

# Detecção de Faces em Imagens baseada na Identificação da Pele e dos Olhos

Fernando J. S. Carvalho<sup>1</sup>, João Manuel R. S. Tavares<sup>2</sup>

fjc@isep.ipp.pt, tavares@fe.up.pt

<sup>1</sup>ISEP – Instituto Superior de Engenharia do Porto

<sup>2</sup>FEUP – Faculdade de Engenharia da Universidade do Porto

**Resumo.** Uma das principais áreas de desenvolvimento do domínio da Visão Computacional é a da detecção de faces em imagens. Das várias metodologias existentes nesta área, apresentam-se duas neste artigo: uma baseada na detecção de zonas de pele e uma segunda baseada num modelo protótipo deformável para detecção dos olhos e extração das suas características. A detecção de zonas de pele permite segmentar imagens de faces considerando regiões nas quais os valores dos seus pixéis, em termos de probabilidade, se assemelham aos exibidos pela pele. No entanto, como esta metodologia não produz informação suficiente para concluir um processo de detecção de faces em imagens com elevada confiança, usa-se em complemento um modelo protótipo deformável para os olhos, que adequadamente posicionado próximo de pequenas regiões detectadas no interior dos segmentos previamente identificados como associados a pele, permite detectar a presença dos olhos e consequentemente validar ou não a existência de uma face. Este artigo apresenta a descrição das referidas metodologias e de alguns resultados experimentais obtidos a partir de implementações desenvolvidas em *Matlab*.

**Palavras-chave:** Detecção de pele, Detecção de faces, Emparelhamento, Modelos protótipo deformáveis.

**Abstract.** One of the main research areas in Computational Vision concerns the detection of faces in images. Among several existing methodologies, two are presented in this paper: one is based on skin detection and the other on deformable templates for eyes segmentation and their features extraction. The detection of skin regions allows the segmentation of faces in images through the identification of regions in which the pixels' values are, in terms of probability, similar with the ones usually exhibited by the skin. However, since from this methodology does not result sufficient information to conclude the detection process with a high success rate, a deformable template for the eyes is used in a complementary way. This template, adequately positioned near small regions detected inside the skin segments previously identify, allows the detection of the eyes and consequently validates or not the existence of a face in the input image. This paper presents the description of the above mentioned methodologies as well as some experimental results obtained using *Matlab*.

**Keywords:** Skin detection, Face detection, Template matching, Deformable template.

## 1 Introdução

Nas últimas três décadas, a investigação dedicada à Visão Computacional tem procurado reproduzir computacionalmente algumas das funcionalidades mais interessantes do sistema de visão humano. Por exemplo, considerando imagens de faces, são inúmeras as metodologias computacionais existentes para a detecção e o reconhecimento de faces, a extracção de informação facial, a análise de expressões faciais e a reconstrução de faces. Usualmente, essas metodologias consideram que a face humana pode ser interpretada globalmente, ou então, examinando algumas das suas características faciais mais representativas, como olhos, boca, sobrancelhas e nariz. A escolha adequada da metodologia computacional que satisfaça as exigências de uma determinada área depende muito da forma como a face se apresenta no cenário em causa, sendo comum considerar-se como factores de decisão variáveis como a luminosidade, escala, rotação, oclusão parcial e existência de óculos, bigode ou barba.

Em (Yang, Kriegman, & Ahuja, 2002), pode-se encontrar, um estudo sobre algumas das metodologias dedicadas à área ligada à detecção de faces em imagens estáticas. No trabalho referido, as metodologias existentes encontram-se divididas em quatro categorias: 1) baseadas no conhecimento, 2) em características invariantes, 3) na aparência, 4) e no emparelhamento de modelos.

No presente trabalho, pretende-se detectar faces presentes em imagens usando essencialmente duas metodologias distintas e complementares: a primeira, usa um algoritmo de detecção de zonas de pele, e pertence à categoria das metodologias baseadas em características invariantes; a segunda, usa um modelo protótipo deformável para os olhos, e enquadra-se na categoria das metodologias baseadas no emparelhamento de modelos.

O objectivo principal do algoritmo de detecção de zonas de pele usado na primeira metodologia, ver, por exemplo, (Rademacher, 2001) e (Carvalho & Tavares, 2005), consiste em extrair segmentos da imagem original formados por pixéis cujos valores se assemelham ao valor usualmente exibido pela pele. A detecção desses pixéis, e a consequente definição da região de um segmento de pele, baseia-se no uso de um modelo estatístico implementado com base no pressuposto que a cor da pele segue uma distribuição Normal (*Gaussiana*). Deste modo, o modelo obtido, detecta todos os pixéis de uma imagem que apresentam elevada probabilidade de pertencerem a uma determinada zona de pele e consequentemente, no caso presente, a prováveis regiões faciais. Para tal, baseia-se no cálculo prévio da média e covariância realizado a partir de um conjunto de imagens de treino das quais são extraídas diversas amostras de pele de regiões faciais de vários indivíduos.

No entanto, a presença de uma zona de pele num determinado segmento não garante por si só a existência de uma face no mesmo. Tal, problema ocorre porque: por um lado, zonas de pele da imagem original que pertencem a diferentes partes do corpo, como mãos, braços e pernas, poderão ser igualmente detectáveis, considerando-se a típica homogeneidade da cor exibida pela pele; e por outro, determinados objectos estranhos presentes na mesma imagem poderão de igual modo ser considerados como pele, em particular no caso das intensidades exibidas por alguns dos seus pixéis se identificarem com as intensidades detectadas pelo modelo.

Deste modo, para validar a existência de uma face em cada um dos segmentos de pele previamente determinados, seleccionam-se apenas os mais susceptíveis de conterem faces. Considerando-se como critério selectivo a localização no interior de cada segmento de pelo menos três regiões “ocas”, duas relativas à localização dos olhos e uma outra relativa à localização da boca. Nesta fase, são ainda removidas do interior dos segmentos de pele determinados as regiões de dimensão reduzida.

Seguidamente, de modo a concluir o processo de detecção de faces em imagens, como complemento ao uso do algoritmo de detecção de zonas de pele e pós-processamento dos segmentos de pele determinados, considera-se uma segunda metodologia computacional baseada no uso de um modelo protótipo deformável para os olhos. Este modelo, uma vez posicionado em torno das regiões “ocas” previamente localizadas, deverá emparelhar na imagem original com a característica facial que mais se assemelha à sua geometria considerada no modelo, no caso presente, equivalente aos olhos.

Alguns exemplos de trabalhos na área da detecção de faces em imagens que usam modelos protótipo deformáveis para os olhos podem ser verificados, por exemplo, em (Yuille, Hallinan, & Cohen, 1992), (Carvalho & Tavares, 2007a) e (Carvalho & Tavares, 2007b). Assim, em (Carvalho & Tavares, 2006), usa-se um modelo protótipo deformável simplificado na detecção da íris do olho e extracção das suas características, e em (Wu, Liu, & Zha, 2003), é usado um modelo protótipo deformável na detecção da pálpebra dos olhos, permitindo detectar o olho na condição de aberto ou fechado.

Os modelos protótipo deformáveis são parametrizados, facto que lhes confere a capacidade de alterarem a sua escala, posição e orientação; isto é, sofrerem transformações rígidas que ocorrem usualmente durante o processo de interacção dinâmica entre esses modelos e a imagem original, sendo em Visão Computacional este processo conhecido por emparelhamento de um modelo protótipo (*template matching*).

Para controlar o referido processo de emparelhamento do protótipo numa imagem, é utilizado um algoritmo de optimização. Usualmente, o gradiente descendente, actualizando-se iterativamente os parâmetros considerados na construção do modelo usado, até que seja mínimo o custo energético traduzido por uma função matemática que caracteriza a sua deformação.

Na condução do processo de emparelhamento são usados campos de energia obtidos a partir da imagem original; tais como: vales de intensidade, picos de intensidade, orlas de intensidade e intensidades em tons de cinzento.

Neste artigo, descrevem-se as metodologias consideradas e apresentam-se alguns exemplos de resultados obtidos experimentalmente, usando implementações das mesmas em plataforma *MatLab*.

O presente artigo encontra-se organizado da seguinte forma: na secção seguinte, apresenta-se a metodologia baseada na detecção de zonas de pele; na secção 3, descreve-se a metodologia baseada no modelo protótipo deformável para os olhos; na secção 4, apresentam-se alguns resultados experimentais obtidos; e finalmente, na última secção, são indicadas algumas conclusões e perspectivas de trabalho futuro.

## 2 Metodologia baseada na Detecção de Zonas de Pele

Nesta secção, apresenta-se um novo formato de cor de representação de imagens, descreve-se um modelo para detecção de zonas de pele em imagens baseado numa função de probabilidade, descreve-se o processo adoptado para a detecção de zonas de pele em imagens e, finalmente, apresenta-se um método heurístico auxiliar que, baseado na informação previamente adquirida, permite seleccionar apenas os segmentos de pele de interesse.

### 2.1 Conversão de imagens de *RGB* para *YCbCr*

Numa primeira fase, as imagens de faces originalmente no formato *RGB* são convertidas para o formato *YCbCr* (*luminance-chrominance*), permitindo assim reduzir os efeitos perturbadores da luminosidade nas mesmas.

O formato *YCbCr* é frequentemente conhecido pelo espaço de cores puras, sendo representado a partir de três componentes: *Y*, que representa o brilho da imagem; *Cb*, que deriva do azul primário *B* subtraído de *Y*; e *Cr*, que deriva do vermelho primário *R* subtraído de *Y* (Fig. 1).

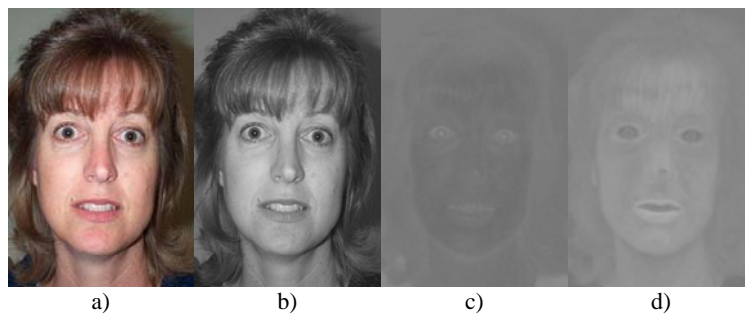


Figura 1 – Imagem de uma face no formato *RGB* (a) e imagens das componentes que constituem o correspondente espaço de cores puras associado: *Y* (b), *Cb* (c), e *Cr* (d).

Frequentemente, é posteriormente usado um filtro de média do tipo passa baixo com uma máscara de nove píxeis de dimensão de forma a eliminar algum ruído existente nas componentes *Cb* e *Cr*.

### 2.2 Modelo de Pele baseado numa Função de Probabilidade de Distribuição Normal

Na construção do modelo estatístico usado neste trabalho para a detecção de zonas de pele em imagens, usaram-se 16 imagens de amostras de zonas de pele, cada uma das quais com 46x46 píxeis de dimensão, extraídas de várias regiões faciais de 8 indivíduos de raça branca e de ambos os sexos (Fig. 2).

Normalmente, o número de amostras considerado é meramente ocasional. Contudo, um número elevado de amostras e com considerável diversidade tendem a melhorar a definição do modelo construído.



Figura 2 – 16 amostras de zonas de pele, com 46x46 pixéis de dimensão, consideradas neste trabalho para a construção do modelo estatístico a usar na detecção de zonas de pele em imagens.

Considerando-se as componentes  $Cb$  e  $Cr$ , duas variáveis aleatórias de média  $\mu$  e covariância  $C$ , onde:

$$\mu = E \{x\}, \quad (1)$$

$$C = E \left\{ (x - \mu)(x - \mu)^T \right\}, \quad (2)$$

e sendo:

$$x = (Cr \ Cb)^T, \quad (3)$$

pode-se construir a função de probabilidade que segue uma Distribuição Normal:

$$P(Cr, Cb) = e^{-\frac{1}{2}(x-\mu)^T C^{-1}(x-\mu)}. \quad (4)$$

Na Fig. 3, representa-se graficamente a função de probabilidade (Eq. 4) considerada neste trabalho, que resulta da atribuição dos valores de intensidade na gama de tons de cinzento às variáveis aleatórias  $Cr$  e  $Cb$ .

Em função das amostras consideradas na construção do modelo, os valores de intensidade média obtida destas variáveis foram respectivamente 143 e 113 pixéis. Estes valores identificam e localizam o ponto de máxima probabilidade; isto é, o ponto no qual é máxima a probabilidade de um pixel de uma imagem pertencer a uma região de pele.

### 2.3 Detecção de Zonas de Pele

Com base no modelo construído, para qualquer imagem original, o algoritmo de detecção de zonas de pele permite gerar a imagem correspondente aos níveis de probabilidade calculados em todos os pixéis da imagem original (Fig. 4a). Na imagem de probabilidade resultante os pixéis com maior probabilidade têm valores próximos

do branco, já os pixéis de valores próximos do preto correspondem aos pixéis de menor probabilidade.

Uma vez que a imagem obtida contempla uma gama ampla de probabilidades, o processo de definição das regiões de pele numa imagem original conclui-se após a binarização da mesma (Fig. 4b). Para esse efeito, é usado um conjunto de limiares de probabilidade (sistema de limiares adaptativo), a partir dos quais são determinadas as correspondentes imagens binárias e a energia associada a cada uma. A menor variação de energia entre dois limiares de probabilidade consecutivos permite encontrar o limiar óptimo; usualmente, é considerado o menor valor entre dois limiares consecutivos.

Finalmente, contabiliza-se o número de diferentes regiões de pele detectadas que corresponde ao número de segmentos a considerar na fase seguinte.

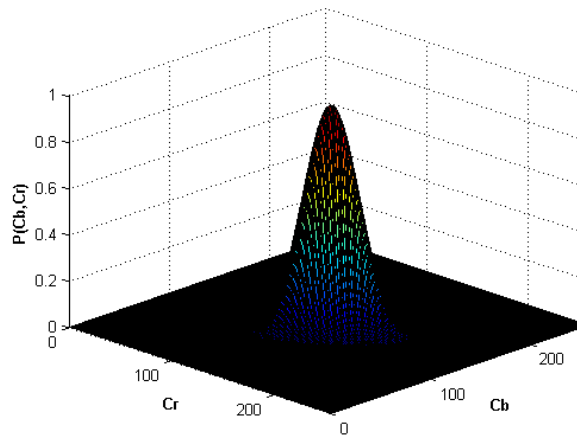


Figura 3 – Representação da função de probabilidade considerada (Eq. 4) que segue uma distribuição normal obtida usando um conjunto de 16 amostras de zonas de pele.

## 2.4 Detecção de faces

Com base na informação adquirida a partir do algoritmo de detecção de zonas de pele numa imagem, considera-se a possibilidade de prever, ainda que com baixa confiança, se os segmentos de pele determinados contêm ou não faces. Tal necessidade surge devido ao facto do algoritmo de segmentação de zonas de pele não detectar no interior dos segmentos características faciais como olhos e boca. Deste modo, com o objectivo de filtrar apenas os segmentos de pele detectados que poderão ser associados a faces, introduz-se um método heurístico auxiliar que contabiliza o número de regiões “ocas” existentes no interior de cada um dos segmentos. Como critério de selecção, considera-se que no interior de cada segmento determinado deverão existir pelo menos três regiões “ocas” (Figs. 4b e 4c): duas relativas aos olhos e uma terceira relativa à boca. Normalmente, a maioria dos segmentos previamente determinados pelo algoritmo de detecção de zonas de pele não cumprem este critério, e como tal, são eliminados do processamento seguinte.

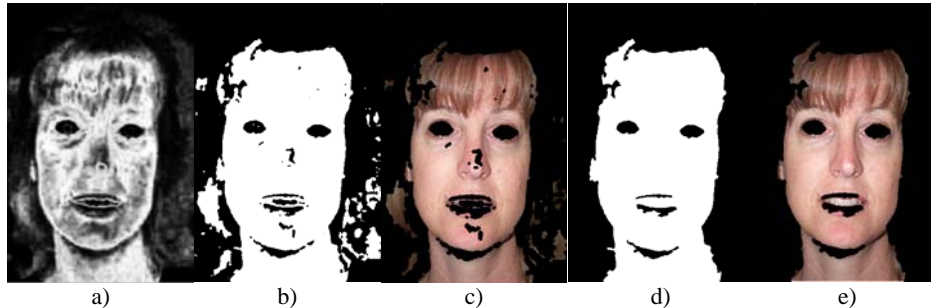


Figura 4 – Segmentação de uma face numa imagem baseada na detecção de pele: imagem de probabilidade (a); imagem binária representativa dos segmentos identificados como pele (b); imagem original representativa dos segmentos identificados (c); segmento da imagem binária filtrado e identificado como potencial face (d); e segmento da imagem original filtrado e identificado como potencial face (e).

De modo a conferir uma maior consistência ao procedimento heurístico considerado, previamente em cada segmento de pele identificado é usado um filtro de forma a eliminar as regiões “ocas” que apresentam um número reduzido de pixéis (Figs. 4d e 4e); no caso presente, inferior ou igual a 500 pixéis em virtude da dimensão das imagens usadas ser de 342x512 pixéis.

Finalmente, considera-se ainda que o algoritmo de detecção de zonas de pele associado ao método heurístico introduzido não garante ainda com elevada confiança a correcta identificação de uma face em cada um dos segmentos determinados. Pois existe sempre a possibilidade de outros segmentos que não correspondam a faces apresentarem no seu interior um número bastante considerável de regiões “ocas”. Assim, considera-se o uso de uma metodologia complementar que se baseia no uso de um modelo protótipo deformável para os olhos que é apresentada na secção seguinte. Para este efeito, na imagem de faces original devem ser identificadas as características faciais que correspondam às localizações das regiões “ocas” determinadas em cada segmento de pele (Fig. 5).

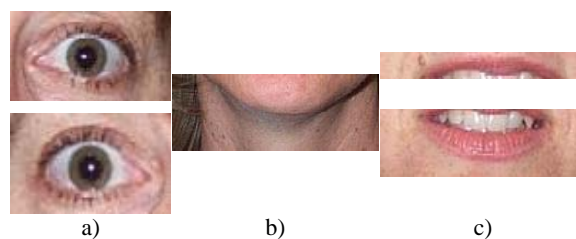


Figura 5 – Imagens representativas das características faciais a identificar na imagem original: segmentos relativos aos olhos (a); segmento relativo ao queixo (b); e segmentos relativos à boca (c).

O processo de segmentação secundário anteriormente referido consiste assim em isolar um outro tipo de segmentos de pele, designados de secundários, extraídos da imagem original que, entre outras, deverão conter características faciais de interesse,

como olhos e boca. Seguidamente, nestes segmentos secundários procura-se emparelhar um modelo protótipo deformável com geometria adequada; permitindo assim, com base no resultado do emparelhamento verificado, validar ou não a existência de uma face em cada um destes segmentos secundários.

### 3 Metodologia baseada num Protótipo Deformável para os Olhos

Nesta secção, apresenta-se a constituição geométrica do modelo protótipo deformável considerado para os olhos, define-se a função de energia usada no processo de optimização, apresenta-se o método de optimização que conduz o emparelhamento do modelo protótipo num segmento secundário previamente determinado e, finalmente, descreve-se a estratégia seguida na actualização dos parâmetros usados no modelo protótipo.

#### 3.1 Modelo Protótipo Deformável

O modelo protótipo deformável usado neste trabalho para os olhos (Fig. 6), é controlado por um conjunto de onze parâmetros e é geometricamente definido por: uma circunferência, que define o contorno da íris e controlada por três parâmetros, o raio  $r$  e o centro da íris  $(x_c, y_c)$ ; e duas parábolas, uma côncava (Eq. 10) e uma segunda convexa (Eq. 11), que definem o contorno superior e inferior do olho respectivamente e são controladas por um conjunto de sete parâmetros, o centro do olho  $(x_e, y_e)$ , dois pontos  $P_1$  e  $P_2$ , localizados no interior da córnea óptica, o ângulo de rotação  $t$ , a largura do olho  $2b$ , e as alturas do contorno superior e inferior do olho ao seu centro,  $a$  e  $c$ .

A orientação do modelo protótipo é controlada por três parâmetros: o ângulo de rotação  $t$ , e os pontos de orientação  $P_1$  e  $P_2$ . Por sua vez, a escala é gerida por quatro parâmetros,  $a$ ,  $b$ ,  $c$ , e  $r$ . Finalmente, a posição é conduzida por dois pontos, aos quais correspondem quatro parâmetros relativos às coordenadas  $(x_e, y_e)$  e  $(x_c, y_c)$ .

Para determinar a orientação do modelo protótipo, são considerados dois vectores unitários:

$$\vec{e}_1 = (\cos(t), \text{sen}(t)), \quad (5)$$

$$\vec{e}_2 = (-\text{sen}(t), \cos(t)). \quad (6)$$

Assim, qualquer ponto  $\vec{u}$ , pode ser representado no espaço 2D em função das suas coordenadas  $(x, y)$ , dado que:

$$\vec{u} = x \vec{e}_1 + y \vec{e}_2, \quad (7)$$

sendo que, desta forma, em torno do ponto de coordenadas  $(x_e, y_e)$  os pontos  $P_1$  e  $P_2$  são representados pelas equações:



$$\overline{u}_1 = (x_e + P_1 \cos(t), y_e + P_1 \sin(t)), \quad (8)$$

$$\overline{u}_2 = (x_e - P_2 \cos(t), y_e - P_2 \sin(t)). \quad (9)$$

As equações seguintes definem respectivamente o contorno superior e inferior do olho:

$$y = y_e + a - \frac{a}{b^2}(x - x_e)^2, \quad (10)$$

$$y = y_e - c + \frac{c}{b^2}(x - x_e)^2. \quad (11)$$

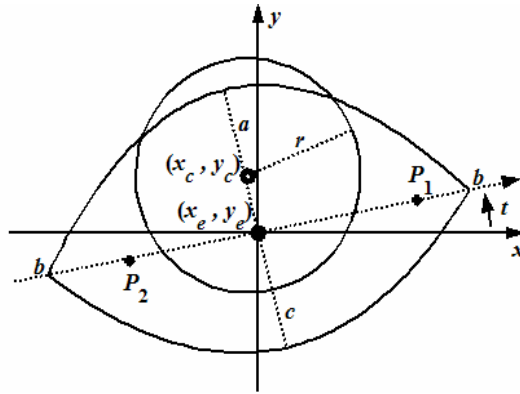


Figura 6 – Modelo protótipo deformável usado para detectar os olhos em segmentos secundários previamente obtidos na imagem original, controlado por onze parâmetros:  $a$ ,  $b$ ,  $c$ ,  $x_e$ ,  $y_e$ ,  $x_c$ ,  $y_c$ ,  $r$ ,  $P_1$  e  $P_2$ .

### 3.2 Emparelhamento

Para conduzir o processo de emparelhamento do modelo protótipo usado num dado segmento secundário previamente determinado, são extraídos da imagem original associada um conjunto de campos de energia, obtidos a partir da aplicação de adequados operadores de imagem que realçam determinadas características do objecto a identificar. Como já referido, consideram-se neste trabalho quatro campos de energia: vales de intensidade, picos de intensidade, orlas de intensidade e intensidades em tons de cinzento (Fig. 7).

A essência do emparelhamento desejado reside no facto dos campos de energia atraírem o modelo protótipo usado para as características associadas aos olhos em imagens, com base na direcção e na norma dos vectores gradiente de cada um. Nomeadamente, o campo de energia correspondente aos vales de intensidade atrai a

circunferência do modelo para a íris; isto porque, este campo tem um gradiente elevado no interior da íris. Por seu lado, o campo das orlas de intensidade, obtido neste trabalho usando o detector de orlas de *Canny*, (Canny, 1986) e (Tavares, 1995), apresenta um gradiente elevado nas zonas de grande contraste de intensidades; isto é, nos limites da íris e nas zonas do contorno do olho, servindo essencialmente para realizar pequenos ajustes ao modelo em emparelhamento. Por outro lado, o campo de energia correspondente aos picos de intensidade apresenta um gradiente elevado na parte branca do olho, a córnea óptica, permitindo assim controlar a rotação do modelo através da atracção dos pontos  $P_1$  e  $P_2$  para o centro da córnea. Finalmente, o campo de energia das intensidades em tons de cinzento contém informação geral sobre a distribuição do brilho no interior da íris e da córnea óptica, atraindo o modelo para essas zonas de maior relevo.

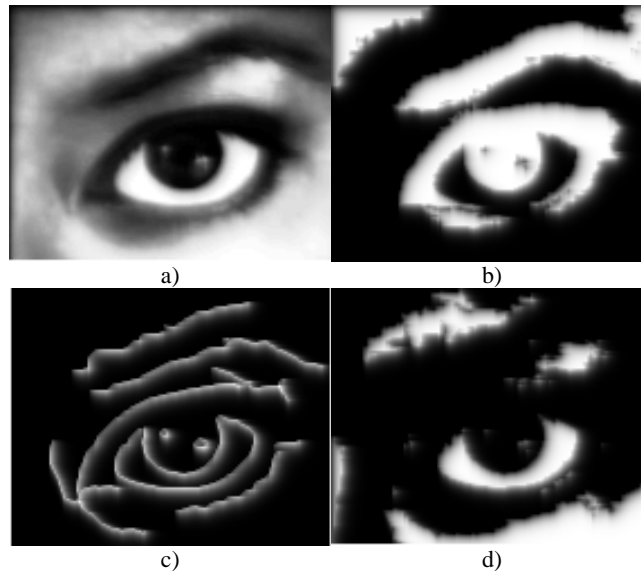


Figura 7 – Imagens representativas dos quatro campos de energia usados: intensidades em tons de cinzento (a), vales de intensidade (b), orlas de intensidade (c), picos de intensidade (d) (retiradas de (Carvalho & Tavares, 2007b)).

### 3.3 Método de Emparelhamento

#### 3.3.1 Função de Energia

A energia total  $E$  é uma função implementada a partir da soma de um conjunto de primitivas que estabelece a ligação entre os campos de energia previamente referidos e os parâmetros do modelo deformável usado, sendo representada por:

$$E = E_v + E_e + E_p + E_i + E_{priori}, \quad (12)$$

onde  $E_v$  é a energia dos vales de intensidade,  $E_e$  é a energia das orlas de intensidade,  $E_p$  é a energia dos picos de intensidade,  $E_i$  é a energia da imagem representada em tons de cinzento e  $E_{priori}$  é a energia interna do modelo.

Durante o processo de emparelhamento,  $E_v$  deve ser maximizada no interior da circunferência do modelo, e o seu valor corresponde à soma de todas as intensidades no interior da mesma dividida pela respectiva área:

$$E_v = -\frac{C_1}{A_1} \iint_{R_c} \varphi_v(\bar{u}) dA, \quad (13)$$

onde  $C_1$  representa uma constante de peso multiplicativa a determinar por experimentação,  $A_1$  é a área da circunferência e  $R_c$  é o domínio interior da mesma.

Também a energia  $E_e$  deve ser maximizada e o seu valor corresponde à soma de todas as intensidades localizadas sobre os contornos da circunferência e das parábolas divididas pelos respectivos comprimentos:

$$E_e = -\frac{C_2}{L_1} \int_{C_b} \varphi_e(\bar{u}) ds - \frac{C_3}{L_2} \int_{P_b} \varphi_e(\bar{u}) ds, \quad (14)$$

onde  $C_2$  e  $C_3$  representam constantes de peso multiplicativas a determinar experimentalmente,  $L_1$  é o comprimento dos arcos das parábolas,  $L_2$  é o perímetro da circunferência, e  $C_b$  e  $P_b$  representam os contornos da circunferência e das parábolas, respectivamente.

Por outro lado, a energia  $E_i$  é constituída por dois termos distintos: o primeiro, refere que a soma das intensidades representativas do brilho no interior da circunferência dividida pela respectiva área deve ser minimizada; por outro lado, a soma do mesmo tipo de intensidades na zona da córnea óptica, dividida pela respectiva área, deve ser maximizada. Desta forma, temos:

$$E_i = \frac{C_4}{A_1} \iint_{R_c} \varphi_i(\bar{u}) dA - \frac{C_5}{A_2} \iint_{R_s} \varphi_i(\bar{u}) dA, \quad (15)$$

onde  $C_4$  e  $C_5$  representam constantes de peso multiplicativas a definir em função da aplicação,  $A_1$  e  $A_2$  representam as áreas da circunferência e da córnea óptica, respectivamente, e  $R_c$  e  $R_s$ , os correspondentes domínios.

Por outro lado, o posicionamento dos pontos  $P_1$  e  $P_2$  permite determinar a energia  $E_p$ , que deve ser minimizada nesses mesmos pontos usando:

$$E_p = C_6 \left( \varphi_p(\bar{u} + P_1 \bar{e}_1) + \varphi_p(\bar{u} - P_2 \bar{e}_1) \right), \quad (16)$$

onde  $C_6$  representa uma constante de peso multiplicativa a definir.

Finalmente, a energia interna  $E_{priori}$  garante que o modelo não se deforma até à situação limite de fechar-se sobre si mesmo ou se expanda para lá do domínio da imagem. Deste modo, considera-se:

$$E_{priori} = \frac{K_1}{2} \left( (x_e - x_c)^2 + (y_e - y_c)^2 \right) + \frac{K_2}{2} \left( P_1 - \frac{1}{2}(r+b) \right)^2 + \frac{K_2}{2} \left( P_2 + \frac{1}{2}(r+b) \right)^2 + \frac{K_3}{2} (b-2r)^2, \quad (17)$$

onde  $K_1$ ,  $K_2$  e  $K_3$  representam constantes de peso multiplicativas também a definir experimentalmente.

### 3.3.2 Algoritmo de Optimização

O processo de emparelhamento do modelo usado numa imagem baseia-se na actualização dos seus parâmetros dinamicamente e iterativamente ao longo de sete fases, usando o método de optimização gradiente descendente. Como exemplo, o raio  $r$  da íris do olho é iterativamente calculado e actualizado da seguinte forma:

$$r_{new} = r_{old} + dt \times r'(t), \quad (18)$$

considerando que:

$$r'(t) = \frac{dr}{dt} = -\frac{\partial E}{\partial r}, \quad (19)$$

onde  $r_{new}$  e  $r_{old}$  são respectivamente os valores dos raios entre duas iterações consecutivas,  $dt$  é o passo de tempo, e finalmente,  $r'(t)$  é a taxa de variação do raio no tempo, usualmente traduzido pelo simétrico da variação da energia total de deformação considerando a variação do parâmetro em causa.

Assim, em cada iteração é determinada a energia total  $E$ , considerando um novo conjunto de parâmetros previamente actualizados na iteração anterior; encerrando-se o processo de emparelhamento quando o valor absoluto do erro medido entre duas somas de energia consecutivas for inferior a um valor predefinido. Desta forma, dependendo do valor final da energia de emparelhamento, considera-se que o modelo se ajusta ou não satisfatoriamente ao olho. Caso esse valor seja baixo, o emparelhamento obtido permite a sua detecção correcta e, considerando-se o conjunto de parâmetros resultantes da última iteração realizada, extraem-se as correspondentes características do olho.

### 3.3.3 Actualização de Parâmetros

Seguidamente, descreve-se o processo de emparelhamento implementado que permite a actualização dinâmica dos parâmetros do modelo protótipo deformável em função da estratégia adoptada. Deve-se salientar que outras estratégias a considerar poderão ser igualmente funcionais.

Primeiramente define-se um conjunto de valores iniciais, de forma a compatibilizar a geometria do modelo usado à geometria característica dos olhos. Para tal, considerou-se neste trabalho as seguintes relações iniciais entre parâmetros:  $x_e = x_c$ ,  $y_e$

$= y_c, P_1 = r, P_2 = -r, t = 0, a = r, b = 3r - r/4, c = r$ , sendo os valores das constantes multiplicativas considerados nulos. Posteriormente, os valores atribuídos às mesmas em cada fase de emparelhamento são determinados empiricamente com base num conjunto de ensaios experimentais, nos quais se analisou a influência desejada de cada campo de energia na variação dos parâmetros do modelo.

Assim, na primeira fase de emparelhamento, incrementa-se a constante de peso multiplicativa  $C_1$  e minimiza-se  $E = E_v$ , o que permite posicionar o modelo em torno do centro da íris, por actualização dos parâmetros  $x_c$  e  $y_c$ , mantendo o raio  $r$  constante.

Já na segunda fase, à energia total  $E$  anterior, adiciona-se  $E_i$  e  $E_e$ , incrementando-se as constantes  $C_2$  e  $C_4$ , o que permite determinar a posição e a escala correcta da íris, por actualização dos parâmetros  $x_c, y_c$  e  $r$ .

Na fase seguinte, incrementa-se as constantes  $C_1, C_2$  e  $C_4$ , com o objectivo de fazer um ajuste fino da circunferência do modelo à íris do olho, actualizando apenas o respectivo raio  $r$ .

Já na quarta fase de emparelhamento, a mais delicada de todo o processo computacional, procura-se minimizar  $E = E_p$ , anulando-se as constantes anteriores e incrementando apenas a constante  $C_6$ . Este procedimento permite orientar o modelo através da actualização do ângulo de rotação  $t$  e da localização dos pontos  $P_1$  e  $P_2$ .

Uma vez na quinta fase do processo interactivo, escala-se as parábolas do modelo nas suas posições correctas, considerando-se a minimização da soma das energias  $E_p, E_i$  e  $E_e$ . Os parâmetros actualizados nesta fase são: os pontos  $P_1$  e  $P_2$ , o ângulo de rotação  $t$ , as coordenadas do centro do olho  $x_e$  e  $y_e$ , e a largura do olho  $b$ .

Na fase seguinte, minimiza-se a mesma soma de energias considerada na fase anterior, actualizando-se a largura do olho  $b$  e as respectivas alturas  $a$  e  $c$ .

Finalmente, na sétima e última fase, usando a soma de todas as energias consideradas, actualizam-se em simultâneo todos os parâmetros do modelo.

No final, o emparelhamento diz-se adequado se a energia total determinada no emparelhamento é mínima e a relação entre os coeficientes determinados na sétima e última fase é aceitável; isto é, se após a deformação verificada no emparelhamento, o modelo apresenta uma geometria compatível com a inicial.

Na Fig. 8, apresenta-se um exemplo que ilustra o correcto emparelhamento do modelo protótipo deformável considerado neste trabalho numa imagem que contém o olho como característica facial a detectar.

## 4 Resultados Experimentais

Nesta secção são apresentados alguns resultados obtidos, usando implementações desenvolvidas em *Matlab* das metodologias anteriormente apresentadas, em três ensaios experimentais.

Por questões de simplicidade, nos ensaios considerados, na fase de emparelhamento do modelo protótipo deformável nos diversos segmentos secundários previamente determinados, seleccionaram-se apenas aqueles que contêm efectivamente o olho esquerdo e o olho direito.

Nos resultados experimentais apresentados nas Tabelas 1, 2 e 3, considera-se que a origem do referencial considerado localiza-se no vértice inferior esquerdo de cada um

dos segmentos secundários que contém um olho, sendo os valores das coordenadas apresentados relativos a esse mesmo referencial.

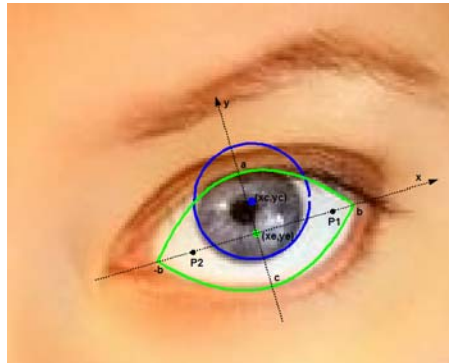


Figura 8 – Imagem representativa do emparelhamento dinâmico do modelo protótipo deformável do olho numa imagem (retirado de (Carvalho & Tavares, 2007a)).

Nos três ensaios realizados verifica-se que, por um lado, o algoritmo de detecção de zonas de pele permite segmentar a totalidade da face presente, exibindo-a totalmente num dos segmentos obtidos e, por outro, que é adequado o emparelhamento do modelo protótipo deformável usado nos segmentos secundários que contém um olho, o que permite detectar e validar com elevada confiança a presença de uma face (Figs. 9e, 10e e 11e).

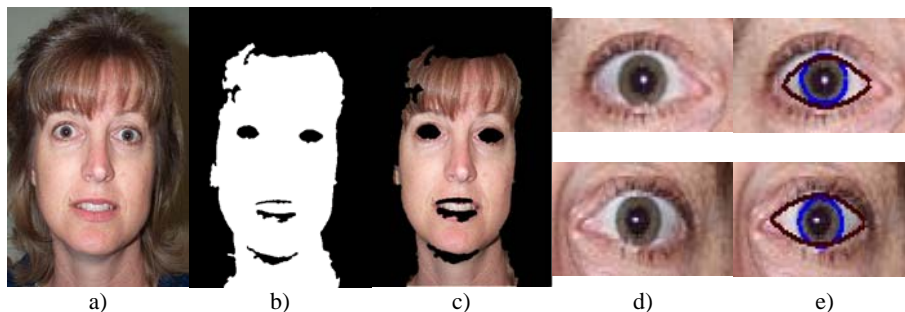


Figura 9 – Imagens dos resultados obtidos no 1º ensaio: imagem original (a); imagem binária correspondente ao segmento de pele que contém cinco características faciais (b); imagem original correspondente às zonas de pele globalmente identificadas no segmento seleccionado (c); segmentos secundários da imagem original representativos das características faciais olho esquerdo e olho direito (d); imagem com o melhor emparelhamento obtido do modelo protótipo deformável nos segmentos secundários correspondentes ao olho esquerdo e ao olho direito (e).

Após um emparelhamento bem sucedido, verifica-se que os valores finais dos parâmetros do modelo protótipo deformável usado traduzem adequadamente a similaridade existente entre o olho esquerdo e o olho direito de cada face.

A análise de resultados obtidos, permite ainda verificar que os valores finais dos parâmetros do modelo protótipo usado são normalmente muito distintos dos valores

iniciais atribuídos. Contudo, a geometria inicial do mesmo mantém-se o que, em conjunto com o valor de energia determinado, indica que o emparelhamento foi bem sucedido.

Tabela 1 – Resultados do 1º ensaio correspondentes à actualização dos parâmetros do modelo no final de cada fase de emparelhamento.

Parâmetros		Fases de Actualização dos Parâmetros do Modelo no Olho Esquerdo							
	Unidades	0	1	2	3	4	5	6	7
<i>t</i>	Radianos	0.00	-	-	-	0.01	0.09	-	0.02
<i>a</i>	Píxeis	11.00	-	-	-	-	-	11.01	11.46
<i>b</i>		30.25	-	-	-	-	20.81	20.69	20.19
<i>c</i>		11.00	-	-	-	-	-	10.29	9.59
<i>P<sub>1</sub></i>		11.00	-	-	-	20.80	21.26	-	21.24
<i>-P<sub>2</sub></i>		11.00	-	-	-	11.68	12.27	-	12.80
<i>xe</i>		20.00	42.72	42.28	-	41.03	41.87	-	42.75
<i>ye</i>		20.00	23.15	22.68	-	20.46	21.59	-	20.88
<i>xc</i>		20.00	42.72	42.28	-	-	-	-	42.13
<i>yc</i>		20.00	23.15	22.68	-	-	-	-	23.14
<i>r</i>		11.00	-	11.00	11.75	-	-	-	10.07
Parâmetros		Fases de Actualização dos Parâmetros do Modelo no Olho Direito							
	Unidades	0	1	2	3	4	5	6	7
<i>t</i>	Radianos	0.00	-	-	-	-0.02	-0.02	-	-0.03
<i>a</i>	Píxeis	13.00	-	-	-	-	-	10.34	10.63
<i>b</i>		35.75	-	-	-	-	21.83	22.02	23.39
<i>c</i>		13.00	-	-	-	-	-	9.62	10.36
<i>P<sub>1</sub></i>		13.00	-	-	-	15.19	18.22	-	17.64
<i>-P<sub>2</sub></i>		13.00	-	-	-	15.09	11.30	-	12.32
<i>xe</i>		20.00	43.76	44.39	-	44.97	41.51	-	42.72
<i>ye</i>		20.00	23.07	24.25	-	23.86	23.80	-	24.27
<i>xc</i>		20.00	43.76	44.39	-	-	-	-	43.59
<i>yc</i>		20.00	23.07	24.25	-	-	-	-	23.51
<i>r</i>		13.00	-	10.00	10.40	-	-	-	10.22

A título de exemplo, no primeiro ensaio realizado foram considerados diferentes valores iniciais para o raio do olho esquerdo e para o raio do olho direito, respectivamente 11 e 13 píxeis, tendo-se obtido após emparelhamento do modelo 10.07 e 10.22 píxeis para esses raios. Verificou-se assim que os valores iniciais atribuídos aos parâmetros do modelo, ainda que inicialmente distintos para cada olho, permitiram obter emparelhamentos satisfatórios e valores de energia idênticos que confirmam a esperada similaridade entre a íris do olho esquerdo e a íris do olho direito. Tal facto mostra a existência de alguma flexibilidade existente na escolha dos parâmetros iniciais do modelo desde que estejam de acordo com o domínio de acção dos campos de energia considerados.



Figura 10 – Imagens dos resultados obtidos no 2º ensaio: imagem original (a); imagem binária correspondente ao segmento de pele identificado que contém seis características faciais (b); imagem original correspondente às zonas de pele globalmente identificadas no segmento seleccionado (c); segmentos secundários da imagem original representativos das características faciais olho esquerdo e olho direito (d); imagem com o melhor emparelhamento obtido do modelo protótipo deformável nos segmentos secundários da imagem original correspondentes ao olho esquerdo e ao olho direito (e).

Tabela 2 – Resultados do 2º ensaio correspondentes à actualização dos parâmetros do modelo no final de cada fase de emparelhamento.

Parâmetros		Fases de Actualização dos Parâmetros do Modelo no Olho Esquerdo							
	Unidades	0	1	2	3	4	5	6	7
$t$	Radianos	0.00	-	-	-	0.02	0.02	-	-0.10
$a$	Pixels	12.00	-	-	-	-	-	15.41	14.46
$b$		33.00	-	-	-	-	15.25	17.05	17.48
$c$		12.00	-	-	-	-	-	5.64	5.44
$P_1$		12.00	-	-	-	14.65	16.41	-	15.62
$-P_2$		12.00	-	-	-	13.10	11.13	-	12.56
$xe$		23.00	34.58	34.64	-	36.43	35.07	-	35.96
$ye$		23.00	31.80	30.26	-	23.90	24.71	-	24.68
$xc$		23.00	34.58	34.64	-	-	-	-	34.51
$yc$		23.00	31.80	30.26	-	-	-	-	30.37
$r$		12.00	-	10.00	10.26	-	-	-	10.20
Parâmetros		Fases de Actualização dos Parâmetros do Modelo no Olho Direito							
	Unidades	0	1	2	3	4	5	6	7
$t$	Radianos	0.00	-	-	-	0.08	0.13	-	0.01
$a$	Pixels	12.00	-	-	-	-	-	14.56	15.05
$b$		33.00	-	-	-	-	15.36	17.73	17.47
$c$		12.00	-	-	-	-	-	3.98	4.37
$P_1$		12.00	-	-	-	15.46	15.58	-	15.39
$-P_2$		12.00	-	-	-	12.84	10.45	-	11.54
$xe$		23.00	43.15	42.01	-	42.05	41.59	-	41.56
$ye$		23.00	31.14	30.08	-	25.38	25.39	-	24.83
$xc$		23.00	43.15	42.01	-	-	-	-	41.92
$yc$		23.00	31.14	30.08	-	-	-	-	30.27
$r$		12.00	-	10.00	9.51	-	-	-	9.73



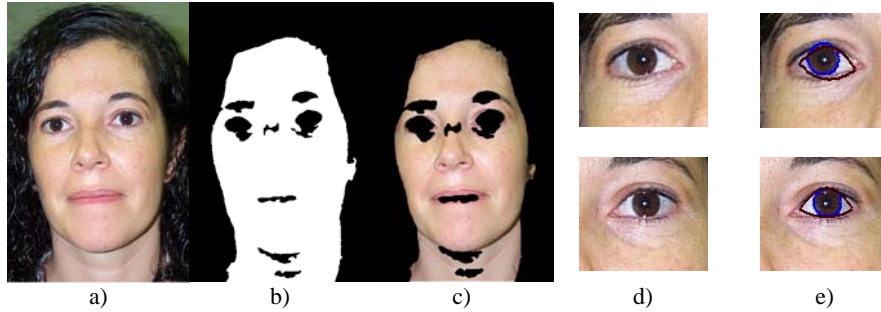


Figura 11 – Imagens dos resultados obtidos no 3º ensaio: imagem original (a); imagem binária correspondente ao segmento de pele identificado que contém oito características faciais (b); imagem original correspondente às zonas de pele globalmente identificadas no segmento seleccionado (c); segmentos secundários da imagem original representativos das características faciais olho esquerdo e olho direito (d); imagem com o melhor emparelhamento obtido do modelo protótipo deformável nos segmentos secundários da imagem original correspondentes ao olho esquerdo e ao olho direito (e).

Tabela 3 – Resultados do 3º ensaio correspondentes à actualização dos parâmetros do modelo no final de cada fase de emparelhamento.

Parâmetros		Fases de Actualização dos Parâmetros do Modelo no Olho Esquerdo							
	Unidades	0	1	2	3	4	5	6	7
$t$	Radianos	0.00	-	-	-	0.12	- 0.09	-	- 0.11
$a$	Pixels	16.00	-	-	-	-	-	13.70	15.00
$b$		44.00	-	-	-	-	21.98	22.26	22.52
$c$		16.00	-	-	-	-	-	10.68	10.49
$P_1$		16.00	-	-	-	16.53	20.16	-	19.16
$-P_2$		16.00	-	-	-	21.34	14.66	-	17.15
$xe$		45.00	49.04	50.80	-	59.28	51.48	-	54.46
$ye$		45.00	54.59	53.87	-	47.77	48.68	-	49.06
$xc$		45.00	49.04	50.80	-	-	-	-	51.33
$yc$		45.00	54.59	53.87	-	-	-	-	54.89
$r$		16.00	-	13.00	12.31	-	-	-	12.27
Parâmetros		Fases de Actualização dos Parâmetros do Modelo no Olho Direito							
	Unidades	0	1	2	3	4	5	6	7
$t$	Radianos	0.00	-	-	-	0.11	0.03	-	0.00
$a$	Pixels	16.00	-	-	-	-	-	18.57	18.58
$b$		44.00	-	-	-	-	22.64	23.69	23.70
$c$		16.00	-	-	-	-	-	7.07	7.05
$P_1$		16.00	-	-	-	18.99	19.85	-	19.73
$-P_2$		16.00	-	-	-	13.19	12.97	-	12.96
$xe$		45.00	61.90	61.39	-	59.69	59.25	-	59.38
$ye$		45.00	61.59	61.12	-	52.10	55.16	-	54.99
$xc$		45.00	61.90	61.39	-	-	-	-	61.05
$yc$		45.00	61.59	61.12	-	-	-	-	61.08
$r$		16.00	-	13.00	13.07	-	-	-	12.36

## 5 Conclusões e Perspectivas de Trabalho Futuro

Os resultados experimentais obtidos, alguns apresentados neste artigo, permitem concluir que a metodologia baseada no algoritmo de detecção de zonas de pele mostra-se eficaz na segmentação das diferentes zonas de pele existentes na imagem original, bem como permite a posterior determinação da localização de características faciais presentes no interior de cada segmento de pele obtido.

Já a metodologia baseada no modelo protótipo deformável para os olhos, usada neste trabalho de forma complementar à metodologia de detecção de zonas de pele para a detecção das faces em imagens, no que diz respeito à sua aplicabilidade mostra-se bastante mais sensível do que a anterior. Por um lado, o emparelhamento considera um conjunto bastante alargado de parâmetros em actualização que, uma vez interdependentes, exigem considerável sucesso em cada uma das sete fases de processamento consideradas. Por outro lado, uma deficiente definição de um dos campos de energia, os grandes condutores de todo o processo de emparelhamento do protótipo numa imagem, mostra-se suficiente para impedir o correcto emparelhamento do mesmo à característica facial pretendida.

Em conclusão, podemos considerar que o uso complementar das duas metodologias consideradas neste trabalho para a detecção de faces em imagem, permite identificar e validar faces em imagens de forma bastante satisfatória e com elevada confiança.

Como trabalho futuro, relativamente à metodologia baseada no uso do modelo protótipo deformável para os olhos, considera-se importante o uso de filtros de imagem mais eficientes que, uma vez aplicados à imagem original, permitam definir de forma substancialmente melhorada os campos de energia a usar no processo de emparelhamento do protótipo de maneira a tornar este processo mais robusto e flexível.

Finalmente, de modo a usar os modelos protótipo deformáveis no seguimento de olhos em sequências de imagens, pretende-se aumentar a velocidade de convergência do método de optimização usado, definindo-se o passo de tempo usado no gradiente descendente com base numa técnica de multipasso.

## Agradecimentos

Este trabalho foi parcialmente desenvolvido no âmbito do Projecto de Investigação “Segmentação, Seguimento e Análise de Movimento de Objectos Deformáveis (2D/3D) usando Princípios Físicos” financiado pela FCT - Fundação para a Ciência e a Tecnologia, com a referência POSC/EEA-SRI/55386/2004.

## Bibliografia

Canny, J. (1986). *A Computational Approach to Edge Detection*. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 8, no. 6, pp. 679-698.

- Carvalho, F. & Tavares, J. (2005). *Metodologias para identificação de faces em imagens: introdução e exemplos de resultados*. Paper presented at Congresso de Métodos Numéricos em Engenharia 2005. In Granada, España.
- Carvalho, F. & Tavares, J. (2006). *Two Methodologies for Iris Detection and Location In Face Images*. Presented at CompIMAGE – Computational Modelling of Objects Represented in Images: Fundamentals, Methods and Applications. Coimbra, Portugal, pp. 129-134.
- Carvalho, F. & Tavares J. (2007a). *Eye detection using a deformable template in static images*. Presented at VIPimage – I ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing. pp. Porto, Portugal, pp. 209-215.
- Carvalho, Fernando J. & Tavares, João Manuel R. S. (2007b). *Detecção e Extração de Características do Olho em Imagens usando um Modelo Protótipo Deformável*. Apresentado em DSM 2007 - Conferência Nacional de Dinâmica de Sistemas Multicorpo. Universidade do Minho, Guimarães, pp. 303-307.
- Rademacher, D. (2001). *Face detection*. In Colorado School of Mines - Introduction to computer Vision and Image. «<http://egweb.mines.edu/eges512/projects/face/Rademacher.PDF>»
- Tavares, J. (1995). *Obtenção de Estrutura Tridimensional a Partir de Movimento de Câmara*. Dissertação de Mestrado. Faculdade de Engenharia: Universidade do Porto.
- Wu, Y., Liu, H. & Zha, H. (2003). *A New Method of Human Eyelids Detection Based on Deformable Templates*. Presented at Sino-Korea Symposium on Intelligent Systems. In Guangzhou, China, pp. 49-54.
- Yang, M., Kriegman, David J. & Ahuja, N. (2002). *Detecting Faces in Images: A Survey*. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 24, no. 1, pp. 34-58.
- Yuille, A., Hallinan, P. & Cohen, D. (1992). *Feature Extraction from Faces Using Deformable Templates*. International Journal of Computer Vision. Vol. 8, no. 2, pp. 99-111.