D. Harwood [11] [12]. É um método muito eficiente que atribui um valor a cada pixel de uma imagem, através da limirização dos pixels vizinhos considerando o resultado como um número binário. Devido ao poder discriminativo que possui e a sua simplicidade computacional, este operador de texturas tem sido utilizado em diversas aplicações, principalmente como um classificador de características humanas [19].

Ele pode ser visto como uma abordagem unificadora para os modelos tradicionalmente divergentes de análises de texturas. Talvez a propriedade mais importante do operador LBP em aplicações do mundo real é a robustez com mudanças constantes em escala de cinza, por exemplo, provocadas por variações de iluminação.



**Figura 2.6:** Calcular o valor LBP de cada pixel [16]

O operador LBP usa os oito vizinhos de cada pixel, utilizando o valor do pixel central como um limite. Um código LBP é produzido para cada pixel, multiplicando os valores da vizinhança por uma máscara de pesos, e somando-se obtém-se o resultado, como mostra a figura 2.6, resultado que vamos considerar como sendo um número binário. Após termos o resultado em binário de todos os pixels vamos efetuar o histograma com  $2^8 = 256$  valores que vai ser usado como descritor.

Na abordagem com o operador LBP para a classificação de texturas, as ocorrências dos códigos LBP são colecionadas para um histograma. A classificação é depois atribuída pela simples computação de semelhanças entre os histogramas.

Entretanto, se considerarmos uma abordagem semelhante para a representação de resultados de características humanas vamos perder a informação espacial e portanto temos que codificar a textura guardando a localização. Uma maneira de alcançar este objetivo é usar o descritor de texturas LBP para construir vários descritores locais da região periocular e combinar todos num descritor global.

Para esta abordagem vou então basear-me na metodologia proposta para descrição de faces [15] em que a imagem é dividida em regiões locais e são extraídos os descritores LBP de cada região independentemente. Os descritores são depois concatenados e formam um descritor global da face, como mostrado na figura 2.7.

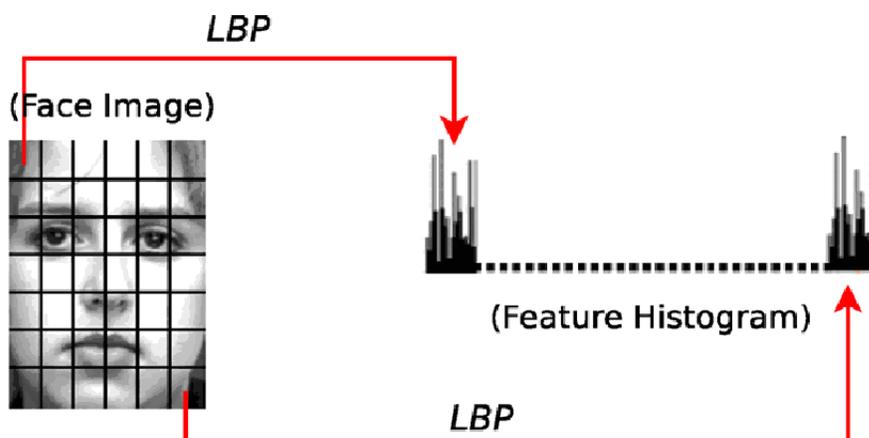


Figura 2.7: Divisão da face em descritores LBP locais [15]

Para aplicarmos esta abordagem à região periocular vamos então dividir a imagem em trinta e cinco regiões locais e para calcular as semelhanças entre imagens diferentes vamos usar a distância euclidiana entre os diferentes descritores LBP.

### 2.1.3 *Histogram of Oriented Gradients*

O método HOG, é um descritor de características utilizado na Visão Computacional e processamento de imagens com o objetivo de detecção e reconhecimento de objetos [4]. Este método é processado numa grelha densa de células com a mesma área e usa uma sobreposição do contraste local normalizado para uma melhor precisão.

O pensamento fundamental por trás dos descritores HOG é que a aparência e forma do objeto local pode ser descrito pela distribuição da intensidade de gradientes ou das direções das arestas. A implementação destes descritores pode ser obtida dividindo a imagem em pequenas regiões ligadas, chamadas células, e para cada célula construímos um histograma de orientações dos gradientes para os pixels dentro da célula.

A combinação destes histogramas representa o descritor. Para melhorar a precisão, os histogramas locais podem ser normalizados pelo contraste, calculando a intensidade através de uma região maior da imagem, chamada de bloco, e então usar este valor para normalizar todas as células dentro do bloco. Esta normalização resulta numa melhor invariância a mudanças na iluminação e sombreamento.

O descritor HOG mantém algumas vantagens importantes sobre os descritor de outros métodos. Como o descritor HOG opera em células localizadas, o método mantém invariância a transformações geométricas e fotométricas, com exceção de orientação a objetos.

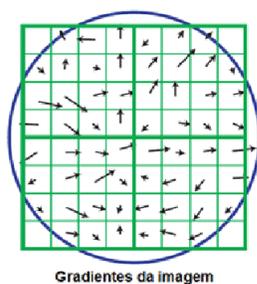


Figura 2.8: *Gradientes de uma célula* [9]

## 2.2 Método proposto

Utilizando a abordagem descrita anteriormente, precisamos de identificar a região periocular de cada imagem, algo que era efetuado manualmente no método base. Para o conseguirmos automaticamente, utilizamos o método descrito por Viola e Jones [18], devido à sua elevada fiabilidade, tanto em termos de tempo de processamento como de taxas de erro.

Após a deteção da região periocular, vamos então fazer a extração de características e a comparação entre diferentes imagens.

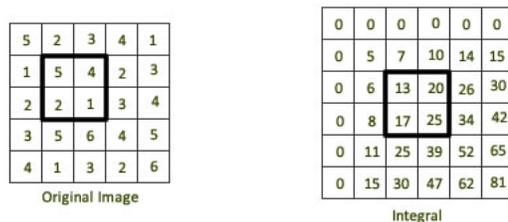
### 2.2.1 *Viola and Jones*

O princípio básico deste algoritmo é analisar uma sub-janela capaz de detetar o objeto pretendido numa determinada imagem. A abordagem padrão para processar a imagem seria redimensionar a imagem para diferentes tamanhos e depois correr o detetor de tamanho fixo pelas imagens resultantes. Esta abordagem revelou-se computacionalmente muito pesada devido ao tempo perdido para efetuar os cálculos nas imagens de diferentes tamanhos.

Contrariamente à abordagem padrão, este método redimensiona o alcance do detetor invés de redimensionar a imagem e corre o detetor várias vezes pela imagem, alterando o tamanho por cada vez. À primeira vista podemos suspeitar que ambas as abordagens vão consumir o mesmo tempo, mas Viola-Jones desenvolveram um detetor invariante à escala que executa o mesmo número de cálculos independentemente do tamanho. Este detetor é construído usando uma imagem chamada "imagem integral" e algumas características retangulares que lembram as *Haar wavelets* [7].

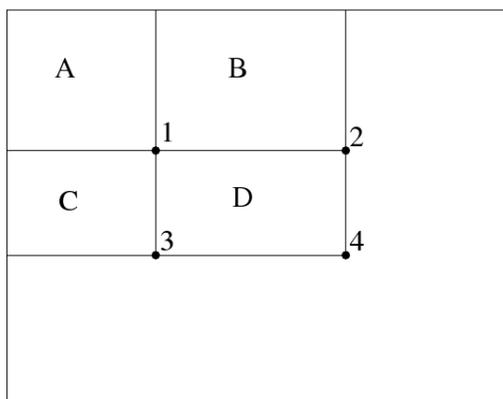
#### **Detetor invariante à escala**

O primeiro passo deste algoritmo consiste em transformar cada imagem numa imagem integral, tornando cada pixel na soma de todos os pixels acima e à esquerda do pixel em questão. Isto é representado na figura 2.9.



**Figura 2.9:** Construção da imagem integral [17]

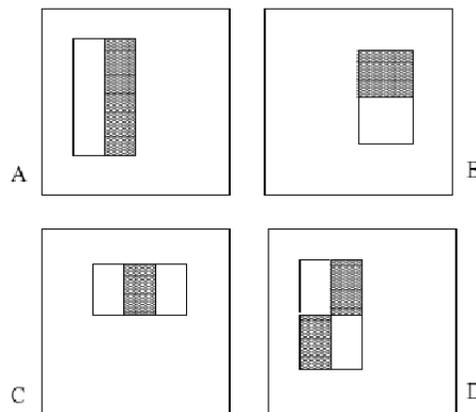
Isto permite o cálculo da soma de todos os pixels dentro de qualquer retângulo usando apenas quatro valores. Estes valores são os pixels na imagem integral que coincidem com os cantos dos retângulos na imagem original. Isto é demonstrado na figura 2.10.



**Figura 2.10:** Cálculo das somas [18]

Na figura 2.10 podemos ver que, para efetuarmos o somatório dos pixels do retângulo  $D$  podemos usar quatro pontos de referência. O valor da imagem integral no ponto 1 é o somatório dos pixels no retângulo  $A$ . O valor no ponto 2 é  $A + B$ , no ponto 3 é  $A + C$  e no ponto 4 é  $A + B + C + D$ . O somatório dos pixels dentro do retângulo  $D$  pode ser efetuado através de  $4 + 1 - (2 + 3)$ .

Foi agora demonstrado como é que o somatório dos pixels dentro dos retângulos de tamanho arbitrário pode ser calculado em tempo constante. O detetor de faces Viola-Jones analisa uma sub-janela usando características consistindo em dois ou mais retângulos. Os diferentes tipos de retângulos são mostrados na figura 2.11.



**Figura 2.11:** *Diferentes tipos de caraterísticas [18]*

Cada caraterística resulta num valor que é calculado fazendo a subtração do somatório dos retângulos brancos pelo somatório dos retângulos pretos. Viola-Jones descobriram que um detetor com uma resolução base de  $24 * 24$  pixels produz resultados satisfatórios, podendo produzir um total de 160.000 caraterísticas diferentes. De seguida vamos tentar estabelecer um padrão nas caraterísticas do objeto que procuramos na imagem e construir inteligentemente uma malha de caraterísticas capaz de detetar o objeto pretendido.

### O algoritmo AdaBoost modificado

Como referido anteriormente podem ser calculadas aproximadamente 160.000 caraterísticas diferentes no detetor com uma resolução base de  $24 * 24$  pixels. Entre estas caraterísticas, algumas vão dar valores consistentemente altos quando estiverem por cima do objeto pretendido. Para encontrar essas caraterísticas Viola-Jones usou uma versão modificada do algoritmo AdaBoost desenvolvido por Freund and Schapire [5].

AdaBoost é um algoritmo para aumentar a aprendizagem automática, capaz de construir um classificador forte através de uma combinação ponderada de classificadores fracos. Um classificador fraco apenas classifica corretamente pouco mais de metade dos casos. Para corresponder esta terminologia com a teoria apresentada, cada característica é considerada como sendo um classificador fraco. Um classificador fraco é matematicamente descrito como:

$$h(x, f, p, \theta) = \begin{cases} 1 & \text{if } pf(x) > p\theta \\ 0 & \text{otherwise} \end{cases} \quad (2.20)$$

Onde  $x$  é a sub-janela de  $24 * 24$  pixels,  $f$  é a característica,  $p$  a polaridade e  $\theta$  é o limite que decide se  $x$  deve ser classificado como positivo ou negativo. Apenas uma pequena quantidade das 160.000 características diferentes são potenciais classificadores fracos, sendo o AdaBoost modificado para escolher as melhores.

Uma parte importante do AdaBoost modificado é a determinação da melhor característica, polaridade e limite. Parece não haver uma solução inteligente para este problema e Viola-Jones sugeriu um método de simples força bruta. Isto quer dizer que a determinação de cada novo classificador fraco envolve avaliar cada característica em todos os exemplos de treino para encontrar a melhor característica em termos de desempenho. Espera-se que isto consuma a maioria do tempo durante o treino.

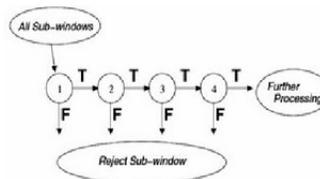
A melhor característica em termos de desempenho é escolhida através do erro ponderado que produz. Este erro pesado é uma função com os pesos pertencentes ao conjunto de treino. O peso de um exemplo corretamente classificado é decrementado e o peso de um exemplo mal classificado mantém-se constante. Como resultado é mais "dispendioso" para a segunda característica (no classificador final) classificar erradamente um exemplo também classificado erradamente pela primeira característica, do que um exemplo classificado corretamente. Uma interpretação alternativa é que a segunda característica é forçada a concentrar-se mais nos exemplos classificados erradamente pela primeira.

### Classificador em cascata

O princípio básico do algoritmo de detecção de faces de Viola-Jones é o de correr o detetor várias vezes pela mesma imagem, cada vez com um tamanho diferente. Mesmo que uma imagem contenha uma ou mais faces é óbvio que seriam avaliadas excessivas sub-janelas que seriam negativas. Esta conclusão levou a uma nova formulação do problema: Em vez de encontrar faces, o algoritmo devia descartar não-faces.

O pensamento por trás deste algoritmo é que é mais rápido descartar não-faces do que encontrar uma face. Com este pensamento, cresceu a necessidade por um classificador em cascata.

Um classificador em cascata é composto por etapas, cada uma contendo um classificador forte. O trabalho de cada etapa é determinar se uma sub-janela é definitivamente uma não-face ou talvez uma face. Quando uma sub-janela é classificada como não-face por uma determinada etapa ela é imediatamente descartada. Quando uma sub-janela é classificada como uma possível face por uma determinada etapa ela passa para a etapa seguinte da cascata. Quanto mais etapas uma sub-janela passar, mais provável é ela conter uma face. O conceito é ilustrado na figura 2.12.



**Figura 2.12:** Demonstração do funcionamento da cascata [6]

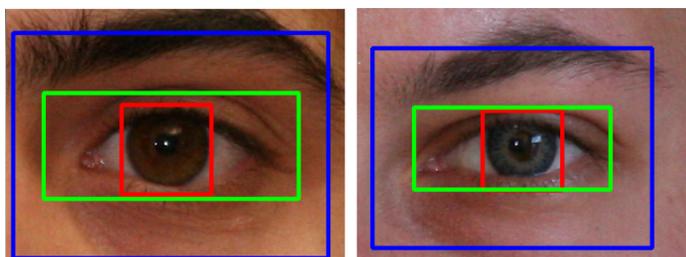
Num classificador com apenas uma etapa aceitamos falsos negativos para reduzir a taxa de falsos positivos. No entanto, para as primeiras etapas num classificador por etapas os falsos positivos não são considerados um problema porque esperamos que as etapas seguintes os descartem. Portanto Viola-Jones aconselha aceitar-se muitos falsos positivos nas etapas iniciais, conseqüentemente o total de falsos negativos na etapa final do classificador deve ser muito pequeno.

## 2.2.2 Extração de características

Para estimarmos a região periocular na imagem usamos dois detectores com cascatas diferentes, uma cascata treinada para detectar a íris e a outra para detectar o olho. Estimamos a região periocular como sendo uma região retangular, tendo como centro do retângulo o centro do olho. Calculamos o tamanho da mesma usando como medida o raio da íris, o comprimento é sete vezes o raio da íris e a altura é cinco vezes o raio da íris.

A detecção da região periocular é representada na figura 2.13, sendo a região contornada a vermelho onde foi detectada a íris, contornada a verde encontra-se a região onde detetou o olho e a região contornada a azul representa a região periocular estimada.

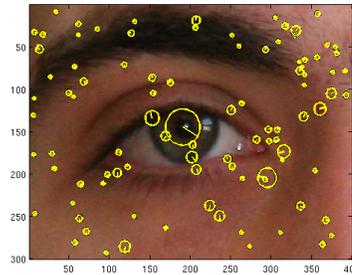
Após a detecção da região periocular, extraiu-se informação através da análise de pontos e texturas resultando num conjunto de características locais e globais.



**Figura 2.13:** Método de detecção: íris (vermelho), olho (verde) e região periocular (azul)

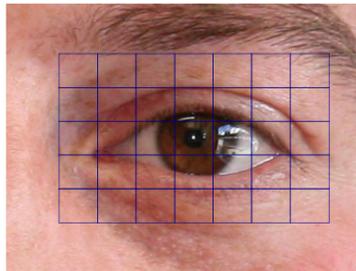
Nesta abordagem são implementados três métodos de extração de características produzindo o conjunto das mesmas, que são usadas para efetuar a comparação entre diferentes imagens, sendo testado também a fusão destas características.

O primeiro método utilizado para extrair as características locais da imagem foi o SIFT, que devolve como características o total de pontos chave encontrados na imagem, sendo cada ponto chave constituído por um descritor, uma orientação e a localização do ponto chave na imagem, como podemos ver na figura 2.14.



**Figura 2.14:** Exemplo de pontos chave detetados

De seguida, utilizou-se um método para extração de características globais, LBP. É um método muito eficiente que atribui um valor a cada pixel de uma imagem, através da limirização dos pixels vizinhos considerando o resultado como um número binário, construindo um histograma da região que analisou. Para o aplicarmos à região periocular vamos então dividir a imagem em trinta e cinco regiões locais, como exemplifica a figura 2.15, concatenando todos os histogramas num só descritor final.



**Figura 2.15:** Exemplo das regiões selecionadas para o HOG e LBP

Por fim, para extrair o mesmo tipo de características do segundo método, utilizamos o método HOG, que é processado numa grelha densa de células com a mesma área e usa uma sobreposição do contraste local normalizado para uma melhor precisão. Para o aplicarmos à região periocular vamos então dividir a imagem nas trinta e cinco regiões locais que dividimos para o método anterior, também concatenando todos os histogramas num só descritor final.

### 2.2.3 Comparação de imagens

Para efetuar a comparação entre duas imagens diferentes foram extraídas características locais e globais de cada imagem, nas locais foram extraídos os pontos chaves através do SIFT e nas globais foram extraídos dois histogramas, um através do LBP e outro através do HOG.

Para fazer a comparação entre os pontos chave encontrados em cada imagem, usamos o esquema de comparação do SIFT, que já foi descrito anteriormente, devolvendo o total de correspondências entre as duas imagens.

Para efetuar a comparação dos histogramas produzidos pelo LBP e pelo HOG, vou basear-me nas semelhanças dos descritores correspondentes. Vou calcular a distância entre os descritores de cada imagem através da distância Euclidiana:

$$d(q, p) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (2.21)$$

Após este processo, ficamos com três classificações diferentes para a comparação entre as duas imagens. Para atribuir uma classificação única, fizemos a fusão das características através da fórmula matemática:

$$\alpha = \beta_1 * i_1 + \beta_2 * i_2 + \beta_3 * i_3 + \beta_4 \quad (2.22)$$

Onde  $\beta$  é o vetor de pesos para atribuir a cada classificação e  $i$  é o vetor com as classificações dos três métodos (SIFT, HOG e LBP). Caso  $\alpha < 0$  a comparação é atribuída à classe *intra*, o que significa que as imagens pertencem à mesma pessoa. Caso contrário, a comparação é atribuída à classe *inter*, tratando-se então de imagens de pessoas diferentes.

## Capítulo 3

# Experiências e Resultados

### 3.1 Conjunto de dados

No decorrer do trabalho, para treinar e testar o método que apresentei foram usadas imagens provenientes da base de dados UBIRIS.v2 [14], fornecidas pelo Professor Doutor Hugo Proença, orientador deste projeto.

O total de imagens fornecido foi de três mil e seiscentas e vinte imagens, tendo sido excluídas muitas destas por não conterem área suficiente para analisar, ou seja, na imagem não se encontrava a totalidade da região periocular, casos como os visualizados na figura 3.1.



**Figura 3.1:** *Exemplo de imagens excluídas*

Com as restantes imagens dividi-as num total de sete conjuntos para efetuar os testes. Estas foram separadas manualmente para representar restrições bem distintas para o sistema e podermos posteriormente obter algumas conclusões.

De seguida, apresentamos os sete conjuntos:

1. *All* - Neste conjunto é exposto um total de oitocentas e cinquenta e nove imagens, criando um conjunto o mais diversificado possível para treino do método proposto, como se pode verificar nas imagens representadas na figura 3.2.



**Figura 3.2:** *Exemplo de imagens do conjunto conjunto ALL*

2. *Frontal* - Este conjunto apresenta duzentas e uma imagens, nas quais o sujeito fixa o olhar na câmara sem qualquer rotação da cabeça, como representado na figura 3.3.



**Figura 3.3:** *Exemplo de imagens do conjunto Frontal*

3. **Hair** - Este conjunto contém um total de trinta e nove imagens, em que existe oclusão da região periocular por parte do cabelo do indivíduo, constando alguns exemplos na figura 3.4.



**Figura 3.4:** Exemplo de imagens do conjunto *Hair*

4. **Rotation** - Conjunto com um total de sessenta e seis imagens, em que as imagens foram obtidas com o sujeito a efetuar uma rotação com a cabeça, como se pode observar na figura 3.5.



**Figura 3.5:** Exemplo de imagens do conjunto *Rotation*

5. **Glasses** - Conjunto composto por duzentas e dezoito imagens, em que todos os sujeitos utilizam óculos, como demonstrado na figura 3.6.



**Figura 3.6:** Exemplo de imagens do conjunto *Glasses*

6. *Temporal* - Este conjunto apresenta cento e quarenta e nove imagens, em que o sujeito não fixa o olhar na câmara, estando o olhar direcionado para o lado temporal, como demonstrado na figura 3.7.



**Figura 3.7:** Exemplo de imagens do conjunto *Temporal*

7. *Nasal* - Conjunto com um total de cento e quarenta e quatro imagens, em que o sujeito não fixa o olhar na câmara, estando o olhar direcionado para o lado nasal, como pode observar-se na figura 3.8.



**Figura 3.8:** Exemplo de imagens do conjunto *Nasal*

## 3.2 Medidas de desempenho

As medidas de desempenho utilizadas neste projeto foram as seguintes:

1. **Area Under Curve (AUC)** - representa a probabilidade de um classificador atribuir a uma amostra aleatória negativa (classe *Intra*) um valor mais baixo que a uma amostra aleatória positiva (classe *Inter*), correspondendo esta probabilidade ao cálculo do integral da área abaixo das curvas ROC. A AUC está relacionada pelo coeficiente de *Gini* ( $G$ ) [8] pela fórmula  $G + 1 = 2AUC$  onde:

$$G_1 = 1 - \sum_{k=1}^n (X_k - X_{k-1})(Y_k + Y_{k-1}) \quad (3.1)$$

2. **Equal Error Rate (EER)** - é a taxa em que as rejeições e as aceitações são iguais, sendo facilmente obtido este valor a partir da curva ROC. Normalmente, o método com o EER mais baixo é o que tem melhor desempenho.
3. **Decidibilidade (DEC)** - esta é definida pela separação entre as médias de duas distribuições de classes diferentes (*Intra* e *Inter*), a dividir pela raiz quadrada da média das respetivas variâncias, representada pela seguinte equação:

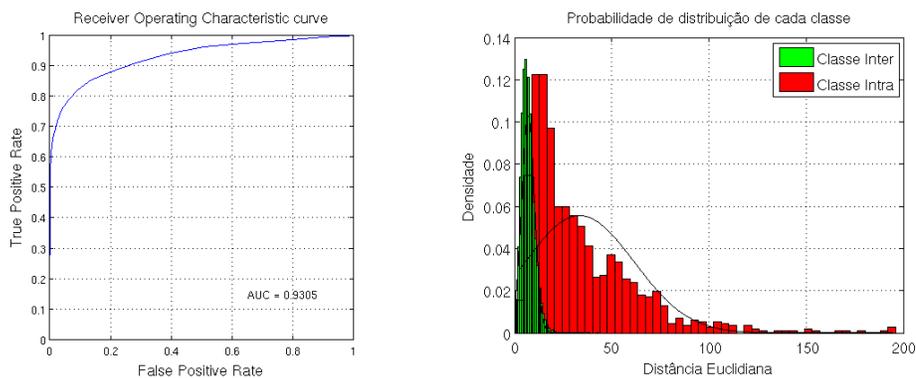
$$Decidibilidade = \frac{\mu_{inter} - \mu_{intra}}{\sqrt{\frac{1}{2}(\sigma_{inter}^2 + \sigma_{intra}^2)}} \quad (3.2)$$

### 3.3 Análise individual de cada método

Nesta secção vamos analisar e testar cada método, utilizando o primeiro conjunto de dados anteriormente descrito, para podermos visualizar como se comporta cada método num conjunto de imagens bastante diversificado.

#### 3.3.1 Resultados do SIFT

Usando o método SIFT para a extração e comparação entre as imagens do conjunto *All*, obtivemos uma AUC de 0.928, com um EER de 0.151 e uma decidibilidade de 1.328, valores que estão representados na figura 3.9.



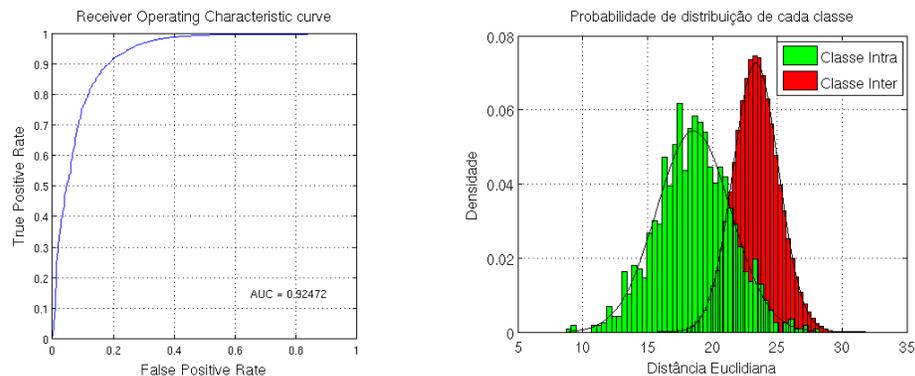
**Figura 3.9:** Curva ROC e distribuição de probabilidades por classe do SIFT

Estes valores surpreenderam-nos bastante devido a não esperarmos uma AUC tão elevada, mas podem facilmente ser explicados com uma boa parametrização do SIFT, resultado de testes com configurações bem distintas. Os valores alterados da configuração predefinida foram o *edge threshold* para 9.6, o *peak threshold* para 0.5 e o *distance ratio* para 1.75.

Com estes resultados podemos analisar que o método SIFT é um bom detetor de pontos chave, obtendo boas comparações a partir dos mesmos, tendo a vantagem de ser invariantes à escala e de terem uma boa tolerância a pequenas rotações.

### 3.3.2 Resultados do HOG

Usando o método HOG para a extração e comparação entre as imagens do conjunto *All*, obtivemos uma AUC de 0.925, com um EER de 0.145 e uma decidibilidade de 2.021, valores que estão representados na figura 3.10.



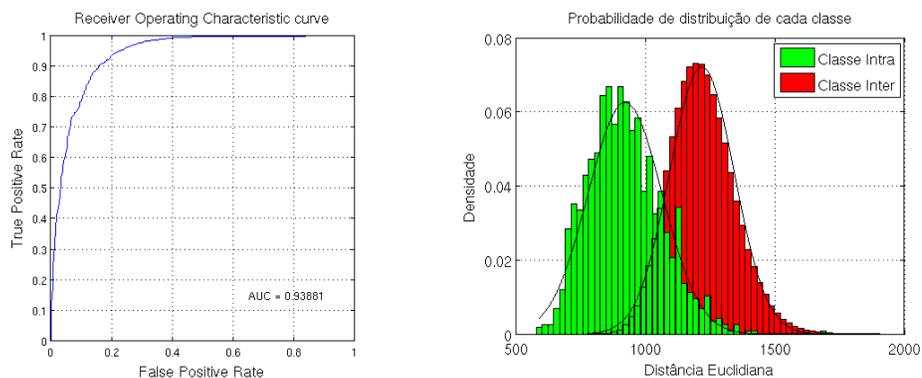
**Figura 3.10:** Curva ROC e distribuição de probabilidades por classe do HOG

Com estes dados podemos confirmar que o HOG é um método descritor de orientações bastante poderoso como já antes referido, fornecendo boas características globais quando aplicado sobre a região periorcular.

A partir destes resultados podemos analisar que o método HOG obteve boas comparações a partir dos histogramas de orientações da região estimada pelo detetor, apesar de ser muito sensível a uma má estimativa por parte do detetor.

### 3.3.3 Resultados do LBP

Usando o método LBP para a extração e comparação entre as imagens do conjunto *All*, obtivemos uma AUC de 0.939, com um EER de 0.133 e uma decidibilidade de 2.209, valores que estão representados na figura 3.11.



**Figura 3.11:** Curva ROC e distribuição de probabilidades por classe do LBP

Com estes dados podemos confirmar que o LBP é um método descritor de texturas bastante poderoso, revelando-se o método com melhores resultados obtidos. Sendo assim, revelou-se como o melhor detetor de características globais.

A partir destes resultados podemos analisar que o método LBP obteve boas comparações a partir dos histogramas extraídos da região estimada pelo detetor, apesar de também revelar-se muito sensível no caso de uma má estimativa por parte do detetor.

### 3.3.4 Cálculo do vetor de pesos

Utilizando a comparação entre duas imagens descrita no método proposto (secção 2.2.3 ), recorreremos à seguinte equação:

$$\alpha = \beta_1 * i_1 + \beta_2 * i_2 + \beta_3 * i_3 + \beta_4 \quad (3.3)$$

Como descrito anteriormente,  $\beta$  é o vetor de pesos atribuído a cada uma das classificações e  $i$  é o vetor com as classificações dos três métodos (SIFT, HOG e LBP). De forma a obtermos os melhores pesos ( $\beta$ ) para cada um dos métodos, foram usadas as características extraídas do conjunto *All*, devido a ser o conjunto de imagens com as características mais diversificadas.

Efetuámos o cálculo do vetor recorrendo a uma regressão linear para encontrar o valor esperado das características e atribuir-lhes o melhor peso, através da equação que se segue:

$$\beta_i = \alpha + \gamma X_i + \epsilon_i \quad (3.4)$$

Em que  $\beta_i$  é a variável dependente que se pretende alcançar,  $\alpha$  é uma constante, que representa a interceção da reta com o eixo vertical,  $\gamma$  é outra constante, que representa o declive da reta,  $X_i$  é a variável independente, que representa o fator explicativo na equação e  $\epsilon_i$  é a variável que inclui todos os fatores residuais mais os possíveis erros de medição.

O comportamento da mesma é aleatório, devido à natureza dos fatores que encerra. Para que a equação possa ser aplicada, os erros devem satisfazer determinadas hipóteses, serem variáveis normais, com a mesma variância  $\sigma^2$  (desconhecida), independentes e independentes da variável explicativa  $X$ .

Para calcular os fatores  $\alpha$  e  $\gamma$  usamos as equações 3.5 e 3.6.

$$\hat{\alpha} = \frac{\sum X^2 \sum Y - \sum (XY) \sum X}{n \sum X^2 - (\sum X)^2} \quad (3.5)$$

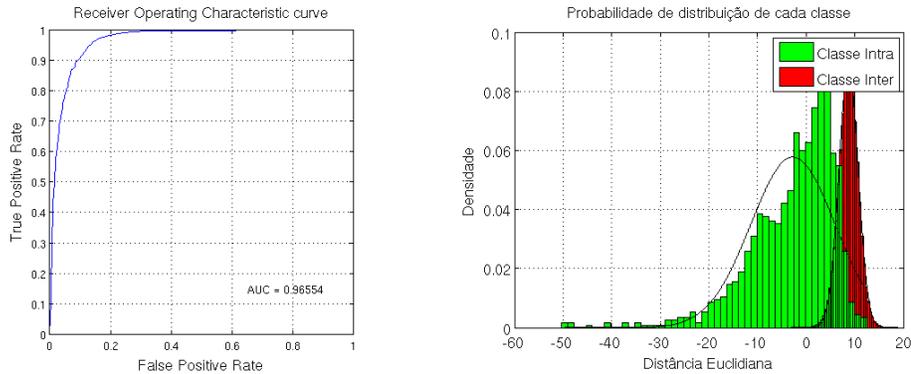
$$\hat{\gamma} = \frac{n \sum (XY) - \sum X \sum Y}{n \sum X^2 - (\sum X)^2} \quad (3.6)$$

Definindo  $\bar{X} = \frac{\sum X}{n}$  e  $\bar{Y} = \frac{\sum Y}{n}$ , temos que  $\hat{\alpha}$  e  $\hat{\gamma}$  se relacionam pela seguinte equação

$$\hat{\alpha} = \bar{Y} - \hat{\gamma} \bar{X} \quad (3.7)$$

### 3.3.5 Resultados da fusão dos métodos

Usando a fusão dos métodos para a extração e comparação entre as imagens do conjunto *All*, obtivemos uma AUC de 0.965, com um EER de 0.097 e uma decidibilidade de 1.810, valores que estão representados na figura 3.12.



**Figura 3.12:** Curva ROC e distribuição de probabilidades por classe da fusão

Podemos constatar com a análise destes dados que a fusão dos três métodos apresenta um grande aumento na AUC, piorando a decibilidade em relação aos HOG e LBP e diminuindo bastante o EER, melhorando significativamente os resultados obtidos pelos métodos individualmente.

Para explicar uma melhoria tão significativa no método proposto, vamos analisar a correlação entre os três métodos através dos coeficiente de correlação de *Pearson*, que mede a correlação e a direção da mesma (negativa ou positiva) entre duas variáveis.

Este coeficiente, normalmente representado por  $\rho$ , assume valores entre  $-1$  (correlação perfeita negativa, quando uma aumenta a outra diminui) e  $1$  (correlação perfeita positiva, aumenta as duas em simultâneo). Quando o coeficiente é próximo de  $0$ , significa que as variáveis estão pouco correlacionadas.

Quando combinamos métodos diferentes, procuramos obter uma baixa correlação entre as variáveis. O coeficiente é calculado através da seguinte equação:

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X) \cdot \text{var}(Y)}} \quad (3.8)$$

onde  $x_1, x_2, \dots, x_n$  e  $y_1, y_2, \dots, y_n$  são os valores medidos de ambas as variáveis.

Calculamos  $\bar{x}$  e  $\bar{y}$ , média aritmética das variáveis, através da equação 3.9 e 3.10, respetivamente.

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i \quad (3.9)$$

$$\bar{y} = \frac{1}{n} \cdot \sum_{i=1}^n y_i \quad (3.10)$$

Calculando a tabela de correlação dos três métodos para o conjunto *All*, o primeiro conjunto de imagens descrito anteriormente, obtivemos a tabela de correlação 3.1, que consiste no coeficiente da correlação de *Pearson*, entre os resultados obtidos em cada um dos métodos.

	SIFT	HOG	LBP
SIFT	1.000	-0.193	-0.152
HOG	-0.193	1.000	0.544
LBP	-0.152	0.544	1.000

**Tabela 3.1:** Tabela de correlação entre os métodos

Analisando a tabela 3.1 podemos concluir que existe pouca correlação entre as diferentes características, principalmente entre as resultantes do SIFT e os dois outros métodos (HOG e LBP), o que permitiu melhorar significativamente os resultados em relação a cada método. Podemos visualizar as melhorias obtidas com a fusão através da tabela 3.2.

Método	AUC	EER	DEC
Scale Invariant Feature Transform	0.928	0.151	1.328
Histogram of Oriented Gradients	0.925	0.145	2.021
Local Binary Pattern	0.939	0.133	2.209
Fusão dos métodos	0.965	0.097	1.810

**Tabela 3.2:** Resultados dos diferentes métodos

Vamos de seguida analisar os melhores e os piores resultados obtidos para cada uma das classes, verificando assim as vantagens e desvantagens deste método proposto (secção 2.2.3) e examinar os fatores de variabilidade que mais influenciam os resultados.

As classes existentes são as classes *Intra* (quando a comparação efetuada é entre imagens da mesma pessoa) e *Inter* (quando a comparação efetuada é entre imagens de pessoas diferentes).

Dispomos as imagens resultantes como mostrado na figura 3.13 para melhor visualização e compreensão de cada coeficiente ( $\alpha$ ) do método proposto obtido.

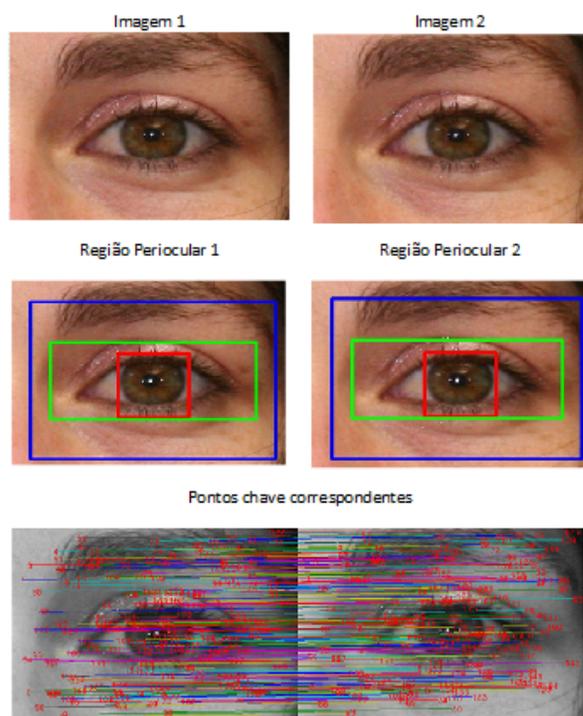
Imagem 1	Imagem 2
Deteções do Viola and Jones na imagem 1	Deteções do Viola and Jones na imagem 2
Pontos chave correspondentes entre as duas imagens	

**Figura 3.13:** Disposição das melhores e piores comparações de cada uma das classes

Onde a *Imagem 1* e a *Imagem 2* são as duas imagens em que foram extraídas características e foi calculado o coeficiente ( $\alpha$ ) do método proposto, *Deteções do Viola and Jones na Imagem 1* e *Deteções do Viola and Jones na Imagem 2* são figuras que representam as regiões detetadas pelo método Viola and Jones (secção 2.2.1) e *Pontos chaves correspondentes entre as duas imagens* equivale a uma figura a ilustrar as correspondências de pontos chave que o método SIFT devolveu para a comparação representada.

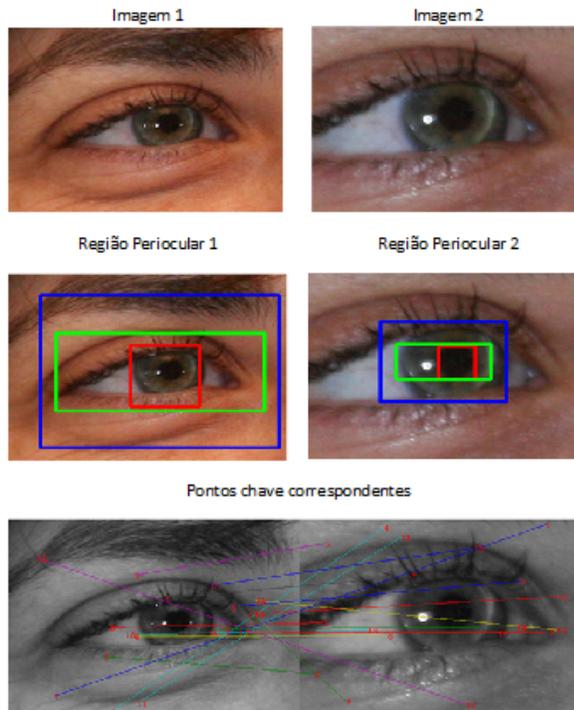
No método de extração de características proposto (secção 2.2.2), o método SIFT é aplicado sobre toda a imagem para localizar pontos chave e os restantes métodos (HOG e LBP) são aplicados apenas na região periorcular reconhecida pelo detetor.

### Melhor resultado da classe Intra



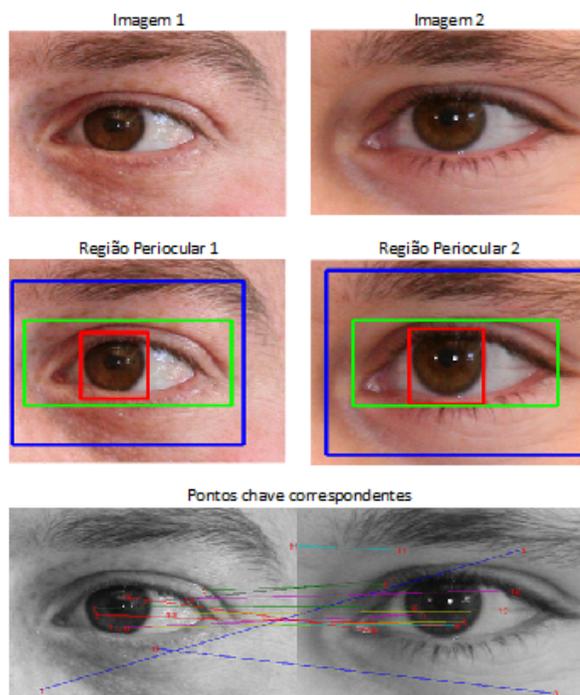
**Figura 3.14:** Melhor resultado da classe Intra para o conjunto All

A figura 3.14 representa a melhor comparação entre imagens diferentes da mesma pessoa no conjunto *All*, efetuada através do método de comparação proposto (secção 2.2.3), tendo um coeficiente negativo ( $\alpha = -50.308674$ ) distante de 0, obtendo também o melhor coeficiente em todos os conjuntos de dados na classe *Intra*. Analisando a figura 3.14 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas regiões muito semelhantes que permitiram um bom desempenho dos métodos HOG e LBP. Também efetuou bastantes correspondências nos pontos chave como podemos observar. Com este coeficiente ( $\alpha = -50.308674$ ) a comparação representada seria corretamente atribuída à classe que pertence, *Intra*, significando que o sistema reconhece que estamos a comparar duas imagens da mesma pessoa.

**Pior resultado da classe Intra**

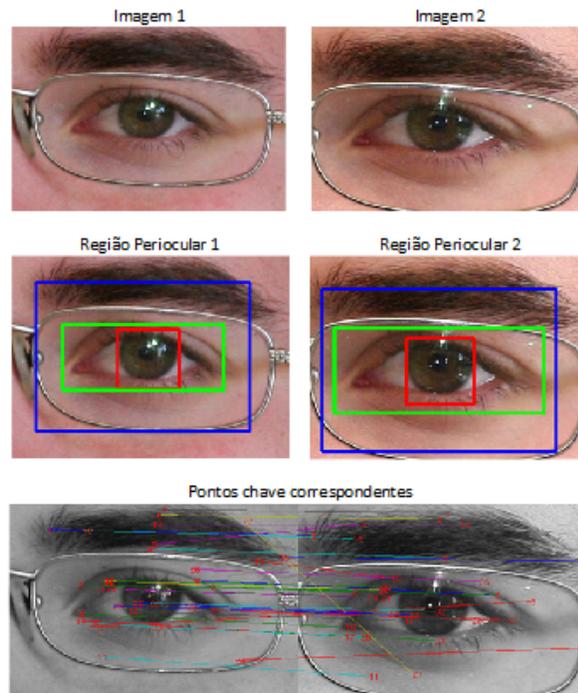
**Figura 3.15:** *Pior resultado da classe Intra para o conjunto All*

A comparação representada na figura 3.15, efetuada através do método de comparação proposto (secção 2.2.3), teve um coeficiente positivo ( $\alpha = 12.471781$ ) distante de 0, representando o pior resultado obtido quando comparamos imagens diferentes da mesma pessoa no *conjunto All*. Analisando a figura 3.15 podemos visualizar que o detetor efetuou uma má estimativa da região periocular na segunda imagem da comparação, indicando a região periocular na imagem quando ela não está contida na mesma, o que torna impossível os restantes métodos obterem o resultado esperado para esta classificação. Com este coeficiente ( $\alpha = 12.471781$ ) a comparação representada seria erradamente atribuída à classe *Inter*, significando que o sistema não reconhece que estamos a comparar duas imagens da mesma pessoa.

**Melhor resultado da classe Inter**

**Figura 3.16:** Melhor resultado da classe *Inter* para o conjunto *All*

A figura 3.16 representa a melhor comparação entre imagens de pessoas diferentes no conjunto *All*, efetuada através do método de comparação proposto (secção 2.2.3), tendo um coeficiente positivo ( $\alpha = 2.792231$ ) ligeiramente distante de 0. Analisando a figura 3.16 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas boas regiões que permitiram um bom desempenho dos métodos HOG e LBP. Também efetuou poucas correspondências nos pontos chave como podemos observar. Com este coeficiente ( $\alpha = 2.792231$ ) a comparação representada seria corretamente atribuída à classe que pertence, *Inter*, significando que o sistema reconhece que estamos a comparar duas imagens de pessoas diferentes.

**Pior resultado da classe Inter**

**Figura 3.17:** *Pior resultado da classe Inter para o conjunto All*

A comparação representada na figura 3.17, efetuada através do método de comparação proposto (secção 2.2.3), teve um coeficiente negativo ( $\alpha = -3.155713$ ) distante de 0, representando o pior resultado obtido quando comparamos imagens de diferentes pessoas no *conjunto All*. Analisando a figura 3.17 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas boas regiões, mas que obtiveram um mau desempenho dos métodos HOG e LBP, muito provavelmente devido à presença de óculos na região periocular. Com este coeficiente ( $\alpha = -3.155713$ ) a comparação representada seria erradamente atribuída à classe *Intra*, significando que o sistema não reconhece que estamos a comparar duas imagens de pessoas diferentes.

### 3.4 Testes ao método proposto

Como o conjunto *All* já foi testado anteriormente e estando o vetor de pesos otimizado para o mesmo, vamos agora analisar a eficácia do método de comparação proposto (secção 2.2.3) utilizando os restantes conjuntos de dados, verificando o comportamento do mesmo sobre diferentes condições.

Com os conjuntos de dados anteriormente referidos (secção 3.1) procurámos representar diversos fatores de variabilidade comuns quando obtemos imagens da região periocular, representando restrições bem distintas para a correta classificação de cada comparação entre duas imagens diferentes.

As imagens foram separadas manualmente para representar restrições bem distintas para o sistema e poderemos posteriormente obter algumas conclusões.

Para analisarmos cada conjunto iremos visualizar o gráfico da curva ROC que representa a AUC dos quatro métodos implementados (SIFT, HOG, LBP e fusão dos três métodos anteriores).

Ao lado da curva ROC visualizaremos o histograma da probabilidade de distribuição dos coeficientes obtidos ( $\alpha$ ) de cada uma das classes (*Intra* e *Inter*) através do método de comparação proposto (secção 2.2.3).

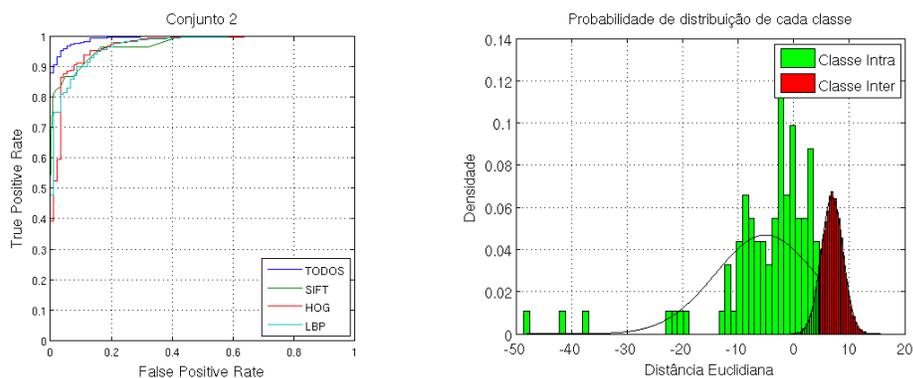
Para melhor visualizarmos os resultados obtidos por cada método vamos representar numa tabela a AUC, o EER e a decidibilidade, permitindo comparar mais facilmente o desempenho individual de cada método.

Concluindo a análise dos conjuntos de dados, vamos mostrar os melhores e os piores resultados obtidos em cada conjunto, verificando assim as vantagens e desvantagens deste método proposto (secção 2.2.3) e examinar os fatores de variabilidade que mais influenciam os resultados.

Dispomos as diversas imagens resultantes das comparações como anteriormente (secção 3.3.5) para melhor visualização e compreensão de cada coeficiente obtido.

### 3.4.1 Análise ao conjunto *Frontal*

Usando a fusão dos métodos para a extração e comparação entre as imagens do conjunto *Frontal*, obtivemos uma AUC de 0.994, com um EER de 0.044 e uma decidibilidade de 1.826, valores que estão representados na figura 3.18.



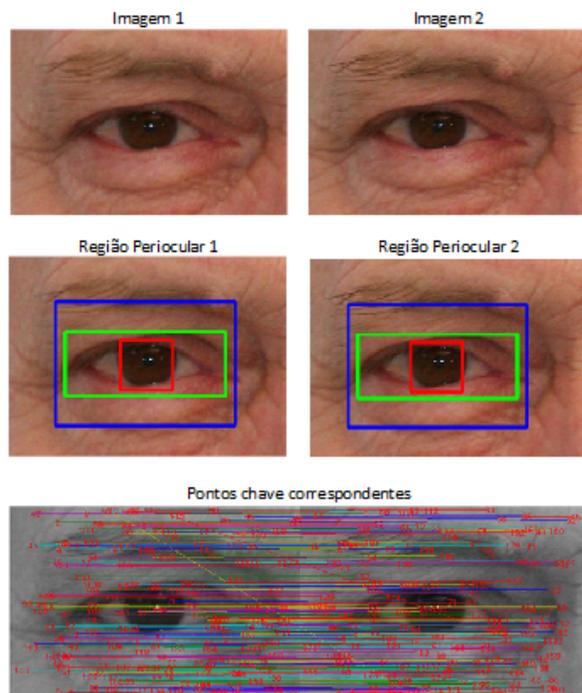
**Figura 3.18:** Curva ROC e distribuição de probabilidades do conjunto *Frontal*

Podemos observar que este conjunto obteve excelentes resultados, algo que já esperávamos por tratar-se de um conjunto de imagens com muitos poucos fatores de variabilidade. Vamos verificar o comportamento de cada um dos métodos individualmente na tabela 3.3.

Método	AUC	EER	DEC
Scale Invariant Feature Transform	0.973	0.103	1.299
Histogram of Oriented Gradients	0.970	0.091	2.623
Local Binary Pattern	0.970	0.096	2.570
Fusão dos métodos	0.994	0.044	1.827

**Tabela 3.3:** Resultados dos diferentes métodos no conjunto *Frontal*

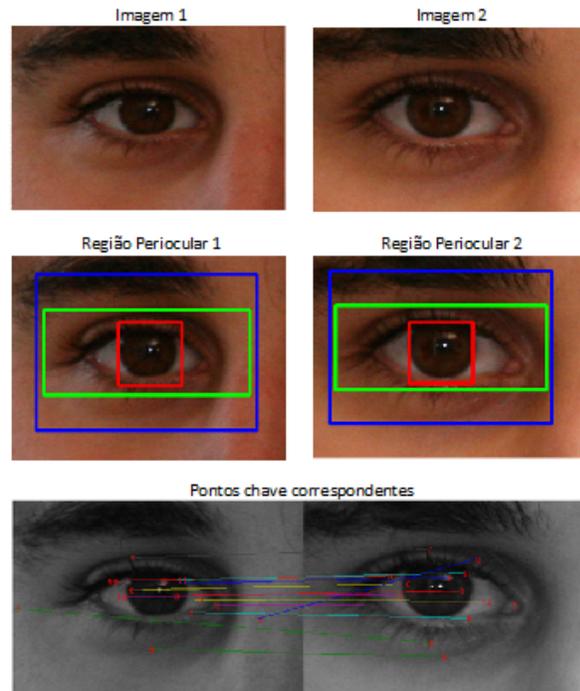
Baseando nestes resultados, vamos individualmente analisar os melhores e os piores resultados obtidos para cada uma das classes.

**Melhor resultado da classe Intra**

**Figura 3.19:** *Melhor resultado da classe Intra para o conjunto Frontal*

A figura 3.19 representa a melhor comparação entre imagens diferentes da mesma pessoa no conjunto *Frontal*, efetuada através do método de comparação proposto (secção 2.2.3), tendo um coeficiente negativo ( $\alpha = -48.576363$ ) distante de 0. Analisando a figura 3.19 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas regiões muito semelhantes. Tendo o sujeito características muito salientes, podemos visualizar que todos os métodos tiveram bons resultados. Com este coeficiente ( $\alpha = -48.576363$ ) a comparação representada seria corretamente atribuída à classe que pertence, *Intra*, significando que o sistema reconhece que estamos a comparar duas imagens da mesma pessoa.

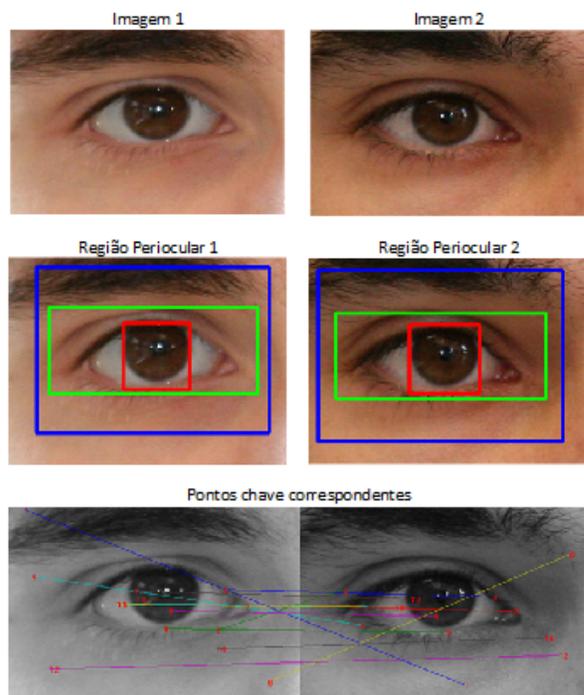
## Pior resultado da classe Intra



**Figura 3.20:** *Pior resultado da classe Intra para o conjunto Frontal*

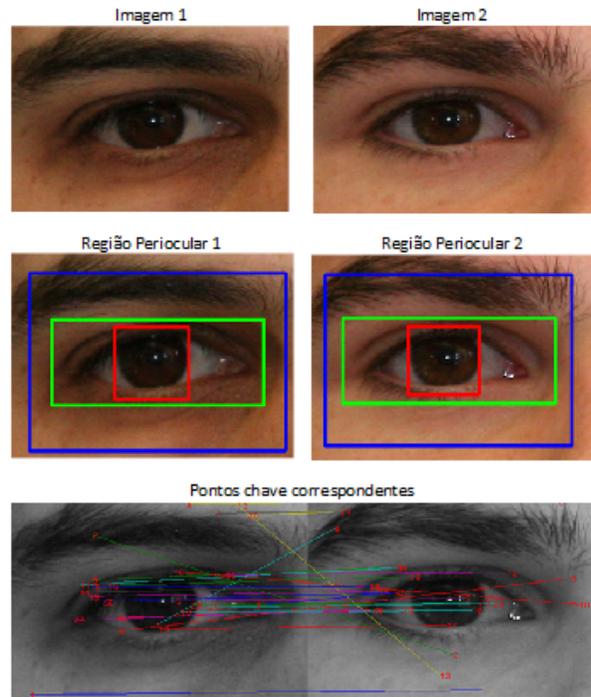
A comparação representada na 3.20, efetuada através do método de comparação proposto (secção 2.2.3), teve um coeficiente positivo ( $\alpha = 4.616366$ ) distante de 0, representando o pior resultado obtido quando comparamos imagens diferentes da mesma pessoa no conjunto *Frontal*. Analisando a figura 3.20 podemos visualizar que o detetor efetuou uma estimativa aceitável da região periocular, mas gerando nas duas imagens a ser comparadas regiões pouco semelhantes devido ao raio da íris ser muito elevado na primeira imagem da comparação, não permitindo um bom desempenho dos métodos HOG e LBP. Podemos também verificar que obtivemos poucas correspondências nos pontos chave. Com este coeficiente ( $\alpha = 4.616366$ ) a comparação representada seria corretamente atribuída à classe que pertence, *Intra*, significando que o sistema reconhece que estamos a comparar duas imagens da mesma pessoa.

### Melhor resultado da classe Inter



**Figura 3.21:** Melhor resultado da classe *Inter* para o conjunto *Frontal*

A figura 3.21 representa a melhor comparação entre imagens de pessoas diferentes no conjunto *Frontal*, efetuada através do método de comparação proposto (secção 2.2.3), tendo um coeficiente positivo ( $\alpha = 1.791219$ ) ligeiramente distante de 0. Analisando a figura 3.21 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas boas regiões que permitiram um bom desempenho dos métodos HOG e LBP. Também efetuou poucas correspondências nos pontos chave como podemos observar. Com este coeficiente ( $\alpha = 1.791219$ ) a comparação representada seria corretamente atribuída à classe que pertence, *Inter*, significando que o sistema reconhece que estamos a comparar duas imagens de pessoas diferentes.

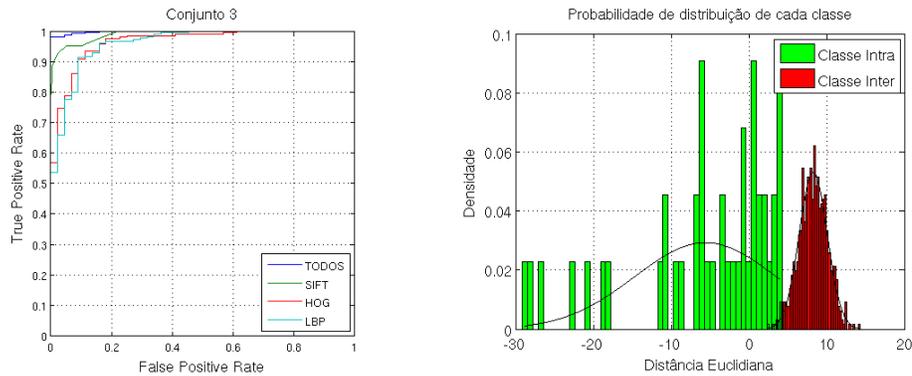
**Pior resultado da classe Inter**

**Figura 3.22:** *Pior resultado da classe Inter para o conjunto Frontal*

A comparação representada na 3.22, efetuada através do método de comparação proposto (secção 2.2.3), teve um coeficiente negativo ( $\alpha = -0.325078$ ) distante de 0, representando o pior resultado obtido quando comparamos imagens de diferentes pessoas no conjunto *Frontal*. Analisando a figura 3.22 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas boas regiões, mas que obtiveram um desempenho pouco satisfatório dos métodos HOG e LBP, efetuando também muitas correspondências nos pontos chaves, o que resultou num coeficiente ( $\alpha$ ) ligeiramente menor que 0. Com este coeficiente ( $\alpha = -0.325078$ ) a comparação representada seria erradamente atribuída à classe *Intra*, significando que o sistema não reconhece que estamos a comparar duas imagens de pessoas diferentes.

### 3.4.2 Análise ao conjunto *Hair*

Usando a fusão dos métodos para a extração e comparação entre as imagens do conjunto *Hair*, obtivemos uma AUC de 0.999, com um EER de 0.014 e uma decidibilidade de 2.104, valores que estão representados na figura 3.23.



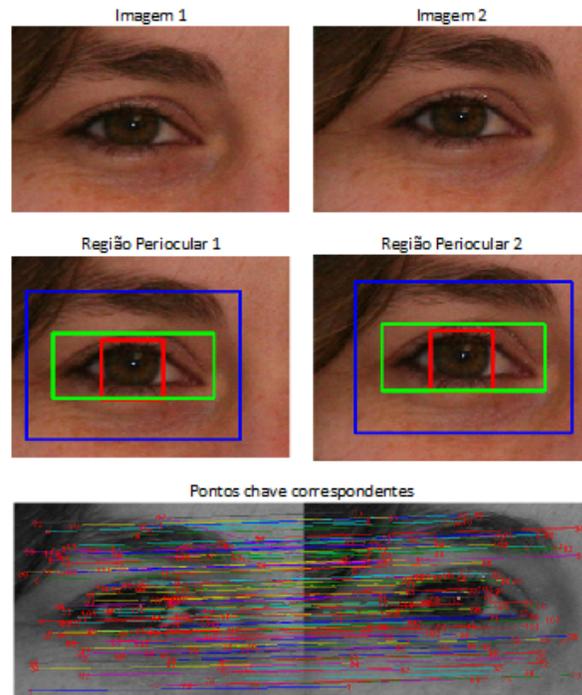
**Figura 3.23:** Curva ROC e distribuição de probabilidades do conjunto *Hair*

Podemos observar que este conjunto obteve excelentes resultados, mas que infelizmente temos que considerar pouco significativos devido ao tamanho da amostra. Outro fator que influenciou os resultados foi que o fator de variabilidade (cabelo) que queríamos testar e analisar encontrar-se poucas vezes na região periocular estimada pelo detetor. Com estes fatos em consideração, vamos verificar o comportamento de cada um dos métodos individualmente na tabela 3.4.

Método	AUC	EER	DEC
Scale Invariant Feature Transform	0.991	0.049	1.762
Histogram of Oriented Gradients	0.966	0.097	2.068
Local Binary Pattern	0.961	0.091	2.424
Fusão dos métodos	0.999	0.014	2.104

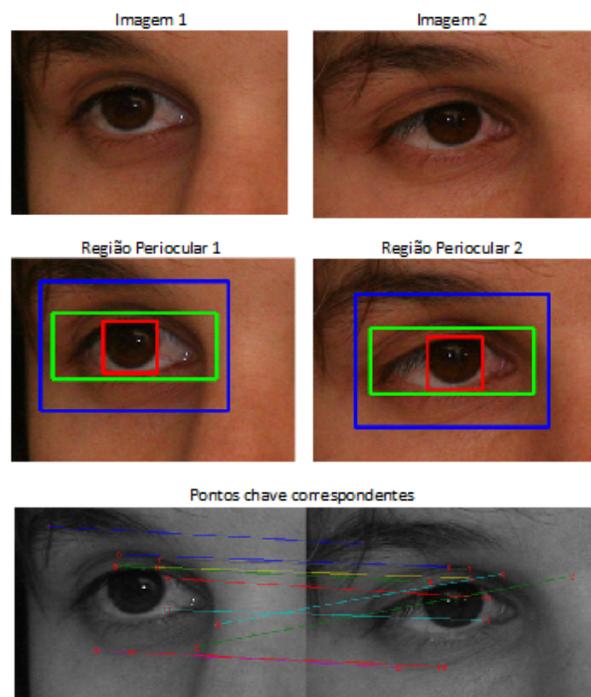
**Tabela 3.4:** Resultados dos diferentes métodos no conjunto *Hair*

Baseando nestes resultados, vamos individualmente analisar os melhores e os piores resultados obtidos para cada uma das classes.

**Melhor resultado da classe Intra**

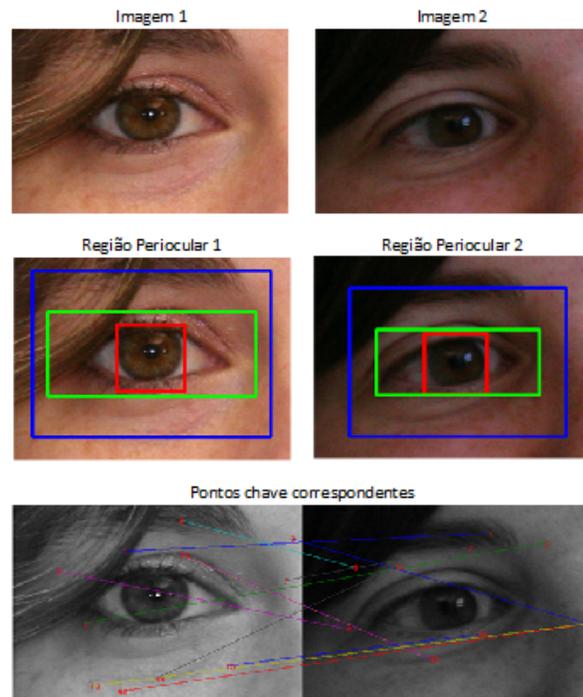
**Figura 3.24:** *Melhor resultado da classe Intra para o conjunto Hair*

A figura 3.24 representa a melhor comparação entre imagens diferentes da mesma pessoa no conjunto *Hair*, efetuada através do método de comparação proposto (secção 2.2.3), tendo um coeficiente negativo ( $\alpha = -29.083982$ ) distante de 0. Analisando a figura 3.24 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas regiões muito semelhantes, contendo as duas uma região de cabelo muito semelhante, que permitiram um bom desempenho dos métodos HOG e LBP. Também efetuou bastantes correspondências nos pontos chave como podemos observar. Com este coeficiente ( $\alpha = -29.083982$ ) a comparação representada seria corretamente atribuída à classe que pertence, *Intra*, significando que o sistema reconhece que estamos a comparar duas imagens da mesma pessoa.

**Pior resultado da classe Intra**

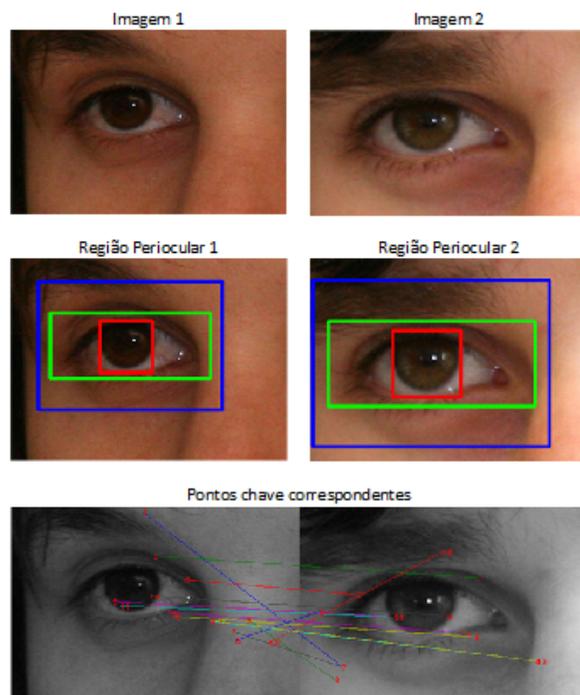
**Figura 3.25:** *Pior resultado da classe Intra para o conjunto Hair*

A comparação representada na 3.25, efetuada através do método de comparação proposto (secção 2.2.3), teve um coeficiente positivo ( $\alpha = 4.319599$ ) distante de 0, representando o pior resultado obtido quando comparamos imagens diferentes da mesma pessoa no conjunto *Hair*. Analisando a figura 3.25 podemos visualizar que o detetor efetuou uma estimativa aceitável da região periocular, mas que devido a uma rotação da cabeça, os métodos HOG e LBP obtiveram maus resultados. Podemos então presumir que esta comparação obteve um resultado mau devido a um fator que pretendíamos analisar noutro conjunto de dados (rotação da cabeça). Com este coeficiente ( $\alpha = 4.319599$ ) a comparação representada seria erradamente atribuída à classe *Inter*, significando que o sistema não reconhece que estamos a comparar duas imagens da mesma pessoa.

**Melhor resultado da classe Inter**

**Figura 3.26:** Melhor resultado da classe *Inter* para o conjunto *Hair*

A figura 3.26 representa a melhor comparação entre imagens de pessoas diferentes no conjunto *Hair*, efetuada através do método de comparação proposto (secção 2.2.3), um coeficiente positivo ( $\alpha = 5.525783$ ) distante de 0, tendo sido o melhor resultado obtido de todos os conjuntos. Analisando a figura 3.26 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas boas regiões que permitiram um bom desempenho dos métodos HOG e LBP. Efetuou poucas correspondências nos pontos chave e podemos verificar que nesta imagem o fator de variabilidade pretendido (cabelo) contribuiu para diferenciar as duas imagens. Com este coeficiente ( $\alpha = 5.525783$ ) a comparação representada seria corretamente atribuída à classe que pertence, *Inter*, significando que o sistema reconhece que estamos a comparar duas imagens de pessoas diferentes.

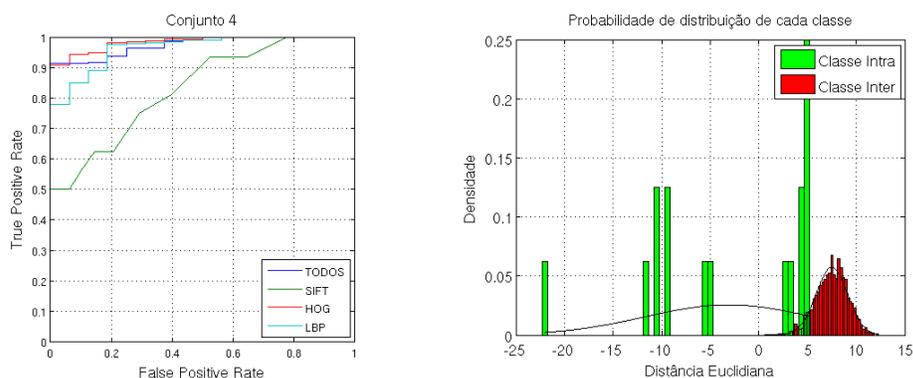
**Pior resultado da classe Inter**

**Figura 3.27:** *Pior resultado da classe Inter para o conjunto Hair*

A comparação representada na 3.27, efetuada através do método de comparação proposto (secção 2.2.3), teve um coeficiente positivo ( $\alpha = 2.299124$ ) distante de 0, representando o pior resultado obtido quando comparamos imagens de diferentes pessoas no conjunto *Hair*. Analisando a figura 3.27 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas boas regiões, mas que obtiveram um mau desempenho dos métodos HOG e LBP, muito provavelmente devido a um fator que pretendíamos analisar noutro conjunto de dados (rotação da cabeça). Com este coeficiente ( $\alpha = -3.155713$ ) a comparação representada seria erradamente atribuída à classe *Intra*, significando que o sistema não reconhece que estamos a comparar duas imagens de pessoas diferentes.

### 3.4.3 Análise ao conjunto *Rotation*

Usando a fusão dos métodos para a extração e comparação entre as imagens do conjunto *Rotation*, obtivemos uma AUC de 0.976, com um EER de 0.088 e uma decidibilidade de 1.712, valores que estão representados na figura 3.28.



**Figura 3.28:** Curva ROC e distribuição de probabilidades do conjunto *Rotation*

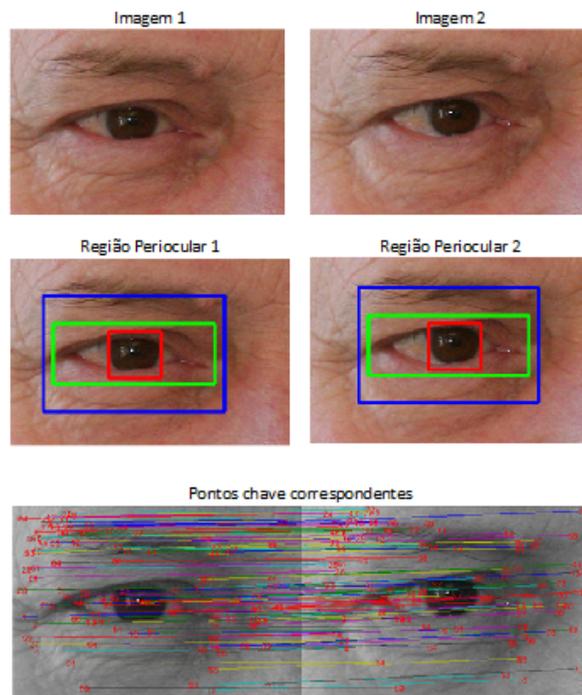
Podemos observar que este conjunto obteve bons resultados, mas que infelizmente temos que considerar pouco significativos devido ao tamanho da amostra. Muitas das imagens também tinham pouca rotação da cabeça, o que melhorou os resultados deste conjunto. Com estes fatos em consideração, vamos verificar o comportamento de cada um dos métodos individualmente na tabela 3.5.

Método	AUC	EER	DEC
Scale Invariant Feature Transform	0.832	0.292	1.259
Histogram of Oriented Gradients	0.985	0.062	2.495
Local Binary Pattern	0.965	0.125	2.397
Fusão dos métodos	0.976	0.086	1.713

**Tabela 3.5:** Resultados dos diferentes métodos no conjunto *Rotation*

Baseando nestes resultados, vamos individualmente analisar os melhores e os piores resultados obtidos para cada uma das classes.

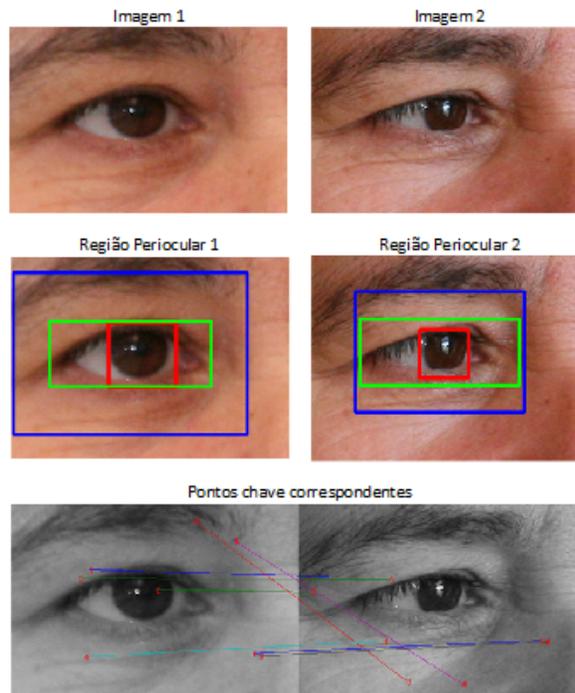
### Melhor resultado da classe Intra



**Figura 3.29:** Melhor resultado da classe Intra para o conjunto *Rotation*

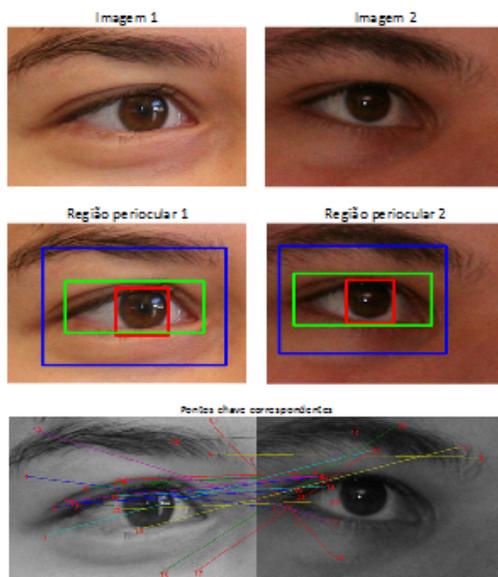
A figura 3.29 representa a melhor comparação entre imagens diferentes da mesma pessoa no conjunto *Rotation*, efetuada através do método de comparação proposto (secção 2.2.3), tendo um coeficiente negativo ( $\alpha = -22.324137$ ) distante de 0. Analisando a figura 3.29 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas regiões muito semelhantes que permitiram um bom desempenho dos métodos HOG e LBP. Também efetuou bastantes correspondências nos pontos chave como podemos observar. Podemos verificar que o coeficiente ( $\alpha$ ) não é muito alterado por pequenas rotações da cabeça. Com este coeficiente ( $\alpha = -22.324137$ ) a comparação representada seria corretamente atribuída à classe que pertence, *Intra*, significando que o sistema reconhece que estamos a comparar duas imagens da mesma pessoa.

## Pior resultado da classe Intra



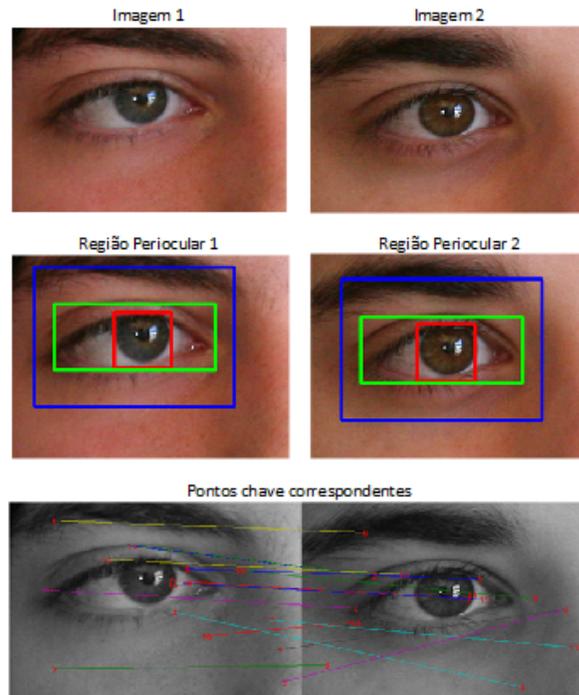
**Figura 3.30:** *Pior resultado da classe Intra para o conjunto Rotation*

A comparação representada na 3.30, efetuada através do método de comparação proposto (secção 2.2.3), teve um coeficiente positivo ( $\alpha = 5.180470$ ) distante de 0, representando o pior resultado obtido quando comparamos imagens diferentes da mesma pessoa no conjunto *Rotation*. Analisando a figura 3.30 podemos visualizar que o detetor efetuou uma estimativa pouco razoável da região periocular na segunda imagem da comparação. Podemos então ver, que com uma rotação elevada da cabeça todos os métodos dão maus resultados. Com este coeficiente ( $\alpha = 5.180470$ ) a comparação representada seria erradamente atribuída à classe *Inter*, significando que o sistema não reconhece que estamos a comparar duas imagens da mesma pessoa.

**Melhor resultado da classe Inter**

**Figura 3.31:** Melhor resultado da classe *Inter* para o conjunto *Rotation*

A figura 3.31 representa a melhor comparação entre imagens de pessoas diferentes no conjunto *Rotation*, efetuada através do método de comparação proposto (secção 2.2.3), tendo um coeficiente positivo ( $\alpha = 3.222420$ ) distante de 0. Analisando a figura 3.31 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas boas regiões que permitiram um bom desempenho dos métodos HOG e LBP. Apesar de algumas falsas correspondências nos pontos chave, podemos verificar que nesta imagem o fator de variabilidade pretendido (rotação da cabeça) contribuiu para aumentar o coeficiente ( $\alpha$ ). Com este coeficiente ( $\alpha = 3.222420$ ) a comparação representada seria corretamente atribuída à classe que pertence, *Inter*, significando que o sistema reconhece que estamos a comparar duas imagens de pessoas diferentes.

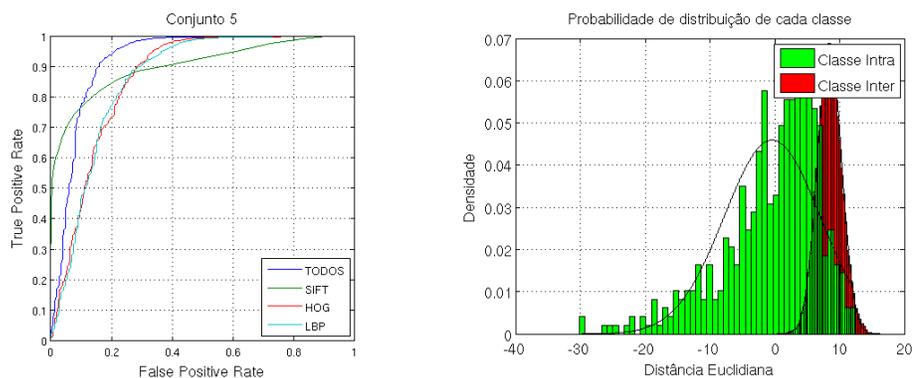
**Pior resultado da classe Inter**

**Figura 3.32:** *Pior resultado da classe Inter para o conjunto Rotation*

A comparação representada na 3.32, efetuada através do método de comparação proposto (secção 2.2.3), teve um coeficiente positivo ( $\alpha = 0.628068$ ) próximo de 0, representando o pior resultado obtido quando comparamos imagens de diferentes pessoas no conjunto *Rotation*. Analisando a figura 3.32 podemos visualizar que o detector efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas boas regiões, que obtiveram um desempenho razoável dos métodos HOG e LBP, apesar do fator que pretendíamos analisar (rotação da cabeça) ter influenciado um pouco a aproximação do coeficiente ( $\alpha$ ) ao valor 0. Podemos também ver que houve muitas correspondências entre pontos chave incorretas. Com este coeficiente ( $\alpha = 0.628068$ ) a comparação representada seria corretamente atribuída à classe *Intra*, significando que o sistema reconhece que estamos a comparar duas imagens de pessoas diferentes.

### 3.4.4 Análise ao conjunto *Glasses*

Usando a fusão dos métodos para a extração e comparação entre as imagens do conjunto *Glasses*, obtivemos uma AUC de 0.924, com um EER de 0.134 e uma decidibilidade de 1.639, valores que estão representados na figura 3.33.



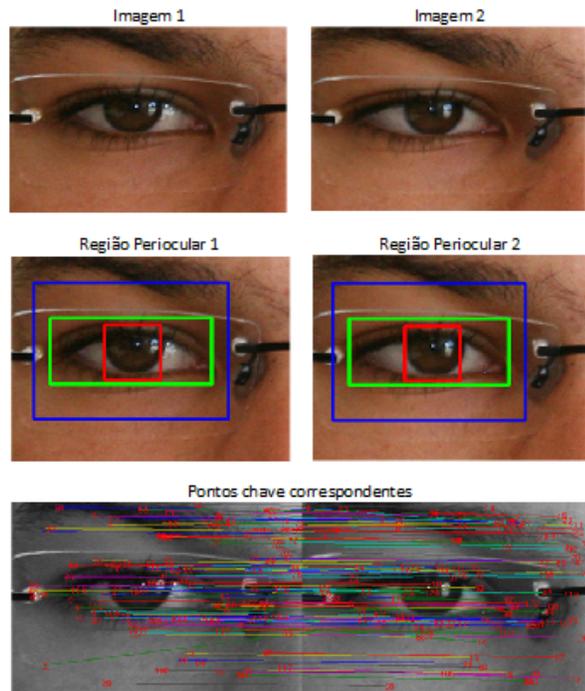
**Figura 3.33:** Curva ROC e distribuição de probabilidades do conjunto *Glasses*

Podemos observar que este conjunto foi o que obteve os piores resultados, tratando-se de um conjunto de imagens com um fator de variabilidade (óculos) bastante comum. Vamos verificar o comportamento de cada um dos métodos individualmente na tabela 3.6.

Método	AUC	EER	DEC
Scale Invariant Feature Transform	0.905	0.828	1.314
Histogram of Oriented Gradients	0.863	0.218	1.570
Local Binary Pattern	0.860	0.210	1.519
Fusão dos métodos	0.924	0.140	1.639

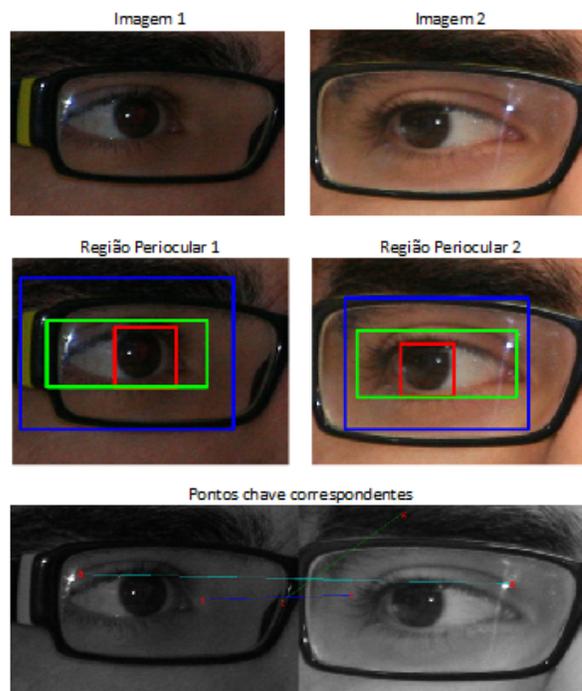
**Tabela 3.6:** Resultados dos diferentes métodos no conjunto *Glasses*

Baseando nestes resultados, vamos individualmente analisar os melhores e os piores resultados obtidos para cada uma das classes.

**Melhor resultado da classe Intra**

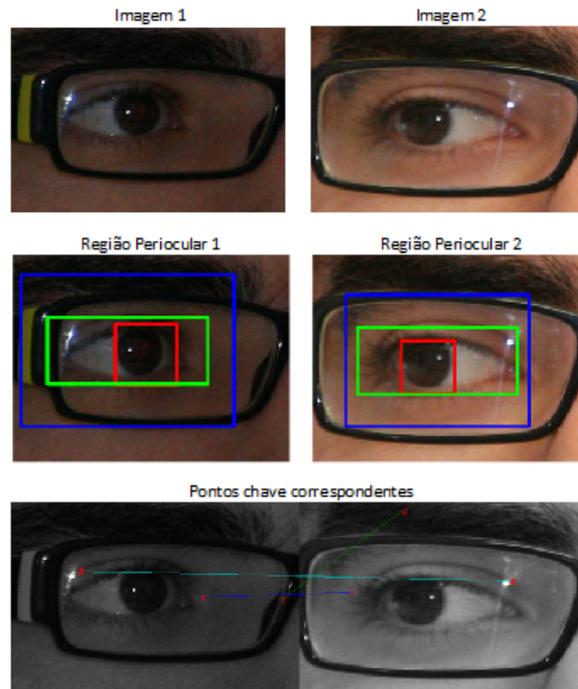
**Figura 3.34:** Melhor resultado da classe Intra para o conjunto Glasses

A figura 3.34 representa a melhor comparação entre imagens diferentes da mesma pessoa no conjunto *Glasses*, efetuada através do método de comparação proposto (secção 2.2.3), tendo um coeficiente negativo ( $\alpha = -30.105810$ ) distante de 0. Analisando a figura 3.34 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas regiões muito semelhantes que permitiram um bom desempenho dos métodos HOG e LBP. Também efetuou bastantes correspondências nos pontos chave como podemos observar. Temos que considerar que teve influência sobre este coeficiente ( $\alpha$ ) os óculos estarem maioritariamente fora da região periocular estimada e não estarem a fazer nenhum reflexo. Com este coeficiente ( $\alpha = -30.105810$ ) a comparação representada seria corretamente atribuída à classe que pertence, *Intra*, significando que o sistema reconhece que estamos a comparar duas imagens da mesma pessoa.

**Pior resultado da classe Intra**

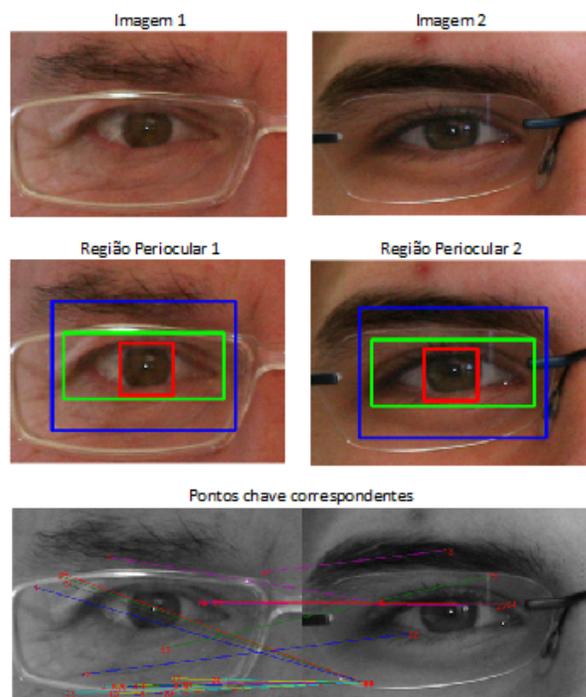
**Figura 3.35:** *Pior resultado da classe Intra para o conjunto Glasses*

A comparação representada na 3.35, efetuada através do método de comparação proposto (secção 2.2.3), teve um coeficiente positivo ( $\alpha = 12.522429$ ) distante de 0, representando o pior resultado obtido quando comparamos imagens diferentes da mesma pessoa no conjunto *Glasses*, sendo também o pior resultado de todos os conjuntos. Analisando a figura 3.35 podemos visualizar que a estimativa da região periocular na primeira imagem da comparação, apanhou grande parte dos óculos e na segunda quase não se encontram na região periocular. De salientar também que são olhos diferentes do sujeito (olho direito e olho esquerdo), olhando os dois para o lado temporal o que dificulta ainda mais o processo. Com este coeficiente ( $\alpha = 12.522429$ ) a comparação representada seria erradamente atribuída à classe *Inter*, significando que o sistema não reconhece que estamos a comparar duas imagens da mesma pessoa.

**Melhor resultado da classe Inter**

**Figura 3.36:** *Melhor resultado da classe Inter para o conjunto Glasses*

A figura 3.36 representa a melhor comparação entre imagens de pessoas diferentes no conjunto *Glasses*, efetuada através do método de comparação proposto (secção 2.2.3), tendo um coeficiente positivo ( $\alpha = 4.740531$ ) ligeiramente distante de 0. Analisando a figura 3.36 podemos visualizar que a estimativa da região periocular na primeira imagem da comparação, apanhou grande parte dos óculos e na segunda quase não se encontram na região periocular. De salientar que os sujeitos estão a olhar em direções diferentes, o que facilita um pouco o processo. Com este coeficiente ( $\alpha = 4.740531$ ) a comparação representada seria corretamente atribuída à classe que pertence, *Inter*, significando que o sistema reconhece que estamos a comparar duas imagens de pessoas diferentes.

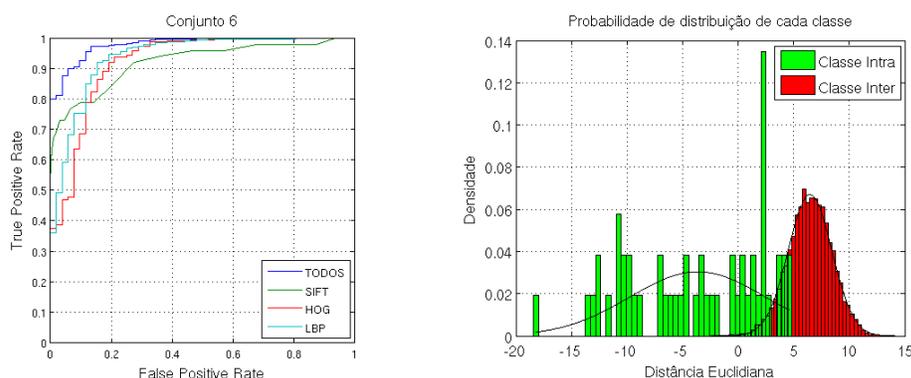
**Pior resultado da classe Inter**

**Figura 3.37:** *Pior resultado da classe Inter para o conjunto Glasses*

A comparação representada na 3.37, efetuada através do método de comparação proposto (secção 2.2.3), teve um coeficiente positivo ( $\alpha = 0.287444$ ) próximo de 0, representando o pior resultado obtido quando comparamos imagens de diferentes pessoas no conjunto *Glasses*. Analisando a figura 3.37 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas boas regiões, tendo um bom desempenho dos métodos HOG e LBP. Apenas o método SIFT obteve piores resultados porque encontrou muitas correspondências, mas o coeficiente final foi positivo, significando que para este conjunto não houve nenhum falso positivo. Com este coeficiente ( $\alpha = 0.287444$ ) a comparação representada seria corretamente atribuída à classe *Inter*, significando que o sistema reconhece que estamos a comparar duas imagens de pessoas diferentes.

### 3.4.5 Análise ao conjunto *Temporal*

Usando a fusão dos métodos para a extração e comparação entre as imagens do conjunto *Temporal*, obtivemos uma AUC de 0.981, com um EER de 0.090 e uma decidibilidade de 2.223, valores que estão representados na figura 3.38.



**Figura 3.38:** Curva ROC e distribuição de probabilidades do conjunto *Temporal*

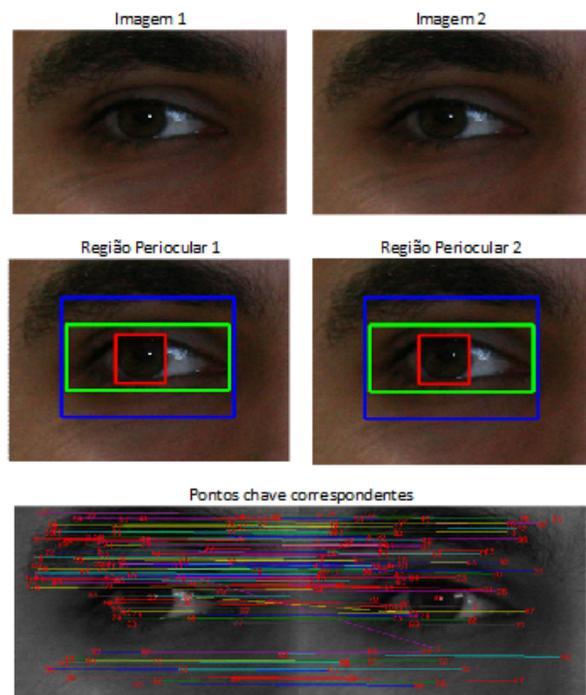
Podemos observar que este conjunto obteve bons resultados, algo que pretendíamos atingir por tratar-se de um conjunto de imagens que pretendia testar o detetor implementado. Vamos verificar o comportamento de cada um dos métodos individualmente na tabela 3.7.

Método	AUC	EER	DEC
Scale Invariant Feature Transform	0.923	0.876	1.705
Histogram of Oriented Gradients	0.923	0.154	2.050
Local Binary Pattern	0.940	0.135	2.233
Fusão dos métodos	0.981	0.090	2.223

**Tabela 3.7:** Resultados dos diferentes métodos no conjunto *Temporal*

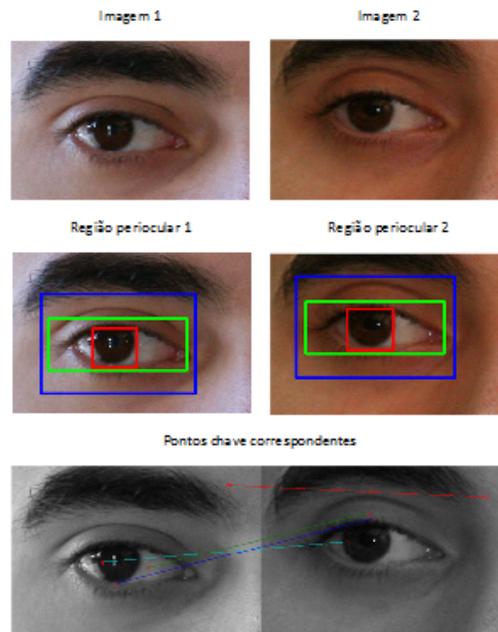
Baseando nestes resultados, vamos individualmente analisar os melhores e os piores resultados obtidos para cada uma das classes.

### Melhor resultado da classe Intra



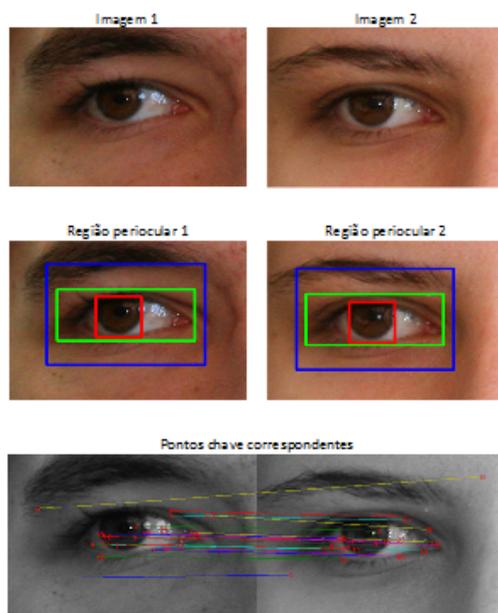
**Figura 3.39:** Melhor resultado da classe *Intra* para o conjunto *Temporal*

A figura 3.39 representa a melhor comparação entre imagens diferentes da mesma pessoa no conjunto *Temporal*, efetuada através do método de comparação proposto (secção 2.2.3), tendo um coeficiente negativo ( $\alpha = -18.421555$ ) distante de 0. Analisando a figura 3.39 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas regiões muito semelhantes que permitiram um bom desempenho dos métodos HOG e LBP. Também efetuou bastantes correspondências nos pontos chave como podemos observar. Esta imagem é um exemplo do bom resultado que conseguimos obter com o detetor implementado quando o olho movimentar-se. Com este coeficiente ( $\alpha = -18.421555$ ) a comparação representada seria corretamente atribuída à classe que pertence, *Intra*, significando que o sistema reconhece que estamos a comparar duas imagens da mesma pessoa.

**Pior resultado da classe Intra**

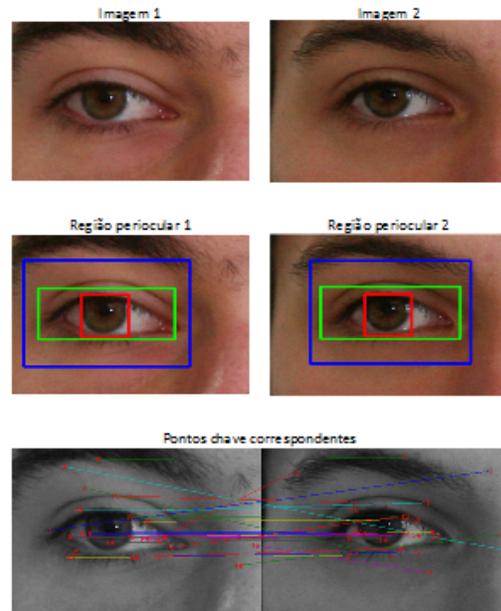
**Figura 3.40:** *Pior resultado da classe Intra para o conjunto Temporal*

A comparação representada na 3.40, efetuada através do método de comparação proposto (secção 2.2.3), teve um coeficiente positivo ( $\alpha = 4.849111$ ) distante de 0, representando o pior resultado obtido quando comparamos imagens diferentes da mesma pessoa no conjunto *Temporal*. Analisando a figura 3.40 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas regiões muito semelhantes, mas que devido à grande diferença de iluminação foi obtido um mau desempenho dos métodos HOG e LBP. Devido ao mesmo motivo, o método SIFT também fez muito poucas correspondências. Com este coeficiente ( $\alpha = 12.471781$ ) a comparação representada seria erradamente atribuída à classe *Inter*, significando que o sistema não reconhece que estamos a comparar duas imagens da mesma pessoa.

**Melhor resultado da classe Inter**

**Figura 3.41:** Melhor resultado da classe *Inter* para o conjunto *Temporal*

A figura 3.41 representa a melhor comparação entre imagens de pessoas diferentes no conjunto *Temporal*, efetuada através do método de comparação proposto (secção 2.2.3), tendo um coeficiente positivo ( $\alpha = 1.073602$ ) ligeiramente distante de 0. Analisando a figura 3.41 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas boas regiões que permitiram um bom desempenho dos métodos HOG e LBP. Também efetuou poucas correspondências nos pontos chave como podemos observar. Esta imagem é um exemplo do bom resultado que conseguimos obter com o detetor implementado quando o olho movimentar-se. Com este coeficiente ( $\alpha = 1.073602$ ) a comparação representada seria corretamente atribuída à classe que pertence, *Inter*, significando que o sistema reconhece que estamos a comparar duas imagens de pessoas diferentes.

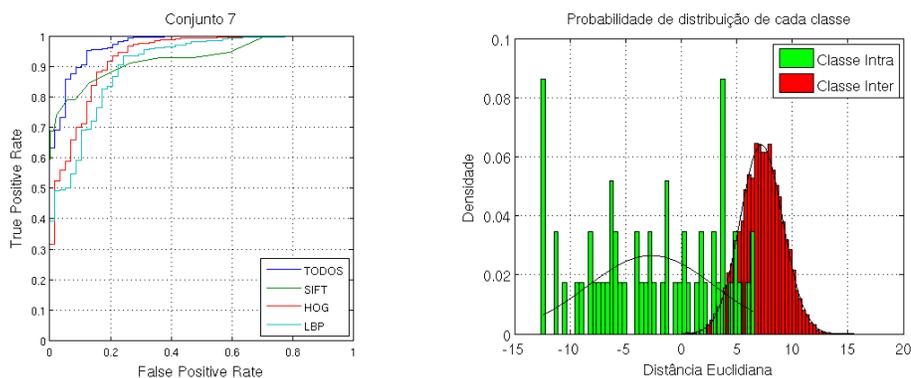
**Pior resultado da classe Inter**

**Figura 3.42:** *Pior resultado da classe Inter para o conjunto Temporal*

A comparação representada na 3.42, efetuada através do método de comparação proposto (secção 2.2.3), teve um coeficiente negativo ( $\alpha = -2.418993$ ) distante de 0, representando o pior resultado obtido quando comparamos imagens de diferentes pessoas no conjunto *Temporal*. Analisando a figura 3.42 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas boas regiões, mas que obtiveram um pouco satisfatório no desempenho dos métodos HOG e LBP. A correspondência de pontos também foi elevada, produzindo em geral maus resultados apesar da boa estimativa do detetor. Com este coeficiente ( $\alpha = -2.418993$ ) a comparação representada seria erradamente atribuída à classe *Intra*, significando que o sistema não reconhece que estamos a comparar duas imagens de pessoas diferentes.

### 3.4.6 Análise ao conjunto *Nasal*

Usando a fusão dos métodos para a extração e comparação entre as imagens do conjunto *Nasal*, obtivemos uma AUC de 0.972, com um EER de 0.098 e uma decidibilidade de 2.296, valores que estão representados na figura 3.43.



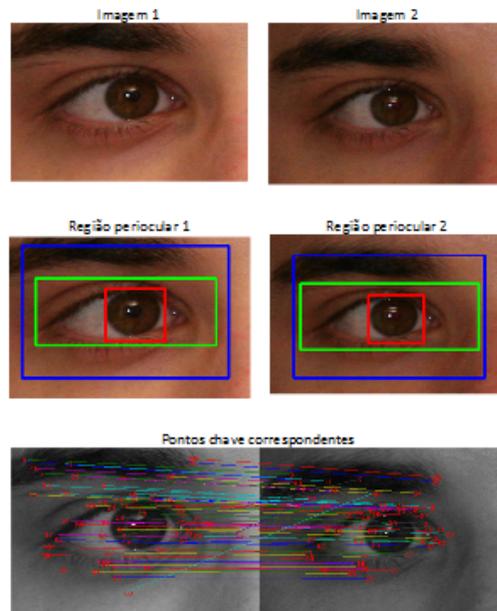
**Figura 3.43:** Curva ROC e distribuição de probabilidades do conjunto *Nasal*

Podemos observar que este conjunto obteve bons resultados, algo que pretendíamos atingir por tratar-se de um conjunto de imagens que pretendia testar o detetor implementado. Vamos verificar o comportamento de cada um dos métodos individualmente na tabela 3.8.

Método	AUC	EER	DEC
Scale Invariant Feature Transform	0.930	0.149	1.805
Histogram of Oriented Gradients	0.931	0.155	2.160
Local Binary Pattern	0.909	0.172	1.920
Fusão dos métodos	0.972	0.098	2.296

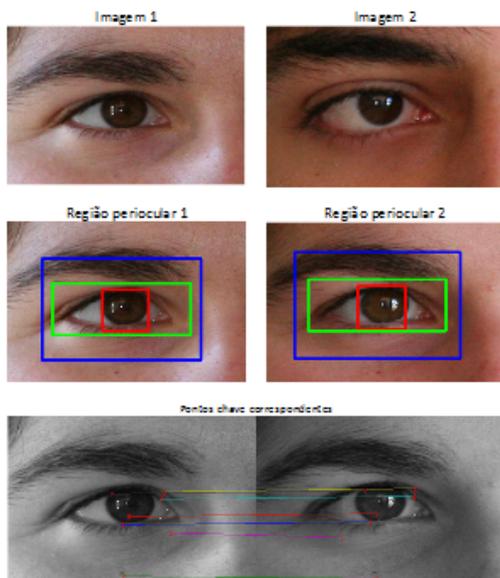
**Tabela 3.8:** Resultados dos diferentes métodos no conjunto *Nasal*

Baseando nestes resultados, vamos individualmente analisar os melhores e os piores resultados obtidos para cada uma das classes.

**Melhor resultado da classe Intra**

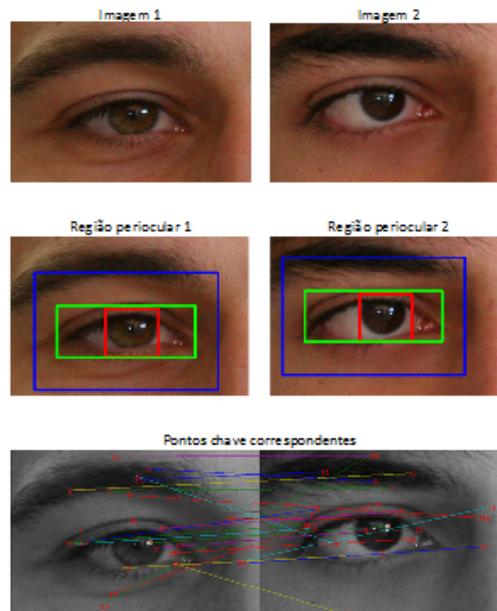
**Figura 3.44:** Melhor resultado da classe Intra para o conjunto Nasal

A figura 3.44 representa a melhor comparação entre imagens diferentes da mesma pessoa no conjunto *Nasal*, efetuada através do método de comparação proposto (secção 2.2.3), tendo um coeficiente negativo ( $\alpha = -12.609316$ ) distante de 0. Analisando a figura 3.44 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas regiões muito semelhantes que permitiram um bom desempenho dos métodos HOG e LBP. Também efetuou bastantes correspondências nos pontos chave como podemos observar. Esta imagem é um exemplo do bom resultado que conseguimos obter com o detetor implementado quando o olho movimentar-se. Com este coeficiente ( $\alpha = -12.609316$ ) a comparação representada seria corretamente atribuída à classe que pertence, *Intra*, significando que o sistema reconhece que estamos a comparar duas imagens da mesma pessoa.

**Pior resultado da classe Intra**

**Figura 3.45:** *Pior resultado da classe Intra para o conjunto Nasal*

A comparação representada na 3.45, efetuada através do método de comparação proposto (secção 2.2.3), teve um coeficiente positivo ( $\alpha = 6.603622$ ) distante de 0, representando o pior resultado obtido quando comparamos imagens diferentes da mesma pessoa no conjunto *Nasal*. Analisando a figura 3.45 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas regiões muito semelhantes, mas que devido à grande diferença de iluminação foi obtido um mau desempenho dos métodos HOG e LBP. Devido ao mesmo motivo, o método SIFT também fez muito poucas correspondências. Com este coeficiente ( $\alpha = 6.603622$ ) a comparação representada seria erradamente atribuída à classe *Inter*, significando que o sistema não reconhece que estamos a comparar duas imagens da mesma pessoa.

**Melhor resultado da classe Inter**

**Figura 3.46:** *Melhor resultado da classe Inter para o conjunto Nasal*

A figura 3.46 representa a melhor comparação entre imagens de pessoas diferentes no conjunto *Nasal*, efetuada através do método de comparação proposto (secção 2.2.3), um coeficiente positivo ( $\alpha = 2.332287$ ) distante de 0. Analisando a figura 3.46 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas boas regiões que permitiram um bom desempenho dos métodos HOG e LBP. Também efetuou poucas correspondências nos pontos chave como podemos observar. Esta imagem é um exemplo do bom resultado que conseguimos obter com o detetor implementado quando o olho movimentar-se. Com este coeficiente ( $\alpha = 2.332287$ ) a comparação representada seria corretamente atribuída à classe que pertence, *Inter*, significando que o sistema reconhece que estamos a comparar duas imagens de pessoas diferentes.

### Pior resultado da classe Inter

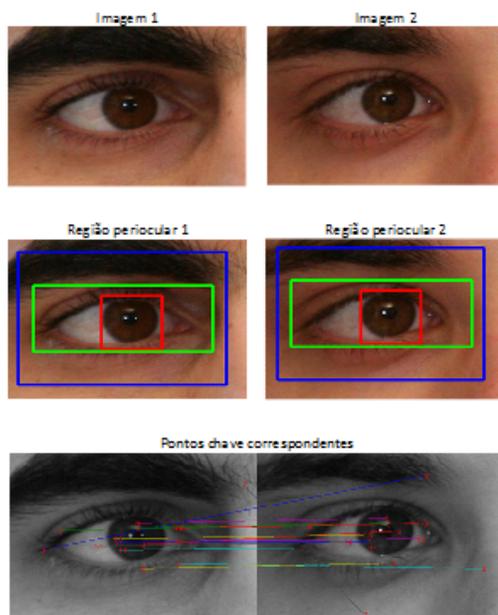


Figura 3.47: Pior resultado da classe Inter para o conjunto Nasal

A comparação representada na 3.47, efetuada através do método de comparação proposto (secção 2.2.3), teve um coeficiente negativo ( $\alpha = 0.135608$ ) distante de 0, representando o pior resultado obtido quando comparamos imagens de diferentes pessoas no conjunto *Nasal*. Analisando a figura 3.47 podemos visualizar que o detetor efetuou uma boa estimativa da região periocular gerando nas duas imagens a ser comparadas boas regiões, obtendo um bom desempenho dos métodos HOG e LBP. Apenas o método SIFT obteve piores resultados porque encontrou muitas correspondências, mas o coeficiente final foi positivo, significando que para este conjunto não houve nenhum falso positivo. Com este coeficiente ( $\alpha = 0.135608$ ) a comparação representada seria corretamente atribuída à classe *Inter*, significando que o sistema reconhece que estamos a comparar duas imagens de pessoas diferentes.

### 3.5 Síntese dos resultados obtidos

Efetuada a análise do método proposto para cada um dos conjuntos de dados podemos agora visualizar os resultados obtidos para cada um dos diferentes fatores de variabilidade. Para uma melhor compreensão e visualização dos resultados representaremos os valores obtidos com o método proposto na tabela 3.9.

Conjuntos de dados	AUC	EER	DEC
<i>All</i>	0.966	0.097	1.810
<i>Frontal</i>	0.994	0.044	1.827
<i>Hair</i>	0.999	0.014	2.105
<i>Rotation</i>	0.976	0.086	1.713
<i>Glasses</i>	0.924	0.140	1.639
<i>Temporal</i>	0.981	0.090	2.223
<i>Nasal</i>	0.972	0.098	2.296

**Tabela 3.9:** Resultados dos diferentes conjuntos

Em suma, podemos afirmar que os resultados obtidos no conjunto *All* são os mais aproximados da realidade devido aos vários fatores de variabilidade presentes no conjunto de dados. Analisando o pior resultado da classe *Intra* no conjunto anteriormente referido podemos verificar o coeficiente ( $\alpha$ ) indesejado que pode resultar de uma má estimativa da região periocular.

Confirmamos que com uma má estimativa da mesma por parte do detetor, todos os resultados obtidos na execução do método proposto não iriam corresponder aos esperados, gerando um coeficiente ( $\alpha$ ) indesejado, demonstrando assim uma total dependência do bom funcionamento do detetor por parte do método proposto.

Analisando imagens capturadas com o sujeito a olhar para a câmara verificámos que o método proposto devolveu bons resultados que podemos demonstrar baseados no coeficiente ( $\alpha$ ) atingido no melhor resultado da classe *Intra* do conjunto *All* e nos resultados obtidos quando analisámos o conjunto *Frontal*, conjunto baseado em imagens semelhantes às descritas.

Visualizando os resultados obtidos na análise dos conjuntos *Hair* e *Rotation*, obtivemos bons resultados o que nos surpreendeu devido aos fatores de variabilidade que tentamos demonstrar nestas amostras, presença de cabelo na região periorcular e rotação da cabeça, respetivamente. Verificando o número total de amostras em cada conjunto ( trinta e nove no conjunto *Hair* e sessenta e seis no conjunto *Rotation*) podemos afirmar que a amostra não é significativa, por isso os dados não são tão fiáveis como os obtidos nos restantes conjuntos, o que não permitiu tirar conclusões destes resultado.

Para analisar a influência destes fatores de variabilidade, podemos observar os piores casos de cada uma das classes (*Intra* e *Inter*) nos dois conjuntos( conjunto *Hair* e *Rotation*) avaliando assim alguns potenciais casos problemáticos no método proposto. A partir da análise destes casos podemos verificar que o método proposto tem um bom desempenho quando o sujeito apresenta uma ligeira rotação da cabeça, mas quando a rotação é acentuada ele obtém muitos coeficientes ( $\alpha$ ) interpretados erradamente como pertencente à classe *Inter*, resultando num valor elevado de falsos negativos, como está demonstrado na tabela 3.10.

Conjuntos de dados	Total	Verdadeiros negativos	Falsos positivos	Falsos negativos	Verdadeiros positivos
<i>All</i>	367653	366447	40	553	613
<i>Frontal</i>	19899	19805	3	26	65
<i>Hair</i>	702	656	2	6	38
<i>Rotation</i>	2079	2063	0	8	8
<i>Glasses</i>	23435	22938	11	217	269
<i>Temporal</i>	10877	10824	1	23	29
<i>Nasal</i>	10152	10093	1	23	35
Total	434797	432826	58	856	1057

**Tabela 3.10:** Representação das matrizes de confusão dos conjuntos

Visualizando a tabela 3.10 podemos verificar que o método proposto tem um número elevado de falsos negativos na maioria dos conjuntos e apresenta um número reduzido de falsos positivos, o que consideramos ser bastante positivo.

Observando os resultados conseguidos quando analisado o conjunto *Glasses* verificamos que este obteve os piores resultados entre todos os conjuntos demonstrando como pretendíamos que os óculos são o fator de variabilidade que influencia mais o resultado nos conjuntos analisados.

Quando analisamos imagens em que o sujeito está a utilizar óculos temos que lidar com diferentes condicionantes, principalmente os reflexos da luz na lente dos óculos e quando os óculos estão sob a região periocular estimada pelo detetor. Como representado na análise do conjunto *Glasses* vemos que todos os métodos obtiveram piores resultados do que em qualquer outro conjunto.

Com os resultados obtidos nos conjuntos *Temporal* e *Nasal* podemos verificar que apesar do olho não estar a olhar para a câmara, ele na maioria dos casos foi corretamente detetado, tendo o detetor devolvido uma boa estimativa da região periocular.

Com o detetor a devolver bons resultados apesar do deslocamento de um dos pontos de referência (íris), as regiões periculares estimadas são muito semelhantes, permitindo um bom reconhecimento através do método proposto.



# Capítulo 4

## Conclusão e trabalho futuro

### 4.1 Conclusão

Os objetivos propostos para este projeto foram atingidos com sucesso, confirmando que a região periocular é das que mais informação discriminante entre indivíduos contém.

Foi implementada a abordagem [13] proposta para este projeto, o que permitiu-nos avaliar o desempenho de cada um dos métodos utilizado por esta ( SIFT, HOG e LBP) e assim concluir e perceber quando é que cada um dos mesmos produz piores resultados, para analisarmos quais os fatores de variabilidade que mais influenciam os diferentes métodos.

Na abordagem anteriormente descrita não foi referido como efetuaram a combinação dos diferentes resultados dos três métodos utilizados, o que nos levou a optar por efetuar a comparação através do método descrito (secção 2.2.3).

Como referido anteriormente, a abordagem em que nos baseávamos não efetuava a segmentação automaticamente, tendo as imagens de ser segmentadas manualmente. Conseguimos implementar o método descrito por Viola e Jones [18] efetuando automaticamente a estimativa da região periocular através da localização da íris e do olho.

Este detetor veio a revelar-se um ótimo acrescento, conseguindo obter boas estimativas da região periocular mesmo sem a íris estar no centro do olho, gerando regiões muito semelhantes na maioria dos casos entre imagens diferentes da mesma pessoa.

Com o método proposto neste projeto para estimar a região periocular ao analisar uma imagem, extrair as características da região obtida e efetuar a comparação entre imagens distintas, podemos afirmar que obtivemos bons resultados nas comparações obtidas.

Analisando todos os dados que obtivemos na realização deste projeto, considerámos os óculos como o fator de variabilidade que mais influenciou o correto desempenho do método proposto, ou seja, os reflexos produzidos pelas lentes e a existência de óculos que estão sob a região periocular estimada pelo detetor prejudica o método proposto.

Outros fatores que influenciaram o desempenho do método proposto foi o nível de iluminação a que a pessoa está sujeita, a presença de cabelo sob a região periocular estimada e a rotação da cabeça quando esta é muito acentuada.

Concluímos que nos conjuntos de dados em que o sujeito está a olhar para o lado (temporal ou nasal), ou seja, quando a íris não se encontra no centro do olho, também existe uma descida na eficiência do método mas que esta foi minimizada pelo detetor que implementámos, permitindo estimar regiões perioculares semelhantes.

## 4.2 Trabalho futuro

Sendo que as experiências aqui efetuadas demonstram um bom potencial biométrico da região periocular, seria de considerar a continuação da análise e desenvolvimento do método proposto.

Numa perspetiva de trabalho futuro podemos considerar uma adaptação do método proposto, investigando-se a hipótese de segmentar a estimativa da região periocular nos vários objetos que nela se encontram, por exemplo, conseguir reconhecer a íris, a sobrancelha, a esclera, a pupila, as pestanas e a pele existente na região periocular.

Depois de obtermos a máscara de segmentação, iríamos poder analisar cada elemento separadamente, possivelmente cada um com o método que mais se adequasse, aplicando métodos descritores de texturas e orientações em cada uma das regiões segmentadas. Também poderíamos aplicar uma limitação na aceitação dos pontos chave, apenas aceitando os que são das mesmas regiões.

Para melhor compreender e exemplificar este processo, representamos os vários passos na figura 4.1.

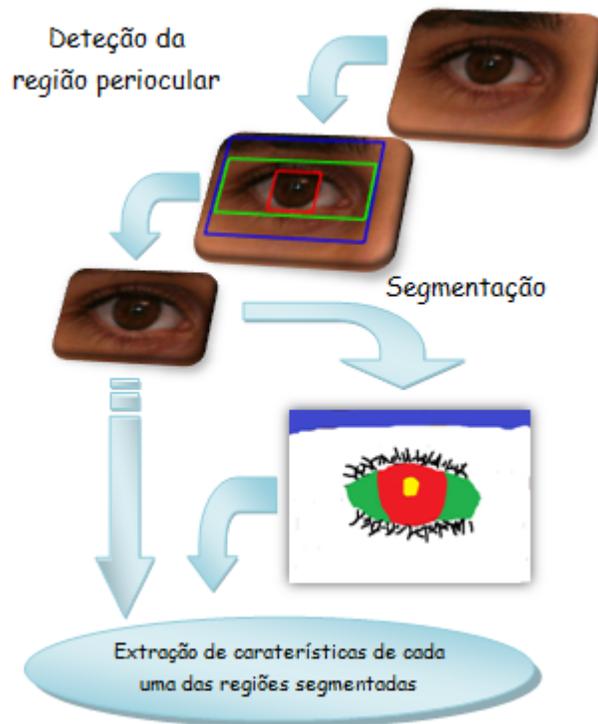


Figura 4.1: Perspetiva de trabalho futuro



# Apêndice A

## Ambiente de desenvolvimento

### A.1 Linguagens de desenvolvimento

O código deste projeto foi desenvolvido sobre duas plataformas:

1. **C++** - foi utilizado o editor de texto Kate (versão 3.5.5) e para compilar o código foi usado o compilador da *GNU Compiler Collection* (GCC) para C++, sendo o código desenvolvido para esta linguagem o correspondente ao detetor utilizado neste projeto.
2. **Matrix Laboratory (MatLab)** - foi utilizado este poderoso software (MatLab versão R2010b) de cálculo para efetuar o restante código, ou seja, todo o processo de extração de características e de comparação entre diferentes imagens.

### A.2 Bibliotecas utilizadas

Para efetuar este projeto também foram importadas duas bibliotecas:

1. **OpenCV** - utilizámos a biblioteca *Open Source Computer Vision* (OpenCV), compatível com C++, que é uma biblioteca de funções em tempo real para visão computacional.
2. **VLFeat** - esta biblioteca implementa alguns algoritmos de visão computacional com interface para MatLab.



# Bibliografia

- [1] Brian C. Becker. In *Computer vision sift lecture*, 2009.
- [2] Jeffrey S. Beis and David G. Lowe. Shape indexing using approximate nearest-neighbour search in high-dimensional spaces. In *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)*, CVPR '97, pages 1000–, Washington, DC, USA, 1997. IEEE Computer Society.
- [3] Matthew Brown and David Lowe. Invariant Features from Interest Point Groups. In *In British Machine Vision Conference*, pages 656–665, 2002.
- [4] W T Freeman and M Roth. Orientation histograms for hand gesture recognition. *Gesture*, 12(TR-94-03):296–301, 1994.
- [5] Yoav Freund and Robert E. Schapire. Experiments with a new boosting algorithm. In *International Conference on Machine Learning*, pages 148–156, 1996.
- [6] CVIS Research Group. Robust real-time face detection by paul viola and michael jones. <http://utarcvis.blogspot.com/2007/09/robust-real-time-face-detection-by-paul.html>.
- [7] Alfred Haar. Zur Theorie der orthogonalen Funktionensysteme. *Mathematische Annalen*, 69(3):331–371, September 1910.
- [8] David J. Hand and Robert J. Till. A simple generalisation of the area under the roc curve for multiple class classification problems. *Mach. Learn.*, 45:171–186, October 2001.

- [9] David G. Lowe. Object recognition from local scale-invariant features. In *Proceedings of the International Conference on Computer Vision-Volume 2 - Volume 2, ICCV '99*, pages 1150–, Washington, DC, USA, 1999. IEEE Computer Society.
- [10] David G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, 60:91–110, November 2004.
- [11] T. Ojala, M. Pietikainen, and D. Harwood. A comparative study of texture measures with classification based on feature distributions. 29(1):51–59, January 1996.
- [12] Timo Ojala, M Pietikainen, and David Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1):51–59, 1996.
- [13] Unsang Park, Arun Ross, and Anil K. Jain. Periocular biometrics in the visible spectrum: a feasibility study. In *Proceedings of the 3rd IEEE international conference on Biometrics: Theory, applications and systems, BTAS'09*, pages 153–158, Piscataway, NJ, USA, 2009. IEEE Press.
- [14] H. Proenca, S. Filipe, R. Santos, J. Oliveira, and L.A. Alexandre. The UBIRIS.v2: A database of visible wavelength images captured on-the-move and at-a-distance. *IEEE Trans. PAMI*, 32(8):1529–1535, August 2010.
- [15] Caifeng Shan, Shaogang Gong, and Peter W. McOwan. Facial expression recognition based on local binary patterns: A comprehensive study. *Image Vision Comput.*, 27:803–816, May 2009.
- [16] Mäenpää T. *The local binary pattern approach to texture analysis - extensions and applications*. PhD thesis, 2003. Dissertation. Acta Univ Oul C 187, 78 p + App.
- [17] Utkarsh. Integral images in opencv. <http://www.aishack.in/2010/07/integral-images/>.
- [18] Paul Viola and Michael J. Jones. Robust real-time face detection. *Int. J. Comput. Vision*, 57:137–154, May 2004.

- 
- [19] Xiaoyu Wang, Tony X. Han, and Shuicheng Yan. An hog-lbp human detector with partial occlusion handling. In *ICCV*, pages 32–39. IEEE, 2009.
- [20] Andrew P. Witkin. Scale-Space Filtering. In *8th Int. Joint Conf. Artificial Intelligence*, volume 2, pages 1019–1022, Karlsruhe, August 1983.
- [21] Damon L. Woodard, Shrinivas Pundlik, Philip Miller, Raghavender Jillela, and Arun Ross. On the fusion of periocular and iris biometrics in non-ideal imagery. In *Proceedings of the 2010 20th International Conference on Pattern Recognition, ICPR '10*, pages 201–204, Washington, DC, USA, 2010. IEEE Computer Society.
- [22] D.L. Woodard, S.J. Pundlik, J.R. Lyle, and P.E. Miller. Periocular region appearance cues for biometric identification. pages 162–169, 2010.